TILASTOMATEMATIIKKA

Harjoitusviikon 2 tehtävät, kevät 2024

Harjoituksen teemoja ovat:

(i) Kertymäfunktio, pistetodennäköisyysfunktio ja tiheysfunktio

(ii) Diskreetti satunnaismuuttuja

(iii) Jatkuva satunnaismuuttuja

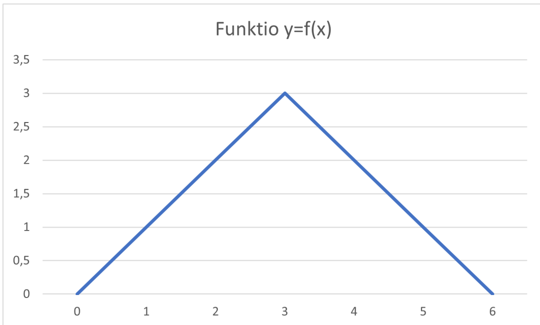
(iv) Tyypillisimmät jakaumat, jotka kantavat omaa nimeä

Tehtävien laskemista voi rytmittää esimerkiksi niin, että alkuviikosta käy läpi tehtäviä 1-5, joista tehtävät 1 ja 2 liittyvät satunnaismuuttujien luokitteluun ja loput tehtävät liittyvät diskreetteihin satunnaismuuttujiin. Tehtävät 6-9 liittyvät jatkuviin satunnaismuuttujiin. Toki myös muunlainen rytmittäminen on mahdollista. Alkuviikon harjoituksissa suositellaan laskettavaksi **tehtävät 2 ja 3** ja loppuviikon harjoituksissa **tehtävät 6 ja 8**.

1. Ilmoita kussakin seuraavista tapauksista, onko kyse jatkuvasta vai diskreetistä satunnaismuuttujasta. Määrää myös satunnaismuuttujan arvojoukko, mikäli se on mahdollista. Voidaanko mallintamisessa käyttää hyväksi jotain tunnettua jakaumaa?

   a) Vikojen lukumäärä neliömetrillä satunnaisesti valitussa paperirullassa.
   b) Kemikaalin konsentraatio liuoksessa.
   c) Liian pitkien pulttien osuus satunnaisesti valitussa pultteja sisältävässä laatikossa.
   d) Virheiden lukumäärä 1000 satunnaisesti valitussa rivissä ohjelmointikoodia.
   e) Satunnaisesti valitun metallilevyn murtolujuus.
   f) Elektronisen komponentin elinikä.

2. Kuvaan 1 on piirretty satunnaismuuttujiin liittyviä kuvioita.



Kuva 1: Satunnaismuuttujiin liittyviä kuvioita

   a) Voivatko kuviot esittää jotain satunnaismuuttujaa sellaisenaan? Jos ei, niin miten voit skaalata kuvioita, että niistä tulee keskeisiä satunnaismuuttujia kuvaavia funktioita? Mikä on näin saatujen satunnaismuuttujien arvojoukko ja tyyppi?

   b) Millä todennäköisyydellä satunnaismuuttujat ovat suurempaa kuin 2?

   c) Mikä on kummankin satunnaismuuttujan todennäköisin arvo?

3. Eräällä tehtaalla valmistetuista mikrosiruista 2 % on viallisia. Poimitaan satunnaisotannalla 100 mikrosirua tarkastukseen. Millä todennäköisyydellä tarkastuksessa löydetään 5 viallista mikrosirua?

   a) Laske todennäköisyys tilanteeseen sopivan satunnaismuuttujan avulla.

   b) Kuinka monta viallista mikrosirua löydetään keskimäärin? Millä todennäköisyydellä viallisten määrä on korkeintaan 4?

   c) Laske kysytty todennäköisyys Poisson-jakauman avulla.

   d) Simuloi kysyttyä todennäköisyyttä Excelillä. Käytä BINOM.INV- ja COUNTIF-funktioita sekä kurssilla esitettyä käänteismuunnosmenetelmää. Simuloi esimerkiksi 10000 näytettä binomijakaumasta. Voit ottaa mallia vaikkapa videosta

      https://www.youtube.com/watch?v=5uF0L6nImn4

      kohdasta 23:15 eteenpäin.

4. Erittäin laaja havupuualue oli neulaskadon takia jäämässä tuotantotavoitteestaan. Tuottoennusteen tekemiseksi tietty osa metsää jaettiin hehtaarin suuruisiin ruutuihin ja laskettiin neulaskadon vaivaamien puiden lukumäärä ruutua kohti. Laskennan tuloksena havaittiin neulaskadon vaivaamien puiden lukumäärä/ruutu olevan Poisson-jakautuneen ja sellaisten ruutujen, joissa neulaskadon vaivaamia puita ei ollut lainkaan, osuuden olevan 7 %.

   a) Mikä oli oli keskimääräinen neulaskadon vaivaamien puiden lukumäärä ruutua kohti?

   b) Millä todennäköisyydellä ruudussa esiintyi ainakin 2 neulaskadon vaivaamaa puuta?

