ECML Statistics 2018

1. Suppose the average stopping distance of tires of Brand A is 55 feet with a standard deviation of 5.2 feet, and that the average stopping distance of tires of Brand B is 51 feet with a standard deviation of 4.9 feet. The stopping distances of tires of each brand are approximately Normally distributed. A braking test is conducted using 50 cars. Half of the cars are randomly assigned to use Brand A tires and the other half to use Brand B tires. Given a Normal model is appropriate for the sampling distribution of the difference between sample means, which of the following gives the probability that the experiment results in a sample mean stopping distance for Brand A that is less than the sample mean stopping distance for Brand B?

(A) $P\left(z < \frac{0-(55-51)}{\sqrt{\frac{5.2^2}{50}-\frac{4.9^2}{50}}}\right)$

(B) $P\left(z < \frac{(55-51)-0}{\sqrt{\frac{5.2^2}{25}+\frac{4.9^2}{25}}}\right)$

(C) $P\left(z < \frac{(55-51)-0}{\sqrt{\frac{5.2}{25}+\frac{4.9}{25}}}\right)$

(D) $P\left(z < \frac{0-(55-51)}{\sqrt{\frac{5.2^2}{25}+\frac{4.9^2}{25}}}\right)$

(E) $P\left(z < \frac{0-(55-51)}{\sqrt{\frac{5.2^2}{25}}+\sqrt{\frac{4.9^2}{25}}}\right)$

2. A factory has 20 assembly lines producing a popular toy. To inspect a representative sample of 100 toys, quality control staff randomly selected 5 toys from each assembly line's output. Was this design a simple random sample (SRS)?

(A) Yes, it was an SRS because the toys were selected at random.
(B) Yes, it was an SRS because each toy produced had an equal chance to be selected.
(C) Yes, it was an SRS because a stratified sample is a type of simple random sample.
(D) No, it was not an SRS because not all combinations of 100 toys could have been chosen.
(E) No, it was not an SRS because toys do not come off the assembly line at random.

3. If batting averages follow a bell-shaped distribution, arrange the following in ascending order:

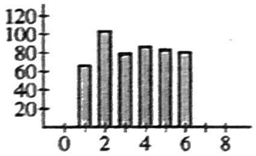
I. A batting average that has a z-score of 1.
II. A batting average at the 80th percentile of the distribution.
III. A batting average that marks the third quartile (Q_3) of the distribution.

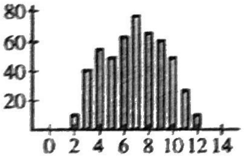
(A) I, II, III
(B) III, I, II
(C) II, I, III
(D) II, III, I
(E) III, II, I

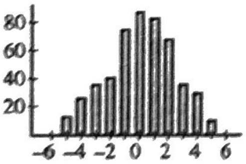
4. A machine that fills soda cans works so that the distribution of the amount of liquid in the cans follows a Normal model with a mean of 12.1 ounces. The label on the cans claims that they each contain 12 oz. of liquid. Management wants to assure that only 1% of cans are "under-filled", that is, contain less than the amount claimed on the label. With what standard deviation does the filling machine need to operate in order to achieve this goal?

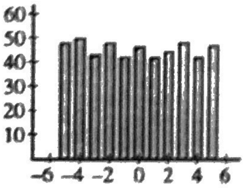
(A) 0.0324 oz.
(B) 0.0388 oz.
(C) 0.0429 oz.
(D) 0.0780 oz.
(E) 1.0000 oz.

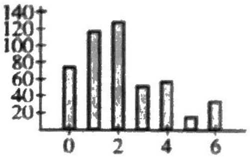
5. For a single roll of a fair die each of the outcomes (1, 2, 3, 4, 5, or 6) is equally likely. A red die and a green die are rolled simultaneously and the difference of the outcomes (red-green) is computed. This is repeated for a total of 500 rolls of the pair of dice. Which of the following graphs best represents the most reasonable distribution of the differences?

