УНИВЕРЗИТЕТ У НИШУ
ЕКОНОМСКИ ФАКУЛТЕТ НИШ

Мр Наташа М. Папић-Благојевић

КОМПАРАТИВНА АНАЛИЗА КЛАСИЧНЕ ИНФЕРЕНЦИЈЕ И БАЈЕСОВОГ ПРИСТУПА У ОБРАДИ ЕКОНОМСКИХ ПОДАТАКА

Докторска дисертација

Ниш, 2014. год.

UNIVERSITY OF NIŠ
FACULTY OF ECONOMICS

Nataša M. Papić-Blagojević

COMPARATIVE ANALYSIS OF CLASSICAL INFERENCE AND BAYESIAN APPROACH IN THE PROCESSING OF ECONOMIC DATA

Doctoral dissertation

Niš, 2014.

Ментор: Др Винко Лепојевић, ванредни професор
Универзитет у Нишу
Економски факултет у Нишу

____

Чланови комисије:

1. Др Вера Ђорђевић, редовни професор
Универзитет у Нишу
Економски факултет у Нишу

____

2. Др Мирко Савић, редовни професор
Универзитет у Новом Саду
Економски факултет у Суботици

____

Датум одбране:____________________________

НАУЧНИ ДОПРИНОС ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Научни допринос докторске дисертације огледа се у комплетној компаративној анализи класичне статистичке инференције и Бајесовог приступа уз анализу резултата примењених метода на одабраним економским подацима. Помоћу метода класичне и Бајесове статистике оцењени су резултати истраживања и на тај начин проверена тврдња да Бајесове оцене представљају унапређење класичних оцена у смислу њихове прецизности. Такође, анализом одабраних временских серија, представљена је могућност предвиђања кретања будућих појава помоћу прецизнијих Бајесових метода. Примењена методологија је детаљно објашњена уз очекивање да ће понуђена решења оцењивања и предвиђања будућих кретања бити од користи и за даља истраживања на тржишту Републике Србије.

SCIENTIFIC CONTRIBUTION OF DOCTORAL DISSERTATION

The scientific contribution of the doctoral dissertation is reflected in the complete comparative analysis of classical statistical inference and Bayesian approach together with analysis of the results of applied methods to selected economic data. The results of the research were estimated by using methods of classical and Bayesian statistics and thus the claim that the Bayesian estimates represents an improvement of classical estimates in terms of their accuracy was verified. Furthermore, by analyzing the selected time series, the possibility of forecasting the future events by using accurate Bayesian methods was presented. The applied methodology is explained in detail with the expectation that offered solutions for estimation and prediction of future trends will be useful for further researches on the market of Republic of Serbia.

ИЗЈАВА МЕНТОРА О САГЛАСНОСТИ ЗА ПРЕДАЈУ

УРАЂЕНЕ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Овим изјављујем да сам сагласан да кандидат **НАТАША ПАПИЋ-БЛАГОЈЕВИЋ** може да преда Реферату за последипломско образовање Факултета урађену докторску дисертацију под називом **КОМПАРАТИВНА АНАЛИЗА КЛАСИЧНЕ ИНФЕРЕНЦИЈЕ И БАЈЕСОВОГ ПРИСТУПА У ОБРАДИ ЕКОНОМСКИХ ПОДАТАКА**, ради организације њене оцене и одбране.

(Потпис ментора: проф. др Винко Лепојевић)

STATEMENT OF MENTOR'S CONSENT FOR SUBMISSION OF COMPLETED DOCTORAL DISSERTATION

Hereby**,** I declare that I agree that the candidate **NATAŠA PAPIĆ-BLAGOJEVIĆ**, can submit completed doctoral dissertation to the officer for postgraduate education of the Faculty under the name of: **COMPARATIVE ANALYSIS OF CLASSICAL INFERENCE AND BAYESIAN APPROACH IN THE PROCESSING OF ECONOMIC DATA** for the purpose of its evaluation and defense.

(Mentor's signature: prof. dr Vinko Lepojević)

И З Ј А В А

Под пуном материјалном и моралном одговорношћу изјављујем да је приложена докторска дисертација резултат сопственог научног истраживања и да је коришћена литература на адекватан начин цитирана, без преузимања идеја, резултата и текста других аутора на начин којим се прикрива оригиналност извора. У потпуности преузимам одговорност за спроведено истраживање, анализу, интерпретацију података и закључке.

Својеручни потпис:

____

У Нишу, дана _________________године

S T A T E M E N T

With due material and moral responsibility, hereby I declare that the doctoral dissertation is the result of personal scientific research and that the references used are cited adequately without use of ideas, results and texts of other authors in the way that hides the source's originality. I take the full responsibility for conducted research, analysis, data interpretation and conclusions.

Signature:

____

Niš, _______________________

КОМПАРАТИВНА АНАЛИЗА КЛАСИЧНЕ ИНФЕРЕНЦИЈЕ И БАЈЕСОВОГ ПРИСТУПА У ОБРАДИ ЕКОНОМСКИХ ПОДАТАКА

Апстракт

Током претходног века, па и дуже, једна од основних расправа међу статистичарима водила се око значења појма вероватноће. Суштински, општа сагласност око аксиома вероватноће постоји, али се и даље поставља питање око њене улоге у процесу закључивања. Ту почиње и раздвајање два основна приступа – класичног и Бајесовог.

Бајесова анализа је била доминантна у литератури током 19. века, међутим, у 20. веку класична статистика добија на популарности због ригорозније интерпретације вероватноће. Последњих неколико година, интересовање за Бајесове методе поново расте. Контроверзе око примене ова два приступа у конкретним околностима и данас су присутне, што је и разлог постојања две супротстављене школе. Док једни сматрају да класичне методе показују боље перформансе у решавању проблема, други проналазе аргументе који иду у прилог Бајесове статистике. Управо тај дуалитет доприноси мишљењу да је неопходно истражити могућности које и један и други приступ пружају у обради података.

Предмет истраживања докторске дисертације је анализа различитих методолошких приступа статистичког закључивања и могућност њихове примене у обради конкретних економских података. Да би се одговорило на предмет истраживања, у дисертацији је презентован историјски развој два различита приступа закључивању, сагледане карактеристике класичне и Бајесове статистике, могућности примене Бајесових метода у обради економских података и, на крају, изведени закључци о оправданости примене одабраних метода и модела у економији. Оценом резултата примењеног метода максималне веродостојности, као класичног метода, и Бајесових метода на примеру једног предузећа, као и компаративном анализом резултата добијених применом оба приступа у анализи кретања кључних финансијских показатеља у Републици Србији и неких макроекономских показатеља у Европској Унији, потврђена је основна хипотеза да

Бајесов приступ статистичком закључивању има предност у односу на класичну инференцију и да је његова примена у економији оправдана.

Кључне речи**:** вероватноћа, класична инференција, методе оцењивања, Бајесове методе, априорне информације, компаративна анализа

Научна област**:** Економија

Ужа научна област**:** Економска статистика, примена математичких и статистичких метода у економским истраживањима

УДК: 33:519.2

COMPARATIVE ANALYSIS OF CLASSICAL INFERENCE AND BAYESIAN APPROACH IN THE PROCESSING OF ECONOMIC DATA

Abstract

Over the past century or more, one of the fundamental debates among statisticians has been about the meaning of the concept of probability. Essentially, there is a general consensus on the axioms of probability but its role in the process of reasoning is still questionable. It is where the separation into two basic approaches begins - classical and Bayesian. Bayesian analysis was dominant in the literature during the 19th century; however, in the 20th century classical statistics gained popularity due to the rigorous interpretation of probability. The last few years, interest in Bayesian methods has been growing again.

The controversies around the application of these two approaches in the specific circumstances are still present today. It is the reason for the existence of two opposing schools. While some believe that classical methods show better performance in solving problems, others find the arguments in favor of Bayesian statistics. This duality contributes to the opinion that it is necessary to explore the possibilities that both approaches provide in the data processing.

The subject of the doctoral dissertation research is the analysis of different methodological approaches of statistical inference and the possibility of their use in the processing of specific economic data. In order to respond to the subject of the study, this dissertation presents the historical development of two different approaches to inference, perceived characteristics of classical and Bayesian statistics, application possibilities of Bayesian methods in the processing of economic data, and finally, conclusions regarding the feasibility of the selected methods and models in economics. By evaluating the results of the applied maximum likelihood method- as the classical method- and Bayesian method on the example of one company, as well as a comparative analysis of the results obtained using both approaches in the analysis of trends of key financial indicators in the Republic of Serbia and some macroeconomic indicators in the European Union, the basic hypothesis that Bayesian approach to statistical inference has the advantage over classical inference was confirmed, as well as was the fact that its application in economics is justified.

***Key words*:** probability, classical inference, estimation methods, Bayesian methods, prior information, comparative analysis

***Scientific field*:** Economics

***Field of Academic Expertise*:** Economic statistics, the application of mathematical and statistical methods in economic research

UDC: 33:519.2

СПИСАК ТАБЕЛА:

СПИСАК СЛИКА:

САДРЖАЈ

УВОД

Развој различитих приступа статистичком закључивању и процесу одлучивања обележио је двадесети век, тако да се у том смислу може извршити подела на три категорије статистичког закључивања: теорија оцењивања и тестирања хипотеза Фишера, Нојмана, Пирсона и других; Бајесове дедуктивне процедуре и Валдов приступ теорији одлучивања.

Бројни су документи истраживачког карактера, књиге и чланци на ову тему, којима се потврђује присутност метода статистичког закључивања у примењеним научним дисциплинама. Са математичке и методолошке тачке гледишта различити приступи су детаљно и концизно описани. Велики број стручних чланака разматра основне ставове и нуди суштинска објашњења, говори о могућностима примене различитих метода у реалним околностима, итд. У том смислу је тешко уочити јасне концептуалне разлике које су неопходне за сагледавање међусобног односа између приступа, претпоставки на којима су засновани, као и њихове практичне импликације. Из тог разлога, предмет истраживања дисертације је компаративна анализа статистичке инференције где ће се кроз детаљну теоријску анализу објаснити разлике које постоје и са појмовног и са филозофског становишта између појединих приступа.

Различити приступи статистичком закључивању произилазе, с једне стране, из става да се статистичка анализа базира на оценама добијеним на основу података из узорка, што одговара традиционалном схватању док, с друге стране, постоје схватања која узимају у обзир и другачије видове информација, где су раније стечено искуство и могуће последице такође релевантни за доношење одлука, па се сматра да је разматрање ширег концепта информација, а не само података из узорка, знатно пожељније. Већина статистичара је свесна овог дуалитета који доводи до потребе за даљим разјашњавањем карактеристика појединих приступа статистичком закључивању, уз реално сагледавање могућности њихове примене у обради актуелних економских података. На тим идејама је и дефинисана тема докторске дисертације, као и предмет њеног истраживања.

Предмет истраживања докторске дисертације је анализа различитих методолошких приступа статистичког закључивања и могућност њихове примене у обради конкретних економских података. Са предметом истраживања уско је повезан и **проблем истраживања** који је првенствено усмерен на примену Бајесове вероватноће у економији. Из тог разлога, у раду ће бити извршена анализа класичне инференције и Бајесовог приступа, прво са теоријског становишта, а затим и у практичном смислу. У новије време је, у значајној мери, дошло до интензивнијег развоја и примене Бајесовог закључивања у статистици. Не тако давно, примена Бајесових метода у решавању сложених проблема била је права реткост. Међутим, расположивост одговарајућих компјутерских алгоритама и присутност софтверских решења омогућили су истраживачима из различитих области практичну примену Бајесовог статистичког закључивања.

Класична статистика са својим начелима заузима значајно место у обради података те јој, као таквој, треба посветити посебну пажњу уз јасно указивање на њене главне предности и недостатке. Као основни циљ класичног приступа поставља се примена података x како би се дефинисала непозната вредност параметра θ. Код класичног приступа параметри се посматрају као непознате константе.

Класична статистичка инференција има широку примену у бројним областима истраживачког рада, међутим, у новије време се оспоравају неки њени теоријски и практични аспекти. Наиме, класична статистика је базирана на идеји да је вероватноћа често ограничавајућа; за присталице овог приступа је уобичајено да додељују вероватноћу догађају који се понавља, док је неизвесност присутна због случајног карактера догађаја (алеаторна неизвесност). Међутим, вероватноћа се не би смела додељивати догађајима код којих неизвесност постоји због недостатка знања (когнитивна неизвесност).[1]

Посебан допринос у развоју класичне статистике дао је Роналд Фишер чија оцена максималне веродостојности, тестови значајности, анализа варијансе и планирање експеримената и даље доминирају једним делом примењене статистике. Посматрано из данашње перспективе, може се рећи да се најинтензивнији развој класичне статистике одиграо током двадесетих и тридесетих година прошлог века. Управо у том периоду, Нојман и Пирсон су дефинисали принципе за тестирање хипотеза и предложили су метод за интервал или област поверења, што је проистекло из њихове идеје о тестовима значајности. Као резултат, одређени су интервали, односно области поверења. Фишер је више био заинтересован за оцењивање него за тестирање хипотеза, па је изградио интервал оцењивања који је био заснован на идеји фидуцијалне вероватноће, што је водило до тзв. фидуцијалних интервала.

Данас можемо рећи да је класична статистичка инференција заправо непризнат спој приступа који је заснован на p-вредности за који се залагао Фишер и приступа заснованог на нивоу α за који су се залагали Нојман и Пирсон. Сматра се да се ниво значајности, приказан преко p-вредности у Фишеровом тесту значајности, односи на вероватноћу посматраних података која представља крајње вредности нулте хипотезе. Овако дефинисана p-вредност има когнитивну улогу, тако што обезбеђује оцену индуктивног доказа против H_0 у појединачном експерименту. У Нојман-Пирсоновом тестирању хипотеза тежи се минимизирању грешке типа II, или β (погрешно неодбацивање истините нулте хипотезе), која је повезана са грешком типа I, или α (погрешно одбацивање истините нулте хипотезе). Међутим, до минимизације грешке може доћи само у ситуацијама понављања узорковања, а не у појединачним експериментима и као таква не представља средство за прикупљање доказа, већ модел одговарајућег понашања.[2] Према

[1] O'Hagan, T. 2004. „Dicing with the unknown“, *Significance Magazine* 1(3), pp. 132

[2] Hubbard, R., Bayarri, M.J., Berk, K.N. and Carlton, M.A. 2003. „Confusion over measures of evidence (p’s) versus errors (α’s) in classical statistical testing“, *The American Statistician* 57, pp. 175

неким схватањима, за већину истраживача закључивање у класичној статистици зависи од догађаја који се још нису десили, односно догађаја који се чине мање важни у односу на податке који су заиста прикупљени.[3]

Такође, процедуре које се користе код класичног закључивања понекад зависе од карактера података, мада је у одређеним ситуацијама тешко одредити који су то аспекти података релевантни. То значи да ће прецизност доказа изведеног на основу експеримента који се понавља зависити од дугорочне успешности примењене процедуре.[4]

Класична статистика користи распоред узорка као основу за доношење закључака о параметру. Из Бајесове перспективе, то је закључивање уназад. За разлику од класичне статистике, Бајесова статистика се у целини заснива на апостериорној расподели параметара за вредности које су заиста настале. Било који каснији закључак, као што је Бајесова оцена или Бајесов интервал поверења, рачунају се из апостериорне расподеле. Отуда, оцена или интервал поверења зависе од стварно насталих вредности. Бајесово закључивање је закључивање унапред.[5]

За разлику од класичног приступа, који захтева да параметри буду непознате константе, Бајесов приступ параметре посматра као непознате случајне промењиве. Расподела вероватноће, позната као априорна расподела, заснива се на ранијем искуству и представља расподелу параметара пре избора узорка који се оцењује. За израчунавање апостериорне расподеле користи се Бајесово правило. Апостериорна расподела не говори о томе каква ће бити наша уверења, већ о томе како ће се мењати под утицајем нових информација.

[3] Wagenmakers, E.J., Lee, M.D., Lodewyckx, T. and Iverson, G. 2008. „Bayesian versus frequentist inference“, In H. Hoijtink, I. Klugkist, & P. A. Boelen (Eds.). *Bayesian Evaluation of Informative Hypotheses*, 181-207. Springer, New York, pp. 6

[4] Berger O.J., Brown D.L. and Wolpert L.R. 1994. „A Unified conditional Frequentist and Bayesian test for fixed and sequential simple hypothesis testing“, *The Annals of Statistics* 22 (4), pp. 1795

[5] Bolstad, M.W. 2007. *Introduction to Bayesian Statistics, Second Edition*, John Wiley&Sons, Inc., pp. 162

На основу кратких уводних напомена о основним карактеристика, али и о разликама између класичне и Бајесове статистике јасно је да су ово подручје током његовог историјског развоја, али и у делу савремене примене, пратиле бројне критике и контроверзе. Ипак, присуство контроверзи и конструктивног критицизма допринело је изузетном развоју овог подручја чији смо сведоци данас.

Из свега наведеног проистиче и **циљ истраживања** докторске дисертације који ће обухватити:

- приказ историјског развоја два различита приступа закључивању;
- сагледавање основних карактеристика класичне статистике, као и њених предности и недостатака;
- представљање елемената Бајесове статистике уз истицање њених савремених карактеристика;
- могућност примене Бајесових метода у обради економских података, и
- извођење закључака о оправданости примене одабраних метода и модела у економији.

Полазећи од основног предмета, проблема и циља истраживања, као **главна хипотеза** докторске дисертације поставља се следећа:

Бајесов приступ статистичком закључивању има предност у односу на класичну инференцију и његова примена у економији је оправдана. У складу са тим, уз примену одговарајућих софтверских решења на одабраним подацима, вршиће се испитивање оправданости ове тврдње.

На основу постављеног циља истраживања, дефинишу се и **помоћне хипотезе**:

- Бајесов приступ је сложенији и рачунски захтевнији у уобичајеним ситуацијама, с обзиром да је потребно оценити априорну вероватноћу и испитати њену осетљивост;
- оцене добијене на основу Бајесовог приступа представљају унапређење особина оцена добијених класичним приступом у смислу њихове прецизности;
- развој одговарајућих рачунарских алата доприноси ефикаснијој примени Бајесове статистике у односу на класично закључивање;
- оцена поузданости одабраних метода предвиђања захтева квалитативну процену резултата добијених применом Бајесове статистике.

Поред уводног дела и закључка, докторска дисертација обухвата још четири структурне целине.

Први део дефинише статистичко закључивање и уводи у различите приступе статистичком закључивању. Сматра се да се било који статистички поступак који користи информације у циљу описивања практичних ситуација, може схватити као инференцијални поступак, а скуп таквих поступака као *статистичко закључивање*. Први део се, такође, бави и евалуацијом схватања појма статистичког закључивања. Док су савремени тумачи теорије вероватноће углавном усмерени на процес закључивања и доношења одлука, дотле је највећи део статистичке теорије, као организоване целине, развијен током последњих осамдесет година. Све до тог времена, није постојала јасна разлика између појединих приступа статистичком закључивању. Тек у периоду између 1920. до 1935. године први пут се помиње термин *класично закључивање*. Развој техника и процедура, који је започет у том периоду и настављен до данашњих дана, произвео је сложену методологију статистичког закључивања о којој ће и бити речи у раду.

Други део докторске дисертације се фокусира на класичну инференцију, основне циљеве и методе примене. Класична статистика, као што је већ поменуто, води порекло из радова Фишера, Нојмана и Пирсона. Назив „класична“ се примењује, пре свега, због њене традиционалне улоге као формалног статистичког метода и нешто дужег постојања у односу на Бајесово закључивање. До појаве класичних ставова у статистици дошло је са развојем Закона великих бројева и Централне граничне теореме, али и развојем концепта расподеле узорка и принципа најмањих квадрата. У овом делу ће се показати да класични приступ захтева да параметри буду непознате константе, јер се обично полази од претпоставке да би се у идеалним околностима експеримент могао понављати бесконачно много пута, а као резултат процеса оцењивања добила би се прецизна вредност параметара. Са становишта класичне статистике, разликује се неколико врста закључака који се могу извести о параметру: тачкасто оцењивање, интервално оцењивање и тестирање хипотеза. Код тачакастог оцењивања појединачна статистика се израчунава на основу података из узорка, тако да представља случајну променљиву, док је њена расподела, расподела узорка. Од метода тачкастог оцењивања, прво ће бити приказан метод момената, као један од најстаријих метода за проналажење тачкастих оцена. Реч је о веома једноставној процедури за налажење оцене једног или више параметара из популације. Други метод који ће бити приказан је метод максималне веродостојности који је установио Фишер и њему ће, кроз функцију веродостојности и оцену максималне веродостојности, бити посвећена нешто већа пажња. Други део класичног закључивања односи се на интервално оцењивање. Област или интервал оцењивања, обично се изражава као област или интервал поверења. Идеја потиче од Нојмана, иако се неки наговештаји наизглед сличног концепта појављују још у радовима Лапласа почетком 19. века. Посебну пажњу треба посветити и проблему тестирања статистичких хипотеза, као важном делу статистичког закључивања. У овом делу биће објашњена Нојман-Пирсонова фундаментална лема, као метод расуђивања у чијој основи се налази количник веродостојности. Применом Нојман-Пирсонове теорије тестирања хипотеза развија се један општи метод за проналажење најбољег (најмоћнијег) теста нулте хипотезе у односу на алтернативну хипотезу.

Трећи део указује на главне карактеристике Бајесовог приступа, објашњава Бајесову теорему и њену рекурзивну примену, као основу на којој почива Бајесова статистика. Бајесово закључивање је, у основи, инференцијална процедура која узима у разматрање како априорне информације, тако и податке из узорка. Априорне информације могу проистећи из акумулираног знања и то најчешће из ранијих посматрања сличних ситуација. У вези са априорним информацијама, биће посебно објашњене различите врсте ових информација и то информативне, неинформативне и коњуговане априорне информације, као и стање априорног незнања. За разлику од класичног, код Бајесовог приступа претпоставља се да је параметар случајна величина која може бити описана расподелом вероватноће. Таква расподела се назива *априорна расподела*, а тумачи се као лично уверење о параметру које постоји пре прикупљања података. Бајесов приступ користи посматране податке за корекцију априорне расподеле у *апостериорну расподелу*. Апостериорна расподела унапређује знање о параметрима и обезбеђује вероватноће за параметре које се могу искористити у сврху даљег истраживања у различитим областима. У овом делу ће бити објашњен и концепт емпиријских Бајесових метода које користе информације из података како би се одредила априорна вероватноћа, па тиме нарушавају основне премисе Бајесове статистике. Емпиријске методе Бајесовог приступа нарочито су популарне у макроекономским истраживањима. Концепт интервалног оцењивања може се проширити и на Бајесову област. Бајесов интервал, аналоган класичном интервалу поверења, назива се интервал поузданости. Интервал поузданости је фиксан, док је θ случајна величина. Ово је у супротности са класичним интервалом, где је интервал случајан, а θ фиксан параметар. Такође, у истраживачком смислу, пожељно би било имати такав метод тестирања хипотеза који је директан, који подржава велики број хипотеза и који даје предност априорној информацији. Сви ови критеријуми могу бити испуњени применом Бајесовог приступа тестирању хипотеза, а посебно ће бити објашњени у оквиру *Тестирања хипотеза из угла Бајесове статистике*.

Четврти део је фокусиран на саму срж истраживања, где ће се на основу изведених теоријских импликација, применом статистичких пакета IBM SPSS *Version* 21, IBM SPSS Amos *Version* 21 и RATS, паралелно анализирати примена класичне и Бајесове статистике на економским подацима. Компарацијом одабраних класичних и Бајесових метода на подацима једног предузећа и подацима преузетим са сајтова Народне банке Србије и Eurostat-a, кроз практичну примену претходно образложених приступа статистичком закључивању, извешће се коначан закључак о посматраној проблематици.

Рад се завршава закључним разматрањима и литературом.

1. СПЕЦИФИЧНОСТИ СТАТИСТИЧКОГ ЗАКЉУЧИВАЊА И ИСТОРИЈСКИ АСПЕКТ ИНФЕРЕНЦИЈЕ

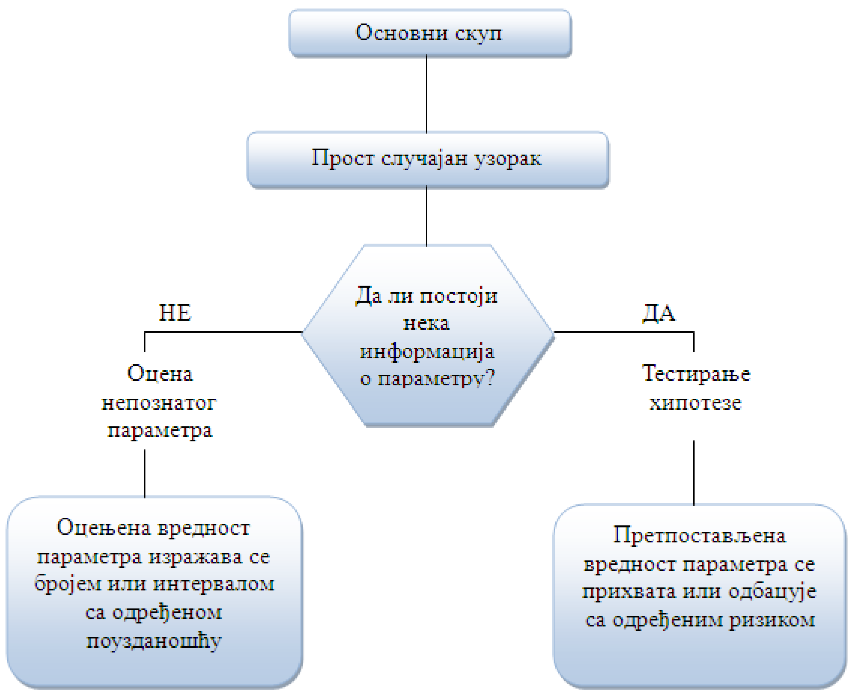
Статистичке методе су се показале као веома корисне научницима приликом тумачења резултата експеримената, истовремено им помажући у самом пројектовању експеримента. У општем смислу, сврха статистичке анализе је да организује скуп података на начин који одговара његовој структури. У одређеним ситуацијама се чини да је цео поступак врло једноставан, тако да се може стећи утисак да уопште није реч о статистици, док у неким супротним околностима поступак делује веома сложено, тако да је неопходно потражити помоћ професионалног статистичара.

Докторска дисертација анализара начине на које статистичари долазе до закључака на основу података добијених експериментом. У складу са тим, примарни циљ рада је анализа врсте закључивања коју статистичари називају *статистичо закључивање*. У том смислу, фокус ће бити на фундаменталним концептима истраживања класичног и Бајесовог закључивања. Статистички закључак се заснива на појму вероватноће, тако да је за боље разумевање фундамента инференције неопходно и боље познавање теорије вероватноће.

1.1 Основне поставке статистичког закључивања

Статистичко закључивање представља поступак доношења закључака о вредностима параметара основног скупа на основу информација добијених из узорка.[6]

Под статистичким закључивањем подразумева се и моделирање посматраних података или података из експеримента као вредности случајних промењљивих, како би се створио оквир из ког ће се извући индуктивни закључци о механизму који је довео до података.[7]



Слика 1. Области статистичког закључивања[8]

[6] Жижић, М., Ловрић, М. и Павличић, Д. 1999. *Методи статистичке анализе*, Економски факултет, Београд, стр. 171

[7] Young, G.A. and Smith, R.L. 2005. *Essentials of statistical inference*, Cambridge University Press, Cambridge, pp.1

[8] Жижић, М., Ловрић, М. и Павличић, Д. 1999. *Методи статистичке анализе*, Економски факултет, Београд, стр. 172

Циљ који је се при томе поставља односи се на процену параметра θ на основу посматраних података x, односно појединачне вредности θ_i. Избор поступка за који ћемо се определити зависи од расположивих информација о непознатом параметру основног скупа пре избора узорка (Слика 1). Ако не располажемо подацима на основу којих бисмо могли да претпоставимо вредност одређеног параметра скупа, ову вредност ћемо оценити применом неке од метода статистичког закључивања. У том смислу, идентификују се три основна типа закључивања: тачкасто оцењивање, интервално оцењивање и тестирање хипотеза. Тачкасто и интервално оцењивање се у литератури често налазе под заједничим називом *статистичко оцењивање*.

Код тачкастог оцењивања, појединачна вредност се израчунава на основу података x и користи се као оцена за θ_i. Код интервалног оцењивања, обезбеђује се скуп вредности за које се са великом вероватноћом претпоставља да садрже праву, али непознату вредност θ_i. Будући да се нумеричка вредност параметра оцењује на основу информације из узорка, не постоји потпуна сигурност у тачност донетог закључка па се, из тог разлога, закључак оцењивања прихвата са поузданошћу мањом од 100%.

У случају да је нека од особина основног скупа позната или се претпоставља њена вредност, примењује се поступак *тестирања хипотеза*. Смисао тестирања хипотеза састоји се у испитивању да ли је дошло до промене вредности параметра, односно да ли је полазна претпоставка прихватљива или није. У том смислу, поставља се хипотеза за θ_i и испитује се вероватноћа остварења хипотезе кроз процену да ли подаци x подржавају или не подржавају тако формулисану хипотезу. Другим речима, испитујемо да ли информација из узорка противречи или подржава почетно уверење о карактеристици основног скупа. С обзиром да је питање исправности донетог закључка увек отворено, претпоставка се прихвата или одбацује уз одређени ризик грешке. Циљ који је се при томе поставља односи се на процену параметра θ на основу посматраних података

Поред основног циља статистичког закључивања, који се односи на процену непознатог параметра на основу посматраних података, могу се поставити и други циљеви који би, пре свега, били усмерени на предвиђање вредности случајне променљиве која још није предмет посматрања, али чија расподела зависи од θ, али и на испитавање прихватљивости разматраног модела.

Када су у питању методе статистичког закључивања, најопштија подела која се врши јесте подела на *параметарске* и *непараметарске* методе. Основни критеријум ове поделе је строгост полазних претпоставки које морају бити испуњене да би примена одређеног модела била оправдана.

Највећи допринос развоју параметарског статистичког закључивања дали су Лаплас, Гаус и Фишер, те да би се потпуније схватила модерна теорија статистичког закључивања, неопходно је сагледати и развојни пут који је инференција прошла како би стигла до оквира у којима данас егзистира.

1.1.1 Идеје Лапласа и Гауса као зачетника статистичког закључивања

Почетну фазу историјског развоја статистичког закључивања, обележили су Лаплас (Pierre Simon Laplace), Гаус (Carl Frederich Gauss) и Фишер (Ronald Aylmer Fisher), на чијим идејама се и развила модерна теорија вероватноће.

У исто време док је радио на инверзној вероватноћи, Лаплас је развио и методе статистичког закључивања засноване на директној вероватноћи. У том периоду, проблеми у примењеној статистици искључиво су се односили на питања из демографије (стопе морталитета и учесталост рађања мушке деце), као и на проблеме из природних наука (дистрибуција грешака и закони природе). Из тог разлога, Лапласово креирање теорије тестирања и оцењивања, која је обухватила релативне фреквенције, аритметичку средину и линеарне моделе, сасвим је логичан наставак поменутог процеса. На основама Лапласовог принципа, развиле су се теорије тестирања, оцењивања и предвиђања, за дате моделе и посматрања.

Међутим, постоји извесна недоследност у Лапласовој теорији оцењивања. За биномне и мултиномне расподеле користио је највероватнију вредност као оцену, док у моделу мерења грешке уводи нови критеријум за оцену локације параметра, односно минимизира апостериорни очекивани губитак, користећи апсолутно одступање као функцију губитка. Оправдање за овај поступак се налази у томе да је апсолутно одступање природна мера квалитета оцене. Међутим, увођење функције губитка показало се као озбиљна грешка, која је отежала развој објективне теорије статистичког закључивања све до данашњих дана. То је уједно и почетак поделе између статистичког закључивања и теорије одлучивања. Период који је обележио рад Лапласа, трајао је све до 1799. год., али је и по његовом окончању остао отворен проблем аритметичке средине.

Друга развојна фаза почела је 1809-1810. године са решењем проблема аритметичке средине, који је омогућио развој два најважнија алата у статистици, развој нормалане расподеле као расподеле опсервација, као и нормалне расподеле као апроксимације расподели средине у великим узорцима. Гаус је 1809. године поставио питање: Да ли постоји расподела грешке која води до аритметичке средине, као оцене локације параметра у складу са принципом инверзне вероватноће? Гаус није поновио Лапласову грешку увођења функције губитка, већ користи највероватнију вредност параметра који се оцењује. Изједначавањем апостериорног модуса са аритметичком средином, добијена је функционална једначина са нормалном расподелом као решењем. Тиме је нормалана расподела дефинисана као математичка конструкција, при чему Гаус није упоређивао нову расподелу грешке са посматрањем.

Као непосредну реакцију на Гаусове резултате, Лаплас је имао две примедбе:[9]

- Уколико је расподела грешке нормална, тада је и апостериорна расподела нормална, а апостериорна средина и медијана једнаке; отуда, метод најмањих квадрата следи из Лапласовог метода оцене, као посебног случаја.
- Ако расподела грешке има коначну варијансу, али је и даље непозната, тада централна гранична теорема даје образложење метода.

[9] Hald, A. 2004. *A History of Parametric Statistical Inference from Bernoulli to Fisher 1713 to 1935*, University of Copenhagen, Department of applied mathematics and statistics, Copenhagen, pp. 4

Дакле, испрва су и Гаус и Лаплас користили инверзну вероватноћу за извођење метода најмањих квадрата.

1.1.2 Допринос Фишера интензивној примени статистичке инференције

Почетком прошлог века, теорија статистичког закључивања обухватала је велики број ad hoc метода, од којих су неке биле крајње контрадикторне, док је теорија малих узорака била тек у развојној фази. Нека важна питања на која је требало наћи одговор била су:[10]

- Како изабрати између директних и инверзних метода вероватноће?
- Како изабрати између различитих функција губитка?
- Како изабрати између различитих статистика које ће се применити у методи аналогије?
- Како пронаћи границе вероватноће за параметре на основу метода директне вероватноће?

Проблеми, као такви, били су предмет расправе, а на већину њих одговор је дао Фишер у периоду између 1922 и 1936. године.

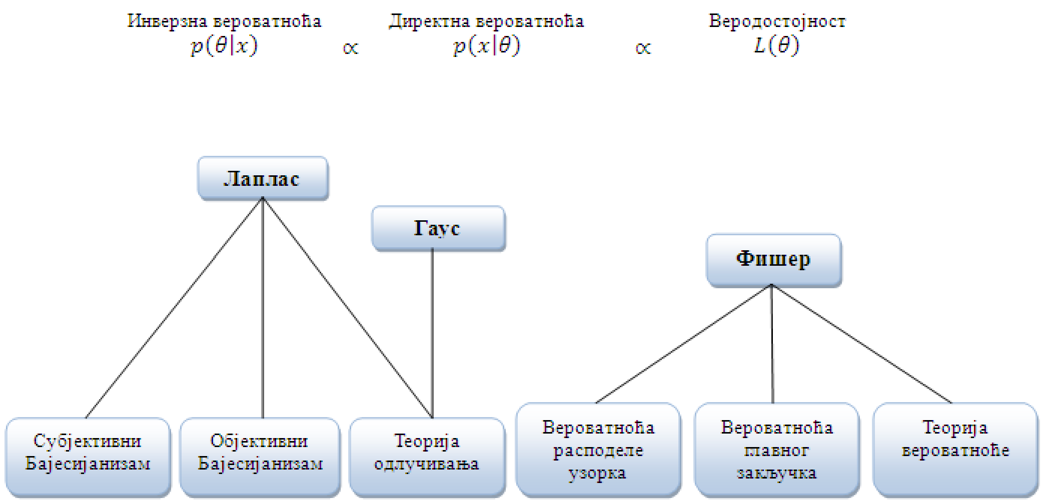
Фишер је објаснио нове статистичке идеје и технике на јасан и крајње убедљив начин, што је довело до прихватања његове теорије у веома кратком временском периоду, не само међу математичким статистичарима, већ и међу истраживачима уопште. Велики део математичке статистике од 1922. године састојао се у разради Фишерове идеје, како у теорији тако и у пракси.

Узимајући у обзир модел и извршена посматрања, Фишер је приметио да се све информације о параметрима налазе у функцији веродостојности. Тиме је доказао и асимптотску оптималност оцена изведених из ове функције, односно максималну веродостојност оцена. Базирајући своју теорију закључивања на функцији

[10] Hald, A. 2004. *A History of Parametric Statistical Inference from Bernoulli to Fisher 1713 to 1935*, University of Copenhagen, Department of applied mathematics and statistics, Copenhagen, pp. 5

веродостојности, избегао је произвољност коју су увели Лаплас и Гаус преко функције губитка и претпоставке о коначној варијанси.

Због фундаменталних односа између апостериорне густине и функције веродостојности, многи Фишерови резултати су идентични онима са Лапласове математичке тачке гледишта, једино што је неопходно ново тумачење. Међутим, тај дуг Лапласу, Фишер никада није признао. Слика 2. приказује како су идеје Лапласа, Гауса и Фишера одредиле статистичку теорију данас.



Слика 2. Допринос развоју статистичке теорије[11]

Фишеров основни мотив био је развој логике индуктивног закључивања, а централно место његовог расуђивања заузимао је принцип поновљеног узорковања. На тај начин је настојао да скрене пажњу од априорне претпоставке Бајесове школе. У Фишеровом приступу закључивању, главну улогу имао је концепт вероватноће и, већ споменути, принцип максималне веродостојности. У суштини, веродостојност мери вероватноћу да ће параметру θ бити додељене различите вредности, под претпоставком понављања експеримента и поновног посматрања стварних података.

[11] Hald, A. 2004. *A History of Parametric Statistical Inference from Bernoulli to Fisher 1713 to 1935*, University of Copenhagen, Department of applied mathematics and statistics, Copenhagen, pp. 7

Фишер је сматрао да је предмет статистичких метода редукција података, која се може постићи ако се подаци посматрају као случајан узорак из претпостављене бесконачне популације, чија је расподела одређена са неколико параметара. Проблем је поделио у три врсте, које је дефинисао на следећи начин:[12]

1. *Проблем спецификације* – овај корак подразумева формирање модела и, као такав, подлеже расправама, слично моделима који су засновани на практичном искуству у сродним ситуацијама. Фишер овај корак приписује логици закључивања. Може се чак рећи да је у питању индуктивно закључивање, јер је реч о преласку са конкретних података на математичке моделе, што овај корак и чини идуктивним.
2. *Проблем оцењивања* – обухвата избор метода обрачуна из статистике узорка, који су креирани за оцену вредности параметра претпостављене популације. Сама формулација проблема оцењивања захтева одговарајући математичко-статистички модел. У овом кораку важну улогу има математичка теорија што се, на пример, односи на увођење појма довољности и извођење математичких предлога који су уско повезани са овим појмом.
3. *Проблем расподеле* - обухвата расправе о расподели статистике изведене из узорка или о било којим функцијама величине чија је расподела позната. У овом кораку ситуација је јасна, јер су апстрактна образложења исплатива у оцењивању само ако на крају доводе до конкретних нумеричких резултата. Најбољи модел нема употребну вредност уколико не резултира у одређеној криви расподеле.

Јасно је да је, када је дефинисан *проблем спецификације*, лакше одредити који су параметри потребни да би се одредила популација из које је узорак извучен; такође, када је познат *проблем оцењивања*, може се одредити како би се, на најбољи могући начин, из узорка могле израчунати оцене ових параметара; и, на крају, уколико је познат *проблем*

[12] Lenhard, J. 2006. „Models and Statistical Inference: The Controversy between Fisher and Neyman–Pearson“, *Oxford University Press on behalf of British Society for the Philosophy of Science*, Vol. 57, pp. 74

расподеле, могуће је одредити тачан облик расподеле у различитим узорцима на основу изведених статистика, па је, самим тим, теоријски аспект било ког дела података у потпуности разјашњен.[13]

Фишеров највећи допринос је у томе што је по први пут обезбеђено мерило оптималности за статистичко оцењивање, опис оптимума који се може постићи у конкретном проблему оцењивања и техника максималне веродостојности која производи оцене за θ које су близу идеалне у смислу тих мерила.

1.1.3 Улога Бајеса у развоју статистичког закључивања

Назив *Бајесова статистика* потиче још из 18. века и везује се за свештеника, влч. Томаса Бајеса (Thomas Bayes).[14] Бајес је преминуо 1761. године, али је његово највеће дело „Есеј о решавању проблема у доктрини о шансама“ (енгл. „*An essay towards solving a problem in the doctrine of chances*”), објављено постхумно у *Филозофским трансакцијама Краљевског друштва* (енгл. *Philosophical Transactions of the Royal Society*) 1763. године.

Теорема презентована у раду, данас је позната као Бајесова теорема. Бајес је показао како се *инверзна вероватноћа* може користити за израчунавање вероватноће претходних догађаја на основу догађаја који су уследили као последица претходних. Његове методе су усвојили Лаплас и други научници у 19. веку, али су драстично изгубиле на значају почетком двадесетог века. Заправо је иронично да је кључна ствар у есеју, Бајесов доказ за спорну претпоставку да је непозната вероватноћа униформне априорне расподеле, погрешно интерпретирана, укључујући Пирсона (Karl Pearson), Фишера, Џефриса (Harold Jeffreys) и Хокинга (Ian Hacking). Бајесов есеј је један од најтеже читљивих радова у историји статистике, па стога не чуди што је често био предмет неразумевања.

---

[13] Fisher, R.A. 1922. „On the Mathematical Foundations of Theoretical Statistics”, *Philosophical Transactions of the Royal Society of London, Series A,* Vol. 222, pp. 314

[14] Новија литература и аутори новије генерације залажу се да се уместо досадашњег изговора „*Бајес*“ и „*Бајесова статистика*“ користи „*Бејс*“ и „*Бејсова статистика*“, због велшког порекла познатог статистичара. Аутор се определио за досадашњи приступ овом питању, те ће у даљем тексту бити коришћени термини „*Бајес*“ и „*Бајесова статистика*“.

Проблем који је разматрао Бајес заправо је, у модерној терминологији, проблем оцењивања параметра θ у биномној расподели, где долази до решења које данас називамо *Бајесово решење*, под претпоставком да θ има униформну априорну густину за (0,1). Ова претпоставка, која се назива и *Бајесов постулат*, у самом раду је врло контроверзна (не сама Бајесова теорема, која је само основна изјава о условној вероватноћи). Неки аутори су задржали мишљење, иако модерни научници то оспоравају, да Бајес због личног незадовољства претпоставком, рад није објавио током живота. Без обзира на (не)истинитост те тврдње, већи део рада је посвећен оправдавању претпоставке за шта је Бајес и дао аргумент. Међутим, Бајесов аргумент је тешко генерализовати у другим ситуацијама, где би се могла применити Бајесова статистика.

У то време, утицај Бајесовог рада био је незнатан и већи део онога што данас називамо *Бајесова статистика* развијен је независно од Бајеса, под утицајем француског математичара Лапласа. Као резултат тога, Лаплас је 1812. године објавио рад „Аналитичка теорија вероватноће“ (франц. *Th´eorie Analytique des Probabilit´es*), иако је највећи део написан током седамдесетих и осамдесетих година 18. века. Лаплас је применио *принцип недовољног разлога* како би оправдао униформну априорну густину: „Не постоји ни један разлог због ког бисмо сматрали да је једна вредност за θ вероватнија у односу на другу, па из тог разлога треба користити униформну априорну расподелу“.[15] Једини недостатак овог аргумента огледа се у томе да, уколико применимо принцип недовољног разлога на θ_2 на пример, априорна вредност ће бити другачија у односу на резултат за θ који бисмо добили применом истог принципа. Аргумент који је применио Бајес био је много адекватнији и заиста је водио до униформне априорне вредности за θ, пре него до трансформације за θ, али само код одређених модела.

У време развоја више модерних теорија о статистичком закључивању, почев са радовима Галтона (Francis Galton) и Пирсона крајем деветнаестог века, Бајесове идеје су изгубиле на значају. Велики допринос томе дао је и Фишер, који се током целокупне своје каријере

[15] Young, G.A. and Smith, R.L. 2005. *Essentials of statistical inference*, Cambridge University Press, Cambridge, pp. 40

оштро противио примени Бајесових метода. У својој књизи „Статистичке методе за истраживаче" (енгл. „*Statistical Methods for Research Workers*"), издатој 1925. године, практично је започео искључење Бајесових метода из статистичке праксе за наредних педесет година. У првом издању, експлицитно је изразио своје лично уверење да теорија инверзне вероватноће, на којој почива Бајесова теорема, има грешку те је, из тог разлога, пожељно њено потпуно одбацивање. Истовремено каже: „Без обзира што морамо одбацити ову поставку, неопходно је и указати на Бајесову величину у разумевању проблема који би требало решити, приказивањем његовог могућег експерименталног решења и, коначно, много јаснијим, у односу на касније ауторе, приказом слабости аксиоматског метода".[16]

Тек средином прошлог столећа, поново се јавило интересовање за Бајесове методе, пре свега од стране научника као што су де Финети (Bruno de Finetti), Џефрис, Саваж (Leonard Jimmie Savage) и Линдли (Dennis Victor Lindley), али и многих других. Њихов главни допринос се огледа у томе што су развили комплетан метод статистичког закључивања заснован на Бајесовој теореми.

У том периоду, ситуација се окренула у корист Бајесове статистике, када је Џефрис, професор са Кембриџа, 1939. године објавио књигу „Теорија вероватноће" (енгл. *Theory of Probability*). Џефрис је познат по великом доприносу примењеној математици, геофизици и астрономији, али и по посвећености научном закључивању. Упркос називу, књига „Теорија вероватноће" заправо је расправа о Бајесовим методама. Следећи традицију Лапласа, Џефрис је веровао да априорна расподела треба да буде неинформативна, колико год је то могуће па је, да би то достигао, предложио општу формулу која је данас позната као *Џефрисова априорна формула*. Међутим, његови аргументи нису убедили скептике. Фишер је, на пример, у прегледу своје књиге навео да је грешку уочио већ на првој страни где се примењује Бајесова формулација, па је тиме одбацио целокупан Џефрисов рад.

[16] Fisher, R.A. 1970. *Statistical methods for Research Workers, Fourteenth Edition*, Oliver & Boyd, Edinburgh, pp. 21

Једна од одлика аргумената које су дали Лаплас и Џефрис је да често за резултат имамо оно што се назива неправилна априорна вероватноћа. С друге стране, већина модерних следбеника Бајесове статистике нема проблем са неправилним априорним расподелама, иако код сложених проблема реално постоји опасност да ће неправилна априорна густина довести до неправилне апостериорне густине, што би свакако требало избећи.

Док је Џефрис развијао своју теорију, 1933. године Нојман (Jerzy Neyman) и Егон Пирсон (Egon Pearson), објавили су теорију о тестирању хипотеза, у којој су избегли било какво позивање на Бајесове идеје. Фишер, такође, није подржао Нојманов приступ, али су се сложили око тога да Бајесове идеје нису добре.

Отприлике у исто време, Бруно де Финети у Италији и Леонард Саваж у Сједињеним америчким државама, развили су алтернативан приступ Бајесовој статистици заснован на субјективној вероватноћи. У Великој Британији, водећи заступник овог приступа био је Линдли. Према де Финетију, Саважу и Линдлију, само логички конзистентна теорија вероватноће, а самим тим и статистике, базирана је на личном уверењу, где је понашање сваког појединца усмерено ка максимизацији очекиване корисности у складу са личним схватањем вероватноће остварења различитих исхода. На тај начин су одбацили не само целу класичну (не-Бајесову статистику), већ и приступ о неинформативној априорној вероватноћи Лапласа и Џефрија. Веровали су да је једино субјективан избор, прави начин избора априорне расподеле, при чему су занемарили чињеницу да ће то различите статистичаре довести до различитих закључака из истог скупа података. Контроверзан део де Финетијеве и Саважове теорије је тврдња да би сви пробабилистички и статистички искази требали да се заснивају на субјективној вероватноћи.

Из перспективе данашње статистике, Бајесове и не-Бајесове методе успешно коегзистирају у највећем броју случајева. Неки модерни теоретичари се снажно залажу за Бајесов приступ, али је највећи број савремених интересовања за Бајесове методе ипак проистекао из знатно прагматичнијих схватања: код веома сложених модела који обухватају хиљаде опсервација и стотине параметара, Бајесове методе се могу ефикасно рачунарски применити.

1.2 Различити приступи статистичком закључивању

Било који статистички поступак који користи информације да би се добио опис практичних ситуација (преко модела вероватноће) је поступак закључивања, а проучавање таквих поступака назива се *статистички закључак* или *статистичка инференција*.[17]

Поступак познат под називом статистичка теорија одлучивања бави се доношењем одлука на основу статистичког знања које се распростире на неизвесност обухваћену проблемом одлучивања. Ова неизвесност се, у највећем броју случајева, посматра као нумеричка величина и означава са θ.

Проблем одлучивања подразумева избор између неколико могућих праваца деловања, што повлачи за собом одговарајуће последице које се могу искористити за тестирање исправности примењених процедура. С друге стране, закључивање се ослања на степен уверења, што може, али и не мора имати одговарајуће консеквенце. Тиме се, наравно, отежава постизање компромиса између питања закључивања и одлучивања.[18]

Процес одлучивања даље развија циљеве закључивања, јер се дескриптивна функција закључивања проширује предложеним правилима понашања. Из тог разлога се може очекивати да ће процес доношења одлука много више варирати у односу на расположиве информације, које треба обухватити и обрадити, него процес закључивања. Уопштено говорећи, закључивање користи податке из узорка и остале информације које описују конкретну ситуацију, на пример, априорне информације, док процес одлучивања увећава инференцијално знање о посматраној ситуацији кроз укључивање процене последица.

[17] Barnett, V. 1999. *Comparative Statistical Inference, Third Edition*, John Wiley & Sons, Ltd, Chichester, West Sussex, England, pp. 13

[18] Smith, C.A.B. 1965. „Personal Probability and Statistical Analysis", *Journal of the Royal Statistical Society, Series A*, Vol. 128, No. 4, pp. 472

На основу дефинисане разлике између закључивања и доношења одлука, могуће је извршити и основну поделу различитих приступа статистици, с тим што ће прва два бити и предмет даљег проучавања.

Први приступ, ***класична статистика***, потиче из радова Фишера, Нојмана, Пирсона и других и обухвата технике *тачкастог* и *интервалног оцењивања*, *тестове значајности* и *тестирање хипотеза*.

Сам назив *класична статистика* проистекао је из традиционалне улоге овог статистичког метода који се много раније појавио и чија је примена много шира у односу на примену Бајесовог закључивања или теорије одлучивања. Међутим, у литератури се могу наћи и другачији називи који одговарају термину класичне статистике као, на пример, теорија узорковања, фреквентистичка или стандардна статитистика и слично.

У том смислу, класична статистика је усмерена на „фреквентистичко" схватање вероватноће, где се подаци из узорка представљају преко њихове вероватноће, а одређени критеријуми постављају на основу расподеле узорка како би се оцениле перформансе примењених техника. На пример, може се очекивати да тачкасте оцене буду непристрасне или конзистентне, тестирање хипотеза подразумева одређивање области (не)одбацивања нулте хипотезе, док се на општем нивоу, тамо где је то могуће, подаци приказују као *довољна статистика*. Самим тим, методе одражавају агрегатне мере сопствене поузданости или тачности преко одређивања стандардних грешака или ефикасности.

Код класичног приступа, тежи се *Фишеровој редукцији*, која подразумева неколико корака:[19]

- дефинисање функције веродостојности,
- редукцију на довољну статистику,
- дефинисање функције за довољну статистику која има расподелу која зависи само од посматраног параметра,

[19] Cox, D.R. 2006. *Principles of statistical inference*, Cambridge University Press, pp. 24

- обртање расподеле како би се дефинисала ограничења за параметар на произвољном скупу нивоа вероватноће,
- формална или неформална примена условне расподеле података како би се проценила адекватност примењене формулације.

Као што је већ споменуто, класична статистика се заснива на директној примени података из узорка (подаци су резултат статистичког истраживања) у доношењу закључака о θ. Међутим, подаци из узорка нису једини вид информација који је релевантан за статистичко истраживање. Други приступи статистици су и дефинисани тако да узимају у обзир и друге врсте информација, тако да сваки покушај испитивања различитих приступа субјекту посматрања мора почивати на обухватању ширег појма информација, а не само података из узорка.

Као додатак узорачким информацијама, обично су на располагању још две врсте информација:

(1) знање о могућим последицама одлуке - ово знање се веома често може квантитативно изразити одређивањем губитка који би могао настати за сваку појединачну одлуку и за различите могуће вредности θ;
(2) априорне информације - то су информације о θ које произилазе из других извора, а не из статистичког истраживања; априорне информације обично проистичу из претходног искуства у сличним ситуацијама које су обухватале сличан параметар θ.

Други статистички приступ, ***Бајесов приступ***, тежи примени априорних информација. Априорне информације су модификоване са подацима из узорка преко поновљене примене Бајесове теореме како би се проценило сазнање о конкретној ситуацији. Вероватноћа поново има кључну улогу. Изведени закључци се приказују преко *апостериорне расподеле* вероватноће, која је сада мерило прецизности. Питање довољности је поново актуелно, али сада не као метод узорковања који би се користио за добијање података из узорка. Овај приступ се не ослања само на „фреквентистичко“

схватање појма вероватноће; субјективна интерпретација је готово неизбежна, а вероватноће се посматрају као условне вероватноће.

Кључну улогу у развоју Бајесових метода имао је појам *кохеренције*. Реч је о појму којим се прецизно жели изразити оно што се захтева од појединца у ситуацијама у којима би требао да реагује рационално или доследно, у условима неизвесности. Из тог разлога, оно што данас зовемо Бајесова статистика, можемо назвати и кохерентна статистика, јер је кохеренција (повезаност) суштина онога чиме се Бајесијанце баве – кохеренција постигнута кроз примену рачуна вероватноће.

Трећи приступ, ***теорија одлучивања***, проистиче из радова Валда (Abraham Wald). У складу са именом, приступ је посебно креиран како би обезбедио правила деловања, односно правила одлучивања у условима неизвесности. Приступ, као такав, обухвата процену последица алтернативних акција изражену преко математичке теорије корисности у облику губитка или функције губитка. Вредност било ког правила одлучивања за предвиђену акцију на основу података из узорка (и било које априорне информације) мери се преко очекиваног губитка или ризика. Основни циљ је одабир правила одлучивања са минималним ризиком, док појам прихватљивости правила одлучивања, као и целокупне групе правила одлучивања заузима централно место у проучавању оптималности. Минимакс принцип је један од основних за избор оптималног правила одлучивања.

Код теорије одлучивања, ниједно посебно схватање појма вероватноће није нужно укључено. Закључци о вероватноћи су обично засновани на „фреквентистичком“ схватању, мада укључивање априорних информација подразумева и прихватање субјективног схватања.

Бајесова анализа и теорија одлучивања природно иду заједно, делом због њиховог заједничког циља да користе неекспериментане изворе информација, а делом и због неких дубљих теоријских веза. Док теорија одлучивања не зависи или не захтева примену Бајесових метода или субјективну интерпретацију вероватноће, проучавање оптималних правила одлучивања је поједностављено кроз Бајесов приступ. Усвајање априорних расподела даје нам могућност да у многим ситуацијама ограничимо обим прихватљивих правила одлучивања, без обзира на то да ли било која априорна расподела одговара или не одговара неком реалном проблему.

Основне карактеристике наведена три приступа могу се сажето приказати и кроз одговарајући табеларни приказ (Табела 1.), док ће предмет даљег разматрања биће искључиво класична и Бајесова статистика и њихове карактеристике.

Табела 1. Основне карактеристике три приступа

Приступ	Функције	Појам вероватноће	Релевантне информације
Класични приступ	Функција закључивања (претежно)	„фреквентистичко“ схватање	Подаци из узорка
Бајесов приступ	Функција закључивања	Субјективни „степен уверења“; могуће фреквенције- компоненте које треба протумачити	Подаци из узорка Априорне информације
Теорија одлучивања	Функција одлучивања	„фреквентистичко“ схватање (Субјективни „степен уверења“, када су укључене и априорне информације)	Подаци из узорка Последични губици или користи (Априорне информације)

Извор: Barnett, V. 1999. *Comparative Statistical Inference, Third Edition*, John Wiley & Sons, Ltd, Chichester, West Sussex, England, pp. 18

Оваква подела приступа, са једне стране, представља корисну основу за дефинисање појединих појмова, критеријума и техника док, са друге стране, једноставност поделе може водити у неосновану сигурност. На тај начин се лако може створити уверење о постојању строгих, добро дефинисаних разлика између приступа у погледу њихових функција и захтева по питању релевантних информација уз, такође, јединствен став о значају интерних критеријума у оквиру сваког појединачног приступа. Тиме се отвара питање исправности извршене поделе.

