MTTTP5

Lisätehtäviä ratkaisuineen luentomonisteen lukuihin 2 - 4 liittyen

1. Laitosneuvostoon valitaan 2 professoria, 4 muuta henkilökuntaan kuuluvaa jäsentä sekä 4 opiskelijaa. Laitosneuvostoon on pyrkimässä 5 professoria, 7 muuta henkilökuntaan kuuluvaa sekä 8 opiskelijaa. Kuinka monta erilaista laitosneuvostoa voi ehdokkaista muodostua?

2. Olet tulossa kotiin. Avainnipussasi on 5 avainta, joista yksi on kotiavaimesi. Valitset satunnaisesti avaimen. Jollei ovi aukea tällä avaimella, niin valitset jäljellä olevista uuden avaimen ja koetat avata oven, jne. Määritellään X = sen kerran järjestysnumero, jolla ovi aukeaa. Määritä X:n jakauma sekä E(X), Var(X). Laske todennäköisyys sille, että joudut yrittämään ainakin 4 kertaa.

3. Esimerkin 2 tilanne, mutta palautat avain nippuun ennen seuraavan valintaa. Määritä X:n jakauma. Laske todennäköisyys sille, että joudut yrittämään ainakin 4 kertaa.

4. Olkoon X:n tiheysfunktio

$$f(x) = \begin{cases} cx, \text{ kun } 0 \le x \le 2, \\ 0, \text{ muulloin} \end{cases}$$

Määritä vakio c ja piirrä tiheysfunktion kuvaaja.
Laske lisäksi $P(X \le 1)$ ja $P(1 \le X \le 2)$.

5. Sijoitetaan 1000 euroa. Sijoitusvaihtoehtoja on kolme, joissa tuotto ajatellaan muodostuvat annettujen todennäköisyyksien mukaisesti seuraavasti: Vaihtoehto 1) 10000 euron voitto todennäköisyydellä 0.15 ja 1000 euron tappio todennäköisyydellä 0.85, 2) 1000 euron voitto todennäköisyydellä 0.50, 500 euron voitto todennäköisyydellä 0.30 ja 500 euron tappio todennäköisyydellä 0.2, 3) 400 euron varma voitto. Millä vaihtoehdolla on suurin tuoton odotusarvo? Entä suurin hajonta? Mitä vaihtoehtoa suosittelet sijoittajalle?

6. Myyjän vuositulot ovat $6000 + 8% myytyjen tuotteiden arvosta. Myytyjen tuotteiden arvoa voidaan kuvata satunnaismuuttujana, jonka odotusarvo on $600000 ja keskihajonta $180000. Määrittele satunnaismuuttuja Y = myyjän vuosiansiot. Laske E(Y) ja Var(Y).

7. Kokeessa vastataan valitsemalla väittämistä oikea vaihtoehto. Kokeessa on 10 tehtävää, joissa jokaisessa on 3 vaihtoehtoa, joista yksi on oikein. Oikeasta vastauksesta vastaaja saa 2 pistettä ja väärästä -3 pistettä. Arvaaja vastaa täysin satunnaisesti jokaiseen kysymykseen. Mikä on arvaajan saaman pistemäärän odotusarvo? Entä varianssi? Jos tenttiin osallistuisi 100 opiskelijaa ja kaikki arvaisivat vastauksen jokaiseen kysymykseen, niin mitä lukua lähellä olisi tenttipistemäärän keskiarvo?

8. Tarkastellaan kolmilapsisia perheitä. Oletetaan, että syntyvistä lapsista 51,2 % on poikia ja syntyvän lapsen sukupuoli on riippumaton sisarusten sukupuolesta. Määritellään X = poikien lukumäärä perheessä. Määritä X:n jakauma. Jos laskettaisiin Suomen kaikista kolmilapsisista perheistä poikien lukumäärän keskiarvo, niin mitä lukua lähellä se olisi? Miksi?

