

2006 January Palm Harbor Invitational – Statistics Individual
Page 1 of 7

For each question below, NOTA means “None of These Answers.” Attached should be a chi-square table and a z-table. Mark answers on the scantron provided.

1. What is the mean of the following set: {1, 2, 3}?
A. 1 B. 2 C. 3 D. 6 E. NOTA
2. Let set X contain the first thirteen nonzero perfect square numbers from the population of all nonzero perfect square numbers. If the standard deviation of set X can be written in the form: $\frac{\sqrt{A}}{B}$, where A and B are integers and the fraction is completely rationalized and is in simplest, radical form, what is the value of A + B?
A. 2899 B. 6281 C. 11002 D. 12560 E. NOTA
3. If a set has a mean of 8 and a variance of 4, what will happen to the mean and standard deviation of that set if 4 is added to each member of the set? Answers are in the form: Mean, Standard Deviation.
A. 8, 4 B. 12, 4 C. 8, 2 D. 12, 2 E. NOTA
4. Sarah and Jessica are sworn enemies. They never want to be compared to each other. They dislike each other so much that they even take different standardized tests. Jessica took the SAT and got a 1910, while Sarah took the ACT and got a 28. If the national averages are 1530 and 22 for the SAT and ACT, respectively, and the variances are 99225 and 24.74 for the SAT and ACT, respectively, who performed better when compared to the nation, assuming that both the SAT and the ACT have a normal distribution?
A. Sarah C. They performed equally well E. NOTA
B. Jessica D. Cannot Be Determined
5. If the $P(A) = .3$, the $P(B) = .5$, the $P(A | B^C) = .2$, and the $P(B | A^C) = \frac{3}{7}$, then what is the $P[(A \cap B^C) \cup (A^C \cap B)]$?
A. .2 B. .3 C. .4 D. .5 E. NOTA
6. On planet X, the Minister of Pain wanted to sample the people who are scared by his name. First he separated the planet into Chu-Pu regions and then took a simple random sample of the Rawars in each Chu-Pu. He also took another SRS of the Chocolate that are present in each Rawar. Finally he sampled every fourth person in the Chocolate. Which sampling technique did the Minister of Pain NOT use?
A. Cluster Sampling C. Stratified Sampling E. NOTA
B. Multistage Sampling D. Systematic Sampling

2006 January Palm Harbor Invitational – Statistics Individual
Page 2 of 7

7. After his sampling, the Minister of Pain from the previous question found that the proportion of people who are scared by his name is .978 when the population proportion is known to be .999. Which of the following correctly explains the variation in the results?
- A. Hawthorne Effect C. Human Mistake E. NOTA
B. Hidden Bias D. Placebo Effect
8. When performing a confidence interval for the average number of seeds in a tomato, Jenn the Eliminator forgot what X% confidence interval she used. She only remembers that she sampled 28 people, the population standard deviation is eight, and the interval produced was $(\sqrt{110}, \sqrt{140})$. What is the value of X, the percent confidence of her interval?
- A. 33 B. 34 C. 44 D. 67 E. NOTA
9. Beth is a Twister_Hoop fanatic. Every night, she plays for at least four hours. With all this practice, one would think she was the Twister_Hoop champion; however, recently she lost to the diabolical Jessica. Beth wants to perform an experiment on the Twister_Hoop board, which consists of a spinner that can land on green, blue, or red, to see if the board is rigged. In her experiment, Beth will spin the board and record the number of times the spinner lands on each color. If each of the spins is independent, which distribution best describes Beth's experiment?
- A. Binomial C. Poisson E. NOTA
B. Geometric D. Multinomial
10. If the correlation between two variables is equal to zero, then which of the following are true?
- I. The least squares regression line is a horizontal line.
II. There is no relationship between the data.
III. The two variables are independent of each other.
- A. I only B. I and II C. I and III D. I, II, III E. NOTA
11. In a very unique binomial distribution, the standard deviation is equal to the mean. Now Timmy forgot what the probability of success for each trial was because he had malaria. The only thing he could remember was that there were ten fixed observations and that the probability for success is greater than zero. What is the probability of success for any one trial?
- A. $\frac{1}{11}$ B. $\frac{1}{2}$ C. $\frac{10}{11}$ D. Does Not Exist E. NOTA

2006 January Palm Harbor Invitational – Statistics Individual
Page 3 of 7

12. In Bizarro land, a very strange die has a varying probability. When six dice are rolled, you can expect to get no threes 25% of the time, one three $\frac{1}{3}$ of the time, two threes $\frac{1}{6}$ of the time, and three or four threes $\frac{1}{24}$ of the time, and five or six threes $\frac{1}{12}$ of the time. Well, Bizarro Dood believes that his crazy dice are seriously flawed. When he rolled all six at the same time 240 times, he got 0, 1, 2, 3, 4, 5, 6 threes 61, 97, 36, 11, 12, 20, 3 times respectively. On which interval does the p-value lie for the null hypothesis that his dice are standard?

