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 a conclusion in significance tests? Circle your choice and write the letter in the blank below.

   A. We have strong evidence that the alternative hypothesis is true.

   B. We do NOT have strong evidence that the alternative hypothesis is true.

   C. We have strong evidence that the null hypothesis is wrong.

   D. We do NOT have strong evidence that the null hypothesis is wrong.

   E. None of the above.

1. ______________

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

   (a) ___ A random and representative sample of size 100 can be used to make inferences about the population.

   (b) ___ We roll a fair die n times and each time we record the number facing up. As n gets larger, the distribution of the number facing up in one roll will get closer to a normal distribution.

   (c) ___ In one-way ANOVA, the predictor is quantitative but the response is categorical.

   (d) ___ In a one-way ANOVA table, the degrees of freedom for Group must be smaller than the degrees of freedom for Error, as long as each treatment has at least two replications.

   (e) ___ In one-way ANOVA, the test statistic is a number that represents how many times larger the variability between groups is than the variability within groups.

   (f) ___ In one-way ANOVA, the test statistic is a number we find in the F-table, using the correct degrees of freedom for the numerator and denominator.

   (g) ___ For $g = 6$ groups, we should use individual confidence level 99% to get Bonferroni family (simultaneous) confidence level 94%.

   (h) ___ When doing multiple comparisons, both Tukey's and Fisher's methods have the same family confidence level 95%, but different individual confidence levels.

   (i) ___ One-way ANOVA contains one factor with different levels while two-way ANOVA contains two factors with different combinations of levels.

   (j) ___ In two-way ANOVA, we always test for interaction first because if interaction effect is NOT significant, then we do NOT even look at main effects.

3. Past experience indicates that 60% of students in Riverside High School like playing video games. We take a random sample of 200 students from Riverside High school and record how many of them like playing video games.

   (a) (10 points) Based on the past experience, what is the probability that more than 70% of the students in the random sample like playing video games?

   (b) (10 points) It is recorded that 45% of the students from the sample like playing video games. Find a 95% confidence interval for the proportion of students in Riverside High School who like playing video games.

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

   (d) (5 points) Is the true value of the proportion (60%) contained in your confidence interval? If not, what might have happened?

4. An experimenter randomly allocated 125 male turkeys to five treatment groups: control and treatment A, B, C and D. There were 25 birds in each group, and the mean results were 2.182, 2.777, 2.946, 2.858, and 2.934, respectively. We want to test if one or more of the treatments differs from the control.

   (a) (5 points) Denote the group means by μ_0, μ_A, μ_B, μ_C and μ_D, respectively. State the null hypothesis and the alternative hypothesis for ANOVA test.

   (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)	2.54	(vii)	0.015
Error	(ii)	(v)	0.79		
Total	(iii)	(vi)			

Table 1: ANOVA table

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

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

   (d) (5 points) Which of the following computational formulae is correct for finding the margin of error in multiple comparisons using Bonferroni method with 95% family confidence? Circle your choice and write the letter in the blank below.

      A. $t_{0.0025,120} \cdot 1.594 \cdot \sqrt{\frac{1}{25} + \frac{1}{25}}$
      B. $t_{0.005,120} \cdot 1.594 \cdot \sqrt{\frac{1}{25} + \frac{1}{25}}$
      C. $t_{0.0025,120} \cdot 0.889 \cdot \sqrt{\frac{1}{25} + \frac{1}{25}}$
      D. $t_{0.005,120} \cdot 0.889 \cdot \sqrt{\frac{1}{25} + \frac{1}{25}}$
      E. None of the above.

(d) ______________

(e) (5 points) In multiple comparisons, there is a method called Scheffe's method. The experimenter employs Scheffe's method and finds the margin of error is 0.78. Using this margin of error, 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.

(f) (5 points) Figure 1 displays Tukey simultaneous 95% confidence intervals for multiple comparisons. Which treatments differ from the control?

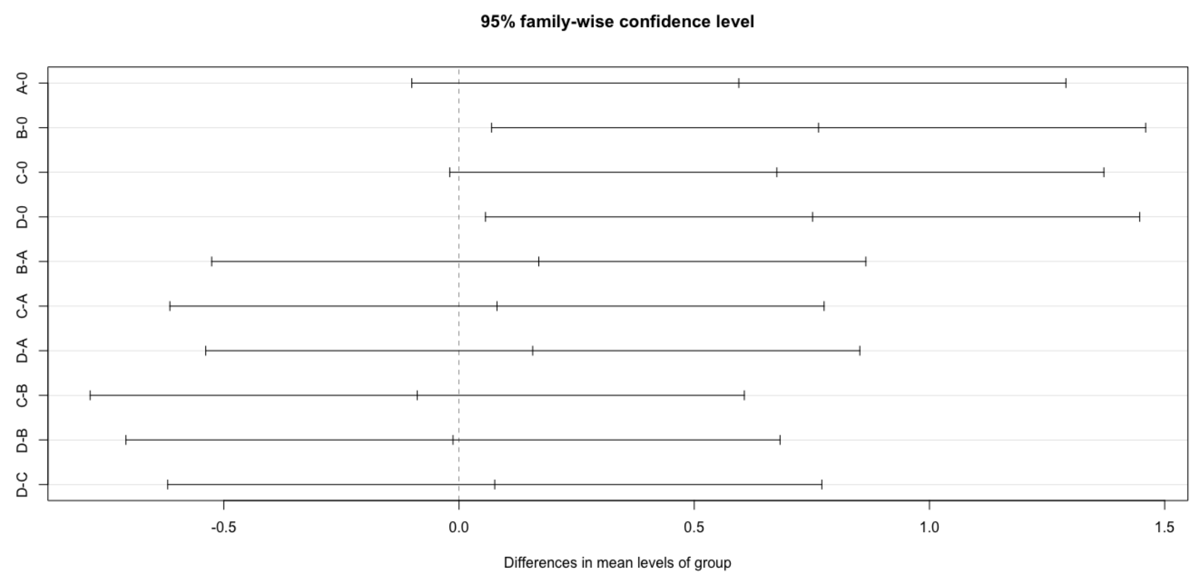
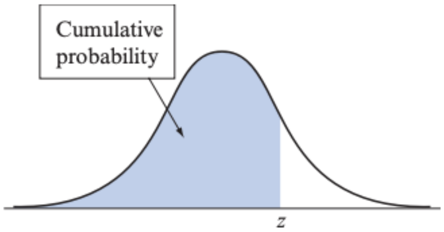


Figure 1: Tukey Simultaneous 95% CIs ("0" denotes Control)

5. (5 points) In two-way ANOVA, factor A has 3 levels (A_1, A_2 and A_3) and factor B has 2 levels (B_1 and B_2). Draw an interaction plot that leads us to expect NO interactions between factor A and B, significant difference due to factor A but NO significant difference due to 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

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

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

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