Aalto-yliopisto

MS-A0502 Todennäköisyyslaskennan ja tilastotieteen peruskurssi — Gripenberg, Kiiski
Harjoitus 2
3–7.11.2014, viikko 45

P1. Keppi, jonka pituus on 4 m, taitetaan kahtia täysin satunnaisesti valitusta kohdasta ja muodostetaan suorakulmainen kolmio, jonka kateetteina ovat syntyneet palaset. Kolmion pinta-ala on satunnaismuuttuja A.

(a) Määritä A:n odotusarvo.
(b) Onko hyvä idea laskea pinta-alan A odotusarvon approksimaatio lähtemällä oletuksesta, että keppi katkeaa katkaisukohdan odotusarvon kohdalla?

Ratkaisu: Jos keppi katkaistaan kohdasta X niin $X \sim \mathrm{Uniform}(0,4)$ ja $A = \frac{1}{2}X(4-X)$. Satunnaismuuttujan X tiheysfunktio on tietenkin $f_X(x) = \frac{1}{4}$ kun $x \in [0,4]$ ja 0 muuten.

(a) Ei ole hyvä idea lähteä laskemaan satunnaismuuttujan A kertymäfunktiota tai tiheysfunktiota koska suoraan saadaan

$$\mathrm{E}(A) = \mathrm{E}\big(\tfrac{1}{2}X(4-X)\big) = 2\mathrm{E}(X) - \tfrac{1}{2}\mathrm{E}(X^2) = 2\int_0^4 \tfrac{1}{4}x\,\mathrm{d}x - \tfrac{1}{2}\int_0^4 x^2\tfrac{1}{4}\,\mathrm{d}x$$
$$= \frac{2}{8}\Big/_0^4 x^2 - \frac{1}{24}\Big/_0^4 x^3 = 4 - \frac{8}{3} = \frac{4}{3}.$$

(b) Nyt $\mathrm{E}(X) = 2$ ja arvolla 2 saadaan pinta-alaksi $\frac{1}{2}\cdot 2\cdot(4-2) = 2$, joka itse asiassa on pinta-alan ääriarvo, joten on erittäin huono idea ottaa odotusarvo katkaisukohdaksi.

P2. Lähdet lomalle Ologaan, jossa aurinkoisen päivän todennäköisyys on p ja lisäksi olet päättänyt että tulet heti kotiin kun olet viettänyt siellä $r \geq 1$ aurinkoista päivää. Johda lauseke todennäköisyydelle, että vietät Ologassa n päivää (ja tämä todennäköisyys on siis negatiivisen binomijakauman pistetodennäköisyysfunktion arvo pisteessä n parametreillä p ja r eli $f_{\mathrm{NegBin}(p,r)}(n)$) seuraavan päättelyn kautta olettaen, että tapahtumat *"Päivä on aurinkoinen"* ovat riippumattomia:

- Olkoon A tapahtuma *"Vietät Ologassa n päivää"*, B tapahtuma *"n:s päivä on aurinkoinen"* ja C tapahtuma *"$n-1$ ensimmäisten päivien aikana on ollut $r-1$ aurinkoista päivää"*. Mitkä näistä tapahtumista ovat toisistaan riippumattomia?
- Esitä tapahtuma A tapahtumien B ja C avulla.
- Määritä $\Pr(B)$ ja $\Pr(C)$.
- Määritä $\Pr(A)$.

Ratkaisu: (a) Tapahtumat B ja C ovat riippumattomia tehdyn oletuksen nojalla mutta tapahtumat A ja C tai A ja B eivät ole riippumattomia.

(b) $A = B \cap C$ eli jos n:s päivä on aurinkoinen ja edellisten päivien aikana on ollut $r-1$ aurinkoista päivää, niin päivän n jälkeen lähdet kotiin.

(c) Oletuksen mukaan $\Pr(B) = p$. Koska tapahtumat *"päivä on aurinkoinen"* ovat riippumattomia ja esiintyvät todennäköisyydellä p niin $\Pr(C)$ saadaan binomijakaumasta ja on $\binom{n-1}{r-1}p^{r-1}(1-p)^{n-1-r+1}$.