1.3 Критике и контроверзе које су обележиле историјски развој

Расправа која је присутна између Бајесове и класичне статистике више од два века, резултат је два различита приступа научном и статистичком начину размишљања. Сматра се да је Бајесова статистика пре свега погодна за индивидуалне истраживаче или групе истраживача који теже да на основу свих расположивих информација изведу што боље закључке. Према једној дефиницији чак се тврди да је класична статистика заправо Бајесијански покушај да се све уради добро или довољно добро у односу на било коју априорну расподелу.[20]

У пракси, границе између приступа нису довољно јасне. Може се десити да не постоји ни једно очигледно средство које би указало да ли је посматрана функција, функција одлучивања или функција закључивања, нити да ли одређено тумачење вероватноће имплицитно одговара или не одговара било ком методу или приступу. Такође, не може се тврдити ни да постоји било која објективна основа за процену вредности интерних критеријума као што је, на пример, непристрасност тачкастих оцена, минимакс принцип за избор оптималног правила одлучивања или примена коњуговане априорне расподеле.

Многи аспекти изучавања статистичког закључивања и теорије одлучивања морају укључити и личну процену појединаца који унапређују или развијају различите приступе. То заправо значи да критеријуми треба да буду сразмерно произвољни у зависности од личних ставова.

Већ је споменуто да се наведена класификација приступа може окарактерисати као сувише једноставна, али би се то могло протумачити и као став класичне статистике. Наиме, овај приступ би се могао схватити и као почетни облик на ком су изграђени делови других приступа који укључују средства за обраду априорних информација и/или корисности, као и податке из узорка. Неуспех класичних метода у прилагођавању априорним информацијама или консеквенцама у било ком смислу, није проистекао из

[20] Cox, D.R. 2006. „Frequentist and Bayesian Statistics: A critique", *Statistical problems in particle physics, astrophysics and cosmology*: 1–4. *Imperial College Press*, London, pp. 3

њихове немогућности да препознају ове алтернативне облике информација. Реч је, пре свега, о свесном избору оснивача класичног приступа. Њихов основни аргумент је да такве информације ретко могу бити објективне или довољно детаљне за изградњу званичне теорије статистичког закључивања. Уместо тога, циљ им је био да развију теорију која ће имати универзалну примену, која ће бити ослобођена субјективних процена и заснована на информацијама које ће увек егзистирати у мерљивом облику (подаци из узорка). Ово схватање, заједно са зависношћу од фреквентистичког схватања вероватноће, изазвало је највеће критике класичне статистике.

Као што се може доказати да чак ни функција поступка није одмах очигледна, може се приказати и природа наведене критике преко примера тестирања хипотеза. Са једне стране, може се рећи да је реч о инференцијалној процедури, пошто је реч о исказу о вредности неког параметра и оцени вероватноће која је повезана са подесношћу самог исказа. Међутим, код тестирања је готово неизбежно предузети неке од алтернативних акција у појединим ситуацијама јер исход тестирања стимулише одређену акцију. У том смислу, реч је о поступку одлучивања.

Такође, тврдња о објективности класичног приступа која је заснована на фреквентистичком схватању вероватноће није широко прихваћена. У многим критикама се тврди да је тумачење резултата тестирања хипотеза или информација добијених на основу интервала поверења заправо засновано на субјективном схватању вероватноће па је, у том смислу, тумачење класичне статистике нетачно. Из практичнијег угла гледано, замерке се упућују и због произвољног избора нивоа значајности теста и због принципа извођења закључака на основу двосмерних тестова.

Бајесов приступ такође нису заобишле критике, пре свега због неопипљивости априорних информација. Замерка је упућена због тога што приступ захтева примену априорних информација чак и када су нејасне, непрецизно дефинисане, субјективне или чак ни не постоје. У том смислу, Фишер је одиграо значајну улогу, јер је одбацио читав концепт инверзне вероватноће на којем почива Бајесово закључивање.

Критички осврти нису усмерени једино на утврђивање јаснијих граница између појединих приступа. Расправе се воде и око тумачења појмова и релевантности критеријума у оквиру посматраних приступа. Код Бајесовог закључивања, на пример, најчешће дискусије су биле на тему описа априорног незнања и тумачења појма вероватноће.

Теорија одлучивања је, такође, подлегла критикама. Пре свега је доведен у питање избор појма корисности, као одговарајућег појма за изражавање последица. Постојање критика и контроверзи, као и у већини других области, и у области статистичког закључивања и теорије одлучивања је неминовно. У основи већине расправа налази се питање зависности од личног мишљења, ставова и процена јер, шта је за једну особу произвољно, за другу може бити очигледно.

Другачије посматрано, постојање контроверзи је и пожељно, јер оне доводе до развоја конструктивног и критичког начина размишљења о предмету посматрања уз сталну поновну процену. Конструктивни критицизам је основа даљег развоја. То се свакако односи и на статистичко закључивање и на теорију одлучивања јер смо сведоци изузетног развоја који се догодио у току прошлог века у условима сталних расправа и преиспитивања.

Наравно, немогуће је и очекивати неки општи, заједнички став око тога шта чини један прави и исправан приступ статистичком закључивању. Склоности ка једном или другом приступу увек морају укључити и субјективан начин размишљања. Док једни заузимају чврст став како би доказали неизбежност једног приступа или његову супериорност у односу на други, други настоје да помире приступе и да их прикажу као различите (или двоструке) аспекте заједничког циља.

Уопштено говорећи, Бајесова статистика је доминирала статистичком праксом током деветнаестог века, док је 20. век обележила класична статистика. Поставља се питање шта се може очекивати у 21. веку? Чињеница је да у данашње време и научници и статистичари имају пред собом много веће скупове података које треба обрадити, са милионима појединачних података и хиљадама параметара које треба одједном размотрити. Класична статистика је више креирана за проблеме малог обима, за највише неколико стотина података и неколико параметара. Бајесово правило је данас веома популаран начин размишљања, с обзиром да је Бајесова статистика направила снажан помак у последњих двадесет година. Технички развој и компјутерска примена Бајесових метода подржали су овај тренд, а као коначан резултат се очекује приближавање Бајесове и класичне статистике.

Различити начини размишљања о важности појединих приступа, још више истичу потребу за даљим испитивањем карактеристика, пре свега, Бајесове и класичне статистике, како би се на крају извели закључци о оправданости или неоправаданости предности у примени Бајесовог приступа статистичком закључивању у последњих неколико деценија.

2. ТРАДИЦИОНАЛНИ ПРИСТУП СТАТИСТИЧКОМ ЗАКЉУЧИВАЊУ

Статистика, као организована и званична научна дисциплина има релативно кратку историју у поређењу са традиционалним наукама. Двадесети век је обележен најинтензивнијим развојем статистике, где је већ после врло кратког времена створено значајно наслеђе и у теоријском и у практичном смислу. Време у ком се стварала традиција, истовремено је послужило и за развој супротних идеја и ставова. У том процесу, неоспорно је да су начела и методе *класичне инференције* имала традиционалну улогу; главни ривал, нарочито током последњих педесет година, је Бајесово закључивање. Овај приступ, као и мноштво различитих индивидуалних ставова, већ дуже време представљају изазов за савремену теорију и праксу.

2.1 Класична статистичка инференција

Током двадесетог века развила се нова статистичка техника која је била довољно моћна да би се могла применити и у другим научним дисциплинама. Ова *класична* статистичка техника није имала очигледно субјективно тумачење па је, последично, постала доминантна филозофија о вероватноћи међу математичким статистичарима и научницима уопште. Енглески статистичар и генетичар, Фишер, дао је посебан допринос у развоју класичне статистике; оцена максималне веродостојности, тестови значајности, анализа варијансе и планирање експеримената и даље, добрим делом, доминирају пољем примењене статистике.

Заједно са Фишером, Нојман и Пирсон дефинисали су класичну статистичку парадигму: *фреквентизам*. Применили су концепт релативне фреквенције који подразумева понављање експеримента како би се измерила пропорција појаве пожељних резултата. Ова пропорција, под условом да се експеримент понови довољан број пута, назива се *вероватноћа*. Ипак, док су Нојман и Пирсон свој рад презентовали као једну од могућих алтернатива, дотле је Фишер тврдио да је његова статистика званично решење проблема индуктивног закључивања.[21]

Иако се кључни развој класичне статистике одиграо у 20. веку, свакако је могуће открити развојни пут различитих и значајних класичних ставова још од самог почетка 19. века. Важни алати вероватноће, као што су Закон великих бројева, Централна гранична теорема, али и кључни појам расподеле узорка и принцип најмањих квадрата, допринели су њеном развоју. Такође, још у радовима Лапласа, Гауса и других, присутни су методи који одговарају класичној статистици.

Класична статистика се развила брзо, као резултат настојања да се пронађу одговори на значајна питања и развије поуздана основа за статистичка испитивања и анализу података. Већина активности је настала крајем 19. века и настављена је све до данашњих дана. Највећи део основне теорије оцењивања и тестирања хипотеза настао је у периоду од 1920-1935. године, кроз два посебна приступа. Са једне стране, Нојман и Пирсон су се највише фокусирали на израду принципа за тестирање хипотеза и предложили су метод за интервал или област поверења, што је проистекло из њихове идеје о тестовима значајности. Као резултат тога, дефинисани су интервали, односно области поверења. Са друге стране, Фишер је више био заинтересован за оцењивање него за тестирање хипотеза, па је изградио интервал или област оцењивања, али са другачије тачке гледишта.

У том периоду, већи део основа класичног закључивања, захваљујући овим различитим приступима, погрешно је тумачен, иако се њихов рад није у потпуности разликовао, ни по

[21] Vallverdú, J. 2008. „The False Dilemma: Bayesian vs. Frequentist“, *E-Logos, Electronic Journal for Philosophy*, ISSN 1211-0442, pp. 3.

питању тумачења, ни по питању примене. Пример којим се то потврђује односи се на разлику између тестирања хипотеза и тестирања значајности, која није била поштеђена интерне контроверзе: Фишер је био нарочито критичан по питању Нојманових и Пирсонових схватања различитих проблема, посебно интервалног оцењивања. Литература из тридесетих година прошлог века на најбољи начин осликава овај дијалог чији утицај није попустио све до данашњих дана. Фишеров концепт *фидуцијалне вероватноће* кључни је елеменат у расправи.

Фишер је предложио метод заснован на идеји фидуцијалне вероватноће, што је водило до тзв. фидуцијалних интервала. Фидуцијални аргумент представља неку врсту енигме у класичној статистици, јер је у супротности са основним принципом предлагања расподеле вероватноће као инференцијалног исказа о параметру θ. Фишер је понекад користио термин *апостериорна фидуцијална расподела* за θ, али је одлучно порицао да је слична или готово иста као Бајесова апостериорна расподела.

Илустративни пример који даје у свом раду „Инверзна вероватноћа" (енгл. „*Inverse probability*") из 1930. године, говори о томе да се пример може дати пре него што се узме у обзир на који начин се његов логични садржај разликује од одговарајуће изјаве о вероватноћи која се добија из познатих априорних вероватноћа. У многим околностима, случајне расподеле узорачке статистике T могу се директно израчунати из посматрања где ће бити представљене преко једног параметра, док ће статистика T бити оцена до које се дошло применом метода максималне веродостојности. Уколико је T статистика континуираних варијација, а P вероватноћа да ће вредност статистике T бити мања од било које наведене вредности, онда говоримо о односу у облику:

$$P = F(T,\theta) \tag{1}$$

Ако се за P сада одреди нека вредност као што је, на пример, 0.95, добија се однос између статистике T и параметра θ, тако да је T 95% вредност која одговара датом θ. Такав однос подразумева објективну чињеницу да ће у 5% узорака T премашити 95%, што је вредност која одговара стварној вредности параметра θ у популацији из које је извучен.

Штавише, било која вредност за T обично ће бити посебна вредност за θ. Стога, ову вредност можемо и назвати „фидуцијалних 5%, вредност за θ" што одговара датој статистици T. У случају да T расте са θ за све могуће вредности, однос се може изразити преко изјаве да ће стварна вредност за θ бити мања од фидуцијалних 5%, што је вредност која одговара посматраној вредности за T у тачно 5 од 100 испитивања.

Израдом табеле са одговарајућим вредностима, након израчунавања T, може се сазнати шта је фидуцијалних 5%, вредност за θ, као и да ће права вредност за θ бити мања од ове вредности у само 5% испитивања. То је коначна изјава о вероватноћи непознатог параметра θ, која је тачна без обзира на било које претпоставке у вези са њеном априорном расподелом.

Табела 2. Фидуцијална вероватноћа

Фидуцијалних 5% ρ	95% r	Фидуцијалних 5% ρ	95% r	Фидуцијалних 5% ρ	95% r
-0.995055	-0.968551	-0.761594	+0.145340	+0.761594	+0.989816
-0.993963	-0.961623	**-0.716298**	+0.270475	**+0.800499**	**+0.991770**
-0.992632	-0.953179	-0.664037	+0.388574	+0.833655	+0.993335
-0.991007	-0.942894	-0.604368	+0.496089	+0.861723	+0.994593
-0.989027	-0.930375	-0.537050	+0.590725	+0.885352	+0.995608
…	…	…	…	…	…

Извор: Fisher, R.A. 1930. „Inverse probability", *Mathematical Proceedings of the Cambridge Philosophical Society*, Volume 26, Issue 04, pp. 533

У Табели 2. са r је представљена корелација која потиче из само једног пара посматрања, а са ρ корелација у популацији из које је узорак извучен; однос између ρ и 95%, вредност за r дата је у табели.

Из табеле можемо очитати са 95%, r за било које дато ρ, односно фидуцијалних 5%, ρ за свако дато r. Дакле, ако је на пример вредност r = 0.991770 добијена из узорка, фидуцијалних 5% ρ биће 0.800499. Вредност ρ може да буде само мања од -0.716298 у случају да r премаши тачку од 95%, што се дешава само једном у двадесет покушаја. У том смислу, ρ има вероватноћу да само једном у 20 случајева буде мање од -0.716298. На исти начин може се наћи било који друга вредност у фидуцијалној расподели за ρ или, уопште, фидуцијална расподела параметра θ за дату статистику T може се изразити на следећи начин:

$$df = -\frac{\partial}{\partial \theta} F(T, \theta) d\theta, \tag{2}$$

док је расподела статистике за дату вредност параметра:

$$df = \frac{\partial}{\partial \theta} F(T, \theta) dT. \tag{3}$$

Фишер је сматрао да је фидуцијална вероватноћа општија и кориснија за практичну примену, јер ће у пракси сви узорци дати различите вредности, а самим тим и различите фидуцијалне расподеле и различите расподеле инверзне вероватноће. Међутим, узимајући у обзир да се очекује да ће фидуцијалне вредности бити различите у сваком случају, при чему су закључци о вероватноћи у одређеном односу са таквим варијабилитетом, извештаји о инверзној вероватноћи заиста имају другачије значење за сваки појединачан узорак, осим ако се посматрана статистика заправо не дешава на исти начин.[22]

Нешто касније, преиспитивањем појма фидуцијалне вероватноће, Бајесове и фидуцијалне методе су прихваћене у одређеним ситуацијама као дуални методи закључивања, где ће се прве користити у случају када постоје априорне информације, а друге у случају када априорних информација нема. Појам вероватноће је био потпуно идентичан класичном схватању вероватноће из радова најранијих аутора. Међутим, фидуцијална расподела вероватноће[23] никада није јасно дефинисана, нити од стране Фишера, нити од стране

[22] Fisher, R.A. 1930. „Inverse probability“, *Mathematical Proceedings of the Cambridge Philosophical Society*, Volume 26, Issue 04, pp. 535

[23] Претпоставимо да је $\tilde{\theta}(\mathrm{X})$, као оцена максималне веродостојности за θ, непрекидна променљива и да има функцију расподеле $F_\theta\left[\tilde{\theta}(x)\right]$. Према Фишеру, $F_\theta(\tilde{\theta})$ може се користити у одређеним околностима како би се извели закључци о вероватноћи параметра θ који ће зависити од стварне вредности, $\theta(x)$, добијене из

каснијих заговорника овог принципа. Познато је да је Фишер ограничио њену употребу на обраду узорка са непрекидном расподелом, што је ограничавајуће из статистичког угла, али је зато математички посматрано погодно за одређене методе предложене за израду фидуцијалне расподеле.

У савременим околностима велики део практичне статистике користи класичан приступ који је, готово у целини, заснован на појмовима, критеријумима и методама који су први пут презентовани у доба Нојмана, Пирсона и Фишера. Основна идеја се током година даље развила и постала широко примењива, па се таква тенденција може очекивати и у годинама које долазе.

2.1.1 Значај вероватноће у одређивању карактеристика класичне инференције

Током више од једног века међу статистичарима се води расправа о значењу вероватноће. Практично, сагласност око аксиома вероватноће постоји, али се суштински основно питање односи на улогу вероватноће у сфери закључивања.[24] Самим тим, дефинисање карактеристика класичне инференције уско је повезано са дефинисањем основног појма вероватноће.

Појам вероватноће, по правилу, везујемо за реализацију експеримента или опсервације; међутим, у литератури су присутна и другачија схватања појма вероватноће. Док су нека тумачења врло кратка и усмерена на детаљније разматрање неког одређеног става, дотле друга усвајају шире погледе који се односе на упоредне анализе различитих теорија и ставова. Једна од основнијих подела вероватноће је, према Барнету (1999), подела на класично, фреквентистичко, логичко и субјективно схватање:

података x. На основу тога се добија изведена расподела вероватноће, односно **фидуцијална расподела.** (Barnett, 1999, pp. 124)

[24] La Valle, S.M. 2006. *Planning Algorithms,* Cambridge University Press, pp. 484

Табела 3. Различита схватања вероватноће

Схватања вероватноће	Основне карактеристике
Класично	Разматрање симетрије; једнако вероватни исходи
Фреквентистичко	Емпиријски карактер; релативне фреквенције у околностима које се понављају
Логичко	Објективан карактер; унутрашње логичко мерило импликација
Субјективно (лично)	Лични „ниво уверења“; индивидуална процена рационалног или кохерентног понашања

Основна идеја *класичног схватања вероватноће* односи се на могућност препознавања ситуација код којих постоји коначан број једнако могућих исхода који обухватају све могућности. У таквим ситуацијама, вероватноћа појединачног догађаја *А* се описује као однос између броја исхода који подржавају или имплицирају догађај *А* и укупног броја могућих исхода. Овакво схватање појма вероватноће побудило је интересовање код многих математичара, али је било очигледно да оно носи са собом и извесне проблеме. Наиме, ограничавање само на ситуације из којих простичу једнако вероватни исходи, значајно редукује примену концепта вероватноће. Термин „једнако вероватни“ је сам по себи споран, као и препознавање једнако вероватних исхода, па класично схватање вероватноће има примену у веома малом броју случајева, као што су случајеви бацања новчића и коцке, праведна подела карата и слично. Покушаји да се прошири опсег примене изгледају прилично неубедљиво, јер се веома често фреквентистичко схватање вероватноће имплицитно уводи у аргумент.

Фреквентистичко схватање појма вероватноће је много чешће било предмет расправа, али и примена у односу на друга схватања. Реч је о концепту који представља основу и оквир класичне статистичке методологије. Из тог разлога, о фреквентистичком схватању вероватноће се говори као о основном ставу о вероватноћи у практичној примени статистике у различитим дисциплинама. Основи елеменат овог схватања је да концепт вероватноће треба да буде објективан и, при томе, ослобођен утицаја било каквих личних фактора уз истовремену могућност практичне примене у експериментима. Тиме се прави кључна разлика између фреквентистичког и осталих схватања вероватноће, нарочито субјективног схватања вероватноће, пошто се сад говори о практичним ситуацијама у којима разумно треба дефинисати појам вероватноће. Под разумним се пре свега мисли на могућност примене овог концепта у околностима које су ослобођене утицаја неопипљивих, лоше дефинисаних или немерљивих фактора карактеристичних за субјективно схватање вероватноће, уз истовремено могућност емпиријске провере. Према овом схватању, вероватноћа се може дефинисати и применити у ситуацијама које се могу изнова и изнова понављати под идентичним условима.

Фреквентистичко схватање појма вероватноће почива на двема карактеристикама исхода проистеклих из поновљених реализација, а које могу бити предмет посматрања. Као прво, чињеница је да исходи варијају од једног до другог понављања на непредвидив начин; тиме се објашњава појам *случајне варијације*. Као друго, примећено је да се, дугорочно посматрано, појављује извесна правилност. Та правилност се објашњава на примеру појединачног догађаја A, који је предмет посматрања. Као поновљене опсервације узимају се случајеви у којима се догодио догађај A. Тада се однос између броја наступања догађаја A, n_A, и укупног броја понављања, n (релативна фреквенција догађаја А, n_A/n) све више приближава некој граничној вредности како $n\to\infty$. Уколико са $P(A)$ означимо вероватноћу догађаја A, онда ће нам она представљати границу релативне фреквенције n_A/n.

Логичко схватање вероватноће у супротности је са класичним и фреквентистичким схватањем. Развијено је у циљу сагледавања стварног света, односно да би послужило као модел за практичне проблеме, од игара картама до медицинских истраживања. Практични мотив простиче из суштинске природе овог схватања и повезан је са потребом да буде примењен. У том смислу, вероватноћа се посматра као средство којим се квантификује неизвесност уочена у практичним појавама. Код логичког схватања, вероватноћа се посматра као концепт који мења и проширује спектар примене формалне логике. Појам вероватноће се уводи како би се изразио степен имплицираности догађаја B који је омогућен реализацијом догађаја A. Вероватноћа је увек условна и представља рационални степен уверења у догађај B.

Субјективно схватање вероватноће обезбеђује оквир за обраду, најчешће нумерички изражених, нивоа уверења које појединци испољавају у различитим ситуацијама и то на основу релевантних индивидуалних искустава. Ово схватање је уско повезано са акумулираним личним искуством које појединац примењује у конкретним ситуацијама. У свим околностима, вероватноћа мери појединачни степен уверења у реализацију одређене могућности. За разлику од логичког схватања, субјективно схватање вероватноће поставља рационалне нивое уверења, као јединствену неизбежну меру подршке коју пружају преовлађујуће околности неког исхода који је предмет интересовања, независно од посматрача.

На основу свега наведеног, могуће је извести и две основне карактеристике класичног приступа статистици. Као прво, овај приступ прихвата, као једини мерљив облик релевантих информација, *податке из узорка* и у складу с тим, као основу за процену и спровођење статистичких процедура, усваја дугорочно понашање под претпостављеним сличним околностима. Друга карактеристика проистиче из прве; наиме, из наведених разлога фреквентистичко схватање вероватноће је једино прихватљиво схватање у оквиру ког се дефинише, тумачи и мери успех примењених процедура класичне статистике. Самим тим се потврђује доминатна примена овог схватања појма вероватноће у односу на класично, логичко и субјективно схватање.

2.1.2 Традиционална улога класичног закључивања у статистици

Примена класичног и Бајесовог метода статистичког закључивања полази од исте почетне тачке – одређеног статистичког модела. У сваком појединачном случају, основни проблем се односи на одређивање параметра или параметара модела на основу добијених информација. Ове информације су традиционално развијене у најпогоднијем облику, док се параметри статистичког модела не могу директно посматрати. Уместо тога, закључци о параметрима модела морају да се изведу на основу тога колико се пута појавио неуспех; односно, параметри морају бити највише у складу са скупом посматраних неуспеха насталих током времена.[25]

[25] Smith, R.L. 1982. „The Bayesian Inference Method and Its Application to Reliability Problems“, *NRL Memorandum Report 4903*, Texas Research Institute, pp. 4

Када се говори о традиционалној улози класичне инференције у статистици, пре свега се мисли на раније порекло и ширу примењивост овог статистичког метода у односу на Бајесову статистику или теорију одлучивања.

Развој и примена концепта вероватноће може се сагледати још од самих почетака, у радовима из 16. века, кроз примену у играма на срећу, преко проблема бацања коцке развијеног средином 17. века, па све до јаснијег прецизирања и дефинисања током наредних 150-200 година. У том периоду, нагласак је био на класичном схватању вероватноће, заснованом на идеји једнако вероватних исхода. Фреквентистичко схватање, иако у том периоду у почетној фази развоја, није на прави начин дефинисано све до средине 19. века. Алтернативни приступи концепту вероватноће, логичко и субјективно схватање, свој пуни израз и примену доживели су током 20. века, нарочито у другој половини.

Са друге стране, у литератури се ретко могао наћи добро организован приступ општој теорији статистичког закључивања. Међутим, ситуација се мења у периоду између 1920-1935 год. развојем *класичног закључивања*. На прелазу у нови век, па све до 1920. године, интересовање је било усмерено на примену вероватноће и статистике на проблеме из области биологије, али и на растућу потребу за једним организованим моделом у агоркултури и индустријској сфери. Под утицајем Еџворта (Francis Edgeworth), Галтона, Карла Пирсона и других, наступио је период детаљног изучавања принципа статистичке анализе и њихове логичке основе. Ово је уродило плодом у радовима Фишера, Нојмана и Егона Пирсона, који су у великој мери одговорни за развој организоване теорије оцењивања и тестирања хипотеза, заснованим на идејама веродостојности и довољности. Реч је о периоду интензивног развоја појмова и алата, који је настављен све до данашњих дана, на чијим основама је и створена модерна методологија статистичког закључивања утемељена у радовима Фишера, Нојмана и Пирсона.

Интересовање за друге приступе статистичком закључивању није заживело све до тридесетих година прошлог века, чак и тада у светлу супротстављених приступа класичној инференцији.

Фишер је 1922. године објавио есеј „Теоријска статистика на математичким основама " (енгл. „*On the Mathematical Foundations of Theoretical Statistics*“) који се може сматрати једним од темеља математички оријентисане статистике. У есеју је презентовано мноштво нових концепата, а један од њих је ниво значајности који је Фишер увео у циљу успостављања логике статистичког закључивања. Изградња бесконачних хипотетичких популација игра централну улога у овој логици. Прецизније речено, за Фишера спецификација бесконачне популације представља суштински корак у успостављању (параметарског) модела и, истовремено, увођење математичког модела у статистичко закључивање.[26]

Деценију касније, Нојман и Пирсон развили су оно што је данас познато као Нојман-Пирсонова теорија статистичког тестирања. Година формирања њихове теорије је, највероватније, 1933. када је објављен есеј „Проблеми најефикаснијих тестова статистичких хипотеза“ (енгл. „*On the Problem of the Most Efficient Tests of Statistical Hypotheses*“). Овде је реч о такозваној фундаменталној леми која представља математичку тврдњу о постојању одређеног оптималног метода тестирања. Лема, као таква, чини окосницу Нојман-Пирсонове теорије и, по мишљењу аутора, истовремено оправдање за Фишеров старији приступ. Нојман-Пирсонова теорија је тесно повезана са класама модела које су предуслов за математичко размишљање, па је, самим тим, и сасвим јасно поштовање које су Нојман и Пирсон исказивали према Фишеру.

2.2 Подаци из узорка као извор информација

Једина квантитативна информација која се експлицитно примењује у оквиру класичног приступа статистици јесу *подаци из узорка*. Примена априорних информација о параметру θ нема званичну улогу у овом приступу, али њихово постојање може утицати на избор статистичке процедуре која би се примењивала у конкретним ситуацијама.

[26] Lenhard, J. 2006. „Models and Statistical Inference: The Controversy between Fisher and Neyman–Pearson“, *Oxford University Press on behalf of British Society for the Philosophy of Science*, Vol. 57, pp. 70

Ставови о карактеристима других релевантних информација су јасно описани и изложени у раним радовима Нојмана и Пирсона, у којима се препознаје и прихвата значај априорних информација, али превасходно као алата којим ће се створити квантитативна основа за статистичку анализу. Фишер је такође изложио слична схватања о овом питању, али је био много оштрији у погледу одбијања било какве примене метода инверзне вероватноће за руковање априорним информацијама.

Код класичног приступа и примене података из узорка за успостављање одговарајућих статистичких процедура за потребе истраживања, као основни циљ се поставља примена података x како би се дефинисала непозната вредност параметра θ. У том смислу може се посматрати посебна функција $\tilde{\theta}(X)$ за случајну променљиву X. Таква функција се назива *статистика* или *оцењена вредност*, а стварна вредност коју узима, $\tilde{\theta}(x)$, је одговарајућа *оцена*.

Избор одговарајуће статистике зависи од параметра популације за који постоји интересовање. Вредност тог параметра је непозната, а крајњи циљ - оцена његове вредности. У том смислу је неопходно направити разлику између два појма: *оцене* и *оцењене вредности*.

По неким схватањима, сматра се да постоји техничка разлика између *оцене*, као функције случајних променљивих и *оцењене вредности*, као појединачног броја. То је разлика између процеса (оцене) и резултата тог процеса (оцењене вредности).[27]

Појединачни податак x, који је предмет посматрања, само је једна од могућих реализација експеримента, како би се дошло до оцењене вредности $\tilde{\theta}(x)$. Читав распон могућих вредности може проистећи за $\tilde{\theta}(x)$, у зависности од стварног исхода x, од којих је свака са одговарајућом вредношћу за његову вероватноћу дату преко $p_\theta(x)$. Добијена расподела вероватноће се назива *узорачка расподела* посматране статистике. $\tilde{\theta}(x)$ се посматра као опсервација трансформисане случајне променљиве $\tilde{\theta}(X)$, где је X случајна променљива са

[27] Hildebrand, D. and Ott, A.L. 1998. *Statistical Thinking for Managers*, Brooks/Cole, New York, pp. 263

функцијом густине вероватноће $p_\theta(x)$. Узорачка расподела је кључан елемент за било коју оцену која ће се извести о статистици $\tilde{\theta}(X)$ и на основу које ће се донети закључци о θ.

Посебна вредност $\tilde{\theta}(x)$ се чини као карактеристична вредност за $\tilde{\theta}(X)$ која може настати понављањем истих ситуација. Особине $\tilde{\theta}(X)$ су дугорочне; наиме, било која оцена вероватноће је искључиво интерпретативна у смислу фреквентистичког схватања вероватноће.

Примена $\tilde{\theta}(x)$ обично доводи до редукције података. На пример, може се разматрати средина скупа од n независних опсервација како би се оценила средина расподеле. У овом случају, n елемената биће сумарно приказано преко информације о једном елементу. Међутим, у другој крајности, $\tilde{\theta}(x)$ може бити оригиналан податак из узорка.

За класични приступ је такође карактеристично да захтева да параметри, које треба оценити, буду непознате константе. Стога, врши се апроксимација непознатих параметара на основу које се одређује приближна вредност расподеле вероватноће. За параметре се каже да су непознате константе, јер се обично полази од основне претпоставке да би се у идеалним околностима експеримент могао понављати бесконачно много пута, а кроз процес оцењивања би се дошло до прецизне вредности параметара. То значи, уколико се број понављања неограничено повећа оцена непознатих параметара конвергира, са вероватноћом 1, ка стварној вредности. Ово тумачење је познато под називом *класична интерпретација вероватноће* или *класична статистика*.

2.3 Улога и значај оцењивања у класичном статистичком закључивању

Теорија оцењивања, поред теорије тестирања статистичких хипотеза, представља посебан облик теорије одлучивања. Са становишта Бајесове статистике, апостериорна расподела параметара посматраних података даје потпун закључак тако што сумира очекивања о параметру након анализирања података. Међутим, са становишта класичне статистике, разликује се неколико врста закључака који се могу извести о параметру: тачкасто оцењивање, интервално оцењивање и тестирање хипотеза. Овакви закључци о параметрима захтевају вероватноћу која је израчуната на основу узорачке расподеле података, за фиксирану вредност непознатих параметара. Вероватноће су засноване на свим могућим случајним узорцима који се могу појавити, али нису условне у односу на стварни узорак који је настао.

Први вид класичног закључивања који је предмет разматрања у докторској дисертацији јесте тачкасто оцењивање, где је појединачна статистика израчуната из података из узорка, тако да представља случајну променљиву, док је њена расподела узорачка расподела. Уколико је њена узорачка расподела центрирана близу стварне, али непознате вредности параметра θ, и уколико нема превелик распон, статистика се може искористити за оцену параметра.

Нека је θ параметар популације. Тачкаста оцена $\tilde{\theta}$ параметра популације θ је функција информације из узорка која даје број који се назива **тачкаста оцена**. На пример, средина узорка $\overline{\mathrm{X}}$ је тачкаста оцена средине популације μ, а вредност коју $\overline{\mathrm{X}}$ претпоставља за дати скуп података назива се тачкаста оцена.

У статистичкој анализи, тачкасто оцењивање параметара популације има веома значајну улогу. У проучавању реалних проблема, полази се од случајног узорка величине n који се узима из укупне популације. Почетни корак у статистичкој анализи посматраних података је идентификовање расподеле вероватноће која је карактеристична за ове информације. Пошто параметри расподеле дефинишу њихове карактеристике, неопходно је одредити параметре. Оцена параметара ће се вршити на основу података из случајног узорка, тако да је од изузетног значаја да оцена параметара популације буде што је могуће боља. Самим тим оцене, као такве, обезбеђују и бољу и прецизнију статистичку анализу.

Значај тачкастих оцена се огледа и у чињеници да је велики број статистичких формула заснован на њима. На пример, тачкасте оцене средине и стандардне девијације користе се у израчунавању интервала поверења и у многим обрасцима за тестирање хипотеза.

Постоји велики број расположивих метода за оцену стварне вредности параметра за коју постоји интересовање. Три најпопуларније су: метод момената, метод максималне веродостојности и Бајесов метод. У даљем тексту биће објашњени метод момената и метод максималне веродостојности као класични методи оцењивања.

2.3.1 Метод момената у функцији оцене стварне вредности параметра

Метод момената је један од најстаријих предложених метода тачкастог оцењивања. Реч је о веома једноставној процедури за налажење оцене једног или више параметара из популације. Примена метода момената има дугу историју, где је кроз периоде оштрих разматрања његових особина у односу на друге поступке оцењивања, ипак опстао као ефикасно средство које је лако за имплементацију.[28]

Метод момената је најчешће предмет критике зато што није јединствено дефинисан, што значи да његова примена подразумева неопходност избора међу могућим оцењеним вредностима како би се пронашла она која најбоље одговара подацима који су предмет анализе. Ипак, предност метода момената је у његовој једноставности, као и потреби да се испуни само једна претпоставка која се односи на то да моменти нижег реда посматране расподеле описују постојећа посматрања.[29]

Претпоставимо да се подаци састоје од независних посматрања неке расподеле. Моменти расподеле представљаће функције делова за θ, тако да је једино потребно да се изједначи довољно момената узорка са одговарајућим моментима популације како би се добиле једначине за оцену θ. Добијене оцењене вредности ће бити пристрасне.

Нека је $\mu_k' = E\left[X^k\right]$ k-ти моменат случајне променљиве X (за који претпостављамо да постоји) и нека је $m_k' = (1/n)\sum_{i=1}^{n} X_j^k$ k-ти моменат узорка. Тада је оцењена вредност μ_k',

[28] Bowman, K.O. and Shenton, L.R. 1985. „Method of moments“, *Encyclopedia of Statistical Sciences*, Volume 5, John Wiley & Sons, New York, pp. 467

[29] Gillard, J.W. and Iles, T.C. 2005. „Method of Moments Estimation in Linear Regression with Errors in both Variables“, *Cardiff University School of Mathematics Technical Paper*, Cardiff, pp. 4

добијена применом метода момената $m_k^{'}$. Метод момената је заснован на изједначавању момената узорка са одговарајућим моментима расподеле популације и изведен је на претпоставци да моменти узорка треба да обезбеде добре оцене одговарајућих момената популације. С обзиром да су моменти популације $\mu_k^{'} = h_k(\theta_1, \theta_2, ..., \theta_l)$ често функције параметара популације, можемо изједначити одговарајуће моменте популације и узорка.

Изаберимо као оцене оне вредности параметара популације које су решења једначине $\mu_k^{'} = m_k^{'}$, k = 1,2, ..., l. Овде је $\mu_k^{'}$ функција параметара популације. На пример, нека је први моменат популације $\mu_1^{'} = E(X)$, а први моменат узорка је $\overline{X} = \sum_{i=1}^{n} X_i / n$. Дакле, моменат оцене $\mu_1^{'}$ је $\overline{X}$. Ако је k =2, тада су други моменти популације и узрока $\mu_2^{'} = E(X^2)$ и $m_2^{'} = (1/n)\sum_{i=1}^{n} X_j^2$, респективно.

У основи, следећа процедура метода момената, у којој се полази од претпоставке да је потребно оценити l параметара (на пример, θ = (θ_1, ..., θ_l)), може се користити у проналажењу тачкастих оцена за параметре популације:[30]

1. Проналажење l момената популације, $\mu_k^{'}$, k = 1, 2, ..., l. $\mu_k^{'}$ ће садржати један или више параметара θ_1, ..., θ_l.
2. Проналажење одговарајућих l момената узорка, $m_k^{'}$, k = 1, 2, ..., l. Број момената узорка биће једнак броју параметара који ће се оцењивати.
3. Из система једначина, $\mu_k^{'} = m_k^{'}$, k = 1, 2, ..., l, решених за параметар θ = (θ_1, ..., θ_l), добиће се моменат оцењене вредности за $\tilde{\theta}$.

На основу изнетог може се закључити да метод момената одређује оцењене вредности непознатих параметара изједначавајући одговарајуће моменте узорка и моменте

[30] Ramachandran, M.K. and Tsokos, P.C. 2009. *Mathematical Statistics with Applications*, Elsevier Academic Press, pp. 228

популације. Овим методом се могу утврдити оцењене вредности када се то не постиже применом других метода или када су процедуре примене других метода значајно сложеније. У поређењу са другим методама, метод момената је једноставан за примену и поседује одређене пожељне карактеристике.

2.3.2 Метод максималне веродостојности као средство за оцењивање непознатих параметара

Сматра се да је од изузетног значаја имати на располагању метод који ће бити опште примењљив у одређивању статистичких оцена са добрим карактеристикама. Један од таквих метода, који има веома важну улогу у оцењивању параметара, дефинисан од стране Роналда Фишера, је метод максималне веродостојности. Назив, као такав, установио је Фишер 1922. године. Иако је метод момената једноставан за примену, обично не даје довољно добре оцене посматраних параметара. Са друге стране, код метода максималне веродостојности, покушава се пронаћи вредност за стварне параметре из којих ће, уз највећу вероватноћу, проистећи подаци који су, заправо, и предмет посматрања. У већини практичних ситуација, метод максималне веродостојности је погодан за велике групе података те, као такав, припада реду најразноврснијих метода за прилагођавање параметарских статистичких модела подацима.

Метод максималне веродостојности се може применити у оцени непознатих параметара или функција непознатих параметара, које одговарају заједничкој функцији густине случајног узорка. Примењена процедура води до оцене параметра θ кроз максимизацију функције веродостојности параметара, за дате исходе посматраног случајног узорка.

Као основне особине метода максималне веродостојности наводе се:[31]

1. Уколико је $\tilde{\theta}$ довољна статистика за θ, јединствено решење оцене веродостојности је функција за $\tilde{\theta}$.

[31] Lindgren, W.B. 1993. *Statistical theory*, *Fourth Edition*, Chapman & Hall/CRC, Boca Raton, Florida, pp. 277

2. Уколико постоји ефикасна оцена за θ, метод максималне веродостојности ће је извести.

3. Под одређеним условима регуларности, метод максималне веродостојности је асимптотски ефикасан, при чему оцена веродостојности има нормалну расподелу са варијансом $1/I(\theta)$.

Идеја метода максималне веродостојности је да се за оцену параметра θ изабере вредност за коју је вероватноћа реализације изабраног реализованог узорка највећа.

2.3.2.1 Функција веродостојности као функција параметра θ

Функција веродостојности, у функционалној форми, идентична је заједничкој функцји густине случајног узорка. Међутим, постоји извесна разлика између ове две функције. Док се, са једне стране, заједничка функција густине тумачи као функција исхода случајне променљиве $(x_1, x_2, ..., x_n)$, за дате вредности параметара (Θ_1, Θ_2,...., Θ_k), дотле се функција веродостојности тумачи као функција вредности параметара (Θ_1, Θ_2,...., Θ_k), за дате вредности исхода случајне променљиве $(x_1, x_2, ..., x_n)$.[32]

Нека је (X_1, X_2,...., X_n) случајан узорак са заједничком густином или функцијом вероватноће $f_\theta(x_1, x_2, ..., x_n)$, где је $\theta \in \Theta$ непознат параметар.

Функција веродостојности за реализован узорак (x_1,x_2,...., x_n) је функција параметра θ дефинисана на Θ помоћу

$$L(\theta; x_1, x_2, ..., x_n) = f_\theta(x_1, x_2, ..., x_n), \tag{4}$$

где је L функција за фиксне вредности узорка.

[32] Mittelhammer, C.R. 1996. *Mathematical Statistics for Economics and Business*, Springer-Verlag New York Inc., pp. 464

Назив „функција веродостојности“ потиче из веће вероватноће да ће θ, за које $f(x_k;\theta)$ има већу вредност бити стварна вредност параметра θ у односу на θ за које $f(x_k;\theta)$ има мању вредност те би, у том случају, била вероватнија и појава x.

У дискретном случају $L(\theta; x_1, x_2,..., x_n)$ је управо вероватноћа реализације $X_1 = x_1,..., X_n = x_n$. У апсолутно непрекидном случају, потребан нам је константан производ $f_\theta(x_1, x_2,..., x_n)$ који би приказао величину потребну за реализацију $(x_1, x_2,..., x_n)$, а како би дошли до истог тумачења функције веродостојности.

Нека је $(X_1, X_2,...., X_n)$ случајан узорак из расподеле $\{p(x_k,\theta), \theta \in \Theta\}$, за обележје дискретног типа, односно $\{f(x,\theta), \theta \in \Theta\}$, ако је обележје X апсолутно непрекидног типа. У том случају, функција веродостојности се дефинише као функција од $\theta \in \Theta$,[33]

$$L(\theta) = \prod_{k=1}^{n} p(X_k, \theta), \tag{5}$$

односно

$$L(\theta) = \prod_{k=1}^{n} f(X_k, \theta). \tag{6}$$

За произвољни реализовани узорак $(x_1, x_2,..., x_n)$, при фиксираној вредности параметра θ, функција

$$L(\theta) = \prod_{k=1}^{n} p(x_k, \theta), \tag{7}$$

односно

[33] Петровић, Љ. 2006. *Теоријска статистика. Теорија статистичког закључивања*, Економски факултет, Београд, стр. 102

$$L(\theta) = \prod_{k=1}^{n} f(x_k, \theta), \tag{8}$$

одређује расподелу случајног узорка (X_1, X_2,...., X_n).

Важно је истаћи да функцију веродостојности, иако зависи од посматраних вредности узорка $x = (x_1, x_2, ..., x_n)$, треба посматрати као функцију параметра θ. У дискретном случају, $L(\theta; x_1, x_2, ..., x_n)$ пружа вероватноћу посматраног узорка $x = (x_1, x_2, ..., x_n)$, за дато θ. Отуда, функција веродостојности је статистика која зависи од посматраног узорка $x = (x_1, x_2, ..., x_n)$.

2.3.2.2 Оцена максималне веродостојности као оцењена вредност непознатог параметра

Прво детаљно објашњење оцене максималне веродостојности дао је Фишер почетком двадесетих година прошлог века, предлажући оцењивање θ на основу вредности $\hat{\theta}$ за коју функција веродостојности има максимум, за дате податке x. Оцењена вредност $\hat{\theta}$ се назива *оцена максималне веродостојности*.

У суштини, смисао је да вредност параметра, под чијим утицајем ће посматрани подаци имати највишу вероватноћу, буде најбоља оцењена вредност за θ.

Оценом максималне веродостојности сматрају се оне вредности параметра које максимизирају функцију веродостојности у односу на параметар θ, односно

$$L(\hat{\theta}; x_1, ..., x_n) = \max_{\theta \in \Theta} L(\theta, x_1, ..., x_n) \tag{9}$$

где је Θ скуп могућих вредности параметра θ.

Када су у питању особине оцене максималне веродостојности $\hat{\theta}$, као најважније наводе се следеће:[34]

1. Под општим условима, $\hat{\theta}$ је конзистентна оцена, из чега проистиче да је асимптотски непристрасна, али је у малим узорцима најчешће пристрасна.
2. Уколико постоји појединачна довољна статистика, $\hat{\theta}$ је њена функција.
3. $\hat{\theta}$ је асимптотски потупно ефикасна оцена.
4. $\hat{\theta}$ је асимптотски нормално дистрибуирана.

Са једне стране се мора водити рачуна о потенцијално слабим карактеристикама $\hat{\theta}$ за мале узорке док, са друге стране, наведене особине доприносе веома важној улози метода максималне веродостојности у класичном тачкастом оцењивању. Трећа и четврта особина директно воде до тестова за велике узорке или до оцене интервала поверења за θ.

Теорема 1. Довољност и оцена максималне веродостојности.[35] Нека је $(X_1, X_2,...., X_n)$ случајан узорак из расподеле дискретног или непрекидног типа и нека је $\tilde{\theta}(X_1, X_2,...., X_n)$ довољна статистика за параметар $\theta \in \Theta$. Уколико постоји јединствена оцена максималне веродостојности $\hat{\theta}$ за θ, тада је $\hat{\theta}$ (неконстантна) функција за $\tilde{\theta}$. Уколико оцена максималне веродостојности за θ постоји, али није јединствена, онда је могуће пронаћи оцену максималне веродостојности $\hat{\theta}$ која је функција за $\tilde{\theta}$.

Доказ. Према теореми факторизације

$$L(\theta; x_1,..., x_n) = h(x_1,..., x_n) g_\theta(\tilde{\theta}(x_1,..., x_n)), \tag{10}$$

за сваки $(x_1, x_2,...., x_n)$, θ и неке h и g_θ. Уколико постоји јединствена оцена $\hat{\theta}$ која максимизира $L(\theta)$, такође максимизира и $g_\theta(\tilde{\theta}(x_1,..., x_n))$ и, отуда, $\hat{\theta}$ је функција за $\tilde{\theta}$.

[34] Barnett, V. 1999. *Comparative Statistical Inference, Third Edition*, John Wiley & Sons, Chichester, West Sussex, England, pp. 155

[35] Rohatgi, V.K. 1984. *Statistical Inference*, John Wiley & Sons, New York, pp. 679

Уколико оцена максималне веродостојности за θ постоји, али није јединствена, бирамо посебну оцену максималне веродостојности $\hat{\theta}$, која је функција за $\tilde{\theta}$, из скупа свих оцена максималне веродостојности.

Према *принципу веродостојности*, у извођењу закључака или доношењу одлука о θ након посматрања x, све релевантне информације из експеримента су садржане у функцији веродостојности за посматрано x. Осим тога, две функције веродостојности садржаће исте информације о θ уколико су пропорционалне једна према другој као функције од θ. Оцена максималне веродостојности задовољава у потпуности принцип веродостојности.[36]

Принцип веродостојности има и неколико карактеристичних ограничења.[37] Као прво, у извођењу експеримента, од изузетног је значаја узети у обзир све вредности x које се могу појавити; самим тим се морају размотрити и одговарајуће фреквентистичке мере. Ситуација код секвенцијалне анализе је врло слична, с тим што се на одговарајућем нивоу мора донети одлука о томе да ли ће се или неће посматрање наставити. У извођењу одлука ове врсте, није довољно само имати информације о функцији веродостојности за θ, већ су потребне и допунске информације за податке који су били предмет посматрања до одређеног тренутка. Наредни проблем се односи на предвиђање будућих посматрања са намером да се претпостави будућа вредност за X.

Овде се поново могу појавити информације о подацима које превазилазе функцију веродостојности за θ. Заправо, принцип веродостојности биће примењен у свим наведеним околностима уколико ће θ садржати све непознанице које су релевантне за проблем, укључујући допунске променљиве X, а не само непознате параметре модела. На крају, једно од ограничења принципа веродостојности је и у томе што не показује како ће се функција веродостојности користити у доношењу одлука или закључака о θ.

[36] Kvam, P.H. and Vidakovic, B. 2007. *Nonparametric Statistics with Applications to Science and Engineering*, John Wiley & Sons, New Jersey, pp. 41

[37] Berger, O.J. 1980. *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag New York Inc., pp. 32

2.3.3 Интервал поверења у класичном статистичком закључивању

За већину практичних проблема тачкасто оцењивање није најпогодније, пошто није позната величина грешке која настаје, нити вероватноћа са којом се може тврдити да је оцењена вредност $\tilde{\theta}$ блиска стварној вредности параметра θ. Та карактеристика тачкастог оцењивања је, истовремено, и недостатак овог метода. Интервално оцењивање има улогу да процени скуп, односно подскуп, одговарајућег реалног простора унутар кога се може сматрати да ће се наћи права вредност параметра.

Код интервалног оцењивања такође се полази од закона вероватноће $f(x;\theta)$, $\theta \in \Omega$ у ком егзистира непознати параметар θ. Међутим, сада се на основу случајног узорка $X_1, X_2,....,$ X_n конструише *случајан интервал* који са унапред датом вероватноћом садржи праву вредност непознатог параметра θ. Случајан интервал је коначан или бесконачан интервал чија је бар једна граница случајна променљива.

Код класичног приступа, уобичајен назив који се користи за област или интервал оцењивања јесте *област поверења* или *интервал поверења*. Суштина класичног приступа односи се на формирање интервала поверења параметара који представљају непознате константе. Идеја о овим интервалима везује се за крај тридесетих година прошлог века и радове Нојмана, мада постоје и извесни наговештаји да сличан појам спомиње Лаплас 1814. год. Упоредо са развојем ове идеје развија се и концепт фидуцијалних региона и фидуцијалних интервала захваљујући Фишеру. У практичном смислу, ова два приступа су се поклопила око проблема једног параметра на који су првобитно и примењени, али је прошло неко време пре него што је основна разлика у ставовима и крајњим резултатима постала очигледна.

Када се говори о интервалима поверења, полази се од идеје о *једностраном* интервалу поверења за параметар θ. Претпоставимо да желимо да искористимо податке x за одређивање доње границе за θ. То није могуће урадити са неком сигурношћу, све док се не усвоји нека проста граница $-\infty$. Међутим, може се дефинисати статистика са следећим особинама:

$$P\{T_1(X) \le \theta\} \ge 1 - \alpha, \qquad \text{за свако } \theta. \tag{11}$$

Уколико је α близу нуле, тада је вероватније да ће T_1 (X) бити мање од θ. У складу с тим, T_1 се назива доња интервална граница за θ са нивоом поверења 1 - α. Алтернативно, $[T, \infty)$ је доњи једнострани 1 - α интервал поверења за θ.

У практичном смислу, посматране податке x заменићемо у $T(x)$, како би се добила нумеричка вредност за доњу границу поверења. Надаље, важно је препознати одговарајуће тумачење вероватноће изјаве поверења, а то је да ће дугорочно посматрано ниво поверења 1 - α за вредности $T(x)$ бити нижи од θ.

На сличан начин, могуће је одредити и горњу границу интервала поверења. У том случају се комбинује T_1 са T_2 при нивоу поверења 1 - α .

Дефиниција. Нека је $(X_1, X_2, ..., X_n)$ случајан узорак из распореда са законом вероватноће $f(x;\theta)$, $\theta \in \Omega$, где је θ непознат параметар. Нека су даље $T_1 = T_1(X_1, X_2, ..., X_n)$ и $T_2 = T_2(X_1, X_2, ..., X_n)$ две статистике за које важи $T_1 \le T_2$. Тада случајан интервал (T_1, T_2) представља *интервал поверења* за непознати параметар θ са *нивоом поверења* 1 - α, $(0 < \alpha < 1)$, или 100 (1 - α)% интервал поверења за θ, ако је за свако θ из Θ:

$$P\{\, T_1(X) \le \theta \le T_2(X)\} \ge 1 - \alpha. \tag{12}$$

Уобичајено је да се ова неједнакост пише као једнакост:

$$P\{\, T_1(X) \le \theta \le T_2(X)\,\} = 1 - \alpha. \tag{13}$$

То значи да је случајан интервал (T_1, T_2) интервал поверења за θ са нивоом поверења 1 - α који садржи праву вредност параметра θ, ако је вероватноћа најмање 1 - α.

Статистике $T_1 = T_1(X_1, X_2, ..., X_n)$ и $T_2 = T_2(X_1, X_2, ..., X_n)$ представљају доњу и горњу границу поверења случајног интервала. Случајни интервал (T_1, T_2) код кога су крајеви случајне промењиве, назива се *двострани* интервал поверења.

За један исти параметар θ може постојати више интервала поверења са истим нивоом поверења 1 - α. Дужина интервала поверења одређује меру прецизности оцењивања. Ниво поверења се може повећати једноставним узимањем интервала поверења веће дужине. Интервал $-\infty < \theta < \infty$ говори о томе да се вредност за θ налази негде на реалној правој и да има ниво поверења 1. У пракси бисмо желели да подесимо ниво за дати фиксан број 1 - α (0 < α <1). Обично се узима ниво поверења 0,95 или 0,99 и ако је могуће конструише интервал који ће бити што је могуће ужи међу свим интервалима са истим нивоом. Овакав интервал је пожељнији, пошто је више информативан.

Општи поступак конструкције најкраћег интервала поверења подразумева да се полази од случајне променљиве:[38]

$$T_\theta(\theta) = T(X_1, X_2, ..., X_n; \theta) \tag{14}$$

која зависи од θ и елемената узорка X_1, X_2,...., X_n, са потпуно одређеним распоредом независним од θ. Нека су бројеви $\lambda_1 = \lambda_1(\alpha)$ и $\lambda_2 = \lambda_2(\alpha)$ изабрани тако да је:

$$P(\lambda_1 \le T_\theta(\theta) \le \lambda_2) \ge 1 - \alpha,$$

и нека из ове неједнакости следи еквивалентна неједнакост:

$$P\left[L_1(X_1, X_2, ..., X_n) \le \theta \le L_2(X_1, X_2, ..., X_n)\right] \ge 1 - \alpha,$$

тада је $\left[L_1(X_1, X_2, ..., X_n), L_2(X_1, X_2, ..., X_n)\right]$, $100(1-\alpha)\%$ интервал поверења за θ. Бројеви λ_1 и λ_1 могу се изабрати на више начина.

Уколико се тражи најкраћи интервал поверења, тада се ови бројеви бирају тако да разлика L_1-L_2 буде минимална. Тако добијен интервал поверења је најкраћи интервал поверења са нивоом 1-α, заснован на случајној променљивој $T_\theta(\theta)$. У овом случају, важно је нагласити на којој случајној променљивој је заснован интервал, пошто може постојати

[38] Милошевић, В. 1995. *Теорија статистичког закључивања*, Научна књига, Београд, стр. 132

друга случајна промењьива $T_{\theta}^{'}(\theta)$ помоћу које би се чак могао добити краћи интервал. Из тог разлога се не тврди да ће одабрана процедура, уколико је успешно спроведена, сигурно довести до интервала са нивоом поверења 1-α који ће бити најкраћи међу свим интервалима тог нивоа.

Дефиниција.[39] Случајна промењьива $T(X,\theta)$ која је функција X=(X_1, X_2,...., X_n) и θ и чија је расподела независна од θ назива се *главна тачка* (енгл. *pivot*).

На пример, у узорковању из нормалне расподеле, уколико је варијанса позната, $(\overline{X}-\mu)\sqrt{n/\sigma}$ је логичан избор за главну тачку. Уколико је σ непознато, онда ће главна тачка бити $(\overline{X}-\mu)\sqrt{n/S}$. Уколико је потребно одредити интервал поверења за варијансу, главна тачка биће $\sum_{1}^{n}(X_i-\mu)^2/\sigma^2$, док ће у случају када је μ непознато главна тачка бити S^2/σ^2.

Процедуре које се користе за проналажење најкраћих интервала поверења заснованих на главним тачкама, такође су применљиве у проналажењу интервала који минимизира очекивану дужину.

2.4 Нојман-Пирсонова теорија тестирања хипотеза као основа класичног статистичког тестирања

Проблем тестирања статистичких хипотеза представља други део статистичког закључивања, где се информације из узорка користе како би се испитала прихватљивост неких тврђења или претпоставки. Тачност ових тврђења и претпоставки може се утврдити једино обухватањем свих јединица посматрања, али пошто то обично није случај, онда се на основу узорка и применом тестирања статистичких хипотеза испитује њихова прихватљивост.

