TRANSPORTA UN SAKARU INSTITŪTS

Jeļena AFANASJEVA

TRANSPORTA LOĢISTISKO PROCESU MODEĻU NOVĒRTĒŠANA BALSTOTIES UZ STATISTIKAS INTENSĪVAJĀM DATORA METODĒM

PROMOCIJAS DARBA KOPSAVILKUMS

RĪGA – 2006

TRANSPORTA UN SAKARU INSTITŪTS

Jeļena AFANASJEVA

TRANSPORTA LOĢISTISKO PROCESU MODEĻU NOVĒRTĒŠANA BALSTOTIES UZ STATISTIKAS INTENSĪVAJĀM DATORMETODĒM

PROMOCIJAS DARBA KOPSAVILKUMS
izvirzīts inženierzinātņu doktora zinātniskā grāda iegūšanai

Zinātņu nozare "Transports un satiksme"
apakšnozare "Telemātika un loģistika"

Zinātniskais vadītājs:
Dr.habil.sc.ing., profesors
Aleksandrs Andronovs

RĪGA – 2006

UDK 656:004
Af 055

Transporta un sakaru institūts

Afanasjeva J .

Af 055 Transporta loģistisko procesu modeļu novērtēšana balstoties uz statistikas intensīvajām datormetodēm. Promocijas darba kopsavilkums. – Rīga: Transporta un sakaru institūts, 2006. – 47 lpp.

ISBN 9984-9865-2-7

PROMOCIJAS DARBS

IZVIRZĪTS TRANSPORTA UN SAKARU INSTITŪTĀ

INŽENIERZINĀTŅU DOKTORA ZINĀTNISKĀ GRĀDA

IEGŪŠANAI

Promocijas darba aizstāvēšana notiks 2006. gada 4. jūlijā plkst. 15:00 Transporta un sakaru institūta promocijas padomē Lomonosova ielā 1 auditorijā 4-710, Rīgā, Latvijā, tālr. (+371) 7100661, fakss (+371) 7100660.

OFICIĀLIE RECENZENTI:

Dr.habil.sc.ing., profesors Jevgēnijs Kopitovs
Transporta un sakaru institūts

Dr.habil.sc.ing., profesors Jurijs Paramonovs
Rīgas Tehniskā Universitāte

Dr.habil.sc.ing., profesors Vjačeslavs Melass
Sanktpēterburgas Valsts Universitāte

APSTIPRINĀJUMS

Es apstiprinu, ka esmu izstrādājusi doto promocijas darbu, kurš iesniegts izskatīšanai Transporta un sakaru institūtā inženierzinātņu doktora zinātniskā grāda iegūšanai. Promocijas darbs nav iesniegts nevienā citā universitātē vai institūtā zinātniskā grāda iegūšanai.

2006. gada 15. maijā   Jeļena Afanasjeva

Promocijas darbs uzrakstīts angļu valodā, tas sastāv no sešām daļām, iekļauj 19 zīmējumus, 140 formulas un 26 tabulas, kopumā darbs aptver 135 lapas. Bibliogrāfijā ietverti 91 informācijas avoti.

ANOTĀCIJA

Ir izstrādāts promocijas darbs „Transporta loģistiko procesu modeļu novērtēšana, balstoties uz statistikas intensīvajām datormetodēm”, autore Jeļena Afanasjeva. Darba zinātniskais vadītājs Dr.habil.sc.ing., profesors Aleksandrs Andronovs.

Darbs ir veltīts mūsdienu statistikas metožu pielietošanai transporta loģistisko modeļu novērtēšanai. Galvena uzmanība ir pievērsta intensīvas datormetodes resamplinga izmantošanai. Tā ir neparametriskā metode, kas dod visefektīvākos sistēmu radītāju novērtējumus mazo ieejas izlašu gadījumā. Pētījumu veica trijos galvenos virzienos: 1) loģistisko modeļu novērtēšana un prognozēšana; 2) transporta līdzekļu drošības un efektivitātes novērtēšana; 3) loģistisko sistēmu krājumu pārvaldes procesu analīze.

Katrai no apskatītajām problēmām attiecīgā matemātiskā modeļa ietvaros, bija piedāvāti resamplinga procedūras pielietošanas algoritmi, izveidotas formulas resamplinga un klasiskās pieeju efektivitātes pētīšanai. Kā efektivitātes kritēriji tika izmantotas novērtējumu: nobīde, dispersija un vidēja kvadrātiskā kļūda. Iegūti skaitliskie rezultāti pierāda piedāvātās pieejas efektivitāti. Izdarīti secinājumi un dotas rekomendācijas par nosacījumiem, izpildot kuros piedāvātā pieeja ir visefektīvākā. Iegūtie rezultāti ir vispārīgi, jo tos var pielietot arī citām priekšmeta nozarēm.

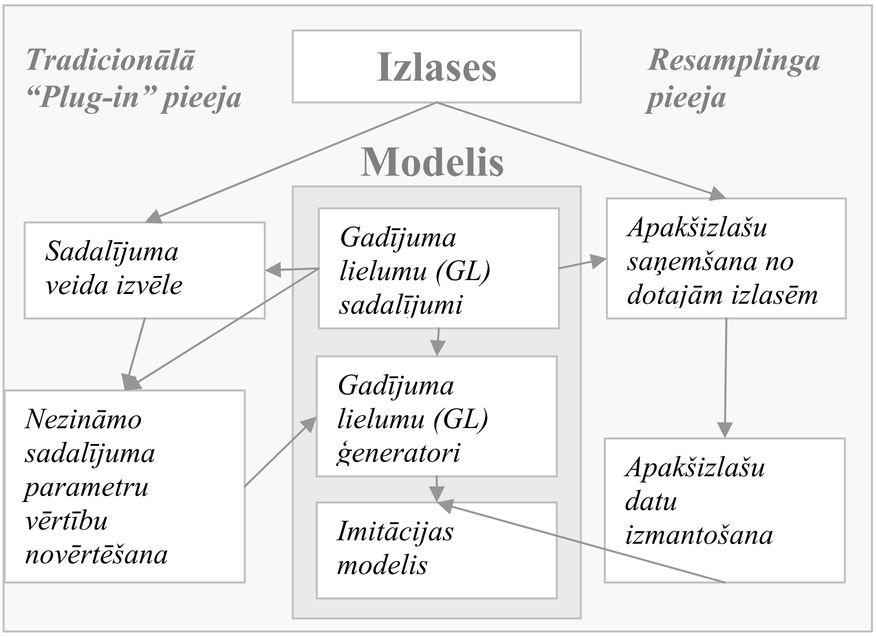
SATURS

1. Darba aktualitāte

Transports Latvijā ir strauji progresējoša nozare, kas atzīta par vienu no visprioritārākajām. Tai ir nepieciešami plaši un dārgi zinātniskie pētījumi, kas attiecas uz informācijas tehnoloģiju, matemātiskā aparāta, ekonomikas un loģistikas metožu piesaisti.

Šajā promocijas darbā vislielākā uzmanība tiek veltīta mūsdienu statistikas metožu pielietošanai transporta loģistikas procesu plānošanā un organizēšanā. Pētījumi norisinās trīs galvenajos virzienos: transporta plūsmas prognozēšana un novērtēšana, transporta līdzekļu drošības un efektivitātes novērtēšana, loģistikas sistēmu krājumu pārvalde.

Izmantojot varbūtiskos modeļus praktisko uzdevumu risināšanā, mēs bieži sastopamies ar tādu problēmu, ka nepietiek statistikas datu, uz kuru pamata jānovērtē nezināmie varbūtisko modeļu parametri. Tādā situācijā tradicionālo parametrisko metožu novērtēšanas izmantošana ir ļoti apgrūtināta. Tomēr intensīvās datormetodes palīdz risināt šo problēmu. Šīs skaitļošanas statistikas metodes parādījās salīdzinoši nesen. Tās paredz daudzreizēju vienu un to pašu datu izmantošanu dažādās kombinācijās, kas nodrošina pilnīgāku statistikas datu izmantošanu. Šīs pieejas svarīga īpašība ir tās neparametriskums. Neparametriskums ļauj izvairīties no kļūdām, kas raksturīgas tradicionālajām metodēm, kas saistītas ar hipotēžu pārbaudi par gadījuma lielumu sadalījumu. Galvenā uzmanība darbā pievērsta intensīvajai datormetodei *resamplings (resampling).* Tradicionāla un resamplinga pieeju salīdzinoša shēma imitācijas modelēšanas procedūrai ir ilustrēta 1. att. Šeit ir parādīta resamplinga neparametriskums, ka gadījuma lielumi nav ģenerēti ar speciāliem ģeneratoriem, bet ir ņemti tieši no sākuma izlasēm. Resamplings ļauj novērtēt sarežģīto sistēmu varbūtiskos raksturojumus, pamatojoties uz salīdzinoši maza apjoma statistikas datiem. Tā kā šī metode ir salīdzinoši jauna, tad kļūst aktuāla problēma par šīs pieejas pētījumu efektivitāti dažādos matemātiskajos modeļos un praktiskajos uzdevumos.



1. att. Tradicionālas un resamplinga pieeju salīdzināšanas shēma

2. Darba mērķis un uzdevumi

Pētījuma galvenais mērķis ir resampling metodes pielietošanas algoritmus dažādu matemātisko modeļu analīzei izstrādāšana, šīs pieejas pielietošanas efektivitātes novērtēšana, kā arī tās izmantošana praktisko uzdevumu risināšanā transporta loģistisko procesu modeļu novērtēšanas uzdevumos.

Darba galvenie uzdevumi ir šādi:

- Pētīt galvenās intensīvās datora metodes (IDM) statistikā un nozares, kur tās izmantot.
- Izstrādāt resamplinga metodes izmantošanas algoritmus regresijas modeļu parametru novērtēšanai.
- Pētīt resamplinga metodes efektivitāti regresijas modeļiem.
- Izskatīt resamplinga metodes izmantošanas iespēju transporta plūsmas prognozēšanai un novērtēšanai, pamatojoties uz regresijas modeļiem.

- Izmantojot resamplinga metodi, piedāvāt algoritmus rindošanas sistēmu raksturojumu novērtēšanai.
- Izstrādāt algoritmus izskatītas metodes efektivitātes novērtēšanai rindošanas procesu statistiskos uzdevumos.
- Izskatīt resamplinga metodes izmantošanas iespēju transporta sistēmās, lai novērtētu transporta līdzekļu drošību un efektivitāti, pamatojoties uz rindošanas procesu teoriju.
- Piedāvāt resamplinga metodes izmantošanas metodiku divu atjaunošanas procesu salīdzināšanai.
- Novērtēt piedāvātās pieejas efektivitāti, divu atjaunošanas procesu salīdzināšanas uzdevumam.
- Izmantot resamplinga pieeju krājumu pārvaldes uzdevumos, pamatojoties uz divu atjaunošanas procesu salīdzināšanas modeļa.
- Izstrādāt programmatūras kompleksu jaunas metodes īstenošanai un tās efektivitātes novērtēšanai.
- Izmantojot resamplinga metodoloģiju, veikt transporta nozares dažādu modeļu novērtēšanu un izmantot izstrādātus algoritmus novērtējumu efektivitātes aprēķināšanai.

3. Tēmas izstrādājuma pakāpe

Intensīvas datora metodes (IDM) vai skaitļošanas statistikas metodes sāka īpaši attīstīties, jaudīgai datortehnikai parādoties. IDM apvieno tādas metodes, kā krustveida pārbaude (cross-validation), saliktā naža (jackknife) metode, butrsrepa (bootstrap) metode, resamplinga metode. Šīs metodes ļauj risināt plašu uzdevumu klāstu. Saliktā naža metodi piedāvājis M. Kenujs (Quenouille) 1949. gadā, kā novērtējumu, kas ir kombinācija no novērtējuma, aprēķināta pēc visiem datiem un novērtējumiem, aprēķinātiem pēc datu daļas. 1979. gadā B. Efrons piedāvāja butrsrepa metodi, kas faktiski bija saliktā naža metodes vispārinājums. 1979. gadā V. Ivnickis piedāvāja izmantot resamplinga novērtējumus sistēmu drošībai ar imitācijas modeli. C. Vu 1979. gadā izskatīja resamplinga-pieejas pielietojumu regresijas modeļu parametru novērtēšanai. No 1995. gada resamplinga un butstrepa pieejas tiek pētītas profesora A. Andronova vadībā. Izskatītas vienkāršais un hierarhiskais resampling-pieejas, kā arī šo pieeju pielietošana drošuma rindošanas procesu un optimizācijas teorijās. Vēlāk, ar J. Merkurjeva un M. Fiošina piedalīšanos tika izskatītas resamplinga pieejas summām, daļēji zināmiem sadalījumiem un ticamības intervālu uzbūvēšanai. Ar J. Afanasjevas piedalīšanos veikti pētījumi resamplinga pieejas pielietojumam kārtas statistiku sadalījumu novērtējumam [4], regresijas modeļu parametru novērtēšanai [6], [5], atjaunošanas teorijā, masveida apkalpošanas teorijā [2], [9]. Daži pētījumi ir prezentēti šajā promocijas darbā. J. Afanasjeva apskatīja arī resamplinga metodes pielietošanu atjaunošanas procesu salīdzināšanas uzdevumos [2], šīs metodes izmantošanu krājumu pārvaldes teorijas uzdevumiem [8], kas veido veselu promocijas darba daļu. Pašlaik pētniecības darbi šajā nozarē turpinās.

4. Pētījumu metodoloģija un uzdevumi

Promocijas darbā par teorētisko un metodisko pamatu ņemtas matemātiskās un skaitļošanas statistikas nozares, kā arī transporta, loģistikas, modelēšanas un datorzinātņu klasiskie darbi. Ņemti vērā zinātnes attīstības virzieni šajās nozarēs.

Darbā izmantotas grāmatas minētajās nozarēs, periodisko izdevumu tematiskie materiāli, starptautisko konferenču materiāli un statistikas krājumi.

Visi secinājumi balstās uz gadījuma procesu, varbūtību teorijas un matemātiskās statistikas teorijas klasisko aparātu izmantošanu.

