3Y03 - PROBABILITY AND STATISTICS FOR ENGINEERING

WS19 **Lecture 35**

Last Time ANOVA for Single Factor Experiments

a treatment levels, n per treatments (total $N = na$)

F-test: $H_0 : \mu_1 = \ldots = \mu_a$ (AKA $\tau_1 = \ldots = \tau_a = 0$)

$H_1 : \mu_i \neq \mu_j$ for at least one pair of treatment means; $\tau_i \neq 0$ for at least one (hence at least two) treatments effects

at level α: reject H_0 if $f_0 > f_{\alpha, a-1, N-a}$

$$F_0 = \frac{MS_{Treatments}}{\underbrace{MS_E}_{\text{unbiased est. of } \sigma^2 = V(Y_{ij}) = V(\varepsilon_{ij})}} = \frac{SS_{Treatments}/(a-1)}{SS_E/(N-a)}$$

Variation due to differences in treatments (between treatments)

Random variation (within treatments)

To "reject H_0" does NOT indicate which claim ($\mu_i = \mu_j$, or $\tau_i = 0$) is false!!!

What else do we get out of the ANOVA setup?

MS_E unbiased est. of $\sigma^2 = V(\varepsilon_{ij}) = V(Y_{ij})$

Each $\hat{\mu}_i = \bar{Y}_{i\cdot}$ is an unbiased estimator for μ_i

So $$T = \frac{\bar{Y}_{i\cdot} - \mu_i}{\sqrt{\frac{MS_E}{n}}} \sim \text{t-distr. with } N-a \text{ d.o.f.}$$
(= # d.o.f. of SS_E)

Thus a $100(1-\alpha)\%$ C.I. for each individual μ_i is given by

$$\bar{y}_{i\cdot} \pm \underbrace{t_{\frac{\alpha}{2}, N-a}\sqrt{\frac{MS_E}{n}}}$$

ME does NOT depend on i if all treatments same size $= n$

& we could also (in principle) do t-tests on individual treatment means.

Example – see sheet a 95% C.I. for μ_2 (mean of treatment 2)

For differences between means:

$\bar{Y}_{i\cdot} - \bar{Y}_{k\cdot}$ is an unbiased est. of $\mu_i - \mu_k$ &

$$V(\bar{Y}_{i\cdot} - \bar{Y}_{k\cdot}) = V(\bar{Y}_{i\cdot}) + V(\bar{Y}_{k\cdot}) = \frac{\sigma^2}{n} + \frac{\sigma^2}{n} = \frac{2\sigma^2}{n}.$$

Similar to above we get a $100(1-\alpha)\%$ C.I. for difference $\mu_i - \mu_k$:

$$\bar{y}_{i\cdot} - \bar{y}_{k\cdot} \pm \underbrace{t_{\frac{\alpha}{2}, N-a}\sqrt{\frac{2MS_E}{n}}}.$$

Again ME is same for all pairs if all treatments same size $= n$.

Example – see sheet (95% C.I. for difference between μ_1 & μ_2.)

Multiple Comparisons after ANOVA (F-test) says

reject $H_0 : \mu_1 = \mu_2 = \ldots = \mu_a$

for H_1 : at least one pair $\mu_i \neq \mu_k$ $(i \neq k)$

→ We do all paired tests $H_0 : \mu_i = \mu_k$ only after

$H_1 : \mu_i \neq \mu_k$

H_0 from F-test rejected → this drops that EER

P(Type I Error somewhere)

Recall For any test $H_0: \vartheta = \vartheta_0$
$H_1: \vartheta \neq \vartheta_0$ at level α,

we reject H_0 if $\hat{\theta}$ does NOT lie in an interval

observed value $\vartheta_0 \pm ME$ where

ME is Margin of Error for $100(1-\alpha)\%$ CI for ϑ

So use the ANOVA $100(1-\alpha)\%$ C.I. setup to do all our paired tests (different from 10-2). Why?

- Use what you've already got!
- If # observations per treatment is consistent ($=n$) the ME is same for all tests (as opposed to 10-2 where each test uses 2 sample variances)

(& even if they're not, the only different input is n_i, n_k, not S_i^2, S_k^2 — use MS_E for all).

i.e. $H_0 : \mu_i = \mu_k$ at level α
$H_1 : \mu_i \neq \mu_k$

Reject H_0 if $|\bar{Y}_{i\cdot} - \bar{Y}_{k\cdot}| > \underbrace{t_{\frac{\alpha}{2}, N-a}\sqrt{\frac{2MS_E}{n}}}$.

Here this critical value (ME) is called Fisher's Least Significant Difference (LSD)

Example (sheet)

$MS_E = 71.5385$, $n = 9$, say $\alpha = 0.05$

From earlier (see sheet) $LSD = t_{0.025,24}\sqrt{\frac{2(71.53)}{9}}$

$= 8.23.$

Now test

① $H_0 : \mu_1 = \mu_2$: $|32.4 - 61.3| = 28.9 > 8.23$

② $H_0 : \mu_1 = \mu_3$: $|32.4 - 39.7| = 7.3 < 8.23$

③ $H_0 : \mu_2 = \mu_3$: $|61.3 - 39.7| = 21.6 > 8.23$

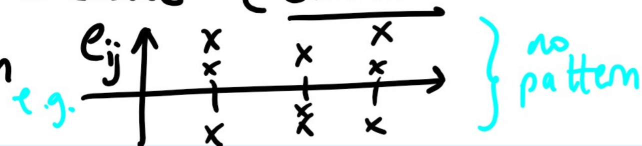
So reject H_0 in ① & ③ — there is a significant difference between μ_1 & μ_2 and μ_2 & μ_3 but not between μ_1 & μ_3.