9. Olkoon $X \sim Bin(n, 6/13)$ ja $Y = 2X - n + 10$. Millä n:n arvolla $E(Y) = 0$? Laske saamallasi n:n arvolla $Var(Y)$.

10. Olkoon $X \sim N(0, 4)$. Hahmottele X:n tiheysfunktion kuvaajaa. Arvioi kuvasi avulla $P(X > 2)$. Laske $P(X > 2)$.

11. Olkoon $X \sim N(1, 9)$. Laske $P(X < 2)$.
Määritä 10X:n jakauma ja laske $P(10X < 2)$. Piirrä vastaavat kuviot.

12. Oletetaan, että älykkyysosamäärä on jakautunut populaatiossa $N(100, 225)$-jakauman mukaan. Minkä rajojen sisällä on keskimmäinen puolikas (so. alarajan alapuolelle jää 25% ja ylärajan yläpuolelle 25%). Jos älykköjen kerhoon pääsee vain 0.1% älykkäimmästä päästä, niin mikä on kerhoon pääsijöiden
alaraja?

13. Kahdeksantoistavuotias Tiina käy lääkärintarkastuksessa. Lääkäri mittaa Tiinan pituuden saaden arvoksi 174 cm. Lääkäri kertoo, että Tiinan ikäisillä nuorilla naisilla pituus vaihtelee normaalijakauman mukaisesti ja Tiinaa pidempiä on 7 %. Lisäksi lääkäri kertoo, että alle 156 cm pitkiä tyttöjä on myös 7%. Määritä kyseisen normaalijakauman parametrit, kun oletetaan lääkärin tietämys oikeaksi.

14. Miespotilaan verikoe osoitti kolesterolipitoisuuden olevan 310 mg/dl. Tohtori sanoi, että keskimääräinen kolesteroliarvo hänen ikäisellään miehellä on 200 mg/dl ja ainoastaan 2.5%:lla on kyseinen arvo suurempi kuin 310. Laske kolesteroliarvon keskihajonta, kun oletetaan muuttuja normaalisti jakautuneeksi sekä tohtorin tietämys oikeaksi. Selvitä lisäksi millaiset mahdollisuudet on löytää mies, jolla arvo on alle 100.

15. Oletetaan, että sokeripussien paino vaihtelee normaalijakauman, jonka odotusarvo on 1000g ja keskihajonta 20 g, mukaisesti. Valitaan sokeripussien joukosta satunnaisesti 4 pussia. Millä todennäköisyydellä näiden pussien yhteispaino on korkeintaan 3900 g?

16. Tarkastellaan tietyn matkan juoksuaikaa. Oletetaan, että A:n juoksuaika $X \sim N(30, 8)$ ja B:n juoksuaika $Y \sim N(34, 8)$. Laske todennäköisyys sille, että B:n voittaa A:n (olet. X ja Y riippumattomiksi).

17. Olkoon $X \sim N(20, 9)$. Laske $P(19 \leq X \leq 21)$. Tehdään 10 alkion satunnaisotos kyseisestä jakaumasta. Tarkastellaan siis riippumattomia satunnaismuuttujia $X_i \sim N(20, 9)$, i = 1, ..., 10. Määritellään $\overline{X} = (X_1 + ... + X_{10})/ 10$. Määritä $\overline{X}$:n jakauma ja laske $P(19 \leq \overline{X} \leq 21)$.

18. Sekoitat tavallisen 52 kortin korttipakan ja pyydät ystävääsi vetämään sieltä kortin sekä korttia katsomatta arvaamaan onko se pata, risti, ruutu vai hertta. Palautat kortin pakkaan ja toistat koetta 100 kertaa. Olkoon X = ystäväsi oikeiden arvausta lukumäärä 100 toistossa. Laske todennäköisyys, että $X \geq 30$.

