2009 National Convention Statistics Solutions

1. **D**. The only way to show a cause and effect relationship is to perform a controlled experiment.
2. **C**. Using the z-score formula z = (raw-mean)/standard deviation, the two z scores are -1.5 and 2.5. The probability that a weight is less than 120 is .0668 and the probability the weight is more than 200 is .0062. The sum of the those two values is the solution.
3. **B**. The formula for a two proportion confidence interval is $(p_1 - p_2) \pm z\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$. Assuming a positive difference and plugging the numbers in gives $(.75 - .62) \pm 1.645\sqrt{\frac{(.75)(.25)}{100} + \frac{(.62)(.38)}{150}} = .13 \pm .0965611731$, which produces (.0334388269, .2265611731). When both values are rounded to six decimal places, the answer is produced.
4. **C**. Assign each person a two digit number 01-10. Reading from left to right, the first five numbers are 07, 05, 10, 02 and 01. The five people with those numbers are the solution.
5. **C**. The sum of the minimum, median and maximum is 100. Q_1 is the average of the first two numbers in the data set, so the second number is 15 because the sum of the first two numbers is 25. Q_3 is the average of the last two numbers in the data set, so the fourth number is 45 because the sum of the last two numbers is 105. 100+15+45 = 160.
6. **D**. The formula that it takes at least n trials before a success is $(1-p)^n$. Plugging the numbers in gives $(1-.82)^4 = .18^4 = .00104976$.
7. **A**. The formula for margin of error is $z\sqrt{\frac{p(1-p)}{n}}$. Plugging the numbers in gives $.03 = 1.96\sqrt{\frac{.83(.17)}{n}} \rightarrow n = 602.28 \approx 603$.
8. **B**. Compute two z scores using the chart: $2.02 = \frac{93 - mean}{sd}$ and $-1.27 = \frac{60 - mean}{sd}$. Solving for the standard deviation gives a value of $\frac{3300}{329}$. Plugging the standard deviation in and solving for the mean gives the solution.
9. **E**. All of the conditions are true.
10. **C**. Find the mean, which is 10. Subtract the mean from each value, square the differences and add them up, which is a total of 252. Divide by (n-1), or 6, and the result is 42. That is the variance. The square root of that is the solution.

11. **D**. $P(A'|B') = \frac{P(A' \cap B')}{P(B')} \rightarrow \frac{11}{20} = \frac{P(A' \cap B')}{\frac{5}{11}} \rightarrow P(A' \cap B') = \frac{1}{4}$. Therefore,

$P(A \cup B) = \frac{3}{4}$. So, $P(A \cup B) = P(A) + P(B) - P(A \cap B) \rightarrow \frac{3}{4} = \frac{10}{33} + \frac{6}{11} - P(A \cap B)$.

Therefore, $P(A \cap B) = \frac{13}{132}$. So, $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{13}{132}}{\frac{6}{11}} = \frac{13}{72}$

12. **D**. The standard deviation of a binomial distribution is $\sqrt{np(1-p)}$. When you plug the numbers in, you get $\sqrt{75(.38)(.62)} = 4.20$.

13. **B**. So far, Ara's grade in the class is 93(.3)+85(.25)+86(.2) = 66.35. Therefore, Ara must earn a minimum of 23.65 on the final to finish with an average of at least 90. So, $23.65 = .25(\textit{final})$. When you solve for the final exam grade, you get the solution.

14. **A**. P(X=2, Y=2) = .05 = .1(P(Y=2)). So, P(Y=2) = .5. Therefore, P(Y=3) = .2. P(X=3, Y=1) = .15 = P(X=3)(.3). So P(X=3) = .5. Therefore, P(X=4) = .2. So, P(X=4, Y=3) = .04.

15. **B**. A multi-stage sample takes an SRS of each full subgroup of a population and then takes an SRS of the remaining subgroups.

16. **C**. When you put the numbers in a Venn diagram, there are three openings. They represent the three positions of students who take exactly two of the languages. Let a= Spanish and Japanese, b=Spanish and French, and c= French and Japanese. You produce three equations: a+b=12, b+c=9 and a+c=11. Solving those equations gives you values of a=7, b=5 and c=4. To find the total number of seniors, there are 40 who take Spanish, 12 who take French only, 15 who take Japanese only, and 4 who take French and Japanese. Those four numbers add to the solution.

17. **D**. Working off of the previous problem, there are 7 students who take Spanish and Japanese and 3 who take all three classes. So of the 29 Japanese students, 10 take Spanish.

