6. Sovellus simulointiin

Olkoot prosessit $\{Y_n\}$, $\{\underline{X}_n\}$ ja funktiot Λ_n ja Λ kuten aiemminkin. Oletetaan tässä kappaleessa, että $d = 1$.

Olkoon A Borel-joukko, jolle

(6.1) $\quad \lim n^{-1} \log \mathbb{P}(\underline{X}_n \in A) \doteq -\gamma < 0.$

Tarkastelun kohteena tulee olemaan todennäköisyys

$$\alpha = \alpha_{n_0} = \mathbb{P}(\underline{X}_{n_0} \in A),$$

missä n_0 on kiinteä, mutta suuri. Tällöin $\mathbb{P}(\underline{X}_{n_0} \in A)$ tulee olemaan pieni; se lähestyy nollaa eksponentiaalisesti vauhdilla, kun $n_0 \to \infty$.

Sovelluksissa on tarvetta löytää tarkka approksimaatio α:lle. Eräs vaihtoehto on simulointi. Tällöin tuotetaan tietokoneella riippumattomia 'havaintoja' $\underline{X}_{n_0}$:n jakaumasta. Olkoot nämä

$$\underline{X}_{n_0}^{(1)}, \ldots, \underline{X}_{n_0}^{(N)}.$$

Estimoidaan α otoksesta asettamalla

$$\hat{\alpha} = N^{-1} \sum_{i=1}^{N} 1(\underline{X}_{n_0}^{(i)} \in A).$$

Selvästi

$$\mathbb{E}(1(\underline{X}_{n_0}^{(i)} \in A)) = \alpha, \quad \mathrm{Var}(1(\underline{X}_{n_0}^{(i)} \in A)) = \alpha(1-\alpha).$$

Nähdään, että estimaattorin $\hat{\alpha}$ odotusarvo ja varianssi ovat

$$\mathbb{E}(\hat{\alpha}) = \alpha, \quad \mathrm{Var}\,\hat{\alpha} = \frac{\alpha(1-\alpha)}{N}$$

Otoskoko N valitaan siten, että suhteellinen hajonta

$$\frac{\sqrt{\mathrm{Var}\,\hat{\alpha}}}{\mathbb{E}(\hat{\alpha})} = \sqrt{\frac{1-\alpha}{\alpha N}} \leq p_0,$$

missä p_0 on kiinnitetty (ja säätelee estimaatin suhteellista virhettä). Vaikki olla esim. $p_0 = 0.1$. Nähdään, että sopiva otoskoko on

$$(4.3) \qquad N = \frac{1-\alpha}{\alpha p_0^2} .$$

Kerta luku on siis vakio $\cdot\, e^{\mathcal{H} n_0}$, missä $-\mathcal{H}$ on (4.1):n raja-arvo. Ongelmaksi muodostuu suuri otoskoon tarve, jos n_0 on suuri.

Ongelmaa voidaan lähestyä esimerkiksi kohdennetun simuloinnin avulla. Olkoon $t_0 \in \mathcal{D}(\Lambda)$ ja $\mathbb{P}_{t_0}$ Y_n:n konjugaattijakauma (tai mitta, kts. kohta 3.2),

$$\mathbb{P}_{t_0}(Y_n \in B) = \mathbb{E}\left(e^{t_0 Y_n - n\Lambda_n(t_0)}\, 1(Y_n \in B)\right),$$

jolloin

$$(4.2) \quad \mathbb{P}(Y_n \in B) = \mathbb{E}_{t_0}\left(e^{-t_0 Y_n + n\Lambda_n(t_0)}\, 1(Y_n \in B)\right), \quad B \in \mathcal{B}.$$

Ajatuksena on tuottaa riippumattomia havaintoja muuttujasta (käyttäen Y_n:lle jakaumaa $\mathbb{P}_{t_0}$)

$$e^{-t_0 Y_n + n\Lambda_n(t_0)}\, 1(Z_n \in A).$$

Olkoot $Z_1, \dots, Z_N$ tällaisia. Estimoidaan α:aa asettamalla

$$\hat{\alpha}(t_0) = N^{-1} \sum_{i=1}^{N} Z_i .$$

Yhtälön (6.2) nojalla

$$\mathbb{E}_{t_0}(\hat{\alpha}(t_0)) = \alpha \quad \text{ja} \quad \mathrm{Var}_{t_0} \hat{\alpha}(t_0) = \frac{\eta_n(t_0) - \alpha^2}{N},$$

missä

$$\eta_n(t_0) = \mathbb{E}_{t_0}\left(e^{-2t_0 Y_n + 2n\Lambda_n(t_0)} 1(\bar{X}_n \in A)\right).$$

