

Insinöörimatematiikka: Todennäköisyyslaskenta

Demonstraatio 4, 7.2.2023

1. Oletetaan tunnetuksi, että riippumattomille satunnaismuuttujille X ja Y pätee $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. Miten tästä seuraa, että riippumattomille satunnaismuuttujille $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$?

Mallivastaus:

$$\begin{aligned}\text{Var}(X + Y) &= \mathbb{E}((X + Y)^2) - \mathbb{E}(X + Y)^2 \\ &= \mathbb{E}(X^2 + 2XY + Y^2) - (\mathbb{E}(X) + \mathbb{E}(Y))^2 \\ &= \mathbb{E}(X^2) + 2\mathbb{E}(XY) + \mathbb{E}(Y^2) - (\mathbb{E}(X)^2 + 2\mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(Y)^2) \\ &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 + \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 \\ &= \text{Var}(X) + \text{Var}(Y)\end{aligned}$$

2. Yritys valmistaa Led-lamppuja, joiden kestoikä on likimain normaalisti jakautunut, odotusarvon ollessa $\mu = 25000$ ja keskihajonnan $\sigma = 4000$ (aikayksikkönä tunti).

- a) Millä todennäköisyydellä lamppu kestää yli 28000 tuntia?
- b) Millä todennäköisyydellä lamppu kestää alle 20000 tuntia?
- c) Mikä on kesto aika, jonka ylittää vain 1% lampuista?

Ohje: Matlabissa standardoidun normaalijakauman kertymäfunktion käänteisfunktio saadaan muodossa

`norminv(p)`.

Mallivastaus: Olkoon T kesto aikaa edustava satunnaismuuttuja.

- a)

$$\begin{aligned}\mathbb{P}(T \geq 28000) &= 1 - \mathbb{P}(T < 28000) = 1 - \Phi\left(\frac{28000 - 25000}{4000}\right) \\ &= 1 - 0.773373 \dots = 0.226627 \dots\end{aligned}$$

- b)

$$\mathbb{P}(T < 20000) = \Phi\left(\frac{20000 - 25000}{4000}\right) = 0.10565 \dots$$

- c) On selvitettävä sellainen luku t , että

$$\begin{aligned}\mathbb{P}(T > t) = 0.01 &\Leftrightarrow 1 - \Phi\left(\frac{t - 25000}{4000}\right) = 0.01 \Leftrightarrow \Phi\left(\frac{t - 25000}{4000}\right) = 0.99 \\ &\Leftrightarrow \frac{t - 25000}{4000} = \Phi^{-1}(0.99) \\ &\Leftrightarrow t = 25000 + 4000\Phi^{-1}(0.99) = 34305.4 \dots\end{aligned}$$

3. Merkitään klaavaa 1:llä ja kruunaa 2:lla painottamattoman kolikon heitossa. Kun kolikkoa heitetään n kertaa. Olkoon X_i i :nnessä heiton tulos (joko 1 tai 2) sekä $\bar{X}_n$ kaikkien heittojen tulosten otoskeskiarvo.

Käytä suurten lukujen lakia selvittääksesi kuinka monta heittoa n tarvitaan, että $\mathbb{P}(|\bar{X}_n - \mathbb{E}(\bar{X}_n)| < 0.01) \geq 0.99$

Mallivastaus: Odotusarvo $\mathbb{E}(X_i) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 2 = \frac{3}{2}$, $\mathbb{E}(X_i^2) = \frac{1}{2} \cdot 1^2 + \frac{1}{2} \cdot 2^2 = \frac{5}{2}$, ja $\text{Var}(X_i) = \mathbb{E}(X_i^2) - \mathbb{E}(X_i)^2 = \frac{5}{2} - \left(\frac{3}{2}\right)^2 = \frac{1}{4}$.

