Statistics 151 Practice Final 1s

Instructions:
1. This is a closed book exam.
2. You may use the STAT 151 formula sheets and tables provided and a NON-PROGRAMMABLE calculator only.
3. You have 120 minutes to complete the exam.
4. The exam consists of 45 multiple choice questions worth 1 mark each.

---

1. Event A occurs with probability 0.2. Event B occurs with probability 0.8. If A and B are disjoint (mutually exclusive) then
a) *P*(A and B) = 0.16. **b) *P*(A or B) = 1.0**. **c)** *P*(A and B) = 1.0. **d)** *P*(A or B) = 0.16.

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) = 0.2 + 0.8 - 0 = 1.$$

2. A researcher states that the time it takes an adult to complete a certain paper and pencil maze is negatively associated with age. This means
a) above average values of age tend to accompany above average values of time to complete the maze.
b) below average values of age tend to accompany below average values of time to complete the maze.
c) below average values of age can be accompanied by either above or below average values of time to complete the maze.
d) above average values of age tend to accompany below average values of time to complete the maze.

Use the following to answer question 3:

A television station is interested in predicting whether voters in its listening area are in favour of federal funding for abortions. It asks its viewers to phone in and indicate whether they support or are opposed to this. Of the 2241 viewers who phoned in, 1574 (70.24%) were opposed to federal funding for abortions.

3. The sample is
a) all viewers listening to the show.
b) the 2241 viewers who phoned in.
c) the 1574 viewers who are opposed.
d) the 667 viewers who are in favour.

4. The one-sample *t* statistic from a sample of *n* = 19 observations for the two-sided test of
H_0: $\mu = 6$, H_a: $\mu \neq 6$
has the value *t* = 1.93. Based on this information
a) we would reject the null hypothesis at $\alpha = 0.10$.
b) $0.025 < P\text{-value} < 0.05$.
c) we would reject the null hypothesis at $\alpha = 0.05$.
d) All of the above.

$$p-value = 2P(t_{18} > 1.93) \in (0.05, 0.10).$$

Use the following to answer question 5:

In a large Midwestern university (the class of entering freshmen being on the order of 6000 or more students), an SRS of 100 entering freshmen in 1999 found that 20 finished in the bottom third of their high school class. Admission standards at the university were tightened in 2000. In 2001 an SRS of 100 entering freshmen found that 10 finished in the bottom third of their high school class. Let p_1 and p_2 be the proportion of all entering freshmen in 1999 and 2001, respectively, who graduated in the bottom third of their high school class.

5. A 99% confidence interval for $p_1 - p_2$ is
a) 0.10 ± .083. **b)** 0.10 ± .098. **c) 0.10 ± .129.** **d)** 0.10 ± .050.

$$C.I. \text{ for } \hat{p}_1 - \hat{p}_2 \text{ has the form: Estimate} \pm \{\text{C.V.} \times \text{S.E.(Estimate)}\}$$

$$C.V. = z^*_{0.005} = 2.575, \quad S.E.(\hat{p}_1 - \hat{p}_2) = 0.05$$

The margin of error is thus 2.575(0.05)=0.12875.

Use the following to answer questions 6-8:

A researcher is studying treatments for agoraphobia with panic disorder. The treatments are to be the drug Imipramine at the doses 1.5 mg per kg of body weight and 2.5 mg per kg of body weight. There will also be a control group given a placebo. Thirty patients were randomly divided into three groups of 10 each. One group was assigned to the control, and the other two groups were assigned to the two treatments. After 24 weeks on treatment, each of the subject's symptoms were evaluated through a battery of psychological tests, where high scores indicate a lessening of symptoms. Assume the data for the three groups are independent, and the data are approximately normal. The means and standard deviations of the test scores for the three groups are given below.

Mean test score	Std. dev. in score	Group
75.70000	12.605554	Control
84.10000	18.441800	Dose = 1.5
102.40000	20.823064	Dose = 2.5

An ANOVA *F* test was run on the data. Below are a portion of the results.

