A N S W E R K E Y

Name:

STAT1051: QUIZ 10 Take Home

Many people in the world are very poor. The salary of a professional like a college teacher in the Philippines might be $400 per month, yet manufactured goods such as appliances and automobiles cost as much there as in the United States. In spite of the impoverishment of the country and the general disdain of the rich for the commoners as manifested in the corrupt judicial system and in the government at all levels, efforts are made to address the medical needs of the rural population. A fascinating account of the success of implementing the principles of public health is given in the book, *An American Doctor's Odyssey*, a best seller in 1936 and written by Victor Heiser. He is given credit for saving the lives of two million people during his tenure as a public health doctor in the Philippines and as a program director for the Rockefeller Foundation. In his opinion, the best medicine is preventive medicine. Dr. Heiser died in 1972 at the age of 99.

In an effort to treat common public heath conditions in rural Philippines, a special advertising campaign is conducted to announce the availability of free medical and dental examinations, free anti-rabies vaccines, free anti-dengue vaccines, free denture, and free circumcisions. However, not everyone takes advantage of these free services. Based on a survey, 236 residents of a rural community were interviewed and they were asked if they were aware of each of the five programs and if they utilized them. The following are the results of the survey.

Service	Medical and Dental	Antirabies	Antidengue	Denture	Circumcision
Aware	74	47	49	15	51
Utilized	45	15	29	2	12

A two parameter fixed effects linear model where SSE=254.9003 was asserted in which the response variable is the number of people who utilized a program and the explanatory variable is the number of people who were aware of a given program.

$\sum_{i=1}^{5} x_i = 236$ $\sum_{i=1}^{5} x_i^2 = 12912$ $\sum_{i=1}^{5} x_i y_i = 6098$ $\sum_{i=1}^{5} y_i = 103$
$\sum_{i=1}^{5} y_i^2 = 3239$ $SS_{xx} = 1772.8$ $SS_{xy} = 1236.4$

FOR THE FOLLOWING ANSWERS: Note $\bar{y} = 20.6$ $\bar{x} = 47.2$ x=number of those who were aware y=number of those who utilized the services.

1. **Answer 1** *(5 pts) Find* $\widehat{\beta}_0 = \bar{y} - \widehat{\beta}_1 \bar{x} = 20.6 - 0.6974278(47.2) = -12.3185921$
2. **Answer 2** *(5 pts) Find* $\widehat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}} = 0.6974278$
3. **Answer 3** *(5 pts) Write the fitted equation of a two parameter linear fixed effects model for* $\widehat{E[y_i]} = -12.3185921 + 0.6974278 x_i$

4. **Answer 4** *(30 pts) Test the hypothesis,* $H_0 : \beta_1 = 0$ *vs* $H_1 : \beta_1 \neq 0$ *at a level of significance,* $\alpha = .05$*. Complete the ANOVA table.*

Source of Variation	df	Sum of Squares	Mean Sum of Squares	F statistic
Mean	A=**1**	2121.8		
Regression	B=**1**	862.2997	G=**862.2997k**	
Residual Error	C=**3**	E=**254.9003**	H=**84.96676**	I=**10.14867**
Total	D=**5**	F=**3239**		

Give $F_{1,3,.05}$*=* ***J=10.12796***

Reject or cannot reject H_0***?*** *K=Reject Why? L=* ***Because*** $10.14867 > 10.12796$*;* ***therefore, reject the null hypothesis. But it is not a convincing rejection.***

Note that the p-value is: 0.04987537

5. **Answer 5** *(5 pts) What is* $s^2 =$ ***84.96676***

6. **Answer 6** *(5 pts) Assume that the fitted model is:* $\widehat{E[y_i]} = -10 + 0.55x_i$ *What is the estimated utilization of a program if 1000 people learned about it?* $\widehat{E[y_i]} = -10 + .55(1000) = 540$

7. The 95% confidence interval of the expected number of people to visit a clinic in a community of 100 people in which everyone was canvassed is (18, 96). If you were to supervise the clinic, how many people would you prepare your staff to treat? Explain.