19. Oletetaan, että paketoimattomien suklaalevyjen paino on normaalijakautunut siten, että odotusarvo on 508 g ja keskihajonta 4 g. Eräästä erästä, josta ei tiedetä, onko levyt punnittu paketoituina vai paketoimattomina tehdään 4 alkion otos ja saadaan otoskeskiarvoksi 516 g. Voitko tehdä päätelmän, että suklaalevyt on punnittu paketoimattomina? Perustele päätelmäsi.

Ratkaisuja

Ilmoita mahdollisista virheistä Leppälälle

1. Erilaisia laitosneuvostoja on $\binom{5}{2}\binom{7}{4}\binom{8}{4} = 24500$.

2. Ks. http:/ / www.sis.uta.fi/ tilasto/ mtttp2/ syksy2014/ luentorunko.pdf esim. 2.3.10 s. 10 ja esim. 3.2.4 s. 18. $P(X \geq 4) = P(X=4) + P(X=5) = 2/5 = 0{,}4$.

3. $P(X = k) = (4/5)^{k-1}(1/5)$, k = 1, 2, ...

$P(X \geq 4) = 1 - P(X<4) = 1 – (P(X=1) + P(X=2) + P(X=3)) = 1 – ((1/5) + (4/5)(1/5) + (4/5)^2(1/5)) = 0{,}512$

4. c = ½, $P(X \leq 1) = 1 \cdot (1 \cdot 1/2)/2 = 1/4$ (kolmion pinta-alana) ja

$P(1 \leq X \leq 2) = 1 - P(X \leq 1) = 1 - 0{,}25 = 0{,}75$.

5. Määritellään X = tuotto.

1) $E(X) = 0.15 \cdot 10000 - 0.85 \cdot 1000 = 650$

$Sd(X) = \sqrt{0.15(10000-650)^2 + 0.85(-1000-650)^2} \approx 3928$

2) $E(X) = 0.5 \cdot 1000 + 0.3 \cdot 500 - 0.2 \cdot 500 = 550$

$Sd(X) = \sqrt{0.5(1000-550)^2 + 0.3(500-550)^2 + 0.2(-500-550)^2} \approx 568$

3) $E(X) = 400$

$Sd(X) = 0$.

Vaihtoehdossa 1) suurin tuoton odotusarvo, mutta myös suurin riski, koska hajonta suurin. Varman tuoton saa vaihtoehdolla 3).

6. Olkoon X = myytyjen tuotteiden arvo, jolloin $Y = 6000 + 0.08X$.

$E(Y) = E(6000 + 0.08X) = 6000 + 0.08\,E(X) = 6000 + 0.08 \cdot 600000 = 54\,000$,

$Var(Y) = Var(6000 + 0.08X) = 0.08^2\,Var(X) = 0.08^2 \cdot 180000^2 = 14400^2$.

7. Kokonaispistemäärä $Y = X_1 + \dots + X_{10}$, missä X_i = i. kysymyksestä saatu pistemäärä. $E(X_i) = (1/3) \cdot 2 + (2/3) \cdot (-3) = -4/3$ ja

$Var(X_i) = \frac{1}{3} \cdot \left(2-(-\frac{4}{3})\right)^2 + \frac{2}{3} \cdot \left(-3-(-\frac{4}{3})\right)^2 = 150/27$, i = 1, 2, ..., 10.

Nyt $E(Y) = E(X_1 + \dots + X_{10}) = E(X_1) + \dots + E(X_{10}) = 10 \cdot (-4/3) \approx -13.33$ ja

$Var(X_1 + \dots + X_{10}) = Var(X_1) + \dots + Var(X_{10}) = 10 \cdot 150/27 \approx 55.56$.

Toisin: Olkoon X = oikeitten vastausten lkm. $X \sim Bin(10, 1/3)$, $E(X)=10/3$, $Var(X) = 20/9$. Kokonaispistemäärä $Y = 2X-3(10-X)= 5X - 30$, joten $E(Y) = 5E(X) - 30 = 50/3 - 30$,

$Var(Y) = 5^2\,Var(X) = 500/9$.

