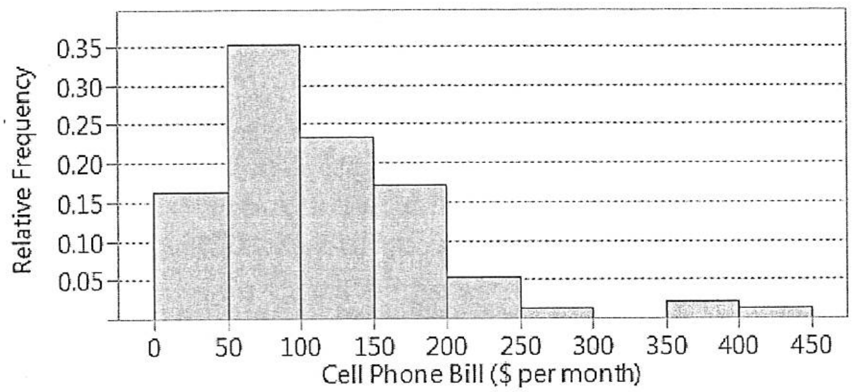
STATISTICS
SECTION I
Time --- 1 hour and 30 minutes
Number of questions --- 40
Percent of total grade --- 50

Directions: Solve each of the following problems, using the available space for scratch work. Decide which is the best of the choices given and fill in the corresponding box on the answer sheet. Do not spend too much time on any one problem.

1. A survey was administered to a nationwide random sample of 199 adults who were asked to report their monthly cell phone bill (in dollars). A histogram of the results is shown below.



Which of the following statements must be <u>false</u>?

T (A) The distribution of monthly cell phone bills is skewed to the right.

T (B) The mean monthly cell phone bill is higher than the median monthly cell phone bill. mean $\bar{x}$ → pulled toward tail

T (C) The number of respondents whose monthly cell phone bill is less than \$50 is approximately 32. $.16(199) = 31.84$

T (D) There were no reported monthly cell phone bills between \$300 and \$350.

F (E) Sixteen percent of respondents reported monthly cell phone bills of more than \$50.

$1 - .16 = (84\%)$

2. A college professor has received comments on his teaching evaluations that his handwriting on the board is too small in large classrooms. He wants to determine if this is affecting his students' grades, so one semester he selects a random sample of his students and takes note of whether each student in this sample is sitting in the front half or the back half of the classroom, then compares the course averages of these two groups at the end of the semester. He discovered that students who sit in the front half of the room have significantly higher final course averages. Which of the following is the most appropriate conclusion based on the results of this study?

(A) Where a student chooses to sit in the classroom has no effect on how well they do in the course.

(B) Students who sit at the back of the classroom do not perform as well as those who sit at the front because they cannot see what is written on the board.

(C) Students who sit at the back of the classroom do not perform as well as those who sit at the front because they are texting and otherwise inattentive.

(D) Students who sit at the front of the classroom are more serious students.

(E) None of the above conclusions are appropriate.

association $\neq$ causation

---

3. The completion time for male Olympic marathon runners is approximately normally distributed with a mean of 140.59 minutes and a standard deviation of 8.01 minutes. The completion time for female Olympic marathon runners is approximately normally distributed with a mean of 156.14 minutes and a standard deviation of 9.62 minutes. A runner's time was recorded as 157.90 minutes but their gender was not recorded. Given that it is unusual for an observation to fall more than two standard deviations from the mean in an approximately normal distribution, which of the following statements is true?

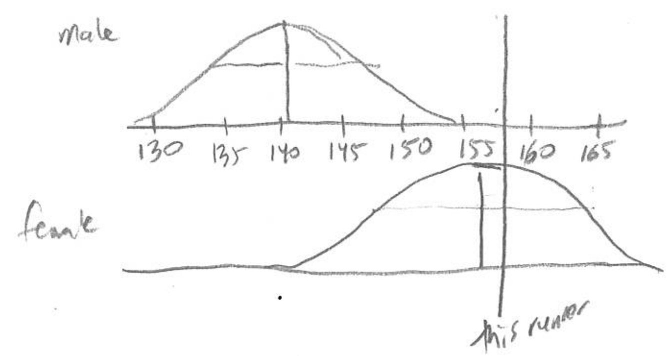
(A) The time of 157.90 minutes would be considered unusual for both male and female Olympic marathon runners.

(B) The time of 157.90 minutes would not be considered unusual for either male or female Olympic marathon runners.

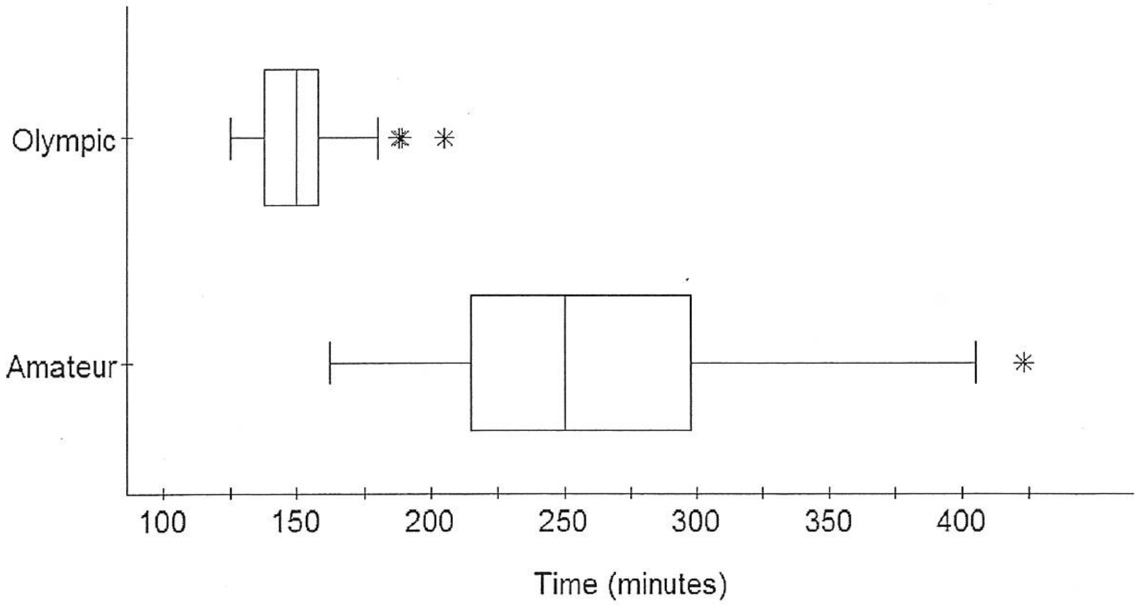
(C) The time of 157.90 minutes would be considered unusual for male but not for female Olympic marathon runners.

