STAC50 Assignment 3 Solution

(Total : 76 pts)

Q 1. (Total: 18 pts)

(a) (2 pts) $H_o: \mu = 20$ vs $H_a: \mu < 20$

(b) (6 pts: 2 pts for distribution of $\bar{X}$ and 4 pts for probability of type I error, α)

$\bar{X}$ is normally distributed with mean $E(\bar{X}) = \mu$ = 20 when H_0 is true, and standard deviation $\frac{\sigma}{\sqrt{n}} = \frac{4}{\sqrt{20}} = .8944$. Thus under this null distribution, α = P($\bar{X}$ ≤ 17.92) = P(Z ≤ -2.3256) = Φ(-2.3256) = 0.01.

(c) (6 pts: 2 pts for distribution and 4 pts for probability of Type II error, β) When μ = 16.7, $\bar{X}$ is normally distributed with mean 16.7 and standard deviation .8944, so β(16.7) = P($\bar{X}$ > 17.92 when μ = 16.7) = P(Z > [17.92-16.7]/.8944) = P(Z > 1.36) = 1 – Φ(1.36) = .0869.

(d) (4 pts) Replace 17.92 by c, where c satisfies $\frac{c-20}{.8944} = -1.645$ (since P(Z ≤ -1.645) = .05). Thus c = 18.53. Increasing α necessarily decreases β (=0.02).

Q. 2. (Total: 10 pts)

(a) (3 pts: 1 pt each)Treatments: Caliper 1, Caliper 2

Experimental unit: inspectors, Design: Randomized Block design (paired design).

(b) (5 pts) H0: $\mu_d = 0$, d=y1-y2 (paired t-test)

```
> t.test(diff)
        One Sample t-test
data:  diff
t = 0.4318, df = 11, p-value = 0.6742
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.001024344  0.001524344
sample estimates:
mean of x
 0.00025
```

Fail to reject H0 with alpha=0.05

(d)(2 pts)(-0.001024344, 0.001524344)

Q. 3 (Total: 10 pts)

In the experiment, a=5, $n_1=n_2=\ldots n_a$ = 13. Recall that SStr = $\sum_i n_i(\bar{y}_{i.} - \bar{y}_{..})^2$ and SSE = $\sum_i \sum_j (y_{ij} - \bar{y}_{i.})^2$. The

grand mean $\bar{y}_{..} = \frac{\bar{y}_{1.} + \bar{y}_{2.} + \bar{y}_{3.} + \bar{y}_{4.} + \bar{y}_{5.}}{5} = 11.336$

So SStr =2.28176.

SSE = $(n_1-1)s_1^2+(n_2-1)s_2^2+(n_3-1)s_3^2+(n_4-1)s_4^2+(n_5-1)s_5^2$=7.9488

Based on SStr and SSE, one can construct the ANOVA table. F_0 = 4.036, and pvalue = 0.004.

The conclusion is that the treatment means are not all equal.

Q. 4 (Total: 15 pts)

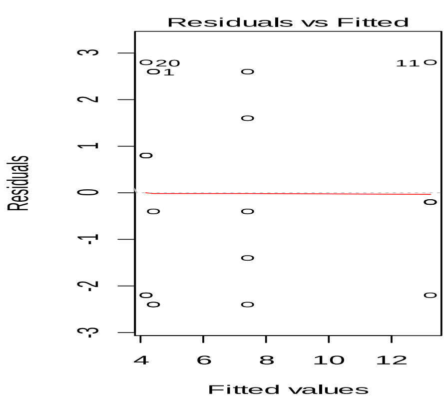
(a) (5 pts): We reject the null hypothesis with 5% level of significance. Not all average amount of defects present the same for all four designs.

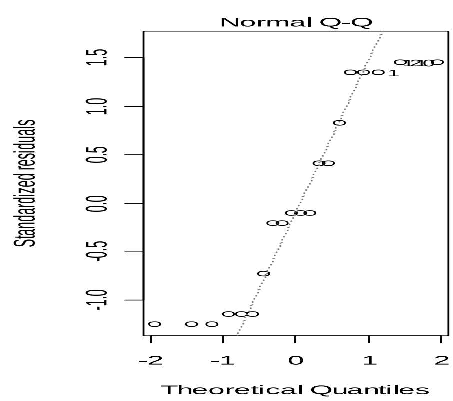
```
 Analysis of Variance Table

Response: defect
                  Df Sum Sq Mean Sq F value    Pr(>F)
factor(design)     3  264.2  88.067  19.041 1.574e-05 ***
Residuals         16   74.0   4.625
---

 Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

(b) (5 pts)QQ plot reveals some departure from normality. This implies that the normality assumption is not valid and the result from ANOVA is questionable.





(c) (5 pts): If they compute Kruskal Wallis test statistic which is same as the computer output then give 3 bonus marks in addition to R output.

```
> kruskal.test(defect, factor(design))


        Kruskal-Wallis rank sum test

data:  defect and factor(design)

Kruskal-Wallis chi-squared = 13.3467, df = 3, p-value = 0.003944
```

The conclusion from Kruskal-Wallis is consistent with that from ANOVA in this problem.

Q. 5 (total: 23 pts)

(a) (5 pts) There is a significant treatments difference.

```
Analysis of Variance Table

Response: Y
               Df Sum Sq Mean Sq F value   Pr(>F)
as.factor(Trt)  3 6026.8 2008.94   6.967 0.002154 **
Residuals      20 5767.0  288.35
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

(b) (5 pts) Using R command, “TukeyHSD” is fine too. Should specify which pairs have the significant difference.

```
> tapply(Y, Trt, mean)
          a         A          b         B
  77.83333 107.00000  65.66667  95.16667
   > qtukey(0.95, 4, 20)
   [1] 3.958293
   CD = qtukey/sqrt(2)*sqrt(MSE(1/6+1/6))=27.44
```

Absolute value of pair of treatment mean differences

a-A	a-b	a-B	A-b	A-B	b-B
29.16667(*)	12.16667	17.33333	41.33333(*)	11.83333	29.5 (*)

(c) (3 pts) Let C1, C2 and C3 be respectively the contrasts for “Hormone I vs Hormone II", “Low Level vs High Level" and ‘Equivalence of Level". Then

$$\begin{aligned} \mathbf{C}_1^t\mathbf{C}_2 &= 1*1+1*(-1)+(-1)*1+(-1)*(-1)=1-1-1+1=0, \\ \mathbf{C}_1^t\mathbf{C}_3 &= 1*1+1*(-1)+(-1)*(-1)+(-1)*1=1-1+1-1=0, \\ \mathbf{C}_2^t\mathbf{C}_3 &= 1*1+(-1)*(-1)+1*(-1)+(-1)*1=1+1-1-1=0. \end{aligned}$$

Hence the three contrasts are orthogonal to each other.

(d)(10pts) 6pt for Contrast anova table and 4 pts for interpretation.

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
C1	1	864.000000	864.000000	3.00	0.0989
C2	1	5162.666667	5162.666667	17.90	0.0004
C3	1	0.166667	0.166667	0.00	0.9811

From the table, Contrast C_1 is close to significant but not significant (p-value = 0.0989), this tells us that the average effect of hormone I and the average effect of hormone II on are not different from each other; Contrast C_2 is very significant (p-value = 0.0004), this shows that the average effect for high levels of hormones and the average effect for low levels of hormones are quite different; Contrast C_3 is not significant at all (p-value = 0.9811), so the difference between the high-level and low-level of hormone I is the same as that between the high-level and low-level of hormone II.