Stat 322 (F17) - Assignment #1

SOLUTIONS (/40)

*(Due ~~Wed. Oct. 4~~ Fri. Oct. 6 at **12:00 pm (noon)** in appropriate STAT 322 slot in assignment box #15 outside the Math Tutorial Centre (MC 4066/4067). Electronic submissions or in-class submissions will not be accepted under any circumstances.)*

1) In response to a complaint that a particular property tax assessor (A) was biased, a study was conducted in which eight properties were randomly selected and each was assessed by assessor A and by an impartial assessor (B). The assessments (10^5 dollars) are shown in the table below.

Property	A	B
1	36.3	35.1
2	48.4	46.8
3	40.2	37.3
4	54.7	50.6
5	28.7	29.1
6	42.8	41.0
7	36.1	35.3
8	39.0	39.1

a) **[3]** Describe how replication, randomization and blocking (if relevant) are employed in this study.

Replication: Each treatment was replicated eight times (i.e., eight properties).

Randomization: Eight sample properties were randomly selected from population of properties. As well, order of allocation of treatment Assessor A, B) would also need to be randomized.

Blocking: Each property constitutes a block, such that each assessor assesses the same property and inferences are drawn from the differences between the assessments for each property (i.e., the differences in the response variate within each block).

b) **[2]** Give the response model that describes the experimental design of this study.

$$Y_{ij} = \mu + \tau_i + \beta_j + R_{ij} \quad R_{ij} \sim G(0,\sigma) \quad ind. \quad i = 1,2 \quad j = 1,2,\ldots,8$$

c) **[5]** Is there a difference in assessed values between assessors A and B? **Answer this question** by calculating manually (i.e. <u>without</u> using dedicated functions in R, such as *anova*, *t.test*, *confint*, etc.) a 95% confidence interval for the difference in treatment effects, $\tau_1 - \tau_2$ (you may use R to perform arithmetic operations and to obtain standard deviations and means, using the *sd* and *mean* functions, respectively).

95% confidence interval for $\tau_1 - \tau_2$:

$$\hat{\tau}_1 - \hat{\tau}_2 \pm t * \frac{s_d}{\sqrt{n_d}}$$

where $t* = t_{.025,7df}$ = 2.365

s_d = standard deviation of the differences

n_d = number of differences (i.e., number of properties (blocks) in this case)

$$= 1.488 \pm 2.365 \frac{1.491}{\sqrt{8}} = 1.488 \pm 1.246 = (0.242, 2.734)$$

Since the interval for the mean difference does not contain 0, we can conclude that there is a significant difference in the mean assessment between the two assessors. More specifically, we can conclude that, on average, assessor A assesses properties in the population of interest at a greater value than does assessor B.

d) **[3]** Verify the confidence interval obtained in c) by using the *t.test* function in R.

```
> A
[1] 36.3 48.4 40.2 54.7 28.7 42.8 36.1 39.0
> B
[1] 35.1 46.8 37.3 50.6 29.1 41.0 35.3 39.1
> t.test(A,B, paired=TRUE)

        Paired t-test

data:  A and B
t = 2.8211, df = 7, p-value = 0.02573
alternative hypothesis: true difference in means not equal to 0
95 percent confidence interval:
 0.2407052 2.7342948 (consistent with that obtained in c) above)
sample estimates:
mean of the differences
                 1.4875
```

e) **[2]** Note the small sample sizes. For confidence intervals and hypothesis tests for $\tau_1 - \tau_2$, what assumptions are required about the differences in assessments between the two assessors for each property?

Based on our model in b) above, we must assume that the residuals (and therefore the within block differences) are approximately normal for *t* – based methods of inference to be valid.

f) **[5]** Suppose the data were (incorrectly) analyzed as having come from a sampling protocol in which 16 properties were randomly selected, and each assessor was randomly assigned to assess eight of these properties each. Answer the same question posed in c) by (manually) performing a (two-sided) hypothesis test. As with all hypothesis tests, be sure to include the null and alternative hypothesis, value of test statistic (with degrees of freedom), p-value (using the *pt* function in R), and conclusion in the context of the study.

