Biology 2244A/Statistics 2244A
Final Exam (25 marks, 40% to 63%, subject to flexible grading)
Monday, Dec 14, 2020; 7:00 pm to 10:10 pm

Student name: _____

1. There are **25 multiple choice questions** on **19 printed pages** in this Test booklet.
2. This is **a closed-book test** and must be **completed individually.**
3. You are allowed to use a hand calculator or an on-screen calculator and scrap paper. Other aids are not permitted.
4. Mark **one** choice for each multiple-choice question. There is only **one** correct answer. Answer all questions; there is no penalty for incorrect answers.

Caution: You might have personal experience or knowledge (unrelated to this course) of some of the scenarios used in a test question. *You should not use your own personal experiences/knowledge with the scenarios to select or justify your answer*. A combination of understanding of the course material (i.e. from lecture and lab materials) and the information in the question is all that should be necessary to select the correct answer.

Question 1.

Dairy cows often receive hormones to increase milk production. After the first hormone injection, a test herd of **81** randomly selected cows increased their daily production from **47** oz. to **56** oz. of milk, on average. The standard deviation of these increases was **18.0** oz.

```
>
> pt(q = 0.9, df = 80)
[1] 0.8145893
> pt(q = 0.95, df = 80)
[1] 0.8275133
> pnorm(q = 0.9, mean = 0, sd = 1)
[1] 0.8159399
> pnorm(q = 0.95, mean = 0, sd = 1)
[1] 0.8289439
>
```

If we assume this herd can be viewed as an SRS of a population of similar dairy cows, what is the margin of error of a **90% confidence interval** for the average increase in milk productions for cows injected with this hormone?

A. 1.629
B. 1.632
C. 1.655
D. 1.659

Question 2.

Intelligence Quotients (IQ) are normally distributed with mean **100** and standard deviation **15**.

```
>
> qnorm(p = 0.25, mean = 0, sd = 1)
[1] -0.6744898
> qnorm(p = 0.5, mean = 0, sd = 1)
[1] 0
> pnorm(q = 0.5, mean = 0, sd = 1)
[1] 0.6914625
> pnorm(q = 1, mean = 0, sd = 1)
[1] 0.8413447
> pnorm(q = 2, mean = 0, sd = 1)
[1] 0.9772499
>
```

What is the approximate probability that the **mean IQ** of four (**4**) randomly selected people will be between **85** and **115**?

A. 0.50
B. 0.68
C. 0.84
D. 0.95
E. 0.997

Question 3.

The growth of children from early childhood through adolescence generally follows a linear pattern. Data on the heights of American girls, from **9** to **13** years old, were compiled and the least squares regression line was obtained as $y = 80 + 6x$, where y is height in centimeters and x is age in years. All of the conditions for regression inference were met.

Which choice best describes the height of a **160** cm, **12** year-old American girl **relative to her age**?

A. Above average height for her age
B. Below average height for her age
C. Average height for her age
D. Not enough information to determine a valid answer.

Question 4.

The following is a one factor ANOVA table with some values missing.

Source	df	SS	MS	F
Group		50		2
Residual	30			
Total	35			

What is the total sum of squares?

A. 100
B. 125
C. 150
D. 175
E. 200

Question 5.

Suppose the heights of teenage Canadian boys, in inches, are normally distributed. I randomly select an SRS of 40 teenage Canadian boys, measure their heights, and construct a confidence interval based on the measurements.

I now wish to design a follow-up study. Which of the following choices will likely result in a new confidence interval that has a **smaller margin of error** than the original study? In each case, assume that no other changes are made to the methodology.

I. Measure the heights of 100 Canadian boys.
II. Only sample 15 year-old Canadian boys.
III. Decrease the confidence level.

A. I only
B. II only
C. I & II
D. I & III
E. I, II & III

Question 6.

An unscrupulous professor asks six essay questions on an exam. Before grading each paper, he rolls a 6-sided die and grades only the question corresponding to the number shown on the die. He assigns the grade for that essay as a grade for the entire exam. He does not tell his students about this practice and thus leads them to believe that he is grading all of their work.

Which choice best describes this grading scheme as a method of estimating the grade that a student actually deserves?

A. Biased and precise
B. Biased and imprecise
C. Unbiased and precise
D. Unbiased and imprecise

This question was typed into OWL incorrectly and the correct answer was not shown. Everyone received credit for this question

Question 7.