- A. $p > .1$ B. $.05 < p < .01$ C. $.01 < p < .005$ D. $p < .01$ E. NOTA

13. Below is a computer printout for women’s shoe size regressed on women’s salary, in thousands of dollars.

Predictor	Coeff	StDev	T	D
Constant	87.2	2.04	.95	.783
Shoe Size	-7.4	1.51	24.1	.002
Rsq = 79.8%		Rsq (adj) = 80.0%		

What is the approximate value of the correlation between shoe size and salary of women?

- A. .798 B. .893 C. -.798 D. -.893 E. NOTA

14. Mrs. Linder is in a happy mood today because she knows that there is an 80% chance that any random student will have his or her homework today. She will give candy to the first X number of people that have her homework. Once she comes across someone without it, she will stop handing out candy. Which best describes the previous situation?

- A. Binomial C. Poisson E. NOTA
B. Geometric D. Multinomial

15. Given the following table:

10	A
35	152

Let A be an integer that allows the table to show complete independence. What is the units digit of A?

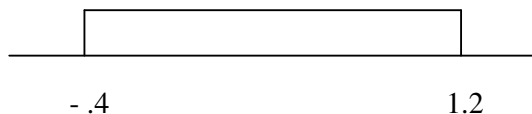
- A. 2 B. 4 C. 6 D. 8 E. NOTA

2006 January Palm Harbor Invitational – Statistics Individual
Page 4 of 7

Use the following information for question 16 – 17:

Students at a rehab clinic for videogames are addicted to at least one and up to and three videogames: Grand Theft Auto 3, Banjo-Kazooie, and Warcraft. Now ten people are addicted to all three games. Twenty four are hooked on Grand Theft Auto, thirty three can't live with out Banjo-Kazooie, and 20 are addicted to Warcraft. Only four people are urging to play only Grand Theft Auto, twelve people lust for Banjo-Kazooie, and seven people yearn for Warcraft, respectively.

16. Assume that a randomly chosen person is addicted to Warcraft. What is the probability that he is also addicted to Grand Theft Auto?
- A. $11/45$ B. $2/9$ C. $11/20$ D. $1/2$ E. NOTA
17. A person randomly selected is not addicted to Banjo-Kazooie. What is the probability that this person is addicted to Warcraft?
- A. $2/3$ B. $8/15$ C. $8/45$ D. $2/9$ E. NOTA
18. How many of the following are qualitative variables?
- | | | | |
|----------------|--------|------------------|-----------------|
| Zip Code | Weight | Number of Bills | Favorite game |
| Types of Bills | Speed | Area Code | Glass Thickness |
| House Number | Height | Percentage Yield | Skin Texture |
- A. 3 B. 4 C. 6 D. 7 E. NOTA
19. If $P(X | Y) = 0$ and $P(X)$ and $P(Y)$ are positive, then events X and Y could be:
- A. Mutually Exclusive C. Both A and B E. NOTA
B. Independent D. Neither A nor B
20. Given the following probability distribution:



What is the probability that X is greater than .6?