Luonteva kontrolli estimointivaiheelle on nytkin vaatia, että

$$\frac{\sqrt{\mathrm{Var}_{t_0} \hat{\alpha}(t_0)}}{\mathbb{E}_{t_0}(\hat{\alpha})} = \sqrt{\frac{\eta_n(t_0) - \alpha^2}{\alpha^2 N}} \leq p_0.$$

Tarvittava otoskoko on

$$N(t_0) = \frac{\eta_n(t_0) - \alpha^2}{\alpha^2 p_0^2}.$$

Tämä voi olla merkittävästi (4.3):a pienempi, jos t_0 valitaan sopivasti.

Otoskoko riippuu t_0:sta vain suureen $\eta_n(t_0)$ kautta, joten tavoitteena on minimoida $\eta_n(t_0)$:aa. Seuraavassa t_0 valitaan asymptoottisen kriteerin avulla, kun $n \to \infty$.

Koska varianssi on ei-negatiivinen, niin

$$\eta_n(t_0) \geq \mathbb{E}_{t_0}\left(e^{-t_0 Y_n + n\Lambda_n(t)} 1(\bar{X}_n \in A)\right)^2$$

$$= \mathbb{P}(\bar{X}_n \in A)^2.$$

Siispä

(6.4). $\liminf n^{-1} \log \eta_n(t_0) \geq -2\gamma$,

kts. (6.1).

Konjugaattimuunnos $\mathbb{P}_{t_0}$ on <u>tehokas</u>, jos

$$\lim n^{-1} \log \eta_n(t_0) = -2\gamma.$$

Tarkastellaan tehokkaita konjugaattimuunnoksia seuraavin oletuksin:

1) $\Lambda(t) < \infty$, $\forall t$, ja $\Lambda'(t)$ $\exists$, $\forall t$, ja $\Lambda(t) = \lim \Lambda_n(t)$

2) $\Lambda'(0) < 0$

3) $A = (a_1, a_2)$, missä $0 < a_1 < a_2 \leq \infty$.

Oletukset tehdään mukavuussyistä, lievemmät kyllä riittäisivät.

Lemma 6.1. Oletetaan, että $\{\bar{X}_n\}$ toteuttaa SPP:n vauhtifunktiolla I. Silloin

$$(6.5)\quad \liminf_{n\to\infty} n^{-1}\log \eta_n(t_0) \geq \Lambda(t_0) - \inf_{x\in \mathring{A}} \{t_0 x + I(x)\}$$

ja

$$(6.6)\quad \limsup_{n\to\infty} n^{-1}\log \eta_n(t_0) \leq \Lambda(t_0) - \inf_{x\in \bar{A}} \{t_0 x + I(x)\}.$$

Todistus. Selvästi

$$n^{-1}\log \eta_n(t_0) = n^{-1}\log \mathbb{E}\left(e^{-\langle t_0, Y_n\rangle + n\Lambda_n(t_0)} 1(\bar{X}_n\in A)\right)$$

$$= \Lambda_n(t_0) + n^{-1}\log \mathbb{E}\left(e^{-\langle t_0, Y_n\rangle} 1(\bar{X}_n\in A)\right).$$

Väitetyt tulokset seuraavat lauseen 5.2 jälkimmäisestä muodosta. $\square$

Seuraus 6.1. Konjugaattijakauma parametrilla t_0 on tehokas, jos

$$(6.7)\quad -\Lambda(t_0) + \inf_{x\in \bar{A}} \{t_0 x + I(x)\} \geq 2\inf_{x\in \mathring{A}} I(x).$$

Oletetaan lisäksi, että $I(\mathring{A}) = I(\bar{A})$. Jos konjugaattijakauma parametrilla t_0 on tehokas, niin

$$(6.8)\quad -\Lambda(t_0) + \inf_{x\in \mathring{A}} \{t_0 x + I(x)\} \geq 2\inf_{x\in \bar{A}} I(x).$$

Todistus. Jos (6.7) toteutuu, niin lemman 6.1 nojalla

$$\limsup n^{-1} \log \eta_n(t_0) \le \Lambda(t_0) - \Lambda(t_0) - 2 \inf_{x \in \mathring{A}} I(x) \le -2I(\mathring{A}).$$