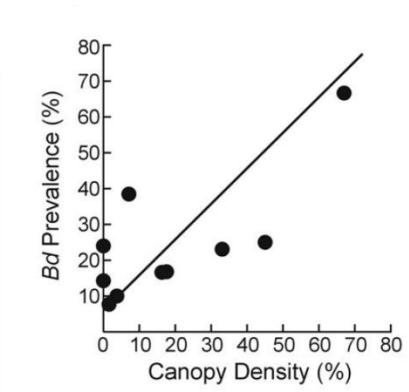
The concentration of the active ingredient in a drug is advertised as **1.5 mg / tablet**. Regulators understand that the concentration will naturally vary somewhat from pill to pill. But, they suspect that the drug maker is actually cheating to save money and mean concentration is lower than advertised. To explore this, they take an SRS of 75 tablets and calculate a hypothesis test. They obtain a *P*-value of 0.54. Based on these results, they decide whether to accuse the drug maker of wrong doing.

If the mean concentration is actually lower than **1.5 mg/tablet**, which choice is most likely the best description of the regulators' decision?

A. Type I error
B. Type 2 error
C. Type 3 error
D. Correct decision

Question 8.

Guilherme Becker et al (2013) studied the abundance of the chytrid fungus, *Batrachochytrium dendrobatidis* (abbreviated *Bd*) in the Adirondack Mountains of New York State. They selected a large property in the mountains known as the Archer and Anna Huntington Wildlife Forest as a sampling frame. Within this sampling frame, they selected an SRS of 10 ponds. At each pond, they measured the canopy density (higher density means thicker forest cover) and the prevalence of *Bd* in green frogs that they caught from the pond (percent of frogs infected). Their paper includes the following figure:



In this figure, dots represent individual ponds and the line represents a regression analysis (slope = 0.765, $P = 0.001$).

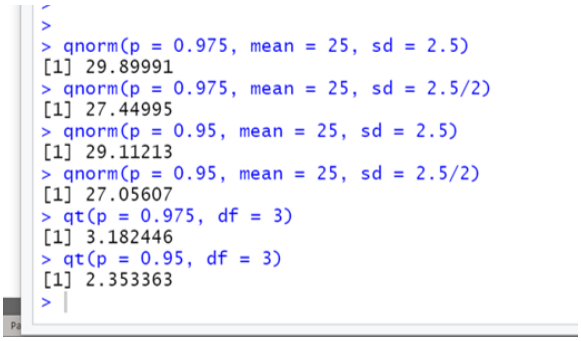
Assuming that the conditions for this inference are all met, how many of the following conclusions are **valid**?

I. Increasing canopy density causes an increase in *Bd* prevalence.
II. Increasing *Bd* prevalence causes an increase in canopy density.
III. Increasing canopy density causes a decrease in *Bd* prevalence.

A. None of the conclusions are valid
B. One of the conclusions is valid
C. Two of the conclusions are valid.
D. All of the conclusions are valid.

Question 9.

Suppose the heights of male German Shepherds follow a normal distribution with mean **25** inches and standard deviation **2.5** inches. Suppose four (**4**) unrelated male German Shepherds were entered into a dog show.



```
>
> qnorm(p = 0.975, mean = 25, sd = 2.5)
[1] 29.89991
> qnorm(p = 0.975, mean = 25, sd = 2.5/2)
[1] 27.44995
> qnorm(p = 0.95, mean = 25, sd = 2.5)
[1] 29.11213
> qnorm(p = 0.95, mean = 25, sd = 2.5/2)
[1] 27.05607
> qt(p = 0.975, df = 3)
[1] 3.182446
> qt(p = 0.95, df = 3)
[1] 2.353363
>
```

Approximately how tall (in inches) must one (**1**) of these dogs be for it to be considered among the tallest **5%** of all such dogs?

A. 27.35 inches
B. 27.45 inches
C. 29.11 inches
D. 29.90 inches
E. Not enough information to determine a valid answer.

Question 10.

The *P*-value for a two-sided hypothesis test is **0.14**.

Which of the following statements is **true**?

A. The probability that H_0 is true is 0.14.
B. If H_0 is true, the probability of failing to reject H_0 is 0.14.
C. If H_0 is true, the probability of observing a test statistic at least as extreme as the value observed is 0.14.
D. If H_0 is true, the probability of observing a test statistic at least as extreme as the value observed is 0.14 in each direction, making a total of 0.28.
E. The probability of making an incorrect decision regarding whether or not to reject H_0 is 0.14.

