Statistics 1040, Sections 007 & 009, Quiz 1 (20 Points)

Friday, August 31, 2007

Your Name: _____

Question 1: Controlled Experiments/Observational Studies I (**14 Points**)

(*hypothetical*) Does regularly taking vitamin C help protect people against flu?

A _____ was conducted to answer this question. The _____ were 500 volunteering college students, assigned _____ to two groups of 250 students. The students in the _____ took regularly a tablet of vitamin C, whereas those in the _____ took an identically looking and tasting pill, called _____. Neither participating students nor personell administrating drugs to them knew who was taking which pill, in other words, it was a _____ experiment. After a couple of months, the numbers of flu cases in both groups were compared ...

Fill the gaps in the paragraph above using the most appropriate words from the following list:

placebo
double–blind
haphazardly
treatment group
observational study
randomly
single–blind
vaccine
confounding factor
objects
control group
controlled experiment
subjects
polio

Please turn over!

Question 2: Controlled Experiments/Observational Studies II (**6 Points**)

In 1990, four passengers were killed by crashes on commuter airlines, compared to 39 killed on scheduled carriers (like United, TWA, and so forth). **True** or **false**? Circle your answer and explain: the data show that if you have to fly, it is safer to do so on a commuter airline.

Statistics 1040, Sections 007 & 009, Quiz 2 (20 Points)

Friday, September 7, 2007

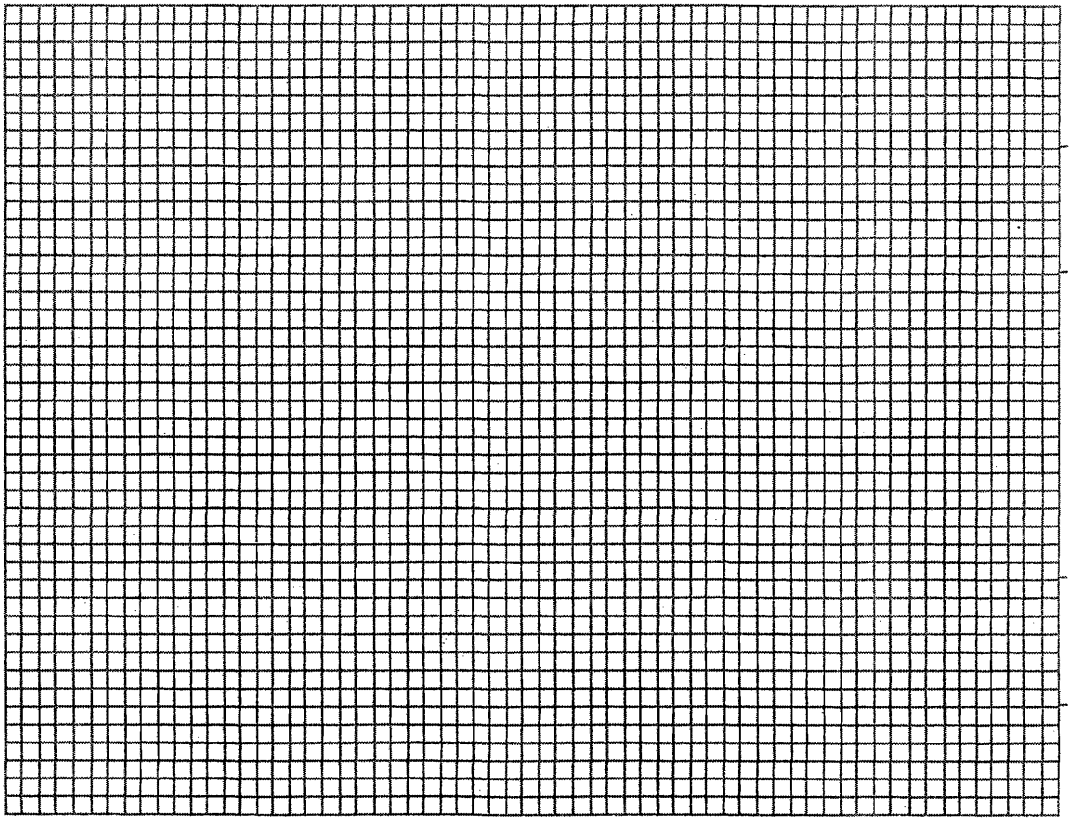
Your Name: _____

Question 1: Histograms (**14 Points**)

The following table is for the gestational age of 1210 babies:

Gestational Age	Number
230-250	47
250-270	206
270-290	731
290-310	199
310-330	27

Draw a histogram for these data on the graph paper provided. Make sure to label the axes.



