

LATVIJAS UNIVERSITĀTE  
FIZIKAS, MATEMĀTIKAS UN OPTOMETRIJAS FAKULTĀTE  
MATEMĀTIKAS NODAĻA

**OPERACIONĀLĀ RISKĀ ZAUDĒJUMU  
STATISTISKĀ MODELĒŠANA, IZMANTOJOT  
BEIJESA STATISTIKAS METODES**

MATEMĀTIKA STATISTIKA PROGRAMMAS  
BAKALaura DARBS

Autore: Linda Mičule

Studentes apliecības nr. Im18052

Darba vadītājs: pētniece Dr. math. Māra Delesa-Vēliņa

RĪGA 2022

## Anotācija

Bakalaura darbā tiek apskatīta operacionālā riska zaudējumu statistiskā modelēšana, izmantojot Beijesa metodi. Darbā tiek aprakstīta standarta zaudējumu sadalījuma pieceja, Beijesa metode, saistītie apriorie sadalījumi, kā arī riskam pakļautās vērtības novērtēšana gan ar Montekarlo simulāciju, gan ar aproksimācijas formulu palīdzību. Praktiskajā daļā tiek pielietota Beijesa metode, lai novērtētu parametrus simulētiem datiem trīs darbā minētajiem saistītajiem sadalījumiem. Kā arī tiek aplūkots datu piemērs, kur apriorie parametri tiek novērtēti ar Beijesa metodi, un gan ar Montekarlo simulāciju, gan aproksimācijas formulu palīdzību tiek aprēķināta riskam pakļautā vērtība. Rezultāti tiek salīdzināti ar maksimālā ticamības intervāla parametru novērtējumiem.

Visi aprēķini tiek veikti programmā *Rstudio*.

Atslēgvārdi: operacionālais risks, zaudējumu sadalījuma pieceja, Beijesa metode, riskam pakļautā vērtība, Montekarlo simulācijas.

## Abstract

The bachelor thesis deals with statistical modelling of operational risk losses using the Bayesian inference. The thesis describes the standard loss distribution approach, the Bayesian inference, conjugate prior distributions and the estimation of the value-at-risk using both Monte Carlo simulations and approximation formulas. In the practical part, the Bayesian method is applied to the simulated data to estimate the three conjugate prior distributions mentioned in the paper. An example of data is also considered where the prior parameters are estimated by the Bayesian method and the value-at-risk is calculated by both Monte Carlo simulations and approximation formulas. The results are compared with the maximum likelihood parameter estimates.

All calculations are done in *Rstudio*.

Key words: operational risk, Loss Distribution Approach, Bayesian inference, Value-at-Risk, Monte Carlo Simulation.

# Satura rādītājs

Apzīmējumu saraksts	4
Ievads	5
<b>1. Standarta zaudējumu sadalījuma pieceja</b>	<b>7</b>
1.1. Zaudējumu biežuma sadalījumi . . . . .	8
1.2. Zaudējumu smaguma sadalījumi . . . . .	9
1.3. Analītiskās formas VaR tuvinājums . . . . .	11
<b>2. Beijesa metode</b>	<b>13</b>
<b>3. Saistītie apriorie sadalījumi</b>	<b>16</b>
3.1. Puasona– gamma . . . . .	16
3.2. Lognormālais– normālais . . . . .	18
3.3. Pareto– gamma . . . . .	19
3.4. Parametru ierobežošana . . . . .	21
<b>4. Strukturālo aprioro parametru subjektīvais novērtējums</b>	<b>22</b>
4.1. Puasona– gamma parametru novērtējums . . . . .	22
4.2. Lognormālā–normālā sadalījumu parametru novērtējums . . . . .	24
4.3. Pareto– gamma sadalījumu parametru novērtēšana . . . . .	27
<b>5. Kapitāla aprēķini, izmantojot Montekarlo simulācijas</b>	<b>30</b>
5.1. Montekarlo procedūras soļi . . . . .	30
5.2. VaR aprēķina piemērs . . . . .	31
<b>Secinājumi</b>	<b>41</b>
<b>Izmantotā literatūra un avoti</b>	<b>42</b>
<b>1. Pielikums. Programmas <i>Rstudio</i> kods</b>	<b>43</b>
	<b>53</b>

## Apzīmējumu saraksts

LDA – zaudējumu sadalījuma pieeja (angl. *loss distribution approach*),

$I_{[a,b]}(x)$  – indikātorfunkcija. Ja  $x \in [a, b]$ , tad  $I = 1$ , pretējā gadījumā  $I = 0$ ,

$\times$  – reizinājuma zīme,

$\propto$  – proporcionālītātes zīme,

$o(1)$  – funkcija, kas tiecas uz 0.

## Ievads

Operacionālais risks ir iespēja ciest zaudējumus prasībām neatbilstošu vai nepilnīgu iekšējo procesu norises, cilvēku un sistēmu darbības vai arī ārējo apstākļu ietekmes dēļ, ieskaitot juridisko risku, bet neieskaitot stratēģisko un reputācijas risku [1]. To veido šādi notikumu veidi

1. iekšējie krāpnieciskie darījumi,
2. ārējie krāpnieciskie darījumi,
3. neatbilstoša nodarbinātības prakse un darbavietas drošība,
4. nekorekta attieksme pret klientiem, neatbilstoši produkti un komercdarbības prakse,
5. materiālo aktīvu bojājumi,
6. komercdarbības pārrāvumi un sistēmu kļūdas (defekti),
7. nepilnības izpildes, piegādes un procesa vadībā.

Bāzeles komiteja 2007. gadā pieņēma lēmumu par jaunajām kapitāla prasībām. Papildus tirgus un kredītriskam, bankām ir jāatspoguļo arī operacionālais risks. Tas nozīmē, ka ir jāizvērtina, cik kapitāla ir nepieciešams, lai segtu operacionālo risku, izmantojot vienu no trīs pieejām :

- pamatrādītāju pieeja,
- standartizētā pieeja,
- attīstītās mērīšanas pieeja (*angl. - advanced measurement approach, AMA*).

Šīs pieejas parāda dažādus riska jūtīguma līmeņus. Atbilstoši pamatrādītāja pieejai pašu kapitāla prasība attiecībā uz operacionālo risku ir 15% no bankas trīs pēdējo gadu vidējiem bruto ieņēmumiem. Standartizētā pieeja atšķiras ar to, ka tiek izmantoti dažādi procentuāli koeficienti ienākumiem no dažādām bankas darbības jomām. AMA, kā viselastīgākā pieeja operacionālā riska kvantitatīvai noteikšanai, ļauj bankai izveidot savu iekšējo operacionālā riska modeli un mērīšanas sistēmu, kas ir pielīdzināma tirgus riska standartiem.

Visatbilstošākie standarti šajā darbā aplūkotajiem jautājumiem ir tādi, ka operacionālā riska mērs ir  $VaR$  ar ticamības līmeni 99.9% viena gada laikā un ka mēra aprēķināšanas pieejai jāietver potenciāli nopietni zaudējumu astes notikumi.

Vēsturiskie iekšējie operacionālā riska zaudējumu dati ir ierobežotas spējas prognozēt turpmāko rīcību, turklāt bankām nav pietiekami daudz iekšējo datu, lai korekti novērtētu zema biežuma notikumus ar lielu ietekmi. Vēsturiskie ārējie dati ir grūti izmantojami dažādu apjomu un citu faktoru dēļ. Turklāt iekšējiem un ārējiem datiem ir izdzīvošanas novirze, jo parasti nav datu par visiem bankrotējošajiem uzņēmumiem.

Šajā darbā tiks apskatītas divas metodes operacionālā riska aprēķinos : standarta zaudējumu sadalījuma pieeja, kuras pamatā ir operacionālo zaudējumu varbūtību sadalījuma modelēšana, izmantojot bankas iekšējos un ārējos datus, un Beijesa secinājumu metode, kas ir statistisks metode, kura ir labi piemērota ekspertu viedokļu un vēsturisko datu apvienošanai.

**Darba mērķis** ir izpētīt un pielietot Beijesa metodi operacionālā riska zaudējumu biežuma un smaguma parametru novērtēšanā.

**Darba uzdevumi:**

1. iepazīties ar standarta zaudējumu sadalījuma pieeju un Beijesa metodi,
2. nostiprināt iegūtās zināšanas, veicot praktiskus uzdevumus,
3. salīdzināt Beijesa metodes novērtētos parametrus ar maksimālās ticamības novērtējumiem,
4. veikt Montekarlo simulācijas, lai varētu novērtēt riskam pakļauto vērtību, kā arī izmantot aproksimācijas formulas  $VaR$  novērtēšanā,
5. apkopot iegūtos rezultātus, veikt secinājumus.

Darbs sastāv no 5 nodaļām, kur pirmā nodaļa ir veltīta standarta zaudējumu sadalījuma pieejai, otrā Beijesa metodei, trešajā tiek apskatīti saistītie apriorie sadalījumi, ceturtajā strukturālo aprioro parametru subjektīvais novērtējums un piektajā kapitāla aprēķini, izmantojot Montekarlo simulācijas.

# 1. Standarta zaudējumu sadalījuma pieeja

Pārdomāta un profesionāla operacionālā riska pārvaldība sniedz uzņēmumam iespēju izvairīties no iespējamiem zaudējumiem, tos samazināt un būt gataviem izsvērtai reakcijai krīzes situācijās. Viens no šādiem pārvaldības modeļiem ir zaudējumu sadalījuma pieeja jeb LDA. Tā sastāv no kopējo zaudējumu summas sadalīšanas zaudējumu smaguma komponentēs, t.i., individuālajās zaudējumu summās, un biežuma komponentēs, t.i., zaudējumu skaitā. Pēc tam kopējos zaudējumus iegūst, apvienojot šo informāciju.

## 1.1. definīcija. Standarta LDA [4]

### 1. Zaudējumu smaguma process

Zaudējumu smagumi  $(X_k)_{k \in \mathbb{N}}$  ir pozitīvi, neatkarīgi un vienādi sadalīti gadījuma lielumi, kas raksturo katra zaudējuma lielumu.

### 2. Zaudējumu biežuma process:

Zaudējumu notikumu skaits  $N(t)$  laika intervālā  $[0, t]$  ir gadījuma lielums,  $t \geq 0$ . Rezultējošais skaitīšanas process  $N(t)_{t \geq 0}$  ir ģenerēts, izmantojot nenegatīvu gadījuma lielumu virki  $(T_n)_{n \geq 1}$ , kura apmierina

$$0 \leq T_1 \leq T_2 \leq \dots$$

un

$$N(t) = \sup\{n \geq 1 : T_n \leq t\}, t \geq 0.$$

### 3. Tiek pieņemts, ka zaudējumu smaguma un biežuma procesi ir neatkarīgi.

### 4. Kopējais zaudējumu process ir

$$Z(t) = \sum_{i=1}^{N(t)} X_i, t \geq 0. \quad (1.1)$$

Ievēro, ka netiek pieprasīts, lai  $X_i$  būtu galīga vidējā vērtība un/vai dispersija.

## 1.2. definīcija. (Riskam pakļautā vērtība VaR [4])

Tiek pieņemts, ka  $G_t$  ir kopējais zaudējumu sadalījums. Tad riska vērtība VaR laika periodā  $t$  pie ticamības līmeņa  $\kappa$  ir definēta kā  $\kappa$  – kvantile zaudējumu sadalījumam:

$$\text{VaR}_t(\kappa) = G_t^{-1}(\kappa), \quad \kappa \in (0, 1),$$

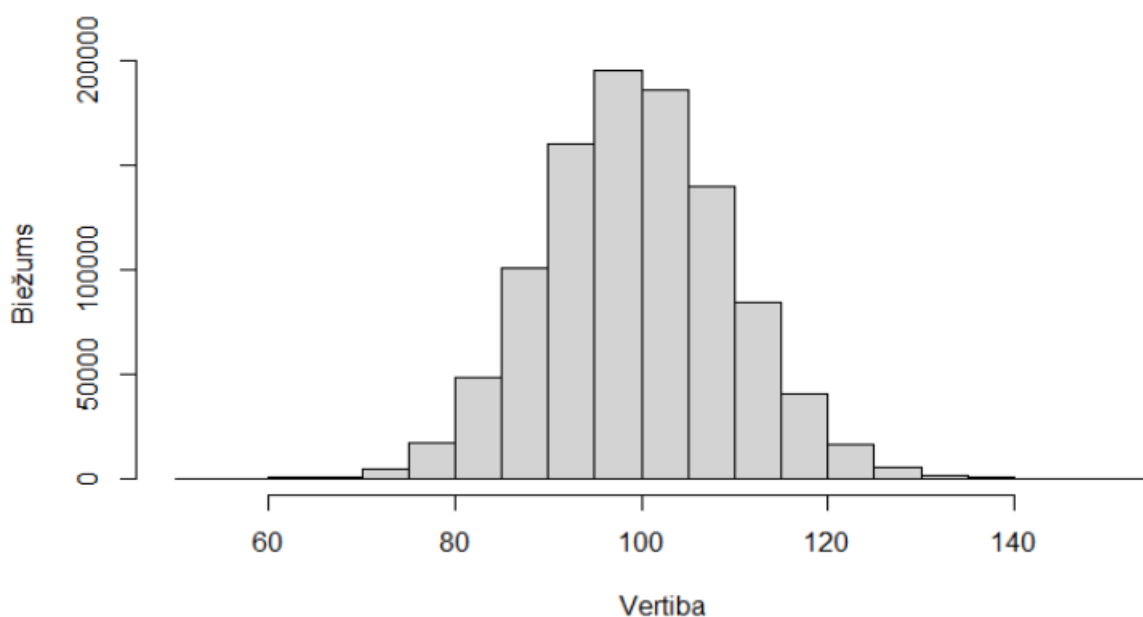
kur  $G_t^{-1}(\kappa) = \inf\{x \in \mathbb{R} : G_t(x) \geq \kappa\}$ ,  $0 < \kappa < 1$ . Pieņem, ka  $G_t$  ir stingri pieaugošs un nepārtraukts.

## 1.1. Zaudējumu biežuma sadalījumi

Šajā apakšnodaļā tiks aplūkoti standarta LDA zaudējumu biežuma sadalījumi, pēc kuriem visbiežāk ir sadalīts zaudējumu biežums.

**1. piemērs.** *Puasona LDA.* Šeit  $(N(t))_{t \geq 0}$  ir homogēns Puasona process ar intensitāti  $\lambda > 0$  un blīvuma funkciju

$$P(N(t) = n) = p_t(n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n \in \mathbb{N}.$$

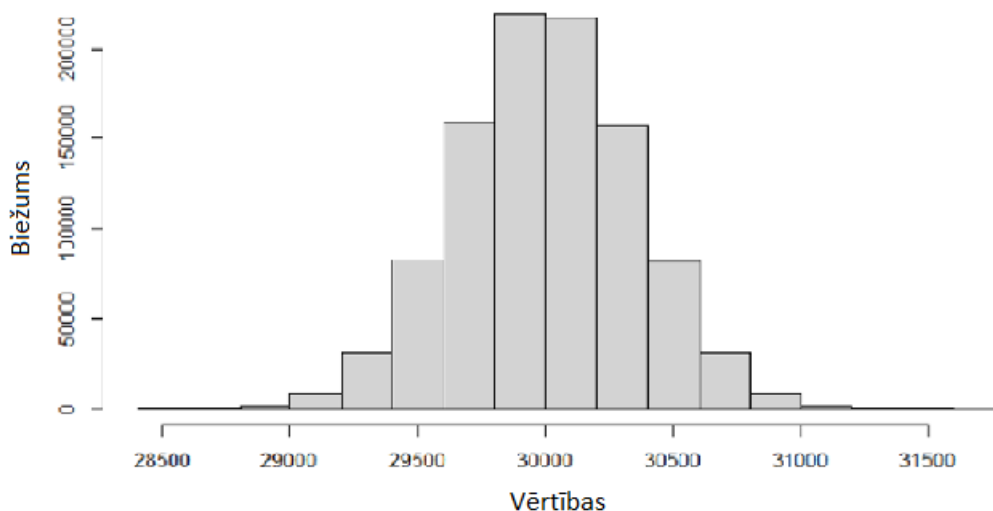


1. att.: Histogramma Puasona sadalījumam ar  $\lambda = 100$  simulētiem datiem (simulāciju skaits  $B = 10^6$ )

**2. piemērs.** *Negatīvais binomiālais LDA.* Šeit  $(N(t))_{t \geq 0}$  ir negatīvais binomiālais process ar parametriem  $\beta, \gamma > 0$ , kur blīvuma funkcija ir

$$P(N(t) = n) = p_t(n) = \binom{\gamma + n - 1}{n} \left(\frac{\beta}{\beta + t}\right)^\gamma \left(\frac{t}{\beta + t}\right)^n, \quad n \in \mathbb{N}.$$

Negatīvais binomiālais sadalījums ir gamma un Puasona sadalījumu apvienojums, t.i., tas var būt skatīts kā Puasona sadalījums, kura parametrs  $\lambda$  ir gamma sadalījuma gadījuma vērtība.



2. att.: Histogramma negatīvajam binomiālajam sadalījumam, kur  $\gamma = 10000$  un  $\beta = 0.25$  (simulāciju skaits  $B = 10^6$ )

## 1.2. Zaudējumu smaguma sadalījumi

Sākumā tiks aplūkota subeksponenciālo sadalījumu definīcija, kas ir speciāla sadalījumu klase ar tā sauktajām smagajām astēm.