Question 11.

The following are three hypothetical residual plots that could be generated based on a linear regression.

Which of the <u>following statements</u> is **true**?

I. Plot 1 is the most likely to indicate the regression inference is valid.
II. Plots 2 & 3 indicate that the shape of the relationship between X and Y is non-linear.
III. Plot 3 indicates that the equal variance condition is not met.

A. I only
B. III only
C. I & II
D. I & III
E. I, II & III

Question 12.

A major contracting company is trying to get a sense of the preferred brick colour for houses in London, Ontario, to inform their future housing development projects. They have already surveyed three hundred (**300**) potential home buyers to get a sense of what buyers might prefer, but they suspect that some of the observed preferences may reflect how common brick colours already are (i.e. buyers may prefer colours that seem unique). Consequently, the company would like to survey the distribution of brick colours in houses that currently exist in London. The company selects fifty (**50**) houses to visit and records the colour of the bricks on the house.

Assuming the appropriate conditions are met, <u>which statistical procedure</u> would be most appropriate to address the following research question: Is the number of home buyers who prefer a red house proportionate to the number of red houses available?

A. Large sample test for p
B. Large sample test for $p_1 - p_2$
C. t-test for mu
D. t-test for $mu_1 - mu_2$

Question 13.

On Twitter (a social media platform), there is a friendly debate over what the 'best' programming language is for data analysis. Two of the dominant languages that battle back and forth for ease of use are Python and R. To get to the bottom of which of these two languages have a higher ranking for ease of use, a Twitter user created a survey that collected data from post-secondary school students who had taken one course in either Python, or, R (but not both). The survey asked the students to rank the programming language they learned in terms of ease of use on a Likert-style scale from **1** (extremely difficult) to **5** (extremely easy); a value of **3** was interpreted as neutral.

The survey creator separated the data into a sample from students who learned Python, and a sample who learned R. She computed the proportion from each sample who ranked the programming language as 4 or 5 (i.e. who found the program easy to some degree), and conducted a hypothesis test for the claim that there is no difference in proportion of students who find the language they learned easy when comparing Python students to R students. She obtained a *P*-value for the hypothesis test of **0.073** (assume that the model for the hypothesis test used was valid).

What choice best describes the <u>test statistic</u> that was used to calculate the *P*-value?

A. p-bar = $\text{p-hat}_1 - \text{phat}_2$
B. $t = (\text{x-bar}_1 - \text{x-bar}_2) / SE_{(\text{x-bar1} - \text{x-bar2})}$
C. $SE = \text{sqrt}[(\text{p-bar} * \text{q-bar} / n_1) + \text{p-bar} * \text{q-bar} / n_2]$
D. $Z = (\text{p-hat}_1 - \text{p-hat}_2) / SE_{(\text{p-hat1} - \text{p-hat2})}$
E. Z = (x-bar – mu) / (sigma / sqrt(n))

Question 14.

The mechanism by which people view TV shows and movies has changed dramatically over the last few years. Rather than purchasing a cable package with a series of channels, viewers purchase subscriptions to 'video on demand' companies like "Crave TV", or "Netflix", etc. Each such company offers a different selection of shows and movies for viewing. It's unclear whether the demographics of viewers who subscribe to such companies differ in some systematic way; one obvious characteristic that may vary across company subscribers is age.

A polling company conducted a representative online survey (i.e. it can be treated like an SRS) of **890** North American TV viewers and asked their age group and their favorite video on demand service. The data from their survey are summarized in the table below.

		Service			
		Amazon Prime	*Crave TV*	*Disney+*	*Netflix*
Age group	25 y and younger	70	85	40	35
	26 to 40 y	55	95	75	80
	41 to 65 y	50	25	60	110
	66 y and older	35	10	25	40

```
> prop.test(x = 255, n = 890, conf.level = 0.98)

        1-sample proportions test with continuity correction

data:  255 out of 890, null probability 0.5
X-squared = 161.39, df = 1, p-value < 2.2e-16
alternative hypothesis: true p is not equal to 0.5
98 percent confidence interval:
 0.2520952 0.3235601
sample estimates:
        p
0.2865169

> prop.test(x = 255, n = 890 - 255, conf.level = 0.98)

        1-sample proportions test with continuity correction

data:  255 out of 890 - 255, null probability 0.5
X-squared = 24.214, df = 1, p-value = 8.62e-07
alternative hypothesis: true p is not equal to 0.5
98 percent confidence interval:
 0.3565683 0.4482737
sample estimates:
        p
0.4015748

>
```

