Sovellettu todennäköisyyslaskenta B

Antti Rasila

4. lokakuuta 2007

1. Moniulotteiset todennäköisyysjakaumat
   - Johdanto
   - Kaksiulotteiset satunnaismuuttujat
   - Kaksiulotteisen diskreetin jakauman pistetodennäköisyysfunktio
   - Kaksiulotteisen jatkuvan jakauman tiheysfunktio
   - Kaksiulotteisten jakaumien kertymäfunktiot
   - Kaksiulotteisten jakaumien reunajakaumat

2. Satunnaismuuttujien riippumattomuus

3. Kaksiulotteisen jakauman yleinen odotusarvo
   - Reunajakaumien odotusarvot

Johdanto

- Yksi satunnaismuuttuja ei riitä kuvaamaan kaikkia satunnaisilmiöitä.
- Jos ilmiöön liittyy useita satunnaisia tekijöitä, ovat näiden tekijöiden väliset riippuvuudet tilastollisen mallintamisen kannalta erityisen mielenkiintoisia.
- Näiden riippuvuuksien mallintaminen tapahtuu yhteisjakauman avulla.
- **Esimerkki 1:** Kahden nopan heiton yhteisjakauma.
- **Esimerkki 2:** Miten tupakoinnin määrä (1) ja kesto (2) vaikuttuvat keuhkosyöpään (3).

Kaksiulotteiset satunnaismuuttujat

- Olkoot X ja Y satunnaismuuttujia, joiden otosavaruudet ovat R ja S.
- Toisin sanoen:

$$X\colon R \to \mathbb{R}$$
$$Y\colon S \to \mathbb{R}$$

- Olkoon $R \times S$ otosavaruuksien R ja S karteesinen tulo:

$$R \times S = \{(r, s) \,|\, r \in R, s \in S\}$$

- Satunnaismuuttujien X ja Y järjestetty pari (X, Y) määrittelee kaksiulotteisen satunnaismuuttujan:

$$(X, Y) : R \times S \to \mathbb{R}^2$$

Kaksiulotteisen diskreetin jakauman pistetodennäköisyysfunktio

- Reaaliarvoinen funktio

$$f_{XY} : \mathbb{R}^2 \longrightarrow \mathbb{R}$$

määrittelee diskreettien satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyysfunktion, jos

1. $f_{XY}(x, y) \geq 0$ kaikilla x, y,
2. $\sum_x \sum_y f_{XY}(x, y) = 1$,
3. $\Pr(X = x \cap Y = y) = f_{XY}(x, y)$,

- Olkoon $A \subset \mathbb{R}^2$ jokin tapahtuma. Tällöin

$$\Pr((X, Y) \in A) = \sum_{(x,y) \in A} \sum f_{XY}(x, y)$$

Esimerkki

- Heitetään (virheetöntä) kolikkoa kolme kertaa. Olkoon X saatujen kruunien kokonaismäärä ja Y kahdessa ensimmäisessä heitossa saatujen klaavojen määrä.
- Nyt X saa jonkin arvoista $0, 1, 2, 3$ ja Y jonkin arvoista $0, 1, 2$.
- Karteesinen tulo $S_{XY} = S_X \times S_Y$ on joukko

$$S_{XY} = \{(0,0), (0,1), \dots, (3,1), (3,2)\}.$$

- Joukon S_{XY} pistetodennäköisyydet voidaan helposti laskea, ja ne on käytännöllistä esittää kaksiulotteisena taulukkona.

Kaksiulotteisen jatkuvan jakauman tiheysfunktio

- Reaaliarvoinen jatkuva funktio

$$f_{XY} : \mathbb{R}^2 \to \mathbb{R}$$

määrittelee jatkuvien satunnaismuuttujien X ja Y yhteisjakauman tiheysfunktion, jos

1. $f_{XY}(x, y) \geq 0$ kaikilla x, y,
2. $$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{XY}(x, y)\, dy\, dx = 1,$$
3. $$\Pr(a \leq X \leq b \cap c \leq Y \leq d) = \int_a^b \int_c^d f_{XY}(x, y)\, dy\, dx$$