(d) Koska B ja C ovat riippumattomia ja $A = B \cap C$ niin

$$\Pr(A) = \Pr(B\cap C) = \Pr(B)\Pr(C) = \binom{n-1}{r-1}p^{r-1}(1-p)^{n-1-r+1}p = \binom{n-1}{r-1}p^r(1-p)^{n-r}.$$

P3. Kalasaalissa kalojen paino on $\mathrm{N}(\mu, \sigma^2)$-jakautunut. Kaikki kalat joiden paino on korkeintaan a syötetään kissoille. Määritä jäljelle jäävien kalojen painojakauman mediaani ja esitä vastuksesi parametrien μ, σ, a ja $\mathrm{N}(0,1)$-jakuman kertymäfunktion $F_{\mathrm{N}(0,1)}$ ja sen käänteisfunktion avulla.
Huom! *Koska kysytään mediaania, sinun ei tarvitse määrittää jäljelle jäävien kalojen jakauman tiheysfunktiota, riittää, että muistat, että puolet jäljelle jäävistä kaloista ovat mediaania painavampia ja puolet kevyempiä.*

Ratkaisu: Jos X on kalan paino, niin oletuksen mukaan $X \sim \mathrm{N}(\mu, \sigma^2)$ jolloin $\dfrac{X-\mu}{\sigma} \sim \mathrm{N}(0,1)$. Koska $X \leq a$ täsmälleen silloin kun $\dfrac{X-\mu}{\sigma} \leq \dfrac{a-\mu}{\sigma}$ niin todennäköisyys, että kala painaa vähemmän kuin a on kertymäfunktion määritelmän nojalla

$$F_{\mathrm{N}(0,1)}\left(\frac{a-\mu}{\sigma}\right).$$

Jos jäljelle jäävien kalojen painojen mediaani on b niin mediaanin määritelmän (ja X:n jatkuvuuden) nojalla tämä tarkoittaa sitä, että

$$\Pr(a < X < b) = \Pr(X > b).$$

Tästä seuraa, että

$$\Pr(X > b) = \frac{1}{2}\Pr(X > a) = \frac{1}{2}(1 - \Pr(X \leq a)),$$

jolloin

$$\Pr(X \leq b) = 1 - \frac{1}{2}(1 - \Pr(X \leq a)) = \frac{1}{2}(1 + \Pr(X \leq a)).$$

Koska $X \leq b$ täsmälleen silloin kun $\dfrac{X-\mu}{\sigma} \leq \dfrac{b-\mu}{\sigma}$ niin

$$\frac{b-\mu}{\sigma} = F^{-1}_{\mathrm{N}(0,1)}\left(\frac{1}{2}\left(1 + F_{\mathrm{N}(0,1)}\left(\frac{a-\mu}{\sigma}\right)\right)\right),$$

jolloin mediaaniksi tulee

$$\mu + \sigma F^{-1}_{\mathrm{N}(0,1)}\left(\frac{1}{2}\left(1 + F_{\mathrm{N}(0,1)}\left(\frac{a-\mu}{\sigma}\right)\right)\right).$$

P4. Komiteaan kuuluu 6 naista ja 4 miestä. Komitean jäsenistä valitaan satunnaisesti jäseniä, yksi kerrallaan, työryhmään kunnes työryhmään kuuluu 3 naista. Mikä on todennäköisyys, että työryhmään kuuluu silloin 5 jäsentä eli 3 naista ja 2 miestä?
Vihje: Mikä on todennäköisyys, että kun on valittu 4 jäsentä työryhmään niin 2 naista ja 2 miestä ovat tulleet valituiksi?

