STA207 Homework 2

Juanjuan Hu, Zhen Zhang

January 18, 2016

21.7

(a)

The age of subjects is suspected to be highly correlated to the reduction of lipid level (response variable). Therefore, blocking the subjects into groups according to age can provide more precise results.

(b)

$\bar{Y}_{1.} = 0.5167$, $\bar{Y}_{2.} = 0.6067$, $\bar{Y}_{3.} = 0.67$, $\bar{Y}_{4.} = 1.1567$, $\bar{Y}_{5.} = 1.27$

$\bar{Y}_{.1} = 1.11$, $\bar{Y}_{.2} = 0.992$, $\bar{Y}_{.3} = 0.43$

$\bar{Y}_{..} = 0.844$

According to the formula $e_{ij} = Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}$, we can calculate the residuals as:

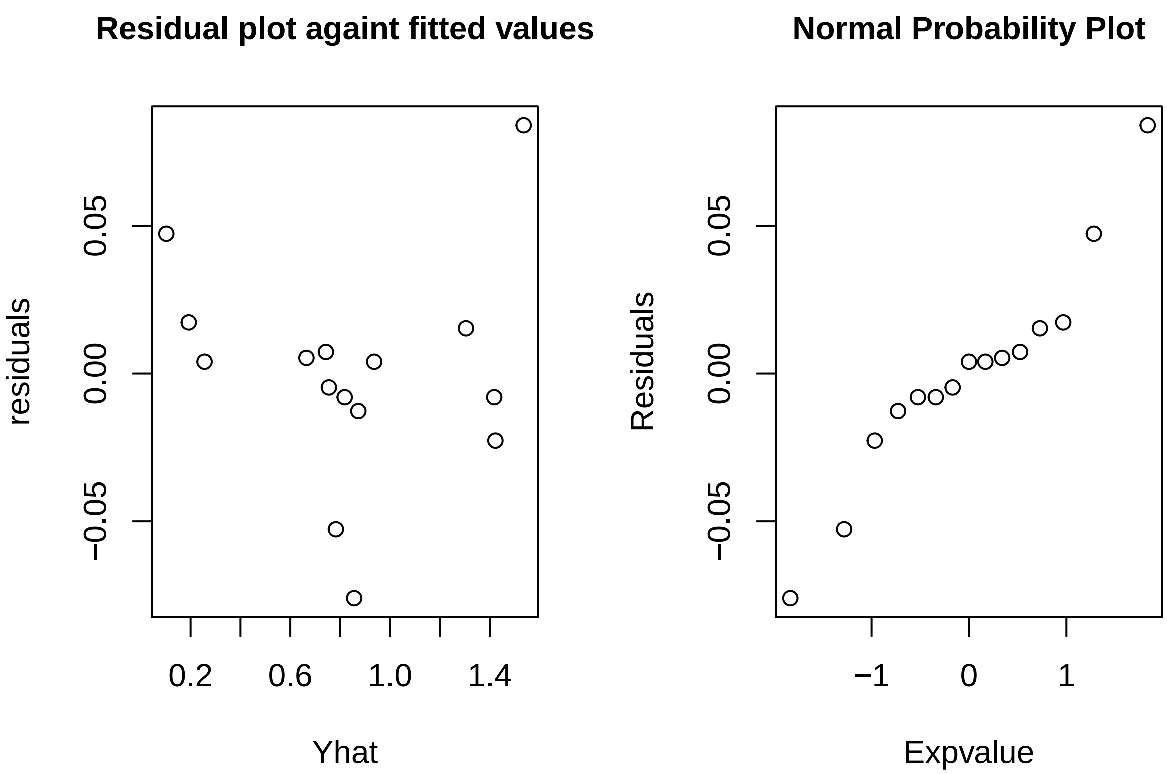
$e_{11} = -0.0527$, $e_{12} = 0.0053$, $e_{13} = 0.0473$

$e_{21} = -0.0127$, $e_{22} = -0.0047$, $e_{23} = 0.0173$

$e_{31} = 0.004$, $e_{32} = -0.008$, $e_{33} = 0.004$

$e_{41} = -0.0227$, $e_{42} = 0.0153$, $e_{43} = 0.0073$

$e_{51} = 0.084$, $e_{52} = -0.008$, $e_{53} = -0.076$



From the above residual plots, we find that there is no evidence of a curvilinear pattern there and no indication of the existence of unequal error variances. In addition, the right plot does not suggest any strong departures from a normal error distribution.

(c)

The plot suggests that there is no important interaction effects between blocks and treatments on the resonses since the reponse curves are almost parallel. In other words, it supports the appropriateness of the no-interaction assumption.

