Simulacija rijetkih događaja

Vazdar, Vedrana

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO–MATEMATIČKI FAKULTET
MATEMATIČKI ODSJEK

Vedrana Vazdar

SIMULACIJA RIJETKIH DOGAĐAJA

Diplomski rad

Voditelj rada:
prof. dr. sc. Bojan Basrak

Zagreb, 2016.

Ovaj diplomski rad obranjen je dana _____ pred ispitnim povjerenstvom u sastavu:

1. _____, predsjednik
2. _____, član
3. _____, član

Povjerenstvo je rad ocijenilo ocjenom _____.

Potpisi članova povjerenstva:

1. _____
2. _____
3. _____

Sadržaj

Uvod

Analiza iznimno rijetkih događaja, odnosno precizna procjena njihovih vjerojatnosti, od velike je važnosti u mnogim područjima. Primjerice, imovina osiguravajućeg društva, tj. odnos između nakupljenih premija i zahtjeva za isplatom, modelira se stohastičkim procesom. Rijetki događaji od važnosti su vjerojatnost propasti u nekom zadanom vremenu i vjerojatnost ultimativne propasti, pri čemu pod propast podrazumijevamo da je iznos zahtjeva za isplatom veći od sakupljenog kapitala od premija. U telekomunikacijskim mrežama s komutacijom paketa, međuspremnici koji pohranjuju dijelove paketa su ograničene veličine, radi smanjenja varijacije kašnjenja prilikom slanja, primjerice, videa u realnom vremenu. U slučaju preljeva međuspremnika dolazi do gubitka podataka. Protok podataka kroz međuspremnik modelira se redovima čekanja te je od velike važnosti znati vjerojatnost gubitka. Komunikacijske sustave često karakteriziraju rijetki događaji, poput nedostupnosti sustava (vjerojatnost da neki produkt ne radi u određenom vremenu), čija je vjerojatnost tipično manja od 10^{-5}.

Iznimno rijetki događaji zanimljivi su jer se često povezuju s katastrofalnim posljedicama. Možemo govoriti o prirodnim katastrofama kao što su jak potres, tsunami ili udar asteroida ili o društvenima kao što su teroristički čin, krah burze i slično. Precizna procjena vjerojatnosti rijetkih događaja stoga može biti od životne važnosti. U navedenim je primjerima veći problem modelirati ponašanje tih procesa, odnosno pretpostaviti vjerojatnosnu distribuciju uključenih varijabli. U ovom radu pretpostavljat ćemo da nam je distribucija modela poznata. Vjerojatnosti koje nas zanimaju tada je teoretski moguće izračunati, primjerice, numeričkim metodama, no takve procjene uglavnom su računalno vrlo zahtjevne ako želimo dobiti određenu preciznost. Stoga se primjena puno više oslanja na stohastičke simulacije.

Osnovni Monte Carlo algoritam pokazuje se problematičnim u slučaju rijetkih događaja, zbog velike relativne pogreške tog procjenitelja. Pokušamo li, primjerice, procijeniti vjerojatnosti reda veličine 10^{-9}, u prosjeku bismo trebali izvesti milijardu simulacija samo da bi se događaj jednom pojavio, a ako bismo htjeli neki razuman pouzdani

interval, broj se penje i do 10^{18} simulacija, što postaje kompjuterski neizvedivo u razumnom vremenu. U ovom radu prikazat će se neke tehnike kojima se ovaj problem može izbjeći, a posebna pozornost pridana je uzorkovanju po važnosti (*importance sampling*).

Rad se sastoji se od tri poglavlja. U prvom poglavlju navode se neki općeniti rezultati iz teorije vjerojatnosti i slučajnih procesa koji se koriste kasnije u radu. Također je objašnjen problem kod simulacije rijetkih događaja osnovnim Monte Carlo algoritmom, te su obrađene neke tehnike redukcije varijance procjenitelja, s naglaskom na uzorkovanje po važnosti. U drugom i glavnom poglavlju nalazi se pregled osnovnih algoritama za simulaciju rijetkih događaja u slučaju distribucija lakog i teškog repa. Dokazana su razna teorijska svojstva svih algoritama, a kroz cijelo poglavlje protežе se mnoštvo primjera koji čitatelju olakšavaju razumijevanje algoritama. U trećem poglavlju pojašnjena je implementacija algoritama na nekoliko stvarnih primjera, a zatim su izneseni rezultati simulacija u MATLAB-u.

Rad se temelji na poglavlju o simulaciji rijetkih događaja u knjizi *Stochastic simulation* Assmusena i Glynna [4], te donosi pregled osnovnih algoritama za simulaciju rijetkih događaja. Važno je napomenuti da je ovo područje koje se stalno širi, te stoga postoji mnogo algoritama koji u radu nisu obrađeni. Ipak, većina njih temelji se na idejama prikazanima u ovom radu, posebno uzorkovanju po važnosti.

Poglavlje 1

Opis problema i motivacija

1.1 Uvodni rezultati

U ovom dijelu nalaze se neki rezultate iz teorije vjerojatnosti, mjere i slučajnih procesa koje ćemo koristiti u drugom poglavlju.

Definicija 1.1. *Neka su μ i ν pozitivne konačne mjere na izmjerivom prostoru $(\Omega, \mathscr{F})$. Ako postoji funkcija $g : \Omega \to (0, \infty)$, za koju vrijedi*

$$\mu(A) = \int_A g d\nu, \quad \forall A \in \mathscr{F},$$

tada g zovemo Radon-Nikodymovom derivacijom od μ u odnosu na ν i označavamo s $\frac{d\mu}{d\nu}$.

Teorem 1.2. *[Kompozicija Radon-Nikodymovih derivacija] Neka su μ i ν konačne pozitivne mjere i neka je λ konačna na izmjerivom prostoru $(\Omega, \mathscr{F})$. Ako postoje Radon-Nikodymove derivacije $\frac{d\nu}{d\mu}$ i $\frac{d\lambda}{d\nu}$, tada postoji $\frac{d\lambda}{d\mu}$ i*

$$\frac{d\lambda}{d\mu} = \frac{d\lambda}{d\nu}\frac{d\nu}{d\mu}, \quad \mu - g.s.$$

Dokaz. Ovo je standardan rezultat teorije mjere, a dokaz se može naći primjerice u Folland [7]. □

Napomena 1.3. *Ako su F i $\tilde{F}$ dvije neprekidne funkcije distribucije, postoji izmjerivi prostor $(\Omega, \mathscr{F})$ i vjerojatnosti $\mathbb{P}$ i $\tilde{\mathbb{P}}$ na njemu koje su jedinstveno određene s F i $\tilde{F}$. Označimo li sa f i $\tilde{f}$ pripadne funkcije gustoće, Radon-Nikodymovu derivaciju od $\mathbb{P}$ u odnosu na $\tilde{\mathbb{P}}$ možemo označiti s $\frac{dF}{d\tilde{F}}$, gdje je $\frac{dF}{d\tilde{F}}(x) = \frac{f(x)}{\tilde{f}(x)}$.*

Propozicija 1.4. *Neka su f i g funkcije gustoće. Tada vrijedi:*

$$\mathbb{E}_f \log \frac{g(X)}{f(X)} = \int \log \frac{g(x)}{f(x)} f(x)dx \leq 0.$$

Ako jednakost vrijedi, tada f i g definiraju istu vjerojatnosnu mjeru.

Dokaz. Po Jensenovoj nejednakosti vrijedi:

$$\begin{aligned}\mathbb{E}_f \log \frac{g(X)}{f(X)} &\leq \log\left(\mathbb{E}_f \frac{g(X)}{f(X)}\right) = \log\left(\int_{\{f>0\}} \frac{g(x)}{f(x)} f(x)dx\right) \\ &= \log\left(\int_{\{f>0\}} g(x)dx\right) \leq \log 1 = 0,\end{aligned}$$

gdje je $\{f > 0\} = \{x : f(x) > 0\}$. Ako jednakost vrijedi, tada iz druge nejednakosti mora biti $\int_{\{f>0\}} g(x)dx = 1$, a iz Jensenove nejednakosti $g(X)/f(X) = c$, $\mathbb{P}_f$-g.s. Sada iz $\int_{\{f>0\}} cf(x)dx = 1$, slijedi da je konstanta $c = 1$, odnosno $g(X) = f(X)$, $\mathbb{P}_f$-g.s. Dakle za svaki A,

$$\begin{aligned}\int_A g(x)dx &= \int_{A\cap\{f>0\}} g(x)dx \\ &= \mathbb{E}_f\left[\frac{g(X)}{f(X)}; X \in A \cap \{f > 0\}\right] \\ &= \mathbb{P}(X \in A \cap \{f > 0\}) = \int_A f(x)dx,\end{aligned}$$

što znači da definiraju istu mjeru. □

Propozicija 1.5. *[Čebiševljeva nejednakost kovarijanci] Neka je X realna slučajna varijabla i f i g rastuće funkcije. Tada je*

$$\mathbb{E}[f(X)g(X)] \geq \mathbb{E}f(X)\mathbb{E}g(X)$$

.

Dokaz. Vidi Thorisson [11]. □

Propozicija 1.6. *[Waldova jednakost] Neka je $(Y_n : n \geq 1)$ niz pozitivnih nezavisnih jednako distribuiranih integrabilnih slučajnih varijabli i označimo $\mu = \mathbb{E}Y_1$. Neka je T vrijeme zaustavljanja takvo da je $\mathbb{E}T < \infty$. Tada je $S_T = Y_1 + \cdots + Y_T$ integrabilna i vrijedi*

$$\mathbb{E}S_T = \mu\mathbb{E}T.$$

Dokaz. Vidi [12]. □

Definicija 1.7. *Neka je $(\Omega, \mathscr{F}, \mathbb{P})$ vjerojatnosni prostor, X slučajna varijabla na tom prostoru i $\mathscr{G}$ σ-podalgebra od $\mathscr{F}$. Tada je uvjetno očekivanje od X uz dano $\mathscr{G}$, $\mathscr{G}$-izmjeriva funkcija Y za koju vrijedi*

$$\mathbb{E}(Y\mathbb{1}_G) = \mathbb{E}(X\mathbb{1}_G), \quad \forall G \in \mathscr{G}.$$

Slučajnu varijablu Y označavamo s $\mathbb{E}[X|\mathscr{G}]$. Ako su X i Z slučajne varijable na istom vjerojatnosnom prostoru, uvjetno očekivanje $\mathbb{E}[X|\sigma(Z)]$, gdje je $\sigma(Z) = \{Z^{-1}(B) : B \in \mathscr{B}\}$, označavamo s $\mathbb{E}[X|Z]$.

Definicija 1.8. *Uvjetnu varijancu slučajne varijable X s obzirom na Y definiramo kao*

$$\mathrm{Var}(X|Y) = \mathbb{E}[(X - \mathbb{E}[X|Y])^2|Y]. \tag{1.1}$$

Odmah se vidi da vrijedi

$$\mathrm{Var}(X|Y) = \mathbb{E}[X^2|Y] - (\mathbb{E}[X|Y])^2. \tag{1.2}$$

Propozicija 1.9. *Neka su X i Y slučajne varijable na vjerojatnosnom prostoru $(\Omega, \mathscr{F}, \mathbb{P})$ neka je $\mathscr{G} \subset \mathscr{F}$.*

(a) Ako je $\mathscr{H} \subset \mathscr{G}$, vrijedi

$$\mathbb{E}[\mathbb{E}[X|\mathscr{G}]|\mathscr{H}] = \mathbb{E}[X|\mathscr{H}] \quad g.s.$$

Posebno, uzmemo li $\mathscr{G} = \{\emptyset, \Omega\}$ i $\mathscr{H} = \sigma(Y)$, dobivamo

$$\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]. \tag{1.3}$$

(b) Vrijedi

$$\mathrm{Var}(X) = \mathbb{E}[\mathrm{Var}(X|Y)] + \mathrm{Var}(\mathbb{E}[X|Y]). \tag{1.4}$$

Dokaz. (a) Neka je $Y = \mathbb{E}[X|\mathscr{G}]$ i $Z = \mathbb{E}[Y|\mathscr{H}]$. Tada je Y $\mathscr{G}$-izmjeriva i za svaki $G \in \mathscr{G}$ vrijedi $\mathbb{E}[Y\mathbb{1}_G] = \mathbb{E}[X\mathbb{1}_G]$. Također, Z je $\mathscr{H}$-izmjeriva (pa i $\mathscr{G}$-izmjeriva zbog $\mathscr{H} \subset \mathscr{G}$) pa za svaki $H \in \mathscr{H}$ vrijedi $\mathbb{E}[Z\mathbb{1}_H] = \mathbb{E}[Y\mathbb{1}_H] = \mathbb{E}[X\mathbb{1}_H]$. Sada po definiciji uvjetnog očekivanja slijedi tvrdnja.

(b) Iz (1.3) i (1.2) slijedi

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mathbb{E}[\mathbb{E}[X^2|Y]] - (\mathbb{E}[\mathbb{E}[X|Y]])^2 \\ &= \mathbb{E}[\mathrm{Var}(X|Y)]] + \mathbb{E}[(\mathbb{E}[X|Y])^2] - (\mathbb{E}[\mathbb{E}[X|Y]])^2 \\ &= \mathbb{E}[\mathrm{Var}(X|Y)] + \mathrm{Var}(\mathbb{E}[X|Y])\end{aligned}$$

□

Lema 1.10. *Neka je $z(x) : \mathbb{R} \to (0, 1]$ funkcija za koju vrijedi $z(x) \to 0$, kad $x \to \infty$. Neka su $Z(x) : \mathbb{R} \to [0, 1]$ slučajne varijable za koje vrijedi $z(x) = \mathbb{E}Z(x)$. Tada su sljedeće dvije tvrdnje ekvivalentne:*

$$a) \quad \limsup_{x\to\infty} \frac{\mathrm{Var}Z(x)}{z(x)^{2-\varepsilon}} = 0, \quad \forall \varepsilon > 0$$

$$b) \quad \liminf_{x\to\infty} \frac{|\log \mathrm{Var}Z(x)|}{|\log z(x)^2|} \geq 1.$$

Dokaz. Pretpostavimo da vrijedi a). Tada za svaki $\delta > 0$ postoji broj $N(\delta)$ takav da:

$$\frac{\mathrm{Var}Z(x)}{z(x)^{2-\varepsilon}} < \delta, \quad \text{za} \quad x > N(\delta).$$

Posebno, uzmimo $0 < \delta < 1$. Iz činjenica da su $0 \leq z(x), \mathrm{Var}Z(x) \leq 1$ i logaritam rastuća funkcija negativna na (0,1) te nejednakosti trokuta, za $x > N(\delta)$ vrijedi:

$$\begin{aligned} \mathrm{Var}Z(x) &< \delta z(x)^{2-\varepsilon}, \\ |\log \mathrm{Var}Z(x)| &> |\log \delta + \log(z(x)^{2-\varepsilon})| \\ \frac{|\log \mathrm{Var}Z(x)|}{|\log z(x)^{2-\varepsilon}|} &> 1 - \frac{|\log \delta|}{|\log z(x)^{2-\varepsilon}|} \end{aligned}$$

Kako gornja nejednakost vrijedi za svaki $\varepsilon > 0$, puštanjem $\varepsilon \to 0$ i uzimanjem $\liminf_{x\to\infty}$, zbog $|\log z(x)| \to \infty$ slijedi b).
Pretpostavimo sad da vrijedi b). Tada za svaki $\varepsilon > 0$ postoji broj $N(\varepsilon)$ takav da je

$$\frac{|\log \mathrm{Var}Z(x)|}{|\log z(x)^2|} \geq 1 - \varepsilon, \quad \text{za} \quad x > N(\varepsilon).$$

Eksponencijalna i logaritamska funkcija su rastuće i logaritam je negativan na (0,1) pa slijedi da za $x > N(\varepsilon)$,

$$\begin{aligned} \log(\mathrm{Var}Z(x)) &\leq (1-\varepsilon)\log(z(x)^2) \\ \mathrm{Var}Z(x) &\leq z(x)^{2(1-\varepsilon)} \\ \frac{\mathrm{Var}Z(x)}{z(x)^{2-2\varepsilon}} &\leq 1 \\ \frac{\mathrm{Var}Z(x)}{z(x)^{2-\varepsilon}} &\leq z(x)^{-\varepsilon} \end{aligned}$$

Puštanjem $\limsup_{x\to\infty}$, zbog $z(x) \to 0$ i pozitivnosti izraza na lijevoj strani slijedi tvrdnja.