(A)


(B)


(C)


(D)


(E)


6. A manager wishes to determine the relationship between the number of miles (in hundreds of miles) his sales representatives travel per month and the amount of sales (in thousands of dollars) per month. He finds the average miles traveled among the representatives is 6.78 hundred miles with a standard deviation of 4.99 hundred miles per month. The average amount of sales is 57.22 thousand dollars with a standard deviation of 17.7 thousand dollars per month. The correlation between the two variables is 0.68. Create a linear model for predicting the amount of sales from the number of miles a sales representative travels per month.

(A) $\hat{y} = 40.87 + 2.412x$
(B) $\hat{y} = 2.412 + 40.87x$
(C) $\hat{y} = 55.92 + 0.1917x$
(D) $\hat{y} = -131.23 + 2.412x$
(E) $\hat{y} = 52.61 + 0.68x$

7. Consider the random experiment of selecting three attendees randomly at a football game in which Penn State is playing Notre Dame. Identify whether the attendee is a Penn State fan or a Notre Dame fan. Since the game is a home game, suppose that 80% of the fans attending the game are Penn State fans, while 20% are Notre Dame fans. Let the random variable X = the number of Penn State fans selected. What are the mean and standard deviation of the random variable X?

(A) $\mu_x = 0.6, \sigma_x = 0.69$
(B) $\mu_x = 0.6, \sigma_x = 0.48$
(C) $\mu_x = 2.4, \sigma_x = 0.23$
(D) $\mu_x = 2.4, \sigma_x = 0.48$
(E) $\mu_x = 2.4, \sigma_x = 0.69$

8. In a particular country, it is known that 40% of the residents have blue eyes, 35% of the residents have brown eyes, and 25% of the residents have green eyes. A student carries out a study to determine whether, in terms of color, the eyes of dolls manufactured in that country are representative of the residents of the country. The student takes a random sample of 40 brands of doll, and finds that 10 of them have blue eyes, 19 of the have brown eyes, and 11 of them have green eyes. She then carries out the appropriate significance test. Which of the following gives the correct degrees of freedom, test statistic, and p-value for the test?

(A) $df = 2, \chi^2 = 5.007, p - value = 0.082$
(B) $df = 2, \chi^2 = 4.136, p - value = 0.126$
(C) $df = 4, \chi^2 = 4.136, p - value = 0.388$
(D) $df = 4, \chi^2 = 5.007, p - value = 0.287$
(E) $df = 2, \chi^2 = 3.765, p - value = 0.152$

9. An automobile manufacturer tries two distinct assembly procedures. In a sample of 350 cars coming off the line using the first procedure there are 28 with defects, while a sample of 500 autos from the second line shows 32 with defects. The manufacturer decides to conduct a two-proportion z-test to determine if there is a significant difference in the proportion of defective autos between the two assembly procedures. Which of the following is the appropriate test statistic?

(A) $z = \frac{.08-.064}{\sqrt{\frac{.08^2}{350}+\frac{.064^2}{500}}}$

(B) $z = \frac{.08-.064}{\sqrt{\frac{(.144)(.856)}{850}+\frac{(.144)(.856)}{850}}}$

(C) $z = \frac{.08-.064}{\sqrt{\frac{(.08)(.92)}{350}\frac{(.064)(.936)}{500}}}$

(D) $z = \frac{.08-.064}{\sqrt{\frac{(.0706)(.9294)}{850}+\frac{(.0706)(.9294)}{850}}}$

(E) $z = \frac{.08-.064}{\sqrt{\frac{(.0706)(.9294)}{350}+\frac{(.0706)(.9294)}{500}}}$

10. A box contains 2 red balls and 1 white ball. Two balls are randomly chosen without replacement. If both balls are red, Player A wins \$5 from player B. Otherwise Player B wins \$2 from Player A. Which of the following is the best interpretation of the expected value of the gain for Player A?

(A) Neither player has the advantage because the expected value is 0.
(B) Player A has the advantage because the expected value is positive.
(C) Player B has the advantage because the expected value is positive.
(D) Player A has the advantage because the expected value is negative.
(E) Player B has the advantage because the expected value is negative.

11. In 1991, the Environmental Protection agency (EPA) adopted what is known as the Lead and Copper Rule, which defines drinking water as unsafe if the concentration of lead is 15 parts per billion (ppb) or greater. With μ denoting the mean concentration of lead, a water system monitoring lead levels might use lead level measurements from a sample of water specimens to test:

$$H_0\text{: } \mu \geq 15 \text{ ppb versus } H_a\text{: } \mu < 15 \text{ ppb}.$$

Which of the following is the best interpretation of a Type I error in this situation?