(d)

The Turkey's interaction model is: $Y_{ij} = \mu_{..} + \rho_i + \tau_j + D\rho_i\tau_j + \epsilon_{ij}$

$H_0 : D = 0$

$H_1 : D \neq 0$

$\hat{\rho}_1 = -0.3273$, $\hat{\rho}_2 = -0.2373$, $\hat{\rho}_3 = -0.174$, $\hat{\rho}_4 = 0.3127$, $\hat{\rho}_5 = 0.426$

$\hat{\tau}_1 = 0.266$, $\hat{\tau}_2 = 0.148$, $\hat{\tau}_3 = -0.414$

$SSTO = 2.7586$, $SSBL = 1.4188$, $SSTR = 1.3203$

$\hat{D} = 0.2727$

$SSBLTR^* = 0.0093$, $MSBLTR^* = SSBLTR^*/1 = 0.0093$

$SSRem^* = SSTO - SSBL - SSTR - SSBLTR^* = 0.0102$, $MSRem^* = SSRem^*/7 = 0.00146$

$F^* = MSBLTR^*/MSRem^* = \frac{0.093}{0.0146} = 6.3613$

$F(0.99, 1, 7) = 12.2464$

Therefore, the decision rule is: if $F^* \geq F(0.99, 1, 7)$, then reject H_0; otherwise, accept H_0. Here, $6.3613 \leq 12.2464$, so we cannot reject H_0. In other words, we conclude that there is no interation effects. The p-value is 0.0397, which is greater than 0.01, leading to the same conclusion.

R code:

```
Yij = c(0.73, 0.67, 0.15, 0.86, 0.75, 0.21, 0.94, 0.81, 0.26, 1.4, 1.32, 0.75, 1.62, 1.41, 0.78)
Yhat = 0.844
SSTO =sum((Yij-Yhat)^2)
rho = c(-0.3273, -0.2372, -0.174, 0.3127, 0.426)
SSBL = 3*sum(rho^2)
tau = c(0.266, 0.148, -0.414)
SSTR = 5*sum(tau^2)
Dhat = (sum(Yij*(rep(rho, each = 3))*(rep(tau, 5))))/((SSBL/3)*(SSTR/5))
SSBLTR_star = sum((Dhat*(rep(rho, each = 3))*(rep(tau, 5)))^2)
SSRem_star = SSTO - SSBL - SSTR - SSBLTR_star
F_star = (SSBLTR_star/1)/(SSRem_star/7)
threshold = qf(0.99, 1, 7)
pVal = 1-pf(F_star, 1, 7)
```

21.8

(a)

Assume the randomized block model is appropriate, we have:

$SSTO = 2.7586$, $SSBL = 1.4188$, $SSTR = 1.3203$, $SSE = SSTO - SSBL - SSTR = 0.0195$

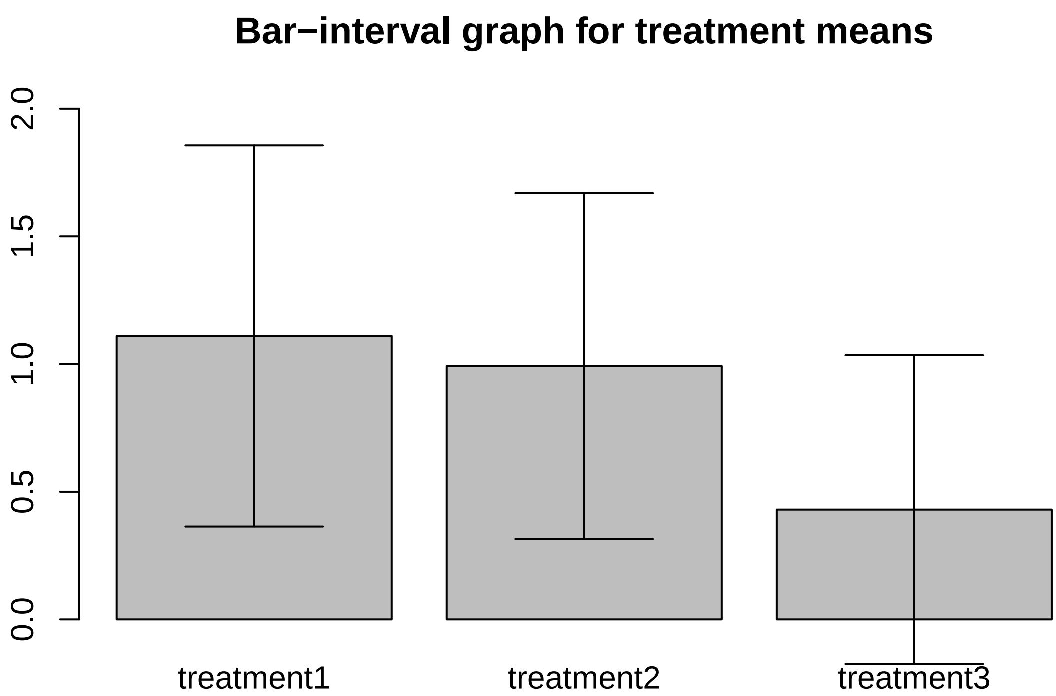
$df(SSTO) = 3 * 5 - 1 = 14$, $df(SSBL) = 5 - 1 = 4$, $df(SSTR) = 3 - 1 = 2$, $df(SSE) = 14 - 4 - 2 = 8$