### 1.3. definīcija. Subeksponenciālais sadalījums [7]

$(X_k)_{k \in \mathbb{N}}$  ir neatkarīgi un vienādi sadalīti gadījuma lielumi ar sadalījuma funkciju  $F(x)$ . Tiks apzīmēts  $\bar{F}(x) = 1 - F(x)$ ,  $x \geq 0$ , kas ir  $F$  aste un  $\bar{F}^{n*}(x) = 1 - F^{n*}(x) = P(X_1 + \dots + X_n > x)$ , kas ir  $n$ -kārtīga konvolūcijas aste. Tad  $F(x)$  ar suportu  $(0, \infty)$  ir subeksponenciālā sadalījuma funkcija, ja

$$\lim_{x \rightarrow \infty} \frac{\bar{F}^{n*}(x)}{\bar{F}(x)} = \lim_{x \rightarrow \infty} \frac{P(X_1 + \dots + X_n > x)}{P(X_1 > x)} = n, \quad \forall n \geq 2.$$

Šo sadalījumu īpašība ir tāda, ka  $n$  subeksponenciālo gadījuma lielumu summas astei ir tāda pati varbūtība kā to maksimālās vērtības astei, t.i.,

$$\lim_{x \rightarrow \infty} \frac{P(X_1 + \dots + X_n > x)}{P(\max(X_1, \dots, X_n) > x)} = 1 \quad \forall n \geq 2.$$

Tas nozīmē, ka  $n$  neatkarīgu gadījuma lielumu smagumu summa, visticamāk, būs liela, jo viens no vienādojuma locekļiem ir liels, vai, uzsverot operacionālo risku, nopietni kopējie zaudējumi galvenokārt ir saistīti ar vienu lielu zaudējumu, nevis uzkrāto nelielu, neatkarīgu zaudējumu summu. Šim uzskatam vajadzētu ietekmēt operacionālo risku pārvaldību.

**Biežāk lietotie zaudējumu smaguma sadalījumi no subeksponenciālās klases ar suportu**  
(0, ∞) [4]

Nosaukums	Sadalījuma funkcija	Parametri
Lognormālais sad.	$F(x) = \Phi\left(\frac{\ln(x)-\mu}{\sigma}\right)$	$\mu \in \mathbb{R}, \sigma > 0$
Veibulla	$F(x) = 1 - e^{-(x/\theta)^\tau}$	$\theta > 0, 0 < \tau < 1$
Pareto	$F(x) = 1 - \left(1 + \frac{x}{\theta}\right)^{-\alpha}$	$\alpha, \theta > 0$

Katra LDA modeļa mērķis ir noteikt kopējo zaudējumu sadalījumu, ko standarta LDA gadījumā var raksīt kā

$$\begin{aligned} G_t(x) &= P(S(t) \leq x) \\ &= \sum_{n=0}^{\infty} p_t(n) P(S(t) \leq x | N(t) = n) \\ &= \sum_{n=0}^{\infty} p_t(n) F^{n*}(x), \quad x \geq 0, \quad t \geq 0, \end{aligned}$$

kur  $F(x) = P(X_k \leq x)$  ir  $X_k$  sadalījuma funkcija, un  $F^{n*}(x) = P(\sum_{i=1}^n X_i \leq x)$  ir  $n$ -reizes  $F$  konvolūcija ar  $F^{1*} = F$  un  $F^{0*} = I_{[0, \infty)}$ .

Lielākajai daļai zaudējuma smagumu un biežuma sadalījumu  $G_t$  funkciju nevar aprēķināt analītiski. Tuvināšanas metodes šīs problēmas atrisināšanai piedāvā Panjera rekursijas, Monte-karlo simulācijas un FFT (ātrās Furjē transformācijas) metodes. Šo metožu trūkums ir tāds, ka rezultāts atgriež melno kasti, un mijiedarbība starp dažādiem modeļa parametriem un to ietekme uz gala rezultātu, t.i.,  $Var$ , ir interpretējama tikai ar plašu jutīguma analīzi.

Tā kā gan pamatkapitāls, gan ekonomiskais kapitāls ir balstīts uz ļoti lielu kopējā zaudējumu sadalījuma  $G_t$  kvantili, dabiska  $Var$  novērtēšanas metode ir balstīta uz asimptotiskās astes un kvantiles novērtēšanu. Tā vietā, lai ņemtu vērā visu sadalījumu, ļoti lielam  $X$  ir jākoncentrējas uz labās puses asti  $P(Z(t) > x)$ .

Pēdējā gadījumā iegūtais rezultāts norāda, ka standarta LDA ar subeksponenciāliem zaudējumu smagumiem, ievērojot vājas regularitātes nosacījumus, katram fiksētam  $t \geq 0$

$$\bar{G}_t(x) \sim EN(t)\bar{F}(x), \quad x \rightarrow \infty, \quad (1.2)$$

kur  $EN(t)$  ir sagaidāmais zaudējumu biežums,  $\bar{F}(x) = 1 - F(x)$  un  $\bar{G}_t(x) = 1 - G_t(x)$  ir kopējo zaudējumu un smaguma astes sadalījumi. Simbols  $\sim$  nozīmē to, ka labās un kreisās

puses dalījums tiecas uz 1, t.i.,  $\lim_{x \rightarrow \infty} \bar{G}_t(x)/\bar{F}(x) = EN(t)$  visiem fiksētiem  $t > 0$ .

### 1.3. Analītiskās formas VaR tuvinājums

Ņemot vērā sakarību (1.2), var iegūt *VaR* izteiksmi, kas ir pielietojama pie ļoti augsta ticamības līmeņa.

Izmantojot (1.2), Klauss Bokers un Klaudija Kluppelberga piedāvā asimtotisku VaR formulu [4], kura tiks apskatīta nākamajā teorēmā.

#### 1.1. teorēma. Analītiskā VaR [4]

Tiek apskatīts standarta LDA modelis fiksētam  $t > 0$  un ar subeksponenciālu zaudējuma smagumu ar sadalījuma funkciju  $F$ . Pieņem, ka astes novērtējums (1.2) ir spēkā. Tad

$$\text{VaR}_t(\kappa) = F^{-1} \left( 1 - \frac{1 - \kappa}{EN(t)} (1 + o(1)) \right), \quad \kappa \rightarrow 1. \quad (1.3)$$

Šis rezultāts ir nozīmīgs divu iemeslu dēļ. Pirmkārt, tas norāda, ka *VaR* pie augsta ticamības līmeņa ir atkarīgs tikai no astes, nevis no visa zaudējumu smaguma sadalījuma. Tādēļ, ja interesē tikai *VaR* aprēķināšana, tad visas  $F$  sadalījuma funkcijas modelēšana ir lieka. Otrkārt, tā kā biežums izteiksmē (1.3) tiek iekļauts tikai ar  $EN(t)$ , arī nav nepieciešams ģenerēt specifisku skaitīšanas procesu. Tā vietā pietiek ar zaudējumu biežuma vidējās vērtības novērtēšanu. Tā rezultātā pārmērīgajai izkliedei, ko modelē negatīvais binomiālais sadalījums, asimptotiski nav ietekmes uz *VaR*.

Lai iegūtu pirmās kārtas tuvinājumu *VaR* konkrētam LDA modelim, pietiek ar kombināciju (1.3) un subeksponenciālā svaru sadalījuma  $F$  asti. Turklāt, pat slēgtas formas atrisinājums asimtotiskam VaR ir pieejams (sk. 2. tabulu).

Problēma atrast zaudējumu smaguma sadalījumu, kas precīzi raksturo empīriskos zaudējumu datus, nav triviāls uzdevums, un atbilstošu smaguma un biežuma sadalījuma parametru noteikšana ir neatņemama katra LDA modeļa sastāvdaļa. Piemērotību ir iespējams pārbaudīt ar sadalījuma testiem. Zaudējumu smagumam tiek izmantots Kolmogorova–Smirnova tests, savukārt, zaudējumu biežumu sadalījuma piemērotību pārbauda  $H_1$  kvadrātā tests. Abiem testiem  $H_0$ : abas datu kopas ir no viena un tā paša nepārtrauktā sadalījuma. Hipotēze tiek noraidīta pie  $p$ -vērtības  $< 0.05$ .

**Pirmās kārtas  $\text{VaR}_t(\kappa)$  aproksimācija, kad  $\kappa \rightarrow 1$  kopējam zaudējuma sadalījumam plaši izmantotiem smaguma sadalījumiem. Šeit  $EN(t) = \lambda t$  Puasona sadalījumam un**

**$EN(t) = \gamma t / \beta$  negatīvajam binomiālajam sadalījumam [4]**

Nosaukums	$\text{VaR}_t(\kappa)$
Lognormālais	$\exp \left[ \mu - \sigma \Phi^{-1} \left( \frac{1-\kappa}{EN(t)} \right) \right]$
Veibulla	$\theta \left[ \ln \left( \frac{EN(t)}{1-\kappa} \right) \right]^{\frac{1}{\tau}}$
Pareto	$\theta \left[ \left( \frac{EN(t)}{1-\kappa} \right)^{1/\alpha} - 1 \right]$

## 2. Beijesa metode

Beijesa metodes nosaukums ir dots par godu Tomasam Beijesam (1607.- 1761.), kurš pierādīja, ka nezināmiem notikumiem var noteikt varbūtības robežas. Tomēr Pjērs Simons Laplass (1749.- 1827.) ieviesa to, ko tagad sauc par Beijesa teorēmu. Metodi izmanto, lai risinātu problēmas debesu mehānikā, medicīnas statistikā, ticamības noteikšanā un jurisprudencē. Sākotnējo Beijesa metodi, kura izmantoja aprioru, kas apmierināja Laplasa nepietiekama iemesla principu, sauca par inverso varbūtību, jo tā secina atpakaļgaitā no novērojumiem uz parametriem vai no sekām uz cēloņiem [6].

Astoņdesmitajos gados strauji pieauga Beijesa metodes pētniecība un pielietojums, galvenokārt pateicoties Markova ķēdes Montekarlo metodes atklāšanai, kas novērsa daudzas skaitļošanas problēmas, un pieaugošajai interesei par nestandarta, kompleksiem pielietojumiem. Neraugoties uz Beijesa metodes pētījumu pieaugumu, lielākajā daļā bakalaura līmeņa programmu joprojām tā netiek apskatīta [3]. Tomēr šī metode ir plaši atzīta un izmantota, piemēram, mašīnmācīšanās jomā.

Beijesa metode ir statistikas metode, kas ir labi piemērota, lai operacionālo datu analīzē iekļautu ekspertu viedokļus. Šī metode ļauj veikt strukturālo modelēšanu, kurā ekspertu viedokļi tiek iekļauti analīzē, norādot modeļa parametru sadalījumus (tā sauktos aprioros sadalījumus). Kad dati kļūst pieejami, tos atjaunina. Jebkurā brīdī eksperts var pārvērtēt iepriekšējos sadalījumus, ņemot vērā jaunas informācijas pieejamību (piemēram, kad tiek ieviesta jauna bankas politikas kontrole), kas šo informāciju iekļaus modelī.

Šīs nodaļas teorijas izklāsts ir veidots saskaņā ar [10].

Pieņem, ka  $\mathbf{X} = (X_1, X_2, \dots, X_n)$  ir gadījuma novērojumu vektors, kura blīvuma funkcija ir  $h(\mathbf{X}|\boldsymbol{\theta})$  pie dotā parametru vektora  $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_K)$ . Izmantojot šo metodi, gan novērojumi, gan parametri tiek uzskatīti par gadījuma lielumiem. Tad Beijesa teorēmu var formulēt kā

$$h(\mathbf{X}, \boldsymbol{\theta}) = h(\mathbf{X}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) = \hat{\pi}(\boldsymbol{\theta}|\mathbf{X})h(\mathbf{X}), \quad (2.1)$$

kur  $\pi(\boldsymbol{\theta})$  ir parametru blīvuma funkcija jeb apriorais sadalījums,  $\hat{\pi}(\boldsymbol{\theta}|\mathbf{X})$  ir parametru blīvuma funkcija, ņemot vērā novērojumu datus  $\mathbf{X}$  jeb aposterioro sadalījumu,  $h(\mathbf{X}, \boldsymbol{\theta})$  ir novēroto datu un parametru kopējā blīvuma funkcija,  $h(\mathbf{X}|\boldsymbol{\theta})$  ir novērojumu blīvuma funkcija, un  $h(\mathbf{X})$  ir  $\mathbf{X}$  marginālā blīvuma funkcija. Vienādojumu (2.1) arī var izteikt kā

$$h(\mathbf{X}) = \int h(\mathbf{X}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}.$$

Vienkāršības labad, tiks aplūkoti tikai nepārtraukti  $\pi(\boldsymbol{\theta})$ . Ja  $\pi(\boldsymbol{\theta})$  ir diskrēts sadalījums, tad integrēšana tiek mainīta uz summēšanu,  $h(\mathbf{X}) = \sum h(\mathbf{X}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$ .

Mērķis operacionālā riska kontekstā ir novērtēt nākotnes sadalījumus zaudējumu biežumam un smagumam nākotnes novērojumiem  $X_{n+1}$ , kas ir atkarīgi no visas pieejamās informācijas  $\mathbf{X} = (X_1, X_2, \dots, X_n)$ . Pieņem, ka gadījuma vektors  $\mathbf{X}$  un  $X_{k+1}$ , nosacīti uz parametru  $\boldsymbol{\theta}$ , ir neatkarīgi,  $X_{n+1}$  blīvuma funkcija ir  $f(X_{n+1}|\boldsymbol{\theta})$ . Tad  $X_{n+1}$  nosacītā blīvuma funkcija pie dotajiem  $\mathbf{X}$  ir

$$f(X_{n+1}|\mathbf{X}) = \int f(X_{n+1}|\boldsymbol{\theta}) \times \hat{\pi}(\boldsymbol{\theta}|\mathbf{X})d\boldsymbol{\theta}. \quad (2.2)$$

Parasti tiek pieņemts (un tas ir pieņemts visos turpmāk minētajos piemēros), ka  $X_1, X_2, \dots, X_n, X_{n+1}$  ir neatkarīgi (nosacīti uz parametriem  $\boldsymbol{\theta}$ ) un vienādi sadalīti. Izmantojot (2.1), aposteriorais sadalījums var tikt izteikts kā

$$\hat{\pi}(\boldsymbol{\theta}|\mathbf{X}) = h(\mathbf{X}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})/h(\mathbf{X}). \quad (2.3)$$

$h(\mathbf{X}|\boldsymbol{\theta})$  sadalījums ir novērojumu nosacītā blīvuma funkcija. Šeit  $h(\mathbf{X})$  ir normalizācijas konstante, tātad aposterioro sadalījumu var uzskatīt kā novērojumu nosacīto blīvuma funkciju apvienotu ar apriorā sadalījuma informāciju. Operacionālā riska gadījumā tiek veikti trīs soļi:

1. Apriorais sadalījums  $\pi(\boldsymbol{\theta})$  tiek novērtēts ar scenāriju analīzi (ekspertu viedoklis ar atsauci uz ārējiem datiem).
2. Tad apriorais sadalījums jāsalīdzina ar novērotajiem datiem, izmantojot formulu (2.3), lai iegūtu aposterioro sadalījumu  $\hat{\pi}(\boldsymbol{\theta}|\mathbf{X})$ .
3. Tālāk tiek izmantota (2.2) formula, lai aprēķinātu prognozējamo sadalījumu  $X_{n+1}$  pie dotajiem  $\mathbf{X}$ .

Beijesa secinājumu pieeja noved pie optimāliem novērtējumiem tādā nozīmē, ka tiek minimizēta vidējā kvadrātiskā prognozes kļūda.

**Iteratīvā aprioru atjaunošanas procedūra.** Ja novērojumi  $X_1, \dots, X_n$ , nosacīti pie  $\boldsymbol{\theta}$ , ir neatkarīgi un vienādi sadalīti ar blīvuma funkciju  $f(x|\boldsymbol{\theta})$ , tad nosacītā blīvuma funkcija var tikt izteikta kā

$$h(\mathbf{X}|\boldsymbol{\theta}) = \prod_{i=1}^n f(X_i|\boldsymbol{\theta}).$$

Apriorais sadalījums pēc  $k$  novērojumiem tiek rēķināts kā  $\hat{\pi}_k(\boldsymbol{\theta}|X_1, \dots, X_k)$ , tad no (2.3) formulas

$$\hat{\pi}_k(\boldsymbol{\theta}|X_1, \dots, X_k) \propto \pi(\boldsymbol{\theta}) \prod_{i=1}^k f(X_i|\boldsymbol{\theta}) \propto \hat{\pi}_{k-1}(\boldsymbol{\theta}|X_1, \dots, X_{k-1})(X_k|\boldsymbol{\theta}). \quad (2.4)$$

No (2.4) var viegli redzēt, ka atjaunošanas procedūra, kura no apriora rēķina aposterioru, var tikt veikta iteratīvi. Lai aprēķinātu aposterioro sadalījumu pēc  $k$  novērojumiem, ir nepieciešams aposteriorais sadalījums, rēķināts pēc  $k-1$  novērojumiem, un  $k$ -tais novērojums. Tātad zaudējumu vēsturiskie dati daudzu gadu griezumā nav nepieciešami, padarot modeli saprotamāku un ērtāku un ļaujot ekspertiem katrā solī koriģēt aprioru. Formāli pēc  $k-1$  novērojuma aprēķināto aposterioro sadalījumu var uzskatīt kā  $k$ -tā novērojuma aprioro sadalījumu. Praksē sākotnēji tiek sākts ar aprioro sadalījumu  $\pi(\boldsymbol{\theta})$ , kas tiek aprēķināts tikai ar ekspertu viedokļiem un ārējiem datiem. Pēc tam, izmantojot 2.3, tiek aprēķināts aposteriorais sadalījums  $\hat{\pi}(\boldsymbol{\theta}|\mathbf{X})$ , kad tiek novēroti faktiskie dati. Ja ir iemesls (piemēram, bankā ieviesta jauna kontroles politika), tad šo aposterioro sadalījumu var koriģēt eksperts un turpmākajos novērojumos uzskatīt par aprioro sadalījumu.

### 3. Saistītie apriorie sadalījumi

Saistītie sadalījumi ir ļoti noderīgi, izmantojot Beijesa secinājumu metodi. Teorētiskais pamats saistītajiem sadalījumiem galvenokārt tiek balstīts uz [8].