- Olkoon $A \subset \mathbb{R}^2$ jokin tapahtuma. Tällöin

$$\Pr((X, Y) \in A) = \iint_A f_{XY}(x, y)\, dy\, dx$$

Kaksiulotteisten jakaumien kertymäfunktiot

- Satunnaismuuttujien X ja Y yhteisjakauman kertymäfunktio F_{XY} määritellään seuraavasti:

$$F_{XY}(x, y) = \Pr(X \leq x \cap Y \leq y)$$

- Diskreetille jakaumalle:

$$F_{XY}(x, y) = \Pr(X \leq x \cap Y \leq y) = \sum_{x_i \leq x} \sum_{y_i \leq y} f_{XY}(x_i, y_i)$$

- Jatkuvalle jakaumalle:

$$F_{XY}(x, y) = \Pr(X \leq x \cap Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(u, v)\, dv\, du$$

Kaksiulotteisten jakaumien reunajakaumat

- Yksittäisten satunnaismuuttujien X ja Y jakaumia sanotaan yhteisjakauman *reunajakaumiksi*.
- Diskreetissä tapauksessa satunnaismuuttujien X ja Y reunajakaumien pistetodennäköisyysfunktiot saadaan seuraavasti:

$$f_X(x) = \Pr(X = x) = \sum_y f_{XY}(x, y)$$

$$f_Y(y) = \Pr(Y = y) = \sum_x f_{XY}(x, y)$$

- Vastaavasti jatkuvassa tapauksessa satunnaismuuttujien X ja Y reunajakaumien tiheysfunktiot saadaan seuraavasti:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y)\, dy$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x, y)\, dx$$

Satunnaismuuttujien riippumattomuus, johdanto

- Olkoot satunnaismuuttujien X ja Y kertymäfunktiot $F_X(x) = Pr(X \leq x)$ ja $F_Y(y) = Pr(Y \leq y)$.
- Merkitään $A = \{X \leq x\}$ ja $B = \{Y \leq y\}$.
- Nyt $F_X(x) = Pr(A)$ ja $F_Y(y) = Pr(B)$, sekä

$$A \cap B = \{s \in S \,|\, X(s) \leq x \wedge Y(s) \leq y\}.$$

- Siten $F_{XY}(x, y) = Pr(A \cap B)$. Tapahtumat A ja B ovat riippumattomia, jos $Pr(A \cap B) = Pr(A)Pr(B)$.
- Tällöin

$$F_{XY}(x, y) = Pr(A \cap B) = Pr(A)Pr(B) = F_X(x)F_Y(y).$$

- **Määritelmä:** Satunnaismuuttujat X ja Y ovat riippumattomia, jos ja vain jos

$$F_{XY}(x, y) = F_X(x)F_Y(y)$$

kaikilla $x, y \in \mathbb{R}$.

Satunnaismuuttujien riippumattomuus (yleisesti)

- Kaksi satunnaismuuttujaa ovat riippumattomia, jos ja vain jos seuraavat yhtäpitävät ehdot toteutuvat:
  1. $f_{XY}(x, y) = f_X(x) f_Y(y)$
  2. $F_{XY}(x, y) = F_X(x) F_Y(y)$
- Tässä $F_X(x)$ ja $F_Y(y)$ ovat reunajakaumien kertymäfunktiot ja f_X, f_Y vastaavasti tiheysfunktiot.
- Yleisesti satunnaismuuttujille $X_1, X_2, \ldots, X_n$, joiden yhteisjakauman pistetodennäköisyys- tai tiheysfunktio on $f(x_1, x_2, \ldots, x_n)$ ja reunajakaumien pistetodennäköisyys- tai tiheysfunktio ovat $f_i(x_i), i = 1, 2, \ldots, n$ (sekä vastaavat kertymäfunktiot merkittynä F_i:llä):
  1. $f(x_1, x_2, \ldots, x_n) = f_1(x_1) f_2(x_2) \cdot \cdots \cdot f_n(x_n)$
  2. $F(x_1, x_2, \ldots, x_n) = F_1(x_1) F_2(x_2) \cdot \cdots \cdot F_n(x_n)$