(D) The time of 157.90 minutes would be considered unusual for female but not for male Olympic marathon runners.

(E) We do not have enough available information to make a determination if the time is unusual.



4. The boxplots below show the completion times in minutes for marathon runners. The runners included both amateur and Olympic athletes.



Which measures of center and spread would be best to use when comparing the distribution of marathon completion times for the two groups of runners?

(A) Mean and interquartile range (IQR)

(B) Mean and standard deviation

(C) Median and interquartile range (IQR) due to presence of skew and outliers

(D) Median and standard deviation

(E) Median and range

5. Suppose that Mary, Gary, and Larry have been asked to randomly sample 100 patients' visit records from a file of 2000 patient records and record the amount of time the attending physician spent with each patient. Gary suggests the time of year during which the patient was seen may be a confounding variable and strongly advises dividing the year into quarters (Jan – Mar, Apr – Jun, Jul – Sep, Oct – Dec) and randomly selecting 25 patient records from each quarter. Which of the following best describes Gary's sampling plan?

(A) A simple random sample

(B) A systematic random sample

(C) A stratified random sample when you believe groups are different and want to include some from each

(D) A cluster sample ← when you believe groups are the same so you can just take one group as the sample

(E) A convenience sample

---

6. A statistics student at Pleasantville High School (PHS) looked at seat belt use by drivers. Customers were observed driving into a local convenience store. After the drivers left their cars, the student asked each driver several questions about seat belt use. In all, 80% of the drivers said that they always use seat belts. However, the student observed that only 61.5% of these same drivers were actually wearing a seat belt when they pulled into the store parking lot. Which of the following best explains the difference in the two percentages?

(A) The difference is due to sampling variability. We shouldn't expect the results of a sample to match the truth about the population every time.

(B) The difference is due to response bias. Drivers who don't use seat belts are likely to lie and say they do use seat belts.

(C) The difference is due to undercoverage bias. The study included only customers of the convenience store and did not include all drivers in the population.

(D) The difference is due to nonresponse bias. Drivers who don't use seat belts are less likely to respond to the student's questions.

(E) The difference is due to voluntary response bias. Drivers are able to volunteer to participate in this survey.

7. In a standard normal distribution, approximately what percentage of the observations are between 2 and 3 standard deviations from the mean?

(A) 2.1%

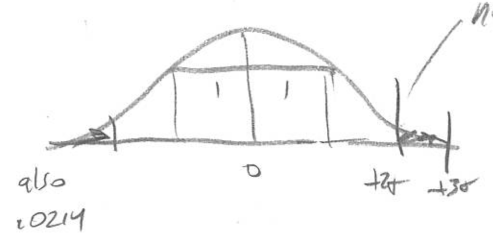
(B) 4.3%

(C) 13.5%

(D) 27%

(E) 95%

normalcdf(2,3,0,1) = .0214

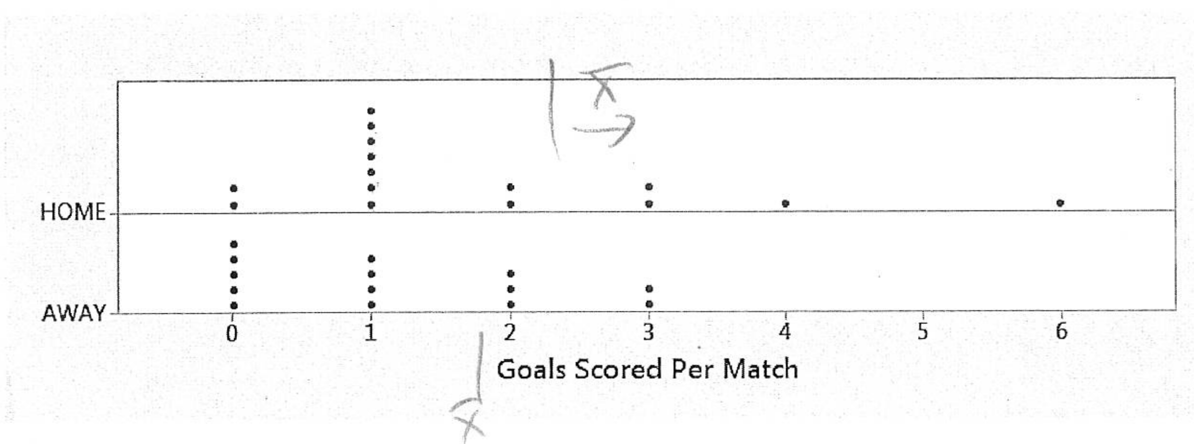


also .0214

$2 \times .0214 = .0428$

---

8. The distribution of number of goals scored per match in the 15 home and 14 away matches so far this season for one team in the English Premier Soccer League is shown below.



Which of the following statements is true?

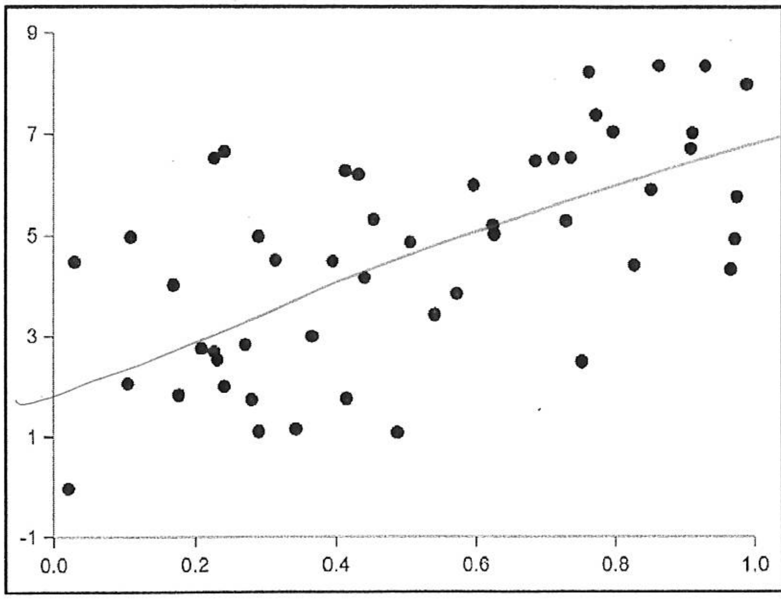
✓ (A) The mean number of goals scored in AWAY matches is less than the mean number of goals scored in HOME matches. skew pulls $\bar{x}$ towards itself