**3.1. definīcija.** Ar funkciju  $F$  apzīmē blīvuma funkciju klasi  $f(\mathbf{X}|\boldsymbol{\theta})$ , kas ir indeksēta pēc  $\boldsymbol{\theta}$ . Klasi  $U$ , kas sastāv no apriorajām blīvuma funkcijām  $\pi(\boldsymbol{\theta})$ , sauc par konjugētu klasei  $F$ , ja aposteriorais sadalījums  $\hat{\pi}(\boldsymbol{\theta}|\mathbf{X}) = f(\mathbf{X}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})/f(\mathbf{X})$ , kur  $f(\mathbf{X}) = \int f(\mathbf{X}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}$ , ir  $U$  funkciju klasē visiem  $f \in F$  un  $\pi \in U$  [5].

Formāli, ja sadalījumu klasē  $U$  ietilpst visas sadalījumu funkcijas, tad tā ir saistīta ar jebkuru klasi  $F$ . Tomēr, lai modelis ir praktiski noderīgs un izmantojams, ir svarīgi, lai  $U$  ir pēc iespējas mazāks, kamēr tas satur reālus sadalījumus. Tālāk tiks apskatīti  $F$  un  $U$  saistīti pāri: Puasona– gamma, lognormālais– normālais, Pareto– gamma, kuri ir visnoderīgākie, lai modelētu zaudējumu biežuma un smaguma procesus operacionālajā riskā. Visos iepriekš minētajos pāros apriorais un aposteriorais sadalījums ir tāda paša tipa un aposteriorā sadalījuma parametri ir viegli novērtējami, izmantojot apriorā sadalījuma parametrus un novērojumus.

#### 3.1. Puasona– gamma

Puasona sadalījums tiek bieži izmantots, modelējot zaudējumu biežumu operacionālajā riskā. Pieņem, ka dotajam  $\lambda$  novērojumi  $\mathbf{N} = (N_1, \dots, N_n)$  ir neatkarīgi gadījuma lielumi no Puasona sadalījuma ( $Poisson(\lambda)$ ) ar blīvuma funkciju

$$f(N|\lambda) = e^{-\lambda} \frac{\lambda^N}{N!}, \quad \lambda \geq 0,$$

un  $\lambda$  apriorais sadalījums ir Gamma ( $Gamma(\alpha, \beta)$ ) ar blīvuma funkciju

$$\pi(\lambda|\alpha, \beta) = \frac{(\lambda/\beta)^{\alpha-1}}{\Gamma(\alpha)\beta} e^{-\lambda/\beta}, \quad \lambda > 0, \alpha > 0, \beta > 0.$$

Tas nozīmē, ka  $\boldsymbol{\theta}$  ir atkarīgs no  $\lambda$  un  $\mathbf{X}$  ir atkarīgs no  $\mathbf{N}$  vienādojumā (2.3). Pie dotajiem  $\lambda$   $N_1, \dots, N_n$  ir nosacīti neatkarīgi, un to nosacītā blīvuma funkcija ir

$$h(\mathbf{N}|\lambda) = \prod_{i=1}^n e^{-\lambda} \frac{\lambda^{N_i}}{N_i!}.$$

Tad, izmantojot (2.3) formulu, aposteriorais sadalījums ir

$$\hat{\pi}(\lambda|\mathbf{N}) \propto \frac{(\lambda/\beta)^{\alpha-1}}{\Gamma(\alpha)\beta} e^{-\lambda/\beta} \prod_{i=1}^n e^{-\lambda} \frac{\lambda^{N_i}}{N_i!} \lambda^{\hat{\alpha}-1} e^{-\lambda/\hat{\beta}},$$

kas ir  $Gamma(\hat{\alpha}, \hat{\beta})$ , t.i. tāds pats kā apriorais sadalījums ar novērtētiem parametriem  $\hat{\alpha}$  un  $\hat{\beta}$ , kuri ir doti kā

$$\hat{\alpha} = \alpha + \sum_{i=1}^n N_i,$$

$$\hat{\beta} = \beta / (1 + \beta \times n).$$

Nosacītā matemātiskā cerība  $E[N_{n+1}|\mathbf{N}]$ , ņemot vērā iepriekšējos notikumus, kas šajā gadījumā ir aposteriorā sadalījuma vidējā vērtība, ir izsakāma ar vienādojumu

$$E[N_{n+1}|\mathbf{N}] = E[\lambda|\mathbf{N}] = \hat{\alpha} \times \hat{\beta} = \beta \times \frac{\alpha + \sum_{i=1}^n N_i}{1 + \beta \times n} = \omega \bar{N} + (1 - \omega)\lambda_0,$$

kur

$$\bar{N} = \frac{1}{n} \sum_{i=1}^n N_i \text{ ir } \lambda \text{ novērtējums, izmantojot tikai novērojumu skaitu,}$$

$$\lambda_0 = \alpha \times \beta \text{ ir } \lambda \text{ novērtējums, izmantojot tikai aprioro sadalījumu,}$$

$$\omega = \frac{n}{n+1/\beta} \text{ ir ticamības svars intervālā } [0, 1), \text{ kas apvieno } \lambda_0 \text{ un } \bar{N}.$$

Palielinoties gadu skaitam  $n$ , palielinās ticamības pakāpe  $\omega$  un otrādi. Tas nozīmē, jo vairāk novērojumu ir dots, jo lielāku ticamības pakāpi var piešķirt novērtējumam, kas ir balstīts uz novērojumu skaitu, kamēr zemāka ticamības pakāpi piešķir novērtējumam, kuru ietekmē eksperta viedoklis. Kā arī, jo lielāks ir eksperta viedokļa svārstīgums  $\beta$ , jo lielāka ticamības pakāpe ir dota novērojumam.

Rekursīvi aposteriorā sadalījuma aprēķins ir vienkāršs. Tiek ņemti vērā ikgadējie novērojumi  $N_1, \dots, N_k, \dots$ , kur  $N_k$  ir novēroto notikumu skaits  $k$ -tajā gadā. Pieņem, ka apriorais sadalījums  $\pi(\lambda|\alpha, \beta)$ ,  $Gamma(\alpha, \beta)$  ir dots sākotnēji, tad aposteriorais sadalījums  $\hat{\pi}_k(\lambda|N_1, \dots, N_k)$  pēc  $k$ -tā gada ir sadalīts pēc  $Gamma(\hat{\alpha}_k, \hat{\beta}_k)$  ar  $\hat{\alpha}_k = \alpha + \sum_{i=1}^k N_i$  un  $\hat{\beta}_k = \beta / (1 + \beta \times k)$ . Jāievēro, ka

$$\hat{\alpha}_k = \hat{\alpha}_{k-1} + N_k, \quad \hat{\beta}_k = \hat{\beta}_{k-1} / (1 + \hat{\beta}_{k-1}). \quad (3.1)$$

Tādējādi tiek izveidota rekursīva shēma, kurā aposteriorā sadalījuma parametrus aprēķina, pamatojoties uz iepriekšējajiem novērojumiem un aposteriorā sadalījuma parametriem, kas aprēķināti tieši pirms šī novērojuma.

### 3.2. Lognormālais– normālais

Lognormālais sadalījums  $LN(\mu, \sigma)$  tiek bieži lietots, lai aprēķinātu zaudējumu smagumu operacionālajam riskam. Tiek pieņemts, ka  $\mathbf{X} = (X_1, \dots, X_n)$ , nosacīti uz  $\mu$  un  $\sigma$ , ir neatkarīgs gadījuma lielums, sadalīts pēc  $LN(\mu, \sigma)$  un ar blīvuma funkciju

$$f(x|\mu, \sigma) = \frac{1}{X\sqrt{2\pi}\sigma^2} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}.$$

Tātad,  $Y_i = \ln X_i$ ,  $i = 1, \dots, n$  ir sadalīti pēc normālā sadalījuma  $N(\mu, \sigma)$ . Tiek pieņemts, ka parametrs  $\sigma$  ir zināms, un  $\mu$  apriorais sadalījums ir normālais sadalījums  $N(\mu_0, \sigma_0)$  ar blīvuma funkciju

$$\pi(\mu|\mu_0, \sigma_0) = \frac{1}{\sigma_0\sqrt{2\pi}} e^{-\frac{(\mu - \mu_0)^2}{2\sigma_0^2}}.$$

Tas nozīmē, ka  $\theta$  ir atkarīgs no  $\mu$  vienādojumā (2.3). Novērtējumu nosacītā blīvuma funkcija ir

$$h(\mathbf{Y}|\mu, \sigma) = \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(Y_i - \mu)^2}{2\sigma^2}}.$$

Tad, izmantojot (2.3) formulu, aposteriorais sadalījums var tikt uzrakstīts kā

$$\hat{\pi}(\mu|\mathbf{Y}) \propto \frac{e^{-\frac{(\mu - \mu_0)^2}{2\sigma_0^2}}}{\sigma_0\sqrt{2\pi}} \prod_{i=1}^n \frac{e^{-\frac{(Y_i - \mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \propto e^{-\frac{(\mu - \hat{\mu}_0)^2}{2\hat{\sigma}_0^2}},$$

kas ir normālais sadalījums  $N(\hat{\mu}_0, \hat{\sigma}_0)$  ar atjaunotiem parametriem

$$\hat{\mu}_0 = \mu_0 + \omega \sum_{i=1}^n Y_i / (1 + n \times \omega),$$

$$\hat{\sigma}_0^2 = \sigma_0^2 / (1 + n \times \omega),$$

kur  $\omega = \sigma_0^2 / \sigma^2$ .

$Y_{n+1}$  nosacītā matemātiskā cerība ir

$$E[Y_{n+1}|\mathbf{X}] = E[\mu|\mathbf{X}] = \hat{\mu}_0 = \frac{\mu_0 + \omega \sum_{i=1}^n Y_i}{1 + n \times \omega} = \omega \bar{Y} + (1 - \omega)\mu_0,$$

kur

$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$  ir  $\mu$  novērtējums, iegūts no novērotajiem zaudējumiem,

$\mu_0$  ir  $\mu$  novērtējums, aprēķināts, izmantojot aprioro sadalījumu,

$\omega = \frac{n}{n+\sigma^2/\sigma_0^2}$  ir uzticamības svars  $[0, 1)$ , kas apvieno  $\mu_0$  un  $\bar{Y}$ .

Palielinoties novērojumu skaitam, palielinās uzticamības pakāpe  $\omega$  un otrādi. Mazāku uzticamības pakāpi piešķir novērtējumam, kas ir balstīts uz ekspertu viedokli. Kā arī lielāka svārstība ekspertu viedoklī  $\sigma_0^2$  norāda uz lielāku uzticamības pakāpi novērojumiem, un lielāks novērojumu svārstīgums  $\sigma^2$  norāda uz augstāku uzticamības pakāpi ekspertu viedoklim.

Aposteriorais sadalījums var tikt izrēķināts rekursīvi. Tiek aplūkoti novērojumi  $Y_1, \dots, Y_k, \dots$ . Pieņem, ka apriorais sadalījums  $\pi(\mu|\mu_0, \sigma_0)$ ,  $N(\mu_0, \sigma_0)$ , ir novērtēts sākotnēji, tad posteriorais sadalījums  $\hat{\pi}_k(\mu|Y_1, \dots, Y_k)$  pēc  $k$ -tā gada ir normāli sadalīts kā  $N((\hat{\mu}_0)_k, (\hat{\sigma}_0)_k)$  ar

$$(\hat{\mu}_0)_k = (\mu_0 + \omega \sum_{i=1}^k Y_i) / (1 + k \times \omega),$$

$$(\hat{\sigma}_0)_k = \sigma_0^2 / (1 + k \times \omega),$$

kur  $\omega = \sigma_0^2 / \sigma^2$ . Var izteikt

$$(\hat{\mu}_0)_k = \frac{(\hat{\mu}_0)_{k-1} + [(\hat{\sigma}_0^2)_{k-1} / \sigma^2] \times Y_k}{1 + [(\hat{\sigma}_0^2)_{k-1} / \sigma^2]},$$

$$(\hat{\sigma}_0^2)_k = \frac{(\hat{\sigma}_0^2)_{k-1}}{1 + [(\hat{\sigma}_0^2)_{k-1} / \sigma^2]}.$$
(3.2)

Tas nozīmē, ka posteriorā sadalījuma parametrus var aprēķināt, balstoties uz iepriekšējajiem novērojumiem un posteriorā sadalījuma parametriem, kuri tika aprēķināti iepriekšējā solī.

### 3.3. Pareto– gamma

Vēl viens nozīmīgs smaguma sadalījums, kas ir noderīgs, lai pielāgotu astes sadalījumu noteiktai robežvērtībai  $L > 0$ , ir Pareto sadalījums ar blīvuma funkciju

$$f(x|\xi) = \frac{\xi}{L} \left(\frac{x}{L}\right)^{-\xi-1},$$

kur  $x \geq L$  un  $\xi > 0$ . Ja  $\xi > 1$ , tad vidējā vērtība ir  $L\xi/(\xi - 1)$ , citos gadījumos vidējā vērtība neeksistē. Pieņem, ka gadījuma vektors  $\mathbf{X} = (X_1, \dots, X_n)$ , nosacīti uz  $\xi$ , ir neatkarīgs un sadalīts pēc Pareto sadalījuma. Astes parametram  $\xi$  apriorais sadalījums ir  $Gamma(\alpha, \beta)$  ar blīvuma funkciju

$$\pi(\xi|\alpha, \beta) \propto \xi^{\alpha-1} e^{-\xi/\beta}.$$

Izmantojot formulu (2.3), aposteriorais sadalījums

$$\hat{\pi}(\xi|\mathbf{X}) = \xi^n e^{-(\xi+1)\sum_{i=1}^n \ln\left(\frac{X_i}{L}\right)} \times \xi^{\alpha-1} e^{\left(-\frac{\xi}{\beta}\right)} \xi^{\hat{\alpha}-1} e^{-\frac{\xi}{\hat{\beta}}}.$$

Tas ir  $Gamma(\hat{\alpha}, \hat{\beta})$  ar atjaunotiem parametriem

$$\begin{aligned}\hat{\alpha} &= \alpha + n, \\ \hat{\beta}^{-1} &= \beta^{-1} + \sum_{i=1}^n \ln\left(\frac{X_i}{L}\right).\end{aligned}$$

Apsteriorā sadalījuma nosacīto matemātisko cerību visiem  $\xi$  var izteikt kā

$$\hat{\xi} = E[\xi|\mathbf{X}] = \hat{\alpha} \times \hat{\beta} = \frac{\alpha + n}{\frac{1}{\beta} + \sum_{i=1}^n \ln\left(\frac{X_i}{L}\right)} = \omega \hat{\xi}^{MLE} + (1 - \omega)\xi_0,$$

kur

$\hat{\xi}^{MLE} = \frac{1}{n} \sum_{i=1}^n \ln\left(\frac{X_i}{L}\right)$  ir  $\xi$  maksimālās ticamības novērtējums, izmantojot novērotos zaudējumus,

$\xi_0 = \alpha \times \beta$  ir  $\xi$  novērtējums, izmantojot aprioro sadalījumu,

$\omega = \left[\sum_{i=1}^n \ln\left(\frac{X_i}{L}\right)\right] \times \left[\sum_{i=1}^n \ln\left(\frac{X_i}{L}\right) + \frac{1}{\beta}\right]^{-1}$  ir uzticamības svars intervālā  $[0, 1)$ , kas apvieno  $\xi_0$  un  $\hat{\xi}^{MLE}$ .

Apsteriorais sadalījums var tikt rēķināts rekursīvi. Tiek aplūkoti novērojumu  $X_1, X_2, \dots, X_k, \dots$ . Pieņem, ka apriorajam sadalījumam  $\pi(\xi|\alpha, \beta)$ ,  $Gamma(\alpha, \beta)$ , ir dots. Tad apsteriorais sadalījums  $\hat{\pi}_k(\xi|X_1, \dots, X_k)$  pēc  $k$ -tā gada ir  $Gamma(\hat{\alpha}_k, \hat{\beta}_k)$  ar parametriem  $\hat{\alpha}_k = \alpha + k$  un  $\hat{\beta}_k^{-1} = \beta^{-1} + \sum_{i=1}^k \ln(X_i/L)$  jeb

$$\hat{\alpha}_k = \hat{\alpha}_{k-1} + 1, \quad \hat{\beta}_{k-1}^{-1} + \ln\left(\frac{X_k}{L}\right). \quad (3.3)$$

Ir svarīgi atcerēties, ka apriorā un apsteriorā sadalījuma  $\xi$  ir Gamma sadalīts un formāli definēts kā  $\xi > 0$ . Tātad pastāv galīga varbūtība, ka  $P[\xi \leq 1] > 0$ , kas noved pie prognozēto sadalījumu bezgalīgiem vidējiem lielumiem, tas ir,  $E[X_i] = \infty$  un  $E[X_{n+1}|\mathbf{X}] = \infty$ . Ja bezgalīga vidējā vērtība nav vēlama, tad  $\xi$  jāierobežo  $\xi > 1$ .

### 3.4. Parametru ierobežošana

Praksē bieži ir nepieciešams ierobežot parametrus. Piemēram, dotajiem novērojumiem  $\mathbf{X} = (X_1, \dots, X_n)$  var tikt izvēlēts lognormālais sadalījums  $LN(\mu, \sigma)$ , un piekārtots apriorais sadalījums parametram  $\mu$ , kuram jābūt normālam  $N(\mu_0, \sigma_0)$ . Tomēr, ja zina, ka  $\mu$  nevar būt negatīvs, tad  $N(\mu_0, \sigma_0)$  tiek ierobežots tikai ar nenegatīvām vērtībām. Vēl viens piemērs var būt Pareto– gamma gadījums, kur apriorais sadalījums parametram  $\xi$ ,  $Gamma(\alpha, \beta)$ , ir definēts visiem  $\xi > 0$ . Bet, ja nav vēlams pieļaut bezgalīgus lielumus nākotnes zaudējumiem, tad parametru  $\xi$  jāierobežo kā  $\xi > 1$ . Šos gadījumus var viegli atrisināt, izmantojot aprioru–aposterioru sadalījumu saīsinātās versijas. Tiek pieņemts, ka  $\pi(\theta)$  ir apriorais sadalījums, kura aposteriorais sadalījums ir  $\hat{\pi}(\theta|\mathbf{X}) = h(X|\theta)\pi(\theta)/h(\mathbf{X})$ , kur  $\theta$  nav ierobežots. Ja parametrs ir ierobežots kā  $a \leq \theta \leq b$ , tad var pieņemt, ka apriorais sadalījums

$$\pi^{tr}(\theta) = \frac{\pi(\theta)}{P[a \leq \theta \leq b]} \times I_{a \leq \theta \leq b},$$

$$P[a \leq \theta \leq b] = \int_a^b \pi(\theta) d\theta,$$

visiem  $a$  un  $b$ , kuriem  $P[a \leq \theta \leq b] > 0$ . Šeit  $I_{a \leq \theta \leq b} = 1$ , ja  $a \leq \theta \leq b$  un vienāds ar nulli citos gadījumos.  $P[a \leq \theta \leq b]$  ir kā normalizātors, un tādējādi aposteriorais sadalījums ir uzrakstāms kā

$$\hat{\pi}(\theta|\mathbf{X}) = \frac{\hat{\pi}(\theta|\mathbf{X})}{P[a \leq \theta \leq b|\mathbf{X}]} I_{a \leq \theta \leq b},$$

$$P[a \leq \theta \leq b|\mathbf{X}] = \int_a^b \hat{\pi}(\theta|\mathbf{X}) d\theta.$$

Ir acīmredzami, ka, ja  $\pi(\theta)$  ir saistītais apriors, tad arī  $\pi^{tr}(\theta)$  ir saistīts apriors.