Using R for summary calculations (not necessary):

```
> mean(A)-mean(B)

[1] 1.4875

> s_A=sd(A)

> s_B=sd(B)

> sigma_hat=sqrt((7*(s_A^2)+7*(s_B^2))/14)

> SE=sigma_hat*sqrt(1/8+1/8)

> SE

[1] 3.722192

t=mean(A-B)/SE

> t

[1] 0.3996302
```

$H_0 : \tau_1 - \tau_2 = 0$

$H_a : \tau_1 - \tau_2 \neq 0$

$$t = \frac{\hat{\tau}_1 - \hat{\tau}_2 - (\tau_1 - \tau_2)}{SE(\hat{\tau}_1 - \hat{\tau}_2)} = \frac{1.4875}{3.7222} = 0.400$$

p-value = 0.6955 (in R: > 2*(1-pt(0.3996302,14)) yields 0.6954592)

Do not reject H_0. There is no significant difference in mean assessed values of the two assessors.

g) **[3]** Verify the hypothesis test results obtained in f) by using the *t.test* function in R.

```
> t.test(A,B, paired=FALSE, var.equal=TRUE)


        Two Sample t-test


data:  A and B
t = 0.3996, df = 14, p-value = 0.6955
alternative hypothesis: true difference in means is not equal to
0
95 percent confidence interval:
 -6.495807  9.470807
sample estimates:
mean of x mean of y
  40.7750   39.2875
```

h) **[2]** Was your conclusion in incorrectly analyzing the data in this way different from that in c)? Briefly explain why.

Yes, the conclusion was different. The p-value increased from 0.02753 for the correct analysis to 0.6955 for the incorrect analysis (based on no blocking), resulting in an invalid conclusion of no significant difference.

The reason for this large discrepancy is that the variation in the values of the assessed properties is accounted for by employing blocking, resulting in a relative small estimate of σ, whereas by assuming no blocking, that variation is now included in the random variation, thus yielding an increased estimate of σ, and associated increases in p-values and wider confidence intervals.

2) A major greeting card company has negotiated a deal with a credit card company to include a coupon for a greeting card set in every cardholder's monthly statement. The greeting card company has come up with 4 coupon designs that it wishes to test to determine which coupon is most effective in generating greeting card set sales. After sending out the 4 different coupons to randomly selected credit card customers in a trial run, a random sample of 8 customers who redeemed their coupons is selected from each coupon design (8 customers for each design), and the dollar value of the card set ordered for each customer is recorded. The data are provided in the following table.

	Coupon Design			
Customer	**#1**	**#2**	**#3**	**#4**
1	86	123	75	33
2	119	204	243	88
3	77	145	98	65
4	127	143	123	109
5	107	98	34	88
6	123	202	76	68
7	111	209	69	88
8	44	156	89	88
mean	**99**	**160**	**101**	**78**
std. dev.	**28**	**41**	**63**	**23**

a) **[3]** Define the factor, factor levels, treatment, and response variate for this study.

Factor: Coupon Design

Factor Levels: #1, #2, #3, #4.

Treatment: Coupon Design Type

Response Variate: Dollar value of card set ordered.

b) **[2]** Describe how replication, randomization and blocking (if relevant) are employed in this study.

Replication: Each coupon design (treatment) was given to eight customers.

Randomization: The eight customers were selected randomly from population of interest. Each of the four treatments (coupon design) were randomly assigned to 8 customers.

Blocking: Blocking was not employed in this study.

c) **[2]** Give the response model that describes the experimental design of this study.

$$Y_{ij} = \mu + \tau_i + R_{ij} \quad R_{ij} \sim G(0,\sigma) \quad ind. \quad i = 1,2,3,4 \quad j = 1,2,...,8$$

d) **[5]** Use the means and standard deviations of the sales for the different coupon designs to generate manually an ANOVA table for this data. Use this table to test the hypothesis that there is no difference in mean sales between the 4 coupon designs. Be sure to include the appropriate hypotheses, value of the test statistic, p-value (using the *pf* function), and conclusion in the context of the study.

Overall mean(from summaries provided) = 109.5

SS(Treatment) = 8(99-109.5)2+8(160-109.5)2+8(101-109.5)2+8(78-109.5)2

= 29800

SS(Residual) = pooled sum of squares from within treatment standard deviations:

= 282(8-1)+ 412(8-1)+ 632(8-1)+ 232(8-1)

= 48741

Source:	df	Sum Sq	Mean Sq	F value	p-value
treatment	3	29800	9933	5.706	0.0035
Residuals	28	48741	1741		

---

test statistic: F = 5.706

p-value = 0.0035(from R: > 1 - pf(5.706,3,28))

Reject H_0 (p-value <.05). There is a significant difference in the mean sales between the different coupons. (Or, in the context of the model parameters, there is a significant difference in the treatment effects of the different coupons on sales.)

e) **[3]** Use R to verify that the values in the ANOVA table and subsequent *F*-test that you calculated by hand in d) are correct (Note: there may be some discrepancy due to round-off error)

```
Analysis of Variance Table

          Df Sum Sq Mean Sq F value   Pr(>F)
treatment  3  29587    9862  5.6636 0.003665 **
Residuals 28  48758    1741
---
```