Tällöin $\mathbb{E}(\overline{X}_n) = \frac{3}{2}$ ja suurten lukujen lain mukaan

$$\mathbb{P}\left(\left|\overline{X}_n - \frac{3}{2}\right| < 0.01\right) \geq 1 - \frac{1/4}{n \cdot 0.01^2}.$$

Tällöin ratkaistavaksi jää epäyhtälö

$$1 - \frac{1/4}{n \cdot 0.01^2} \geq 0.99$$

$$\Leftrightarrow n \geq \frac{1/4}{0.01^3} = 250000.$$

4. Hyllyissä on tilaa kirjoille yhteensä 150 m. Yhden kirjan paksuus on keskimäärin $\mu = 3$ cm ja paksuuden keskihajonta $\sigma = 0,8$ cm. Arvioi millä todennäköisyydellä 5200 kirjaa ei mahdu hyllyihin.

Ohje: Jos X_i on kirjan i paksuus, on yhteispaksuus $X_1 + \dots + X_{5200}$. Käytä keskeistä raja-arvolauseetta.

Mallivastaus: Jos merkitään $P = X_1 + \dots + X_{5200}$ (yhteispaksuus, on $\mathbb{E}(P) = 5200 \cdot 0.03 = 156$ ja $\text{Var}(P) = 5200 \cdot 0.008^2 = 0.3328$. Tällöin

$$\mathbb{P}(P > 150) = 1 - \mathbb{P}(P \leq 150) \approx 1 - \Phi\left(\frac{150 - 0.03 \cdot 5200}{0.008\sqrt{5200}}\right) = 1 - 1.23151 \cdot 10^{-25} \approx 1.$$

5. Kun heitetään noppaa 1000 kertaa, saadaan 180 kertaa silmäluku 6. Arvioi tällaisen tapahtuman todennäköisyyttä, mikäli noppa on painottomaton.

Ohje: Voit esimerkiksi määritellä satunnaismuuttujan X joka saa kaksi arvoa: 0, jos silmäluku on ≤ 5 ja 1, jos silmäluku on 6 ja soveltaa keskeistä raja-arvolauseetta.

Vaihtoehtoisesti todennäköisyyttä voi arvioida binomijakaumaa käyttäen.

Mallivastaus: Tehtävässä on helpointa käyttää binomitodennäköisyyttä. Jos satunnaismuuttuja S_6 merkitsee silmälukujen 6 määrää, on

$$\mathbb{P}(S_6 = 180) = \binom{1000}{180} \left(\frac{1}{6}\right)^{180} \left(\frac{5}{6}\right)^{1000-180} = 0.0175348 \dots$$

Mikäli halutaan käyttää keskeistä raja-arvolauseetta ja normaalijakaumaa, voidaan merkitä $S = X_1 + \dots + X_{1000}$, laskea $\mathbb{E}(X_i) = \frac{1}{6} \cdot 1 + \frac{5}{6} \cdot 0 = \frac{1}{6}$, $\mathbb{E}(X_i^2) = \frac{1}{6} \cdot 1^2 + \frac{5}{6} \cdot 0^2 = \frac{1}{6}$, $\text{Var}(X_i) = \mathbb{E}(X_i^2) - \mathbb{E}(X_i)^2 = \frac{1}{6} - \left(\frac{1}{6}\right)^2 = \frac{5}{36}$. Jos silmälukuja 6 on tasan 180, saa S arvon $180 \cdot 1 + (1000 - 180) \cdot 0 = 180$.

Tällöin tulee arvioitavasi $\mathbb{P}(S = 180)$, mutta mikäli tätä tahdotaan arvioida jatkuvalla jakaumalla, on tehtävä kokonaislukukorjaus ja arvioitava todennäköisyyttä

$$\begin{aligned} \mathbb{P}(179.5 < S < 180.5) &= \mathbb{P}(S < 180.5) - \mathbb{P}(S < 179.5) \\ &\approx \Phi\left(\frac{180.5 - \frac{5}{36} \cdot 1000}{\sqrt{\frac{5}{36} \cdot 1000}}\right) - \Phi\left(\frac{179.5 - \frac{5}{36} \cdot 1000}{\sqrt{\frac{5}{36} \cdot 1000}}\right) \\ &= 0.0178511 \dots \end{aligned}$$

6. Olkoon X normaalijakautuman $N(\mu, \sigma^2)$ mukainen satunnaismuuttuja. Määritä todennäköisyydet $\mathbb{P}(|X - \mu| > 3\sigma)$ ja $\mathbb{P}(|X - \mu| > 5\sigma)$.

