פרק ב-3

הסטטיסטיקה של ניתוח שונות.

MSE הוא אמד חסר הטיה לשונות

כבר ראינו שMSE הוא הכללה של Sp בריבוע.

$$E[MS_E] = E[\frac{SS_E}{N-a}] = \frac{1}{N-a}E[\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i\bullet})^2] =$$

$$\frac{1}{N-a}E[\sum_{i=1}^{a}\sum_{j=1}^{n}{y_{ij}}^2 - \frac{1}{n}\sum_{i=1}^{a}{y_{i\bullet}}^2]$$

$$\boxed{y_{ij} = \underbrace{\mu + \tau_i}_{\mu_i} + \tilde{\varepsilon}_{ij}}$$

$$\frac{1}{N-a}E[MS_E] = \frac{1}{N-a}E[\sum_{i=1}^{a}\sum_{j=1}^{n}(\mu + \tau_i + \tilde{\varepsilon}_{ij})^2 - \frac{1}{n}\sum_{i=1}^{a}\left(\sum_{j=1}^{n}(\mu + \tau_i + \tilde{\varepsilon}_{ij})\right)^2] = \ldots$$

$$\ldots = \frac{1}{N-a}[N\mu^2 + n\sum_{i=1}^{a}{\tau_i}^2 + N\sigma^2 - N\mu^2 - n\sum_{i=1}^{a}{\tau_i}^2 - a\sigma^2] = \sigma^2$$

תחת H0 – גם MStreatments הוא אמד חסר הטיה לשונות

$$E[MS_{Treatments}] = E[\frac{SS_{Treatments}}{a-1}] = E[\frac{n\sum_{i=1}^{a}(\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet})^2}{a-1}] = \ldots = \sigma^2 + \frac{n\sum_{i=1}^{a}\tau_i^2}{a-1}$$

תחת H0:

$$E[MS_{Treatments}] = \sigma^2$$

תחת H1:

$$E[MS_{Treatments}] > \sigma^2$$

לקראת פילוג הסטטיסטי F0

$$y_{ij} \sim N(\mu + \tau_i, \sigma^2)$$

$$\boxed{F_0 = \frac{SS_{Treatments}/(a-1)}{SS_E/(N-a)} = \frac{MS_{Treatments}}{MS_E}}$$

$$\frac{SS_{Total}}{\sigma^2} \sim \chi^2_{N-1}$$

לא רק תחת H0

$$\frac{SS_E}{\sigma^2} \sim \chi^2_{N-a}$$

$$H_0: \quad \frac{SS_{Treatments}}{\sigma^2} \sim \chi^2_{a-1}$$

האם SStreatments ו SSe בלתי תלויים?

משפט קוקרן (Cocharan)

משפט: יהיו $Z_i \sim NID(0,1)$ $i=1,\ldots,\nu$ ובנוסף:

$$\sum_{i=1}^{\nu} Z_i^2 = Q_1 + Q_2 + \ldots + Q_s \qquad s \le \nu$$

כאשר מספר דרגות החופש של Q_i הוא ν_i.

אז $Q_1, Q_2, \ldots, Q_s$ הם משתנים מקריים χ^2 ב"ת

עם דרגות חופש $\nu_1, \ldots, \nu_s$,

אם ורק אם $\nu = \nu_1 + \ldots + \nu_s$

יישום משפט קוקרן לניתוח שונות חד-כווני

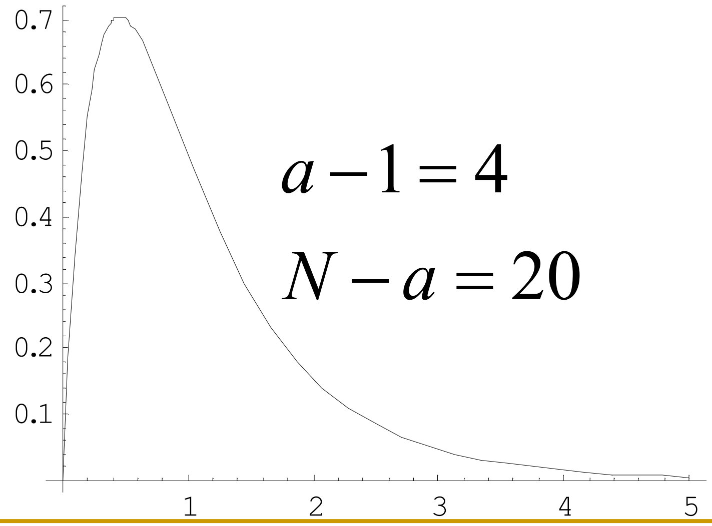
$$SS_{Total} = SS_{Treatments} + SS_E$$

$$\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{\bullet\bullet})^2 \qquad n\sum_{i=1}^{a}(\bar{y}_{i\bullet}-\bar{y}_{\bullet\bullet})^2 \qquad \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i\bullet})^2$$

