

TILTA1 Matemaattinen tilastotiede

Tutustumista R-ohjelmiston käyttöön 2

viikko 37, 2007

- Käytetään aineistoa verkosta: http://mtl.uta.fi/tilasto/mtt06/Datat/kaivos_omn.dat, jossa on muuttuja *days* ($n=109$). Aineistossa on aikajärjestyksessä pahojen kaivosonnettomuuksien väliajat päivinä ajanjaksolta 6.12.1875–29.5.1951.
 - Tutustu aineistoon: laske aineistosta perustunnuslukuja ja piirrä histogrammi.
 - Oletetaan, että todelliset luokkarajat ovat: -0.5, 49.5, 99.5, 149.5, 199.5, 249.5, 299.5, 349.5, 399.5, 499.5, 599.5, 699.5, 799.5, 899.5, 999.5, 1999.5
 - Piirrä koko aineistosta histogrammi ja muodosta vastaava frekvenssitaulukko.
 - Piirrä erilliset histogrammit 56 ensimmäisestä ja 53 viimeisestä havainnosta.
 - Vertaa kohdissa (i) ja (ii) tehtyjä kuvioita.
 - Piirrä kuvio muuttujan *days* empiirisestä kertymäfunktioista.
 - Laske myös
 - empiirisen kertymäfunktion arvo pisteessä 60 ja
 - empiirisen jakauman arvo $P_{109}(149.5,199.5)$ empiirisen kertymäfunktion avulla.
- Tutustutaan normaalijakaumaan ja siitä generointiin.
 - Generoi 100 satunnaislukua $N(0,1)$ -jakaumasta ja piirrä tuloksista histogrammi. Kokeile generointia myös $N(50,5^2)$ -jakaumasta.
 - Mikä on $N(0,1)$ -jakauman *tiheysfunktion* arvo, kun satunnaismuuttuja $X=0$? Entä mikä on $N(30,8^2)$ -jakauman tiheysfunktion arvo kun $X=20$? Piirrä tiheysfunktioiden kuvaajat molempien jakaumien tapauksissa.
 - Mikä on $N(0,1)$ -jakauman *kertymäfunktion* arvo, kun $X=0$? Entä mikä on $N(30,8^2)$ -jakauman kertymäfunktion arvo, kun $X=20$? Piirrä kertymäfunktioiden kuvaajat molemmille jakaumille.
 - Laske 95 %:n kvantiili $N(0,1)$ -jakaumasta ja $N(3,1)$ -jakaumasta.
- Tutustutaan binomijakaumaan ja siitä generointiin.
 - Laske $P(45 < X < 55)$, kun $X \sim \text{Bin}(100,0.5)$.
 - Generoi 10 lukua binomijakaumasta, kun $X \sim \text{Bin}(50,0.2)$.
- (Kirjan/luentorungon luvun 1 harjoitus 4.) Heitä harhatonta noppaa 60, 120, 240, 480, 960 ja 2000 kertaa ja laske eri silmälukujen suhteelliset frekvenssit eri heittosarjoissa. Piirrä myös suhteellisten frekvenssien histogrammat. Miten heittojen lkm:n n kasvattaminen vaikuttaa suhteellisiin frekvensseihin?
- (Kirjan/luentorungon luvun 1 harjoitus 9.)
 - Heitetään samanaikaisesti kahta noppaa ja olkoon tulos silmälukujen summa. Olkoot kaikki 36 alkeistapausta ovat yhtä todennäköisiä. Osoita, että tuloksen tn -jakauma on:

Tulos	2	3	4	5	6	7	8	9	10	11	12
$36*tn$	1	2	3	4	5	6	5	4	3	2	1
 - Heitä kahta noppaa 100 kertaa. Vertaa tuloksen havaittuja frekvenssejä odotettuihin frekvensseihin.
- Heitetään kolmea tavallista noppaa. Muodosta R:llä kaikki eri tulosvaihtoehdot (käytä *expand.grid*-funktioita). Laske R:llä todennäköisyys, että heitolla saadaan kolme samaa silmälukua.