

Review Day 4

Have your pink Notation Quiz (Practice) out and available

Explain what each of the following represent, and when they would be used.

1. \hat{p} Sample proportion (proportion predicted by sample)
2. p Population proportion (true)
3. p_0 Parameter used in null hypotheses
4. p -value Probability of obtaining a sample statistic at least as extreme as the observed statistic

5. z standard normal value

6. z^* upper p critical value

7. $-z^*$ lower p -critical value

8. H_0 Null Hypotheses

9. H_a Alternative hypotheses

10. n sample size

11. N Population size

12. $n\hat{p}$ Number of successes observed in sample
(counts)

13. $n\hat{q}$ Number of failures observed in sample
(counts)

14. np_0 Expected # of successes (counts) in sample

15. nq_0 Expected # of failures in sample
(counts)

16. $\sqrt{\frac{\hat{p}\hat{q}}{n}}$ Std. error of sample proportion

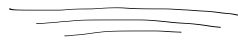
17. $\sqrt{\frac{p_0q_0}{n}}$ Std. deviation of parameter, P ,
(in one-sample z test)

18. $z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$ MOE (in CI for p)

19. $\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$ CI for one-sample z interval

20. $\frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$ Z statistic
for one-sample z test

Turn in what
you have



Partner LCC
(by last name)

2. Check your Review of Inference

- 12 MCQ
- 2 FRQ

HW lottery

3. Survey to establish
April Review Options

- for those taking AP Exam

4. Name that Test - Set B

- Do all 10
- Collaborate to finalize

Set B

Name that Test! (or Interval)

1. During a flu vaccine shortage in the United States, it was believed that 45 percent of vaccine-eligible people received flu vaccine. The results of a survey given to a random sample of 2,350 vaccine-eligible people indicated that 978 of the 2,350 people had received flu vaccine.
 - (a) Construct a 99 percent confidence interval for the proportion of vaccine-eligible people who had received flu vaccine. Use your confidence interval to comment on the belief that 45 percent of the vaccine-eligible people had received flu vaccine.

One-sample z interval for p

2. Product advertisers studied the effects of television ads on children's choices for two new snacks. The advertisers used two 30-second television ads in an experiment. One ad was for a new sugary snack called Choco-Zuties, and the other ad was for a new healthy snack called Apple-Zuties.

For the experiment, 75 children were randomly assigned to one of three groups, A, B, or C. Each child individually watched a 30-minute television program that was interrupted for 5 minutes of advertising. The advertising was the same for each group with the following exceptions.

- The advertising for group A included the Choco-Zuties ad but not the Apple-Zuties ad.
- The advertising for group B included the Apple-Zuties ad but not the Choco-Zuties ad.
- The advertising for group C included neither the Choco-Zuties ad nor the Apple-Zuties ad.

After the program, the children were offered a choice between the two snacks. The table below summarizes their choices.

Group	Type of Ad	Number Who Chose Choco-Zuties	Number Who Chose Apple-Zuties
A	Choco-Zuties only	21	4
B	Apple-Zuties only	13	12
C	Neither	22	3

- (a) Do the data provide convincing statistical evidence that there is an association between type of ad and children's choice of snack among all children similar to those who participated in the experiment?

Chi-square test of independence

3. High cholesterol levels in people can be reduced by exercise, diet, and medication. Twenty middle-aged males with cholesterol readings between 220 and 240 milligrams per deciliter (mg/dL) of blood were randomly selected from the population of such male patients at a large local hospital. Ten of the 20 males were randomly assigned to group A, advised on appropriate exercise and diet, and also received a placebo. The other 10 males were assigned to group B, received the same advice on appropriate exercise and diet, but received a drug intended to reduce cholesterol instead of a placebo. After three months, posttreatment cholesterol readings were taken for all 20 males and compared to pretreatment cholesterol readings. The tables below give the reduction in cholesterol level (pretreatment reading minus posttreatment reading) for each male in the study.