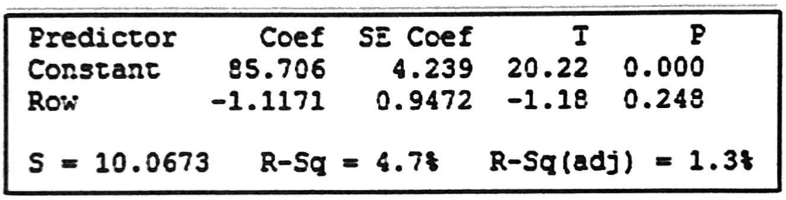
(A) The drinking water is determined to be unsafe when in actuality it is safe.
(B) The drinking water is determined to be safe when in actuality it is unsafe.
(C) The drinking water is determined to be unsafe when in actuality it is unsafe.
(D) The drinking water is determined to be safe when in actuality it is safe.
(E) A Type I error could not be committed in this situation.

12. We test the hypotheses that $H_0: p = .35$ versus $H_a: p > .35$. Suppose it is actually true that the true population proportion is 26%. With which sample size and level of significance will our test have the greatest power?

(A) $\alpha = 0.01, n = 250$
(B) $\alpha = 0.03, n = 250$
(C) $\alpha = 0.01, n = 400$
(D) $\alpha = 0.03, n = 400$
(E) The power of the test will always be the same regardless of sample size and α.

Use the following information to answer #13-14:

Many people believe that students learn better if they sit closer to the front of the classroom. To investigate, an AP Statistics teacher randomly assigned 30 students to seat locations in his classroom for a particular chapter and recorded the test score for each student at the end of the chapter. The explanatory variable in this experiment is which row the student was assigned (Row 1 is closest to the front and Row 7 is the farthest away). He then put the data into statistical analysis software with the results shown below:



```
Predictor       Coef   SE Coef       T       P
Constant      85.706     4.239   20.22   0.000
Row          -1.1171    0.9472   -1.18   0.248

S = 10.0673    R-Sq = 4.7%    R-Sq(adj) = 1.3%
```

13. One particular student in the study who sat in Row 5 earned a score of 90 on the chapter test. What is the value of this student's residual score according to the resulting least squares regression equation?

(A) 80.1
(B) -9.9
(C) 9.9
(D) 4.3
(E) -4.3

14. Which of the following gives the 95% confidence interval estimate for the true slope of the least squares regression line relating row number and test score?

(A) -1.1171 ± (1.180)*(0.9472)
(B) -1.1171 ± (1.701)*(0.9472)
(C) -1.1171 ± (2.048)*(0.9472)
(D) -1.1171 ± (1.180)*(10.0673)
(E) -1.1171 ± (1.960)*(10.0673)

15. For purposes of making budget plans for staffing, a college reviewed student's year in school and area of study. Of the students, 23% are seniors, 25% are juniors, 25% are sophomores, and the rest are freshmen. Also, 40% of the seniors major in the area of humanities, as did 39% of the juniors, 40% of the sophomores, and 36% of the freshmen. A student is selected at random. About what percentage of students who are humanities majors are upperclassmen (seniors or juniors)?

(A) 19%
(B) 31%
(C) 39%
(D) 49%
(E) 73%

16. A survey of families revealed that 59% of all families eat turkey at holiday means, 45% eat ham, and 12% have both turkey and ham to eat at holiday meals. What is the probability that a family selected at random had neither turkey nor ham at their holiday meal?

(A) 0.02
(B) 0.08
(C) 0.16
(D) 0.27
(E) 0.92

17. A randomized block design will be used in an experiment to compare two lotions that protect people from getting sunburned. Which of the following should guide the formation of the blocks?

(A) Participants in the same block should receive the same lotion.
(B) Participants should be randomly assigned to the blocks.
(C) Participants should be kept blind as to which block they are in.
(D) Participants in each block should be as similar as possible with respect to how easily they get sunburned.
(E) Participants in each block should be as different as possible with respect to how easily they get sunburned.

