11 Raja-arvolauseita ja approksimaatioita

- Tässä luvussa esitellään sellaisia kuuluisia todennäköisyysteorian raja-arvolauseita, joita sovelletaan usein tilastollisessa päättelyssä.
- Näiden raja-arvolauseiden tunteminen kuuluu jokaisen tilastotieteilijän yleissivistykseen.
- Päätulosten todistuksia ei voida käydä läpi tämän kurssin puitteissa.

11.1 Suurten lukujen laki

- Suurten lukujen laki (engl. *law of large numbers*) kertoo, että otoskeskiarvo suppenee kohti vastaavaa odotusarvoa, kun otoskoko kasvaa rajatta.
- Suurten lukujen lakeja on olemassa lukuisia sen mukaan, minkätyyppisiä objekteja tarkastellaan ja minkälaisia oletuksia niiden yhteisjakaumasta tehdään.
- Suuren lukujen laeista on olemassa sekä heikkoja versioita (sana heikko viittaa stokastiseen suppenemiseen) että vahvoja versioita (sana vahva viittaa melkein varmaan suppenemiseen).

Stokastinen suppeneminen

Määritelmä

Jono satunnaismuuttujia $X_1, X_2, \dots$ suppenee stokastisesti eli konvergoi stokastisesti (engl. *converges in probability*) kohti satunnaismuuttujaa Y, jos

$$P(|X_n - Y| \geq \epsilon) \xrightarrow[n\to\infty]{} 0,$$

kaikilla $\epsilon > 0$.

- Voitaisiin yhtäpitävästi vaatia, että $P(|X_n - Y| > \epsilon) \to 0$ kaikilla $\epsilon > 0$, tai $P(|X_n - Y| \leq \epsilon) \to 1$ kaikilla $\epsilon > 0$.
- Stokastista suppenemista merkitään usein seuraavaan tapaan,

$$X_n \xrightarrow{\mathrm{p}} Y, \quad \text{tai} \quad \operatorname{plim}_{n\to\infty} X_n = Y.$$

Melkein varma suppeneminen

Määritelmä

Jono satunnaismuuttujia $X_1, X_2, \ldots$ suppenee (eli konvergoi) melkein varmasti (engl. *converges almost surely*) kohti satunnaismuuttujaa Y, jos $X_n(\omega) \to Y(\omega)$ perusjoukon osajoukossa, jonka tn on yksi, eli jos

$$P(\lim_{n\to\infty} X_n = Y) = 1.$$

- Melkein varmaa suppenemista merkitään esim. seuraavilla tavoilla,
$$X_n \xrightarrow{\text{m.v.}} Y, \quad \text{tai} \quad X_n \xrightarrow{\text{a.s.}} Y.$$
- On mahdollista osoittaa, että melkein varmasta suppenemisesta seuraa stokastinen suppeneminen, eli että
$$X_n \xrightarrow{\text{a.s.}} Y \qquad \Rightarrow \qquad X_n \xrightarrow{\text{p}} Y \tag{1}$$

Heikko suurten lukujen laki *i.i.d.*-jonolle

- Lyhenne i.i.d. on todennäköisyysteoriassa hyvin yleinen. Se tulee sanoista independent, identically distributed eli riippumattomat ja samoin jakautuneet.
- Todistimme jaksossa 6.1 Tšebyševin epäyhtälön avulla ns. heikon suurten lukujen lain (engl. *weak law of large numbers, WLLN*).
- Jos tätä heikkoa suurten lukujen lakia sovelletaan *i.i.d.*-jonoon satunnaismuuttujia $X_1, X_2, \ldots$, niin se toteaa, että mikäli $\sigma^2 = \mathrm{var}(X_i)$ on äärellinen, niin n ensimmäisen satunnaismuuttujan aritmeettinen keskiarvo suppenee stokastisesti kohti vastaavaa odotusarvoa $\mu = EX_i$, eli

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow{\mathrm{p}} \mu$$

Vahva suurten lukujen laki *i.i.d.*-jonolle

Lause (Vahva suurten lukujen laki)

(Engl. *strong law of large numbers, SLLN.*) Olkoon $X_1, X_2, \dots$ jono riippumattomia ja samoin jakautuneita satunnaismuuttujia, joiden odotusarvo $\mu = EX_1 \in \mathbb{R}$. Tällöin keskiarvojen

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

muodostama jono suppenee melkein varmasti kohti arvoa μ, eli $\bar{X}_n \xrightarrow{\text{a.s.}} \mu$.