□

1.2 Monte Carlo

Pretpostavimo da želimo procijeniti vjerojatnost nekog događaja A, $z = \mathbb{P}(A)$. Ideja je iz modela simulirati N nezavisnih uzoraka i pri svakoj realizaciji bilježiti je li se A dogodio ili ne. Neka je Z_i Bernoullijeva slučajna varijabla:

$$Z_i = \begin{cases} 1, & \text{ako se A dogodio u i-tom uzorku} \\ 0, & \text{inače.} \end{cases}$$

Tada je *osnovni Monte Carlo procjenitelj* za z

$$\hat{z} = \frac{1}{N} \sum_{i=1}^{N} Z_i. \tag{1.5}$$

Zakon velikih brojeva garantira nam da $\hat{z} \to \mathbb{E}Z$, odnosno $\hat{z} \to \mathbb{P}(A)$ kad $N \to \infty$. Primijetimo da je $\mathbb{E}Z_i = z$ i $\sigma^2 = \mathrm{Var}Z_i = z(1-z)$, odnosno $\mathbb{E}\hat{z} = z$ ($\hat{z}$ je nepristran procjenitelj) i $\mathrm{Var}\hat{z} = z(1-z)/\sqrt{N}$. Prirodno je pitati se koliko je zapravo $\hat{z}$ blizu točne vrijednosti z, odnosno koliko je naša procjena točna. Pogrešku Monte Carlo procjene možemo izraziti na nekoliko načina. Slučajne varijable $Z_1, \cdots, Z_N$ su nezavisne i jednako distribuirane, pa po centralnom graničnom teoremu vrijedi:

$$\frac{\hat{z} - z}{\sigma} \sqrt{N} \xrightarrow{\mathscr{D}} \mathscr{N}(0,1). \tag{1.6}$$

Slijedi da za velike uzorke pogreška $\hat{z} - z$ ima asimptotski normalnu distribuciju $\mathscr{N}(0, \sigma^2)$. Označimo sa Φ funkciju distribucije jedinične normalne razdiobe. Tada broj z_α za koji vrijedi $\Phi(z_\alpha) = \alpha$ zovemo α-kvantilom jedinične normalne razdiobe. Iz (1.6) slijedi da je asimptotska vjerojatnost događaja:

$$\left\{ z_{\alpha/2}\sigma/\sqrt{N} < \hat{z} - z < z_{1-\alpha/2}\sigma/\sqrt{N} \right\}$$

jednaka $\Phi(z_{1-\alpha/2}) - \Phi(z_{\alpha/2}) = (1 - \alpha/2) - \alpha/2 = 1 - \alpha$. Dakle, interval

$$I_\alpha = (\hat{z} - z_{1-\alpha/2}\sigma/\sqrt{N}, \hat{z} - z_{\alpha/2}\sigma/\sqrt{N}) \tag{1.7}$$

je asimptotski $(1-\alpha)\%$ pouzdani interval za z, zbog $\mathbb{P}(z \in I_\alpha) \to 1 - \alpha$, za $N \to \infty$. U praksi je varijanca σ^2 nepoznata i potrebno ju je procijeniti iz uzorka. Uzoračka varijanca

$$s^2 := \frac{1}{N-1} \sum_{i=1}^{N} \left(Z_i - \hat{z}\right)^2$$

je nepristran procjenitelj za σ^2 i $s^2 \xrightarrow{\mathbb{P}} \sigma^2$. Slijedi da je

$$(\hat{z} - z_{1-\alpha/2}s/\sqrt{N}, \hat{z} - z_{\alpha/2}s/\sqrt{N}) =: \hat{z} \pm z_{1-\alpha/2}s/\sqrt{N}$$

asimptotski $(1-\alpha)\%$ pouzdani interval za z. Preciznost simulacije sada se može izraziti u terminima širine polovice asimptotski pouzdanog intervala. Razlikujemo dva kriterija za točnost procjene. Pod pojmom *apsolutna točnost* ε podrazumijevamo da računamo vrijednost z do na točnost ε, odnosno da zahtijevamo da je apsolutna pogreška $|z-\hat{z}| < \varepsilon$. Pod pojmom *relativna točnost* ε podrazumijevamo da računamo vrijednost z do na točnost $\varepsilon|z|$, odnosno da zahtijevamo da je apsolutna pogreška $|z-\hat{z}|/|z| < \varepsilon$. Iz (1.7) slijedi da su veličine uzorka potrebne za postizanje redom apsolutne i relativne točnosti ε,

$$N \approx \frac{\sigma^2 z_{1-\alpha/2}^2}{\varepsilon^2}, \quad N \approx \frac{\sigma^2 z_{1-\alpha/2}^2}{\varepsilon^2 z^2}. \tag{1.8}$$

Najčešće nas zanima 95% pouzdani interval za vrijednost koju procjenjujemo, što zbog $z_{1-\alpha/2} = 1.96$ postaje $\hat{z} \pm 1.96\sigma/\sqrt{N}$, odnosno potrebne veličine uzorka za zadanu preciznost ε postaju $N \approx \sigma^2 1.96^2/\varepsilon^2$ i $N \approx \sigma^2 1.96^2/\varepsilon^2 z^2$.

Pretpostavimo sad da je $z = \mathbb{P}(A) \leq 10^{-3}$, odnosno da je A rijedak događaj. U slučaju rijetkih događaja puno bolju informaciju o grešci daje nam relativna pogreška u odnosu na apsolutnu. Primjerice, ako je $z \approx 10^{-3}$, a naša procjena $\hat{z} = 2 \cdot 10^{-3}$, tada je apsolutna pogreška 10^{-3} 'mala', dok je relativna pogreška čak 100%. Iz prethodnih razmatranja vidimo da je veličina uzorka potrebna da se dobije relativna točnost ε, obrnuto proporcionalna 'učestalosti događaja', odnosno što je događaj A rjeđi, potreban je veći uzorak da bismo postigli zadanu preciznost. Uzmimo primjerice, da želimo odrediti $z \approx 10^{-3}$ do na točnost od $\varepsilon = 0.01$ uz 95% pouzdani interval. Iz (1.8) slijedi da je veličina uzorka potrebna za postizanje relativne točnosti ε, $N \approx 3.84 \cdot 10^5 \sigma$. Za $\sigma^2 = z(1-z)$ kao u našem primjeru, $N = 3.84 \cdot 10^8$, a za postizanje dodatne decimale točnosti potreban nam je još 100 puta veći uzorak. Pokušamo li procijeniti još manje vjerojatnosti, primjerice reda 10^{-9}, ovakva simulacija postaje praktički nemoguća.

Uočimo da velik utjecaj na pogrešku procjene ima varijanca procjenitelja σ^2. Preciznije, što je varijanca manja, pouzdani interval je uži, odnosno procjena je točnija. Također, s manjom varijancom, smanjuje se i veličina uzorka potrebna za dobivanje određene točnosti procjene. Stoga ćemo u sljedećem dijelu razmotriti problem pronalaženja procjenitelja za z manje varijance od opisanog.

1.3 Tehnike redukcije varijance

Cilj redukcije varijance je pronaći alternativni procjenitelj vjerojatnosti $z = \mathbb{P}(A)$, koji ima puno manju varijancu od osnovnog Monte Carlo procjenitelja $\hat{z}$. Kažemo puno manju, jer smanjenje varijance nema veliku svrhu ako nije značajno. Da bismo to pokazali, pretpostavimo da smo našli alternativni procjenitelj za $\mathbb{P}(A)$, $\tilde{z}$, čija je varijanca 25% manja od varijance Monte Carlo procjenitelja, odnosno $\tilde{\sigma}^2 = 0.75\sigma^2$. Ako su N i $\tilde{N}$ veličine uzoraka potrebne da se postigne relativna točnost ε, vrijedi

$$\varepsilon = \frac{\sigma z_{1-\alpha/2}}{z\sqrt{N}} = \frac{\tilde{\sigma} z_{1-\alpha/2}}{z\sqrt{\tilde{N}}} \quad \Rightarrow \quad \tilde{N} = \frac{N\tilde{\sigma}^2}{\sigma^2} = 0.75\tilde{N}.$$

Dakle, pod pretpostavkom da je kompjutersko vrijeme izvršavanja oba algoritma jednako, uspjeli smo smanjiti veličinu uzorka potrebnu za postizanje točnosti ε za samo 25%. Kako je u praksi redukcija varijance često algoritamski složeniji postupak i zahtijeva više resursa, ne postignemo li značajno smanjenje varijance cijeli se postupak zapravo ne isplati. U nastavku ćemo se upoznati s nekoliko metoda redukcije varijance pomoću kojih ćemo kasnije konstruirati efikasne procjenitelje rijetkih događaja.

Uzorkovanje po važnosti

Neka je X neprekidna slučajna varijabla s funkcijom gustoće $f(x)$ i pretpostavimo da želimo izračunati $z = \mathbb{P}\{X \in A\}$. Neka je $g(x)$ funkcija gustoće takva da je $g(x) > 0$ za sve $x \in A$ za koje je $f(x) > 0$. Tada je:

$$z = \mathbb{E}_f[\mathbb{1}_{\{X\in A\}}] = \int \mathbb{1}_{\{x\in A\}} f(x)dx = \int \left[\frac{\mathbb{1}_{\{x\in A\}} f(x)}{g(x)}\right] g(x)dx = \mathbb{E}_g[\mathbb{1}_{\{X\in A\}} L(X)] \tag{1.9}$$

gdje je $L(x) = f(x)/g(x)$ omjer vjerodostojnosti i $\mathbb{E}_g$ označava da uzorkujemo iz gustoće g. Ova jednakost predlaže sljedeću metodu procjene koju zovemo *uzorkovanje po važnosti* (kraće IS, prema eng. *importance sampling*). Iz gustoće g uzorkujemo $X_1, \cdots, X_N$ i definiramo $\tilde{Z}_i = L(X_i)Z_i$, gdje je $Z_i = \mathbb{1}_{\{X_i\in A\}}$. Iz (1.9) slijedi da je $\mathbb{E}_g[\tilde{Z}_i] = z$. Tada je nepristran procjenitelj za z dan sa:

$$\hat{z} = \frac{1}{N}\sum_{i=1}^{N} \tilde{Z}_i = \frac{1}{N}\sum_{i=1}^{N} Z_i L(X_i). \tag{1.10}$$

Označimo sa $\sigma_g^2 = \mathrm{Var}_g \tilde{Z} = \mathrm{Var}_g\big(ZL(X)\big)$. U praksi je ona nepoznata, odnosno procjenjujemo je iz uzorka sa:

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Z_i L(X_i) - \hat{z})^2.$$

Očito za g možemo odabrati bilo koju funkciju koja zadovoljava traženi uvjet. Stavimo li $g(x) = g^*(x) = f(x)/z$, za $x \in A$ i $g^*(x) = 0$ inače, dobivamo

$$\tilde{Z}_i = Z_i L(X_i) = \mathbb{1}_{\{x_i \in A\}} \frac{f(x_i)}{g(x_i)} = z, \quad \mathbb{P}_f\text{-g.s.}$$

S obzirom da je varijanca konstante 0, uz g^* se postiže maksimalna redukcija varijance. Iako je očito nemoguće u praksi uzorkovati iz g^*, jer zahtijeva znanje o z kojeg zapravo pokušavamo procijeniti, gornje razmatranje daje nam dobru predodžbu o tome kakve g tražimo. Ideja uzorkovanja po važnosti je odabrati g tako da se većinom uzorkuje iz onog dijela prostora koji najviše pridonosi z, tj. koji je 'važan', pa stoga i naziv *uzorkovanje po važnosti*.

Usporedimo sada uzorkovanje po važnosti s osnovnom Monte Carlo procjenom. Prisjetimo se, veličina uzorka potrebna da bismo dobili relativnu preciznost ε, je $N \approx 1.96^2\sigma^2/\varepsilon^2 z^2$, gdje je σ^2 varijanca slučajnih varijabli iz modela. U slučaju Monte Carlo procjenitelja, $\sigma^2 = z(1-z)$, pa za vrlo male z imamo $N \approx 1.96^2/\varepsilon^2 z$. U slučaju IS-procjenitelja, dobivamo $N \approx 1.96^2\sigma_g^2/\varepsilon^2 z^2$. Redukcijski faktor je, dakle, z/σ_g^2 i bit će značajan ako je $\sigma_g^2 \ll z$. To nam sugerira da bi dobar odabir distribucije g bio onaj za koju je drugi moment od $ZL(X)$ značajno manji od z.

Generalizirajmo sada ovaj primjer na problem računanja očekivanja neke slučajne varijable. Pretpostavimo da želimo odrediti $z = \mathbb{E}Z$, gdje je Z slučajna varijabla na vjerojatnosnom prostoru $(\Omega, \mathscr{F}, \mathbb{P})$. Tada je

$$z = \int_\Omega Z(\omega)\mathbb{P}(d\omega).$$

Pretpostavimo sada da za neku drugu vjerojatnost $\tilde{\mathbb{P}}$ na $(\Omega, \mathscr{F})$ vrijedi da postoji Radon-Nikodymova derivacija L, takva da je

$$Z(\omega)\mathbb{P}(\omega) = Z(\omega)L(\omega)\tilde{\mathbb{P}}(\omega), \quad \omega \in \Omega.$$

Kao gore, vrijedi:

$$z = \mathbb{E}Z = \int_\Omega Z(\omega)\mathbb{P}(dw) = \int_\Omega Z(\omega)L(\omega)\tilde{\mathbb{P}}(d\omega) = \tilde{E}[ZL]. \tag{1.11}$$

Da bismo odredili z, u svjetlu gornjeg primjera, možemo umjesto iz $\mathbb{P}$ uzorkovati $Z_1L_1, \cdots, Z_NL_N$ iz distribucije $\tilde{\mathbb{P}}$ te procijeniti z kao $\hat{z} = \sum_{i=1}^{N} Z_iL_i$. Uzorkovanje iz drugačije distribucije zovemo još i promjenom mjere, a $\tilde{\mathbb{P}}$, odnosno pripadnu funkciju distribucije, nazivamo distribucijom važnosti. Očito se za $\tilde{\mathbb{P}}$ može uzeti bilo koja vjerojatnost koja zadovoljava traženi uvjet, no nama su zanimljive samo one mjere koje smanjuju varijancu procjenitelja $\hat{z}$. Sljedeći teorem govori o optimalnoj distribuciji važnosti.

Teorem 1.11. *Neka za $\mathbb{P}^*$ vrijedi*

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = \frac{|Z|}{\mathbb{E}Z}, \quad \mathbb{P}^*(d\omega) = \frac{|Z|}{\mathbb{E}Z}\mathbb{P}(d\omega), \quad \textit{odnosno} \quad L^* = \frac{\mathbb{E}|Z|}{|Z|}$$

Tada IS-procjenitelj ZL^* uz $\mathbb{P}^*$ ima manju varijancu nego IS-procjenitelj ZL uz bilo koju drugu $\tilde{\mathbb{P}}$. Ako je $Z \geq 0$ $\mathbb{P}$-g.s., onda je $\mathbb{P}^*$-varijanca od ZL^* 0.

Dokaz. Iz Cauchy-Schwartzove nejednakosti slijedi:

$$\mathbb{E}^*\big[(ZL^*)^2\big] = \mathbb{E}^*\big[(\mathbb{E}|Z|)^2\big] = (\mathbb{E}|Z|)^2 = (\tilde{\mathbb{E}}\big[|Z|L\big])^2 \leq \tilde{\mathbb{E}}\big[(ZL)^2\big],$$

pa procjenitelj uz $\mathbb{P}^*$ ima manji drugi moment, a time i manju varijancu, naime $\mathrm{Var}(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2$. Ako je $Z \geq 0$ $\mathbb{P}$-g.s., tada je $L^* = \mathbb{E}Z/Z$ pa slijedi

$$\mathrm{Var}^*(ZL^*) = \mathrm{Var}^*(\mathbb{E}Z) = \mathrm{Var}^*(z) = 0,$$

jer je varijanca konstante 0. □

Primijetimo da se optimalna distribucija nikada ne može implementirati u praksi, budući da nam je potrebno znanje o z, koji upravo i pokušavamo procijeniti. No, gornji teorem daje nam dobru ideju o tome iz kakve bismo otprilike distribucije trebali simulirati. Ako je $Z = \mathbb{P}(A)$, iz teorema 1.11 slijedi da je

$$\mathbb{P}^*(d\omega) = \frac{\mathbb{1}_{\{\omega \in A\}}}{\mathbb{P}(A)}\mathbb{P}(d\omega),$$

odnosno da je $\mathbb{P}^*(\cdot) = \mathbb{P}(\cdot|A)$. Dakle, za što veću redukciju varijance, htjeli bismo simulirati iz distribucije $\tilde{\mathbb{P}}$ koja je što bliža uvjetnoj distribuciji uz uvjet da se dani događaj A dogodio. Pogledajmo za kraj još jedan ne sasvim realan, ali zanimljiv primjer.

Primjer 1.12. *Pretpostavimo da želimo simulirati očekivanje slučajne varijable X koja ima eksponencijalnu distribuciju s parametrom* 1. *Prisjetimo se, gustoća slučajne varijable $X(\lambda)$ s eksponencijalnom distribucijom s parametrom λ je $f(x) = \lambda e^{-\lambda x}$.*

Tada je $z = \mathbb{E}X = \int_0^\infty xe^{-x}dx$. *Primijenimo li uzorkovanje po značajnosti uz eksponencijalnu distribuciju s nekim drugim parametrom* λ, *procjenitelj postaje:*

$$\tilde{Z}(\lambda) = X(\lambda)\frac{e^{-X(\lambda)}}{\lambda e^{-\lambda X(\lambda)}} = \frac{1}{\lambda}X(\lambda)e^{-(1-\lambda)X(\lambda)}.$$

Sada imamo:

$$\mathbb{E}Z(\lambda)^2 = \int_0^\infty \frac{1}{\lambda^2}x^2e^{-2(1-\lambda)}\lambda e^{-\lambda x}dx = \begin{cases} \frac{2}{\lambda(2-\lambda)^3}, & 0 < \lambda < 2, \\ \infty, & \lambda \geq 2. \end{cases} \tag{1.12}$$

Dakle, $\mathrm{Var}Z(\lambda) = \infty$ *za* $\lambda \geq 2$. *To nam govori da neke promjene mjere mogu biti jako loše. Pretpostavimo da je neki parametar* λ^* *optimalan u smislu minimiziranja varijance. Intuitivno, pomičemo li se u "krivom smjeru", od* λ^*, *a ne prema* λ^*, *varijance raste i može postati beskonačna. To će rezultirati lošom procjenom z, odnosno prevelikim i besmislenim pouzdanim intervalima. U ovom primjeru, optimalan* λ *je* $\lambda^* = 1/2$ *(minimizator drugog momenta u (1.12)) pa uzimamo li distribuciju važnosti s parametrom* $\lambda > 1$, *pomičemo se u "krivom smjeru". Kako je* $\mathbb{E}Z(\lambda) = 1$ *za svaki* λ, *za optimalni* $\lambda^* = 1/2$, $\mathrm{Var}Z(\lambda^*) = 2/\lambda^*(2-\lambda^*)^3 - 1 = 5/27$, *u usporedbi s varijancom Monte Carlo procjenitelja* $\mathrm{Var}Z(1) = 1$. *Vidimo i da za* $\lambda < \lambda^*$, *za* $\lambda(2-\lambda)^3 < 1$ *nema više redukcije varijance i* $\mathrm{Var}Z(\lambda) \to \infty$ *kad* $\lambda \to 0$.

