TILASTOLLISEN PÄÄTTELYN JATKOKURSSI / NIEMI

9.8.2007

1. Oletetaan T_n, $n = 1, 2, \dots$, on jono θ:n estimaattoreita, jolle pätee

$$\sqrt{n}\,(T_n - \theta) \xrightarrow[w]{} N(0, \tau^2). \quad (n \to \infty)$$

Estimoitava $g(\theta)$. Tarkastele $g(\theta)$:n estimaattorijonon $g(T_n)$, $n = 1, 2, \dots$, asymptoottista jakaumaa tilanteessa $g'(\theta) = 0$, $g''(\theta) \neq 0$ [ohje: Taylor-sarja].

2. Osoita, että eksponentiaalisen perheen jakaumille uskottavuusyhtälöllä on yksikäsitteinen ratkaisu (1-ulotteinen parametri, kanoninen parametrisointi).

3. Estimaattorijonon konsistenssi. Anna määritelmä ja kerro miksi konsistenssi on tilastotieteessä tärkeä asia.

Muodosta parametrin

$$p = P(X \le x_0)$$

x_0 kiinteä

konsistentti estimaattorijono. Perustele konsistenssiominaisuuden voimassaolo.

4. LR-testin testisuureen määrittäminen tapauksessa $\underline{X} \sim N(\mu, 1)$

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

Tarkastele/havainnollista asiaa myös graafisesti logaritmoidun uskottavuusfunktion kuvaajan avulla.

Tilastollisen päättelyn jatkokurssi / NIEMI

1. Estimaattorijonon konsistenssi

a) Esitä estimaattorijonon konsistenssin määritelmä

b) Miksi konsistenssi on tärkeä ominaisuus

c) Määritä parametrin

$$p = F\{\underline{X} \leq x_0\}$$

konsistentti estimaattorijono (perustelut!)

Tässä: x_0 on annettu (kiinteä!) reaaliluku ja F on tunnettu = tarkasteltavan satunnaismuuttujan $\underline{X}$ kertymäfunktio.

2. Kaksisuuntainen luokitus, ts. tilanne, jossa havaintoyksiköt luokitellaan kahden kriteerin / muuttujan mukaan.

Esimerkiksi kohderyhmän henkilöt hiusten värin ja silmien värin mukaan (Ks. oheinen liite).

a) Esitä ilmiötä kuvaava satunnaismuuttuja / todennäköisyysmalli

b) Tarkastellaan hypoteesia

H_0: luokitukset ovat riippumattomat.

Hypoteesia voidaan testata LR-testillä (uskottavuusosamäärätesti). Määritä testisuureen asymptoottisen jakauman ($= \chi^2$-jakauma) vapausasteiden lukumäärä (perustelut!)

3) Suurimman uskottavuuden estimaattori ($=$ MLE)

a) Määritä MLE tapauksessa $X \sim \text{Bernoulli}(\theta)$, kun käytettävissä on $\perp\!\!\!\perp$ obs. $x_1, \ldots, x_n$.

Esimerkin tapauksessa:

b) Onko MLE harhaton (perustelut)

c) Arvioi MLE:n varianssin hyvyyttä (perustelut)

4)) Rao-Blackwell -menetelmä harhattoman estimaattorin varianssin pienentämiseksi

a) Formuloi Rao-Blackwell-lause

b) Anna esimerkki tilanteesta, jossa Rao-Blackwell menetelmällä ei varianssia ole mahdollista enää pienentää

c) Olkoon $X_1, \dots, X_n$ II otos jatkuvasta jakaumasta ja $(X_{(1)}, \dots, X_{(n)})$ vastaava järjestystunnusluku. Laske

$$E(X_1 \mid (X_{(1)}, \dots, X_{(n)}))$$

ja

$$E(\bar{X} \mid (X_{(1)}, \dots, X_{(n)})).$$

5) Olkoon $\underline{X} \sim N(\theta, \sigma^2)$.
Tarkastele θ^2:n estimaattori-
jonon $(\overline{\underline{X}}_n)^2$ $\left[\overline{\underline{X}}_n = \frac{1}{n}(\underline{X}_1 + \dots + \underline{X}_n)\right]$
$[\underline{X}_1, \dots, \underline{X}_n \perp\!\!\!\perp$ otos$]$. asymptoottista
jakaumaa $(n \to \infty)$, kun
a) $\theta \neq 0$
b) $\theta = 0$

Perustelut (Asymptoottisen normaalisuuden säilymistä muunnoksissa koskevat yleiset tulokset voi olettaa tunnetuiksi).