Source	df	Sums of squares	Mean square	*F*-statistic
Treatment	2	3727.8	1863.9	6.00
Error	27		310.87	
Total	29			

6. The denominator degrees of freedom for the ANOVA *F* test are
a) 29. **b)** 2. **c)** 3. **d) 27**.
n-g=30-3=27.

7. The pooled standard deviation is
a) 17.63. **b)** 310.87. **c)** 8393.5. **d)** 17.29.

$$\sqrt{MSE} = \sqrt{310.87} = 17.63$$

8. The value of the ANOVA *F* statistic is
a) 6.00. **b)** 12.00. **c)** 0.44. **d)** less than 0.01.

$$F_0^* = \frac{MST}{MSE} = \frac{1863.9}{310.87} \approx 6.00$$

Use the following to answer question 9:

An old saying in golf is "You drive for show and you putt for dough." The point is that good putting is more important than long driving for shooting low scores and hence winning money. To see if this is the case, data on the top 69 money winners on the PGA tour in 1993 are examined. The average number of putts per hole for each player is used to predict their total winnings using the simple linear regression model

Average 1993 winnings = $\alpha + \beta$(average number of putts per hole)

This model was fit to the data using the method of least squares. The following results were obtained from statistical software.

$R^2 = 0.081$
$s =$ 281,777

Variable	Parameter Estimate	Std. Err. of Parameter Est.
Constant	7,897,179	3,023,782
Avg. Puts	–4,139,198	1,698,371

9. The correlation between 1993 winnings and average number of putts per hole is
a) –0.285. **b)** 0.285. **c)** 0.081. **d)** –0.081.

$$|r| = \sqrt{R^2} = 0.285.$$ Since the slope is negative, correlation is negative.

Use the following to answer question 10:

An agricultural researcher plants 100 plots with a new variety of corn. The average yield for these plots is $\overline{x}$ = 150 bushels per acre. Assume that the yield per acre for the new variety of corn follows a normal distribution with unknown mean μ and standard deviation $\sigma = 20$ bushels per acre. A 90% C.I. for μ is (146.71, 153.29).

10. Which of the following would produce a confidence interval with a smaller margin of error than the 90% confidence interval given above?
a) Plant 500 plots rather than 100.
b) Compute a 99% confidence interval rather than a 90% confidence interval. The increase in confidence indicates that we have a better interval.
c) Plant only 25 plots rather than 100 because 25 are easier to manage and control.
d) None of the above.

Increasing sample size decreases sampling variability.

11. In their advertisements, the manufacturers of a diet pill would like to claim that taken daily, their pill will produce a mean weight loss of **more than 10** pounds in one month. In order to determine if this is a valid claim, they hire an independent testing agency, which then selects 25 people to take the pill daily for a month. The agency should be testing the null hypothesis H_0: $\mu = 10$ and the alternative hypothesis

a) H_a: $\mu \neq 10 \pm \frac{\sigma}{\sqrt{n}}$. **b)** H_a: $\mu \neq 10$. **c)** H_a: $\mu < 10$. **d) H_a: $\mu > 10$.**

Use the following to answer question 12:

In a particular game, a fair die is tossed. If the number of spots showing is six you win \$6, if the number of spots showing is five you win \$3, and if the number of spots showing is four you win \$1. If the number of spots showing is one, two, or three you win nothing. You are going to play the game twice.

12. The probability that you will win over \$10 in total on the two plays of the game is
a) 1/4. **b) 1/36**. **c)** 1/6. **d)** 1/3.

P(win over \$10)= P(roll 6 both rolls)

13. If the knowledge that an event A has occurred implies that a second event B cannot occur, the events A and B are said to be
a) disjoint. **b)** the sample space. **c)** independent. **d)** collectively exhaustive.

Use the following to answer question 14:

Suppose we wish to predict the mean profits (in hundreds of thousands of dollars) for all companies that had sales (in hundreds of thousands of dollars) of 500. We use statistical software to do the prediction and obtain the following output.

Sales	Predict	Stdev.Mean Predict	95% C.I.	95% P.I.
500	–130.4	59.3	(–248.5, –12.3)	(–1066.4, 805.6)

14. Suppose we wish to predict the profits (in hundreds of thousands of dollars) for a company that had sales (in hundreds of thousands of dollars) of 500. A 95% interval for this prediction is
a) (–248.5, –12.3). **b)** 500 ± 59.3. **c) (–1066.4, 805.6).** **d)** –130.4 ± 59.3.

The interval is a prediction interval for a single company.