Pētījuma procesā tika analizēti dažādi transporta loģistikas nozares piemēri, kuriem tika ilustrēta piedāvātās metodikas izmantošana. Piemēros izmantoti gan reāli, gan hipotētiski dati, kas pilnībā atspoguļo izmantojamās pieejas efektivitāti un praktisko uzdevumu specifikāciju. Kā izmantojamās metodes efektivitātes kritērijus izvēlējās novērtējuma nobīdi, dispersiju un vidējo kvadrātisko kļūdu. Tika analizētas efektivitātes mainīšanas tendences atkarībā no dažādiem faktoriem, kas ļauj lemt par iespēju izmantot metodi praktisko uzdevumu risināšanā.

Lai risinātu dotos uzdevumus, darbā izmantotas gan analītiskas, gan eksperimentālas metodes. Ar analītiskām metodēm iegūtas analītiskas izteiksmes metodes, lai analizētu efektivitāti dažādās situācijās. Ar eksperimentālām metodēm aprēķināti izmantojamās metodikas efektivitātes kritēriji konkrētiem skaitliskajiem piemēriem, kas ļauj izdarīt secinājumus par dažādu faktoru ietekmi uz rezultātiem.

Piedāvātās metodikas izmantošanas rezultāti tiek salīdzināti ar klasisko pieeju izmantošanas rezultātiem. Tiek izdarīti secinājumi par nepieciešamajiem noteikumiem, pie kuriem resamplinga pieeja ir efektīvāka.

5. Darba zinātniskā novitāte

Var runāt par diviem darba zinātniskās novitātes aspektiem. Pirmkārt, no matemātiskā viedokļa. Tā ir resampling metodes izmantošana dažādos statistiskos uzdevumos un tās izmantošanas efektivitātes analīze.

Otrkārt, no praktiskā viedokļa. Resamplinga metodes izmantošana transporta sistēmās līdz šim šādā nostādnē vēl nav analizēta.

Šajā promocijas darbā statistikas metožu izmantošana transporta loģistikas procesu plānošanā un organizēšanā ir analizēta trīs galvenajos virzienos: 1) transporta plūsmas prognozēšana un novērtēšana; 2) transporta līdzekļu drošības un efektivitātes novērtēšana; 3) krājumu pārvalde loģistiskas sistēmās.

Pirmajā uzdevumā, kas saistīts ar transporta plūsmas prognozēšanu, risināšanai par matemātisko pamatu ņemts regresijas modelis. Šajā pētījuma daļā resamplinga metode tika izmantota, lai novērtētu regresijas modeļa parametrus. Efektivitātes analīzes izmantojamie algoritmi un aprēķina algoritmi tiek uzskatīti par jauniem zinātniskiem atklājumiem.

Pētījuma otrajā daļā, kas saistīta ar transporta līdzekļu drošības un efektivitātes novērtēšanu, par matemātisko modeli ņemta rindošanas sistēma ar bezgalīgu apkalpošanas aparātu skaitu. Tika izskatīta situācijas resamplinga metodes izmantošana, izveidotas efektivitātes analīzes formulas. Tā ir pētījuma otrās daļas zinātniskā novitāte.

Resamplinga metode pētījuma pēdējā daļā tika izmantota, lai risinātu krājumu pārvaldes teoriju uzdevumus. Par analīzes matemātisko aparātu kļuvuši atjaunošanas procesi. Tika risināts deficīta varbūtības novērtējuma uzdevums konkrētā krājumu pārvaldes uzdevumā, kuru matemātiski varēja izskaidrot kā divu atjaunošanas procesu salīdzināšanas uzdevumu. Līdzās šī gadījuma algoritmu izmantošanai, resamplinga metodē tika iegūtas arī izteiksmes, lai novērtētu šīs pieejas efektivitāti. Tie veido pēdējās daļas zinātnisko novitāti.

6. Darba praktiskā vērtība, realizācija un praktiskais pielietojums

Praktiskā vērtība ir piedāvāto pieeju izmantošana transporta loģistikas praktisko uzdevumu risināšanā. Pirmais apskatītais uzdevums ir transporta procesu, kā dažu makroekonomisko rādītāju modeļu prognozēšana un novērtēšana. Šajā nodaļā ilustrācijas dēļ tiek risināti šādi uzdevumi:

- Pasažieru plūsmas pieprasījuma prognozēšana aviācijas transportam Eiropas Savienības valstīs uz noteiktu gadu, atkarībā no dažādiem faktoriem: valsts teritorija, iedzīvotāju skaits, viena iedzīvotāja vidējā mēnešalga, iekšzemes kopprodukts uz vienu iedzīvotāju.
- Transporta pakalpojumu pieprasījuma prognozēšana dažādos reģionos, atkarībā no urbanizācijas, izglītības līmeņa un algas līmeņiem.

Izskatāmā uzdevumu otra grupa saistīta ar transporta līdzekļu efektivitātes un drošības novērtēšanu. Šajā nodaļā tiek sniegts šāds praktiskā uzdevuma risinājuma piemērs:

- Pastāv aviokonstrukciju bojājumu plūsma. Ir statistika par bojājumu rašanos līdz bīstamas situācijas attīstībai. Nepieciešams novērtēt varbūtību, ka izskatāmajā laika intervālā nenotiks atteikums.

Trešā šīs metodikas izmantošanas nozare ir krājumu pārvaldes teorija. Pieeja tiek ilustrēta ar šāda uzdevuma piemēru:

- Ir aplūkots kādas preces izmantošanas process. Ir zināmas kāds ir šo preču sākotnējās krājums, kā arī statistika par katras preces vienības intervāliem starp piegādi un izeju no ierindas. Nepieciešams novērtēt varbūtību, ka dotais pieprasītais izstrādājums neizraisīs deficītu, kā arī atrast optimālu krājuma līmeni pie dotajiem izmantošanas pieauguma parametriem, glabāšanas izdevumiem un zaudējumiem deficīta dēļ.

Iegūtie skaitļi ļauj spriest par apskatāmās pieejas efektivitāti izklāstīto uzdevumu risināšanā. Rezultātu analīze ļauj izdarīt secinājumus par iespējamību lietderīgi izmantot resamplinga metodi analoģiskos transporta loģistikas uzdevumos. Iegūti rezultāti ir vispārīgi, tāpēc tos var pielietot arī citām priekšmeta nozarēm.

7. Publikācijas

Darba rezultāti tika apskatīti desmit zinātnisko žurnālu publikācijās, starptautisko konferenču krājumos un zinātnisko darbu krājumos. Publikācijās apskatītām problēmām bija referāti starptautiskās zinātniskajās konferencēs Latvijā, Polijā, Izraēlā un Francijā.

8. Darba struktūra

Pirmajā nodaļā apskatīta statistikas metožu izmantošana transporta loģistikas procesu plānošanā un organizēšanā. Otrajā nodaļā apskatītas mūsdienu statistikas intensīvās datora metodes (IDM). Trešajā nodaļā ir izpildīts resamplinga metodes pielietošanas apskats dažādiem matemātiskiem modeļiem. Katra nākamā nodaļa apskata resamplinga metodes izmantošanu konkrēta uzdevuma risināšanā. Tie ir prognozēšanas uzdevumi, kas balstās uz transporta procesu regresijas modeļiem, transporta līdzekļu drošības un efektivitātes novērtēšanu un izmantošanas efektivitāti, un krājumu pārvaldi. Katrā nodaļā ir saturīgs uzdevumu izklāsts, matemātiskais modelis, tradicionālo pieeju analīze dotā uzdevuma risināšanai un piedāvātās pieejas motivācija. Tālāk tiek izklāstīts piedāvātās pieejas saturs ar algoritmiem un izteiksmēm efektivitātes aprēķināšanai. Katra nodaļa beidzas ar skaitlisku pieejas efektivitātes analīzi, kas balstās uz hipotētiskiem datiem, kuri pilnībā atspoguļo apskatāmā uzdevuma specifiku un īpatnības. Tālāk seko izmantošanas piemērs no transporta nozares, kas satur uzdevuma nostādni, piedāvātās pieejas izmantošanas algoritmu un skaitliskos rezultātus. Katras nodaļas beigās ir izdarīti secinājumi, atbilstoši piedāvātās pieejas lietderīgai izmantošanai apskatītajā situācijā. Veikts salīdzinājums ar tradicionālo metožu izmantošanas rezultātiem, sniegtas rekomendācijas, pie kādiem noteikumiem piedāvātās pieejas izmantošana sniedz labākus rezultātus. Darba nobeigums satur secinājumus un literatūras sarakstu.

9. Pētījuma galveno rezultātu apraksts

9.1. Statistiskas metodes transporta loģistisko procesu plānošanā un organizēšanā

Matemātisko metožu un informācijas tehnoloģiju pielietošanas nepieciešamība bija izsaukta ar pārvadājumu apjomu strauju palielināšanos, ar tehnoloģisko procesu sarežģītības pieaugumu, ar saistību nepārtrauktu pastiprināšanu starp dažādām transporta sistēmas daļām. Ir nepieciešams izstrādāt tādu matemātisko aparātu, kas ļaus analizēt transporta sistēmu darbības efektivitāti, pārvadājumu procesu modelēšanu, prognozēt pārvadājumu pieprasījumu, izstrādāt transporta sistēmas optimālus attīstības variantus. Loģistika savukārt ir viens no vadošajiem virzieniem Eiropas Savienības transporta sektora mūsdienu zinātniskajā un tehnoloģiskajā attīstībā. Tā ir ekonomikas sfēra, kas vel joprojām attīstās, kā arī un jaunais zinātniskais virziens.

Dotajā darbā tiek piedāvāts mūsdienu statistikas metožu pielietojums dažu uzdevumu risināšanai transporta loģistisko procesu plānošanai un organizēšanai. Šīs problēmas aktualitāte ir saistīta ar to, ka pārvaldīt sarežģītas ekonomikas sistēmas ir neiespējami bez uzticamas statistiskas informācijas par pētāmiem objektiem.

Analizējot attiecīgos modeļus transporta jomā, var atrisināt sekojoša veida problēmas: efektīvāko maršrutu plānošana labākai transporta un pasažieru plūsmas organizēšanai, statistikas analīze par ceļu negadījumiem, lai atrastu vissnozīmīgos to ietekmējošus faktorus; simulācijas programmas izstrādāšana, ceļu virsmas stāvokļa novērtēšanai, kā arī attiecīga remonta prognozēšana, transportēšanas pieprasījuma prognozēšana dažādiem transporta veidiem, pasažieru plūsmu maršrutizācija, ceļu tīkla vai aviolīniju tīkla izveide, transporta līdzekļu parka krājumu plānošana utt. Šeit pieļautās kļūdas var izraisīt neefektīvu ražošanas fondu izmantošanu un pieprasījumu apmierinājumu nepilnā apjomā.

Bieži vien datu iegūšana ir sarežģīts uzdevums. Aprakstītā situācija ir ļoti raksturīga dažu transporta veidu drošuma uzdevumiem. Sakarā ar paaugstinātu drošumu aviācijas tehnikā ir ļoti maza atteikumu statistika.

Dotajā darbā ir apskatītas transporta plūsmu prognozēšanas un novērtēšanas problēmas, transporta līdzekļu drošības un efektivitātes novērtēšana, transporta loģistisko sistēmu krājumu pārvalde. Tiek pielietoti sekojoši matemātiskie modeļi un statistiskās metodes: regresijas modeļi, nezināmo parametru novērtēšanas metodes, statistisko hipotēžu pārbaude, atjaunošanas teorija, krājumu pārvaldes teorija, gadījuma procesu teorija, masveida apkalpošanas teorija. Tiek pielietotas resamplinga metodes nezināmo sistēmu parametru novērtēšanai, kas tiek salīdzinātas ar tradicionālām ieejām.

9.2. Mūsdienu statistikas intensīvas datoru metodes

Klasisko (tradicionālo) statistisko metožu attīstība bija apturēta ar iespējamu rēķināšanas apjomu, kuru agrāk bija iespējams izpildīt tikai ar rokām. Tāpēc klasiskie pamatrezultāti bija saistīti ar asimptotiskām pieejam, lielajiem izlašu apjomiem. Ļoti maziem izlašu apjomiem bija iespējams izmantot parasto pārlasi. Ir palicis atklāts jautājums, kad izlases apjoms ir vidējs. Pirmās izmaiņas notika ar datoru ieviešanu. Tad šī problēma gandrīz vairs nepastāvēja. Tai sekoja skaitļošanas statistikas metožu vai intensīvo datormetožu (IDM) attīstība.

Tradicionālās statistikas metodes un tradicionālie statistikas modeļi ir bieži balstīti uz daudziem pieņēmumiem (lineārām atkarībām, sadalījumu klasēm, nemainīgumu utt.). No vienas puses, šādi pieņēmumi vienkāršo modeli un padara iespējamu analītisko metožu pielietošanu. Bet, no otras puses, praktiskos uzdevumos šādi pieņēmumi var novest pie nepilnīga un nepareiza modeļa. Izmantojot šādas metodes, iegūtie rezultāti ir neprecīzi.

Ir grūti izmantot tradicionālās metodes sarežģītu sistēmu, nestacionāru sistēmu, gadījuma, kad sadalījuma klases atšķiras no klasiskajām, analīzei. Šādu sistēmu analītisko modeļu izstrāde ir sarežģīta un pat bieži neiespējama, tāpēc šajos gadījumos ir labāk izmantot IDM.

Jāņem vērā, ka IDM nav tik precīzas kā klasiskās metodes. Tā ir patiesība tikai tādā gadījumā, ja tiek izmantoti klasisko metožu pieņēmumi. Gadījumā, ja klasisko metožu pieņēmumi netiek realizēti, IDM ir daudz efektīvākas. Tāpēc, var teikt, ka IDM var risināt daudzas problēmas, kuras agrāk nevarēja pienācīgi atrisināt.

Nedrīkst aizmirst potenciālo risku, izmantojot IDM. Ir jāsaprot, ka lielais aprēķinu skaits negarantē, ka informācija tiek lietota pareizi un ka rezultāti būs labāki. Daudzos gadījumos klasisko metodi var izmantot šādā uzdevumā, un tā dod precīzāku rezultātu nekā mēs varam iegūt pēc veselas dienas aprēķiniem, izmantojot IDM.