- A. $3/8$ C. $3/4$ E. NOTA
B. $3/5$ D. More Information Needed

2006 January Palm Harbor Invitational – Statistics Individual
Page 5 of 7

21. A recent study claimed that after giving an IQ test to four different species (cats, dogs, fish, and snakes), 72% of the cats scored in the genius range. Of the 160 dogs that took the test, 80% were considered genius. Only twelve out of sixteen snakes could be called genius based in their performance on the test. Based on the above results, which of the species test is the smartest, if smartness is determined by which species have the most geniuses?
- A. Cats C. Snakes E. NOTA
B. Dogs D. More Information Needed
22. On a test in Spanish, Alice got a 103, while Akash got a 49 because he had salmonella. Alice was in the 99.8 percentile while Akash was placed back at the 2.44 percentile. What is the tenths place of the variance for the class average on the Spanish test?
- A. 1 B. 3 C. 5 D. 7 E. NOTA
23. How many of the following are true?
- a. The control group may receive either the placebo or the treatment being tested
 - b. Blocking is used to reduce bias in an experiment
 - c. A well designed experiment always has a single treatment but may test that treatment at different levels
 - d. Association generally implies causation in an experiment
- A. 0 B. 1 C. 2 D. 3 E. NOTA
24. Given the word STATISTICS, how many different rearrangements of the letters are possible, if the order of the letters does matter and the letters must be used?
- A. 3,628,800 B. 604,800 C. 100,800 D. 50,400 E. NOTA
25. Given that S_1 (the S stands for distribution) has a mean of 3.14 and a standard deviation of 2.71, that S_2 has a mean of 1.41 and a standard deviation of 1, and that $S_3 = [3(S_1 + 2S_2) - 4.1]$, what is units place of the variance of S_3 ?
- A. 0 C. 7 E. NOTA
B. 2 D. More Information Needed
26. If a least squares regression line has the form: $Y = aX + b$, where a and b are constants, and that the means of X and Y are 31 and 14 respectively, what is the tenths place of a + b, assuming that one point used to make the least squares regression line was (16, 44) and its residual was 3.7?
- A. 2 B. 3 C. 5 D. 6 E. NOTA

2006 January Palm Harbor Invitational – Statistics Individual
Page 6 of 7

27. What of the following is the best approximation of the area under a normal curve that is within seven standard deviations of the mean?
- A. .68 B. .95 C. .997 D. 1 E. NOTA
28. Theoretically, which of the following methods is most accurate to gather data in order to compare the number of bad hair days and the magnitude of smelly feet of Americans?
- A. Sample Survey C. Census E. NOTA
B. Experiment D. Observational Study
29. The standard M&M&M bag weighs 2.13 pounds. Now the distribution for the weight of an M&M&M is normal with a mean of .071 pounds per M&M&M with a standard deviation of .023 pounds. What is the probability that 32 M&M&Ms in a bag will exceed the weight of a standard bag of M&M&M?
- A. .138 B. .425 C. .552 D. .891 E. NOTA
30. Regress X (the first twenty prime numbers) on Y (the first twenty multiples of seven). The points should appear as (2, 7), (3, 14), (5, 21), (7, 28), etc. What is the sum of all the residuals for all points? (Rounded to the nearest integer)
- A. -7 B. -2 C. 1 D. 13 E. NOTA

2006 Palm Harbor University Invitational – Statistics Solutions

1) The mean is $(1+2+3)/3 = 2$.

B

2) Set $X = \{1, 4, 9, 16, 25, 36, 49, 64, 81, 100, 121, 144, 169\}$

Mean $X = 63$. Sum of $(x - \text{mean})^2 = 37674$. For standard deviation, use $n - 1$ degrees of freedom, so $13 - 1$, or 12.

3) The mean will increase by 4, so 12, and nothing will happen to the standard deviation, so 2. The answer is 12, 2.

D

$$\text{Standard deviation} = \sqrt{\frac{\sum(x - \text{Mean})^2}{n - 1}} = \sqrt{\frac{37674}{12}} = \frac{\sqrt{12558}}{2}. \quad \mathbf{D}$$

4) Since both the SAT and the ACT have a normal distribution, you can just standardize both scores to the entire population, which is the nation, to see which person did better.

Use the formula: $z = (X - \mu) / \sigma$, where X is the score, σ is the standard deviation of the population and μ is the mean of the population, to standardize the scores. Remember to change the variance to standard deviation!

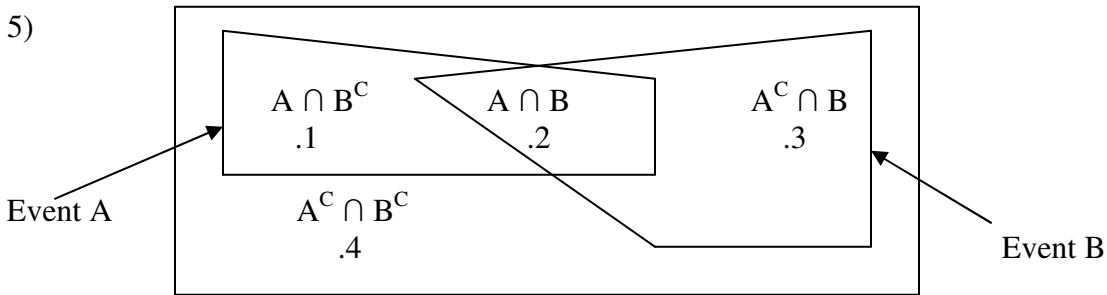
Sarah: $(28 - 22) / 4.97... = 1.206289...$

Jessica: $(1910 - 1530) / 315 = 1.206359...$

Since Jessica has the higher z-score, she did better overall.