X (B) The range of goals scored for AWAY matches is greater than the range of goals scored for HOME matches. $range_{away} = 3-0=3$, $range_{home} = 6-0=6$

X (C) The shape of the distribution of HOME goals scored is left-skewed. = right skewed

X (D) The median number of goals scored for AWAY matches is more than the median number of goals scored for HOME matches. $MEDIAN_{away} = 1$, $MEDIAN_{home} = 1$

X (E) The number of AWAY matches in which the team scored three or more goals is three. (2)



b positive so r positive medium strength

9. Which of the following is the most likely value of the linear correlation coefficient between the two variables depicted in the scatterplot above?

(A) −0.95

(B) −0.60

(C) 0

(D) 0.60

(E) 0.95

---

10. In the weeks leading up to the 2016 presidential election, a random sample of 1001 likely voters in northern California was selected by a major polling organization. Each person was asked who he/she planned to vote for in the upcoming presidential election. Of those surveyed, 125 claimed to still be undecided about their preferred candidate. To which of the following populations can the results of this survey be safely generalized?

(A) All likely voters in California

(B) Only undecided voters in northern California

(C) Only registered voters in northern California

(D) Only likely voters in northern California

(E) Only the 1001 likely voters in northern California that were surveyed

11. Christy operates a business called *Dinner to Go!* that provides home-cooked meals for people that don't have time to cook for themselves. Each meal consists of a main dish and two side dishes. Bob operates a similar business across town called *Let Me Feed You.* They decide to combine their operations. Both of them will continue working independently from their own locations, but Christy will provide the side dishes and Bob will provide the main dish. The mean amount of time it takes Christy to make the side dishes for one day is 4.2 hours with a standard deviation of 0.8 hours. The mean amount of time it takes Bob to make the main dishes for one day is 5.1 hours with a standard deviation of 1.1 hours. The price customers are charged will be based in part on the total time spent preparing the meals. Find the mean and standard deviation for the total amount of time necessary to make the side dishes and main dish for one day.

(A) mean = 9.30 hours, standard deviation = 1.90 hours

(B) mean = 9.30 hours, standard deviation = 1.85 hours

(C) mean = 9.30 hours, standard deviation = 1.36 hours

(D) mean = 4.65 hours, standard deviation = 1.90 hours

(E) mean = 4.65 hours, standard deviation = 1.36 hours

✓ means follow the algebra

variances always add ✓

$T = M + S$

$\mu_T = \mu_M + \mu_S$
$= 5.1 + 4.2$
$= 9.3$ hrs

$\sigma_T^2 = \sigma_M^2 + \sigma_S^2$
$\sigma_T^2 = (1.1)^2 + (.8)^2$
$\sigma_T = \sqrt{(1.1)^2 + (.8)^2}$
$= 1.36$ hrs

---

12. Consider two events A and B in a probability experiment. Define A^C to be the complement of event A and B^C to be the complement of event B. If $P(A \cup B) = 0.75$, $P(A \cap B^C) = 0.25$, and $P(A^C \cap B) = 0.35$, then $P(B) =$

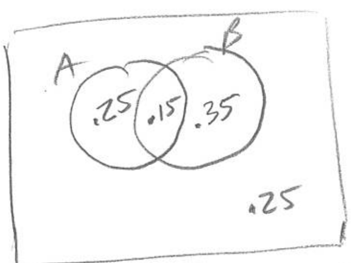
(A) 0.35

(B) 0.50

(C) 0.60

(D) 0.80

(E) The probability cannot be determined from the given information.

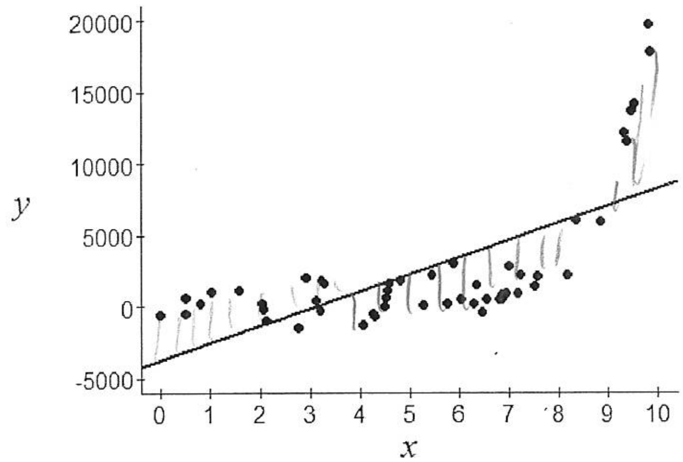


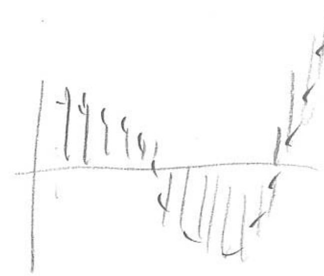
$P(B) = .15 + .35 = .5$

$.75 - .25 - .35 = .15$

$1 - .75 = .25$

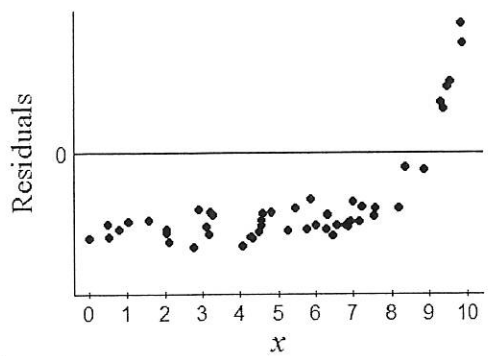
13. Consider the scatterplot and least-squares regression line shown here.



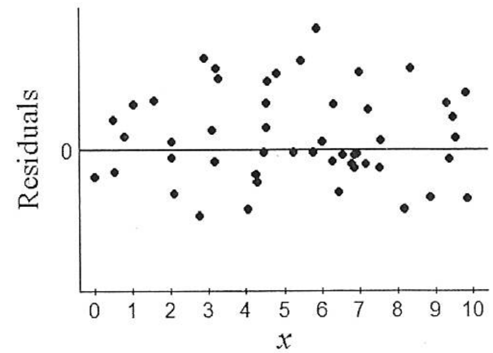


Which of the following could represent a residual plot for this fit?