Definicija 1.13. *Neka je* X *slučajna varijabla za koju vrijedi* $\mathbb{E}e^{\theta X} < \infty$. *Tada uzorkovanje po važnosti uz distribuciju*

$$\mathbb{P}_\theta(X \in dx) = e^{\theta x - \kappa(\theta)}\mathbb{P}(X \in dx),$$

gdje je $\kappa(\theta) = \log(\mathbb{E}\left[e^{\theta X}\right])$, *zovemo eksponencijalnom promjenom mjere.*

Uvjetni Monte Carlo

U ovoj metodi redukcije varijance osnovni Monte Carlo procjenitelj Z zamjenjujemo sa $Z_{cond} = \mathbb{E}[Z|\mathscr{G}]$, gdje je $\mathscr{G}$ sigma algebra kao u definiciji 1.7. Radi jednostavnosti bavit ćemo se procjeniteljem $Z_{cond} = \mathbb{E}[Z|W]$, gdje je W neka slučajna varijabla na istom vjerojatnosnom prostoru kao Z. Primijetimo da zbog b) dijela propozicije 1.9 vrijedi

$$\begin{aligned} \sigma^2 &= \mathrm{Var}(Z) = \mathrm{Var}(\mathbb{E}[Z|W]) + \mathbb{E}[\mathrm{Var}(Z|W)] \\ &= \sigma^2_{cond} + \mathbb{E}[\mathrm{Var}(Z|W)] \geq \sigma^2_{cond}, \end{aligned} \tag{1.13}$$

pa uvjetna Monte Carlo metoda uvijek smanjuje varijancu procjenitelja. Problem je pronaći slučajnu varijablu W za koju se gornje uvjetno očekivanje može izračunati. Ideja je tada simulirati $W_1, \ldots, W_n$ te procijeniti traženu vjerojatnost kao

$$\bar{Z}_{cond} = \sum_{i=1}^{n} \mathbb{E}[X|W_i].$$

Primjer 1.14. *Neka su X_1 i X_2 nezavisne slučajne varijable s distribucijom F i pretpostavimo da želimo procijeniti vjerojatnost $z = \mathbb{P}(X_1 + X_2 > x)$. Tada je Monte Carlo procjenitelj $Z = \mathbb{1}_{\{X_1+X_2>x\}}$. Jedna ideja koja se odmah nameće je za W uzeti X_1, čime dobivamo*

$$Z_{cond} = \mathbb{P}(X_1 + X_2 > x|X_1) = \mathbb{P}(X_2 > x - X_1) = 1 - F(x - X_1).$$

U sljedećem poglavlju često ćemo koristiti sljedeće funkcije i njihova svojstva.

Definicija 1.15. *Za slučajnu varijablu X definiramo funkciju $\hat{F} = \hat{F}_X$ sa*

$$\hat{F}[s] = \mathbb{E}\left[e^{sX}\right]$$

i nazivamo je funkcijom izvodnicom momenata.

Propozicija 1.16. *Označimo sa $\kappa(s) = \log \hat{F}[s]$. Tada je κ konveksna funkcija.*

Dokaz. Neka je $u \in [0, 1]$ i označimo radi jednostavnosti $v = 1 - u$. Iz Hölderove nejednakosti slijedi:

$$\begin{aligned}
\kappa(su + tv) &= \log \mathbb{E}\left[e^{(su+tv)X}\right] \\
&= \log \mathbb{E}\left[(e^{sX})^u (e^{tX})^v\right] \\
&\leq \log \left[\left(\mathbb{E}\left[e^{sX}\right]\right)^u \left(\mathbb{E}\left[e^{tX}\right]\right)^v\right] \\
&= u\kappa(s) + v\kappa(t),
\end{aligned}$$

što dokazuje konveksnost. □

Poglavlje 2

Algoritmi

2.1 Problemi efikasnosti

U ovom poglavlju detaljno razmatramo problem procjene $z = \mathbb{P}(A)$, kad je z manji od 10^{-3}, odnosno A je rijedak događaj. Opisat ćemo nekoliko algoritama te razmotriti njihovu efikasnost u slučajevima distribucija lakog i teškog repa.
U prethodnim poglavljima objasnili smo neke teškoće koje se javljaju prilikom simulacije rijetkih događaja. Razmotrimo ih detaljnije na sljedećem primjeru.

Primjer 2.1. *Neka je X slučajna varijabla s funkcijom gustoće $f(x)$ i promotrimo problem procjene vjerojatnosti z da se X nalazi u nekom skupu A,*

$$z = \mathbb{P}\{X \in A\} = \int_{-x}^{x} \mathbb{1}_{\{x \in A\}} f(x)dx = \mathbb{E}[\mathbb{1}_{\{X \in A\}}]. \tag{2.1}$$

Monte Carlo pristup problemu bio bi, za zadani N, uzorkovati $X_1, \cdots, X_N$ iz gustoće f, te procijeniti traženu vjerojatnost kao

$$\hat{z}_N = \frac{1}{N} \sum_{n=1}^{N} I_n, \quad \textit{gdje je } I_n = \mathbb{1}_{\{X_n \in A\}}.$$

Jasno je da su očekivanje i varijanca procjenitelja, $\mathbb{E}[\hat{z}_N] = z$ i $\sigma^2_{\hat{z}_N} = \dfrac{z(1-z)}{N}$.
Relativna pogreška definira se kao omjer standardne devijacije i očekivanja procjenitelja. U ovom slučaju:

$$RE(\hat{z}_N) = \frac{\sqrt{z(1-z)}}{z\sqrt{N}} \sim \frac{1}{\sqrt{z}\sqrt{N}} \to \infty, \quad \textit{kad} \quad z \to 0.$$

Dakle, relativna pogreška je neograničena za rijetke događaje.
Pretpostavimo sad da želimo postići relativnu preciznost od 10%*, u terminima relativne širine polovice, primjerice* 95%*-tnog intervala. Po centralnom graničnom teoremu, procjena za* $(1-\alpha)\cdot 100\%$ *pouzdani interval za* z *je* $\hat{z}_N \pm z_{\alpha/2}\sqrt{z(1-z)}/\sqrt{N}$*, gdje je* $z_{\alpha/2}$*,* $1-\alpha/2$ *kvantil jedinične normalne razdiobe. Tada mora vrijediti*

$$1.96\frac{\sqrt{z(1-z)}}{\hat{z}_N\sqrt{N}} \leq 0.1.$$

Zanima nas koliko velik uzorak trebamo imati da bismo postigli odgovarajuću preciznost. Kako $\hat{z}_N \to z$ *kad* $N \to \infty$*, vidimo da je* $N \approx 100 \cdot 1.96^2 \cdot (1-z)/z$*, odnosno* N *je proporcionalan s* $1/z$*. Što je manji* z*, treba nam veći* N*, a već za* $z = 10^{-6}$*, da bismo dobili preciznost od samo* 10% *(jedna decimala), potreban nam je uzorak veličine* $3.84 \cdot 10^8$*. Za dodatnu decimalu preciznosti, potreban nam je još 100 puta veći uzorak, a pokušamo li procijeniti vjerojatnosti reda* 10^{-9} *dolazimo do veličine uzorka reda* 10^{13}*. Stoga ovakva simulacija nije samo neefikasna, već i praktički nemoguća.*

Uvedimo sad i formalno naš problem. Neka je $A(x)$ familija rijetkih događaja, gdje je $x \in (0,\infty)$ ili $x \in \mathbb{N}$. Označimo sa $z(x) := \mathbb{P}A(x)$ i neka $z(x) \to 0$ za $x \to \infty$. Za svaki x neka je $Z(x)$ nepristran procjenitelj za $z(x)$, odnosno $\mathbb{E}Z(x) = z(x)$. Algoritam definiramo kao familiju slučajnih varijabli $Z(x)$ koje zadovoljavaju gornje svojstvo.
Htjeli bismo da algoritmi imaju ograničenu relativnu pogrešku, ili barem da ona raste vrlo sporo kad $x \to \infty$, odnosno da veličina uzorka potrebna da se postigne zadana relativna pogreška sporo raste što je događaj rjeđi.
Pod *ograničenom relativnom pogreškom* podrazumijevamo da

$$\limsup_{x\to\infty} \frac{\mathrm{Var}Z(x)}{z(x)^2} < \infty. \tag{2.2}$$

Takav algoritam imat će svojstvo da veličina uzorka N, potrebna da se ostvari zadana preciznost, ostaje ograničena za $x \to \infty$.
Nešto slabije svojstvo je *logaritamska efikasnost*: $\mathbb{V}arZ(x) \to 0$ tako brzo da je za svaki $\varepsilon > 0$

$$\limsup_{x\to\infty} \frac{\mathrm{Var}Z(x)}{z(x)^{2-\varepsilon}} = 0, \tag{2.3}$$

što je po lemi 1.10 ekvivalentno s

$$\liminf_{x\to\infty} \frac{|\log \mathrm{Var}Z(x)|}{|\log z(x)^2|} \geq 1. \tag{2.4}$$

Iako je logaritamska efikasnost slabiji koncept, u praktičnom smislu razlika je vrlo mala, a često ju je lakše dokazati nego ograničenost relativne pogreške. Gornje tvrdnje obično se, umjesto za $\mathrm{Var}Z(x)$, dokazuju za njenu gornju među $\mathbb{E}Z(x)^2$.

Spomenimo još da se rijedak događaj A može uklopiti u ovo formalno okruženje na više načina. Primjerice, ako je $A = \{S_n > nx\}$, pri čemu je S_n suma n nezavisnih jednakodistribuiranih slučajnih varijabli, A može biti rijedak zbog posljedica zakona velikih brojeva (n je velik i x je veći od očekivanja) ili zbog toga što je x tako velik da se nx nalazi u repu distribucije od S_n. Stoga oba limesa $x \to \infty$ i $n \to \infty$ mogu biti relevantna.

Velik dio teorije simulacije rijetkih događaja temelji se na uzorkovanju po važnosti kao načinu za dizajniranje efikasnih algoritama. Podsjetimo se, u prethodnom poglavlju pokazali smo da je optimalna promjena mjere uvjetna distribucija uz dani rijetki događaj A; $\tilde{\mathbb{P}}(B) = \mathbb{P}(B \mid A)$. No, kao što smo i prije razmatrali, nije praktično simulirati iz ove distribucije, s obzirom da je $z = \mathbb{P}(A)$ nepoznat. Ideja je dakle odabrati mjeru $\tilde{\mathbb{P}}$ što sličniju $\mathbb{P}(\cdot \mid A)$.

U nastavku ćemo obraditi neke algoritme za simulaciju rijetkih događaja koji imaju ograničenu relativnu pogrešku ili su logaritamski efikasni. Prije toga, pogledajmo gornja razmatranja na jednostavnom primjeru.

Primjer 2.2. *Neka je X geometrijska slučajna varijabla na $\{1, 2, \dots\}$ s vjerojatnošću uspjeha π, tj. $\mathbb{P}(X = n) = \pi(1-\pi)^{n-1}$. Promatramo*

$$z = \mathbb{P}(X \leq m) = \sum_{n=1}^{m} \pi(1-\pi)^{n-1} = 1 - (1-\pi)^m, \tag{2.5}$$

gdje $\pi = \pi(x)$ i $m = m(x)$ oboje ovise o parametru x, takvom da $z = z(x) \to 0$ za $x \to \infty$. Dakle, $(1-\pi)^m \to 1$, tj. $m\pi \to 0$ iz čega slijedi da je $z \sim m\pi$.
Kako bi efikasno simulirali z, u svjetlu uzorkovanja po važnosti X ćemo simulirati iz $\tilde{\mathbb{P}}$, odnosno odgovarajuće geometrijske distribucije s parametrom $\tilde{\pi}$. Postavlja se pitanje, na koji način izabrati $\tilde{\pi}$.
Procjenitelj, simuliran iz $\tilde{\mathbb{P}}$ je

$$Z := \mathbb{1}_{\{X \leq m\}} \frac{\pi(1-\pi)^{X-1}}{\tilde{\pi}(1-\tilde{\pi})^{X-1}}.$$

Slijedi da je:

$$\begin{aligned}\mathbb{E}Z^2 &= \frac{\pi^2}{\tilde{\pi}^2}\tilde{\mathbb{E}}\left[\left(\frac{1-\pi}{1-\tilde{\pi}}\right)^{2(X-1)}; X \le m\right] = \frac{\pi^2}{\tilde{\pi}^2}\sum_{n=1}^{m}\tilde{\mathbb{E}}\left[\left(\frac{1-\pi}{1-\tilde{\pi}}\right)^{2(n-1)}; X = n\right] \\ &= \frac{\pi^2}{\tilde{\pi}^2}\sum_{n=1}^{m}\left(\frac{1-\pi}{1-\tilde{\pi}}\right)^{2(n-1)}\tilde{\mathbb{P}}(X=n) = \frac{\pi^2}{\tilde{\pi}}\sum_{n=1}^{m}\frac{(1-\pi)^{2(n-1)}}{(1-\tilde{\pi})^{n-1}} \\ &= \frac{\pi^2}{\tilde{\pi}}\frac{(1-\pi)^{2m}/(1-\tilde{\pi})^m - 1}{(1-\pi)^2/(1-\tilde{\pi}) - 1}.\end{aligned} \tag{2.6}$$

Stavimo li $\tilde{\mathbb{E}}X = m$*, kako je* $\mathbb{E}X = 1/\pi$*, dobivamo* $\tilde{\pi} = 1/m$*. Iz* $\pi \sim z/m$ *i razmatranja na početku primjera te (2.6) dobivamo*

$$\tilde{\mathbb{E}}Z^2 \sim \frac{z^2/m^2}{1/m}\frac{1/e^{-1}-1}{1/(1-1/m)-1} \sim \frac{z^2}{m}\frac{e-1}{1/m} = z^2(e-1). \tag{2.7}$$

Naime, $\left(1-\frac{1}{m}\right)^m \sim e^{-1}$ *i* $\frac{1}{1-1/m} - 1 \sim \frac{1}{m}$ *za velike* m*. Dakle, u ovom slučaju procjenitelj ima ograničenu relativnu pogrešku.*
Ograničenost relativne pogreške postiže se i uzimanjem $\tilde{\pi} = c/m$*, za neki* c*, jer je tada* $\tilde{\mathbb{E}}Z^2 \sim z^2(e^c-1)/c^2$*. Optimalan* c *postiže se minimiziranjem* $(e^c-1)/c^2$*, što daje* $c^* = 1.59$*. Primijetimo da je uz optimalan* c *redukcija varijance minimalna;* $(e-1)c^{*2}/(e^{c^*}-1) = 1.12$*.*

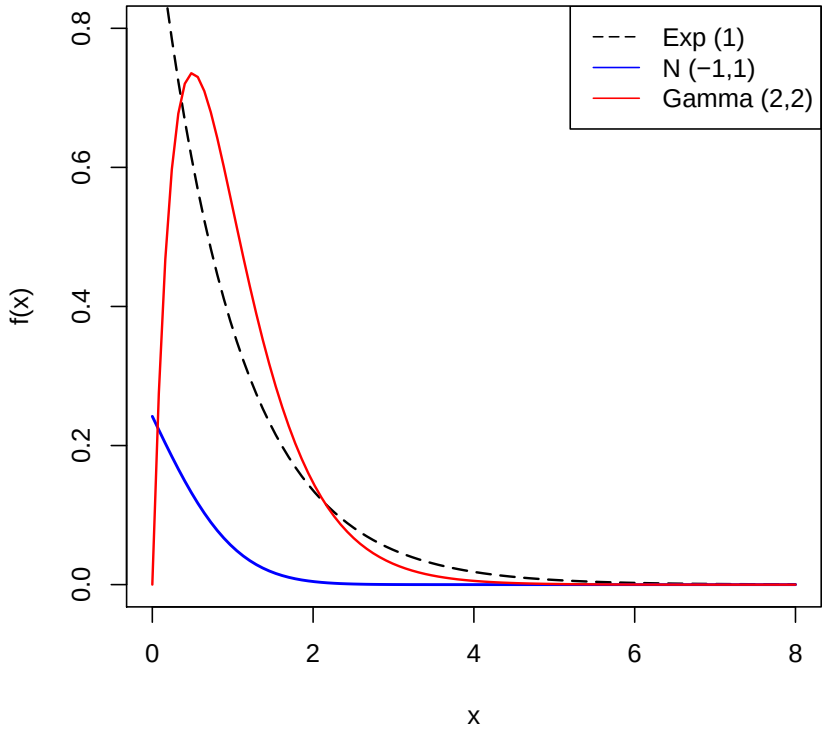
2.2 Distribucije lakog repa

U ovom dijelu razmotrit ćemo neke algoritme za simulaciju rijetkih događaja u slučaju distribucija lakog repa.

Definicija 2.3. *Kažemo da je funkcija distribucije* F *distribucija lakog desnog repa, ako je za barem jedan* $s > 0$*,* $\mathbb{E}\left[e^{sX}\right] < \infty$*.*

Intuitivno, distribucije lakog repa su one kojima repovi opadaju eksponencijalno ili brže. Najpoznatiji primjeri su eksponencijalna i gama distribucija kojima ćemo se baviti u nastavku. Na slici 2.1 prikazane su funkcije gustoće eksponencijalne distribucije s parametrom $\lambda = 1$, normalne distribucije s očekivanjem -1 i varijancom 1, te gama distribucije s parametrima $\alpha = 2$ i $\beta = 2$.

Neka su $X_1, X_2, \ldots$ nezavisne, jednako distribuirane slučajne varijable s funkcijom distribucije F. Prisjetimo li se definicije 1.15 i propozicije 1.16, funkciju izvodnicu momenata od X_i označit ćemo s $\hat{F}[s] = \mathbb{E}e^{sX_i}$ i funkcija $\kappa(s) = \log \hat{F}[s]$ je konveksna.