5. Valmistajan ilmoituksen mukaan halvoista elektronisista komponenteista AA5 keskimäärin 5 % ei täytä spesifikaatioita. Ostaja osti AA5:ttä monen tuhannen kappaleen erissä ja teki seuraavan päätöksentekomallin erän hyväksymiseksi vastaanottotarkastuksen yhteydessä: Erästä valittiin satunnaisesti 10 komponenttia, jotka testattiin. Jos kaikki testatut komponentit toteuttivat spesifikaatiot, niin erä hyväksyttiin. Jos ainakin 2 testatuista komponenteista ei toteuttanut spesifikaatioita, erä hylättiin. Jos testatuista täsmälleen yksi ei toteuttanut spesifikaatioita, otettiin uusi 10 yksilön satunnainen näyte, ja erä hyväksyttiin, jos uudessa näytteessä kaikki komponentit täyttävät spesifikaatiot. Jos valmistajan väite pitää paikkansa, niin millä todennäköisyydellä

   a) erä hyväksytään ensimmäisen näytteen perusteella?

   b) lopullinen päätös tehdään eli erä hylätään tai hyväksytään ensimmäisen näytteen perusteella?

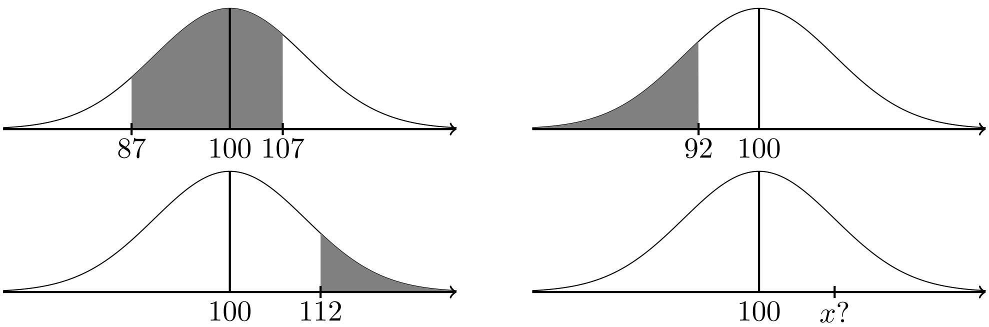
   c) erä hyväksytään?

6. Tarkastellaan satunnaismuuttujaa $X \sim N(\mu, 10^2)$, jonka tiheysfunktion profiileja on piirretty Kuvaan 2.

   a) Laske väritettyjen alueiden pinta-alat.

   b) Missä pisteessä x pätee
      (i)$\mathbb{P}(X \leq x) = 0.93$?
      (ii)$\mathbb{P}(X \leq x) = 0.38$?

   c) Olkoot $X_1, X_2, \ldots, X_{10}$ riippumattomia ja samalla tavalla jakautuneita kuin X. Millä todennäköisyydellä muuttujien maksimi on suurempaa kuin 107? Toisin sanoen, laske todennäköisyys $\mathbb{P}(\max_{1 \leq i \leq 10} X_i > 107)$.



Kuva 2: Tehtävän 6 tiheysfunktiot

7. Eloonjäämisfunktio S (*survival function*) ilmoittaa todennäköisyyden, että henkilö on elossa tietyn ajan kuluttua. Oletetaan, että syöpäpotilaan eloonjäämisfunktio on muotoa $S(t) = 1 - F_X(t)$, missä t on syöpädiagnoosista kulunut aika vuosina ja X on elinajan ilmoittava satunnaismuuttuja, jonka oletetaan noudattavan Weibull-jakaumaa parametreilla $\alpha = 0.98$ ja $\beta = 0.30$.

   a) Esitä Weibull-jakauman tiheysfunktion lauseke (löytyy esimerkiksi luentomonisteesta).

   b) Laske todennäköisyys, että henkilö on elossa 5 vuotta syöpädiagnoosin jälkeen.

   c) Mikäli henkilö on elossa 5 vuotta syöpädiagnoosin jälkeen, niin millä todennäköisyydellä hän elää vielä 5 vuotta?

8. Oletetaan, että komponentin elinaika on eksponenttijakautunut ja että elinajan odotusarvo on 20 päivää.

   a) Esitä eksponenttijakauman tiheysfunktion ja kertymäfunktion lauseke (löytyy esimerkiksi luentomonisteesta). Miten parametri liittyy odotusarvoon?

   b) Millä todennäköisyydellä komponentti kestää vähintään 30 päivää?

   c) Jos komponentti on kestänyt 30 päivää, niin millä todennäköisyydellä komponentti kestää vielä toiset 30 päivää?

   d) Varastossa on 10000 komponenttia. Kuinka monta komponenttia keskimäärin kestää vähintään 30 päivää?

9. Valmistetaan metallikuulia, joiden halkaisijaa X pystytään kontrolloimaan. Mikä on kuulien tilavuuden jakauma (kertymäfunktio ja tiheysfunktio), kun oletetaan, että

   a) $X \sim \text{Tas}(a,b)$?

   b) $X \sim N(\mu,\sigma^2)$?

   Vihje: Tilavuus V on X:n muunnos $V = h(X)$ jollekin funktiolle h. Tilavuuden kertymäfunktio voidaan ratkaista X:n jakaumasta hyödyntämällä yhtäsuuruutta

$$\mathbb{P}(V \leq v) = \mathbb{P}(X \leq h^{-1}(v)),$$

   missä h^{-1} on h:n käänteisfunktio.