$$N-1=a\cdot n-1 = a-1 + a\cdot(n-1)=an-a$$

פילוג הסטטיסטי F

$$F_0 = \frac{SS_{Treatments}/(a-1)}{SS_E/(N-a)} = \frac{MS_{Treatments}}{MS_E} = \frac{\chi^2_{a-1}\sigma^2 / a-1}{\chi^2_{N-a}\sigma^2 / N-a} \sim F_{a-1,N-a}$$



מתי נדחה H0?

$$F_0 = \frac{MS_{Treatments}}{MS_E} > F_{1-\alpha,\ a-1, N-a}$$

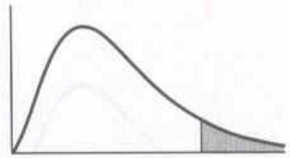
נזכר בקשר בין התפלגות t ו – F.

$$F_{1,m} \overset{d}{=} t^2_m$$

$$P(t_{\alpha/2} \le t_m \le t_{1-\alpha/2}) = 1-\alpha =$$
$$P(|t_m| \le t_{1-\alpha/2}) =$$
$$P(|t_m|^2 \le t_{1-\alpha/2}{}^2) =$$
$$P(t_m{}^2 \le t_{1-\alpha/2}{}^2) =$$
$$P(F_{1,m} \le t_{1-\alpha/2}{}^2)$$

דוגמא מהטבלאות

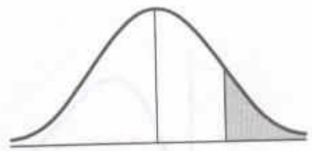
TABLE 4 Cutoff points for the F distribution, right-hand tail probabilities



Denominator df	α	Numerator df 1	2	3	4	5	6	7
	.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91
	.050	161.45	199.50	215.71	224.58	230.10	233.99	236.77
1	.025	647.79	799.50	864.16	899.58	921.85	937.11	948.22
	.010	4052.2	4999.5	5403.4	5624.6	5763.6	5859.0	5928.4
	.001	405284	500000	540379	562500	576405	585937	592873
	.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35
	.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35
2	.025	38.51	39.00	39.17	39.25	39.30	39.33	39.36
	.010	98.50	99.00	99.17	99.25	99.30	99.33	99.36
	.001	998.50	999.00	999.17	999.25	999.30	999.33	999.36
	.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27
	.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89
3	.025	17.44	16.04	15.44	15.10	14.88	14.73	14.62
	.010	34.12	30.82	29.46	28.71	28.24	27.91	27.67
	.001	167.03	148.50	141.11	137.10	134.58	132.85	131.58
	.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98
	.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09
4	.025	12.22	10.65	9.98	9.60	9.36	9.20	9.07
	.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98
	.001	74.14	61.25	56.18	53.44	51.71	50.53	49.66
	.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37
	.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88
5	.025	10.01	8.43	7.76	7.39	7.15	6.98	6.85
	.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46
	.001	47.18	37.12	33.20	31.09	29.75	28.83	28.16
	.100	3.78	3.46	3.29	3.18	3.11	3.05	3.01
	.050	5.99	5.14	4.76	4.53	4.39	4.28	4.21

Casella and Berger

TABLE 2 Cutoff points for Student's t distribution, right-hand tail probabilities



df	.25	.20	.15	.10	.05	.025	.02	.01	.005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169

2.571^2 = 6.61

מבחן t להשוואת תוחלות כמקרה פרטי של ניתוח שונות חד-כווני.

- במבחן t, ‏a=2 ו - N=n1+n2.

$$t=\frac{(\overline{X}-\overline{Y})}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t_{n_1+n_2-2}$$

$$t^2=\frac{(\overline{X}-\overline{Y})^2}{S_p^{\,2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\sim F_{1,n_1+n_2-2}$$

$$t^2=\ldots\quad\ldots=F_0=\frac{SS_{Treatments}/a-1}{SS_E/N-2}=\frac{SS_{Treatments}}{SS_E/n_1+n_2-2}\sim F_{1,n_1+n_2-2}\equiv\left(t_{n_1+n_2-2}\right)^2$$