Slika 2.1: Funkcije gustoća nekih distribucija lakog repa

Pretpostavljamo da je F distribucija lakog repa, odnosno da je za neke $s > 0$, $\hat{F}[s] < \infty$. Prisjetimo se, eksponencijalna promjena mjere definirana je u 1.13 kao

$$F_\theta(dx) = \frac{e^{\theta x}}{\hat{F}[\theta]} F(dx).$$

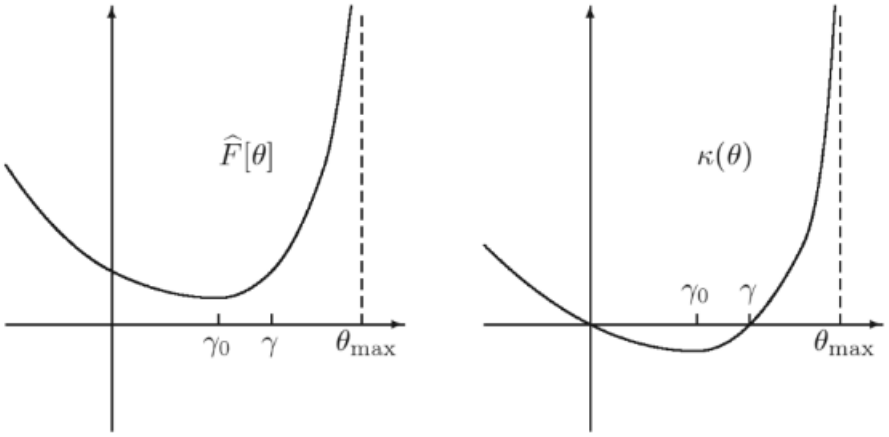
Omjer vjerodostojnosti $(d\mathbb{P}/d\mathbb{P}_\theta)_n$ iz uzorkovanja po važnosti za $X_1, \ldots, X_n$ tada je

$$L_{n,\theta} := L_n(F|F_\theta) = \prod_{k=1}^{n} \frac{\hat{F}[\theta]}{e^{\theta X_k}} = e^{-\theta S_n} \hat{F}[\theta]^n = e^{-\theta S_n + n\kappa(\theta)},$$

gdje je $S_n = X_1 + \cdots + X_n$. Ako je Z_n $\sigma(X_1, \cdots, X_n)$-izmjeriva slučajna varijabla, primjerice $Z_n = \sum_{i=1}^{n} X_i$ kao u uvodnom primjeru, tada je po (1.11)

$$\mathbb{E}Z_n = \mathbb{E}_\theta\big[Z_n L_{n,\theta}\big] = \mathbb{E}\big[Z_n e^{-\theta S_n + n\kappa(\theta)}\big]. \tag{2.8}$$

Označimo sa $\theta_{max} = \sup\{\theta : \hat{F}[\theta] < \infty\}$. Za distribucije lakog repa vrijedi $0 < \theta_{max} \leq \infty$. Promatrat ćemo one distribucije F za koje vrijedi da $\hat{F}[\theta] \uparrow \infty$, kad $\theta \uparrow \theta_{max}$. Za bolju predodžbu pogledajmo graf funkcije izvodnice momenata $\hat{F}[\theta]$ i $\kappa(\theta)$ u slučaju slučajne varijable s negativnim očekivanjem. S γ_0 označili smo rješenje $\hat{F}'[\gamma_0] = \kappa'(\gamma_0) = 0$, odnosno vrijednost koja minimizira $\hat{F}$ i κ, a s γ rješenje od $\kappa(\gamma) = 0$, odnosno $\hat{F}[\gamma] = 1$, različito od nule.



Važno je uočiti da je,

$$\mu_\theta := \mathbb{E}_\theta X = \mathbb{E}\left[\frac{Xe^{\theta X}}{\hat{F}[\theta]}\right] = \frac{\hat{F}'[\theta]}{\hat{F}[\theta]} = \kappa'(\theta). \tag{2.9}$$

Stoga je $\mu_\theta < 0$, kad je $\theta < \gamma_0$, $\mu_\theta > 0$ za $\theta > \gamma_0$ i $\mu_{\gamma_0} = 0$.
Označimo li s $\hat{F}_\theta[s] = \mathbb{E}_\theta[e^{sX}]$ i $\kappa_\theta(s) = \log \hat{F}_\theta(s)$, primijetimo da vrijedi:

$$\kappa_\theta(s) = \log \mathbb{E}_\theta\left[e^{sX}\right] = \log \mathbb{E}\left[\frac{e^{sX}e^{\theta X}}{\hat{F}[\theta]}\right] = \log \mathbb{E}\left[e^{(s+\theta)X}\right] - \log \hat{F}[s] = \kappa(s+\theta) - \kappa(s). \tag{2.10}$$

Siegmundov algoritam

Promotrimo sada sljedeću diskretnu slučajnu šetnju $(X_n, n \in \mathbb{N})$, $S_n = X_1 + \cdots + X_n$. Neka je F distribucija od X_i i pretpostavimo da F nije koncentrirana na $(-\infty, 0]$ te da je $\mathbb{E}X_i < 0$. Želimo procijeniti

$$z(x) := \mathbb{P}(\tau(x) < \infty), \quad \text{gdje je} \quad \tau(x) = \inf\{n : S_n > x\}$$

u slučaju da je x jako velik, odnosno $z(x)$ mali. Kao što ćemo vidjeti, ovaj problem ima veliku primjenu u računanju tzv. vjerojatnosti propasti (eng. *ruin probabilities*).

Implementiramo li uzorkovanje po važnosti s eksponencijalnom promjenom mjere, iz (2.8) dobivamo

$$\begin{aligned} z(x) = \mathbb{P}(\tau(x) < \infty) &= \mathbb{E}_\theta \big[L_{\tau(x),\theta}; \tau(x) < \infty \big] \\ &= \mathbb{E}_\theta \Big[e^{-S_{\tau(x)} + \tau(x)\kappa(\theta)}; \tau(x) < \infty \Big]. \end{aligned} \tag{2.11}$$

Pretpostavimo da postoji $\gamma > 0$, takav da je $\hat{F}[\gamma] = 1$, odnosno $\kappa(\gamma) = 0$. Znamo da je κ konveksna, $\kappa(0) = 0$ i $\kappa'(0) = \mathbb{E}X < 0$ po pretpostavci problema. Uzimamo li u obzir distribucije F za koje $\hat{F}[\theta] \uparrow \infty$ za $\theta \uparrow \theta_{max}$, jer je κ padajuća u 0 i konveksna, slijedi da postoji $\gamma > 0$, takav da je $\kappa(\gamma) = 0$, odnosno $\hat{F}[\gamma] = 1$. Sada mora vrijediti da je κ rastuća u γ, odosno da je $\kappa'(\gamma) = \mathbb{E}_\theta X \geq 0$ tj. $\mathbb{P}_\theta(\tau(x) < \infty) = 1$. Primijetimo da to odgovara razmatranjima u prvom poglavlju, kako je optimalna distribucija važnosti ona koncentrirana na područje u kojem se nalazi naš događaj. Za takav θ, (2.11) postaje

$$z(x) = \mathbb{P}(\tau(x) < \infty) = \mathbb{E}_\theta L_{\tau(x),\theta} = \mathbb{E}_\theta \Big[e^{-\theta S_{\tau(x)} + \tau(x)\kappa(\theta)} \Big]. \tag{2.12}$$

Stoga $z(x)$ možemo procijeniti osnovnom Monte Carlo metodom, simulirajući $Z(x) = L_{\tau(x),\theta}$ iz $\mathbb{P}_\theta$.
Pokazat će se da je γ upravo optimalna vrijednost za θ. Stavimo li $\xi(x) = S_{\tau(x)} - x$, u ovom slučaju (2.12) postaje

$$z(x) = \mathbb{P}(\tau(x) < \infty) = \mathbb{E}_\gamma e^{-\gamma S_\tau(x)} = e^{-\gamma x} \mathbb{E}_\gamma e^{-\gamma \xi(x)} \tag{2.13}$$

Dokazat ćemo da ovaj algoritam ima ograničenu relativnu pogrešku. Za dokaz je ključna *Cramér Lundbergova aproksimacija* koju ovdje iskazujemo u odgovarajućoj formi, a dokaz se može naći u Asmussen [1].

Teorem 2.4. *[Cramér Lundbergova aproksimacija] $z(x) \sim Ce^{-\gamma x}$, kad $x \to \infty$, gdje je $C = \mathbb{E}_\gamma e^{-\gamma \xi(\infty)}$.*

Teorem 2.5. *Algoritam $Z(x) = e^{-\gamma x} e^{-\gamma \xi(x)}$, simuliran iz P_γ ima ograničenu relativnu pogrešku.*

Dokaz. Po Cramér Lundbergovoj aproksimaciji

$$z(x) = \mathbb{E}_\gamma Z(x) \sim Ce^{-\gamma x}$$

i

$$\mathbb{E}_\gamma Z(x)^2 = e^{-2\gamma x}\mathbb{E}_\gamma e^{-2\gamma\xi(x)} \sim C_1 e^{-2\gamma x}.$$

Po Jensenovoj nejednakosti, $(\mathbb{E}_\gamma Z(x))^2 < \mathbb{E}_\gamma Z(x)^2$, odnosno $C^2 < C_1$ pa stoga

$$\mathrm{Var}_g Z(x) \sim C_1 e^{-2\gamma x} - \left(Ce^{-\gamma x}\right)^2 \sim C_2 e^{-2\gamma x},$$

gdje je $C_2 = C_1 - C^2 > 0$. Relativna pogreška sada je

$$\frac{\sqrt{\mathrm{Var}_g Z(x)}}{z(x)} \sim \frac{C_2^{1/2} e^{-\gamma x}}{Ce^{-\gamma x}} = C_3,$$

uz $C_3 = \sqrt{C_2}/C$. Kako C_3 ne raste s $x \to \infty$, relativna pogreška je ograničena. □

Pogledajmo kako izgleda ova optimalna promjena mjere na jednostavnom primjeru.

Primjer 2.6. *Neka je F distribucija normalne razdiobe s negativnim očekivanjem i varijancom 1, odnosno $N(-\mu, 1), \mu > 0$. Tada je*

$$\begin{aligned}
\hat{F}[s] = \mathbb{E}\left[e^{sX}\right] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{sx} e^{-\frac{(x+\mu)^2}{2}} dx \\
&= \left[z = x + \mu\right] \\
&= \frac{1}{\sqrt{2\pi}} e^{-s\mu} \int_{-\infty}^{\infty} e^{sz - \frac{z^2}{2}} dz \\
&= e^{-s\mu} e^{\frac{s^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{(z-s)^2}{2}} dz \\
&= e^{-s\mu + \frac{s^2}{2}}
\end{aligned}$$

Sada γ koji minimizira $\hat{F}$ mora zadovoljavati $-\mu\gamma + \gamma^2/2 = 0$, iz čega slijedi $\gamma = 2\mu$ jer uzimamo pozitivno rješenje. Slijedi

$$\hat{F}_\gamma[s] = \mathbb{E}_\gamma\left[e^{sX}\right] = \mathbb{E}\left[e^{sX}/e^{-\gamma x}\right] = \hat{F}[s+\gamma] = e^{\mu s + s^2/2}.$$

Dakle, $F_\gamma \sim N(\mu, 1)$.

Optimalna promjena mjere u Siegmundovom algoritmu

Promotrimo sada algoritam uzorkovanja po važnosti uz neku proizvoljnu distribuciju važnosti $\tilde{F}$, ne nužno iz eksponencijalne familije. To znači da simuliramo $X_1, \ldots, X_{\tau(x)}$ iz $\tilde{F}$ i definiramo procjenitelj $Z(x)$ sa

$$Z(x) := L_{\tau(x)}(F|\tilde{F}) = \frac{dF}{d\tilde{F}}(X_1) \cdots \frac{dF}{d\tilde{F}}(X_{\tau(x)}), \tag{2.14}$$

gdje je $dF/d\tilde{F}$ Radon-Nikodymova derivacija u svjetlu napomene 1.3.

Teorem 2.7. *Algoritam dan u (2.14) je logaritamski efikasan ako i samo ako je* $\tilde{F} = F_\gamma$.

Dokaz. Dovoljnost je dokazana u teoremu 2.5 za jaču tvrdnju - ograničenost relativne pogreške. Pretpostavimo da $\tilde{F} \neq F_\gamma$, ali da je algoritam u (2.14) logaritamski efikasan. Po kompoziciji Radon-Nikodymovih derivacija (Teorem 1.2) slijedi

$$\begin{aligned}\tilde{\mathbb{E}}Z(x)^2 &= \tilde{\mathbb{E}}L^2_{\tau(x)}(F|\tilde{F}) = \tilde{\mathbb{E}}\big[L^2_{\tau(x)}(F|F_\gamma)L^2_{\tau(x)}(F_\gamma|\tilde{F})\big] \\ &= \mathbb{E}_\gamma\big[L^2_{\tau(x)}(F|\tilde{F})L_{\tau(x)}(F_\gamma|\tilde{F})\big] = \mathbb{E}_\gamma\left[\prod_{i=1}^{\tau(x)}\frac{dF_\gamma}{d\tilde{F}}(X_i)\left(\frac{dF}{dF_\gamma}(X_i)\right)^2\right] \\ &= \mathbb{E}_\gamma e^{K_1+\cdots+K_{\tau(x)}},\end{aligned}$$

gdje je

$$\begin{aligned}K_i &= \log\left(\frac{dF_\gamma}{d\tilde{F}}(X_i)\left(\frac{dF}{dF_\gamma}(X_i)\right)^2\right) = \log\left(\frac{dF_\gamma}{d\tilde{F}}(X_i)\right) + 2\log\left(\frac{dF}{dF_\gamma}(X_i)\right) \\ &= -\log\frac{d\tilde{F}}{dF\gamma}(X_i) - 2\gamma X_i.\end{aligned}$$

Treća jednakost slijedi iz definicije eksponencijalne promjene mjere i γ. Označimo li

$$\varepsilon' = -\mathbb{E}_\gamma \log\frac{d\tilde{F}}{dF_\gamma}(X_i),$$

iz propozicije 1.4 slijedi da je $\varepsilon' > 0$. Sada imamo

$$\mathbb{E}_\gamma K_i = \varepsilon' - 2\gamma\mathbb{E}_\gamma X_i = \varepsilon' - 2\gamma\mu_\gamma,$$

gdje je $\mu_\gamma = \mu_{F_\gamma} > 0$, zbog razmatranja prije teorema 2.5. Kako su $K_1, \cdots, K_{\tau(x)}$ nezavisne i jednako distribuirane, iz Jensenove nejednakosti i Waldove jednakosti (1.6) slijedi

$$\tilde{\mathbb{E}}Z(x)^2 \geq e^{\mathbb{E}_\gamma(K_1+\cdots+K_{\tau(x)})} = e^{(\varepsilon'-2\gamma\mu_\gamma)\mathbb{E}_\gamma\tau(x)}.$$

Po zakonu velikih brojeva $\mathbb{E}_\gamma\tau(x)/x \to 1/\mu_\gamma$, pa pomoću Cramér Lundbergove aproksimacije (2.4) slijedi da za ε'' takav da $0 < \varepsilon'' < \varepsilon'$, $0 < \varepsilon < \varepsilon''/\gamma\mu_\gamma$ vrijedi:

$$\begin{aligned}\liminf_{x\to\infty}\frac{\tilde{\mathbb{E}}Z(x)^2}{z(x)^{2-\varepsilon}} &= \liminf_{x\to\infty}\frac{\tilde{\mathbb{E}}Z(x)^2}{C^{2-\varepsilon}e^{-2\gamma x+\varepsilon\gamma x}}\\ &\geq \liminf_{x\to\infty}\frac{e^{\mathbb{E}_\gamma\tau(x)(\varepsilon'-2\gamma\mu_\gamma)}}{C^{2-\varepsilon}e^{-2\gamma x+\varepsilon\gamma x}}\\ &= \liminf_{x\to\infty}\frac{e^{x(\varepsilon''/\mu_\gamma-2\gamma)}}{C^{2-\varepsilon}e^{-2\gamma x+\varepsilon\gamma x}}\\ &= \liminf_{x\to\infty}C^{\varepsilon-2}e^{x(\varepsilon''/\mu_\gamma-\varepsilon\gamma)}=\infty.\end{aligned}$$

Dakle, relativna pogreška je neograničena, odnosno algoritam nije logaritamski efikasan, što je u kontradikciji s pretpostavkom.

□

Složeni Poissonov proces

Jedan od važnih primjera rijetkih događaja su vjerojatnosti propasti u teoriji rizika. Osnovni model u modeliranju rizika osiguranja je Cramér-Lundbergov model koji pretpostavlja sljedeću situaciju. Neka tražene isplate osiguranja dolaze po Poissonovom procesu $(N_t, t \geq 0)$ s parametrom λ. Neka su traženi iznosi isplata $X_1, X_2, \ldots$ nenegativne nezavisne jednakodistribuirane slučajne varijable s distribucijom F i pretpostavimo da je N_t nezavisan od $X_1, X_2, \ldots$. Tada možemo promatrati ukupno potraživanje u nekom vremenu, primjerice u jednoj godini (broj isplata u godini dana N_1 označit ćemo s N i primijetimo da vrijedi $N \sim P(\lambda)$) i ono je dano s

$$C = \sum_{i=1}^{N} X_i.$$

Rijedak događaj koji ćemo promatrati je $A(x) = \{C > x\}$ za velike x, odnosno vjerojatnost da je ukupan broj traženih isplata iznimno velik što je vrlo nepogodno za osiguravajuće društvo. Pokazat ćemo da je i u ovom slučaju uzorkovanje po važnosti uz eksponencijalnu promjenu mjere logaritamski efikasno, pod nekim uvjetima na distribuciju F. Primijetimo da pod gornjim pretpostavkama, uz iste oznake kao u

prijašnjim razmatranjima, vrijedi

$$\begin{aligned}\kappa(\alpha) &= \log \mathbb{E}[e^{\alpha C}] = \log \mathbb{E}[e^{\alpha(X_1+\cdots+X_N)}] \\ &= \log \sum_{n=0}^{\infty} \mathbb{E}[e^{\alpha(X_1+\cdots+X_n)}]\mathbb{P}(N=n) \\ &= \log \sum_{n=0}^{\infty} \frac{(\mathbb{E}e^{\alpha X_1})^n \lambda^n}{n!} e^{-\lambda} = \log(e^{\lambda \hat{F}[\alpha]} e^{-\lambda}) \\ &= \lambda(\hat{F}[\alpha] - 1).\end{aligned}$$

Uz eksponencijalnu promjenu mjere s distribucijom važnosti $F_\theta(dx) = e^{\theta x} F(dx)/\hat{F}[\theta]$, iz (2.10) slijedi

$$\kappa_\theta(\alpha) = \kappa(\alpha+\theta) - \kappa(\theta) = \lambda(\hat{F}[\alpha+\theta] - \hat{F}[\theta]) = \lambda_\theta(\hat{F}_\theta[\alpha] - 1),$$

gdje je $\lambda_\theta = \lambda \hat{F}[\theta]$. Uz ovu notaciju, klasična Esscherova aproksimacija (vidi Jensen [8]) za $z(x)$ je

$$z(x) = \mathbb{P}(C > x) \sim \frac{e^{-\theta x + \kappa(\theta)}}{\theta\sqrt{2\pi\lambda \hat{F}''[\theta]}}, \quad x \to \infty. \tag{2.15}$$

U svjetlu prijašnjih razmatranja o optimalnoj promjeni mjere, $\theta = \theta(x)$ ćemo odrediti kao rješenje od $\mathbb{E}_\theta C = x$, odnosno $x = \kappa'(\theta) = \lambda \hat{F}'[\theta]$.
Esscherova aproksimacija (2.15) zahtjeva neke uvjete na gustoću f, koji su detaljno opisani u [8]. Pokazuje se da je bilo koji od sljedeća dva uvjeta dovoljan.