15. A statistic is said to be unbiased if
a) it is used for only honest purposes.
b) both the person who calculated the statistic and the subjects whose responses make up the statistic were truthful.
c) the survey used to obtain the statistic was designed so as to avoid even the hint of racial or sexual prejudice.
d) the mean of its sampling distribution is equal to the true value of the parameter being estimated.

Use the following to answer question 16:

A sample was taken of the verbal GRE scores of 20 applicants to graduate school at a large Midwestern university. Below are the scores. For convenience, the data are ordered.

280	310	340	350	370	410	420	420	420	470
490	510	520	520	600	610	670	720	750	770

16. The third quartile for the applicant scores is
a) 480. **b)** 640. **c)** 340. **d) 605.**

$$Q_3 = \frac{600+610}{2} = 605.$$

Use the following to answer question 17:

Using data on the appraised value of homes, a real estate agent computes the least-squares regression line for predicting a home's value in 2002 from its value in 1992. The equation of the least-squares regression line is
y = \$22,000 + 1.6x
where y represents a home's value in 2002 and x is the value in 1992.

17. Suppose Joe owns a home that was worth \$100,000 in 1992. What would be the predicted value of his home in 2002?
a) \$160,000.
b) \$122,000.
c) cannot be determined from the information given. We also need to know the correlation.
d) \$182,000.

$$\hat{y} = 22000 + 1.6(100000) = 182000$$

18. The heights (in inches) of males in the United States are believed to be normally distributed with mean μ. The average height of a random sample of 50 American adult males is found to be $\bar{x}$ = 69.72 inches, and the standard deviation of the 50 heights is found to be s = 4.23. The standard error of $\bar{x}$ is

a) 0.731. **b) 0.598**. **c)** 0.084. **d)** 0.357.

$$S.E.(\bar{X}) = s/\sqrt{n} = 0.598$$

Use the following to answer question 19:

An SRS of 100 flights of a large airline (call this airline 1) showed that 64 were on time. An SRS of 100 flights of another large airline (call this airline 2) showed that 80 were on time. Let p_1 and p_2 be the proportion of all flights that are on time for these two airlines.

19. Is there evidence of a difference in the on-time rate for the two airlines? To determine this, you test the hypotheses

H_0: $p_1 = p_2$, H_a: $p_1 \neq p_2$

The P-value of your test of the hypotheses given is

a) between .01 and .001.
b) between .10 and .05.
c) below .001.
d) between .05 and .01.

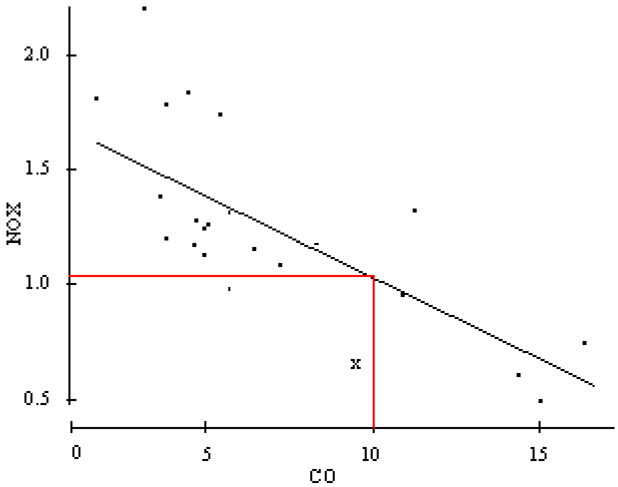
$$z_0^* = \frac{Estimate - Hypothesized\ value}{S.E._0(Estimate)} = \frac{\hat{p}_1 - \hat{p}_2}{S.E._0(\hat{p}_1 - \hat{p}_2)} = \frac{-0.16}{0.0635} = -2.52.$$

$$p - value = 2P(Z > 2.52) = 2(0.0059) = 0.0118.$$

Use the following to answer question 20:
In a recent article it was reported that in a random sample of children in grades 2–4, a significant *negative* relationship was found between the amount of time watching television and test scores.

20. This is an example of
a) an observational study.
b) the establishing of a causal relationship through correlation.
c) a block design, with grades as blocks.
d) an experiment.

21. Consider the following scatterplot of amounts of CO (carbon monoxide) and NOX (nitrogen oxide) in grams per mile driven in the exhausts of cars. The least-squares regression line has been drawn in the plot. The least-squares line would predict that a car that emits 10 grams of CO per mile driven would emit how many grams of NOX per mile driven?