Datoru statistikas nozare šodien ietver sevī lielu skaitu metožu. Šīs metodes ļauj apskatīt datus no dažādām perspektīvām un analizēt dažādas datu apakškopas. Tā kā ievaddatu dažādo kombināciju skaits var būt ļoti liels, var pieprasīt lielu aprēķinu skaitu.

Montekarlo (Monte-Carlo) methods, pirmās metodes, kas pieprasa lielu aprēķinu skaita realizēšanu, kas parādījušies jau 60-tajos gados. Šīs metodes prezentē nosacījumu randomizācijas izmēģināšanu pie principā neliela pārtēriņa, turklāt paši izmēģinājumi var būt ne tikai fiziski eksperimenti, bet arī aprēķini uz datora. Metodes Montekarlo izmanto, lai pārbaudītu noteiktu statistikas hipotēzi, nosakot kādas statistikas nozīmīguma līmeni, kas domāta datu savākšanai, kā arī, lai novērtētu sistēmas nezināmos parametrus.

Randomizācijas (Randomization) metodes is Krustveida pārbaude (*Cross-validation)* un *saliktā naža(Jackknife) metodes.*

Pārkārtošanās, apakškopu veidošana vai citas manipulācijas ar modeļiem nevar mums sniegt papildus informāciju, taču, pārkārtojot datu kopu, varam iegūt informāciju, piemēram, cik nejauša ir datu kopa salīdzinājumā ar hipotēzi. Tā ir randomizācijas jēga. *Krustveida pārbaude* ir process, kurā kopa tiek iedalīta „aprēķina kopā" un „pārbaudes kopā". Aprēķina kopu lieto, lai aprēķinātu interesējošos parametrus, bet pārbaudes kopu lieto, lai analizētu modeļa kvalitāti. *Saliktā naža* metodi 1949. gadā ieviesa M. Kenujs (M. Quenouille). Domāts bija pakāpeniski izslēgt no pētīšanas pa vienam novērojumam, apstrādāt visu atlikušo informāciju un paredzēt rezultātu ar izslēgšanas metodi. Iegūtais kopskaits šādā veidā pa visiem punktiem, kas nesakrīt, nes sevī informāciju par pārvietošanu, kuru var izmantot. Dž. Tjukijs (J. Tukey), kas ir aktīvs šīs metodes realizētājs, nosaucis to par saliekamā naža metodi.

Butstrapa (Bootstrap) metodi ieviesa B. Efrons (B. Efron) 1977. gadā, un to paplašināja citi autori. Metodes būtība ir tieša esošo statistikas datu izmantošana modelēšanas procesā. Tā ļauj vairākkārt apstrādāt vienu datu dažādas daļas. Butstraps – neparametriskā metode, kas sākotnēji parādījās, lai novērstu nobīdi, kas saistīta ar izlases mazo apjomu. Lai risinātu šo problēmu, jāizvēlas tādi modeļi, kuru rezultāti ir maz atkarīgi no patiesās situācijas. Tas ir neparametriskās statistikas paņēmiens, kurš sekmējis tādu pieeju attīstību kā *neparametriskās un robustuma metodes*. Pēc tam noskaidrojās, ka to var izmantot, lai novērtētu izlases dispersiju, konfidencialitātes intervālu izveidi un hipotēžu pārbaudi.

Trace-Driven simulation metodi izmanto modeļa pārbaudei jeb validācijai, salīdzinot ar reālu modeli. Trace-Driven simulation ir aplūkots laikā 1998.-2001. gadam. J.Kleijnena (J.Kleijnen) darbos. Šis autors pielietoja šo metodi transporta modeļa pārbaudei jeb validācijai.

9.3. Resamplinga metode statistikā un tās attīstīšana

Ļoti grūti izveidot robežu starp dažādām IDM. Dažādos literatūras avotos šo metodi apvieno ar nosaukumu resamplings.

Resamplings ir IDM, kas var būt izmantots statistikā un modelēšanā. Tā ir alternatīva metode imitācijas sistēmu modelēšana, kuru efektīvi var pielietot mazu izlašu apjoma gadījumos un, kad tikai viena izlase ir pieejama dažiem mainīgiem. Šī metode sistēmu modelēšanai bija agrāk aprakstīta A. Andronova rakstos. Bija parādīts, ka resamplinga metode ļauj izvairīties no novērtējumu novirzes un samazināt to dispersiju.

Resamplinga metodes pielietošanas pamatideja ir balstīta uz faktu, ka imitācijas modelēšanas procesa gadījuma lielumu vērtības nav izstrādātas ar speciāliem ģeneratoriem, bet tiek uzreiz ņemtas no dotām izlašu populācijām. Var gadīties, ka sākuma izlases ir nevienāda vai pārāk maza apjoma, lai īstenotu kvalitatīvu modelēšanu. Minēto grūtību var pārvarēt ar resampling-pieejas pielietošanu, kas paredz vienam gadījuma lielumam vienu un to pašu datu daudzkārtēju izmantošanu dažādās kombinācijās ar citu gadījuma lielumu datiem.

Tiek apskatīta zināma funkcija ϕ no vairākiem neatkarīgiem savā starpā gadījuma lielumiem $X_1, X_2, ..., X_m$: $\phi(X_1, X_2, ..., X_m)$. Ir pieņemts, ka gadījuma lieluma X_i sadalījuma funkcija $F_i(.)$ ir nezināma, bet katram gadījuma lielumam ir zināma attiecīga izlases populācija:

$$H_i = \{X_{i1}, X_{i2}, \dots, X_{in_i}\}, \ |H_i| = n_i, \ i=1,..,m. \qquad (1.)$$

Mums uzdevums ir atrast matemātiskas cerības novērtējumu:

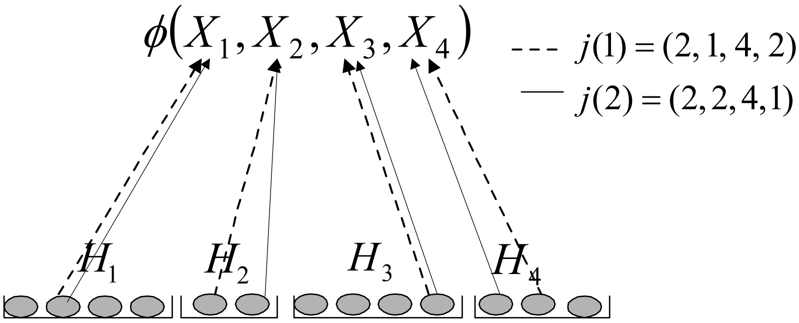
$$\Theta = E(\phi(X_1, X_2, \dots, X_m)). \qquad (2.)$$

Izmantojot šo pieeju, funkcijas ϕ argumenti X_i tiek paņemti gadījuma veidā no attiecīgām izlasēm $\{H_i\}$. Citiem vārdiem $j(l) = (j_1(l), \dots, j_m(l))$, $l = 1,2,\dots$, ir gadījuma izlases no $\{H_i\}$. Parasti runā, ka $j(l)$ un $X(l) = (X_{1j_2(l)}, X_{2j_2(l)}, \dots, X_{mj_m(l)})$ - tas ir l-tā resamplinga izlase. Tad resamplinga novērtējums parametram Θ ir:

$$\Theta^* = \frac{1}{r}\sum_{l=1}^{r}\phi(X(l)). \qquad (3.)$$

Parasti skaitlis r – resamplinga procedūras atkārtojumu skaits ir ievērojami mazāks, nekā visas iespējamās sākumdatu kombinācijas.

Resamplinga metodes izmantošana ir ilustrēta 2. att. Šeit var redzēt, ka vektors $j(l)$ satur to elementu izlases indeksus, kas bija izvelētas l-tā realizācijā. Kad l=1, j(1)=(2,1,4,2), tas nozīmē, ka no pirmās izlases $\{H_1\}$ bija otrais elements, no otrās izlases $\{H_2\}$ bija pirmais elements, no trešās izlases $\{H_3\}$ ceturtais elements un no ceturtās izlases $\{H_4\}$ otrais elements.



2. att. Resamplinga procedūras piemērs

Elementu izvēle resamplinga izlasē kādā realizācijā ir konservatīva. Mēs pieņemam, ka visām izlasēm ir viens un tas pats marģinālais sadalījums $\{F_i\}$, kas ir iegūstams no ieejas sadalījuma *X*(*l*), *l*=1,2,.... Izlasēm $(X(l), X(l'))$ ir vienāds kopīgs sadalījums $l \neq l'$. Acīmredzams, ka tādam konservatīvam plānam mūs interesējošais novērtējums ir nenobīdīts: $E\Theta^* = \Theta$.

Konservatīvais plāns paredz daudzveido resamplinga eksistenci: parastais resamplings, kontrolēts un hierarhiskais resamplings, stohastisko procesu resamplings, resamplings regresijas modeļa parametru novērtēšanai. Neviena no šīm metodēm neizmanto visu iespējamo kombināciju pilnu pārlasi.

Apskatīsim dažādas resamplinga procedūras, kur efektivitātes kritērijs ir dispersija.

Ievadīsim sekojošus momentu apzīmējumus:

$$\mu_\nu = E(\phi(X(l))^\nu), \quad \nu = 1,2,\ldots \tag{4.}$$

$$\mu_{11} = E(\phi(X(l)), \phi(X(l'))), \quad l \neq l'. \tag{5.}$$

Tad resamplinga novērtējuma dispersija ir:

$$Var(\Theta^*) = E(\Theta^{*2}) - \mu^2, \;\; E(\Theta^{*2}) = \frac{1}{r}\left(\mu_2 + (r-1)\mu_{11}\right). \tag{6.}$$

Atzīmēsim, ka jaukta momenta vērtība μ_{11} ir atkarīga no elementu izvēles shēmas resamplinga izlasē (ar atkārtojumiem, bez atkārtojuma)

9.4. Transporta procesu regresijas modeļi

Viena no svarīgākajām ilgtermiņa plāna izstrādes sākumprocedūrām ir prognozēšana. Prognozes ļauj noteikt, kādus procesus ir lietderīgi attīstīt paātrināti, ļauj sagatavot to attīstības alternatīvas un tos objektīvi salīdzināt.

Prognozēšanas modeļu matemātiskais aparāts ir laika rindas un regresijas teorijas statistikas metodes analīzes.

Aplūkosim gadījumu, kad prognoze ir veikta ar regresijas modeļa palīdzību un aprakstīsim resamplinga metodi, lai novērtētu regresijas parametrus [6].

Lineāras regresijas modelis ir viens no populārākajiem statistikas modeļiem. Tam ir sekojoša forma:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}, \tag{7.}$$

kur $\mathbf{X}$ - $n \times m$ izmēra matrice no neatkarīgiem mainīgiem, n - novērojumu skaits, m - neatkarīgo mainīgu skaits, $\mathbf{Y}$ - $n \times 1$ izmēra vektors no atkarīgiem mainīgiem, $\mathbf{Z}$ - $n \times 1$ izmēra vektors, kam komponenti Z_1, Z_2, ... , Z_n - neatkarīgie vienādi sadalīti gadījuma lielumi ar vidējo viens un dispersiju σ^2, tad $Cov(\mathbf{Z}) = \sigma^2\mathbf{I}_n$, un $\boldsymbol{\beta}$ - $m \times 1$ izmēra vektors no regresijas parametriem, kurus jānovērtē.

Klasiskā (tradicionālā) pieeja

Klasiska pieeja parametra $\boldsymbol{\beta}$ mazāko kvadrātu novērtējumam (MKN) ir labi zināma:

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Y}. \tag{8.}$$

Matemātiskā statistikā ir pierādīts, ka lineāras regresijas ietvaros vislabākais, jeb efektīvākais novērtējums ir mazāko kvadrātu novērtējums.

Novērtējums (8.) palīdz prognozēt atkarīga mainīga Y_d vērtību izvēlētājam novērojumam *d,* prezentētam ar vektoru $\mathbf{x}_d$ izmēra $1\times m$ no neatkarīgu mainīgo vērtībām, izmantojot sekojošu formulu:

$$\hat{Y}_d = \mathbf{x}_d \cdot \hat{\boldsymbol{\beta}}. \tag{9.}$$

Novērtējumu (8.) var sekmīgi pielietot, ja izlasē nav novērojumu kas traucē. Pretējā gadījumā saņemta novērtējuma kvalitāte nav tik laba, tās varētu būt nobīdīts. Lai uzlabotu iegūtā novērtējuma kvalitāti piedāvāts izmantot resamplinga pieeju, klasiskas pieejas vietā. Bija pētīta piedāvātas metodes efektivitāti, ņemot novērtējuma nobīdi, kā efektivitātes kritēriju.

Resamplinga pieeja

Resamplinga pieeja darbojās sekojoši. Lai $\mathbf{X} = \left(\mathbf{x}_1^T\ \mathbf{x}_2^T \dots \mathbf{x}_n^T\right)^T$, kur $\mathbf{x}_i$ ir i-tā matrices **X** rinda, kas atbilst i-tajam novērojumam, k - veselais skaitlis $m \leq k < n$, un $N = \{1, 2, \dots, n\}$ - veselu skaitļu kopa 1, 2, ... , n. Mēs izpildām ciklu ar r soļiem. Tekošā solī (piemēram. l-tajā) veidojam resamplinga izlasi, izvelējoties (bez atkārtošanās) k skaitļu no N: $J(l,\ 1)$, $J(l,\ 2)$, ... , $J(l,\ k)$. Tad vektors $\mathbf{J}(l) = (J(l,\ 1)\ J(l,\ 2)\ \dots\ J(l,\ k))$ definē novērojumu numurus (matricu **X** un **Y** rindas), kas bija izvēlētās l-tajā resamplinga izlasē. Izveidojam l-to resamplinga izlasi:

$$\mathbf{X}(\mathbf{J}(l)) = \left(\mathbf{x}_{J(l,1)}^T\ \mathbf{x}_{J(l,2)}^T \dots \mathbf{x}_{J(l,k)}^T\right)^T,$$

$$\mathbf{Y}(\mathbf{J}(l)) = \left(Y_{J(l,1)}\ Y_{J(l,2)} \dots Y_{J(l,k)}\right)^T.$$

Aprēķinām modeļa parametra $\boldsymbol{\beta}$ novērtējumu $\boldsymbol{\beta}^*$ pēc formulas (8.):

$$\boldsymbol{\beta}^*(\mathbf{J}(l)) = \left(\mathbf{X}(\mathbf{J}(l))^T \mathbf{X}(\mathbf{J}(l))\right)^{-1} \mathbf{X}(\mathbf{J}(l))^T \mathbf{Y}(\mathbf{J}(l)). \tag{10.}$$

Tad, pēc r atkārtojamiem iegūstam secība no novērtējumiem:

$$\boldsymbol{\beta}^*(\mathbf{J}(1)), \boldsymbol{\beta}^*(\mathbf{J}(2)), \dots, \boldsymbol{\beta}^*(\mathbf{J}(r)). \tag{11.}$$

Tad katras komponentes aritmētiskais vidējais no šīs secības elementiem dod mums resamplinga parametra $\boldsymbol{\beta}$ novērtējumu:

$$\boldsymbol{\beta}^*(\mathbf{J}) = \frac{1}{r}\sum_{l=1}^{r} \boldsymbol{\beta}^*(\mathbf{J}(l)). \tag{12.}$$

Šajā nodaļā ir pētīts klasiskās un resamplinga novērtējumu robustums gadījumam, kad modelis ir ar traucējumiem, ņemot novērtējuma nobīdi, kā efektivitātes kritēriju.

