Stat 350B mt2 Solutions.

1. (a) RCB. Treatment: diet; blocking variable: age

(b) $e_{ij} = y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..} = 1.62 - 1.21 - 1.27 + 0.93 = 0.07$

(c) $y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$, $i = 1, \cdots, 3$, $j = 1, \cdots, 4$.

$\hat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..}$, $\hat{\alpha}_1 = 1.21 - 0.93 = 0.28$

$\hat{\alpha}_2 = 1.07 - 0.93 = 0.14$, $\hat{\alpha}_3 = 0.50 - 0.93 = -0.43$

Lower fat content in diet, more reduction in plasma lipid level.

(d) $SSTr = 4 \sum (\hat{\alpha}_i)^2 = 4\{0.28^2 + 0.14^2 + (-0.43)^2\} = 1.13$

(e) $H_0: \alpha_1 = \alpha_2 = \alpha_3 = 0$ H_a: At least one $\alpha_i \neq 0$. $i = 1, 2, 3$.

$$F = \frac{SSTr/(t-1)}{SSE/(t-1)(b-1)} = \frac{1.13/2}{0.013/6} \approx 260$$

$df_1 = 2$, $df_2 = 6$. p-value $< .001$ $\Rightarrow$ reject H_0

OR: $260 > F_{0.05}(2, 6) = 5.14$ $\Rightarrow$ reject H_0

There are significant differences in the mean reduction of plasma lipid for the three diets.

Solutions:

2 (a) $RoomA = \begin{cases} 1 & \text{if Rooms} = 1 \\ 0 & \text{else} \end{cases}$ (5 or 6 rooms)

$RoomB = \begin{cases} 1 & \text{if Rooms} = 2 \\ 0 & \text{else} \end{cases}$ (7 or 8 rooms)

(b) $Price = \beta_0 + \beta_1 RA + \beta_2 RB + \beta_3 SF + \beta_4 RA*SF + \beta_5 RB*SF + \varepsilon$

For houses w/ 5-6 rooms, $RA = 1$, $RB = 0$.

$price = (\beta_0 + \beta_1) + (\beta_3 + \beta_4) SF + \varepsilon$

slope estimate: $\hat{\beta}_3 + \hat{\beta}_4 = -0.245 + .27 = .025$

For houses w/ 7-8 rooms, $RA = 0$, $RB = 1$

$price = (\beta_0 + \beta_2) + (\beta_3 + \beta_5) SF + \varepsilon$

$\hat{\beta}_3 + \hat{\beta}_5 = -.245 + .275 = .03$

For houses w/ 9-10 rooms, $RA = RB = 0$.

$Price = \beta_0 + \beta_3 SF + \varepsilon$.

$\hat{\beta}_3 = -.245$.

(c) H_0: $\beta_4 = \beta_5 = 0$.

H_a: Not both β_4 & β_5 equal 0.

extra S.S. = 3842.5 − 3143.8

or 7723.7 − 7025.0

or 0.3 + 698.3 = 698.7 or 698.6

$$F = \frac{698.7/2}{3143.8/16} = 1.778 < F_{0.05}(2,16) = 3.63$$

⇒ fail to reject H_0

There is not enough evidence that the linear relationships between price and square feet are different for the 3 categories of houses, that is, the linear relationships between price and square feet does not depend on the # of rooms.

3. (a) The average number of man-hours needed for steam drum type is 2093.35 more than that for mud drum type, after controlling for boiler capacity, boiler design pressure, and boiler type.

(Or: when holding the other variables (X_1, X_2, X_3) constant.)

(b) $R^2 = \frac{SSR}{SSTO} = \frac{SSTO - SSE}{SSTO}$

For $y' = y$, $R^2 = 0.903$

For $y' = \sqrt{y}$. $R^2 = 0.924$

For $y' = \ln(y)$, $R^2 = 0.880$

Choose the model w/ largest R^2 $\Rightarrow$ $\sqrt{y}$ transformation is the best choice.

Or. you may also use $R_a^2 = 1 - \frac{SSE/(n-k-1)}{SSTO/(n-1)}$

here $n = 35$

$k = 4$.

Since all 4 choices have the same df's for SSE and SSTO, we would end up with the same choice: $\sqrt{y}$ transformation.