B

5)



Using the information given, the above picture is then formed. Just add $(A \cap B^c)$ and $(A^c \cap B)$, or $.1 + .3$, to get .4.

C

6) Separating the planet into different regions is stratified sampling. The multiple SRS's, first in the Chu-Pu and then in the Rawar, signifies multistage sampling. Sampling every fourth person is systematic sampling. The only choice not used is cluster sampling, when one samples everything within a certain region. A

7) Natural variation in results caused by sampling is called **sampling error**. In other words, you will never get the exact proportion every time you do a sample. You will always be a little off. E

8) Two equations are made based off of the information:

$$(1) \quad \bar{x} + z^* (8/28)^{1/2} = 140^{1/2}$$

$$(2) \quad \bar{x} - z^* (8/28)^{1/2} = 110^{1/2}$$

Solving yields: $\bar{x} = 11.16\dots$, $z^* = .4445\dots$

Looking at the z-table, you can deduce that a .4445... corresponds to 34% confidence interval. B

9) This is a multinomial distribution because there are more than two outcomes. There are three. If there were two, it would be binomial. Also, each spin is independent and the probability for each outcome is equal and constant, which allows for this distribution to be multinomial. D

10) I. True

II. False; there is no **LINEAR** relationship between the data.

III. True

Therefore, C

11) Mean = $n \cdot p$, Standard deviation = $\sqrt{np(1-p)}$, where n is the number of trials and p is the probability of success.

$$10p = \sqrt{10p(1-p)}, \text{ solving for } p \text{ yields } 1/11. \quad \underline{A}$$

12) This is a Goodness for Fit Test. The chart will calculate the chi square statistic.

# of 3's	0	1	2	3	4	5	6
Expected(E)	60	80	40	10	10	24	24
Observed(O)	61	97	36	11	12	20	3
$(O - E)^2/E$.0167	3.1625	.4	.1	.4	0	14.45
Sum of $(O - E)^2/E$	18.97916						

Use the chi square table, 18.97916, to look for the chi square statistic with $(n - 1)$ degrees of freedom. "n" is the number of categories. So the degrees of freedom is $7 - 1$, or 6. looking at the chi square table, the p-value is very close to .005, which $< .01$. **D**

13) "Rsq" in the data refers to the coefficient of determination. Take the square root to find the correlation: $.798^{1/2} = .893$; however, upon closer inspection, you will see that the "Coeff" for shoe size is negative, which indicates a negative correlation. Thus, the correlation is $-.893$. **D**

14) This is a geometric distribution since the probability of a student having his homework is constant, a student having his homework is independent of the next, and she will stop at the first "success," which in this case is when a student does not have his homework. **B**

15) For complete independence, each cell should be the expected value of that cell.

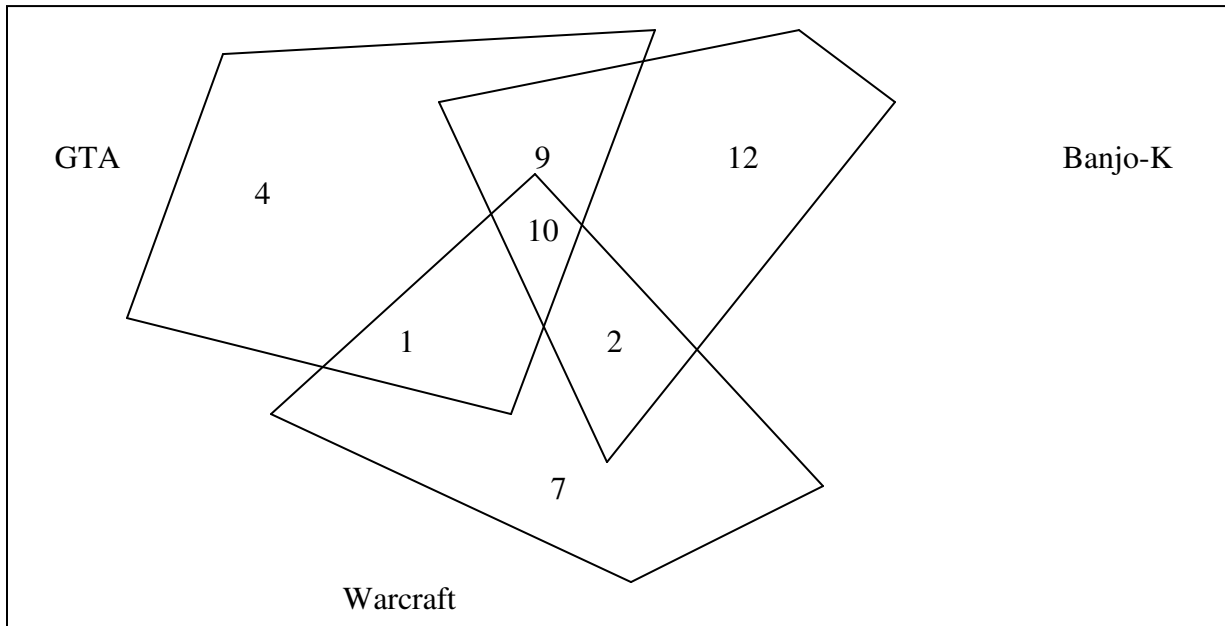
$$\text{Expected Value} = (\text{Row Total} \times \text{Column Total}) / \text{Table Total}$$