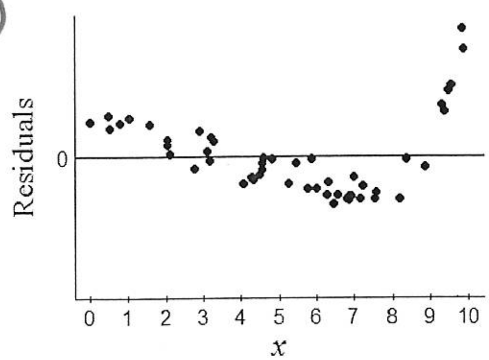
(A)



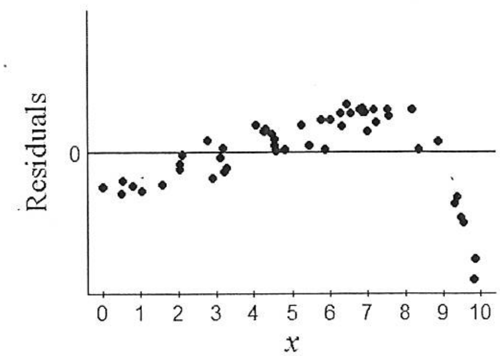
(B)



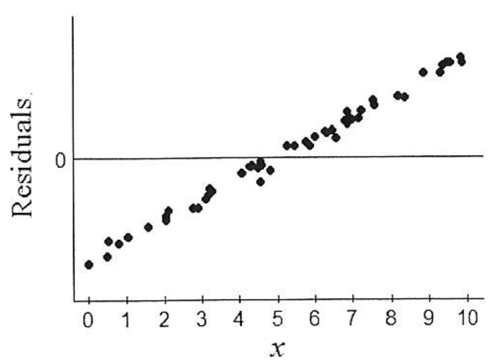
(C)



(D)



(E)



14. An experiment is to be conducted to determine which detergent (A or B) has the least effect on the discoloration of fabrics after repeated washings. In previous studies it was determined that fabric type is associated with discoloration after repeated washings while fabric color is not associated with discoloration. This experiment would best be done

(A) by blocking on fabric type

(B) by blocking on fabric color

(C) by blocking on detergent type

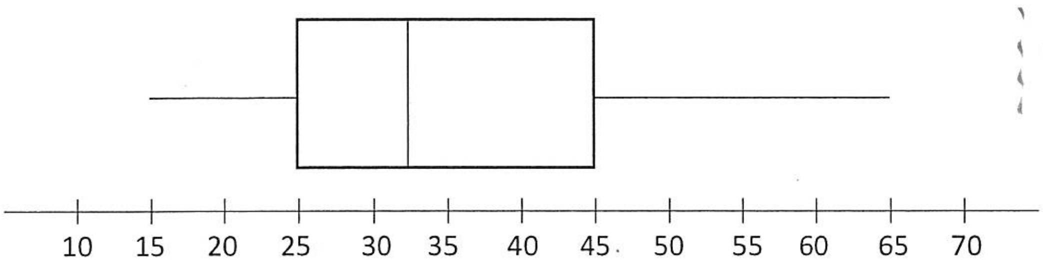
(D) by blocking on discoloration rating

(E) by using a completely randomized design

BLOCK on traits of the subjects that may produce different results

---

15. Consider the boxplot for a data set of $n = 50$ observations given below.



outlier

If the maximum value of 65 is changed to 70, would the value of 70 be considered an outlier?

(A) Yes, because 70 would be greater than $Q_3 + 1.5(Q_3 - Q_1)$.

(B) No, because 70 would not be greater than $Q_3 + 1.5(Q_3 - Q_1)$.

(C) Yes, because 70 would be more than 3 standard deviations above the mean.

(D) No, because 70 would not be more than 3 standard deviations above the mean.

(E) It cannot be determined from the given information.

upper fence $= Q3 + 1.5 IQR$
$= 45 + 1.5(45-25)$
$= 45 + 1.5\ (20)$
$= 45 + 30$
$= 75$

change of max point no effect on Q3 or IQR

16. Jack and Jill want to know whether running up a hill once a day improves lung capacity and decided to conduct an experiment. Both Jack and Jill measured their lung capacity before the experiment. Through random allocation, it was decided that Jack would not exercise for six weeks while Jill ran up the hill once a day for the same six week period. At the end of six weeks, Jack and Jill again measured their lung capacity and compared the measure to the initial value for potential improvement. What key feature of a well-designed experiment did Jack and Jill not incorporate?

(A) Random assignment of subjects to treatments

(B) Replication within each treatment group

(C) Stratifying by factors that affect the response variable

(D) A control group for comparison

(E) This experiment includes all key features of a well-designed experiment

2 meanings for replication:
1) within one experiment:
replication = large enough groups for randomization to spread unknown factors between groups

2) beyond one experiment:
replication = repeat experiments to verify repeatability of results.

---

17. A medical research study is interested in predicting the blood pressure (in mmHg) for subjects exposed to noise at various volumes (in decibels). A least-squares regression line was fit to data collected from 20 subjects who were randomly assigned to receive a certain noise exposure level. The equation of the line is

$$\hat{y} = -10 + 0.20x$$

where x is noise exposure and $\hat{y}$ is the predicted blood pressure. Which of the following gives the best interpretation of the slope of the least-squares regression line?

(A) There is an increase of 0.20 mmHG in the predicted blood pressure of a subject for every increase of 1 decibel in noise exposure volume. "on average"

(B) There is an increase of 0.20 decibels in the predicted noise exposure volume of a subject for every increase of 1 mmHg in blood pressure.

(C) There is a decrease of 10 mmHG in the predicted blood pressure of a subject for every increase of 1 decibel in noise exposure volume.

(D) There is a decrease of 10 decibels in the predicted noise exposure volume of a subject for every increase of 1 mmHg in blood pressure.

(E) Approximately 20% of the variability in blood pressure is predicted by its linear relationship with noise exposure volume.