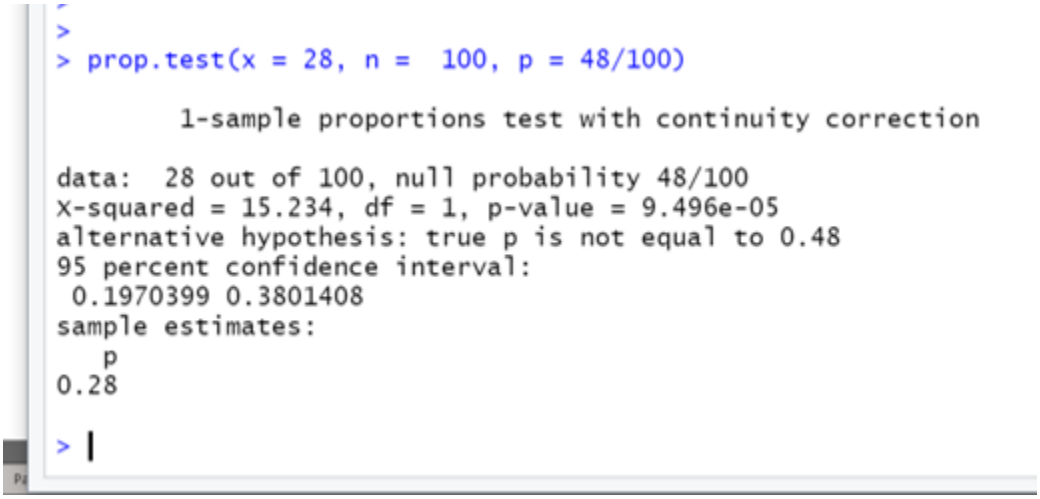
What is the approximate margin of error for a **98%** confidence interval for the proportion of viewers who prefer Netflix?

A. 0.036
B. 0.046
C. 0.072
D. 0.092

Question 15.

Surveys that ask school-aged children about their drinking, smoking, and illicit drug use are conducted on a regular basis across Canada. For research on drug use by New Brunswick students, researchers surveyed an SRS of **100** students about their prior drug use. They observed that **48%** of students had consumed alcohol and **28%** had tried cannabis at least once during the previous year.

The data are analyzed by Senior Statistician, Dr. Stan Darderor, who runs the following procedure in R.



Based on this analysis, Dr. Darderor reports that the 95% confidence interval for the difference in the proportion of students who have consumed alcohol and the proportion who have tried cannabis is (0.197, 0.380)

What is the <u>most appropriate feedback</u> to offer Dr. Darderor?

A. The analysis is correct – good job.
B. Actually, you should not conduct that analysis because you've violated some of the conditions, so your confidence interval is not valid.
C. Actually, you calculated a confidence interval for the proportion of students who consumed alcohol.
D. Actually, you calculated a confidence interval for the proportion of students who tried cannabis.

Question 16.

Van Dongen *et al* (2005) conducted a study to determine the impact of sleep deprivation on cognitive function (i.e. ability to remember, think critically, etc.). They obtained a sample of **48** participants who were screened initially to ensure that they had no previous sleep-related disorders. At the beginning of the study, the participants all slept for **8** hours for the first two nights. After the first two nights, the participants were each randomly assigned to one of four (**4**) sleep deprivation conditions:

- Sleep for only 8 hours per night for two weeks (9 participants)
- Sleep for only 6 hours per night for two weeks (13 participants)
- Sleep for only 4 hours per night for two weeks (13 participants)
- Sleep for 0 hours per night for three (3) consecutive nights (13 participants)

For all conditions, participants spent their days in a common laboratory setting, where sleep periods always ended at 7:30 am. Each day, the participants completed a series of tests that measured their working memory, and a set of questionnaires about their perceived level of sleepiness. The researchers concluded that sleep deprivation to 4 or 6 hours reduced cognitive function equivalently to three days without sleep, but participants were unaware of this negative impact.

The data are imported into R and stored in a data frame name *sleep_study*. The data frame includes columns named *cog_funct* (numeric assessment of cognitive function at the end of the experiment) and *sleep_treat* (the experimental condition assigned to each participant)

```
> m <- aov(cog_funct ~ sleep_treat, data = sleep_study)
> TukeyHSD(m)
```

Assuming the data are formatted correctly, how many statistical comparisons are calculated by the R function TukeyHSD() in the code above

A. 2
B. 3
C. 4
D. 6
E 10

Question 17.