- Oletuksena tarvitaan ainoastaan se, että odotusarvon täytyy olla olemassa (reaalilukuna).
- Jos jono $(\bar{X}_n)$ toteuttaa vahvan suurten lukujen lain, niin se tietenkin toteuttaa myös heikon suurten lukujen lain.

11.2 Jakaumasuppeneminen

Määritelmä

Olkoon $X_1, X_2, \ldots$ jono satunnaismuuttujia, joiden kertymäfunktiot ovat $F_1, F_2, \ldots$. Olkoon Y satunnaismuuttuja, jonka kertymäfunktio on G, ts. kaikilla x

$$F_n(x) = P(X_n \leq x), \qquad G(x) = P(Y \leq x).$$

Jono (X_n) suppenee jakaumaltaan (engl. *converges in distribution* tai *converges in law*) kohti Y:tä, mikäli

$$\lim_{n \to \infty} F_n(x) = G(x)$$

kaikissa rajajakauman kertymäfunktion G jatkuvuuspisteissä x.

Kommentteja jakaumasuppenemisesta

- Usein jakaumasuppenemisessa rajajakauma on jatkuva jakauma (esim. $N(0,1)$).
- Jatkuvan jakauman kertymäfunktio on jatkuva funktio koko reaaliakselilla. Jos rajajakauma on jatkuva, niin kertymäfunktioiden jonon pitää supeta jokaisessa reaaliakselin pisteessä,

$$\lim_{n\to\infty} F_n(x) = G(x), \qquad \text{kaikilla } x \in \mathbb{R}.$$

- Merkitsemme jakaumasuppenemista $X_n \xrightarrow{\mathrm{d}} Y$. Nuolen yläindeksi d tulee sanasta *distribution*.
- Jos rajajakauma on jokin tuttu jakauma, kuten $N(0,1)$, niin jakaumasuppenemista voidaan merkitä siten, että satunnaismuuttujan sijasta käytetään rajajakauman tunnusta, esim.

$$X_n \xrightarrow{\mathrm{d}} N(0,1).$$

Jakaumasuppenemisen hyödyntäminen

- Tilastotieteessä jakaumasuppenemista käytetään usein jakaumien approksimointiin.
- Jos tiedetään, että

$$X_n \xrightarrow{\mathrm{d}} Y, \qquad \text{kun } n \to \infty$$

  niin jollakin äärellisellä indeksin n arvolla voidaan satunnaismuuttujan X_n jakaumaa approksimoida rajamuuttujan Y jakaumalla, mitä voidaan merkitä symbolisesti

$$X_n \overset{\mathrm{d}}{\approx} Y.$$

- Edellisessä merkinnässä Y:n tilalla voidaan käyttää sen jakauman tunnusta.

Esimerkki: asymptoottinen luottamusväli

- Jos rajamuuttujalla Y on jatkuva jakauma, niin suppenemisesta $X_n \xrightarrow{\mathrm{d}} Y$ seuraa esimerkiksi, että

$$P(X_n \in I) \underset{n\to\infty}{\longrightarrow} P(Y \in I),$$

  kun $I \subset \mathbb{R}$ on mikä tahansa väli.
- Tämän ansiosta suurilla n

$$P(X_n \in I) \approx P(Y \in I).$$

- Usein tilastollisessa päättelyssä joudutaan tarkkojen luottamusvälien sijasta soveltamaan tähän ajatukseen perustuvia asymptoottisia luottamusvälejä.
- Tämän takia jakaumasuppeneminen on tärkeä käsite (frekventistisessä) tilastotieteessä.