A: $f(x) \sim c_1 x^{\alpha-1} e^{\delta x}, \quad x \to \infty$

B: $f(x) = q(x)e^{h(x)}$, gdje je $q(x)$ ograničena i $h(x)$ konveksna na intervalu oblika $[x_0, x^*)$, gdje je $x^* = \sup\{x : f(x) > 0\}$. Također mora vrijediti $\int_0^\infty f(x)^m dx < \infty$, za neki $m \in (1,2)$.

Teorem 2.8. *Pretpostavimo da vrijedi ili A ili B. Tada je procjenitelj* $Z(x) = e^{-\theta C + \kappa(\theta)} \mathbb{1}_{\{C>x\}}$ *za* $z(x)$ *logaritamski efikasan.*

Dokaz. Ovdje donosimo kratke crte dokaza, a svi detalji dani su u Jensen [8].

Vrijedi sljedeća heuristika:

$$\begin{aligned}\mathbb{E}_\theta\left[e^{-2\theta(C-x)}; C > x\right] &\sim \int_0^\infty e^{-2\theta\sqrt{\lambda\hat{F}_\theta''}y}\frac{1}{\sqrt{2*\pi}}e^{-y^2/2}dy \\ &= \left[z = -2\theta\sqrt{\lambda\hat{F}_\theta''}y\right] \\ &= \frac{1}{2\theta\sqrt{2\pi\lambda\hat{F}_\theta''}}\int_0^\infty e^{-z}e^{-z^2/(8\theta^2\lambda\hat{F}_\theta'')}dz \\ &\sim \frac{1}{2\theta\sqrt{2\pi\lambda\hat{F}_\theta''}}\int_0^\infty e^{-z}dz = \frac{1}{2\theta\sqrt{2\pi\lambda\hat{F}_\theta''}}.\end{aligned}$$

Stoga dobivamo,

$$\mathbb{E}_\theta Z(x)^2 = e^{-2\theta x+2\kappa(\theta)}\mathbb{E}_\theta\left[e^{-2\theta(C-x)}; C > x\right] \sim \frac{e^{-2\theta x+2\kappa(\theta)}}{2\theta\sqrt{2\pi\lambda\hat{F}_\theta''}}.$$

Usporedimo li to s (2.15), da bismo dobili logaritamsku efikasnost lako se vidi da treba vrijediti

$$\log(\theta\sqrt{\hat{F}''[\theta]}) = o(\theta x - \kappa(\theta)). \tag{2.16}$$

U slučaju **A**, $\theta_{max} = \delta$ i integriranjem se dobije

$$\hat{F}[\theta] \sim \frac{c_1}{(\delta-\theta)^\alpha}, \quad \hat{F}'[\theta] \sim \frac{c_2}{(\delta-\theta)^{\alpha+1}}, \quad \hat{F}''[\theta] \sim \frac{c_3}{(\delta-\theta)^{\alpha+2}}, \quad \theta \uparrow \theta_{max}.$$

Kako je $\theta = \theta(x) \uparrow \theta_{max} = \delta < \infty$, za $x \to \infty$ dobivamo

$$\theta x - \kappa(\theta) = \theta\lambda\hat{F}'[\theta] - \lambda(\hat{F}[\theta] - 1) \sim \frac{c}{(\delta-\theta)^{\alpha+1}}.$$

Sada se lako dobije (2.16).
U slučaju **B** vrijedi $\theta_{max} = +\infty$ i možemo pisati

$$\begin{aligned}\theta x - \kappa(\theta) = \theta\kappa'(\theta) - \int_0^\theta \kappa'(s)ds &= \int_0^\theta s\kappa''(s)ds \\ &\geq \theta\int_0^\theta \frac{s}{\theta}ds\int_0^\theta \frac{\kappa''(s)}{\theta}ds = \frac{\theta}{2}(\kappa'(\theta) - \kappa'(0)),\end{aligned}$$

gdje smo iskoristili formulu za parcijalnu integraciju i propoziciju 1.5. U [8] je pokazano da je $\hat{F}'''[\theta]$ reda veličine $\kappa'(\theta)^2\kappa(\theta)$, što je $O(\kappa'(\theta)^3)$. Kako je $\log(\theta\kappa'(\theta)^{3/2}) = o(\theta\kappa'(\theta))$, slijedi (2.16). □

Proširimo sada Cramér Lundbergov model na sljedeći način. Promatramo proces pristizanja zahtjeva za isplatom

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

gdje su veličine zahtjeva X_i nezavisne jednako distribuirane slučajne varijable s distribucijom F. Pretpostavljamo i da su one nezavisne od vremena dolazaka zahtjeva T_n, danih sa

$$T_0 = 0, \quad T_n = W_1 + \cdots + W_n, \quad n \geq 1,$$

gdje su W_n nezavisna jednako distribuirana pozitivna vremena čekanja. Tada je proces broja zahtjeva $N(t) = \#\{n \geq 1 : T_n \leq t\}$ nezavisan od X_i. Pretpostavimo također da se premije konstantno akumuliraju po stopi c, odnosno po linearnoj funkciji $p(t) = ct$, te da je početna imovina osiguravajućeg društva u. Tada proces rizika definiramo kao

$$U(t) = u + p(t) - C(t), \quad t \geq 0.$$

U kontekstu rijetkih događaja zanimljivo je promatrati vjerojatnost ultimativne propasti, odnosno događaja

$$z(u) = \mathbb{P}(\inf_{t \geq 0} U(t) \leq 0) = \mathbb{P}(\tau < \infty), \quad \text{za} \quad \tau = \inf\{t > 0 : U(t) < 0\}.$$

S obzirom na konstrukciju procesa rizika U, propast se može dogoditi jedino u vremenima dolaska zahtjeva $t = T_n$, budući da je funkcija $U(t)$ linearno rastuća na intervalima $[T_n, T_{n+1})$. Vjerojatnost propasti stoga možemo izraziti pomoću vremena čekanja W_n, veličine zahtjeva X_n i rate premije c na sljedeći način:

$$\begin{aligned}
\left\{ \inf_{t>0} U(t) < 0 \right\} &= \left\{ \inf_{n \geq 1} U(T_n) < 0 \right\} \\
&= \left\{ \inf_{n \geq 1} [u + p(T_n) - S(T_n)] < 0 \right\} \\
&= \left\{ \inf_{n \geq 1} \left[u + cT_n - \sum_{i=1}^{n} X_i \right] < 0 \right\}.
\end{aligned}$$

U zadnjem koraku iskoristili smo činjenicu da je $N(T_n) = \#\{i \geq 1 : T_i \leq T_n\} = n$ g.s., budući da smo pretpostavili da su $W_i > 0$. Stavimo li

$$Z_n = X_n - cW_n, \quad S_n = Z_1 + \cdots + Z_n, \quad n \geq 1, \quad S_0 = 0,$$

vjerojatnost propasti možemo pisati i kao

$$z(x) = \mathbb{P}\big(\inf_{n\geq 1}(-S_n) < -u\big) = \mathbb{P}\big(\sup_{n\geq 1} S_n > u\big) = \mathbb{P}(\tau(u) < \infty),$$

za $\tau(u) = \inf\{n : S_n > u\}$.

Primjer 2.9. *Pretpostavimo sada da u osiguravajućem društvu zahtjevi za isplatom dolaze po Poissonovom procesu s parametrom λ, a veličine zahtjeva X_i su nezavisne eksponencijalno distribuirane slučajne varijable s parametrom ν. Tada su vremena čekanja do dolaska zahtjeva W_i eksponencijalno distribuirana s parametrom λ i nezavisna su od X_i. Vjerojatnost propasti sada možemo procijeniti Siegmundovim algoritmom uz odgovarajuću distribuciju važnosti. Funkcija izvodnica momenata od $Z_i = X_i - cW_i$ dana je sa*

$$\begin{aligned}\hat{F}[s] &= \mathbb{E}\big[e^{sZ_i}\big] = \mathbb{E}\big[e^{sX_i}\big]\mathbb{E}\big[e^{-scW_i}\big] \\ &= \int_0^\infty \lambda e^{(s-\lambda)x}dx \int_0^\infty \nu e^{(-sc-\nu)x}dx \\ &= \frac{\lambda}{\lambda - s}\frac{\nu}{\nu + sc}.\end{aligned}$$

Optimalan parametar distribucije važnosti γ, opet ćemo odrediti kao rješenje od $\hat{F}[\gamma] = 1$, te jednostavnim računom dobivamo $\gamma = \nu - \lambda/c$. Odredimo sad distribuciju važnosti pomoću njene funkcije izvodnice momenata. Iz (2.10) zbog $\hat{F}[\gamma] = 1$ te gornjih razmatranja slijedi

$$\begin{aligned}\hat{F}_\gamma[s] = \hat{F}[s+\gamma] &= \frac{\nu}{\nu - (s+\gamma)}\frac{\lambda}{\lambda + c(s+\gamma)} \\ &= \frac{\nu}{\lambda/c - s}\frac{\lambda}{c(s+\nu)} \\ &= \frac{\nu}{\nu + s}\frac{\lambda}{\lambda - sc} \\ &= \mathbb{E}\big[e^{-sX}\big]\mathbb{E}\big[e^{scW}\big] \\ &= \mathbb{E}\big[e^{-sZ}\big] = \hat{F}[-s].\end{aligned}$$

Zanimljivo, dobili smo da je optimalna distribucija važnosti, distribucija slučajne varijable $Z_i = cW_i - X_i$ gdje je $W_i \sim Exp(\lambda)$, a $X_i \sim Exp(\nu)$, odnosno proces obrnut onom kojeg simuliramo.

Teorija velikih devijacija i problem optimalnog puta

Teorija velikih devijacija općenito se bavi graničnim ponašanjem vjerojatnosti rijetkih događaja. Primjena na eksponencijalnu promjenu mjere ima nekoliko varijanti, a ovdje ćemo razmotriti koncept optimalnog puta. I dalje ćemo se baviti slučajnom šetnjom s distribucijom F takvom da za neki $s > 0$ vrijedi $\hat{F}[\theta] = \mathbb{E}\left[e^{sX}\right] < \infty$ te označiti $\kappa(\theta) = \log \hat{F}[\theta]$. Trebat će nam sljedeća funkcija intenziteta:

$$I(y) = \sup_{\theta \in \Theta}(\theta y - \kappa(\theta)), \quad y \in \{\kappa'(\theta) : \theta \in \Theta\}.$$

Radi jednostavnosti pretpostavit ćemo da je Θ otvoren. Tada $\kappa(\theta)/|\theta| \uparrow \infty$ kad se približavamo rubu od Θ. Slijedi da se supremum postiže za neki $\theta(y)$ koji zadovoljava $\kappa'(\theta(y)) = y$ pa je $I(y) = \theta(y)y - \kappa(\theta(y))$.
Važna tvrdnja za eksponencijalnu promjenu mjere u Siegmundovom algoritmu dana je u sljedećoj propoziciji.

Propozicija 2.10. *Minimum* $\min\limits_{0<y<\infty} \dfrac{I(y)}{y}$ *postiže se za* $y^* = \kappa'(\gamma)$*, gdje je* γ *pozitivno rješenje od* $\kappa(\gamma) = 0$.

Dokaz. Kad $y \downarrow 0$, odosno $\kappa'(\theta(y)) \downarrow 0$, $\kappa(\theta(y))$ teži u neku točku minimuma koja postoji jer je κ konveksna. Također, $\theta(y)$ je rastuća funkcija ograničena odozgo pa slijedi da $I(y)/y \to \infty$ kad $y \downarrow 0$. Budući da je $\theta(y)$ rastuća i κ konveksna, vrijedi da je za velike y, $\kappa(\theta(y)) > 0$, pa iz (2.17) slijedi da je $I(y)/y$ rastuća za velike y pa minimum postoji. Deriviranjem dobivamo

$$\begin{aligned} I'(y) &= \theta'(y)y + \theta(y) - \theta'(y)\kappa'(\theta(y)) = \theta(y) \\ \frac{d}{dy}\frac{I(y)}{y} &= \frac{yI'(y) - I(y)}{y^2} = \frac{y\theta(y) - \theta(y)y + \kappa(\theta(y))}{y^2} \\ &= \frac{\kappa(\theta(y))}{y^2}. \end{aligned} \tag{2.17}$$

Izjednačimo li zadnju jednakost s 0, dobivamo $K(\theta(y^*)) = 0$ iz čega slijedi $\theta(y^*) = \gamma$, jer gledamo samo vrijednosti y takve da je $y > 0$ pa ne uključujemo $\theta(y^*) = 0$. Iz defincije y sada odmah slijedi $y^* = \kappa'(\gamma)$. □

Jedna od glavnih tema u teoriji velikih devijacija je procijeniti vjerojatnost da slučajna šetnja ili neki općenitiji proces slijedi neki atipičan put. Za diskretnu slučajnu šetnju definirajmo put $S^{(n)}$ kao $S^{(n)}(t) = S_{\lfloor nt \rfloor}/n$ za $0 \le t \le 1$, gdje $\lfloor x \rfloor$ označava najmanji cijeli broj manji od x. Zanima nas vjerojatnost da $S^{(n)}$ slijedi put drugačiji od onog

danog zakonom velikih brojeva $\varphi_0(t) = \mu t$. Važan rezultat teorije velikih devijacija, poznat kao Mogulskijev teorem (vidi Bucklew [6]) kaže da uz neke uvjete vrijedi

$$\mathbb{P}(S^{(n)}(\cdot) \in \mathscr{F}) \approx e^{-n \inf\limits_{\varphi \in \mathscr{F}} \int_0^1 I(\varphi'(t))dt} \tag{2.18}$$

za neke podskupove $\mathscr{F}$ neprekidnih puteva takve da $\varphi_0 \notin \mathscr{F}$. U puno primjera postoji jedinstveni put φ^* za koji se gornji minimum postiže i njega zovemo *optimalni put*. Da bismo razumjeli kako slučajna šetnja u našem primjeru dostiže x, optimizirat ćemo ne samo po φ već i po n. Tada x možemo pisati u formi $x = ny$ i uzet ćemo da je $\mathscr{F}$ skup neprekidnih funkcija na $[0, 1]$ takvih da je $\varphi(0) = 0, \varphi(1) = y$. Sada (2.18) postaje

$$\mathbb{P}(S_n \sim x) \approx e^{-\frac{x}{y} \inf\limits_{\varphi \in \mathscr{F}} \int_0^1 I(\varphi'(t))dt}. \tag{2.19}$$

Kako je I konveksna, iz Jensenove nejednakosti slijedi,

$$\int_0^1 I(\varphi'(t))dt \geq I(\int_0^1 \varphi'(t)dt) = I(\varphi(1) - \varphi(0)) = I(y),$$

gdje jednakost vrijedi ako i samo ako je φ' konstanta, odnosno $\varphi(t) = ty$. Dakle, za fiksni n vrijedi

$$\mathbb{P}(S_n \sim x) \approx e^{-\frac{x}{y} I(y)}.$$

Kako je minimiziranje po n isto kao minimiziranje po y sada iz propozicije 2.10 slijedi da je se minimum postiže za $y^* = \kappa'(\gamma)$. Dakle, najvjerojatniji put kojim slučajna šetnja poprima vrijednost veću od x je onaj u kojem je

$$\tau(x) = n = \frac{x}{y} = \frac{x}{\kappa'(\gamma)},$$

i put je linearna funkcija s nagibom $y = \kappa'(\gamma)$, odnosno $S^{(n)}(t) = \kappa'(\gamma)t$. Primijetimo da je zbog (2.9) ovaj put upravo put dan zakonom velikih brojeva slučajne šetnje s distribucijom F_γ iz Siegmundovog algoritma.

Iako su rezultati ovog tipa nešto više heuristični od prethodnih (primijetimo da smo odredili samo očekivanje distribucije važnosti), teorija velikih devijacija jedan je od najpopularnijih pristupa simulaciji rijetkih događaja. Ideja je razumjeti optimalna svojstva optimalne eksponencijalne promjene mjere za jednostavne probleme i zatim ih pomoću rezultata teorije velikih devijacija generalizirati na složenije probleme. Više o primjeni teorije velikih devijacija na simulacije može se pronaći u Bucklew [6].

2.3 Distribucije teškog repa

U teoriji vjerojatnosti distribucijama teškog repa smatraju se one distribucije koje nisu eksponencijalno ograničene, odnosno imaju "teže" repove od eksponencijalne distribucije. Ovdje ćemo promatrati klasu subeksponencijalnih distribucija.