Modeļa ar traucējumiem definīcija

Apskatīsim gadījumu, kad visu novērojumu skaitā ir novērojumi ar traucējumiem. Apzīmēsim "*patiesus*" ("*true*") novērojumus ar indeksu t un ar traucējumiem *"aplama"* (*"false"*) ar indeksu f. Bez vispārības zuduma, varam pieņemt, ka patiesie novērojumi atbilst pirmām n - h **X**, **Y** un **Z** rindām:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_t \\ \mathbf{Y}_f \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \mathbf{X}_t \\ \mathbf{X}_f \end{bmatrix}, \quad \mathbf{Z} = \begin{bmatrix} \mathbf{Z}_t \\ \mathbf{Z}_f \end{bmatrix},$$

kur

$\mathbf{Y}_t = (Y_1\ Y_2\ \ldots Y_{n-h})^T$, $\quad \mathbf{Y}_f = (Y_{n-h+1}\ Y_{n-h+2}\ \ldots\ Y_n)^T$,

$\mathbf{X}_t = \left(\mathbf{x}_1^T\ \mathbf{x}_2^T\ \ldots\ \mathbf{x}_{n-h}^T\right)^T$, $\quad \mathbf{X}_f = \left(\mathbf{x}_{n-h+1}^T\ \mathbf{x}_{n-h+2}^T\ \ldots\ \mathbf{x}_n^T\right)^T$,

$\mathbf{Z}_t = (Z_1\ Z_2\ \ldots Z_{n-h})^T$, $\quad \mathbf{Z}_f = (Z_{n-h+1}\ Z_{n-h+2}\ \ldots\ Z_n)^T$,

$$\begin{gathered} \mathbf{Y}_t = \mathbf{X}_t\boldsymbol{\beta} + \mathbf{Z}_t, \qquad \mathbf{Y}_f = \mathbf{X}_f\boldsymbol{\beta}_f + \mathbf{Z}_f, \\ E(\mathbf{Z}_t) = \mathbf{0}, \quad Cov(\mathbf{Z}_t) = \sigma^2\mathbf{I}_{n-h}, \quad E(\mathbf{Z}_f) = \mathbf{0}, \\ Cov(Z_f) = \sigma^2\mathbf{I}_h, \quad Cov(\mathbf{Z}_t, \mathbf{Z}_f) = \mathbf{0}. \end{gathered} \tag{13.}$$

Šādu modeli sauksim par *modeli ar traucējumiem (disturbed model)*.

Pētīšanas gaitā iegūstam sekojošus novērtējumus mūs interesējošam parametram, izmantojot abas pieejas, gadījumam kad ir tikai viens aplams novērtējums. Modeļa ar traucējumiem klasiskā novērtējuma matemātiskā cerība ir sekojoša:

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} - \frac{1}{1 + \mathbf{x}\left(\mathbf{X}_t^T\mathbf{X}_t\right)^{-1}\mathbf{x}^T}\left(\mathbf{X}_t^T\mathbf{X}_t\right)^{-1}\mathbf{x}^T\mathbf{x}(\boldsymbol{\beta} - \boldsymbol{\beta}_f). \tag{14.}$$

Modeļa ar traucējumiem resamplinga novērtējuma matemātiskā cerība ir sekojoša:

$$\begin{aligned} E(\boldsymbol{\beta}^*) &= \boldsymbol{\beta} - \binom{n}{k}^{-1} \times \\ &\times \left[\sum_{\mathbf{u}\in\mathbf{L}_t(k-1)} \frac{1}{1 + \mathbf{x}\left(\mathbf{X}^T(\mathbf{u})\mathbf{X}(\mathbf{u})\right)^{-1}\mathbf{x}^T} \cdot \left(\mathbf{X}^T(\mathbf{u})\mathbf{X}(\mathbf{u})\right)^{-1}\right]\mathbf{x}^{\mathbf{T}}\mathbf{x}(\boldsymbol{\beta} - \boldsymbol{\beta}_{\mathbf{f}}). \end{aligned} \tag{15.}$$

Varam redzēt, kā abi novērtējumi ir nobīdīti. Skaitliskos piemēros mēs pētīsim šo nobīdi. Darbā ir arī apskatīts gadījums, kad modelī ir vairāki traucējoši novērojumi.

Regresijas modeļa mediānas resamplinga novērtējumi

Šī pieeja bija apskatītā darbā [4].

Apspriedīsim regresijas modeļa mediānas resamplinga novērtējumu iegūšanas procedūru. Mēs nevaram uzreiz pielietot mediānas pieeju secībai (11.), tāpēc, ka tās elementi ir vektori, kurus nav iespējams sakārtot un atrast mediānu. Tāpēc mēs izmantojām šo pieeju prognožu novērtējumiem. Mēs apskatam situāciju, kad gribam iegūt kāda pagātnes vai nākotnes novērojuma atkarīga mainīgā prognozi. Izmantojot resamplinga izlases (11.) d-to elementu $\mathbf{x}_d$ (vektors-rinda) pēc formulas (9.) iegūstam mainīgā Y prognozes vērtības resamplinga izlašu novērtējumu:

$$Y^*(\mathbf{J}(l)) = \mathbf{x}_d \boldsymbol{\beta}^*(\mathbf{J}(l)), l = 1, \dots, r , \tag{16.}$$

Atkārtojot šo procedūru katram secības (11.) elementam, iegūstam secību no Y prognozes vērtības resamplinga izlašu novērtējumiem:

$$Y^*(\mathbf{J}(1)), Y^*(\mathbf{J}(2)), \dots , Y^*(\mathbf{J}(r)) . \tag{17.}$$

Sakārtosim secības (17) elementus augošā kārtībā:

$$Y^*_{d(1)}, Y^*_{d(2)}, \dots, Y^*_{d(s+1)}, \dots Y^*_{d(r)}, \tag{18.}$$

kur r=2s+1 un resamplinga izlases novērtējums, kas izraisa vidējo vērtību $Y^*_{d(s+1)}$ sakārtotajā secībā, tiks saukts par regresijas parametru mediānas resamplinga novērtējumu.

Neskatoties uz faktu, ka vidējais ir nenobīdīts matemātiskās cerības novērtējums, bet mediānas novērtējums ir nobīdīts, pēdējam tomēr ir zināmas priekšrocības. Apskatīsim dažus aspektus: priekšrocības un trūkumus, kas attiecas uz mediānas parametra $\boldsymbol{\beta}$ novērtējumu izmantošanu vidēja (12.) novērtējuma vietā. Pirmais aspekts ir tas, ka nenobīdīšanai, kā kritērijam arī piemīt zināmi trūkumi, tā nenobīdītam novērtējumam varētu būt nepareiza zīme, liela dispersija, tas varētu dot nerobustus rezultātus „trokšņainu" datu gadījumā (kā šajā gadījumā). Otrais aspekts attiecas uz to, ka, ņemot $E(\Theta^* - \Theta)^2$ kā efektivitātes kritēriju nobīdīts novērtējums var būt labāks. Pēdējais aspekts, kas ir mediānas novērtējumu priekšrocība, ir tas, ka viņi ir mediānas nenobīdīti robusti novērtējumi, gadījumam kad mēs novirzāmies no pieņēmumu statistiskās hipotēzes.

Abu pieeju efektivitātes skaitliskā analīze

Lai ilustrētu piedāvātās pieejas efektivitāti apskatīsim sekojošu skaitlisko piemēru par kādas transporta preces patēriņa prognozēšanu. Dati ir paņemti no Dreipera (Draper) un Smita(Smith) grāmatas, kur bija n=9 novērojumu un m=3 neatkarīgo mainīgo jeb faktoru.

Aprēķinājām Mahalanobisa (Mahalonobis) kvadrātisko katra novērojuma attālumu no centra:

D=(2.62 5.281 1.95 0.524 3.686 3.695 1.379 1.205 3.659).

Analizējot šo rezultātu varam secināt, ka novērojumiem ar numuriem 2, 6, 5 un 9 ir vislielākā iespēja būt par traucēkļa kandidātu, jo viņiem šis attālums ir vislielākais.

Skaitliskā abu pieeju efektivitātes pētīšana

Vesela eksperimentu sērija bija izpildīta, lai salīdzinātu klasiskās un resamplinga pieejas parametra $\boldsymbol{\beta}$ novērtējumu kvalitāti. Ja mūsu datos nav aplama novērojuma, tad klasiskie novērtējumi (8.) būs efektīvāki lineāro nenobīdītu novērtējumu klasē. Ja mūsu datos ir apšaubāmi, aplami novērojumi, tad klasiskie novērtējumi varētu būt nobīdīti.

Pieņemsim, ka viens novērojums mūsu datos ir aplams. Aprēķināsim klasisko (8.) un resamplinga novērtējumu (12.) nobīdi, kas ir apzīmēti, atbilstoši: $Bias\boldsymbol{\beta} = \boldsymbol{\beta} - E\hat{\boldsymbol{\beta}}$ un $Bias\boldsymbol{\beta}^* = \boldsymbol{\beta} - E\boldsymbol{\beta}^*$.

Mēs pārbaudījām visus iespējamos variantus, katru reizi ņemot jauno novērojumu, kā aplamo. Attiecīgie skaitliskie rezultāti doti 1. tabulā. Šeit pirmā kolonna satur aplama novērojuma indeksus, otrā kolonna satur novērojamos parametrus, trešā kolonna satur klasisko novērtējumu (8.) nobīdes.

Nākamās četras kolonnas satur resamplinga novērojumu (12.) nobīdi $Bias\boldsymbol{\beta}^*$ Katrai kolonnai ir savs resamplinga izlases apjoms k=3, 5, 6, 7.

1. tabula rāda, ka resamplinga pieeja dod labākus rezultātus, ja 2 vai 6 novērojumi uzskatīti par aplamiem. Atzīmēsim, ka tieši šiem novērojumiem bija vislielākais Mahalanobisa attālums. Tas ļauj mums formulēt sekojošu praktisko rekomendāciju. Pirms regresijas analīzes veikšanas, vajag aprēķināt Mahalanobisa attālumu un piefiksēt tos, kam šis attālums ir vislielākais. Turklāt, ja pēdējie novērojumi varēja būt aplami, tad labāk izmantot resamplinga pieeju regresijas modeļa parametru novērtēšanai.

1. tabula

Klasisko un resamplinga regresijas parametru novērtējumu nobīdes salīdzināšana

i	$\boldsymbol{\beta}$	$Bias\hat{\boldsymbol{\beta}}$	$Bias\boldsymbol{\beta}^*$			
			k=3	k=5	k=6	k=7
1	β_1	0,018	0,268	0,033	0,020	0,015
	β_2	-0,433	-1,500	-0,563	-0,477	-0,436
	β_3	0,263	0,328	0,294	0,278	0,268
2	β_1	1,895	0,989	1,403	0,020	0,015
	β_2	-7,615	-3,900	-5,586	-0,477	-0,436
	β_3	0,158	0,055	0,097	0,278	0,268
3	β_1	0,879	1,561	1,310	1,170	1,027
	β_2	-4,126	-6,844	-5,873	-5,310	-4,729
	β_3	0,397	0,420	0,432	0,425	0,413
4	β_1	-0,350	-1,138	-0,538	-0,452	-0,394
	β_2	1,455	4,757	2,233	1,874	1,635
	β_3	0,042	-0,039	0,026	0,034	0,039
5	β_1	0,171	0,243	0,194	0,178	0,171
	β_2	-1,190	-1,771	-1,363	-1,258	-1,204
	β_3	0,307	0,414	0,339	0,322	0,312
6	β_1	-0,541	0,063	-0,122	-0,266	-0,266
	β_2	3,527	0,678	1,646	2,311	2,311
	β_3	-0,576	-0,364	-0,465	-0,510	-0,510
7	β_1	-0,978	-1,773	-1,395	-1,240	-1,099
	β_2	4,386	7,636	6,125	5,477	4,886
	β_3	-0,172	-0,278	-0,235	-0,210	-0,188
8	β_1	0,056	0,981	0,421	0,259	0,139
	β_2	0,046	-3,401	-1,296	-0,706	-0,262
	β_3	-0,053	-0,058	-0,065	-0,057	-0,054
9	β_1	-1,067	-1,111	-1,224	-1,168	-1,116
	β_2	5,286	5,681	6,013	5,751	5,510
	β_3	-0,410	-0,521	-0,467	-0,444	-0,426

Resamplinga-mediānas novērtējums

Mēs izpildījām citu eksperimentu sēriju ar tiem pašiem datiem

Mēs ņemam 2. novērtējumu, kā aplamo, tāpēc ka tam bija vislielākais Mahalanobisa attālums, tas nozīmē, ka tam bija vislielākās izredzes būt traucēklim. Resamplinga-mediānas novērtējumu prognozētās vērtības ir aprēķinātas pēc formulas (18.). Mēs salīdzinājām šos rezultātus ar tradicionālas pieejas novērtēšanas rezultātiem, ņemot nobīdi, kā efektivitātes kritēriju.