11.3 Keskeinen raja-arvolause

- Olkoon $X_1, X_2, \ldots$ *i.i.d.*-jono, ja $\bar{X}_n$ olkoon n ensimmäisen jonon muuttujan keskiarvo. Merkitään $\mu = EX_1$ ja $\sigma^2 = \operatorname{var} X_1$.
- Tiedämme suurten lukujen lain nojalla, että $\bar{X}_n$ suppenee kohti odotusarvoa μ.
- Lisäksi tiedämme, että

$$E\bar{X}_n = \mu, \qquad \operatorname{var} \bar{X}_n = \frac{\sigma^2}{n}.$$

- Keskiarvon $\bar{X}_n$ jakauma keskittyy yhä tiiviimmin arvon μ ympärille, kun n kasvaa. Keskittymisvauhtia kuvaa tietyllä tavalla keskiarvon $\bar{X}_n$ keskihajonta $\sigma/\sqrt{n}$, joka suppenee nollaa kohti vauhdilla $1/\sqrt{n}$.

Minkä suureen jakauma suppenee keskeisessä raja-arvolauseessa

- Jos keskiarvo $\bar{X}_n$ standardoidaan vähentämällä siitä keskiarvon odotusarvo ja jakamalla keskiarvon keskihajonnalla, saadaan lauseke

$$\frac{\bar{X}_n - E\bar{X}_n}{\sqrt{\operatorname{var}\bar{X}_n}} = \sqrt{n}\,\frac{\bar{X}_n - \mu}{\sigma} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{X_i - \mu}{\sigma}.$$

- Standardoidun keskiarvon odotusarvo on tietenkin nolla, ja sen varianssi on yksi kaikilla n.
- Keskeisen raja-arvolauseen mukaan standardoitu keskiarvo suppenee jakaumaltaan kohti standardinormaalijakaumaa.

Keskeinen raja-arvolause

Lause (Keskeinen raja-arvolause)

(Engl. *central limit theorem, CLT*.) Olkoon $X_1, X_2, \ldots$ jono riippumattomia ja samoin jakautuneita satunnaismuuttujia siten, että $0 < \sigma^2 < \infty$, jossa $\sigma^2 = \operatorname{var} X_1$. Merkitään

$$\mu = EX_1, \qquad \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Tällöin

$$\sqrt{n}\,\frac{\bar{X}_n - \mu}{\sigma} \xrightarrow{\mathrm{d}} N(0,1).$$

Kommentteja

- Keskeisen raja-arvolauseen väite voitaisiin yhtä hyvin muotoilla siten, että

$$\sqrt{n}\,(\bar{X}_n - \mu) \xrightarrow{\mathrm{d}} N(0, \sigma^2).$$

- Standardoinnissa on oleellista vain se, että keskiarvon varianssin $\mathrm{var}(\bar{X}_n)$ riippuvuus otoskoosta n jaetaan pois.
- Yo. versio autta ymmärtämään moniulotteista versiota keskeisestä raja-arvolauseesta.

Moniulotteinen versio keskeisestä raja-arvolauseesta

Lause (Keskeinen raja-arvolause, moniulotteinen versio)

Jos $\mathbf{X}_1, \mathbf{X}_2, \dots$ on *i.i.d.*-jono satunnaisvektoreita, joiden yhteinen odotusarvovektori on $\boldsymbol{\mu}$ ja kovarianssimatriisi on $\boldsymbol{\Sigma}$, niin keskeinen raja-arvolause on voimassa muodossa

$$\sqrt{n}(\bar{\mathbf{X}}_n - \boldsymbol{\mu}) \xrightarrow{\mathrm{d}} N(\mathbf{0}, \boldsymbol{\Sigma}), \tag{2}$$

jossa $\bar{\mathbf{X}}_n$ on n ensimmäisen satunnaisvektorin keskiarvo $\frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i$.

