Johdatus tilastolliseen päättelyyn, erilliskoe 15.11.2012

Taulukkokirjan käyttö ei ole sallittua.

1. Poimimme satunnaisotoksen Poissonin jakaumasta, jonka odotusarvo $\mu > 0$ on tuntematon. Tämän jakauman pistetodennäköisyysfunktio on

$$g(y;\mu) = \mathrm{e}^{-\mu}\,\frac{\mu^y}{y!}, \qquad y = 0, 1, 2, \ldots$$

Kirjoita havaintoja

$$0, 1, 2, 1, 2, 3$$

vastaava uskottavuusfunktio sekä johda huolellisesti perustellen suurimman uskottavuuden estimaatti parametrille μ.

2. Lasket saamastasi aineistosta tietyllä tilasto-ohjelmalla odotusarvoparametrille μ sen kaksisuuntaisen t-luottamusvälin luottamustasolla 0.95 ja saat tulokseksi välin $[0.23, 4.65]$. Sitten tuhoat epähuomiossa aineiston. Harmiksesi kuulet, että sinun olisi oikeastaan pitänyt testata nollahypoteesia, jonka mukaan $\mu = 0$, kun vastahypoteesi on kaksisuuntainen.

Sano kustakin seuraavasta väitteestä, onko se oikein vai väärin. Anna valinnallesi lyhyt perustelu.

a) Luottamustason 0.99 kaksisuuntainen t-luottamusväli sisältyisi väliin $[0.23, 4.65]$.

b) Kun merkitsevyystaso on 0.05, edellä kuvattu testi hylkää nollahypoteesin.

c) Edellä kuvatun testin p-arvo on suurempi kuin 0.05.

3. Olkoon $Y_1, \ldots, Y_n$ satunnaisotos normaalijakaumasta $N(\mu, \sigma^2)$, jossa molemmat parametrit ovat tuntemattomia. Tällöin parametreja on tapana estimoida otoskeskiarvolla $\bar{Y}$ sekä otosvarianssilla S^2, jossa tunnetusti $\frac{n-1}{\sigma^2}\,S^2 \sim \chi^2_{n-1}$.

a) Anna kaava otosvarianssille S^2. Kurssilla opittiin, että estimaattori S^2 on harhaton. Selitä, mitä tämä tarkoittaa.

b) Muodosta varianssiparametrille 95 %:n luottamusväli, kun aineistossa

$$\bar{y} = -0.019, \quad s^2 = 0.0339, \quad n = 9.$$

KÄÄNNÄ!

4. Kahden amerikkalaisen puusammakkolajin (*Hyla chrysoscelis* ja *Hyla versicolor*) kutsuhuutojen kestoja (sekuntia) tutkittiin Lousianassa eri urosyksilöiltä. Havainnoista esitetään yhteenveto seuraavassa taulukossa.

	otoskoko	keston keskiarvo	keston keskihajonta
H. chrysoscelis	43	0.65	0.18
H. versicolor	12	0.54	0.14

Nämä kaksi lajia ovat lähes mahdottomia erottaa toisistaan ulkonäön perusteella. Tutki annetun aineiston perusteella, eroavatko lajit toisistaan kutsuhuudon keston perusteella. Kohtelemme eri lajien kutsuhuutojen jakaumia normaalijakaumina, joilla on sama varianssi mutta mahdollisesti erisuuret odotusarvot.

Valintasi mukaan vastaa **yhteen** kysymyksistä (1) tai (2) tai teoriakysymykseen (3). Jokaisella näistä tavoista voit saada tehtävästä täydet pisteet.

(1) Laske kaksisuuntainen luottamusväli lajien kutsuhuutojen odotusarvojen erotukselle luottamustasolla 0.95. Kerro lisäksi, tukeeko laskemasi luottamusväli sitä päätelmää, että eri lajien kutsuhuudoilla on samat odotusarvot vai sitä päätelmää, että ne ovat erisuuret.

(2) Testaa merkitsevyystasolla 0.05, ovatko näiden kahden lajin kutsuhuutojen odotusarvot samoja.

(3) Totea ensin, minkälaisia oletuksia aineiston jakaumasta tässä ongelmassa tehdään, ja sen jälkeen johda sellaiset kaavat, joihin luvut sijoittamalla saisit vastattua joko kysymykseen (1) tai kysymykseen (2).