Definicija 2.11. *Kažemo da je funkcija distribucije subeksponencijalna ako za svaki* $n \in \mathbb{N}$ *za nezavisne slučajne varijable* $X_1, X_2, \ldots, X_n$ *s distribucijom* F *vrijedi*

$$\frac{\mathbb{P}(X_1 + \cdots + X_n > x)}{\mathbb{P}(X_1 > x)} \to n, \quad x \to \infty. \tag{2.20}$$

Može se pokazati da je ova definicija ekvivalentna s

$$\mathbb{P}\Big(\max_{i \le n} X_i > x | S_n > x\Big) \to 1, \quad x \to \infty.$$

To je glavna intuicija u pozadini distribucija teških repova - najvjerojatniji način da zbroj slučajnih varijabli s teškim repovima postane velik jest da jedna od njih postane velika, za razliku od slučaja lakih repova gdje sve varijable u sumi pridonose njenoj veličini. Sve poznate distribucije teškog repa poput lognormalne, Paretove ili t-distribucije su subeksponencijalne. Posebno ćemo razmatrati sljedeće slučajeve subeksponencijalnih distribucija, koje možemo promotriti na slici 2.2.

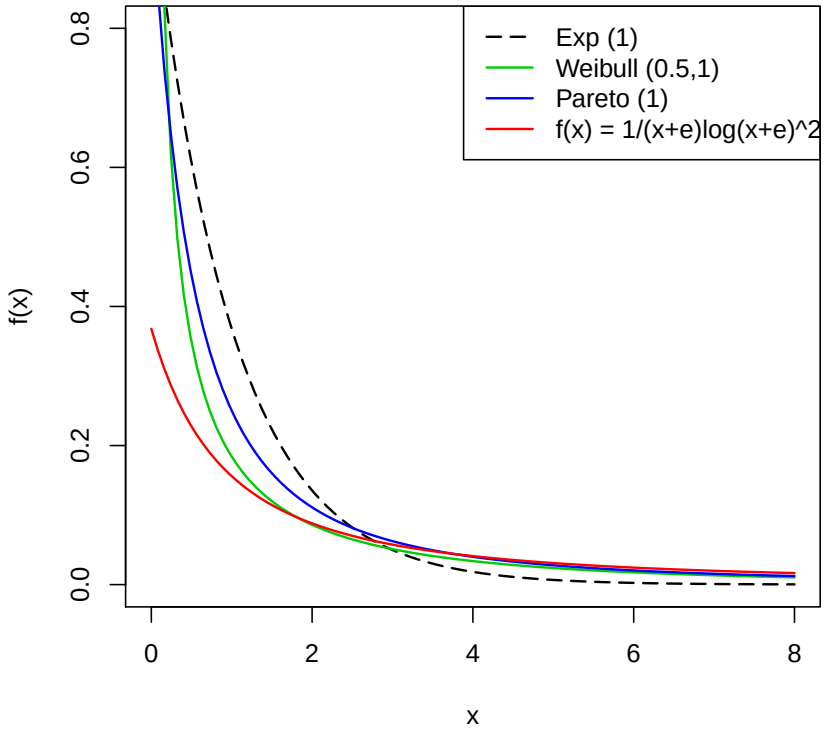
Definicija 2.12. *Kažemo da je distribucija* F ***regularno varirajuća*** *ako vrijedi* $\bar{F}(x) := 1 - F(x) = L(x)/x^\alpha$*, za* $\alpha > 0$ *i* L *takav da vrijedi* $L(tx)/L(x) \to 1$*, kad* $x \to \infty$ *za svaki fiksni* $t > 0$*.*

Važan primjer regularno varirajuće distribucije je **Paretova** distribucija s gustoćom $f(x) = \alpha/(1+x)^{\alpha+1}, 0 < x < \infty$, odnosno repom $1/(1+x)^\alpha$.

Definicija 2.13. *Slučajna varijabla ima* ***lognormalnu distribuciju*** *ako je* $X = e^Y$*, gdje je* Y *normalno distribuirana,* $Y \sim N(\mu, \sigma^2)$*.*

Definicija 2.14. *Slučajna varijabla ima* ***Weibullovu distribuciju*** *ako je* $\bar{F}(x) = e^{-cx^\beta}$*, za neki* $0 < \beta$*. U slučaju* $\beta < 1$ *distribucija je subeskponencijalna.*

Algoritmi za simulaciju rijetkih događaja u slučaju teških repova znatno su manje razvijeni nego u slučaju lakih repova. U nastavku ćemo se baviti samo sljedećim vrlo jednostavnim problemom. Neka su $X_1, X_2, \ldots$ nezavisne jednako distribuirane i nenegativne slučajne varijable sa subeksponencijalnom funkcijom distribucije F. Promatrat ćemo problem procjene vjerojatnosti $z(x) = \mathbb{P}(S_n > x)$ za velike x, gdje je $S_n = X_1 + \cdots + X_n$. Kako je F subeksponencijalna, iz (2.20) slijedi da je



Slika 2.2: Funkcije gustoća nekih distribucija teškog repa

$z(x) \sim n\bar{F}(x)$, no pokazalo se da je takva simulacija u puno slučajeva netočna. Za razliku od algoritama u slučaju lakih repova, koji se također bave računanjem $\mathbb{P}(S_n > x)$, ovdje ćemo se baviti limesom $x \to \infty$, umjesto $n \to \infty$, kako bi se uklopili u definiciju subeksponencijalnih distribucija.

Uvjetni Monte Carlo algoritmi

Osnovni Monte Carlo procjenitelj vjerojatnosti gore opisanog događaja bio bi $Z(x) = \mathbb{1}_{\{S_n > x\}}$, pa je uvjetni Monte Carlo procjenitelj

$$\mathbb{P}(S_n > x|\mathscr{F}), \quad \mathscr{F} \subset \sigma(X_1, \ldots, X_n).$$

Prisjetimo se da smo u (1.13) pokazali da je varijanca uvjetnog Monte Carlo procjenitelja uvijek manja od varijance osnovnog procjenitelja. Htjeli bismo pronaći $\mathscr{F}$ takvu

da je redukcija varijance značajna, odnosno da su algoritmi logaritamski efikasni ili imaju ograničenu relativnu pogrešku. Jedna ideja je, kao u primjeru 1.14, uzeti u obzir ishode varijabli $X_1, \ldots, X_{n-1}$, što daje

$$\tilde{Z}_1(x) = \mathbb{P}(S_n > x | X_1, \ldots, X_{n-1}) = \bar{F}(x - S_{n-1}).$$

Dakle, morali bismo generirati samo $X_1, \ldots, X_{n-1}$. Ipak varijanca je i dalje istog reda veličine kao prije - $\bar{F}(x)$. To se lako vidi jer

$$\mathbb{E}\tilde{Z}_1(x)^2 = \mathbb{E}[\bar{F}(x - S_{n-1})^2] \geq \mathbb{E}[\bar{F}(x - S_{n-1}); X_1 > x] = \mathbb{P}(X_1 > x) = \bar{F}(x),$$

gdje druga jednakost vrijedi jer su X_i pozitivne pa ako je $X_1 > x$, tada je $S_{n-1} > x$, a $\bar{F}(y) = 1$ za $y < 0$. Ovaj algoritam u praksi daje vrlo loše rezultate zbog toga što je prevelika vjerojatnost da neki od $X_1, \ldots, X_{n-1}$ bude prevelik. Assmusen i Binswanger [2] predložili su da se najveći od $X_1, \ldots, X_n$ odbaci i uvjetuje samo na preostale. Dakle, generirali bismo $X_1, \ldots, X_n$, sortirali ih kao

$$X_{(1)} < X_{(2)} < \cdots < X_{(n)},$$

izbacili $X_{(n)}$ te formirali procjenitelj kao

$$\tilde{Z}_2(x) = \mathbb{P}(S_n > x | X_{(1)}, X_{(2)}, \ldots, X_{(n-1)}).$$

Primijetimo li da vrijedi

$$\mathbb{P}(X_{(n)} > x | X_{(1)}, X_{(2)}, \ldots, X_{(n-1)}) = \frac{\bar{F}(X_{(n-1)} \vee x)}{\bar{F}(X_{(n-1)})},$$

i označimo li s $S_{(n-1)} = X_{(1)} + X_{(2)} + \cdots + X_{(n-1)}$, dobivamo

$$\begin{aligned} \mathbb{P}(S_n > x | X_{(1)}, X_{(2)}, \ldots, X_{(n-1)}) &= \mathbb{P}(X_{(n)} + S_{(n-1)} > x | X_{(1)}, X_{(2)}, \ldots, X_{(n-1)}) \\ &= \mathbb{P}(X_{(n)} > x - S_{(n-1)} | X_{(1)}, X_{(2)}, \ldots, X_{(n-1)}) \\ &= \frac{\bar{F}((x - S_{(n-1)}) \vee X_{(n-1)})}{\bar{F}(X_{(n-1)})}. \end{aligned} \tag{2.21}$$

Pokazuje se da je ovaj algoritam logaritamski efikasan u slučaju regularne varijacije i to je prvi primjer logaritamski efikasnog algoritma u slučaju teških repova. Ovdje to nećemo pokazivati (za dokaz vidi Asmussen i dr. [2]), već ćemo obraditi sličan algoritam, poznatiji kao Assmusen-Kroeseov algoritam za koji se pokazalo da daje bolje teorijske i praktične rezultate. Algoritam se bazira na sljedećem rezultatu: uz oznaku $M_n = \max\{X_1, \ldots, X_n\}$ zbog simetrije vrijedi

$$z(x) = \mathbb{P}(S_n > x) = \sum_{i=1}^{n} \mathbb{P}(S_n > x, M_n = X_i) = n\mathbb{P}(S_n > x, M_n = X_n).$$

Dakle, ideja je particionirati u skladu s najvećim $X_i =: M_n$ i zatim uvjetovati na sve $X_j, j \neq i$. Sada iz (2.21) zbog $M_{n-1} = X_{(n-1)}$ i $S_{(n-1)} = S_{n-1}$ u slučaju $M_n = X_n$ slijedi

$$\begin{aligned}\tilde{Z}_3(x) &:= n\mathbb{P}(S_n > x, M_n = X_n | X_1, \ldots, X_{n-1}) \\ &= n\mathbb{P}(S_n > x | M_n = X_n; X_1, \ldots, X_{n-1})\mathbb{P}(M_n = X_n | X_1, \ldots, X_{n-1}) \\ &= n\frac{\bar{F}((x - S_{n-1}) \vee M_{n-1})}{\bar{F}(M_{n-1})}\mathbb{P}(X_n > M_{n-1}) \\ &= n\bar{F}((x - S_{n-1}) \vee M_{n-1}).\end{aligned}$$

Teorem 2.15. *Procjenitelj $\tilde{Z}_3$ ima ograničenu relativnu pogrešku u slučaju regularne varijacije.*

Dokaz. Kako nas zanima vjerojatnost $\mathbb{P}(S_n > x)$ kad je x vrlo velik, razumno je promatrati slučaj $M_{n-1} \leq x/n$. Tada je $S_{n-1} \leq (n-1)x/n$, odnosno $x - S_{n-1} \geq x/n$ pa je $M_{n-1} \vee (x - S_{n-1}) \geq x/n$. Iz definicije regularne varijacije slijedi

$$\begin{aligned}\frac{\mathbb{E}\tilde{Z}_3(x)^2}{\bar{F}(x)^2} &\leq n^2\frac{\bar{F}(x/n)^2}{\bar{F}(x)^2} = n^2\frac{L(x/n)^2/(x/n)^{2\alpha}}{L(x)^2/x^{2\alpha}} \\ &= n^{2+2\alpha}\frac{L(x/n)^2}{L(x)^2} \sim n^{2+2\alpha}.\end{aligned}$$

Iskoristimo li sada $z(x) \sim n\bar{F}(x)$, dobivamo

$$\frac{\mathbb{E}\tilde{Z}_3(x)^2}{z(x)^2} \sim \frac{n^{2+2\alpha}}{n^2} \sim n^{2\alpha},$$

što je ograničeno za $x \to \infty$, pa algoritam $\tilde{Z}_3$ zaista ima ograničenu relativnu pogrešku. □

Može se pokazati i da je u slučaju Weibullove distribucije algoritam $\tilde{Z}_3$ logaritamski efikasan za $\beta < \bar{\beta} = \log(1.5)/\log(2) = 0.585$. Dokaz je nešto kompliciraniji i može se naći u Asmussen i dr. [5].

Algoritmi uzorkovanja po važnosti

I dalje se bavimo istim problemom procjene vjerojatnosti $z(x) = \mathbb{P}(S_n > x)$ za slučajnu šetnju s negativnim očekivanjem i velike x. Za početak ćemo razmotriti uzorkovanje po važnosti u kojem $X_1, \ldots, X_n$ uzorkujemo iz distribucije važnosti $\tilde{F}$. U svojem su radu Asmussen, Binswanger i Højgaard [3] razmatrali slučaj u kojem

$\tilde{F}$ ne ovisi o x. Pokazali su da je dobar izbor za distribuciju važnosti $\tilde{F}$ distribucija dosta težeg repa od F. Formalnije, označimo li s f i $\tilde{f}$ pripadne gustoće, odabrat ćemo $\tilde{f}$ tako da vrijedi:

Uvjet 2.16. *Funkcija K je subeksponencijalna distribucija s gustoćom $k(x) = c^{-1}f^2(x)/\tilde{f}(x)$, gdje je $c = \int_0^\infty f^2/\tilde{f}$ i zadovoljava*

$$\liminf_{x\to\infty} \frac{|\log \bar{K}(x)|}{|2\log \bar{F}(x)|} \geq 1.$$

Sada je procjenitelj za z

$$Z(x) = \prod_{i=1}^{n} \frac{f(X_i)}{\tilde{f}(X_i)} \mathbb{1}_{\{S_n > x\}}. \tag{2.22}$$

Pokazuje se da je taj algoritam logaritamski efikasan.

Propozicija 2.17. *Ako vrijedi uvjet 2.16, gornji algoritam je logaritamski efikasan u slučaju subeksponencijalnih distribucija.*

Dokaz. Vrijedi

$$\begin{aligned}
\tilde{\mathbb{E}}Z^2 &= \tilde{\mathbb{E}}\left[\left(\prod_{i=1}^{n} \frac{f(X_i)}{\tilde{f}(X_i)}\right)^2; S_n > x\right] \\
&= \int \cdots \int_{x_1+\cdots+x_n>x} \frac{f(x_1)^2}{\tilde{f}(x_1)^2} \cdots \frac{f(x_n)^2}{\tilde{f}(x_n)^2} \tilde{f}(x_1)\cdots\tilde{f}(x_n)dx_1\cdots dx_n \\
&= c^n \int \cdots \int_{x_1+\cdots+x_n>x} k(x_1)\cdots k(x_n)dx_1\cdots dx_n \\
&= c^n \mathbb{P}_k(S_n > x) \sim nc^n \bar{K}(x),
\end{aligned}$$

gdje aproksimacija u zadnjem koraku slijedi iz definicije subeksponencijalnih distribucija (2.20). Kako je također iz definicije subekponencijalnih distribucija $z \sim n\bar{F}(x)$ i c ne ovisi o x, iz uvjeta (2.16) slijedi da je

$$\liminf \frac{|\log \mathbb{E}Z^2(x)|}{|\log z(x)^2|} \geq 1,$$

odnosno algoritam $Z(x)$ je logaritamski efikasan. □

Pokazuje se da je jedan dobar primjer za distribuciju važnosti u slučaju subeksponencijalnih distribucija

$$\tilde{f}(x) = \frac{1}{(x+e)\log(x+e)^2}, \quad x \geq 0.$$

Primijetimo da je rep $\bar{\tilde{F}}(x) = 1/\log(x+e)$ vrlo težak (slika 2.2). U [3] je pokazano da u čestim slučajevima, poput Paretove i Weibullove distribucije, ova distribucija važnosti zadovoljava uvjet 2.16 pa stoga uzorkovanje po važnosti uz takav $\tilde{F}$ daje logaritamski efikasnu procjenu za $z(x)$. Već smo ranije razmatrali kako je na skupu $\{S_n > x\}$, tipično $X_{(n)} > x$, a ostali $X_{(i)}$, za $i \leq n-1$ su mali i neosjetljivi na x. Algoritam $Z(x)$ na ovom skupu jednak je umnošku $\prod_{i=1}^n f(X_i)/\tilde{f}(X_i)$. Izaberemo li $\tilde{f}$ s puno težim repom od početne distribucije f, omjer $f(X_{(n)})/\tilde{f}(X_{(n)})$ je malen i smanjuje se kako x raste, dok se za ostale $i \leq n-1$ omjer $f(X_{(i)})/\tilde{f}(X_{(i)})$ uglavnom ne mijenja. Stoga je $Z(x)$ za velike x mala i na skupu $\{S_n > x\}$ ima malu varijancu, što je upravo ono što želimo od algoritma.

Pokazuje se da iako ima lijepa teorijska svojstva, ovaj algoritam daje slabe rezultate u praksi. Juneja i Shahabuddin [9] predložili su efikasniji algoritam, u kojem distribucija važnosti $\tilde{f}$ ovisi o x. Ideja je da distribucija važnosti na $[x_0, \infty)$ ima rep $c_1\bar{F}(x)^{\theta(x)}$ za $\theta(x) \to 0$ kad $x \to \infty$, a na $(0, x_0)$ ima gustoću $c_2 f(x)$ uz neke uvjete na parametre c_1 i c_2. Ovdje ćemo pokazati ideju algoritma u Pareto slučaju, a općenita razmatranja mogu se naći u [9].