Lai pētītu resamplinga izlases apjoma k ietekmi uz novērtējumu īpašībām, mēs mainījām to intervālā $m \leq k < n$. Atzīmēsim, ka šajā piemērā resampling-izlašu skaits r, lai iegūtu resamplinga novērtējumu ir vienāds ar visām k no n kombinācijām. Vispārīgā gadījumā resamplinga izlašu skaits ir krietni mazāks, nekā visas kombinācijas.

Mēs iegūstam atkarīga mainīgā prognozētās vērtības resamplinga-mediānas novērtējumus visiem iespējamiem novērojumiem. Resamplinga-mediānas un klasiskus novērtējumus mēs salīdzinājām ar tīriem novērtējumiem, kas bija aprēķināti bez aplama novērtējuma. Minētais nozīmē, ka izņēmām aplamu novērtējumu no modeļa un izmantojām MKN (8.). Attiecīgie rezultāti ir apkopoti 2. tabulā.

2. tabula

Resamplinga mediānas pieejas novērtējumu nobīdes piemēram

Para-metri	Prognozētā novērtējuma nobīde						
	MKN	resamplinga-mediānas					
		k=3	k=4	k=5	k=6	k=7	k=8
Y_1	0.115	0.617	0.43	0.332	0.139	**0.034**	0.333
Y_2	5.793	7.57	7.441	5.802	**4.85**	**5.037**	**5.631**
Y_3	1.957	5.131	3.165	2.301	**1.702**	**1.682**	1.967
Y_4	0.497	2.571	**0.108**	**0.243**	0.545	0.739	0.772
Y_5	0.129	1.012	0.636	0.282	**0.024**	**0.042**	**0.066**
Y_6	1.481	4.385	3.53	**1.242**	**0.523**	**0.517**	**1.246**
Y_7	1.519	**0.416**	**0.681**	**0.959**	**1.204**	1.572	1.935
Y_8	1.199	1.557	1.523	**0.111**	**0.073**	**0.539**	**0.749**
Y_9	0.499	2.857	0.956	**0.261**	**0.493**	0.543	0.778

Pirmā kolonna satur mūs interesējošo prognozējamo parametru nosaukumus viesiem novērtējumiem. Otrā kolonna satur nobīdi, kā starpību starp klasiskajiem un tīriem novērtējumiem. Nākamajās kolonnās ir apkopoti resamplinga-mediānas pieejas novērtējumu rezultāti dažādiem izlases apjomiem *k*. Tur ir prezentētas nobīdes, kā starpība starp resamplinga novērtējumiem un tīriem novērtējumiem. Analizējot iegūtos rezultātus, varam secināt, kā resamplinga pieeja visiem novērojumiem deva mazāku nobīdes vērtību, nekā klasiskā, ņemot adekvātu resamplinga izlases apjomu. Īpaši labi rezultāti bija iegūti, kad k=6. Piemēram, 7. novērojumam klasiskā pieeja deva nobīdi vienādu ar 1.519, bet resamplinga-mediānas pieeja intervālā (0.416-1.572), atkarībā no resamplinga izlases apjoma.

Piemērs: Pasažieru pārvadājumu ar aviācijas transportu prognozēšanas regresijas modelis

Apskatīsim pasažieru pārvadājumu ar aviācijas transportu prognozēšanas regresijas modeli. Prognoze ir balstīta uz sekojošiem faktoriem:

- x_1-valsts teritorijas plašumu (km^2);
- x_2 – valsts iedzīvotāju kopskaitu (cilv.);
- x_3 – iekšējais valsts kopprodukts uz vienu iedzīvotāju (EUR);
- x_4 –viena iedzīvotāja vidējo mēnešalgu (EUR).

Regresijas modelis būs sekojošais:

$$y=\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3+\beta_4 x_4+\varepsilon,$$

kur,

y – vienas valsts teritorijā transportētu pasažieru skaits viena gada laikā.

ε - gadījuma komponenta, ar normālu sadalījumu ar nulles vidējo un pastāvīgo dispersiju. Sākumdati prognozēšanai, kas ir apkopoti darbā, ir no Eiropas Savienības oficiālās statistikas Web-lapas. Dati bija sadalīti divās grupās: pirmā grupa izmantota modeļa veidošanai, otrā grupa izmantota tikai prognozes modeļa pārbaudei.

Piedāvātā klasiskā modeļa korektuma analīzei bija izmantoti sekojoši kritēriji: Fišera kritērijs hipotēzes pārbaudei par regresijas nenozīmīgumu (ar nozīmīguma līmeni α=5%), Stjudenta kritērijs hipotēzes pārbaudei par i-tā faktora nenozīmīgumu (α=5%), daudzkārtīgas determinācijas koeficients R^2.

Modeļos 2-3 bija nolemts modificēt sākumdatus, izmantojot dažu faktoru kombinācijas. Tiek prognozēta individuālā iedzīvotāju mobilitāte (transportētu pasažieru daļa attiecībā uz iedzīvotāju kopskaitu) katrai valsts y/x_2.

3. tabula

Dažādi prognozēšanas modeļi

Nr. p.k.	Regresijas modelis	R^2
1.	$y=\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3+\beta_4 x_4+\varepsilon$, Koeficienti, sakārtoti pēc nozīmīguma pakāpes: β_3, β_2, β_4.	0.85
2.	$y/x_2=\beta_0+\beta_1 x_4+\beta_2(x_1/x_2)+\beta_3(x_3/x_2)+\beta_4 x_3/(x_2 * \sqrt{x_1})+$ $+\beta_5 x_3/(x_2 * x_1)+\varepsilon$, Koeficienti, sakārtoti pēc nozīmīguma pakāpes: β_1, β_3, β_5, β_4, β_2, β_0.	0.90
3.	(Ar izslēgto Luksemburgu) $y/x_2=\beta_0+\beta_1 x_3+\beta_2 x_4+\beta_3(x_1/x_2)+\beta_4(x_3/x_2)+\beta_5(x_4/x_2)+$ $+\beta_6\sqrt{x_1}+\beta_7 x_3/(x_2 * \sqrt{x_1})+\beta_8 x_3/(x_2 * x_1)+\varepsilon$, Koeficienti, sakārtoti pēc nozīmīguma pakāpes: β_7, β_3 ,β_8, β_2	0.92

Visu modeļu koeficienti tika novērtēti ar klasisko un resamplinga metodēm. Paši modeļi, to parametri un daudzkārtīgas determinācijas koeficienti ir apkopoti 3. tabulā.

Resamplinga pieeja praktiski visur deva salīdzināmus, vairākos gadījumos pat labākus prognozes novērtējumu rezultātus nekā tradicionālā pieeja, ņemot kā efektivitātes kritēriju prognozes novērtējuma nenobīdīšanu. Šī salīdzinājuma procentuālās attiecības ir apkopotas 4. tabulā.

Secinājumi

Apskatīta resamplinga pieeja regresijas modeļa novērtēšanai gadījumā, kad ir traucējoši novērojumi starp visiem novērojumiem. Regresijas parametru novērtējumu nobīde bija aprēķināta ar klasisko un resamplinga pieeju. Apskatīti nosacījumi, kurus izpildot, resamplinga pieeja dod mazāku nobīdes vērtību, nekā klasiskā pieeja. Ja piedāvāto pieeju izmantot modeļiem ar traucējošiem novērojumiem, tā dod labākus rezultātus, nekā tradicionāla pieeja. Skaitlisko rezultātu analīze rāda, kā piedāvāta mediānas resamplinga pieeja arī dot salīdzinoši labus rezultātus, ja izvēlamies novērtējuma EY_d nobīdi, kā efektivitātes kritēriju.

Resamplinga pieejas pielietošana bija ilustrēta, ja veic pasažieru pārvadājumus ar aviācijas transportu prognozēšanas regresijas modelim.

4. tabula

Klasiskās un resamplinga pieejas rezultātu salīdzinājums

Nr. p.k.	Modeļa dati (resamplings/klasiskais)	Pārbaudes dati (resamplings/klasiskais)
1.	24/14 (63/37)%	17/7 (71/29)%
2.	21/17 (55/45)%	15/9 (62.5/37.5)%
3.	25/10 (71/29)%	5/14 (26/74)%

9.5. Par vienu transporta līdzekļu drošības un efektivitātes novērtēšanas uzdevumu

Darba saturīgs apraksts

Drošums - tā ir varbūtība, ka ierīce izpilda savas funkcijas saskaņā ar izvirzītajām prasībām noteiktajā laika intervālā. Kā zināms, transportā šim rādītājam ir liela nozīme. Problēmas, kas saistītas ar transporta līdzekļu drošumu, itin bieži var būt dzīvībai nozīmīgas. Problēmas, kas saistītas ar piegādes drošumu loģistikas ķēdēs, var radīt kompānijai milzīgus finansiālus zaudējumus. Minētais pieprasa vērtējuma un prognozēšanas matemātiskā aparāta piesaisti, kas veicina pareizu pārvaldes risinājumu izstrādi.

Šai problēmai raksturīga ir statistikas sākumdatu apjomu ierobežotība, piemēram, par atteikumiem, kas ierobežo klasisko metožu novērtēšanas izmantošanu. Tas pieprasa jaunu intensīvu datormetožu piesaisti, kas spēj novērst šīs problēmas. Šī pētījuma rezultāti ir publicēti darbos [2], [9].

Matemātiskais modelis un tā analīzes metodes

Apskatīsim rindošanas procesu matemātisko modeli, kas ir raksturīgs drošības teorijai. Modelis paredz divu veidu bojājumus: sākuma un gala-bojājums. Sākuma bojājums parādījās atbilstoši Puassona procesam ar intensitāti λ. Katrs sākuma bojājums pārvēršas par gala bojājumu pēc gadījuma laika *B*. Ja sākuma bojājums parādījās laikā τ_i, tad tas kļūst gala bojājums momentā $B_i + \tau_i$. Gala bojājumi un attiecīgie sākuma bojājumi uzreiz tiek likvidēti. Pieņemsim, ka $\{B_i\}$ ir savstarpēji neatkarīgi, vienādi sadalīti gadījuma mainīgie, kas ir arī neatkarīgi no $\{\tau_i\}$. Lai $F(x)$ ir GL B sadalījuma funkcija. Mūs interesē sākuma bojājumu skaits $X(t)$ laika momentā t (kuri nav pārvērtušies par gala bojājumiem) un gala bojājumu skaits $Y(t)$, kas notikuši laikā t. Lai $\Lambda(t) = EX(t)$ un $\widetilde{\Lambda}(t) = EY(t)$ būtu attiecīgas matemātiskas cerības, $P_i(t) = P\{X(t) = i\}$, $R_i(t) = P\{Y(t) = i\}$ būtu attiecīgie sadalījuma līkumi, $i = 0, 1, \ldots$.

Labi zināms, ka $X(t)$ un $Y(t)$ ir savstarpēji neatkarīgi GL, tad:

$$\Lambda(t) = \lambda \int_0^t (1 - F(x))dx\,,\ \widetilde{\Lambda}(t) = \lambda \int_0^t F(x)dx\,, \qquad (19.)$$

$$P_i(t) = \frac{1}{i!}(\Lambda(t))^i \exp(-\Lambda(t)),\ i = 0, 1, \ldots \qquad (20.)$$

Varbūtību $R_i(t)$ analoģiski var aprēķināt pēc formulas (20.), kur $\Lambda(t)$ ir aizvietots ar $\widetilde{\Lambda}(t)$.

Faktiski intensitāte λ un sadalījuma funkcija $F(x)$ nav zināmas. Mums jānovērtē $\Lambda(t)$, $\widetilde{\Lambda}(t)$, $R_i(t)$ un $P_i(t)$ izmantojot intervālu starp sākuma bojājumu rāšanos izlasi $A_1, A_2, \ldots, A_k$ un izlasi $B_1, B_2, \ldots, B_l$.

Mēs apskatījām divas novērtējuma metodes: tradicionālo un resamplinga. Šī daļa satur apskatīto novērtējumu matemātiskās cerības un dispersijas īpašību pētījumus. Novērtējumu nobīde un dispersija bija izmantotas kā abu pieeju efektivitātes salīdzinošs kritērijs. Paradījām, ka resamplinga novērtējumiem ir dažas priekšrocības, kad izlašu apjomi k un l ir mazi.

Tradicionālās (plug-in) pieejas saturs

Plug-in novērtējumi izmanto novērtējumus $\hat{\lambda}$ un $\hat{F}(t)$ nezināmo λ un $F(t)$ vietā. Šeit $\hat{F}(t)$ ir empīriskā GL B sadalījuma funkcija, kas bija aprēķināta, balstoties uz $B_1, B_2, ..., B_l$ un $\hat{\lambda}$ intensitātes λ punktsveida novērtējumu:

$$\hat{\lambda} = \left(\frac{1}{k} \sum_{i=1}^{k} A_i \right)^{-1}. \tag{21.}$$

Šajā gadījumā mums ir sekojošie $\Lambda(t)$ un $P_i(t)$ novērtējumi:

$$\hat{\Lambda}(t) = \hat{\lambda} \int_0^t \left(1 - \hat{F}(x)\right) dx, \tag{22.}$$

$$\hat{P}_i\ (t) = \frac{1}{i!} \hat{\Lambda}(t)^i \exp\left(- \hat{\Lambda}(t)\right),\ i = 0,\ 1,\ \ldots\ . \tag{23.}$$

Novērtējumus parametriem $\widetilde{\Lambda}(t)$ un $R_i(t)$ var saņemt pēc analoģijas. Lai pētītu šo novērtējumu statistiskas īpašības, ir nepieciešams zināt sekojošu GL sadalījumus $\hat{\lambda}$ un $\int_0^t (1 - \hat{F}(u))du$. Ieguvām formulas arī plug-in novērtējumu (22.), (23.) matemātiskai cerības un dispersijai. Novērtējuma $\hat{\Lambda}(t)$ vidēja kvadrātiskā kļūda ir sekojoša: $MSE\hat{\Lambda}(t) = Var\hat{\Lambda}(t) + \left(E\hat{\Lambda}(t) - \Lambda(t)\right)^2$.