Please turn over!

Question 2: Observational Studies / Controlled Experiments (**6 Points**)

For each of the following studies, determine whether the study in question was a **randomized controlled experiment** or an **observational study** (circle the correct answer).

- Twenty male employees and twenty female employees participate in research designed to compare "attitudes towards the Social Security System" of men and women. Each individual responds to a series of questions on a survey. Mean scores are computed for men and for women.

  randomized controlled experiment observational study

- A researcher wants to learn whether regularly taking zinc supplements may reduce the risk of getting a cold. Volunteers in this study chose to (or chose not to) take a zinc supplement.

  randomized controlled experiment observational study

- A researcher wants to learn about whether computer simulations help students better understand statistical concepts. She puts the names of 20 volunteers into a box and randomly draws the names of 10 people who will use computer simulations to learn statistical concepts. The other 10 study participants will use a conventional approach, without computer simulations, to learn the same concepts.

  randomized controlled experiment observational study

Statistics 1040, Sections 007 & 009, Quiz 3 (20 Points)

Friday, September 14, 2007

Your Name: _____

Question 1: The Average and the SD I (**10 Points**)

Here is a list of numbers:

0.7 1.6 9.8 3.2 5.4 0.8 7.7 6.3 2.2 4.1
8.1 6.5 3.7 0.6 6.9 9.9 8.8 3.1 5.7 9.1

1. Without doing any arithmetic, guess whether the average is around (i) 1, (ii) 5, or (iii) 10. **Circle your answer and explain.**

2. Without doing any arithmetic, guess whether the SD is around (i) 1, (ii) 3, or (iii) 6. **Circle your answer and explain.**

Question 2: The Average and the SD II (**10 Points**)

A study on college students found that the men had an average weight of about 66 kg and an SD of about 9 kg. The women had an average weight of about 55kg and an SD of about 9 kg (Note that 1 kg = 2.2 lb).

1. Just roughly, what percentage of the men weighted between 57 kg and 75 kg?
   Answer: _____ %
   Fill in your answer and explain.

2. If you took the men and women together, would the SD of their weights be (i) smaller than 9 kg, (ii) just about 9 kg, or (iii) bigger than 9 kg?
   Circle your answer and explain.

Statistics 1040, Sections 007 & 009, Quiz 4 (20 Points)

Friday, September 21, 2007

Your Name: _____

Question 1: Normal Approximation for Data (**20 Points**)

The Graduate Record Examination (GRE) is a test taken by college students who intend to pursue a graduate degree in the United States. For around 428,000 examinees who took the General GRE Test in 2001–02, the mean for the verbal ability portion of the exam was around 470 and the standard deviation was around 125. We can assume that the histogram follows a normal curve. **Show your work!**

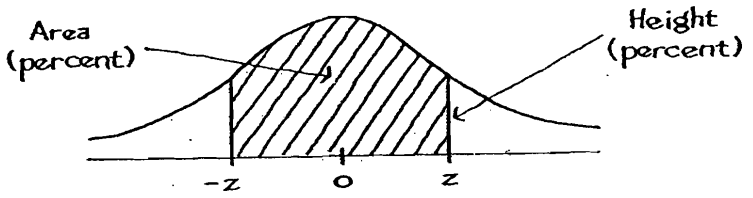
- (**7 Points**) The percentage of examinees who scored **less than 420** on the GRE test is roughly _____ %.

- (**7 Points**) The percentage of examinees who scored **between 370 and 645** is about _____ %.

- (**6 Points**) In order to be among the top 4%, a student must have obtained a minimum GRE score of about _____.

Please turn over!

