Consultation prior to the exam

Mon 11th Nov. Rob Haillandet C448
10.30 - 1, 2 - 4

Wed 13th Nov. Sharon Cown C444
9.30 - 11, 11.30 - 1, 2.15 - 3.45

KAREN ?? no set days but
not any Thursday.

Lecture notes p. 148 MINITAB output correct.
Will correct my overhead for web.

Not doing Ch. 12—reading exercise.
Not examinable.

Everything in Ch's 1 - 11 is EXAM POSSIBLE.
(a) \[ \text{Source} \quad \text{DF} \quad \text{SS} \quad \text{MS} \quad \text{F} \]

\[ \begin{align*}
\rightarrow & \quad \text{Length} \quad 4 \quad 44024 \quad 11006 \quad 10.48 \\
\text{Error} & \quad 30 \quad 31497 \quad 1049.9 \\
\text{Total} & \quad 34 \quad 75521
\end{align*} \]

\( H_0: \mu_1 = \mu_2 = \ldots = \mu_5 \)

\( H_1: \text{not } H_0 \)

\( F = 10.48 \) and under \( H_0, F \approx F_{(4,30)} \)

\( F_{(4,30)}(0.99) = 6.125 < 10.48 \)

- Reject \( H_0 \) and conclude that length affects stiffness.

(b) \( \psi = \frac{\mu_4 + \mu_6 + \mu_8}{3} - \frac{\mu_{10} + \mu_{12}}{2} \)

\[ \hat{\psi} = \frac{333.2 + 368.1 + 375.1 - 407.4 + 437.2}{3} = -63.5 \]

\[ \text{Var}(\hat{\psi}) = \frac{\left(\frac{1}{3}\right)^2}{3} \times 38^2 + \frac{\left(\frac{1}{3}\right)^2}{2} \times 25^2 \]

\[ s^2 = 1049.9 \]

\[ \text{SE}(\hat{\psi}) = 11.1798 = \sqrt{\frac{38^2}{3} + \frac{25^2}{2}} \]

\[ 95\% \text{ CI for } \psi = -63.5 \pm t_{0.025, 30} \times 11.1798 \]

\[ = (-86.33, -40.67) \]

(c) \( \text{RSS} = \sum_{i=1}^{6} \hat{e}_i^2 = 6 \left[ (29.5)^2 + (20.83)^2 + (44.9)^2 + (26.07)^2 \right] \)

\[ = 23465.5 \]
Question 2:
(a) An increase of 1 unit in length gives an increase of 12.4 units in stiffness index.

(b) \[ r = \sqrt{\frac{R_{yy}}{100}} = 0.753 \] (note: \( r > 0 \)) as slope > 0

(c) 95% PI for St. index of plate 12 in. long:

\[ \text{length} = 12; \quad \text{index} = 285.27 + 12.365 \times 12 \]

\[ \mu(12) = 433.65 \]

\[ s_e(\mu(12)) = \sqrt{\frac{s^2}{n} + (x-\bar{x})^2 \cdot \text{Var}(\beta)} \]

\[ = \sqrt{\frac{31.48^2}{35} + (12-8)^2 (1.882)^2} \]

\[ = 9.2187 \]

So:

\[ 433.65 \pm 0.975 \sqrt{32 + 9.2187^2} \]

\[ = 433.65 \pm 2.0345 \times 9.2187 \]

\[ = (366.9, 500.4) \]

(d) \( H_0: \mu(x) = \alpha + \beta x \)

\( H_1: \mu(x) \text{ arbitrary} \)

\[ E(Y|x) = \mu(x) \quad \text{full} \]

\[ E(Y|x) = \mu(x) \quad \text{(testing adequacy of a regression model)} \]

\[ F = \frac{(RSS_0 - RSS_1)/\text{df}_0 - \text{df}_1}{RSS_1/\text{df}_1} \]

\[ = \frac{(32711 - 31497)/(33-30)}{31497/30} \]

\[ = 0.3854 \]
Under null distribution of $F$ is $F_{(3,30)}$
and $F_{(3,30)}(0.95) = 2.922 > 0.3851$

\[ \therefore \text{Retain } H_0. \]

The simple linear regression model is better. It is

- simpler (fewer parameters estimated)
- explains almost as much variation as the one-way model,

\[ 56.7\% \text{ vs } \frac{44024}{75521} \times 100 \approx 58.3\%. \]
Question 3:

\[ (4 + 3 + 5 + 3 = 15) \]

(a) Score

![Graph with verbal and none categories]

(b) \( H_0: \) no interaction

\( H_1: \) there is interaction

\[ \text{F-statistic} = \frac{52.9.3}{15.8} = 33.5 \]

Under \( H_0 \), \( F \) \( \approx \) \( F_{(2,10)} \) and \( F_{(2,10)} (0.999) = 13.29 \)

\[ < 33.5 \]

\( \therefore \) Reject \( H_0 \). The effect of interaction is significant.

