Final Exam- Applied Statistics- Third Year- English Section- First Semester, 2018- Time Allowed 3 hours

FROM D

Answer the following two sets of questions, mark your final answers on the provided electronic answer sheet. Show all your preliminary work in the provided answer booklet . (15 points)

Do not forget to mark the form number in the provided space on the electronic sheet

FORM D

MCQ: Mark the answer that best fits your answer

Use the following to answer questions 1-4

$x_1 = 100$ $n_1 = 400$ $x_2 = 150$ $n_2 = 300$

1. . The pooled estimated proportion to test $(P_2 - P_1) = 0$ is:
   a. .25 b. .75 c. .50 d. .357 .
2. In (1) above, the standard error for the difference in the two populations is:
   a. .0013 b. .0316 c. .0366 d. .025
3. The upper bound for a 95% confidence interval to estimate $(P_2 - P_1)$ is:
   a. .309 b .322 c. .059 d. .071
4. The margin of error when constructing 95% CI is:
   a. .322 b. .0366 c. .357 d. .0717
5. To test that the population variance = 25 against that it does not equal 25, and a sample of size 16 yields a variance of 20, the chi square value for the above test is:
   a. 12 b. 25 c. .80 d. 12.8
6. In (5) above the degrees of freedom is :
   a. 20 b. 15 c. 14 d. 16
7. The upper critical value for the test in (5) above at $\propto = 5\%$ *is* :
   a. 27.48 b. 25 c. 10.85 d. 6.26
8. The lower critical value for the test in (5) above at $\propto = 5\%$ *is* :
   a. 27.48 b. 25 c. 10.85 d. 6.26

Use the following contingency table and answer questions 9-11

9. A researcher wants to test if "income" and Education Level" are correlated or not, the null hypothesis is that the two variables are:
   a. correlated b. independent
   c. categorical d. identical.

Education level	*Income*		*Total*
	high	*Low*	
University	45	15	60
Secondary	20	30	50
Basic	15	35	50
Total	80	80	160

10. *The critical value (α=5%) to test the above hypothesis is:*
    a. 5.99 b. 7.61 c. 12.59 d. 11.07
11. *The Chi-Square value for testing the null hypothesis is:*
    a.20.2 b..86 c. 25 d 60

True/ False Questions

1. The ratio of two variances follows the F distribution .
2. The F variable has two parameters, the degrees of freedom for the variance in the numerator and the degrees of freedom for the denominator.
3. The F value with 10, 8 degrees of freedom and $\alpha=.05$ is 3.35
4. When $n_1=10$, $n_2=7$, then the critical value for the t-test to test that the difference in two population means equals to = 0 at $\alpha=.05$ is 1.740.
5. The z test for testing the difference in two population proportions is always valid only when sample sizes are large.
6. Under the homogeneity of variance assumption, the pooled variance is an estimate for the two population variances.
7. If a confidence interval for the difference in the two population means contains zero, then the null hypothesis of zero difference between the two population means is rejected at any level of significance.
8. In a simple regression ANOVA table, the $SS_{regression}$ not always = $MS_{regression}$.
9. In a simple regression ANOVA table, with n=14, the critical value for testing the fitness of the model at $\alpha = 5\%$ equals 4.60
10. Testing that the correlation coefficient in the population =0 is equivalent to testing that the slope coefficient in the population =0 .
11. The correlation coefficient and the slope coefficient do not always take the same sign.
12. When testing for the fitness of the simple regression model, the null hypothesis is that the simple regression model $=\overline{y}$.
13. When comparing two or more population means, the one way ANOVA is used.
14. In one way ANOVA, the total variability is partitioned into variability between means and variability within means.
15. In one way ANOVA , when N=16, k=3, then the critical value for the F test statistic at $\alpha =$ 5% is 3.29
16. When testing that two or more population proportions have different proportions, the chi goodness of fit test is used.
17. In ANOVA, the critical value is always to the right with probability $=\frac{\alpha}{2}$.
18. In all ANOVA tables, the Mean squares column (MS) represents variances.
19. Rejection of the null hypothesis for the chi-square test of independence means that the two variables are related.
20. In a chi-square test of independence with variable 1 consists of 4 category and variable 2 has 3 categories, then the degrees of freedom for the test statistic are 6 df.