Mallivastaus: Voidaan tarkastella komplementaarista todennäköisyyttä:

$$\begin{aligned}\mathbb{P}(|X - \mu| \leq 3\sigma) &= \mathbb{P}(-3\sigma \leq X - \mu \leq 3\sigma) \\ &= \mathbb{P}(-3\sigma + \mu \leq x \leq 3\sigma + \mu) = F(3\sigma + \mu) - F(-3\sigma + \mu) \\ &= \Phi\left(\frac{3\sigma + \mu - \mu}{\sigma}\right) - \Phi\left(\frac{-3\sigma + \mu - \mu}{\sigma}\right) \\ &= \Phi(X < 3) - \Phi(X < -3) = 0.9973002039 \dots\end{aligned}$$

Yllä F tarkoittaa normaalijakauman $N(\mu, \sigma^2)$ kertymäfunktia. Tällöin

$$\mathbb{P}(|X - \mu| > 3\sigma) = 1 - \mathbb{P}(|X - \mu| \leq 3\sigma) = 0.002699796063 \dots$$

Samoin

$$\mathbb{P}(|X - \mu| \leq 5\sigma) = \Phi(5) - \Phi(-5) = 0.9999994267 \dots,$$

josta

$$\mathbb{P}(|X - \mu| > 5\sigma) = 1 - \mathbb{P}(|X - \mu| \leq 5\sigma) = 0.0000005733031438 \dots \approx 5.7 \cdot 10^{-7}$$

7. Satunnaismuuttuja X saa arvot 1, 2, 3 ja 4 todennäköisyyksillä $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$ ja $\frac{1}{8}$. Laske entropia $H(X)$ (yksikkönä bitti)

Mallivastaus:

$$H(X) = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{4} \log_2 \frac{1}{4} - \frac{1}{8} \log_2 \frac{1}{8} - \frac{1}{8} \log_2 \frac{1}{8} = \frac{7}{4}$$

8. Satunnaismuuttujien X ja Y mahdolliset arvot ovat joukossa $\{1, 2, 3, 4\}$ ja näillä on seuraavan taulukon mukainen yhteisjakauma:

$\begin{matrix} \backslash \\ x \end{matrix}$	1	2	3	4
1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	0
2	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$
3	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
4	$\frac{1}{4}$	0	0	0

Määritä todennäköisyydet $\mathbb{P}(X = i)$ ja $\mathbb{P}(Y = i)$, kun $i \in \{1, 2, 3, 4\}$ ja laske entropiat $H(X)$, $H(Y)$, $H(X | Y)$ sekä $I(X | Y)$. Ohje: Esim. Todennäköisyys slle, että Y saa arvon 1 lasketaan 1. rivin todennäköisyyksien summana (miksi?)

Mallivastaus: $\mathbb{P}(X = i)$ saadaan i :nnen sarakkeen todennäköisyyksien summana ja $\mathbb{P}(Y = i)$ vastaavasti rivisummina. Tämä seuraa esimerkiksi siitä miten kokonaistodennäköisyys lasketaan. Esim.

$$\begin{aligned}\mathbb{P}(X = 1) &= \mathbb{P}(X = 1 | Y = 1)\mathbb{P}(Y = 1) + \mathbb{P}(X = 1 | Y = 2)\mathbb{P}(Y = 2) \\ &+ \mathbb{P}(X = 1 | Y = 3)\mathbb{P}(Y = 3) + \mathbb{P}(X = 1 | Y = 4)\mathbb{P}(Y = 4) \\ &= \mathbb{P}(X = 1 \wedge Y = 1) + \mathbb{P}(X = 1 \wedge Y = 2) \\ &+ \mathbb{P}(X = 1 \wedge Y = 3) + \mathbb{P}(X = 1 \wedge Y = 4)\end{aligned}$$

Näin ollen todennäköisyydet muuttujan X arvoille 1, 2, 3 ja 4 ovat $\frac{1}{2}$, $\frac{1}{4}$, $\frac{5}{32}$ ja $\frac{3}{32}$. Vastaavasti Y todennäköisyydet saada arvoja 1, 2, 3 ja 4 ovat $\frac{1}{4}$, $\frac{1}{4}$, $\frac{1}{4}$ ja $\frac{1}{4}$.

