January 2007 Statistics Solutions

1. **C**. The symbols used for the mean and standard deviation are for statistics, not parameters. This indicates that the students represent a sample.
2. **B**. This is a geometric situation. Three failures followed by a success. The result is $\left(\frac{7}{8}\right)^3\left(\frac{1}{8}\right)=\frac{343}{4096}$.
3. **A**. The formula for a confidence interval is $\hat{p}\pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. Plugging the numbers in gives

$$\frac{90}{300}\pm 1.96\sqrt{\frac{.3(.7)}{300}}=\frac{3}{10}\pm .0518567257=(.2481432743,.3518567257)=(.248143,.351857)$$

4. **C**. To find the mean, multiply X(P(X)) and add the results. The solution is $2(.16)+6(.21)+12(.27)+15(.13)+21(.23)=.32+1.26+3.24+1.95+4.83=11.6$.
5. **D**. To find the probability of at least one win, calculate the probability that Barry loses all three races. Then do 1-P(0 wins) to find the solution. The answer is $1-\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)\left(\frac{4}{7}\right)=1-\frac{2}{21}=\frac{19}{21}$.
6. **A**. When the researcher doesn't know which groups the subjects are in, it is a blinding situation.
7. **D**. This is the definition of a cluster sample, to survey everyone in a particular area. The area chosen is through a SRS, but the sample itself is a cluster sample.
8. **B**. There are 48 parks with both types of fields. So the answer is $\frac{48}{182}=\frac{24}{91}$.
9. **D**. The margin of error formula for a proportion is $z\sqrt{\frac{p(1-p)}{n}}$. Plugging the numbers in gives $.06=1.96\sqrt{\frac{.6(.4)}{n}}$. Solving for n gives a value of 256.1067, which rounds up to 257.
10. **C**. Do two z scores, one with each raw score. The results are $\frac{490-540}{20}=-2.5$ and $\frac{600-540}{20}=3$. From the z score chart, the corresponding percentages are .0062 and .9987, respectively. The difference between them is .9925, or 99.25%.
11. **E**. To find the new mean, multiply by 1.5 and subtract 9. 62(1.5)-9= 84. The new standard deviation is 6(1.5)= 9. The answer is (84, 9).
12. **A**. The coefficient of determination is $r^2$. To find the correlation coefficient, we use the formula for slope. $m=r\frac{s_y}{s_x}$. Plugging the numbers in gives $\frac{72}{175}=r\left(\frac{4}{7}\right)$. Solving for r gives r= .72. And r squared is the answer.
13. **B**. To find the standard deviation, first find the mean of the variable, which is 16.04. Subtract the mean from each value, square the differences and multiply the squared differences by their probabilities, then add those results to get a sum of 109.5984. Take the square root of the sum to get the answer.

14. **A**. The mean is the lowest because of the skew, then median.

15. **A**. The formula for $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$. Plugging the numbers in gives $\frac{7}{13} = \frac{P(A \cap B)}{\frac{39}{100}}$. Therefore, $P(A \cap B) = .21$. The venn diagram breaks down into A only = .36, A and B= .21, B only = .18, and neither =.25. So, the answer comes from $P(A' \mid B') = \frac{P(A' \cap B')}{P(B')} = \frac{.25}{.61} = \frac{25}{61}$.

16. **C**. Type the data into the calculator list. The first quartile is 4, the third quartile is 14 and the mean is 9. The discriminant of a quadratic equation is $B^2 - 4AC$. So the result is $14^2 - 4(4)(9) = 196 - 144 = 52$.

17. **B**. The formula for chi square is $\sum \frac{(obs - \exp)^2}{\exp}$. The expected grades are 14 A's, 17.5 B's, 17.5 C's, 14 D's and 7 F's. The complete solution for this problem is $\chi^2 = \frac{(11-14)^2}{14} + \frac{(21-17.5)^2}{17.5} + \frac{(26-17.5)^2}{17.5} + \frac{(9-14)^2}{14} + \frac{(3-7)^2}{7} = \frac{9}{14} + \frac{7}{10} + \frac{289}{70} + \frac{25}{14} + \frac{16}{7} = \frac{334}{35}$.

18. **C**. Using the z score chart, the z score for this problem is 1.08. Plugging the numbers into the z score formula gives $1.08 = \frac{53-48}{sd} \Rightarrow sd = 4.62962963 = 4.63$.

19. **E**. None of the situations in a, b or c change the value of the correlation coefficient.

20. **A**. Solving the first joint probability, .4(y=2) = .2, so (y=2) = .5. Then (y=3) = .2 because the sum of the probability must equal one. Solving the second joint probability, (x=3)(.3) = .045, so (x=3) = .15. Then (x=4) = .25, again because the sum has to be one. Therefore P(x=4, y=2) = (.25)(.5) = .125.

21. **D**. Median and IQR are not affected by outliers.

22. **C**. For the mean 3(31)-2(24)= 45. For the standard deviation, they are 9 and 4 after the scalar. $\sqrt{9^2 + 4^2} = \sqrt{97}$.

23. **B**. You either look at the vertical or horizontal ratio to find n. The horizontal ratio is 3.5. The vertical ratio is 1.6. Either way the resulting proportion gives you the answer.

24. **D**. Every sample will give different statistics and those differences are the sampling error from the population parameter.

25. **A**. The formula for residual=observed – predicted. The predicted value is 2.25(74)+12= 178.5. 175-178.5= -3.5.

26. **D**. Statement three is not true.

27. **A**. There are two z-scores. 65-mean= -.62 SD and 88-mean= 1.26 SD. Eliminating the mean and solving for the standard deviation gives 23= 1.88 SD. Solving for SD gives 12.23404255, which rounds to 12.23.

28. **B**. Find the mean, which is 11. Subtract the mean from each value and square the differences. Add them up and the total is 420. Divide by (n-1), or 9 and take the square root. The result is the answer when simplified.

29. **C**. 60% are males, so 40% are female. 80% of males go to games, so 20% don't. 60% of females don't go, so 40% do. Therefore, the percent of students who go to games are .6(.8) + .4(.4) = .64. The probability the student is female is $\frac{.16}{.64} = \frac{1}{4}$.

30. **B**. The formula for standard deviation of a binomial distribution is $\sqrt{np(1-p)}$. Plugging the numbers in gives $\sqrt{150(.72)(.28)} = 5.499090834 = 5.50$.