Sovellettu todennäköisyyslaskenta B

Antti Rasila

3. marraskuuta 2007

Esimerkki (Laininen) 1/2

- Palkin paksuudeksi ($n = 10$) mitattiin (tuumaa):
  2.11 2.02 2.20 1.98 2.14 1.99 2.09 2.19 2.23 2.10
- Oletetaan, että palkin paksuus $X \sim N(\mu, \sigma)$.
- Otoksesta voidaan laskea $\bar{x} = 2.105$ ja $s = 0.08786$. Vapausasteparametri $\nu = 10 - 1 = 9$.
- Taulukosta saadaan $t_{0.025} = 2.262$, joten voidaan laskea

$$\mu = 2.105 \pm 2.262 \cdot \frac{0.08786}{\sqrt{10}}.$$

- Siis 95% luottamusväli on [2.04,2.17] tuumaa.

Esimerkki (Laininen) 2/2

- Varianssin luottamusväli muodostetaan 9-vapausasteiseen χ^2-jakauman avulla.
- Taulukosta (tai tietokoneella) saadaan $\chi^2_{0.975} = 2.70$ ja $\chi^2_{0.025} = 19.0$.
- Sijoitus kaavaan antaa 95% luottamusväliksi

$$\frac{9}{19.0} \cdot 0.0879^2 \leq \sigma^2 \leq \frac{9}{2.70} \cdot 0.0879^2,$$

eli [0.0037,0.0258].

Bernoulli-jakauman odotusarvon luottamusväli 1/2

- Olkoon havainnot $X_1, \ldots, X_n$ yksinkertainen satunnaisotos Bernoulli-jakaumasta parametrilla p. Eli

$$X \sim \mathrm{Bernoulli}(p).$$

- Suhteellinen frekvenssi

$$\hat{p} = \frac{f}{n}$$

on Bernoulli-jakauman *odotusarvoparametrin p harhaton estimaattori*.
- Olkoon

n = havaintojen lukumäärä
$z_{\alpha/2}$ = häntätodennäköisyyttä $\alpha/2$ vastaava piste standardoidusta normaalijakaumasta $N(0,1)$.

Bernoulli-jakauman odotusarvon luottamusväli 2/2

- Suhteellisen frekvenssin $\hat{p}$ odotusarvo ja varianssi ovat:

$$E(\hat{p}) = p$$
$$Var(\hat{p}) = \frac{pq}{n}$$

- Suhteellinen frekvenssi $\hat{p}$ noudattaa suurissa otoksissa approksimatiivisesti normaalijakaumaa.
- Bernoulli-jakauman *odotusarvoparametrin p approksimatiivinen luottamusväli luottamustasolla* $(1-\alpha)$ on

$$\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right).$$

Otoskoon määrääminen 1/2

- Ennakkotiedon mukaan Bernoulli-jakauman odotusarvoparametrin arvoksi oletetaan p. Kuinka suuri otos on otettava, jotta p:lle voidaan muodostaa $(1-\alpha)$-luottamusväli, jonka pituus on $2A$?
- Parametrin p luottamusväli luottamustasolla $(1-\alpha)$ on

$$\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right).$$

- Ennakkotiedon perusteella oletetaan $\hat{p} = p$.

Otoskoon määrääminen 2/2

- Jotta luottamusvälin pituus olisi $2A$, on oltava

$$z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} = A.$$

josta voidaan ratkaista *tarvittava otoskoko n*:

$$n = \left(\frac{z_{\alpha/2}\sqrt{p(1-p)}}{A}\right)^2.$$

- Huomaa, että otoskoko saavuttaa maksiminsa

$$n = \left(\frac{z_{\alpha/2}}{2A}\right)^2,$$

kun $p = 1/2$.