11.4 Normaaliapproksimaatio

- Kun (X_i) on *i.i.d.*-jono, niin keskeiseen raja-arvolauseeseen avulla usein approksimoidaan äärellisellä, kiinteällä n

$$\sqrt{n}\,\frac{\bar{X}_n - \mu}{\sigma} \overset{\mathrm{d}}{\approx} N(0,1).$$

- Tämä on sama asia kuin approksimaatio

$$\bar{X}_n \overset{\mathrm{d}}{\approx} N(\mu, \frac{\sigma^2}{n}).$$

(Teeskentele, että edellä saatiin täsmällinen jakaumatulos, kerro vakiolla $\sigma/\sqrt{n}$ ja lopuksi lisää μ.)

- Tämä on sama asia kuin approksimaatio

$$\sum_{i=1}^{n} X_i \overset{\mathrm{d}}{\approx} N(n\mu, n\sigma^2).$$

Jakauma-approksimaatiot otoskeskiarvon ja summan jakaumille

Huomaa, miten approksimaatioiden

$$\bar{X}_n \overset{\mathrm{d}}{\approx} N(\mu, \frac{\sigma^2}{n}), \qquad \sum_{i=1}^{n} X_i \overset{\mathrm{d}}{\approx} N(n\mu, n\sigma^2).$$

normaalijakauman parametrit saadaan approksimoitavan suureen odotusarvosta ja varianssista:

$$E\bar{X}_n = \mu, \qquad \operatorname{var} \bar{X}_n = \frac{1}{n}\sigma^2$$

$$E\sum_{1}^{n} X_i = n\mu, \qquad \operatorname{var} \sum_{1}^{n} X_i = n\sigma^2.$$

Suuren otoskoon jakauma-approksimaatioita

- Keskeinen raja-arvolause takaa, että nämä normaaliapproksimaatiot (eli normaaliset approksimaatiot tai normaalijakauma-approksimaatiot) saadaan mielivaltaisen tarkoiksi, kun otoskoko n valitaan riittävän suureksi.
- Milloin otoskoko n on riittävän suuri?
- Tämä asia riippuu toisaalta halutusta tarkkuudesta ja approksimaation käyttötarkoituksesta ja toisaalta satunnaismuuttujien X_i yhteisen jakauman luonteesta.
- Symmetrisille ja yksihuippuisille jatkuville jakaumille saavutetaan pienehköllä (muutaman kymmenen) otoskoolla useisiin tarpeisiin riittävä tarkkuus, mutta vinojen jakaumien kohdalla vastaavaan tarkkuuteen tarvittava otoskoko voi olla monta kertaluokkaa suurempi.

Esimerkki: binomijakauman normaalinen approksimaatio

- Olkoot $X_1, X_2, \ldots, X_n$ riippumattomia $\mathrm{Bernoulli}(p)$-jakaumaa noudattavia sm:ia, jossa $0 < p < 1$. Tällöin

$$E\bar{X}_n = p, \qquad \operatorname{var} \bar{X}_n = \frac{1}{n} p(1-p).$$

- Normaaliapproksimaatiolla saadaan

$$\bar{X}_n \overset{\mathrm{d}}{\approx} N(p, \frac{1}{n} p(1-p)), \quad \text{ja} \quad \sum_{i=1}^{n} X_i \overset{\mathrm{d}}{\approx} N(np, np(1-p)).$$

- Tässä tapauksessa pätee tarkka tulos

$$\sum_{i=1}^{n} X_i \sim \mathrm{Bin}(n, p).$$

Numeerinen esimerkki binomijakauman approksimoinnista

- Esimerkiksi, kun $n = 25$ ja $p = 0.6$, niin $np = 15$ ja $np(1-p) = 6$.
- Kun $Z \sim N(0,1)$ ja Φ on standardinormaalijakauman kertymäfunktio, niin saadaan approksimaatio

$$P(\sum_1^n X_i \leq 13) \approx P(15 + \sqrt{6}\,Z \leq 13) = P(Z \leq \frac{13-15}{\sqrt{6}})$$
$$= \Phi\left(\frac{13-15}{\sqrt{6}}\right) \approx 0.207.$$