18. In auto racing, a pit crew claims that its mean pit stop time (for 4 new tires and fuel) is less than 13 seconds. Assume the standard deviation of all pit stop times is 0.2 seconds. In order to investigate this claim, a random sample of 36 pit stop times for this particular pit crew is taken, the sample mean is calculated, and a one-sample z-test for the mean is performed. To the nearest thousandth of a second, which of the following is the highest value of the sample mean possible that would still result in supporting the pit crew's claim at the 10% level of significance?

(A) 12.744 sec.
(B) 12.922 sec.
(C) 12.950 sec.
(D) 12.957 sec.
(E) 12.963 sec.

19. You are running a political campaign and wish to estimate, with 95% confidence, the population proportion of registered voters who will vote for your candidate. The candidate wishes your estimate to be accurate within 3% of the true population proportion. Which of the following is the minimum sample size of potential voters needed if no preliminary estimate for p is available?

(A) 11
(B) 251
(C) 752
(D) 1068
(E) 2135

20. The following gives the five-number-summary for a random sample of the ages (in years) of attendees at a concert:

Min = 12 $Q_1 = 19$ Med = 26 $Q_3 = 39$ Max = 72

Which of the following gives the minimum age of an attendee who would be considered a high outlier at this concert based on the given statistics?

(A) 56
(B) 59
(C) 69
(D) 72
(E) 102

Tie-Breaker: Carly commutes to work, and her commute time is dependent on the weather. When the weather is good, the distribution of her commute times is approximately normal with mean 18 minutes and standard deviation 1.5 minutes. When the weather is not good, the distribution of her commute times is approximately normal with mean 26 minutes and standard deviation 3 minutes. Suppose the probability that the weather will be good tomorrow is 0.8. Find the probability that Carly's commute time tomorrow will be greater than 22 minutes. Round to four decimal places.

ECML Statistics Answers 2018

1. D
2. D
3. E
4. C
5. C
6. A
7. E
8. B
9. E
10. B
11. B
12. D
13. C
14. C
15. D
16. B
17. D
18. D
19. D
20. D

Answer Key

ECML Statistics 2018

1. Suppose the average stopping distance of tires of Brand A is 55 feet with a standard deviation of 5.2 feet, and that the average stopping distance of tires of Brand B is 51 feet with a standard deviation of 4.9 feet. The stopping distances of tires of each brand are approximately Normally distributed. A braking test is conducted using 50 cars. Half of the cars are randomly assigned to use Brand A tires and the other half to use Brand B tires. Given a Normal model is appropriate for the sampling distribution of the difference between sample means, which of the following gives the probability that the experiment results in a sample mean stopping distance for Brand A that is less than the sample mean stopping distance for Brand B?

$$\sigma_{A-B} = \sqrt{\frac{5.2^2}{25} + \frac{4.9^2}{25}}$$

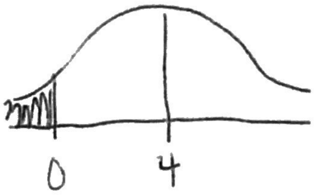
(A) $P\left(z < \frac{0-(55-51)}{\sqrt{\frac{5.2^2}{50} - \frac{4.9^2}{50}}}\right)$

(B) $P\left(z < \frac{(55-51)-0}{\sqrt{\frac{5.2^2}{25} + \frac{4.9^2}{25}}}\right)$

(C) $P\left(z < \frac{(55-51)-0}{\sqrt{\frac{5.2}{25} + \frac{4.9}{25}}}\right)$

[✓ (D)] $P\left(z < \frac{0-(55-51)}{\sqrt{\frac{5.2^2}{25} + \frac{4.9^2}{25}}}\right)$

(E) $P\left(z < \frac{0-(55-51)}{\sqrt{\frac{5.2^2}{25}} + \sqrt{\frac{4.9^2}{25}}}\right)$



2. A factory has 20 assembly lines producing a popular toy. To inspect a representative sample of 100 toys, quality control staff randomly selected 5 toys from each assembly line's output. Was this design a simple random sample (SRS)?

