Tilastomatematiikka

Loppukoe 11.04.2019

1. Lähetetään 4 bittiä digitaalisessa viestintäkanavassa. Yksittäinen bitti välittyy joko oikein tai virheellisenä muista biteistä riippumatta. Olkoot

$$A_i = \text{"}i\text{:s bitti välittyy virheellisenä"}, \quad i = 1, 2, 3, 4,$$

tapahtumia.

a) Määrää tilanteeseen sopiva otosavaruus ja luettele suotuisat alkeistapahtumat tapahtumille

$$A = A_1 \cap A_2 \cap A_3 \cap A_4 \quad \text{ja} \quad B = (A_1 \cap A_2) \cup (A_3 \cap A_4)$$

b) Määrää a)-kohdassa annettujen tapahtumien A ja B todennäköisyydet, jos yksittäinen bitti välittyy virheellisenä todennäköisyydellä p.

2. Elektronisen komponentin elinikä tunteina on satunnaismuuttuja X, jonka tiheysfunktio on

$$f_X(x) = \tfrac{1}{2000}e^{-\frac{1}{2000}x}, \quad x > 0.$$

a) Valmistaja antaa komponentin takuuiäksi 500 tuntia. Millä todennäköisyydellä komponentti kestää takuuajan?

b) Jos komponentti on ollut ehjänä käytössä 500 tuntia, niin millä todennäköisyydellä komponentti kestää toiset 500 tuntia?

3. Analysoitiin katalyytin vaikutusta erään kemiallisen prosessin keskimääräiseen saantoon. Katalyytti 1 on käytössä, mutta katalyytti 2 on halvempi, joten sitä olisi syytä käyttää edellyttäen, ettei sen käyttö muuta prosessin saantoa. Muotoile sopivat hypoteesit ja testaa ne riskitasolla 5%, kun analyysi antoi seuraavan havaintoaineiston

	n	$\overline{x}$	s
Katalyytti 1	8	92.26	2.39
Katalyytti 2	8	92.73	2.98

Kannattaako katalyyttiä vaihtaa tämän aineiston perusteella?

4. Tutkittiin transistorin kantavirran y [mA] ja kanta-emitterijännitteen x [V] välistä yhteyttä. Saatiin seuraavat mittaukset

x	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
y	0.734	0.886	1.04	1.19	1.35	1.50	1.66	1.81	1.97	2.12

a) Määrää havaintoja vastaava regressiosuora. Piirrä havaintopisteet ja regressiosuora samaan koordinaatistoon.

b) Paljonko mallin mukaan kasvaa kantavirta, kun jännite kasvaa yhden voltin?

c) Mikä on mallin mukaan kantavirta, kun jännite on 0 [V]? Miten tulkitset tuloksen?

5. Määrää c ja muuttujien X ja Y välinen kovarianssi, kun satunnaisvektorin (X, Y) arvojoukko on $S_{X,Y} = \{(x, y) \,|\, x = 1, 2,\ y = 1, 2\}$ ja pistetodennäköisyydet ovat muotoa $p_{xy} = P(\{X = x\} \cap \{Y = y\}) = c(x + y)$ kaikilla $(x, y) \in S_{X,Y}$.

Tilastomatematiikka

Loppukoe 08.06.2019

1. Lähetetään 4 bittiä digitaalisessa viestintäkanavassa. Yksittäinen bitti välittyy joko oikein tai virheellisenä muista biteistä riippumatta. Olkoon

$$S = \{x_1x_2x_3x_4 \,|\, x_i \in \{0,1\},\ i = 1,2,3,4\}$$

satunnaiskoetta kuvaava otosavaruus ja olkoot

$$A_i = \text{"}i\text{:s bitti välittyy virheellisenä"}, \quad i = 1,2,3,4,$$

tapahtumia.

a) Ovatko tapahtumat A_i toisensa poissulkevia? (2p)

b) Luettele suotuisat alkeistapahtumat tapahtumille (4p)

$$A = \overline{A_1 \cup A_2 \cup A_3 \cup A_4} \quad \text{ja} \quad B = (A_1 \cup A_2) \cap (A_3 \cup A_4)$$

2. Satunnaismuuttuja X on *Pareto*-jakautunut parametreilla $m > 0$ ja $\alpha > 0$, jos sen kertymäfunktio on

$$F_X(x) = 1 - \left(\frac{m}{x}\right)^{\alpha}, \quad x > m.$$

a) Mitä arvoja X saa? Laske odotusarvo $E(X)$, kun $\alpha > 1$.

b) Jos X kuvastaa tietyssä populaatiossa kuukausittaisia ansiotuloja ja $\alpha = 1.2$ sekä $m = 500$ [€], niin millä todennäköisyydellä satunnaisesti valittu populaation edustaja tienaa vähintään 4000 €?