Vastaus: $\frac{2}{7}$

Ratkaisu: Jos lopputulos on se, että työryhmään kuuluu 5 jäsentä, niin työryhmään kuuluu 2 naista ja 2 miestä kun on valittu 4 jäsentä ja viidenneksi valitaan nainen. Eli jos

$$\begin{aligned} A &= \{\text{"Työryhmään kuuluu 5 henkilöä"}\} \\ B &= \{\text{"4:stä ensin valituista 2 ovat naisia ja 2 miehiä"}\} \\ C &= \{\text{"5. valituista on nainen"}\}. \end{aligned}$$

niin $A = B \cap C$ jolloin saadaan

$$\Pr(A) = \Pr(B \cap C) = \Pr(C|B)\Pr(B).$$

Nyt

$$\Pr(B) = \frac{|\text{"Valitaan 2 naista 6:sta"}| \cdot |\text{"Valitaan 2 miestä 4:stä"}|}{|\text{"Valitaan 4 henkilöä 10:stä"}|} = \frac{\binom{6}{2} \cdot \binom{4}{2}}{\binom{10}{4}}.$$

Jos ensin on valittu 2 naista ja 2 miestä niin jäljellä on 4 naista ja 2 miestä joten tässä tapauksessa todennäköisyys että viidenneksi valittu on nainen on

$$\Pr(C|B) = \frac{4}{6},$$

ja saadaan

$$\Pr(A) = \frac{\binom{6}{2} \cdot \binom{4}{2}}{\binom{10}{4}} \cdot \frac{4}{6} = \frac{2}{7}.$$

Jos haluaa laskea suotuisten vaihtoehtojen osuus kaikista vaihtoehdoista niin voidaan ajatella, että kaikki komitean jäsenet järjestetään jonoon ja tämä voi tapahtua $10!$:lla eri tavalla. (Koska varmasti 3 naista on tullut valituksi kun on valittu 7 henkilöä riittäisi muodostaa kaikki 7:n henkilön jonot, mutta tällä ei juuri ole vaikutusta laskuihin.) Kuten edellä todettiin niin jos kolmas nainen tulee valituksi kun valitaan viides jäsen niin neljän ensimmäisen joukossa on 2 naista ja 2 miestä. Näiden naisten (ja silloin myös miesten) valinnan järjestysnumerot voidaan valita $\binom{4}{2}$ eri tavalla. Sitten voidaan valita 3 naista $6 \cdot 5 \cdot 4$ eri tavalla ja 2 miestä $4 \cdot 3$ eri tavalla ja lopuksi voidaan valita jonon loppupää $5!$ eri tavalla koska siina sukupuolella ei ole merkitystä enää. Kaiken kaikkiaan saadaan

$$\binom{4}{2} \cdot 6 \cdot 5 \cdot 4 \cdot 4 \cdot 3 \cdot 5!,$$

eri vaihtoehtoa ja todennäköisyydeksi tulee

$$\frac{\binom{4}{2} \cdot 6 \cdot 5 \cdot 4 \cdot 4 \cdot 3 \cdot 5!}{10!} = \frac{4 \cdot 3 \cdot 6 \cdot 5 \cdot 4 \cdot 4 \cdot 3}{2 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6} = \frac{2}{7}.$$

P5. Jos numerot $0, 1, 2, \ldots, 9$ laitetaan satunnaisesti järjestykseen, mikä on todennäköisyys, että mikään pariton numero ei tule heti toisen parittoman numeron jälkeen?

Vastaus: $\frac{1}{42}$

Ratkaisu: Voimme laittaa 10 numeroa järjestykseen $10!$:lla er tavalla.

Parittomia numeroita on 5 kappaletta ja kahden parittoman numeron välillä täytyy olla vähintaan yksi parillinen numero ja siihen tarvitaan 4 parillista numeroa. Viidennen parillisen numeron voimme laittaa ensimmäiseksi, viimeiseksi tai kahden parittoman luvun väliin, eli meillä on 6 eri vaihtoehtoa. Koska voimme järjestää sekä parilliset että parittomat numerot $5!$:lla eri tavalla niin meillä on $6 \cdot 5! \cdot 5!$ eri mahdollisuutta järjestää numerot siten, että mikään pariton numero ei tule heti toisen parittoman numeron jälkeen.

Klassisen todennäköisyyden määritelmän mukaan kysytty todennäköisyys on

$$\frac{6 \cdot 5! \cdot 5!}{10!} = \frac{1}{42}.$$