18. At a local high school with over 2,000 students, 65% of the student body participates in a school sponsored extracurricular activity. Suppose that the school newspaper is doing an article on student participation in extracurricular activities and surveys a random sample of 100 students about their participation. Let $\hat{p}$ represent the proportion of students who participate in an extracurricular activity in a sample of size 100. Why is the sampling distribution of $\hat{p}$ approximately normal?

(A) Because np and $n(1-p)$ are both sufficiently large.

(B) Because the sample was randomly selected.

(C) Because the sample size is less than 10% of the population size.

(D) Because the sampling distribution of a sample proportion is always approximately normal.

(E) The sampling distribution of $\hat{p}$ is not approximately normal because the survey question has only two possible answers, and a binary variable cannot be approximately normally distributed.

---

19. You simulate tossing a fair coin 10,000 times. Which of the following statements is true?

$EV = np = 10000(0.5)$ for binomial

(A) It is guaranteed that you will observe exactly 5,000 heads.

(B) The proportion of the 10,000 simulated tosses resulting in heads is expected to be 0.5.

(C) The chance of toss number 1,000 resulting in heads depends on the results of the previous 999 tosses.

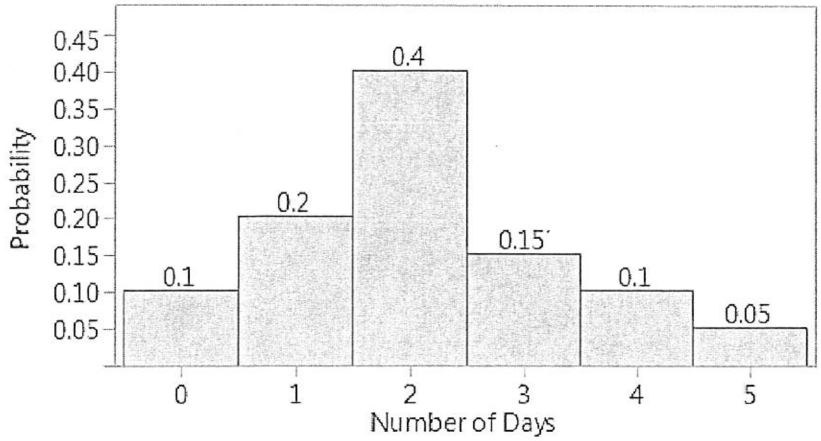
(D) If you get a string of 15 tails in a row the probability that the next toss will be a head is greater than 0.5.

(E) All of the above statements are true.

20. When the stock market closes higher than it opened at the beginning of the day, the stock market is said to have "closed up." The histogram below displays the probability distribution for the number of days the stock market "closes up" in a given week. In the long run, what is the average number of days per week the stock market "closes up?"

mean

expected value = probability weighted average



(A) 2.0

(B) 2.1

(C) 2.2

(D) 2.5

(E) 3.0

$(0)(0.1)+(1)(.2)+(2)(.4)+(3)(.15)+(4)(.1)+(5)(.05)$

$= 2.1$

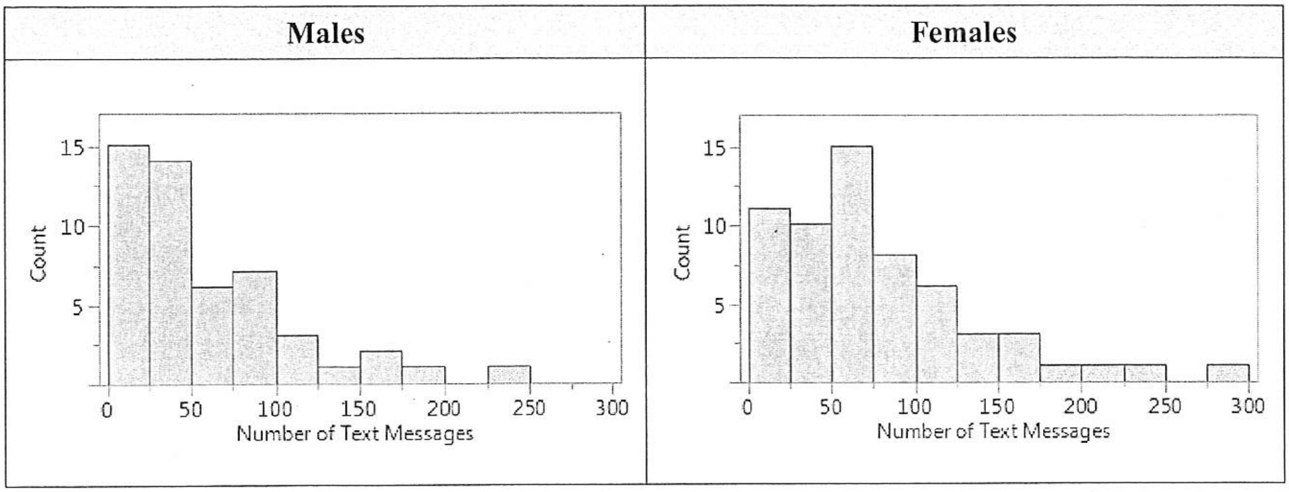
or

L1	L2
0	.1
1	.2
2	.4
3	.15
4	.1
5	.05

1Varstats L1, L2

$\bar{x} = 2.1$

21. For their final project, a group of statistics students at a large university investigated their belief that females at their school text more than males at their school. They asked a random sample of 110 students – 50 males (M) and 60 females (F) – from their school to record the number of text messages sent and received over a two-day period. Histograms of their data are shown below.



Would it be appropriate to use this data to perform a two-sample mean t-interval addressing these students' belief?

(A) No, because the histograms are quite skewed indicating it is not reasonable to assume that the populations are normally distributed.

(B) No, because the data are not from independent random samples.

(C) No, because $n_M\hat{p}_M$, $n_M(1-\hat{p}_M)$, $n_F\hat{p}_F$, and $n_F(1-\hat{p}_F)$ are not all sufficiently large.

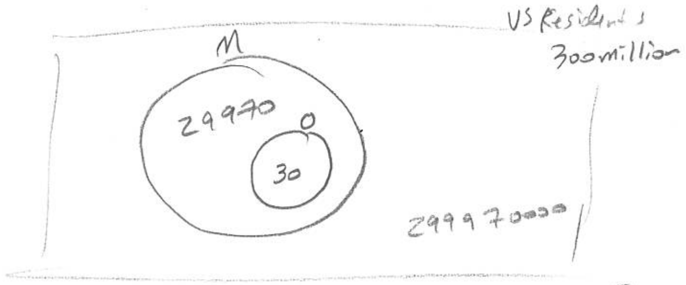
(D) No, because n_M and n_F are not equal.