[39] Rohatgi, V.K. 1976. *An Introduction to Probability Theory and Mathematical Statistics*, John Wiley & Sons, pp. 480

Хипотеза је тврдња која наговештава да стварна расподела вероватноће, која описује природан варијабилитет у посматрању, припада одговарајућем подскупу фамилије могућих расподела вероватноће. Другачије се може рећи да хипотеза ставља до знања да стварна вредност параметра θ припада одговарајућем подскупу параметарског простора Ω.

Теорија тестирања хипотеза се бави проблемом: „Да ли је дато посматрање у складу са неком наведеном хипотезом или није?“ Статистички тест хипотезе је правило које придружује свако могуће посматрање једној од две категорије: „сагласан са хипотезом која је предмет посматрања“ или „није сагласан са посматраном хипотезом“.

Нека је $(X_1, X_2, ..., X_n)$ случајан узорак из расподеле популације f_θ, $\theta \in \Omega$, где је функционална форма f_θ позната, осим вредности параметра θ. У многим практичним проблемима извођач експеримента је заинтересован за тестирање оправданости тврдње о непознатом параметру θ. Проблем ове врсте се најчешће тумачи као проблем *тестирања хипотеза*.

Поступак тестирања хипотеза може се дефинисати и као правило којим се ближе одређује:[40]

- За које вредности узорка ће се донети одлука да се прихвата хипотеза H_0 као истинита.
- За које вредности узорка ће се хипотеза H_0 одбацити и прихватити хипотеза H_1 као истинита.

Подскуп узорачког простора за који ће хипотеза H_0 бити одбачена назива се *област одбацивања* или *критична област*, док се област комплементарна области одбацивања назива *област прихватања*.

[40] Casella, G. and Berger, R.L. 2002. *Statistical Inference*, *Second Edition*, Duxbury Thomson Learning, pp. 374

Проблем тестирања хипотеза се може описати на следећи начин: За дати узорак $x = (x_1, x_2,...., x_n)$ потребно је пронаћи функцију одлуке која ће нас одвести до одлуке о прихватању или одбацивању нулте хипотезе. Функција одлуке мора бити тако одабрана да дели n-димензионални простор узорка на два дисјунктна скупа. Другим речима, поделом простора $\mathcal{R}_n$ на два дисјунктна скупа C_0 и C_1, у случају када је $x \in C_0$, одбацујемо нулту хипотезу H_0: $\theta \in \Omega_0$ (и прихватамо алтернативну H_1) и ако је $x \in C_1$, прихватамо нулту хипотезу H_0 да је $X \sim f_\theta$, $\theta \in \Omega_0$.

Уколико се користи наведена процедура, постоји могућност појаве две врсте грешака. Једна, која може одбацити нулту хипотезу H_0 када је она суштински тачна, назива се *грешка прве врсте*, друга, која прихвата H_0 када је нулта хипотеза погрешна, назива се *грешка друге врсте*.

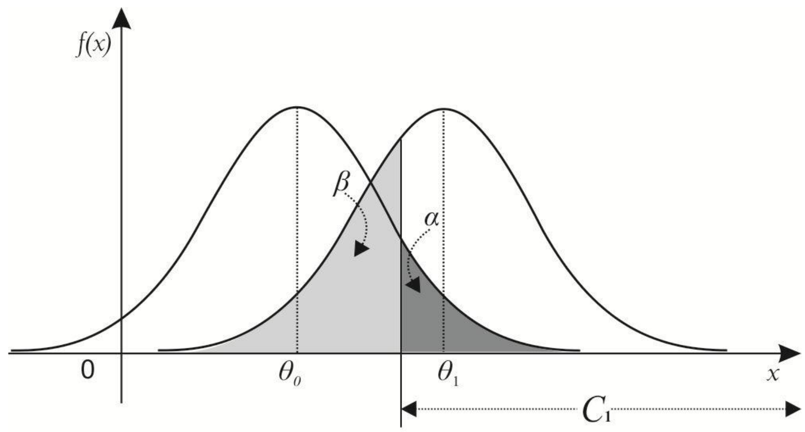
Вероватноћа грешке прве врсте је функција од θ и означава се са:

$$\alpha(\theta) = P\left[(X_1, X_2, ..., X_n) \in C_1 \middle| \theta \in \Omega_0\right] = P(H_1 | H_0)\,. \tag{15}$$

Вероватноћа грешке друге врсте је у општем случају функција од θ и означава се са:

$$\beta(\theta) = P\left[(X_1, X_2, ..., X_n) \in C_0 \middle| \theta \in \Omega_1\right] = P(H_0 | H_1)\,. \tag{16}$$

Идеално би било када би вероватноће обе врсте грешке биле што је могуће мање. Међутим, постоји извесна зависност између вероватноћа ове две врсте грешака, пошто је ризик β одређен и *нивоом значајности теста*. За одређени узорак, било које смањење вероватноће грешке прве врсте α, довешће до повећања вероватноће грешке друге врсте β и обрнуто (Слика 3). Једини начин да се смањи вероватноћа једне грешке, а да се при томе не узрокује повећање вероватноће друге грешке, јесте повећање обима узорка n. Тако се за дати ризик α са порастом величине узорка смањује ризик β, односно повећава се јачина теста. То значи да ће се са великим узорцима повећати вероватноћа одбацивања неистините нулте хипотезе.



Слика 3. Зависност вероватноћа грешке прве и друге врсте[41]

Посматрањем вероватноћа $\alpha(\theta)$ и $\beta(\theta)$ може се уочити да вредности $\alpha(\theta)$ опадају само ако се област C_1 смањује. У случају смањења области C_1, аутоматски би дошло до повећања области C_0 и вредности $\beta(\theta)$ и обрнуто.

Процедуре које се користе у пракси делују у правцу ограничења вероватноће грешке прве врсте на неки унапред утврђени ниво α (обично 0.01 или 0.05) и минимизирање вероватноће грешке друге врсте, с обзиром да не постоји тест, односно функција одлуке која ће делити n-димензионални простор на делове C_0 и C_1 тако да вероватноће $\alpha(\theta)$ и $\beta(\theta)$ буду истовремено минималне.

Након фиксирања вероватноће α на неки одређени ниво, проналази се класа тестова односно области C_1, за коју важи:

$$\alpha(\theta) \le \alpha, \quad \forall \theta \in \Omega_0 \quad (0 < \alpha < 1). \tag{17}$$

Област C_1 назива се област одбацивања нулте хипотезе или *критична област теста*, а област C_0 област прихватања нулте хипотезе. Најбоља критична област величине α је она критична област чија је вероватноћа грешке друге врсте мања од вероватноће грешке

[41] Петровић, Љ. 2006. *Теоријска статистика. Теорија статистичког закључивања*, Економски факултет, Београд, стр. 144

друге врсте ма које друге критичне области чија је размера мања или једнака α. Најбољи тест са нивоом поверења α је тест чија је критична област најбоља критична област размере α.

Пошто је

$$P\left[(X_1, X_2, ..., X_n) \in C_0 \middle| \theta \in \Omega_1\right] + P\left[(X_1, X_2, ..., X_n) \in C_1 \middle| \theta \in \Omega_1\right] = 1$$

може се извести да је:

$$P\left[(X_1, X_2, ..., X_n) \in C_1 \middle| \theta \in \Omega_1\right] = 1 - \beta(\theta). \tag{18}$$

Приказана вероватноћа назива се *јачином теста* и означава се са:

$$C(\theta) = P\left[(X_1, X_2, ..., X_n) \in C_1 \middle| \theta \in \Omega_1\right], \quad \forall \theta \in \Omega_1,$$

а може се приказати и преко вероватноће грешке друге врсте, односно:

$$C(\theta) = 1 - \beta(\theta), \quad \forall \theta \in \Omega_1. \tag{19}$$

Даље се комбинују функције $\alpha(\theta)$ и $\beta(\theta)$ у једну функцију која се назива *функција јачине теста* и означава се са:

$$C(\theta) = \begin{cases} \alpha(\theta), & \forall \theta \in \Omega_0, \\ \\ 1 - \beta(\theta), & \forall \theta \in \Omega_1. \end{cases} \tag{20}$$

Важно је нагласити да се јачина теста $C(\theta)$ израчунава само за θ из Ω_1, док је функција јачине теста дефинисана за свако θ из Ω. Пожељно је да вредности $C(\theta)$ за $\theta \in \Omega_1$ буду блиске 1, како би тест био добар. Функција јачине теста има важну улогу у процени квалитета статистичких тестова, нарочито у поређењу више критичних области које се могу користити за тестирање нулте против алтернативне хипотезе.

Такође, може се дефинисати и функција оперативне карактеристике теста, која представља вероватноћу прихватања нулте хипотезе, без обзира која је хипотеза тачна. Функција $K(\theta) = 1 - C(\theta),\ \ \theta \in \Omega$ представља функцију оперативне карактеристике и за свако θ из Ω она даје вероватноћу прихватања хипотезе H_0. График ове функције је крива оперативне карактеристике или *OC* крива.

У пракси постоје различите методе које се користе у поступку тестирања хипотеза. Реч је о методама које су се показале као корисне у различитим околностима и које су испољиле своје предности код посматрања различитих аспеката одређених проблема.

Математичка теорија статистичког закључивања развила се током двадестих и тридесетих година прошлог века, пре свега захваљујући Роналду Фишеру (1890-1962), Џерзију Нојману (1894-1981) и Егону Пирсону (1895-1980). Теорија тестирања хипотеза је једна од најчешће коришћених квантитативних методологија која је нашла примену у скоро свим сферама људског живота.[42]

Нојман-Пирсонова теорија тестирања хипотеза, која је изворно утемељена као покушај унапређења Фишеровог приступа, представља основу класичног статистичког тестирања. Овим приступом се дефинишу две супротстављене хипотезе, нулта хипотеза (H_0) и алтернативна хипотеза (H_1). Дефинисањем алтернативне хипотезе направљена је разлика између Фишерове и Нојман-Пирсонове методологије.

Фишер је користио неку врсту каузалне, опште алтернативе приликом израчунавања *p* вредности, али никада није јасно дефинисао нити користио одређену алтернативну хипотезу. Централно место у његовом схватању индуктивног закључивања заузимала је нулта хипотеза H_0.

Истраживање би требало започети постављањем нулте хипотезе за узорак који је узет из претпостављене бесконачне популације са познатом узорачком расподелом. Нулта

[42] Lehmann, E.L. 1993. „The Fisher, Neyman-Pearson Theories of Testing Hypotheses: One Theory or Two?", *Journal of the American Statistical Association*, 88, pp. 1242

хипотеза биће оповргнута или одбачена уколико оцене узорка одступају од средине узорачке расподеле више од предвиђеног мерила, односно нивоа значајности.[43] Према Фишеру[44] „уобичајено је и погодно за истраживаче да одаберу 5% за стандардни ниво значајности, у смислу да ће бити спремни да игноришу све резултате који не успевају да постигну овај стандард". Према томе, модел тестирања значајности усмерен је на одбијање нулте хипотезе на нивоу $p \leq 0.05$, те пошто Фишер сматра да p вредности представљају индуктивни доказ против нулте хипотезе, важиће да што је мања p вредност, већа је тежина поменутог доказа.

Према дефиницији, p вредност $p(X)$ је статистика теста која задовољава $0 \leq p(x) \leq 1$ за сваку узорачку вредност x. Мала вредност $p(X)$ даје доказ да је хипотеза H_1 тачна, док се за p вредност каже да је *валидна* за свако $\theta \in \Theta_0$ и за свако $0 \leq \alpha \leq 1$, односно:

$$P_\theta(p(X) \leq \alpha) \leq \alpha. \tag{21}$$

Уколико је $p(X)$ валидна p вредност, лако се може конструисати тест нивоа α који је заснован на $p(X)$. Тест којим се одбацује H_0 ако и само ако је $p(X) \leq \alpha$ је тест нивоа α пошто је $P_\theta(p(X) \leq \alpha) \leq \alpha$. Предност приказивања резултата теста преко p вредности јесте у томе што се може изабрати ниво α за који се сматра да је погодан, а затим упоредити p вредност са α како би се утврдило да ли такви подаци воде до прихватања или одбацивања H_0. Отуда, p вредност приказује резултате теста као континуиране вредности, пре него као дихотомну одлуку „прихвата се H_0" или „одбацује се H_0".

Теорема 2.[45] Нека је $T(X)$ таква статистика теста да велике вредности за T дају доказ да је хипотеза H_1 тачна. За сваку вредност из узорка x, дефинише се:

[43] Hubbard, R., Bayarri, M.J., Berk, K.N. and Carlton, M.A. 2003. „Confusion over measures of evidence (p's) versus errors (α's) in classical statistical testing", *The American Statistician* 57, pp. 172

[44] Fisher, R.A. 1966. *The Design of Experiments, Eight Edition*, Oliver and Boyd, Edinburgh, pp.13

[45] Casella, G. and Berger, R.L. 2002. *Statistical Inference, Second Edition*, Duxbury Thomson Learning, pp. 397

$$p(X) = \sup_{\theta \in \Theta_0} P_\theta(T(X) \geq T(x)). \tag{22}$$

Отуда, $p(X)$ је валидна p вредност.

Фишер се залагао за коришћење p вредности за процену исхода теста, па је код његовог тестирања p вредност базичнији појам од нивоа α. Технички, ниво α је једноставно правило одлучивања о томе које вредности p ће довести истраживача до тога да одбаци нулту хипотезу. Уколико је p вредност мања од или једнака α, нулта хипотеза се одбацује. Ниво α заправо одређује који ће подаци довести до одбацивања H_0, а који подаци неће узроковати одбијање. Област одбацивања нивоа α се дефинише као скуп свих тачака података које имају p вредност мању од или једнаку α; p вредност представља одличну меру обима у ком подаци не противрече моделу (велике p вредности не противрече моделу). Ако је изабран ниво α, уколико се деси било каква слична контрадикторност, ниво α мора бити мали. Са друге стране, чак и без одређене алтернативе, уколико је ниво α премали, тест ће изгубити смисао, па ће бити изузетно тешко одбити тест без ваљаног разлога. Разуман став би био да ниво α никада не би требало да буде изабран, већ да научник треба једноставно да процени доказе оличене у p вредности.[46]

Код Нојман-Пирсонове теорије тестирања хипотеза одабира се нулта хипотеза која се тестира против алтернативне хипотезе. За разлику од Фишерове идеје претпостављене бесконачне популације, резултати Нојман-Пирсонове теорије заснивају се на претпоставци о поновљеном случајном узорковању из дефинисане популације. Из тог разлога, примена Нојман-Пирсонове теорије је погодна у ситуацијама у којима поновљено случајно узорковање има смисла, као што су, на пример, експерименти контроле квалитета.

Нојман-Пирсонова теорија настоји да пронађе најбољи ниво α, где је α вероватноћа одбијања хипотезе H_0 када је она истинита. Област одбацивања је скуп вредности података који доводе до одбацивања нулте хипотезе, па код H_0 вероватноћа области одбацивања

[46] Christensen, R. 2005. „Testing Fisher, Neyman, Pearson, and Bayes“, *The American Statistician*, 59(2), pp. 121

мора бити α. Као најбољи тест је дефинисан онај са највећом моћи, где моћ представља највећу вероватноћу одбацивања H_0 када је H_1 истинита. Дефинисањем нивоа α, као вероватноће одбацивања нулте хипотезе када је истинита, ставља се акценат на поновљено узорковање, тако да Закон великих бројева сугерише да ће α током времена довести до погрешне одлуке, под условом да је нулта хипотеза истинита у свим узорцима. Пре него што се узму у обзир расположиви подаци, ова дефиниција је прихватљива, међутим, њена релевантност након сагледавања података постаје упитна.[47]

Пошто се код Нојман-Пирсонове теорије тестирања хипотеза испитује зависност између вероватаноћа грешака прве и друге врсте за статистике теста код којих је вероватноћа грешке прве врсте мања или једнака некој константи α, у практичним ситуацијама унапред се фиксира ниво значајности теста α, па се међу свим критичним областима са овим нивоом значајности бира она област C_1 за коју је вероватноћа грешке друге врсте β најмања.

У свом излагању у есеју „Проблеми најефикаснијих тестова статистичких хипотеза“ из 1933. године, Нојман и Пирсон су покушали да пронађу недостатке у логици тестирања фокусирајући се на две кључне тачке. Као прво, критиковали су асиметрично посматрање нулте хипотезе као недостатак Фишерове логике тестирања, инсистирајући на томе да се већ у почетној фази тестирања уведе симетричност. То би значило да модел треба да обухвати две супротстављене хипотезе (нулту против алтернативне), како би се на основу посматраних података донела одлука која је хипотеза пожељнија. Као друго, настојали су да одаберу вредност функције веродостојности или количник веродостојности, као величину коју треба максимизирати. То је био један од начина за успостављање симетричности приликом посматрања две хипотезе.

Теоријска основа Нојман-Пирсонових схватања формирана је преко тзв. фундаменталне леме, која представља израз систематског приступа ова два аутора. Из њихове анализе две врсте статистичке грешке, изведени су и појмови *величине* и *моћи*. Величина одговара

[47] Christensen, R. 2005. „Testing Fisher, Neyman, Pearson, and Bayes“, *The American Statistician*, 59(2), pp. 122

нивоу значајности, док моћ одговара аналогној величини за грешку друге врсте. Приликом супротстављања две хипотезе, прво је потребно поправити *величину* теста, а затим оптимизирати његову *моћ*. Ова размишљања су претворена у фундаменталну лему: у случају једноставне дихотомије хипотеза, за било коју могућу величину, постоји јединствено најмоћнији тест те величине.[48]

Нојман-Пирсонова фундаментална лема се уобичајено дефинише као метод расуђивања у чијој основи се налази количник веродостојности. Као таква, даје нам општи метод за проналажење најмоћнијег теста нулте хипотезе против алтернативне хипотезе.

Нека је X непрекидна случајна променљива са законом вероватноће $\{f(x,\theta),\theta\in\Omega\}$. Посматрајући проблем тестирања H_0: $\theta = \theta_0$ против H_1: $\theta = \theta_1$, закључујемо да обележје X има расподелу $f(x;\theta_0)$ и расподелу $f(x;\theta_1)$. Даље нека је $(X_1, X_2,..., X_n)$ случајан узорак променљиве X и нека су $L_0 = L_0(x_1, x_2,..., x_n;\theta_0)$ и $L_1 = L_1(x_1, x_2,..., x_n;\theta_1)$ функције веродостојности овог узорка. Уколико су C_0 и C_1 делови простора C, важиће да је $\frac{L_0(x_1, x_2,..., x_n;\theta_0)}{L_1(x_1, x_2,..., x_n;\theta_1)} \le k$ за сваку тачку $(x_1, x_2,..., x_n)\in C_1$ и $\frac{L_0(x_1, x_2,..., x_n;\theta_0)}{L_1(x_1, x_2,..., x_n;\theta_1)} > k$ за сваку тачку $(x_1, x_2,..., x_n)\in C_0$, где је k одређен тако да вероватноћа грешке прве врсте буде α.

Теорема 3. Нојман-Пирсонова фундаментална лема. Уколико је C_1 критична област величине α, односно $\alpha(\theta_0) = P\left[(X_1, X_2,..., X_n)\in C_1 \middle| H_0\right] = \alpha$, таква да за $k \ge 0$ важи:

$$\frac{L_0}{L_1} \le k \text{ унутар области } C_1$$

$$\text{и } \frac{L_0}{L_1} > k \text{ изван области } C_1, \tag{23}$$

[48] Lenhard, J. 2006. „Models and Statistical Inference: The Controversy between Fisher and Neyman–Pearson“, *Oxford University Press on behalf of British Society for the Philosophy of Science*, Vol. 57, pp. 80

тада је C_1 најбоља критична област величине α за тестирање нулте хипотезе H_0: $\theta = \theta_0$ против алтернативне хипотезе H_1: $\theta = \theta_1$. За одговарајући тест се каже да је најбољи тест величине α или униформно најјачи тест величине α за тестирање нулте хипотезе H_0: $\theta = \theta_0$ против алтернативне хипотезе H_1: $\theta = \theta_1$.

Тестирање помоћу количника веродостојности може се вршити и у случајевима када постоје „сметајући" параметри, који су присутни у моделу, али немају директан утицај на закључивање. Присуство таквих параметара неће утицати на примену метода веродостојности, али би могло водити до примене другачијег теста.[49]

Нојман-Пирсонова лема се може применити и код закона вероватноће који садржи више параметара, с тим што је важно да и нулта и алтернативна хипотеза буду просте. При томе, хипотезе H_0 и H_1 не морају да се односе на параметре. Приликом тестирања две просте хипотезе применом Нојман-Пирсонове леме, постојање критичне области је загарантовано. Поред тога, горенаведена теорема обезбеђује средство за одређивање најбоље критичне области. Међутим, ова теорема даје само облик области одбацивања, док стварна област одбацивања зависи од конкретне вредности α.

У реалним околностима, ретко се дешава да се тестирају две просте хипотезе. Такође, не постоји неки општи резултат у облику Теореме 2. за сложене хипотезе. Ипак, приликом тестирања нулте хипотезе H_0: $\theta = \theta_0$ против алтернативне хипотезе H_1: $\theta > \theta_0$, може се узети одређена вредност $\theta_1 > \theta_0$, а затим пронаћи најмоћнији тест за H_0: $\theta = \theta_0$ против H_1: $\theta > \theta_0$. Уколико овај тест не зависи од одређене вредности θ_1, тада се за такав тест каже да је униформно најмоћнији тест за за H_0: $\theta = \theta_0$ против H_1: $\theta > \theta_0$.

[49] Casella, G. and Berger, R.L. 2002. *Statistical Inference, Second Edition*, Duxbury Thomson Learning, pp. 378

3. БАЈЕСОВА СТАТИСТИКА

Статистичка анализа обједињује становишта и искуство два основна статистичка филозофска правца – фреквентизма и Бајесијанизма. Основна разлика између два погледа на статистичку теорију и праксу огледа се у различитом тумачењу појма вероватноће.[50] Као што само име каже, заговорници фреквентистичке статистике прихватају фреквентистичко тумачење, по ком је вероватноћа остварења једног догађаја граница његове дугорочне релативне фреквенције. Строго придржавање овог тумачења није увек могуће у пракси. На пример, приликом проучавања ретких догађаја, често нису на располагању велики узорци података, па у таквим случајевима присталице фреквентистичке статистике прибегавају теоријским резултатима. Са друге стране, Бајесова статистика своје становиште о вероватноћи тумачи на субјективан начин, по ком је вероватноћа субјективна и представља степен уверења који се коригује у тренутку када информације или подаци постану расположиви.[51]

Са појмом вероватноће је уско повезан појам неизвесности, где је по присталицама фреквентистичког приступа извор неизвесности случајно инхерентан у реализацији случајних променљивих. Расподела вероватноће променљивих није под утицајем неизвесности. Супротно, Бајесова статистика расподелу вероватноће посматра као неизвесну, истовремено сматрајући да подлеже променама у тренутку када нове

[50] Rachev, T.S., Hsu, J.S.J., Bagasheva, B.S. and Fabozzi, F.J. 2008. *Bayesian Methods in Finance*, John Wiley & Sons, Inc., New Jersey, pp. 1.

[51] Појам субјективне вероватноће проистиче из аргумената о рационалности преференција заговорника овог схватања. Сам појам датира још из тридесетих година прошлог века из радова де Финетија и Ремзија, а даљи развој доживљава преко Саважа и Линдлија.

информације постану познате. Самим тим, незвесност је индиректно повезана са кориговањем вероватноће. Уверења о вероватноћи заснована на постојећем знању презентују се као *априорна вероватноћа*, док *апостериорна вероватноћа* представља коригована уверења.

Бајесова статистика обезбеђује сврсисходан метод за унапређење уверења, а све у циљу добијања нових информација. Главни циљ проучавања јесте предвиђање природе будућих података на основу ранијег искуства. У том смислу је неопходно конструисати модел вероватноће који ће садржати непознате параметре. Параметри су заправо механизам на основу ког се врши оцена будућег понашања.

За разлику од класичног, фреквентистичког приступа који захтева да параметри буду непознате константе које треба оценити, али које се не сматрају случајним промењивима, Бајесов приступ параметре посматра као непознате случајне промењиве. Расподела вероватноће, позната као *априорна расподела*, више се заснива на ранијем искуству и представља расподелу параметара пре избора узорка који се оцењује. Експеримент се спроводи коришћењем расположивих података, којих има коначан број, при чему се примењује Бајесово правило за израчунавање *апостериорне расподеле* као расподеле вероватноће над параметарским простором. Апостериорна расподела не говори о томе каква ће бити наша уверења, већ о томе како ће се мењати под утицајем нових информација.

У овом делу рада биће приказане опште поставке теорије које развија тзв. Бајесовска школа, као и главне карактеристике Бајесовог приступа, уз објашњење Бајесове теореме и њене рекурзивне примене, као основе на којој почива Бајесова статистика.

3.1 Бајесов приступ статистичком закључивању

Током последњих година, па и деценија, Бајесов приступ је постао изузетно популаран у креирању статистичких модела за решавање реалних проблема. Бајесово закључивање се схвата као аналитички метод који комбинује информације прикупљене из експеримената са претходним знањем које је важно за извођење експеримента. Тиме се комбинују све нове информације које су на располагању са априорним информација, те се тако ствара основа за примену одговарајуће статистичке процедуре.

Класични приступ статистичком закључивању, који је у претходном делу детаљно објашњен, заснован је једино на случајном узорку. То значи да ако расподела вероватноће зависи само од скупа параметара θ, применом класичног приступа биће изведени закључци о θ само на основу узорка $X_1, X_2, ..., X_n$. Код овог приступа, анализира се једино скуп узорачких вредности, међутим, мора се истовремено замислити шта би се догодило уколико би се извукао велики број случајних узорака из популације. На пример, уколико се посматра нормалан узорак са познатом варијансом, 95% интервал поверења за средину популације μ биће приказан преко случајног интервала $(\bar{X} - 1.96\sigma/\sqrt{n}, \bar{X} + 1.96\sigma/\sqrt{n})$. То значи да уколико се узорци више пута узимају из популације, најмање 95% случајних интервала ће садржати стварну средину μ. Класични приступ закључивању не узима у обзир ни једну априорну информацију која можда постоји из ранијих истраживања. Често се дешава да су истраживачи суочени са проблемом за који је карактеристичан један скуп податка, па је потребно одредити вредност параметра у тренутку узимања података. Основно питање које се провлачи јесте која је најбоља оцена параметра која се може одредити на основу једне априорне информације. Статистички приступ који примењује априорно знање, најчешће субјективно, као додатак доказима из узорка како би се оценили параметри популације познат је под називом ***Бајесов приступ***.

Бајесовим приступом обезбеђује се метод за кориговање неизвесности под утицајем нових доказа. За податке се и даље претпоставља да воде порекло из расподеле која припада познатој параметарској фамилији. Међутим, Бајесово закључивање подразумева субјективно тумачење вероватноће. Док класични приступ занемарује априорно знање, дотле Бајесов приступ прикључује ово знање подацима који су предмет посматрања како би се кориговала вероватноћа. Након што су подаци прикупљени, мишљење се може променити. Применом Бајесовог правила, израчунаће се апостериорна расподела, заснована на априорним знањима и доказима из података. Сви закључци биће изведени израчунавањем одговарајуће статистике апостериорне расподеле.

Бајесов приступ тежи оптималном спајању информација из два извора: (1) знања познатог из теорије или мишљења формираног на почетку истраживања у априорном облику; и (2) информација садржаних у подацима у облику функције веродостојности. У суштини, априорна расподела репрезентује почетна уверења, с обзиром да су информације у подацима изражене преко функције веродостојности. Комбиновањем априорне расподеле и функције веродостојности добија се апостериорна расподела. Код Бајесовог приступа, параметри се посматрају као случајне променљиве у смислу да им је могуће доделити субјективну расподелу вероватноће која описује уверење о стварној вредности параметра.

Разлози који иду у прилог Бајесовом приступу могли би се дефинисати на следећи начин:[52]

- Применом Бајесовог приступа највећи број закључака изведен је на основу посматраних података - за разлику од класичног приступа, узимају се у обзир само посматрани подаци, а не и неки други подаци; такође, примена Бајесовог приступа не захтева разматрање узорачке расподеле.
- Из угла Бајесовог приступа, оправдано је говорити о вероватноћи да ће се вредност параметра наћи у одређеном интервалу или о вероватноћи да је хипотеза тачна – често се дешава да су закључци класичног приступа нетачни.
- Бајесов приступ обезбеђује прикладан модел за имплементацију научног метода – априорна расподела вероватноће може се користити за изражавање почетних уверења о посматраној популацији, при чему ће се прикупити сви релевантни подаци, док ће апостериорна расподела вероватноће одражавати нова, кориговања уверења о параметрима популације под утицајем новоприкупљених података.

[52] Ramachandran, M.K. and Tsokos, P.C. 2009. *Mathematical Statistics with Applications*, Elsevier Academic Press, pp. 561

- Сви закључци о параметру изводе се давањем одговарајућег прегледа апостериорне расподеле вероватноће.

Иако Бајесов приступ, због свих теоријских и рачунских изазова пред којима се налази, има релативно ограничену примену, најновија достигнућа Бајесове анализе у комбинацији са растућим рачунарским могућностима чине Бајесове методе све практичнијим и популарнијим.

3.1.1 Структурне особине Бајесовог закључивања

Статистичко закључивање може се посматрати кроз две димензије, где свака од димензија представља два могућа избора. Први је избор извођење закључака или доношење одлука, док се други избор односи на препознавање да ли је нагласак на директној или инверзној вероватноћи.

Слика 4. је шематски приказ ове две димензије, где се на *X* оси полази од закључака, који су на левој страни, и иде до одлука на десној страни, док се на *Y* оси полази од директне вероватноће и иде до инверзне вероватноће.

На слици се види да се према Фишеровом схватању тестови значајности заснивају више на директној вероватноћи, док је фидуцијална вероватноћа налик инверзној вероватноћи. У случају Нојмана и Пирсона, присутна је извесна разлика, пошто Пирсон никада није истицао значај своје улоге у доношењу одлука, већ је тумачење више заснивао на личном доприносу у подстицању нечијих идеја.

Део који захтева посебну пажњу је де Финетијева субјективна Бајесова теорија заснована на инверзној вероватноћи.

Слика 4. Шематски приказ основних школа закључивања[53]

Италијански математичар, актуар и Бајесијанац, Бруно де Финети (1906-1985), проценио је да ће до 2020. године Бајесов приступ статистичком закључивању у потпуности превладати над класичним приступом. Без обзира да ли ће се његова предвиђања обистинити, нема сумње да Бајесова статистика све више добија на значају.[54] Данас када се говори о примени Бајесове статистике у различитим областима истраживања, мисли се на један моћан статистички метод који у статистичкој анализи заузима равноправно место са класичном методологијом.

Бајесова статистика обрађује вероватноћу исказа преко Бајесове теореме. У овој поставци, вероватноћа је субјективна и условна, за разлику од онога што се тврди да је циљ класичног гледишта. Сама Бајесова теорема је изведена као директна последица једног од аксиома теорије вероватноће, па се, као таква, веома често процењује на основу његових последица, а не на основу његовог порекла. Да би се то разјаснило непходно је размотрити битне структурне особине Бајесовог закључивања које омогућавају његову ефикасну употребу и евалуацију.

[53] Senn, S. 2011. „You May Believe You Are a Bayesian But You Are Probably Wrong“, *RMM*, Volume 2, Special Topic: Statistical Science and Philosophy of Science, pp. 57

[54] Senn, S. 2003. „Bayesian, Likelihood, and Frequentist Approaches to Statistics, A comparison of methods“, *Applied Clinical Trials*, pp. 35

Бајесово тумачење поузданости моделовања значајно се разликује од класичног тумачења. Параметри модела се посматрају као непознате константе, али ова терминологија има различито значење за класичне и Бајесове статистичаре. Према класичном схватању, непозната константа је мање дисперзна скаларна величина чија вредност није одређена, док Бајесијанци представљају „непознати“ аспект преко функције густине вероватноће, па се потом математички модел параметра третира као случајна променљива. Такође, Бајесијанци настоје да нагласе да параметар није заправо променљива у смислу промена, али да његова права вредност једноставно није доступна (у експерименту коначне величине). Номенклатура „случајна величина“ је и уведена да би се направила ова разлика.[55]

Као што је већ споменуто, према класичном схватању, статистички параметри модела се посматрају као непознате константе. Методе закључивања дају оцењене вредности параметара. Ове оцењене вредности ће бити дистрибуиране, односно зависиће од непознате стварне вредности параметара. Отуда знатан део класичних статистичких закључака доприноси развоју статистичких особина оцењених вредности (пристрасност, ефикасност, итд.). Једна од потврда тога је да и поузданост извештаја није директно повезана са параметарима модела. Друга особина класичног закључивања је да закључци често зависе од експерименталних процедура, као што су правила заустављања.

За разлику од класичног, према Бајесовом схватању, параметри модела су случајне величине које се директно описују. Поступак закључивања се спроводи модификовањем априорне расподеле параметара (функција густине вероватноће) преко веродостојности узорка како би се добила апостериорна расподела. Тиме је, у односу на класично закључивање, параметарски простор непосредно доступан. Интервали поверења су развијени кроз интегрисање апостериорне густине и директно приказивање важећих исказа о вероватноћи параметара модела. По правилу, конкретни исходи експеримента се реализују без последица – правило заустављања је неинформативног карактера. Бајесове

[55] Smith, R.L. 1982. „The Bayesian Inference Method and Its Application to Reliability Problems“, *NRL Memorandum Report 4903*, Texas Research Institute, pp. 5

методе пружају могућност интегрисања претходног искуства (кроз априорне информације) са оним што је научено до посматраног тренутка. Овакав опис чини Бајесово закључивање атрактивном алтернативом, међутим, кључна потешкоћа је у избору одговарајуће априорне информације. Једна могућност је игнорисање ових информација и дефинисање такозваног *априорног незнања* (априорно игнорисање). Међутим, априорно незнање је најчешће сматрано неприкладним и самим тим је постало основа за наставак дебате.

Бајесова статистика, чију је основу поставио Бајес пре готово 250 година, данас представља модеран приступ статистичком закључивању. Расправа око правог значења Бајесових идеја и даље се води, али, упркос томе, Бајесове методе су нашле своју примену у различитим областима истраживања, као што су актуарске науке, биостатистика, економија, образовање, психологија, друштвене науке итд. Управо из тог разлога, неопходно је реално сагледати могућности примене Бајесове статистике у савременим околностима, уз сагледавање основних елемената и њихових карактеристика, које је и чине подобним алатом за сложену статистичку анализу.

3.1.2 Основне идеје на којима је заснован Бајесов приступ

Један од основних механизама учења је асимилирање информација које долазе из спољашњег окружења, а затим и ажурирање постојеће базе знања помоћу тих информација.[56] Овај механизам се налази у средишту Бајесовог учења. Доношење одлука на основу Бајесовог приступа подразумева ревидирање уверења под утицајем нових, доступних података. Са Бајесове тачке гледишта, вероватноће се тумаче као степени уверења, па се такав процес учења састоји се од преиспитивања вероватноће. Бајесова теорема пружа формално средство за активирање таквог механизма, пошто представља једноставан израз комбиновања знања о расподели параметара модела и информација о параметрима садржаним у подацима.

[56] Rachev, T.S., Hsu, J.S.J., Bagasheva, B.S. and Fabozzi, F.J. 2008. *Bayesian Methods in Finance*, John Wiley & Sons, Inc., New Jersey, pp. 6

Као основне идеје на којима се темељи Бајесов приступ статистици, наводе се следеће:

- С обзиром да постоји неизвесност у погледу стварне вредности параметара, параметри ће бити посматрани као случајне променљиве.

- Правила вероватноће се користе како би се директно извели закључци о параметрима; тврдње о параметрима треба тумачити као „степене слободе“, а априорна расподела мора бити субјективна. Мишљење појединца се дефинише као релативна мера, а затим се додељује свакој могућој вредности параметра. На тај начин се одређује колико је вероватна свака вредност параметра пре посматрања података.

- Након добијања података, применом Бајесове теореме преиспитују се уверења у вези параметра. На овај начин долази се до апостериорне расподеле која даје релативну вредност која је додељена сваком параметру након анализирања података. Апостериорна расподела проистиче из два извора: априорне расподеле и посматраних података.

Такође, сматра се да је Бајесова теорема једини подобан начин за измену личних уверења о параметрима посматраних података.

Бајесова анализа у суштини представља независну парадигму за статистику. У том смислу је неопходно навести и седам главних аргумената[57] којима се пружа подршка Бајесовој анализи:

1. ***Расположивост значајних априорних информација***. У одређеним деловима статистичког проблема могу се појавити грешке приликом укључивања априорних информација које, последично, могу довести до безначајних, па чак и апсурдних закључака.

[57] Berger, O.J. 1980. *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York, pp. 118

Већина следбеника класичног приступа би се сложила око примене поузданих и значајних априорних информација, тако да је утицај овог размишљања на општу прихваћеност Бајесове парадигме нејасан. С друге стране, једна од предности прихватања Бајесовог гледишта односи се на могућност препознавања тренутка *када* значајне априорне информације постају расположиве. Такође, када постану расположиве, Бајесов приступ указује на то како их је могуће искористити на адекватан начин.

2. ***Одређивање* н*еизвесности на основу вероватноће***. Задатак статистике је да обезбеди информације или закључке о неизвесним величинама и да дâ одговоре о величини неизвесности. Језик неизвесности је вероватноћа и једино Бајесов приступ доследно користи овај језик како би директно одговорио на неизвесност.

Као потврда наведеног може се, на пример, узети тестирање статистичких хипотеза. Хипотезе су неизвесне, а резултати Бајесове анализе биће једноставно извештаји о претпостављеној *вероватноћи* хипотеза (на основу података и априорних информација). Супротно, класични приступ ће обезбедити „вероватноће грешке прве врсте или грешке друге врсте“ или „ниво значајности (p вредност)“ који су, у најбољем случају, индиректно повезани са *вероватноћом хипотеза*.

3. ***Условно становиште***. Код Бајесовог приступа од суштинског значаја је условна анализа посматраних података, насупрот утврђивања средње вредности свих могућих података код класичног приступа,

4. ***Кохеренција и рационалност***. Уколико је дефинисана функција губитка преко теорије корисности и априорна расподела за θ, онда је потребно оценити ефекат a преко Бајесовог очекиваног губитка, а правило одлучивања δ преко Бајесовог ризика. Ово претпоставља постојање губитка и априорне вероватноће, па се у том смислу не сматра правим аргументом који иде у прилог примени Бајесове анализе.

Упркос ограничењима и недостацима у погледу развоја рационалности и кохеренције, важно је истаћи да се на овај начин обезбеђује веома моћан доказ којим се потврђује истинитост Бајесовог приступа.

5. ***Еквиваленција класичног оптималног и Бајесовог правила***. Када је класични оптимални принцип задат, очекивана је редукција размишљања о класама статистичких процедура које су прихватљиве према овом принципу. У теорији одлучивања, на пример, очекује се да се узмимају у обзир само прихватљива правила одлучивања. У специјалним случајевима, у тестирању просте против просте хипотезе, у жељи да мала вероватноћа грешке буде принцип оптималности, потребно је редуковати размишљање о категорији „најмоћнијег" теста. У оваквим ситуацијама, категорија „прихватљивог" правила одлучивања у сагласности је са Бајесовим правилима одлучивањима. Ова подударност јасно сугерише да је потребно извршити избор из расположивих, прихватљивих правила, а на основу априорне информације.

6. ***Оперативне предности Бајесове анализе***. На основу наведеног би се могло рећи да је Бајесова анализа у потпуности разумљива, али и да је, истовремено, доста сложена. Бајесова анализа подразумева дефинисање априорне расподеле (за потребе теорије одлучивања) и функције губитка. У суштини, све док се не види како Бајесова анализа функционише, веома је тешко извршити поређење њене оперативне ефективности са истом класичног карактера. Прави начин за поређење би био пролажење кроз различите ситуације уз истовремену примену ове две парадигме.

Упркос томе, корисно је навести неке од значајнијих оперативних предности Бајесове анализе:

- Довођењем посматраних података у жељено стање, значајно је поједностављена анализа (узима се у обзир само посматрана функција вероватноће). Предности оваквог приступа су:

  \- избор реалних модела је једноставнији пошто постоји мања потреба за моделима који захтевају посебна класична израчунавања;

  \- олакшан је рад са робустношћу свих врста, с обзиром да варијације модела не изазивају суштинске промене у потребним израчунавањима;

- дозвољена је примена правила заустављања (енгл. *stopping rules*), јер као што наводе Едвардс (Ward Edwards), Линдман (Harold Lindman) и Саваж:[58] „Ирелевантност правила заустављања у статистичком закључивању враћа једноставност и слободу у обликовању експеримената... Многи извођачи експеримената би желели да имају слободу у прикупљању података...“. Класична статистика не дозвољава да се експеримент прекине када се појаве неки неочекивани докази.

- различити облици цензурисања података не представљају проблем за Бајесову анализу, али праве озбиљне проблеме класичној статистици.

- За Бајесову неинформативну априорну анализу очекује се да буде једноставна и успешна. Уколико се жели избећи априорно субјективно одређивање, најбољи избор је Бајесова анализа.

- Бајесова анализа за резултат има коначну, апостериорну расподелу за непознато θ и на основу тога је могуће добити одговор на велики број питања. На пример, поред оцењивања θ могуће је добити и прецизне величине за оцену (или, алтернативно, Бајесов скуп поузданости за θ), што је у супротности са класичном статистиком где се, у овом сегменту, јављају проблеми. Други пример се односи на тестирање хипотеза, где се Бајесова вероватноћа може користити код било ког броја хипотеза (на основу апостериорне расподеле), док код класичног приступа, тестирање хипотеза постаје све теже са повећањем броја хипотеза.

- Бајесова анализа је одлична алтернатива за статистичке процедуре које се примењују код великих узорака. Када су узорци веома велики Бајесове процедуре су, готово увек, еквивалентне класичним процедурама, а и често су подесније за ограничавање, нарочито код узорака мале величине (где обично многе класичне процедуре великих узорака немају успеха).

[58] Edwards, W., Lindman, H. and Savage, L.J. 1963. „Bayesian Statistical Inference for Psychological Research”, *Psychological Review*, Vol. 70, No. 3, pp. 239

7. *Објективност и недовољна научна поузданост*. Најчешћа критика следбеника класичне статистике, када је у питању Бајесова анализа, односи се на то да различита претходна знања обично дају различите одговоре, под претпоставком мањка објективности. Ипак, како би се постигла објективност, најбоље је определити се за Бајесову анализу која би била заснована на неинформативним априорним подацима.

Наведени аргументи, појединачно посматрани, нису довољан доказ за оправданост примене Бајесове статистике, међутим, кроз њихово целовито сагледавање може се обезбедити јака основа за прихватање Бајесовог приступа статистици.

Једна од основних карактеристика Бајесове статистике јесте и то што се заснива на примени једног алата, Бајесове теореме, којом се ревидира наше уверење о датим подацима. У концептуалном смислу, реч је о једноставнијем начину за доношење закључака. Из тог разлога се може рећи да Бајесова перспектива нуди бројне предности у односу на традиционалну, фреквентистичку перспективу:[59]

- „Објективност" фреквентистичке статистике је добијена без узимања у обзир претходног знања о процесу који се мери. Ипак, у научним истраживањима обично постоји неко претходно знање о посматраном процесу који се мери, тако да одбацивањем ове информације губе се значајни подаци. Бајесова статистика користи оба извора информација: априорне информације о процесу и информације о процесу које су садржане у подацима, а користе се у комбинацији применом Бајесове теореме.

- Бајесов приступ даје директне исказе о вероватноћи параметара, што је много корисније од исказа поверења које обезбеђује фреквентистичка статистика. То је и веома убедљив разлог за коришћење Бајесове статистике, пошто ће се фреквентистички интервал поверења тумачити као интервал вероватноће.

[59] Bolstad, M.W. 2007. *Introduction to Bayesian Statistics,* Second Edition, John Wiley&Sons, Inc., Hoboken, New Jersey, pp. 21

- Бајесова статистика има јединствени алат, Бајесову теорему, која се користи у свим околностима, што је у супротности са фреквентистичким процедурама, које захтевају много различитих алата.

- Бајесове методе често надмашују фреквентистичке методе, чак и када се оцена врши по фреквентистичким критеријумима.

- Бајесова статистика се на директан начин бави параметрима, тако што их маргинализује од заједничке апостериорне расподеле.

- Применом Бајесове теореме може се предвидети расподела будућих опсервација, што није увек једноставно урадити применом класичне статистике.

Предности примене Бајесове статистике, као и њене основне карактеристике, добро су познате статистичарима, међутим, примена Бајесових метода у пракси наилазила је на бројне потешкоће. У последњих неколико година, развијени су компјутерски алгоритми за извлачење случајно изабраног узорка из апостериорне расподеле, без потребе да се изврши потпуна процена. Апостериорну расподелу је могуће довести до жељене прецизности узимањем довољно великог случајног узорка. На овај начин се отклањују недостаци Бајесове статистике, тако да ју је сада могуће користити у пракси за решавање проблема са великим бројем параметара, као и за расподеле општих узорака и опште априорне расподеле.

3.2 Предности примене Бајесове парадигме

Бајесове статистичке методе пружају комплетну парадигму и за статистичко закључивање и за доношења одлука у условима неизвесности. Бајесове методе углавном се заснивају на строгим математичким основама којима се обезбеђује кохерентна методологија која омогућава укључивање релевантних почетних информација, а која решава многе проблеме са којима се суочавају традиционалне статистичке методе. Бајесова парадигма се заснива на тумачењу вероватноће као условне мере неизвесности која најближе одговара употреби термина „вероватноће“ у уобичајеном речнику. Статистичко закључивање о посматраној величини описује се као модификација неизвесности његове вредности у складу са расположивим доказима, док Бајесова теорема прецизира како би требало извршити наведену модификацију. Бајесова парадигма пружа једноставна правила, заснована на законима вероватноће, за управљање неизвесношћу. Анализом сложених проблема са више различитих, али међусобно делујућих извора неизвесности, добијају се проблеми који су резултат математичке манипулације. Закони вероватноће се затим примењују како би се извели закључци о вероватноћи било које величине или скупа величина која су предмет интересовања.[60]

Бајесове методе могу се применити на комплексне, значајно структуриране проблеме, који су углавном недоступни традиционалним статистичким методама. У научном извештавању и одлучивању, често се као важан конкретан случај наводи да је једини прихватљив податак онај који се може извести из расположивих документованих података. За разлику од већине других грана математике, класичним методама статистичког закључивања недостаје аксиоматска основа, па, као последица тога, анализа истих података, применом различитих процедура, може довести до некомпатибилних резултата. Насупрот томе, Бајесов приступ статистичком закључивању чврсто је заснован на аксиоматским темељима који обезбеђују логичку структуру и гарантују узајамну конзистентност предложених метода, па се може рећи да Бајесове методе стварају целовиту парадигму за статистичко закључивање.

Практичне последице усвајања Бајесове парадигме су далекосежне. Сматра се да Бајесове методе:[61]

- смањују статистички закључак до проблема у теорији вероватноће, чиме се смањује потреба за потпуно новим концептом, и

[60] West, M. and Harrison, J. 1997. *Bayesian Forecasting and Dynamic Models, Second Edition*, Springer-Verlag, New York, Inc. pp. 20

[61] Bernardo, J.M. 2001. „Bayesian Statistics“. In *Encyclopedia of Life Support Systems* (EOLSS), UNESCO, pp. 4

- служе за дискриминацију класичних статистичких техника, тако што пружају логично оправдање за неке или доказују логичку недоследност других.

Такође, као две најзначајније предности примене Бајесове парадигме издвајају се:[62]

- неизвесност се изражава преко расподеле вероватноће и статистички закључак може бити аутоматизован, тако да прати концептуално једноставан рецепт оличен у Бајесовој теореми, и

- доступне априорне информације су доследно укључене у статистички модел који описује податке.

Са друге стране, главна последица је математичка потреба да се опишу све неизвесности присутне у проблему помоћу расподеле вероватноће. Конкретно, непознати параметри у моделима вероватноће морају имати заједничку расподелу вероватноће која описује доступне информације о њиховим вредностима, што се често сматра карактеристичним елементом Бајесовог приступа. Важно је поново приметити да се параметри у оквиру Бајесове парадигме, за разлику од класичне статистике, третирају као случајне променљиве. Тиме није дат опис њихове варијабилности (параметри су обично фиксне непознате величине), већ опис неизвесности о њиховим правим вредностима. Код Бајесове парадигме, закључак је у принципу једноставан, јер је логички статус будућег посматрања и параметра θ исти, пошто су оба случајне променљиве. Ово је у супротности са инференцијалним методама класичног приступа, које су углавном теже, пошто будуће посматрање и параметар имају другачији статус; први је случајна променљива, а други фиксна вредност.[63]

Код Бајесове парадигме, процес учења из података се систематски спроводи кроз примену Бајесове теореме која комбинује расположиве априорне информације и информације

[62] Vidakovic, B. 2011. *Statistics for Bioengineering Sciences With MATLAB and WinBUGS Support*, Springer Science+Business Media, LLC, pp. 281

[63] Young, G.A. and Smith, R.L. 2005. *Essentials of statistical inference*, Cambridge University Press, Cambridge, pp. 52

добијене из података како би се добила апостериорна расподела. Међутим, у многим ситуацијама се дешава да су доступне априорне информације о посматраној величини или сувише нејасане, да би се оправдао напор потребан да се она формализује у облику расподеле вероватноће, или превише субјективне да би биле корисне у научном смислу или за доношење одлука. Из тог разлога је веома важно идентификовати математички облик „неинформативе“ априорне информације, која ће имати минималан ефекат на апостериорни закључак.

Суштински, једна од предности Бајесове парадигме је и могућност употребе стручних априорних информација. Укључивањем стручног априорног мишљења подаци који су предмет анализе добиће на значају и у стварним животним проблемима одлучивања, али и у проблемима статистичког закључивања са узорцима мале величине или са бесконачним димензионим параметарским простором.

Бајесова парадигма омогућава интеграцију посматраних података и априорних знања. У случају када нема података или постоји веома мало података, апостериорна расподела је једнака или веома близу априорне расподеле. У случају када постоји много података, а посебно у случају када се функција веродостојности разликује од априорне расподеле, апостериорна расподела доминира над функцијом веродостојности. Иако је Бајесов приступ веома интуитиван и појмовно разуман, примена може да буде тешка у зависности од избора параметара. Поред тога, веома је важно питање како ће се Бајесово ажурирање вршити у ситуацији која захтева примену скупљих рачунских модела и метода тестирања. Уколико постоји само неколико посматраних података, тада ће апостериорна расподела бити веома слична априорној расподели, па ће општи закључак бити да се није дошло до нових сазнања. У случају када постоји много података, пожељнија је примена метода максималне веродостојности који може бити једноставнији и лакши за примену од формулисања априорне расподеле.

Једна од предности Бајесовог приступа је да се добија комплетна апостериорна расподела оцењена за параметре као део поступка закључивања, а не само тачкаста оцена параметара. За сложене моделе, поприлично је тешко добити добре оцене параметара применом класичног приступа. Критика која се упућује Бајесовој статистици односи се на формулацију априорне расподеле. У идеалном случају, априорна расподела треба да се изведе на основу субјективне процене и претходног искуства. У пракси, априорне вероватноће се често бирају из фамилије расподела што чини израчунавање апостериорне расподеле прихватљивијим. Ове фамилије се називају коњуговане фамилије, а априорне расподеле коњуговане априорне расподеле. У таквим случајевима, извођење Бајесовог ажурирања не подразумева комплексну интеграцију: апостериорна расподела је из исте фамилије као и априорна са параметрима који се могу добити из априорних параметара и података. У Табели 4. приказане су неке од коњугованих априорних расподела које су карактеристичне за Бајесову статистику.

Табела 4. Коњуговане априорне расподеле повезане са различитим узорачким расподелама

Узорачка расподела	Коњугована априорна расподела
Биномна	Вероватноћа успеха је бета
Негативна биномна	Вероватноћа успеха је бета
Пуасонова	Средња вредност је гама
Експоненцијална са средином (1/λ)	λ је гама
Нормална расподела са познатом варијансом и непознатом средином	Средња вредност је нормална
Нормална расподела са непознатом варијансом и познатом средином	Варијанса је инверзна гама

Извор: Swiler, L.P. 2006. „Bayesian Methods in Engineering Design Problems“, *Sandia National Laboratories*, Albuquerque, New Mexico, pp. 11

Коњуговане расподеле су повезане са функцијом веродостојности која одговара датом проблему. На тај начин се долази до структурних карактеристика, као што су фактори који су заједнички за веродостојност и условну априорну расподелу. Многе коњуговане фамилије садрже довољно својстава својих чланова тако да се свако априорно уверење из широког спектра расположивих може сасвим адекватно користити за потребе Бајесовог закључивања.

3.3 Значај Бајесове методологије за избор априорне вероватноће

Бајесова методологија подразумева избор априорне расподеле вероватноће и примену Бајесове теореме како би се извела апостериорна расподела вероватноће. Израчунавање апостериорне расподеле вероватноће посматраног модела је захтевно, поготово код сложених модела (Han and Carlin, 2001), али се ове потешкоће успешно могу превазићи (Carlin and Chib, 1995; Chib, 1995; Meng and Wong, 1996). Применом Бајесове теореме могуће је оценити непознате параметре, одредити интервале поверења за непознате параметре и тестирати хипотезе за те параметре. Овако једноставан поступак није могуће спровести применом метода класичне статистике, пошто се код класичног приступа не полази од Бајесове теореме. У том делу се огледа и једна од кључних предности Бајесове над класичном статистиком.

3.3.1 Бајесова теорема као основа Бајесове статистике

У основи Бајесове теореме налази се појам условне вероватноће. У ситуацији када је познато да се догађај B реализовао, поставља се питање како ова информација утиче на закључивање о другом догађају A. На тај начин се долази до дефиниције условне вероватноће.

Дефиниција. Нека је B догађај за који важи $P(B) > 0.$ Тада се вероватноћа догађаја A, под условом да се догађај B већ реализовао, дефинише са:

$$P(A|B) = \frac{P(AB)}{P(B)} \tag{24}$$

и назива се *условна вероватноћа* догађаја A, за дато B. Условна вероватноћа није дефинисана када је $P(B) = 0.$

Слично, када је $P(A) > 0$ добија се:

$$P(B|A) = \frac{P(AB)}{P(A)}.$$

Уколико су A и B два искључива и коначна догађаја, а C је било који други неизвестан догађај, тада је:

$$P(C) = P(A)P(C|A) + P(B)P(C|B) \tag{25}$$

У примени у Бајесовој статистици, једначина (25) подразумева извесна проширења.

Поред условне вероватноће, за дефинисање Бајесове теореме важан је и појам статистичке независности. Статистичка или стохастичка независност догађаја је посебан случај у ком је условна вероватноћа догађаја A за дато B једнака вероватноћи догађаја A, односно $P(A|B) = P(A)$ када је $P(B) > 0$. Ова тврдња суштински значи да реализација догађаја B није утицала на вероватноћу догађаја A.

Дефиниција. У случају када се посматрају два догађаја, за догађаје A и B се каже да су *статистички независни* ако и само ако је

$$P(AB) = P(A)P(B) \tag{26}$$

Условне вероватноће нису дефинисане када је $P(A)$ или $P(B) = 0$, али независност јесте.

Дефинисањем појмова *условне вероватноће* и *статистичке независности догађаја* створила се основа за разјашњење Бајесовог правила, по коме је и названа целокупна теорија која је на њему заснована.

Бајесова теорема дефинисана је пре 250 година, постхумним објављивањем рада влч. Томаса Бајеса „Есеј о решавању проблема у доктрини о шансама“, и као таква представља основу на којој је изграђен читав метод Бајесовог закључивања.

Примена Бајесове теореме захтева пажљиву анализу проблема. Први задатак се огледа у идентификовању догађаја који чине узорачки простор. Након што су догађаји дефинисани, потребно је утврдити способност одабране процедуре да изврши предвиђање на основу датих података и, на крају, изразити једно или више питања у облику који одговара Бајесовој теореми.

Бајесова теорема заправо полази од правила множења вероватноћа. Уколико су A и B два догађаја, тада Бајесова теорема тврди

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

и

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}.$$

Занимљива примена Бајесове теореме је у изразу субјективне вероватноће. Претпоставимо да је појединац заинтересован за догађај B и да формира субјективно мишљење о вероватноћи реализације догађаја B. У том случају, вероватноћа $P(B)$ се назива априорна вероватноћа. Затим, уколико појединац набави додатну информацију – конкретно, да се догађај A остварио – то може довести до промене првобитног става о вероватноћи реализације догађаја B. Пошто је познато да се догађај A реализовао, погодна вероватноћа за B је сада условна вероватноћа догађаја B у односу на догађај A и назива се апостериорна вероватноћа. Посматрано на овај начин, Бајесова теорема би се могла схватити и као механизам за кориговање априорне вероватноће у апостериорну вероватноћу у тренутку када информација о реализацији догађаја A постане расположива.

