A N S W E R K E Y

Stat 1111: Mock–Final Examination

1. Suppose $X \sim U(3, 8)$; find a such that $P(X \geq a) = .6$. ANS: a=5

2. A manufacturing company wants to determine the difference of the productivity of the day shift and the productivity of the night shift. A sample of 12 day shift workers reveals an average output of 74.3 parts per hour with a sample standard deviation of 16 parts per hour. For a sample of 10 night shift workers, the average output is 69.7 parts per hour with a sample standard deviation of 18 parts per hour. Respond to a vice-president's claim that the day shift workers are more productive than the night shift workers at a level of significance of .05.

   ANS:

   (a) $H_0 : \mu_1 - \mu_2 = 0$ vs $H_1 : \mu_1 - \mu_2 > 0$ at the level of significance of .05.

   (b)

Sample 1	Sample 2
$n_1 = 12$	$n_2 = 10$
$\bar{x}_1 = 74.3$	$\bar{x}_2 = 69.7$
$s_1 = 16$	$s_2 = 18$

   $S_p^2 = 286.6$

   (c) $t_{\nu;\alpha} = 1.725$

   (d) $T = .6345$

   (e) Is $T = .6345 > 1.725$? NO

   (f) Cannot reject null hypothesis

3. A random sample of n=25 observations is drawn from a normal population with sample mean equal to 10 and a population standard deviation equal to 9. Find c such that P($\bar{x} \geq c$)=.17.

   ANS: c=11.728

4. If $X_i \sim N(\mu, \sigma^2)$ and the X_i's are iid, what is $P(\mu \in (\bar{x} - \frac{s}{\sqrt{n}} t_{n-1,\alpha/2}, \bar{x} + \frac{s}{\sqrt{n}} t_{n-1,\alpha/2})) = 1 - \alpha$

5. Complete the following statement: The truth or falsity of a statistical hypothesis is never known with certainty unless we : ANS: Given in lecture notes.

6. A farmer claims that the average yield of corn of variety A exceeds the average yield of variety B by at least 12 bushels per acre. To test this claim, 6 one acre plots of each variety are planted and grown under similar conditions. Variety B yielded, on the average, 116.7 bushels per acre with a sample standard deviation of 8 bushels per acre, while variety A yielded, on the average, 127.8 bushels per acre with a sample standard deviation of 6 bushels per acre. Test the farmer's claim using a .05 level of significance.

   $H_0 : \mu_2 - \mu_1 = 12$ vs $H_1 : \mu_2 - \mu_1 > 12$

(a)

Sample 1	Sample 2
$n_1 = 6$	$n_2 = 6$
$\bar{x}_1 = 116.7$	$\bar{x}_2 = 127.8$
$s_1 = 8$	$s_2 = 6$

(b) $S_p^2 = 50$

(c) $t_{\nu;\alpha} = 1.812$

(d) $T = -.220$

(e) Is $T = -.220 > 1.812$? NO

(f) Cannot reject null hypothesis

7. A random sample of 50 mathematics grades showed a mean of 75 and a standard deviation of 10. What is the 95% confidence interval for the population mean, μ. ANS: 95% C.I. of μ is (72.17,77.828)

8. 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.

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

   (a)

Sample 1	Sample 2
$n_1 = 12$	$n_2 = 15$
$\bar{x}_1 = 78$	$\bar{x}_2 = 74$
$s_1 = 6$	$s_2 = 8$

   (b) $S_p^2 = 51.68$

   (c) $t_{\nu;\alpha} = 1.708$

   (d) $T = 1.4366$

   (e) Is $T = 1.4366 > 1.728$? NO

   (f) Cannot reject null hypothesis

9. 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? ANS: Given in lecture notes.