18. **D**. The formula for χ^2 is $\sum \frac{(obs - \exp)^2}{\exp}$. Plugging the values in gives

$$\frac{(32-27)^2 + (25-27)^2 + (20-27)^2 + (35-27)^2 + (28-27)^2 + (22-27)^2}{27} =$$

$$\frac{25+4+49+64+1+25}{27} = \frac{168}{27} = \frac{56}{9}.$$

19. **B**. The number of degrees of freedom for a goodness of fit test is (n-1), or 5 in this case.

20. **B**. Power is the measure of the ability of a hypothesis test to detect a difference between the hypothesized mean and the true mean.

21. **E**. You are not told in the problem that the variables X and Y are independent. Therefore, you can not calculate the standard deviation of (X-Y).

22. **A**. To find the standard deviation, multiply X(P(X)) and add them up. The mean is 15.37. Subtract the mean from each value, square the differences, and multiply those by their corresponding probabilities. The sum is 8.8131. The square root of this number is the solution. When it is rounded to two decimal places, the answer is produced.

23. **C**. The coefficient of determination is r^2.

24. **D**. 60% of the drivers are boys, so 18% of students are boy drivers. Therefore, there are 32% who are not boy drivers. So (.7)(x) = .32. Therefore, x = 16/35. The final solution is $\dfrac{\frac{16}{35}}{\frac{7}{10}} = \dfrac{32}{49}$.

25. **B**. The formula for a t confidence interval is $\bar{x} \pm t\dfrac{s}{\sqrt{n}}$. Plugging the numbers in produces $143 \pm 1.711\dfrac{(13.2)}{\sqrt{25}} = 143 \pm 4.51704 = (138.483, 147.517)$.

26. **D**. The formula for the line of best fit is $y - \bar{y} = r\dfrac{s_y}{s_x}\left(x - \bar{x}\right)$. Plugging the values in produces $y - 82 = (.72)\left(\dfrac{8}{5}\right)(x - 53) \rightarrow y - 82 = 1.152(x - 53)$. That leads to the final solution when you solve for y.

27. **A**. The solution is the sum of four binomial probabilities. The probability of getting four, five, six and seven questions correct out of ten with a probability of success of .25. This produces ${}_{10}C_4\left(\dfrac{1}{4}\right)^4\left(\dfrac{3}{4}\right)^6 + {}_{10}C_5\left(\dfrac{1}{4}\right)^5\left(\dfrac{3}{4}\right)^5 + {}_{10}C_6\left(\dfrac{1}{4}\right)^6\left(\dfrac{3}{4}\right)^4 + {}_{10}C_7\left(\dfrac{1}{4}\right)^7\left(\dfrac{3}{4}\right)^3 = \dfrac{14661}{65536}$.

28. **D**. The formula is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. In this case, the solution is $\dfrac{26}{52} + \dfrac{40}{52} - \dfrac{20}{52} = \dfrac{46}{52} = \dfrac{23}{26}$.

29. **B**. First you must transform the data by changing the standard deviation from 10 to 6, or multiplying by $\dfrac{3}{5}$. When you multiply the original mean by $\dfrac{3}{5}$ you get 40.8. To get to the new mean of 75, you must add 34.2. So the transformation equation is $y = \dfrac{3}{5}x + 34.2$. When you plug 80 into the equation for x, the new score is the solution.

30. **D**. To find the numerical contribution, subtract the mean from the value, square the difference and divide by (n-1). In this case, $10 - 6 = 4,\ 4^2 = 16,\ and\ \dfrac{16}{8} = 2$.

Tiebreak 1. $\mathbf{\dfrac{1}{10}}$ or **.1**. The only integers that have an odd number of factors are perfect squares. There are 10 perfect squares between 1 and 100. Therefore, the answer is $\dfrac{10}{100} = \dfrac{1}{10}$.

Tiebreak 2. **.454** or $\mathbf{\dfrac{227}{500}}$. When two events are independent, $P(A \cap B) = P(A)P(B)$. So, $P(A \cap B) = .58 \times .35 = .203$. Then, $P(A \cup B) = .35 + .58 - .203 = .727$. $P(A' \cap B')$ is the part of the Venn diagram outside of $P(A \cup B)$. So, $P(A' \cap B') = .273$. Therefore, the final answer is .727-.273 = .454.

Tiebreak 3. $\mathbf{\frac{\sqrt{15}}{2}}$. The standard deviation of a geometric distribution is $\sqrt{\frac{1-p}{p^2}}$.

Plugging in the value of p gives $\sqrt{\frac{1-.4}{(.4)^2}} = \sqrt{\frac{.6}{.16}} = \sqrt{\frac{15}{4}} = \frac{\sqrt{15}}{2}$.