A) 10.0.
B) 1.7.
C) 2.2.
D) 1.1.

22. A study found a correlation of $r = -0.61$ between the gender of a worker and his or her income. You may correctly conclude

a) women earn less than men on the average.
b) an arithmetic mistake was made. Correlation must be positive.
c) this is incorrect because *r* makes no sense here.
d) women earn more than men on the average.

Linear correlation, r, can only be calculated between two numerical variables. Gender is not a numerical variable.

23. Birth weights at a local hospital have a normal distribution with a mean of 110 oz. and a standard deviation of 15 oz. The proportion of infants with birth weights above 125 oz. is

a) 0.841. **b)** 0.341. **c)** 0.500. **d) 0.159.**

$$P(X > 125) = P(Z > 1) = 0.1587$$

Use the following to answer question 24:

An SRS of 100 labourers who use the services of a national temporary employment agency found that in the past year the average number of days worked by these labourers was $\bar{x}$ = 107 days, with standard deviation s = 45 days. Assume the distribution of the number of days worked in the population of labourers using this employment agency is approximately normal, with mean μ. Are these data evidence that μ has lowered from the value of 120 days of 5 years ago? To determine this, we test the hypotheses

H_0: $\mu = 120$, H_a: $\mu < 120$

24. Based on the data, the value of the one-sample *t* statistic is

a) -2.89. **b)** -2.38. **c)** -3.41. **d)** -2.67.

$$t_0^* = \frac{Estimate - Hypothesized\ \ value}{S.E.(Estimate)} = \frac{107-120}{45/\sqrt{100}} = -2.89$$

25. There are three children in a room—ages 3, 4, and 5. If a four-year-old child enters the room, the

- **a)** mean age and variance will stay the same.
- **b)** mean age and variance will increase.
- **c) mean age will stay the same but the variance will decrease.**
- **d)** mean age will stay the same but the variance will increase.

26. A 99% confidence interval for the mean μ of a population is computed from a random sample and found to be 6 ± 3. We may conclude

- **a) if we took many, many additional random samples and from each computed a 99% confidence interval for μ, approximately 99% of these intervals would contain μ.**
- **b)** there is a 99% probability that μ is between 3 and 9.
- **c)** there is a 99% probability that the true mean is 6 and there is a 99% chance that the true margin of error is 3.
- **d)** all of the above.

Use the following to answer question 27:

A random sample of 79 companies from the Forbes 500 list (which actually consists of nearly 800 companies) was selected, and the relationship between sales (in hundreds of thousands of dollars) and profits (in hundreds of thousands of dollars) was investigated by regression. The following simple linear regression model was used

Average Profits = $\alpha + \beta$(Sales)

This model was fit to the data using the method of least squares. The following results were obtained from statistical software.

$R^2 = 0.662$
$s = 466.2$

Variable	Parameter Est.	Std. Err. of Parameter Est.
Constant	–176.644	61.16
Sales	0.092498	0.0075

27. A 90% confidence interval for the slope β in the simple linear regression model is (approximately)
a) –0.09 ± 0.0075. **b)** –0.09 ± 0.012. **c) 0.09 ± 0.012.** **d)** 0.09 ± 0.0075.

$C.I.$ for β has the form: Estimate $\pm\{\text{C.V.}\times\text{S.E.(Estimate)}\}$

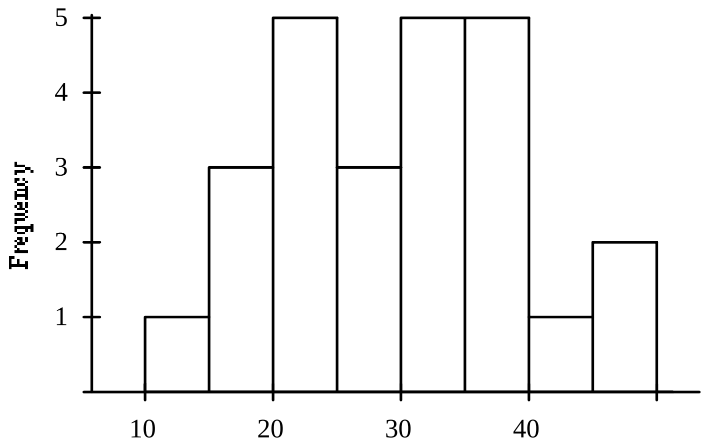
$C.V. = t^*_{n-2,\alpha/2} = t^*_{77,0.05} \approx 1.664$ or 1.671 (from t-tables)

$Estimate = b = 0.09,\ S.E.(b) = 0.0075$

$\Rightarrow 0.09 \pm 0.012$

Comment: (a) and (b) can be eliminated just by looking at the estimate. (d) can be eliminated as well because the standard error is not equal to the margin of error.