   Answer 7 *(5 pts) I would prepare for 96 people to enter the clinic because the upper limit of the 95% CI suggests that I will be roughly 95% confident that the number of guests will be 96 and therefore I will be confident that the clinic will be adequately staffed and have adequate stores to accommodate the crowd of visitors.*

8. **Answer 8** *(15 pts) Explain whether or not the model is a good model.*

   *(a) M=****The plot of the data indicates a linear relationship between x and y.***

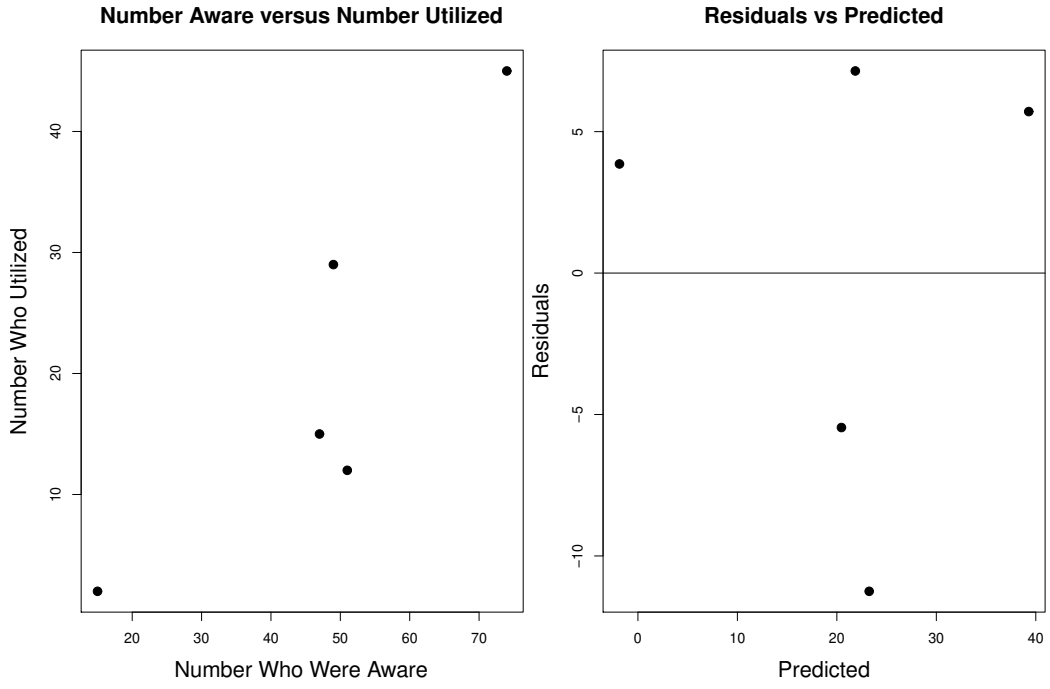
   *(b) N=****The results of the survey indicate that the more people who are aware of a service, the more will come to receive the services which is what one would expect so the theory makes sense.***

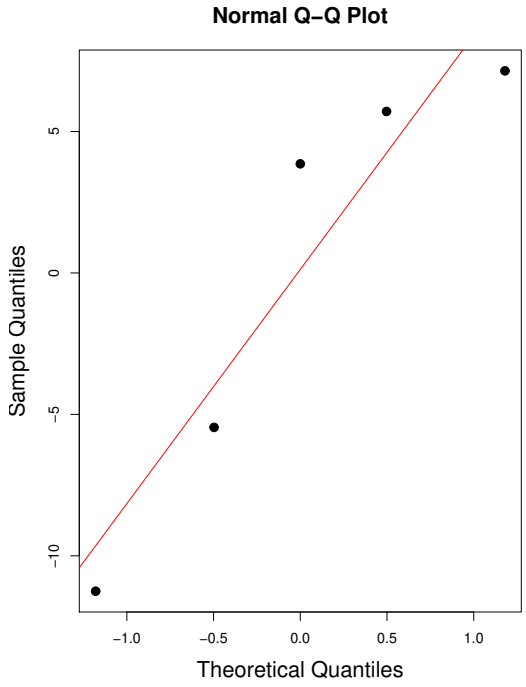
   *(c) O=****The null hypothesis can be rejected but only by thinnest of margins***