Chi-square Distribution Table

d.f.	.995	.99	.975	.95	.9	.1	.05	.025	.01
1	0.00	0.00	0.00	0.00	0.02	2.71	3.84	5.02	6.63
2	0.01	0.02	0.05	0.10	0.21	4.61	5.99	7.38	9.21
3	0.07	0.11	0.22	0.35	0.58	6.25	7.81	9.35	11.34
4	0.21	0.30	0.48	0.71	1.06	7.78	9.49	11.14	13.28
5	0.41	0.55	0.83	1.15	1.61	9.24	11.07	12.83	15.09
6	0.68	0.87	1.24	1.64	2.20	10.64	12.59	14.45	16.81
7	0.99	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21
11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.72
12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22
13	3.57	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69
14	4.07	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14
15	4.60	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58
16	5.14	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00
17	5.70	6.41	7.56	8.67	10.09	24.77	27.59	30.19	33.41
18	6.26	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81

F Distribution: Critical Values of *F* (5% significance level)

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.91	245.36	246.46	247.32	248.01
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.42	19.43	19.44	19.45
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.71	8.69	8.67	8.66
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.87	5.84	5.82	5.80
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.64	4.60	4.58	4.56
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.96	3.92	3.90	3.87
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.53	3.49	3.47	3.44
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.24	3.20	3.17	3.15
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.03	2.99	2.96	2.94
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.86	2.83	2.80	2.77
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.74	2.70	2.67	2.65
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.64	2.60	2.57	2.54
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.55	2.51	2.48	2.46
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.48	2.44	2.41	2.39
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.42	2.38	2.35	2.33
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.37	2.33	2.30	2.28
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.33	2.29	2.26	2.23
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.29	2.25	2.22	2.19
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.26	2.21	2.18	2.16
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.22	2.18	2.15	2.12

٤٣ احصاء تطبيقي

Tanta University *Final Exam. (first term 2019-2020)*

Faculty of Commerce (English Section) *Applied Statistics*

Choose the Correct answer for the following Questions

In order to estimate the average time spent on the computer terminals per student at a local university, data were collected for a sample of 100 business students over a one-week period. Assume the population standard deviation is 1.8 hours. (answer Q1:Q3)

1. The standard error of the mean is: a. 7.50 b. 0.18 c. 2.00 d. 0.20

2. With a 0.95 probability, the margin of error is approximately a. 0.39 b. 1.96 c. 0.35 d. 1.64

3. If the sample mean is 9 hours, then the 95% confidence interval is

a. 8.65 to 9.35 hours b. 7.36 to 10.64 hours c. 7.80 to 10.20 hours d. 8.61 to 9.39 hours

*** A new brand of chocolate bar is being market tested. Five hundred of the new chocolate bars were given to consumers to try. The consumers were asked whether they liked or disliked the chocolate bar. You are given their responses below. (answer Q4:Q7)

Response	Liked	Disliked	Total
Frequency	300	200	500

4. The point estimate for the proportion of people who liked the chocolate bar is:

a. 0.75 b. 0.25 c. 0.60 d. 1.25

5. Construct a 95% confidence interval for the true proportion of people who liked the chocolate bar.

a. 0.56 to 0.64 b. 0.451 to 0.549 c. 0.451 to 0.79 d. 0.71 to 0.79

6. With a .95 probability, how large of a sample needs to be taken to provide a margin of error of 3% or less?

a. 1068 b. 801 c. 108 d. 1025

7. If a hypothesis is not rejected at the 5% level of significance, it

a. will also not be rejected at the 1% level b. will always be rejected at the 1% level

c. will sometimes be rejected at the 1% level d. None of these alternatives is correct.