(E) Yes, because all of the conditions for this inference procedure have been met.

22. Of the approximately 300 million current U.S. residents, only about 30,000 have ever appeared in a movie. Only 30 of the current U.S. residents having appeared in a movie have ever won an Oscar for Best Actor, which is an award presented annually to the American actor who delivered the most outstanding performance in a leading movie role. A person is randomly selected from the current U.S. residents. Let M represent the event the selected individual has appeared in a movie. Let O represent the event the selected individual has ever won an Oscar for Best Actor. Order the following probabilities from smallest to largest.

I. P(O)
II. P(M)
III. P(M|O)
IV. P(O|M)

(A) I, III, II, IV
(B) IV, II, I, III
(C) IV, III, I, II
(D) I, II, IV, III
(E) IV, I, III, II



$I\ P(O) = \frac{30}{300{,}000{,}000} = 1 \times 10^{-7}$

$II\ P(M) = \frac{30000}{300{,}000{,}000} = 1 \times 10^{-4}$

$III\ P(M|O) = \frac{P(M \cap O)}{P(O)} = \frac{30}{30} = 1$

$IV\ P(O|M) = \frac{P(O \cap M)}{P(M)} = \frac{30}{30000} = .001$

$I < II < IV < III$

---

23. A polling agency wants to estimate the percentage of voters in favor of repealing and replacing the Affordable Care Act, and it wants to provide a margin of error of no more than 1.8 percentage points with 99% confidence. Which of the following should be used to determine the number of respondents (n) the agency must poll?

(A) It cannot be determined from the given information.

(B) $1.96\sqrt{\frac{(0.5)(0.5)}{n}} \le 0.018$

(C) $2.576\sqrt{\frac{(0.5)(0.5)}{n}} \le 0.018$

(D) $2.576\sqrt{\frac{(0.5)(0.5)}{n}} \le 0.036$

(E) $2.576\sqrt{\frac{(0.99)(0.01)}{n}} \le 0.018$

proportion so

$CI = \hat{p} \pm \underbrace{z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}}_{ME}$

z^* for 99%:

$z^* = \text{invNorm}(.005, 0, 1)$
$= (-)\ 2.576$

for worst case, should use (0.5) for $\hat{p}$

one sample = χ^2 test of independence

24. A random sample of 50 students was taken by the principal of a large high school to see if a relationship exists between gender and attitude concerning a proposed mandatory community service requirement. The data appear in the table below.

(association)

	Female	Male	Total
Approve	11	8	19
Disapprove	5	13	18
Undecided	6	7	13
Total	22	28	50

Which of the following gives the null hypothesis in a test to determine if a relationship exists between gender and attitude concerning a proposed mandatory community service requirement?

(A) $p_{\text{Female}} = p_{\text{Male}}$

(B) $p_{\text{Approve}} = p_{\text{Disapprove}} = p_{\text{Undecided}}$

(C) Gender of a student is independent of opinion on the proposal.

(D) Gender of a student is not independent of opinion on the proposal.

(E) The distribution of opinions is the same for each gender.

---

25. A waiter kept track of how much he received in tips from each table he served over the course of the past year. The distribution of these tip amounts is approximately normal with a mean of \$6.50 and a standard deviation of \$2.10. The top 20% of his tips are greater than what amount?

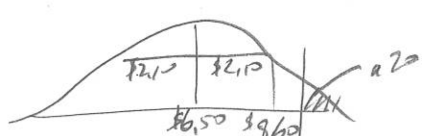
(A) \$2.52

(B) \$3.81

(C) \$7.80

(D) \$8.27

(E) \$9.19



? = invNorm(.80, 6.50, 2.10)
= \$8.267

26. A test of $H_0: \mu = 10$ vs. $H_a: \mu < 10$ is conducted based on a random sample of 15 observations. Assume that the conditions for inference have been met. What does it mean to use a significance level of $\alpha = 0.05$ for the test?

(A) The test procedure would result in an incorrect decision for 5% of all possible random samples of 15.

(B) If $\mu < 10$, the test procedure would result in a correct decision for 5% of all possible random samples of 15.

(C) If $\mu < 10$, the test procedure would result in an incorrect decision for 5% of all possible random samples of 15.

(D) If $\mu = 10$, the test procedure would result in a correct decision for 5% of all possible random samples of 15.

(E) If $\mu = 10$, the test procedure would result in an incorrect decision for 5% of all possible random samples of 15.

	Ho T	Ho F
R	I $\alpha = .05$	
NR		II

$\alpha = P(\text{type I error})$
$= P(\text{rejecting Ho} \mid \text{Ho is true})$
$= P(\text{incorrect decision} \mid \mu = 10)$

---

27. The table below classifies vehicles by their country of origin and type.

		Type of Vehicle						
		Hybrid	SUV	Sedan	Sports	Truck	Wagon	Total
Country of Origin	Asia	3	25	94	17	8	11	158
	Europe	0	10	78	23	0	12	123
	USA	0	25	90	9	16	7	147
	Total	3	60	262	49	24	30	428

Which of the following statements is true?

(A) Of the six vehicle types, sedans have the highest proportion that are made in Asia. Sedan $\frac{94}{262}$ .36, Hybrid $\frac{3}{3} = 1.0$

(B) For each of the three countries of origin, over half of the vehicles produced are sedans. $\frac{94}{158}$ (.59), $\frac{78}{123}$ (.63), $\frac{90}{147}$ (.61)

(C) Of the six vehicle types, SUVs have the lowest proportion of vehicles from Europe. no hybrid $\frac{0}{3} = 0$

(D) The proportion of trucks is higher among vehicles made in Asia than among vehicles made in the USA. $\frac{16}{147}$ (11%) $\frac{8}{158}$ (5%)

(E) More than half of the vehicles were either an SUV or a hybrid. $\frac{3+60}{428} = .15$

28. The 600 students enrolled in an introductory statistics course at a large university were surveyed and asked to report how many hours they participate in sports or other physical exercise in a typical week. Selected summary statistics are provided below.