__LIITE:__ __Tehtävä 2.__ Esimerkki 2-suuntaisesta luokittelusta.

Table B.5 Observed frequencies of people with a particular hair and eye colour

	Hair colour				
Eye colour	Black	Brunette	Red	Blond	Total
Brown	68	119	26	7	220
Blue	20	84	17	94	215
Hazel	15	54	14	10	93
Green	5	29	14	16	64
Total	108	286	71	127	592

TILASTOLLISEN PÄÄTTELYN JATKOKURSSI

TENTTI 20.5.2008

1. Oletetaan $X \sim \text{Poisson}(\lambda)$.

a) Määritä keskeisen raja-arvolauseen mukainen $\overline{X}_n$:in asymptoottinen jakauma.

b) Määritä (varianssin stabiloiva) muunnos $g(\cdot)$, s.e.

$$\sqrt{n}\,(g(\overline{X}_n) - g(\lambda)) \xrightarrow[L]{} N(0, c),$$

missä c on λ:sta riippumaton vakio

$\overline{X}_n$ = n:n suuruisen // otoksen otoskeskiarvo.

2. Oletetaan, että

$$Y_i \sim N(\alpha, \sigma^2), \quad i = 1, \dots, n \quad //$$

σ^2 tunnettu

Muodosta hypoteesin

$$H_0 : \alpha = \alpha_0$$

LR-testi, osoita, että tässä tapauksessa saadaan testisuureen tarkka jakauma ym.

3. Oletetaan

$$\sqrt{n}\,(T_n - \theta) \xrightarrow[L]{} N(0, \tau^2)$$

Estimoitava $g(\theta)$ (g riittävän säännöllinen funktio.) Määritä $g(\theta)$:n estimaattorin $g(T_n)$ ~~asymptoottisen~~ asymptoottinen jakauma, kun $g'(\theta) = 0$ & $g''(\theta) \neq 0$.

Ohje: Taylor-sarja kehitelmä.

4. Olkoon $\bar{X} \sim N(\theta, \sigma^2)$, $\theta \neq 0$, σ^2 tunnettu. Osoita, että

$$\delta_n = (\bar{X}_n)^2 - \sigma^2/n \quad (\sigma^2 \text{ tunnettu})$$

on θ^2:n

a) harhaton ja

b) asymptoottisesti normaalinen estimaattori.

c) Määritä asympt. jakauma.

Poisson-jakauma: $X \sim \mathrm{Poisson}(\lambda)$:

$$P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, \ldots$$

Normaali-jakauma: $X \sim N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

TILASTOLLISEN PÄÄTTELYN JATKOKURSSI 12/6 2008

1. θ:n suhteen minimaalisen tyhjentävän tunnusluvun konstruktio aliluokkamenetelmällä kun $X \sim \text{Bernoulli}(\theta)$. Esitä myös (teoria-) perustelut aliluokkamenetelmän käytölle tässä tilanteessa.

2. LR-testin määrittäminen tapauksessa $X \sim N(\mu, 1)$

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0.$$

Havainnollista LR-testisuureen arvon määräytymistä myös graafisesti logaritmoidun uskottavuusfunktion avulla.

3. a) Formuloi Rao-Blackwell lause. Mikä on Rao-Blackwell menetelmän käyttötarkoitus?

3. b) Olkoon $X \sim \text{Poisson}(\lambda)$.

Sovella Rao–Blackwell menetelmää λ:n harhattomalle estimaattorille $\bar{X}$ ja λ:n suhteen tyhjentävälle tunnusluvulle $T = X_1 + \dots + X_n$ ($\underset{\sim}{X} = (X_1, \dots, X_n)$ $\perp\!\!\!\perp$ otos). Mitä tulos kertoo $\bar{X}$:n ominaisuuksista λ:n estimaattorina.