Теорема 4. Бајесова формула.[64] Нека су A_1, A_2,..., A_n, међусобно искључиви догађаји и нека је B неки други догађај. Вероватноћу догађаја A_i, када је дат догађај B, можемо наћи применом Бајесове теореме

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(B)}, \quad i = 1, 2, \ldots n \tag{27}$$

Уколико се, у складу са формулом потпуне вероватноће, именилац $P(B)$ замени са:

[64] Ивковић, З. 1992. *Математичка статистика*, Научна књига, Београд, стр. 10

$$P(B) = \sum_{k=1}^{n} P(A_k)P(B|A_k),$$

следи да је

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{k=1}^{n} P(A_k)P(B|A_k)}, \ i = 1, 2, ..., n. \quad (28)$$

Приказана на овај начин, Бајесова теорема има широку примену која не изазива никакве контроверзе. Пошто Бајесова теорема има кључну улугу у Бајесовом закључивању, неопходно је проширити њено значење.

Бајесова формула се још назива и формула вероватноћа хипотеза (или узорака) јер A_1, A_2,...,A_n, можемо посматрати као различите узорке који могу довести до реализације догађаја B. Уместо да се посматрају догађаји A_i, посматраће се скуп хипотеза H_1, ..., H_k. Једна и само једна од ових хипотеза може бити тачна. Сада догађај A представља посматрани исход и може се сматрати *подацима из узорка*. Вероватноће $P(H_i)$ представљају *априорне вероватноће* различитих хипотеза и формирају други извор информација. Вероватноће $P(A|H_i)$ (i = 1, 2, ..., k), за посматрани догађај A, представљају веродостојност за податке из узорка.

Бајесову теорему је сада могуће реинтерпретирати. Коригована оцена приказана кроз *апостериорне вероватноће* $P(H_i|A)$ (i = 1, 2, ..., k) за различите хипотезе биће дефинисана са:

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{\sum_{j=1}^{k} P(H_j)P(A|H_j)}, \ i = 1, 2, ..., k. \quad (29)$$

Појмови *априорна* и *апостериорна* односе се на експеримент, пошто се пре извођења експеримента сматрало да су вероватноће хипотеза $P(H_i)$, а након спроведеног експеримента, утврђено је да се реализовао догађај A. Преко Бајесове формуле могу се одредити вероватноће $P(H_i|A)$ реализација појединих хипотеза које су довеле до реализације догађаја A.

Суштина Бајесовог закључивања се изражава кроз став да су неопходна два извора информација, априорне вероватноће и подаци из узорка (приказани у облику веродостојности) како би се на основу њих одредиле апостериорне вероватноће H_i за дато A, што је пропорционално производу априорних вероватноћа H_i и веродостојности за A када је H_i истинито. Именилац у једначини 29. је само константа нормализације која зависи од i, али његово одређивање има одређени значај и може бити проблематично.

На тај начин је априорна информација о конкретној ситуацији проширена подацима из узорка, како би се добио опис вероватноће посматране ситуације. У том погледу, Бајесово закључивање је инференцијално. Тиме се истиче да је постојеће знање у потпуности описано скупом апостериорних вероватноћа, $P(H_i|A)$ (i = 1, 2, ..., k).

Да би се сагледала директна последица једначине 29, којом долази до преправљања основних принципа Бајесовог закључивања, полази се од претпоставке да се посматрају две хипотезе, H_i и H_j. Количник њихових апостериорних вероватноћа, тзв. *апостериорни унакрсни количник*[65] (енгл. *posterior odds ratio*), изражава се са:

$$\frac{P(H_i|A)}{P(H_j|A)} = \frac{P(H_i)}{P(H_j)} \cdot \frac{P(A|H_i)}{P(A|H_j)}, \tag{30}$$

а представља производ *априорног унакрсног количника* (енгл. *prior odds ratio*) и *количника веродостојности* (енгл. *likelihood ratio*).

[65] Barnett, V. 1999. *Comparative Statistical Inference, Third Edition*, John Wiley & Sons Ltd. Chichester, West Sussexs, pp. 204

Једначина 29. представља уобичајен облик Бајесове теореме за дискретне расподеле. Аналогна овој једначини јесте једнакост где се одабира одређени статистички модел како би се утврдио континуиран распон могућих вредности параметра λ:[66]

$$P(\lambda|A) = \frac{P(\lambda)P(A|\lambda)}{\int P(\lambda)P(A|\lambda)d\lambda}. \tag{31}$$

Примена Бајесове теореме се може изразити кроз неколико корака:[67]

1. дефинише се подскуп догађаја на основу датог проблема,
2. одређује се вероватноћа за догађаје дефинисане у првом кораку,
3. израчунавају се потпуне вероватноће,
4. примењује се Бајесова теорема за израчунавање вероватноће за решење проблема.

Оваква процедура је врло корисна за истраживаче који су се већ суочили са сличним проблемима, међутим, већина реалних ситуација не прати овај поступак. Понекад је потребно вратити се на претходни корак и преиспитати почетне дефиниције, док је у неким случајевима корисно написати Бајесову теорему пре одређивања вероватноћа.

3.3.2 Рекурзивна примена Бајесове теореме

Уколико се Бајесова теорема напише у облику:

$$P(A|BC) = \frac{P(A|C)P(B|AC)}{P(B|C)} \tag{32}$$

A би означавало тврдњу о непознатој појави, *B* тврдњу која садржи информацију о непознатој појави, а *C* тврдњу о додатној информацији. $P(A|C)$ се назива *априорна*

[66] Smith, R.L. 1982. „The Bayesian Inference Method and Its Application to Reliability Problems“, *NRL Memorandum Report 4903*, Texas Research Institute, pp. 16

[67] Newbold, P., Carlson L.W. and Thorne, B. 2002. *Statistics for Business and Economics*, Prentice Hall, New Jersey, pp. 120

вероватноћа, $P(A|BC)$ *апостериорна вероватноћа* и $P(B|AC)$ *веродостојност*. Априорна вероватноћа појаве, пре него што су информације прикупљене, коригована је вероватноћом о датим информацијама које се односе на посматрану појаву. То нас је одвело до апостериорне вероватноће непознате појаве под условом да су информације расположиве.

Вероватноћа $P(B|C)$ у имениоцу може бити посматрана и као константа нормалности. Уколико су догађаји A_1, A_2,..., A_n, међусобно искључиви, тада је вероватноћа:

$$P(B|C) = \sum_{j=1}^{n} P(A_k|C)P(B|A_kC) \tag{33}$$

а Бајесова формула поприма следећи облик:

$$P(A_i|BC) = \frac{P(A_i|C)P(B|A_iC)}{\sum_{j=1}^{n} P(A_k|C)P(B|A_kC)} \quad \text{за } i = 1, 2, ..., n \tag{34}$$

Према томе, константа $c = \sum_{j=1}^{n} P(A_k|C)P(B|A_kC)$ се понаша као константа нормалности, јер је $\sum_{i=1}^{n} P(A_i|BC) = 1$, што проистиче из следећег:

Уколико догађаји A_1, A_2,..., A_n нису само међусобно искључиви, већ и потпуни, што значи да их додатна информација C условљава, при чему је само један догађај тачан (уколико је само један тачан, сви остали су нетачни), тада важи

$$P(A_1 + A_2 + ... + A_n|C) = \sum_{i=1}^{n} P(A_i|C) = 1. \tag{35}$$

Веома често се ова константа занемарује, па у том случају Бајесова теорема гласи

$$P(A_i|BC) \propto P(A_i|C)P(B|A_iC) \tag{36}$$

где је са $\propto$ означена пропорционалност.

Отуда следи да је *апостериорна вероватноћа* пропорционална производу *априорне вероватноће* и *веродостојности*.

Уколико је информација о непознатој појави A дата као производ $B_1 B_2 ... B_n$ догађаја $B_1, B_2, ..., B_n$ преко Бајесове формуле (в. 36) утврђујемо:

$$P(A|B_1 B_2 ... B_n C) \propto P(A|C) P(B_1 B_2 ... B_n | AC)$$

и у случају независности догађаја, а на основу:

$$P(A_1 A_2 ... A_n | C) = P(A_1|C) P(A_2|C) ... P(A_n|C)$$

добијамо:

$$P(A|B_1 B_2 ... B_n C) \propto P(A|C) P(B_1|AC) P(B_2|AC) ... P(B_n|AC). \tag{37}$$

Према томе, у случају када су информације независне, Бајесова теорема се може *рекурзивно*[68] применити. Информација B_1 дата са (36) гласи

$$P(A|B_1 C) \propto P(A|C) P(B_1|AC).$$

Ова апостериорна вероватноћа је примењена на априорну вероватноћу како би се анализирала информација B_2 па добијамо

$$P(A|B_1 B_2 C) \propto P(A|B_1 C) P(B_2|AC)$$

У циљу анализирања информације B_k, рекурзивна примена Бајесове теореме би дала

$$P(A|B_1 B_2 ... B_k C) \propto P(A|B_1 B_2 ... B_{k-1} C) P(B_k|AC), \ k \in \{2, ..., n\} \tag{38}$$

[68] Koch, K.R. 2007. *Introduction to Bayesian Statistics, Second Edition*, Springer-Verlag, Berlin, pp. 16

са

$$P(A|B_1C) \propto P(A|C)P(B_1|AC).$$

Овај резултат је сагласан са обрасцем 37. Анализирањем информација од B_1 до B_n стање знања A о непознатој појави је успешно кориговано, што је еквивалентно процесу учења уз коришћење додатних информација.

3.4 Априорне информације као извор података

Код Бајесовог метода оцењивања параметри се интерпретирају као случајне променљиве које имају одређену расподелу $\pi = \pi(\theta)$, која се назива априорна расподела. Расподела $\pi(\theta)$ изражава степен личног (субјективног) убеђења статистичара о расподели случајне променљиве и не може се емпиријски проверити, јер се једино могу регистровати вредности обележја X, а не и параметра θ.[69]

Кључна питања важна за дефинисање априорне расподеле односе се на:

- информације које ће бити саставни део априорне расподеле и
- особине настале апостериорне расподеле.

Уз добро идентификоване параметре и велики узорак, избор априорне расподеле имаће мање ефекте на апостериорне закључке. Уколико је узорак мали или расположиви подаци пружају само индиректне информације о посматраним параметрима, априорна расподела постаје значајнија.[70]

[69] Петровић, Љ. 2006. *Теоријска статистика. Теорија статистичког закључивања*. Економски факултет, Београд, стр. 129

[70] Gelman, A., Carlin, B. J., Stern, S.H. and Rubin, B.D. 2004. *Bayesian Data Analysis, Second Edition*, Chapman&Hall/CRC, pp. 1

Бајесово закључивање омогућава да априорно знање утиче на изведене закључке. У том смислу, априорне информације нису само од изузетне користи због укључивања постојећег знања, већ су и важан предуслов за рационално закључивање. Бергер сматра да „уколико различите априорне информације доведу до суштински различитих одговора, не може се тврдити да постоји само један одговор. Тада би било исправније признати да постоји научна неизвесност, са закључком који ће зависити од априорних уверења“.[71] Априорне информације су корисне, неопходне и информативне и сматрају се саставним делом рационалног закључивања.[72]

3.4.1 Основа проблема примене априорних информација

Од времена Лапласа, примена теорије вероватноће наилазила је на потешкоће, пре свега због проблема који су се јављали у вези схватања примене априорних информација. У реалним проблемима одлучивања или закључивања, често се располаже априорним информацијама које су од изузетног значаја за посматрани проблем. Самим тим, не узимање у обзир априорне информације могло би довести до апсурдних или погрешних резултата. Током дугог низа година, класична школа мишљења, коју заступа већина статистичара, потпуно је одбацивала коришћење априорне вероватноће, осим у случају када се априорне информације састоје од фреквентних података, што у пракси није чест случај.[73]

Са развојем теорије одлучивања, овај проблем је добио нови значај. Као што је познато, развој теорије одлучивања започет је захваљујући Валду са циљем да се пронађе нова статистичка основа која би подразумевала општост, али и избегавање претпостављене грешке. Међутим, врло брзо се установило да је једина конзистентна процедура

[71] Berger, O.J. 1980. *Statistical Decision Theory and Bayesian Analysis,* Springer-Verlag New York Inc., pp. 125

[72] Wagenmakers, E.J., Lee, M.D., Lodewyckx, T. and Iverson, G. 2008. „Bayesian versus frequentist inference“, In H. Hoijtink, I. Klugkist, & P. A. Boelen (Eds.). *Bayesian Evaluation of Informative Hypotheses*, 181-207. Springer, New York., pp. 22

[73] Jaynes, T.E. 1968. „Prior Probabilities“, *IEEE Transactions On Systems Science and Cybernetics*, Vol. sec-4, No. 3, pp. 228

укључивања информација у процес одлучивања идентична примени Бајесове теореме и да је, у случају када су дате функција губитка, узорачка расподела и узорак, једина рационална основа за избор међу прихватљивим одлукама априорна вероватноћа.

Тако је у модерној теорији одлучивања статистичка пракса достигла ниво где се проблем априорне вероватноће више не може игнорисати нити занемаривати. У савременим проблемима пројектовања, контроле квалитета, операционих истраживања, итд., читав проблем се не може превести у математичком смислу све док се не научи како да се пронађе априорна вероватноћа која описује априорну информацију. У ствари, у неким од најважнијих проблема, априорна информација је једина расположива информација, тако да и одлуке морају бити у потпуности засноване на њој. У одсуству било ког принципа за дефинисање априорне расподеле, такви проблеми се не могу математички посматрати.

Априорне информације су контроверзан аспект Бајесових метода. Бајесове методе захтевају да се расположива априорна информација о θ квантитативно изрази, без обзира да ли садржи једну компоненту или је мултидимензионална, да ли је информација специфична или обимна или, у крајњем случају, у суштини не постоји. У том смислу, прави се разлика између *одсуства априорних информација*, *значајног априорног знања* и, међукатегорије, *нејасног априорног знања*.[74] У раду ће бити објашњена и категорија *коњугованих априорних информација*.

3.4.2 Одсуство априорних информација или априорно незнање

Бајесове статистичке процедуре код великих узорака дају одговоре који су веома слични одговорима до којих се долази применом класичне статистике. Из тог разлога, многи статистичари сматрају да разлика између Бајесових и класичних метода и није толико важна у пракси. Ипак, код малих узорака, примењене процедуре доводе до различитих одговора, где Бајесово решење зависи од прихваћених априорних информација које се могу посматрати и као предност и као недостатак, у зависности од тога колико се сматра

[74] Barnett, V. 1999. *Comparative Statistical Inference, Third Edition*, John Wiley & Sons Ltd. Chichester, West Sussexs, pp. 219

да априорна густина репрезентује стварну априорну информацију и циљ истраживања.[75] Априорна густина $p(\theta)$ не зависи од података, већ садржи све расположиве информације о θ које нису повезане са подацима. Другим речима, априорна густина сумира сва априорна знања о θ, пре него што су прикупљени подаци.[76]

Да би се Бајесове методе могле применити у закључивању неопходно је квантитативно изразити априорно знање, односно, потребна је нумеричка спецификација стања *априорног незнања* или, другачије речено, *одсуства априорних информација*. Појам априорног незнања једно је од најспорнијих питања у Бајесовој статистици.

Стање потпуног незнања није расподела вероватноће. У том случају се сваком могућем исходу додељује исти степен уверења да априорне информације не постоје.[77]

Питањем априорног незнања посебно се бавио Џефрис који је, како је већ поменуто, у својој књизи „Теорија вероватноће“ извршио анализу примене Бајесових метода. Између осталог, осврнуо се на Бајес-Лапласов принцип недовољног разлога, по ком се сматра да уколико не постоји разлог да се више верује остварењу једне хипотезе него друге, вероватноће ће бити једнаке. Џефрис сматра да правило по ком су вероватноће остварења хипотеза једнаке није израз неког уверења, нити закључак изведен из претходног искуства, већ је само формалан начин за изражавање незнања.[78]

Џефрис је знатно проширио овај принцип од априорног незнања о дискретном скупу хипотеза до случаја априорног незнања о стално промењивом параметру θ у параметарском простору Ω. Очигледно проширење за једнодименезионални параметар је додељивање једнаких априорних вероватноћа густине за свако $\theta \in \Omega$. Отуда, за локацију параметра, μ, где је параметарски простор $(-\infty, \infty)$, одабраће се да $\pi(\mu)$ буде константно.

[75] Young, G.A. and Smith, R.L. 2005. *Essentials of statistical inference*, Cambridge University Press, Cambridge, pp. 24

[76] Koop, G. 2003. *Bayesian Econometrics*, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, England, pp. 3

[77] Norton, J.D. 2008. „Ignorance and Indifference“, *Philosophy of Science* 75, pp. 45

[78] Jeffreys, H. 1961. *Theory of Probability*, *Third Edition*, Clarendon Press, Oxford, pp. 33-34

Уколико је параметар θ мултидимензионалан, априорно незнање треба изразити у виду претпоставке о независности комопоненти параметра θ и могућности избора априорних расподела за сваку од њих. Отуда, за параметар θ пар (μ,σ) одражава локацију и скалу, где се за μ може доделити униформна расподела $(-\infty, \infty)$, док је за σ (које треба да буде позитивно) потребно одабрати неки други облик априорне расподеле.

Како би олакшао избор одговарајуће расподеле, Џефрис предлаже следеће принципе:[79]

- Када је $\Omega = (-\infty, \infty)$, треба користити униформну априорну расподелу за θ. Подршку за овај избор проналази у указивању да уколико, по овој основи, постоји априорно незнање о θ, исто ће важити и за било коју линеарну функцију о θ.

- Када је $\Omega = (0, \infty)$, треба користити априорну расподелу која је сразмерна $1/\theta$, јер то значи да на сличан начин постоји незнање о било ком θ^{α} $(\alpha \neq 0)$.

Априорно незнање такође може бити представљено и преко одговарајуће стандардне расподеле са великом дисперзијом; реч је о такозваној равномерној или дифузној расподели.

Дефинисањем априорног незнања ствара се могућност примене Бајесових метода у случајевима када не постоји априорно знање о θ. Осим тога, применом Бајесовог приступа изражава се повезаност између података и параметра θ, без обзира на априорну информацију о θ. Отуда, ова априорна информација се поставља на ниво незнања или неинформативности, како би подаци говорили сами за себе.

Савремене присталице Бајесове статистике и даље, добрим делом, поштују и примењују Џефрисове предлоге уз претпоставку о независности различитих компоненти θ. Наравно, питање које је увек отворено јесте: како наставити даље уколико се ништа унапред не зна о параметру θ? Можда ово питање не би ни требало да буде од суштинског значаја, пошто

[79] Barnett, V. 1999. *Comparative Statistical Inference*, *Third Edition*, John Wiley & Sons Ltd. Chichester, West Sussexs, pp. 221

је чињеница да се неретко дешава да уплив података измени априорне информације које тиме постају безначајне у посматраном контексту.

3.4.3 Неинформативне априорне информације

У многим случајевима, априорна уверења могу бити нејасна и, отуда, тешка за превођење у информативну априорну информацију. Из тог разлога, постоји потреба да се неизвесност о параметрима модела изрази без битног утицаја на апостериорно закључивање о параметру. У том смислу користе се *неинформативне* априорне информације, које се још у литератури срећу и под називом *нејасне* или *дифузне* априорне информације.[80]

Често се дешава да не постоји довољно информација да би се изабрала априорна густина за непознате параметре или се расположива информација тешко преводи у израз вероватноће; ипак, потребна је априорна информација да би се применила Бајесова теорема. У овом случају можемо покушати да пронађемо неинформативне априорне информације које ће омогућити да се изведе Бајесов закључак уз минимално увођење вештачких информација. Међутим, питање је како пронаћи неинформативну априорну информацију? Једна од могућности је дефинисање мере информација, нпр. ентропије и одређивање априорне информације која минимизира/максимизира ову меру нпр. максималну ентропију.[81]

Ентропија је мера неизвесности. Принцип максималне ентропије се примењује за извођење априорне функције густине која садржи априорне информације о непознатим параметрима, али изван тога они су неизвесни колико је то могуће.[82] Извођењем експеримента или мерења за добијање информација о непознатој појави, уклања се неизвесност која је постојала пре експеримента или вршења мерења. Неизвесност која је

[80] Rachev, T.S., Hsu, J.S.J., Bagasheva, B.S. and Fabozzi, F.J. 2008. *Bayesian Methods in Finance*, John Wiley & Sons, Inc., New Jersey, pp. 25

[81] Scales, J.A. and Tenorio, L. 2001. „Tutorial: Prior information and uncertainty in inverse problems“, *Geophysics*, Vol. 66, No. 2, pp. 393

[82] Koch, K.R. 2007. *Introduction to Bayesian Statistics, Second Edition*, Springer-Verlag, Berlin, pp. 57

елиминисана извођењем експеримента одговара информацијама до којих се дошло извршењем експеримента.

Основни циљ Бајесовог закључивања је да, преко апостериорне расподеле за θ, изрази удружене информације добијене из априорне расподеле и података из узорка. Априорна информација доводи до повећања вредности, тако да апостериорна расподела више одступа од веродостојности. По том основу, може се очекивати да ће одсуство априорних информација довести до тога да апостериорна расподела буде управо сразмерна веродостојности, што се и дешава уколико се примењује априорна униформна расподела. У случајевима када постоје одређене априорне информације које се мењају под утицајем информација из узорка, апостериорна расподела се поново своди на нормализовану веродостојност. У случајевима овог типа, може се говорити да постоји *нејасно априорно знање* о θ.

У том смислу, неинформативна априорна информација могла би се користити у одсуству априорног информисања или у случају када постоји неслагање о могућим вредностима хипотеза или параметара. Такође се може користити у поређењу са више информативним априорним информацијама, као један аспект анализе осетљивости апостериорних закључака у односу на априорне информације.[83]

Најједноставније речено, априорна информација је нејасна, слаба или неинформативна уколико је априорна расподела константна за скуп вредности θ, где веродостојност није занемарљива вредност за прикупљене податке. Отуда су закључци о θ независни од априорне расподеле и изведени су на основу апостериорне расподеле, која је заправо норамализована верзија функције веродостојности.

Неинформативне информације у литератури се срећу и под називом *објективне априорне информације* и као такве су део објективне Бајесове анализе. Примена објективних,

[83] Congdon, P. 2001. *Bayesian Statistical Modelling*, John Wiley and Sons Ltd., pp. 4

односно неинформативних априорних информација у Бајесовој анализи подлеже критици, која се може представити на следећи начин:[84]

1. Неинформативне априорне информације су информације које не постоје. Како их је онда могуће дефинисати?
2. Објективна Бајесова анализа је ad hoc, па отуда није ништа боља од ad hoc субјективне Бајесове анализе коју покушава да замени.
3. Боље је покушати са применом априорних информација, него покушавати да се пронађу неинформативне априорне информације.
4. Постоји велики број неинформативних априорних информација за посматрани проблем; за коју се определити?
5. Неинформативне априорне информације су обично неприкладне, а неприкладне информације обично немају толико смисла као квантитативно изражавање уверења.
6. Уколико θ има униформну расподелу због недостатка информација, онда би ово требало да важи за било коју један на један функцију.
7. Зашто би неинформативна апирорна информација зависила од модела података?
8. Какав је утицај претходних ставова на кохеренцију и принцип веродостојности?

Основни смисао увођења објективне априорне информације јесте да се одреди апостериорна расподела која ће више зависити од података него од објективне информације. Осим тога, према објективној Бајесовој анализи, немогуће је дефинисати неинформативне априорне информације на неограниченом параметарском простору, јер максимална ентропија не мора да буде коначна.

[84] Ghosh, J.K, Delampady, M. and Samanta, T. 2006. *An Introduction to Bayesian Analysis, Theory and Methods*, Springer Science+Business Media, LLC, New York, pp. 147

3.4.4 Информативне априорне информације као значајне информације

Трећа могућност, када је реч о априорним информацијама, односи се на случај када су априорне информације *значајне*, у смислу да доводе до тога да апостериорна расподела значајно одступа од функције веродостојности. Овакве априорне информације још се називају и *информативне* априорне информације. Као и код појма нејасног априорног знања, и овде је реч о вези између количине информација из података и обиму априорног знања, а не само о одразу искључиво априорног знања. Када се априорна информација квантитативно изрази преко априорне расподеле, не постоје формална ограничења за примену Бајесових метода, па ће закључци бити засновани на апостериорној расподели.

Априорна уверења су информативна када значајно мењају информације садржане у подацима из узорка, тако да се закључци о параметрима модела изводе на основу апостериорне расподеле и на основу расподеле података различити. Најчешћи приступ који се користи за представљање информативних априорних уверења је избор расподеле са непознатим параметром и одређивање параметара који ће одражавати ова уверења.

Основна препорука за коришћење информативних априорних информација јесте то што представљају најбоље додатне информације о посматраној појави. Међутим, проблем са коришћењем ових информација је у томе што различити људи користе различите додатне информације или их тумаче на другачији начин, тако да се често чини да су информативне априорне информације субјективне.[85]

Применом информативних априорних информација у извођењу Бајесових метода, истичу се три важна фактора:[86]

- Коначан закључак значајно зависи од претпостављеног облика априорне расподеле. Тиме се поново наглашава колико је важно да априорна расподела буде тачна и да

[85] Downey, A.B. 2012. *Think Bayes, Bayesian Statistics Made Simple*, *Version* 1.0., Green Tea Press, Needham, Massachusetts, pp. 30

[86] Barnett, V. 1999. *Comparative Statistical Inference*, *Third Edition*, John Wiley & Sons Ltd. Chichester, West Sussexs, pp. 229

у потпуности одражава расположиве априорне информације. Питање квалитета и валидности априорне расподеле срж је примене и критике Бајесовог приступа.

- Математичко одређивање апостериорне расподеле од изузетне је важности.
- Неопходно је тумачење апостериорне расподеле у односу на априорну расподелу.

Како би се обезбедила реализација другог и трећег фактора у одговарајућим околностима, потребно је испитати кључни концепт коњугованих фамилија априорних расподела које олакшавају математичке прорачуне и обезбеђују поређење значајности података из узорка и априорних информација.

3.4.5 Коњуговане априорне информације

Главни технички проблем у Бајесовој анализи је проналажење апостериорне расподеле на основу веродостојности и априорне расподеле. Апостериорна расподела је сразмерна производу веродостојности и априорне расподеле вероватноће. У случају када веродостојност само коригује априорне параметре и не мења њихову функционалну форму, за такве априорне параметре кажемо да су коњуговани са веродостојношћу. Коњугација је популарна због математичке погодности; једном када се препозна коњуговани пар веродостојност/априорна расподела, апостериорна расподела се може одредити без интеграције.

Априорне информације одражавају све информације којима истраживач располаже пре него што добије податке које жели да укључи у истраживање. Отуда, априорне информације могу имати било какав облик. Међутим, уобичајено је да се одаберу одређење класе априорних информација које се лако тумаче и/или олакшавају израчунавања. Коњуговане априорне информације обично имају обе предности.

Коњугована априорна расподела је таква да, у комбинацији са веродостојношћу, доноси апостериорну расподелу која припада истој класи расподела. Коњугована априорна расподела има исту функционалну форму као и функција веродостојности, што заправо значи да се априорна информација може тумачити на исти начин као и информација функције веродостојности. Другим речима, може се рећи да априорна информација проистиче из скупа замишљених података који је део истог процеса који генерише стварне податке.[87]

Да би се извео закључак о параметру θ, који је показатељ фамилије расподела $\mathscr{P} = \{p_\theta(x); \theta \in \Omega\}$, полази се од претпоставке да је априорна расподела за θ члан неке параметарске фамилије расподела, P, са особином, у односу на $\mathscr{P}$, па је и апостериорна расподела за θ такође члан P. У том случају се каже да је P фамилија *коњугованих априорних расподела* у односу на $\mathscr{P}$.

Према дефиницији[88] класа априорних расподела P за θ, назива се *коњугована* за узорачки модел $p(y|\theta)$ ако је

$$p(\theta) \in \mathsf{P} \Rightarrow p(\theta|y) \in \mathsf{P}. \tag{39}$$

Коњугована априорна расподела олакшава израчунавање апостериорне расподеле, али можда није одраз стварних априорних информација. Ипак, мешавине коњугованих априорних расподела су веома флексибилне и рачунски прилагодљиве.

Пре развоја компјутера, коњуговане априорне информације су се интензивно користиле, пре свега због наведене рачунске погодности. Данас је присутна општа сагласност да једноставна коњугована анализа има ограничену практичну вредност, с обзиром да јој је, у односу на веродостојност, способност моделовања ограничена.

На крају, сагледавајући важност априорне информације у Бајесовој статистици, основни приступи избору априорне расподеле могу се свести на:[89]

[87] Koop, G. 2003. *Bayesian Econometrics*, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, England, pp. 18

[88] Hoff, D.P. 2009. *A First Course in Bayesian Statistical Methods*, Springer Science + Business Media LLC, New York, pp. 38

- физичко резоновање (Бајес) – сматра се сувише рестриктивним у практичној примени;
- равномерна или униформна расподела, укључујући неприкладне априорне информације (Лаплас, Џефрис) - најраспрострањенији метод у пракси, али је теоријско оправдање за овакав приступ и даље извор аргумената;
- субјективне априорне информације (де Финети, Саваж) - користе се у одређеним ситуацијама, као што су прогноза времена (чак и ако не постоји, априорна информација тежи да буде део формалне Бајесове анализе са веродостојношћу и апостериорном расподелом) и за одређене врсте пословних апликација, где је априорна информација веома важна, па се сматра да је изузетно значајно од клијента добити право субјективно мишљење; ипак, тешко се користи за рутинске статистичке анализе;
- коњуговане априорне расподеле - у пракси се најчешће користе само за једноставна израчунавања.

3.5 Поступак одређивања апостериорне расподеле

Априорна расподела, која изражава априорно знање о параметру θ са функцијом густине $\pi(\theta)$, кључни је део Бајесовог закључивања преко које се, у комбинацији са расподелом вероватноће нових података, долази до *апостериорне расподеле*. Уколико се апостериорна функција густине вероватноће за параметар θ означи са $\pi(\theta|x)$, применом Бајесове теореме добија се:[90]

[89] Young, G.A. and Smith, R.L. 2005. *Essentials of statistical inference*, Cambridge University Press, Cambridge, pp. 42

[90] Barnett, V. 1999. *Comparative Statistical Inference*, *Third Edition*, John Wiley & Sons Ltd. Chichester, West Sussexs, pp. 209

$$\pi(\theta|x) = \frac{p_\theta(x)\pi(\theta)}{\int_\Omega p_\theta(x)\pi(\theta)}. \tag{40}$$

Апостериорна расподела, $\pi(\theta|x)$, сматра се потпуним описом онога што је познато о параметру θ на основу априорних информација и података из узорка и као таква представља директнији облик закључивања од оног који нуди класична статистика.

Апостериорна расподела, која се представља као кориговaна расподела вероватноће над параметарским простором, користи се за израчунавање расподеле вероватноће за будуће догађаје, а на основу претходног искуства. За разлику од класичног приступа, ово је прецизна расподела, али садржи субјективни елеменат који се изражава преко априорне расподеле. У Бајесовој статистици, израчунавање апостериорне расподеле обично захтева примену одговарајућих нумеричких метода, међу којима су се Монте Карло методе показале као најефикасније. Ова техника је фреквентистичка, у смислу да се ослања на произвољно велику понуду независних случајних бројева како би се добила жељена прецизност.[91]

Фамилија метода Марковљеви ланци Монте Карло (енгл. *Markov chain Monte Carlo* (MCMC)) често је класификована као Бајесово израчунавање или Бајесова апостериорна симулација, јер њене методе имају Бајесове принципе.[92] Принцип MCMC метода је једноставан: случајним одабиром узорка из циљне расподеле вероватноће дизајнира се Марковљев ланац чија је стационарна расподела циљна расподела. Дугом симулацијом таквог ланца, циљна расподела се може добро проценити.

Код примене овог приступа, тежи се оцењивању апостериорне расподеле $\pi(\theta|x)$, макар у смислу маргиналних компоненти $\pi(\theta_1|x), \pi(\theta_2|x), \ldots \pi(\theta_k|x)$. Након тога се изграђује Марковљев ланац са простором Ω, на основу ког се лако врши симулација и који садржи

[91] Koski, T. and Noble, M.J. 2009. *Bayesian Networks, An Introduction*, John Wiley & Sons, Ltd., pp. 18

[92] Васић, В. 2003. „Фамилија метода Марковљеви ланци Монте Карло у анализи некомплетних података“, *Економски анали*, бр. 159, стр. 147

наведене компоненте као равнотежну расподелу. Потом, уколико је узорак узет из овог ланца након што је већ покренут, успешно ће се вршити узорковање из посматране апостериорне расподеле $\pi(\theta|x)$.

За изградњу одговарајућег Марковљевог ланца користе се различите стратегије: Метрополис-Хејстингс, Гибсово узорковање, узорковање дела, савршено узорковање, као и многе специјализоване технике, а у наставку ће бити детаљеније објашњене две које су се показале као најуспешније у процењивању значајних карактеристика за $\pi(\theta|x)$.

3.5.1 Алгоритам за Гибсово узорковање као средство за изградњу Марковљевог ланца

Назив *Гибсов алгоритам* или *Гибсов узорак* проистиче из аналогије статистичке филозофије, где се слични методи користе за генерисање стања из Гибсове расподеле. Гибсов узорак је посебан случај Метрополис-Хејстингс узорка код ког се насумична вредност увек прихвата. Основни задатак се односи на изградњу Марковљевог ланца чије ће вредности конвергирати ка циљној дистрибуцији.

У Гибсовом узорку, само један параметар се мења током времена, док су сви остали фиксиране вредности. Параметар је случајно узет из условне функције густине вероватноће, расподеле вероватноће једног параметра, за дате остале параметре, $\pi\,(\theta_i|x;\, \theta_j,\, j\neq i)$. Такве условне расподеле су далеко лакше за симулацију него сложене заједничке расподеле и обично имају једноставне форме (често су нормална, инверзна χ^2 или нека друга уобичајена априорна расподела). Дакле, врши се симулација n узастопних случајних променљивих на n основу униваријантних условних расподела, пре него генерисање једног n-димензионалног вектора користећи тоталну заједничку расподелу.[93]

Нека је $\theta=(\theta_1,\theta_2,...,\theta_n)$ и $\pi(\theta|x)=\pi(\theta_1,\theta_2,...,\theta_k|x)$. Поступак итерације започиње са одабраним вредностима $\theta_1^{(0)},\theta_2^{(0)},...,\theta_k^{(0)}$ уз ажурирање θ_i у свакој фази. На тај начин се

[93] Walsh, B. 2004. „Markov Chain Monte Carlo and Gibbs Sampling“, *Lecture Notes for EEB* 581, *Version* 26, pp. 16

извлачи $\theta_1^{(1)}$ из $\pi(\theta|x;\theta_2^{(0)},...,\theta_n^{(0)})$, $\theta_2^{(1)}$ из $\pi(\theta_2|x;\theta_1^{(1)},\theta_3^{(0)},\theta_4^{(0)},...,\theta_k^{(0)})$, $\theta_3^{(1)}$ из $\pi(\theta_3|x;\theta_1^{(1)},\theta_2^{(1)},\theta_4^{(0)},...,\theta_k^{(0)})$ и тако даље. Уколико се овај процес настави кроз, на пример, N итерација и понови m пута имаћемо m понављања $\theta^{(N)}=(\theta_1^{(N)},\theta_2^{(N)},...,\theta_k^{(N)})$. Отуда, за велико N, поновљене вредности $\theta_{ij}^{(N)}$ $(j=1,2,...,n)$ представљају случајан узорак из $\pi(\theta_i|x)$ $(i=1,2,...,k)$. Процес померања из $\theta^{(0)}$ у $\theta^{(1)}$ и даље у $\theta^{(N)}$ је процес Марковљевог ланца.

Основна претпоставка по којој ради Гибсов алгоритам јесте стварање могућности за лако извлачење случајног узорка из условних функција густине вероватноће. Отуда је Гибсов алгоритам посебан случај Метрополис-Хејстингс алгоритма.

3.5.2 Метрополис-Хејстингс алгоритам у Бајесовој анализи

Један од проблема примене Монте Карло интеграције јесте у прибављању узорка из неке сложене расподеле вероватноће $p(x)$. Покушаји да се реши овај проблем су корени MCMC метода. Конкретно, ови покушаји се односе на покушаје математичких физичара да интегришу веома сложене функције методом случајног узорка и Метрополис-Хејстингс алгоритмом.

Метрополис-Хејстингс алгоритам користи другачије итеративне шеме у односу на Гибсово узорковање, где се следећа итерација или прихвата или одбацује у складу са прописаним механизмом вероватноће. Овај алгоритам симулира узорке из расподеле вероватноће тако што користи тоталну заједничку функцију густине и независне предложене расподеле за сваку од променљивих понаособ. Да би се дошло до пуне формулације Метрополис-Хејстингс алгоритма, неопходно је поћи од дефинисања Метрополис алгоритма.

Пођимо од претпоставке да је циљ извлачење узорака из неке расподеле $p(\theta)$ где је $p(\theta)$ = $f(\theta)/K$, где константа нормализације K не мора бити позната, а истовремено може бити тешка за израчунавање. Метрополис алгоритам генерише низ извлачења из ове расподеле на следећи начин:[94]

1. Почиње се са било којом вредношћу θ_0 која задовољава $f(\theta_0) > 0$.

2. Коришћењем текуће вредности за θ, узима се „кандидат" узорак θ^* из неке расподеле $q(\theta_1, \theta_2)$ који представља вероватноћу враћања вредности θ_2 за дату претходну вредност θ_1. Ова расподеле је позната и као *предложена* или *кандидат* расподела. Једино ограничење расподеле $q(\theta_1, \theta_2)$ у Метрополис алгоритму је у томе што је симетрична, тј. $q(\theta_1, \theta_2) = q(\theta_2, \theta_1)$.

3. За кандидата θ^* потребно је израчунати количник густине (θ^*) и текућих вредности (θ_{t-1}):
$$\alpha = \frac{p(\theta^*)}{p(\theta_{t-1})} = \frac{f(\theta^*)}{f(\theta_{t-1})}. \qquad (41)$$
Пошто се количник $p(x)$ разматра према двема различитим вредностима, константа нормализације K се поништава.

4. Уколико се повећа густина $(\alpha > 1)$, прихвата се вредност кандидата $(\theta_t = \theta^*)$ и враћа се на други корак. Уколико се смањи густина $(\alpha < 1)$, онда се са вероватноћом α прихвата вредност кандидата, а у осталим случајевима одбацује и враћа се на други корак.

Метрополис узорковање може се сумирати на следећи начин:

$$\alpha = \min\left(\frac{f(\theta^*)}{f(\theta_{t-1})}, 1 \right) \qquad (42)$$

[94] Walsh, B. 2004. „Markov Chain Monte Carlo and Gibbs Sampling", *Lecture Notes for EEB* 581, *Version* 26, pp. 7

уз прихватање вредности кандидата са вероватноћом α. На тај начин се генерише Марковљев ланац $(\theta_0, \theta_1, ..., \theta_k, ...)$, као вероватноће прелаза од θ_t до θ_{t+1} који зависи само од θ_t, а не од $(\theta_0, ..., \theta_{t-1})$. Пролазећи кроз довољан број корака, ланац се приближава својој расподели, а узорци из вектора $(\theta_{k+1}, ..., \theta_{k+n})$ су узорци из $p(x)$.

Хејстингс (W. Keith Hastings) је 1970. године[95] уопштио Метрополис алгоритам применом функције вероватноће прелаза $q(\theta_1, \theta_2) = \Pr(\theta_1 \to \theta_2)$ и стављањем вероватноће прихватања за вредност кандидата као:

$$\alpha = \min\left(\frac{f(\theta^*)\, q(\theta^*, \theta_{t-1})}{f(\theta_{t-1})\, q(\theta_{t-1}, \theta^*)}, 1 \right). \tag{43}$$

На тај начин добија се *Метрополис-Хејстингс алгоритам*.

Метрополис-Хејстингс алгоритам и алгоритам за Гибсово узорковање могу се користити у Бајесовој анализи за регресионе проблеме, у проучавању комбинованих линеарних модела, итд.

3.6 Улога емпиријског Бајесовог приступа у анализи података

Како је већ споменуто, Бајесова анализа зависи од априорне расподеле параметара посматраног модела. Низ параметара и априорних расподела сачињава хијерархијски модел. У неком тренутку, хијерархија мора да стане заједно са свим преосталим априорним параметрима за које се претпоставља да су познати. Уместо да се бави претпоставкама, *емпиријски Бајесов приступ* користи прикупљене податке како би оценио параметре у завршној фази, а затим даље наставио као да су априорни параметри били познати.[96] Априорне расподеле могу бити параметарске и непараметарске, у зависности од

[95] Hastings, W.K. 1970. „Monte Carlo sampling methods using Markov chains and their applications“, *Biometrika*, Vol. 57, No. 1, pp. 100

[96] Carlin, B.P. and Louis, T.A. 2009. *Bayesian Methods for Data Analysis, Third Edition*, Chapman & Hall/CRC, pp. 225

самих непознатих параметара, па је уобичајено да се прави разлика између параметарске (ПЕБ) и непараметарске емпиријске Бајесове анализе (НЕБ). За емпиријску Бајесову анализу карактеристична је процена априорних вредности параметара из маргиналних расподела података. Пошто се процењују априорне вредности параметара, може се применити уобичајен поступак, уколико су ове вредности фиксиране на почетку.[97]

Емпиријске Бајесове методе могу се користити са било којим моделом и посебно су популарне код хијерархијских априорних модела, међутим, ове методе су и критиковане због тога што рачунају податке два пута. Подаци се прво користе за избор априорних хиперпараметарских вредности, а након што се одаберу ове вредности, подаци се користе други пут у стандардној Бајесовој анализи.

Емпиријским Бајесовим методама врши се оцењивање априорних хиперпараметара из података, а не субјективан избор вредности или њихово коришћење као неинформативних вредности. За избор априорних хиперпараметара може се израчунати маргинална вероватноћа, па ће се у емпиријској Бајесовој анализи користити они хиперпараметри који достигну највећу граничну вероватноћу. Међутим, претраживање свих могућих априорних хиперпараметара може бити веома захтевно, па се, сходно томе, емпиријски Бајесови методи често користе за један или два кључна априорна хиперпараметра.[98]

У емпиријском Бајесовом приступу оцена хиперпараметара априорне расподеле из текућег скупа података врши се уз помоћ метода максималне веродостојности или метода узорка. Израчунавање параметара априорне дистрибуције из тренутног скупа података и коришћење ових податка за израчунавање вероватноће у Бајесовој једначини крши теорему, пошто априорна расподела треба да зависи само од параметара, а не од података.

---

[97] Young, G.A. and Smith, R.L. 2005. *Essentials of statistical inference*, Cambridge University Press, Cambridge, pp. 38

[98] Koop, G. 2003. *Bayesian Econometrics*, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, England, pp. 188

Као резултат, могу се добити некохерентне оцене, па се проблем у већини случајева превазилази применом MCMC метода за процену параметара.[99]

Иако се принципи емпиријског Бајесовог приступа могу користити за хијерархијске моделе било ког броја нивоа, ради једноставности може се приказати стандардни двоетапни модел.

Полази се од претпостављене веродостојности $f(y|\theta)$ за посматране податке y с обзиром на вектор непознатих параметара θ и апирорну вероватноћу за θ са условном функцијом густине $G(\theta)$ и одговарајућом густином $g(\theta|\eta)$, где је η вектор хиперпараметара. Са познатим η користи се Бајесово правило како би се израчунала апостериорна расподела:[100]

$$p(\theta|y,\eta) = \frac{f(y|\theta)g(\theta|\eta)}{m(y|\eta)}, \tag{44}$$

где $m(y|\eta)$ означава маргиналну расподелу за y,

$$m(y|\eta) = \int f(y|\theta)g(\theta|\eta)d\theta, \tag{45}$$

или, општије, $m_G(y) = \int f(y|\theta)dG\theta.$

Уколико је η непознато, Бајесов приступ ће прихватити хипераприорну расподелу $h(\eta)$ и израчунати апостерирону расподелу:

[99] Swiler, L.P. 2006. „Bayesian Methods in Engineering Design Problems“, *Sandia National Laboratories*, Albuquerque, New Mexico, pp. 9

[100] Carlin, B.P. and Louis, T.A. 2009. *Bayesian Methods for Data Analysis, Third Edition*, Chapman & Hall/CRC, pp. 225

$$p(\theta|y) = \frac{f(y|\theta)g(\theta|\eta)h(\eta)d\eta}{\int\int f(y|u)g(u|\eta)h(\eta)dud\eta} = \int p(\theta|y,\eta)h(\eta|y)d\eta. \tag{46}$$

Из наведеног се може видети да је апостериорна расподела мешавина условне апостериорне расподеле (44) за дату фиксирану вредност η и маргиналне апостериорне расподеле за η.

У емпиријској Бајесовој анализи користи се маргинална расподела (45) за оцену η преко $\hat{\eta} \equiv \hat{\eta}(y)$. У том случају ће закључивање бити засновано на оцењеној апостериорној расподели $p(\theta|y,\hat{\eta})$. Назив „емпиријска Бајесова анализа“ проистиче из чињенице да се користе подаци за оцену хиперпараметра η.

На овај начин је заправо описан *параметарски емпиријски Бајесов приступ* (ПЕБ). Дакле, претпостављамо да расподела $g(\theta|\eta)$ узима параметарски облик, тако да је потребно само изабрати вредности за η како би се потпуно одредила оцењена апостериорна расподела. Због веће флексибилности, може се спровести *непараметарски емпиријски Бајесов приступ* (НЕБ) код ког $G(\theta)$ има непознату форму. Овај метод прво покушава да представи апостериорну средину у смислу маргиналне расподеле, а затим користи податке како би се директно оценило Бајесово правило.

Резултати Бајесове емпиријске анализе имају карактеристике које су аналогне фреквентистичким тачкастим оценама и интервалима поверења. Уобичајена тачкаста оцена је средина апостериорне густине или *апостериорна средина*.

Нека је θ вектор са k елемената, $\theta = (\theta_1,...,\theta_k)'$. Апостериорна средина за било који елеменат θ биће:[101]

[101] Koop, G. 2003. *Bayesian Econometrics*, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, England, pp. 7

$$E(\theta_i|y) = \int \theta_i\, p(\theta|y)d\theta. \tag{47}$$

Осим у неким једноставним случајевима, овај интеграл није могуће израчунати аналитички, већ је неопходна рачунарска подршка.

Поред тачкасте оцене, обично је пожељно представити и меру степена неизвесности која је повезана са тачкастом оценом. Најчешћа оваква мера је *апостериорна стандардна девијација*, која представља квадратни корен из *апостериорне варијансе*, а може се израчунати на следећи начин:

$$\operatorname{var}(\theta_i|y) = E(\theta_i^2|y) - \{E(\theta_i|y)\}^2,$$

што захтева израчунавање друге очекиване вредности:

$$E(\theta_i^2|y) = \int \theta_i^2\, p(\theta|y)d\theta.$$

Све апостериорне функције које се могу израчунати у Бајесовој анализи имају следећи облик:

$$E[g(\theta)|y] = \int g(\theta)\, p(\theta|y)d\theta, \tag{48}$$

где је $g(\theta)$ *фунцкија од интереса*.

Све ове функције у себи садрже интеграле, као и маргинална веродостојност и предиктивна густина. Интеграли садржани у Бајесовој анализи обично се оцењују применом метода симулације за који је неопходна примена одговарајућег софтверског пакета.

Емпиријске Бајесове методе користе информације из података како би одредили априорну вероватноћу, па самим тим нарушавају основне премисе Бајесових метода. Познато је да априорна информација може бити чисто субјективна, али и поред тога, емпиријске Бајесове или хијерархијске априорне информације често користе макроекономисти у својим истраживањима, као и сви они који су заинтересовани за практичне алате који се добро показују у пракси. У случају када смо ограничени информацијама о подацима у односу на број параметара, улога априорне инфорамације постаје све значајнија, а тиме и примена Бајесове анализе у истраживању.

3.7 Бајесов метод интервалног оцењивања

Бајесова анализа нуди алтернативу класичном тестирању хипотеза и оцењивању интервала поверења заснованом на претпоставци о нормалности. По класичном схватању параметар θ је непозната константа и тренутно расположиви подаци се користе за оцену његове стварне вредности. Бајесов метод интервалног оцењивања посматра параметар θ као случајну променљиву са законом вероватноће $\pi(\theta)$ који изражава субјективно убеђење о распореду непознатог параметра.

Бајесова статистичка анализа обезбеђује апостериорну расподелу за параметре засновану на априорној расподели, скупу података и моделу. Апостериорна расподела унапређује знање о параметрима и обезбеђује вероватноће за параметре које се могу искористити у сврху даљег истраживања у различитим областима. Исто тако, Бајесов приступ комбинује претходне информације о параметру са анализом узорачког скупа података коју тек треба извести.

Укупну статистику из апостериорне расподеле или целокупну расподелу, могуће је искористити за оцену параметара. Код Бајесовог статистичког закључивања, не постоји потреба за индиректним тумачењем резулатата, као што чине следбеници класичне статистике уз понављање експеримента. Логика Бајесове статистичке анализе обезбеђује вероватноћу за параметре, на основу датих података, што је у супротности са логиком класичне статистичке анализе која обезбеђује вероватноће за скуп података, на основу датих параметара.

У многим ситуацијама се показало да је корисније применити интервално оцењивање уместо тачкастог оцењивања за одређивање стварне вредности параметра популације θ.

Овакве интервале у класичној статистици називамо интервалима поверења. Бајесов интервал, аналоган интервалу поверења, назива се *интервал поузданости*.

Ако је (X_1, X_2,...., X_n) случајан узорак са законом вероватноће $f(x|\theta)$, а параметар θ се сматра случајном променљивом са априорном расподелом $\pi(\theta)$, тада случајне променљиве $T_1(X_1, X_2,...., X_n)$ и $T_2(X_1, X_2,...., X_n)$ за које је

$$P\{T_1(X_1, X_2,..., X_n) \leq \theta \leq T_2(X_1, X_2,..., X_n)\} \geq 1-\alpha$$

формирају $100(1-\alpha)\%$ Бајесов интервал поузданости за непознати параметар θ који се означава са

$$(T_1(X_1, X_2,..., X_n), T_2(X_1, X_2,..., X_n)).$$

Дефиниција.[102] Нека је (X_1, X_2,...., X_n) случајан узорак из расподеле са функцијом вероватноће (или густине) $f(\theta|x)$ и нека Θ има априорну функцију вероватноће (или густине) $\pi(\theta)$. Нека је $h(\theta| x_1, x_2,...., x_n)$ апостериорна расподела од Θ за дато $X_1 = x_1$, $X_2 = x_2$, ..., $X_n = x_n$. Тада се било који интервал [$T_1(x_1, x_2,...., x_n)$, $T_2(x_1, x_2,...., x_n)$] који задовољава услов

$$\int_{T_1}^{T_2} h(\theta|x_1, x_2,..., x_n)d\theta \geq 1-\alpha, \tag{49}$$

назива *Бајесов интервал поузданости* са нивоом поверења 1 - α за непознати параметар θ.

$T_1(x_1, x_2,...., x_n)$ и $T_2(x_1, x_2,...., x_n)$ представљају доњу и горњу границу Бајесовог интервала поузданости. Наравно, на сличан начин се може дефинисати и једнострани Бајесов интервал или 1 - α ниво поверења доњих и горњих Бајесових граница. Јасно је да ће бити

[102] Rohatgi, V.K. 1984. *Statistical Inference*, John Wiley & Sons, New York, pp. 700

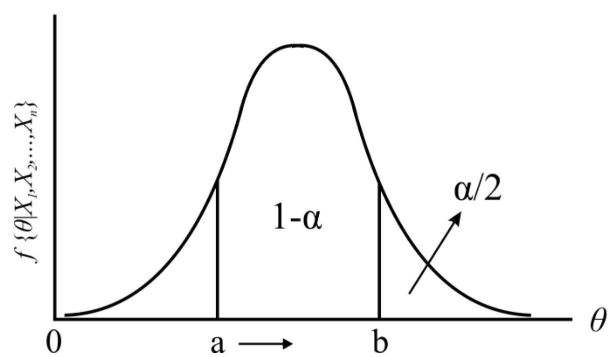
изабран најкраћи Бајесов интервал поузданости међу свим интервалима поузданости са нивоом поверења 1 - α.

Дефиниција. Интервал поузданости 100(1-α)% за θ је интервал (a, b), такав да је

$$P(a \leq \theta \leq b\ (x_1, x_2, ..., x_n) \geq (1\text{-}\alpha)100\%, \tag{50}$$

где је α позитивна величина која узима вредност између 0 и 1, а (x_1, x_2,...., x_n) су вредности узорка.

Може се приметити да је дефиниција прочитана уназад, тако да се са најмање (1-α)100% поузданости може тврдити да ће се стварна вредност параметра θ наћи између a и b, за дату информацију из узорка.



Слика 5. Интервал поузданости за θ[103]

С обзиром да је условна расподела од θ, за дато (X_1, X_2,...., X_n), заправо расподела вероватноће, оправдано је говорити о вероватноћи да ће се θ наћи у интервалу (a, b). Када су подаци проучени, интервал поверења је фиксан, док је θ случајна величина. Ово је у супротности са класичним интервалом поверења, где је интервал случајан, а θ фиксан параметар. У класичном смислу се каже да ће, дугорочно посматрано, 100(1-α)% свих

[103] Ramachandran, M.K. and Tsokos, P.C. 2009. *Mathematical Statistics with Applications,* Elsevier Academic Press, San Diego, California, pp. 580

оваквих интервала садржати стварну вредност параметра θ. Код Бајесовог приступа се каже да је вероватноћа најмање 100(1-α)% да се θ налази унутар одређеног интервала (a, b).

Као и код класичног приступа, пожељно је минимизирати ширину интервала поузданости. То захтева избор само тачака са највишим вредностима у густини $f(\theta|x_1, x_2,...., x_n)$, као што је приказано на Слици 6.

Претходна дефиниција може се другачије формулисати применом апостериорне расподеле од θ.

Дефиниција. *Интервал поузданости* 100(1-α)% за θ је интервал (a, b), такав да је:

1. $\int_{\theta}^{b} f(\theta|x_1, x_2,.., x_n)d\theta \geq 1-\alpha,$ ако је θ непрекидно и апостериорна функција густине вероватноће за θ је $f(\theta|, x_2,.., x_n),$

(51)

2. $\sum^{b} f(\theta|x_1, x_2,..., x_n) \geq 1-\alpha,$ ако је θ дискретно.

Ова процедура се може приказати и кроз неколико корака који заправо представљају процедуру за одређивање Бајесовог интервала поузданости:[104]

1. Узима се у обзир да је θ случајна променљива са априорном функцијом расподеле вероватноће (или заједничком функцијом густине) $\pi(\theta)$.

2. Коригује се априорна расподела $\pi(\theta)$ применом Бајесове теореме. Апостериорна расподела за θ проналази се помоћу формуле

[104] Ramachandran, M.K. and Tsokos, P.C. 2009. *Mathematical Statistics with Applications,* Elsevier Academic Press, San Diego, California, pp. 582

$$\pi(\theta|x) = \begin{cases} \dfrac{f(x|\theta)\pi(\theta)}{\int f(x|\theta)\pi(\theta)d\theta}, & \text{у непрекидном случају,} \\[2ex] \dfrac{f(x|\theta)\pi(\theta)}{\sum f(x|\theta)\pi(\theta)}, & \text{у дискретном случају.} \end{cases}$$

3. Проналазе се две величине a и b такве да је

$$\int_a^b \pi(\theta|x)d\theta \geq 1-\alpha, \qquad \text{у непрекидном случају,}$$

$$\sum_{\theta=a}^{b} \pi(\theta|x) \geq 1-\alpha, \qquad \text{у дискретном случају.}$$

Напомена: Величине a и b проналазимо тако да је

$$\int_{-\infty}^{a} \pi(\theta|x)d\theta = \alpha/2, \qquad \text{у непрекидном случају,}$$

$$\sum_{\theta\leq a} \pi(\theta|x) = \alpha/2, \qquad \text{у дискретном случају}$$

и

$$\int_{b}^{\infty} \pi(\theta|x)d\theta = \alpha/2, \qquad \text{у непрекидном случају,}$$

$$\sum_{\theta\geq b} \pi(\theta|x) = \alpha/2, \qquad \text{у дискретном случају.}$$

(1-α)100% интервал поузданости за θ је интервал (a,b).