Jos taas $\mathbb{P}_{t_0}$ on tehokas ja $I(\mathring{A}) = I(\bar{A})$, niin

$$-2I(\bar{A}) = \lim n^{-1} \log \eta_n(t_0)$$

$$\ge \Lambda(t_0) - \inf_{x \in \mathring{A}} \{ t_0 x + I(x) \}$$

alarajan (6.5) nojalla. Tämä on juuri (6.8). $\square$

Olkoon nyt erityisesti $I = \Lambda^*$. Oletetaan, että

$$\Lambda^*(\bar{A}) = \Lambda^*(\mathring{A}).$$

Koska $\Lambda^*(\Lambda'(0)) = 0$, on Λ^* kasvava alueessa $x \ge \Lambda'(0)$, joten

$$\Lambda^*(A) = \Lambda^*(a_1).$$

Seurauksen 6.1 soveltamiseksi, on selvitettävä millä parametrin t arvoilla

(6.9) $$-\Lambda(t) + \inf_{x \in \bar{A}} \{ tx + \Lambda^*(x) \} \ge 2\Lambda^*(a_1).$$

Olkoon aluksi $t \ge 0$. Tällöin (6.9):n vasen puoli on

$$f(t) := -\Lambda(t) + ta_1 + \Lambda^*(a_1).$$

Selvästi f on konkaavi ja yhtälöllä

$$f'(t) = -\Lambda'(t) + a_1 = 0$$

on ratkaisu, jos $a_1 \in \text{ri}\, D(\Lambda^*)$ (Lemma 3.5). Oletetaan, että näin on ja olkoon

$$(6.10) \qquad \Lambda'(t_0) = a_1.$$

Tällöin

$$f(t_0) = -\Lambda(t_0) + t_0 a_1 + \Lambda^*(a_1) = 2\Lambda^*(a_1).$$

Siis (6.7) toteutuu ja parametri t_0 antaa tehokkaan simulointijakauman. Jos $\Lambda'(t) \neq a_1$, ei t_0 toteuta ehtoa (6.7). Tällöin nimittäin f on joko aidosti kasvava tai aidosti vähenevä pisteessä t, joten

$$f(t) < \sup_{u \in \mathbb{R}} \{f(u)\} = f(t_0)$$

(Lemma 3.11). Samoin (6.7) ei toteudu, jos $t_0 < 0$.

Oletetaan vielä lisäksi, että $\{Y_n\}$ on satunnaiskulku. Voidaan näyttää, että Λ on aidosti konveksi, joten yhtälöllä (6.1) on korkeintaan yksi ratkaisu. Olkoon

$$M(t) = \mathbb{E}(e^{tY_n}) = \mathbb{E}(e^{t\xi})^n$$

Konjugaattimuunnoksen jälkeen

$$M_{t_0}(t) = \mathbb{E}_{t_0}(e^{tY_n}) = \mathbb{E}(e^{(t_0+t)Y_n - n\Lambda(t_0)})$$
$$= \frac{\mathbb{E}(e^{(t+t_0)Y_n})}{\mathbb{E}(e^{t_0 Y_n})} = \left(\frac{\mathbb{E}(e^{(t+t_0)\xi})}{\mathbb{E}(e^{t_0\xi})}\right)^n.$$

Tämä on satunnaisluvun momentit generoiva funktio, joten konjugaattimuunnoksen jälkeen Y_n on edelleen satunnaisluku. Lisäyksen jakauma on ξ:n konjugaattijakauma parametrilla t_0. (päättely perustuu tunnettuun tulokseen: jos momentit generoiva funktio on äärellinen jossain avoimessa joukossa, niin se määrää jakauman yksikäsitteisesti; M ja M_{t_0} edellä ovat momentit generoivia funktioita).

7. Brycin esityslause vauhtifunktiolle

Jos $\{Z_n\}$ toteuttaa SPP:n vauhtifunktiolla Λ^*, niin

$$\Lambda^*(x) = \sup\{\langle t, x\rangle - \Lambda(t);\ t \in \mathbb{R}^d\}$$

$$= \sup\{f(x) - \Lambda_f;\ f \in (\mathbb{R}^d)^*\},$$

missä

(7.1) $$\Lambda_f = \limsup n^{-1} \log \mathbb{E}(e^{n f(Z_n)})$$

ja

$$(\mathbb{R}^d)^* = \{f : \mathbb{R}^d \to \mathbb{R};\ f \text{ lineaarinen}\}.$$

Tämä tulos yleistyy (lähes) kaikille konvekseille vauhtifunktioille. Jos nimittäin $\{Z_n\}$ toteuttaa SPP:n konveksilla vauhtifunktiolla ja

$$\Lambda(t) = \limsup n^{-1} \log \mathbb{E}(e^{n\langle t, Z_n\rangle}) < \infty,\ \forall t \in \mathbb{R}^d,$$

niin $I = \Lambda^*$. Perusteluja esitetään harjoituksissa.