Resamplinga pieejas saturs

Resamplinga pieeja paredz parastu imitācijas modelēšanu, kurai ir viena atšķirība, tā neizmanto gadījuma lielumu ģeneratorus. Saskaņā ar to, nepieciešamie elementi tiek iztverti gadījuma veidā tieši no dotajām izlasēm $\{A_1, A_2, ..., A_k\}$ un $\{B_1, B_2, ..., B_l\}$. Lai $k \le l$. Mēs izpildām r neatkarīgu imitācijas modelēšanas procesa realizāciju. Veicot q-to realizāciju, izlasām elementus no $\{A_1, A_2, ..., A_k\}$ bez atkārtojuma, veidojam secību no laika intervāliem starp sākuma bojājumu rašanos $A(q) = \{A_{i_1(q)}, A_{i_2(q)}, ..., A_{i_{N_t}(q)}\}$. Aprēķinām $\tau_i(q) = \sum_{u=1}^{i} A_{i_u(q)}$, i = 1,2, …, $N_t(q)$, kur $N_t(q)$ ir sākuma bojājumu skaits par laiku t realizācijā q:

$$N_t(q) = \begin{cases} \max\{j : \tau_j(q) \le t\} & ja \quad \tau_k(q) \ge t, \\ k & citādi. \end{cases} \tag{24.}$$

Analoģiski izveidojām secību $B(q) = \{B_{j_1(q)}, B_{j_2(q)}, ..., B_{j_{N_t}(q)}\}$ no intervāliem līdz sākuma bojājumu pārveidošanas uz gala bojājumiem. Tad aprēķinām secību $\{\tau_1(q) + B_{j_1(q)}, \tau_2(q) + B_{j_2(q)}, ..., \tau_{N_t}(q) + B_{j_{N_t}(q)}\}$ no gala bojājumu rašanās momentiem q-ajā realizācijā.

Lai $\zeta_j(t)$ ir indikācijas funkcija notikumam: "j-tais sākuma bojājums noticis, bet nav pārveidojies par gala bojājumu līdz laika momentam":

$$\zeta_{j,q}(t) = \begin{cases} 1 & ja \quad \tau_j(q) \le t < \tau_j(q) + B_{j_j(q)}, \\ 0 & citādi. \end{cases} \tag{25.}$$

Tad sākuma bojājumu skaitu $\mathrm{X}_q(t)$, kas nav pārveidojušies pār gala bojājumiem līdz laikam t, realizācijai q var būt aprēķināt sekojoši:

$$\mathrm{X}_q(t) = \sum_{j=1}^{N_t(q)} \zeta_{j,q}(t) = \sum_{j=1}^{k} \zeta_{j,q}(t) . \tag{26.}$$

Resamplinga $\Lambda(t)$ novērtējums ir sekojošs:

$$\Lambda^*(t) = \frac{1}{r} \sum_{q=1}^{r} X_q(t) . \tag{27.}$$

Analoģiski var aprēķināt sākuma bojājumu skaitu $Y_q(t)$, kas ir pārveidojušies par gala bojājumiem līdz laikam *t,* realizācijai q , attiecīgi aizvietojot formulā (26.) funkciju $\zeta_{j,q}(t)$. Resamplinga novērtējumu $\widetilde{\Lambda}^*(t)$ var iegūt pēc formulas (27.), aizvietojot funkciju $X_q(t)$ ar $Y_q(t)$.

Tagad mums jāaprēķina resamplinga novērtējumus varbūtībām $P_i(t)$ un $R_i(t)$. Lai $\Phi_i(X(t))$ ir indikācijas funkcija notikumam $\{X(t)=i\}$. Resamplinga novērtējums varbūtībai $P_i(t)$ ir sekojošs:

$$P_i^*(t) = \frac{1}{r}\sum_{q=1}^{r} \Phi_i(\mathrm{X}_q(t)). \tag{28.}$$

Resamplinga novērtējumu $R_i^*(t)$ var aprēķināt pēc formulas (28), aizvietojot $\Phi_i(X(t))$. Aprēķināsim resamplinga novērtējumu matemātisko cerību. Ir skaidrs, kā $EP_i^*(t) = E\Phi_i(t)$.

Tādējādi, resamplinga novērtējuma $\Lambda^*(t)$ matemātisko cerību $E\Lambda^*(t)$ var aprēķināt sekojoši:

$$E\Lambda^*(t) = q_1 \sum_{j=1}^{k} j \frac{(\lambda t)^j}{j!} \cdot \exp(-\lambda t) + q_1 k \cdot \sum_{j=k+1}^{\infty} \frac{(\lambda t)^j}{j!} \cdot \exp(-\lambda t). \tag{29.}$$

Bija izveidota formula arī novērtējuma $P_i^*(t)$ matemātiskai cerībai $EP_i^*(t)$:

$$\begin{aligned} EP_i^*(t) = \sum_{j=i}^{l} \frac{(\lambda t)^j}{j!} \cdot \exp(-\lambda t) \cdot \binom{j}{i} \cdot q_1{}^i \cdot (1-q_1)^{j-i} + \\ + \binom{l}{i} \cdot q_1{}^i \cdot (1-q_1)^{l-i} \cdot \sum_{j=l+1}^{\infty} \frac{(\lambda t)^j}{j!} \cdot \exp(-\lambda t),\ i = 0,1,2,..., \end{aligned} \tag{30.}$$

kur $\binom{j}{i}$ ir binomiālais koeficients; q_1 – varbūtība, ka līdz momentam t fiksēts sākuma bojājums paliks kā sākuma.

Novērtējumu $R_i^*(t)$ un $\tilde{\Lambda}^*(t)$ matemātiskās cerības $ER_i^*(t)$ un $E\tilde{\Lambda}^*(t)$ var aprēķināt analoģiski. Bija saņemti arī resamplinga novērtējumu (28.) dispersijas $Var\Lambda^*(t)$ izteikumi. Vidējo kvadrātisko kļūdu var aprēķināt pēc formulas: $MSE\Lambda^*(t) = Var\Lambda^*(t) + \left(E\Lambda^*(t) - \Lambda(t)\right)^2$.

Pieeju efektivitātes Skaitliskā analīze
Piemērs: "Trijstūra sadalīts laiks starp sākuma bojājumu pārveidošanu par gala bojājumu".

Lai sākuma bojājumi rodas saskaņā ar Puassona sadalījumu ar intensitāti λ=0.4; Trijstūra sadalīts laiks starp sākuma bojājuma pārveidošanu par gala bojājumu ar parametru a=2.

5. tabulā ir apkopotas tradicionālas (Trad) un resamplinga (Res) pieeju matemātiskās cerības $E\hat{P}_i(t)$ un $EP_i^*(t)$, kas fiksētas laika momentā t=5, ar dažādiem resamplinga izlašu apjomiem l, salīdzinājumā ar īstām attiecīgo notikumu varbūtībām.

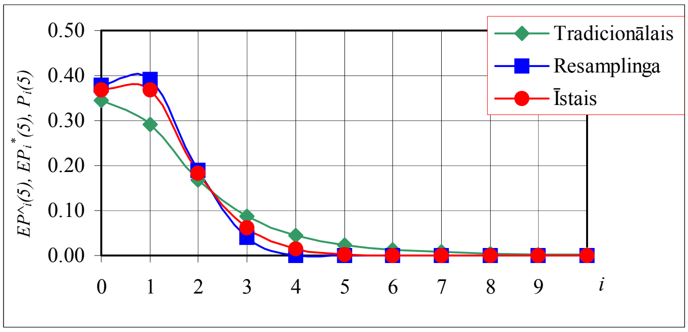
Mēs apskatam gadījumu, kad ieejas izlases apjomi l un k ir vienādi. Redzam, ka ar l palielināšanos, nobīde samazinās un resamplinga pieejas novērtējumam pavisam pazūd.

Mēs varam viegli salīdzināt šos novērtējumus, analizējot 3. un 4. attēlus. 4. attēlā ir redzams, ka īstas varbūtības līkne un resamplinga pieejas matemātiskā cērība veido vienu līkni, ja l=8.

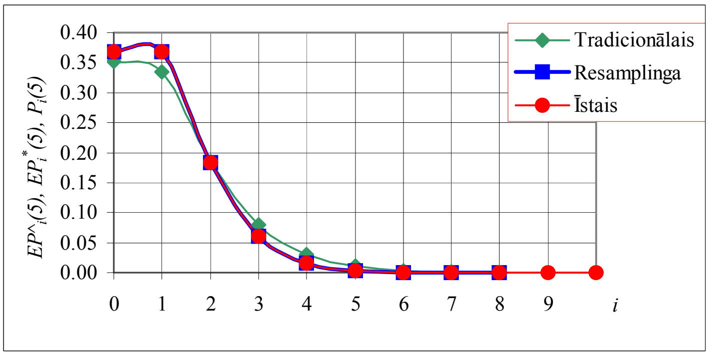
5. tabula

Plug-in un resamplinga novērtējumu matemātiskas cerības: $E\hat{P}_i(5)$, $EP_i^*(5)$

i	*l=3*		*l=4*		*l=5*		*l=8*		Īstā
	Plug.	Res.	Plug.	Res.	Plug.	Res.	Plug.	Res.	
0	.346	.379	.348	.370	.350	.368	.352	.368	.368
1	.291	.392	.307	.374	.317	.369	.334	.368	.368
2	.169	.189	.176	.189	.180	.186	.186	.184	.184
3	.088	.040	.087	.058	.085	.062	.080	.061	.061
4	.045		.041	.009	.037	.014	.031	.015	.015
5	.024		.019		.016	.002	.011	.003	.003
6	.013		.010		.007		.004		.001
7	.008		.005		.003		.001		
8	.005		.003		.002		.001		
9	.003		.001		.001				
10	.002		.001						



3. att. Plug-in un resamplinga novērtējumu matemātiskas cerības: $E\hat{P}_i(5)$**,** $EP_i^*(5)$**,** ***l*=3, *t*=5.**



4. att. Plug-in un resamplinga novērtējumu matemātiskās cerības: $E\widehat{P}_i(5)$**,** $EP_i^*(5)$**,** ***l*=8, *t*=5.**

6. tabulā ir apkopotas novērtējumu matemātiskās cerības $E\hat{\Lambda}(t)$, $E\Lambda^*(t)$, dispersijas $Var\hat{\Lambda}(5)$, $Var\Lambda^*(5)$ un vidējās kvadrātiskās kļūdas $MSE\ \hat{\Lambda}(5)$, $MSE\ \Lambda^*(5)$.

Pētniecības procesā tika apskatīti dažādi piemēri ar dažādiem sadalījumu veidiem un parametru vērtībām, bet rezultātiem tendences līdzinājās tikko aprakstītai situācijai. Visās situācijās resamplinga pieeja deva salīdzinošus, vai pat labākus rezultātus, nekā tradicionālā. Tas ir īpaši redzams, kad izlases apjoms pārsniedza 3, jo mazāks izlases apjoms nevarēja dot ticamus rezultātus resamplinga pieejai. Ar resamplinga pieeju mēs nevarējām iegūt varbūtības rindas garumam lielākam nekā l (resamplinga izlases apjoms), kas varētu notikt reālā dzīvē.

Resamplinga pieeja vairākos gadījumos dod mazāku nobīdi, nekā tradicionālā nobīde, bet kad resamplinga izlases apjoms ir pietiekami liels (lielāks pār 8) nobīde pavisam pazūd.

6. tabula

Plug-in un resamplinga novērtējumu $\Lambda(t)$ matemātiskās cerības, dispersijas un vidējās kvadrātiskās kļūdas

i	l=3	l=4	l=5	l=6	l=7	l=8
$E\hat{\Lambda}(5)$	1.41	1.32	1.25	1.21	1.19	1.16
$E\Lambda^*(5)$	0.89	0.96	0.99	0.997	0.99	0.99
$Var\hat{\Lambda}(5)$	1.52	0.79	0.51	0.38	0.30	0.24
$Var\Lambda^*(5)$	0.58	0.55	0.49	0.43	0.36	0.31
$MSE\ \hat{\Lambda}(5)$	1.69	0.89	0.57	0.42	0.34	0.27
$MSE\ \Lambda^*(5)$	0.59	0.55	0.49	0.43	0.36	0.31

Secinājumi

Piedāvātā resampling pieeja ir laba alternatīva tradicionālajam dotajam drošības uzdevumam. Tas bija īpaši pamanāms ar izlases apjomu palielināšanos, kad resamplinga novērtējumu konverģence reālām parametru vērtībām notika ātrāk nekā tradicionālai. Vienīgais piedāvātās pieejas trūkums, kā mēs nevaram iegūt vērtības $EP_i^*(t)$, ja $i > l$. Tādos gadījumos ir labāk izmantot tradicionālu pieeju. Labākais veids ir apvienot šīs divas pieejas un izmantot resamplinga novērtējumus, ja $i \le l$ un tradicionālos novērtējumus, ja $i > l$. Mēs varam izmantot tādu kombināciju ar speciālu normalizāciju, lai varbūtību summa būtu vienāda ar 1.

Svarīgi atzīmēt, ka apskatīts gadījums, kad sadalījums bija izvēlēts pareizi. Reālajā dzīvē bieži vien mēs sastopamies ar situācijām, kad pieejamās izlases ir pārāk mazas (kā šajā gadījumā), un mēs nevaram pareizi izvēlēties sadalījumus. Tad mēs varam izveidot pavisam citu modeli, kura rezultātiem vispār nevar ticēt. Šādās situācijās īpaši efektīvi var izmantot resamplinga pieeju. Tā kā resamplinga pieeja ir konkurētspējīga pat sliktākās situācijās, tās priekšrocības minētajā gadījumā ir acīmredzamas.

9.6. Loģistisko sistēmu krājumu pārvalde

Problēmas apraksts un uzdevuma matemātiskā nostādne

Ja krājumi tiek apskatīti kā vadības objekts loģistikās sistēmās, tad galvenie jautājumi ir: krājuma optimāla līmeņa aprēķins, risku krājumu deficīta rašanās.