У дискретном случају, једноставан начин за проналажење интервала поверења најмање дужине је да се све вредности за θ поређају од највероватнијих до најмање вероватних (у циљу давања значаја апостериорној расподели), а затим се вредности за θ смештају у интервал све док кумулативна апостериорна вероватноћа скупа не надмаши $(1-\alpha)100\%$. Такав интервал се назива *интервал највише апостериорне густине* (енгл. *highest posterior density interval* (HPD)).

Дефиниција.[105] Нека је $\pi(\theta|x)$ апостериорна густина за θ из $\Theta \subset R$. Област $C=C(x)$ се назива област *највише апостериорне густине* садржине p ако је

(а) $\int_C \pi(\theta|x)d\theta = p,$

(52)

(б) за било које $\theta \in C$ и $\theta^* \notin C$, имамо да је

$$\pi(\theta|x) \geq \pi(\theta^*|x).$$

HPD области нису обавезно гранични интервали. Уколико је апостериорна густина унимодална, тада ће HPD област бити интервал.

Међутим, узимајући у обзир да апостериорна расподела није увек унимодална, HPD области могу да садрже неколико неповезаних сетова. То може деловати нелогично са класичног становишта, али се мора схватити као указивање на неодређеност, било у подацима или у априорним информацијама, о могућим вредностима θ.[106]

[105] Knight, K. 2000. *Mathematical Statistics*, Chapman & Hall/CRC, Boca Raton, Florida, pp. 350

[106] Marin, J.M. and Robert, C.P. 2007. *Bayesian Core: A Practical Approach to Computational Bayesian Statistics*, Science+Business Media, LLC, pp. 27

Дефиниција.[107] Претпоставимо да је апостериорна густина за θ унимодална. Тада је HPD интервал за θ интервал

$$C = \{\theta : \pi(\theta | X = x) \geq k\}, \tag{53}$$

где је k одабрано тако да је

$$P(C | X = x) = 1 - \alpha. \tag{54}$$

Нека је ($X_1, X_2, \ldots, X_n$) случајан узорак из популације са нормалном расподелом $\mathcal{N}(\theta, \sigma^2)$, са познатом σ^2, где θ има нормалну априорну расподелу. Апостериорна расподела ће такође бити нормална у облику $\mathcal{N}(\mu_1, \sigma^2)$, где су μ_1 (које зависи од ($x_1, x_2, \ldots, x_n$)) и σ_1^2 израчунати. У овом случају се HPD интервал, нормални апостериорни интервал и интервал са једнаким крајевима поклапају и узимају следећи облик:

$$(\mu_1 - \sigma_1 z_{\alpha/2},\ \mu_1 + \sigma_1 z_{\alpha/2}).$$

У случају ограничења да априорна варијанса $\sigma_0^2 \to \infty$ (дифузна или *неправилна априорна расподела* за θ), тада је:

$$\mu_1 \to \bar{x}, \qquad \sigma_1^2 \to \frac{\sigma^2}{n},$$

тако да се Бајесов интервал у потпуности слаже са класичним интервалом поверења, иако је интерпретација ових интервала потпуно другачија.

За вектор θ, може се узети у обзир HPD скуп поузданости, посебно уколико је апостериорна расподела унимодална. У супротном, могуће је конструисати интервале поузданости за сваки елеменат.

[107] Ghosh, J.K, Delampady, M. and Samanta, T. 2006. *An Introduction to Bayesian Analysis, Theory and Methods*, Springer Science+Business Media, LLC, New York, pp. 49

Интервали поузданости су једноставни за израчунавање за разлику од интервала поверења, чија израда захтева кључне величине или инверзију фамилије тестова.

Нека је Θ скуп параметара и нека посматрана случајна промењлива X има функцију расподеле $f_\theta(x)$. Посматрајмо сада θ као случајну промењливу са законом вероватноће $\pi(\theta)$. Тада $f_\theta(x)$ можемо сматрати условном функцијом расподеле за X. Када је X=x постоји могућност дефинисања заједничке расподеле за X и θ, маргиналне расподеле за X и, такође, условне расподеле за θ.

Заједничка расподела дата је са

$$f(x,\theta) = \pi(\theta) f_\theta(x), \tag{55}$$

а маргинална расподела за X са

$$g(x) = \begin{cases} \sum \pi(\theta) f_\theta(x) & \text{ако је } \pi \text{ функција вероватноће,} \\ \\ \int \pi(\theta) f_\theta(x) d\theta & \text{ако је } \pi \text{ функција густине.} \end{cases} \tag{56}$$

Условна расподела за θ, за реализовану вредност x, дата је са:

$$h(\theta|x) = \frac{\pi(\theta) f_\theta(x)}{g(x)}, \qquad g(x) > 0. \tag{57}$$

Једном када је израчуната апостериорна расподела од Θ, лако је наћи функције $T_1(X)$ и $T_2(X)$ за које важи:

$$P\{\, T_1(X) < \theta < T_2(X)\} \geq 1 - \alpha, \tag{58}$$

где

$$P\{\, T_1(X) < \theta < T_2(X) |\, X = x \} = \begin{cases} \displaystyle\int_{T_1}^{T_2} h(\theta|x)\,d\theta, \\ \\ \displaystyle\sum_{T_1}^{T_2} h(\theta|x), \end{cases} \tag{59}$$

у зависности од тога да ли је h функција густине или функција вероватноће.

Неки од разлога који иду у прилог Бајесовом интервалном оцењивању су: (1) већина закључака је условљена посматраним подацима; за разлику од традиционалног приступа, не мора се водити рачуна о сетовима података, већ само о онима који су предмет посматрања; (2) са Бајесове тачке гледишта, сасвим је оправдано говорити о вероватноћи да ће се вредност параметра наћи у неком одређеном интервалу, рецимо (0.2, 0.6), или о вероватноћи да је хипотеза тачна. Са друге стране, веома често се дешава да су класични закључци погрешни. Бајесова статистика обезбеђује погодан модел за имплементацију научног метода.

3.8 Тестирање хипотеза из угла Бајесове статистике

Бајесов приступ тестирању хипотеза развио је Џефрис као кључни део његовог програма научног закључивања. Иако је Џефрис своје методе, по угледу на Фишера, назвао *тестовима значајности*, очигледно је да је у питању грешка, пошто су Џефрисови циљеви били потпуно другачији. У свом раду, Џефрис се бавио поређењем резултата до којих се долази применом две конкурентске научне теорије. У том смислу, статистички модели су уведени како би се приказала вероватноћа података према свакој од ове две теорије, док је Бајесова теорема коришћена за израчунавање апостериорне расподеле да је једна од теорија тачна.

Контроверзе око примене Бајесових и не Бајесових процедура тестирања у значајној мери су присутне у литератури и у том смислу је потребно истаћи неколико кључних тачака:[108]

1. не може се очекивати да ће *p* вредност бити једнака апостериорној вероватноћи да је нулта хипотеза тачна;

2. класични тестови готово уобичајено одбацују нулту хипотезу код веома великих узорака, док код Бајесовог фактора то није случај; из тог разлога Бајесов фактор данас има широку примену уз коришћење Бајесовог информационог критеријума за апроксимацију;

3. Бајесов фактор, као и уопште Бајесове процедуре, следи принцип веродостојности; као резултат тога, у околностима где се случајеви могу јављати узастопно, Бајесов фактор се може примењивати без забринутости да ће се јавити потреба за непланираном анализом података;

4. Бајесов фактор се подједанако лако може применити и код неуклапајућих, као и код уклапајућих модела (енгл. *nested*); супротно, примена класичних тестова значајности код неуклапајућих модела веома је тешка;

5. класични тестови значајности су развијени ради потребе поређења два модела, међутим практична анализа података обично укључује много више модела, макар имплицитно; у таквим околностима, примена вишеструких класичних тестова како би се дошло до најбољег модела, може дати погрешне резултате; укључивањем неизвесности у модел и применом Бајесовог фактора, могуће је избећи овај проблем.

Поступак тестирања код класичног приступа заснива се на контроли две врсте грешака – грешака прве и друге врсте. Класични поступци тестирања ограничавају грешку прве

[108] Kass, R.E. and Raftery, A.E. 1995. „Bayes factor", *Journal of the American Statistical Association*, Vol. 90, No. 430, pp. 789

врсте на ниво α и минимизирају грешку друге врсте. Уколико је грешка друге врсте неприхватљиво висока, биће смањена повећањем величине узорка.

Бајесов приступ тестирању хипотеза за просте хипотезе је прилично једноставан. Поступак одлучивање између две хипотезе за дати скуп података x скраћује се израчунавањем апостериорне вероватноће. Уколико је доступна функција губитка, применом Бајесовог правила минимизираће се очекивана вредност функције губитка у односу на апостериорну расподелу. У одсуству функције губитка, вероватноће грешке прве и друге врсте су од малог интереса за Бајесов приступ.

Бајесов метод тестирања статистичких хипотеза, полази од тестирања просте нулте хипотезе H_0: $\theta = \theta_0$ против просте алтернативне хипотезе H_1: $\theta = \theta_1$. Поред информације о параметру θ, која се добија из случајног узорка $(X_1, X_2,...., X_n)$, узима се у обзир и априорна информација о параметру θ која је садржана у закону вероватноће априорног параметра Θ:[109]

Θ	θ_0	θ_1
$\pi(\theta)$	$\pi(\theta_0)$	$1-\pi(\theta_0)$

Нека је $\pi(\theta)$ априорна расподела вероватноће на Θ. Пошто у овом случају постоје само две вредности за Θ, θ_0 и θ_1, тада је Бајесов ризик:

$$R(\pi,d) = E_\theta R(\theta,d(X))$$

$$= \begin{cases} \int_\Theta R(\theta,d)\pi(\theta)d\theta, & \text{ако је } \pi \text{ функција густине вероватноће,} \\ \sum_\Theta R(\theta,d)\pi(\theta), & \text{ако је } \pi \text{ заједничка функција вероватноће,} \end{cases}$$

[109] Милошевић, В. 1995. *Теорија статистичког закључивања*, Научна књига, Београд, стр. 172

$$= \begin{cases} \int_{\Theta_0} b(\theta)\pi(\theta)P_\theta\{d(X)=a_1\}d\theta + \int_{\Theta_1} a(\theta)\pi(\theta)P_\theta\{d(X)=a_0\}d\theta, \\ \qquad \text{ако је } \pi \text{ функција густине вероватноће,} \\ \sum_{\Theta_0} b(\theta)\pi(\theta)P_\theta\{d(X)=a_1\} + \sum_{\Theta_1} a(\theta)\pi(\theta)P_\theta\{d(X)=a_0\}, \\ \qquad \text{ако је } \pi \text{ заједничка функција вероватноће.} \end{cases} \tag{60}$$

Бајесово решење је правило одлучивања које минимизира Бајесов ризик $R(\pi, d)$. У том смислу, ограничићемо пажњу на случај где и H_0 и H_1имају тачно по једну тачку, $\Theta_0 = \{\theta_0\}$, $\Theta_1 = \{\theta_1\}$. Нека је $\pi(\theta_0) = \pi_0$ и $\pi(\theta_1) = 1 - \pi_0 = \pi_1$. Тада је

$$R(\pi, d) = b\pi_0 P_{\theta_0}\{d(X)=a_1\} + a\pi_1 P_{\theta_1}\{d(X)=a_0\}, \tag{61}$$

где је $b(\theta_0) = b, a(\theta_1) = a; (a, b > 0)$.

Теорема 5.[110] Нека је $X = (X_1, X_2, ..., X_n)$ случајна променљива дискретног (непрекидног) типа са заједничком функцијом вероватноће f_θ, $\theta \in \Theta = \{\theta_0, \theta_1\}$. Нека је $\pi(\theta_0) = \pi_0$, $\pi(\theta_1) = 1 - \pi_0 = \pi_1$ априорна вероватноћа заједничке функције на Θ. Бајесово решење за тестирање H_0: $X \sim f_{\theta_0}$ против $H_1 : X \sim f_{\theta_1}$ уз коришћење функције губитка, одбациће H_0 ако је

$$\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)} \geq \frac{b\pi_0}{a\pi_1}.$$

Доказ. Потребно је пронаћи d које минимизира:

$$R(\pi, d) = b\pi_0 P_{\theta_0}\{d(X)=a_1\} + a\pi_1 P_{\theta_1}\{d(X)=a_0\}.$$

[110] Rohatgi, V.K. 1976. *An Introduction to Probability Theory and Mathematical Statistics*, John Wiley & Sons, New York, pp. 462

Сада је:

$$R(\pi, d) = E_\theta R(\theta, d)$$

$$= E\{E_\theta\{L(\theta, d)|X\}\}$$

па је довољно да минимизира $E_\theta\{L(\theta, d)|X\}$.

Апостериорна расподела за θ дата је са:

$$h(\theta|x) = \frac{\pi(\theta) f_\theta(x)}{\sum_\theta f_\theta(x)\pi(\theta)}$$

$$= \frac{\pi(\theta) f_\theta(x)}{\pi_0 f_{\theta_0}(x) + \pi_1 f_{\theta_1}(x)}$$

$$= \begin{cases} \dfrac{\pi_0 f_{\theta_0}(x)}{\pi_0 f_{\theta_0}(x) + \pi_1 f_{\theta_1}(x)}, & \text{ако је } \theta = \theta_0, \\[2em] \dfrac{\pi_1 f_{\theta_1}(x)}{\pi_0 f_{\theta_0}(x) + \pi_1 f_{\theta_1}(x)}, & \text{ако је } \theta = \theta_1. \end{cases}$$

Отуда:

$$E_\theta\{L(\theta, d(X))|X = x\} = \begin{cases} b\,h(\theta_0|x), & \theta = \theta_0, d(X) = a_1, \\[2em] a\,h(\theta_1|x), & \theta = \theta_1, d(X) = a_0. \end{cases}$$

Следи да се H_0 одбацује јер је $d(X) = a_1$ ако је:

$$b\,h(\theta_0|x) \le ah(\theta_1|x)$$

што је у случају ако и само ако је:

$$b\pi_0 f_{\theta_0}(x) \le a\pi_1 f_{\theta_1}(x),$$

као доказ.

Код Нојман-Пирсонове леме фиксира се $P_{\theta_0}\{d(X)=a_1\}$, као вероватноћа одбацивања H_0 када је H_0 истинито и минимизира се $P_{\theta_1}\{d(X)=a_0\}$, као вероватноћа прихватања H_0 када H_0 није истинито. Код Бајесовог теста више не постоји фиксиран ниво α за $P_{\theta_0}\{d(X)=a_1\}$. Уместо тога дозвољено је да се узима било која вредност све дотле док је $R(\pi,d)$, дефинисано у (61), минимално.

Теорему 3. је могуће уопштити у случају сложених одлука. Нека је X случајна променљива са заједничком функцијом вероватноће f_θ, где θ може узети било коју од k вредности θ_1, θ_2,...., θ_k. Проблем је у проучавању x и доношењу одлуке која је од наведених вредности θ_i тачна вредност за θ. Нека је H_i: $\theta=\theta_i$, i = 1, 2, ..., k. Претпоставимо да је $\pi(\theta_i)=\pi_i$, i = 1, 2, ..., k, $\sum_1^k \pi_i = 1$, априорна расподела вероватноће на $\Theta\{\theta_1, \theta_2,...., \theta_k\}$. Нека је

$$L(\theta_i,d)=\begin{cases} 1, & \text{уколико } d \text{ одабере } \theta_j,\ j\neq i, \\ \\ 0, & \text{уколико } d \text{ одабере } \theta_i. \end{cases}$$

Проблем је пронаћи одлуку d која минимизира $R(\pi,d)$. Бајесово решење прихвата H_i: $\theta=\theta_i$, i = 1, 2, ..., k, ако је:

$$\pi_i f_{\theta_i}(x) \ge \pi_j f_{\theta_j}(x), \qquad \text{за свако } j\neq i,\ j = 1, 2, ..., k.$$

У случају тестирања нулте и алтернативне хипотезе у облику H_0:$\theta \in \Theta_0$ против H_1:$\theta \in \Theta_1$, где су Θ_0 и Θ_1 подскупови са реалне праве, Бајесова процедура тестирања може се приказати и на другачији начин.

Нека је $(X_1, X_2,...., X_n)$ узорак из популације са функцијом расподеле $f_\theta(x)$. Код Бајесовог тестирања хипотеза рачунају се следеће апостериорне вероватноће:

$$\alpha_0 = P(\theta \in \Theta_0 | x_1,..., x_n) \tag{62}$$

и

$$\alpha_1 = P(\theta \in \Theta_1 | x_1,..., x_n)\,. \tag{63}$$

Ако је $\alpha_0 > \alpha_1$, нулту хипотезу прихватамо и ако је $\alpha_0 < \alpha_1$, нулту хипотезу одбацујемо. Процедура за Бајесово тестирање наведене хипотезе, може се укратко приказати на следећи начин:

Нека је $\pi(\theta)$ априорна расподела. Такође,

$$\pi_0 = P(\theta \in \Theta_0) \tag{64}$$

и

$$\pi_1 = P(\theta \in \Theta_1) \tag{65}$$

Количник π_0/π_1 назива се *априорни случајни коефицијент*.[111] Количник α_0/α_1 из образаца (62) и (63) назива се *апостериорни случајни коефицијент*.

Апостериорни случајни коефицијент је количник апостериорних вероватноћа датих података за нулту и алтернативну хипотезу. Апостериорни случајни коефицијент се

[111] Ramachandran, M.K., Tsokos, P.C. 2009. *Mathematical Statistics with Applications*, Elsevier Academic Press, San Diego, California, pp. 585

користи у процесу одлучивања за тестирање хипотеза. Сада се могу израчунати α_0 и α_1 применом Бајесове теореме. То значи,

$$\alpha_0 = P(\theta \in \Theta_0 | x_1,...,x_n) = \begin{cases} \int_{\Theta_0} f(\theta | x_1,....,x_n) d\theta, & \text{у непрекидном случају} \\ \sum_{\theta \in \Theta_0} f(\theta | x_1,....,x_n), & \text{у дискретном случају.} \end{cases}$$

Слично, (66)

$$\alpha_1 = P(\theta \in \Theta_1 | x_1,...,x_n) = \begin{cases} \int_{\Theta_1} f(\theta | x_1,....,x_n) d\theta, & \text{у непрекидном случају} \\ \sum_{\theta \in \Theta_1} f(\theta | x_1,....,x_n), & \text{у дискретном случају.} \end{cases}$$

Уколико је случајни коефицијент $(\alpha_0/\alpha_1) < 1$ H_0 се одбацује. Хипотеза H_0 се прихвата уколико је $(\alpha_0/\alpha_1) > 1$.

Количник апостериорног и априорног коефицијента често се назива *Бајесов фактор* и зависи само од података из узорка. Уколико се претпостави једнака вероватноћа за просту и алтернативну хипотезу, апостериорни коефицијент биће једнак количнику веродостојности.[112]

[112] Swiler, L.P. 2006. „Bayesian Methods in Engineering Design Problems“, *Sandia National Laboratories*, Albuquerque, New Mexico, pp. 8

У случају тестирања нулте хипотезе H_0 против алтернативне хипотезе H_1 Бајесов фактор биће означен са BF_{01}, где ће нулта хипотеза бити у бројиоцу. Уколико се резултати пореде са стандардним тестом количника веродостојности, погодно је да се уместо тога нулта хипотеза стави у именилац и да се као Бајесов фактор користи BF_{10}. Када се посматра више хипотеза, уобичајено је да се за Бајесов фактор пише BF_{jk} за H_j против H_k.[113]

Бајесов фактор у корист H_0, као однос између одговарајућег апостериорног и априорног коефицијента, може се представити на следећи начин:[114]

$$BF_{01}^{\pi}(x) = \frac{\dfrac{P(\theta \in \Theta_0 | x_1,...,x_n)}{P(\theta \in \Theta_1 | x_1,...,x_n)}}{\dfrac{P(\theta \in \Theta_0)}{P(\theta \in \Theta_1)}} = \frac{\dfrac{\alpha_0}{\alpha_1}}{\dfrac{\pi_0}{\pi_1}}. \tag{67}$$

У случају када су хипотезе просте, тј. H_0: $\theta = \theta_0$ против H_1: $\theta = \theta_1$, и када је априорна расподела $\pi(\theta_0) = \pi_0$ и $\pi(\theta_1) = \pi_1 = 1 - \pi_0$, Бајесов фактор у корист H_0 постаје количник веродостојности:

$$BF_{01}^{\pi}(x) = \frac{\dfrac{P^{\theta|X}(\theta \in \Theta_0)}{P^{\theta|X}(\theta \in \Theta_1)}}{\dfrac{P^{\theta}(\theta \in \Theta_0)}{P^{\theta}(\theta \in \Theta_1)}} = \frac{\dfrac{f(x|\theta_0)\pi_0}{f(x|\theta_1)\pi_1}}{\dfrac{\pi_0}{\pi_1}} = \frac{f(x|\theta_0)}{f(x|\theta_1)}. \tag{68}$$

Бајесов фактор мери релативне промене у априорним коефицијентима када су прикупљени докази. Као централни појам Бајесовог тестирања, Бајесов фактор се може користити за Бајесово тестирање без узимања у обзир специфичног губитка већ, на пример, коришћењем Џефрисове скале доказа.

---

[113] Kass, R.E. and Raftery, A.E. 1995. „Bayes factor”, *Journal of the American Statistical Association*, Vol. 90, No. 430, pp. 777

[114] Anscombe, F.J. 1961. „Bayesian statistics“, *The American Statistician*, 15(1), pp. 11

У Табели 5. дате су практичне смернице за Бајесово тестирање хипотеза у зависности од вредности Бајесовог фактора.

Табела 5. Тумачење Бајесовог фактора

Бајесов фактор *BF*	**Тумачење**
$BF > 1$	Доказ подржава H_0
$1 > BF > 10^{-1/2}$	Благ доказ против H_0
$10^{-1/2} > BF > 10^{-1}$	Значајан доказ против H_0
$10^{-1} > BF > 10^{-3/2}$	Јак доказ против H_0
$10^{-3/2} > BF > 10^{-2}$	Веома јак доказ против H_0
$10^{-2} > BF$	Одлучујући доказ против H_0

Извор: Young, G.A. and Smith, R.L. 2005. *Essentials of statistical inference*, Cambridge University Press, Cambridge, pp. 74

Оваква интерпретација Бајесовог фактора понекад може бити спорна, али суштина је у томе да се величина Бајесовог фактора схвати као сама по себи смислена мера. Управо та тачка гледишта представља праву полазну тачку за прављење разлике између класичног и Бајесовог приступа тестирању хипотеза.

Такође, потребно је истаћи да за просту нулту хипотезу, где је θ једнако некој одређеној вредности, неће бити лако применити Бајесову процедуру. Отуда, за разлику од класичног приступа, Бајесов приступ ће у већини случајева разматрати сложене хипотезе.

Претпоставимо сада да су хипотезе H_0 и H_1 сложене. Како би се израчунао Бајесов фактор, треба да буде позната потпуна априорна расподела за θ, што значи да није довољно ако су познате априорне расподеле π_0 и π_1, већ је потребно знати и априорну густину за θ за хипотезе H_0 и H_1.

Претпоставимо даље да θ има априорну густину $g_0(\theta)$, $\theta \in \Theta_0$ уколико је H_0 тачна и $g_1(\theta)$, $\theta \in \Theta_1$ уколико је H_1 тачна.

У овом случају, Бајесов фактор од H_0 у односу на H_1, дефинише се као:[115]

$$BF_{01} = \frac{\int_{\Theta_0} f(x|\theta) g_0(\theta) d\theta}{\int_{\Theta_1} f(x|\theta) g_1(\theta) d\theta}, \tag{69}$$

где је $g_i(\theta)$ априорна функција густине вероватноће од θ по Θ_i тако да је:

$$\int_{\Theta_i} g_i(\theta) d\theta = 1.$$

Уколико је H_0 проста хипотеза, H_0: $\theta = \theta_0$, а H_1 сложена хипотеза, H_1: $\theta \neq \theta_0$, Бајесов фактор се може написати у следећем облику:

$$BF_{01} = \frac{f(x|\theta_0)}{\int_{\Theta_1} f(x|\theta) g_1(\theta) d\theta}. \tag{70}$$

Све се може и општије приказати. Претпоставимо да имамо два параметарска модела M_1 и M_2 за податке X и да два модела имају одговарајуће параметаре θ_1 и θ_2. Уколико су априорне густине $\pi_i(\theta_i)$, $i = 1, 2$ за параметре у моделима, маргинална расподела за X биће:

$$p(x|M_i) = \int f(x|\theta_i, M_i) \pi_i(\theta_i) d\theta_i, \quad i = 1, 2,$$

а Бајесов фактор биће количник:

$$BF = \frac{p(x|M_1)}{p(x|M_2)}. \tag{71}$$

Тумачење Бајесовог фактора биће исто као у горенаведеној табели, али ће стварни обрачун зависити од априорних густина.

[115] Ghosh, J.K, Delampady, M. and Samanta, T. 2006. *An Introduction to Bayesian Analysis, Theory and Methods*, Springer Science+Business Media, LLC, New York, pp. 43

У случају када је број промењљивих мали, могу се користити стохастичке методе за процену маргиналне вероватноће или неки од критеријума процене као што су Акаике информациони критеријум (енгл. *Akaike Information Criterion*, *AIC*) или Бајесов информациони критеријум (енгл. *Bayesian Information Criterion*, *BIC*).

Акаике се залагао да се за дату класу конкурентских модела за скуп података, одабере модел који минимизира:[116]

$$AIC = (-2)\log(\text{maksimalna verodostojnost}) + 2(\text{broj parametara}) \tag{72}$$

Акаике инфорамциони критеријум се дефинише као оцена (-2) очекиване log веродостојности модела чији су параметри одређени методом максималне веродостојности. У овом изразу „log" представља природни логаритам. Једноставна процедура којом се одабира модел са минималним *AIC* међу скупом модела одређује минималну *AIC* оцену. Увођење *AIC* критеријума помогло је у препознавању значаја моделирања у статистици, док су многе практично корисне статистичке процедуре развијене као минималне *AIC* процедуре.

Сматра се да што је мања вредност ове статистике, модел је бољи. *AIC* критеријум мери колико је добро уклапање мерено по максималној log веродостојности у односу на комплексност модела, мерено бројем параметара.[117]

Примена Акаике информационог критеријума образлаже се на два начина. Прво образложење се заснива на предиктивном аргументу. Уколико је предиктивна расподела условљена једним моделом и оцењеним параметарима тог модела, тада ће *AIC* одабрати модел који даје најбољу апроксимацију. Међутим, оваква предиктивна расподела је неисправна, пошто не укључује неизвесност у погледу вредности параметара и облика модела.

[116] Akaike, H. 1978. A Bayesian Analysis of the Minimum AIC Procedure, *Annals of the Institute of Statistical Mathematics*, 30, Part A, pp. 9

[117] Tableman, M. and Kim, J.S. 2004. *Survival Analysis Using S*, *Analysis of Time-to-Event Data*, Chapman & Hall/CRC, pp. 107

Друго образложење за примену Акаике информационог критеријума је Бајесијанско. Сам Акаике објашњава да је поређење модела засновано на овом критеријума асимптотски једнако оном које је засновано на Бајесовом фактору.

У складу са тим, метод који може да пружи грубу процену доказа у корист једног модела у односу на други, без узимања у обзир било које априорне расподеле, jесте *Бајесов информациони критеријум* или другачије *Шварцов критеријум* (engl. *Schwarz criterion*). Према овом критеријуму, модел са највећом апостериорном вероватноћом јесте онај модел који минимизира:[118]

$$BIC = (-2)\log(\text{maksimalna verodostojnost}) + (\log n)(\text{broj parametara}), \tag{73}$$

где је n величина узорка која се користи за израчунавање оцене максималне веродостојности. Поређењем образаца (72) и (73) може се закључити да *BIC* има тенденцију да фаворизује једноставније моделе од оних који су изабрани применом *AIC* критеријума.

Бајесов информациони критеријум се базира на резултату код ког се, за велике величине узорка n, приближавање -2 log *BF* објашњава са:[119]

$$\Delta BIC = W - (p_2 - p_1)\log n, \tag{74}$$

где је p_i број параметара у моделу M_i, $i = 1, 2$, а *W статистика количника веродостојности*:

$$W = -2\log\frac{\sup_{\theta_1} f(x|\theta_1, M_1)}{\sup_{\theta_2} f(x|\theta_2, M_2)}. \tag{75}$$

[118] Kass, R.E. and Raftery, A.E. 1995. „Bayes factor”, *Journal of the American Statistical Association*, Vol. 90, No. 430, pp. 790

[119] Young, G.A. and Smith, R.L. 2005. *Essentials of statistical inference*, Cambridge University Press, Cambridge, pp. 77

Груба апроксимација Бајесовом фактору, која не зависи од априорних расподела параметара у посматрана два модела, биће дата са:

$$BF \approx \exp(-\frac{1}{2}\Delta BIC). \tag{76}$$

Шварцов критеријум или *BIC* лак је за примену и може се користити за сумирање научних резултата.

Важно је истаћи да Бајесов фактор пореди једну теорију са другом, отуда се неколико Бајесових фактора може израчунати како би се одредила релативна подршка различитим теоријама. Бајесови фактори варирају у складу са претпоставкама, али не могу варирати по жељи. Такође, претпоставке подлежу суду јавности и о њима се може јавно расправљати. У том смислу, Бајесов фактор је објективан. Код Нојман-Пирсоновог закључивања, закључак зависи од истраживача и његовог виђења хипотезе и ова питања не подлежу увиду јавности.[120]

Бајесов фактор има много и предности и недостатака. Једна од главних предности јесте чврста логичка основа, која нуди велику флексибилност. Као што је већ поменуто, предност у израчунавању, као и развој рачунарске технике значајно су прошириле корисност Бајесових метода, тако да се Бајесов фактор сада може израчунати за велики број различитих модела.

Основни недостатак Бајесовог фактора је његова осетљивост на претпоставке у параметарским моделима и избор апироне информације. *BIC* би се могао користити за извештавање о резултатима истраживања као подршка другим анализама.

И поред свега наведеног, Бајесов фактор није општеприхваћен, чак ни код следбеника Бајесове статистике. За оне који верују да једини рационалан приступ статистици лежи у пуној спецификацији субјективних априорних вероватноћа, Бајесов фактор изгледа као

[120] Dienes, Z. 2011. „Bayesian Versus Orthodox Statistics: Which Side Are You On?”, *Perspectives on Psychological Science*, 6(3), pp. 284

начин избегавања дела проблема. Формална спецификација Бајесовог правила одлучивања не може да зависи само од Бајесовог фактора.

4. МОГУЋНОСТИ ПРИМЕНЕ КЛАСИЧНЕ И БАЈЕСОВЕ СТАТИСТИКЕ У ОБРАДИ ЕКОНОМСКИХ ПОДАТАКА

Дуги низ година, може се рећи и деценија, класична статистика је имала незаменљиву улогу у истраживањима у односу на Бајесове методе. Као два главна разлога за вишедеценијску превласт, наводе се априорне информације и поступак израчунавања.

Када је у питању први разлог, многи истраживачи су стављали приговор на коришћење *субјективне* априорне информације у наводно *објективној* економској науци. Расправа о улози априорних информација у статистичкој науци је врло дуго трајала и још увек је присутна. Осврт на ову проблематику је направљен у тачки 3.4, где је указано на суштинске проблеме, али и суштински значај Бајесовог приступа. Укратко, већина Бајесијанаца тврди да читав процес изградње модела може да обухвати огромну количину информација које нису повезане са подацима (нпр. мора се одабрати модел који ће се користити, променљиве које ће бити укључене, критеријуми за поређење модела или оцењивање параметара, емпиријски резултати који ће бити приказани, итд).

Бајесов приступ је веома јасан по питању тога како тачно треба користити информације које нису повезане са подацима. Штавише, ако је априорна информација доступна, пожељније је имати више информација него мање. Као последњу линију одбране, Бајесијанци су развили неинформативне априорне информације за бројне класе модела. То значи да Бајесов приступ омогућава коришћење априорних информација уколико истраживач жели да их користи. Без обзира на то какво је мишљење истраживача о априорној информацији, оно ни на који начин не треба да буде препрека за усвајање Бајесових метода.

Други и, историјски гледано, значајнији разлог због ког је класична статистика имала предност у примени био је поступак израчунавања. Бајесову анализу је рачунски било тешко или чак немогуће урадити за све, већ само за неколико специфичних класа модела. Рачунарска револуција омогућила је превазилажење овог проблема и довела до процвата Бајесових метода у многим областима.

Бајесову парадигму карактерише неколико предности у односу на класичну, као што је кохеренција саме парадигме која простиче из систематичне примене Бајесовог правила, концепт субјективне вероватноће, општи карактер Бајесових метода које не захтевају посебне услове регуларности, дефинисање појма интервала поверења, као и тестирање.[121] Сви ови елементи подржани развојем рачунарске технологије омогућили су примену Бајесове анализе у предвиђању будућих кретања економских појава.

Да би се одговорило на захтеве предмета истраживања докторске дисертације, спроведена је анализа коришћењем различитих економских података, полазећи од примера једног, конкретног предузећа, да би се кроз наредне примере извршила анализа података преузетих са сајтова Eurostat и Народне банке Србије. Резултати су презентовани кроз три различита примера, од којих је први обрађен у статистичким пакетима IBM SPSS *Version* 21 и IBM SPSS Amos *Version* 21, а други и трећи у статистичком пакету RATS *Version* 8.0 и 8.3.

4.1 Компарација метода максималне веродостојности и Бајесове статистике на примеру једног предузећа

Постоји више рачунарских софтверских пакета који се успешно могу користити за Бајесову анализу у одређеним класама модела. По сложености, модели се веома разликују, па је употреба једних врло једноставна и омогућава истраживачу да лако дође до жељених оцена и тест статистике, док други захтевају од истраживача да поседује програмерске вештине. Бајесова инференција наводи истраживаче да размишљају о смислу модела,

[121] Poirier, D. 2008. *Bayesian Econometrics*, In S.N. Durlauf and L.E. Blume (eds), The New Palgrave Dictionary of Economics, Palgrave Macmillan, pp. 772

односно веродостојности, као и о априорним информацијама које су важне за емпиријска питања која су предмет разматрања. Одабрани статистички пакети имају све карактеристике неопходне за упоредну анализу класичне и Бајесове инференције, па ће у даљој анализи бити објашњења и суштина њиховог функционисања у задатим оквирима.

Потреба за компаративном анализом метода максималне веродостојности, као метода класичне статистике, и Бајесовог метода проистекла је из самог циља истраживања, али и из тврдње да *оцене добијене на основу Бајесовог приступа представљају унапређење особина оцена добијених класичним приступом у смислу њихове прецизности*. Да би се испитала оправданост ове тврдње, два метода су упоређена обрадом података једног предузећа.

Како је већ споменуто (в. 2.3.2), метод максималне веродостојности има широку и значајну примену у одређивању статистичких оцена са добрим карактеристикама. Његова примена, као и примена Бајесових метода, не ограничава се само на поље економије, већ се може успешно имплементирати и у психологији, медицини, биологији, туризму, итд. Преглед актуелне литературе, потврђује свеприсутност ових метода у различитим пољима истраживања.

Тако, на пример, Лемон, Браун, Штангер-Хол и Лемон (Lemmon, Brown, Stanger-Hall and Lemmon, 2009) проучавају ефекат двосмислених података, односно недостајућих вредности на истраживања у биологији, применом метода максималне веродостојности и Бајесовог метода. Вард (Ward, 2008) у свом раду пореди Бајесове моделе и моделе максималне веродостојности за оцењивање еколошких модела, при чему класични метод доследно фаворизује једноставније моделе посматране популације у односу на Бајесове критеријуме.

Са друге стране, Флури и Шепард (Flury and Shephard, 2011) говоре о широкој примени ових метода у микроекономији, макроекономији и финансијској економетрији. При томе, илуструју примену класичних и Бајесових метода на четири економска проблема, где долазе до закључка да, уколико се посматрају заједно, ови методи имплицирају да извођење симулација на основу економских модела даје могућност да се одреди веродостојност. Пит, Силва, Ђордани и Кон (Pitt, Silva, Giordani and Kohn, 2012) у свом раду се баве развојем методологије за Бајесово закључивање за опште моделе временских серија помоћу Марковљевих ланаца Монте Карло (MCMC) симулација са оценама веродостојности. Фернандез-Виљаверде и Рубио-Рамирез (Fernández-Villaverde and Rubio-Ramírez, 2007) показују како се спроводи класично оцењивање на основу веродостојности у динамичним макроекономским моделима. Истовремено описују како се може користити аутпут за оцењивање параметара структуралних модела који се односе на преференције и технологију, као и за упоређивање различитих економија. Оба задатка се могу реализовати применом и класичних и Бајесових метода.

Бајесово оцењивање се такође може примењивати код решавања неких проблема који се најчешће појављују у традиционалној статистици. На пример, добијање оцена за немогуће параметре, пружање помоћи у идентификацији модела (Kim, Suh, Kim, Albanese and Langer, 2013), добијање прецизнијих оцена параметара (Depaoli, 2013) и помагање у ситуацијама у којима су на располагању само узорци мале величине (Zhang, Hamagami, Wang, Grimm, and Nesselroade, 2007) .

4.2 Примена одабраних модела истраживања

Да би се извршило поређење метода класичне и Бајесове статистике, узети су подаци компаније CaliVita Int., представништва за Републику Србију Fitco d.o.o., Нови Сад.[122] Одабрани подаци се односе на 252 производа који се налазе у продајном асортиману предузећа (в. Прилог 1). За потребе истраживања, производи су разврстани према: називу производа, врсти производа, произвођачу, цени, као и оствареној продаји у периоду од јануара до јуна 2014. године и обрађени у статистичком пакету IBM SPSS *Version* 21.

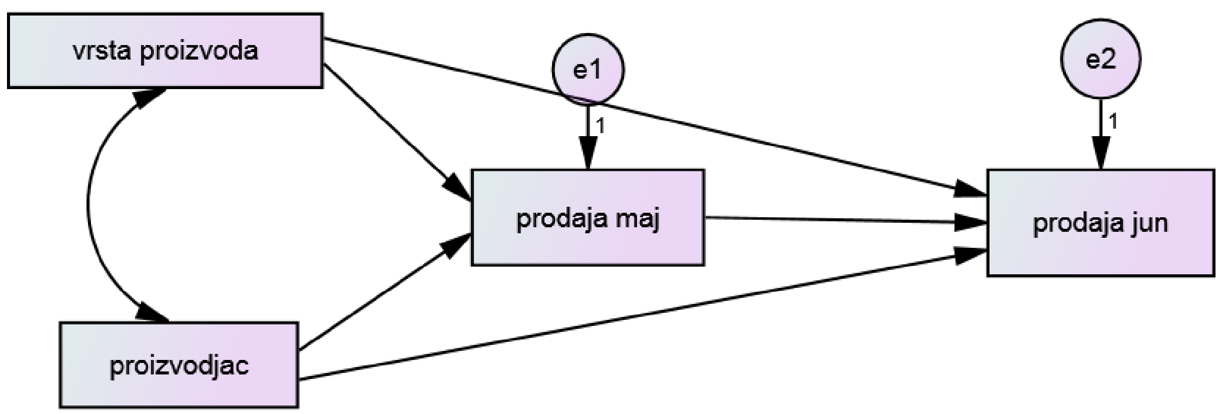
Узорак од 252 производа се сматра довољним за успешно обављање истраживања с обзиром да Бајесова статистика није заснована на великим узорцима (Schoot and Depaoli,

[122] http://www.calivita.rs/

2014). Многи чланци такође приказују предности Бајесове статистике у смислу малог скупа података (нпр. Zhang et al., 2007; Lee and Song, 2004).

Примена софтверског пакета IBM SPSS Amos *Version* 21 омогућила је оцену прикупљених података применом метода максималне веродостојности и Бајесовог метода, а након тога и извођење закључака поређењем резултата спроведеног истраживања. IBM SPSS Amos примењује општи приступ анализи података познат као структурално моделирање једначина (енгл. *Structural Equation Modeling*-SEM), а такође је познат и као анализа коваријансе структуре или узрочно моделирање.

Да би се извршила анализа, одабрани су подаци *продаја мај* и *продаја јун* као посматране, ендогене променљиве, које су условљене са две посматране, егзогене променљиве: *врста производа* и *произвођач*. На Слици 6. приказан је модел за одабране податке.



Слика 6. Структурални модел[123]

Овако дефинисан модел указује на потребу испитивања утицаја врсте производа и произвођача на продају у мају и јуну месецу. Истовремено се испитује и међусобни утицај, односно корелација између променљивих врста производа и произвођач.

[123] Истраживање аутора

4.2.1 Резултати истраживања добијени применом модела максималне веродостојности

Применом одабраних процедура за оцену посматраних података методом максималне веродостојности добијени су резултати приказани у Табели 6.

Табела 6. Оцењене вредности применом метода максималне веродостојности

Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
pr_maj	<---	proiz	-3,730	0,792	-4,709	***	
pr_maj	<---	vrpr	-0,941	1,502	-0,626	0,531	
pr_jun	<---	proiz	-1,090	,465	-2,346	0,019	
pr_jun	<---	pr_maj	0,730	,035	20,581	***	
pr_jun	<---	vrpr	-0,033	,845	-0,039	0,969	

Standardized Regression Weights: (Group number 1 - Default model)

			Estimate
pr_maj	<---	proiz	-0,308
pr_maj	<---	vrpr	-0,041
pr_jun	<---	proiz	-0,097
pr_jun	<---	pr_maj	0,784
pr_jun	<---	vrpr	-0,002

Means: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
vrpr	6,996	0,142	49,276	***	
proiz	6,246	0,269	23,200	***	

Intercepts: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
pr_maj	69,800	10,095	6,914	***	
pr_jun	11,827	6,191	1,910	0,056	

Covariances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
proiz <--> vrpr	3,926	0,654	6,000	***	

Correlations: (Group number 1 - Default model)

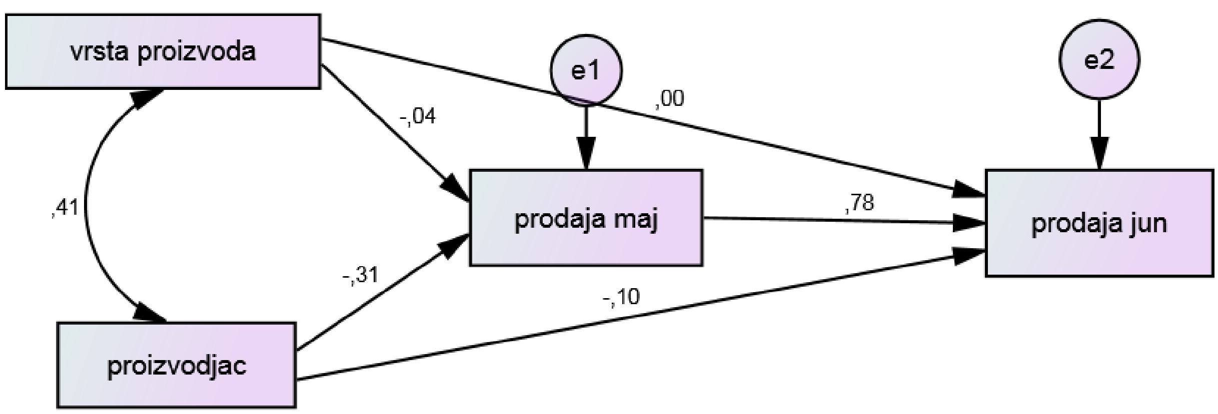
	Estimate
proiz <--> vrpr	0,409

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
proiz	18,193	1,624	11,203	***	
vrpr	5,060	0,452	11,203	***	
e1	2386,501	213,029	11,203	***	
e2	754,074	67,312	11,203	***	

Извор: Истраживање аутора

Након израчунавања стандардизованих оцена (енгл. *Standardized Estimates*), на дијаграму ће бити приказане следеће вредности:

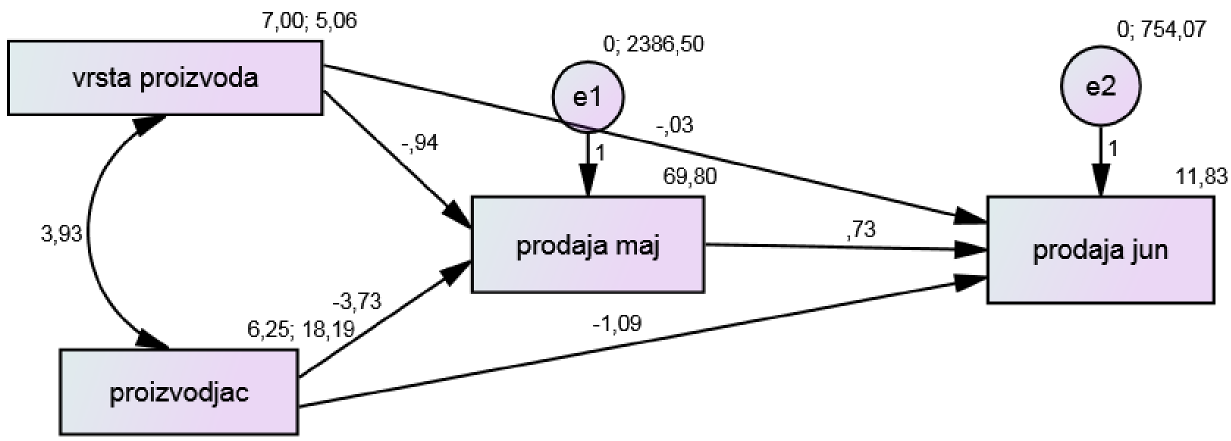


Слика 7. Структурални модел са стандардизованим оценама[124]

[124] Истраживање аутора

Вредност 0,41 представља корелацију између *врста производа* и *произвођач*. Вредности -0,04; -0,31; -0,10; 0,00 (према Табели 6. ова вредност износи -0,002) и 0,78 су станардизовани регресиони пондери (енгл. *Standardized Regression Weights*).

Уколико се изврши израчунавање нестандардизованих оцена (енгл. *Unstandardized Estimation*), вредности добијених резултата биће следеће:



Слика 8. Структурални модел са нестандардизованим оценама[125]

Вредности са Слике 8. презентоване су и у Табели 6, где су распоређене према значењу. Из табеле и са графикона може се очитати оцењена вредност коваријансе између *врста производа* и *произвођач* у висини од 3,93. Одмах поред коваријансе, у табели је у *S.E.* колони приказана оцена стандардне грешке коваријансе у висини од 0,654. Оцена 3,93 је посматрање приближно нормално дистрибуиране случајне променљиве центриране око коваријансе популације са стандардном девијацијом од 0,654. Поред стандардне грешке налази се *C.R.* колона која представља вредности критичног рациа који се утврђује као количник оцењене коваријансе и њене стандардне грешке (*C.R.*=3,926/0,654=6,00). Колона *P* показује приближну двосмерну *p* вредност за тестирање нулте хипотезе да је вредност параметра у популацији нула. Ознака *** говори о статистичкој значајности односа посматраних променљивих.

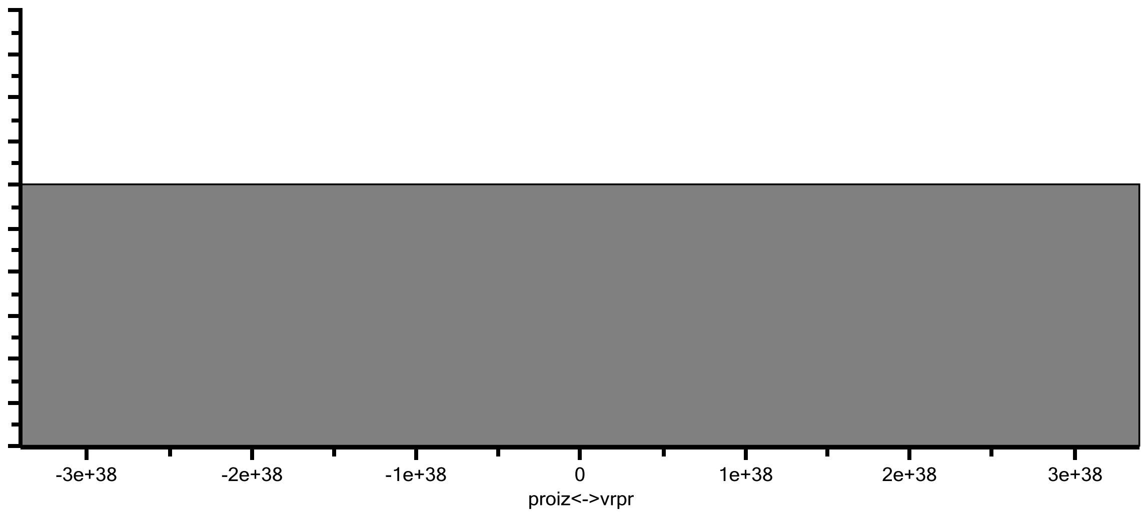
[125] Ibid.

4.2.2 Резултати истраживања добијени применом Бајесовог метода

Како би се спровело истраживање на одабраним подацима применом Бајесових метода оцењивања, прво је неопходно одабрати одговарајућу априорну расподелу. У многим случајевима, одабрана априорна расподела садржаће веома мало информација, тако да ће се закључци изводити само на основу података. Такве информације називају се *дифузне* или *неинформативне* априорне информације (в. 3.4.3).

У ужем смислу, међутим, ниједна априорна расподела није увек потпуно неинформативна, чак ни униформна расподела коју IBM SPSS Amos користи као подразумевану за сваки параметар, с обзиром да свака априорна расподела носи са собом бар неке информације. Како величина скупа података расте, тако докази из података преплављују ове информације, па се утицај априорне расподеле смањује. Осим у случају ако су модел и априорна расподела противречни подацима, може се приметити да се одговори добијени из Бајесове анализе мењају уколико се априорна расподела мења. У анализи коју је аутор спровео, видеће се да промена априорне расподеле утиче на резултате истраживања у смислу побољшања њихове прецизности.

Код Бајесовог оцењивања се верује да сваки параметар има априорну расподелу која укључује и (не)извесност о вредности тог параметра. Неинформативна априорна расподела се користи како би се приказао велики удео неизвесности у популацији којој припада посматрани параметар. У обрађеном примеру, првобитно одабрана априорна расподела јесте униформна априорна расподела која има карактер неинформативне расподеле, а резултати добијени применом ове расподеле приказани су у Табели 7. Слика 9. презентује изглед униформне расподеле само за параметре *врста производа-произвођач*, док се и за остале посматране параметре добијају слична графичка решења.



Слика 9. Униформна априорна расподела за *врста производа- произвођач*[126]

Табела 7. Оцењене вредности применом Бајесовог метода са униформном расподелом

	Mean	S.E.	S.D.	C.S.	Skewness	Kurtosis	Min	Max
Regression weights								
pr_maj<--vrpr	-0,945	0,01	1,521	1,000	-0,006	0,022	-7,908	5,381
pr_jun<--vrpr	-0,051	0,006	0,858	1,000	-0,015	0,06	-3,564	3,994
pr_maj<--proiz	-3,721	0,006	0,801	1,000	0,003	0,014	-7,02	-0,066
pr_jun<--proiz	-1,084	0,003	0,471	1,000	0,007	0,03	-3,144	0,842
pr_jun<--pr_maj	0,73	0	0,036	1,000	0,005	0,027	0,583	0,898
Means								
proiz	6,247	0,003	0,273	1,000	-0,006	0,024	5,055	7,442
vrpr	6,995	0,001	0,144	1,000	0,012	0,041	6,403	7,723
Intercepts								
pr_jun	11,897	0,039	6,274	1,000	0,012	0,067	-15,053	40,312
pr_maj	69,774	0,074	10,227	1,000	0,013	0,036	24,371	118,019
Covariances								
vrpr<->proiz	4,036	0,005	0,683	1,000	0,258	0,137	1,368	7,302
Variances								
vrpr	5,205	0,003	0,472	1,000	0,374	0,265	3,719	7,808

[126] Ibid.

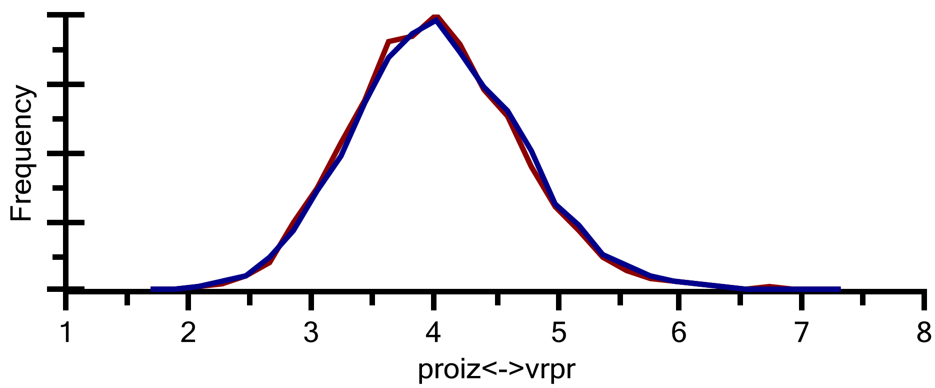
proiz	18,712	0,014	1,706	1,000	0,399	0,301	13,201	27,83
e1	2452,702	1,515	223,096	1,000	0,391	0,368	1705,556	3696,974
e2	779,255	0,684	70,795	1,000	0,355	0,213	546,868	1136,97

Извор: Истраживање аутора

Сваки ред табеле описује маргиналну апостериорну расподелу за сваки параметар модела. Прва колона, означена са *Mean*, означава апостериорну средину која је центар или просек апостериорне расподеле. Ова вредност се може користити као Бајесова тачкаста оцена параметра која је заснована на подацима и априорној расподели. За униформну априорну расподелу апостериорна средина биће близу вредности оцене добијене методом максималне веродостојности, што се потврђује и у нашем примеру где је апостериорна средина 4,036 за коваријансу *врста производа-произвођач*, а оцена добијена методом максималне веродостојности 3,926.

Колона *C.S.* представља укупну статистику конвергенције и како се вредност ове статистике приближава 1,000, што је случај и у нашем примеру, нека већа прецизност се не може постићи даљим узорковањем, па се може стати са итерацијама. Осим ове статистике и дијаграмом апостериорне расподеле може се проверити конвергенција Бајесових MCMC метода које се користе у поступку итерације.

Да би се проценило да ли су успешно идентификоване значајне карактеристике апостериорне расподеле коваријансе *врста производа-произвођач* неопходно је симултано приказати две оцене расподеле, једну добијену из прве трећине акумулираних узорака и другу добијену из последње трећине, што је и приказано црвеном линијом на следећој слици:



Слика 10. Апостериорна расподела[127]

Са графичког приказа се грубо може оценити да је апостериорна расподела центрирана близу вредности 4, што одговара вредности апостериорне средине овог параметра *Mean* која износи 4,036. На овај начин приказан је полигон фреквенција за расподелу коваријансе *врста производа-произвођач* за преко 80.500 узорака.

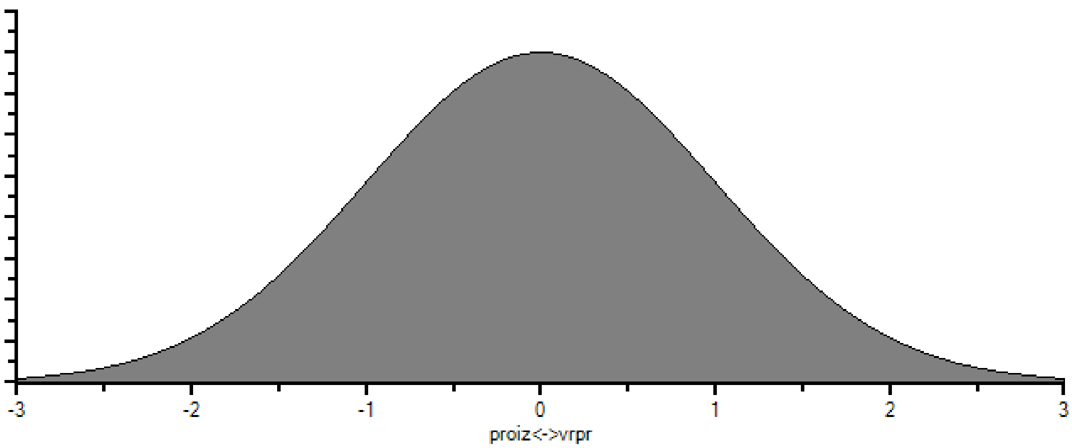
Да би се извео закључак о резултатима добијеним применом метода максималне веродостојности и Бајесовог метода заснованог на униформној, неинформативној расподели неопходно је упоредити апостериорну стандардну девијацију Бајесове статистике, односно стандардну девијацију расподеле (у ознаци *S.D.*) са стандардном грешком класичне статистике (у ознаци *S.E.*), која представља корисну меру неизвесности. Као што је случај са оцењеном вредности апостериорне средине, компарацијом Табела 6 и 7, може се закључити и да су вредности стандардних грешака осталих параметара добијених применом метода класичне статистике вредносно врло блиски вредностима стандардне девијације истих параметара добијених применом Бајесових метода. На овај начин је потврђено да се, у случају да су априорне информације дифузне, односно неинформативне, резултати класичне и Бајесове статистике веома мало разликују.