~~(A)~~ Yes, it was an SRS because the toys were selected at random.
~~(B)~~ Yes, it was an SRS because each toy produced had an equal chance to be selected.
~~(C)~~ Yes, it was an SRS because a stratified sample is a type of simple random sample.
[✓ (D)] No, it was not an SRS because not all combinations of 100 toys could have been chosen.
(E) No, it was not an SRS because toys do not come off the assembly line at random.

3. If batting averages follow a bell-shaped distribution, arrange the following in ascending order:

I. A batting average that has a z-score of 1. $z = 1$
II. A batting average at the 80th percentile of the distribution. $z = 0.84$
III. A batting average that marks the third quartile (Q_3) of the distribution. $z = 0.67$

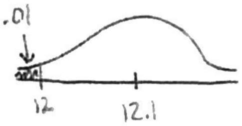
(A) I, II, III
(B) III, I, II
(C) II, I, III
(D) II, III, I
[✓ (E)] III, II, I

III, II, I

$\mu = 12.1$

4. A machine that fills soda cans works so that the distribution of the amount of liquid in the cans follows a Normal model with a mean of 12.1 ounces. The label on the cans claims that they each contain 12 oz. of liquid. Management wants to assure that only 1% of cans are "under-filled", that is, contain less than the amount claimed on the label. With what standard deviation does the filling machine need to operate in order to achieve this goal?

(A) 0.0324 oz.
(B) 0.0388 oz.
[✓ (C)] 0.0429 oz.
(D) 0.0780 oz.
(E) 1.0000 oz.



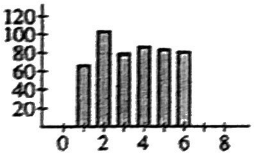
$$-2.326 = \frac{12-12.1}{\sigma}$$

$z^* = -2.326$

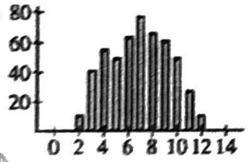
$\sigma = 0.43$ oz.

5. For a single roll of a fair die each of the outcomes (1, 2, 3, 4, 5, or 6) is equally likely. A red die and a green die are rolled simultaneously and the difference of the outcomes (red-green) is computed. This is repeated for a total of 500 rolls of the pair of dice. Which of the following graphs best represents the most reasonable distribution of the differences?

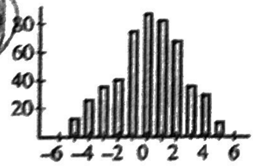
(A)



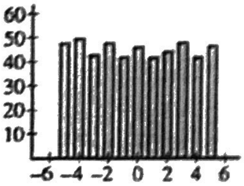
(B)



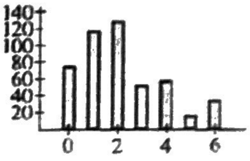
[✓ (C)]



(D)



(E)



Red

X_1	1	2	3	4	5	6
$P(x)$	1/6	1/6	1/6	1/6	1/6	1/6

Green

X_2	1	2	3	4	5	6
$P(x_2)$	1/6	1/6	1/6	1/6	1/6	1/6

Red-Green

$X_1 - X_2$	-5	-4	-3	-2	-1	0	1	2	3	4	5
$P(X_1 - X_2)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Red

Green

	1	2	3	4	5	6
	0	1	2	3	4	5
2	-1	0	1	2	3	4
3	-2	-1	0	1	2	3
4	-3	-2	-1	0	1	2
5	-4	-3	-2	-1	0	1

x

6. A manager wishes to determine the relationship between the number of miles (in hundreds of miles) his sales representatives travel per month and the amount of sales y (in thousands of dollars) per month. He finds the average miles traveled among the representatives is 6.78 hundred miles with a standard deviation of 4.99 hundred miles per month. The average amount of sales is 57.22 thousand dollars with a standard deviation of 17.7 thousand dollars per month. The correlation between the two variables is 0.68. Create a linear model for predicting the amount of sales from the number of miles a sales representative travels per month.