(c) \( \text{LSD} = \frac{3.389}{\sqrt{2}} \sqrt{\frac{s^2}{n_1 + n_2}} \)

\[ = \frac{3.49}{\sqrt{2}} \sqrt{15.8 \left( \frac{1}{10} + \frac{1}{10} \right)} \]

\[ = 3.2698 \]

\[ \mu_{20} - \mu_{40}: |23.167 - 28.333| > \text{LSD} \quad \therefore \text{significant} \]

\[ \mu_{20} - \mu_{60}: |23.167 - 12.833| > \text{LSD} \quad \therefore \]

\[ \mu_{40} - \mu_{60}: |28.333 - 12.833| > \text{LSD} \quad \therefore \]

When no reinforcement is used, mean scores for all isolation times are significantly different.
(d) $Y_{ij}$ = test score of a student given reinforcement level $i$ and isolation team level $j$.

Assume:

- $Y_{ij} = \mu + \alpha_i + \beta_j + (\alpha \beta)_{ij} + \text{error}$

  $\sum \alpha_i = 0$,
  $\sum \beta_j = 0$,
  $\text{Var}(Y_{ij}) = \text{Var(\text{error})} = \sigma^2$,

- $(\alpha \beta)_{ij} = 0$ \text{, } $Y_{ij}$ is normally distributed,

- $(\alpha \beta)_{ij} = 0$\text{, observations are independent.}
Question 4:  
(a) response is test score

\[(3+1+2+1+5 = 12)\]  
expl: reinforcement 2
- isolation time 3

(b) experimental units: students

(c) 2 factor, 2 x 3 factorial experiment, completely randomised design.

(d) 6 treatments \(\text{(6 replications)}\)

(e) randomised block design  
 gender
- randomly select 18 boys, 18 girls
- randomly allocate the 18 boys to 6 groups \(\text{same for girls}\)

- within each group, randomly assign students to each treatment so same no. of students in each treatment.
Question 5:

(a) \[ \hat{g}(x_1, x_2) = 86.8 - 0.123x_1 \\
\quad + 5.09x_2 + 0.0709x_2^2 - 0.001x_1x_2 \]
\[ \Rightarrow \hat{g}(3200, 57) = 86.8 - 0.123(3200) \\
\quad + 5.09(57) + 0.0709(57^2) - 0.001(3200)(57) \]
\[ = 31.28 \]

(b) \[ H_0: \mu(x_1, x_2) = \mu \]
\[ H_1: \mu(x_1, x_2) = \alpha + \beta_1x_1 + \beta_2x_2 + \beta_3x_2^2 + \beta_4x_1x_2 \]

ANOVA table:

\[
\begin{array}{c|c|c|c|c|c|}
\hline
\text{Source} & \text{df} & \text{SS} & \text{MS} & \text{F} \\
\hline
\text{Regression} & 4 & 5896.6 & 1474.15 & 64.41 \\
\text{Residual} & 26 & 595.1 & 22.889 & \\
\text{Total} & 30 & 6491.7 & \\
\hline
\end{array}
\]

\[ 6.125 < F_{0.05, 26} (0.999) < 6.589 \Rightarrow \text{64.41 in } RR. \]

\[ \Rightarrow \text{reject } H_0 \text{ — the model is useful.} \]
(c) \[ R^2 = \left( 1 - \frac{593.1}{6491.7} \right) \times 100\% = 90.83\% \]

\[ s = \sqrt{22.889} = 4.7842 \]

The model explains 90.8\% of the variation of the observed values of the dependent variable.

The standard error of the dependent variable about the expected value for all \((x_1, x_2)\) pairs is 4.78.

(d) larger residual SS.

