Math 119 Name: ____________________________

FINAL

(LESSONS 33-43: HYPOTHESIS TESTING; CORRELATION AND REGRESSION)
MATH 119 – FALL 2019 – KUNIYUKI
120 POINTS TOTAL

No notes or books allowed. A scientific calculator is allowed. Simplify as appropriate.
You do not have to reduce fractions. For example, 10/20 does not have to be rewritten as ½.

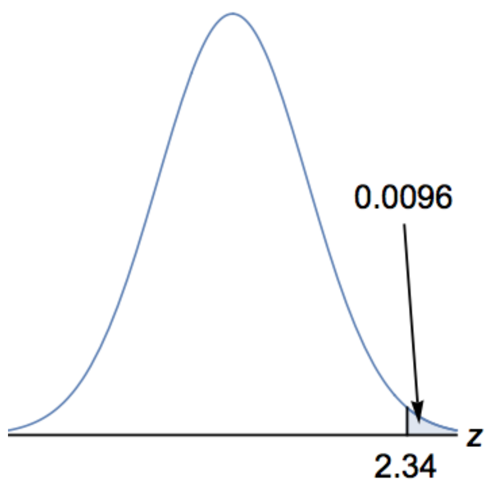
THE FORMULA SHEETS ARE ON A SEPARATE HANDOUT.

If you round off in the middle, round off to at least five significant digits.

1) (28 points). A particular governor wants to know if a majority of likely voters approve of the governor. (A majority means more than 50%.) Test the claim that a majority of likely voters approve of the governor at the 0.05 significance level. Warning: Write percents as decimals.

Let p = the proportion of likely voters who approve of the governor.

Use these hints about the z distribution:



• a) Write the **setup** for the hypothesis test. The setup will include H_0, H_1, identifying which is the claim, and the significance level. (5 points)

• b) Is this test **two-tailed, right-tailed, or left-tailed**? (2 points)

We gather **sample data**. We randomly select 1000 likely voters in a poll.
537 of them approve of the governor.

- c) Find the **sample proportion** $\hat{p}$. (3 points)

- d) **Verify** that normal approximations are appropriate in this problem. (2 points)

- e) Compute the z **test statistic** for our sample and round it off to two decimal places. Show all work, as in class! (7 points)

- f) Find the corresponding ***P*-value**. (2 points)

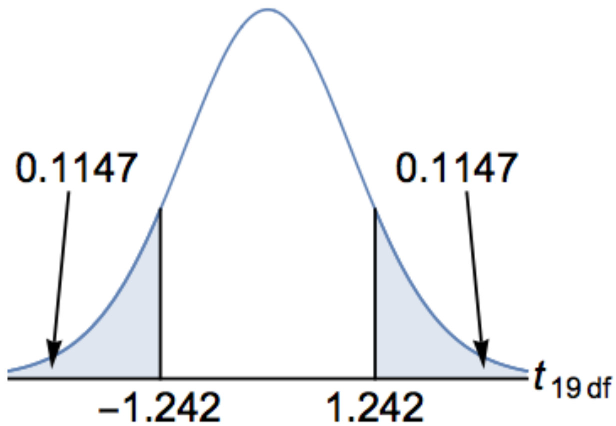
- g) **Decide** whether or not to reject H_0. (2 points)

- h) Write our **conclusion** relative to the claim. (5 points)

2) (23 points). A pharmaceutical company makes a pill that is supposed to be 500 micrograms (mcg) by mass. The company claims that the average mass of its pills is 500 mcg. Test this claim at the 0.10 significance level. Assume that the pills (by mass) are approximately normally distributed.

Let μ = the mean mass of pills produced by the company.

Use these hints about the *t* distribution on 19 degrees of freedom (df):



• a) Write the **setup** for the hypothesis test. The setup will include H_0, H_1, identifying which is the claim, and the significance level. (5 points)

• b) Is this test **two-tailed, right-tailed, or left-tailed**? (2 points)

We gather **sample data**. We randomly select 20 pills produced by the company. The sample mean pill mass is 495 mcg and the sample standard deviation (SD) is 18 mcg.

• c) Compute the *t* **test statistic** for our sample and round it off to three decimal places. Show all work, as in class! (6 points)

• d) Find the corresponding ***P*-value**. (3 points)

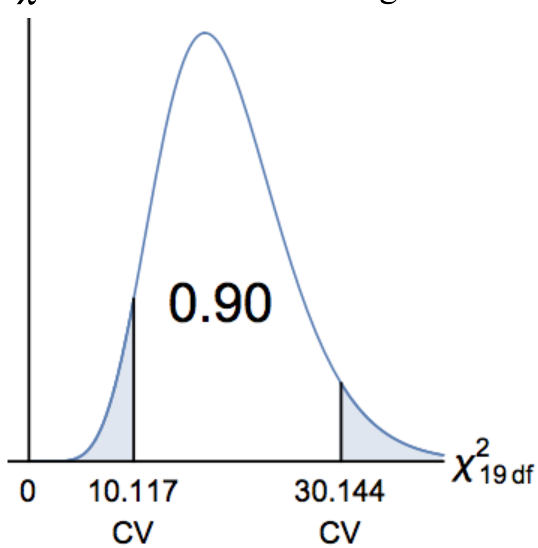
• e) **Decide** whether or not to reject H_0. (2 points)