Jos kaikki tenttiin osallistujat arvaisivat, niin tenttipistemäärän keskiarvo olisi lähellä Y:n odotusarvoa ole lukua -13.3.

8. $X \sim Bin(3, 0.512)$, $E(X) = 3 \cdot 0.512 = 1.536$. Poikien lukumäärän keskiarvon pitäisi olla lähellä tätä lukua, jos oletus poikien syntymätodennäköisyydestä pitää paikkansa.

9. Jos $X \sim Bin(n, p)$, niin $E(X) = np$, $Var(X) = np(1-p)$, joten tässä

$E(X) = n \cdot 6/13$, $Var(X) = n \cdot (6/13) \cdot (7/13)$.

$E(Y) = E(2X - n + 10) = 2E(X) - n + 10 = 2 \cdot n \cdot (6/13) - n + 10$, joka on 0, jos n = 130.

Tällöin $Var(Y) = Var(2X - n + 10) = 4Var(X) = 4 \cdot 130 \cdot (6/13) \cdot (7/13)$.

10. Kuvaaja symmetrinen nollan suhteen. Yhden hajonnan päässä nollasta (-2 ja 2 välissä) n. 68 %, kahden hajonnan päässä n. 95 %.
Arvio $P(X > 2) \approx 0.5\text{-}0.68/2 = 0.16$.
$P(X > 2) = 1 - P(X \leq 2) = 1 - P((X - 0)/2 < (2 - 0)/2) = 1 - \Phi(1) \approx 1 - 0.8413 = 0.1587$.

11. $P(X < 2) = P((X - 1)/3 < (2 - 1)/3) = \Phi(1/3) \approx \Phi(0.33) \approx 0.6293$.
$Y = 10X$, $E(Y) = 10\, E(X) = 10 \cdot 1 = 10$ ja $Var(Y) = Var(10X) = 10^2 \cdot Var(X) = 10^2 \cdot 9$, joten $Y = 10X \sim N(10, 900)$.
Tällöin $P(10X < 2) = P(Y < 2) = P((Y - 10)/30 < (2 - 10)/30) = \Phi(-0.27) = 1 - \Phi(0.27) = 1 - 0.6064 = 0.3936$.

12. Merkitään näitä kysyttyjä rajoja a, b, c. Tällöin $P(a \leq X \leq b) = 0.5$, $P(X \leq a) = 0.25$, $P(X \leq b) = 0.75$, $P(X \leq c) = 0.999$.
$P(X \leq b) = P((X\text{-}100)/15 \leq (b - 100)/15) = \Phi((b - 100)/15) = 0.75$, jolloin taulukosta $(b - 100)/15 = 0.67$, joten $b \approx 110$. Koska normaalijakauma on symmetrinen odotusarvon suhteen, $a \approx 100 - 10 = 90$. $P(X \leq c) = P((X\text{-}100)/15 \leq (c - 100)/15) = \Phi((c - 100)/15) = 0.999$, jolloin taulukosta $(c - 100)/15 = 3.09$, joten $c \approx 146$.

13. Merkitään X = nuoren naisen pituus. Oletetaan, että X noudattaa normaalijakaumaa odotusarvona μ ja keskihajontana σ. Oletuksen (lääkärin tietämyksen) mukaan $P(X > 174) = 0{,}07$ ja $P(X < 156) = 0{,}07$, joista saadaan (yhtälöpari)
$\Phi((174 - \mu)/\sigma) = 1 - 0{,}07 = 0{,}93$
$\Phi((156 - \mu)/\sigma) = 0{,}07$.
Normaalijakauman kertymäfunktion taulukon avulla saadaan, että $(174 - \mu)/\sigma \approx 1{,}48$ ja $(156 - \mu)/\sigma \approx -1{,}48$ joista (yhtälöparista) σ ja μ voidaan ratkaista. Saadaan (tässä välivaiheet jätetty kirjaamatta) $\sigma \approx 6{,}08$ ja $\mu \approx 165$. Voi myös käyttää hyväkseen normaalijakauman symmetrisyyttä ja laskea ensin odotusarvon, joka on suoraan $(174 + 156)/2$.

