

## Johdatus tilastolliseen päättelyyn Harjoitus 3 (12. 4.–16. 4. 2010)

1. Olkoot  $y_1, \dots, y_n$  reaalilukuja ja  $\bar{y} = \sum_{i=1}^n y_i/n$  niiden aritmeettinen keskiarvo. Näytä, että

$$\sum_{i=1}^n y_i^2 = \sum_{i=1}^n (y_i - \bar{y})^2 + n\bar{y}^2.$$

(Luennot sovelsimme tätä muotoa  $y_i = x_i - \mu$  oleviin lukuihin.)

2. Tarkastellaan riippumatonta toistokoetta, jossa ”onnistumistodennäköisyys” on  $0 \leq \theta \leq 1$ , toistojen lukumäärä  $n = 10$  ja on havaittu  $t = 6$  onnistumista. Palauta luennoilta mieleen tätä havaintoa vastaava uskottavuusfunktio ja määritä 50 %:n ja 10 %:n uskottavuusvälit parametrille  $\theta$ . (Riittää antaa välien päätepisteet yhden desimaalin tarkkuudella; käytä muistiinpanojen sivun 11 kuvaa apuna.)
3. Olkoon  $X_1, \dots, X_n$  riippumaton otos normaalijakaumasta  $N(\mu, 1)$ . Kuinka suuri on otoskoon  $n$  oltava, jos halutaan, että havaintojen  $x_1, \dots, x_n$  perusteella muodostettava 95 %:n luottamusväli  $\mu$ :lle on pituudeltaan (noin) 0.1 eli muotoa  $(\hat{\mu} - 0.05, \hat{\mu} + 0.05)$ ? Entä kuinka suureksi  $n$  on valittava, jos halutaan kymmenkertainen tarkkuus eli luottamusvälin pituudeksi vain (noin) 0.01?

[Muista, että jos  $Z \sim N(0, 1)$ , niin  $P(Z > 1.96) \approx 0.025$ .]

4. Halutaan selvittää tietynmerkkisen savukkeen keskimääräinen hääpitoisuus  $\mu$ . Tätä varten mitattiin 16:sta satunnaisesti valitusta savukeyksilöstä hääpitoisuudet. Mittaustulokset vaihtelevat savukkeiden luonnollisesta eroista ja mittausvirheistä johtuen, ja oletamme, että vaihtelua voidaan pitää normaalisti jakautuneena. Pitoisuuksien keskiarvoksi otoksessa saatiin  $\bar{x} = 9.7$  milligrammaa ja keskihajonnaksi  $s = 1.4$  milligrammaa. Mikä on  $\mu$ :n suurimman uskottavuuden estimaatti ja mikä on sen 95 %:n luottamusväli? Miten luottamusväli muuttuisi, jos otoskoko olisikin ollut 16:n sijasta 61?

5. [Tämä tehtävä käsitellään vasta 4. harjoituksissa.]

Henkilö on saanut puutavaraliikkeeltä tuumanpaksuista tuppeensahattua koivulautaa. Liike lupaa, että laudan paksuuden toleranssit ovat  $\pm 0.05$  tuumaa, minkä tulkitsemme tarkoittavan sitä, että paksuuden keskihajonta  $\sigma$  on korkeintaan 0.05 tuumaa. Henkilö tekee saamastaan lautaerästä 10 paksuusmittausta, saaden tulokset 0.89, 0.92, 0.96, 0.96, 0.99, 1.02, 1.03, 1.06, 1.08 ja 1.09 tuumaa. Laske näiden mittausten keskihajonta  $s$  ja testaa  $\chi^2$ -testillä sekä merkitsevyytasolla 0.05 liikkeen esittämän väitteen paikkansapitävyyttä. Voimme olettaa, että paksuuden vaihtelut noudattavat kyllin tarkasti normaalijakaumaa.

**Kääntöpuolella** on Studentin  $t$ -jakauman ja  $\chi^2$ -jakauman taulukot tehtäviä 4 ja 5 varten. Vasemman asteikon luvut  $f$  ja  $m$  tarkoittavat vapausasteiden määrää. Itse taulukon luvut ovat ns. prosenttipisteitä, jotka kertovat sen kohdan ao. jakaumassa, josta lähtien jakauman oikean hännän todennäköisyysmassa on (vaaka-asteikolla annettu luku)  $\alpha$ .

**TABLE 2**  
Student(*f*) distribution: right-tail percentage points  $t_{\alpha}^f$ ;  $\alpha = P(t \geq t_{\alpha}^f)$

