ANSWERS

STAT112: Mock-Midterm Examination

1. On an examination in psychology, 12 students in one class had a sample mean grade of 78 with a sample standard deviation of 6, while 15 students in another class had a sample mean grade of 74 with a sample standard deviation of 8. Using a level of significance of .05, determine whether the first group is superior to the second group.

   Answer 1 $H_0 : \mu_1 - \mu_2 = 0$ *vs* $H_1 : \mu_1 - \mu_2 > 0$

   (a) Is $T = 1.4366 > 1.728$*? NO*

   (b) Cannot reject null hypothesis

2. What is the difference between these two estimators: $\widehat{E[y_i]}$ and $\hat{y}_i$ where y_i is the response variable of a linear model?

   Answer 2 $\widehat{E[y_i]}$ *is an estimate of the average, while* $\hat{y}_i$ *is an estimate of a particular value.*

3. Given the samples, $\mathcal{S}_1 = \{6\ 2\ 1\ 3\ 3\}$ and $\mathcal{S}_2 = \{5\ 3\ 2\ 4\ 2\}$, test the hypothesis: $H_0 : \mu_1 = \mu_2$ vs $H_1 : \mu_1 \neq \mu_2$ at the .05 level of significance.

   Answer 3 $H_0 : \mu_1 - \mu_2 = 0$ *vs* $H_1 : \mu_1 - \mu_2 \neq 0$

   (a) Is $T = -.196 > 2.306$*? NO*

   (b) Cannot reject null hypothesis

4. Using the same two samples of problem 3, find the 95% CI of $\mu_2 - \mu_1$.

   Answer 4 *(a)* $a = 3.2 - 3 - sqrt(2.598) * sqrt(1/5 + 1/5) * 2.306 - 2.15$

   (b) $b = 3.2 - 3 + sqrt(2.598) * sqrt(1/5 + 1/5) * 2.306 = 2.55$

   (c) 95% C.I. of $\mu_1 - \mu_2$ *is (-2.15,2.55)*

5. Let t_0 represent a particular value of Student's t distribution. Find the value of t_0 which will make the following statement true: $P(t \leq -t_0 \text{ or } t \geq t_0) = .10$ with df=12.

   Answer 5 $t_{12,.05} = 1.782$

6. A random sample of 36 drinks from a soft-drink machine has an average content of 8 ounces with a standard deviation of .48 ounces. The vendor claims that the machine gives 7.5 ounces per cup. Are the customers being given too much? That is, test the hypothesis: $H_0 : \mu = 7.5$ ounces vs $H_1 : \mu > 7.5$ ounces at the .05 level of significance.

Answer 6 *(a)* $H_0 : \mu_= 7.50$ *vs* $H_1 : \mu > 7.5$

(b) $6.25 > 1.684$*? Yes, reject null hypothesis.*

7. Given the following data of the rate by which laborers per 100 quit their jobs and of their hourly wage, a linear fixed effects model was proposed where the response variable is the rate of attrition and the explanatory variable is the average wage.

Rate of Attrition	1.4	.7	2.6	3.4	1.7	1.7	1.0	.5
Average Wage	\$8.2	10.35	6.18	5.37	9.94	9.11	10.59	13.29
Age	36	45	25	20	44	40	46	55
Rate of Attrition	2.3	1.9	1.4	1.8	2.0	2.0	3.8	
Average Wage	\$7.5	6.43	8.83	10.93	8.80	7.99	5.549	
Age	35	21	33	27	39	47	39	

(a) Write the two parameter fixed effects linear model where the response variable is rate of attrition and the explanatory variable is average wage:

Answer 7 $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ *where* $\epsilon_i \sim N(0, \sigma^2)$

(b) Compute $\hat{\beta}_0$=

Answer 8 *4.8619*

(c) Compute $\hat{\beta}_1$=

Answer 9 *-.3465*

(d) When the wage is \$7.50 per hour, what is the estimate of the expected rate of attrition?

Answer 10 *2.26245*

(e) Find the 95% confidence interval for that predicted sample average, $E[\widehat{E[y_i]}]$.

Answer 11 *(1.957,2.568)*

(f) Test the hypothesis, $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$ at a level of significance, $\alpha = .05$.
$F_{r-1,n-r,\alpha}$= 4.667

(g) Give s=

Answer 12 *.4862*

(h) Compute the 95% confidence interval for β_1.

Table 1: Analysis of Variance for Fitting Regression

Source of Variation	df	Sum of Squares	Mean Sum of Squares	F statistic
Mean				
Regression				34.81647
Residual Error				
Total				

Answer 13 *(-.4732,-.2196)*

(i) An age*wage interaction term is added to the model. Write this model:

Answer 14 $y_i = \beta_0 + \beta_1 wage + \beta_2 wage * age + \epsilon_i$ *where* $\epsilon_i \sim N(0, \sigma^2)$

(j) The resulting abbreviated ANOVA is given in Table 2

Table 2: Analysis of Variance for Fitting Regression

Source of Variation	df	Sum of Squares	Mean Sum of Squares	p-value
wage	1	8.2507	39.8792	3.86e-05
age*wage	1	0.5906	2.8547	0.1169
Residual Error	12	2.4827		

i. Is the interaction term significant at a level of significance of .05? Explain.

Answer 15 *No, because* $\alpha = .05 \not> .1169 = p - value$

ii. The estimate of the interaction coefficient is: 0.016743. (The estimate of β_0 is: 10.317015 and the estimated wage coefficient is: -1.640321.) Interpret what this estimate of the interaction coefficient means to someone who never took statistics.

Answer 16 *While the rate of attrition is inversely related to wage, the combination of age and wage contributes only a little positive effect on the rate of attrition.*

8. One die is tossed 180 times with the following results:

Face	1	2	3	4	5	6
Count	28	36	36	30	27	23

Is the die fair? Use a .01 level of significance.

Answer 17 *(a) Is* $X^2 = 4.46667 > 15.09$*?NO*

(b) Cannot reject the null hypothesis that the dice are uniformly distributed.

9. An association of homeowners of a residential sub-division requested from the local jurisdiction painted crosswalks to be placed across a busy street where there are a school bus and municipal bus stops arguing that the crosswalks will improve the safety of their children and pedestrians. The state traffic engineer argued against the proposal saying that out of 54,750,000 vehicles passing by that spot in the past ten years, only two accidents occurred and none with pedestrians. Therefore, based on accident statistics spending $200 for painted crosswalks is not justified. Is the statistical argument convincing? Explain.

Answer 18 *This is an example of lying with statistics. For a measely $200, an effort to prevent a catastrophic event.*