ANSWERS AND EXPLANATIONS

Practice Examination 1

Section I: Multiple Choice Solutions

1. B
   IQR = Q3 − Q1. Using the 1.5 IQR Rule:
   Q1 − 1.5 IQR = 119.5 − 1.5(16) = 95.5, which is less than the minimum value *and* Q + 1.5 IQR = 135.5 + 1.5(16) = 159.5, which is less than the maximum value. Therefore, there is at least one outlier, the maximum value.

2. D
   $P(A) + P(B) = 0.69$; $P(A \cup B) = 0.68$; $0.69 \neq 0.68$ and thus the two events are not mutually exclusive.
   $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.69 - P(A \cap B) = 0.68$; therefore, $P(A \cap B) = 0.01$.
   The two events are not independent. If they were, then either $P(B|A) = P(B)$ *or* $P(A)P(B) = P(A \cap B)$ would have to be true. But,
   $P(B|A) = \dfrac{P(A \cap B)}{(A)} = \dfrac{0.01}{0.43} = 0.023$ which $\neq P(B)$, 0.26. And
   $P(A)P(B) \neq P(A \cap B)$ since $0.11118 \neq 0.01$.

3. C
   The current belief is that the average family income is $45,000. Thus the null hypothesis should set μ equal to 45,000. The first-time home buyers are trying to show that the average family income is less than $45,000; therefore, the alternative hypothesis should set μ less than 45,000.

4. A
   With a mean of 14.1 and a standard deviation of 0.04, the curve should have a peak at 14.1, and the change in curvature should occur 0.04 unit on either side of 14.1. *At least 14 oz* means 14 oz or more. Therefore, the shading should be to the right of 14.

5. D
   The upper quartile of set *Y* is equivalent to the median of set *X*. Therefore, approximately 50% of the data values in set *X* are greater than approximately 75% of the data values in set *Y*.

6. C
   The data are blocked by grocery item; therefore, the matched-pairs design is appropriate for this test.

7. B
   Since 20% of the subscribers also subscribe to broadband, 20%, or two of the ten digits, should represent broadband subscribers.

8. C
The profit for each shirt is \$10 − \$4 = \$6. The expected number of shirts sold can be calculated as follows:
$0(0.02) + 1(0.15) + 2(0.18) + 3(0.21) + 4(0.14) + 5(0.08)$
$+ 6(0.08) + 7(0.04) + 8(0.03) + 9(0.02) + 10(0.05) = 3.78.$
Therefore, the expected profit is $(3.78)(\$6) = \$22.68$.

9. A
For small sample sizes, the shape of the distribution will mimic that of the population. The mean of the sampling distribution will equal that of the population, and the standard deviation of the sampling distribution will equal that of the population divided by the square root of 5 in this case.

10. D
For independent variables,
$\mu_{A+B} = \mu_A + \mu_B = 65 + 154 = 219$ and
$\sigma_{A+B} = \sqrt{\sigma_A^2 + \sigma_B^2} = \sqrt{5.75^2 + 8.02^2} \approx 9.87.$

11. E
Cause-and-effect relationships can only be determined by a controlled randomized experiment.

12. B
All least squares regression lines are in the form $\hat{y} = a + bx$, where a is the constant (constant coefficient = 21.84) and b is the coefficient of the explanatory variable (weight coefficient = 0.037).

13. D
There are a fixed number of observations in a binomial setting, not a geometric setting.

14. B
The formula for the construction of the confidence interval is
$$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}},$$
where $\hat{p}_1 = \frac{66}{200}$, $\hat{p}_2 = \frac{12}{140}$, $n_1 = 200$, $n_2 = 140$, and $z^* = 1.96$.

15. E
If each of the values is multiplied by 9, then the mean and standard deviation will also be multiplied by 9.

16. B
For the given equation, the intercept is 382.1, and the slope is −12.25. The intercept yields the predicted value for the dependent variable when the independent variable assumes a value of 0. Since the slope is negative, the slope will give the decrease in value of the dependent variable for every unit increase in the independent variable.

17. E

The appropriate significance test would be a chi-square test of independence; however, the conditions have not been met. Two of the six expected counts are 4.731 and 0.828; this violates the condition that no more than 20% of the expected counts can be less than 5. Additionally, all expected counts should be greater than 1.

18. D

Since the slope coefficient is positive, the correlation coefficient will be positive. The correlation coefficient, r, is the square root of R-squared.

19. C

A p-value is the probability of obtaining results like those from your sample (or more extreme) if the null hypothesis is true. Thus, if the two roommates spend equal amounts on groceries, you could get results like those sampled (or more extreme) 6.7% of the time. Since $p = 0.067 < 0.10$, the results are significant at the 0.10 level.