Use the following to answer question 28:
The following histogram represents the distribution of acceptance rates (percent accepted) among 25 business schools in 1995.



28. What percentage of the schools has an acceptance rate of fewer than 15%?
a) 16% **b)** 1% **c) 4%** **d)** 12%

1 out of 25!

Use the following to answer question 29:
A company runs a three-day workshop on strategies for working effectively in teams. On each day a different strategy is presented. Forty-eight employees of the company attend the workshop. At the outset, all 48 are divided into 12 teams of four. The teams remain the same for the entire workshop. Strategies are presented in the morning. In the afternoon, the teams are presented with a series of small tasks, and the number completed successfully using the strategy taught that morning is recorded for each team. The mean number of tasks completed successfully by all teams each day and the standard deviation is computed. The results are given below.

Day (strategy)	Means	Std.Dev.
1	17.25	7.10
2	17.64	14.14
3	17.25	14.03

The researchers did an ANOVA *F* test of the data and obtained the following results.

Source	Sums of squares	Mean square	*F*-ratio
Day	1.36	0.68	0.0045
Error	5321.71	152.05	
Total	5323.08		

29. The *P*-value of the ANOVA *F* test is
a) below 0.01.
b) larger than 0.10.
c) between 0.10 and 0.05.
d) between 0.05 and 0.01.

Note: The F-distribution is a strictly positive distribution. The probability that it takes on a value greater than 0.0045 is virtually 1.

Use the following to answer questions 30-31:

A food company is developing a new breakfast drink, and their market analysts are currently working on preliminary taste-testing studies. To help with their marketing strategy, they were first interested in whether preference for the new product was related to a person's gender. There were 100 male and 100 female volunteers available for the taste test. Both the males and females tasted the product and rated the flavour on a scale of 1 to 10, 1 being "very unpleasant" and 10 being "very pleasant." The mean rating for males was $\overline{x}_1$ = 6.4, with a standard deviation s_1 = 1.5. The mean rating for females was $\overline{x}_2$ = 7.0, with a standard deviation s_2 = 2.0. Let μ_1 and μ_2 represent the mean ratings we would observe for the populations of males and females, respectively, and assume our samples can be regarded as samples from these populations.

30. Suppose the researcher had wished to test the hypotheses
H_0: $\mu_1 = \mu_2$, H_a: $\mu_1 \neq \mu_2$
The P-value for the test is (use the conservative value for the degrees of freedom)

a) **between 0.05 and 0.01.**
b) below 0.01.
c) between 0.10 and 0.05.
d) larger than 0.10.

$$t_0^* = \frac{Estimate - Hypothesized\ value}{S.E.(Estimate)} = \frac{-0.6}{0.25} = -2.40$$

Using the conservative $df = \min(n_1 - 1, n_2 - 1) = 99$,

$$p - value = 2P(t_{99} > 2.4) \in (0.01, 0.02)$$

31. A 90% confidence interval for $\mu_1 - \mu_2$ is (use the conservative value for the degrees of freedom)
a) –0.6 ± 0.32. **b)** –0.6 ± 0.5. **c)** –0.6 ± 0.25. **d)** **–0.6 ± 0.42**.

$C.I.$ for $\mu_1 - \mu_2$ has the form: Estimate $\pm$ {C.V. $\times$ S.E.(Estimate)}

$$C.V. = t^*_{99,0.05} \approx 1.660 \text{ or } 1.664 \text{ (from t-tables)}$$

$$\Rightarrow -0.6 \pm 0.42.$$

Use the following to answer question 32:

The distribution of actual weights of 8 oz. wedges of cheddar cheese produced by a certain company is normal with mean 8.1 oz. and standard deviation 0.1 oz.

