Answer Key Points Available: 150

Show your work wherever some is required. Do not assume I will know what you did.
You may use three 3"x5" cards, Tables III and IV, and a graphing calculator where appropriate.

1. Which of our main topics this semester was not a branch of statistics? (No explanation needed.)

   Probability

2. On a survey, a respondent gets distracted and enters their age as 35 instead of 37. What is our term for this kind of error? (No explanation needed.)

   Non-sampling

3. The following data are the annual total nursing salaries ($ hundreds) in the subset of 34 nursing homes in New Mexico which are located in rural areas. ("*A Comparison of Financial Performance, Organizational Characteristics, and Management Strategy Among Rural and Urban Nursing Facilities*", Journal of Rural Health, Winter 1992, pp 27-40.) Construct a 5 category frequency table for the data. (Note the data are sorted!)

1288	1498	1914	1925	1988	1995	2008	2051	2066	2088	2143	2245	2459	2782	2784	2810	3025
3224	3547	3622	3672	3720	3866	3995	4029	4156	4166	4173	4406	4729	5173	5257	5933	6059

   1. 5 Categories
   2. range= $6059 - 1288 = 4771$
   3. $\frac{4771}{5} = 954.2 \to 954 \to 955$
   4. $1288 - .5 = 1287.5$

Categories	Frequency	Relative Freq.	Cumulative Freq.	CRF
1287.5 – 2242.5	11	0.324	11	0.324
2242.5 – 3197.5	6	0.176	17	0.500
3197.5 – 4152.5	8	0.235	25	0.735
4152.5 – 5107.5	5	0.147	30	0.882
5107.5 – 6062.5	4	0.118	34	1.000

4. As a batch of parts comes off an assembly line, a random part is selected for quality control inspection. After that, every tenth part is selected until the desired total of 50 parts is achieved. What type of sampling is this? (No explanation needed.)

Systematic

5. A card trick us performed using a special deck with cards numbered 1 through 50. One card is drawn at random, and laid aside. A volunteer from the audience then draws a second card. Is the number on the second card independent of the number on the first card? Explain in **one sentence**.

No, since the card chosen second cannot be the same as the card laid aside.

6. A large online role playing game is played by millions of players worldwide. The game developers are interested in the choice of players to play on the "good" or "evil" side. They plan to gather 15,000 players from the US (where there are almost 200,000 total players) at random and record which side they chose. They want to see if the proportion of all US players of this game that would choose the good side is different from 0.5. Answer only the following questions about this setting.

   (a) Construct the hypotheses step of the appropriate test for this setting.

   H_0: $p = 0.5$
   H_a: $p \neq 0.5$
   Where: p is the proportion of all US players of this game who would choose the good side.

   (b) What would the conditions step look like in this problem?

   We have independent, random observations from a binomial experiment and there are enough trials to use the Central Limit Theorem. (reasonable)

   (c) Suppose that the P-value in this setting was 0.0638 based on a negative test statistic. What conclusion would you make at the 0.05 level? (You do not need to verify this P-value.)

   We do not have enough evidence ($P = 0.0638$) to suggest that the proportion of all US players of this game who would choose the good side is not 0.5.

7. A study in the April 2016 issue of the Journal of Developmental and Behavioral Pediatrics considered the effect on children of being raised in a household in which the parents were the same gender. One variable measured was the emotional difficulties of the children (all 6-17 years old). Among 95 children included in the study (all with same-gender parents), the average score on the emotional difficulty inventory was 1.97 pts. on a 5 point scale with a standard deviation of 0.73 pts. On the scale used in the paper, higher scores indicate more emotional difficulties. In the same paper, they found that for 95 children of parents with different genders, the average score was 1.93 pts out of 5 with a standard deviation of 0.69. Does this data provide enough evidence, at the 0.01 level, to conclude that the average emotional difficulty rating for all children of same gender parents is not the same as the average rating for children of parents with different genders?