## 4. Strukturālo aprioro parametru subjektīvais novērtējums

Kopumā apriorā sadalījuma strukturālie parametri var tikt novērtēti subjektīvi, izmantojot tikai ekspertu viedokli (tīrā Beijesa pieeja) vai izmantojot datus (empīriskā Beijesa pieeja). Tīrajā Beijesa pieejā apriorais sadalījums ir definēts subjektīvi (tikai ekspertu viedoklis). Berģers [2] ir uzskaitījis vairākas metodes apriorā sadalījuma novērtēšanā. Viena no tām ir dotās funkcionālās formas saskaņošanas metodi. Tās darbības mērķis ir atrast apriorā sadalījuma parametrus, pieņemot kādu funkcionālo formu šim sadalījumam, lai tas atbilstu apriorajiem pieņēmumiem (momenti, kvantiles, u.c.), cik vien tuvu tas ir iespējams.

Turpmāk, izmantojot dotās funkcionālās formas saskaņošanas metodi, tiks aplūkoti apriorā sadalījuma parametru novērtējumi Puasona– gamma, Pareto– gamma un lognormālā– normālā sadalījumu pāriem.

### 4.1. Puasona– gamma parametru novērtējums

Tiek pieņemts, ka ikgadējais operacionālā riska zaudējumu biežums  $N$  ir sadalīts pēc Puasona sadalījuma  $Poisson(\lambda)$  un ka  $\lambda$  apriorais sadalījums  $\pi(\lambda|\alpha, \beta)$  ir  $Gamma(\alpha, \beta)$ . Tad  $E[N|\lambda] = \lambda$  un  $E[\lambda] = \alpha \times \beta$ . Eksperts var novērtēt sagaidāmo notikumu skaitu, bet viņš nevar būt pārliecināts par šo novērtējumu. Var uzskatīt, ka eksperta novērtējums ir matemātiskās cerības labākais novērtējums, t.i.,

$$E[E[N|\lambda]] = E[\lambda].$$

Ja eksperts nosaka  $E[\lambda]$  un nenoteiktība, ka "īstā"  $\lambda$  vērtība nākamajam gadam ir intervālā  $[a, b]$  ar varbūtību  $P[a \leq \lambda \leq b] = p$  (ērtības labad parasti tiek izmantots  $p = 2/3$ ), tad vienādojumi

$$\begin{aligned} E[\lambda] &= \alpha \times \beta, \\ P[a \leq \lambda \leq b] &= \int_a^b \pi(\lambda|\alpha, \beta) d\lambda = F_{\alpha, \beta}[b] - F_{\alpha, \beta}[a] \end{aligned} \tag{4.1}$$

var tikt aprēķināti, lai novērtētu strukturālos parametrus  $\alpha, \beta$ . Šeit  $F_{\alpha, \beta}[y]$  ir Gamma kumulatīvā sadalījuma funkcija

$$F_{\alpha, \beta}[y] = \int_0^y \frac{x^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha} e^{-\frac{x}{\beta}} dx.$$

**3. piemērs.** *Piemērā izmantotie parametri tiks ņemti no raksta [8].*

*Eksperts ir noteicis, ka  $E[\lambda] = 0.5$  un  $P[0.25 \leq \lambda \leq 0.75] = 2/3$ . Atrisinot vienādojumus (4.1), izmantojot R studio komandu `uniroot()` [9], var novērtēt apriorā sadalījuma parametrus  $\alpha \approx 3.407$ ,  $\beta \approx 0.147$ . Pieņem, ka banka pirmajā gadā nav cietusi zaudējumus (pēc tā kā apriorais sadalījums tika novērtēts). Tad, izmantojot (3.1), aposteriorā sadalījuma parametri tiek aprēķināti 15 gadu griezumā (skatīt 3. tabulā).*

3. tabula

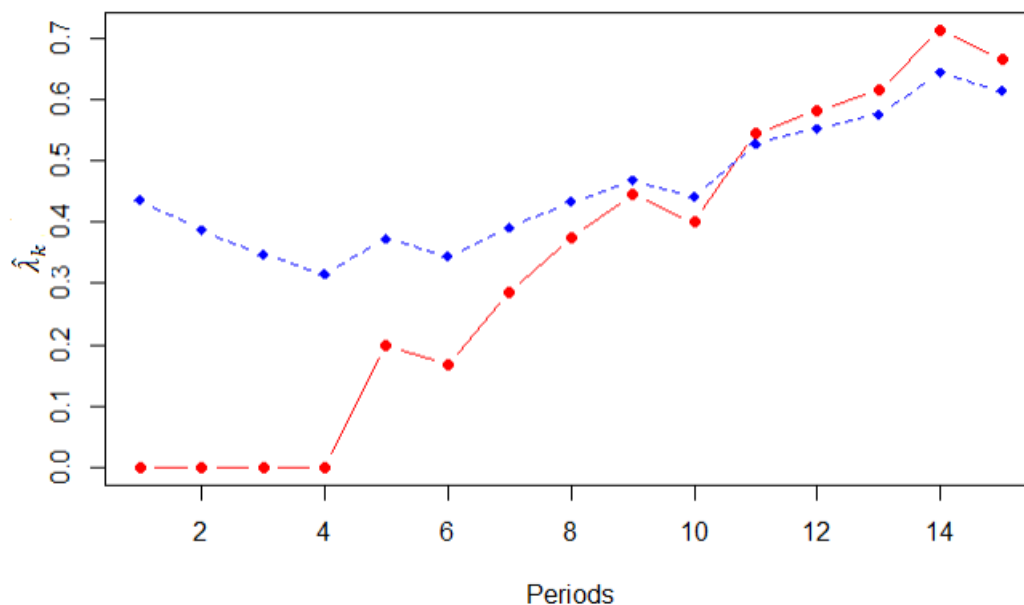
**Apriorā un aposteriorā sadalījuma novērtētie parametri 15 gadu griezumā**

Gads	$\hat{\alpha}_k$	$\hat{\beta}_k$	$\hat{\lambda}_k$
1	3.41	0.13	0.44
2	3.41	0.11	0.39
3	3.41	0.10	0.35
4	3.41	0.09	0.32
5	4.41	0.08	0.37
6	4.41	0.08	0.34
7	5.41	0.07	0.39
8	6.41	0.07	0.43
9	7.41	0.06	0.47
10	7.41	0.06	0.44
11	9.41	0.06	0.53
12	10.41	0.05	0.55
13	11.41	0.05	0.58
14	13.41	0.05	0.64
15	13.41	0.05	0.61

*Tātad, sākot no ekspertu norādītā apriorora, novērojumi regulāri atjauno aposterioro sadalījumu.*

*3. attēlā tiek atspoguļots novērtējums  $\hat{\lambda}_k = \hat{\alpha}_k \times \hat{\beta}_k$ ,  $k = 1, \dots, 15$ , kad ikgadējais notikumu skaits  $N_k$  ir simulēts no Puasona sadalījuma ar  $\lambda = 0.6$ ,  $N = (0, 0, 0, 0, 1, 0, 1, 1, 1, 0, 2, 1, 1, 2, 0)$ . Maksimālais ticamības novērtējums tiek aprēķināts pēc formulas  $\tilde{\lambda}_k = \frac{1}{k} \sum_{i=1}^k N_i$ .*

*Aptuveni pēc 8 gadiem abi novērtējumi atrodas ļoti tuvu viens otram. Tikai pēc 12 gadiem abi novērtējumi konverģē uz īsto vērtību  $\lambda = 0.6$ . Tas notiek tāpēc, ka bankai nebija zaudējumu pirmos četrus gadus. Tā kā sākotnēju eksperti noteica, ka  $\lambda = 0.5$ , šis piemērs parāda, ka*



3. att.: Beijesa metodes (zilā līnija) un maksimālā ticamības intervāla (sarkanā līnija)  $\hat{\lambda}_k$  novērtējums 15 gadu griezumā. Ikgadējo notikumu skaits tika simulēts no Puasona sadalījuma ar  $\lambda = 0.6$

sākotnēji neprecīzs aprioru novērtējums var tikt koriģēts ar jauna perioda novērojumu skaitu, tiklīdz tie kļūst pieejami.

Pēc 14 gadiem novērojumi atkal sāk attālināties. Tas notiek tāpēc, ka 12., 13. un 14. gadā notikumu skaits bija 1, 1 un 2 attiecīgi. Kā rezultāts maksimālās ticamības novērtējums kļūst augstāks nekā īstā vērtība, kamēr Beijesa novērtējums ir stabilāks (plūstošāks, vienmērīgāks) attiecībā uz gadiem, kuros notika zaudējumi.

## 4.2. Lognormālā-normālā sadalījumu parametru novērtējums

Tiek pieņemts, ka operacionālo zaudējumu smagums  $\mathbf{X}$  ir modelēts pēc lognormālā sadalījuma  $LN(\mu, \sigma)$ . Tad dotajiem  $\mu$  un  $\sigma$  zaudējumu matemātiskā cerība ir

$$E[X|\mu, \sigma] = M(\mu, \sigma) = e^{\mu + \frac{1}{2}\sigma^2},$$

un šo zaudējumu  $q$ -tā kvantile ir

$$Q_q(\mu, \sigma) = e^{\mu + \sigma Z_q},$$

kur  $Z_q$  ir standartnormālā  $q$ -tā kvantile. Pieņem, ka  $\sigma$  ir zināms, un  $\mu$  apriorais sadalījums ir  $N(\mu_0, \sigma_0)$ . Tad  $M(\mu, \sigma)$  ir sadalīts pēc  $LN(\mu + \frac{1}{2}\sigma^2, \sigma_0)$ , un kvantile  $Q_q(\mu, \sigma)$  ir sadalīta pēc  $LN(\mu_0 + \sigma Z_q, \sigma_0)$ .

Eksperti var piedāvāt, viņuprāt, labāko novērtējumu zaudējumu matemātiskajai cerībai  $E[M(\mu, \sigma)]$  un nenoteiktības intervālu  $[a, b]$  tādu, ka īstie sagaidāmie zaudējumi ir šajā intervālā ar varbūtību  $P[a \leq M \leq b] = 2/3$ . Tad vienādojumi var tikt atrisināti, lai atrastu  $\mu_0, \sigma_0$

$$E[M] = e^{\mu_0 + \frac{1}{2}\sigma^2 + \frac{1}{2}\sigma_0^2},$$

$$P[a \leq M \leq b] = \Phi\left[\frac{\ln b - \frac{1}{2}\sigma^2 - \mu_0}{\sigma_0}\right] - \Phi\left[\frac{\ln a - \frac{1}{2}\sigma^2 - \mu_0}{\sigma_0}\right]. \quad (4.2)$$

Šeit  $\Phi[x]$  ir standartnormālā kumulatīvā sadalījuma funkcija.

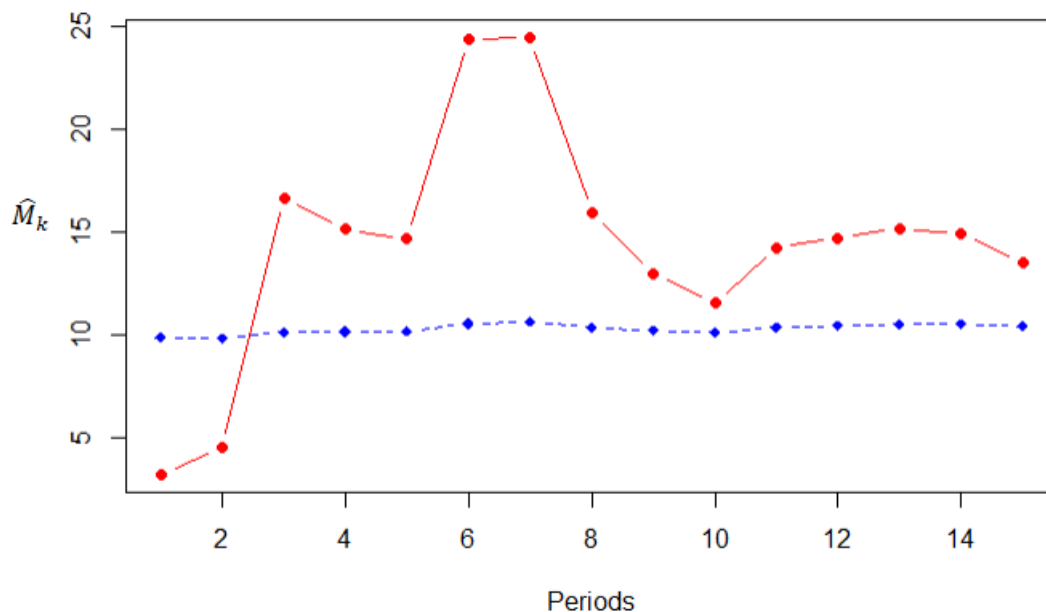
**4. piemērs.** Piemērā izmantotie parametri tiks ņemti no raksta [8].

Ir dots, ka  $\sigma = 2$ , un ekspertu novērtējums ir  $E[M] = 10$ , kā arī  $P[a \leq M \leq 12] = 2/3$ . Izmantojot (4.2) formulas programmā R – studio, var aprēķināt  $\mu_0 \approx 0.28$  un  $\sigma_0 \approx 0.21$ . Vēlāk, izmantojot 3.2 formulas, var izrēķināt aposterioros parametrus  $(\hat{\mu}_0)_k, (\hat{\sigma}_0)_k$ , kad ir zināmi novērojumi  $X_k, k = 1, 2, \dots, 15$

4. tabula

**Novērtētie aposteriorā sadalījuma parametri un matemātiskā cerība**

Periods	$(\hat{\mu}_0)_k$	$(\hat{\sigma}_0)_k$	$M_k$
1	0.27	0.21	9.88
2	0.26	0.21	9.83
3	0.30	0.21	10.16
4	0.30	0.21	10.18
5	0.30	0.20	10.20
6	0.34	0.20	10.57
7	0.35	0.20	10.66
8	0.32	0.20	10.38
9	0.31	0.20	10.24
10	0.30	0.20	10.15
11	0.32	0.20	10.39
12	0.33	0.20	10.46
13	0.34	0.20	10.54
14	0.34	0.20	10.55
15	0.33	0.19	10.44



4. att.: Beijesa metodes (zilā līnija) un maksimālā ticamības intervāla (sarkanā līnija)  $\hat{M}_k$  novērtējums 15 gadu griezumā. Ikgadējo zaudējumu smagums tika simulēts no  $LN(0.3, 2)$

4. attēlā tiek atspoguļots  $\hat{M}_k = e^{(\hat{\mu}_0)_k + \frac{1}{2}\sigma^2 + \frac{1}{2}\hat{\sigma}_0)_k}$ ,  $k = 1, \dots, 15$ , kad ikgadējais zaudējumu smagums ir simulēts no  $LN(0.3, 2)$ . Maksimālās ticamības metodes novērtējums tika aprēķināts pēc formulas  $\tilde{M}_k = e^{\hat{\mu}_k^* + \frac{1}{2}\sigma^2}$ , kur  $\hat{\mu}_k^*$  ir normālā sadalījuma parametrs zaudējumu smagumam, kurš tika novērtēts ar maksimālās ticamības metodi pie fiksēta  $\sigma = 2$ .

Beijesa metodes novērtējums ir stabilāks, tas ir ap īsto vērtību  $M = 10$ , kamēr maksimālās ticamības novērtējums svārstās. Kad zaudējumu smagums strauji pieaug,  $\tilde{M}_k$  arī pieaug, bet  $\hat{M}_k$  pieaugums ir minimāls. 10. gadā vērtības gandrīz satiekas, bet tad seko straujš zaudējumu smaguma pieaugums, un maksimālās ticamības novērtējums atkal pieaug, attālinoties no Beijesa novērtējuma.

### 4.3. Pareto– gamma sadalījumu parametru novērtēšana

Tiek pieņemts, ka operacionālo zaudējumu smagums  $X$ , kas pārsniedz robežvērtību  $L$ , ir modelēts pēc Pareto sadalījuma  $Pareto(\xi)$ . Tad dotajam  $\xi$

$$E[X|\xi] = \mu(\xi) = L\xi/(\xi - 1),$$

kur  $\xi > 1$ , un  $q$ -tā kvantile ir

$$Q_q(\xi) = Le^{-\frac{\ln(1-q)}{\xi}},$$

kur  $\xi > 0$ .