- Kun käytetään binomijakaumaa, eikä sen approksimaatiota, saadaan tulos

$$P(\sum_1^n X_i \leq 13) \approx 0.268.$$

Approksimaation parantaminen jatkuvuuskorjauksella

- Näin pienellä otoskoolla normaalinen approksimaatio ei ole kovin tarkka, mutta sitä voidaan parantaa käyttämällä ns. jatkuvuuskorjausta.
- Koska summa on kokonaislukuarvoinen, niin seuraavat tapahtumat ovat samoja,

$$\{\sum_{1}^{n} X_i \leq 13\} = \{\sum_{1}^{n} X_i \leq 13 + 0.5\}.$$

- Kun normaaliapproksimaatiossa yläraja 13 korvataan luvulla 13.5, niin saadaan jo alkuperäistä approksimaatiota paljon tarkempi approksimaatio

$$P(\sum_{1}^{n} X_i \leq 13) \approx P(Z \leq \frac{13.5 - 15}{\sqrt{6}}) \approx 0.270.$$

Kokonaislukuarvoisten satunnaismuuttujien summa

- Jos satunnaismuuttujat X_i ovat kokonaislukuarvoisia, niin myös niiden summa $S = \sum_{i=1}^{n} X_i$ on kokonaislukuarvoinen.
- Sen yhteydessä on mielekästä tarkastella häntätodennäköisyyksiä vain kokonaislukuarvoisissa pisteissä.
- Jos x ja y ovat kokonaislukuja, niin

$$\{S \geq x\} = \{S \geq x - \frac{1}{2}\}$$

$$\{S \leq y\} = \{S \leq y + \frac{1}{2}\}$$

$$\{x \leq S \leq y\} = \{x - \frac{1}{2} \leq S \leq y + \frac{1}{2}\}$$

Jatkuvuuskorjaus

- Kun kokonaislukuarvoisten satunnaismuuttujien summan diskreettiä jakaumaa approksimoidaan normaalijakaumalla, niin saavutetaan parempi tarkkuus, mikäli tällaiset kokonaislukuarvoisen satunnaismuuttujan kokonaislukuarvoiset rajat korvataan tähän tapaan kokonaislukujen puoliväleissä olevilla arvoilla.
- Tämä idea on nimeltään jatkuvuuskorjaus (engl. *continuity correction*).
- Jos approksimoidaan otoskeskiarvon jakaumaa, niin diskreetin jakauman arvojoukkoon (hilapisteet) sattuvat pisteet vastaavasti korvataan hilapisteiden puolivälissä olevilla arvoilla.

Jatkuvuuskorjaus *i.i.d.*-muuttujien summalle

- $S = \sum_1^n X_i$ ja X_i:t ovat kokonaislukuarvoisia.
- Jatkuvuuskorjaus johtaa summalle S approksimaatioihin

$$P(S \geq x) = P(S \geq x - \frac{1}{2}) \approx 1 - \Phi\left(\frac{x - \frac{1}{2} - n\mu}{\sqrt{n}\,\sigma}\right)$$

$$P(S \leq y) = P(S \leq y + \frac{1}{2}) \approx \Phi\left(\frac{y + \frac{1}{2} - n\mu}{\sqrt{n}\,\sigma}\right)$$

$$P(x \leq S \leq y) = P(x - \frac{1}{2} \leq S \leq y + \frac{1}{2}) \approx \Phi\left(\frac{y + \frac{1}{2} - n\mu}{\sqrt{n}\,\sigma}\right) - \Phi\left(\frac{x - \frac{1}{2} - n\mu}{\sqrt{n}\,\sigma}\right)$$

11.5 Deltamenetelmä

Deltamenetelmän (engl. *delta method*) avulla voidaan normaaliapproksimaatiota käyttää myös sileille funktioille otoskeskiarvoista.