A university professor conducted a research study investigating the relationship between employment and academic performance of students; her interest was on being able to advise new students on how to balance working with school. In particular, she wanted to use the data so she could advise individual students when asked, what grade should I expect to get on a test, based on how much time I spend working at my job?

The professor's study involved collecting data about two variables from a sample of students: (i) the number of hours they worked at a paying job during the week before a midterm test, and (ii) the grade that they received on the midterm test. She collected the following data:

	Work less than 15 hours per week	Work 15 or more hours per week
Passed the midterm	111	28
Failed the midterm	15	3

```
> pbinom(q = 2, prob = 18/157, size = 10)
[1] 0.9023852
> pbinom(q = 1, prob = 18/157, size = 10)
[1] 0.67909
> pbinom(q = 2, prob = 15/111, size = 10)
[1] 0.8572313
> pbinom(q = 2, prob = 139/157, size = 10)
[1] 1.083688e-06
>
```

Of an SRS of ten (**10**) students, what is the probability that two (**2**) or more **failed the midterm**?

A. 0.00
B. 0.10
C. 0.14
D. 0.32
E 0.68

Question 18.

A controversial issue in sports is the use of the "instant replay" for making decisions on plays that are difficult to call by an official. A survey of **102** professional athletes in four sports was conducted, asking them if they felt "instant replays" should be used to decide close calls. The results are as follows:

	Use of Instant Replay	
	Favor	Oppose
Football	22	2
Baseball	18	6
Basketball	15	26
Hockey	3	10

The consulting company that collected these data now wants to investigate whether an athlete's attitude is related to how often past instant replays have turned out in their favour. Investigating this question will involve watching many hours of recorded sporting events, so the company decides to select a small sub-sample from the original survey respondents.

Which of the following sampling schemes meets the **criteria of a binomial setting**?

I. Selecting an SRS of 4 of the 22 football players
II. Selecting an SRS of 4 of the 102 athletes
III. Randomly selecting one athlete who plays each sport (total n = 4)

A. I only
B. II only
C. I and II
D. I and III
E II and III

Question 19.

Prof. Rubin has a recurring nightmare. Each time he has this dream, the exact situation may change but the overall idea is that he is taking an exam for which he is completely unprepared. In one instance, he is given a multiple-choice test which consists of **10** questions. Each question has four (**4**) possible answers and for each question, Prof. Rubin chooses an answer randomly.

Which of the following pairs of events are independent?

I. Correct answer for question 1; correct answer for question 2
II. Correct answer for question 1; Passing the test (5 or more correct answers in total)
III. Correct answer for question 1; incorrect answer for question 1

A. I only
B. II only
C. III only
D. I and II
E I and III

Question 20.

Each person in a random sample of males and females was asked to state his/her gender and preferred colour. The resulting frequencies are shown below.

	Preferred Colour		
	Red	Blue	Green
Male	3	11	6
Female	17	11	2

Which of the following statements is **FALSE**?

A. The probability that a male respondent prefers the colour blue is 0.55.
B. Of respondents who prefer the colour green, the probability of being female is 0.25.
C. The probability that a respondent prefers blue is 0.44.
D. The probability of preferring blue is higher for male respondents than or female respondents.
E The probability that a respondent is male and prefers the colour red is 0.15.

Question 21.

An advocacy group is interested in whether or not Canadians support the recent legalization of Marijuana. First, they obtain a complete list of all **308** federal voting districts. Then, they select a random sample of **25** voters from each district.

Which choice best describes the sampling method used?

A. Stratified sample
B. Cluster sample
C. Systematic sample
D. Multistage sample

Question 22.

A group of statistical analysts discover that they share an interest in playing the board game, Monopoly. They organize a weekly game on Friday evenings and call their group Statistical Professionals who Adore Monopoly (SPAM). The game of Monopoly relies heavily on rolling dice and the number of times the dice are rolled in the course of a game varies between approximately 100 and 300.

To make sure that the dice are not tampered with, one of the more competitive players, Dr. Anne Nova, keeps records of all official dice rolls each week and analyzes the probability that they could result from fair dice. For the sake of transparency, Dr. Nova also posts the weekly tallies of dice rolls to the SPAM website.

