

TILASTOMATEMATIIKKA

Harjoitusviikko 4, kevät 2024

Suosittelaa laskettavaksi **tehtävät 1 ja 2**. Huomaa, että tämä harjoitus edellyttää osin *Excel*:n (tms. laskentaohjelmiston) käyttöä. Välivaiheiden ja perustelujen on oltava näkyvillä ohjelmiston käytöstä huolimatta.

1. Tutkitaan ohessa olevan Excel-tiedoston dataa.
 - a) Piirrä datan histogrammi Excelin Insert-valikon valmiilla Histogram-toiminnolla. Maa-laa ensin simulaation data. Sen voi tehdä näppärästi valitsemalla ensimmäinen solu ja käyttämällä näppäinyhdistelmää "CTRL+SHIFT+Nuoli alas". Miltä histogrammi näyttää? Tutki myös datan tilastoja *Data Analysis Toolpakin Descriptive Statistics*-työkalulla. Mikä on havaintojen lukumäärä, keskiarvo, mediaani, keskihajonta ja keskiarvon keskivirhe? Mistä jakaumasta arvelet datan olevan peräisin?
 - b) Laske odotusarvon piste-estimaatti ja 95 % symmetrinen 2-suuntainen luottamusväli. Mitä voit sanoa estimaatin tarkkuudesta?
2. Lauri Markkanen on kuluvalle kaudella heittänyt 193 vapaaheittoa, joista 169 on mennyt koriin (tilanne 26.1.2024).
 - a) Määrää vapaaheiton onnistumisprosentin 95 % luottamusväli.
 - b) Kuinka suuri vapaaheittojen määrän täytyy olla, jotta onnistumisprosentin virhemarginaali olisi 2 prosenttiyksikköä suuntaansa?
3. Lentokoneen siipipäällysteen niiteille on asetettu laatuvaatimukseksi vähintään 400 N (Newton) odotettavissa oleva vetolujuus ainakin 99 % varmuudella. Laatu testattiin aika ajoittamatta 20 niitin umpimähkäinen otos ja vetorasittamalla niitit poikki. Erään otoksen vetorasitustulokset olivat
401, 405, 390, 385, 399, 360, 378, 408, 410, 365, 400, 402, 383, 388, 389, 403, 395, 397, 402, 391.
Minkä rajan yläpuolella niittien odotettavissa oleva vetolujuus on mittaustulosten mukaan 99 % varmuudella? Toteutuiko niittien laatuvaatimus? Niittien vetolujuutta voidaan pitää normaalijakautuneena suureena.
4. Eräässä prosessissa biopolttoaineen valmistuksessa biomassaa ruokosokerin muodossa muutetaan polttoaineeksi katalyyttisissä reaktioissa. Yhdessä vaiheessa prosessia saadaan piltuutta 3 olevia hiiliketjuja. Eräs katalyytti tuotti seuraavat saannot [gal]
5.57 5.76 4.18 4.64 7.02 6.62 6.33 7.24
5.57 7.89 4.67 7.24 6.43 5.59 5.39
 - a) Perehdy normaalisuuden graafiseen tarkasteluun Excelillä piirtämällä QQ-kuvio¹. Katso esimerkiksi Youtube-video
Normal probability plot
ja piirrä QQ-kuvio videossa esitetyllä tavalla. Mitä voit sanoa datan normaalisuudesta QQ-kuvion perusteella?
 - b) Laske saannon odotusarvon ja varianssin suurimman uskottavuuden estimaatit.
 - c) Laske saannon odotusarvon 95 % luottamusväli.

¹Q-Q plot, normal probability plot

5. Osakkeiden hinnan mallittamisessa käytetään kuuluisaa Black-Scholesin mallia, jonka mukaan osakkeen hinta ajanhetkellä t on satunnaismuuttuja

$$H(t) = H(0) \cdot \exp\left(\left(b - \frac{1}{2}\sigma^2\right)t + \sigma W(t)\right),$$

missä $H(0)$ on osakkeen hinta ajanhetkellä $t = 0$, $b, \sigma > 0$ ovat vakioita ja $W(t) \sim N(0, t)$ on satunnaismuuttuja. Oletetaan, että $H(0) = 100$, $b = 0.1$ ja $\sigma = 0.3$ (tällä mallilla osakkeen hinta on lognormaalijakautunut).

- a) Laske osakkeen hinnan odotusarvo ajanhetkellä $t = 1$ (**Vihje:** neliöi odotusarvossa esiintyvä eksponentti ja käytä integroinnissa sopivaa sijoitusta).
- b) Määrää ajanhetkellä $t = 1$ osakkeen hinnalle väli, jolla hinta on 95% todennäköisyydellä, eli väli $[a_1, a_2]$, jolle $P(a_1 \leq H(1) \leq a_2) = 0.95$, laskemalla normaalijakauman avulla sellainen luku r_0 , että $P(-r_0 \leq W(1) \leq r_0) = 0.95$.