• f) Write our **conclusion** relative to the claim. (5 points)

3) (19 points). An article on the pharmaceutical company from 2) claims that the population standard deviation (SD) of the company's pills (by mass) is 25 mcg. Test this claim at the 0.10 significance level. Assume that the pills (by mass) are approximately normally distributed. Use **the traditional (classical) method** of hypothesis testing.

Let σ = the standard deviation (SD) of the company's pills (by mass).

Use these hints about the χ^2 distribution on 19 degrees of freedom (df):



• a) Write the **setup** for the hypothesis test. The setup will include H_0, H_1, identifying which is the claim, and the significance level. (5 points)

• b) Is this test **two-tailed, right-tailed, or left-tailed**? (2 points)

We gather **sample data**. As in 2), we randomly select 20 of the company's pills. The sample standard deviation (SD) is 12.5 mcg.

• c) Compute the χ^2 **test statistic** for our sample and round it off to three decimal places. Show all work, as in class! (5 points)

• d) **Decide** whether or not to reject H_0. (2 points)

• e) Write our **conclusion** relative to the claim. (5 points)

4) (2 points). We reject a H_0 that is true. What type of error have we made? Box in one:

Type I error　　　　Type II error

5) (21 points). According to a 2017 Gallup Poll, 51% of registered voters in California are Democrats, 30% are Republicans, and 19% are neither. Let's assume that these proportions are correct. The student newspaper at a college wants to test the claim that the distribution of party preferences among registered voters at their college is the same as for California. Test the newspaper's claim at the 0.05 significance level.

• a) Write the **setup** for the hypothesis test. The setup will include H_0, H_1, identifying which is the claim, and the significance level. (5 points)

Let p_D = the proportion of registered voters at the college who are Democrats.

Let p_R = the proportion of registered voters at the college who are Republicans.

Let p_N = the proportion of registered voters at the college who are neither.

• b) Is this test **two-tailed, right-tailed, or left-tailed**? (2 points)

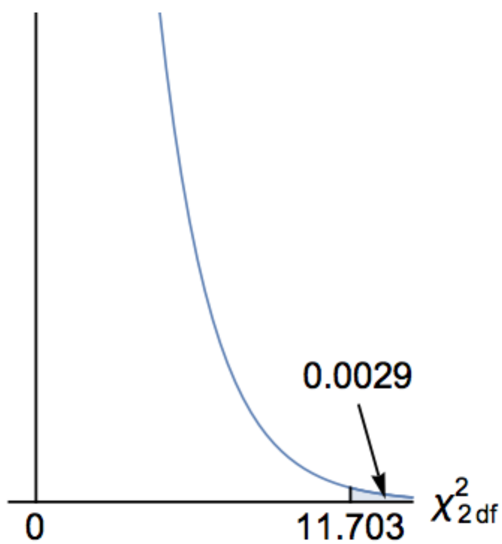
We gather **sample data**. We randomly sample 200 registered voters at the college. Among those voters, 120 are Democrats, 60 are Republicans, and 20 are neither.

• c) Write the **Observed (*O*) Table** and the **Expected (*E*) Table**. Note that each of the *E* values is at least 5, so we may apply the methods of this Lesson. (7 points)

Party preference	**Observed (*O*) Frequencies**	**Expected (*E*) Frequencies**
Democrats		
Republicans		
Neither		

• d) The χ^2 **test statistic** is about 11.703. **Decide** whether or not to reject H_0. (2 points)

Use these hints about the χ^2 distribution on 2 degrees of freedom (df):



• e) Write our **conclusion** relative to the claim. (5 points)

6) (14 points). An article claims that student participation in an exam preparation program is independent of whether or not students pass the exam. Test the article's claim at the 0.05 significance level.

