TILASTOMATEMATIIKKA

Harjoitusviikko 3, kevät 2024

Harjoituksen teemoja ovat

(i) Diskreetin jakauman odotusarvo, varianssi ja keskihajonta

(ii) Jatkuvan jakauman odotusarvo, varianssi ja keskihajonta

(iii) Odotusarvon ja varianssin ominaisuudet

(iv) Keskeinen raja-arvolause

(v) Normaalijakauma-approksimaatio

(vi) Jatkuvuuskorjaus

Tehtävät 1-5 liittyvät teemoihin (i)-(iii), joita käsitellään pääasiassa alkuviikon harjoituksissa. Teemoja (iv)-(vi) käsitellään tehtävissä 6-9, jotka on ensisijaisesti tarkoitettu loppuviikon tehtäviksi. Tehtäviä voi toki ratkoa haluamassaan järjestyksessä. Suositellaan laskettavaksi alkuviikosta **tehtävät 2 ja 3** alkuviikosta ja loppuviikosta **tehtävät 6 ja 8**.

1. Tarkastellaan diskreettiä satunnaismuuttujaa, jonka pistetodennäköisyydet ovat

x	0	1	2	3
$\mathbb{P}(X=x)$	1/8	3/8	3/8	1/8

a) Määrää X:n odotusarvo ja varianssi.

b) Olkoon $X \sim \mathrm{Bin}(3, 1/2)$. Määrää X:n pistetodennäköisyydet. Mitä havaitset? Käytä kaavakokoelmaa X:n odotusarvon ja varianssin määräämiseen.

d) Simuloi Excelillä annettua jakaumaa *Data Analysis Toolpakin Random number generation*-toiminnolla generoimalla kokoa 10000 oleva otos jakaumasta. Poimi kunkin arvon esiintymiskerrat *COUNTIF*-komennolla. Mitä saat simuloiduiksi pistetodennäköisyyksiksi? Voit katsoa mallia vaikkapa videosta

https://www.youtube.com/watch?v=emtaBhH7kf0.

2. Määrää vakio c, kun tiedetään, että X on satunnaismuuttuja, jonka

a) pistetodennäköisyydet ovat

x	1	2	3	4	5
$\mathbb{P}(X=x)$	c	$2c$	$3c$	$2c$	c

Laske X:n odotusarvo ja keskihajonta. Mitä intuitiosi sanoo odotusarvosta?

b) tiheysfunktio on

$$f_X(x) = \begin{cases} c(3-|x-3|), & 0 \le x \le 6, \\ 0, & \text{muulloin.} \end{cases}$$

Laske X:n odotusarvo ja keskihajonta. Mitä intuitiosi sanoo odotusarvosta?

3. Olkoot X_1, X_2 riippumattomia ja samalla tavalla jakautuneita satunnaismuuttujia. Laske muuttujan $X = \max\{X_1, X_2\}$

a) odotusarvo ja varianssi, kun $X_i \sim \mathrm{Bin}(1, p)$ (Bernoulli-jakautuneita).

b) odotusarvo ja varianssi, kun $X_i \sim \mathrm{Tas}(0, 1)$.

c) odotusarvo (Excelillä) simuloimalla, kun $X_i \sim N(0, 1)$. Käytä käänteisfunktiomenetelmää Excelin NORM.S.INV-toiminnolla. Simuloi esimerkiksi 10000 näytettä kummastakin muuttujasta X_i ja laske saamiesi maksimien keskiarvo. Vertaa harjoitusviikon 2, tehtävä 3. **Haaste:** laske odotusarvo $\mathbb{E}(X) = \frac{1}{\sqrt{\pi}}$ analyyttisesti.

