Practice Exam 1

STA 3024 Spring 2023

Class #: 16898 (Zheng)

Name: ____________________

UFID: ____________________

Instructions:

1. This examination contains 8 pages, including this page.
2. You have **50 minutes** to complete the exam.
3. The total score is 105. The extra 5 points serve as a buffer, so the highest score you can get is 100.
4. Write your answers clearly and legibly on the exam. Answers without sufficient work shown will not receive full credit.
5. You may use a scientific calculator. Do not share a calculator with anyone.
6. The last two pages contain a formula sheet and the z-table. You may not use any other resources including lecture notes, books, or other students.
7. Please sign the below Honor Code statement.

In recognition of the UF Student Honor Code, I certify that I will neither give nor receive unauthorized aid on this examination.

Signature: ____________________

1. (5 points) Which of the following statements CANNOT be an interpretation of the confidence interval $(-1.3, 4.4)$? Circle your choice and write the letter in the blank below.

   A. We are 95% confident that the true parameter is captured by the interval $(-1.3, 4.4)$.

   B. We are 95% confident that $(-1.3, 4.4)$ includes the true value of the parameter.

   C. The probability of the true parameter falling between -1.3 and 4.4 is 95%.

   D. The procedure used to compute the confidence interval requires that the probability of the confidence interval including the true parameter is 95%.

   E. None of the above.

1. ______________

2. (30 points) Are the following statements true or false? You do not need to give reasons.

   (a) ___ When making inference about one group mean, if we somehow know the original distribution is a normal distribution, we do not need to check if the sample size n is larger than or equal to 30.

   (b) ___ Small p-values support H_a and lead us to reject H_0 and determine the results are statistically significant.

   (c) ___ ANOVA determines if there is difference in population means for several groups by comparing variability between groups to variability within groups.

   (d) ___ In a ANOVA test the null hypothesis is $H_0 : \mu_1 = \mu_2 = \mu_3$. Then the alternative hypothesis can be $H_a : \mu_1 \neq \mu_2 \neq \mu_3$.

   (e) ___ In one-way ANOVA, we find a significant difference if the test statistic is very large.

   (f) ___ Tukey's Honest Significant Difference has family confidence 95% and individual confidence smaller than 95%.

   (g) ___ Fisher's Least Significance Difference has individual confidence 95%. When there are three simultaneous comparisons, it has family confidence larger than 85%.

   (h) ___ Suppose the group means are $35, 27, 41, 33, 40$, respectively. If the margin of error is 5.1, we will have four pairwise comparisons that are NOT significant.

   (i) ___ In two-way ANOVA, we always test for interaction first because it is easier than testing main effects.

   (j) ___ In two-way ANOVA, it is okay that the response variable only takes finite values.

3. A food factory uses an automatic canning machine to pack canned food. The standard weight of each can is 500 grams. It is necessary to check the working condition of the machine every day. 10 cans were sampled today, and their weights were measured and the summary statistics were $\bar{X} = 502$ and $S = 6.5$. Suppose the can weight manufactured by the automatic canning machine follows a normal distribution with mean μ and stand deviation σ (both are unknown). We want to see if the machine works normally.

   (a) (5 points) Write down the null and alternative hypotheses with respect to μ.

   (b) (10 points) Find a 95% confidence interval for the μ. (Here are some values you might find useful: $t_{9,0.05} = 1.833$, $t_{9,0.025} = 2.262$, $t_{10,0.05} = 1.812$, $t_{10,0.025} = 2.228$)

   (c) (5 points) Suppose the confidence interval found in part (b) is (x, y). Interpret it.

   (d) (5 points) Based on the confidence interval found in part (b), do you think the machine works normally? Why?

4. The market research department of a chain of hamburger restaurants wants to compare the mean monthly sales of hamburgers under three different marketing strategies. It randomly assigns 15 restaurants to the three groups, five per group. The sample means for the three groups were 17.864, 15, and 12.136.

(a) (5 points) State one of the assumptions we need for conducting ANOVA test and how we can check that assumption.

(b) (10 points) Fill out the ANOVA table below. Write down numbers in the blank below the table. You do not need to show your calculation process.

Source	df	SS	MS	F	p-value
Group	(i)	(iv)	41.00	1.02	0.389
Error	(ii)	(v)	(vii)		
Total	(iii)	(vi)			

Table 1: ANOVA table

(i)__________ (ii)__________ (iii)__________ (iv)__________ (v)__________
(vi)__________ (vii)__________

(c) (5 points) Which of the following quantities is equal to the p-value in the ANOVA table? Circle your choice and write the letter in the blank below.

A. $F_{2,12,0.05}$
B. $F_{2,12,0.389}$
C. $F_{2,14,0.05}$
D. $F_{2,14,0.389}$
E. $P(F \geq 1.02)$ where F follows F distribution with degrees of freedom 2 and 12
F. $P(F \leq 1.02)$ where F follows F distribution with degrees of freedom 2 and 12
G. $P(F \geq 1.02)$ where F follows F distribution with degrees of freedom 2 and 14
H. $P(F \leq 1.02)$ where F follows F distribution with degrees of freedom 2 and 14

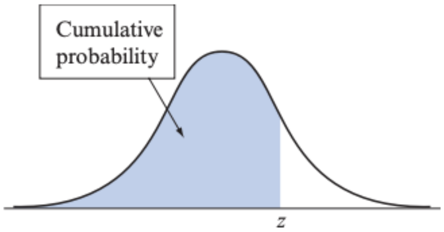
(c) ____________

(d) (5 points) What conclusion can we draw from the ANOVA table?

(e) (10 points) Compute the margin of error for comparing all mean monthly sales under three marketing strategies using Bonferroni method with 97% family confidence. (Here are some values you might find useful: $t_{12,0.005} = 3.055$, $t_{14,0.005} = 2.977$, $t_{2,0.005} = 9.925$, $t_{12,0.01} = 2.681$, $t_{14,0.005} = 2.624$, $t_{2,0.01} = 6.965$)

(f) (5 points) Using the margin of error in part (e), do we find any pair(s) of means that are significantly different? If yes, which pair(s)? If no, explain how you come to such conclusion.

5. (5 points) In two-way ANOVA, factor A has 4 levels (A_1, A_2, A_3 and A_4) and factor B has 3 levels (B_1, B_2 and B_3). Draw an interaction plot that leads us to expect NO interactions between factor A and B, but significant difference due to both factor A and factor B.



Cumulative probability for z is the area under the standard normal curve to the left of z

Table A **Standard Normal Cumulative Probabilities (*continued*)**

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Case	Parameter	Estimator	Standard Error	Sampling Distribution
one mean	μ	$\bar{x}$	$s/\sqrt{n}$	t_{n-1}
mean of matched pairs difference	μ_d	$\bar{x}_d$	$s_d/\sqrt{n}$	t_{n-1}
difference of two independent means	$\mu_1 - \mu_2$	$\bar{x}_1 - \bar{x}_2$	$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	t with df between: $\min(n_1 - 1, n_2 - 1)$ and $n_1 + n_2 - 2$
one proportion	p	$\hat{p}$	CI: $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ ST: $\sqrt{\frac{p_0(1-p_0)}{n}}$	z
difference of two independent proportions	$p_1 - p_2$	$\hat{p}_1 - \hat{p}_2$	CI: $\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ ST: $\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}$	z

$$SST = \sum_{i=1}^{g} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{\cdot\cdot})^2$$

$$SSG = \sum_{i=1}^{g} n_i (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2$$

$$SSE = \sum_{i=1}^{g} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i\cdot})^2$$