MS-A050X, Todennäköisyyslaskennan ja tilastotieteen peruskurssi
Matematiikan ja systeemianalyysin laitos
Aalto-yliopisto

Tentti
4. huhtikuuta 2017
J Tölle

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 ja ETUNIMET (tikkukirjaimin)
- Opiskelijanumero
- Koulutusohjelma ja vuosikurssi
- Päivämäärä ja nimikirjoitus

Sallitut apuvälineet: Mellinin tilastolliset taulukot, laskin ja a4-muistilappu (käsin kirjoitettu, tekstiä vain toisella puolella, oikeassa yläkulmassa oma nimi)

T1 Oleta, että X on satunnaismuuttuja, joka saa arvot -2, -1, 0, 1, 2; jokaisen niistä $\frac{1}{5}$ todennäköisyydellä. Määrittele $Y = X^2$.

(a) Määritä X:n ja Y:n yhteisjakauma ja Y:n jakauma, **(2p)**

(b) määritä $E(X)$ ja $E(Y)$, **(2p)**

(c) näytä, että $\mathrm{Cov}(X, Y) = 0$, **(1p)**

(d) todista, että X ja Y eivät ole riippumattomia. **(1p)**

T2 Erään yliopiston matematiikan ensimmäisen vuosikurssin ideaalikoko on 90 opiskelijaa. Tässä yliopistossa tiedetään kokemuksesta, että keskimäärin vain 20% opiskelemaan valituista opiskelijoista – toisistaan riippumatta – todella aloittaa opinnot. Tämän kokemuksen perusteella tämä yliopisto ottaa opiskelemaan vuosittain uutta 400 matematiikan opiskelijaa.

(a) Kirjoita ja perustele kaava, joka kuvaa tarkan todennäköisyysjakauman niiden ensimmäisen vuoden opiskelijoiden lukumäärästä, jotka todella aloittavat matematiikan opinnot. (Arvoa ei tarvitse laskea tähän). **(2p)**

(b) Mitkä ovat odotusarvo ja keskihajonta opinnot todella aloittavien ensimmäisen vuoden matematiikan opiskelijoiden lukumäärälle? **(2p)**

(c) Laske approksimoitu todennäköisyys sille, että enemmän kuin 90 opiskelijaa aloittaa ensimmäisellä matematiikan vuosikurssilla. **(2p)**

Huomaa: *Jos X on diskreetti satunnaismuuttuja, laske $P(X = i)$ jatkuvana approksimaationa seuraavalla tavalla $P\left(i - \frac{1}{2} < X < i + \frac{1}{2}\right)$.*

T3 Uudentyyppisten autonrenkaiden valmistaja lupaa, että sen renkaiden kestoikä on keskimäärin 40 000 kilometriä. Tämän väitteen todentamiseksi testataan 12 renkaan otanta. Testattavien renkaiden kestoiät olivat seuraavat (tuhansissa kilometreissä):

Tire no.	1	2	3	4	5	6	7	8	9	10	11	12
Lifetime	36	40	33	38	42	35	37	41	36	37	33	36

(a) Mitä täytyy olettaa, että väite voidaan testata järkevästi t-testillä? **(3p)**

(b) Kuvitellaan, että oletuksesi todennetaan oikeiksi. Luo nollahypoteesi ja vaihtoehtoinen hypoteesi, ja testaa onko väite totta 5% luottamustasolla. **(3p)**

Aputuloksia: *Havaintojen summa (tuhansissa kilometreissä) ja neliöiden summa (miljoonissa neliökilometreissä) ovat:*

$$\sum_{i=1}^{12} x_i = 444 \quad ja \quad \sum_{i=1}^{12} x_i^2 = 16\,518.$$

T4 Ovatko seuraavat väittämät tosia? Vastaus: **1** = tosi, **x** = ei voida tietää annetun tiedon perusteella, tai **2** = epätosi. Perustele vastauksesi.

(a) $A = \emptyset$ ($\emptyset$ osoittaa *tyhjän joukon*) jokaiselle tapahtumalle A, niin että $P(A) = 0$. **(1p)**

(b) Jos X:n ja Y:n yhteisjakauman tiheysfunktio on

$$f(x,y) := \begin{cases} xy, & \text{kun } 0 \leq x \leq \sqrt{2},\ 0 \leq y \leq \sqrt{2} \\ 0, & \text{muissa tapauksissa,} \end{cases}$$

sitten X ja Y ovat riippumattomat. **(1p)**

(c) Oleta, että X_1, X_2, ..., X_n ovat satunnaismuuttujia, jotka ovat riippumattomat ja joilla on identtiset Bernoullin jakaumat. Silloin summalla $X_1 + \ldots + X_n$ on approksimatiivisesti normaali jakauma. **(1p)**

(d) Keräämme kaksi näytettä $(x_1, \ldots, x_n)$, ja $(y_1, \ldots, y_n)$. Haluaisimme testata erotusten joukon $d_i = x_i - y_i$, $1 \leq i \leq n$ nollahypoteesin $H_0 : \mu = 0$ ja vaihtoehtoisen hypoteesin $H_1 : \mu \neq 0$ mukaan. Tässä tapauksessa kaksisuuntainen parittainen t-testi on hyvä valinta tilastolliseksi hypoteesitestiksi. **(1p)**

(e) Testin p-arvo osoittaa ehdollisen todennäköisyyden – ehdolla, että nollahypoteesi on tosi – että sattumanvarainen testisuure saa arvot, jotka ovat vähintään niin poikkeuksellisia kuin aineistosta lasketulla suureella. **(1p)**

(f) Todennäköisyys, että nollahypoteesi on tosi, lasketaan $1 - \alpha$, jos α osoittaa hylkäämisvirheen. **(1p)**