14. Merkitään X = kolesteroliarvo. Oletetaan, että $X \sim N(200, \sigma^2)$ ja $P(X \leq 310) = 0.975$, joten $\Phi((310 - 200)/\sigma) = 0.975$, josta $(310 - 200)/\sigma = 1.96$ ja $\sigma = 56.12$. $P(X \leq 100) = \Phi((100 - 200)/56.12) = \Phi(-1.78) = 1 - \Phi(1.78) = 1 - 0.9625 = 0.0375$.

15. Olkoon X_i = sokeripussin paino, $i = 1, 2, 3, 4$. Tällöin yhteispaino $Y = X_1 + X_2 + X_3 + X_4$. Nyt Y on normaalisti jakautunut odotusarvona $4 \cdot 1000$ g ja keskihajontana $\sqrt{4} \cdot 20$ g = 40 g. $P(Y \leq 3900) = \Phi((3900 – 4000)/40) = \Phi(-2{,}5) = 1 - \Phi(2{,}5) \approx 1 – 0{,}9938 = 0{,}0062$.

16. $Y - X \sim N(34\text{-}30, 8\text{+}8)$, joten $P(Y - X < 0) = \Phi((0 - 4)/4) = \Phi(-1) = 1 - 0.8413 = 0.1587$.

17. $P(19 \le X \le 21) = \Phi((21 - 20)/\sqrt{3}) - \Phi((19 - 20)/\sqrt{3}) = \Phi(1/\sqrt{3}) - \Phi(-1/\sqrt{3}) = \ldots =$ 0.2586.
$\overline{X} \sim N(20, 9/10)$, joten $P(19 \le \overline{X} \le 21) = \Phi((21 - 20)/0.949) - \Phi((19 - 20)/0.949) =$ $\Phi(1.05) - \Phi(-1.05) = \ldots = 0.7062$.

18. $X \sim Bin(100, 0.25)$, joten $E(X) = 100 \cdot 0.25 = 25$, $Var(X) = 100 \cdot 0.25 \cdot (1 - 0.25)$ $= 18.75$. Binomijakaumaa voidaan approksimoida normaalijakaumalla.
Tässä $X \overset{\text{likimain}}{\sim} N(25, 18.75)$, joten $P(X \ge 30) \approx 1 - \Phi((30 - 25)/4.33) = 1 - \Phi(1.15) =$ $1 - 0.8749 = 0.1251$. Voidaan myös ajatella, että binomijakaumasta $P(X \ge 30) = 1 -$ $P(X \le 29)$, jota sitten approksimoidaan normaalijakaumalla ja saadaan $P(X \ge 30) \approx$ $1 - \Phi((29 - 25)/4.33) = 1 - \Phi(0.92) = 0.1788$.
Jos käytetään jatkuvuuskorjausta $P(X \ge 30) \approx 1 - \Phi((29.5 - 25)/4.33) = 0.1515$.
Jos kysytty todennäköisyys lasketaan binomijakaumasta, saadaan 0.1495 (laskettu excelillä) .

19. Olkoon X = levyn paino paketoimattomana. Tällöin $\overline{X} \sim N(508, 4^2/4)$, joten $P(\overline{X} \ge 516) = 1 - P(\overline{X} \le 516) = 1 - \Phi((516-508)/2) = 1 - \Phi(4) \approx 0$. Siis on hyvin epätodennäköistä, että satunnaisesti paketoimattomien levyjen joukosta tehdyn 4 suklaalevyn otoksen keskiarvo on suurempi kuin 516. Tehdään päätelmä, että suklaalevyt on punnittu paketoituina.