Pieņemsim, ka mums ir divi neatkarīgi atjaunošanas procesi $\{X_i, i=1,2,...\}$ un $\{Y_i, i=1,2,...\}$, kur $\{X_i\}$ un $\{Y_i\}$ - secības no nenegatīviem neatkarīgiem GL, katrai secībai savs kopīgais sadalījums. Lai $D_m = \sum_{i=1}^{m} X_i$ un $S_m = \sum_{i=1}^{m} Y_i$ ir attiecīgo procesu m-to atjaunojumu laika momenti. Secību $\{X_i\}$ un $\{Y_i\}$ sadalījuma funkcijas nav zināmas, bet attiecīgas izlases ar apjomiem n_X un n_Y ir pieejamas. Pētījuma mērķis ir novērtēt varbūtību $P\{D_m > S_k\}$, kur $n_X \geq 2m$ un $n_Y \geq 2k$.

Šai problēmai ir daudz pielietojumu, piemēram, krājumu pārvaldes teorijā var apskatīt sekojošu uzdevumu. Pieņemsim, ka krājuma sākotnējais līmenis ir vienāds ar K, kur K ir zināms veselais skaitlis. Krājuma līmenis pieaug saskaņā ar piedāvājumu un samazinās, saskaņā ar pieprasījumu. Ja pieprasījums pārsniedz piedāvājumu, tad rodas deficīts. Pētījuma mērķis ir novērtēt deficīta trūkuma varbūtību m-tā pieprasījuma momentā.

Aprakstīto piemēru var apskatīt atjaunošanas procesu terminos. Lai pieprasījums atbilst pirmajam atjaunošanas procesam $\{X_i, i=1,2,...\}$ un m-tā atjaunojumu moments atbilst m-tajam pieprasījumam. Lai piedāvājums atbilstu otrajam atjaunošanas procesam $\{Y_i, i=1,2$ un m-tā atjaunojumu moments atbilstu m-tās preces vienības piedāvājumam. Tad mūs interesējošā varbūtība ir notikumam $D_m>S_{m-K}$, ka m-tais pieprasījums būs vēlāk nekā $m-K$-tais piedāvājums. Ir paredzēts arī ka, krājuma sākotnējais līmenis ir zināms. Mēs gribam pētīt deficīta trūkuma varbūtības dažādu pieeju novērtējumu īpašības.

Formāla uzdevuma nostādne

Apskatīsim šo problēmu formāli. Lai mums secību $\{X_i\}$ un $\{Y_i\}$ sadalījuma funkcijas $F_X^1(x)$ un $F_Y^1(x)$, bet attiecīgas izlases

$H_X = \{X_1, X_2, \ldots, X_{n_X}\}$ un $H_Y = \{Y_1, Y_2, \ldots, Y_{n_Y}\}$ ir pieejamas, kur $|H_X| = n_X$ un $|H_Y| = n_Y$.

Mēs vēlamies uzzināt m-tā un $m-K$ -tā atjaunojumu laika momentus: $D_m = \sum_{i=1}^{m} X_i$, $S_{m-K} = \sum_{i=1}^{m-K} Y_i$. Mūsu uzdevums ir novērtēt varbūtību $P\{D_m > S_{m-K}\}$, ka m-tais pieprasījumu procesa $\{X_i\}$ atjaunojums notiks vēlāk, nekā $m-K$ -tais piedāvājuma procesa $\{Y_i\}$ atjaunojums.

Apspriedīsim indikācijas funkciju $\Psi(\mathbf{x}, \mathbf{y})$, kur $\mathbf{x} = (x_1, x_2, \ldots, x_{m_X})$ un $\mathbf{y} = (y_1, y_2, \ldots, y_{m_Y})$ ir reālu skaitļu vektors:

$$\Psi(\mathbf{x}, \mathbf{y}) = \begin{cases} 1 & if \sum_{i=1}^{m_X} x_i > \sum_{i=1}^{m_Y} y_i, \\ 0 & citādi. \end{cases} \qquad (31.)$$

Pieņemsim, ka mums ir divi GL vektori $\mathbf{X} = (X_1, X_2, \ldots X_{m_X})$ un $\mathbf{Y} = (Y_1, Y_2, \ldots Y_{m_Y})$, $m_X = m$, $m_Y = m - K$. Tad mūs interesējošo varbūtību varētu prezentēt ar indikācijas funkcijas matemātisko cerību: $\Theta = E(\Psi(\mathbf{X}, \mathbf{Y}))$. Mēs novērtēsim Θ, izmantojot dažādas pieejas: klasisko un resamplinga. Klasiskā pieeja ir parametriskā, ko daudzi zina. Mēs apskatīsim alternatīvo neparametrisko resamplinga pieeju.

Klasiskā pieeja

Klasiskā pieeja apskatāmās varbūtības novērtēšanai ir parametriskā. Tā paredz sadalījumu parametru punktsveida novērtējumu, ja ieejas izlases H_i $i=\{X, Y\}$ sadalījuma likums ir zināms.

Piemērs: Eksponenciālais sadalījums

Apskatīsim piemēru, kad GL X un Y ir eksponenciāli sadalīti ar parametriem λ un v attiecīgi. Kā ir zināms, eksponenciāli sadalītai GL summai ir Erlanga sadalījums. Mūs interesē varbūtība $\Theta = P\{D_{m_X} > S_{m_Y}\}$.

Klasiskā pieeja paredz punktsveida novērtējumus izmantot vērtību λ un v vietā:

$$\hat{\Theta} = \sum_{i=0}^{m_X-1} \frac{\hat{\nu}^{m_Y}}{(\hat{\lambda}+\hat{\nu})^{m_Y+i}} \frac{\hat{\lambda}^i}{i!} \prod_{p=0}^{i-1} (m_Y + p), \qquad (32.)$$

kur $\prod_{p=0}^{-1} = 1$, $\hat{\lambda} = n_X / D_{n_X}$ un $\hat{\nu} = n_Y / S_{n_X}$.

Tagad varam aprēķināt novērtējuma $\hat{\Theta}$ matemātisko cerību un dispersiju.

Piemērs: Normālais sadalījums

Apskatīsim tagad gadījumu, kad GL X un Y ir normālais sadalījums, attiecīgi N(μ_X,σ_X) un N(μ_Y,σ_Y). Deficīta trūkuma reāla varbūtība ir: $\Theta = P\{D_{m_X} > S_{m_Y}\} = P\{D_{m_X} - S_{m_Y} > 0\}$.

Lai novērtētu šo varbūtību, klasiskā pieeja paredz parametru $\boldsymbol{\mu} = (\mu_X, \mu_Y)$, $\boldsymbol{\sigma} = (\sigma_X, \sigma_Y)$ punktsveida novērtējumu pēc pieejamām izlases populācijām. Mērķa novērtējums ir sekojošs:

$$\hat{\Theta} = \hat{\Theta}(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\sigma}}) = 1 - \Phi\left(\frac{0 - (m_X \hat{\mu}_X - m_Y \hat{\mu}_Y)}{\sqrt{m_X \hat{\sigma}_X^2 + m_Y \hat{\sigma}_y^2}} \right). \qquad (33.)$$

Tagad varam aprēķināt novērtējuma $\hat{\Theta}$ matemātisko cerību $E\hat{\Theta}$, dispersiju $Var\,\hat{\Theta}$ un vidējo kvadrātisko kļūdu $MSE\hat{\Theta} = Var\,\hat{\Theta} + (\Theta - E\,\hat{\Theta})^2$.

Resamplinga pieeja

Šī metode, atšķirībā no tradicionālās pieejas, neparedz sadalījumu parametru novērtējumu, vai empīrisko sadalījuma funkciju konstruēšanu, lai atrastu mērķa rādītājus. Alternatīvi, mēs izmantojam sākumdatus dažādās kombinācijās, kas ļauj iegūt nenobīdītus novērtējumus un samazināt to dispersiju. Resamplinga pieeja paredz sekojošu soļu izpildi. Mēs, gadījuma veidā izvēlamies m_X elementus un izlases H_X un m_Y elementus no izlases H_Y. Elementi tiek paņemti bez atkārtojumiem, atkārtosim, ka $n_X \geq 2m_X$, $n_Y \geq 2m_Y$. Tad attiecīgas funkcijas $\Psi(\mathbf{x}, \mathbf{y})$ vērtības aprēķinām pēc formulas (31.). Pēc tam izvēlētus elementus atgriežam attiecīgajās izlasēs.

Atkārtojam šo procedūru r reizēs (r realizācijas). Lai $j_d^i(l)$, d=1,..,m_i izlases H_i, $i \in \{X, Y\}$ elementi l-tajā realizācijā. Tad iegūstam sekojošus vektorus:

$$\mathbf{X}(l) = (X_{j_1^X(l)}, X_{j_2^X(l)}, \ldots, X_{j_{m_X}^X(l)}),\ \mathbf{Y}(l) = (Y_{j_1^Y(l)}, Y_{j_2^Y(l)}, \ldots, Y_{j_{m_Y}^Y(l)}).$$

Resamplinga novērtējums Θ^*, kas ir vidējais aritmētiskais pēc r realizācijām, ko var aprēķināt pēc formulas (3.), aizvietojot $\phi(X(l))$ ar $\Psi(\mathbf{X}(l), \mathbf{Y}(l))$.

Šis novērtējums ir nenobīdīts. Tagad meklēsim šā novērtējuma dispersiju. Novērtējuma momentus un dispersiju var aprēķināt pēc formulām (4.-6.), aizvietojot funkciju $\phi(X(l))$ ar $\Psi(\mathbf{X}(l), \mathbf{Y}(l))$.

Lai novērtētu novērtējuma dispersiju, sākumā mums vajag atrast jauktā momenta μ_{11} izteikumu, kas ir sastopams formulā (5.). Lai aprēķinātu momentu μ_{11}, α-pāru notācija tiek ieviesta.

Apzīmēsim $W_i(l)$, l=1,...,r, $i \in \{X, Y\}$, izlases H_i apakškopu, kas bija izmantota, veidojot vektoru $\mathbf{X}(l)$ un $\mathbf{Y}(l)$ vērtības attiecīgi, $W_i(l) \subset H_i$. Apzīmēsim $M_i=\{0,1,...,m_i\}$, $M=M_X \times M_Y$. Lai $\boldsymbol{\alpha}=(\alpha_X, \alpha_Y)$ ir M kopas elements: $\boldsymbol{\alpha} \in M$. Mēs sakām, ka $W_i(l)$ un $W_i(l')$ veido $\boldsymbol{\alpha}$-pāri, tad un tikai tad, ja $W_i(l)$ un $W_i(l')$ ir α_i kopīgu elementu: $|W_i(l) \cap W_i(l)| = \alpha_i$.

Lai $A_{ll'}(\boldsymbol{\alpha})$ apzīmē notikumu "apakškopas ($\mathbf{X}(l)$,$\mathbf{Y}(l)$) un ($\mathbf{X}(l')$,$\mathbf{Y}(l')$) veido $\boldsymbol{\alpha}$-pāri", bet $P_{ll'}(\boldsymbol{\alpha})$ ir notikuma $P_{ll'}(\boldsymbol{\alpha})=P\{A_{ll'}(\boldsymbol{\alpha})\}$ varbūtība. Tā kā visas realizācijas l=1,..,r ir statistiski ekvivalentas, mēs izlaižam apakšējo indeksu ll' un rakstam $P(\boldsymbol{\alpha})$.

Lai

$$\mu_{11}(\boldsymbol{\alpha}) = E\big(\Psi(\mathbf{X}(l), \mathbf{Y}(l))\Psi(\mathbf{X}(l'), \mathbf{Y}(l')) \mid A_{ll'}(\boldsymbol{\alpha})\big), \tag{34.}$$

tad

$$\mu_{11} = \sum_{\boldsymbol{\alpha} \in M} P(\boldsymbol{\alpha})\mu_{11}(\boldsymbol{\alpha}). \tag{35.}$$

Varam secināt, ka mums sākumā jāaprēķina $P\{\boldsymbol{\alpha}\}$ un $\mu_{11}(\boldsymbol{\alpha})$ visiem $\boldsymbol{\alpha} \in M$. Varbūtība $P\{\boldsymbol{\alpha}\}$ varētu būt aprēķināta pēc hiperģeometriskā sadalījuma**.** Formulas momentiem $\mu_{11}(\boldsymbol{\alpha})$, $\forall \boldsymbol{\alpha} \in M$ tika iegūtas.

Pieejas efektivitātes skaitliskā analīze. Piemērs: Normālais sadalījums

Apskatīsim gadījumu, kad GL X un Y ir normālais sadalījums ar parametriem $\mu_X = \mu_Y = 2$, $\sigma_X = \sigma_Y = 1$. Lai mūsu izlašu apjomi ir $n = n_X = n_Y$. Mēs apskatam jau minēto varbūtību m-tās preces vienības pieprasījuma momentā, atkarībā no krājuma sākotnēja līmeņa K=0..3. Visi aprēķini ir izpildīti ar r = 1000 realizācijām.

Mēs gribam salīdzināt resamplinga novērtējuma dispersiju ar klasisko vidējo kvadrātisko kļūdu. Resamplinga pieejai dispersija un vidējā kvadrātiskā kļūda sakrīt, jo novērtējums ir nenobīdīts, atšķirībā no klasiskās pieejas.

7. tabulā ir dati, lai salīdzinātu resamplinga novērtējumu dispersiju $Var\,\Theta^*$ ar klasisko dispersiju $Var\,\hat{\Theta}$, nobīdi $Bias\,\hat{\Theta}$ un vidējo kvadrātisko kļūdu $MSE\,\hat{\Theta}$. 7. tabulā ir parādīta rezultātu izmaiņas tendence, atkarībā no dažādiem sākuma izlašu n, pieprasītās preces vienības numura m un krājuma sākotnēja līmeņa K.