$$\text{For A: } [(10 + A) \times (A + 182)] / (227 + A) = A$$

$$\rightarrow A^2 + 192A + 1820 = 227A + A^2$$

$$\rightarrow A = 52 \quad \textbf{\underline{A}}$$

16) From the information, the following Venn diagram is formed



Since it is given the person is addicted to Warcraft, the person is being picked from a sample of 20 (total addicted to Warcraft). There are only 10 people who are addicted to both Warcraft and GTA, so the probability is $10/20 = 1/2$. **C**

17) Only 12 people are not addicted to Banjo-K. Of them, 8 are addicted to Warcraft. Thus, the probability is $8/12$, or $2/3$. **A**

18) The following are qualitative variables: Zip Code, Favorite game, Types of Bills, Area Code, House Number, Skin Texture. This is a total of 6. **C**

19) Events A and B could only be mutually exclusive, since, by definition, two mutually exclusive events can never occur at the same time. They are NOT independent because by definition, $P(X | Y) = P(X)$, if they are independent. **A**

20) Since this is a probability function, the area under the curve has to be equal to one. Thus, since the width is 1.6 ($1.2 - (-.4)$), the height must be $5/8$. Now the distance between .6 and 1.2 is $3/5$. So the probability of greater than .6 is $3/5 * 5/8 = 3/8$. **A**

21) Since only percentages are known for cats, it is impossible to say which has more geniuses because 72%, for example, could mean 72 out of 100 or 7200 out of 10000. Thus, more information is needed to determine which has the most geniuses. **D**

22) The key is to set up 2 equations that revolve around standardizing scores, as in question 4.

Using the formula: $z = (X - \mu) / \sigma$, the following equations are obtained.

For Alice: $(103 - \mu) / \sigma = 2.88$ (Z score for 99.8 percentile)

For Akash: $(49 - \mu) / \sigma = -1.97$ (Z score for 2.44 percentile)

Solving yields the following: $\mu = 70.934\dots$, $\sigma = 11.134\dots$. Thus, the answer is: A

- 23) I) False; The control group only receives the placebo.
II) False; **Blocking** is used to reduce **variation**
III) False; A well designed experiment may have more than one treatment
IV) False; Association does NOT imply causation. This is the common misconception about correlation.

Since none are true, the answer is A. A

24) Total Number of Rearrangements = $\frac{10!}{3!3!2!} = 50,400$. D

25) Since it is not known that S_1 and S_2 are independent of each other, it is impossible to find the variance of S_3 because it is required to add the variances of S_1 and S_2 to find the variance of S_3 . Therefore, the question is impossible to solve without further information. D

26)

- a) "the means of X and Y are 31 and 14 respectively"
b) the regression line will always go through the point (X mean, Y mean)
a) and b) yield the equation: $14 = 31a + b$ (1)

c) "one point used to make the least squares regression line was (16, 44) and its residual was 3.7"

- d) a residual of a point is: residual = observed y - predicted y
c) and d) yields the equation: $3.7 = 44 - (16a + b)$ (2)

Solving (1) and (2) yields: $a = 1.155\dots$, $b = -2.814\dots$ $a + b = -20.6595\dots$ D

27) Anything within seven standard deviations covers about 99.99999% of the data. So, the approximate area is 1. C

28) THEORETICALLY, the most accurate method for gathering data would be to go to every person in the population and get their numbers (number of bad hair days and the magnitude of smelly feet). This is also called a census. C