Assumptions with ANOVA

Within treatment errors: $\varepsilon_{ij} \sim N(0, \sigma^2)$

To check Normality, do Normal prob. plot of residuals $e_{ij} = y_{ij} - \bar{y}_{i\cdot}$ & look for (approx.) straight line.

To check common variance, plot e_{ij} against e.g. treatment levels or treatment means (<u>should</u> be irrelevant) — Should see no pattern



We also assumed <u>for convenience</u> that all treatments have same size i.e. $n_i = n$ for all i.

WE DID NOT NEED THIS!

For "unbalanced experiments" everything works, but some formulas more complicated:

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n_i}\left(Y_{ij}^2\right) - N\bar{Y}_{\cdot\cdot}^2 \qquad SS_{Treatments} = \sum_{i=1}^{a}\left(\frac{Y_{i\cdot}^2}{n_i}\right) - N\bar{Y}_{\cdot\cdot}^2$$

In $100(1-\alpha)\%$ C.I.s:

For single means μ_i: $\quad \bar{y}_{i\cdot} \pm t_{\frac{\alpha}{2}, N-a}\sqrt{\frac{MS_E}{n_i}}$

For differences between pairs of means: $\quad \bar{y}_{i\cdot} - \bar{y}_{k\cdot} \pm \underbrace{t_{\frac{\alpha}{2}, N-a}\sqrt{MS_E\left(\frac{1}{n_i}+\frac{1}{n_k}\right)}}$

For tests on pairs of means: $\quad$ critical value = LSD

(Much simpler than in 10-2 for doing tests on all the different pairs, <u>once</u> we've done the ANOVA!)

Source: A. Parenti, L. Guerrini, P. Masella, S. Spinelli, L. Calamai,
P. Spugnoli (2014). "Comparison of Espresso Coffee Brewing Techniques,"
Journal of Food Engineering, Vol. 121, pp. 112-117.

Description: Comparison of foam index (Y, in %) for 3 methods of brewing espresso
Method 1 = Bar Machine (BM), **a=3**
Method 2 = Hyper-Espresso Method (HIP),
Method 3 = I-Espresso System (IT).

9 replicates/treatment. **n=9**

Variables: foamIndx (response)
method (factor)

method	1		2		3	
	y_{1j}		y_{2j}		y_{3j}	
foamIndx	36.64	1342.4896	70.84	5018.31	56.19	3157.32
	39.65	1572.1225	46.68	2179.02	36.67	1344.69
	37.74	1424.3076	73.19	5356.78	35.35	1249.62
	35.96	1293.1216	57.78	3338.53	40.11	1608.81
	38.52	1483.7904	48.61	2362.93	33.52	1123.59
	21.02	441.8404	72.77	5295.47	37.12	1377.89
	24.81	615.5361	65.04	4230.2	37.33	1393.53
	34.18	1168.2724	62.53	3910	32.68	1067.98
	23.08	532.6864	54.26	2944.15	48.33	2335.79

y_1.	**291.6**	**y_2.**	**551.7**	**y_3.**	**357.3**	**y..**	**1200.6**
ȳ_1.	**=291.6/9**	**ȳ_2.**	**=551.7/9**	**ȳ_3.**	**=357.3/9**	**ȳ..**	**=1200.6/27**
	32.4		**61.3**		**39.7**		**44.4666667**
$(y\_1.)^2$	**85030.56**	**$(y\_2.)^2$**	**304373**	**$(y\_3.)^2$**	**127663**	**$(\bar{y}..)^2$**	**1977.28444**

$SS_{Treatments}$ **4065.18** $= \sum\sum \frac{y_{i\cdot}^2}{9} - 27\bar{y}_{\cdot\cdot}^2$
SS_E **1716.9192**
SS_T **5782.0992** $= \sum\sum y_{ij}^2 - 27\bar{y}_{\cdot\cdot}^2$

$\Sigma(y_{ij})^2$ = 59168.7792

ANOVA Table

Source of Variation	SS	d.o.f.	MS
Treatments	4065. ...	2	2032.59
Error	1716. ...	24	71.5383
Total	5782...	26	

$= F_0 = 28.41$ Compare to $f_{0.01,2,24} = 5.61$
$28.41 > 5.61$
So reject H_0
Test at $\alpha = 0.01$ level

95% C.I. for mean of tr. 2
$61.3 \pm t_{0.025,\underbrace{24}_{N-a}} \sqrt{\frac{71.5383}{9}}$
$= 61.3 \pm 5.82$

95% C.I. for difference $\mu_1 - \mu_2$:
$32.4 - 61.3 \pm t_{0.025,24}\sqrt{\frac{2(71.5)}{9}}$
$= -28.9 \pm 8.23$