7. tabula

Klasisko $\hat{\Theta}$ un resamplinga Θ^* novērtējumu eksperimentālie rezultāti

		K=0	K=1	K=2	K=3
n=10 m=5	$Var\,\hat{\Theta}$	.061	.043	.015	.002
	$Bias\,\hat{\Theta}$	0	.028	.029	.013
	$MSE\,\hat{\Theta}$	.061	.044	.015	.002
	$Var\,\Theta^*$	.087	.055	.014	.001
n=10 m=4	$Var\,\hat{\Theta}$	.053	.032	.006	---
	$Bias\,\hat{\Theta}$	0	.021	.018	---
	$MSE\,\hat{\Theta}$	.053	.033	.007	---
	$Var\,\Theta^*$	.069	.039	.005	---
n=12 m=6	$Var\,\hat{\Theta}$	.06	.045	.019	.004
	$Bias\,\hat{\Theta}$	0	.028	.033	.019
	$MSE\,\hat{\Theta}$	.06	.046	.02	.004
	$Var\,\Theta^*$	.085	.058	.02	.002

Analizējot tabulas rezultātus, varam secināt, ka abu pieeju dispersija un attiecīga vidējā kvadrātiskā kļūda samazinās ar izlašu apjomu n, m, un krājuma sākotnēja līmeņa K palielināšanos. Resamplinga novērtējumu dispersija gandrīz vienmēr ir tuvu tradicionālai. Tomēr resamplinga novērtējumi ir nenobīdīti. Ņemot vidējo kvadrātisko kļūdu, kā novērtējumu efektivitātes kritēriju, resamplinga dod labākus rezultātus, īpaši lielām K vērtībām.

Tiek apskatīts gadījums, kad GL X un Y ir eksponenciālais sadalījums ar parametriem λ=0.3, ν=0.7. Atkal ir pielietotas divas pieejas, lai kā iepriekšējā piemērā novērtētu to pašu varbūtību. Efektivitātes kritēriji arī bija tādi paši. Resamplinga pieejai ir labi rezultāti, dažreiz labāki, nekā tradicionālai pieejai.

Piemērs no loģistikas jomas

Apskatīsim vidējā ienākuma no ienākošām preču vienībām iegūšanas piemēru ar sekojošiem datiem:

- K- krājuma sākotnējais līmenis;
- $f(k)$ – sākuma krājuma izmaksas funkcija (piemēram, $f(k)=b_0+b_1\cdot k$);
- c_d – ienākums, ja viens pieprasījums ir apmierināts;
- c_s –sods, par savlaicīgu pieprasījuma nepiegādi.

Vidējais ienākums no m preču vienību pieprasījuma ir sekojošais:

$$\Pi(m,K)=m\cdot c_d-\left[f(K)+c_s\cdot\sum_{i=1}^{m}P_i(K)\right], \qquad (36.)$$

kur $P_i(K)$ - varbūtība, ka ar krājuma sākotnējo līmeni K i-tās preču vienības pieprasījuma momentā deficīta nebūs.

Resamplinga pieeja dod nenobīdītus novērtējamās funkcijas novērtējumus, tāpēc šī novērtējuma matemātiskā cerība sakrīt ar šīs funkcijas īsto vērtību. Agrāk iegūtās deficīta trūkuma varbūtības i-tās preču vienības pieprasījuma brīdī: 1-$P_i(K)$.

Vidējais zaudējums no m preču vienību apmierinājuma procesā ir:

$$\Xi(m,K)=f(K)+c_s\cdot\sum_{i=1}^{m}P_i(K). \qquad (37.)$$

Tādējādi: $\Pi(m,K)=m\cdot c_d-\Xi(m,K)$. Lai atrastu optimālo krājuma sākotnējo līmeni vajag palielināt (maksimizēt) ienākumus vai samazināt (minimizēt) zaudējumus:

$$\max_K \Pi(m,K) = m \cdot c_d - \Xi(m,K) \text{ vai } \min_K \Pi(m,K) = f(K) + c_s \cdot \sum_{i=1}^{m} P_i(K).$$

Analizēsim šo funkciju izmaiņas no K ar fiksētiem pārējiem parametriem $c_d=2$, $c_s=5$, $b_0=0$, $b_1=0.2$, ar normālo pieprasījumu $N(2,1)$, un normālo piedāvājumu $N(2.5,0.2)$.

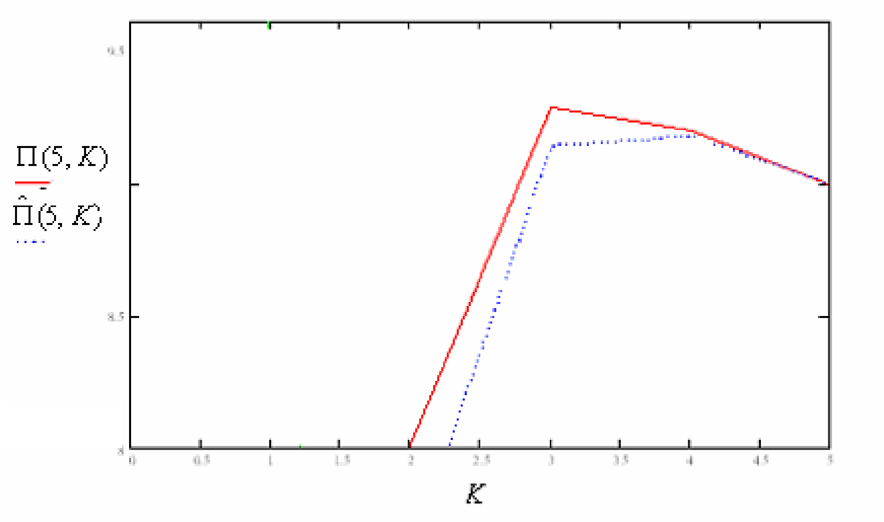
Skaidrs, ka reālajā situācijā pieprasījuma un piedāvājuma sadalījumi nav zināmi. Tomēr mēs varam tos novērtēt, balstoties uz pieejamām izlasēm, kā arī novērtēt mērķa deficīta varbūtību un vidējos ienākumus no apmierināta pieprasījuma. Šajā nodaļā bija aprakstītas tradicionāla "plug-in" un resamplinga metodes. Bija atzīmēts, ka ar maziem ieejas izlašu apjomiem labāk izmantot resamplinga pieeju, jo tā dod nenobīdītus parametru novērtējumus ar mazāku vidējo kvadrātisko kļūdu, nekā tradicionālai pieejai. Parādīsim kāda ir klasiskās pieejas nobīde, salīdzinot ar īstām vērtībām, aprēķinot vidējos ienākumus. 8. tabulā ir apkopoti sekojušu funkciju rezultāti: Π- reālā ienākumu vērtība un $\hat{\Pi}$ to klasiskais novērtējums ar krājuma sākotnēja līmeņa K un laika (m-tās preču vienības pieprasījuma laika brīdis) izmaiņām.

8. tabula

Vidējo ienākumu novērtējuma nobīde no reālas vērtības

		K=0	K=1	K=2	K=3	K=4
m=1	$\Pi(1,K)$	-2.816	1.8	1.6	1.4	1.2
	$\hat{\Pi}(1,K)$	-2.872	1.8	1.6	1.4	1.2
m=3	$\Pi(3,K)$	-9.723	2.781	5.444	5.4	5.2
	$\hat{\Pi}(3,K)$	-11.197	2.709	5.375	5.4	5.2
m=5	$\Pi(5,K)$	-17.624	0.462	8.02	9.285	9.197
	$\hat{\Pi}(5,K)$	-19.9	0.34	7.604	9.144	9.18
m=7	$\Pi(7,K)$	-26.139	-4.109	8.532	12.572	13.129
	$\hat{\Pi}(7,K)$	-28.814	-3.995	7.653	11.954	12.945

Šie rezultāti liecina, ka meklējot optimālo krājuma sākotnējo apjomu, mēs varam to noteikt nepareizi, klasiskā novērtējuma nobīdes dēļ. 5. att. ir parādīts, ka 5. preču vienības pieprasījuma momentā krājuma sākotnēja līmeņa apjoms bija 3 vienības. Tajā pašā laikā, klasiskais šā parametra novērtējums parādīja optimālo krājuma sākotnējo preču vienību skaitu, kas vienāds ar 4 vienībām. Tā kā resamplinga pieeja dod nenobīdītus parametru novērtējumus, tas ļauj izvairīties no klasisko novērtējumu atzīmētiem trūkumiem.



5. att. Vidējo ienākumu optimālā vērtība atkarībā no krājuma sākotnēja līmeņa *K*

Secinājumi

Resamplinga pieeju var sekmīgi pielietot, lai iegūtu atjaunošanas procesa parametru novērtējumus. Iegūtās formulas ļauj aprēķināt resamplinga un klasiskā novērtējumu pieeju dispersijas. Skaitliskie piemēri rāda piedāvātās pieejas efektivitāti, kā kritēriju ņemot vidējo kvadrātisko kļūdu. Resamplinga pieeja var kļūt par labu alternatīvu tradicionālai pieejai, kā arī to var pielietot augsti drošo un dārgo krājumu pārvaldei.

Nobeigums

1. Darbs ir veltīts transporta loģistiko procesu organizācijas un pārvaldes pilnveidošanas jautājumiem, balstoties uz intensīvu datormetožu izmantošanu statistikā. Tās aktualitātes iemesls ir tas, ka jebkura procesa adekvātais apraksts ir nepieciešams nosacījums, lai efektīvi organizētu pārvaldi.
2. Resamplingam, kas ir viena no galvenajām intensīvajām datormetodēm statistikā, attiecībā uz tradicionālām metodēm ir sekojošas priekšrocības: tā ir neparametriska, tāpēc tas ir brīvs no kļūdām gadījumu lielumu sadalījuma likuma izvēlē; tas balstās uz tiešas apskatīta procesa imitācijas, tāpēc var būt pielietots situācijās, kad procesa analītiskais apraksts un analīze ir neiespējamas.
3. Resamplinga praktiskais pielietojums atšķiras ar sevišķu vienkāršību, tātad darba galvenais mērķis bija pētīt tā efektivitāti uzdevumiem, kas ir tipiski transportam un loģistikai. Tās veido darba teorētisko daļu. Analizēt resamplinga pieejas efektivitāti, salīdzinot ar tradicionālām pieejām uz matemātiskās statistikas sekojošo uzdevumu piemēra:
   - ❑ daudzdimensiju lineāras regresijas parametru novērtējumu robustums;
   - ❑ masveida apkalpošanas sistēmas ar neaprobežotu serveru skaitu, kas plaši pielietota drošības teorijā un apdrošināšanas efektivitātes rādītāju novērtēšana;
   - ❑ divu dažādu atjaunošanas procesu salīdzināšana.

   Analītiskie pētījumi un eksperimentu rezultāti liecina par to, ka resamplinga pieeja ir labāka, ja ir mazs izlašu apjoms.
4. Resamplinga pieejas lietderība ir ilustrēta ar sekojošiem praktiskiem piemēriem no transporta un loģistikas jomas.
   - ❑ Eiropas Savienības valstu aviācijas transporta pieprasījumu prognozēšana, kā funkcija no tādiem faktoriem, kā teritorijas plašums, iedzīvotāju skaits iekšējais valsts kopprodukts uz vienu iedzīvotāju, vidēja viena iedzīvotāja mēnešalga.
   - ❑ Preču ar bojājumu attīstīšanos un uzkrāšanu, drošības prognozēšana.
   - ❑ Optimāla krājuma sākotnēja līmeņa aprēķināšana, ar gadījuma pieprasījumiem un piedāvājumiem.

5. Saņemtie teorētiskie rezultāti ir jauni un tie nekad iepriekš neparādījās literatūrā. Pasažieru pārvadājumu ar aviācijas transportu pieprasījuma prognozēšanas metodika un rezultāti bija pieprasīti atskaites iekļaušanai 2. Izglītības un zinātnes ministrijas projektu: Matemātisko metožu izstrādāšana un novērtēšana pasažieru un krāvu plūsmu prognozēšanai Baltijas reģionā.
6. Darba rezultāti ir referēti deviņās Starptautiskajās zinātniskajās konferencēs un semināros un publicēti desmitos rakstos. Četrās publikācijās pretendente ir vienīgā autore.

Publikācijas, kurās autore piedalījusies

1. Afanasyeva H. *Statistical Analysis of Air Traffic in Latvian Region* //In proceedings of The Second International Conference Simulation, Gaming, Training and Business Process Reengineering in Operations. -Riga: RTU, 2000, pp. 125-129.
2. Afanasyeva H. *The Resampling-estimator of Queuing Length Nonstationary Distribution for the Queuing System M/G/∞.* //Transport and Telecommunication, Vol. 3. N1, - Riga: TTI, 2002, pp. 89-94.
3. Андронов А.М., Афанасьева Е.Н. *Компьютеризация преподавания дисциплин прикладной математики*, Transport and Telecommunication, - Riga, TTI, 2004. pp. 5-8.
4. Andronov A., Afanasyeva H. *Resampling Based Non-Parametric Statistical Inferences about Distribution, Moments and Quantiles of Order Statistics*. //In transactions of *The XX1V International Seminar on Stability Problems for Stochastic Models,* -Riga: TTI, 2004, pp. 300-307.
5. Afanasyeva H. *Resampling Median Estimators for Linear Regression Model*. //In Proceedings of RelStat'04 Transport and Telecommunication, -Riga: TTI, 2005, pp. 90-94.
6. Afanasyeva H., Andronov A. *On robustness of resampling estimators for linear regression models.* //In proceedings of the International Symposium on Stochastic Models in Reliability, Safety and Logistics, Beer Sheva, Israel, 2005, pp. 6-11.
7. Afanasyeva H. *Resampling-approach to a task of comparison of two renewal processes*. //In proceedings of the 12th International Conference on Analytical and Stochastic Modelling Techniques and Applications, -Riga: RTU, 2005, pp. 94-100.
8. Afanasyeva H. *A task of the storage control theory in transport systems using resampling-method.* //In proceedings

of the 5-th International Conference “Transport Systems Telematics”, Katowice-Ustron, Poland, 2005, pp. 13-21.
9. Andronov A., Afanasyeva H. and Fioshin M. *Statistical Estimation for a Failure Model with the Accumulation of Damages*. //In proceedings of The international Conference on Degradation, Damage, Fatigue and Accelerated Life Models in Reliability Testing, Angers, France, 2006, pp. 75-81.
10. Afanasyeva H., Andronov A. *On robustness of resampling estimators for linear regression models.* //Communications in Dependability and Quality Management: An international Journal, Vol. 9, No. 1, 2006, pp. 5-11.