3. Testattiin 20 komponentin kestävyyttä [MPa] vetokokeessa. Keskimääräiseksi vetokuormaksi, jolla komponentti meni rikki, saatiin 13.71 [MPa] ja keskihajonnaksi 3.55 [MPa]. Oletetaan, että vetokuorma on normaalijakautunut. Ylittääkö vetokuorma otoksen perusteella 12 [MPa]? Testaa sopivat hypoteesit riskitasolla 5%.

4. Tarkastellaan havaintoja

x	0	1	2	3	4
y	-0.5	-0.5	1.5	3.5	7

a) Piirrä havaintoja vastaava sirontakuvio. Mitä voit sanoa muuttujien välisestä riippuvuudesta sen perusteella? (2p)

b) Muodosta malli $y = a + bx^2$ havaintojen välille. Määrää kertoimet a ja b pienimmän neliösumman menetelmällä. (4p)

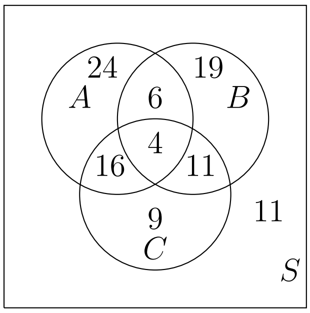
5. Määrää vakio c ja kovarianssimatriisi satunnaisvektorille (X,Y), jonka arvojoukko on $S_{X,Y} = \{(i,j) \,|\, i = 1,2,\ j = 1,2\}$ ja pistetodennäköisyydet ovat muotoa

$$p_{ij} = P(\{X = i\} \cap \{Y = j\}) = c \cdot i \cdot j \quad \text{kaikilla } (x,y) \in S_{X,Y}.$$

Tilastomatematiikka

Loppukoe 12.09.2019

1. Kuvaan 1 on merkitty eräiden tapahtumien todennäköisyyksiä *prosentteina.*



Kuva 1: Tapahtumia ja niiden todennäköisyyksiä prosentteina

Laske todennäköisyydet

$$P(A), \quad P(B \cap C), \quad P(B \cup C), \quad P(A|B), \quad P(B \cup C|A) \quad \text{ja} \quad P(A|B \cup C).$$

2. Tehdas valmistaa pultteja, joiden tavoitepituus on 6 senttiä. Oletetaan, että pulttien pituus on normaalijakautunut satunnaismuuttuja parametreilla $\mu = 6$ ja $\sigma = 0.06$.

   a) Pulttien laatua pidetään riittävänä, jos niiden pituus on toleranssirajojen 5.9 senttiä ja 6.1 senttiä välissä. Kuinka suuri osa pulteista täyttää tämän laatuvaatimuksen?

   b) Pulttien laatua halutaan parantaa säätämällä hajontaa σ. Mikä hajonnan on oltava, jotta 99 % pulteista on toleranssirajojen sisällä?

3. Matkapuhelimen akun valmiusajalle on asetettu 200 tunnin takuu. Tutkittiin 5000 akkua, joista 15 lopetti toiminnan ennen 200 tuntia. Tukevatko nämä koetulokset väitettä, että korkeintaan 0.2 prosenttia yrityksen valmistamista akuista ei kestä takuuaikaa? Muotoile sopivat hypoteesit ja testaa ne luottamustasolla 95 %.

4. Tarkastellaan bakteeriviljelmän kasvua, kun x on aika vuorokausissa viljelyksestä ja y on bakteerien lukumäärä tuhansissa. Saatiin seuraavat havainnot

x	3	6	9	12	15	18
y	115	147	239	356	579	864

   a) Piirrä suureita x ja $\ln y$ vastaava sirontakuvio. Mitä voit sanoa muuttujien välisestä riippuvuudesta sen perusteella? (2p)

   b) Muodosta malli $y = a \cdot \mathrm{e}^{bx}$ havaintojen välille. Määrää kertoimet a ja b soveltamalla pienimmän neliösumman menetelmää suureisiin x ja $\ln y$. (4p)

5. Anna esimerkki kaksiulotteisesta jakaumasta ja laske muuttujien välinen kovarianssi antamallesi jakaumalle.

Tilastomatematiikka

Loppukoe 4.11.2019

1. Olkoon $S = \{a, b, c, d, e, f, g, h\}$ otosavaruus sekä $A = \{a, c, e, g\}$ ja $B = \{a, b, c, d, e\}$ tapahtumia. Laske todennäköisyydet

$$P(A \cap B), \quad P(A|B) \quad \text{ja} \quad P(A|\overline{B}),$$

kun oletetaan, että kaikki alkeistapahtumat ovat yhtä todennäköisiä.