(Link: http://journals.lww.com/jrnldbp/Fulltext/2016/04000/Same_Sex_and_Different_Sex_Parent_Households_and.1.aspx)

1) **Hypotheses**:

H_0: $\mu_1 = \mu_2$

H_a: $\mu_1 \neq \mu_2$

Where: μ_1 is the mean score on the inventory for all children (6-17) with same-sex parents and μ_2 is the mean score on the inventory for all children (6-17) of parents with different genders.

2) **Conditions**:We have independent, random observations from 2 populations and each sample has enough observations that we can use the Central Limit Theorem. (random?)

3) **Rejection Region**:Reject H_0 if $TS > 2.576$ or $TS < -2.576$ since $\frac{1}{2}\alpha = 0.005$

4) **Test Statistic**:

$$TS = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{n}}} = \frac{1.97 - 1.93}{\sqrt{\frac{0.73^2}{95} + \frac{0.69^2}{95}}} = 0.388129 \to 0.39$$

5) **P-value**:$P = 2 \cdot P(Z > 0.39) = 2 \cdot (0.5 - 0.1517) = 2 \cdot 0.3483 = 0.6966$

6) **Conclusion**:We do not have enough evidence ($P = 0.6966$) to suggest that the children of same-gender parents and children with parents of different genders differ with respect to their mean score on the emotional difficulty inventory.

8. Administrators of a large school district are attempting to study the effect that a new magnet school is having on students. To get a baseline for comparison, 12 students who applied for the magnet school and were not selected in the random lottery are tested to see the increase in their test scores over the course of the year in their original school. For these 12 students, the mean increase in standardized science test score was 3.000 points with a standard deviation of 3.977 points. Construct a 95% confidence interval for the mean increase in points on the test for all students not selected in the magnet school lottery.

1) **Parameter**: μ is the mean test score improvement for all students not selected in the magnet school lottery

2) **Conditions**: We have independent, random observations from a normally distributed population with an unknown population variance.

3) **Point Estimate**: $\bar{X} = 3.000$ pts.

4) **Margin of Error**:

$$t_{11} \cdot \frac{3.977}{\sqrt{n}} = 2.201 \cdot \frac{3.977}{\sqrt{12}} = 2.52688$$

5) **Confidence Interval**:

$$3.000 \pm 2.52688 = (0.47312, 5.52688)$$

6) **Conclusion**: We are 95% confident that the mean test score improvement for all students not selected in the magnet school lottery is between 0.473 and 5.527 pts.

9. Would you feel confident saying that the mean improvement was positive, based on your interval? (No explanation needed.)
Yes

10. Suppose we wanted to conduct a test to see if we have sufficient evidence to conclude that there is a difference between the mean fat content of sandwiches at Burger King and the mean fat content of sandwiches at McDonalds.

    (a) What would a Type I error be? (No more than one sentence.)

    The fat content does not differ, but we believe that it does.

    (b) What would a Type II error be? (No more than one sentence.)

    The fat content differs, but we cannot conclude this.

11. Suppose that the grades at a particular large public high school are normally distributed with an average of 86.4% and a standard deviation of 8.2%.

    (a) What is the probability that a random student would have a grade less than 90%?

    $$Z = \frac{90 - 86.4}{8.2} = 0.44 \rightarrow 0.1700 \rightarrow P(X < 90) = 0.5 + 0.1700 = 0.6700$$

    (b) If the bottom 20% of students are recommended for summer school, what would the cutoff be for the recommendation?

    $$0.5 - 0.2 = 0.3 \rightarrow Z = -.84 = \frac{X - 86.4}{8.2} \rightarrow \bar{X} = 79.5\%$$

    (c) What is the probability that a random group of 20 students would have an average grade of less than 85%?

    $$Z = \frac{85 - 86.4}{\frac{8.2}{\sqrt{20}}} = -0.76 \rightarrow 0.2764 \rightarrow P(\bar{X} < 85) = 0.5 - 0.2764 = 0.2236$$