32. There is only a 5% chance that the average weight of the sample of five of the cheese wedges will be below
a) 8.20 oz. **b)** 7.98 oz. **c)** 7.94 oz. **d)** **8.03 oz.**

$$X \sim N(\mu_X = 8.1, \sigma_X = 0.1)$$

$$\overline{X} \sim N(\mu_{\overline{X}} = 8.1, \sigma_{\overline{X}} = \sigma_X / \sqrt{n} = 0.04472)$$

The question asks, what is 'a' so that $P(\overline{X} < a) = 0.05$?

From N(0,1) table, we know that $P(Z < -1.645) = 0.05$.

$$-1.645 = \frac{a - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{a - 8.1}{0.04472}.$$

$$\Rightarrow a = 8.03$$

33. The scores on a university examination are normally distributed with a mean of 62 and a standard deviation of 11. If the bottom 2.5% of students will fail the course, what is the lowest mark that a student can have and still be awarded a passing grade?
a) 57 **b) 41** **c)** 44 **d)** 62

$X \sim N(\mu_X = 62, \sigma_X = 11)$

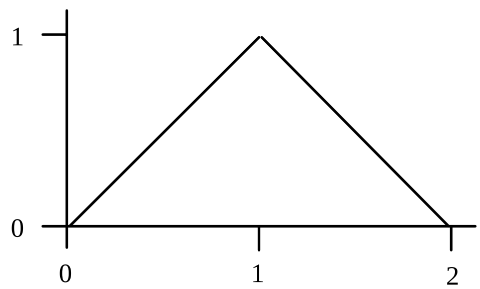
The question asks, what is 'a' so that $P(X < a) = 0.025$?

From N(0,1) table, we know that $P(Z < -1.96) = 0.05$.

$$-1.96 = \frac{a - \mu_X}{\sigma_X} = \frac{a - 62}{11}.$$

$\Rightarrow a = 40.44$

34. Suppose X is a continuous random variable taking values between 0 and 2 and having the probability density function below.



$P(X = 1)$ is
A) 0.00.
B) 0.25.
C) 0.50.
D) 1.00.

Use the following to answer question 35:

A newspaper conducted a state-wide survey concerning a proposal to raise taxes in order to prevent budget cuts to education. The newspaper took a random sample (assume it is an SRS) of 1200 registered voters and found that 580 would vote to raise taxes. Let p represent the proportion of registered voters in the state that would vote to raise taxes.

35. A 98% confidence interval for p is
a) 0.483 ± 0.249. **b) 0.483 ± 0.034**. **c)** 0.483 ± 0.014. **d)** 0.483 ± 0.024.

$C.I.$ for p has the form: Estimate $\pm$ {C.V. $\times$ S.E.(Estimate)}

$C.V. = z^*_{0.01} \approx 2.33$ (from z-table)

$Estimate = \hat{p} = 0.4833$. $S.E.(\hat{p}) = 0.0144$

$\Rightarrow 0.4833 \pm 0.034.$

Use the following to answer question 36:

A radio talk show host with a large audience is interested in the proportion $\hat{p}$ of adults in his listening area that think the drinking age should be lowered to 18. To find this out he poses the following question to his listeners: “Do you think that the drinking age should be reduced to 18 in light of the fact that 18-year-olds are eligible for military service?” He asks listeners to phone in and vote “yes” if they agree the drinking age should be lowered and “no” if they do not.

36. You are told that the proportion $\hat{p}$ of those who phoned in and answered yes is $\hat{p}$ = 0.70 and the standard error $SE_{\hat{p}}$ of the proportion is 0.0459. The number of people who phoned in

a) cannot be determined from the information given. **b)** is 50. **c) is 100**. **d)** is 200.

$$S.E.(\hat{p}) = 0.0459 = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.7(0.3)}{n}}$$

$$\Rightarrow n = \left(\frac{0.7(0.3)}{0.0459^2}\right) = 99.67 \Rightarrow 100$$

37. An event *A* will occur with probability 0.5. *A*n event *B* will occur with probability 0.6. The probability that both *A* and *B* will occur is 0.1. The conditional probability of *A* given *B* is

a) 1/6. **b)** cannot be determined from the information given. **c)** 0.2. **d)** 0.3.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.1}{0.6} = 1/6$$

38. How large a sample *n* would you need to estimate *p* with margin of error 0.01 with 95% confidence? Use the guess *p* = 0.5 as the value for *p*.