4. Olkoon $\underset{\sim}{X}$ satunnaisvektori, jolle on voimassa

$$P\left\{\underset{\sim}{X} = \begin{bmatrix}1\\0\end{bmatrix}\right\} = \theta_1, \quad P\left\{\underset{\sim}{X} = \begin{bmatrix}0\\1\end{bmatrix}\right\} = \theta_2, \quad P\left\{\underset{\sim}{X} = \begin{bmatrix}0\\0\end{bmatrix}\right\} = 1 - \theta_1 - \theta_2$$

(kolme vaihtoehtoista arvoa) & $\theta_1 > 0$, $\theta_2 > 0$, $\theta_1 + \theta_2 < 1$.

a) Määritä $\underset{\sim}{X}$:n varianssi/kovarianssimatriisi.

b) Määritä keskeisen raja-arvolauseen mukainen $\underset{\sim}{\bar{X}}$:n rajajakauma ($\underset{\sim}{\bar{X}}_n$ = otoskeskiarvo, kun $\underset{\sim}{X}_i \sim \underset{\sim}{X}$ $\perp\!\!\!\perp$, $i = 1, \dots, n$).

21.10.2008

Tilastollisen päättelyn jatkokurssi:

1. Olkoon $\underset{\sim}{X}$ satunnaisvektori (2-dim), jolle on voimassa

$$P\{\underset{\sim}{X} = \begin{bmatrix}1\\0\end{bmatrix}\} = \theta_1, \quad P\{\underset{\sim}{X} = \begin{bmatrix}0\\1\end{bmatrix}\} = \theta_2$$

$$P\{\underset{\sim}{X} = \begin{bmatrix}0\\0\end{bmatrix}\} = 1 - \theta_1 - \theta_2$$

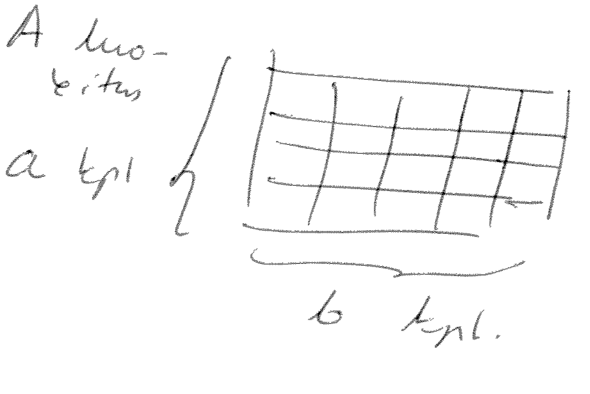
(kolme vaihtoehtoista arvoa) &

$$\theta_1 \geq 0, \quad \theta_2 \geq 0, \quad \theta_1 + \theta_2 \leq 1$$

a) Määritä $\underset{\sim}{X}$:n varianssi/kovarianssimatriisi.

b) Määritä keskeisen raja-arvolauseen mukainen $\underset{\sim}{\bar{X}}_n$:n raja-jakauma ($\underset{\sim}{\bar{X}}_n$ = otoskeskiarvo, kun $\underset{\sim}{X}_i \sim \underset{\sim}{X}$ ⊥⊥, $i = 1, \ldots, n$).

2. Kaksisuuntainen luokitus



(Esim. A ⟷ hiustenväri
B ⟷ silmien väri)

a) Ilmiöstä kuvaava todennäköisyys-malli/jakauma?

b) Esitä hypoteesi

H₀: A-luokitus on riippumaton B-luokituksesta

parametrisessa muodossa.

c) Määritä vastaavan uskottavuus-osamäärätestin asymptoottinen jakauma (tukeutuen yleisiin teoriatuloksiin). Vapausasteiden lukumäärä on perusteltava!

3. a) Formuloi Rao-Blackwell lause. Mikä on Rao-Blackwell menetelmän käyttötarkoitus.

b) Olkoon $\underline{X} \sim N(\mu, 1)$. Käytä Rao-Blackwell menetelmää

μ:n harhattomalle estimaattorille $\overline{X}$ ja μ:n suhteen tyhjentävälle tunnusluvulle $T = X_1 + \dots + X_n$ ($X_1, \dots, X_n$ iid otos). Tulkitse tulos $\overline{X}$:n Minimum Variance Unbiased Estimator -ominaisuuden kannalta.