5. Eräässä Gregor Mendelin (1822–1884) tekemässä hernekasvien risteytyskokeessa saatiin tulokseksi 556 kasvia. Näistä 423 kasvin herneet olivat sileitä ja 133 kasvin herneet ryppyisiä. Testaa merkitsevyystasolla 0.05, ovatko havaitut frekvenssit yhteensopivia sen (Mendelin teorioiden mukaisen) hypoteesin kanssa, että ko. kokeessa saadaan todennäköisyydellä 3/4 siletä herneitä ja todennäköisyydellä 1/4 ryppyisiä herneitä tuottava kasvi.

Taulukko 1: t-jakauman u-yläkvantiileja $t_\nu(u)$ eri vapausasteluvun ν arvoilla. Tässä $u = P(X > t_\nu(u))$, kun $X \sim t_\nu$.

$\nu \backslash u$	0.05	0.025	0.01
1	6.314	12.706	31.821
2	2.920	4.303	6.965
3	2.353	3.182	4.541
4	2.132	2.776	3.747
5	2.015	2.571	3.365
6	1.943	2.447	3.143
7	1.895	2.365	2.998
8	1.860	2.306	2.896
9	1.833	2.262	2.821
10	1.812	2.228	2.764
11	1.796	2.201	2.718
12	1.782	2.179	2.681
13	1.771	2.160	2.650
14	1.761	2.145	2.624
15	1.753	2.131	2.602
16	1.746	2.120	2.583
17	1.740	2.110	2.567
18	1.734	2.101	2.552
19	1.729	2.093	2.539
20	1.725	2.086	2.528
21	1.721	2.080	2.518
22	1.717	2.074	2.508
23	1.714	2.069	2.500
24	1.711	2.064	2.492
25	1.708	2.060	2.485
26	1.706	2.056	2.479
27	1.703	2.052	2.473
28	1.701	2.048	2.467
29	1.699	2.045	2.462
30	1.697	2.042	2.457

$\nu \backslash u$	0.05	0.025	0.01
31	1.696	2.040	2.453
32	1.694	2.037	2.449
33	1.692	2.035	2.445
34	1.691	2.032	2.441
35	1.690	2.030	2.438
36	1.688	2.028	2.434
37	1.687	2.026	2.431
38	1.686	2.024	2.429
39	1.685	2.023	2.426
40	1.684	2.021	2.423
41	1.683	2.020	2.421
42	1.682	2.018	2.418
43	1.681	2.017	2.416
44	1.680	2.015	2.414
45	1.679	2.014	2.412
46	1.679	2.013	2.410
47	1.678	2.012	2.408
48	1.677	2.011	2.407
49	1.677	2.010	2.405
50	1.676	2.009	2.403
51	1.675	2.008	2.402
52	1.675	2.007	2.400
53	1.674	2.006	2.399
54	1.674	2.005	2.397
55	1.673	2.004	2.396
56	1.673	2.003	2.395
57	1.672	2.002	2.394
58	1.672	2.002	2.392
59	1.671	2.001	2.391
60	1.671	2.000	2.390

Taulukko 2: Standardinormaalijakauman $N(0,1)$ yläkvantiileja z_u. Tässä $u = P(Z > z_u)$, kun $Z \sim N(0,1)$.

0.05	0.025	0.01	0.005	0.001
1.645	1.960	2.326	2.576	3.090

Taulukko 3: χ^2_ν-jakauman alakvantiileja $q_\nu(u)$ eri vapausasteluvun ν arvoilla. Tässä $u = P(X < q_\nu(u))$, kun $X \sim \chi^2_\nu$.

$\nu \backslash u$	0.001	0.005	0.01	0.025	0.05	0.95	0.975	0.99	0.995	0.999
1	0.000	0.000	0.000	0.001	0.004	3.841	5.024	6.635	7.879	10.828
2	0.002	0.010	0.020	0.051	0.103	5.991	7.378	9.210	10.597	13.816
3	0.024	0.072	0.115	0.216	0.352	7.815	9.348	11.345	12.838	16.266
4	0.091	0.207	0.297	0.484	0.711	9.488	11.143	13.277	14.860	18.467
5	0.210	0.412	0.554	0.831	1.145	11.070	12.833	15.086	16.750	20.515
6	0.381	0.676	0.872	1.237	1.635	12.592	14.449	16.812	18.548	22.458
7	0.598	0.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278	24.322
8	0.857	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955	26.124
9	1.152	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589	27.877
10	1.479	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188	29.588