Primjer 2.18. *Neka je $f(x) = \alpha/(1+x)^{1+\alpha}$ i promatrajmo distribuciju važnosti $\tilde{f}(x) = \tilde{\alpha}/(1+x)^{1+\tilde{\alpha}}$, za $\tilde{\alpha} := \tilde{\alpha}(x) \to 0$ kad $x \to \infty$. Definiramo li standardno procjenitelj uzorkovanja po važnosti kao u (2.22), iz propozicije 2.17 slijedi da je njegov drugi moment $c^n\mathbb{P}_k(S_n > x)$, gdje je*

$$c = \int_0^\infty \frac{\alpha^2/(1+x)^{2\alpha+2}}{\tilde{\alpha}/(1+x)^{\tilde{\alpha}+1}}dx = \frac{\alpha^2}{\tilde{\alpha}(\tilde{\alpha}-2\alpha)} \sim \frac{\alpha}{2\tilde{\alpha}}, \quad i$$

$$k(x) = \frac{1}{c}\frac{f^2(x)}{\tilde{f}(x)} \sim 2\alpha(1+x)^{-(2\alpha-\tilde{\alpha})-1},$$

pa ćemo $\mathbb{P}_k$ označiti sa $\mathbb{P}_{2\alpha-\tilde{\alpha}}$. Ograničimo li sada $\mathbb{P}_{2\alpha-\tilde{\alpha}}(S_n > x)$ odozdo i odozgo s

$$\mathbb{P}_{2\alpha-\varepsilon}(S_n > x) \sim \frac{n}{x^{2\alpha-\varepsilon}} \quad i \quad \mathbb{P}_{2\alpha+\varepsilon}(S_n > x) \sim \frac{n}{x^{2\alpha+\varepsilon}},$$

i pustimo li $\varepsilon \to 0$, slično kao u dokazu propozicije 2.17 dobije se da je algoritam logaritamski efikasan, ako $\tilde{\alpha}(x)$ zadovoljava $\log\tilde{\alpha}(x)/\log x \to 0$ kad $x \to \infty$. Pitanje koje se postavlja je kako odabrati optimalan $\tilde{\alpha}$. Logično bi bilo uzeti onaj $\tilde{\alpha}(x)$ koji minimizira varijancu procjenitelja $Z(x)$, odnosno minimizira

$$\tilde{\mathbb{E}}Z^2 \sim c^n\mathbb{P}_{2\alpha-\tilde{\alpha}}(S_n > x) \sim \left(\frac{\alpha}{2\tilde{\alpha}}\right)^n \frac{1}{x^{2\alpha-\tilde{\alpha}}}.$$

Pokazuje se da je to $\tilde{\alpha}^*(x) = n/\log x$.

Spomenimo za kraj još jednu zanimljivu ideju koja se pokazuje najboljom u praksi. Sjetimo li se rezultata

$$\mathbb{P}(S_n > x) = n\mathbb{P}(S_n > x, M_n = X_n),$$

gdje je $M_n = \max\{X_1 \dots X_n\}$, ideja uzorkovanja po važnosti je sljedeća. Simuliramo $X_1, \dots, X_{n-1}$ iz gustoće f te njima nezavisnu X_n iz gustoće $\tilde{f}$. Tada je procjenitelj

$$Z(x) = n\frac{f(X_n)}{\tilde{f}(X_n)}\mathbb{1}_{\{S_n > x, M_n = X_n\}}.$$

Može se pokazati da je ovaj algoritam logaritamski efikasan za subeksponencijalne F (Asmussen i Kroese [5]).

Poglavlje 3

Simulacije

U ovom poglavlju proučit ćemo praktična svojstva algoritama opisanih u drugom poglavlju na nekim stvarnim primjerima.

Siegmundov algoritam

U drugom poglavlju bavili smo se diskretnom slučajnom šetnjom $(X_n, n \in \mathbb{N})$, $S_n = X_1 + \cdots + X_n$, gdje su $X_1, \ldots, X_n$ nezavisne jednako distribuirane slučajne varijable s distribucijom F i negativnim očekivanjem $\mathbb{E}X_i < 0$. Podsjetimo se, Siegmundov algoritam rješava problem procjene vjerojatnosti

$$z(x) := \mathbb{P}(\tau(x) < \infty), \quad \text{gdje je} \quad \tau(x) = \inf\{n : S_n > x\}$$

u slučaju da je x velik, odnosno kad je ta vjerojatnost vrlo mala. Za Siegmundov algoritam pronašli smo optimalnu distribuciju važnosti uz koju je on logaritamski efikasan (Teorem 2.7).
Razmatranja u drugom poglavlju vode do sljedećeg simulacijskog postupka.

1. Odredi rješenje $\gamma > 0$ od $\hat{F}[\gamma] = 1$, odnosno $\kappa(\gamma) = 0$.
2. Postavi $S_0 = 0$. Sve dok je $S_n < x$ stavi

   $$S_{n+1} = S_n + X_{n+1}, \quad n = 0, 1, 2, \ldots$$

   gdje su $X_1, X_2, \ldots$ generirane iz distribucije važnosti F_γ. Neka je $\tau(x)$ najmanji n za koji je $S_n \geq x$.
3. Stavi $Z = e^{-\gamma S_{\tau(x)}}$.
4. Ponovi korake 2 i 3 N puta te procijeni $z(x)$ kao prosjek nezavisnih replikacija $Z_1, \ldots, Z_N$.

Za numeričku simulaciju prvo ćemo promatrati slučajnu šetnju u kojoj su skokovi normalno distribuirani s negativnim očekivanjem $-\mu$ i varijancom 1, odnosno $X_i \sim N(-\mu, 1), \mu > 0$. U primjeru 2.6 pokazali smo da je optimalna distribucija važnosti u ovom slučaju upravo $N(\mu, 1)$. Algoritam je implementiran u MATLAB-u, a simulacije su izvedene za dvije različite vrijednosti parametra $\mu = -0.5, -1$ te za sedam različitih vrijednosti x. Rezultati su dani u tablici 3.1, gdje su oznake redom $z(x)$ - procijenjena vjerojatnost, 95% CI - širina polovice 95% pouzdanog intervala i RE - relativna pogreška.

μ	x	$z(x)$	95% CI	RE
-0.5	3	2.7863×10^{-2}	7.7947×10^{-5}	0.0014273
-0.5	5	3.7682×10^{-3}	1.0483×10^{-5}	0.0014194
-0.5	8	1.8802×10^{-4}	5.2278×10^{-7}	0.0014186
-0.5	15	1.7133×10^{-7}	4.7933×10^{-10}	0.0014274
-0.5	20	1.1558×10^{-9}	3.2078×10^{-12}	0.001416
-0.5	30	5.2439×10^{-14}	1.4623×10^{-16}	0.0014227
-1	3	7.9491×10^{-4}	4.3046×10^{-6}	0.0027629
-1	5	1.4442×10^{-6}	7.8700×10^{-8}	0.0027804
-1	8	3.5924×10^{-8}	1.9481×10^{-10}	0.0027668
-1	15	2.9927×10^{-14}	1.6248×10^{-16}	0.002770
-1	20	1.3502×10^{-18}	7.3452×10^{-21}	0.0027755
-1	30	2.8057×10^{-27}	1.5217×10^{-29}	0.0027672

Tablica 3.1: Siegmundov algoritam

Neovisno o veličini x-a algoritam se čini vrlo preciznim, pouzdani intervali su uski (širina je uvijek nekoliko redova veličine manja od procjene) čak i za iznimno male vjerojatnosti reda 10^{-27}. Primjećujemo razliku u relativnoj pogrešci za različite parametre μ, no zanimljivo je da za isti parametar relativna pogreška ne raste sa x, već pokazuje vrlo malu varijabilnost.

μ	x	$z(x)$	95% CI	RE
-0.5	3	2.7820×10^{-2}	1.0193×10^{-3}	0.01869
-0.5	5	3.9000×10^{-3}	3.8632×10^{-4}	0.05054
-0.5	8	1.4000×10^{-4}	7.3332×10^{-5}	0.26724

Tablica 3.2: Osnovni Monte Carlo algoritam

Za usporedbu, u tablici 3.2 dani su rezultati osnovnog Monte Carlo algoritma za neke kombinacije parametara μ i x. S obzirom da su sve simulacije rađene uz 10^5 replikacija, vjerojatnosti manje od 10^{-5} osnovni Monte Carlo algoritam u prosjeku bi procjenio s 0 pa one nisu prikazane. Primjećujemo da su i Siegmundov i Monte Carlo algoritam dali sličnu procjenu za parametre $\mu = -0.5$, $x = 3$ kad je vjerojatnost reda 10^{-2}, no Monte Carlo procjenitelj ima oko deset puta veću relativnu pogrešku i puno širi 95% pouzdani interval. Također, vidimo da se, kako se vjerojatnost koju treba procijeniti smanjuje, procjene ova dva algoritama sve više razlikuju, a relativna pogreška Monte Carlo procjenitelja raste do čak 26% u slučaju vjerojatnosti reda 10^{-4}.

Prisjetimo se sada proširenog modela rizika osiguravajućeg društva opisanog u primjeru 2.9. Pretpostavit ćemo da su veličine zahtjeva za isplatom X_i nezavisne eksponencijalno distribuirane slučajne varijable s parametrom ν, te da su vremena čekanja između zahtjeva W_i nezavisne jednako distribuirane varijable s eksponencijalnom distribucijom s parametrom λ. Također pretpostavljamo konstantnu stopu premije c i početnu imovinu osiguravajućeg društva u. Podsjetimo se, problem koji smo rješavali bio je procijeniti vjerojatnost propasti osiguravajućeg društva, odnosno

$$\mathbb{P}(\tau(u) < \infty), \quad \text{za} \quad \tau(u) = \inf\{n : S_n > u\},$$

gdje je $S_n = \sum_{i=1}^{n} Z_i$, za $Z_i = X_i - cW_i$. Može se pokazati (vidi Mikosch [10]) da je u opisanom slučaju eksponencijalno distribuiranih zahtjeva ova vjerojatnost egzaktno dana s

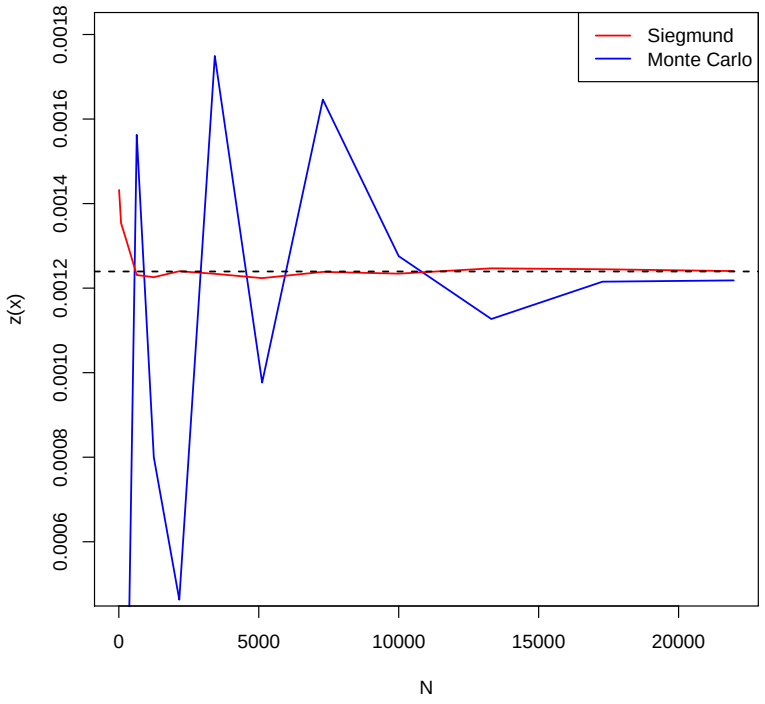
$$z(x) = qe^{-\gamma x}, \quad \text{za} \quad q = \frac{\lambda}{c\nu}. \tag{3.1}$$

U primjeru 2.9 odredili smo optimalnu distribuciju važnosti iz Siegmundovog algoritma - to je distribucija slučajne varijable $Z_i = cW_i - X_i$, za $X_i \sim \text{Exp}(\nu)$ i $W_i \sim \text{Exp}(\lambda)$. Sada možemo usporediti procjenu Siegmundovog algoritma $\hat{z}(x)$ s egzaktnom vrijednošću $z(x)$. U tablici 3.3 dani su rezultati simulacija za različite parametre λ, ν i u, dok je stopa premije c postavljena na 2. Čini se da točnost Siegmundovog algoritma u ovom slučaju dosta ovisi o parametrima modela. Za $\lambda = 1$ procjena je vrlo dobra i egzaktna vrijednost je uvijek unutar 95% pouzdanog intervala, no za $\lambda = 0.8$ algoritam ne daje zadovoljavajuću procjenu. Primijetimo da u tim slučajevima egzaktna vrijednost nije u 95% pouzdanom intervalu, koji je relativno uzak, odnosno Siegmundov algoritam je presiguran.

Na slici 3.1 crvenom bojom prikazano je kretanje procjene Siegmundovog algoritma za $z(x)$ ovisno o broju simulacija N, a plavom bojom kretanje procjene osnovnog Monte Carlo algoritma. Vjerojatnost koju smo u ovom slučaju procjenjivali odgovara parametrima $\nu = 1$, $\lambda = 1$ i $u = 12$ i označena je crnom iscrtkanom linijom.

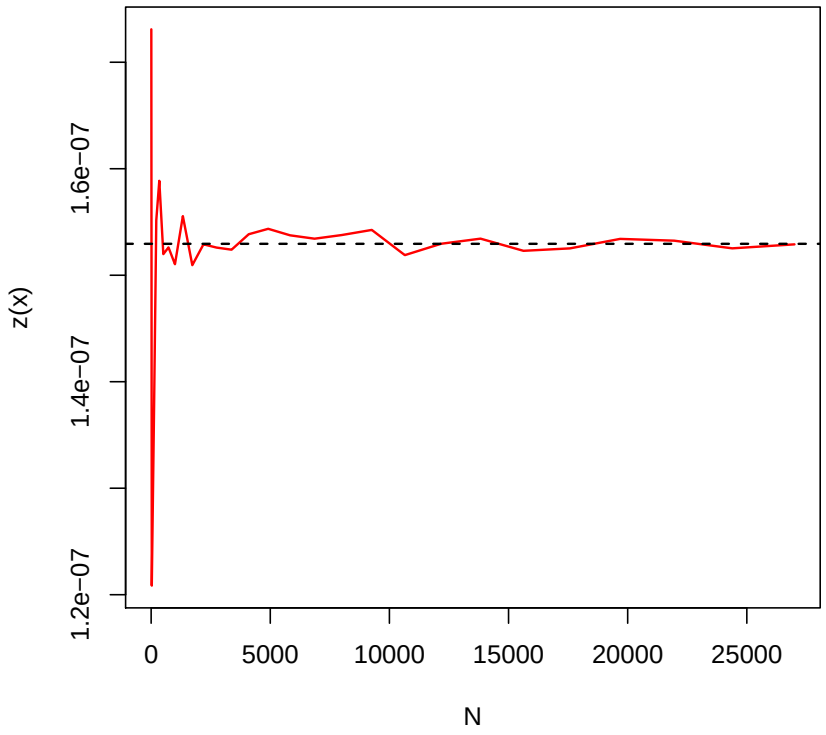
ν	λ	u	$z(x)$	$\hat{z}(x)$	95% CI	RE
1	1	15	2.7654×10^{-4}	2.7752×10^{-4}	4.4425×10^{-6}	0.008167
1	1	25	1.8633×10^{-6}	1.8710×10^{-6}	3.0071×10^{-8}	0.008199
1	1	50	6.9440×10^{-12}	7.0691×10^{-12}	1.1019×10^{-13}	0.007953
1	0.8	15	4.9364×10^{-5}	6.2389×10^{-5}	9.8497×10^{-7}	0.008055
1	0.8	25	1.2236×10^{-7}	1.5656×10^{-7}	2.4090×10^{-9}	0.007851
1	0.8	50	3.7430×10^{-14}	4.7553×10^{-14}	7.4641×10^{-16}	0.008009
0.6	1	65	1.2529×10^{-3}	1.2595×10^{-3}	5.7902×10^{-6}	0.002346
0.6	1	100	3.7833×10^{-5}	3.7830×10^{-5}	1.7846×10^{-7}	0.002407
0.6	0.8	65	1.5069×10^{-6}	1.7167×10^{-6}	1.1869×10^{-8}	0.003528
0.6	0.8	100	1.3741×10^{-9}	1.5620×10^{-9}	1.0801×10^{-11}	0.003528

Tablica 3.3: Siegmundov algoritam, vjerojatnost propasti



Slika 3.1: Graf procjene $z(x)$ u ovisnosti o broju simulacija N

Vidimo da Siegmundov algoritam puno brže kovergira od Monte Carlo algoritma, koji tek nakon 20000 replikacija daje približno točnu procjenu. S obzirom na veliku skalu grafa na slici 3.1, konvergenciju Siegmundovog algoritma možemo bolje promotriti na slici 3.2. Primjećujemo da algoritam daje vrlo dobru procjenu već za oko 2000 replikacija.



Slika 3.2: Graf procjene $z(x)$ u ovisnosti o broju simulacija N

Složeni Poissonov proces

U nastavku ćemo promotriti problem procjene vjerojatnosti

$$z(x) = \mathbb{P}(C \geq x) = \mathbb{P}(X_1 + \cdots + X_N \geq x),$$

gdje su $X_1, X_2, \ldots$ nezavisne jednako distribuirane slučajne varijable s distribucijom F, i $N \sim \text{Poi}(\lambda)$ nezavisna od X_i. Prisjetimo se, u drugom poglavlju razmatrali smo ovaj problem u kontekstu osiguranja - varijable X_i predstavljale su zahtjeve za isplatom, N broj isplata u npr. godini dana, a $z(x)$ vjerojatnost da je broj traženih isplata u godini dana iznimno velik. Uz određene uvjete iznijeli smo logaritamski efikasan algoritam za procjenu vjerojatnosti $z(x)$ (vidi Teorem 2.5). Konkretno, razmatranja u poglavlju 2 vode nas do sljedećeg simulacijskog postupka.