$\bar{x} = 6.78 \quad S_x = 4.99$

$\bar{y} = 57.22 \quad S_y = 17.7 \qquad r = 0.68$

[✓ (A)] $\hat{y} = 40.87 + 2.412x$
~~(B)~~ $\hat{y} = 2.412 + 40.87x$
~~(C)~~ $\hat{y} = 55.92 + 0.1917x$
(D) $\hat{y} = -131.23 + 2.412x$
~~(E)~~ $\hat{y} = 52.61 + 0.68x$

$b_1 = 0.68\left(\frac{17.7}{4.99}\right) \approx 2.412$

$57.22 = b_0 + 2.412(6.78)$

$b_0 = 40.867$

7. Consider the random experiment of selecting three attendees randomly at a football game in which Penn State is playing Notre Dame. Identify whether the attendee is a Penn State fan or a Notre Dame fan. Since the game is a home game, suppose that 80% of the fans attending the game are Penn State fans, while 20% are Notre Dame fans. Let the random variable X = the number of Penn State fans selected. What are the mean and standard deviation of the random variable X?

(A) $\mu_x = 0.6, \sigma_x = 0.69$
(B) $\mu_x = 0.6, \sigma_x = 0.48$
(C) $\mu_x = 2.4, \sigma_x = 0.23$
(D) $\mu_x = 2.4, \sigma_x = 0.48$
[✓ (E)] $\mu_x = 2.4, \sigma_x = 0.69$

$p = .80 \quad q = .20$

$\mu = 3(.8) = 2.4$

$\sigma = \sqrt{3(.8)(.2)} = 0.6928...$

8. In a particular country, it is known that 40% of the residents have blue eyes, 35% of the residents have brown eyes, and 25% of the residents have green eyes. A student carries out a study to determine whether, in terms of color, the eyes of dolls manufactured in that country are representative of the residents of the country. The student takes a random sample of 40 brands of doll, and finds that 10 of them have blue eyes, 19 of the have brown eyes, and 11 of them have green eyes. She then carries out the appropriate significance test. Which of the following gives the correct degrees of freedom, test statistic, and p-value for the test?

(A) $df = 2, \chi^2 = 5.007, p-value = 0.082$
[✓ (B)] $df = 2, \chi^2 = 4.136, p-value = 0.126$
(C) $df = 4, \chi^2 = 4.136, p-value = 0.388$
(D) $df = 4, \chi^2 = 5.007, p-value = 0.287$
(E) $df = 2, \chi^2 = 3.765, p-value = 0.152$

obs	exp
10	16
19	14
11	10

χ^2 GOF

$df = 2 \qquad \chi^2 = 4.1357$

.45

9. An automobile manufacturer tries two distinct assembly procedures. In a sample of 350 cars coming off the line using the first procedure there are 28 with defects, while a sample of 500 autos from the second line shows 32 with defects. The manufacturer decides to conduct a two-proportion z-test to determine if there is a significant difference in the proportion of defective autos between the two assembly procedures. Which of the following is the appropriate test statistic?

$\hat{p}_1 = \frac{28}{350} = .08$ $\hat{p}_2 = \frac{32}{500}$ $.064$

$\hat{p}$ pooled $= \frac{66}{850} = .07066$

(A) $z = \frac{.08-.064}{\sqrt{\frac{.08^2}{350}+\frac{.064^2}{500}}}$

(B) $z = \frac{.08-.064}{\sqrt{\frac{(.144)(.856)}{850}+\frac{(.144)(.856)}{850}}}$

(C) $z = \frac{.08-.064}{\sqrt{\frac{(.08)(.92)}{350}\ \frac{(.064)(.936)}{500}}}$

(D) $z = \frac{.08-.064}{\sqrt{\frac{(.0706)(.9294)}{850}+\frac{(.0706)(.9294)}{850}}}$

[✓ (E)] $z = \frac{.08-.064}{\sqrt{\frac{(.0706)(.9294)}{350}+\frac{(.0706)(.9294)}{500}}}$

10. A box contains 2 red balls and 1 white ball. Two balls are randomly chosen without replacement. If both balls are red, Player A wins $5 from player B. Otherwise Player B wins $2 from Player A. Which of the following is the best interpretation of the expected value of the gain for Player A?

(A) Neither player has the advantage because the expected value is 0.
[✓ (B)] Player A has the advantage because the expected value is positive.
(C) Player B has the advantage because the expected value is positive.
(D) Player A has the advantage because the expected value is negative.
(E) Player B has the advantage because the expected value is negative.