Täten

$$H(X) = -\left(\frac{1}{2} \log_2 \frac{1}{2} + \frac{1}{4} \log_2 \frac{1}{4} + \frac{5}{32} \log_2 \frac{5}{32} + \frac{3}{32} \log_2 \frac{3}{32}\right) = 1.73861 \dots,$$

$$H(Y) = -\left(\frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{4} \log_2 \frac{1}{4}\right) = 2.$$

Ehdollinen entropia $H(X | Y)$ voidaan laskea kaavan $H(X, Y) - H(Y)$ mukaisesti, kunhan ensin lasketaan

$$\begin{aligned} H(X, Y) &= -\left(\frac{1}{8} \log_2 \frac{1}{8} + \frac{1}{16} \log_2 \frac{1}{16} + \dots \right. \\ &\quad + \frac{1}{16} \log_2 \frac{1}{16} + \frac{1}{8} \log_2 \frac{1}{8} + \dots \\ &\quad + \frac{1}{16} \log_2 \frac{1}{16} + \frac{1}{16} \log_2 \frac{1}{16} + \dots \\ &\quad \left. + \frac{1}{4} \log_2 \frac{1}{4} + \dots\right) = \frac{53}{16} = 3.3125 \end{aligned}$$

Tästä saadaan $H(X | Y) = H(X, Y) - H(Y) = 1.3125 \dots$ ja edelleen $I(X | Y) = H(X) - H(X | Y) = 0.426109 \dots$

9. Jääkiekkjoukkue voittaa kotona pelatessaan todennäköisyydellä $\frac{3}{5}$, mutta vierasjoukkueena pelatessaan vain todennäköisyydellä $\frac{1}{3}$. Pelipaikka on todennäköisyydellä $\frac{1}{4}$ kotihalli ja $\frac{3}{4}$ vieras halli.

Millä todennäköisyydellä joukkue voittaa pelin? Olkoon X satunnaismuuttuja, joka saa arvot {voitto, häviö} ja Y satunnaismuuttuja, joka saa arvot {koti, vieras}. Laske $I(X | Y)$.

Mallivastaus: Voiton todennäköisyys saadaan kokonaistodennäköisyys laske-
malla:

$$\mathbb{P}(\text{voitto}) = \mathbb{P}(\text{voitto} | \text{koti})\mathbb{P}(\text{koti}) + \mathbb{P}(\text{voitto} | \text{vieras})\mathbb{P}(\text{vieras}) = \frac{3}{5} \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{3}{4} = \frac{2}{5}.$$

Nyt siis

$$H(X) = -\left(\frac{2}{5} \log_2 \frac{2}{5} + \frac{3}{5} \log_2 \frac{3}{5}\right) = 0.970951 \dots,$$

$$H(X | \text{koti}) = -\left(\frac{3}{5} \log_2 \frac{3}{5} + \frac{2}{5} \log_2 \frac{2}{5}\right) = 0.970951 \dots$$

ja

$$H(X | \text{vieras}) = -\left(\frac{1}{3} \log_2 \frac{1}{3} + \frac{2}{3} \log_2 \frac{2}{3}\right) = 0.918296.$$

Edellisten odotusarvona saadaan

$$H(X | Y) = \frac{1}{4} H(X | \text{koti}) + \frac{3}{4} H(X | \text{vieras}) = 0.93146 \dots,$$

ja siis

$$I(X | Y) = H(X) - H(X | Y) = 0.0394911 \dots$$