1. Odredi rješenje θ^* od $\lambda \hat{F}'[\theta] = x$.
2. Simuliraj $N \sim \text{Poi}(\lambda \hat{F}[\theta^*])$.
3. Generiraj N nezavisnih jednako distribuiranih slučajnih varijabli $X_1, \ldots, X_N$ s funkcijom izvodnicom momenata

$$\hat{F}_{\theta^*}[t] = \frac{\hat{F}[t + \theta^*]}{\hat{F}[\theta^*]}.$$

   Postavi $C = X_1 + \cdots + X_N$.
4. Generiraj M nezavisnih replikacija od C i procijeni traženu vjerojatnost kao

$$\hat{z(x)} = \frac{1}{M} \sum_{k=1}^{M} \mathbb{1}_{\{C^{(k)} \geq x\}} e^{-\theta^* C^{(k)} + \lambda(\hat{F}[\theta^*] - 1)}$$

Modelirajmo sada na ovaj način konkretnu situaciju u osiguravajućem društvu. Pretpostavimo da osiguravajuće društvo s kapitalom od 10^9 kuna zaprimi u prosjeku λ zahtjeva za isplatom osiguranja godišnje. Pretpostavit ćemo da broj zahtjeva za isplatom u jednoj godini N ima Poissonovu distribuciju s parametrom λ, a veličine zahtjeva $X_1, \ldots, X_N$, mjerene u milijunima kuna, dolaze iz Gama$(2, 1)$ distribucije (s gustoćom $f(x) = xe^{-x}$, $x \geq 0$). Zanima nas vjerojatnost da ukupna vrijednost zahtjeva u jednoj godini bude veća od kapitala s kojim osiguravajuće društvo raspolaže, odnosno

$$z(x) = \mathbb{P}(C \geq x) = \mathbb{P}(X_1 + \cdots + X_N \geq 10^3).$$

Da bismo proveli simulacijski postupak, prvo trebamo izračunati funkciju izvodnicu momenata slučajne varijable X s Gama$(2, 1)$ distribucijom. Ona iznosi:

$$\begin{aligned}\hat{F}[\theta] = \mathbb{E}\left[e^{\theta X}\right] &= \int_0^\infty x e^{\theta x - x} dx \\ &= \frac{1}{\theta - 1} x e^{x(\theta-1)} \Big|_0^\infty - \frac{1}{\theta - 1} \int_0^\infty e^{x(\theta-1)} dx = [\text{za} \quad \theta < 1] \\ &= -\frac{1}{(1-\theta)^2} e^{x(\theta-1)} \Big|_0^\infty \\ &= \frac{1}{(1-\theta)^2}.\end{aligned}$$

Sada je rješenje θ^* od $\lambda \hat{F}'[\theta] = x$, $\theta^* = 1 - (2\lambda/x)^{1/3}$, a parametar Poissonove distribucije važnosti $\lambda_{\theta^*} = \lambda/(1-\theta^*)^2 = (\lambda x^2/4)^{1/3}$. Također, slijedi da je

$$\hat{F}_{\theta^*}[t] = \left(\frac{1-\theta^*}{1-\theta^*-t}\right)^2,$$

a to je upravo funkcija izvodnica momenata Gama$(2, 1-\theta^*)$ distribucije. Dakle, broj zahtjeva N simulirat ćemo iz Poi$((\lambda x^2/4)^{1/3})$ distribucije, a veličinu zahtjeva $X_1, \ldots, X_N$ iz Gama$(2, (2\lambda/x)^{1/3})$ distribucije.

λ	$z(x)$	95% CI	RE
200	1.6416×10^{-43}	1.4168×10^{-44}	0.044035
300	3.0335×10^{-17}	1.9843×10^{-18}	0.033375
400	5.0031×10^{-5}	2.1174×10^{-6}	0.021593

Tablica 3.4: Složeni Poissonov proces

U tablici 3.4 dani su rezultati algoritma za tri različita parametra λ. Primijetimo da je čak i u slučaju procjene vjerojatnosti reda 10^{-43} relativna pogreška samo 4%.

Algoritmi za distribucije teških repova

U slučaju distribucija teškog repa u drugom poglavlju bavili smo se samo sljedećim jednostavnim problemom. Neka su $X_1, X_2, \ldots$ nezavisne jednako distribuirane i nenegativne slučajne varijable sa subeksponencijalnom funkcijom distribucije F. Promatramo problem procjene vjerojatnosti $z(x) = \mathbb{P}(S_n > x)$ za velike x, gdje je $S_n = X_1 + \cdots + X_n$.
Detaljnije smo obradili dva algoritma - jedan koji se temelji na uvjetnoj Monte Carlo metodi i jedan algoritam uzorkovanja po važnosti.
Podsjetimo se, prvi algoritam bio je

$$\tilde{Z}_3(x) := n\bar{F}(M_{n-1} \vee (x - S_{n-1})),$$

gdje smo s M_{n-1} označili $\max\{X_1, \ldots, X_{n-1}\}$, $\bar{F}(x) = 1 - F(x)$ i $x \vee y$ označava $\max\{x, y\}$. Pokazali smo da ovaj procjenitelj ima ograničenu relativnu pogrešku u slučaju regularne varijacije, te da je logaritamski efikasan u slučaju Weibullove distribucije s parametrom $\beta < 0.585$. Simulacijski postupak je sljedeći:

1. Generiraj $X_1, \ldots, X_n$ iz distribucije F.
2. Izračunaj
$$Z = n\bar{F}\left(\left(x - \sum_{j=1}^{n-1} X_j\right) \vee \max_{j \neq n} X_j \right)$$
3. Ponovi korake 1 i 2 N puta te procijeni traženu vjerojatnost kao prosjek N nezavisnih replikacija od Z

$$z(x) = \frac{1}{N} \sum_{k=1}^{N} Z_k$$

Drugi algoritam temelji se na uzorkovanju po važnosti uz neku distribuciju važnosti $\tilde{f}$. Procjenitelj je tada dan s

$$Z(x) = \prod_{i=1}^{n} \frac{f(X_i)}{\tilde{f}(X_i)} \mathbb{1}_{\{S_n > x\}}.$$

Kao primjer dobre distribucije važnosti za subeksponencijalne distribucije naveli smo distribuciju s gustoćom

$$\tilde{f}(x) = \frac{1}{(x+e)\log(x+e)^2}, \quad x \geq 0, \tag{3.2}$$

odnosno funkcijom distribucije $\tilde{F}(x) = 1 - \log(x + e)$ pa ćemo nju koristiti u simulacijama. Simulacijski postupak za ovaj algoritam je sljedeći:

1. Generiraj $X_1, \ldots, X_n$ iz unaprijed određene distribucije $\tilde{F}$.
2. Izračunaj
$$Z = \prod_{i=1}^{n} \frac{f(X_i)}{\tilde{f}(X_i)} \mathbb{1}_{\{S_n > x\}}.$$
3. Ponovi korake 1 i 2 N puta i procijeni traženu vjerojatnost kao prosjek N nezavisnih replikacija od Z
$$z(x) = \frac{1}{N} \sum_{k=1}^{N} Z_k$$

n	x	$z(x)$	95% CI	RE
6	10^3	6.1746×10^{-3}	3.4653×10^{-6}	2.8634×10^{-4}
6	10^6	6.0004×10^{-6}	1.2405×10^{-10}	1.0547×10^{-5}
6	10^9	6.0000×10^{-9}	1.1112×10^{-16}	9.4490×10^{-9}
10	10^3	1.0536×10^{-2}	7.9316×10^{-6}	3.8410×10^{-4}
10	10^6	1.0001×10^{-5}	3.3125×10^{-10}	1.6898×10^{-5}
10	10^9	1.0000×10^{-8}	5.6120×10^{-16}	2.2863×10^{-8}

Tablica 3.5: Pareto ($\alpha = 1$), Algoritam 1

n	x	$z(x)$	95% CI	RE
6	10^3	6.6129×10^{-3}	2.6473×10^{-4}	2.0425×10^{-2}
6	10^6	6.8234×10^{-6}	5.3779×10^{-7}	4.0212×10^{-2}
6	10^9	5.9429×10^{-9}	6.3588×10^{-10}	5.4591×10^{-2}
10	10^3	1.2156×10^{-2}	1.1723×10^{-3}	4.9200×10^{-2}
10	10^6	1.0115×10^{-5}	1.4519×10^{-6}	7.3235×10^{-2}
10	10^9	1.0120×10^{-8}	8.3126×10^{-9}	1.0863×10^{-1}

Tablica 3.6: Pareto ($\alpha = 1$), Algoritam 2

U tablicama 3.5 i 3.6 dani su rezultati simulacija oba algoritma za dvije vrijednosti n i tri vrijednosti x u slučaju Paretove distribucije s parametrom $\alpha = 1$. Sve simulacije rađene su uz 10^5 replikacija. U stupcima se nalaze redom procjena vjerojatnosti $z(x) = \mathbb{P}(S_n > x)$, širina polovice 95% pouzdanog intervala za $z(x)$ (95% CI) i relativna pogreška procjenitelja (RE).

Primjećujemo da su procjene algoritama donekle slične, no prvi algoritam daje puno uže pouzdane intervale i ima (ovisno o parametrima) za nekoliko redova veličine manju relativnu pogrešku. Zanimljivo je da što je događaj rjeđi, procjena drugog algoritma postaje sve nepreciznija, a u slučaju $n = 10$ i $x = 10^9$ dolazimo do relativne pogreške od čak 10%. Iako smo pokazali da je taj algoritam logaritamski efikasan, čini se da u praksi nije pretjerano koristan. S druge strane, prvi algoritam očito se pokazao vrlo uspješnim - širina pouzdanog intervala za nekoliko je redova veličine manja od procjene vjerojatnosti i relativna pogreška je vrlo mala. Zanimljivo je da se što je događaj rjeđi, taj algoritam doima sigurnijim, odnosno relativna pogreška je manja.

Pogledajmo sada rezultate simulacija u slučaju Weibullove distribucije s parametrom $\alpha = 0.5$. Drugi algoritam isprobat ćemo s dvije različite distribucije važnosti: $\tilde{f}$ iz 3.2 (Tablica 3.8) kao u prošlom primjeru i Paretovom distribucijom s parametrom $\alpha = 1$ (Tablica 3.9).

I u ovom slučaju dobivamo slične rezultate kao kod Paretove distribucije. Drugi algoritam, bez obzira na odabranu distribuciju važnosti, opet ima vrlo veliku relativnu pogrešku, ponekad čak 26%. S obzirom da su pouzdani intervali uži i relativna pogreška nešto manja ako za distribuciju važnosti odaberemo Paretovu distribuciju, možda bismo mogli zaključiti da je u ovom slučaju Paretova distribucija nešto pogodniji izbor od gore opisane $\tilde{f}$, no i dalje su rezultati algoritma loši i ne bi se trebali pouzdati u njegovu procjenu. Prvi algoritam daje zadovoljavajuće rezultate u pogledu pouzdanih intervala i relativne pogreške, no možemo primijetiti da su rezultati bili precizniji u Paretovom slučaju, što odgovara teorijskim svojstvima opisanima u drugom poglavlju.

n	x	$z(x)$	95% CI	RE
6	100	5.6026×10^{-4}	4.2476×10^{-6}	0.00387
6	400	1.6700×10^{-8}	5.6814×10^{-11}	0.00174
6	900	6.7660×10^{-13}	1.0712×10^{-15}	0.00081
10	100	1.6874×10^{-3}	1.6747×10^{-5}	0.00506
10	400	3.5922×10^{-8}	2.7794×10^{-10}	0.00395
10	900	1.3080×10^{-12}	2.9902×10^{-15}	0.00117

Tablica 3.7: Weibull ($\beta = 0.5$), Algoritam 1

n	x	$z(x)$	95% CI	RE
6	100	5.6495×10^{-4}	8.5495×10^{-5}	0.07721
6	400	1.4497×10^{-8}	3.8868×10^{-9}	0.13679
6	900	4.9183×10^{-13}	1.6280×10^{-13}	0.16888
10	100	2.2605×10^{-3}	9.9026×10^{-4}	0.22350
10	400	2.5533×10^{-8}	8.9489×10^{-9}	0.17882
10	900	1.3435×10^{-12}	7.0854×10^{-13}	0.26908

Tablica 3.8: Weibull ($\beta = 0.5$), Algoritam 2 ($\tilde{f}$)

n	x	$z(x)$	95% CI	RE
6	100	5.4773×10^{-4}	5.3422×10^{-5}	0.04976
6	400	2.2490×10^{-8}	9.3980×10^{-9}	0.21320
6	900	5.086×10^{-12}	1.5733×10^{-13}	0.15781
10	100	1.6665×10^{-3}	2.4679×10^{-4}	0.07556
10	400	2.6652×10^{-8}	8.4347×10^{-9}	0.16147
10	900	9.5671×10^{-13}	3.0688×10^{-13}	0.16366

Tablica 3.9: Weibull ($\beta = 0.5$), Algoritam 2 (Pareto ($\alpha = 1$))

Bibliografija

[1] Søren Asmussen, *Ruin probabilities*, Advanced series on statistical science & applied probability, World Scientific, Singapore, River Edge (N.J.), 2000, Autre tirage : 2001.

[2] Søren Asmussen i Klemens Binswanger, *Simulation of Ruin Probabilities for Subexponential Claims*, ASTIN Bulletin **27** (1997), 297–318, ISSN 1783-1350.

[3] Søren Asmussen, Klemens Binswanger i Bjarne Højgaard, *Rare events simulation for heavy-tailed distributions*, Bernoulli **6** (2000), br. 2, 303–322.

[4] Søren Asmussen i Peter W. Glynn, *Stochastic simulation : algorithms and analysis*, Stochastic modelling and applied probability, Springer, New York, 2007.

[5] Søren Asmussen i Dirk P. Kroese, *Improved algorithms for rare event simulation with heavy tails*, Adv. in Appl. Probab. **38** (2006), br. 2, 545–558.

[6] James Antonio Bucklew, *Large deviation techniques in decision, simulation, and estimation*, Wiley series in probability and mathematical statistics. Applied probability and statistics, J. Wiley, New York, Chichester, 1990, A Wiley-Interscience publication.

[7] Gerald B. Folland, *Real analysis*, second., Pure and Applied Mathematics (New York), John Wiley & Sons, Inc., New York, 1999, Modern techniques and their applications, A Wiley-Interscience Publication. MR 1681462 (2000c:00001)

[8] Jens Ledet Jensen, *Saddlepoint Approximations*, Oxford statistical science series 16, Oxford University Press, 1995.

[9] Sandeep Juneja i Perwez Shahabuddin, *Simulating heavy tailed processes using delayed hazard rate twisting*, ACM Trans. Model. Comput. Simul. **12** (2002), br. 2, 94–118.

[10] T. Mikosch, *Non-Life Insurance Mathematics: An Introduction with the Poisson Process*, Universitext, Springer Berlin Heidelberg, 2009.

[11] Hermann Thorisson, *Coupling, stationarity, and regeneration*, Probability and its applications, Springer, New York, 2000.

[12] David Williams, *Probability with Martingales.*, Cambridge mathematical textbooks, Cambridge University Press, 1991.

Sažetak

Standardnim Monte Carlo algoritmom teško je precizno procijeniti vjerojatnosti iznimno rijetkih događaja zbog velike relativne pogreške takvih procjenitelja. Cilj ovog diplomskog rada bio je prikazati neke od simulacijskih tehnika kojima se taj problem može zaobići. Naglasak je stavljen na uzorkovanje po važnosti i uvjetne Monte Carlo metode kao tehnike redukcije varijance procjenitelja. Precizno su definirane mjere efikasnosti algoritama poput ograničenosti relativne pogreške i logaritamske efikasnosti. Opisano je nekoliko algoritama koji se uglavnom bave procjenom vjerojatnosti prelaska praga u slučajnim šetnjama u slučajevima distribucija skokova lakog i teškog repa te su dokazana njihova teorijska svojstva efikasnosti. Algoritmi u slučaju distribucija lakog repa temelje se na eksponencijalnoj promjeni mjere u uzorkovanju po važnosti, a posebno je obrađen Siegmundov algoritam za koji je nađena optimalna distribucija važnosti. Također je spomenuta poveznica s teorijom velikih devijacija. Algoritmi u slučaju teških repova znatno su manje razvijeni od onih u slučaju lakih repova, te je za njih često teško pronaći dobru distribuciju važnosti. U ovom radu obrađeni su neki primjeri algoritama u slučaju subeksponencijalnih distribucija (s naglaskom na Paretovu i Weibullovu distribuciju) koji se temelje na uvjetnoj Monte Carlo metodi ili uzorkovanju po važnosti. U radu se navodi više stvarnih primjera primjene ovih algoritama, poput vjerojatnosti propasti u osiguravajućim društvima. Na kraju su algoritmi implementirani u MATLAB-u te uspoređeni u simulacijama. Siegmundov algoritam pokazao se dobrim i u praksi, dok su za slučaj teških repova bolje rezultate dali algoritmi uvjetne Monte Carlo metode nego uzorkovanja po važnosti.

Summary

The probability of exceptionally rare events is difficult to estimate accurately using the standard Monte Carlo algorithm due to a high relative error of such estimators. The aim of this master's thesis is to present some simulation techniques which may overcome this problem, with the emphasis on the importance sampling and conditional Monte Carlo method as variance reduction techniques. Algorithm efficiency measures, such as the bounded relative error and logarithmic efficiency, are precisely defined. Several algorithms are described, mainly concerning hitting times probabilities in light and heavy tailed random walks, and their theoretical efficiency proven. For light tailed distributions, the algorithms are based on the exponential change of measure in the importance sampling. The Siegmund's algorithm is presented, for which the optimal importance distribution has been determined. A brief explanation of large deviations approach to optimal exponential change of measure is also provided. Algorithms for heavy tailed distributions are substantially less developed and it is often difficult to find a good importance distribution. This paper presents some algorithms in the case of subexponential distributions (with the emphasis on Pareto and Weibull distribution), which are based on the conditional Monte Carlo method and importance sampling. Several examples of real life application of the algorithms are discussed, such as ruin probabilities in insurance. Finally, the algorithms have been implemented in MATLAB and compared in simulations. The Siegmund's algorithm has also proven good in practice, whereas in heavy tailed case the conditional Monte Carlo algorithms have shown better performance than the importance sampling ones.

Životopis

Rođena sam 1. ožujka 1993. godine u Zagrebu. Nakon završene osnovne škole Malešnica upisala sam matematičko-informatički smjer u XV. gimnaziji u Zagrebu. 2011. godine upisala sam preddiplomski sveučilišni studij Matematika na Prirodoslovno-matematičkom fakultetu u Zagrebu, a 2014. na istom fakultetu nastavila sam diplomski studij Statistika. Za vrijeme studija aktivno sam sudjelovala u radu udruge Mladi nadareni matematičari "Marin Getaldić", čija sam dopredsjednica od 2014. godine. Dobitnica sam dvije nagrade vijeća Matematičkog odsjeka za izniman uspjeh u studiju i izvannastavna postignuća.