A) *n* = 677.
B) *n* = 9604.
C) *n* = 14406.
D) *n* = 19208

$$n = \left(\frac{z^*_{\alpha/2}}{m}\right)^2 \times 0.25 = \left(\frac{1.96}{0.01}\right)^2 \times 0.25 = 9604$$

Use the following to answer question 39:

A SRS of 20 third grade children is selected in Chicago and each is given a test to measure his or her reading ability. In the sample, the mean score is 64 points and the standard deviation is 12 points.

39. A 99% confidence interval for the population mean score based on these data is

a) 64 ± 4.64 points. **b)** 64 ± 5.62 points. **c)** 64 ± 6.84 points. **d) 64 ± 7.68 points**.

$C.I.$ for μ has the form: Estimate ± {C.V. × S.E.(Estimate)}

$C.V. = t^*_{n-1,\alpha/2} = t^*_{19,0.005} = 2.861$ (from t-table)

$Estimate = \bar{x} = 64.\ S.E.(\bar{X}) = s/\sqrt{n} = 2.6833$

$\Rightarrow 64 \pm 7.68.$

40. The average salary of all female workers is \$35,000. The average salary of all male workers is \$41,000. What must be true about the average salary of all workers?

a) It must be larger than \$38,000.
b) It could be any number between \$35,000 and \$41,000.
c) It must be \$38,000.
d) It must be larger than the median salary.

Use the following to answer question 41:
A noted psychic was tested for ESP. The psychic was presented with 200 cards face down and asked to determine if the card was one of five symbols: a star, cross, circle, square, or three wavy lines. The psychic was correct in 50 cases. Let *p* represent the probability that the psychic correctly identifies the symbol on the card in a random trial.

41. Suppose you wished to see if there is evidence that the psychic is doing better than just guessing. To do this you test the hypotheses

$H_0: p = 0.20$, $H_a: p > 0.20$

The *P*-value of your test is

a) between .01 and .001.
b) between .10 and .05.
c) below .001.
d) between .05 and .01.

$$z_0^* = \frac{Estimate - Hypothesized\ value}{S.E._0(Estimate)} = \frac{\hat{p} - 0.2}{S.E._0(\hat{p})} = \frac{0.05}{0.02828} = 1.77.$$

$$p - value = P(Z > 1.77) = 0.0384.$$

Use the following to answer question 42:
For a Physics course containing 10 students, the maximum point total for the quarter was 200. The point totals for the 10 students are given in the stem plot below.

11	6 8
12	1 4 8
13	3 7
14	2 6
15	
16	
17	9

42. The lowest score in the class as a percentage of the total possible points is
a) 116%. **b)** 90%. **c) 58%.** **d)** 75%.

➔116/200

Use the following to answer question 43:
An automobile insurer has found that repair claims have a mean of \$920 and a standard deviation of \$870. Suppose that the next 100 claims can be regarded as a random sample from the long-run claims process.

43. The probability that the average $\overline{X}$ of the 100 claims is larger than \$1000 is
a) 0.0800. **b) 0.1788**. **c)** 0.8212. **d)** 0.9200.

$$\overline{X} \approx N(\mu_{\overline{X}} = 920, \sigma_{\overline{X}} = \sigma_X / \sqrt{n} = 87)$$

$$P(\overline{X} > 1000) = P(Z > 0.92) = 0.1788$$

44. In a statistical test of hypotheses, we say the data are statistically significant at level α if
a) the *P*-value is larger than α. **b)** $\alpha = 0.05$. **c)** α is small. **d) the *P*-value is less than α.**

45. Twelve runners are asked to run a 10-kilometer race on each of two consecutive weeks. In one of the races the runners wear one brand of shoe and in the other a second brand. The brand they wear in each race is determined at random. All runners are timed and are asked to run their best in each race. The results (in minutes) are given below.

Runner	Brand 1	Brand 2
1	31.23	32.02
2	29.33	28.98
3	30.50	30.63
4	32.20	32.67
5	33.08	32.95
6	31.52	31.53
7	30.68	30.83
8	31.05	31.10
9	33.00	33.12
10	29.67	29.50
11	30.55	30.57
12	32.12	32.20

To determine if there is evidence that the mean time using brand 1 is less than the mean time using brand 2, we would use

a) the two-sample *t* test.
b) the one-sample *t* test.
c) the matched pairs *t* test.
d) any of the above are valid. It is at the experimenter's discretion.