Matemaattisten tieteiden kandiohjelma / MTO

Tilastollinen päättely I

Kurssikoe 8.5.2019 (kesto 2h 30min)

Sallitut apuvälineet: kirjoitusvälineet ja laskin. Tehtäväpaperin ohessa olevia tietoja jakaumista ja taulukoista saa myös käyttää

1. Henkilö A kirjoittaa harjoitustehtävänä simulointiohjelmaa R:llä. Hänen ohjelmassaan on kuitenkin virhe, joka ilmeneee erikoisena virheilmoituksena ohjelman ajoa toistettaessa todennäköisyydellä $\theta \in [0, 1]$. Kunkin toiston voi ajatella olevan riippumaton. Henkilö A ajaa kuitenkin ohjelmaansa n kertaa, selvittääkseen kuinka usein virhe ilmenee. Satunnaismuuttuja X kuvaa niiden kertojen lukumäärää, milloin ohjelma ilmoittaa virheestä näiden n toistojen aikana.

   (a) Muotoile tilanteeseen sopiva tilastollinen malli. Perustele valintasi huolellisesti.

   (b) Johda huolellisesti perustellen parametrin θ suurimman uskottavuuden estimaattori.

   (c) Henkilö A ajoi ohjelmansa 850 kertaa ja näistä 9 kertaa ohjelma ilmoitti virheestä. Mikä on havaittua aineistoa vastaava suurimman uskottavuuden estimaatti parametrille θ?

2. Tehdas tuottaa tenttitehtäviä varten valkoisia palloja, joiden tavoitesäde on 3 cm. Pallojen säteen on havaittu noudattavan likimain normaalijakaumaa. Kontrollitestin yhteydessä tuotetuista palloista poimitaan satunnaisesti n kappaletta palloa, joiden säteet mitataan. Mallinnetaan tilannetta riippumattomilla satunnaismuuttujilla $Y_1, \ldots, Y_n$, jotka noudattavat jakaumaa $N(\mu, \sigma^2)$, missä $\mu \in \mathbb{R}$ on tuntematon parametri ja $\sigma^2 > 0$ on tunnettu. Tarkastellaan parametrin μ luottamusväliä luottamustasolla $1 - \alpha$, missä $\alpha \in (0, 1)$.

   (a) Kumpi seuraavista sopisi tilanteeseen paremmin: z- vai t- luottamusväli? Perustele valintasi huolellisesti.

   (b) Oletetaan, että $n = 16$, $\sigma^2 = 0.04$ ja havaitaan $\overline{y} = 3.14$. Valitaan $\alpha = 0.05$. Laske luottamustason $1 - \alpha$ kaksisuuntainen luottamusväli parametrille μ.

   (c) Mitä voit tulosten perusteella päätellä pallojen todellisesta säteestä ja otoskoon riittävyydestä?

3. Oletetaan, että havaintoja $y_1, \ldots, y_n$ vastaavat satunnaismuuttujat $Y_1, \ldots, Y_n$ ovat riippumattomia ja noudattavat jakaumaa $N(\mu, \sigma^2)$, missä molemmat parametrit ovat tuntemattomia. Tarkastellaan hypoteesiparia $H_0 : \mu = \mu_0$ ja $H_1 : \mu \neq \mu_0$.

   (a) Muodosta tilanteeseen sopiva testisuure. Mitä jakaumaa valitsemasi testisuure noudattaa, kun H_0 pätee? Onko kyseessä yksi- vai kaksisuuntainen testi?

   (b) Oletetaan, että testissä on havaittu p-arvo 0.037. Millaisella kaavalla p-arvo olisi laskettu? Hylätäänkö nollahypoteesi merkitsevyystasolla $\alpha = 0.05$?

   (c) Jos parametrille μ laskettaisiin 99 prosentin luottamusväli, niin kuuluisiko piste μ_0 tälle luottamusvälille, jos on havaittu p-arvo 0.037 kuten edellisessä kohdassa?

4. Pöydällä on neljä ulkoapäin täysin identtistä kulhoa. Henkilö B tietää, että

   - kulhossa 1 on 4 valkoista ja yksi vihreä pallo
   - kulhossa 2 on 3 valkoista ja kaksi vihreää palloa

- kulhossa 3 on 6 valkoista palloa
- ja kulhossa 4 on 2 valkoista palloa, yksi vihreä ja yksi keltainen pallo

Henkilö C ei tiedä, minkälaisia palloja kulhoissa on. He valitsevat umpimähkään yhden kulhoista, mutta valitun kulhon numero θ, joka on siis 1,2,3 tai 4, on tuntematon. Henkilö B nostaa sekoittaen takaisinpanolla valitusta θ kulhosta kaksi kertaa, ja näyttää pallon henkilölle C. Kummallakin kerralla pallo on valkoinen. Esitä parametrille θ priorijakauma, uskottavuusfunktio sekä posteriorijakauma luettelemalla niiden arvot.

