

Ohje: Vastaa lyhyesti ja ytimekkäästi, mutta perustele ratkaisusi. Pelkkä lukuarvo vastauksena ei anna pisteitä. Kokeessa on 4 tehtävää, jokaisesta saa 0–6 pistettä. Merkitse jokaiseen vastauspaperiin:

- Kurssin nimi ja koodi
- SUKUNIMI, ETUNIMET ja OPISKELIJANUMERO (tikkukirjaimin)
- Koulutusohjelma ja vuosikurssi
- Päivämäärä ja nimikirjoitus

Sallitut apuvälineet: laskin ja a4-muistilappu (käsin kirjoitettu, tekstiä vain toisella puolella, oikeassa yläkulmassa oma nimi, ei tarvitse palauttaa)

T1 Tavallista kolikkoa heitetään kolme kertaa ja heittojen tuloksia merkitään satunnaisluvuin X_1, X_2, X_3 , missä 0 tarkoittaa klaavaa ja 1 tarkoittaa kruunaa.

- Luettele satunnaisvektorin (X_1, X_2, X_3) kaikki mahdolliset arvot (8 kpl) ja niiden todennäköisyydet. (1p)
- Täydennä luetteloa siten, että kullakin satunnaisvektorin arvolla lasket, minkä arvon tällöin saavat satunnaismuuttujat $S = X_1 + X_2$ ja $T = X_1 + X_2 + X_3$. (1p)
- Määritä satunnaismuuttujien S ja T jakaumat, kumpikin erikseen taulukkona, ja laske S :n ja T :n odotusarvot. (2p)
- Määritä satunnaismuuttujien S ja T yhteisjakauma taulukkona ja perustele sen nojalla, ovatko S ja T stokastisesti riippumattomat. Määritä lisäksi ehdolliset todennäköisyydet $\Pr(T = 3 | S = 2)$ ja $\Pr(S = 2 | T = 3)$. (2p)

T2 Erään Twitter-viestin arvioidaan olevan joko henkilön 1 tai 2 kirjoittama (todennäköisyyksin 0.9 ja 0.1). Tulkitaan kirjoittajan henkilöllisyys satunnaismuuttujaksi Θ , joka saa arvon 1 tai 2. Henkilön 1 viesteissä kullakin rivillä on kirjoitusvirheiden määrä (muista riveistä riippumatta) diskreetisti tasajakautunut joukossa $\{0, 1, 2, 3, 4, 5\}$. Henkilön 2 viesteissä kullakin rivillä on kirjoitusvirheiden määrä (muista riveistä riippumatta) binomijakautunut parametrein $n = 10$ ja $p = 0.3$.

- Millä todennäköisyydellä viestin ensimmäiseltä riviltä ei löydy kirjoitusvirheitä? Ota huomioon ennakoarvio kirjoittajan henkilöllisyydestä. (2p)
- Viestin ensimmäiseltä riviltä ei löydy kirjoitusvirheitä. Määritä viestin kirjoittajan henkilöllisyyden $\Theta \in \{1, 2\}$ posteriorijakauma ensimmäisen rivin perusteella. (2p)
- Viestin toiselta riviltä löydetään 4 kirjoitusvirhettä. Määritä Θ :n posteriorijakauma molempien rivien havaintojen perusteella. (1p)
- Viestin kolmannelta riviltä löydetään 8 kirjoitusvirhettä. Määritä Θ :n posteriorijakauma tähänastisten havaintojen valossa. (Perusteltu päätelmä riittää, ei tarvitse laskea.) (1p)

T3 Eri matkustuspäivien odotusajat (min) eräällä bussipysäkillä ovat toisistaan riippumattomia ja noudattavat jatkuvan välin $[0, \theta]$ tasajakaumaa tiheysfunktiona

$$f(t|\theta) = \begin{cases} \frac{1}{\theta}, & 0 \leq t \leq \theta, \\ 0, & \text{muuten.} \end{cases}$$

Neljänä eri matkustuspäivänä on havaittu odotusajat $x_1 = 2$, $x_2 = 5$, $x_3 = 5$ ja $x_4 = 7$. Auta Ronaldia, Karlia ja Thomasia estimoimaan parametrin θ arvo näiden havaintojen pohjalta.

- (a) Ronald päättää käyttää suurimman uskottavuuden estimaattia. Laske tämän estimaatin arvo datajoukolle $x = (x_1, \dots, x_4)$. (2p)
- (b) Karl päättelee, että jos matka-aika on tasajakautunut välillä $[0, \theta]$, niin siitä määräytyy myös matka-ajan odotusarvo μ . Hän tulkitsee μ :n estimoitavaksi parametriksi ja käyttää sen estimaattina havaintojen keskiarvoa

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Sitten hän kääntäen päättelee, että havainnot ovat tulleet siitä tasajakaumasta, jonka odotusarvo on $\hat{\mu}$. Laske, mihin tasajakaumaan Karl päätyy. (2p)

- (c) Thomas tulkitsee tuntemattoman parametrin satunnaismuuttujaksi Θ ja valitsee priorijakauman tiheysfunktioiksi

$$f_0(\theta) = \begin{cases} c \cdot \theta^8 \cdot (10 - \theta), & 0 \leq \theta \leq 10, \\ 0, & \text{muuten,} \end{cases}$$

missä $c > 0$ on eräs normalisointivakio. Thomas päättää käyttää estimaattina Θ :n posteriorijakauman moodia eli maksimikohtaa. Määritä Thomasin laskema estimaatti. (Huom: Sinun ei tarvitse ratkaista normalisointivakion c arvoa.) (2p)

T4 Kun satunnaisotannalla valitun 2 000 suomalaisen aikuisen pituudet (cm) mitattiin, saatiin keskiarvoksi 170.9 ja keskihajonnaksi 11.2. Määritä mittaustulosten perusteella kaikkien suomalaisten aikuisten keskipituudelle μ :

- (a) piste-estimaatti, (1p)
- (b) likiarvoinen luottamustason 99% luottamusväli. (2p)
- (c) Jos yo. mittaus uusittaisiin 20 000 suomalaisen otokselle, niin miten paljon kapeampi luottamusväli todennäköisesti saataisiin samalla luottamustasolla? (1p)
- (d) Onko (b)-kohdassa tarpeen olettaa, että suomalaisaikuisen pituudet koko väestössä noudattavat likimain normaalijakaumaa? (1p)
- (e) Selitä, mitä tarkoittaa harhaton estimaattori. Onko (a)-kohdassa raportoimasi piste-estimaatti laskettu harhattomalla estimaattorilla? Perustele. (1p)

Normaalijakauman taulukko

Allaolevaan taulukkoon on koottu lukuarvoja normitetun normaalijakauman kertymäfunktiolle

$$F_Z(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999