*** If $n=25$ $\bar{x}=24.6$ s=10 $H_0:\mu=20$ vs $H_a:\mu>20$ (let $t_{0.05}(24)=2.064$) (answer Q8:Q9)

8. The test statistic is: a. -2.3 b. 0.38 c. 2.3 d. -0.38

9. If the test is done at 95% confidence, the null hypothesis should

a. not be rejected b. be rejected c. None of these alternatives is correct.

d. Not enough information is given to answer this question.

**Salary information regarding male and female employee of a large company is shown below.(answer Q 0:Q15)

	Sample Size	Sample Mean Salary (in $1,000)	Population Variance
Male	64	44	128
Female	36	41	72

21. Based on the results of question (20), if $H_0 : \mu_1 - \mu_2 = 0$ vs $H_a : \mu_1 - \mu_2 \neq 0.$

a. H_0 should be rejected b. H_0 should not be rejected c. H_a should not be rejected d. (a+b+c) are false

***An insurance company selected samples of clients under 18 years of age and over 18 and recorded the number of accidents they had in the previous year. The results are shown below.(answer Q22:Q26)

Under Age of 18	Over Age of 18
$n_1 = 500$	$n_2 = 500$
Number of accidents = 180	Number of accidents = 150

We are interested in determining if the accident proportions differ between the two age groups.

22. let P_u represent the proportion under and P_o the proportion over the age of 18. The null hypothesis is

a. $P_u - P_o \leq 0$ b. $P_u - P_o \geq 0$ c. $P_u - P_o \neq 0$ d. $P_u - P_o = 0$

23. The pooled proportion is: a. 0.305 b. 0.300 c. 0.33 d. 0.45

24. A 95% interval estimate for the difference between the two population proportions is:

a. 0.05 to 0.07 b. 0.002 to 0.118 c. 0.058 t0 0.118 d. -1.96 t0 1.96

25. The test statistic is: a. 2.02 b. 1.96 c. 2.96 d. 3.96

26. At 95% confidence, the conclusion is

a. H_0 should be rejected b. H_0 should not be rejected c. H_a should not be rejected d.(a+b+c) are false

***Two independent samples are drawn from two populations, and the following information is provided.

	n	$\bar{x}$	s
Population 1	34	55	14
Population 2	52	65	18

We want to test the following hypotheses. $H_o: \mu_1 - \mu_2 \geq 0$ $H_a: \mu_1 - \mu_2 < 0$ ($t_{0.05}(84) = 1.663$)

Assume the two populations are normally distributed and have equal variances. (answer Q27:Q29)

27. Determine the degrees of freedom. a. 86 b. 84 c. 81 d. 80

28. Compute the test statistic. a. t= 2.887 b. t=1.663 c. -2.887 d. -1.663

29. At 95% confidence, test the hypotheses.

a. H_0 should be rejected b. H_0 should not be rejected c. H_a should not be rejected d. (a+b+c) are true

30. In an analysis of variance problem if SSE = 120 and SSTR = 80, then SST is

a. 200 b. 40 c. 80 d. 120

Solve these two problems:

Q1:(8 points):

Five hundred randomly selected automobile owners were questioned on the main reason they had purchased their current automobile. The results are given below. (answer Q30: Q33) ($\chi^2_{0.05}(2) = 5.991$)

	Styling	Engineering	Fuel	Economy	Total
	Male	70	130	150	350
	Female	30	20	100	150
		100	150	250	500

1. State the null and alternative hypotheses for a contingency table test.
2. Calculate the χ^2 test statistic.
3. Give your conclusion for this test.

Q2: (15 points):

The following is part of an ANOVA table that was obtained from data regarding three treatments and a total of 15 observations. ($F_{0.05}(2,12) = 3.89$)

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F
Between Treatments	64			
Error (Within Treatments)	96			

1. State the null and alternative hypotheses for a contingency table test.
2. The number of degrees of freedom corresponding to between treatments is
3. The number of degrees of freedom corresponding to within treatments is
4. The mean square between treatments (MSTR) is
5. The computed test statistics is
6. The conclusion of the test is that the means

GOOD LUCK !!!