Aalto-yliopisto

Gripenberg

MS-A0502 Todennäköisyyslaskennan ja tilastotieteen peruskurssi
Tentti ja välikoeuusinta 7.1.2015

Kirjoita jokaiseen koepaperiin nimesi, opiskelijanumerosi ym. tiedot ja minkä kokeen suoritat!
Laskin, I. Mellinin tilastolliset taulukot, Matlab/Octave-funktiolista ja muistiinpanolappu ovat sallittuja apuvälineitä!
Kirjoita välivaiheet näkyviin!
Käytä joko taulukoita tai kirjoita millä Matlab/Octave-komennoilla voisit laskea tarvittavat kertymäfunktioiden tai niiden käänteisfunktioiden arvot ja minkälaisia päätelmiä, esimerkiksi merkitsevyystasojen suhteen, näiden arvojen perusteella voisit tehdä.

Tentti: Tehtävät 1, 4, 6, 7 ja 8.
1. välikoe: Tehtävät 1, 2, 3 ja 4.
2. välikoe: Tehtävät 5, 6, 7 ja 8.

1.

(a) Heitetään virheetöntä noppaa kaksi kertaa. Mikä on todennäköisyys, että silmälukujen summa on 9?

(b) Heitetään virheetöntä noppaa 900 kertaa. Mikä on todennäköisyys, että saadaan 6 korkeintaan 130 kertaa. Käytä normaaliapproksimaatiota.

2. Uurnassa on 3 valkoista ja 2 mustaa palloa. Poimitaan uurnasta satunnaisesti pallo. Jos se on valkoinen niin uurnaan laitetaan takaisin 3 valkoista palloa ja 1 musta pallo ja jos se on musta niin uurnaan laitetaan takaisin 2 mustaa palloa ja 1 valkoinen. Tämän jälkeen poimitaan uurnasta taas satunnaisesti uusi pallo ja menetellään samalla tavalla kuin ensimmäisellä kerralla. Määritä tämän jälkeen uurnassa olevien valkoisten pallojen lukumäärän pistetodennäköisyysfunktio. Käytä puuverkkoa.

3.

Pistetodennäköisyysfunktion f_{XY} arvot on annettu seuraavassa taulukossa:

$f_{XY}(x,y)$		$Y=1$	$Y=2$	$Y=3$
X	0	$\frac{2}{20}$	$\frac{2}{20}$	$\frac{1}{20}$
	1	0	$\frac{3}{20}$	$\frac{2}{20}$
	2	$\frac{1}{20}$	$\frac{1}{20}$	0
	3	$\frac{1}{20}$	$\frac{4}{20}$	$\frac{3}{20}$

(a) Mistä nähdään, että kyseessä todella on pistetodennäköisyysfunktio.

(b) Määritä satunnaismuuttujan X pistetodennäköisyysfunktio (eli määritä sen reunajakauma).

(c) Määritä satunnaismuuttujan $(X|Y = 2)$ pistetodennäköisyysfunktio eli X:n jakauma ehdolla $Y = 2$.

4. Alueen sorasta on 70% hyvää ja 30% huonoa tiettyyn rakennustarkoitukseen. Yhdestä alueen sorakuopasta otettu näyte testattiin testillä, joka todennäköisyydellä 0.8 osoittaa soran hyväksi (ja todennäkösyydellä 0.2 huonoksi), jos sora on hyvää ja todennäköisyydellä 0.9 huonoksi (ja todennäköisyydellä 0.1 hyväksi), jos sora on huonoa. Jos testi näytti että sora oli hyvää, mikä on todennäköisyys että kuopan sora todella on hyvää?

5.

(a) Satunnaismuuttujasta X, joka on Poisson(λ)-jakautunut on saatu havainnot 5, 6, 4, 7 ja 6. Laske momenttimenetelmällä λ:n estimaatti.

(b) Diskreetin satunnaismuuttujan Y jakauma riippuu parametristä $\theta \in [0, \frac{1}{2}]$ siten, että

$$\Pr(Y = j) = \begin{cases} 1 - 2\theta, & j = 0, \\ \theta, & j \in \{-1, 1\}, \\ 0, & \text{muuten.} \end{cases}$$

Satunnaismuuttujasta Y on saatu seuraavat havaintoarvot: 0, 1, -1, 0, 1, 0, -1, 1, -1 ja 0. Määritä θ:n estimaatti suurimman uskottavuuden menetelmällä.

6. Sairaalan poliklinikalle tulee kokemuksen perusteella keskimäärin 9 potilasta tunnissa. Eräänä päivänä, kun keli on ollut erityisen liukas tulee 12 tunnin aikana 130 potilasta. Oletetaan, että potilaiden lukumäärä aikayksikössä noudattaa Poisson-jakaumaa. Testaa merkitsevyystasolla 0.05 nollahypoteesia, jonka mukaan potilaiden lukumäärän odotusarvo kyseisenä päivänä ei ole ainakaan tavanomaista suurempi. Käytä normaaliapproksimaatiota.

7. Havainnot X_i, $i = 1, \dots, 21$ ja Y_i, $i = 1, \dots, 31$ ovat satunnaisotoksia jakaumista $\mathrm{N}(\mu_X, \sigma^2)$ ja $\mathrm{N}(\mu_Y, \sigma^2)$. Havaintoarvoista on laskettu $\overline{x} = 2.8$, $\sum_{i=1}^{21} |x_i - \overline{x}| = 26$, $\sum_{i=1}^{21} (x_i - \overline{x})^2 = 46$, $\overline{y} = 2.0$, $\sum_{i=1}^{31} |y_i - \overline{y}| = 35$ ja $\sum_{i=1}^{31} (y_i - \overline{y})^2 = 81$. Testaa merkitsevyystasolla 0.05 hypoteeseja

(a) $\mathrm{H}_0 : \mu_X = 3.5$, (jolloin sinun ei pidä ottaa huomioon y_i-havaintoarvoja),

(b) $\mathrm{H}_0 : \mu_X = \mu_Y$.

8. Lukiokoulutusta antavien oppilaitosten lukumäärät olivat vuosina 2002–2013 seuraavat:

Vuosi	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Lkm.	484	483	479	471	461	449	449	441	439	433	428	421

Oletetaan, että (ainakin approksimatiivisesti) $Y_j = \beta_0 + \beta_1 x_j + \varepsilon_j$, missä Y_j on lukiokoulutusta antavien oppilaitosten lukumäärä vuonna x_j ja satunnaismuuttujat ε_j ovat $\mathrm{N}(0, \sigma^2)$-jakautuneita ja riippumattomia.

Määritä parametrin β_1 pienimmän neliösumman estimaatti ja testaa merkitsevyystasolla 0.01 nollahypoteesia, että lukiokoulutusta antavien oppilaitosten lukumäärän odotusarvo vuonna 2016 on korkeintaan 390.

Alla olevassa taulukossa on annettu joitakin tunnuslukuja:

$\overline{y}$	s_x^2	s_y^2	s_{xy}
453.17	13	485.97	-78.636