

Kirjoita selvästi jokaiseen vastauspaperiin:

- Kurssin koodi, päivämäärä, kokeen tyyppi (Tentti/Vk1/Vk2)
- Opiskelijanumero + kirjain
- TEKSTATEN sukunimi ja kaikki etunimet
- Koulutusohjelma ja vuosikurssi
- Nimikirjoitus



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

- Jos suoritat tentin, vastaa kysymyksiin 1, 3, 5, 8 (merkitty *).
- Jos suoritat välikokeen 1 erikoisluvalla ensimmäistä kertaa, vastaa kysymyksiin 1–4.
- Jos suoritat välikokeen 2 erikoisluvalla ensimmäistä kertaa, vastaa kysymyksiin 5–8.

Sallitut apuvälineet: Laskin, Mellinin tilastolliset taulukot ja a4-muistilappu (käsinkirjoitettu, tekstiä vain toisella puolella, oikeassa yläkulmassa opiskelijan nimi)

- * Ovatko seuraavat väittämät totta? Vastaa kyllä tai ei. (1 p/kohta)
 - Jatkuvan satunnaisvektorin komponentit ovat jatkuvia satunnaismuuttujia.
 - Diskreettien satunnaismuuttujien yhteisjakauma on diskreetti.
 - Riippumattomien reaaliarvoisten satunnaismuuttujien korrelaatio on nolla.
 - Tapahtuma on satunnaisilmiön mahdollisten realisaatioiden joukon osajoukko.
 - On olemassa satunnaismuuttuja X , jolle $\Pr(X = t) = 0$ kaikilla reaaliluvuilla t .
 - Jos satunnaismuuttujilla X ja Y on sama tiheysfunktio, on niillä sama varianssi.
- Hyvin sekoitetussa urnassa on 5 punaista ja 7 sinistä palloa. Onnetar poimii urnasta yksi kerrallaan kolme palloa side silmillään, palauttamatta poimittuja palloja urnaan. (1 p/kohta)
 - Mikä on todennäköisyys, että ensimmäinen poimittu pallo on punainen?
 - Mikä on todennäköisyys, että kolmas poimittu pallo on punainen?
 - Mikä on todennäköisyys, että jokainen poimittu pallo on punainen?
 - Mikä on todennäköisyys, että kolmas poimittu pallo on punainen, kun tiedetään, että kaksi ensin poimittua palloa olivat sinisiä?
 - Mikä on todennäköisyys, että täsmälleen kaksi punaista palloa tulee poimituksi?
 - Ovatko tapahtumat $A =$ "Ensimmäinen poimittu pallo on punainen" ja $B =$ "Kolmas poimittu pallo on punainen" riippuvat vai riippumattomat?

3. * Eräessä maassa 10 % asukkaista kantaa tietämättään bakteeria, joka tietyissä olosuhteissa saattaa aiheuttaa erään taudin puhkeamisen. Pikatesti paljastaa bakteerin kantajista 95 %, mutta toisaalta se antaa virheellisen positiivisen tuloksen 5 %:lle niistä, jotka eivät kanno ko. bakteeria.
- (a) Mikä on todennäköisyys, että satunnaisesti valitulle henkilölle testi antaa positiivisen tuloksen? (3 p)
- (b) Mikä on todennäköisyys, että positiivisen testituloksen saanut henkilö todella kantaa ko. bakteeria? (3 p)
4. Alla oleva taulukko esittää diskreettien satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyysfunktioita $f_{X,Y}(x, y) = \Pr(X = x, Y = y)$. (1 p/kohta)