Ir pamatoti pieņemt, ka zaudējumu matemātiskā cerība ir galīga, tāpēc astes parametram  $\xi$  jāapmierina nosacījums  $\xi \geq B > 1$ , un līdz ar to ir iespējams izvēlēties  $\xi$  aprioro sadalījumu kā Gamma sadalījumu

$$\pi(\xi|\alpha, \beta) = I_{B \leq \xi} \times \frac{\xi^{\alpha-1} e^{-\frac{\xi}{\beta}}}{(1 - F_{\alpha, \beta}^{(G)}[B]) \times \Gamma(\alpha) \beta^\alpha}, \quad \xi \geq B, \quad \alpha > 0, \quad \beta > 0,$$

kur  $F_{\alpha, \beta}^{(G)}[x]$  ir Gamma kumulatīvā sadalījuma funkcija. Ja eksperti novērtē  $E[\xi]$  un  $P[a \leq \xi \leq b] = p$ , tad divi sekojošie vienādojumi

$$\begin{aligned} E[\xi] &= \alpha \times \beta \frac{1 - F_{\alpha+1, \beta}^{(G)}[B]}{1 - F_{\alpha, \beta}^{(G)}[B]}, \\ P[a \leq \xi \leq b] &= \frac{F_{\alpha, \beta}^{(G)}[b] - F_{\alpha, \beta}^{(G)}[a]}{1 - F_{\alpha, \beta}^{(G)}[B]} \end{aligned} \quad (4.3)$$

var tikt risināti, lai iegūtu strukturālo parametru  $\alpha$  un  $\beta$  novērtējumus.

**5. piemērs.** *Piemērā izmantotie parametri tiks ņemti no raksta [8].*

Tiek pieņemts, ka zemākā astes parametra robeža  $B = 2$ , un ekspertu novērtējums ir  $E[\xi] = 5$ ,  $P[4 \leq \xi \leq 6] = 2/3$ . Tad var pielāgot  $\alpha \approx 23,086$  un  $\beta \approx 0,217$ , un tad, izmantojot (3.3) vienādojumus, var rēķināt aposteriorā sadalījuma parametrus  $\hat{\alpha}_k, \hat{\beta}_k$ , kad novērojumi  $X_k, k = 1, 2, \dots, 15$  ir zināmi.

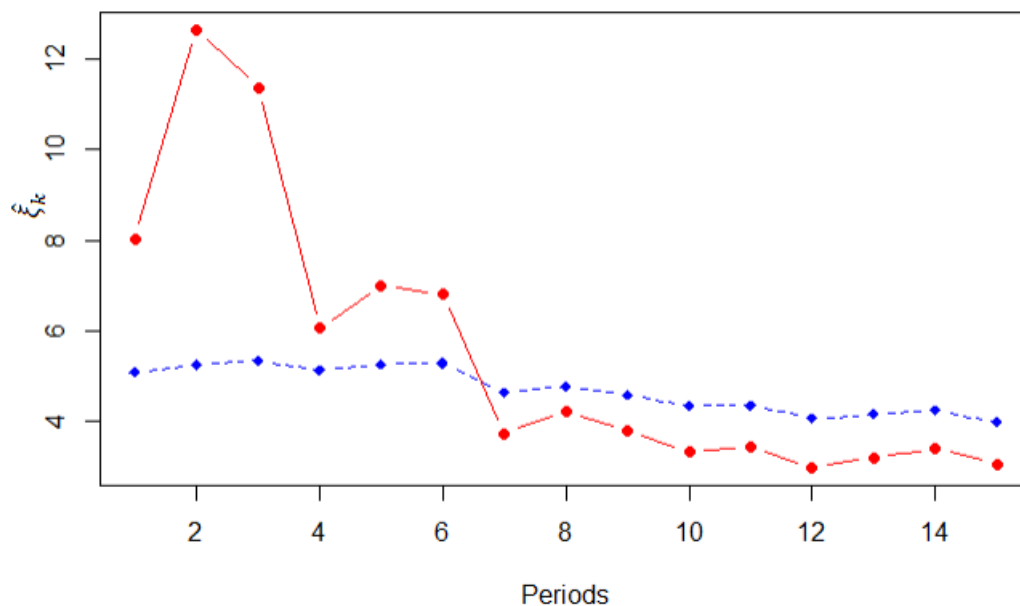
5. tabulā var redzēt sekojošu aposteriora novērtējumu astes parametram.

$$\hat{\xi}_k = \hat{\alpha}_k \times \hat{\beta}_k \times (1 - F_{\hat{\alpha}+1, \hat{\beta}}^{(G)}[B]) / (1 - F_{\hat{\alpha}, \hat{\beta}}^{(G)}[B]), \quad k = 1, 2, \dots,$$

kad zaudējumi  $X_k$  ir simulēti no Pareto sadalījuma ar parametriem  $\xi = 4$  un  $L = 1$ .

## Apriorā un aposteriorā sadalījumu parametru novērtējumi

Periods	$\hat{\alpha}_k$	$\hat{\beta}_k$	$\hat{\xi}_k$
1	24.08	0.21	5.08
2	25.08	0.21	5.25
3	26.08	0.20	5.34
4	27.08	0.19	5.13
5	28.08	0.19	5.27
6	29.08	0.18	5.29
7	30.08	0.15	4.64
8	31.08	0.15	4.77
9	32.08	0.14	4.59
10	33.08	0.13	4.35
11	34.08	0.13	4.35
12	35.08	0.12	4.07
13	36.08	0.12	4.16
14	37.08	0.11	4.25
15	38.08	0.10	4.00



5. att.: Beijesa metodes (zilā līnija) un maksimālā ticamības intervāla (sarkanā līnija)  $\hat{\xi}_k$  novērtējums 15 gadu griezumā. Ikgadējo zaudējumu skaits tika simulēts no  $Pareto(1, 4)$

5. attēlā ir redzams maksimālā ticamības intervāla novērtējums astes parametram

$$\tilde{\xi}_k = \left[ \frac{1}{k} \sum_{i=1}^k \ln(X_i/L) \right]^{-1}.$$

Var novērot, ka Beijesa novērtējums ir stabilāks, kamēr maksimālās ticamības intervāla novērtējums ir diezgan svārstīgs, kad novērojumu skaits ir mazs. Palielinoties novērojumu skaitam, abi novērtējumi kļūst gandrīz vienādi.

## 5. Kapitāla aprēķini, izmantojot Montekarlo simulācijas

Operacionālā riska regulējošā kapitāla aprēķinam ikgadējais zaudējumu sadalījums (jo īpaši tā 99.9% kvantile kā riska rādītājs) ir jānosaka kvantitatīvi katrai bankas riska šūnai (notikumu veidam/ uzņēmējdarbības līnijai).

Tiks apskatīta bankas riska šūna  $j$ . Pieņem, ka pie dotajiem parametriem  $\lambda_j$  un  $\alpha_j$  šūnai ir izvēlēts zaudējumu biežums  $g_j(x|\lambda_j)$  un smagums  $f_j(y|\lambda_j)$ . Kā arī aposteriorie sadalījumi  $\hat{\pi}(\lambda_j|N)$  un  $\hat{\pi}(\alpha_j|X)$  ir novērtēti, izmantojot aprioros sadalījumus. Tad, pamatojoties uz modeli (1.1), šūnas ikgadējo zaudējumu sadalījumu var aprēķināt, izmantojot Montekarlo procedūru.

### 5.1. Montekarlo procedūras soļi

Lai veiktu Montekarlo simulācijas, ir nepieciešams izpildīt piecus darbību soļus [8].

1. solis. Dotajam riskam  $j$  simulēt riska parametrus  $\lambda_j$  un  $\alpha_j$  no tos aposteriorajiem sadalījumiem  $\hat{\pi}(\lambda_j|\mathbf{N})$  un  $\hat{\pi}(\alpha_j|\mathbf{X})$ .
2. solis. Pirmajā solī iegūtajam  $\lambda_j$ , simulēt ikgadējo notikumu skaitu  $N_j$ , izmantojot  $j$ -tās šūnas zaudējumu biežuma sadalījumu  $g(x|\lambda_j)$ .
3. solis. Pirmajā solī iegūtajam  $\alpha_j$ , simulēt zaudējuma smagumus  $X_{j,n}$ ,  $n = 1, \dots, N_j$ , izmantojot  $j$ -tās šūnas zaudējumu smagumu sadalījumu  $f_j(y|\alpha_j)$ . Jāatzīmē, ka visi smagumi  $X_{j,n}$ ,  $n = 1, \dots, N_j$  tiek simulēti no viena un tā paša sadalījuma, t.i., riska profila parametrs  $\alpha^{(j)}$  ir piemērots visam gadam. Tas notiek tāpēc, ka viens no (1.1) modeļa pieņēmumiem ir tāds, ka visi zaudējuma smagumi ir neatkarīgi un vienādi sadalīti dotajiem parametriem  $\lambda_j$  un  $\alpha_j$ .
4. solis. Atrast  $j$ -tās šūnas ikgadējo zaudējumu  $Z_j = \sum_{i=1}^{N_j} X_{j,n}$ .
5. solis. Atkārtot soļus 1.- 4.  $K$  reizes, lai izveidotu ikgadējo zaudējumu  $Z_j(k)$ ,  $k = 1, \dots, K$  izlasi. Tad, izmantojot  $Z_j(k)$ , novērtēt nepieciešamo kvantiļu vērtības.

Ja bankai ir  $J$  riska šūnas un tās ir neatkarīgas, tad bankas kopējie ikgadējie zaudējumi var tikt aprēķināti, atkārtojot 1.- 4. soli katrai šūnai, lai atrastu  $Z_j$ ,  $j = 1, \dots, J$ . Kopējie ikgadējie zaudējumi ir izsakāmi formā

$$Z_{kop} = \sum_{j=1}^J Z_j.$$

Atkārtojot visu procedūru  $K$  reizes, tiks iegūts zaudējumu modelis  $Z_{kop}(k)$ ,  $k = 1, \dots, K$ , kas var tikt izmantots, lai novērtētu visas bankas ikgadējo zaudējumu sadalījumu.

## 5.2. VaR aprēķina piemērs

Pēc tā, kā tika aprēķināts  $Z_{kop}$ , izmantojot Montekarlo procedūru, ir svarīgi novērtēt kvantiles pie izvēlētajā ticamības līmeņa. Tās parāda operacionālā riska  $VaR$  vērtības.

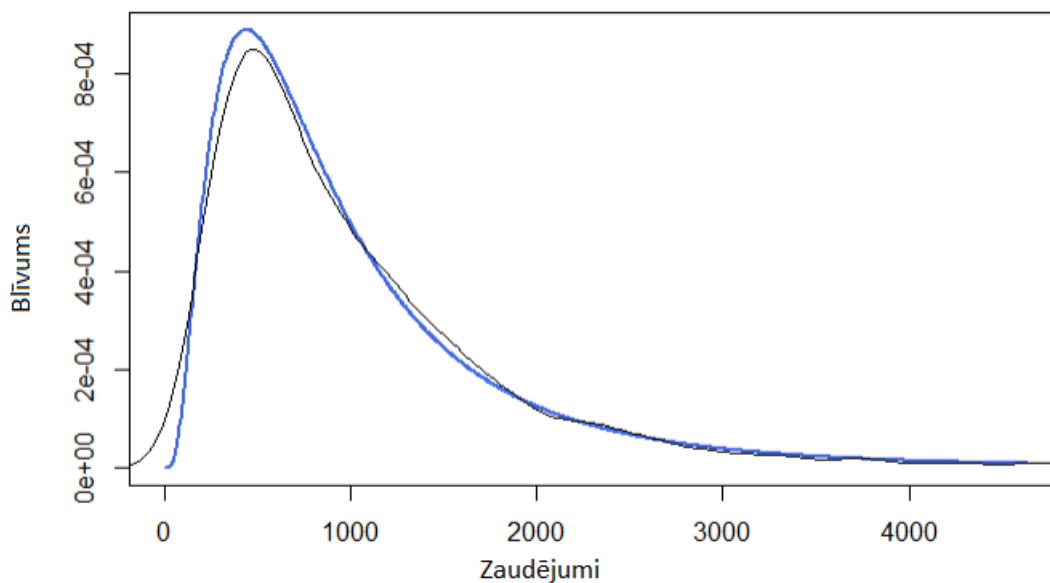
Piemērā tiks izmantota komanda *quantile* no pakotnes *stats* programmā *Rstudio*, lai aprēķinātu  $VaR$ .

Tā kā reālus operacionāla riska datus iegūt nav iespējams banku konfidencialitātes un drošības noteikumu dēļ, tad tiks izmantoti dati no pakotnes *OpVaR* [11] ar nosaukumu *lossdata*, kas ir hipotētisks, simulēts piemērs ar divdimensiju atkarību starp zaudējumu biežumu un smagumu. Atkarības modelēšana šajā darbā netiks apskatīta.

Kopa *lossdata* sastāv no 4 šūnām. Katrai no tām ir 3 kolonnas: zaudējumi, laika zīmogs un periods. Piemērā tiks izmantota trešā datu kopa. Tā sastāv no 1995 zaudējumu smaguma novērojumiem, kuri tika ievākti 40 periodu garumā. Tiks veikti statistikas testi tādi, kā Kolmogorova–Smirnova un  $H_1$  kvadrātā testi attiecīgi zaudējumu smagumam un biežumam, lai pārbaudītu izvēlētajā sadalījuma atbilstību. Sākumā tiks pārbaudīts, vai zaudējumu smagumiem atbilst lognormālais sadalījums.

Kolmogorova–Smirnova divu izlašu tests pārbauda, vai piekārtotais sadalījums atbilst datiem. Trešās datu kopas zaudējumu smagumu sadalījums atbilst lognormālajam sadalījumam, jo  $p$ -vērtība = 0.4116, kas ir lielāka par 0.05, un statistikas vērtība = 0.03, kas arī liecina par atbilstību, tātad nulles hipotēzi par sadalījuma piemērotību nevar noraidīt. Ar maksimālās ticamības metodi tiek novērtēti parametri  $\mu = 6.67$  un  $\sigma = 0.76$ .

6. attēlā var pārliecināties, ka lognormālais sadalījums labi apraksta zaudējumus.



6. att.: Zaudējumu biežuma sadalījuma grafiks. Īstā blīvuma funkcija (melnā) un novērtētā lognormālā sadalījuma blīvuma funkcija (zilā)

Tālāk trešajai šūnai ir jāpārbauda zaudējumu biežuma sadalījums ar Hī kvadrātā testa palīdzību, kas tāpat kā Kolmogorova– Smirnova tests pārbauda, vai izlašu sadalījums atbilst pieņemtajam. Testa  $p$ - vērtība = 0.999. Tas nozīmē, ka zaudējumu biežumam var pieņemt Puasona sadalījumu ar parametru  $\lambda = 49.88$ , kas tika novērtēts ar maksimālās ticamības metodi.

**Beijesa metode** Tā kā zaudējumu biežums ir novērtēts kā Puasona sadalījums, tad ekspertam ir jāizsaka savs viedoklis par parametriem  $E[\lambda]$  un  $P[a \geq \lambda \geq b]$ . Eksperta viedoklis nav pieejams, tāpēc tiks apskatīti 6 dažādi gadījumi šo parametru izvēlei.

6. tabula

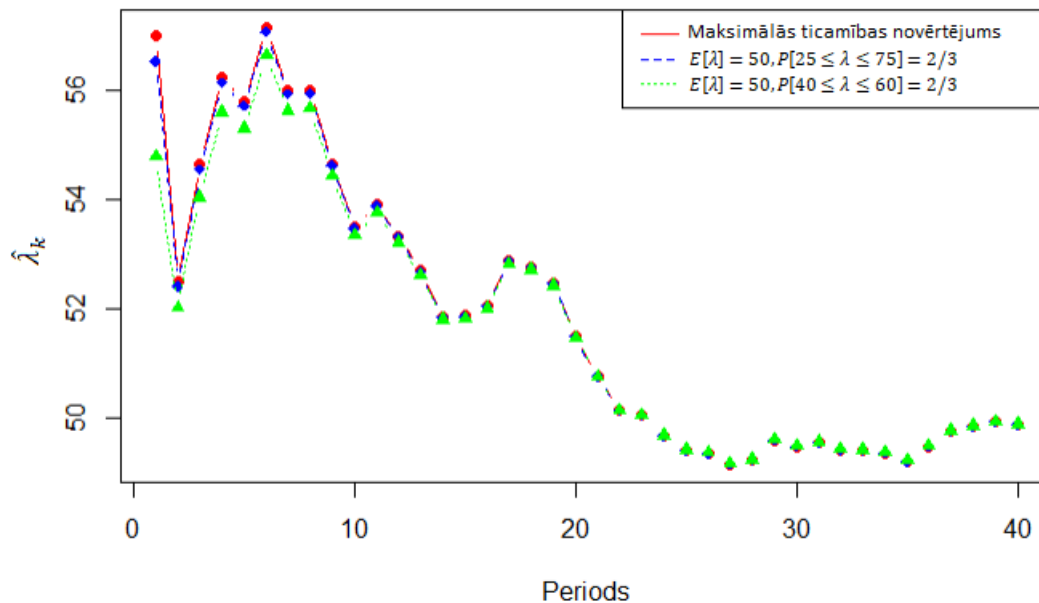
Ekspertu viedokļu varianti par zaudējumu biežumu

$E[\lambda]$	$P$	$a$	$b$
50	2/3	25	75
50	2/3	40	60
40	2/3	20	60
40	2/3	35	45
60	2/3	30	90
60	2/3	50	70

Šādas vērtības tika izvēlētas, jo tās ir tuvu maksimālās ticamības novērtējumam  $\lambda = 49.88$ ,

tāpēc eksperta viedoklis variē ap šo vērtību. Kā arī intervāliem ir divi veidi: ciešāks ap matemātisko cerību un izstieptāks.