Group A (placebo)

Reduction (in mg/dL)	2	19	8	4	12	8	17	7	24	1
----------------------	---	----	---	---	----	---	----	---	----	---

Mean Reduction: 10.20 Standard Deviation of Reductions: 7.66

Group B (cholesterol drug)

Reduction (in mg/dL)	30	19	18	17	20	-4	23	10	9	22
----------------------	----	----	----	----	----	----	----	----	---	----

Mean Reduction: 16.40 Standard Deviation of Reductions: 9.40

Do the data provide convincing evidence, at the $\alpha = 0.01$ level, that the cholesterol drug is effective in producing a reduction in mean cholesterol level beyond that produced by exercise and diet?

Two-sample t test for $\mu_1 - \mu_2$

4. A French study was conducted in the 1990s to compare the effectiveness of using an instrument called a cardiopump with the effectiveness of using traditional cardiopulmonary resuscitation (CPR) in saving lives of heart attack victims. Heart attack patients in participating cities were treated with either a cardiopump or CPR, depending on whether the individual's heart attack occurred on an even-numbered or an odd-numbered day of the month. Before the start of the study, a coin was tossed to determine which treatment, a cardiopump or CPR, was given on the even-numbered days. The other treatment was given on the odd-numbered days. In total, 754 patients were treated with a cardiopump, and 37 survived at least one year; while 746 patients were treated with CPR, and 15 survived at least one year.
- The conditions for inference are satisfied in the study. State the conditions and indicate how they are satisfied.
 - Perform a statistical test to determine whether the survival rate for patients treated with a cardiopump is significantly higher than the survival rate for patients treated with CPR.

two-sample z test for $p_1 - p_2$

5. A bottle-filling machine is set to dispense 12.1 fluid ounces into juice bottles. To ensure that the machine is filling accurately, every hour a worker randomly selects four bottles filled by the machine during the past hour and measures the contents. If there is convincing evidence that the mean amount of juice dispensed is different from 12.1 ounces or if there is convincing evidence that the standard deviation is greater than 0.05 ounce, the machine is shut down for recalibration. It can be assumed that the amount of juice that is dispensed into bottles is normally distributed.

During one hour, the mean number of fluid ounces of four randomly selected bottles was 12.05 and the standard deviation was 0.085 ounce.

- Perform a test of significance to determine whether the mean amount of juice dispensed is different from 12.1 fluid ounces. Assume the conditions for inference are met.

One-sample t test for μ

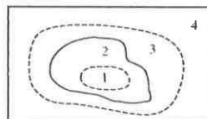
6. One of the two fire stations in a certain town responds to calls in the northern half of the town, and the other fire station responds to calls in the southern half of the town. One of the town council members believes that the two fire stations have different mean response times. Response time is measured by the difference between the time an emergency call comes into the fire station and the time the first fire truck arrives at the scene of the fire. Data were collected to investigate whether the council member's belief is correct. A random sample of 50 calls selected from the northern fire station had a mean response time of 4.3 minutes with a standard deviation of 3.7 minutes. A random sample of 50 calls selected from the southern fire station had a mean response time of 5.3 minutes with a standard deviation of 3.2 minutes.
- (a) Construct and interpret a 95 percent confidence interval for the difference in mean response times between the two fire stations.
- (b) Does the confidence interval in part (a) support the council member's belief that the two fire stations have different mean response times? Explain.

two-sample t interval for $\mu_1 - \mu_2$

7. A study was conducted to determine where moose are found in a region containing a large burned area. A map of the study area was partitioned into the following four habitat types.

- (1) Inside the burned area, not near the edge of the burned area,
- (2) Inside the burned area, near the edge,
- (3) Outside the burned area, near the edge, and
- (4) Outside the burned area, not near the edge.

The figure below shows these four habitat types.



Note: Figure not drawn to scale.

The proportion of total acreage in each of the habitat types was determined for the study area. Using an aerial survey, moose locations were observed and classified into one of the four habitat types. The results are given in the table below.

The proportion of total acreage in each of the habitat types was determined for the study area. Using an aerial survey, moose locations were observed and classified into one of the four habitat types. The results are given in the table below.