10. 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. ANS: $H_0 : \mu_1 - \mu_2 = 0$ vs $H_1 : \mu_1 - \mu_2 \neq 0$

    (a)

Sample 1	Sample 2
$n_1 = 5$	$n_2 = 5$
$\bar{x}_1 = 3$	$\bar{x}_2 = 3.2$
$s_1 = 1.87$	$s_2 = 1.3038$

    (b) $S_p^2 = 2.598$

    (c) $t_{\nu;\frac{\alpha}{2}} = 2.306$

(d) $T = -.196$

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

(f) Cannot reject null hypothesis

11. Using the same two samples of the previous problem , find the 95% CI of $\mu_2 - \mu_1$. ANS: 90% C.I. of $\mu_1 - \mu_2$ is (-2.15,2.55)

12. Let t_o represent a particular value of Student's t distribution. Find the value of t_o which will make the following statement true: $P(t \leq -t_o \text{ or } t \geq t_o) = .10$ with df=12. ANS: 1.782

13. 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. ANS:

    (a) $H_0 : \mu$=7.50 vs $H_1 : \mu > 7.5$

    (b) $t_{n-1,\alpha} = 1.684$

    (c) $T = 6.25$

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

    Given the following data of the rate by which laborers per 100 quit their jobs and of their hourly 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

Rate of Attrition	2.3	1.9	1.4	1.8	2.0	3.8
Average Wage	\$7.5	6.43	8.83	10.93	7.99	5.54

a linear fixed effects model was proposed where the response variable is the rate of attrition and the explanatory variable is the average wage.

14. Write the linear model: ANS: Given in lecture notes.

    $\sum_{i=1}^{15} x_i = 129.0559$, $\sum_{i=1}^{15} x_i^2 = 1179.06$, $\sum_{i=1}^{15} x_i y_i = 218.840$, $\sum_{i=1}^{15} y_i = 28.2$, $\sum_{i=1}^{15} y_i^2 = 64.34$, $SS_{xx} = 68.6448$, $SS_{xy} = -23.7907$, and $SSE = 3.07334$.

15. Compute $\hat{\beta}_0$=4.8619

16. Compute $\hat{\beta}_1$=-.3465

17. When the wage is \$7.50 per hour, what is the estimate of the expected rate of attrition= 2.26245

18. Find the 95% confidence interval for that predicted sample average, $E[\widehat{E[y_i]}]$. (1.957,2.568)

19. 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

Table 1: Analysis of Variance for Fitting Regression

Source of Variation	df	Sum of Squares	Mean Sum of Squares	F statistic
Mean	1	53.016		
Regression	1	8.2453	8.2507	34.81647
Residual Error	13	3.0786	.2364	
Total	15	64.34		

20. Give s=.4862
21. Compute the 95% confidence interval for β_1. (-.4732,-.2196)
22. One dice is tossed 180 times with the following results:

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

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

(a) $X^2_{n-1,\alpha} = 15.09$

(b) $X^2 = 4.46667$

(c) Is $X^2 = 4.46667 > 15.09$?NO

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

23. A random sample of 13 observations is drawn from a normal population. The sample mean, $\bar{x} = 28$ and the sample variance, $s^2 = 9.5$. Find c such that P ($\bar{x} \leq c$)=.95.

ANS: c=29.524

24. Draw a graph of the line: y=3-.75x where $0 \leq x \leq 4$.

BONUS

Sonnet 135 by William Shakespeare

Whoever hath her wish, thou hast thy *Will*,
And *Will* to boot, and *Will* in overplus;
More than enough am I that vex thee still,
To thy sweet will making addition thus.
Wilt thou, whose will is large and spacious,
Not once vouchsafe to hide my will in thine?
Shall will in others seem right gracious,
And in my will no fair acceptance shine?
The sea, all water, yet receives rain still
And in abundance addeth to his store;
So thou, being rich in *Will*, add to thy *Will*
One will of mine, to make thy large *Will* more.
Let no unkind, no fair beseechers kill;
Think all but one, and me in that one *Will*.

1. (2pts) William Shakespeare wrote 154 sonnets. Estimate the number of times that the word, will (including any conjugated or declined forms of it), appears in all of his sonnets. (Hint: every sonnet consists of exactly fourteen lines). ANS: There are fourteen instances of the word, will. $14(154) = 2156$

2. (3pts) What makes the confidence interval which is associated with the above estimate of the population total **INVALID**? The sonnet is not representative of the collection of sonnets. The estimate is therefore biased hence the CI is not valid.