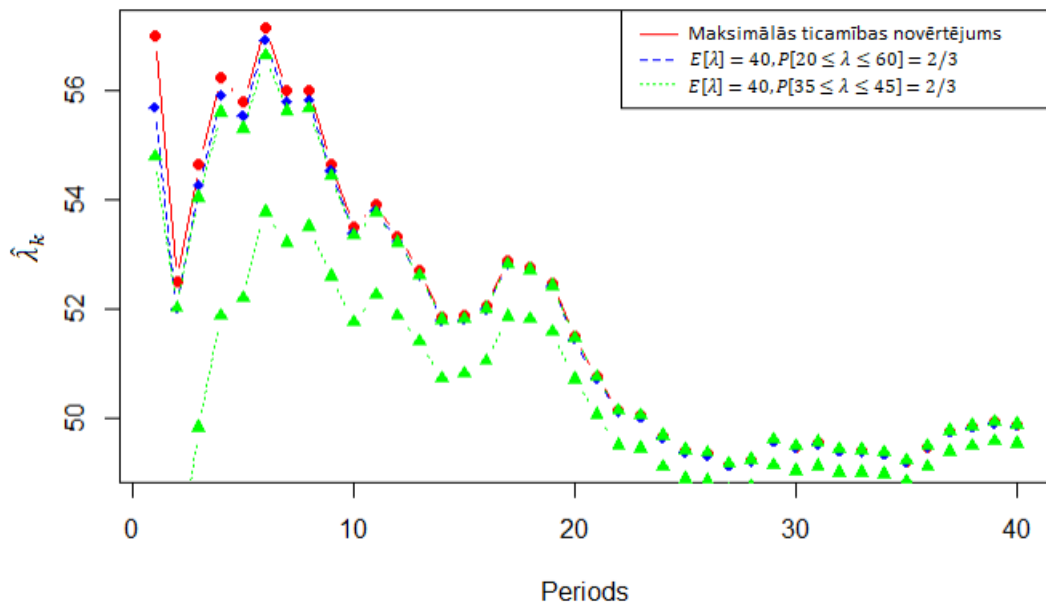
Ar (4.3) formulām tiek izteikti  $\alpha$  un  $\beta$ , izmantojot funkciju *unirroot*. Tad, izmantojot (3.1) formulas, tiek aprēķināti aposteriorā sadalījuma parametri  $\hat{\lambda}_k$  katram eksperta viedokļa veidam katram periodam  $k = 1, \dots, 40$ , tas nozīmē, ka katru gadu ekspertu viedoklis tiks atjaunots, papildinot datus ar jaunā ienākošā perioda datiem. Kā arī izrēķina maksimālo ticamības novērtējumu ar formulu  $\tilde{\lambda}_k = \frac{1}{2} \sum_{i=1}^k N_i$  rezultātu salīdzināšanai.  $\mathbf{N}_i$  ir vektors, kas sastāv no zaudējumu biežuma katrā periodā, kas ir sadalīts pēc *Poisson*( $\lambda = 49.88$ ).



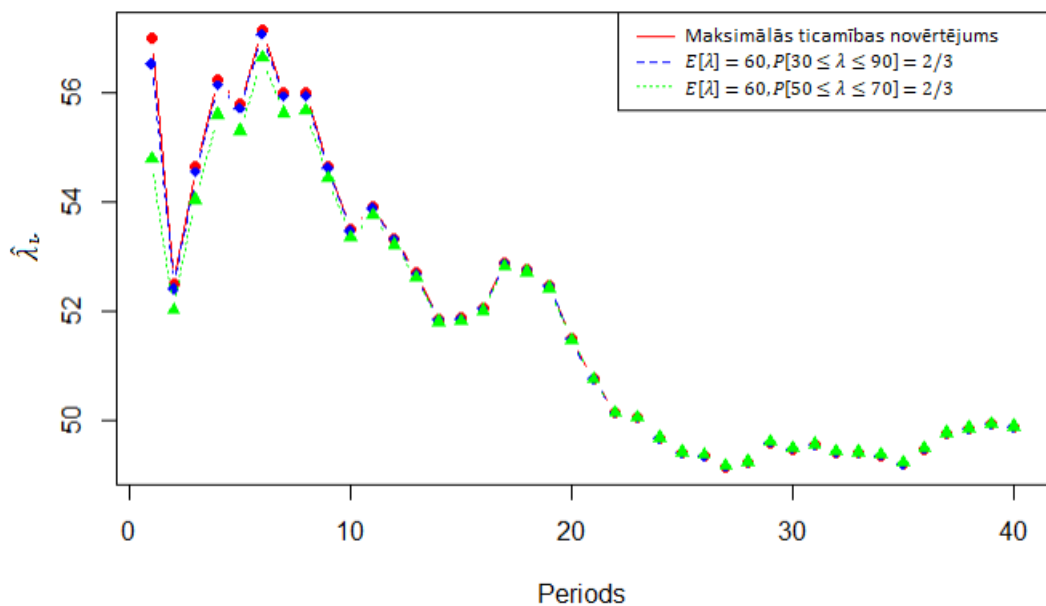
7. att.: Beijesa metodes (zilā līnija –1.variants, zaļā– 2.variants) un maksimālā ticamības intervāla(sarkanā līnija)  $\hat{\lambda}_k$  novērtējums 40 periodiem ar *lossdata* datu zaudējumu biežumu

7. attēlā var redzēt, ka, novērtējot  $E[\lambda]$  ļoti tuvu reālajai vērtībai, abu metožu rezultāti ir gandrīz vienādi, un pēc aptuveni 18 periodiem tie ir vienādi. Plašākā intervāla vērtības ir līdzīgas ar maksimālās ticamības novērtējumu, bet ciešākā sākumā ir mazliet lielākas.

8. attēlā 3.varianta ekspertu viedokļa rezultāti ir gandrīz vienādi ar maksimālās ticamības metodes novērtējumiem. Savukārt, 4. variantā vērtības ir mazākas, bet pamazām arī konverģē uz maksimālās ticamības novērtējumiem. Šī variantā parametri bija  $E[\lambda] = 40$ ,  $P = 2/3$ ,  $a = 35$ ,  $b = 45$ . Maza novirze iespējams rodas, jo intervālā  $[35, 45]$  neiekļūst novērtētā  $\lambda = 49.88$  vērtība. Intervāla maksimums ir mazāks par parametru, līdz ar to arī novērtējums atrodas zemāk.



8. att.: Beijesa metodes (zilā līnija –3.variants, zaļā– 4.variants) un maksimālā ticamības intervāla(sarkanā līnija)  $\hat{\lambda}_k$  novērtējums 40 periodiem ar *lossdata* datu zaudējumu biežumu



9. att.: Beijesa metodes (zilā līnija –5.variants, zaļā– 6.variants) un maksimālā ticamības intervāla(sarkanā līnija)  $\hat{\lambda}_k$  novērtējums 40 periodiem ar *lossdata* datu zaudējumu biežumu

9. attēls ir līdzīgs 7. attēlam pēc savas struktūras. Visas metožu vērtības katrā periodā konverģē uz vienu un ir gandrīz vienādas.

No grafikiem var secināt, ka ne tik svarīga ir  $E[\lambda]$  kā  $[a, b]$  vērtību izvēle. Ja sākumā novērtētā  $\lambda$  vērtība neietilpst intervālā, tad grafiks ir nedaudz novirzīts attiecīgi intervāla attiecībai pret parametru.

Lai novērtētu zaudējumu smagumu ar Beijesa metodi, ir nepieciešams zināt eksperta viedokli par  $E[M]$  un  $P[a \leq M \leq b]$ . Šajā gadījumā  $\sigma$  jau ir dots, un  $\sigma = 0.76$ . Tā kā eksperta viedoklis nav pieejams, tad atkal tiks izvirzīti daži tā varianti.

7. tabula

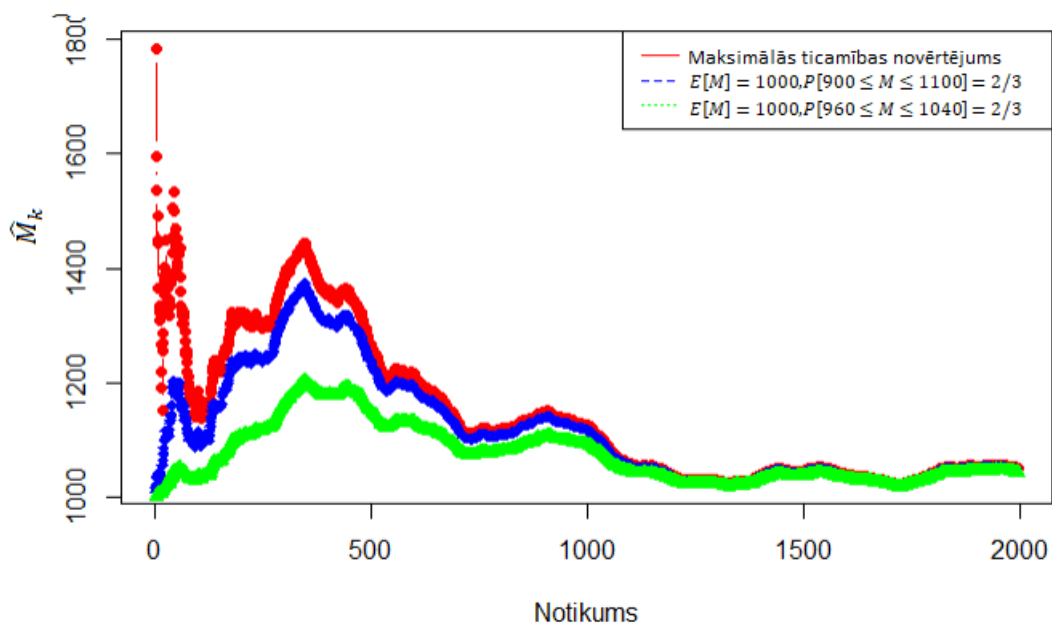
**Ekspertu viedokļu varianti par zaudējumu smagumu**

$E[M]$	$P$	$a$	$b$
1000	2/3	900	1100
1000	2/3	960	1040
1125	2/3	1000	1250
1125	2/3	1100	1150
900	2/3	800	1000
900	2/3	875	925

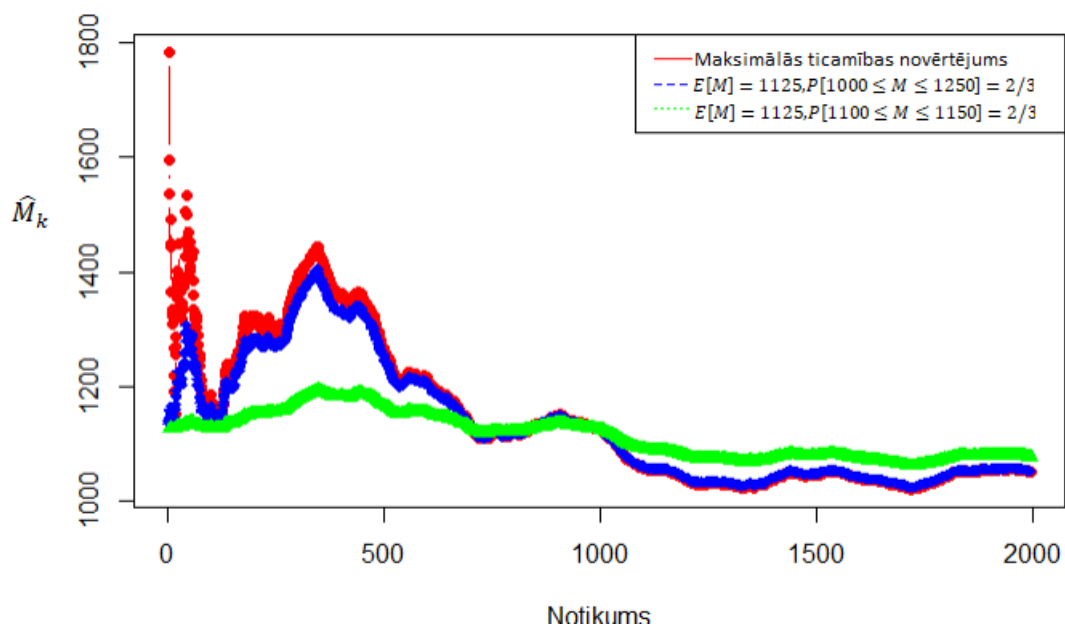
$E[M]$  vērtības tika izvēlētas ap zaudējumu smaguma vidējo vērtību visā datu kopā, kas bija vienāda ar 1002, 51.

Izmantojot (4.2) formulas programmā *Rstudio*, var aprēķināt  $\mu_0$  un  $\sigma_0$  katram ekspertu viedokļu variantam. Vēlāk, izmantojot (3.2) formulas, var izrēķināt aposterioros parametrus  $(\hat{\mu}_0)_k$ ,  $(\hat{\sigma}_0)_k$  arī katram ekspertu viedokļu variantam, kad ir zināmi novērojumi  $X_k$ ,  $k = 1, 2, \dots, 1995$ . Salīdzināšanai  $M$  tika novērtēts arī ar maksimālās ticamības metodes palīdzību. Tās formula ir  $\tilde{M}_k = e^{\mu_k^* + \frac{1}{2}\sigma^2}$ , kur  $\ln X_k \sim N(\mu_k^*, \sigma = 0.76)$ .

10. attēlā visi novērtējumi konverģē uz vienu, īsto vērtību  $M = 1002.51$ . Maksimālās ticamības metodes vērtības sākotnēji ir lielākas nekā Beijesa metodes vērtības, bet tām ir līdzīgs svārstīgums.

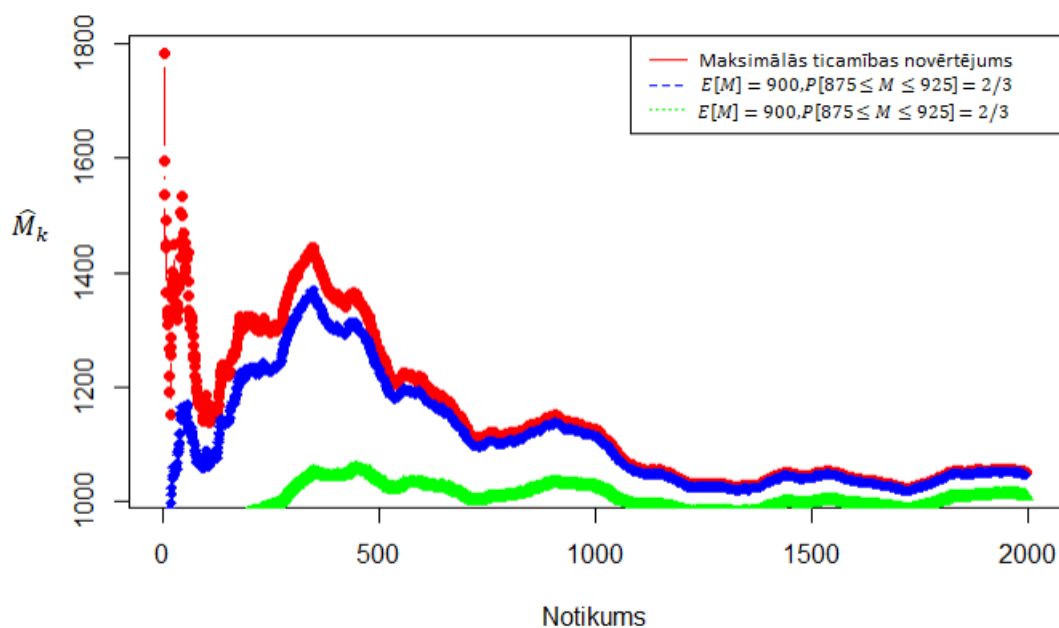


10. att.: Bejesa metodes (zilā līnija –1.variants, zaļā– 2.variants) un maksimālā ticamības intervāla(sarkanā līnija)  $\hat{M}_k$  novērtējums 1995 notikumiem ar *lossdata* datu zaudējumu smagumu



11. att.: Bejesa metodes (zilā līnija –3.variants, zaļā– 4.variants) un maksimālā ticamības intervāla(sarkanā līnija)  $\hat{M}_k$  novērtējums 1995 notikumiem ar *lossdata* datu zaudējumu smagumu

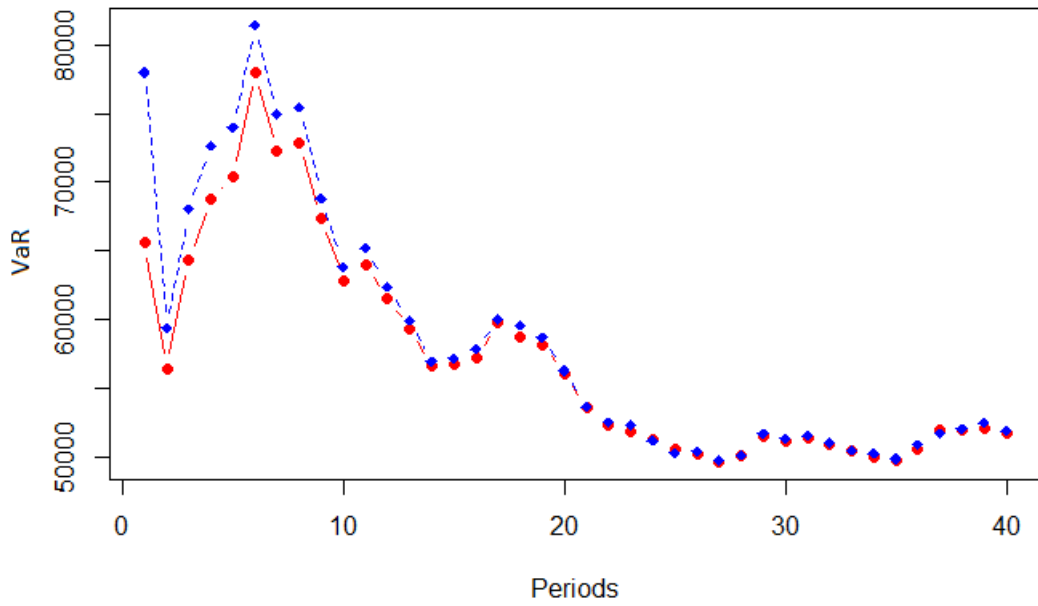
11. attēlā zaļā līnija, kura attēlo Bejesa metodes novērtējumu ar 4.varianta parametriem, atšķiras no pārējām, tā ir vienmērīgāka, bet, izskatās, ka lēnām konverģē uz vērtībām, kuras pieņem pārējās vērtības. 4. variantā intervālā [1100, 1150] neietiplst  $M = 1002.51$ .