- a) Write the **setup** for the hypothesis test. The setup will include H_0, H_1, identifying which is the claim, and the significance level. (5 points)

- b) Is this test **two-tailed, right-tailed, or left-tailed**? (2 points)

We gather **sample data**, summarized in the two-way **Observed (*O*) Table** below:

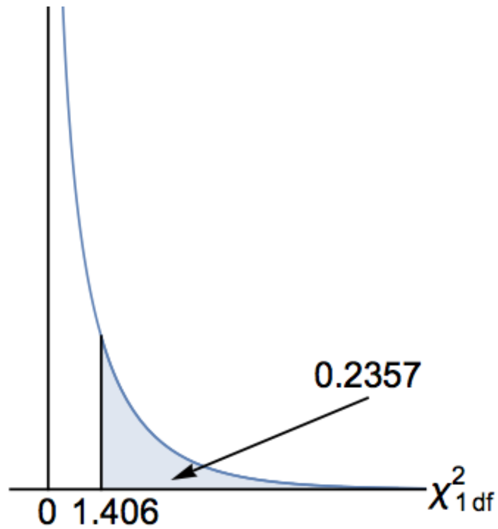
	Pass the exam	Fail the exam	← **Exam result**
Students in the program	175	125	**300**
Students not in the program	325	275	**600**
↑ **Participation status**	**500**	**400**	**900**

The **Expected (*E*) Table** is below:

	Pass the exam	Fail the exam	← **Exam result**
Students in the program	166.667	133.333	**300**
Students not in the program	333.333	266.667	**600**
↑ **Participation status**	**500**	**400**	**900**

• c) The χ^2 **test statistic** is about 1.406. We use 1 df. **Decide** whether or not to reject H_0. (2 points)

Use these hints about the χ^2 distribution on 1 degree of freedom (df):



• d) Write our **conclusion** relative to the claim. (5 points)

7) (2 points). Fill in the blank: If a regression line for sample data is given by

$\hat{y} = 40 + 10x$, then along the regression line, for every increase of 1 unit in *x*,

there is an increase of _____ units in *y*.

8) (1 point). A student scores two standard deviations above the mean on Midterm 1 in a math class. According to the principle of **regression to the mean**, which of the following is the most likely outcome for the student on Midterm 2 in that class? Box in one:

a) The student will score three standard deviations above the mean on Midterm 2.

b) The student will score one standard deviation above the mean on Midterm 2.

c) The student will score two standard deviations below the mean on Midterm 2.

9) (2 points). Given sample bivariate data involving two variables, x and y, we obtain $r = 0.8$ and find the corresponding least squares regression model $\hat{y} = b_0 + b_1 x$. What proportion of the variance of y is accounted for by x and the regression model? Box in the best answer below, based on the class notes and homework:

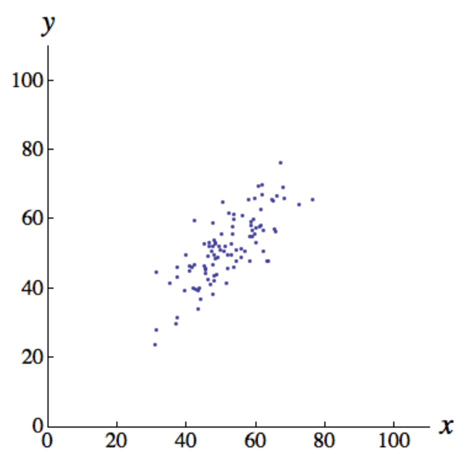
a) 8% b) 16% c) 64% d) 80% e) 160%

10) (8 points). (Matching)
For each variable, the average is 50 and the standard deviation is 10.

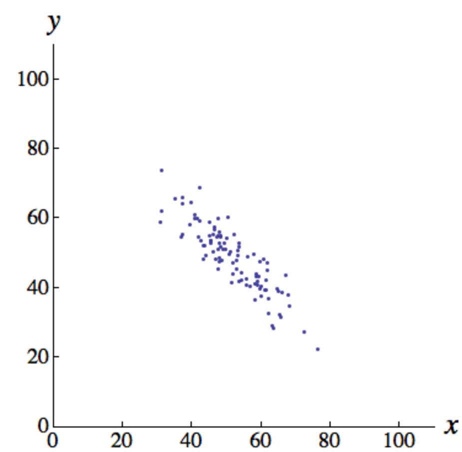
For one of the graphs below, $r = -0.90$.
For one of the graphs below, $r = 0.00$.
For one of the graphs below, $r = 0.80$.
For one of the graphs below, $r = 0.95$.

Fill in the blanks:

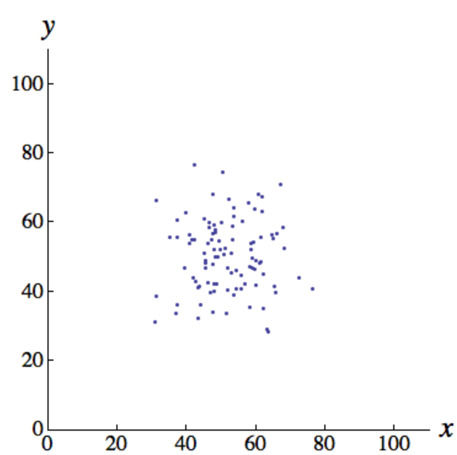
a) r for the graph below is ______.



b) r for the graph below is ______.



c) r for the graph below is ______.



d) r for the graph below is ______.