Estimaattorin valinta

- Oletetaan, että aineistona on yksinkertainen satunnaisotos $X_1, \dots, X_n$ jakaumasta
$$f_X(x; \theta_1, \dots, \theta_r).$$
- Satunnasmuuttujan X siis oletetaan noudattavan jakaumaa $f_X(x; \theta_1, \dots, \theta_r)$ parametreilla $\theta_1, \dots, \theta_r$, mutta parametrien arvoja ei tunneta.
- Tehtävänä on muodostaa parametreille $\theta_1, \dots, \theta_r$ estimaattorit $\hat{\theta}_1, \dots, \hat{\theta}_r$.
- Estimaattorien muodostamiseen on useita eri menetelmiä. Tällä kurssilla käsitellään *momenttimenetelmä* ja *suurimman uskottavuuden menetelmä*.

Momenttimenetelmä

- Useasta jakaumasta $f_X(x; \theta_1, \dots, \theta_r)$ voidaan määrittää teoreettisia momentteja $m_k = E[X^k]$, jotka ovat parametrien $\theta_1, \dots, \theta_r$ funktioita.
- Näiden teoreettisten momenttien estimaattorit $\hat{m}_k$ voidaan määrittää havaintoarvoista kaavalla
$$\hat{m}_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k.$$
- Kun teoreettiset momentit asetetaan yhtäsuuriksi estimaattorien kanssa, saadaan yhtälöryhmä, josta voidaan ratkaista parametrien $\theta_1, \dots, \theta_r$ estimaattorit.
- Jos halutaan määrittää estimaattorit r:lle parametrille, tarvitaan r yhtälöä. Tätä varten tulee määrittää r ensimmäistä teoreettista momenttia.

Esimerkki (Laininen) 1/2

- Huollon kestoajan X tiheysfunktio on
$$f_X(x; \theta_1, \theta_2) = \begin{cases} \theta_1 e^{-\theta_1(x-\theta_2)}, & \text{kun } x \geq \theta_2, \\ 0, & \text{kun } x < \theta_2, \end{cases}$$
missä $\theta_1, \theta_2 > 0$.
- Olkoot $x_1, \dots, x_n$ havaittuja huollon kestoaikoja.
- Parametrien estimoimiseksi momenttimenetelmällä lasketaan integroimalla ensimmäinen ja toinen momentti:
$$m_1 = \int_{\theta_2}^{\infty} x\theta_1 e^{-\theta_1(x-\theta_2)}\,dx = \frac{1}{\theta_1} + \theta_2,$$
$$m_2 = \int_{\theta_2}^{\infty} x^2\theta_1 e^{-\theta_1(x-\theta_2)}\,dx = \theta_2^2 + \frac{2}{\theta_1^2} + \frac{2\theta_2}{\theta_1}.$$

Esimerkki (Laininen) 2/2

- Estimaatit θ_1, θ_2 ratkaistaan yhtälöistä
$$\hat{m}_1 = \frac{1}{\hat{\theta}_1} + \hat{\theta}_2, \qquad \hat{m}_2 = \theta_2^2 + \frac{2}{\theta_1^2} + \frac{2\theta_2}{\theta_1},$$
missä
$$\hat{m}_1 = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \hat{m}_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2.$$
- Ratkaisuksi saadaan
$$\hat{\theta}_1 = \frac{1}{\sqrt{\hat{m}_2 - \hat{m}_1^2}}, \qquad \hat{\theta}_2 = \hat{m}_1 - \sqrt{\hat{m}_2 - \hat{m}_1^2}.$$

Suurimman uskottavuuden menetelmä

- *Suurimman uskottavuuden menetelmän* ajatuksena on, että malli $f_X(x; \theta_1, \dots, \theta_r)$ on paras, jos se tuottaa havaitun otoksen suurimmalla todennäköisyydellä.
- Menetelmässä muodostetaan *likelihood-funktio*
$$L(\theta_1, \dots, \theta_r),$$
joka kuvaa saadun otoksen todennäköisyyttä parametrien $\theta_1, \dots, \theta_r$ funktiona.
- Kun likelihood-funktiota maksimoidaan parametrien $\theta_1, \dots, \theta_r$ suhteen, ovat saadut satunnaismuuttujat $\hat{\theta}_1, \dots, \hat{\theta}_r$ parametrien $\theta_1, \dots, \theta_r$ *suurimman uskottavuuden estimaattorit*.