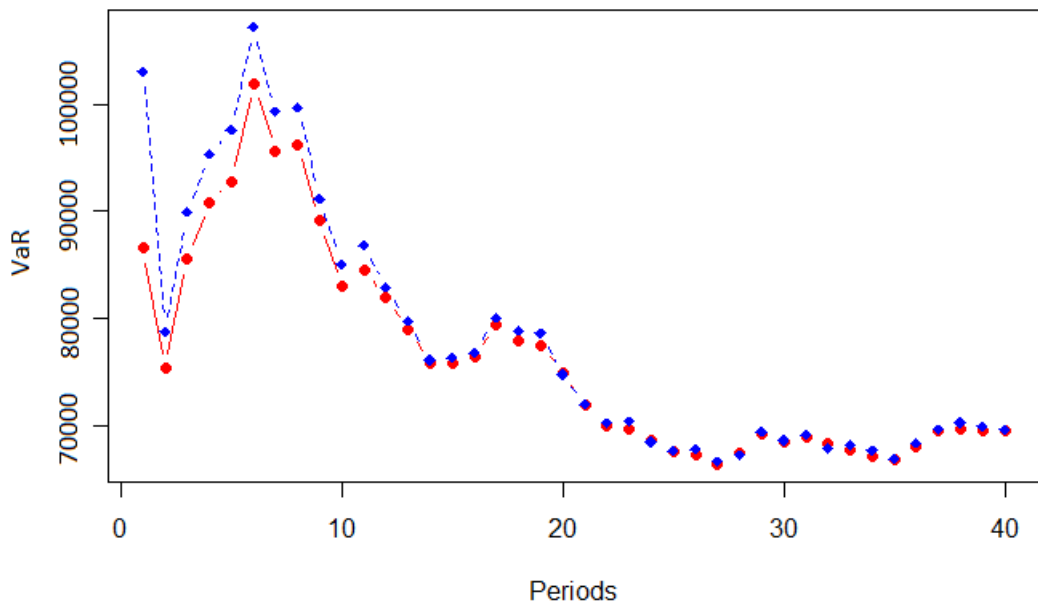
12. att.: Beijesa metodes (zilā līnija –5.variants, zaļā– 6.variants) un maksimālā ticamības intervāla(sarkanā līnija)  $\hat{M}_k$  novērtējums 1995 notikumiem ar *lossdata* datu zaudējumu smagumu

12. attēlā 5.varianta Beijesa metodes  $M$  novērtējums ir līdzīgs ar maksimālās ticamības novērtējumam, ar laiku tie konverģē uz vienu vērtību. Savukārt, 6. variants pieņem mazākas vērtības nekā pārējie. Šī variantā intervāls bija  $[875, 925]$ , kurā neietilpst īstā  $M$  vērtība.

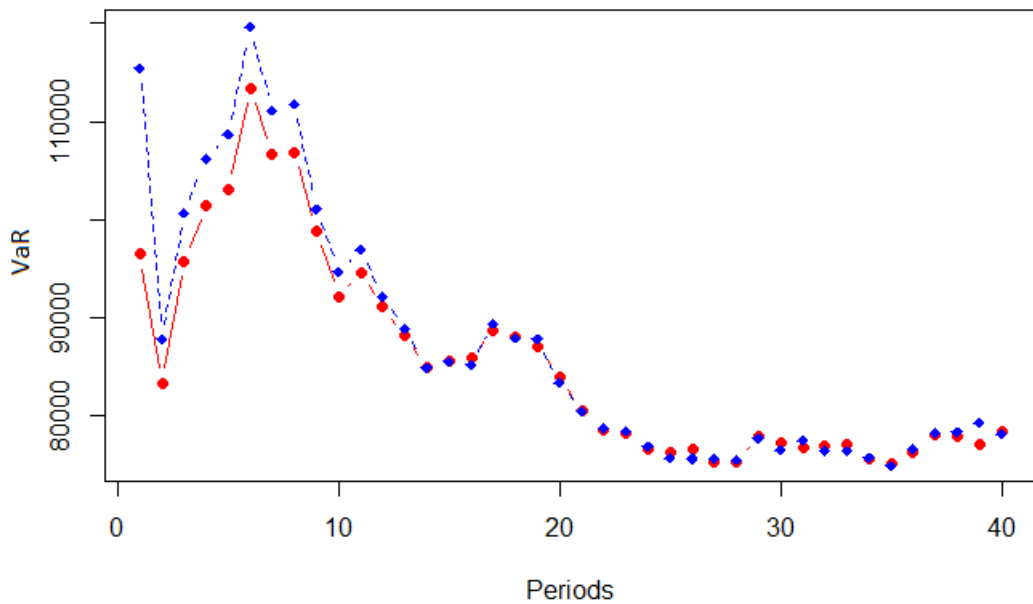
**VaR novērtēšana** Tiks novērtētas *VaR* vērtības katra perioda beigās pie ticamības līmeņiem 50%, 95%, 99% un 0.999. Tam tiks izmantota Montekarlo simulāciju metode, kuras soļi ir aprakstīti 5.1. nodaļā. Tiks apskatīti gan parametri, kuri ir novērtēti ar Beijesa metodi, gan ar maksimālās ticamības metodi. Beijesa metodei tiks ņemti parametri, kuri tika novērtēti ar 1.varianta ekspertu viedokli gan zaudējumu biežumam, gan smagumam.



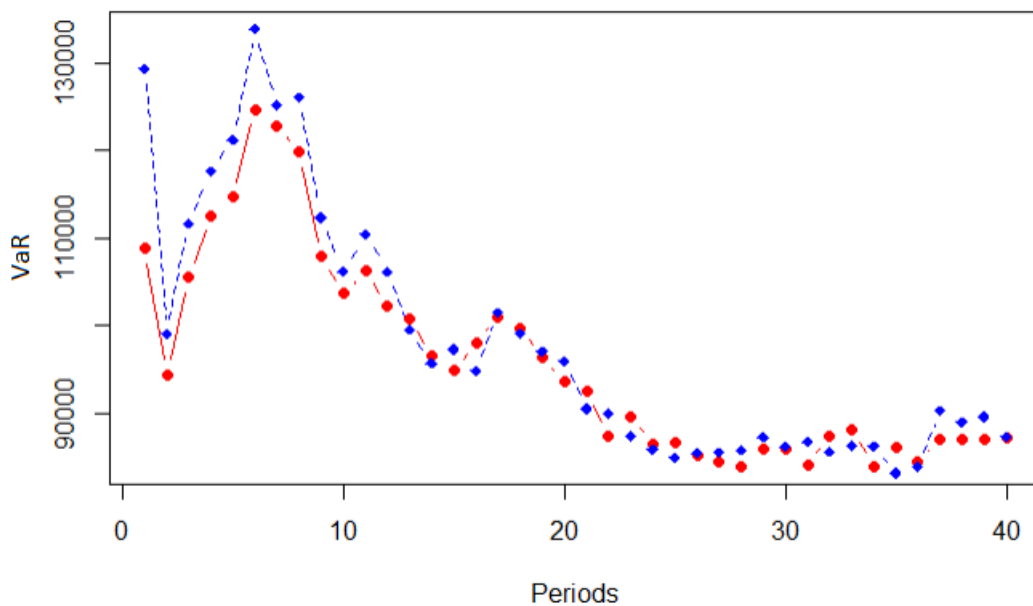
13. att.:  $VaR_{50\%}$  (sagaidāmie zaudējumi) parametriem, kuri tika novērtēti ar Beijesa metodi (sarkanā) un maksimālās ticamības metodi (zilā) katra perioda beigās



14. att.:  $VaR_{95\%}$  (sagaidāmie zaudējumi) parametriem, kuri tika novērtēti ar Beijesa metodi (sarkanā) un maksimālās ticamības metodi (zilā) katra perioda beigās



15. att.:  $VaR_{99\%}$  (sagaidāmie zaudējumi) parametriem, kuri tika novērtēti ar Beijesa metodi (sarkanā) un maksimālās ticamības metodi (zilā) katra perioda beigās



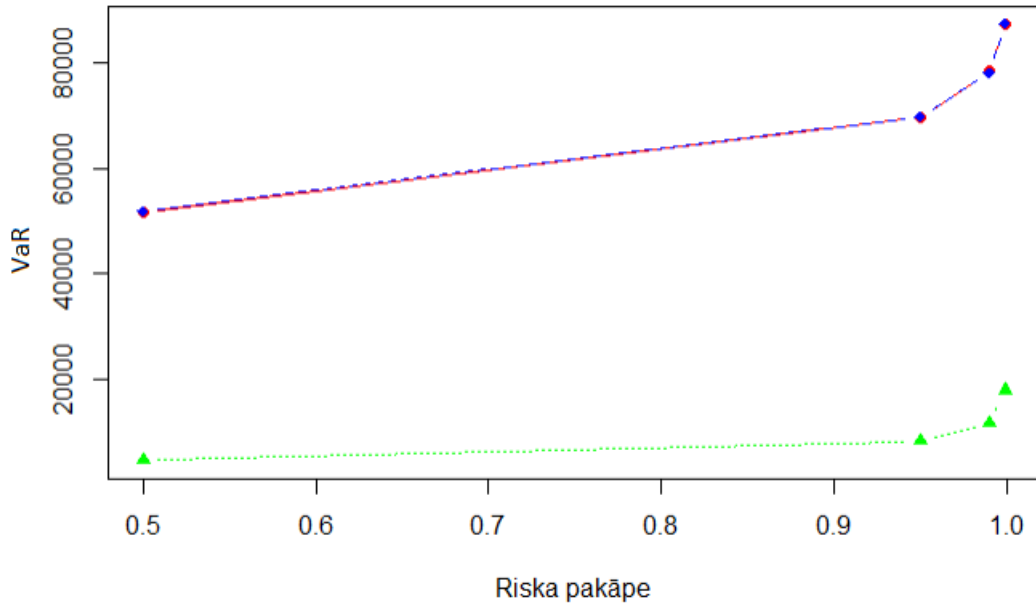
16. att.:  $VaR_{99.9\%}$  (sagaidāmie zaudējumi) parametriem, kuri tika novērtēti ar Beijesa metodi (sarkanā) un maksimālās ticamības metodi (zilā) katra perioda beigās

13., 14., 15., 16. attēlos ir redzamas riskam pakļautās vērtības pie riska līmeņiem 50%, 95%, 99% un 0.999 attiecīgi katra perioda beigās. Kopējais periodu skaits datiem ir 40. Var redzēt, ka grafiki pēc savas struktūras ir līdzīgi. Pieaugot periodu skaitam,  $VaR$  vērtība samazinās, abu metožu rezultāti konverģē uz vienu. Aptuveni līdz 10. periodam  $VaR$  vērtības, kas tika novērtētas ar parametriem, kuri tika novērtēti pēc Beijesa metodes, ir mazākas, salīdzinot ar

maksimālās ticamības metodes vērtībām.

Var novērot, ka, pieaugot riska līmenim, pieaug arī riskam pakļautā vērtība.

Tālāk tiks salīdzinātas aproksimācijas formulas  $VaR$  (skatīt 2. tabulu) vērtības ar Montekarlo simulāciju vērtībām 40. perioda beigās pie četriem riska līmeņiem 50%, 95%, 99% un 99.9%.



**17. att.: Montekarlo simulāciju rezultāti, novērtēti ar Beijesa metodi(sarkanā) un maksimālās ticamības metodi (zilā), un aproksimācijas formulas rezultāti(zaļā) pēdējā perioda beigās**

17. attēlā var redzēt, ka aproksimācijas formulas rezultāti atrodas ļoti tālu no Montekarlo simulāciju rezultātiem. Tas nozīmē, ka aproksimācijas formula diemžēl nav izmantojama.

## Secinājumi

Bakalaura darbā tika apskatīta operacionālā riska zaudējumu statistiskā modelēšana, aprioro parametru novērtēšana, izmantojot Beijesa statistikas metodes. Tika gūts priekšstats par Beijesa metodes pamatprincipiem, tuvāk apskatītas Montekarlo simulācijas un maksimālās ticamības funkcijas dažādiem sadalījumiem.

Beijesa metode tika izmantota, lai novērtētu operacionālā riska zaudējumu smagumu un biežumu, tiktu ņemti vērā ne tikai iekšējie dati, bet arī ekspertu viedoklis.

Kopumā tika apskatīti 3 saistītie apriorie sadalījumi, katram simulēts piemērs, lai gūtu labāku izpratni par Beijesa metodes darbību, un salīdzinātu to ar maksimālās ticamības metodes novērtējumiem. Grafikos varēja novērot, ka parametri, kuri ir novērtēti ar Beijesa metodi, ir daudz vienmērīgāki, kamēr maksimālā ticamības intervāla novērtējumi svārstās ap reālo vērtību pie maza notikumu skaita  $N$ .

Apriorā sadalījuma parametri tika novērtēti arī datu piemēram, izmantojot lognormālo–normālo saistīto sadalījumu pāri zaudējumu smagumam un Puasona–gamma sadalījumu zaudējumu biežumam. Tā kā eksperta viedoklis nebija pieejams, katram sadalījumu pārim tika simulēti 6 eksperta viedokļa varianti. No tiem varēja secināt, ka ne tik svarīgi ir pēc iespējas precīzāk novērtēt aposteriorā sadalījuma parametru, cik svarīgi ir noteikt intervālu  $[a, b]$  tā, lai tajā ietilptu reālā parametra vērtība.

Šim piemēram tika veiktas Montekarlo simulācijas gan parametriem, kuri tika novērtēti ar Beijesa metodi, gan ar maksimālo ticamības metodi katra perioda beigās, lai novērtētu  $VaR$  četros riska līmeņos 50%, 95%, 99%, 99.9%. Abu metožu iegūtie  $VaR$  novērtējumi bija līdzīgi, Beijesa metodes rezultāti bija nedaudz mazāki, salīdzinot ar otru. Tomēr tie konverģēja uz vienu vērtību, palielinoties periodam. Tad rāda, ka Beijesa metode var būt īpaši noderīga, kamēr vēl nav pieejams pietiekams daudzums novērojumu datu vai tie nav uzticami.

Pēdējā perioda beigās tika salīdzināti abu metožu Montekarlo simulāciju  $VaR$  rezultāti ar aproksimācijas formulām. Aproksimācijas formulas sniedza daudz mazākas  $VaR$  vērtības, un tika secināts, ka tās nav izmantojamas praksē.

Kā turpmākos soļus, pētīt tālāk operacionālā riska zaudējumu statistisko modelēšanu, būtu svarīgi iepazīties ar modeli, kas sevī iekļauj ne tikai iekšējos datus un ekspertu viedokli, bet arī ārējos datus.

## Izmantotā literatūra un avoti

- [1] *Finanšu un kapitāla tirgus komisijas 2006. gada 28. jūlija ieteikumi Nr.125 "Operacionālā riska pārvaldīšanas ieteikumi"*. Latvijas Vēstnesis, 2006.
- [2] James O Berger. Prior information and subjective probability. In *Statistical Decision Theory and Bayesian Analysis*, pages 74–117. Springer, 1985.
- [3] José M Bernardo. A bayesian mathematical statistics primer. In *Proceedings of the Seventh International Conference on Teaching Statistics. CD ROM. International Association for Statistical Education, Salvador (Bahia), Brazil*, volume 8, 2006.
- [4] Klaus Böcker and Claudia Klüppelberg. Operational var: a closed-form approximation. *Risk*, pages 90–93, 2005.
- [5] Hans Bühlmann and Alois Gisler. *A course in credibility theory and its applications*, volume 317. Springer, 2005.
- [6] Stephen E Fienberg. When did bayesian inference become "bayesian"? *Bayesian analysis*, 1(1):1–40, 2006.
- [7] Charles M Goldie and Claudia Klüppelberg. Subexponential distributions. *A practical guide to heavy tails: statistical techniques and applications*, pages 435–459, 1998.
- [8] Pavel V Shevchenko and Mario V Wüthrich. The structural modelling of operational risk via bayesian inference: Combining loss data with expert opinions. *arXiv preprint arXiv:0904.1067*, 2009.
- [9] Karline Soetaert. Package rootsolve: roots, gradients and steady-states in r. *Google Scholar*, 2014.
- [10] Yiu-Kuen Tse. *Nonlife actuarial models: theory, methods and evaluation*. Cambridge University Press, 2009.
- [11] Christina Zou, Marius Pfeuffer, Matthias Fischer, Nina Buoni, Kristina Dehler, Nicole Derfuss, Benedikt Graswald, Linda Moestel, Jixuan Wang, and Leonie Wicht. Statistical methods for modelling operational risk.