20. E

$$\$10\left(\frac{1}{8}\right) + \$5\left(\frac{1}{8}\right) - \$5\left(\frac{3}{4}\right) = -\$1.875 \approx -\$1.88 \text{ for one game.}$$

Therefore, the player can expect to lose $(2)(1.875) = \$3.75$ for two games.

21. E

$$z = \left(\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}\right) = \left(\frac{0.9 - 0.876}{\sqrt{\frac{(0.876)(0.124)}{200}}}\right)$$

and we wish to find the probability that more than 180 of the 200 individuals are *not* senior citizens. Since 12.4% are senior citizens, 87.6% are not.

22. C

A t-interval with 19 degrees of freedom should be constructed with the given information, and substitutions should be made into this formula:

$$\bar{x} \pm t^*_{19}\frac{s}{\sqrt{n}} = 5.65 \pm 2.093\frac{1.69}{\sqrt{20}}.$$

23. C

Standardized value for this year's test: $\dfrac{40 - 38}{2} = 1$

Standardized value for last year's test: $\dfrac{35 - 34}{1} = 1$

Therefore, in both years, the student scored 1 standard deviation higher than the mean. Thus, the student scored equally well in both years.

24. D
Since $p = 0.0028 > 0.001$ but $p = 0.0028 < 0.01$, there is evidence of a relationship between distance and diameter at the 0.01 level.

25. E
The mean of the distribution of sample means should have a mean of approximately 12.262 and a standard deviation of approximately $9.610/\sqrt{9}$ or 3.1. It is important to check the scale on each axis and not just look at the shape of the distribution.

26. C
The slope of the least squares regression line is 0.959388. With 19 degrees of freedom and 95% confidence, the critical value is 2.093. The standard error of the slope coefficient, as can be seen in the computer output, is 0.1326.

27. E
Remember, this is a *cumulative* frequency histogram. Only one game is added to the histogram in the last two classes (6.5 and 7).

28. C
In a distribution skewed to the right, the mean is greater than the median. Recall that the mean follows the tail of a skewed distribution.

29. A
A Type I error is the error of rejecting the null hypothesis when the null hypothesis is true. In this case, it means the couple rejects the belief that profits are \$15,000 or less, believing the average profits are more than \$15,000. Since the average profits really are \$15,000 or less, the couple will go bankrupt.

30. B
A confidence interval yields an interval of values that one can be reasonably confident will capture the true parameter of interest. In this case, the parameter of interest is the proportion of seniors *at this school* who will attend an institution of higher learning (eliminate answer choice C). We already know the proportion from our sample (eliminate A and E). The true proportion is a fixed value and will either be captured by the interval or not (eliminate D and E).

31. D
This is a geometric setting with $p = 0.2$ and $x = 3$. There are two failures (0.8^2) before the first success (0.2).

32. B
If each value of the distribution is multiplied by 2, then the standard deviation will also be multiplied by 2. Addition of 5 to each of the values only shifts the distribution and does not change the spread of the distribution.

33. E

This is a binomial setting with $p = 0.014$ and $n = 500$. Therefore, the mean is $(500)(0.014) = 7$ and the standard deviation is $\sqrt{(500)(0.014)(1 - 0.014)} \approx 2.627$.

34. B

Increasing the sample size will decrease the width of a confidence interval. Additionally, decreasing the confidence level will decrease the width of a confidence interval. Therefore, the combination that will most decrease the width of the interval would be a combination of increasing the sample size and decreasing the confidence level.

35. A

Predicted grocery cost $= -33.22 + 44.77(4) = \$145.86$

Actual grocery cost $= \$135$

Residual = actual − fitted $= 135 - 145.86 = -10.86$

36. D

Performing a *one-proportion z-test* with H_0: $p = 0.089$ and H_a: $p > 0.089$, $z = 1.706$, and $p = 0.0440$. The critical region boundary for the 0.05 level is $z = 1.645$. Therefore, since $1.706 > 1.645$, or since $.0440 < 0.5$ ($p < \alpha$), there is evidence that the county's poverty rate is higher than that of the state.

37. B

The peak of the curve is at approximately 30, and the change in curvature occurs at 25 and 35, 5 units (or 1 standard deviation) on either side of the mean. Additionally, from the Empirical Rule we know that 99.7% of the values should fall within 3 standard deviations of the mean.

38. B

For independent variables, $\mu_{A-B} = \mu_A - \mu_B = 5.1 - 4.5 = 0.6$ and $\sigma_{A-B} = \sqrt{\sigma_A^2 + \sigma_B^2} = \sqrt{0.57^2 + 0.63^2} \approx 0.85$.

39. E

No more than 3 means 3 or fewer. This is a binomial setting with $p = 0.176$ and $n = 10$.

40. A

Increasing the sample size improves power, and increasing the value of α also improves power. Therefore, overall, power can best be improved by increasing both the sample size and the α value.