4. Oletetaan: $E(X) = \mu$, $Var(X) = \sigma^2 < \infty$ & $\mu_4 < \infty$.
Mitä osaat sanoa σ:n estimaattorin $\sqrt{S_n^2}$ asymptoottisista ominaisuuksista? Perustelut?

ohje: Tiedetään tuloksesta

$$\sqrt{n}\,(S_n^2 - \sigma^2) \xrightarrow[\mathcal{L}]{} N(0, \mu_4 - \sigma^4)$$

$(n \to \infty)$

Tässä: $s_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$

$s_n'^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$

$\mu_4 = E(\bar{X}^4)$.

5. a) Palauta mieleen (= esitä) tyhjentävyyden määritelmä ja sen tulkinta.

b) Miksi tunnusluvun tyhjentävyys on tärkeä asia?

12.11.2008 Tilastollisen päättelyn jatkokurssi

1. a) Formuloi Rao-Blackwell lause. Mikä on Rao-Blackwell menetelmän käyttötarkoitus?

b) Olkoon $X \sim N(\mu, 1)$. Käytä Rao-Blackwell menetelmää μ:n harhattomalle estimaattorille $\bar{X}$ ja μ:n suhteen tyhjentävälle tunnusluvulle $T = X_1 + \dots + X_n$ ($X_1, \dots, X_n$ ⫫ otos). Tulkitse tulos $\bar{X}$:n Minimum Variance Unbiased Estimator -ominaisuuden kannalta.

2. Oletetaan, että $X \sim \text{Poisson}(\lambda)$.

a) Määritä $\bar{X}_n$:n asymptoottinen normaalijakauma. (2 p)

b) Määritä (varianssin stabiloiva) muunnos $g(\cdot)$, ts. muunnos $g(\cdot)$, (4 p) jolle pätee

$$\sqrt{n}\,(g(\bar{X}_n) - g(\lambda)) \xrightarrow[\mathcal{L}]{} N(0, c)$$

missä c on jokin λ:sta riippumaton vakio.

<u>Huom.</u> $\bar{X}_n = \frac{1}{n}(X_1 + \dots + X_n)$

$X_1, \dots, X_n$ II dtos

$X \sim \text{Poisson}(\lambda) \Rightarrow E(X) = \lambda,\ \text{var}(X) = \lambda.$

<u>3.</u> Oletetaan: $\sqrt{n}\,(T_n - \theta) \xrightarrow{\mathcal{L}} N(0, \tau^2)$

Estimoitava $g(\theta)$, (kun g on riittävän säännöllinen muunnos). Tarkastele $g(\theta)$:n estimaattorin $g(T_n)$ asymptoottista jakaumaa tilanteessa $g'(\theta) = 0$, $g''(\theta) \neq 0$

[Ohje: Taylor sarjatarkastelu].

4. Oletetaan $\underline{X} \sim N(\mu, \sigma^2)$.
Muodosta LR-testisuure (=uskottavuusosamäärätesti) ja kehitä sen esitys muodossa, josta nähdään tarkka jakauma äärellisillä n:n arvoilla, kun

$$H_0 : \sigma^2 = \sigma_0^2 .$$

Miten kriittinen alue määräytyy tässä tapauksessa?

Ohje: μ:n MLE $= \bar{X}$ sekä vapaan että sidotun (H_0:n vallitessa) mallin tapauksessa & $\hat{\sigma}^2_{MLE} = s_n^2$ (vapaan mallin tapauksessa),

$$s_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 .$$

5. Estimaattorijonon konsistenssi.

a) Esitä määritelmä

b) Määritä parametrin

$$p = F(X \le x_0)$$

| x_0 kiinteä, F = X:n kertymäfunktio

konsistentti estimaattori, lähtökohtana X:stä saadut riippumattomat havainnot $X_1, \dots, X_n$.

Perustele konsistenssiominaisuus!