Да би се одговорило на постављену хипотезу да *оцене добијене на основу Бајесовог приступа представљају унапређење особина оцена добијених класичним приступом у*

[127] Истраживање аутора

смислу њихове прецизности, настављено је даље са истраживањем како би се показало да промена априорне расподеле доводи и до промене резултата Бајесове статистике.

У циљу доказивања претходне тврдње, уместо провобитно изабране униформне расподеле, сада се одабира нормална расподела за априорну расподелу, која по својим особинама има карактер информативне априорне расподеле (в. 3.4.4). Априорна расподела која у значајној мери садржи извесност о параметру популације назива се информатина расподела. Ове расподеле садрже искључиво нумеричке информације које су кључне за оцењивање модела и могу имати велики утицај на коначне оцене.[128] На Слици 11. приказана је нормална априорна расподела за одабрани параметар *врста производа-произвођач*.



Слика 11. Нормална априорна расподела за *врста производа- произвођач* [129]

Након тога, израчунате вредности Бајесове статистике са информативном априорном расподелом биће следеће:

Табела 8. Оцењене вредности применом Бајесовог метода са нормалном расподелом

Regression weights	Mean	S.E.	S.D.	C.S.	Skewness	Kurtosis	Min	Max
pr_maj<--vrpr	-0,657	0	0,752	1	0,009	0,016	-4,202	2,761

[128] Schoot, R.V.D. and Depaoli, S. 2014. „Bayesian analyses: where to start and what to report“, *The European Health Psychologist*, Volume 16, Issue 2, pp. 79

[129] Истраживање аутора

pr_jun<--vrpr	-0,075	0	0,591	1	-0,002	-0,006	-2,526	2,289
pr_maj<--proiz	-2,471	0	0,59	1	0,006	0,01	-4,894	-0,143
pr_jun<--proiz	-0,9	0	0,407	1	0,004	0,017	-2,661	0,877
pr_jun<--pr_maj	0,735	0	0,034	1	-0,019	0,035	0,591	0,896

Means

proiz	5,959	0	0,194	1	0	-0,015	5,122	6,828
vrpr	6,837	0	0,128	1	-0,009	0,009	6,246	7,369

Intercepts

pr_jun	10,7	0,04	4,479	1	-0,01	0,021	-9,471	29,141
pr_maj	59,963	0,03	5,757	1	0,021	0,025	34,37	83,481

Covariances

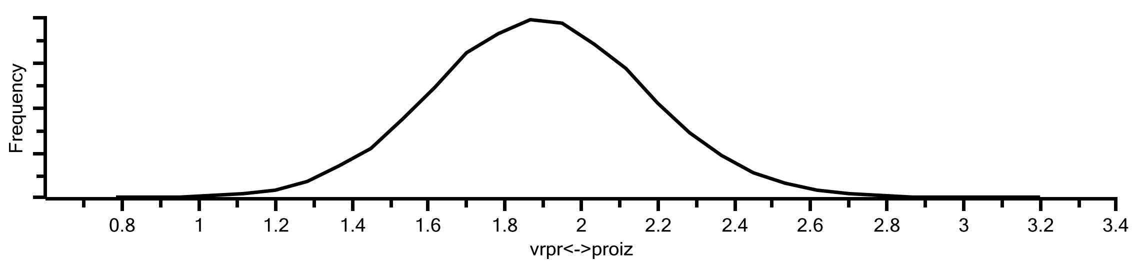
vrpr<->proiz	1,887	0	0,279	1	0,043	0,052	0,574	3,097

Variances

vrpr	4,194	0	0,303	1	0,247	0,116	3,016	5,933
proiz	10,044	0	0,481	1	0,124	0,004	7,973	12,273
e1	2464,09	1,45	203,1	1	0,283	0,144	1779,8	3538,89
e2	765,691	0,49	55,989	1	0,198	0,049	574,69	1053,22

Извор: Истраживање аутора

Вредности израчунате и графички приказане апостериорне расподеле одговараће новодобијеним вредностима, па ће ова расподела бити центрирана близу вредности 1,9, што одговара вредности апостериорне средине овог параметра *Mean* која износи 1,887.



Слика 12. Апостериорна расподела[130]

[130] Истраживање аутора

Да би се извео коначан закључак и потврдила или одбацила хипотеза неопходно је упоредити резултате добијене применом метода класичне и Бајесове статистике. Пореде се вредности *S.E.*, добијене методом максималне веродостојности, и *S.D.*, добијене применом Бајесових метода. Посматрају се резултати презентовани у Табели 6. и Табели 8.

За разлику од случаја када је била одабрана униформна расподела за априорну расподелу и када није постојала значајна разлика између добијених резултата, у случају када је нормална расподела одабрана за априорну, ситуација се мења. Сада је апостериорна средина 1,887 за коваријансу *врста производа-произвођач*, а оцена добијена методом максималне веродостојности 3,926. Ако упоредимо наведене две вредности за све остале параметре, закључак је исти: резулати добијени применом Бајесових метода су прецизнији у односу на резултате добијене применом метода класичне статистике. На тај начин је и ***потврђена*** постављена хипотеза да *оцене добијене на основу Бајесовог приступа представљају унапређење особина оцена добијених класичним приступом у смислу њихове прецизности*.

4.3 Даље могућности компарације Бајесове и класичне статистике на примеру одабраних временских серија

Један од циљева истраживања докторске дисертације јесте анализа односа два приступа статистичком закључивању и могућност примене Бајесових метода у обради економских података у односу на класичну статистику. Такође, у оквиру постављених хипотеза дефинисана је и хипотеза која се односи на оцену поузданости одабраних метода предвиђања кроз квалитативну процену резултата добијених применом Бајесове статистике. У те сврхе, одабрани су подаци са сајта Народне банке Србије[131] и Eurostat-a[132] како би се на наведене захтеве одговорило анализом одабраних временских серија. Применом софтверског пакета RATS[133] (енгл. *Regression Analysis of Time Series*) *Version*

[131] http://www.nbs.rs/

[132] http://epp.eurostat.ec.europa.eu/

[133] https://www.estima.com/ratsdemo.shtml

8.0 и 8.3 истовремено ће бити приказани резултати добијени применом класичних и Бајесових метода и на основу постављених критеријума изведени коначни закључци.

4.3.1 Преглед владајућих ставова

За показатеље који су обухваћени истраживањем, успешно се могу користити Бајесове методе заједно са различитим моделима. Ови модели су развијени како би се обратила пажња на чињеницу да већина питања која су предмет интересовања истраживача укључује више променљивих, па се морају решавати помоћу метода мултиваријационих временских серија. Много различитих модела мултиваријационих временских серија је коришћено у макроекономији, али од рада Симса из 1980. године, „*Макроекономија и реалност*“, векторски ауторегресивни модели (енгл. *Vector Autoregressive* (VAR) *models*) су међу најпопуларнијим.

Векторски ауторегресивни модели су коришћени посебно у економским предвиђањима за сродне јединице посматрања, на пример, запосленост у секторима индустрије или по регионима, за заједничке зависне серије (незапосленост и производња), као и у анализама историјских флуктуација (Ritschl and Woitek, 2000). Ови модели укључују само унапред одређене варијабле као предикторе, избегавајући спецификацију ендогене зависности (Bauwens and Lubrano, 1995).

Развој Бајесовог приступа довео је до увођења информативне Минесота априорне информације (Doan, Litterman and Sims, 1984; Litterman, 1986), као и општијих априорних информација у VAR моделе (Sims and Zha, 1998) и векторске ARMA моделе (Ravishanker and Ray, 1997).

У новије време, осим класичних VAR модела за предвиђање будућих кретања све више се користе Бајесови VAR модели, односно BVAR модели. Тако Бергер и Оштерхолм (Berger and Österholm, 2008), користећи Бајесов VAR приступ, закључују да укључивање новчане масе омогућава боље предвиђање кретања инфлације, док Фишер, Ленца, Пил и Рајхлин (Fischer, Lenza, Pill and Reichlin, 2008) показују да биваријантне монетарне прогнозе, уз помоћ BVAR модела, доносе нове информације у општа макроекономска предвиђања.

Ставрев и Бергер (Berger and Stavrev, 2012) својим чланком доприносе расправи о улози новца у монетарној политици анализом информационог садржаја новца у предвиђању инфлације еврозоне. У ту сврху, извршили су поређење предиктивних перформанси унутар и између различитих класа структурних и емпиријских модела уз помоћ Бајесових и других техника оцењивања. Афонсо и Суза (Afonso and Sousa, 2012) проучавају макроекономске утицаје фискалне политике на GDP применом BVAR модела, анализирајући емпиријске податке из САД-а, Велике Британије, Немачке и Италије. Посебан допринос овог чланка је у коришћењу квартални фискалних података, што је обезбедило добијање прецизнијих резултата када су у питању ефекти фискалне политике.

Водећи се искуством Афонсоа и Сузе, касније ће у раду бити узети у разматрање квартални подаци за одабране серије података са подручја Европске Уније, где ће се такође потврдити да се на основу таквих података добијају прецизнији резултати предвиђања.

Буијан (Bhuiyan, 2012) у свом раду развија Бајесов VAR модел за оцену ефеката шокова монетарне политике у случају Канаде. Бисвас, Синг и Сина (Biswas, Singh and Sinha, 2010) покушали су да конструишу модел за прогнозирање инфлације у индијској економији, применом BVAR модела. Главна предност примене овог модела је укључивање априорне информације како би се побољшале перформансе модела. На основу поређења резултата предвиђања добијених применом VAR и BVAR модела, мерено вредношћу одабране мере прецизности (RMSE), утврђено је да BVAR модел даје боље резултате од VAR модела. До сличних закључака долази и Карајани (Caraiani, 2010), предвиђајући кретање GDP-а у Румунији применом VAR и BVAR модела. Наиме, и у наведеном раду се потврђују наводи из литературе да Бајесове методе дају прецизније резултате од класичних метода.

Бичемин (Beauchemin, 2011) закључује да макроекономски прогнозери располажу са моћним алатом, Бајесовим векторским ауторегресивним моделом, којим се може обухватити сложена и динамична природа одабраних привреда. Миљардо (Migliardo, 2010) у свом раду предлаже примену BVAR модела за испитивање краткорочних ефеката шокова монетарне политике на италијанску привреду. Боићук (Boiciuc, 2014) настоји да својим радом допринесе анализи макроекономских ефеката фискалне политике у земљама источне Европе. При томе, за анализу користи емпиријске податке из Румуније, Бугарске, Чешке и Мађарске за период од првог квартала 2000. до трећег квартала 2013. године. И овде се потврђује да се употребом кварталних података добијају прецизнији резултати истраживања, конкретно резултати о ефектима фискалне политике.

Узимајући у обзир актуелност материје која је предмет истраживања, у раду су коришћене две групе података како би се испитала оправданост примене Бајесових метода у економији. До резултата истраживања дошло се поређењем две врсте BVAR модела, *униваријантног* BVAR *модела* (енгл. *Univariate* BVAR) и *једноставног* BVAR *модела* (енгл. *Simple* BVAR), који су засновани на различитим априорним информацијама и два класична модела, униваријантног OLS модела (енгл. *Univariate* OLS) и OLS VAR модела.

4.3.2 Примењени модели у истраживању - VAR, BVAR и OLS модел

За класични VAR приступ је карактеристичан губитак степени слободе који се експоненцијално смањује у односу на број укључених лагова (енгл. *lags*).[134] Бајесов приступ предлаже решење за овај проблем стављањем акцента на коришћење априорних расподела за параметре, па априорна расподела постаје кључни фактор у BVAR приступу (енгл. *Bayesian Vector Autoregressive*).[135]

Бајесов приступ оцењивању узима у обзир стварну структуру популације као неизвесну и не додаје никакву посебну *тежину* било којој вредности параметра посматраног модела. Уместо тога, поменути приступ укључује ову неизвесност у обрачун у облику априорне расподеле вероватноће за параметре. Степен неизвесности представљен овом априорном расподелом може бити измењен информацијама садржаним у подацима, уколико су два

---

[134] У временским серијама, лагови представљају прошле вредности, односно вредности променљивих које су постојале пре текућег посматрања.

[135] Caraiani, P. 2010. „Forecasting Romanian GDP using a BVAR Model“, *Romanian Journal of Economic Forecasting*, 4/2010, pp. 77

извора информација различита. Као резултат, BVAR приступ даје боља предвиђања од редукованог облика VAR модела оцењеног на класичан начин.[136]

За дати скуп од n променљивих посматране временске серије $y_t = (y_{1t},...,y_{nt})'$, стандардни VAR модел може се написати на следећи начин:[137]

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + \mu + \varepsilon_t, \tag{77}$$

где је $A = \lfloor A_1;....;A_p \rfloor$, A_i матрица коефицијената димензије $(n \times n)$, y_t је $(n \times 1)$ вектор ендогених променљивих које су нестационарне, μ је $(n \times 1)$ вектор константних коефицијената, а ε_t је $(n \times 1)$ је вектор грешака, независно идентичних и нормално распоређених.[138]

Класично оцењивање VAR модела може довести до непрецизних оцена које ће одговарати подацима само због великог броја укључених променљивих. Овај проблем је у литератури познат као проблем „вишег уклапања“ (енгл. *overfitting*). Заправо, број параметара које треба оценити геометријски расте са бројем променљивих (n) и пропорционалан је броју лагова (*p*).[139] Како би се превазишао овај проблем Литерман (Litterman, 1980) предлаже увођење Бајесовог оцењивања, пошто се сматра да је Бајесов приступ користан, јер се не може знати да ли су неки коефицијенти статистички значајни или не. На тај начин је могуће повезати расподелу вероватноће за вектор параметра, а затим оцењене резултате као производ априорне расподеле и информација из података.

[136] Ciccarelli, M. and Rebucci, A. 2003. „Bayesian VARs: A Survey of the Recent Literature with an Application to the European Monetary System”, IMF *Working Papers* 03/102, International Monetary Fund, pp. 3

[137] Lütkepohl, H. 2007. „Econometric Analysis with Vector Autoregressive Models”, EUI *Working papers*, ECO 2007/11, pp. 9

[138] Lütkepohl, H. 1999. „Vector autoregressive analysis”, *Discussion Papers, Interdisciplinary Research Project* 373: Quantification and Simulation of Economic Processes, No. 1999, 31, pp. 3

[139] Ciccarelli, M. and Rebucci, A. 2003. „Bayesian VARs: A Survey of the Recent Literature with an Application to the European Monetary System”, IMF *Working Papers* 03/102, International Monetary Fund, pp. 4

Конкретније, у одређивању априорних информација истиче се значај три статистичке законитости за макроекономске временске серије:[140]

1. већина макроекономских временских серија је дефинисана трендом;
2. чињеница је да новије вредности серија обично садрже више информација о текућој вредности серија, него претходне вредности;
3. чињеница је да претходне вредности посматране променљиве садрже више информација о тренутном стању варијабле, него прошле вредности других варијабли.

Следбеници Бајесовог приступа, могу навести ове законитости додавањем расподеле вероватноће параметрима, на начин да је:

1. средина коефицијената додељена свим лаговима осим првом једнака нули;
2. варијанса коефицијената инверзно зависна од броја лагова; и
3. коефицијент променљиве j у једначини g додељује нижу априорну варијансу од оне коју додељује променљива g.

Стандардне априорне информације имају следеће карактеристике:[141]

1. за детерминистичке променљиве, априорне информације су неинформативне;
2. за лагове ендогених променљивих, априорне информације су независне и нормално распоређене;
3. по правилу, априорна средина за први лаг зависне променљиве у свакој једначини биће један.

[140] Litterman, R.B. 1986. „Forecasting with Bayesian vector autoregressions: Five years of experience“, *Journal of Business and Economic Statistics* 4, pp. 28

[141] Doan, T. 2007. RATS User’s Manual, *Version* 7, Estima, pp. 378

Поред тога, потребно је још одредити и априорну информацију за варијансу. Стандардна грешка оцењеног коефицијента за лаг l променљиве j у једначини i биће представљена помоћу стандардне девијације облика $S(i,j,l)$:[142]

$$S(i,j,l) = \frac{[\gamma g(l) f(i,j)] s_i}{s_j}, \tag{78}$$

где је $f(i,j) = 1$, ако је $i = j$ и $f(i,j) = w_{ij}$, ако је $i \neq j$.

На тај начин, комплетна априорна расподела може бити дефинисана, уколико се одреди вредност за хиперпараметар γ и дефинишу функције $g(l)$ и $f(i,j)$. Хиперпараметар γ познат је у литератури као „укупна затвореност“ (енгл. *overall tightness*) априорне информације. Затвореност једног лага у односу на лаг l одређена је функцијом $g(l)$. Како се дужина лага повећава, тако се повећава и затвореност око априорне средине. Функција $f(i,j)$ одређује затвореност априорне информације променљиве j у односу на променљиву i у једначини за променљиву i.

Бајесово предвиђање засновано на векторској ауторегресији и структурним априорним расподелама представља значајно развојно подручје и важан огранак ширег контекста Бајесовог предвиђања.[143]

За приказивање Бајесовог приступа оцењивању, једначина (77) може се написати у компактнијем облику;

$$y_t = X_t \beta + \varepsilon_t \qquad t = 1, \ldots, T, \tag{79}$$

где је β непознат параметар модела.

[142] Caraiani, P. 2010. „Forecasting Romanian GDP using a BVAR Model“, *Romanian Journal of Economic Forecasting*, 4/2010, pp. 78

[143] West, M. and Harrison, J. 1997. *Bayesian Forecasting and Dynamic Models, Second Edition*, Springer-Verlag, New York, Inc., pp. 626

Бајесово оцењивање претходне једнакости је једноставно и функционише на следећи начин: за дату функцију густине вероватноће података, која је условљена параметрима модела (информацијом садржаном у подацима у облику функције веродостојности),[144]

$$L(Y|\beta,\Sigma) \propto |\Sigma|^{-T/2} \exp\left\{-\frac{1}{2}\sum_t (Y_t - X_t\beta)'\Sigma^{-1}(Y_t - X_t\beta\right\} \tag{80}$$

и заједничку априорну расподелу параметара, $p(\beta,\Sigma)$, заједничка апостериорна расподела параметара која је одређена подацима биће добијена преко Бајесовог правила:[145]

$$p(\beta,\Sigma|Y) = \frac{p(\beta,\Sigma)L(Y|\beta,\Sigma)}{p(Y)} \propto p(\beta,\Sigma)L(Y|\beta,\Sigma), \tag{81}$$

Према дефиницији условне вероватноће, заједничка функција података и параметара, $p(\beta,\Sigma,Y)$, може се написати као:

$$p(\beta,\Sigma,Y) = L(Y|\beta,\Sigma)\,p(\beta,\Sigma) = p(\beta,\Sigma|Y)\,p(Y). \tag{82}$$

За дато $p(\beta,\Sigma|Y)$ маргинална апостериорна расподела зависна од података, $p(\Sigma|Y)$ и $p(\beta|Y)$, може се добити интегрисањем β и Σ из $p(\beta,\Sigma|Y)$, респективно. На крају, локација и дисперзија $p(\Sigma|Y)$ и $p(\beta|Y)$ може се лако анализирати како би се добиле тачкасте оцене параметара и мере прецизности које су упоредиве са оним добијеним применом класичног приступа оцењивању.

У оквиру спроведеног истраживања, осим VAR и BVAR модела биће коришћен и OLS модел и то униваријантни OLS модел и OLS VAR модел.

[144] Функције густине вероватноће биће означене са p када представљају априорну или апостериорну расподелу параметара, а са L када представљају функцију веродостојности.

[145] Ciccarelli, M. and Rebucci, A. 2003. „Bayesian VARs: A Survey of the Recent Literature with an Application to the European Monetary System", IMF *Working Papers* 03/102, International Monetary Fund, pp. 6

OLS модел или ординарни модел најмањих квадрата (енгл. *Ordinary Least Squares*) може се представити на следећи начин:[146]

$$y = X\beta + \varepsilon, \tag{83}$$

где y представља актуелне вредности посматране променљиве, β је вектор регресионог коефицијента, X матрица предикторских променљивих, а ε вектор грешака.

Једноставна једначина OLS модела већ је посредно приказана код BVAR модела (79).

OLS модел се заснива на следећим претпоставкама:[147]

1. грешке су нормално распоређене са средином нула и константном варијансом;
2. грешке су независне; и
3. независне променљиве исказују се без грешке.

OLS модел се сматра најефикаснијим методом када су ове претпоставке задовољене, пошто за резултат има коефицијенте са минималном варијансом за све линеарне непристрасне оцене. Међутим, када подаци не испуњавају претпоставке OLS модела или уколико се јављају одређени проблеми, OLS регресиони коефицијент биће нетачан, што за резултат може имати недовољно добре оцене параметара модела.

OLS оцена биће идентична апостериорној средини која је одређена на основу неинформативних априорних информација. Међутим, не налази се свака апостериорна средина заснована на информативним априорним информацијама између априорне средине и OLS оцена. Разлог лежи у томе што је апостериорна средина матрица пондерисаног просека априорне средине и OLS оцене. Пондерисана матрица не

[146] Schmidheiny, K. 2013. „The Multiple Linear Regression Model“, *Short Guides to Microeconometrics*, Version: 12-9-2013, 20:29, pp. 2

[147] http://forrest.psych.unc.edu/research/vista-frames/pdf/chap08.pdf

подразумева да сваки индивидуални коефицијент лежи између његове априорне средине и OLS оцене.[148]

4.3.3 Преглед одабраних мера прецизности за поређење модела

Модели су упоређени на основу неколико статистика предвиђања којима се мери прецизност и то: средине грешке (енгл. *mean error*), средине просечне грешке (енгл. *mean average error*, MAE), корена средње квадратне грешке (енгл. *root mean square error*, RMSE) и Тејлове U-статистике (енгл. *Theil* U).

RMSE као једна од мера прецизности, којом се такође мери и неизвесност у предвиђању, може се добити на следећи начин:[149]

$$RMSE = \sqrt{\frac{\sum (X_t - F_t)^2}{m}} = \sqrt{\frac{\sum e_t^2}{m}}, \tag{84}$$

где X_t представља актуелне податке у периоду t, F_t предвиђање које је извршено у периоду t применом одређеног модела, e_t је грешка у предвиђању у периоду t, док је m број коришћених метода или опсервација за израчунавање RMSE.

RMSE за поређење посматраних серија може узети вредности у интервалу од 0 до +∞.

Тејлова U-статистика је релативна мера прецизности која се израчунава на следећи начин:[150]

[148] Koop, G. 2003. *Bayesian Econometrics*, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, England, pp. 50

[149] Makridakis, S. and Hibon, M. 1995. „Evaluating Accuracy (or Error) Measures", INSEAD *Working Paper Series*, 95/18/TM, INSEAD, Fontainebleau, France, pp. 3

[150] Ibid., pp. 8

$$U=\sqrt{\sum_{t=1}^{m}\frac{\left(\frac{X_t-F_t}{X_t}\right)^2/m}{\left(\frac{X_t-FN_t}{X_t}\right)^2/m}}, \tag{85}$$

где је FN_t предвиђање које служи као репер, на пример последње расположиве вредности или последње расположиве вредности након што је узет у обзир и сезонски карактер промењљивих.

Уколико је Тејлова U-статистика један, то значи да је прецизност одабраног модела једнака прецизности модела заснованог на *Naïve* предвиђању, које подразумева да се код стационарних временских серија предвиђање врши на основу историјске просечне вредности, док се код нестационарних временских серија за предвиђање користи стварна вредност из претходног периода. Уколико је вредност Тејлове U-статистике нижа од 1, тада је примењени метод бољи, односно уколико је вредност виша од 1, лошији од репер модела.

4.3.4 Резултати истраживања

Први део спроведеног истраживања применом софтверског пакета RATS, обухватио је финансијске показатеље и то: каматне стопе, кредите становништву, кредите привреди, депозите становништва и депозите привреде. На ове податке, који су преузети са сајта Народне банке Србије, примењене су Бајесове методе заједно са одговарајућим моделима.

4.3.4.1 Примена одабраних модела истраживања на податке из привреде Републике Србије

Одабране серије података: каматне стопе, кредити одобрени становништву, кредити одобрени привреди, депозити становништва и депозити привреде, међусобно су упоређени на основу четири модела: униваријантног OLS модела, униваријантног BVAR модела, једноставног BVAR модела и OLS VAR модела.

За моделе мале величине, према препоруци Doan (2007), одабране априорне информације треба да буду симетричне са укупном затвореношћу од $\gamma = 0{,}20$ и релативном тежином од $\omega = 0{,}5$. За стандардне OLS и UVAR моделе релативна тежина пада скоро на 0, док је препоручена вредност у литератури 0,001.[151]

С обзиром да се објашњени модели могу успешно користити у обради финансијских показатеља и предвиђању њиховог будућег кретања, у раду су коришћени подаци са сајта Народне банке Србије[152] и то подаци о кретању каматних стопа на девизне кредите и депозите привреде и становништва, новоодобреним девизним кредитима становништву и привреди, као и новоположеним девизним депозитима становништва и привреде. Обухваћени су подаци од јануара 2011. до марта 2014. године по месецима (в. Прилог 2). Приказани подаци су ради даље обраде трансформисани логартимовањем, тако да су надаље у раду коришћене ознаке: LOG_KAM (логаритамска серија-каматне стопе), LOG_KS (логаритамска серија-кредити становништву), LOG_KP (логаритамска серија-кредити привреди), LOG_DS (логаритамска серија-депозити становништва) и LOG_DP (логаритамска серија-депозити привреде) за посматране серије података.

За посматране моделе одабране су следеће вредности за укупну затвореност и релативну тежину:

Табела 9. Одабране вредности модела

Модел	Укупна затвореност (енгл. *tightness*) γ	Релативна тежина (енгл. *relative weight*) ω
униваријантни OLS модел	2.0	0.001
униваријантни BVAR модел	0.1	0.001
једноставни BVAR модел	0.1	0.5
OLS VAR модела	2.0	1.0

Извор: Истраживање аутора

[151] Ciccarelli, M. and Rebucci, A. 2003. „Bayesian VARs: A Survey of the Recent Literature with an Application to the European Monetary System", IMF *Working Papers* 03/102, International Monetary Fund, pp. 80

[152] http://www.nbs.rs/

Оптималан број лагова за наведене моделе одређен је на основу Акаике и Бајес информационих критеријума. Оба наведена информациона критеријума указују на то да су за ово истраживање најбољи модели засновани на једном лагу, што показују резултати из Табеле 10.

Табела 10. Оптималан број лагова према посматраним информационим критеријумима

Број лагова	AIC	Значајност	BIC	Значајност
0	-470.64899		-463.22728	
1	-615.77470	*	-582.03093	*
2	-599.66527		-565.88583	
3	-498.88048		-512.32498	
4	-314.18664		-473.48379	

Извор: Истраживање аутора

4.3.4.2 Статистика предвиђања одабраних серија података

У наредним табелама приказана је статистика предвиђања за сваку посматрану серију, применом сва четири модела, а поређење је извршено на основу одабраних мера прецизности.

Табела 11. Униваријантни OLS модел – статистика предвиђања за серију LOG_KAM

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0072424	0,0185943	0,0253322	0,9847	15
2	-0,0176627	0,0361322	0,0461807	0,9657	14
3	-0,0323574	0,0511151	0,0620288	0,9374	13
4	-0,0477001	0,0631205	0,0741281	0,9074	12
5	-0,0628306	0,0712068	0,0824671	0,8711	11
6	-0,0724595	0,0763596	0,0890779	0,8447	10
7	-0,0796271	0,0839378	0,09832	0,8339	9
8	-0,0889088	0,0945848	0,1104684	0,8271	8
9	-0,1016977	0,1090697	0,1239454	0,8187	7
10	-0,1192381	0,1220724	0,1361167	0,8034	6
11	-0,135998	0,135998	0,1439432	0,7855	5
12	-0,1389564	0,1389564	0,1428146	0,7648	4

Извор: Истраживање аутора

Табела 12. Униваријантни BVAR модел – статистика предвиђања за серију LOG_KAM

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0072815	0,0184566	0,0251232	0,9766	15
2	-0,0178246	0,035952	0,0459277	0,9604	14
3	-0,0327608	0,0510287	0,0618523	0,9347	13
4	-0,0485275	0,0633527	0,074163	0,9078	12
5	-0,0642368	0,0723686	0,0828902	0,8755	11
6	-0,0742848	0,0779075	0,0899136	0,8526	10
7	-0,0818143	0,0858202	0,0996443	0,8451	9
8	-0,0915362	0,0968859	0,1125491	0,8427	8
9	-0,1048926	0,1119241	0,1270097	0,8389	7
10	-0,123151	0,1256391	0,1402239	0,8276	6
11	-0,1403237	0,1403237	0,1486614	0,8113	5
12	-0,143061	0,143061	0,1473363	0,789	4

Извор: Истраживање аутора

Табела 13. Једноставни BVAR модел – статистика предвиђања за серију LOG_KAM

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0054321	0,0187357	0,0244006	0,9485	15
2	-0,0150289	0,0354744	0,0445826	0,9323	14
3	-0,0297245	0,0485467	0,059604	0,9008	13
4	-0,0459801	0,0606759	0,0710615	0,8698	12
5	-0,06207	0,069909	0,0790634	0,8351	11
6	-0,0725426	0,0758219	0,0863244	0,8186	10
7	-0,0804326	0,0840437	0,0971103	0,8236	9
8	-0,0908997	0,0958002	0,1118149	0,8372	8
9	-0,105772	0,112293	0,128618	0,8495	7
10	-0,1261529	0,1280575	0,1439065	0,8493	6
11	-0,1446715	0,1446715	0,1537927	0,8393	5
12	-0,1457119	0,1457119	0,1503742	0,8052	4

Извор: Истраживање аутора

Табела 14. OLS VAR модел – статистика предвиђања за серију LOG_KAM

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,0037329	0,0214951	0,0261852	1,0179	15
2	0,000278	0,0382944	0,0448051	0,9369	14
3	-0,0079674	0,0498891	0,0551201	0,833	13
4	-0,018448	0,0524521	0,057218	0,7004	12
5	-0,0264065	0,0465364	0,0521659	0,551	11
6	-0,0311899	0,0438732	0,0484904	0,4598	10

7	-0,0336264	0,0491644	0,0550495	0,4669	9
8	-0,0393554	0,0590325	0,0670845	0,5023	8
9	-0,0525293	0,0714293	0,083401	0,5509	7
10	-0,0711236	0,08054	0,0948293	0,5597	6
11	-0,1004976	0,1004976	0,108333	0,5912	5
12	-0,1058538	0,1058538	0,1091369	0,5844	4

Извор: Истраживање аутора

У Табелама 11-14. приказани су резултати прогнозирања за серију LOG_KAM. Модели су упоређени на основу неколико статистика предвиђања, а изведени су следећи закључци:

- код униваријантног OLS модела (Табела 11), код првог корака, корен средње квадратне грешке RMSE износи 0,0253322, а Тејлова U-статистика 0,9847;
- код OLS VAR модела (Табела 14), код првог корака, корен средње квадратне грешке износи 0,0261852, а Тејлова U-статистика 1,0179;
- код униваријантног BVAR модела (Табела 12), RMSE и Тејлова U-статистика износе 0,0251232 и 0,9766;
- једноставни BVAR модел (Табела 13) дао је вредност од 0,0244006 за корен средње квадратне грешке и 0,9485 за Тејлову U-статистику;
- имајући у виду да се бољим сматра онај модел чије су вредности одабраних мера прецизности ниже, а у овом случају су Бајесове методе, односно униваријантни и једноставни BVAR модел, дале боље резултате од два класична модела, закључујемо да је *потврђена помоћна хипотеза* да оцене добијене на основу Бајесовог приступа представљају унапређење особина оцена добијених класичним приступом у смислу њихове прецизности.

Табела 15. Униваријантни OLS модел – статистика предвиђања за серију LOG_KS

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0114879	0,0174406	0,0212028	1,3158	15
2	-0,0221265	0,027386	0,0341362	1,175	14
3	-0,0307603	0,0328591	0,0408285	1,0624	13
4	-0,0366675	0,0367151	0,0458695	1,0062	12
5	-0,041501	0,041501	0,0507272	0,9527	11
6	-0,0462186	0,0462186	0,0553244	0,8941	10

7	-0,051006	0,051006	0,0596359	0,8408	9
8	-0,0555704	0,0555704	0,0637067	0,8055	8
9	-0,061671	0,061671	0,0684466	0,7647	7
10	-0,0671965	0,0671965	0,0732971	0,7354	6
11	-0,0730073	0,0730073	0,0788783	0,7026	5
12	-0,0809192	0,0809192	0,0861726	0,666	4

Извор: Истраживање аутора

Табела 16. Униваријантни BVAR модел – статистика предвиђања за серију LOG_KS

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0092062	0,0127731	0,0161972	1,0052	15
2	-0,0189302	0,0222787	0,027419	0,9438	14
3	-0,0273554	0,0273554	0,0337994	0,8795	13
4	-0,0332175	0,0332175	0,038233	0,8387	12
5	-0,0383574	0,0383574	0,0431681	0,8107	11
6	-0,0436826	0,0436826	0,0486195	0,7857	10
7	-0,049062	0,049062	0,0540257	0,7617	9
8	-0,0538682	0,0538682	0,0586417	0,7415	8
9	-0,0606438	0,0606438	0,0644742	0,7203	7
10	-0,0668609	0,0668609	0,0700943	0,7033	6
11	-0,0736936	0,0736936	0,0773283	0,6888	5
12	-0,083071	0,083071	0,0875142	0,6764	4

Извор: Истраживање аутора

Табела 17. Једноставни BVAR модел – статистика предвиђања за серију LOG_KS

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0080323	0,0118809	0,0153107	0,9501	15
2	-0,0171883	0,0206642	0,0266168	0,9162	14
3	-0,0250463	0,0250463	0,033087	0,861	13
4	-0,0300658	0,0300658	0,0366217	0,8034	12
5	-0,0339993	0,0339993	0,0393644	0,7393	11
6	-0,0389188	0,0389188	0,044275	0,7155	10
7	-0,044218	0,044218	0,0499114	0,7037	9
8	-0,0486759	0,0486759	0,0543407	0,6871	8
9	-0,0551804	0,0551804	0,0598032	0,6681	7
10	-0,0606779	0,0606779	0,0636946	0,6391	6
11	-0,0672107	0,0672107	0,0697799	0,6216	5
12	-0,0782857	0,0782857	0,0819571	0,6334	4

Извор: Истраживање аутора

Табела 18. OLS VAR модел – статистика предвиђања за серију LOG_KS

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0037899	0,0158242	0,0198605	1,2325	15
2	-0,0078977	0,0219067	0,0297773	1,025	14
3	-0,0129837	0,0249288	0,0333122	0,8669	13
4	-0,0168268	0,028083	0,0377562	0,8282	12
5	-0,0184534	0,0292133	0,0370272	0,6954	11
6	-0,0218883	0,0321471	0,0409347	0,6615	10
7	-0,025376	0,0392151	0,0473318	0,6673	9
8	-0,0278102	0,046507	0,0537147	0,6792	8
9	-0,0315731	0,0517643	0,0573651	0,6409	7
10	-0,0320185	0,0512288	0,0537468	0,5392	6
11	-0,0343634	0,0558525	0,0587068	0,5229	5
12	-0,0756462	0,0756462	0,0807504	0,6241	4

Извор: Истраживање аутора

У Табелама 15-18, посматрањем резултата за серију LOG_KS могу се извести следећи закључци:

- код униваријантног OLS модела (Табела 15), корен средње квадратне грешке RMSE износи 0,0212028, а Тејлова U-статистика 1,3158;
- код OLS VAR модела (Табела 18), код првог корака, RMSE износи 0,0198605, а Тејлова U-статистика 1,2325;
- код униваријантног BVAR модела (Табела 16), одабране мере износе 0,0161972 и 1,0052;
- једноставни BVAR модел (Табела 17) дао је вредност од 0,0153107 за корен средње квадратне грешке и 0,9501за Тејлову U-статистику;
- посматрањем вредности Тејлове U-статистике, једноставни BVAR модел је показао најбоље резултате у односу на остале моделе, док су према вредности за RMSE, униваријантни и једноставни BVAR модел дају боље резултате у односу на класичне моделе;
- за серију LOG_KS код једноставног BVAR модела *потврђена* је *помоћна хипотеза* да оцене добијене на основу Бајесовог приступа представљају унапређење особина оцена добијених класичним приступом у смислу њихове прецизности.

Табела 19. Униваријантни OLS модел – статистика предвиђања за серију LOG_KP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0441813	0,0646553	0,0818667	1,072	15
2	-0,0537937	0,075578	0,0949873	0,9831	14
3	-0,0583618	0,0775091	0,1000822	1,0573	13
4	-0,0726897	0,0792388	0,1053942	1,095	12
5	-0,080169	0,0869698	0,111256	1,0437	11
6	-0,0853344	0,0925507	0,1169196	0,9949	10
7	-0,0880102	0,0974862	0,1217725	1,0054	9
8	-0,0960701	0,1062748	0,1292686	0,9602	8
9	-0,1076629	0,1170896	0,1379909	1,0062	7
10	-0,1325815	0,1325815	0,1488487	1,0735	6
11	-0,1517394	0,1517394	0,1631153	1,0308	5
12	-0,168159	0,168159	0,1780416	0,9239	4

Извор: Истраживање аутора

Табела 20. Униваријантни BVAR модел – статистика предвиђања за серију LOG_KP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0225398	0,0658425	0,0760512	0,9958	15
2	-0,0283201	0,0743245	0,0935104	0,9678	14
3	-0,0337809	0,0759171	0,0916301	0,968	13
4	-0,0541152	0,0729887	0,0960302	0,9977	12
5	-0,0638964	0,0773111	0,1035076	0,971	11
6	-0,0694842	0,0852669	0,1110854	0,9452	10
7	-0,0700404	0,0984824	0,1166234	0,9629	9
8	-0,0776823	0,1036921	0,1262896	0,9381	8
9	-0,0889391	0,1086751	0,1306232	0,9525	7
10	-0,1146916	0,1146916	0,1355079	0,9773	6
11	-0,1333353	0,1333353	0,1482564	0,9369	5
12	-0,1472637	0,1472637	0,1682468	0,873	4

Извор: Истраживање аутора

Табела 21. Једноставни BVAR модел – статистика предвиђања за серију LOG_KP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0191647	0,0649151	0,0743164	0,9731	15
2	-0,0233839	0,0725514	0,0907912	0,9397	14
3	-0,0279892	0,071933	0,0871832	0,9211	13
4	-0,0489697	0,0704562	0,0908005	0,9434	12

5	-0,0588714	0,0727876	0,0977865	0,9173	11
6	-0,0645611	0,0792092	0,1053082	0,8961	10
7	-0,0645241	0,0932107	0,1095975	0,9049	9
8	-0,0723767	0,0987744	0,1200266	0,8916	8
9	-0,0844311	0,1037352	0,125205	0,913	7
10	-0,1115523	0,1115523	0,1300955	0,9382	6
11	-0,1325274	0,1325274	0,1468698	0,9281	5
12	-0,1475881	0,1475881	0,169722	0,8807	4

Извор: Истраживање аутора

Табела 22. OLS VAR модел – статистика предвиђања за серију LOG_KP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0261523	0,0593344	0,0751005	0,9834	15
2	-0,0294518	0,0653969	0,078638	0,8139	14
3	-0,0323191	0,0621644	0,0802079	0,8474	13
4	-0,0477383	0,0637626	0,0874674	0,9088	12
5	-0,0541761	0,0697587	0,0904762	0,8487	11
6	-0,0584776	0,0648708	0,0927095	0,7889	10
7	-0,0593654	0,0720312	0,0907031	0,7489	9
8	-0,0668666	0,0780519	0,0985799	0,7323	8
9	-0,07879	0,0894014	0,1124824	0,8202	7
10	-0,1026396	0,1026396	0,1123077	0,8099	6
11	-0,1316774	0,1316774	0,137597	0,8695	5
12	-0,1622243	0,1622243	0,1731701	0,8986	4

Извор: Истраживање аутора

У Табелама 19-22, приказана је статистика предвиђања за серију LOG_KP. Модели су упоређени на основу неколико одабраних мера прецизности и изведени су следећи закључци:

- у Табели 19. за униваријантни OLS модел, код првог корака, мера прецизности RMSE износи 0,0818667, а Тејлова U-статистика 1,072;
- Табела 22. приказује резултате за OLS VAR модел, где је израчуната RMSE мера 0,0751005 , а Тејлова U-статистика 0,9834;
- код униваријантног BVAR модела приказаног у Табели 20, RMSE и Тејлова U-статистика износе 0,0760512 и 0,9958;

- једноставни BVAR модел, приказан у Табели 21, дао је вредност од 0,0743164 за корен средње квадратне грешке и 0,9731за Тејлову U-статистику;
- униваријантни и једноставни BVAR модел дају боље резултате од класичног униваријантног OLS модела, док је једноставни BVAR модел бољи и од OLS VAR модела и према RMSE вредностима и према Тејловој U-статистици.

Табела 23. Униваријантни OLS модел – статистика предвиђања за серију LOG_ DS

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0393922	0,0510007	0,0638919	1,143	15
2	-0,0712428	0,079296	0,0917211	1,3072	14
3	-0,0961586	0,1003824	0,1142834	1,3567	13
4	-0,1143772	0,1165242	0,1322786	1,2847	12
5	-0,1238538	0,1238538	0,1381875	1,5526	11
6	-0,1401944	0,1401944	0,1531685	1,5444	10
7	-0,1535719	0,1535719	0,1675157	1,3978	9
8	-0,1631739	0,1631739	0,1778487	1,3766	8
9	-0,1741718	0,1741718	0,1897644	1,3013	7
10	-0,1838282	0,1838282	0,2008674	1,2348	6
11	-0,1962516	0,1962516	0,2141381	1,2191	5
12	-0,239989	0,239989	0,2424036	1,1356	4

Извор: Истраживање аутора

Табела 24. Униваријантни BVAR модел – статистика предвиђања за серију LOG_DS

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0214063	0,0393933	0,0565735	1,0121	15
2	-0,0418028	0,0558664	0,0730237	1,0408	14
3	-0,0605172	0,0717578	0,0880707	1,0455	13
4	-0,0766247	0,0884306	0,1052845	1,0225	12
5	-0,0806529	0,0878632	0,0978632	1,0995	11
6	-0,0962877	0,0988593	0,1095062	1,1042	10
7	-0,1107705	0,1191318	0,1277025	1,0656	9
8	-0,1220436	0,1250788	0,1375921	1,065	8
9	-0,1362204	0,1362204	0,1525128	1,0459	7
10	-0,1476916	0,1476916	0,1664465	1,0232	6
11	-0,1597256	0,1597256	0,1790507	1,0193	5
12	-0,2077034	0,2077034	0,2101513	0,9845	4

Извор: Истраживање аутора

Табела 25. Једноставни BVAR модел – статистика предвиђања за серију LOG_ DS

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0186526	0,0373835	0,0554734	0,9924	15
2	-0,0376818	0,05203	0,0697163	0,9936	14
3	-0,0559252	0,0679181	0,0843437	1,0013	13
4	-0,0728984	0,0858586	0,1033224	1,0035	12
5	-0,0758571	0,0842685	0,0934004	1,0494	11
6	-0,0912866	0,094168	0,1042496	1,0512	10
7	-0,105431	0,1161912	0,1231989	1,028	9
8	-0,1167878	0,119653	0,1315845	1,0185	8
9	-0,1315009	0,1315009	0,1480109	1,015	7
10	-0,1433777	0,1442379	0,1628926	1,0014	6
11	-0,1545612	0,1545612	0,1732649	0,9864	5
12	-0,2046818	0,2046818	0,2072644	0,971	4

Извор: Истраживање аутора

Табела 26. OLS VAR модел – статистика предвиђања за серију LOG_ DS

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0037807	0,0390414	0,0515806	0,9228	15
2	-0,0194469	0,048818	0,0578525	0,8245	14
3	-0,0314306	0,0684181	0,0767583	0,9112	13
4	-0,0419192	0,0912495	0,1009075	0,98	12
5	-0,0416215	0,1002501	0,107352	1,2062	11
6	-0,0482223	0,1325726	0,140676	1,4185	10
7	-0,049306	0,1445711	0,1707279	1,4246	9
8	-0,0466723	0,127545	0,1462078	1,1317	8
9	-0,043898	0,1431536	0,1727099	1,1844	7
10	-0,0308447	0,1503338	0,1894431	1,1646	6
11	-0,0559925	0,2092447	0,2388607	1,3598	5
12	-0,2292954	0,2292954	0,2321929	1,0878	4

Извор: Истраживање аутора

На основу података о серији LOG_DS, приказаних у Табелама 23-26, могу се извести следећи закључци:

- у Табели 23, код униваријантног OLS модела RMSE, код првог корака, износи 0,0638919, а Тејлова U-статистика 1,143;
- код OLS VAR модела (Табела 26), корен средње квадратне грешке износи 0,0515806, а Тејлова U-статистика 0,9228;

- код униваријантног BVAR модела приказаног у Табели 24, RMSE и Тејлова U-статистика износе 0,0565735 и 1,0121;
- једноставни BVAR модел (Табела 25) дао је вредност од 0,0554734 за корен средње квадратне грешке и 0,9924 за Тејлову U-статистику;
- Према вредностима RMSE и Тејлове U-статистике, једноставни и униваријантни BVAR модел су бољи од униваријантног OLS модела, док су резултати лошији у односу на OLS VAR модел.
- у овом делу истраживања модел класичне статистике, OLS VAR модел, дао је боље резултате од Бајесових модела, па из тог разлога *не можемо прихватити* помоћну хипотезу

Табела 27. Униваријантни OLS модел – статистика предвиђања за серију LOG_DP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0208045	0,0394875	0,0452651	1,0019	15
2	-0,0518645	0,0569181	0,0734107	0,9744	14
3	-0,0827367	0,0855142	0,1018544	0,9774	13
4	-0,1133132	0,1163248	0,1283586	0,9858	12
5	-0,1340124	0,1340124	0,1455947	0,9961	11
6	-0,1476983	0,1476983	0,1560855	1,0001	10
7	-0,1668341	0,1668341	0,1709309	0,9968	9
8	-0,1863837	0,1863837	0,1894755	0,9918	8
9	-0,2103192	0,2103192	0,2129155	0,9784	7
10	-0,2360399	0,2360399	0,241716	0,9615	6
11	-0,2812966	0,2812966	0,2852423	0,9548	5
12	-0,3185134	0,3185134	0,3200915	0,9359	4

Извор: Истраживање аутора

Табела 28. Униваријантни BVAR модел – статистика предвиђања за серију LOG_DP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0186117	0,0380541	0,044087	0,9758	15
2	-0,0471627	0,0540321	0,070718	0,9386	14
3	-0,0753739	0,0799407	0,0972191	0,9329	13
4	-0,1033328	0,1088846	0,1213846	0,9322	12
5	-0,1217904	0,1239685	0,1358149	0,9292	11
6	-0,1335452	0,1335452	0,143571	0,9199	10
7	-0,1510121	0,1510121	0,1557687	0,9084	9

8	-0,1693617	0,1693617	0,1725031	0,903	8
9	-0,1928154	0,1928154	0,1954798	0,8983	7
10	-0,2185385	0,2185385	0,2255516	0,8972	6
11	-0,265966	0,265966	0,2710818	0,9074	5
12	-0,3076802	0,3076802	0,3095534	0,9051	4

Извор: Истраживање аутора

Табела 29. Једноставни BVAR модел – статистика предвиђања за серију LOG_ DP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0164883	0,0378615	0,0440575	0,9751	15
2	-0,0431987	0,0542197	0,0699078	0,9279	14
3	-0,069326	0,0767499	0,0949662	0,9113	13
4	-0,0955528	0,1051524	0,1178157	0,9048	12
5	-0,1112492	0,118843	0,1300661	0,8899	11
6	-0,1212587	0,1219881	0,135128	0,8658	10
7	-0,1368142	0,1368142	0,144352	0,8418	9
8	-0,1537054	0,1537054	0,1591964	0,8333	8
9	-0,1761329	0,1761329	0,1811362	0,8324	7
10	-0,2005257	0,2005257	0,2124671	0,8451	6
11	-0,2543619	0,2543619	0,2621625	0,8775	5
12	-0,3038012	0,3038012	0,3053376	0,8928	4

Извор: Истраживање аутора

Табела 30. OLS VAR модел – статистика предвиђања за серију LOG_ DP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0154744	0,043675	0,0603987	1,3368	15
2	-0,0390992	0,0688974	0,0950945	1,2622	14
3	-0,0598007	0,099671	0,1200578	1,152	13
4	-0,0814077	0,1302361	0,1440261	1,1061	12
5	-0,0830664	0,1495199	0,1589233	1,0873	11
6	-0,08248	0,1617747	0,1664865	1,0667	10
7	-0,0837335	0,17253	0,1765401	1,0295	9
8	-0,0841317	0,1884964	0,1949521	1,0205	8
9	-0,0872395	0,2171537	0,2263756	1,0403	7
10	-0,0821503	0,271171	0,2781253	1,1063	6
11	-0,1554695	0,3366072	0,3460965	1,1585	5
12	-0,3310868	0,3310868	0,333609	0,9754	4

Извор: Истраживање аутора

За серију LOG_DP, приказану у Табелама 27-30, могу се извести следећи закључци:

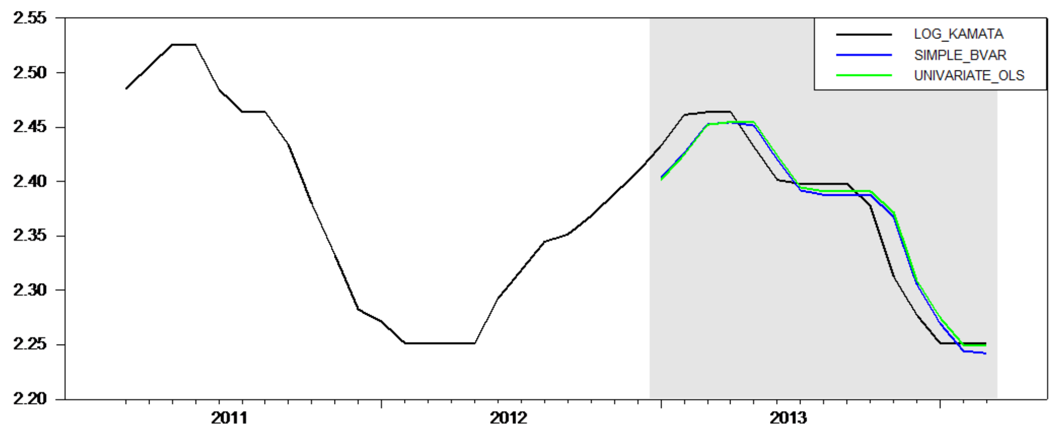
- код униваријантног OLS модела (Табела 27), код првог корака, корен средње квадратне грешке RMSE износи 0,0452651, а Тејлова U-статистика 1,0019;
- код OLS VAR модела, приказаног у Табела 30, RMSE мера прецизности износи 0,0603987, а Тејлова U-статистика 1,3368;
- код униваријантног BVAR модела (Табела 28), RMSE и Тејлова U-статистика износе 0,044087 и 0,9758;
- једноставни BVAR модел (Табела 29) дао је вредност од 0,0440575 за корен средње квадратне грешке и 0,9751за Тејлову U-статистику;
- У Табелама 27-30, за серију LOG_DP, униваријантни и једноставни BVAR модел дају боље резултате од класичних модела: униваријантног OLS и OLS VAR модела, док је једноставни BVAR модел бољи од свих посматраних модела.
- *потврђена* је *помоћна хипотеза* да оцене добијене на основу Бајесовог приступа представљају унапређење особина оцена добијених класичним приступом у смислу њихове прецизности.

На основу извршене анализе добијене применом софтверског пакета RATS, може се закључити да код свих серија једноставни BVAR модел даје најбоље резултате у погледу RMSE и Тејлове U-статистике, једино су нешто лошији резултати од OLS VAR модела код серије LOG_DS, што не умањује квалитет изведених закључака.

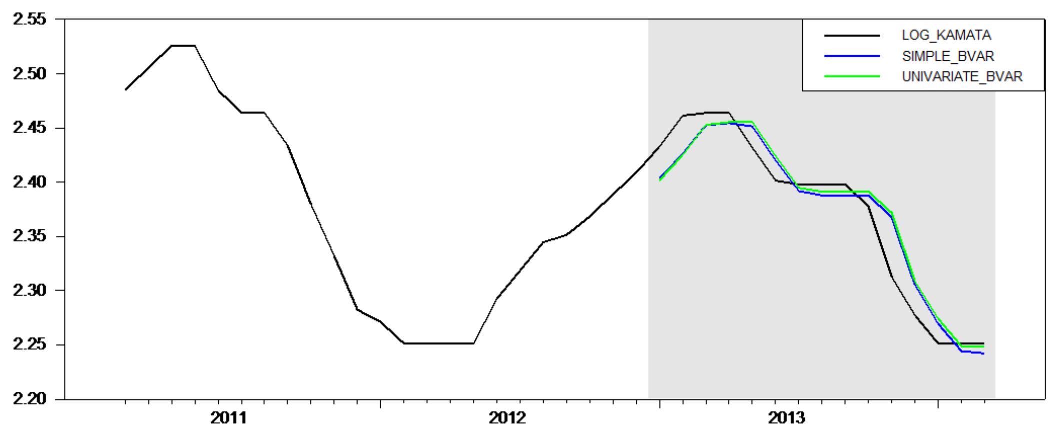
У анализи резултата се показало да је вредност укупне затворености γ један од кључних елемената за тачност предвиђања. Постављање ове вредности на висок ниво (γ = 2.0) за класичне моделе довело је до слабијег учинка, док избор ниских вредности није побољшао перформансе. Са економске тачке гледишта, резултати указују да присуство априорних информација проистеклих било из искуства или из економске теорије, може значајно побољшати прогнозе економских модела.

Други важан елеменат је релативна тежина ω, чије су вредности за појединачне моделе наведене у Табели 9. Постављање ове вредности на ω=0.001 за униваријантни OLS модел и униваријантни BVAR модел, није се показало као најбоља опција. Истовремено, постављањем вредности релативне тежине на висок ниво за OLS VAR модел, ω=1.0, такође се није показало као оптимално решење. Избором неке средње вредности, као што је то учињено за једноставни BVAR модел, ω=0.5, добијени су најбољи резултати.

С обзиром да се једноставни BVAR модел показао као најбољи, на наредним графичким приказима је представљен овај модел у поређењу са оригиналним подацима и са сваким од осталих модела понаособ. Одабрани су подаци за серију LOG_KAM који су приказани у Прилогу 3. Слична графичка решења се добијају и за остале серије.