Mean: 6.5 hours
Standard Deviation: 5.5 hours
First Quartile: 3 hours
Median: 5 hours
Third Quartile: 8 hours

About 300 of the surveyed students typically exercise

(A) less than 6.5 hours each week

(B) more than 3 hours each week

(C) more than 8 hours each week

(D) between 1 and 12 hours each week

(E) between 3 and 8 hours each week

---

29. In a survey conducted by the Gallup organization, a random sample of U.S. adults who work full time were asked to report how many hours they worked in the previous week. Based on the results, a 95% confidence interval for the mean number of hours worked is 46.1 $\pm$ 1.2 hours. Which of the following *could* be the 99% confidence interval based on the same survey results?

(A) 46.1 $\pm$ 0.8

(B) 46.1 $\pm$ 1.2

(C) 46.1 $\pm$ 1.6

(D) 46.1 $\pm$ 6.1

(E) Without knowing the sample size, any of the above intervals could be the 99% confidence interval.

ME for means: $t^* \frac{s}{\sqrt{n}} = ME$

if n is large $t^* \approx z^*$

for 95% $z^* = 1.96$

for 99% $z^* = 2.58$

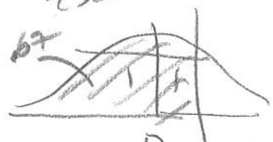
$\frac{s}{\sqrt{n}} = \frac{ME}{t^*} = \frac{1.2}{1.96} = .6122$

for 99% t^* changes.

ME now $(2.58)(.6122) = 1.58$

30. A score at the 67th percentile of an approximately normal distribution is how many standard deviations from the mean?

z score



? = invNorm(.67, 0, 1)

= .4399

(A) 1.84 standard deviations below the mean

(B) 0.75 standard deviations below the mean

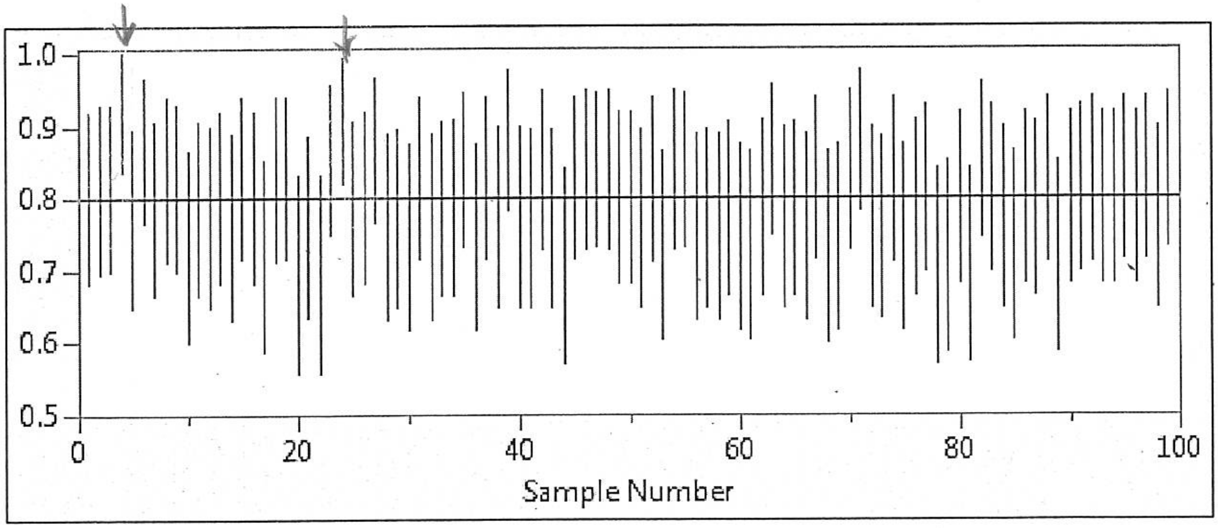
(C) 0.44 standard deviations below the mean

(D) 0.44 standard deviations above the mean

(E) 0.75 standard deviations above the mean

---

31. The graph below shows the results from constructing 100 different confidence intervals of the same confidence level, each based on a different sample of size $n = 75$ from a population in which the population proportion is $p = 0.8$. Each vertical line in the figure below spans the length of a different confidence interval.



Based on these results, what is the most likely value of the confidence level used in constructing these confidence intervals?

2 out of 100 fail to capture the population p

so about 98% confidence level

(A) 75%

(B) 80%

(C) 90%

(D) 99% (closest)

(E) 100%

32. A polling agency surveyed 1,018 randomly selected adults in the United States. Respondents were assigned at random to one of two different versions of a question asking them to estimate the size of Canada's population. Each version is shown below.

experiment

Version A The population of Australia is about 23 million. How many people do you think live in Canada? Version B The population of the U.S. is about 319 million. How many people do you think live in Canada?

The average response from those given version A was about 42 million and the average response from those given version B was about 95 million. The polling agency conducted a large sample test for the difference in two means and calculated a p-value of 0.0004. Assuming that the conditions for inference are met, which of the following is the most appropriate conclusion based on these results?

(A) There is convincing statistical evidence that the difference in the question wording makes no difference in how Americans would respond, on average.

(B) There is convincing statistical evidence that the difference in the question wording causes a difference in how Americans would respond, on average. yes, b/c this is an experiment

(C) There is not convincing statistical evidence to claim that the difference in the question wording causes a difference in how Americans would respond, on average.

(D) There is convincing statistical evidence of a difference in how Americans would respond, but we cannot say that the question wording is the cause of this difference.

(E) No conclusion can be made since there is such a large difference in the populations of Australia and the U.S.

33. A large population has a distribution that is fairly uniform with a mean of 80 and a standard deviation of 12. If all possible samples of size 36 are taken, which of the following best describes the distribution of sample means?

(A) The distribution will also be uniform with a mean of 80 and a standard deviation of 12.

(B) The distribution will be approximately normal with a mean of 80 and a standard deviation of 12.

(C) The distribution will be symmetric but not approximately normal, with a mean of 80 and a standard deviation of 12.

(D) The shape of the distribution will be unknown, but it will have a mean of 80 and a standard deviation of 2.

(E) The distribution will be approximately normal with a mean of 80 and a standard deviation of 2.

$\mu_{\bar{x}} = \mu = 80$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{36}} = 2$

$n > 25$ ~Normal by CLT

---

34. A farmer wants to see if there is a difference in average yield between two varieties of corn (variety A and variety B). In order to compare the two varieties, he prepares a two-acre plot on each of his seven different fields. On each field he plants variety A on a one acre plot and variety B on the one acre plot right next to it. Which acre plot gets which variety is determined at random. At harvest time, he measures the yield (in bushels) on each of the acre plots.