Diskreetti tapaus 1/2

- Olkoon $X_1, \dots, X_n$ yksinkertainen satunnaisotos diskreetistä jakaumasta, jonka pistetodennäköisyysfunktio on
$$f_X(x_i; \theta_1, \dots, \theta_r).$$
- Todennäköisyys otokselle $x_1, \dots, x_n$ on
$$L(\theta_1, \dots, \theta_r) = f_X(x_1; \theta_1, \dots, \theta_r) \cdots f_X(x_n; \theta_1, \dots, \theta_r).$$
- Parametreille $\theta_1, \dots, \theta_r$ tulee löytää sellaiset arvot, että L maksimoituu.

Diskreetti tapaus 2/2

- L:n maksimi löytyy pisteestä, jossa kaikki osittaisderivaatat ovat nollia, eli
$$\frac{\partial L}{\partial \theta_i} = 0, \quad \forall\, i$$
- Funktion L *logaritmimuunnos* $\ln(L)$ saa maksiminsa samassa pisteessä kuin L. Usein tarkastellaankin funktiota $\ln(L)$, koska logaritmimuunnos muuntaa tulomuotoisen likelkihood-funktion summaksi, jota on helpompi derivoida.

Esimerkki

- Tiedonsiirtojärjestelmä siirtää merkkejä 0 ja 1. Merkki 1 voi vaihtua merkiksi 0 todennäköisyydellä p.
- Todennäköisyyden arvioimiseksi lähetetään 10000 kertaa merkki 1, jolloin todetaan, että merkki on vaihtunut 90 kertaa.
- Koska 10000 merkin siirrossa merkki on välittynyt oikein 9910 kertaa ja vaihtunut 90 kertaa, saadaan
$$L(p) = (1-p)^{9910} p^{90}.$$
- Etsitään se parametrin p arvo, jolla $L(p)$ saa maksiminsa:
$$\frac{\partial}{\partial p} L(p) = -9910(1-p)^{9909} p^{90} + 90(1-p)^{9910} p^{89} = 0.$$
- Ratkaistaan
$$\hat{p} = \frac{90}{10000} = 0.009.$$

Jatkuva tapaus

- Olkoon $X_1, \ldots, X_n$ yksinkertainen satunnaisotos jatkuvasta jakaumasta, jonka tiheysfunktio on

$$f_X(x; \theta_1, \ldots, \theta_r).$$

- Koska jatkuvassa tapauksessa kaikkien yksittäisten havaintosarjojen $x_1, \ldots, x_n$ todennäköisyys on nolla, likelihood-funktio

$$L(\theta_1, \ldots, \theta_r) = f_X(x_1; \theta_1, \ldots, \theta_r) \cdots f_X(x_n; \theta_1, \ldots, \theta_r)$$

ei kuvaa havaintosarjan todennäköisyyttä. Tiheysfunktion arvojen tulona sitä voidaan kuitenkin perustellusti käyttää kuten diskreetissä tapauksessa.

Esimerkki

- Oletetaan, että X on tasaisesti jakautunut välille $[0, \theta]$, missä parametri $\theta > 0$ ja X:n tiheysfunktio on

$$f_X(x) = \begin{cases} 1/\theta, & \text{kun } x \in [0, \theta], \\ 0, & \text{muuten.} \end{cases}$$

- Likelihood-funktio on

$$L(\theta) = \begin{cases} 1/\theta^n, & \text{kun } \theta \geq \max\{x_1, \ldots, x_n\}, \\ 0 & \text{muuten.} \end{cases}$$

- Nyt maksimi ei löydy derivoimalla, mutta voidaan päätellä tutkimalla funktiota $L(\theta)$.
- Suurin arvo saadaan selvästi, kun $\hat{\theta} = \max\{x_1, \ldots, x_n\}$.