The ANOVA table is:

Source	df	SS	MS
Block	4	1.4188	0.3547000
Treatment	2	1.3203	0.6601500
Error	8	0.0195	0.0024375
Total	14	2.7586	NA

(b)

From the following bar interval plot, we can see that the means between treatment 1 (Extremely low) and treatment 2 (Fairly low) do not differ much. However, the mean of treatment 3 (Moderately low) differs from other two means substantially.



(c)

H_0: All τ_j equal to zero; H_1: Not all τ_j equals to zero

$F^* = MSTR/MSE = 0.66015/0.0024375 = 270.83$

$F(0.95, 2, 8) = 4.45897$

$p - val = 4.487283e - 08$

The decision rule is: if F^* is greater thatn 4.45897, then reject H_0, otherwise, accept H_1. Here, $270.83 \geq 4.45897$, so we reject H_0, concluding that the mean reduction in lipid level differ for three diets.

The p-value is almost zero, which is less than 0.05, leading to the same conclusion. ## (d) For $L_1 = \mu_{.1} - \mu_{.2}$, the estimated mean and standard deviation are:

$\hat{L}_1 = \bar{Y}_{.1} - \bar{Y}_{.2} = 1.11 - 0.992 = 0.118$

$s^2(\hat{L}_1) = \frac{2MSE}{5} = \frac{2*0.0024375}{5} = 0.00975$, $s(\hat{L}_1) = 0.0312$

The Bonferroni multiplier is: $B = t(0.9875, 8) = 2.7515$

0.95 CI = $(\hat{L}_1 - B * s(\hat{L}_1),\ \hat{L}_1 + B * s(\hat{L}_1))$

Therefore, the 95% confidence interval for L_1 is: (0.0322, 0.2038).

For $L_2 = \mu_{.2} - \mu_{.3}$, the estimated mean and standard deviation are:

$\hat{L}_2 = \bar{Y}_{.2} - \bar{Y}_{.2} = 0.992 - 0.43 = 0.562$

$s^2(\hat{L}_2) = \frac{2MSE}{5} = \frac{2*0.0024375}{5} = 0.00975$, $s(\hat{L}_2) = 0.0312$

The Bonferroni multiplier is: $B = t(0.9875, 8) = 2.7515$

0.95 CI = $(\hat{L}_2 - B * s(\hat{L}_2),\ \hat{L}_2 + B * s(\hat{L}_2))$

Therefore, the 95% confidence interval for L_2 is: (0.4762, 0.6478).

We find that both the CI's for L_1 and L_2 do NOT include zero. It means that the first diet treatment and the second diet treatment have different effects on reduction in lipid level; and the second diet treatment and the third diet treatment have different effects on reduction in lipid level.

(e)

H_0: All ρ_i equal to zero; H_1: Not all ρ_i equals to zero

$F^* = MSBL/MSE = 0.3547/0.0024375 = 145.5179$

$F(0.95, 4, 8) = 3.8379$

$p - val = 1.670994e - 07$

The decision rule is: if F^* is greater thatn 3.8379, then reject H_0, otherwise, accept H_1. Here, $145.5179 \geq 3.8379$, so we reject H_0, concluding that blocking effects are present.

The p-value is almost zero, which is less than 0.05, leading to the same conclusion.

(f)

In this experiment, the response variable is REDUCTION in lipid level. Before being assigned experimental diets, the lipid level of subjects can be considered to be "standard". Imagin we have a control group with standard diet, then the response values within this control group are expected to be around zero. Thus, a control treatment is not needed for this experiment.

21.19

$\hat{E} = \frac{(4*MSBL+4*2*MSE)/(4*3)}{MSE} = \frac{(4*0.3547+8*0.0024375)/12}{0.0024375} = 49.17$

Therefore, the efficiency of the use of blocking is 49.17, indicating a high efficiency.