February 2006 Statistics Regional Solutions

1. **C**. When you put the information into a Venn diagram, you have three unknowns. You let a=Math and History, b= Math and English and c=English and History. By subtraction, a+b=12, b+c=9, and a+c=11. Solving those equations gets you a=7, b=5, and c=4. Therefore, the answer is 5.

2. **B**. The number of people in $(M \cap E')$ is 17, 10 from math only and 7 from the previous solution in #1. The number of people in $(E \cap H)$ is 12, 8 from all three and 4 from the previous solution in #1. Since it is union and all four numbers are unique from each other, you add them together and get 29. Therefore, the probability is $\frac{29}{46}$.

3. **E**. You are not told in the question that sets X and Y are independent. Therefore, you can't find the value of the standard deviation of the sum. Therefore, no solution.

4. **A**. Plugging the data into a graphing calculator produces a mode of 1, a mean of 6 and a median of 7. Putting those solutions into the equation produces $x^2 + 6x - 7 = 0 \rightarrow (x-1)(x+7) = 0$. Solving gets you the answer.

5. **B**. The equation for the least squares regression line is $y - \bar{y} = r\left(\frac{S_x}{S_y}\right)(x - \bar{x})$.

Plugging the information in gives you $y - 62 = .72\left(\frac{5}{8}\right)(x - 93) \rightarrow y - 62 = .45(x - 93)$

$y - 62 = .45x - 41.85 \rightarrow y = .45x + 20.15$.

6. **D**. The formula for a t confidence interval is $\bar{x} \pm t\left(\frac{s}{\sqrt{n}}\right)$. Plugging the values in produces $22.5 \pm (2.365)\left(\frac{7.19}{\sqrt{8}}\right) \rightarrow 22.5 \pm 6.011945597 \rightarrow (16.4880544, 28.5119456)$.

Rounding each to four decimal places gives the answer.

7. **B**. You set up two equations with two unknowns. Using the z-chart, the two z-scores are -1.32 and 1.57. The two equations using the z-score formula are 64-mean= -1.32SD and 93-mean= 1.57SD. Using algebra produces the answer.

8. **A**. It is a binomial distribution. The answer is the difference between two cumulative binomial distributions. The answer is from $binocdf(17,.53,13) - binomcdf(17,.53,8) =$ .5869300208. When rounded, the answer is complete.

9. **C**. First, find the mean of the distribution by multiplying X by P(X) and adding the results, which sums to 5.3. Subtract the mean from each value, square the differences, and multiply them by their corresponding probabilities. Add the results together to get a variance of 5.61. Take the square root to find the standard deviation= 2.368543856, which when rounded gives you the answer.

10. **D**. The first step is to change the standard deviation from 5 to 4, which is done by multiplying the standard deviation by $\frac{4}{5}$. You also multiply the mean by this same value, which gets you $\frac{4}{5}(73) = 58.4$. To get to 80, you must add 21.6. The answer is the results in slope intercept form.

11. **D**. You do a tree diagram, with overall percents 10 and 90. Then you break up into positive test or not. Going left to right is 85, 15, 2 , 98 percent. The overall percent for those who test positive is (.10)(.85)+(.90)(.02)= .103. Those who test positive but don't have the disorder is (.90)(.02)= .018. The result comes from .018/.103.

12. **A**. First you must calculate the raw score using the z-score formula. The z-score is 1.645 at the 5% level. Plugging in the values produces $1.645 = \frac{raw - 450}{\frac{125}{\sqrt{500}}}$. Solving gives you a raw value of 459.1958296. Plugging in this value and the alternative mean into another z-score formula produces a z-score of -.143854382. Using the z-score chart by rounding the z-score to two decimal places (-.14), gives you the answer using the chart appropriately (1-.4443).

13. **C**. Statements one and two are true. Statement three is false. Decreasing standard deviation increases power.

14. **C**. The coefficient of variation is the standard deviation divided by the mean. So the result is 25/100, which reduces to the answer.

15. **B**. Using the formula $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$ gives a result of $P(A \cap B) = 0.3$. Putting all the information into a venn diagram leaves 0.4 outside the circles and inside the venn diagram. Therefore, the solution is 0.4/0.6, which reduces to the answer.

16. **C**. The coefficient of determination is $r^2$. To find r, use the slope formula, $m = r\left(\frac{S_y}{S_x}\right)$. Plugging in the numbers appropriately gives the following $.45 = r\left(\frac{9}{14}\right) \rightarrow r = .7$. Therefore, the solution is $.7^2 = .49$.

17. **A**. Using the z chart, the z for a 90% confidence interval is 1.645. Using the formula $p_1 - p_2 \pm z\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$ gives you $(.75 - .6) \pm (1.645)\sqrt{\frac{.75(.25)}{92} + \frac{.6(.4)}{45}} = .15 \pm .1412343086 =$ $(.0087656914, .2912343086)$ which when rounded gives you the answer.

18. **D**. Using the formula for chi square testing, $\sum \frac{(obs - \exp)^2}{\exp}$ gives the following $\frac{\left(15 - \frac{50}{3}\right)^2 + \left(20 - \frac{50}{3}\right)^2 + \left(25 - \frac{50}{3}\right)^2 + \left(20 - \frac{50}{3}\right)^2 + \left(10 - \frac{50}{3}\right)^2 + \left(10 - \frac{50}{3}\right)^2}{\frac{50}{3}} = 11.$

19. **B**. The only way to show a true relationship is to do a controlled experiment.

20. **A**. residual=actual value-predicted value. When you plug 90 into the equation, you get 87.5 for the predicted value. So residual= 80-87.5= -7.5.

21. **C**. Using the z-score formula, the two z-scores are $\frac{65-74}{6}=\frac{-9}{6}=-1.5$ and $\frac{89-74}{6}=\frac{15}{6}=2.5$. The two percentages for those z scores in Table A are .0668 and .9938 respectively. .9938-.0668= .927.

22. **D**. The sum of the six tests are 489. The formula for the mean is $\frac{489+3x}{9}>86$. Solving for x gives the answer.

23. **D**. The formula for the standard deviation is $\sqrt{np(1-p)}$. Plugging the numbers in gives you $\sqrt{55(.62)(.38)}=\sqrt{12.958}=3.599722212=3.60$

24. **B**. To get four of a kind, you use the formula $\frac{{}_4C_4 \bullet {}_{48}C_1}{{}_{52}C_5}=\frac{1}{54145}$. Then you multiply by 13 because there are 13 different ways to get four of a kind (Ace through 10, J,Q,K). $\frac{1}{54145}\bullet 13=\frac{1}{4165}$.

25. **A**. Because the mean is lower than the median, low scores are dragging the distribution down, therefore the distribution is skewed to the left.

26. **C**. The standard deviation is the square root of the variance. The standard deviation can be zero if all values are the same. Statement three is true.

27. **C**. The definition of the purpose of a residual plot is the answer.

28. **A**. The mean of the X distribution is 2.1 and the mean of the Y distribution is 5.1. Subtracting the mean, squaring the differences and multiplying by the probabilities gives 2.49 for X and 3.29 for Y. Taking the square roots of both numbers gives 1.256174106 for X and 1.346787182 for Y. 4X+5Y= 15.38107211 which rounds to the answer.

29. **E**. The p-value is the probability of a result at least as far out as the result we actually got.

30. **D**. A t-test is used when the standard deviation of the population is unknown.