		x		
		-1	0	+1
y	+1	0	3/8	0
	0	3/8	0	1/8
	-1	0	1/8	0

- (a) Määritä X :n pistetodennäköisyysfunktio $f_X(x)$.
- (b) Laske X :n odotusarvo.
- (c) Määritä X :n ehdolliset pistetodennäköisyydet $f_{X|Y}(x|1)$ tapahtuman $\{Y = 1\}$ sattuessa.
- (d) Laske X :n ehdollinen odotusarvo tapahtuman $\{Y = 1\}$ sattuessa.
- (e) Laske X :n ja Y :n kovarianssi.
- (f) Ovatko X ja Y riippuvat vai riippumattomat?
5. * Ovatko seuraavat väittämät totta? Vastaa kyllä tai ei. (1 p/kohta)
- (a) Harhattoman estimaattorin varianssi on nolla.
- (b) Havaittujen lukuarvojen mediaani voi alittaa lukuarvojen keskiarvon.
- (c) Havaittujen lukuarvojen yläkvartiili voi olla aidosti lukuarvojen mediaania pienempi.
- (d) Tilastollisen testin p-arvo kertoo todennäköisyyden, jolla nollahypoteesi hylätään.
- (e) $t(n)$ -jakauman odotusarvo on nolla aina kun vapausasteiden lukumäärä $n \geq 2$.
- (f) Standardoitua normaalijakaumaa noudattavien riippumattomien satunnaismuuttujien neliöiden summa noudattaa χ^2 -jakaumaa.
6. Kaverisi on kehittänyt algoritmin, joka generoi riippumattomia satunnaislukuja normaalijakaumasta $N(\mu, \sigma^2)$. Hän lähettää sinulle sähköpostitse 30 lukua sisältävän satunnaisotoksen. Saamassasi viestissä olevien lukujen keskiarvo on 10.09 ja otosvarianssi 0.0108.
- (a) Määritä 95 % luottamusväli jakauman odotusarvolle μ . (4 p)

- (b) Onko mahdollista, että parametrin μ arvo kuuluu (a)-kohdassa lasketun luottamusvälin ulkopuolelle? (1 p)
- (c) Kaverisi äiti väittää, että satunnaisluvut on generoitu käyttäen odotusarvoa $\mu = 10.10$. Pitääkö väite paikkansa? (Kyllä / Ei / Kyllä 95 % todennäköisyydellä / Kyllä 5 % todennäköisyydellä / Vastausta ei tunneta annettujen tietojen perusteella) (1 p)
7. Kotitalouden energiankulutusta suhteessa kotitalouden tulotasoon tutkittiin eräässä kaupungissa selvittämällä kahdeksan satunnaisesti valitun kotitalouden vuositulot (keUR) ja vuotuinen energiankulutus (MWh). Selvityksen tulokset on listattu alla.

Kotitalous	1	2	3	4	5	6	7	8
Vuositulot	22.7	34.6	45.4	62.6	68.5	85.1	100.4	108.1
Energiankulutus	52.7	87.9	140.6	146.5	190.4	205.1	263.7	266.6

Taulukon luvuista saadaan vuositulojen keskiarvoksi 65.93 ja otoskeskihajonnaksi 30.76, energiankulutuksen keskiarvoksi 169.19 ja otoskeskihajonnaksi 77.16 sekä mitattujen suureiden väliseksi otoskovarianssiksi 2336.23.

- (a) Laske havaittujen suureiden välinen Pearsonin otoskorrelaatio. (1 p)
- (b) Estimoi pienimmän neliösumman menetelmällä kertoimet β_0 ja β_1 lineaariselle regressiomallille, jossa vasteena on vuotuinen energiankulutus ja syötteenä vuositulot. (2 p)
- (c) Ennusta vuotuinen energiankulutus kotitaloudelle, jonka vuositulot ovat 72 keUR. (1 p)
- (d) Ennusta vuotuinen energiankulutus kotitaloudelle, jonka vuositulot ovat 15 keUR. (1 p)
- (e) Perustele kummalla kohtien (c) ja (d) ennusteista olisi kapeampi luottamusväli. (1 p)
8. * Suuren elektroniikka-alan sopimusvalmistajan tehtailla otettiin käyttöön uusi työsuojeluohjelma, jonka tavoitteena on tehostaa yrityksen tuottavuutta parantamalla työturvallisuutta. Alla on listattu työtapaturmien takia menetetyt viikottaiset henkilötyötunnit (keskiarvoistettuna 1 kk ajalta) kymmenessä samankaltaisessa tehtaassa ennen ja jälkeen työsuojeluohjelman käyttöönoton.

Tehtas	1	2	3	4	5	6	7	8	9	10
Ennen	30.5	18.5	24.5	32	16	15	23.5	25.5	28	18
Jälkeen	23	21	22	28.5	14.5	15.5	24.5	21	23.5	16.5
Erotus	7.5	-2.5	2.5	3.5	1.5	-0.5	-1	4.5	4.5	1.5

Yllä listattujen lukujen keskiarvot ovat 23.15 (ennen), 21.00 (jälkeen) ja 2.15 (erotus). Vastaavat otosvarianssit ovat 36.56 (ennen), 19.17 (jälkeen) ja 9.00 (erotus).

Testaa 5 % merkitsevyystasolla, onko työsuojeluohjelmasta ollut hyötyä. Käytä nollahypoteesina oletusta, että työsuojeluohjelmalla ei ole vaikutusta työtapaturmien takia menetettyjen työtuntien lukumäärään. (6 p)