(d) P=***The plot of residuals versus predicted values shows a random pattern, but the number of observations is rather small. The qq plot shows a straight diagonal line, but it is not a strong one; therefore, normality of the residuals seems acceptable, but normality of the residuals is not very good.***

Final answer based on the four answers given above. Is the model a good model? R= ***All indications support the claim that the model is a good one but it is not a strong one. Probably, for the kind of use which the model will receive, it should be adequate.***

Service	Medical and Dental	Antirabies	Antidengue	Denture	Circumcision
$\widehat{E[y]}$	39	20	22	0	23
residuals	6	-5	7	3	-11





9. From the pigmentation survey of Scottish children (Tocher's data) the following are the numbers of boys and girls from the same district whose hair color falls into each of five classes. This problem is taken from page 87 of Ronald Fisher's book, *Statistical Methods for Research Workers*, published in 1946 by Oliver and Boyd.

   In the following tabulation of the data, the original frequencies are at the bottom of each cell; the expected frequency is in the upper right hand corner within parentheses; and the contribution to the chi-squared test statistic is in the upper left corner.

	Fair	Red	Medium	Dark	Jet Black
Boys	.81455 (614.37) 592	.0407 (116.61) 119	.6811 (825.28) 849	.3016 (516.48) 504	2.968 (27.04) 36
Girls	.9593 (521.62) 544	.0480 (99.18) 97	.8022 (700.71) 677	**.3552 (438.51)** 451	3.4959 (22.95) 14

To test the hypothesis at $\alpha = .05$ that the distribution of children according to color of hair is independent of their sex, find the following:

(a) **Answer 9** *(10 pts) Complete the table by filling in the blanks.* *XXX*=$\frac{(451-438.51)^2}{438.51}$ = .3552 *YYY*=$\frac{955(1783)}{3883}$ = 438.51

(b) **Answer 10** *(5 pts) Degrees of freedom=* ***4***

(c) **Answer 11** *(5 pts) The appropriate X^2 quantile=* $X^2_{4,.05} = \mathbf{9.487729}$

(d) **Answer 12** *(5 pts) Knowing that the test statistic* $X^2 = 10.467$*, can the hypothesis be rejected or not rejected?* ***Reject*** *Give your decision with justification.* ***Because*** $X^2 = 10.467 > 9.487729 = X^2_{4,.05}$ *There appears, therefore, a difference in the color of hair by sex of the child. (We surmise that specifically more boys than girls have jet black hair, because it is that color which contributes the most to the X^2 test statistic.)*
(Note: p-value=0.03324834)

BONUS The distribution of colors of 400 M & M's are as follows:

Color	Brown	Yellow	Red	Orange	Green	Blue
Observed	84	79	105	49	36	47
Expected	**66.667**	66.667	66.667	66.667	66.667	66.667
Deviation	**17.333**	12.333	38.333	-17.667	-30.667	-19.667

Test the hypothesis, at the .05 level of significance, that the distribution of colors is uniform.

(a) **Answer 13** ***BONUS*** *(2 pts) Complete the table: xxx=****66.6667*** *yyy=****17.333***

(b) **Answer 14** ***BONUS*** *(2 pts)* $X^2_{n-1,\alpha} = X^2_{5,.05} = \mathbf{11.07050}$

(c) **Answer 15** ***BONUS*** *(2 pts)* $X^2 = \sum_{i=1}^{n} \frac{(observed_i - expected_i)^2}{expected_i} = \frac{300.44}{66.667} + \frac{152.11}{66.667} + \frac{1469.44}{66.667} + \frac{312.11}{66.667} + \frac{940.44}{66.667} + \frac{386.77}{66.667} = \mathbf{53.42}$

(d) **Answer 16** ***BONUS*** *(2 pts) Can the null hypothesis that the colors are uniformly distributed be reject?* ***Yes*** *Explain* ***Because*** $X^2 = 53.42 > 11.07050$ *(Note that* $p - value = P(X^2 > 53.42) = 1 - pchisq(53.42, 5) = 3E{-}10$*)*