X	5	-2
P(x)	$\left(\frac{2}{3}\right)\left(\frac{1}{2}\right)$ $\frac{2}{6} = \boxed{\frac{1}{3}}$	$\boxed{2/3}$

$E(x) = \frac{1}{3}(5) + \left(\frac{2}{3}\right)(-2) = 0.33$

11. In 1991, the Environmental Protection agency (EPA) adopted what is known as the Lead and Copper Rule, which defines drinking water as unsafe if the concentration of lead is 15 parts per billion (ppb) or greater. With μ denoting the mean concentration of lead, a water system monitoring lead levels might use lead level measurements from a sample of water specimens to test:

H_0: $\mu \geq \overset{\text{unsafe}}{15\ ppb}$ versus H_a: $\mu < \overset{\text{safe}}{15\ ppb}$.

Which of the following is the best interpretation of a Type I error in this situation?

(A) The drinking water is determined to be unsafe when in actuality it is safe.
[✓ (B)] The drinking water is determined to be safe when in actuality it is unsafe.
(C) The drinking water is determined to be unsafe when in actuality it is unsafe.
(D) The drinking water is determined to be safe when in actuality it is safe.
(E) A Type I error could not be committed in this situation.

12. We test the hypotheses that $H_0: p = .35$ versus $H_a: p > .35$. Suppose it is actually true that the true population proportion is 26%. With which sample size and level of significance will our test have the greatest power?

Want to Reject Ho

Largest α & largest n

(A) $\alpha = 0.01, n = 250$
(B) $\alpha = 0.03, n = 250$
(C) $\alpha = 0.01, n = 400$
[✓ (D)] $\alpha = 0.03, n = 400$
(E) The power of the test will always be the same regardless of sample size and α.

Use the following information to answer #13-14:

Many people believe that students learn better if they sit closer to the front of the classroom. To investigate, an AP Statistics teacher randomly assigned 30 students to seat locations in his classroom for a particular chapter and recorded the test score for each student at the end of the chapter. The explanatory variable in this experiment is which row the student was assigned (Row 1 is closest to the front and Row 7 is the farthest away). He then put the data into statistical analysis software with the results shown below:

```
Predictor       Coef   SE Coef       T       P
Constant      85.706     4.239   20.22   0.000
Row          -1.1171    0.9472   -1.18   0.248

S = 10.0673    R-Sq = 4.7%    R-Sq(adj) = 1.3%
```

13. One particular student in the study who sat in Row 5 earned a score of 90 on the chapter test. What is the value of this student's residual score according to the resulting least squares regression equation?

(A) 80.1
(B) -9.9
[✓ (C) 9.9]
(D) 4.3
(E) -4.3

$\hat{y} = 85.706 - 1.1171x$

$\hat{y} = 85.706 - 1.1171(5)$

$\hat{y} = 80.1205$

$\text{Resid} = y - \hat{y} = 90 - 80.1205 = 9.8795$

14. Which of the following gives the 95% confidence interval estimate for the true slope of the least squares regression line relating row number and test score?

$n = 30 \quad df = 28$

$t^* = \text{invT}(.975, 28) = 2.048$

(A) -1.1171 ± (1.180)*(0.9472)
(B) -1.1171 ± (1.701)*(0.9472)
[✓ (C)] -1.1171 ± (2.048)*(0.9472)
(D) -1.1171 ± (1.180)*(10.0673)
(E) -1.1171 ± (1.960)*(10.0673)

15. For purposes of making budget plans for staffing, a college reviewed student's year in school and area of study. Of the students, 23% are seniors, 25% are juniors, 25% are sophomores, and the rest are freshmen. Also, 40% of the seniors major in the area of humanities, as did 39% of the juniors, 40% of the sophomores, and 36% of the freshmen. A student is selected at random. About what percentage of students who are humanities majors are upperclassmen (seniors or juniors)?