2. a) Olkoon $X \sim N(100, 10)$ normaalijakautunut satunnaismuuttuja. Laske todennäköisyys $P(X > 85)$.

   b) Olkoon $X \sim N(100, \sigma)$ normaalijakautunut satunnaismuuttuja. Millä hajonnan σ arvolla $P(X > 85) = 0.8$?

3. Tutkittiin katalyytin vaikutusta erään kemiallisen prosessin saantoon ja saatiin seuraavat havainnot prosessin saannolle

Katalyytti 1	91.50	94.18	92.18	95.39	91.79	89.07	94.72	89.21
Katalyytti 2	89.19	90.95	90.46	93.21	97.19	97.04	91.07	92.75

Muotoile sopivat hypoteesit ja testaa ne riskitasolla $\alpha = 0.05$, kun halutaan tutkia, onko katalyytillä vaikutusta prosessin keskimääräiseen saantoon.

4. Eräällä kurssilla saatiin seuraava arvosanajakauma

x	0	1	2	3	4	5
y	49	28	18	10	8	7

missä y on arvosanan x saaneiden opiskelijoiden lukumäärä.

   a) Piirrä suureita x ja $\ln y$ vastaava sirontakuvio. Mitä voit sanoa muuttujien välisestä riippuvuudesta sen perusteella? (2p)

   b) Muodosta malli $y = Ce^{kx}$ havaintojen välille. Määrää kertoimet C ja k soveltamalla pienimmän neliösumman menetelmää suureisiin x ja $\ln y$. (4p)

5. Anna esimerkki kaksiulotteisesta satunnaismuuttujasta $\mathbf{X} = (X, Y)$, jolle

   a) X ja Y ovat riippumattomat,

   b) X ja Y eivät ole riippumattomat.

Kaavoja

$$
\begin{aligned}
&\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)\\
&\mathrm{P}(A \backslash B) = \mathrm{P}(A \cap \bar{B}) = \mathrm{P}(A) - \mathrm{P}(A \cap B)\\
&\mathrm{P}(\bar{A}) = 1 - \mathrm{P}(A)\\
&\mathrm{P}(A|B) = \mathrm{P}(A \cap B)/\mathrm{P}(B)\\
&f(x) = 1/(b-a),\ \mathrm{E}(X) = (a+b)/2,\ \mathrm{Var}(X) = (b-a)^2/12\\
&f(x) = \theta e^{-\theta x},\ \mathrm{E}(X) = 1/\theta,\ \mathrm{Var}(X) = 1/\theta^2\\
&\mathrm{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k},\ \mathrm{E}(X) = np,\ \mathrm{Var}(X) = np(1-p)\\
&\mathrm{P}(X=k) = p(1-p)^{k-1},\ \mathrm{E}(X) = 1/p,\ \mathrm{Var}(X) = (1-p)/p^2\\
&\mathrm{P}(X=k) = \frac{a^k}{k!} e^{-a},\ \mathrm{E}(X) = a,\ \mathrm{Var}(X) = a\\
&\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \text{ (likimain, kun ''}n\text{ on suuri'')}\\
&\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}\\
&\frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{1}{n} + \frac{1}{m}} \sqrt{\frac{(n-1)S_x^2 + (m-1)S_y^2}{n+m-2}}} \sim t_{n+m-2};\\
&r = \frac{s_{xy}}{\sqrt{s_{xx}}\sqrt{s_{yy}}};\ s_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y});\ s_{xx} = s_x^2;\\
&y = a + bx;\quad b = \frac{s_{xy}}{s_{xx}};\quad a = \overline{y} - b\overline{x}\\
&s_r^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - a - bx_i)^2 = \frac{n-1}{n-2}(1-r^2)s_{yy}\\
&\sqrt{n-1}\, s_x \frac{\frac{S_{xY}}{s_{xx}} - \beta}{S_r} \sim t_{n-2};\\
&\mathrm{E}(aX + bY) = a\mathrm{E}(X) + b\mathrm{E}(Y);\quad \mathrm{Var}(X) = \mathrm{E}((X - \mathrm{E}(X))^2);\\
&\sigma_{XY} = \mathrm{Cov}(X,Y) = \mathrm{E}((X - \mathrm{E}(X))(Y - \mathrm{E}(Y))),\quad \sigma_{XX} = \sigma_X^2;\\
&\qquad = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)\\
&\rho(X,Y) = \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y}.
\end{aligned}
$$