The least squares regression is the fitted model with the smallest residual SS.
Question 6: (a) $y_{ijk} = \mu + \alpha_i + B_j + (\alpha B)_{ij} + e_{ijk}$

- $\mu$ = overall mean
- $\alpha_i$ = fixed effect of casing material $i$
- $B_j$ = random effect of bearing source $j$
- $(\alpha B)_{ij}$ = random interaction effect
- $E_{ijk}$ = random error $\sim N(0, \sigma^2_e)$

(b) $\hat{\sigma}^2 = \frac{MS_E}{b} = 0.1113$

$\hat{\sigma}_{\alpha B}^2 = \frac{MS_{\alpha B} - MS_E}{b} = \frac{1.4507 - 0.1113}{6} = 0.2360$

$\hat{\sigma}_B^2 = \frac{MS_B - (\hat{\sigma}^2 + 2\hat{\sigma}_{\alpha B}^2)}{6} = \frac{MS_B - MS_{\alpha B}}{6}$

$= \frac{9.1687 - 1.4507}{6} = \frac{1.2863}{6}$
(c) $H_0$: casing material has no effect ($\alpha = 0$)
$H_1$: not $H_0$

$$F = \frac{0.3523}{1.4507} = 0.2428 < F_{2,8}(0.95) = 4.459$$

:. Retain $H_0$ and conclude that casing material is not significant—has no effect.
Question 7:  
(a) \[ H_0: E(Y|x) = \alpha + \beta_1 x_1 + \beta_3 x_1^2 + \beta_4 x_2 \]
(b) \[ H_1: E(Y|x) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_2 \]

\[ F_{out} = 4 \]

\[ n=31; \quad F = \frac{(185.9-185.9)/1}{185.9/26} = 0 < 4 \]

\[ \therefore \text{accept } H_0 : x_2 \text{ out} \]

\[ H_0: E(Y|x) = \alpha + \beta_1 x_1 + \beta_3 x_1^2 \]
\[ H_1: E(Y|x) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_2 \]

\[ F = \frac{(311.4-185.9)/1}{185.9/27} = 19.23 > 4 \]

\[ \therefore \text{reject } H_0 : \text{ retain } x_2^2 \]

Selected model: \[ \frac{E(Y|x)}{(A)} = \alpha + \beta_1 x_1 + \beta_3 x_1^2 + \beta_4 x_2^2 \]

(b) (i) \[ \text{adj } R^2 = 1 - \frac{\text{RSS}/(n-p)}{\text{TotalSS}/n-1} \]

\[ = 1 - \frac{389.6/29}{810.1/30} \]

\[ = 0.9503 \quad \text{or} \quad 95.03\% \]

\[ C_p = \frac{389.6}{185.9/26} + 2(2-31) = 27.49 \]
Model B: $R^2 \quad \text{adj } R^2 \quad C_p \quad 5$

$x_i^2: 95.9 \quad 95.8 \quad 19.7 \quad 3.37$

- Model B not as good as that using $x_i^2$ only (by all diagnostics)
- $x_i^2$ only not as good as Model A.
- Model A better than Model B.

Question 8: (a) Assumes pre-test no effect

- One-way ANOVA

Therapy influences post-test score.

Model: $\mu(i,x) = \mu_i$

$\mu(i,x)$ = mean post-test score for a person on treatment $i$ with pre-test score $x$.

From the analysis:

$F = 6.71 \rightarrow P-value = 0.004 < 0.05$

$\rightarrow$ The treatment effect is highly significant.
(1) \( \mu(i, x) = \mu + \alpha_i + \beta x \)  
   - Additive ANCOVA

The pre-test and therapy may have an effect on post-test score, but there is no interaction between pre-test score and therapy.

(iii) Treatment: \( F = 70.10 \) \( \Rightarrow P\text{-value} = 0.000 \)  
      - Effect is very highly significant.

(iii) Fit a model with interaction to see if it is significant. If not, then this model is the most appropriate.

(iv) No. In this model pre-test score is significant (P-value = 0.000).

: Student A's model is not appropriate as it ignores pre-test score.

\( \mu(i, x) \)

(c) \( x = 24: \)

A: \( \hat{\mu}(1, 24) = \bar{y}_i = 52.80 \)

B: \( \hat{\mu}(1, 24) = \hat{\mu} + \hat{\alpha}_1 + \hat{\beta} \times 24 \)

\[ 7.233 + 11.7114 + 1.12563 \times 24 = 46.03 \]