Luonnollista on kysyä, onko vauhtifunktiolla yleisemminkin jokin samantyyppinen esitys. Brycin lauseen nojalla näin on. Tarkastellaan tätä seuraavassa.

Lause 7.1. Oletetaan, että $\{Z_n\}$ toteuttaa SPP:n vauhtifunktiolla I. Silloin

$$I(x) = \sup\{f(x) - \Lambda_f;\ f \in C_b(\mathbb{R}^d)\},$$

missä Λ_f on kuten edellä ja

$$C_b(\mathbb{R}^d) = \{f : \mathbb{R}^d \to \mathbb{R};\ f \text{ jatkuva ja rajoitettu}\},$$

Todistus. Todetaan ensin, että $\Lambda_f < \infty$ ja (7.1) pätee raja-arvona. Varadhanin integraalilemman nojalla, $\forall f \in C_b(\mathbb{R}^d)$, ja

$$\Lambda_f = \sup_{x \in \mathbb{R}^d} \{ f(x) - I(x) \}.$$

Harjoitustehtävässä näytetään, että

$$I(x) \geq \sup \{ f(x) - \Lambda_f ; f \in C_b(\mathbb{R}^d) \}, \quad \forall x \in \mathbb{R}^d.$$

Käänteisen epäyhtälön todistamiseksi tarkastellaan kiinteää $x_0 \in \mathbb{R}^d$. Olkoon $M > 0$ suuri ja $\varepsilon > 0$ pieni. Olkoon ensin $I(x_0) < \infty$. Määrätään $\delta > 0$ siten, että

$$\inf \{ I(x) ; |x - x_0| \leq \delta \} \geq I(x_0) - \varepsilon.$$

Valitaan sopiva $a > 0$ ja $f = f_a$ ehdosta

$$f_a(x) = \max \{ -a|x - x_0|, -M \}.$$

Tällöin $f_a \in C_b(\mathbb{R}^d)$. Lisäksi

$$\sup \{ f_a(x) - I(x) ; |x - x_0| \leq \delta \} \leq -I(x_0) + \varepsilon$$

ja

$$\sup \{ f_a(x) - I(x) ; |x - x_0| > \delta \} \leq \max(-a\delta, -M).$$

Kiinnitetään $a = \frac{M}{\delta}$, jolloin

$$\begin{aligned} \Lambda_{f_a} &= \sup \{ f_a(x) - I(x) ; x \in \mathbb{R}^d \} \\ &\leq \max(-I(x_0) + \varepsilon, -M) \\ &= -I(x_0) + \varepsilon, \quad \text{kun oletaan } M \geq I(x_0). \end{aligned}$$

Nähdään, että

$$
\begin{aligned}
I(x_0) &\le -\Lambda_{f_a} + \varepsilon \\
&= f_a(x_0) - \Lambda_{f_a} + \varepsilon \\
&\le \sup\{ f(x_0) - \Lambda_f ;\ f \in C_b(\mathbb{R}^d)\} + \varepsilon .
\end{aligned}
$$

Olkoon lopuksi $I(x_0) = \infty$ ja $\delta > 0$ sellainen, että

$$\inf\{\ I(x) ;\ |x - x_0| \le \delta\} > M .$$

Olkoon f_a kuten edellä. Tällöin

$$\Lambda_{f_a} \le \max(-a\delta, -M)$$

ja

$$f_a(x_0) - \Lambda_{f_a} \ge M ,$$

kun valitaan $a = \frac{M}{\delta}$. Nähdään, että

$$\sup\{ f(x_0) - \Lambda_f ;\ f \in C_b(\mathbb{R}^d)\}$$

$$= \infty = I(x_0). \quad \square$$

Sisällysluettelo

Johdatus suunten poikkeamien teoriaan

Lisälähteitä

Sovelluspainotteinen kirja

Bucklew, J. (1990) Large deviation techniques in Decision, Simulation, and Estimation. Wiley & Sons.

Kohta 4. Collamore, J. (1996) Hitting probabilities and large deviations. Ann. Probab 24, 2065-2078

Kohta 6. Dieker, B. and Mandjes, M. (2005) On asymptotically efficient simulation of large deviations probabilities. Adv. Appl. Prob. 37, 539-552.