Lecture 28 Wednesday April 5

Exam 2: Total number of possible points was 100 (by page: 24, 10, 24, 12, 30).

Distribution of Exam 2 Scores:

Mean	SD	Q1	Median	Q3
81.0	11.8	72.0	81	91.0

Midterm average is entered on your test booklet. This was calculated as follows: $.2(\underline{\text{hw.per}}) + .2(\text{Exam1}) + .2(\text{Exam2}) + \underline{.4(\text{Project1})}$

out of total of 110
whereas 130 pts possible so far

double-weight on Project 1

If you have a question about a score, or a concern about your grade, first **study the solution sheet** for Exam 2, and **check the calculation of your midterm average**. See me in office hour or by appointment starting Thursday April 6 with your grade questions/concerns.

Let's continue (finish) the analysis of the corn yield data.

Tests for Factor Main Effects

The hypotheses for Factor A main effects:

$H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_a = 0$

$H_a :$ Not all $\alpha_i = 0$

The test statistic is:

$$F_{\text{obs}} = \frac{\text{MSA}}{\text{MSE}}$$

Under H_0, the distribution of F_{obs} is $F_{a-1,(n-1)ab}$.

Decision Rule: Reject H_0 at level α if $F_{\text{obs}} > F_{1-\alpha,a-1,(n-1)ab}$

(*)

or if $P = \boxed{Pr(F_{a-1,(n-1)ab} > F_{obs})} < \alpha$

If you reject H_0, proceed to form CIs for contrasts of Factor A level means.

Factor B main effects

Hypotheses, test statistic, null distribution, and decision rule:

$H_0 : \beta_1 = \beta_2 = \ldots = \beta_b = 0$ vs. $H_a :$ Not all $\beta_j = 0$

(*) $Pr(X > F_{obs})$

where X is a r.v. with distribution $F_{a-1,(n-1)ab}$

Factor B main effects

Hypotheses, test statistic, null distribution, and decision rule:

The test statistic is:

$$F_{\text{obs}} = \frac{\text{MSB}}{\text{MSE}}$$

$H_0 : \beta_1 = \beta_2 = \ldots = \beta_b = 0$ vs. $H_a :$ Not all $\beta_j = 0$

Null distr.: $F_{b-1,(n-1)ab}$

Decision Rule: Reject H_0 at level α if

$F_{\text{obs}} > F_{1-\alpha,b-1,(n-1)ab}$, or if $p = Pr(F_{b-1,(n-1)ab} > F_{obs}) < \alpha$

If you reject H_0, proceed to form CIs for contrasts of Factor B level means.

Example. Corn Yield (p. 137) and Lecture 25 for table of means

Test of manure main effects:

$F_{\text{obs}} = 19.208$, $P = 0.01816$. Reject H_0 at level $\alpha = .05$, and proceed to estimate the pairwise comparison of factor level means.

Test of fertilizer main effects:

$F_{\text{obs}} = 17.672$, $P = 0.02258$. Reject H_0 at level $\alpha = .05$, and proceed to estimate the pairwise comparison of factor level means.

Cell Means

Manure (A)	Fertilizer (B) L	H	$\bar{Y}_{i..}$
L	11.3	13.9	12.6
H	14.0	15.1	14.55
$\bar{Y}_{.j.}$	12.65	14.50	$\bar{Y}_{...} = 13.575$

ANOVA table p. 135

$MSE = \frac{44.4}{16} = 2.775$

FACTOR A = Manure. Estimate $L_A = \mu_{2.} - \mu_{1.}$ (Hi v. Low) $= \sum_{i=1}^{a} c_i \mu_{i.}$ $c_1 = -1$, $c_2 = 1$

by $\hat{L}_A = \bar{Y}_{2..} - \bar{Y}_{1..} = 14.55 - 12.60 = 1.95$

and get a formula for $Var(\hat{L}_A)$: $Var(\hat{L}_A) = \frac{\sigma^2}{bn} \sum_{i=1}^{a} c_i^2 = \frac{\sigma^2}{2(5)}((-1)^2 + 1^2) = \frac{\sigma^2}{5}$

So, $\hat{Var}(\hat{L}_A) = \frac{MSE}{5} = \frac{2.775}{5} = .555$

and $s(\hat{L}) = \sqrt{\hat{Var}(\hat{L})} = .7450$

Then an individual confidence interval for L is

$\hat{L} \pm t_{.975,16}\, s(\hat{L})$ which is $1.95 \pm 2.1199(.7450)$, or $(.37, 3.53)$.

B (Fertilizer)

$L_B = \mu_{.2} - \mu_{.1}$ (Hi vs Lo) $= \sum_{j=1}^{b} c_j \mu_{.j}$ $\quad c_1 = -1, c_2 = 1$

$\hat{L}_B = \bar{Y}_{.2.} - \bar{Y}_{.1.} = 14.50 - 12.65 = 1.85$

or $\hat{\beta}_2 - \hat{\beta}_1$

$\hat{Var}(\hat{L}_B) = \dfrac{MSE}{an} \sum_{j=1}^{b} c_j^2 = \dfrac{2.775}{5} = .555$

(work is same as for $\hat{L}_A$ since $a = b$)

95% individual CI

$\hat{L}_B \pm t_{.975, 16}\, s(\hat{L}_B)$ which is 1.85 ± 1.5793

or $(.27, 3.43)$

Same margin of error as for A comparison

138-2 [redacted] because $a = b = 2$

(General procedure)

Estimation of Contrast of Factor Level Means

This is the followup analysis that is appropriate when:

1. Interaction effect is not significant, with P-value $> .2$, and
2. The factor for which we are analyzing contrasts of level means had a significant F test.