Field	1	2	3	4	5	6	7
Variety A	148.2	144.6	149.7	140.5	154.6	147.1	151.4
Variety B	141.5	141.0	144.0	141.2	153.0	141.7	146.8

Assuming that the conditions for inference are met, what statistical test should be used to determine if there is a significant difference in average yield between the two corn varieties?

(A) A chi-square goodness-of-fit test

(B) A chi-square test of independence

(C) A matched-pairs t-test for means

(D) A two-sample t-test for means

(E) A linear regression t-test

35. A 95% confidence interval for $\mu_A - \mu_B$ is given by $(-2.3, 4.5)$. What can we infer from this confidence interval?

(A) This confidence interval implies that the mean for sample A is less than the mean for sample B.

(B) This confidence interval implies that the mean for sample A is equal to the mean for sample B.

(C) With 95% confidence, there is convincing statistical evidence that the mean for population A is less than the mean for population B.

(D) With 95% confidence, there is convincing statistical evidence that the mean for population A is greater than the mean for population B.

(E) There is not convincing statistical evidence that the mean for population A is different from the mean for population B at the 5% significance level.

---

36. An instructor at a large university wanted to see if homework sets containing some review problems are more effective than homework sets containing only problems consisting of new material. She conducted a study by randomly assigning 156 student volunteers to the two homework groups (78 students to each group) and then compared the two groups' average scores on the final exam using a one-tailed, two-sample t-procedure. The p-value for the test was 0.015. The group who had homework sets containing review problems had a higher mean final exam score by 6.3 percentage points. Which of the following gives the best interpretation of the p-value?

(A) There is a 1.5% chance that there is no difference in effectiveness between the two styles of homework sets.

(B) There is a 1.5% chance that homework sets with review problems are more effective than homework sets with only problems over new material.

(C) There is a 1.5% chance of observing a sample mean final exam score for the review problems group at least 6.3 percentage points greater than the sample mean final exam score for the new material only problems group, if there is no difference in effectiveness between the two styles of homework sets.

(D) There is a 1.5% chance of observing a sample mean final exam score for the review problems group at least 6.3 percentage points greater than the sample mean final exam score for the new material only problems group, if homework sets with review problems are more effective than homework sets with all similar problems.

(E) There is a 1.5% chance that the population mean final exam score for all students doing homework sets with review problems is 6.3 percentage points higher than that for all students doing homework sets with only problems consisting of new material.

37. In a recent survey of 1,612 randomly selected U.S. adults, 66% of the men and 77% of the women reported that they use Facebook. Which of the following is the best justification for conducting a hypothesis test to analyze these results?

(A) A hypothesis test is always conducted in any research study.

(B) A hypothesis test is conducted to determine whether 66% and 77% are different.

(C) A hypothesis test is conducted to assess whether the population data are nearly normal.

(D) A hypothesis test is conducted to measure bias due to problems with question wording.

(E) A hypothesis test is conducted to quantify how likely it would be to get such sample results simply by random sampling, if there is no difference in Facebook use between men and women.

---

38. The career center at a large university surveyed a random sample of students who graduated the previous year. Of the 125 surveyed students who majored in engineering, 90 were able to find a job within six months of graduating. Of the 68 surveyed students who majored in business, 51 were able to find a job within six months of graduating. Which of the following expressions gives a 90% confidence interval for the difference in the proportion of engineering and business majors who could find jobs within six months of graduating?

$\hat{p}_E = \frac{90}{125} = .72$

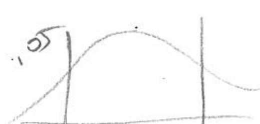
$\hat{p}_B = \frac{51}{68} = .75$

(A) $0.03 \pm 0.05\sqrt{\frac{(0.72)(0.28)}{125} + \frac{(0.75)(0.25)}{68}}$

(B) $0.03 \pm 1.645\sqrt{\frac{(0.72)(0.28)}{125} + \frac{(0.75)(0.25)}{68}}$

90%

~~(C)~~ $0.03 \pm 1.645\sqrt{\frac{(0.731)(0.269)}{125} + \frac{(0.731)(0.269)}{68}}$



(D) $0.03 \pm 1.96\sqrt{\frac{(0.72)(0.28)}{125} + \frac{(0.75)(0.25)}{68}}$

invnorm(.05, 0, 1) =
(-)1.6449
(1.65)

~~(E)~~ $0.03 \pm 1.96\sqrt{\frac{(0.731)(0.269)}{125} + \frac{(0.731)(0.269)}{68}}$

Use the following information to answer questions 39 – 40.
Students in an AP Statistics class wanted to determine if a person's hand span is associated with the number of candies he or she can pick up from a bowl of candy. For the fourteen students in the class, each student's hand span and the number of candies he or she was able to pick up was recorded. A least-squares analysis was conducted on the number of candies versus hand span. The conditions for inference were checked and deemed reasonable. The regression analysis output is given in the table below.

Predictor	Coef	SE Coef	T	P
Constant	-8.74	10.81	-0.81	0.434
Hand span	1.26	0.52	2.44	0.031
S = 3.34	R-Sq = 33.1%			

39. The estimate of the slope of the least-squares regression line using a 95% confidence interval is

(A) $1.26 \pm 2.18(0.52)$

(B) $1.26 \pm 2.44(0.52)$

(C) $1.26 \pm 2.18\left(\frac{0.52}{\sqrt{14}}\right)$

(D) $1.26 \pm 1.96(0.52)$

(E) $1.26 \pm 3.34\left(\frac{0.52}{\sqrt{14}}\right)$

$CI = b \pm t^* s_b$

$= 1.26 \pm (2.18)(.52)$

.95

.025 .025

$t^* = invT(.025, 12) = (-)2.1788$

$df = n-2$

40. A hypothesis test for the slope of the least-squares regression for number of candies versus hand span was conducted at the 5% significance level. Which of the following gives the correct decision for the hypothesis test and the possible error?

(A) Reject the null hypothesis; Type I error

(B) Reject the null hypothesis; Type II error

(C) Fail to reject the null hypothesis; Type I error

(D) Fail to reject the null hypothesis; Type II error

(E) Reject the null hypothesis; No error possible since the conditions for inference were met

p-value = .031, reject Ho

	T	F
R		I
NR	II	