## 1. Pielikums. Programmas *Rstudio* kods

```
1 library(OpVaR)
2 library(dplyr)
3 library(xtable)
4 library(rootSolve)
5 library(Pareto)
6 library(fitdistrplus)
7
8 #Puasona-Gamma funkcija
9 p_g_fun <- function(E_lambda, Pr, a, b, freq_tab, v){
10   funkcija <- function(alpha) {pgamma(b, shape = alpha, scale = E_lambda/
11     alpha) - pgamma(a, shape = alpha, scale = E_lambda/alpha) - Pr}
12
13   alpha <- uniroot(funkcija, upper = 100, lower = 0.5)
14   alpha <- round(alpha$root, 3)
15   beta <- round(E_lambda/alpha, 3)
16
17   alpha_k <- rep(0, length(freq_tab))
18   beta_k <- rep(0, length(freq_tab))
19
20   alpha_k[1] <- alpha + freq_tab[1]
21   beta_k[1] <- round(beta/(1+beta), 3)
22
23   for (i in 2:length(freq_tab)) {
24     alpha_k[i] = alpha_k[i-1] + freq_tab[i]
25     beta_k[i] = beta_k[i-1] / (1 + beta_k[i-1])
26   }
27
28   lambda_jumtinu <- round(alpha_k * beta_k, 3)
29   tab <- data.frame(alpha_k, round(beta_k, 3), lambda_jumtinu)
30
31   if (v==1){return(lambda_jumtinu)}
32   else {return(tab)}
33 }
34 #lognormala-normala funkcija
35 l_n_fun <- function(sigma, E_M, Pr, a, b, not, v){
36   mu_0 <- function(sigma_0) {log(E_M) - 1/2 * sigma^2 - 1/2 * sigma_0^2}
37   funkcija <- function(sigma_0)
```

```

38 {mu_0 = log(E_M) - 1/2 * sigma^2 -1/2 * sigma_0^2
39 pnorm((log(b) - 1/2 * sigma^2 - mu_0)/sigma_0) - pnorm((log(a) - 1/2 *
      sigma^2 -mu_0)/sigma_0) - Pr}
40
41 sigma_0 <- seq(0.01, 10, 0.05)
42
43 sigma_0 <- uniroot(funkcija, upper = 50, lower = 0.01)$root
44 mu_0 <- mu_0(sigma_0)
45
46 mu_0_k <- rep(0, length(not))
47 sigma_0_k <- rep(0, length(not))
48
49 mu_0_k[1] <- (mu_0 + (sigma_0^2 / sigma^2) * not[1]) / (1 + (sigma_0^2 /
      sigma^2))
50 sigma_0_k[1] <- sigma_0^2 / (1 + (sigma_0^2/sigma^2))
51 sigma_0_k[1] <- sqrt(sigma_0_k[1])
52
53 for (i in 2:length(not)) {
54   mu_0_k[i] = (mu_0_k[i-1] + (sigma_0_k[i-1]^2 / sigma^2) * not[i]) / (1
      + (sigma_0_k[i-1]^2 / sigma^2))
55   sigma_0_k[i] = sigma_0_k[i-1]^2 / (1+(sigma_0_k[i-1]^2 / sigma^2))
56   sigma_0_k[i]=sqrt(sigma_0_k[i])
57 }
58 M_0 <- exp(mu_0_k+1/2*sigma_0_k^2+1/2*sigma^2)
59 tab <- data.frame(mu_0_k, sigma_0_k, M_0)
60 if (v==1){ return(M_0)}
61 else if (v==2) {return(tab)}
62 else {return(mu_0_k)}
63
64 }
65
66 #Pareto - gamma funkcija
67 par_g_fun <- function(B, E, Pr, a,b, not, v){
68   model <- function(x){
69     F1 <- x[1] * x[2] * ((1-pgamma(B, shape = x[1] + 1, scale = x[2]))/(1 -
      pgamma(B, shape = x[1], scale = x[2]))) - E
70     F2 <- (pgamma(b, shape = x[1], scale = x[2]) - pgamma(a, shape = x[1],
      scale = x[2])) / (1- pgamma(B, shape = x[1], scale = x[2])) - Pr
71     c(F1=F1, F2=F2)
72   }

```

```

73 x <- multiroot(f = model, start = c(23,0.2))$root
74 alpha <- x[1]
75 beta <- x[2]
76
77 alpha_k <- rep(0,15)
78 beta_k <- rep(0,15)
79 alpha_k[1] <- alpha+1
80 beta_k[1] <- beta^(-1) + log(X_k[1])
81 beta_k[1] <- beta_k[1]^(-1)
82
83 for (i in 2:length(X_k)) {
84   alpha_k[i] = alpha_k[i-1] + 1
85   beta_k[i] = beta_k[i-1]^(-1) + log(X_k[i]/L)
86   beta_k[i] = beta_k[i]^(-1)
87 }
88
89 tail <- rep(0,15)
90 for (i in 1:length(X_k)) {
91   tail[i] = alpha_k[i] * beta_k[i] * (1-pgamma(B, shape = alpha_k[i]+1,
92     scale = beta_k[i]))/(1-pgamma(B, shape = alpha_k[i], scale = beta_k[
93     i]))
94 }
95
96 tab <- data.frame(alpha_k, beta_k, tail )
97
98 if (v==1){return(tail)}
99 else {return(tab)}
100 }
101 #Puasona- gamma piemers
102 lambda <- 0.6
103 c <- c(0, 0, 0, 0, 1, 0, 1, 1, 1, 0, 2, 1, 1, 2, 0)
104
105 lambda_jumtinu <- p_g_fun(E_lambda = 0.5, Pr=2/3, a=0.25, b=0.75, freq_tab
106   = c, v=1)
107 tab <- p_g_fun(E_lambda = 0.5, Pr=2/3, a=0.25, b=0.75, freq_tab = c, v=0)
108 print(xtable(tab, type = "latex"))
109 mle <- rep(0,length(c))

```

```

110 for (i in 1:length(c)) {
111   mle[i] = 1/i * sum(c[1:i])
112 }
113
114 plot(mle, type = "b", pch = 19, col = "red", xlab = "Periods", ylab = "")
115 lines(lambda_jumtinu, pch = 18, col = "blue", type = "b", lty = 2)
116
117
118 #Lognormalais-normalais piemers
119 sigma <- 2
120 set.seed(123)
121 not <- rlnorm(15, 0.3, 2)
122 not <- log(not)
123
124 mle <- rep(0, length(not))
125 mle[1] <- exp(not[1]+1/2*sigma^2)
126 for (i in 2:length(not)) {
127   x <- fitdist(not[1:i], "norm", fix.arg = list(sd=2))
128   x <- x$estimate
129   mle[i] = exp(x+1/2*sigma^2)
130 }
131 M_0 <- l_n_fun(sigma=2, E_M=10, Pr=2/3, a=8, b=12, not=not, v=1)
132
133 tab <- l_n_fun(sigma=2, E_M=10, Pr=2/3, a=8, b=12, not=not, v=2)
134 print(xtable(tab, type = "latex"))
135
136 plot(mle, type = "b", pch = 19, col = "red", xlab = "Periods", ylab = "")
137 lines(M_0, pch = 18, col = "blue", type = "b", lty = 2)
138
139
140 #Pareto-gamma piemers
141 set.seed(120)
142 L <- 1
143 xi <- 4
144 X_k <- rPareto(15, L, xi)
145
146 mle <- rep(0, 15)
147 for (i in 1:length(mle)) {
148   mle[i] = 1/i * sum(log(X_k[1:i]/L))
149   mle[i] = mle[i]^(-1)

```

```

150 }
151
152 tail <- par_g_fun(B=2, E=5, Pr=2/3, a=4, b=6, not = X_k, v=1)
153 tab <- par_g_fun(B=2, E=5, Pr=2/3, a=4, b=6, not = X_k, v=0)
154 print(xtable(tab, type="latex"))
155 plot(mle, type = "b", pch = 19, col = "red", xlab = "periods", ylab = "")
156 lines(tail, pch = 18, col = "blue", type = "b", lty = 2)
157
158
159 #Datu piemers
160 data("lossdat")
161
162 opriskmodel=list()
163 for(i in 1:length(lossdat)){
164   opriskmodel[[i]]=list()
165 }
166
167 for (i in 1:length(opriskmodel)) {
168   opriskmodel[[i]]$freqdist=fitFreqdist(lossdat[[i]], "pois")
169   opriskmodel[[i]]$sevdist=fitPlain(lossdat[[i]], "lnorm")
170 }
171
172 sev_1 <- unlist(goftest(lossdat[[1]], opriskmodel[[1]]$sevdist))
173 sev_2 <- unlist(goftest(lossdat[[2]], opriskmodel[[2]]$sevdist))
174 sev_3 <- unlist(goftest(lossdat[[3]], opriskmodel[[3]]$sevdist))
175 sev_4 <- unlist(goftest(lossdat[[4]], opriskmodel[[4]]$sevdist))
176
177 testi <- data.frame(1:4, round(as.numeric(c(sev_1[2], sev_2[2], sev_3[2], sev_
178   4[2])), 5))
179 colnames(testi) <- c("Datu kopas nr.", "P-v rt ba")
180 testi
181
182 print(xtable(testi, type = "latex"))
183
184 freq_3 <- unlist(goftest(lossdat[[3]], opriskmodel[[3]]$freqdist))
185 testi2 <- data.frame(1, round(as.numeric(freq_3[3]), 4))
186 colnames(testi2) <- c("Datu kopas nr.", "P-v rt ba")
187 testi2
188
189 lambda <- opriskmodel[[3]]$freqdist
190 lambda <- lambda[[2]]

```

```

189
190 ln_vert <- unlist(opriskmodel[[3]]$sevdist)
191 mu <- ln_vert$par5
192 sigma <- ln_vert$par6
193
194 #zaudejumu biezums
195 freq_tab <- lossdat[[3]]
196 freq_tab <- tabulate(freq_tab$Period)
197
198 mle <- rep(0, length(freq_tab))
199 for (i in 1:length(freq_tab)) {
200   mle[i] = 1/i * sum(freq_tab[1:i])
201 }
202
203 lambda_jumtinu_1 <- p_g_fun(E_lambda = 50, Pr = 2/3, a = 25, b = 75, freq_
    tab, v=1)
204 lambda_jumtinu_11 <- p_g_fun(E_lambda = 50, Pr = 2/3, a = 40, b = 60, freq_
    tab, v=1)
205 plot(mle, type = "b", pch = 19, col = "red", xlab = "Periods", ylab = " ")
206 lines(lambda_jumtinu_1, pch = 18, col = "blue", type = "b", lty = 2)
207 lines(lambda_jumtinu_11, pch = 17, col = "green", type = "b", lty = 3)
208
209
210 lambda_jumtinu_2 <- p_g_fun(E_lambda = 40, Pr = 2/3, a = 20, b = 60, freq_
    tab, v=1)
211 lambda_jumtinu_22 <- p_g_fun(E_lambda = 40, Pr = 2/3, a = 35, b = 45, freq_
    tab, v=1)
212 plot(mle, type = "b", pch = 19, col = "red", xlab = "Periods", ylab = " ")
213 lines(lambda_jumtinu_2, pch = 18, col = "blue", type = "b", lty = 2)
214 lines(lambda_jumtinu_22, pch = 17, col = "green", type = "b", lty = 3)
215
216 lambda_jumtinu_3 <- p_g_fun(E_lambda = 60, Pr = 2/3, a = 30, b = 90, freq_
    tab, v=1)
217 lambda_jumtinu_33 <- p_g_fun(E_lambda = 60, Pr = 2/3, a = 50, b = 70, freq_
    tab, v=1)
218 plot(mle, type = "b", pch = 19, col = "red", xlab = "Periods", ylab = " ")
219 lines(lambda_jumtinu_1, pch = 18, col = "blue", type = "b", lty = 2)
220 lines(lambda_jumtinu_11, pch = 17, col = "green", type = "b", lty = 3)
221
222 #zaudejumu smagums

```

```

223 severity <- lossdat[[3]] %>%
224   group_by(Period) %>%
225   summarise(mean(Loss))
226 mean(severity$`mean(Loss)` )
227
228 not <- lossdat[[3]]$Loss
229 not <- log(not)
230
231 mle <- rep(0, length(not))
232 mle[1] <- exp(not[1]+1/2*sigma^2)
233 for (i in 2:length(not)) {
234   x <- fitdist(not[1:i], "norm", fix.arg = list(sd=sigma))
235   x <- x$estimate
236   mle[i] = exp(x+1/2*sigma^2)
237 }
238
239 ml <- l_n_fun(sigma, 1000, 2/3, 900,1100,not, v=1)
240 ml1 <- l_n_fun(sigma, 1000, 2/3, 960, 1040, not, v=1)
241 plot(mle, type = "b", pch = 19, col = "red", xlab = "Notikums", ylab = " ")
242 lines(ml, pch = 18, col = "blue", type = "b", lty = 2)
243 lines(ml1, pch = 17, col = "green", type = "b", lty = 3)
244
245 m2 <- l_n_fun(sigma, 1125, 2/3, 1000, 1250, not, v=1)
246 m22 <- l_n_fun(sigma, 1125, 2/3, 1100,1150, not, v=1)
247 plot(mle, type = "b", pch = 19, col = "red", xlab = "Notikums", ylab = " ")
248 lines(m2, pch = 18, col = "blue", type = "b", lty = 2)
249 lines(m22, pch = 17, col = "green", type = "b", lty = 3)
250
251 m3 <- l_n_fun(sigma, 900, 2/3, 800, 1000, not, v=1)
252 m33 <- l_n_fun(sigma, 900,2/3,875, 925, not, v=1)
253 plot(mle, type = "b", pch = 19, col = "red", xlab = "Notikums", ylab = " ")
254 lines(m3, pch = 18, col = "blue", type = "b", lty = 2)
255 lines(m33, pch = 17, col = "green", type = "b", lty = 3)
256
257 #VaR aprekini
258 proc <- c(.5, .95, .99, .999) #izv las ticam bas interv.
259
260 #Log-norm l sadal juma Montekarlo simul cijas
261 ln_montekarlo <- function(x, mu, sigma){
262   severity <- vector("double", 10000)

```

```

263   for(i in seq_along(severity)){
264     severity[i] =sum(rlnorm(x[i], mu, sigma))
265   }
266 VaR_MK = quantile(severity , proc)
267 return (VaR_MK)
268 }
269
270 #Log-Norm l sad. ar Beijesa MK
271 ln_mk_B <- function(B, N_i, lambda, mu_0_k, sigma){
272   severity1 = vector("double", B)
273   nj=vector("double", length(N_i))
274   mu=vector("double", length(N_i))
275   for (i in 1:B) {
276     for (j in 1:length(N_i)) {
277       Z_j = vector("double", length(N_i))
278       nj[j]= rpois(1, lambda[j])
279       mu[j] = mu_0_k[sum(N_i[1:j])]
280       Z_j[j]=sum(rlnorm(nj[j],mu[j], sigma))
281     }
282     severity1[i] = sum(Z_j)
283   }
284 VaR_MK = quantile(severity1 , proc)
285 return (VaR_MK)
286 }
287
288 #Log-norm l sad. p c aproksim cijas formul m
289 VaR_form <- vector("double", length(proc))
290 xx <- vector("double", length(proc))
291 phi <- vector("double", length(proc))
292
293 ln_formula <- function(lambda, mu, sigma){
294   for (i in 1:length(proc)) {
295     xx[i]=(1 - proc[i]) / (lambda) #apr ina Phi iekavu v rt bu
296     phi[i]=qnorm(xx[i]) # kumulat v sad. f-ja norm lajam sad.
297     VaR_form[i] =exp(mu - sigma * phi[i]) #VaR aprosim cijas funkcija
298   }
299   return (VaR_form)
300 }
301
302 mu_nov <- l_n_fun(sigma, 1000, 2/3, 900,1100,not, v=0)

```

```

303
304 var_b_tab<- data.frame(matrix(0,
305                             nrow = length(freq_tab),
306                             ncol = length(proc)))
307 var_mk_tab <- var_b_tab
308
309 var_beijesa <- function(B, N_i, lambda, mu_0_k, sigma){
310   for (i in 1:length(N_i)) {
311     var_bei = rep(0, length(proc))
312     var_bei = ln_mk_B(B, N_i[1:i], lambda, mu_0_k, sigma)
313     var_b_tab[i, 1:length(proc)] = var_bei
314   }
315   return(var_b_tab)
316 }
317 tabula <- var_beijesa(10000, freq_tab, lambda_jumtinu_1, mu_nov, sigma)
318
319 #Parametri ar MLE
320 mu_mle <- rep(0,length(freq_tab))
321 lambda_mle <- rep(0,length(freq_tab))
322 mu_mle1 <- rep(0,length(not))
323
324 mu_mle1[1] <- not[1]
325 for (i in 2:length(not)) {
326   x <- fitdist(not[1:i], "norm", fix.arg = list(sd=sigma))
327   mu_mle1[i] <- x$estimate
328 }
329
330 for (i in 1:length(freq_tab)) {
331   lambda_mle[i] = 1/i * sum(freq_tab[1:i])
332   mu_mle[i] = mu_mle1[sum(freq_tab[1:i])]
333 }
334
335 var_mk <- function(B, lambda, mu, sigma){
336   for (i in 1:length(lambda)) {
337     x=rpois(B, lambda[i])
338     var_mkk= rep(0, length(proc))
339     var_mkk = ln_montekarlo(x, mu = mu_mle[i], sigma=sigma )
340     var_mk_tab[i, 1:length(proc)] = var_mkk
341   }
342   return(var_mk_tab)

```

```

343 }
344 tabula2 <- var_mk(10000, lambda_mle, mu_mle, sigma)
345
346 for (i in 1:length(proc)) {
347   plot( tabula[,i],ylim=c(min(tabula[,i]), max(tabula2[,i])),type = "b",pch
        = 19, col = "red", xlab = "Periods", ylab = " ")
348   lines(tabula2[,i],pch = 18, col = "blue", type = "b", lty = 2)
349 }
350
351 x <- c(tabula[40,]$X1, tabula[40,]$X2, tabula[40,]$X3, tabula[40,]$X4)
352 y <- c(tabula2[40,]$X1, tabula2[40,]$X2, tabula2[40,]$X3, tabula2[40,]$X4)
353 var_aproks <- ln_formula(lambda, mu, sigma)
354
355 plot( proc,x,ylim=c(min(var_aproks), max(y)),type = "b",pch = 19, col = "
    red", xlab = "Riska pak pe", ylab = " ")
356 lines(proc,y,pch = 18, col = "blue", type = "b", lty = 2)
357 lines(proc,var_aproks , pch = 17, col = "green", type = "b", lty = 3)

```

Bakalaura darbs "Operacionālā riska zaudējumu statistiskā modelēšana, izmantojot Beijesa statistikas metodes" izstrādāts LU fizikas, matemātikas un optometrijas fakultātē.

Ar savu parakstu apliecinu, ka pētījums veikts patstāvīgi, izmantoti tikai tajā norādītie informācijas avoti un iesniegtā darba elektroniskā kopija atbilst izdrukai.

Autors: Linda Mičule

---

(paraksts)

---

(datums)

Rekomendēju darbu aizstāvēšanai:

Vadītājs: Dr.math, pētniece Māra Delesa- Vēliņa

---

(paraksts)

---

(datums)

Recenzents: Rūdolfis Krēgers

---

(paraksts)

---

(datums)

Darbs iesniegts Matemātikas nodaļā 2022.gada 2. jūnijā

---

(Dekāna pilnvarotā persona: vecākā metodiķe Inita Šneidere)

Darbs aizstāvēts bakalaura gala pārbaudījuma komisijas sēdē

2022. gada 9. jūnijā