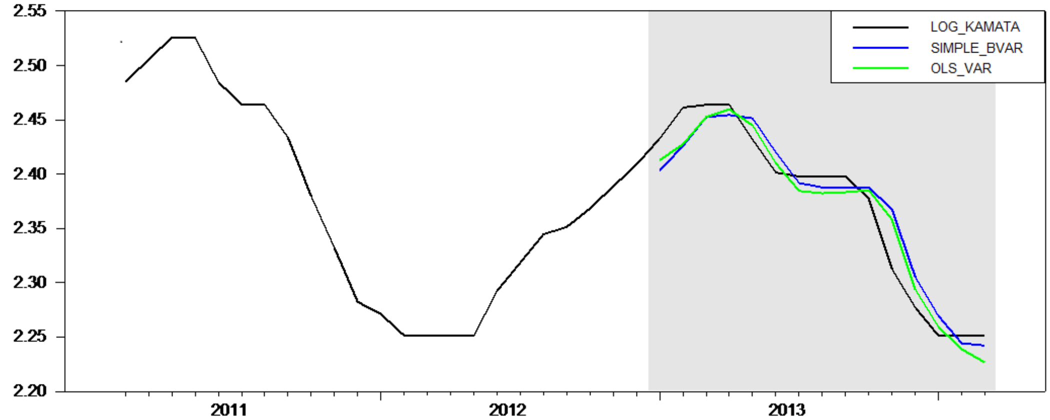
Слика 13. Једноставни BVAR модел у односу на униваријантни OLS модел[153]



Слика 14. Једноставни BVAR модел у односу на униваријантни BVAR модел[154]

[153] Истраживање аутора

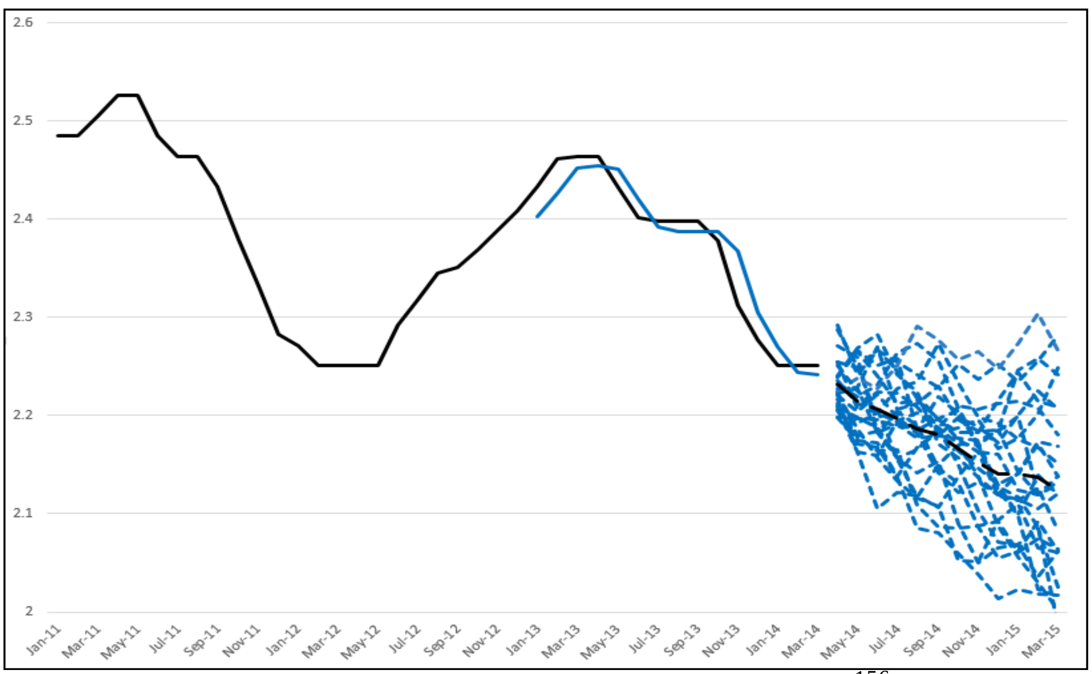
[154] Ibid.



Слика 15. Једноставни BVAR модел у односу на униваријантни OLS VAR модел[155]

4.3.4.3 Симулација кретања каматних стопа

У циљу предвиђања будућег кретања каматних стопа извршено је 25 симулација на основу једноставног BVAR модела, што је приказано на Слици 16.



Слика 16. Симулација кретања каматних стопа[156]

[155] Ibid.

[156] Ibid.

На слици су црном линијом приказани историјски подаци од јануара 2011. до марта 2014. године, плавом линијом предвиђање модела од јануара 2013. до марта 2014. године, а затим су испрекиданим плавим линијама представљене симулације будућих кретања на основу одабраног модела, док је тамна испрекидана линија просечна вредност ових 25 симулација.

На овај начин је кроз анализу временских серија у раду потврђена оправданост примене Бајесових метода у предвиђању кретања економских појава, уз истовремену *потврду главне хипотезе* да *Бајесов приступ статистичком закључивању има предност у односу на класичну инференцију и да је његова примена у економији оправдана.* Такође, с обзиром да у Републици Србији нису расположиви подаци за дужи временски период, Бајесова статистика је свакако и једно од пожељних решења.

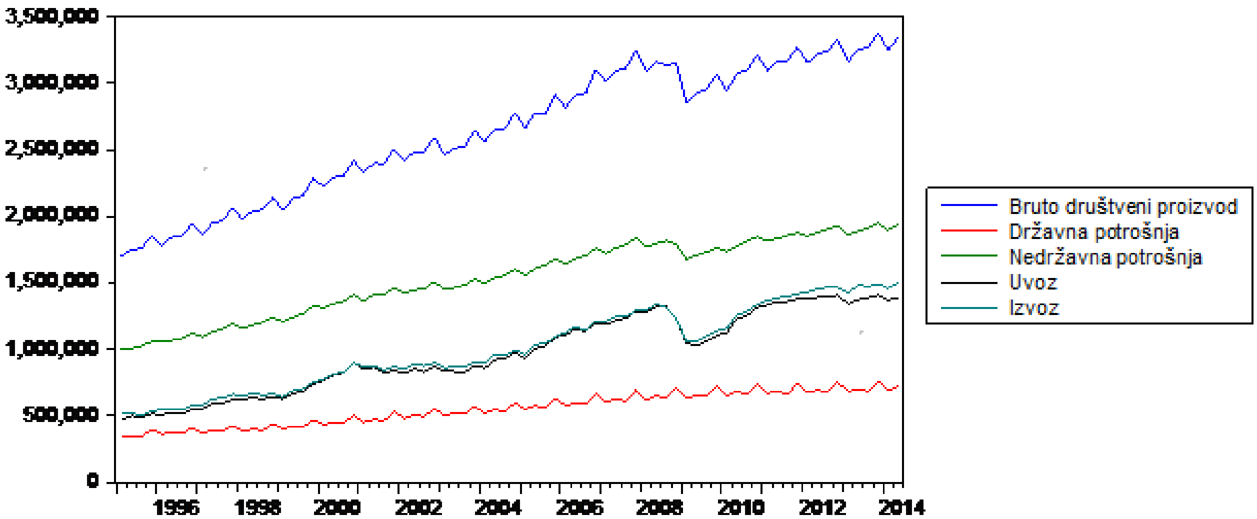
4.3.4.4 Примена одабраних модела истраживања на примеру података са подручја Европске Уније

Други део спроведеног истраживања применом софтверског пакета RATS, извршен је на следећим макроекономским показатељима који се односе на Европску унију:

1. Бруто друштвени производ (енгл. *Gross Domestic Product*-GDP);
2. Државна потрошња (енгл. *Government Expenditure*-GOV);
3. Недржавна потрошња (енгл. *Nongovernment Expenditure*-NONGOV);
4. Извоз (енгл. *Export*-EXP) и
5. Увоз (енгл. *Import*-IMP).

Подаци су изражени квартално и обухватају период од првог квартала 1995. до другог квартала 2014. године, а преузети су са сајта Eurostat.[157] Наведене серије приказане су на Слици 17.

[157] ec.europa.eu/eurostat



Слика 17. Оригинални подаци одабраних серија

Као што се може уочити на слици, сезонски ефекти значајно су изражени код серија бруто друштвеног производа и државне и недржавне потрошње, док се код серија извоза и увоза ови ефекти мање присутни. То потврђују и коефицијенти сезонских ефеката, који се код бруто друштвеног производа и државне потрошње крећу у распонима од 0.977 до 1.029, односно од 0.968 до 1.076, док су распони ових коефицијената значајно нижи код серија увоза и извоза (од 0.983 до 1.019, односно од 0.984 до 1.013), што је приказано у Табели 31.

Табела 31. Коефицијенти сезонских ефеката

Квартал	GDP	Државна потрошња	Недржавна потрошња	Увоз	Извоз
1Q	0,977162851	0,967814	0,98362	0,983495	0,98417
2Q	0,999818431	0,996168	0,996545	1,005696	1,009766
3Q	0,994827448	0,964047	1,001062	0,99248	0,992886
4Q	1,028878647	1,075913	1,019095	1,018684	1,013468

Извор: Истраживање аутора

Да би се искључио утицај сезонских ефеката, извршена је трансформација одабраних серија мултипликативном методом покретних просека, а резултати који су проистекли из тако обрађених података, приказани су у наредним табелама.

4.3.4.5 Статистика предвиђања за макроекономске податке

У табелама које следе, приказана је статистика предвиђања за сваку трансформисану серију, применом одабраних модела, а поређење је извршено на основу одговарајућих мера прецизности. Приказани подаци су ради даље обраде трансформисани логартимовањем, тако да су коришћене следеће ознаке за посматране серије података: LOG_GDP (логаритамска серија-бруто друштвени производ), LOG_GOV (логаритамска серија-државна потрошња), LOG_NONGOV (логаритамска серија-недржавна потрошња), LOG_EXP (логаритамска серија-извоз) и LOG_IMP (логаритамска серија-увоз). За моделе који су предмет посматрања, одабране су вредности за укупну затвореност и релативну тежину приказане у Табели 9.

Табела 32. Униваријантни OLS модел – статистика предвиђања за серију LOG_GDP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,00165517	0,00653605	0,00766972	0,7829	15
2	0,00317071	0,00757344	0,00985900	0,6309	14
3	0,00336463	0,00996864	0,01246608	0,5898	13
4	0,00338882	0,00968614	0,01313904	0,516	12
5	0,00318324	0,01185782	0,01505739	0,4964	11
6	0,00213327	0,00989658	0,01340156	0,4058	10
7	0,00173329	0,01171870	0,01456390	0,3895	9
8	0,00247496	0,01144441	0,01490204	0,3524	8
9	0,00323870	0,01402921	0,01735332	0,3613	7
10	0,00386833	0,01358655	0,01723113	0,3279	6
11	0,00377163	0,01187502	0,01763437	0,3096	5
12	0,00390213	0,00998003	0,01371031	0,2275	4

Извор: Истраживање аутора

Табела 33. Униваријантни BVAR модел – статистика предвиђања за серију LOG_GDP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,00163286	0,00653293	0,0076638	0,7823	15
2	0,00312553	0,00755844	0,00984086	0,6297	14
3	0,00329605	0,00994049	0,01244139	0,5886	13
4	0,00329631	0,00967502	0,01310678	0,5147	12
5	0,00306622	0,01185123	0,01502113	0,4952	11

6	0,00199103	0,00989613	0,01336368	0,4046	10
7	0,00156511	0,01170673	0,01452574	0,3885	9
8	0,00227969	0,01144824	0,01485064	0,3512	8
9	0,00301567	0,01405505	0,01728833	0,3599	7
10	0,00361629	0,01359467	0,01715063	0,3263	6
11	0,00349062	0,01191615	0,01754479	0,308	5
12	0,00359024	0,00999215	0,01358939	0,2255	4

Извор: Истраживање аутора

Табела 34. Једноставни BVAR модел – статистика предвиђања за серију LOG_GDP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,00243787	0,00679301	0,00801572	0,8183	15
2	0,0048104	0,00840432	0,01093351	0,6996	14
3	0,00591636	0,01125068	0,01419252	0,6715	13
4	0,00690706	0,01087439	0,01547062	0,6075	12
5	0,00772789	0,01302046	0,01812721	0,5976	11
6	0,00780655	0,01179069	0,01723478	0,5219	10
7	0,00861277	0,01426214	0,01922308	0,5141	9
8	0,0107736	0,01431201	0,02046085	0,4838	8
9	0,01305147	0,01602574	0,02450612	0,5102	7
10	0,01534045	0,01764177	0,02577405	0,4904	6
11	0,01665502	0,01910097	0,02784185	0,4888	5
12	0,01869147	0,01880629	0,02658821	0,4412	4

Извор: Истраживање аутора

Табела 35. OLS VAR модел – статистика предвиђања за серију LOG_GDP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,00355944	0,00789916	0,00956161	0,9761	15
2	0,00808566	0,01118635	0,01476117	0,9446	14
3	0,01182942	0,01640167	0,02070721	0,9797	13
4	0,01571335	0,01922315	0,02393123	0,9398	12
5	0,01930173	0,02237080	0,02904245	0,9575	11
6	0,02254908	0,02351143	0,03096545	0,9376	10
7	0,02658802	0,02713482	0,03614212	0,9666	9
8	0,03258947	0,03258947	0,03959561	0,9363	8
9	0,03898119	0,03898119	0,04721507	0,9829	7
10	0,04518135	0,04518135	0,05152505	0,9804	6

11	0,04875833	0,04875833	0,05634854	0,9893	5
12	0,05435355	0,05435355	0,05988262	0,9937	4

Извор: Истраживање аутора

У Табелама 32-35, презентовани су резултати прогнозирања за серију LOG_ GDP. На основу одабраних мера прецизности, упоређени су модели и изведени следећи закључци:

- код класичног, униваријантног OLS модела (Табела 32), RMSE мера прецизности износи 0,00766972, а Тејлова U-статистика 0,7829;
- код другог класичног модела, OLS VAR модела (Табела 35), корен средње квадратне грешке RMSE износи 0,00956161, а Тејлова U-статистика 0,9761;
- код униваријантног BVAR модела (Табела 33), RMSE и Тејлова U-статистика износе 0,00766380 и 0,7823;
- други Бајесов модел, једноставни BVAR модел (Табела 34) за ове мере прецизности има вредности 0,00801572 и 0,8183;
- с обзиром да се сматра да је бољи онај модел код ког су вредности одабраних мера прецизности ниже, за серију LOG_ GDP закључујемо да је униваријантни BVAR модел дао боље резултате од два класична модела, док је једноставни BVAR модел дао боље резултате од OLS VAR модела; тиме је *потврђена помоћна хипотеза* да оцене добијене на основу Бајесовог приступа представљају унапређење особина оцена добијених класичним приступом у смислу њихове прецизности.

Табела 36. Униваријантни OLS модел – статистика предвиђања за серију LOG_GOV

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,00174501	0,00736329	0,00865164	0,8135	15
2	0,00269717	0,00801809	0,00991671	0,6424	14
3	0,00301087	0,0109275	0,01377403	0,6331	13
4	0,00300848	0,01011224	0,01291546	0,5119	12
5	0,00301641	0,01228748	0,01599260	0,5183	11
6	0,00230585	0,01033774	0,01403621	0,4172	10
7	0,00236158	0,01364002	0,01617000	0,4174	9
8	0,00365095	0,0126565	0,01522631	0,3489	8
9	0,00553894	0,0149156	0,01973400	0,3880	7

10	0,00650621	0,01436794	0,01921533	0,3466	6
11	0,00557764	0,01423414	0,02035770	0,3431	5
12	0,00531970	0,01134892	0,01649052	0,2655	4

Извор: Истраживање аутора

Табела 37. Униваријантни BVAR модел – статистика предвиђања за серију LOG_GOV

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,00172438	0,0073599	0,00864566	0,8130	15
2	0,00265528	0,0079992	0,00990088	0,6414	14
3	0,00294712	0,0109064	0,01375242	0,6321	13
4	0,00292222	0,01007839	0,01288501	0,5107	12
5	0,00290698	0,0122779	0,01595694	0,5171	11
6	0,00217232	0,01029565	0,01399512	0,416	10
7	0,00220308	0,01362239	0,01612412	0,4162	9
8	0,00346616	0,01262105	0,01515954	0,3474	8
9	0,00532674	0,01489551	0,01964717	0,3863	7
10	0,00626549	0,01437403	0,01910624	0,3446	6
11	0,00530878	0,01420555	0,0202486	0,3412	5
12	0,00502	0,01135866	0,01635548	0,2633	4

Извор: Истраживање аутора

Табела 38. Једноставни BVAR модел – статистика предвиђања за серију LOG_GOV

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,00267544	0,00760789	0,00915928	0,8613	15
2	0,00463966	0,00903716	0,01119508	0,7253	14
3	0,00603186	0,01235988	0,01585019	0,7285	13
4	0,00717216	0,01256223	0,01579122	0,6259	12
5	0,00839232	0,01437488	0,01992925	0,6458	11
6	0,00904059	0,01394528	0,01904651	0,5661	10
7	0,01055446	0,01620617	0,02235965	0,5771	9
8	0,01351323	0,01646215	0,02272120	0,5207	8
9	0,01726516	0,02031261	0,02928742	0,5758	7
10	0,02020575	0,02100462	0,03075582	0,5547	6
11	0,0209532	0,02289056	0,03375885	0,5689	5
12	0,02295381	0,02324412	0,03288861	0,5295	4

Извор: Истраживање аутора

Табела 39. OLS VAR модел – статистика предвиђања за серију LOG_GOV

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,00716026	0,00889541	0,01220235	1,1474	15
2	0,01300173	0,01419923	0,01791635	1,1607	14
3	0,01787216	0,01935583	0,02513020	1,1550	13
4	0,02230128	0,02315517	0,02852465	1,1305	12
5	0,02636272	0,02768734	0,03473453	1,1256	11
6	0,02996701	0,03004774	0,03724331	1,1070	10
7	0,03445994	0,03445994	0,04313660	1,1134	9
8	0,04074684	0,04074684	0,04724414	1,0827	8
9	0,04778653	0,04778653	0,05588380	1,0987	7
10	0,05406162	0,05406162	0,06068851	1,0946	6
11	0,05713512	0,05713512	0,06524673	1,0995	5
12	0,06237975	0,06237975	0,06855217	1,1036	4

Извор: Истраживање аутора

У Табелама 36-39, посматрани су резултати за серију LOG_GOV и изведени су следећи закључци:

- код униваријантног OLS модела (Табела 36), у првом кораку, корен средње квадратне грешке RMSE износи 0,00865164, а Тејлова U-статистика 0,8135;
- код OLS VAR модела (Табела 39), код првог корака, мера прецизности RMSE износи 0,01220235, а Тејлова U-статистика 1,1474;
- код униваријантног BVAR модела (Табела 37), RMSE и Тејлова U-статистика износе 0,00864566 и 0,8130;
- код једноставног BVAR модела (Табела 38) одабране мере прецизности имају вредност 0,00915928 и 0,8613;
- посматрањем вредности RMSE и Тејлове U-статистике, може се закључити да је, као и код серије LOG_GOV, униваријантни BVAR модел дао најбоље резултате од свих посматраних модела, док је једноставни BVAR модел дао боље резултате од OLS VAR модела.

Табела 40. Униваријантни OLS модел – статистика предвиђања за серију LOG_NONGOV

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0036653	0,0066132	0,0075224	1,0499	15
2	-0,0073015	0,0094308	0,0110648	1,0411	14
3	-0,0124335	0,0131195	0,0157623	1,2176	13
4	-0,0174756	0,0180875	0,0205632	1,3415	12
5	-0,0230034	0,0230034	0,0255503	1,5016	11
6	-0,0282525	0,0282525	0,0302386	1,5959	10
7	-0,0330278	0,0330278	0,0349811	1,5677	9
8	-0,0370611	0,0370611	0,0390257	1,4629	8
9	-0,0425224	0,0425224	0,0440919	1,4867	7
10	-0,0481585	0,0481585	0,0506685	1,5067	6
11	-0,0544106	0,0544106	0,0570745	1,5399	5
12	-0,0634264	0,0634264	0,0652875	1,7198	4

Извор: Истраживање аутора

Табела 41. Униваријантни BVAR модел – статистика предвиђања за серију LOG_NONGOV

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0036785	0,0066179	0,0075297	1,0509	15
2	-0,0073292	0,0094458	0,0110850	1,0430	14
3	-0,0124772	0,0131601	0,0158005	1,2205	13
4	-0,0175371	0,0181432	0,0206213	1,3453	12
5	-0,0230839	0,0230839	0,0256311	1,5064	11
6	-0,0283533	0,0283533	0,0303452	1,6015	10
7	-0,0331511	0,0331511	0,0351106	1,5735	9
8	-0,0372120	0,0372120	0,0391821	1,4688	8
9	-0,0427032	0,0427032	0,0442789	1,4930	7
10	-0,0483746	0,0483746	0,0508876	1,5132	6
11	-0,0546597	0,0546597	0,0573284	1,5468	5
12	-0,0637229	0,0637229	0,0655880	1,7277	4

Извор: Истраживање аутора

Табела 42. Једноставни BVAR модел – статистика предвиђања за серију LOG_NONGOV

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	-0,0022210	0,0060302	0,0067471	0,9417	15
2	-0,0042169	0,0076090	0,0090220	0,8489	14

3	-0,0075294	0,0094931	0,0117693	0,9091	13
4	-0,0105475	0,0121471	0,0143594	0,9368	12
5	-0,0139163	0,0143220	0,0166653	0,9794	11
6	-0,0168083	0,0168083	0,0181759	0,9593	10
7	-0,0189885	0,0189885	0,0202719	0,9085	9
8	-0,0198594	0,0198594	0,0213349	0,7998	8
9	-0,0218537	0,0218537	0,0231849	0,7817	7
10	-0,0235895	0,0235895	0,0263869	0,7846	6
11	-0,0263466	0,0263466	0,0292125	0,7882	5
12	-0,0302469	0,0302469	0,0321107	0,8459	4

Извор: Истраживање аутора

Табела 43. OLS VAR модел – статистика предвиђања за серију LOG_NONGOV

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,00675541	0,00735558	0,00918981	1,2826	15
2	0,01173835	0,01173835	0,01362492	1,2820	14
3	0,01416641	0,01416641	0,01661064	1,2831	13
4	0,01615284	0,01646286	0,01927289	1,2573	12
5	0,01671224	0,01671224	0,01900180	1,1168	11
6	0,01805336	0,01805336	0,01952552	1,0305	10
7	0,02010275	0,02010275	0,02143944	0,9608	9
8	0,02353941	0,02353941	0,02413863	0,9049	8
9	0,02614871	0,02614871	0,02711877	0,9144	7
10	0,02886425	0,02886425	0,03083966	0,9170	6
11	0,02918430	0,02918430	0,03080285	0,8311	5
12	0,02949315	0,02949315	0,03036394	0,7998	4

Извор: Истраживање аутора

У Табелама 40-43, приказана је статистика предвиђања за серију LOG_NONGOV. На основу неколико одабраних мера прецизности упоређени су модели и изведени су следећи закључци:

- мере прецизности RMSE и Тејлова U-статистика (Табела 40) за униваријантни OLS модел износе 0,0075224 и 1,0499;
- одабране мере прецизности за OLS VAR модел, приказане у Табели 43, износе 0,00918981 и 1,2826;

- код униваријантног BVAR модела приказаног у Табели 41, корен средње квадратне грешке износи 0,0075297, док Тејлова U-статистика има вредност 1,0509;
- једноставни BVAR модел, приказан у Табели 42, има вредности 0,0067471 и 0,9417 за одабране мере прецизности;
- униваријантни и једноставни BVAR модел дају боље резултате од OLS VAR модела, док је једноставни BVAR модел бољи и од униваријантног OLS модела и према RMSE вредностима и према Тејловој U-статистици.

Табела 44. Униваријантни OLS модел – статистика предвиђања за серију LOG_ EXP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,00723504	0,01994652	0,02442874	0,8848	15
2	0,01343312	0,02547480	0,03733517	0,8273	14
3	0,01718106	0,03724059	0,05025057	0,8137	13
4	0,02077273	0,03908354	0,05818399	0,7805	12
5	0,02388412	0,04858094	0,06858231	0,7726	11
6	0,02376641	0,05092478	0,07073728	0,7396	10
7	0,02623354	0,06052910	0,07933469	0,7352	9
8	0,03115973	0,06276393	0,08316209	0,7037	8
9	0,03658321	0,06678759	0,09176287	0,6970	7
10	0,04274785	0,06701507	0,09204780	0,6569	6
11	0,04764171	0,07517584	0,09825914	0,6494	5
12	0,05473368	0,07539832	0,0961284	0,6066	4

Извор: Истраживање аутора

Табела 45. Униваријантни BVAR модел – статистика предвиђања за серију LOG_EXP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,00711255	0,01992411	0,02440141	0,8838	15
2	0,01318583	0,02550321	0,03727165	0,8259	14
3	0,01680888	0,03731351	0,05016501	0,8123	13
4	0,02027663	0,03903298	0,0580686	0,7789	12
5	0,02326611	0,04873197	0,06845002	0,7711	11
6	0,02302411	0,05106499	0,07059782	0,7382	10
7	0,02536932	0,06079801	0,07918615	0,7338	9
8	0,03017955	0,06300586	0,08296247	0,7020	8
9	0,03549295	0,06696923	0,0915247	0,6952	7

10	0,04155810	0,06708935	0,09171624	0,6545	6
11	0,04636496	0,07508008	0,09788275	0,6469	5
12	0,05340261	0,07502821	0,09560818	0,6033	4

Извор: Истраживање аутора

Табела 46. Једноставни BVAR модел – статистика предвиђања за серију LOG_ EXP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,00934519	0,01981068	0,02430158	0,8802	15
2	0,01770879	0,02505566	0,03724239	0,8252	14
3	0,02355499	0,03575376	0,05004561	0,8104	13
4	0,02908770	0,03917321	0,05797204	0,7776	12
5	0,03404331	0,04395078	0,06849435	0,7716	11
6	0,03596687	0,04914863	0,07018390	0,7338	10
7	0,04031705	0,05551882	0,07874456	0,7297	9
8	0,04666646	0,05805596	0,08279536	0,7006	8
9	0,05349310	0,06275825	0,09240204	0,7018	7
10	0,06024112	0,06136741	0,09417134	0,6720	6
11	0,06534749	0,07111748	0,10155076	0,6711	5
12	0,07133334	0,07551680	0,10121317	0,6387	4

Извор: Истраживање аутора

Табела 47. OLS VAR модел – статистика предвиђања за серију LOG_ EXP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,01953130	0,02480933	0,0328827	1,1910	15
2	0,03747016	0,04036626	0,05488391	1,2161	14
3	0,05247252	0,05403508	0,07537137	1,2204	13
4	0,06680711	0,06744917	0,09049265	1,2139	12
5	0,07935230	0,08059365	0,1078042	1,2145	11
6	0,08944731	0,08982706	0,11678186	1,2211	10
7	0,10202798	0,10223183	0,13240606	1,2269	9
8	0,11678620	0,11678620	0,14359793	1,2151	8
9	0,13267557	0,13267557	0,16071235	1,2207	7
10	0,14815851	0,14815851	0,17119762	1,2217	6
11	0,15976612	0,15976612	0,18514432	1,2236	5
12	0,17427123	0,17427123	0,19495038	1,2301	4

Извор: Истраживање аутора

У Табелама 44-47 приказани су подаци о серији LOG_ EXP. Закључци који се могу извести су следећи:

- вредности RMSE и Тејлове U-статистике, приказане у Табели 44, за униваријантни OLS модел износе 0,02442874 и 0,8848;
- за OLS VAR модел, приказан у Табели 47, одабране мере прецизности износе 0,0328827 и 1,1910;
- код униваријантног BVAR модела (Табела 45), RMSE и Тејлова U-статистика износе 0,02440141 и 0,8838;
- једноставни BVAR модел (Табела 46) има вредност од 0,02430158 за корен средње квадратне грешке и 0,8802 за Тејлову U-статистику;
- за приказану серију LOG_ EXP, оба Бајесова модела дају боље резултате од модела класичне статистике, па се може *прихватити помоћна хипотеза* која говори у прилог примене Бајесових метода.

Табела 48. Униваријантни OLS модел – статистика предвиђања за серију LOG_IMP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,0066376	0,02149673	0,02627804	0,9173	15
2	0,01198399	0,02903469	0,04128562	0,8814	14
3	0,01423567	0,04043461	0,05457232	0,8733	13
4	0,01612112	0,04280401	0,06325829	0,8539	12
5	0,01617202	0,05435762	0,07462631	0,8573	11
6	0,01353428	0,06080417	0,07728340	0,8433	10
7	0,01298136	0,06903872	0,08397527	0,8400	9
8	0,01422677	0,07105016	0,08641833	0,8171	8
9	0,01688893	0,07427552	0,09412289	0,8093	7
10	0,02018836	0,06992756	0,09028549	0,7603	6
11	0,02250766	0,07123353	0,09427286	0,7470	5
12	0,02840175	0,07042633	0,08880592	0,6865	4

Извор: Истраживање аутора

Табела 49. Униваријантни BVAR модел – статистика предвиђања за серију LOG_IMP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,00650812	0,02148320	0,02624972	0,9163	15

2	0,01172224	0,02904388	0,04122297	0,8801	14
3	0,01384051	0,04044436	0,05449484	0,8720	13
4	0,01559318	0,04292015	0,06316605	0,8527	12
5	0,01550645	0,05459171	0,07454457	0,8564	11
6	0,01272863	0,06105839	0,07723041	0,8427	10
7	0,01203881	0,06929217	0,08394806	0,8397	9
8	0,01315267	0,07127497	0,08639645	0,8169	8
9	0,01568857	0,07468234	0,09409458	0,8091	7
10	0,01887917	0,07026894	0,09021005	0,7597	6
11	0,02109340	0,07123560	0,09418272	0,7463	5
12	0,02692421	0,07044750	0,08858492	0,6848	4

Извор: Истраживање аутора

Табела 50. Једноставни BVAR модел – статистика предвиђања за серију LOG_ IMP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,00751138	0,02138761	0,02621555	0,9151	15
2	0,01386408	0,02859260	0,04134684	0,8827	14
3	0,01714986	0,03960075	0,05456533	0,8732	13
4	0,02001374	0,04111702	0,06313410	0,8523	12
5	0,02123574	0,05195574	0,07470191	0,8582	11
6	0,01998096	0,05624438	0,07689206	0,8390	10
7	0,02071049	0,06386708	0,08353778	0,8356	9
8	0,02310852	0,06557067	0,08557313	0,8091	8
9	0,02695716	0,06731457	0,09445146	0,8121	7
10	0,03077258	0,06330976	0,09170397	0,7723	6
11	0,03346544	0,06796724	0,09742029	0,7719	5
12	0,03869316	0,06727465	0,09430758	0,7290	4

Извор: Истраживање аутора

Табела 51. OLS VAR модел – статистика предвиђања за серију LOG_ IMP

Step	Mean Error	Mean Abs Err	RMS Error	Theil U	N Obs
1	0,00831689	0,02637754	0,03303173	1,1530	15
2	0,01933134	0,03858699	0,05547114	1,1843	14
3	0,02930652	0,05407754	0,07566116	1,2107	13
4	0,04001771	0,06329675	0,08943278	1,2073	12
5	0,04947758	0,07281853	0,10655819	1,2242	11
6	0,05711886	0,08009561	0,11353996	1,2389	10

7	0,06688504	0,08993763	0,12603519	1,2607	9
8	0,07964674	0,09299933	0,13239368	1,2519	8
9	0,09406231	0,10274213	0,14697553	1,2637	7
10	0,10816383	0,10816383	0,15127635	1,2740	6
11	0,11845368	0,11845368	0,16346714	1,2952	5
12	0,13344910	0,13344910	0,17037798	1,3171	4

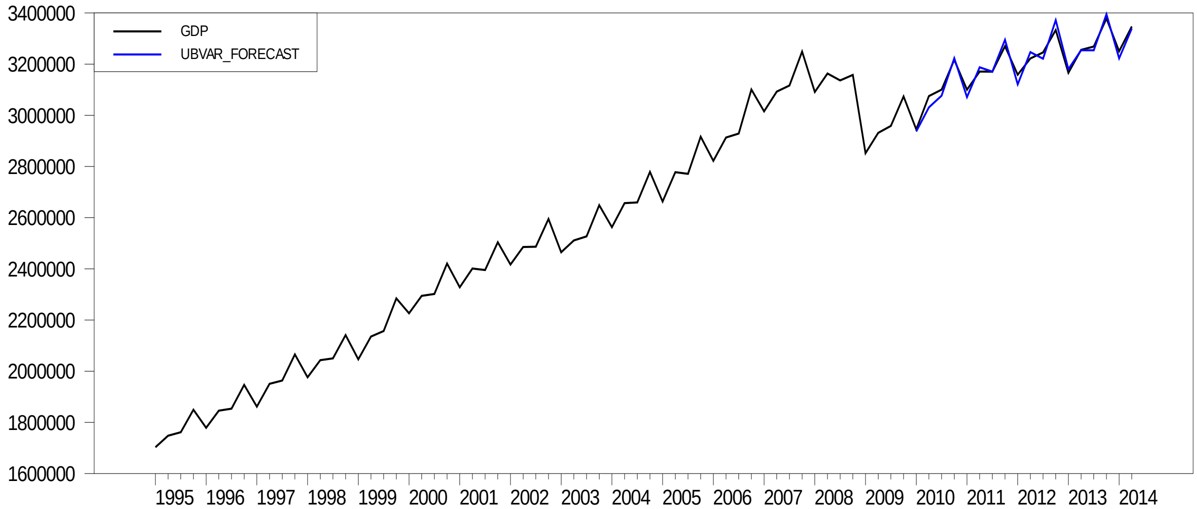
Извор: Истраживање аутора

За серију LOG_ IMP, приказану у Табелама 48-51, могу се извести следећи закључци:

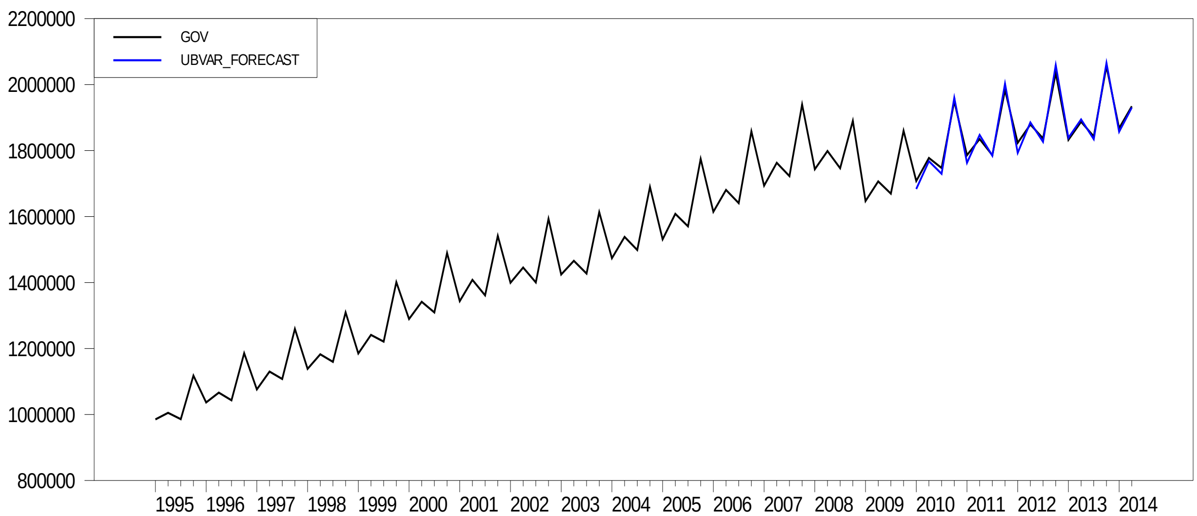
- код униваријантног OLS модела (Табела 48), одабране мере прецизности износе 0,02627804 и 0,9173;
- код другог класичног модела, OLS VAR модела, приказаног у Табели 51, RMSE мера прецизности износи 0,03303173, а Тејлова U-статистика 1,1530;
- у Табели 49 приказан је униваријантни BVAR модел, код ког RMSE и Тејлова U-статистика износе 0,02624972 и 0,9163;
- код једноставног BVAR модела приказаног у Табели 50 вредност одабраних мера прецизности износе 0,02621555 и 0,9151;
- за серију LOG_ IMP, униваријантни и једноставни BVAR модел дају боље резултате од класичних модела: униваријантног OLS и OLS VAR модела;
- *потврђена* је *помоћна хипотеза* да оцене добијене на основу Бајесовог приступа представљају унапређење особина оцена добијених класичним приступом у смислу њихове прецизности.

На основу анализираних резултата, показало се да униваријанти BVAR модел, у већини случајева, даје најбоље резултате у погледу RMSE и Тејлове U-статистике, једино су нешто лошији резултати од униваријантног OLS модела за серију LOG_NONGOV.

С обзиром да се униваријанти BVAR модел показао као најбољи, вредности добијене применом овог модела коришћене су за предвиђање кретања одабраних серија, тако што су након антилогаритмовања предвиђених вредности на посматране серије примењени сезонски коефицијенти. Добијени резултати приказани су сликама 18-22.

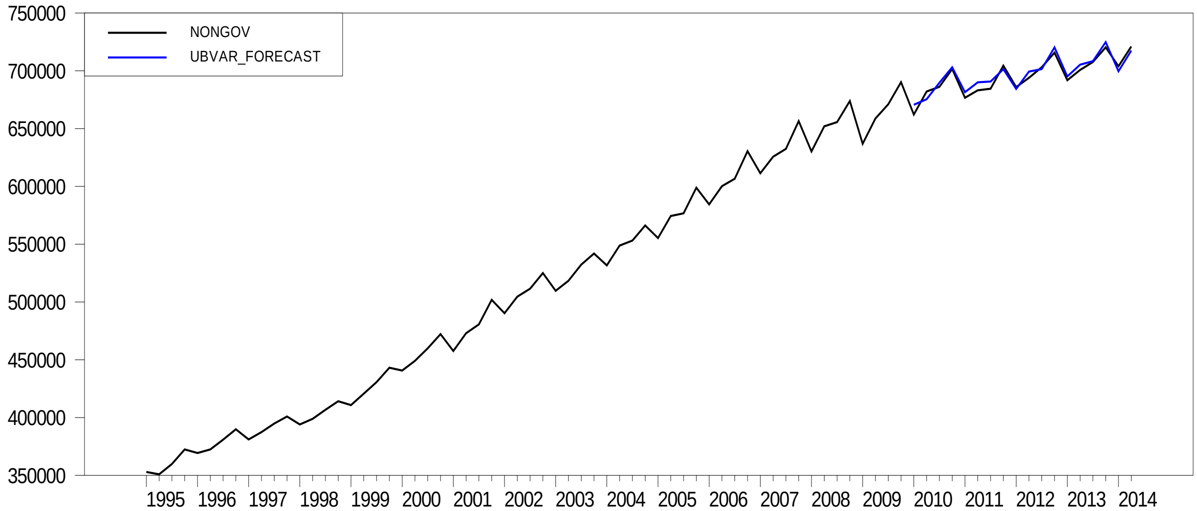


Слика 18. Предвиђање GDP применом униваријантног BVAR модела[158]

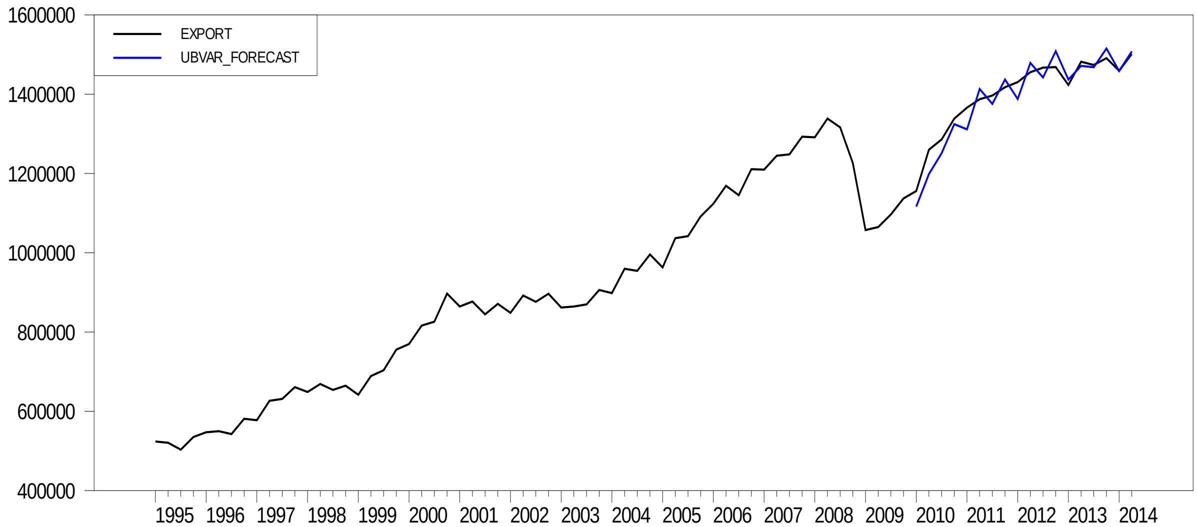


Слика 19. Предвиђање државне потрошње применом униваријантног BVAR модела [159]

[158] Истраживање аутора
[159] Ibid.

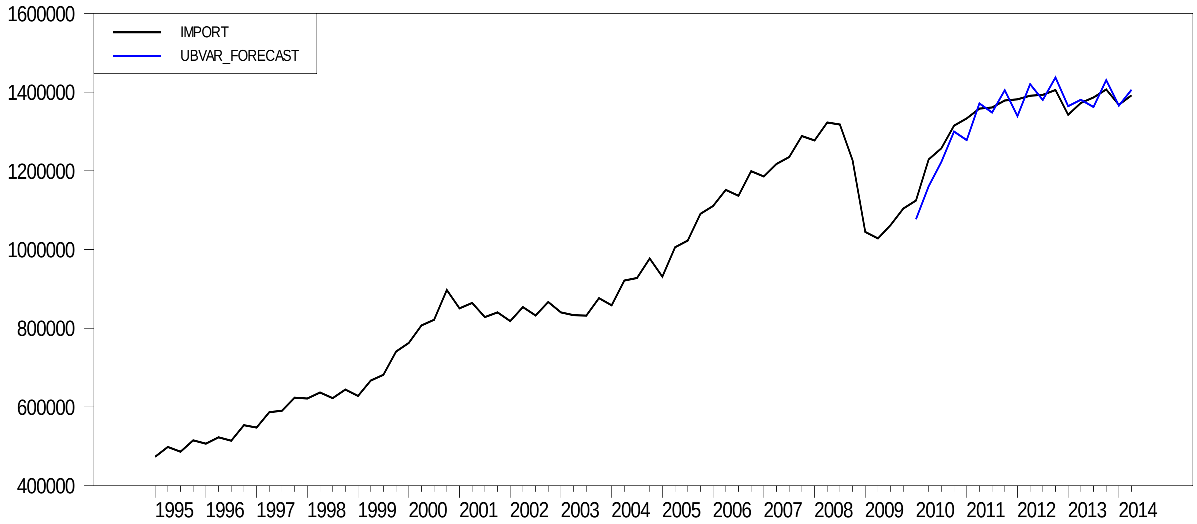


Слика 20. Предвиђање недржавне потрошње применом униваријантног BVAR модела [160]



Слика 21. Предвиђање извоза применом униваријантног BVAR модела [161]

[160] Ibid.
[161] Ibid.



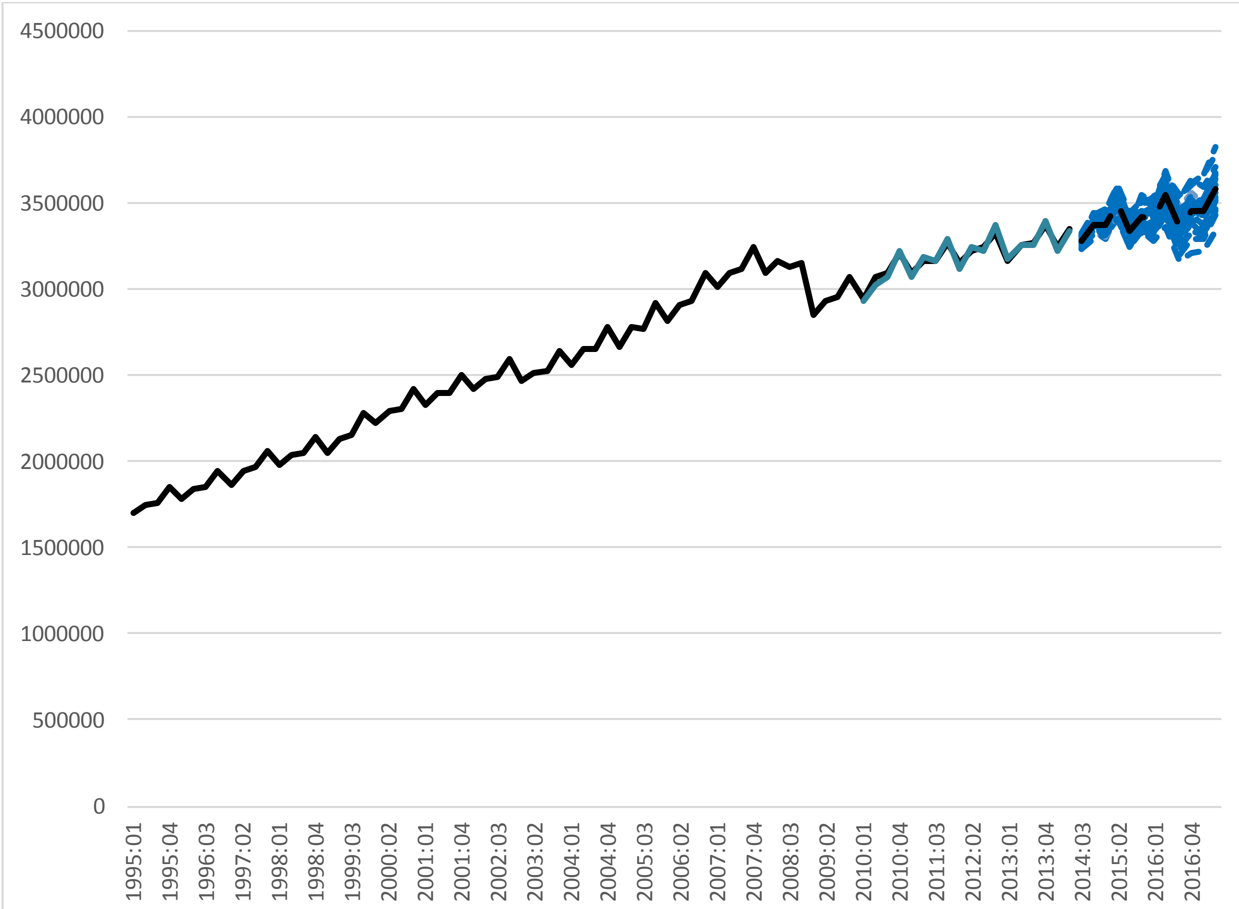
Слика 22. Предвиђање увоза применом униваријантног BVAR модела [162]

У односу на претходни пример, где су одабрани модели истраживања примењени на податке из домаће привреде, у овом примеру, одабрани модели су дали прецизнија предвиђања. Овакав резултат је логична последица примене модела на дужим и стабилнијим временским серијама, због већег броја расположивих информација.

4.3.4.6 Симулација кретања бруто друштвеног производа

Са намером да се изврши предвиђање будућег кретања бруто друштвеног производа извршено је 25 симулација на основу униваријантног BVAR модела. Резултат истраживања је приказан на Слици 23.

[162] Ibid.



Слика 23. Симулација кретања GDP-а применом униваријантног BVAR модела[163]

На слици су црном линијом приказани историјски подаци од првог квартала 1995. год до до другог квартала 2014. године, плавом линијом предвиђање модела од првог квартала 2010. до другог квартала 2014. године, а потом су испрекиданим плавим линијама представљене симулације будућих кретања, док је црна испрекидана линија просечна вредност ових 25 симулација.

Применом два BVAR модела и њиховим поређењем са моделима класичне статистике добијени су резултати, у смислу прецизности предвиђања, који иду у прилог Бајесовим моделима. Резултати јасно показују да BVAR приступ у одређеним испитивањима

[163] Ibid.

превазилази друге приступе, потврђујући опште налазе у литератури и на примеру из Републике Србије и на примеру из Европске Уније.

5. ЗАКЉУЧНА РАЗМАТРАЊА

Предмет истраживања докторске дисертације под називом „*Компаративна анализа класичне инференције и Бајесовог приступа у обради економских података*" је анализа различитих методолошких приступа статистичког закључивања и могућност њихове примене у обради конкретних економских података. Из тог разлога, прво су са теоријског аспекта сагледане карактеристике и перформансе класичне и Бајесове статистике, а затим је извршена компарација одабраних метода на практичним примерима.

Идеја за израду рада проистекла је из честих навода у статистичкој литератури о опречним мишљењима и ставовима присталица класичне и Бајесове инференције. Опречност се развила до те мере да присталице једног правца негирају и одбацују онај други као недовољно адекватан за примену у статистичкој анализи. Посебан изазов за аутора био је испитивање компаративних предности одабраних приступа на примерима из Републике Србије, као и прикупљање одговарајућих података који ће својим карактеристикама задовољити захтеве постављеног истраживања.

Основна хипотеза од које се пошло у раду је да *Бајесов приступ статистичком закључивању има предност у односу на класичну инференцију и да је његова примена у економији оправдана.* Из основне хипотезе изведене су и четири помоћне хипотезе:

1. Бајесов приступ је сложенији и рачунски захтевнији у уобичајеним ситуацијама, с обзиром да је потребно оценити априорну вероватноћу и испитати њену осетљивост;
2. Оцене добијене на основу Бајесовог приступа представљају унапређење особина оцена добијених класичним приступом у смислу њихове прецизности;

3. Развој одговарајућих рачунарских алата доприноси ефикаснијој примени Бајесове статистике у односу на класично закључивање;
4. Оцена поузданости одабраних метода предвиђања захтева квалитативну процену резултата добијених применом Бајесове статистике.

Да би се испитале постављене хипотезе и одговорило на циљеве истраживање, целокупно истраживање је подељено на два дела. У првом делу, који се односио на компарацију метода максималне веродостојности, као метода класичне статистике, и Бајесовог метода, обухваћено је 252 производа који се налазе у продајном асортиману компаније CaliVita International, представништва за Републику Србију Fitco d.o.o., Нови Сад. Изградњом структуралног модела створена је основа за даљу анализу. Пример је јасно показао колико је значајна улога априорне информације у Бајесовом оцењивању, па је тиме **потврђена** прва помоћна хипотеза да је *Бајесов приступ сложенији и рачунски захтевнији у уобичајеним ситуацијама, с обзиром да је потребно оценити априорну вероватноћу и испитати њену осетљивост*. Укључивањем нормалне априорне расподеле, као информативне априорне расподеле, **потврђена** је и друга помоћна хипотеза да *оцене добијене на основу Бајесовог приступа представљају унапређење особина оцена добијених класичним приступом у смислу њихове прецизности*, с обзиром да су се оцене параметара добијене применом Бајесових метода показале као прецизније у односу на класичне оцене.

Класична статистика је дуги низ година имала објективну предност у односу на Бајесов приступ. Оно што је спречавало присталице Бајесове статистике да у први план истакну могућност примене Бајесових метода у обради података, била је реална немогућност примене сложених метода у обради велике количине података. Појава адекватних софтверских решења омогућила је интензивнију примену Бајесових метода у различитим областима истраживања. Опредељење аутора за статистичке пакете IBM SPSS Amos *Version* 21 и RATS *Version* 8.0 и 8.3, који у себи садрже решења за Бајесове методе, **потврдило** је и трећу помоћну хипотезу да *развој одговарајућих рачунарских алата доприноси ефикаснијој примени Бајесове статистике у односу на класично закључивање*.

Закључак о потврди ове хипотезе проистекао је из потврђених осталих хипотеза кроз практичну обраду података.

Компаративна анализа метода класичног и Бајесовог приступа настављена је у наредном делу истраживања којим су обухваћени подаци са сајта Народне банке Србије и Eurostat-a. Потреба за овим делом истраживања, проистекла је из четврте помоћне хипотезе, којом је требало да се испита да ли су резултати *предвиђања* одабраних Бајесових метода прецизнији у односу на резултате добијене применом класичних метода, што је захтевало њихову *квалитативну процену* применом одговарајућих *мера прецизности*.

До резултата истраживања дошло се поређењем две врсте BVAR модела, *униваријантног* BVAR *модела* и *једноставног* BVAR *модела*, који су засновани на различитим априорним информацијама и два класична модела, униваријантног OLS модела и OLS VAR модела. Квалитет резултата добијених применом наведена четири метода оцењен је на основу одабраних мера прецизности и то, пре свега, корена средње квадратне грешке, RMSE, и Тејлове U-статистике, Theil U. Разматрана су два различита примера, где се анализом добијених резултата дошло до закључка да је макар један од посматраних Бајесових модела дао прецизније резултате у односу на класичне моделе.

Спроведено истраживање на финансијским подацима који се тичу кретање каматних стопа на девизне кредите и депозите привреде и становништва, новоодобрене девизне кредите становништву и привреди и новоположене девизне депозите становништва и привреде, за период од јануара 2011. до марта 2014. године, навело је аутора на закључак да је код свих посматраних серија једноставни BVAR модел дао најбоље резултате у погледу RMSE и Тејлове U-статистике, чиме је **потврђена** друга помоћна хипотеза да *оцене добијене на основу Бајесовог приступа представљају унапређење особина оцена добијених класичним приступом у смислу њихове прецизности*.

Други пример из овог дела је обухватио макроекономске показатеље који се односе на Европску унију и то: бруто друштвени производ, државну потрошњу, недржавну потрошњу, извоз и увоз. На основу анализираних резултата, показало се да компаративну предност, у већини случајева, има други Бајесов модел, униваријанти BVAR модел, па је вредност добијена применом овог модела коришћена за предвиђање кретања бруто друштвеног производа. Дужа временска серија, која је обухватила податке од првог квартала 1995. до другог квартала 2014. године, дала је прецизније резултате предвиђања због већег броја расположивих информација. Кроз наведене примере извршена је и квалитативна процена резултата добијених применом Бајесове статистике и оцењена поузданост одабраних метода предвиђања, чиме су испуњени захтеви и четврте помоћне хипотезе.

Разноврсност одабраних примера омогућила је сагледавање примене метода класичне и Бајесове статистике у различитим околностима. Добијени резултати имплицирају извођење и коначног закључка којим се **потврђује** основна хипотеза да *Бајесов приступ статистичком закључивању има предност у односу на класичну инференцију и да је његова примена у економији оправдана.*

Када говоримо о савременим околностима, може се уочити да се одређени број теоретичара снажно залаже за Бајесов приступ, док се, са друге стране, присталице и поштоваоци класичне статистике залажу за примену традиционалних метода у обради података. Ипак, може се закључити да је највећи број савремених интересовања за Бајесову статистику проистекао из чињенице да се код веома сложених модела Бајесове методе могу успешно користити захваљујући одговарајућој софтверској подршци. У годинама које долазе, може се очекивати да ће се наставити са применом Бајесових метода у различитим областима истраживања, те да ће, с обзиром на новије трендове, примена ових метода постати све популарнија и у истраживањима на територији Републике Србије.

ЛИТЕРАТУРА:

1. Afonso, A. and Sousa, R.M. 2012. „The macroeconomic effects of fiscal policy“, *Applied Economics*, 44, 4439–4454.

2. Akaike, H. 1978. A Bayesian Analysis of the Minimum AIC Procedure, *Annals of the Institute of Statistical Mathematics*, 30, Part A, 9-14.

3. Albert, J. 2009. *Bayesian Computation with R, Second Edition*, Springer Science+Business Media LLC, New York.

4. Anscombe, F.J. 1961. „Bayesian statistics“, *The American Statistician*, 15(1), 1-22.

5. Barnett, V. 1999. *Comparative Statistical Inference, Third Edition*, John Wiley & Sons Ltd. Chichester, West Sussexs.

6. Bauwens, L. and Lubrano, M. 1995. „Bayesian and classical econometric modeling of time series“, *Journal of Econometrics*, 69, l–4.

7. Beauchemin, K.R. 2011. „Shocks and the Economic Outlook“, *Economic Commentary*, Number 2011-10

8. Berger, H. and Stavrev, E. 2012. „The information content of money in forecasting euro area inflation“, *Applied Economics*, 44, 4055–4072.

9. Berger, H. and Österholm, P. 2008. „Does money growth granger-cause inflation in the euro area? Evidence from out-of-sample forecasts using Bayesian VARs“, *IMF Working Papers*, WP/08/53, International Monetary Fund.

10. Berger, O.J. 1980. *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag New York Inc.

11. Berger O.J., Brown D.L. and Wolpert L.R. 1994. „ A Unified conditional Frequentist and Bayesian test for fixed and sequential simple hypothesis testing“, *The Annals of Statistics* 22 (4): 1787-1807.

12. Bernardo, J.M. 2001. „Bayesian Statistics“, In *Encyclopedia of Life Support Systems* (EOLSS). UNESCO.

13. Bhuiyan, R. 2012. „Monetary transmission mechanisms in a small open economy: a Bayesian structural VAR approach“, *Canadian Journal of Economics* / *Revue canadienne d'Economique*, Vol. 45, No. 3, 1037-1061.

14. Biswas, D., Singh, S. and Sinha, A. 2010. „Forecasting Inflation and IIP Growth: Bayesian Vector Autoregressive Model“, *Reserve Bank of India Occasional Papers*, Vol. 31, No.2, Monsoon.