Lause (Deltamenetelmä)

Jos

$$\sqrt{n}(X_n - a) \xrightarrow{\mathrm{d}} N(0, \sigma^2),$$

jossa a on vakio, ja funktio g on derivoituva pisteessä a, niin

$$\sqrt{n}(g(X_n) - g(a)) \xrightarrow{\mathrm{d}} N(0, \sigma^2(g'(a))^2). \tag{3}$$

Deltamenetelmän heuristinen perustelu

- Oletuksesta $\sqrt{n}(X_n - a) \xrightarrow{\mathrm{d}} N(0, \sigma^2)$ seuraa, että suurella n sm:n X_n jakauma on voimakkaasti keskittynyt arvon a lähelle, sillä
$$X_n - a = \frac{1}{\sqrt{n}} \sqrt{n}(X_n - a).$$
- Tässä jono $\frac{1}{\sqrt{n}}$ suppenee kohti nollaa ja jono $\sqrt{n}(X_n - a)$ on siinä mielessä stabiili, että sillä on rajajakauma.
- Tämän takia suurilla n tuntuu järkevältä approksimoida satunnaismuuttujaa $g(X_n)$ pisteessä a kehitetyllä ensimmäisen asteen Taylorin polynomilla, josta jäännöstermi voidaan (toivon mukaan) unohtaa.

Deltamenetelmän heuristinen perustelu päättyy

- Tehdään ensimmäisen asteen Taylorin approksimaatio

$$g(X_n) \approx g(a) + g'(a)(X_n - a),$$

  jossa jäännöstermi jätettiin pois.

- Tämän jälkeen

$$\begin{aligned} \sqrt{n}(g(X_n) - g(a)) &\approx g'(a)\sqrt{n}(X_n - a) \\ &\overset{\mathrm{d}}{\approx} g'(a)\, N(0, \sigma^2) \\ &\overset{\mathrm{d}}{=} N(0, \sigma^2\, (g'(a))^2). \end{aligned}$$

Deltamenetelmän soveltaminen

- Deltamenetelmää sovelletaan tavallisesti siten, että äärellisellä, kiinteällä otoskoolla n approksimoidaan

$$\sqrt{n}(g(X_n) - g(a)) \overset{\mathrm{d}}{\approx} N(0, \sigma^2 (g'(a))^2).$$

Esimerkki: jakauma-approksimaatio *log-odds*-suureelle

- Tarkastellaan riippumattomia muuttujia $X_i \sim \mathrm{Bernoulli}(p)$, $i = 1, \ldots, n$, jossa $0 < p < 1$.
- Parametrin p SU-estimaatti on $\bar{X}_n = s/n$, jossa s on onnistumisten lukumäärä.
- Oletetaan, että tahdotaan estimoida ns. *log-odds* -suuretta (vedonlyöntisuhteen logaritmia),

$$\theta = \ln \frac{p}{1-p}.$$

Sen luonteva estimaatti on

$$\hat{\theta}_n = \ln \frac{\bar{X}_n}{1-\bar{X}_n} = \ln \frac{s}{n-s},$$

jonka jakauma ei ole mikään tunnettu jakauma.

Jakauma-approksimaatio *log-odds*-suureelle – esimerkki päättyy

- Valitaan deltamenetelmässä

$$g(u) = \ln \frac{u}{1-u}, \qquad 0 < u < 1,$$

- Koska $\sqrt{n}(\bar{X}_n - p) \xrightarrow{\mathrm{d}} N(0, p(1-p))$, niin helppojen laskujen avulla saadaan approksimaatio

$$\sqrt{n}(\hat{\theta}_n - \theta) \overset{\mathrm{d}}{\approx} N(0, \frac{1}{p(1-p)}).$$

- Tällä perusteella olisi esim. mahdollista johtaa asymptoottinen luottamusväli parametrille θ.