4. Magneettisessa tiedon tallennuksessa oli havaittu virheiden määrän suurta talletusyksikköä kohti noudattavan Poisson-jakaumaa siten, että talletuksessa esiintyi keskimäärin yksi virheellinen bitti 100000 talletettua bittiä kohti. Tieto talletetaan sektoreina, missä sektori käsitti 4096 8-bittistä sanaa.

   a) Kuinka monta tallennusvirhettä esiintyi keskimäärin yhdessä sektorissa?

   b) Kuinka monta sektoria on keskimäärin käytävä läpi jotta ainakin yksi tallennusvirhe löytyy? Sektorit ovat toisistaan riippumattomat.

5. Erään kemikaalin pitoisuus $[g/cm^3]$ tietyssä aineseoksessa on normaalijakautunut satunnaismuuttuja, jolla on odotusarvo μ ja hajonta $\sigma = 0.004\ g/cm^3$. Tehtävänä on arvioida odotusarvoa μ otoksen avulla. Kuinka monta toisistaan riippumatonta näytettä on otettava, jotta näytteestä mitattujen kemikaalipitoisuuksien keskiarvo poikkeaisi μ:stä suuntaan tai toiseen korkeintaan $0.002\ g/cm^3$ todennäköisyydellä 0,9?

6. Eräällä tehtaalla valmistetuista mikrosiruista 2 % on viallisia. Poimitaan satunnaisotannalla 100 mikrosirua tarkastukseen. Laske normaalijakauma-approksimaatiolla todennäköisyys, että tarkastuksessa löydetään 5 viallista mikrosirua. Vertaa harjoitusviikon 2 tehtävään 3.

7. Erään metallikomponentin väsymisrajaa X [MPa] mallinnetaan log-normaalijakaumalla, eli $\ln X \sim N(\mu, \sigma^2)$. On havaittu, että väsymisrajan mediaani Md on 400 [Mpa] ja että $\sigma^2 = 0.15$.

   a) Määrää parametri μ. Käytä hyväksi tietoa
   $$P(X \leq \text{Md}) = 0.5.$$

   b) Mikä on väsymisrajan odotusarvo?

   c) Mitataan 30 komponentin väsymisraja. Laske normaalijakauma-approksimaatiolla todennäköisyys, että väsymisrajan keskiarvo on vähintään 350 [MPa]. Oletetaan, että mittaukset ovat riippumattomia ja samalla tavalla jakautuneita.
   Vihje: Log-normaalijakauman hajonnan kaavan voi kaivaa vaikka Wikipediasta.

8. Oletetaan, että satunnaisesti valittu nuorukainen ostaa suosittua energiajuomaa yhden pullon todennäköisyydellä 0.9 ja kaksi pulloa todennäköisyydellä 0.1.

   a) Laske normaalijakauma-approksimaatiolla todennäköisyys, että 100 nuorukaista ostaa vähintään 115 pulloa.

   b) Laske normaalijakauma-approksimaatiolla, kuinka monta pulloa kauppiaan on varattava myyntiin, jotta 100 nuorukaiselle riittää juomia vähintään 95 % todennäköisyydellä.

   c) Laske edelliset kohdat tarkasti tilanteeseen sopivalla jakaumamallilla. Mitä oletuksia täytyy tehdä, jotta homma onnistuu?

9. Fyysikot käyttävät satunnaiskävelyä diffuusion tai yleisesti partikkelien satunnaisliikkeen mallintamiseen. Partikkeli suorittaa siirtymiä ajanhetkillä $t = 1, 2, \ldots$ ja sen paikka S_n ajanhetkellä n voidaan ajatella olevan siirtymien $X_1, \ldots, X_n$ summa. Oletetaan, että siirtymät ovat riippumattomia ja samalla tavalla jakautuneita niin, että partikkeli voi samalla todennäköisyydellä joko pysyä paikallaan tai siirtyä yhden yksikön verran $(+1)$ oikealle tai vasemmalle (-1). Määrää normaalijakauma-approksimaatiolla todennäköisyys, että 10000 siirtymän jälkeen partikkeli on vähintään 100 yksikköä lähtöpisteestä oikealle.