Muistin tueksi

$$X \sim \mathrm{Bin}(n,p) \Longrightarrow P(X=k) = \binom{n}{k} p^k (1-p)^{n-k},\ k = 0,1,\dots,n;$$
$$EX = np \text{ ja } \mathrm{Var}X = np(1-p).$$

$$X \sim \mathrm{Hyperg}(N,K,n) \Longrightarrow P(X=k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}},\ k = 0,1,\dots,n;$$
$$EX = n\frac{K}{N} \text{ ja } \mathrm{Var}X = n\frac{K}{N}\frac{N-K}{N}\frac{N-n}{N-1}.$$

$$X \sim \mathrm{Geom}_0(p) \Longrightarrow P(X=k) = p(1-p)^k,\ k = 0,1,2,\dots;$$
$$EX = \frac{1-p}{p} \text{ ja } \mathrm{Var}X = \frac{1-p}{p^2}.$$

$$X \sim \mathrm{Geom}_1(p) \Longrightarrow P(X=k) = p(1-p)^{k-1},\ k = 1,2,3,\dots;$$
$$EX = \frac{1}{p} \text{ ja } \mathrm{Var}X = \frac{1-p}{p^2}.$$

$$X \sim \mathrm{Poisson}(\lambda) \Longrightarrow P(X=k) = e^{-\lambda}\frac{\lambda^k}{k!},\ k = 0,1,2,\dots;\quad EX = \lambda \text{ ja } \mathrm{Var}X = \lambda.$$

$$X \sim \mathrm{Tas}(a,b) \Longrightarrow f(x) = \begin{cases} \frac{1}{b-a}, & x \in (a,b), \\ 0, & \text{muulloin}; \end{cases} \quad EX = \frac{a+b}{2} \text{ ja } \mathrm{Var}X = \frac{(b-a)^2}{12}.$$

$$X \sim \mathrm{Exp}(\lambda) \Longrightarrow f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & \text{muulloin}; \end{cases} \quad EX = \frac{1}{\lambda} \text{ ja } \mathrm{Var}X = \frac{1}{\lambda^2}.$$

$$X \sim N(\mu,\sigma^2) \Longrightarrow f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},\ x \in \mathbb{R};\quad EX = \mu \text{ ja } \mathrm{Var}X = \sigma^2.$$

Standardinormaalijakauman yläkvantiileja

α	0.1	0.05	0.025	0.01	0.005
z_α	1.28	1.64	1.96	2.33	2.58

t-jakauman yläkvantiileja

α	0.1	0.05	0.025	0.01	0.005
$t_2(\alpha)$	1.89	2.92	4.30	6.96	9.92
$t_3(\alpha)$	1.64	2.35	3.18	4.54	5.84
$t_4(\alpha)$	1.53	2.13	2.78	3.75	4.60
$t_5(\alpha)$	1.48	2.02	2.57	3.36	4.03
$t_6(\alpha)$	1.44	1.94	2.45	3.14	3.71
$t_7(\alpha)$	1.41	1.89	2.36	3.00	3.50
$t_8(\alpha)$	1.40	1.86	2.31	2.90	3.36
$t_9(\alpha)$	1.38	1.83	2.26	2.82	3.25
$t_{10}(\alpha)$	1.37	1.81	2.23	2.76	3.17
$t_{11}(\alpha)$	1.36	1.80	2.20	2.72	3.11
$t_{12}(\alpha)$	1.36	1.78	2.18	2.68	3.05
$t_{13}(\alpha)$	1.35	1.77	2.16	2.65	3.01
$t_{14}(\alpha)$	1.35	1.76	2.14	2.62	2.98
$t_{15}(\alpha)$	1.34	1.75	2.13	2.60	2.95
$t_{16}(\alpha)$	1.34	1.75	2.12	2.58	2.92
$t_{17}(\alpha)$	1.33	1.74	2.11	2.57	2.90
$t_{18}(\alpha)$	1.33	1.73	2.10	2.55	2.88
$t_{19}(\alpha)$	1.33	1.73	2.09	2.54	2.86
$t_{20}(\alpha)$	1.33	1.72	2.09	2.53	2.85
$t_{21}(\alpha)$	1.32	1.72	2.08	2.52	2.83
$t_{22}(\alpha)$	1.32	1.72	2.07	2.51	2.82
$t_{23}(\alpha)$	1.32	1.71	2.07	2.50	2.81
$t_{24}(\alpha)$	1.32	1.71	2.06	2.49	2.80
$t_{25}(\alpha)$	1.32	1.71	2.06	2.49	2.79
$t_{26}(\alpha)$	1.31	1.71	2.06	2.48	2.78
$t_{27}(\alpha)$	1.31	1.70	2.05	2.47	2.77
$t_{28}(\alpha)$	1.31	1.70	2.05	2.47	2.76
$t_{29}(\alpha)$	1.31	1.70	2.05	2.46	2.76
$t_{30}(\alpha)$	1.31	1.70	2.04	2.46	2.75