15. Boiciuc, I. 2014. „The Macroeconomic Effects of Fiscal Policy. A BVAR Approach“, *Journal of Public Administration*, *Finance and Law*, Special Issue.

16. Bolstad, M.W. 2007. *Introduction to Bayesian Statistics, Second Edition*, John Wiley&Sons, Inc., Hoboken, New Jersey.

17. Bowman, K.O. and Shenton, L.R. 1985. „Method of moments“, *Encyclopedia of Statistical Sciences*, Volume 5, John Wiley & Sons, New York, 467-473.

18. Brenner, D. and Fraser, D.A.S. 1982. „On the foundations: Statistical models and inference”, *Canadian Journal of Statistics* 10 (3): 155–161.

19. Caraiani, P. 2010. „Forecasting Romanian GDP using a BVAR Model“, *Romanian Journal of Economic Forecasting*, 4/2010, 76-87.

20. Carlin, B.P. and Louis, T.A. 2009. *Bayesian Methods for Data Analysis*, *Third Edition*, Chapman & Hall/CRC.

21. Carlin, B.P. and Chib, S. 1995. „Bayesian Model Choice via Markov Chain Monte Carlo Methods", *Journal of the Royal Statistical Society*, *Series B* (*Methodological*), Vol. 57, Issue 3, 473-484.

22. Casella, G. 2008. „Bayesian and Frequentists, Models, Assumptions and Inference", ACCP 37th *Annual Meeting*, Philadelphia, Department of Statistics.

23. Casella, G. and Berger, R.L. 2002. *Statistical Inference*, *Second Edition*, Duxbury Thomson Learning.

24. Chib, S. 1995. „Marginal Likelihood from the Gibbs Output", *Journal of the American Statistical Association*, Vol. 90, No. 432, 1313-1321.

25. Christensen, R. 2005. „Testing Fisher, Neyman, Pearson, and Bayes", *The American Statistician*, 59(2), 121-126.

26. Ciccarelli, M. and Rebucci, A. 2003. „Bayesian VARs: A Survey of the Recent Literature with an Application to the European Monetary System", IMF *Working Papers* 03/102, International Monetary Fund.

27. Congdon, P. 2003. *Applied Bayesian Modelling*, John Wiley and Sons Ltd.

28. Congdon, P. 2001. *Bayesian Statistical Modelling*, John Wiley and Sons Ltd.

29. Cox, D.R. 2006. „Frequentist and Bayesian Statistics: A critique", Statistical problems in particle physics, astrophysics and cosmology: 1–4. *Imperial College Press*, London.

30. Cox, D.R. 2006. *Principles of statistical inference*, Cambridge University Press.

31. Depaoli, S. 2013. „Mixture class recovery in GMM under varying degrees of class separation: Frequentist versus Bayesian estimation", *Psychological Methods*, 18(2), 186-219.

32. Dienes, Z. 2011. „Bayesian Versus Orthodox Statistics: Which Side Are You On?”, *Perspectives on Psychological Science*, 6(3): 274–290.

33. Doan, T. 2007. RATS User’s Manual, *Version* 7, Estima.

34. Doan, T., Litterman, R. and Sims, C. 1984. „Forecasting and conditional projection using realistic prior distributions“, *Econometric Reviews*, 3, 1–100.

35. Downey, A.B. 2012. *Think Bayes*, *Bayesian Statistics Made Simple*, *Version* 1.0., Green Tea Press, Needham, Massachusetts.

36. Ђорђевић, В. 2004. *Статистика у економији*, Економски факултет, Ниш.

37. Edwards, W., Lindman, H. and Savage, L.J. 1963. „Bayesian Statistical Inference for Psychological Research”, *Psychological Review*, Vol. 70, No. 3, 193-242.

38. Fernández-Villaverde, J. and Rubio-Ramírez, J.F. 2007. „Estimating macroeconomic models: A likelihood approach“, *The Review of Economic Studies*, 74(4), 1059-1087.

39. Fienberg, S.E. 2006. „When Did Bayesian Inference Become “Bayesian”?”, *Bayesian Analysis*, International Society for Bayesian Analysis, No.1, 1-40.

40. Fischer, B., Lenza, M., Pill, H. and Reichlin, L. 2008. „Money and monetary policy: the ECB experience 1999–2006, in The Role of Money – Money and Monetary Policy in the Twenty-First Century“, *Proceedings of the 4th ECB Central Banking Conference*, 9–10 November 2006 (Eds.), A. Beyer and L. Reichlin, ECB, Frankfurt, 102–175.

41. Fisher, R.A. 1970. *Statistical methods for Research Workers*, *Fourteenth Edition*, Oliver & Boyd, Edinburgh.

42. Fisher, R.A. 1966. *The Design of Experiments*, *Eight Edition*, Oliver and Boyd, Edinburgh.

43. Fisher, R.A. 1930. „Inverse probability“, *Mathematical Proceedings of the Cambridge Philosophical Society*, Volume 26, Issue 04, 528-535.

44. Fisher, R.A. 1922. „On the Mathematical Foundations of Theoretical Statistics”, *Philosophical Transactions of the Royal Society of London*, Series A, Vol. 222, 309-368.

45. Flury, T. and Shephard, N. 2011. Bayesian inference based only on simulated likelihood: particle filter analysis of dynamic economic models, *Econometric Theory*, 27(05), 933-956.

46. Gau, W.C, Gangopadhyay, A. and Han, Z. 2008. „Interval Estimation of the Credibility Factor“, *Variance*, Casualty Actuarial Society, 2(1): 71-84.

47. Gelman, A., Carlin, B.J., Stern, S.H. and Rubin, B.D. 2004. *Bayesian Data Analysis, Second Edition*, Chapman&Hall/CRC.

48. Geweke, J. 2005. *Contemporary Bayesian Econometrics and Statistics*, John Wiley & Sons, Inc.

49. Ghosh, J.K, Delampady, M. and Samanta, T. 2006. *An Introduction to Bayesian Analysis, Theory and Methods*, Springer Science+Business Media, LLC, New York.

50. Gillard, J.W. and Iles, T.C. 2005. „Method of Moments Estimation in Linear Regression with Errors in both Variables“, *Cardiff University School of Mathematics Technical Paper*, Cardiff, 3-43.

51. Griffiths, T.L. and Yuille, A. 2006. „A primer on probabilistic inference“, *Trends in Cognitive Sciences*, Supplement to special issue on Probabilistic Models of Cognition, 10 (7): 1–11.

52. Хаџивуковић, С. 1979. *Статистика*, Издавачка радна организација „Рад“, Београд.

53. Hadživuković, S. and Nikolić-Đorić, E. 2005. „R.A. Fisher and Modern Statistics“, *International Statistical Institute*, 55th Session.

54. Хаџивуковић, С. 2008. „Статистички модели“, *Статистичка ревија*, 57(3-4): 33-44.

55. Hald, A. 2004. *A History of Parametric Statistical Inference from Bernoulli to Fisher 1713 to 1935*, University of Copenhagen, Department of applied mathematics and statistics, Copenhagen.

56. Hamadicharef, B. 2010. „Frequentist versus Bayesian approaches for AUC confidence intervals bounds“, *10th International Conference on Information Science, Signal Processing and their Applications* (ISSPA).

57. Han, C. and Carlin, B. 2001. „Markov Chain Monte Carlo Methods for Computing Bayes Factors, A Comparative Review”, *Journal of the American Statistical Association*, Vol. 96, Issue 455, 1122-1132.

58. Hastings, W.K. 1970. „Monte Carlo sampling methods using Markov chains and their applications“, *Biometrika*, Vol. 57, No. 1, 97-109.

59. Hildebrand, D. and Ott, A.L. 1998. *Statistical Thinking for Managers*, Brooks/Cole, New York.

60. Hoff, D.P. 2009. *A First Course in Bayesian Statistical Methods*, Springer Science + Business Media LLC, New York.

61. Hubbard, R., Bayarri, M.J., Berk, K.N. and Carlton, M.A. 2003. „Confusion over measures of evidence (p’s) versus errors (α’s) in classical statistical testing“, *The American Statistician* 57: 171-182.

62. Ивковић, З. 1992. *Математичка статистика*, Научна књига, Београд.

63. Jaynes, T.E. 1968. „Prior Probabilities“, *IEEE Transactions On Systems Science and Cybernetics*, Vol. sec-4, No. 3, 227-241.

64. Jeffreys, H. 1961. *Theory of Probability*, *Third Edition*, Clarendon Press, Oxford.

65. Kapadia, A.S., Chan, W. and Moye, L. 2005. *Mathematical Statistics with Applications*, Taylor and Francis Group, LLC.

66. Kass, R.E. and Raftery, A.E. 1995. „Bayes factor", *Journal of the American Statistical Association*, Vol. 90, No. 430, 773-795.

67. Kim, S.Y., Suh, Y., Kim, J.S., Albanese, M., and Langer M.M. 2013. „Single and multipleability estimation in the SEM framework: a non-informative Bayesian estimation approach", *Multivariate and Behavioral Research*, 48(4) , 563-591.

68. Knight, K. 2000. *Mathematical Statistics*, Chapman & Hall/CRC.

69. Koch, K.R. 2007. *Introduction to Bayesian Statistics, Second Edition*, Springer-Verlag, Berlin.

70. Koop, G. 2003. *Bayesian Econometrics*, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, England.

71. Koop, G. and Korobilis, D. 2009. „Bayesian Multivariate Time Series Methods for Empirical Macroeconomics", *Foundations and Trends in Econometrics*, Vol. 3, No. 4, 267–358.

72. Koski, T. and Noble, M.J. 2009. *Bayesian Networks, An Introduction*, John Wiley & Sons, Ltd.

73. Kvam, P.H. and Vidakovic, B. 2007. *Nonparametric Statistics with Applications to Science and Engineering*, John Wiley & Sons, New Jersey.

74. Lamm-Tennant, J., Starks, L. and Stokes, L. 1992. „An Empirical Bayes Approach to Estimating Loss Ratios", *The Journal of Risk and Insurance* 59(3): 425-442.

75. La Valle, S.M. 2006. *Planning Algorithms*, Cambridge University Press.

76. Lee, S.Y. and Song, X.Y. 2004. „Evaluation of the Bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes", *Multivariate Behavioral Research*, 39, 653–686.

77. Lehmann, E.L. 2008. *Reminiscences of a Statistician, The Compani I Kept*. Springer Science+Business Media LLC, New York.

78. Lehmann, E.L. and Romano, P.J. 2005. *Testing Statistical Hypotheses, Third Edition*, Springer Science+Business Media, LLC.

79. Lehmann, E.L. 1993. „The Fisher, Neyman-Pearson Theories of Testing Hypotheses: One Theory or Two?", *Journal of the American Statistical Association*, 88, 1242-1249

80. Lemmon, A.R., Brown, J. M., Stanger-Hall, K., and Lemmon, E.M. 2009. „The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and Bayesian inference", *Systematic Biology*, 58(1), 130-145.

81. Lenhard, J. 2006. „Models and Statistical Inference: The Controversy between Fisher and Neyman–Pearson", *Oxford University Press on behalf of British Society for the Philosophy of Science*, Vol. 57, 69–91.

82. Лепојевић, В. и Јанковић-Милић, В. 2011. „Бајесијанско моделирање у анализи здружених ефеката", *Маркетинг* 3: 171-178.

83. Лепојевић, В., Ђорђевић В. и Јанковић-Милић, В. 2007. „Two-way analysis of variance for random models", *Economic themes* 4: 89-99, Економски факултет Ниш.

84. Лепојевић, В. 2004. „F tests in the three-way analysis of variance", *Economic themes*, 6: 89-99, Економски факултет Ниш.

85. Lindgren, W.B. 1993. *Statistical theory, Fourth Edition*, Chapman & Hall/CRC, Boca Raton, Florida.

86. Lindley, D.V. 1953. „Statistical Inference", *Journal of the Royal Statistical Society*, Series B (Methodological), Vol. 15, No. 1, 30-76.

87. Litterman, R.B. 1986. „Forecasting with Bayesian vector autoregressions: Five years of experience", *Journal of Business and Economic Statistics* 4, 25-38.

88. Litterman, R. 1980. „Techniques for Forecasting with Vector Autoregressions”, University of Minnesota, doctoral dissertation.

89. Lütkepohl, H. 2007. „Econometric Analysis with Vector Autoregressive Models”, EUI *Working papers*, ECO 2007/11, 1-56.

90. Lütkepohl, H. 1999. „Vector autoregressive analysis”, Discussion Papers, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes, No. 1999, 31, 1-25.

91. Makridakis, S. and Hibon, M. 1995. „Evaluating Accuracy (or Error) Measures”, INSEAD *Working Paper Series*, 95/18/TM, INSEAD, Fontainebleau, France, 1-31.

92. Marin, J.M. and Robert, C.P. 2007. *Bayesian Core*: *A Practical Approach to Computational Bayesian Statistics*, Science+Business Media, LLC.

93. Mayo, D.G. and Cox, D.R. 2006. „Frequentist statistics as a theory of inductive inference“, *2nd Lehmann Symposium* - Optimality IMS Lecture Notes - Mongraphs Series.

94. Meng, X. and Wong, W.H. 1996. „Simulating Ratios of Normalizing Constants via a Simple Identity: a theoretical exploration”, *Statistica Sinica*, 6, 831-860.

95. Migliardo, C. 2010. „Monetary Policy Transmission in Italy: A BVAR Analysis with Sign Restriction“, *Czech Economic Review*, 4 (2010): 139–167.

96. Милошевић, В. 1995. *Теорија статистичког закључивања*, Научна књига, Београд.

97. Mittelhammer, C.R. 1996. *Mathematical Statistics for Economics and Business*, Springer-Verlag New York Inc.

98. Newbert, S.L. 2007. „Empirical research on the resource-based view of the firm: An assessment and suggestions for future research“, *Strategic Management Journal* 28 (2): 121-146.

99. Newbold, P., Carlson L.W. and Thorne, B. 2002. *Statistics for Business and Economics*, Prentice Hall, New Jersey.

100. Nguyen, H.T., Rogers, G.S. and Nguyen, T.H. 1989. *Fundamentals of Mathematical Statistics, Volume I: Probability for Statistics*, Springer-Verlag New York Inc.

101. Norton, J.D. 2008. „Ignorance and Indifference“, *Philosophy of Science* 75: pp. 45–68.

102. O'Hagan, T. 2004. „Dicing with the unknown“, *Significance Magazine* 1(3): 132-133.

103. Panchapakesan, S. and Balakrishnan, N. 1997. *Advances in Statistical Decision Theory and Applications*, Birkhauser, Boston.

104. Pestman, R.W. 1998. *Mathematical Statistics*, Walter de Gruyter Gmbh&Co, Berlin. 1998

105. Петровић, Љ. 2003. *Теорија вероватноћа*, Економски факултет, Београд.

106. Петровић, Љ. 2006. *Теоријска статистика. Теорија статистичког закључивања*, Економски факултет, Београд.

107. Pitt, M.K., Silva, R.D.S., Giordani, P., and Kohn, R. 2012. „On some properties of Markov chain Monte Carlo simulation methods based on the particle filter“, *Journal of Econometrics*, 171(2), 134-151.

108. Poirier, D. 2008. *Bayesian Econometrics*, In S.N. Durlauf and L.E. Blume (eds), The New Palgrave Dictionary of Economics, Palgrave Macmillan.

109. Поповић, Б. 2003. *Математичка статистика и статистичко моделовање*, Природно математички факултет, Ниш.

110. Rachev, T.S., Hsu, J.S.J., Bagasheva, B.S. and Fabozzi, F.J. 2008. *Bayesian Methods in Finance*, John Wiley & Sons, Inc., New Jersey

111. Ramachandran, M.K. and Tsokos, P.C. 2009. *Mathematical Statistics with Applications*, Elsevier Academic Press.

112. Raats V.M. and Moors J.J.A. 2003. „Double-checking auditors: a Bayesian approach“, *Journal of the Royal Statistical Society: Series D (The Statistician)* 52(3): 351–365.

113. Ravishanker, N. and Ray, B. 1997. „Bayesian analysis of vector ARMA models using Gibbs sampling“, *Journal of Forecasting*, 16, 177–194.

114. Ritschl, A. and Woitek, U. 2000. „Did monetary forces cause the great depression? A Bayesian VAR analysis for the US economy“, *Discussion Paper Series*, Centre For Economic Policy Research, London.

115. Robert, P.C. 2007. *The Bayesian choice, From Decision-Theoretic Foundations to Computational Implementation*, *Second Edition*, Springer Science + Business Media LLC, New York.

116. Rohatgi, V.K. 1976. *An Introduction to Probability Theory and Mathematical Statistics*, John Wiley & Sons.

117. Rohatgi, V.K. 1984. *Statistical Inference*, John Wiley & Sons, New York.

118. Ross, M.S. 2009. *Probability and Statistics for Engineers and Scientists, Fourth Edition*, Academic Press.

119. Савић, М. 2005. *Пословна статистика*, Економски факултет, Суботица.

120. Scales, J.A. and Tenorio, L. 2001. „Tutorial: Prior information and uncertainty in inverse problems“, *Geophysics*, Vol. 66, No. 2, 389–397.

121. Schmidheiny, K. 2013. „The Multiple Linear Regression Model“, *Short Guides to Microeconometrics*, *Version*: 12-9-2013, 20:29, 1-15.

122. Schoot, R.V.D. and Depaoli, S. 2014. „Bayesian analyses: where to start and what to report“, *The European Health Psychologist*, Volume 16, Issue 2, 75-84.

123. Senn, S. 2011. „You May Believe You Are a Bayesian But You Are Probably Wrong“, *RMM*, Volume 2, Special Topic: Statistical Science and Philosophy of Science, 48-66.

124. Senn, S. 2003. „Bayesian, Likelihood, and Frequentist Approaches to Statistics, A comparison of methods“, *Applied Clinical Trials*, 35-38.

125. Shao, J. 2005. *Mathematical Statistics: Exercises and Solutions*, Springer Science+Business Media, Inc.

126. Shi, N.Z. and Tao, J. 2008. *Statistical Hypothesis Testing*, World Scientific Publishing Co. Pte. Ltd.

127. Silvey, S. D. 1975. *Statistical inference*, Chapman & Hall/CRC, Boca Raton, Florida.

128. Sims, C. 1980. „Macroeconomics and reality“, *Econometrica* 48, 1–48.

129. Sims, C. and Zha, T. 1998. „Bayesian methods for dynamic multivariate models”, *International Economic Review*, 39, 949–968.

130. Smith, R.L. 1982. „The Bayesian Inference Method and Its Application to Reliability Problems“, *NRL Memorandum Report 4903*, Texas Research Institute.

131. Smith, C.A.B. 1965. „Personal Probability and Statistical Analysis“, *Journal of the Royal Statistical Society*, Series A (General), Vol. 128, No. 4, 469-499.

132. Stauffer, B.H. 2008. *Contemporary Bayesian and Frequentist Statistical Research Methods for Natural Resource Scientists*, John Wiley & Sons, Inc., Hoboken, New Jersey.

133. Stigler M.S. 1982. „Thomas Bayes′ Bayesian Inference“, *Journal of the Royal Statistical Society*, Series A, 145: 250–258.

134. Stock, J.H. and Watson, M.W. 2005. „Implications od Dynamic Factor Models for VAR Analysis”, NBER *Working Paper* No. 11467, 1-65.

135. Swiler, L.P. 2006. „Bayesian Methods in Engineering Design Problems“, *Sandia National Laboratories*, Albuquerque, New Mexico.

136. Tableman, M. and Kim, J.S. 2004. *Survival Analysis Using S, Analysis of Time-to-Event Data*, Chapman & Hall/CRC.

137. Tang, Y.C. and Liou, F.M. 2010. „Does firm performance reveal its own causes? The role of Bayesian inference“, *Strategic Management Journal* 31: 39-57.

138. Terrell, R.G., Casella, G. and Olkin, I. 1999. *Mathematical Statistics, A Unified Introduction*, Springer-Verlag New York Inc.

139. Vallverdú, J. 2008. „The False Dilemma: Bayesian vs. Frequentist“, *E-Logos*, *Electronic Journal for Philosophy*, ISSN 1211-0442.

140. Васић, В. 2003. „Фамилија метода Марковљеви ланци Монте Карло у анализи некомплетних података“, *Економски анали*, бр. 159: 147-158.

141. Vidakovic, B. 2011. *Statistics for Bioengineering Sciences, With MATLAB and WinBUGS Support*, Springer Science+Business Media, LLC.

142. Вукадиновић, С. и Поповић, Ј. 2004. *Математичка статистика*, Саобраћајни факултет Универзитета у Београду, Београд.

143. Wagenmakers, E.J., Lee, M.D., Lodewyckx, T. and Iverson, G. 2008. „Bayesian versus frequentist inference“, In H. Hoijtink, I. Klugkist, & P. A. Boelen (Eds.). *Bayesian Evaluation of Informative Hypotheses*, 181-207. Springer, New York.

144. Walsh, B. 2004. „Markov Chain Monte Carlo and Gibbs Sampling“, *Lecture Notes for EEB* 581, *Version* 26.

145. Ward, E.J. 2008. A review and comparison of four commonly used Bayesian and maximum likelihood model selection tools, *Ecological Modelling*, 211(1), 1-10.

146. West, M. and Harrison, J. 1997. *Bayesian Forecasting and Dynamic Models*, *Second Edition*, Springer-Verlag, New York, Inc.

147. Wilks, S.S. 1943. *Mathematical Statistics*, Princeton University Press.

148. Young, G.A. and Smith, R.L. 2005. *Essentials of statistical inference*, Cambridge University Press, Cambridge.

149. Zhang, Z., Hamagami, F., Wang, L., Grimm, K.J., and Nesselroade, J.R. 2007. „Bayesian analysis of longitudinal data using growth curve models“, *International Journal of Behavioral Development*, 31 (4) , 374-383.

150. Жижић, М., Ловрић, М. и Павличић, Д. 1999. *Методи статистичке анализе*, Економски факултет, Београд.

151. http://epp.eurostat.ec.europa.eu/

152. https://www.estima.com/ratsdemo.shtml

153. http://www.nbs.rs/

154. http://www.calivita.rs/

155. http://forrest.psych.unc.edu/research/vista-frames/pdf/chap08.pdf

Прилог 1. Одабрани производи компаније CaliVita International, представништва за Републику Србију Fitco d.o.o., Нови Сад

	Naziv proizvoda	pr_jan	pr_feb	pr_mart	pr_april	pr_maj	pr_jun	vr pr	pr	cena
1	100% Whey Protein Chocolate; 1Kg	4	5	10	30	16	11	8	1	3290
2	100% Whey Protein Strawberry; 1Kg	0	2	2	0	0	3	8	1	3290
3	100% Whey Protein Vanilla; 1Kg	0	5	2	6	12	6	8	1	3290
4	Acidophilus with Psyllium 100 kapsula	31	36	0	28	30	35	9	3	2429
5	AC-zymes, 100 kapsula	51	128	0	51	57	54	9	3	1849
6	Amino Fuel Liq. concentr. 32. oz	145	48	0	66	42	21	8	7	2250
7	Aquabelle hidratantni losion za čišćenje lica	7	3	21	3	3	5	3	2	1439
8	Aquabelle hranjiva krema	11	4	63	2	9	9	3	2	1199
9	Aquabelle regenerativna krema	10	8	26	7	14	2	3	2	1199
10	Arginine PLUS 100 k	18	10	0	12	21	21	2	3	2429
11	Beauty formula, 90 tableta	16	10	0	49	52	32	2	2	2429
12	Bee Power 50 kapsula	57	58	16	60	61	42	7	3	1109
13	Beta karoten PN, 100 tableta	9	125	0	28	25	36	10	4	1159
14	Beta karoten, 25000IU, 100 tableta	0	0	13	0	0	0	10	4	1159
15	C-1000, 100 tableta	0	157	23	0	0	204	10	2	2199
16	C-300 plus, 120 kaps	354	268	50	330	288	136	10	2	1848
17	C-500 Delayed Release 100 tableta	0	109	0	0	0	65	10	3	1539
18	CaliClean Prirodno sredstvo za čišćenje kuhinje	2	6	0	1	0	4	11	2	859
19	CaliClean Prirodno sredstvo za čišćenje kupatil	4	7	0	0	2	1	11	2	859
20	CaliClean Prirodno sredstvo za čišćenje opšte n	5	4	0	0	3	2	11	2	859
21	CaliClean Prirodno sredstvo za čišćenje stakla	3	5	0	0	3	2	11	2	859
22	CaliClean Prirodno sredstvo za čišćenje toaleta	1	4	0	3	3	2	11	2	859
23	CF BCAA 400 GR	13	34	34	32	35	19	8	1	2490
24	CF DAA 180 CAP	0	11	6	11	44	10	8	1	1690
25	CF Liquid Amino 32oz	17	16	21	26	23	23	8	1	2290
26	CF THERMOXTREME HC ULTRA 120 CAPS	6	2	33	25	27	22	8	1	1790
27	CF Vitamin D3	4	8	6	12	10	5	8	1	675

28	CF WHEY AMINO 180 TABS	63	29	2	66	49	40	8	1	2190
29	Chelated Zinc 15.0, 100 komada	63	53	0	92	122	27	5	3	799
30	Cholestone 90 tableta	99	85	0	76	92	72	2	2	1929
31	Chromium Max 500, 100 tableta	45	45	36	63	66	41	4	3	1579
32	Citrimax and Chromium	30	34	0	24	26	32	4	3	1749
33	COQ10 sublingual limun, 30 tableta	263	195	100	244	235	195	1	2	2119
34	COQ10 sublingual; 30 tableta	0	0	78	0	0	0	1	2	2119
35	C-Plus Flavonoid 100 tableta	354	208	8	163	234	301	10	3	919
36	Creatine Monohydrate CF; 500g	179	97	124	111	115	119	8	1	1590
37	Creatine monohydrate; 300+300 gr	5	0	0	38	44	44	8	5	2590
38	C-RX Krema, 150ml	11	2	15	33	42	22	3	2	1649
39	Culevit 180 tableta	20	25	0	21	14	17	2	10	5039
40	Cytogenix Xenadrine Results 60	3	0	0	13	10	1	8	16	2790
41	Cytogenix Xenadrine Ripped 120	45	0	0	2	2	41	8	16	4990
42	Cytogenix Xenadrine Ultra Mix 3.8 G 21 pak	6	0	0	27	32	10	8	16	4990
43	Digestive Enzymes 100 tabs	113	84	73	93	128	91	9	2	2249
44	DYM BCAA 300g	12	0	0	13	38	20	8	8	2790
45	DYM Creatine Mono 500g	37	30	0	127	166	81	8	8	1990
46	DYM Elite Casein Vanilla 2 LBS.	1	0	0	1	3	0	8	8	3190
47	DYM ELITE GOURMET CAPUCCINO 907g	8	0	0	26	6	0	8	8	3490
48	DYM Elite Gourmet Milk Choc, 5 LBS	24	0	0	33	8	0	8	8	7490
49	DYM Elite Gourmet Straw Cream, 5LBS	6	0	0	0	2	0	8	8	7490
50	DYM ELITE MASS COOKIES AND CREAM 2.7kg	4	1	0	5	14	4	8	8	4990
51	DYM ELITE MASS COOKIES AND CREAM 4,5kg	5	0	0	0	9	0	8	8	7390
52	DYM ELITE MASS DOUBLE CHOCOLATE 2,7kg	5	20	0	17	0	11	8	8	4990
53	DYM ELITE MASS VANILLA ICE CREAM 2.7kg	2	0	0	4	0	5	8	8	4990
54	DYM ELITE MASS VANILLA ICE CREAM 4,5kg	2	1	0	16	5	1	8	8	7390
55	DYM ELITE PRIMAL BLUE RASP 920g	2	1	0	1	5	2	8	8	3290
56	DYM ELITE PRIMAL FRUIT	0	12	0	2	11	4	8	8	3290

	PUNCH 920g									
57	DYM Elite Whey Chocolate 4.5Kg	5	0	0	35	27	13	8	8	11900
58	DYM ELITE WHEY XT BLUEBERRY MUFFIN 2kg	6	0	0	6	9	3	8	8	4790
59	DYM ELITE WHEY XT RICH CHOCOCOLATE 1kg	14	0	0	0	1	0	8	8	2490
60	DYM ELITE WHEY XT RICH CHOCOLATE 2kg	25	0	0	43	49	19	8	8	4790
61	DYM ELITE WHEY XT RICH VANILLA 1kg	6	0	0	44	30	0	8	8	2490
62	DYM ELITE WHEY XT RICH VANILLA 2kg	4	0	0	0	0	0	8	8	4790
63	DYM Glutamine 500g	28	74	0	28	33	33	8	8	3990
64	DYM ISO 100 Whey choco 5lbs	12	1	0	25	105	0	8	8	3485
65	DYM ISO 100 Whey vanilla 5lbs	0	8	0	8	9	0	8	8	3485
66	DYM L-Carnitine Liquid Berry 31 Serv	1	0	0	1	0	0	8	8	1990
67	DYM L-Carnitine Liquid Blue Rasp. 31 Serv	18	46	0	78	62	43	8	8	1850
68	DYM SUPER AMINO 4800 450 CT	0	13	0	0	0	31	8	8	3090
69	DYM SUPER AMINO 4800mg 325 kapleta	0	44	0	62	86	37	8	8	2490
70	DYM Super Mass Gainer Chocolate 5.5kg	21	0	0	20	5	3	8	8	5890
71	DYM Super Multi Vitamin, 120 caps	9	0	0	0	2	57	8	8	1990
72	DYM Xpand Xtreme Pump Orange 40 serv	2	3	0	39	43	29	8	8	5390
73	DYM Xpand Xtreme Pump Raspberry 40 serv	0	1	0	39	37	27	8	8	5390
74	Elite Whey Berry, 2.27Kg	0	21	0	53	15	8	8	8	6990
75	Elite Whey Berry, 907g	0	0	0	12	14	10	8	8	3390
76	Elite Whey Cafe Moncha, 907g	0	0	0	11	19	21	8	8	3390
77	Elite Whey Chocolate, 2.27Kg	20	89	0	118	111	89	8	8	6990
78	Elite Whey Chocolate, 907g	16	84	0	35	76	62	8	8	3390
79	Elite Whey Vanilla, 2.27Kg	0	26	0	0	0	47	8	8	6990
80	Elite Whey Vanilla, 907g	17	0	0	28	29	17	8	8	3390
81	Energy and Memory; 90 tabs	60	37	0	38	42	31	1	2	2249
82	Evening Primrose Oil 100 softgels	41	99	49	38	32	33	7	3	1939
83	Figure shaper, 60 kapsula	9	3	0	3	4	2	7	2	4499
84	Full spectrum multivitamin 90 k.	179	209	7	205	176	132	6	4	1099
85	Garlic with 250 mg Parsley 100	54	74	4	63	80	60	7	4	1159

	gel kapsula									
86	GASP INTRAPRO PROTEIN BERRY 907g	20	0	125	0	0	0	8	9	3590
87	GASP INTRAPRO PROTEIN CHOCOLATE 907g	17	0	0	0	0	0	8	9	3590
88	GASP MYOFUSION HYDRO BANANA 2,27kg	23	7	0	12	36	17	8	9	6990
89	GASP MYOFUSION HYDRO CHOCOLATE 2,27kg	7	0	0	58	90	64	8	9	6990
90	GASP MYOFUSION HYDRO CHOCOLATE 907g	51	0	40	75	58	49	8	9	3590
91	GASP MYOFUSION HYDRO COOKIES & CREAM 2,27kg	44	0	0	21	34	18	8	9	6990
92	GASP MYOFUSION HYDRO STRAWBERRY 907g	14	0	31	18	10	14	8	9	3590
93	GASP MYOFUSION HYDRO STRAWBERRY2,27kg	22	0	0	11	0	0	8	9	6990
94	GASP MYOFUSION HYDRO VANILLA 2,27kg	10	0	0	8	28	35	8	9	6990
95	GASP MYOFUSION HYDRO VANILLA 907g	0	0	62	10	35	3	8	9	3590
96	GASP MYOFUSION PROTEIN CHOC MINT 2,27kg	0	0	17	0	0	0	8	9	6990
97	GASP REAL MASS CHOCOLATE 2,7kg	7	3	23	6	11	4	8	9	5290
98	GASP REAL MASS STRAWBERRY 2,7kg	0	0	167	2	2	1	8	9	5290
99	GASP REAL MASS Vanilla 2,7kg	2	0	0	0	1	1	8	9	5290
100	GASP SUPERPUMP MAX BLUE RASP 640g	0	7	0	14	26	7	8	9	5390
101	GASP SUPERPUMP MAX ORANGE 640g	2	34	0	55	36	13	8	9	5390
102	GASP SUPERPUMP MAX PINK LEMONADE 640g	0	2	0	11	27	1	8	9	5390
103	Ginkgo XC, 100 tabl	79	66	0	86	68	66	7	2	1499
104	Glutamine 400g	96	163	85	103	88	70	8	1	1690
105	Green Care, lucerka, 240 tableta	47	35	114	36	37	101	7	3	1709
106	Hand and Body lotion; 1000ml	41	52	0	51	54	62	3	2	1429
107	Immunaid; 180 kapsula	47	63	0	35	47	37	7	2	3629
108	Infant Formula, bebi multivitamin	25	28	35	24	21	23	6	3	1649
109	IntiMoments, 135g	13	15	0	14	13	13	3	2	1339
110	Iron Plus 100 tableta	152	164	0	188	179	167	5	2	1159
111	IS OH YEAH VANILLA TOFFEE FUDGE 12 BARS	7	8	3	3	2	5	8	11	2390
112	IS OhYeah almond fudge brownie 45g 12 bars	4	0	0	7	7	0	8	11	2390

113	IS OhYeah chocolate&caramel 45g 12 bars	3	0	12	4	4	0	8	11	2390
114	IS OhYeah Cookie Caramel crunch 45g 12 bars	2	0	8	3	3	2	8	11	2390
115	IS OhYeah Cookie Caramel crunch 85g 12 bars	4	2	25	25	31	4	8	11	3790
116	IS OhYeah vanilla toffee fudge 45g 12 bars	0	0	4	8	6	1	8	11	2390
117	ISS OH YEAH Peanut butter caramel 45g	2	0	12	4	3	0	8	11	2390
118	ISS OH YEAH RTD C&C 414ml 12kom	6	3	10	6	4	4	8	11	4890
119	ISS OH YEAH RTD Vanila 414ml 12kom	4	3	58	5	8	0	8	11	4890
120	Joint Fuel Liquid 16oz	0	1	0	6	22	6	8	7	2890
121	JOINT Protex 90tabs	78	87	47	70	91	69	2	2	3139
122	Joint Protex Forte, 90 tableta	69	83	155	77	90	72	2	2	4119
123	Kids Greens, 30 tableta	27	0	49	70	68	25	7	2	1890
124	L-Carnitien Liquid Orange, 31serv	25	57	0	3	1	53	8	8	1990
125	L-Carnitine Liquid Vanilla, 31 serb	0	0	0	0	0	0	8	8	1990
126	L-Carnitine Xtreme, 60ct	0	33	0	50	64	13	8	8	1590
127	Lion Kids C Vitamin Chewable 90 tableta	73	92	6	87	108	33	10	3	849
128	Lion Kids Punch with Vitamin D; 90 tableta	157	223	19	145	110	83	6	2	1029
129	Liquid C sa bioflavonoidima	34	10	0	27	22	193	1	2	2028
130	Liquid Chlorophyl, 0,5L	40	49	64	55	57	51	7	2	3295
131	Liquid L-Carnitine; 355 ml	15	41	0	3	27	5	8	5	2090
132	Liver Aid 100 kapsula	15	20	140	18	17	27	2	2	4009
133	L-Lysine Plus 60 caps	30	24	0	29	29	6	2	2	1847
134	Low Deuterium Oxy Crystal 50ml	54	55	82	39	53	48	2	2	5999
135	Lutein 60 kapsula	3	0	26	0	1	1	1	2	2879
136	Lutein Plus, 60 kapsula	37	50	112	57	33	48	1	2	2879
137	Magnezi B6 90 tab	37	72	29	48	42	52	6	2	1708
138	Meal Time 100 tableta	70	76	22	105	105	120	9	3	789
139	Mega B-Complex	42	27	180	51	45	38	10	3	3289
140	Mega Protect 4 Life, 90 tableta	74	50	38	81	75	60	1	3	3159
141	Mega Vitamin E	0	0	37	0	0	0	10	3	1189
142	Mega Zinc 50mg 100 tableta	82	104	0	72	71	74	5	3	969
143	Menopausal Formula 135 kapsula	49	33	0	31	39	28	2	3	3589
144	MSTCH Cell TECH PRO lemon lime 3kg	3	0	0	17	9	0	8	12	6890

145	MSTCH CELL TECH PRO orange 3kg	9	0	0	14	22	4	8	12	6890
146	MTCH Nitro-tech Ps Choco milkshake 1.8kg	0	17	0	29	36	15	8	12	7390
147	MTECH 100% PREMIUM WHEY PROTEIN CHOCOLATE 2,27k	28	88	0	20	37	12	8	12	6090
148	MTECH 100% PREMIUM WHEY PROTEIN VANILLA 2,27kg	1	96	0	6	9	6	8	12	6090
149	Mtech Alpha-Amino Prototype-216 120 tabs	4	0	0	2	3	0	8	12	4590
150	MTECH ANABOLIC HALO PRO SERIES ORANGE 907g	0	1	0	1	7	1	8	12	5790
151	MTECH BCAA HARDCORE 150 CAPSM	0	8	0	3	17	0	8	12	2790
152	MTECH Creakic	1	0	0	4	8	0	8	12	5990
153	Mtech Hydroxycut hardcore X120 liquid-caps	0	19	0	0	0	0	8	12	4990
154	MTECH Leukic	5	0	0	7	8	7	8	12	6490
155	MTECH MASS TECH NEW CHOCOLATE 2,27kg	0	7	0	11	15	7	8	12	5490
156	MTECH Nano Vapor HC Proseries 1Kg, Blueraspb	1	0	0	0	0	0	8	12	5690
157	Mtech Nano X9 Hardcore 180 caps	3	0	0	8	3	0	8	12	5690
158	MTECH Nitro-Tech HC Pro series Cookies cream, 4	0	1	0	0	0	0	8	12	7390
159	MTECH Nitro-Tech HC Pro series Straw banana, 4l	0	5	0	9	5	4	8	12	7390
160	Mtech Smart Protein Bar PB Caramel 63g 9pak	0	0	0	8	8	4	8	12	1900
161	MTECH SMART PROTEIN BAR TRIPLE CHOC 63g 9 BARS	0	0	0	10	6	2	8	12	1900
162	MUTANT MASS CHOCOLATE 2.2kg	6	34	12	9	29	34	8	6	3450
163	MUTANT MASS CHOCOLATE 6.8kg	2	17	58	4	9	11	8	6	8090
164	MUTANT MASS COOKIES & CREAM 2.2kg	1	10	70	4	15	12	8	6	3450
165	MUTANT MASS COOKIES & CREAM 6.8kg	0	6	32	2	5	4	8	6	8090
166	MUTANT MASS VANILLA 2.2kg	1	17	6	4	11	4	8	6	3450
167	MUTANT MASS VANILLA 6.8kg	1	3	60	0	1	0	8	6	8090
168	MUTANT WHEY TRIPLE CHOC. ERUPTION 2.27kg	20	52	221	0	113	21	8	6	5590
169	MUTANT WHEY VANILLA BEAN INFUSION 2.27kg	7	21	23	18	33	10	8	6	5590

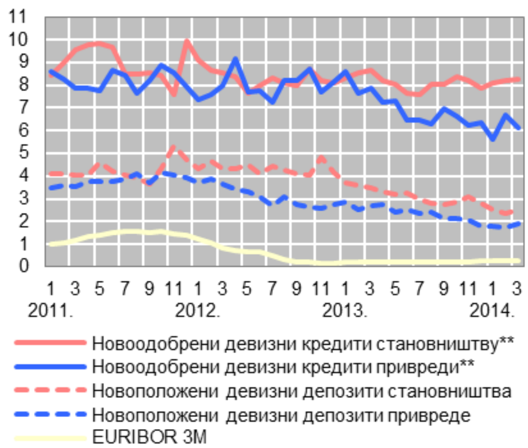
170	New Life, 120 komada	96	109	0	113	103	110	6	2	1690
171	N-Large 2; Čokolada; 1,7 Kg	0	5	0	10	3	5	8	5	3690
172	Noni Juice, 946 ml	232	60	0	206	249	208	7	2	5579
173	Noni; 90 kapsula	48	49	248	30	38	43	7	3	2909
174	Nopalin 200 tableta	110	111	0	122	114	95	7	2	2139
175	NUT ANABOL 5 120caps	0	0	0	15	7	5	8	13	5690
176	NUT Hemo Rage Black Berry 2 lbs	21	11	0	15	8	0	8	13	5990
177	NUT Hemo Rage Black Punch 2 lbs	0	10	0	19	18	0	8	13	5990
178	NUT LIPO 6 black 120 liquid caps	44	0	0	52	107	33	8	13	5690
179	NUT LIPO 6X 120caps	28	0	0	43	5	8	8	13	5690
180	NUT LIPO-6 120caps	34	5	0	6	2	28	8	13	4790
181	Ocean 21; 1L	62	44	9	49	57	42	7	2	4727
182	OH YEAH BAR milk choc. peanut, 12 bars	1	5	0	29	33	0	8	11	2390
183	OH YEAH BAR pb&straw, 12 bars	0	2	0	18	9	7	8	11	2390
184	OH YEAH BAR peanut butter, 12 bars	10	0	0	6	11	5	8	11	2390
185	OH YEAH BAR vanillia caramel 12 bars	16	0	0	0	0	0	8	11	2390
186	Omega 3 Natural Fish Oil 100 gel kapsula	222	0	16	19	0	1	7	2	1826
187	Oregano oil, 30ml	37	44	49	16	31	44	7	2	2179
188	OxyMax 60 ml	125	121	2	143	135	119	2	2	5499
189	Panax Ginseng 100 tableta	17	12	106	93	52	18	7	3	1579
190	Paraprotex 100 tableta	232	234	0	227	221	225	7	2	3285
191	Power mins 100 komada	90	84	23	67	114	80	5	3	1269
192	Pro Selenium 50 mcg 60 tableta	95	119	5	114	104	103	5	3	849
193	ProbioBalance, 60 tableta	118	117	0	80	107	128	9	2	2139
194	Pro-state Power, 60 tableta	48	53	0	50	37	38	2	2	3629
195	Protect 4 Life antioxydant with Gingko 90 tb4	80	68	0	108	109	90	1	3	1099
196	Pure Yucca 100 kapsula	48	9	751	69	65	20	7	3	1539
197	Resveratrol Plus, 60 kapsula	30	29	139	21	29	29	1	2	3519
198	Rhodiolin 60 kapsula	0	0	2	0	0	0	7	2	3079
199	RHODIOLIN; 120 tableta	336	324	145	312	348	311	7	2	3079
200	Samburex 237 ml	4	3	64	26	18	45	7	2	3989
201	Senior Formula; 90 tableta	39	35	68	25	34	35	6	3	1749
202	Shark-Aid 90 tableta	11	18	241	19	29	14	7	3	3398
203	Silk&Shine Shampoo; 250ml	13	53	0	52	36	40	3	2	659
204	SIX STAR PS AMINOMAX	0	0	0	8	10	7	8	14	2290

	ELITE SERIES TROP. FP 0.6									
205	SIX STAR PS CASEIN PROTEIN ELITE SERIES Choco 6	0	0	0	14	21	2	8	14	3190
206	SIX STAR PS CREATINE X3 ELITE SERIES FP 2.5 LB	105	69	0	94	129	92	8	14	2690
207	SIX STAR PS N.O. FURY ELITE SERIES FRUIT PUNCH	8	0	0	11	20	14	8	14	2990
208	SIX STAR PS WHEY ISOLATE ELITE SERIES Choco 680	0	16	0	6	20	3	8	14	3190
209	SIX STAR PS WHEY ISOLATE ELITE SERIES Vanila 68	0	6	0	4	9	2	8	14	3190
210	SIX STAR PS WHEY PROTEIN ELITE SERIES BANANA 2	5	6	0	12	12	29	8	14	3290
211	SIX STAR PS WHEY PROTEIN ELITE SERIES Choco 371	0	9	0	28	26	17	8	14	1890
212	SIX STAR PS WHEY PROTEIN ELITE SERIES CHOCOLATE	31	0	0	46	50	50	8	14	3290
213	SIX STAR PS WHEY PROTEIN ELITE SERIES COOKIES &	20	0	0	16	13	36	8	14	3290
214	SIX STAR PS WHEY PROTEIN ELITE SERIES VANILLA 2	16	6	0	17	25	24	8	14	3290
215	Slim Formula 90tabs	43	30	38	27	50	41	4	2	2739
216	Smokeraid 90 tableta	9	12	0	18	13	5	7	2	2299
217	Spirulina MAX, 60 tableta	9	13	24	10	14	14	7	3	1269
218	Spirulina-Chlorella Plus, 100 tableta	8	7	114	10	10	12	7	2	3157
219	Stres Management B-complex	117	106	0	102	107	95	10	3	1439
220	Strong Bones Calcium/Magnesium 100 kapsula	41	39	0	53	63	39	5	2	1799
221	Strong Bones Calcium/Magnesium 250 kapsula	38	34	0	35	37	36	5	3	2599
222	Strong Bones Plus, 100 kapsula	138	120	4	129	121	96	5	3	849
223	Super Amino Liquid - berry, 32oz	0	23	0	55	45	24	8	8	2590
224	Super Amino Liquid - grape, 32oz	0	0	0	59	25	14	8	8	2590
225	Super Amino Liquid - Orange, 32 oz	0	24	0	46	21	25	8	8	2590
226	Super CoQ10 plus 90	150	127	0	151	174	162	1	2	3886

227	Super Lecithin 100 gel kapsula	115	134	0	96	124	101	7	4	1339
228	Super Mega 50, 90 tableta	99	105	125	101	89	105	6	3	3779
229	Super Soya lecithin, 250 komada	49	52	0	37	50	35	7	4	3039
230	Super Vitamin Packs (30 kom)	77	0	1	20	7	143	10	11	2190
231	Tribulus FUEL; 100 kapsula	7	3	0	12	12	5	8	7	1890
232	Tribulus, 90ct	4	31	0	34	51	21	8	8	2190
233	Trimex WMS 16 OZ	2	6	2	6	5	3	7	2	3809
234	Triple Potency Lecithin 100 gel kapsula	91	73	10	89	106	82	7	3	2799
235	URX 90 kapsula	0	0	321	0	0	0	2	2	2859
236	UV Animal Flex 44 pack	31	31	0	42	39	38	8	15	4290
237	UV Animal Pack 44 pack	63	2	0	43	10	54	8	15	4690
238	UV Animal Stack2 21pack	21	9	0	7	20	11	8	15	5290
239	Vein Care 50+25	18	27	40	31	40	49	3	2	1529
240	Vein Protex, 60 tableta	35	16	133	23	23	23	2	2	3289
241	Virago; 90 tableta	29	48	1	33	30	39	7	2	3049
242	Vital 0 90 tableta	75	50	6	71	41	35	6	2	2139
243	Vital A 90 tableta	78	65	0	74	83	46	6	2	2139
244	Vital AB 90 tableta	17	9	3	14	14	5	6	2	2139
245	Vital B 90 tableta	26	25	0	27	31	21	6	2	2139
246	Vital Man 60 tableta	46	45	0	33	46	39	6	2	1489
247	Vital Woman 60 tableta	24	43	0	30	27	33	6	2	2028
248	Whey Amino 2000; 325 tableta	0	8	4	22	5	0	8	5	3590
249	White Willow 100	32	48	0	24	18	10	7	3	1269
250	XANTHO Plus, 90 kapsula	10	5	0	13	10	55	7	2	3037
251	XSHAPE, 90 kapsula	9	20	0	7	7	4	7	2	3037
252	Zen Tonic	19	16	101	43	21	22	2	2	4490

Прилог 2. Детерминанте инфлације-финансијско тржиште

Графикон IV.1.6. **Кретање каматних стопа на девизне кредите и депозите привреди и становништву***
(просечне пондерисане вредности, на годишњем нивоу, у %)



Извор: НБС и Европска банкарска федерација.

* Искључени револвинг кредити, прекорачења по текућем рачуну и дуг по кредитним картицама.

** У еврима и евроиндексирани.

….а на девизне изворе повећана током Т1.

		Новоодобрени девизни кредити становништву**	Новоодобрени девизни кредити привреди**	EURIBOR 3M	Новоположени девизни депозити становништва	Новоположени девизни депозити привреде
1 **2011**	1 **2011.**	8,45	8,61	1,02	4,11	3,48
2	2	9,02	8,25	1,09	4,10	3,60
3	3	9,54	7,85	1,18	4,04	3,52
4	4	9,77	7,86	1,32	4,04	3,75
5	5	9,85	7,75	1,43	4,60	3,76
6	6	9,65	8,68	1,49	4,20	3,78
7	7	8,50	8,41	1,60	4,02	3,86
8	8	8,49	7,65	1,55	3,94	4,10
9	9	8,55	8,18	1,54	3,65	3,71
10	10	8,42	8,91	1,58	4,32	4,18
11	11	7,60	8,57	1,49	5,34	4,02
12	12	9,96	7,91	1,43	4,74	3,92

1 2012	**1 2012.**	9,12	7,36	1,22	4,32	3,69
2	2	8,65	7,57	1,05	4,69	3,89
3	3	8,55	8,00	0,86	4,32	3,67
4	4	8,36	9,16	0,74	4,30	3,45
5	5	7,65	7,72	0,69	4,47	3,33
6	6	7,99	7,73	0,66	4,11	3,08
7	7	8,30	7,24	0,50	4,42	2,70
8	8	8,11	8,22	0,33	4,28	3,09
9	9	8,00	8,20	0,25	4,09	2,75
10	10	8,70	8,72	0,21	4,03	2,66
11	11	8,19	7,68	0,19	4,81	2,58
12	12	8,08	8,15	0,19	4,15	2,73
1 2013	**1 2013.**	8,33	8,60	0,21	3,69	2,84
2	2	8,53	7,67	0,22	3,58	2,55
3	3	8,67	7,87	0,21	3,51	2,71
4	4	8,23	7,25	0,21	3,31	2,76
5	5	8,05	7,29	0,20	3,19	2,42
6	6	7,62	6,48	0,21	3,27	2,50
7	7	7,58	6,47	0,22	2,96	2,38
8	8	8,03	6,27	0,23	2,79	2,39
9	9	8,06	6,98	0,22	2,77	2,13
10	10	8,35	6,62	0,23	2,84	2,11
11	11	8,21	6,24	0,22	3,11	2,10
12	12	7,88	6,33	0,27	2,82	1,79
1 2014	**1 2014.**	8,09	5,61	0,29	2,52	1,82
2	2	8,21	6,71	0,29	2,38	1,73
3	3	8,26	6,14	0,31	2,54	1,89

Прилог 3. Предвиђање кретања каматних стопа помоћу одабраних модела

	Univariate OLS	OLS VAR	Simple BVAR	Univariate BVAR
01.2013	2,401567292	2,412978575	2,403223031	2,401755833
02.2013	2,424697498	2,427420564	2,425958853	2,425046928
03.2013	2,452231830	2,452615075	2,452478518	2,452775292
04.2013	2,454695050	2,459659944	2,454489772	2,455257216
05.2013	2,454685892	2,444844516	2,451442895	2,455257195
06.2013	2,423834968	2,410073742	2,420076567	2,424194150
07.2013	2,394592916	2,384504049	2,391876900	2,394746046
08.2013	2,391091053	2,382398423	2,387765464	2,391222091
09.2013	2,391091191	2,383302075	2,387834546	2,391222092
10.2013	2,391091135	2,385225164	2,387808745	2,391222092
11.2013	2,371610148	2,358128467	2,367496839	2,371608346
12.2013	2,308789806	2,293941815	2,305341978	2,308350728
01.2014	2,274779008	2,259329525	2,269446767	2,274110742
02.2014	2,249733698	2,238877559	2,244312602	2,248892515
03.2014	2,249727660	2,226507250	2,242151537	2,248892501

Биографија аутора

Наташа Папић-Благојевић рођена је 4.фебруара 1977. године у Зрењанину, где је завршила основну школу „Доситеј Обрадовић“ и средњу Економско-трговинску школу „Јован Трајковић“ као ђак генерације.

На Економском факултету у Суботици, Универзитета у Новом Саду, дипломирала је у марту 2005. године са просечном оценом 8,10. Након завршетка основних студија уписала је последипломске студије на Економском факултету у Београду, Универзитета у Београду, на смеру *Статистичка анализа*. Магистарске студије је завршила са просечном оценом 9,90, а након одбране магистарског рада у јулу 2010. године стекла је звање *магистра статистичких наука*.

Наташа Папић-Благојевић је запослена на Високој пословној школи струковних студија у Новом Саду (некадашњој Вишој пословној школи) од маја 2005. године. У периоду од маја 2005. до октобра 2007. године обављала је послове стручног сарадника - приправника за економске наставне дисциплине, на предмету Пословна статистика. Од октобра 2007. до новембра 2010. године обављала је послове сарадника у настави за економске наставне дисциплине, на предметима Пословна статистика и Квантитативни методи у пословном одлучивању. У новембру 2010. године изабрана је у звање предавача за ужу област Информатика, инжењеринг и квантитативна анализа у економији за предмете Квантитативни методи у пословном одлучивању и Актуарство, на којима и данас изводи наставу.

Коаутор је уџбеника „Квантитативне методе“ који представља основну литературу за наставни предмет Квантитативни методи у пословном одлучивању на Високој пословној школи струковних студија у Новом Саду. Објавила је преко двадесет научних и стручних радова и члан је Статистичког друштва Војводине.

Као део међународне сарадње Високе пословне школе струковних студија у Новом Саду, која подразумева и размену наставника и сарадника, у мају 2013. године боравила је на

„Международной высшей школы управления“ у Санкт-Петербургу, Русија, а у јуну исте године на краковској академији „Andrzej Frycz Modrzewski“, Пољска.

Удата је и мајка двоје деце.

ИЗЈАВА О АУТОРСТВУ

Изјављујем да је докторска дисертација под насловом

Компаративна анализа класичне инференције и Бајесовог приступа у обради економских података

- резултат сопственог истраживачког рада,
- да предложена дисертација, ни у целини, ни у деловима, није била предложена за добијање било које дипломе, према студијским програмима других високошколских установа,
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права, нити злоупотребио/ла интелектуалну својину других лица.

У Нишу, ________________

Аутор дисертације: Наташа Папић-Благојевић

Потпис докторанда:

НПапић-Благојевић

ИЗЈАВА О ИСТОВЕТНОСТИ ШТАМПАНЕ И ЕЛЕКТРОНСКЕ ВЕРЗИЈЕ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Име и презиме аутора: _Наташа Папић-Благојевић_

Студијски програм: ____________________

Наслов рада: Компаративна анализа класичне инференције и Бајесовог приступа у обради економских података

Ментор: _проф. др Винко Лепојевић_

Изјављујем да је штампана верзија моје докторске дисертације истоветна електронској верзији, коју сам предао/ла за уношење у **Дигитални репозиторијум Универзитета у Нишу.**

Дозвољавам да се објаве моји лични подаци, који су у вези са добијањем академског звања доктора наука, као што су име и презиме, година и место рођења и датум одбране рада, и то у каталогу Библиотеке, Дигиталном репозиторијуму Универзитета у Нишу, као и у публикацијама Универзитета у Нишу.

У Нишу, ________________

Аутор дисертације: _Наташа Папић-Благојевић_

Потпис докторанда:

Н.Папић-Благојевић

ИЗЈАВА О КОРИШЋЕЊУ

Овлашћујем Универзитетску библиотеку „Никола Тесла“ да, у Дигитални репозиторијум Универзитета у Нишу, унесе моју докторску дисертацију, под насловом:

Компаративна анализа класичне инференције и Бајесовог приступа у обради економских података

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату, погодном за трајно архивирање.

Моју докторску дисертацију, унету у Дигитални репозиторијум Универзитета у Нишу, могу користити сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons), за коју сам се одлучио/ла.

1. Ауторство

2. Ауторство - некомерцијално

3. <u>Ауторство – некомерцијално – без прераде</u>

4. Ауторство – некомерцијално – делити под истим условима

5. Ауторство – без прераде

6. Ауторство – делити под истим условима

(Молимо да подвучете само једну од шест понуђених лиценци, кратак опис лиценци је у наставку текста).

У Нишу, ________________

Аутор дисертације: Наташа Папић-Благојевић

Потпис докторанда:

Н Папић-Благојевић