Which of the following found on the SPAM website best represents a sampling distribution?

A. A histogram showing all 123 dice rolls from the game on February 14, 2019
B. The standard deviation of an SRS of 62 dice rolls from the game played on June 4, 2018
C. A histogram showing the mean roll for each game between Dec 10, 2011 and Jan 31, 2020
D. A histogram showing the number of fives in the first twenty rolls for each game between Dec 10, 2011 and Jan 31, 2020.
E. A histogram showing the number of times the dice were rolled in each game between Dec 10, 2011 and Jan 31, 2020

Question 23.

In this class you learned about an overall framework of inquiry abbreviated PPDAC. In lab, you also learned and applied a multistage data analysis protocol.

Which of the following choices best describes the relationship between PPDAC and the data analysis protocol?

A. The data analysis protocol is a detailed procedure for completing the D, A, and C steps of PPDAC.
B. The data analysis protocol outlines the same general concepts as PPDAC but in more quantitative detail.
C. P, P, D, A, and C correspond directly to the five steps of the data analysis protocol.
D. The D, A and C steps are all included in the data analysis protocol, but the protocol arranges them in a different order.

Question 24.

A soft drink dispenser has a knob that allows the user to set the volume of soda it will dispense to anywhere between **8** oz. and **28** oz. The manager of the fast food restaurant that owns the dispenser wants to know whether the volume setting is accurate. She understands that the exact volume of soda dispensed will vary slightly each time the machine is used. Also, she is willing to assume that as long as the knob remains unchanged, the dispensed volumes will be normally distributed and that any set of drinks dispensed can be treated as an SRS.

She sets the knob to **20 oz.** and invites five of her closest friends over to the restaurant after closing one day for a double-bubble-trouble soda-drinking extravaganza. She uses the machine to dispense 6 soft drinks and measures the volume of each very carefully with a flask that she borrowed from the lab of a chemist who lives in her neighborhood. She and her five friends each drink one of the six sodas. Then she repeats the procedure of pouring a second round of drinks and measuring the volumes carefully before they are consumed.

After recovering from her stomach-ache, she sits down to analyze the data she collected: n = **12**, mean = **19.8 oz.**, standard deviation = **0.5 oz**.

```
> pnorm(q = -0.2/(0.5/sqrt(12)), mean = 0, sd = 1)
[1] 0.08292833
> pnorm(q = -0.2/0.5, mean = 0, sd = 1)
[1] 0.3445783
> pt(q = -0.2/0.5, df = 11)
[1] 0.3484073
> pt(q = -0.2/(0.5/sqrt(12)), df = 11)
[1] 0.09665353
>
>
```

What is the two-tailed P-value that she calculates?

A. 0.16
B. 0.19
C. 0.69
D. 0.70
E. Even with the assumptions she is willing to make, a valid *P*-value cannot be determined.

Question 25.

A friend of yours is interested horoscopes and wants to test astrological predictions using statistics. Since your friend knows you have taken 2244, he comes to you for advice in planning his study. He explains that he wants to conduct a two-sided test to see whether the proportion of people who identify as left-handed differs between people who are born under the star sign of Aries and those born under the sign of Gemini. (In astrology, a person's "star sign" is determined by the month and date of their birth).

He believes that he will be able to obtain a suitable sample of 25 individuals born under each star sign. He wishes to use a low value of alpha (alpha = 0.01) because there are a lot of skeptics about horoscopes. Also, he believes that if an effect does exist it is likely to be quite small. He wants to know whether the sample size is large enough for his purposes.

After you give your friend some quick pointers about power analysis and R, he comes back with the following calculations:

```
> library(pwr)
> cohen.ES(test = "p", size = "medium")

     Conventional effect size from Cohen (1982)

           test = p
           size = medium
    effect.size = 0.5

> pwr.2p.test(h = 0.5, sig.level = 0.01, n = 25)

     Difference of proportion power calculation for binomial
 distribution (arcsine transformation)

              h = 0.5
              n = 25
      sig.level = 0.01
          power = 0.2095344
    alternative = two.sided

NOTE: same sample sizes

> |
```

Based on the information shown, what is the best advice for you to give your friend?

A. I think the study will work out fine the way you have planned it.
B. Based on these results, you could save yourself some time by using a smaller sample.
C. Based on these results, you should probably try to get a larger sample if you really want to go ahead with this.
D. Sorry, you've done the calculations wrong. There is really nothing I can tell you based on this output.