Habitat Type	Proportion of Total Acreage	Number of Moose Observed
1	0.340	25
2	0.101	22
3	0.104	30
4	0.455	40
Total	1.000	117

- (a) The researchers who are conducting the study expect the number of moose observed in a habitat type to be proportional to the amount of acreage of that type of habitat. Are the data consistent with this expectation? Conduct an appropriate statistical test to support your conclusion. Assume the conditions for inference are met.
- (b) Relative to the proportion of total acreage, which habitat types did the moose seem to prefer? Explain.

Chi-Square test for goodness of fit

8. Investigators at the U.S. Department of Agriculture wished to compare methods of determining the level of *E. coli* bacteria contamination in beef. Two different methods (A and B) of determining the level of contamination were used on each of ten randomly selected specimens of a certain type of beef. The data obtained, in millimicrobes/liter of ground beef, for each of the methods are shown in the table below.

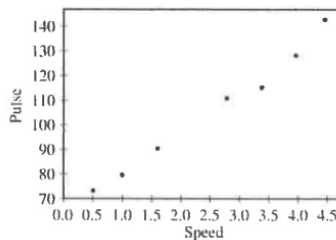
		Specimen									
		1	2	3	4	5	6	7	8	9	10
Method	A	22.7	23.6	24.0	27.1	27.4	27.8	34.4	35.2	40.4	46.8
	B	23.0	23.1	23.7	26.5	26.6	27.1	33.2	35.0	40.5	47.8

Is there a significant difference in the mean amount of *E. coli* bacteria detected by the two methods for this type of beef? Provide a statistical justification to support your answer.

Paired t test for

μ diff

9. John believes that as he increases his walking speed, his pulse rate will increase. He wants to model this relationship. John records his pulse rate, in beats per minute (bpm), while walking at each of seven different speeds, in miles per hour (mph). A scatterplot and regression output are shown below.



Regression Analysis: Pulse Versus Speed					
Predictor	Coef	SE Coef	T	P	
Constant	63.457	2.387	26.58	0.000	
Speed	16.2809	0.8192	19.88	0.000	
S = 3.087		R-Sq = 98.7%		R-Sq (adj) = 98.5%	
Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	1	3763.2	3763.2	396.13	0.000
Residual	5	47.6	9.5		
Total	6	3810.9			

- (c) John wants to provide a 98 percent confidence interval for the slope parameter in his final report. Compute the margin of error that John should use. Assume that conditions for inference are satisfied.

t interval for slope

10.

A researcher believes that treating seeds with certain additives before planting can enhance the growth of plants. An experiment to investigate this is conducted in a greenhouse. From a large number of Roma tomato seeds, 24 seeds are randomly chosen and 2 are assigned to each of 12 containers. One of the 2 seeds is randomly selected and treated with the additive. The other seed serves as a control. Both seeds are then planted in the same container. The growth, in centimeters, of each of the 24 plants is measured after 30 days. These data were used to generate the partial computer output shown below. Graphical displays indicate that the assumption of normality is not unreasonable.

	N	Mean	StDev	SE Mean
Control	12	15.989	1.098	0.317
Treatment	12	18.004	1.175	0.339
Difference	12	-2.015	1.163	0.336

- (a) Construct a confidence interval for the mean difference in growth, in centimeters, of the plants from the untreated and treated seeds. Be sure to interpret this interval.
- (b) Based only on the confidence interval in part (a), is there sufficient evidence to conclude that there is a significant mean difference in growth of the plants from untreated seeds and the plants from treated seeds? Justify your conclusion.

Paired t test for μ_{diff}

Assignment:

Musical Chairs

- You will be randomly assigned two of the 10 questions to do.
- They will be collected. (No lottery) tomorrow.
- You will get feedback on each via a Musical Chair rotation.

Your first question will be done on a pre-determined color of paper

Your 2nd question should be done on your own white paper.

Groups of 5

blue

green

pink

yellow

cream

Set B

blue	#1
green	#3
pink	#7
yellow	#6
cream	#10

For your second question (done on white paper), choose any of the other four