$\alpha \backslash f$	0.25	0.1	0.05	0.025	0.01	0.005
1	1.000	3.078	6.314	12.706	31.821	63.657
2	0.816	1.886	2.920	4.303	6.965	9.925
3	0.765	1.638	2.353	3.182	4.541	5.841
4	0.741	1.533	2.132	2.776	3.747	4.604
5	0.727	1.476	2.015	2.571	3.365	4.032
6	0.718	1.440	1.943	2.447	3.143	3.707
7	0.711	1.415	1.895	2.365	2.998	3.499
8	0.706	1.397	1.860	2.306	2.896	3.355
9	0.703	1.383	1.833	2.262	2.821	3.250
10	0.700	1.372	1.812	2.228	2.764	3.169
11	0.697	1.363	1.796	2.201	2.718	3.106
12	0.695	1.356	1.782	2.179	2.681	3.055
13	0.694	1.350	1.771	2.160	2.650	3.012
14	0.692	1.345	1.761	2.145	2.624	2.977
15	0.691	1.341	1.753	2.131	2.602	2.947
16	0.690	1.337	1.746	2.120	2.583	2.921
17	0.689	1.333	1.740	2.110	2.567	2.898
18	0.688	1.330	1.734	2.101	2.552	2.878
19	0.688	1.328	1.729	2.093	2.539	2.861
20	0.687	1.325	1.725	2.086	2.528	2.845
21	0.686	1.323	1.721	2.080	2.518	2.831
22	0.686	1.321	1.717	2.074	2.508	2.819
23	0.685	1.319	1.714	2.069	2.500	2.807
24	0.685	1.318	1.711	2.064	2.492	2.797
25	0.684	1.316	1.708	2.060	2.485	2.787
26	0.684	1.315	1.706	2.056	2.479	2.779
27	0.684	1.314	1.703	2.052	2.473	2.771
28	0.683	1.313	1.701	2.048	2.467	2.763
29	0.683	1.311	1.699	2.045	2.462	2.756
30	0.683	1.310	1.697	2.042	2.457	2.750
40	0.681	1.303	1.684	2.021	2.423	2.704
60	0.679	1.296	1.671	2.000	2.390	2.660
120	0.677	1.289	1.658	1.980	2.358	2.617
$\infty$	0.674	1.282	1.645	1.960	2.326	2.576

**TABLE 3**  
Chi-Square(*m*) distribution: right-tail percentage points  $\chi_{\alpha}^2$ ;  $\alpha = P(\chi^2 \geq \chi_{\alpha}^2)$

$\alpha \backslash m$	0.995	0.975	0.950	0.900	0.850	0.800	0.750	0.700	0.650	0.600	0.550	0.500	0.450	0.400	0.350	0.300	0.250	0.200	0.150	0.100	0.050	0.025	0.010	0.005	
1	0.0043927	0.0039821	3.84146	5.02389	6.63490	7.87944																			
2	0.0100025	0.050636	5.99147	7.37776	9.21034	10.5966																			
3	0.071721	0.215795	7.81473	9.34840	11.3449	12.8381																			
4	0.206990	0.484419	9.48773	11.1433	13.2767	14.8602																			
5	0.411740	0.831211	11.0705	12.8325	15.0863	16.7496																			
6	0.675727	1.237347	12.5916	14.4494	16.8119	18.5476																			
7	0.989265	1.68987	14.0671	16.0128	18.4753	20.2777																			
8	1.344419	2.17973	15.5073	17.5346	20.0902	21.9550																			
9	1.734926	2.70039	16.9190	19.0228	21.6660	23.5893																			
10	2.15585	3.24697	18.3070	20.4831	23.2093	25.1882																			
11	2.60321	3.81575	19.6751	21.9200	24.7250	26.7569																			
12	3.07382	4.40379	21.0261	23.3367	26.2170	28.2995																			
13	3.56503	5.00874	22.3621	24.7356	27.6883	29.8194																			
14	4.07468	5.62872	23.6848	26.1190	29.1413	31.3193																			
15	4.60094	6.26214	24.9958	27.4884	30.5779	32.8013																			
16	5.14224	6.90766	26.2962	28.8454	31.9999	34.2672																			
17	5.69724	7.56418	27.5871	30.1910	33.4087	35.7185																			
18	6.26481	8.23075	28.8693	31.5264	34.8053	37.1564																			
19	6.84398	8.90655	30.1435	32.8523	36.1908	38.5822																			
20	7.43386	9.59083	31.4104	34.1696	37.5662	39.9968																			
21	8.03366	10.28293	32.6705	35.4789	38.9321	41.4010																			
22	8.64272	10.9823	33.9244	36.7807	40.2894	42.7956																			
23	9.26042	11.6885	35.1725	38.0757	41.6384	44.1813																			
24	9.88623	12.4011	36.4151	39.3641	42.9798	45.5585																			
25	10.5197	13.1197	37.6525	40.6465	44.3141	46.9278																			
26	11.1603	13.8439	38.8852	41.9232	45.6417	48.2899																			
27	11.8076	14.5733	40.1133	43.1944	46.9630	49.6449																			
28	12.4613	15.3079	41.3372	44.4607	48.2782	50.9933																			
29	13.1211	16.0471	42.5569	45.7222	49.5879	52.3356																			
30	13.7867	16.7908	43.7729	46.9792	50.8922	53.6720																			
40	20.7065	24.4331	55.7585	59.3417	63.6907	66.7659																			
50	27.9907	32.3574	67.5048	71.4202	76.1539	79.4900																			
60	35.5346	40.4817	79.0819	83.2976	88.3794	91.9517																			
70	43.2752	48.7576	90.5312	95.0231	100.425	104.215																			
80	51.1720	57.1532	101.879	106.629	112.329	116.321																			
90	59.1963	65.6466	113.145	118.136	124.116	128.299																			
100	67.3276	74.2219	124.342	129.561	135.807	140.169																			

Source: "Tables of the Percentage Points of the  $\chi^2$ -Distribution," *Biometrika*, Vol. 32 (1941), pp. 188-189, by Catherine M. Thompson; reproduced by permission of Professor E. S. Pearson and the Biometrika Trustees.