Esimerkki (Laininen) 1/2

- Pyöreä tappi ja reikä tehdään koneellisesti toisistaan riippumatta. Tapin halkaisija X on tasaisesti jakautunut välille $(20, 21)$ millimetriä ja reiän halkaisija Y välille $(20.5, 22)$ mm. Laske todennäköisyys, että tappi ei mahdu reikään.
- Tiheysfunktiot ovat:

$$f_X(x) = \begin{cases} 1, & \text{kun } 20 < x < 21, \\ 0, & \text{muuten,} \end{cases}$$

$$f_X(x) = \begin{cases} 1/1.5, & \text{kun } 20.5 < x < 22, \\ 0, & \text{muuten.} \end{cases}$$

- Satunnaismuuttujat X ja Y ovat riippumattomia, joten

$$f_{XY}(x, y) = f_X(x) f_Y(y) = 1/1.5, \text{ kun } 20 < x < 21 \text{ ja } 20.5 < y < 22.$$

Muualla tiheysfunktio on nolla.

Esimerkki (Laininen) 2/2

- Tappi ei mahdu reikään, jos $X > Y$.
- Saadaan

$$
\begin{aligned}
Pr(X > Y) &= \int_{20.5}^{21} \int_{20.5}^{x} f_{XY}(x, y)\, dy\, dx \\
&= \int_{20.5}^{21} \int_{20.5}^{x} \frac{1}{1.5}\, dy\, dx \\
&= \frac{2}{3} \int_{20.5}^{21} (x - 20.5)\, dx = 0.083.
\end{aligned}
$$

Esimerkki (Niemi) 1/2

- Olkoon $f_{XY}(x, y) = |xy|e^{-x^2-y^2}$.
- Funktio toteuttaa tiheysfunktiolta vaadittavat ehdot (1) ja (2), joten voidaan ajatella, että se on joidenkin jatkuvien satunnaismuuttujien X, Y yhteisjakauman tiheysfunktio.
- Reunajakaumat ovat

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} |xy|e^{-x^2-y^2}\, dy = |x|e^{-x^2} \int_{-\infty}^{\infty} |y|e^{-y^2}\, dy \\
&= -|x|e^{-x^2} \int_{-\infty}^{0} ye^{-y^2}\, dy + |x|e^{-x^2} \int_{0}^{\infty} ye^{-y^2}\, dy \\
&= \frac{1}{2}|x|e^{-x^2} + \frac{1}{2}|x|e^{-x^2} = |x|e^{-x^2}.
\end{aligned}
$$

Esimerkki (Niemi) 2/2

- Vastaavasti saadaan $f_Y(y) = |y|e^{-y^2}$.
- Koska

$$f_{XY}(x,y) = |xy|e^{-x^2-y^2} = |x|e^{-x^2}|y|e^{-y^2} = f_X(x)f_Y(y),$$

satunnaismuuttujat X, Y ovat riippumattomia.

Kaksiulotteisen jakauman yleinen odotusarvo

- Olkoon $g : \mathbb{R}^2 \to \mathbb{R}$ jatkuva funktio.
- Tällöin diskreetissä tapauksessa:

$$\mathrm{E}(g(X,Y)) = \sum_x \sum_y g(x,y) f_{XY}(x,y)$$

- Vastaavasti jatkuvassa tapauksessa:

$$\mathrm{E}(g(X,Y)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y) f_{XY}(x,y)\, dy\, dx$$

Reunajakaumien odotusarvot

- Diskreetissä tapauksessa:

$$\mathrm{E}(X) = \sum_x \sum_y x f_{XY}(x,y) = \sum_x x \sum_y f_{XY}(x,y) = \sum_x x f_X(x)$$

$$\mathrm{E}(Y) = \sum_x \sum_y y f_{XY}(x,y) = \sum_y y \sum_x f_{XY}(x,y) = \sum_y y f_Y(y)$$

- Vastaavasti jatkuvassa tapauksessa:

$$\begin{aligned}\mathrm{E}(X) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f_{XY}(x,y)\, dy\, dx \\ &= \int_{-\infty}^{+\infty} x \int_{-\infty}^{+\infty} f_{XY}(x,y)\, dy\, dx = \int_{-\infty}^{+\infty} x f_X(x)\, dx\end{aligned}$$