(A) 19%
(B) 31%
(C) 39%
[✓ (D) 49%]
(E) 73%

$$\frac{(.4 \cdot .23)+(.39 \cdot .25)}{(.4 \cdot .23)+(.39 \cdot .25)+(.4 \cdot .25)+(.36 \cdot .27)} = \frac{.1898}{.3867} = \boxed{.49}$$

16. A survey of families revealed that 59% of all families eat turkey at holiday means, 45% eat ham, and 12% have both turkey and ham to eat at holiday meals. What is the probability that a family selected at random had neither turkey nor ham at their holiday meal?

(A) 0.02
[✓ (B) 0.08]
(C) 0.16
(D) 0.27
(E) 0.92

$$1 - [P(\text{ham} \cup \text{turkey})]$$
$$1 - [.59 + .45 - .12]$$
$$1 - .92 = \boxed{.08}$$

17. A randomized block design will be used in an experiment to compare two lotions that protect people from getting sunburned. Which of the following should guide the formation of the blocks?

(A) Participants in the same block should receive the same lotion.
(B) Participants should be randomly assigned to the blocks.
(C) Participants should be kept blind as to which block they are in.
[✓ (D) Participants in each block should be as similar as possible with respect to how easily they get sunburned.]
(E) Participants in each block should be as different as possible with respect to how easily they get sunburned.

18. In auto racing, a pit crew claims that its mean pit stop time (for 4 new tires and fuel) is less than 13 seconds. Assume the standard deviation of all pit stop times is 0.2 seconds. In order to investigate this claim, a random sample of 36 pit stop times for this particular pit crew is taken, the sample mean is calculated, and a one-sample z-test for the mean is performed. To the nearest thousandth of a second, which of the following is the highest value of the sample mean possible that would still result in supporting the pit crew's claim at the 10% level of significance?

(A) 12.744 sec.
(B) 12.922 sec.
(C) 12.950 sec.
[✓ (D) 12.957 sec.]
(E) 12.963 sec.

$H_0: \mu \geq 13$ $\sigma = 0.2$
$H_a: \mu < 13$ $n = 36$

$$z = \frac{\bar{x} - \mu}{\sigma/} \qquad -1.282 = \frac{\bar{x} - 13}{0.2/\sqrt{36}}$$

$= 12.957$

$z^* = 1.96$

19. You are running a political campaign and wish to estimate, with 95% confidence, the population proportion of registered voters who will vote for your candidate. The candidate wishes your estimate to be accurate within 3% of the true population proportion. Which of the following is the minimum sample size of potential voters needed if no preliminary estimate for *p* is available?

~~(A) 11~~
(B) 251
(C) 752
[✓ (D) 1068]
(E) 2135

$$.03 = 1.96\sqrt{\frac{(.5)(.5)}{n}}$$

$n \geq 1068$

20. The following gives the five-number-summary for a random sample of the ages (in years) of attendees at a concert:

Min = 12 $\quad$ $Q_1 = 19$ $\quad$ Med = 26 $\quad$ $Q_3 = 39$ $\quad$ Max = 72

Which of the following gives the minimum age of an attendee who would be considered a high outlier at this concert based on the given statistics?

(A) 56
(B) 59
~~(C) 69~~
[✓ (D) 72]
(E) 102

$Q_3 + 1.5(IQR)$

$39 + 1.5(39-19)$

$39 + 30 = \underline{69}$ ← must be more than

Tie-Breaker: Carly commutes to work, and her commute time is dependent on the weather. When the weather is good, the distribution of her commute times is approximately normal with mean 18 minutes and standard deviation 1.5 minutes. When the weather is not good, the distribution of her commute times is approximately normal with mean 26 minutes and standard deviation 3 minutes. Suppose the probability that the weather will be good tomorrow is 0.8. Find the probability that Carly's commute time tomorrow will be greater than 22 minutes. Round to four decimal places.

$x \sim N(18, 1.5)$ $\qquad$ $y \sim N(26, 3)$

$(x > 22)$ $\qquad$ $P(y > 22)$

$z = \frac{22-18}{1.5} = \frac{4}{1.5} = \frac{8}{3}$ $\qquad$ $z = \frac{22-26}{3} = \frac{-4}{3}$

$P(z > -\frac{4}{3}) = .9088$

$P(z > 8/3) = .0038$

$(.8)(.0038) + (.2)(.9088) = \boxed{.1848}$