Example. Corn Yield 2×2 factorial. Only one pairwise comparison of interest for Manure, and ditto for Fertilizer.

For $a > 2$ or $b > 2$, we are usually interested in multiple pairwise comparisons, or multiple factor level contrasts. Then the multiple comparison procedures for one-way ANOVA can be used with only minor modifications for two-factor studies.

you might want to

Ex. $a = 3$, estimate all three pairwise comparisons plus $L_1 = \frac{\mu_{1.} + \mu_{2.}}{2} - \mu_3$ using Scheffé method.

Two-way ANOVA: Follow-up analysis

Consider the general case of a levels of Factor A, b levels of Factor B.

What to do if factors do not interact (Text: Section 19.8)

If the first step of analysis determines that interactions are insignificant and/or unimportant, then follow-up analysis is based on contrasts of factor level means. If only one comparison is being made, use individual t or F methods. If several comparisons are being made, each of the three multiple comparison methods (Bonferroni, Scheffe, Tukey) can be adapted to the two-way layout; the choice depends on the application.

Ingredients needed for forming CIs for comparisons of factor level means:

1. df(Error) = df(Residuals), and $\hat{\sigma}^2$. From the Residuals row of the two-way ANOVA table.
2. Estimates of the desired contrasts and their SE's.

Factor A contrasts and standard errors:

Let $L = \sum_{i=1}^{a} c_i \mu_{i.}$, with $\sum c_i = 0$; estimate L by

$$\hat{L} = \sum_{i=1}^{a} c_i \overline{Y}_{i..}$$

In a balanced two-way design, each row mean $\overline{Y}_{i..}$ is an average of bn observations, and so has variance $\frac{\sigma^2}{bn}$. So, the variance and estimated variance of $\hat{L}$ are:

$$\mathrm{Var}(\hat{L}) = \frac{\sigma^2}{bn} \sum_{i=1}^{a} c_i^2$$

$$\widehat{\mathrm{Var}(\hat{L})} = \frac{\text{MSE}}{bn} \sum_{i=1}^{a} c_i^2$$

Factor B contrasts and standard errors:

Let $L = \sum_{j=1}^{b} c_j \mu_{.j}$, with $\sum c_j = 0$; estimate L by

$$\hat{L} = \sum_{j=1}^{b} c_j \overline{Y}_{.j.}$$

In a balanced two-way design, each column mean $\overline{Y}_{.j.}$ is an average of an observations, and so has variance $\frac{\sigma^2}{an}$. So, the variance and estimated variance of $\hat{L}$ are:

$$\mathrm{Var}(\hat{L}) = \frac{\sigma^2}{an} \sum_{j=1}^{b} c_j^2$$

$$\widehat{\mathrm{Var}(\hat{L})} = \frac{\text{MSE}}{an} \sum_{j=1}^{b} c_j^2$$

What to do if factors do interact (Text: Section 19.9)

If the first step of analysis determines that interaction is significant and/or practically important, then the follow-up analysis is based on contrasts of cell means.

If only one comparison is being made, use individual t or F methods. If several comparisons are being made, each of the three multiple comparison methods (Bonferroni, Scheffe, Tukey) can be adapted to the two-way layout, whichever is deemed appropriate.

Important for Project ←, last question

Cell mean constrasts, estimates and standard errors:

Let $L = \sum_{i=1}^{a} \sum_{j=1}^{b} c_{ij}\mu_{ij}$, with $\sum \sum c_{ij} = 0$; estimate L by

$$\hat{L} = \sum_{i=1}^{a} \sum_{j=1}^{b} c_{ij}\overline{Y}_{ij.}$$

So, the variance and estimated variance of $\hat{L}$ are:

$$\mathrm{Var}(\hat{L}) = \frac{\sigma^2}{n} \sum_{i=1}^{a} \sum_{j=1}^{b} c_{ij}^2$$

$$\widehat{\mathrm{Var}(\hat{L})} = \frac{\text{MSE}}{n} \sum_{i=1}^{a} \sum_{j=1}^{b} c_{ij}^2$$

Important for Project 2 last question

Interaction Model: Scheffé method for estimating contrasts of cell means

First recall the one-way ANOVA application of Scheffé method (p. 82 of these notes): The critical constant (in one-way ANOVA notation) was $\sqrt{S^2}$, where S^2 is given by:

Let $S^2 = (r-1)F(1-\alpha; r-1, n_T - r)$ (Here the capital "S" refers to critical constant for Scheffé method)

Since the $a \times b$ two-way ANOVA model with interaction effects is equivalent to the one-way ANOVA model with $r = ab$, the Scheffé method is applied exactly as before, and the critical constant is:

$$S = \sqrt{(ab-1)F(1-\alpha; ab-1, abn-ab)}$$