Tables



A NORMAL TABLE

z	*Area*	*z*	*Area*	*z*	*Area*
0.00	0	1.50	86.64	3.00	99.730
0.05	3.99	1.55	87.89	3.05	99.771
0.10	7.97	1.60	89.04	3.10	99.806
0.15	11.92	1.65	90.11	3.15	99.837
0.20	15.85	1.70	91.09	3.20	99.863
0.25	19.74	1.75	91.99	3.25	99.885
0.30	23.58	1.80	92.81	3.30	99.903
0.35	27.37	1.85	93.57	3.35	99.919
0.40	31.08	1.90	94.26	3.40	99.933
0.45	34.73	1.95	94.88	3.45	99.944
0.50	38.29	2.00	95.45	3.50	99.953
0.55	41.77	2.05	95.96	3.55	99.961
0.60	45.15	2.10	96.43	3.60	99.968
0.65	48.43	2.15	96.84	3.65	99.974
0.70	51.61	2.20	97.22	3.70	99.978
0.75	54.67	2.25	97.56	3.75	99.982
0.80	57.63	2.30	97.86	3.80	99.986
0.85	60.47	2.35	98.12	3.85	99.988
0.90	63.19	2.40	98.36	3.90	99.990
0.95	65.79	2.45	98.57	3.95	99.992
1.00	68.27	2.50	98.76	4.00	99.9937
1.05	70.63	2.55	98.92	4.05	99.9949
1.10	72.87	2.60	99.07	4.10	99.9959
1.15	74.99	2.65	99.20	4.15	99.9967
1.20	76.99	2.70	99.31	4.20	99.9973
1.25	78.87	2.75	99.40	4.25	99.9979
1.30	80.64	2.80	99.49	4.30	99.9983
1.35	82.30	2.85	99.56	4.35	99.9986
1.40	83.85	2.90	99.63	4.40	99.9989
1.45	85.29	2.95	99.68	4.45	99.9991

Statistics 1040, Sections 007 & 009, Quiz 5 (20 Points)

Friday, September 28, 2007

Your Name: _____

Question 1: Measurement Error (**8 Points**)

You send a yardstick to a local laboratory for calibration, asking that the procedure be repeated three times. They report the following values:

35.96 inches 36.01 inches 36.03 inches

If you send the yardstick back for a fourth calibration, you would expect to get 36 inches, give or take

(a) .01 inches or so (b) .03 inches or so (c) .06 inches or so

Circle your answer and explain.

Question 2: Correlation (**6 Points**)

A teaching assistant gives a quiz to his section. There are 10 questions on the quiz and no part credit is given. After grading the papers, the TA writes down for each student the number of questions the student got right and the number wrong. The average number of right answers is 6.4 with an SD of 2.0; the average number of wrong answers is 3.6 with the same SD of 2.0.

The correlation coefficient between the number of right answers and the number of wrong answers is

(a) exactly 0 (b) -0.50 (c) +0.50
(d) -1.0 (e) +1.0 (f) -2.0 (g) +2.0
(h) can't tell without the data

Circle your answer and explain.

Please turn over!

Question 3: Correlation (**6 Points**)

For women age 25 and over in the U.S. in 2005, the relationship between age and educational level (years of schooling completed) can be summarized as follows:

average age $\approx$ 50 years, SD $\approx$ 16 years
average ed. level $\approx$ 13.2 years, SD $\approx$ 3.0 years, r $\approx$ -0.20

True or **false**, and explain: As you get older, you become less educated. If this statement is false, what accounts for the negative correlation?

Formulas:

$$\text{avg} = \frac{\text{sum of all numbers}}{\text{how many numbers}}$$

$$\text{SD} = \sqrt{\text{average of [(deviations from avg)}^2\text{]}}$$

Statistics 1040, Sections 007 & 009, Quiz 6 (20 Points)

Friday, October 12, 2007

Your Name: _____

Question 1: Regression **(20 Points)**

A selection of 65 varieties of cereal were tested for calories and sodium (in milligrams) for an one–cup serving. The results can be summarized as follows:

Average sodium = 240 mg;	SD = 131 mg;	
Average calories = 149 calories;	SD = 62 calories;	$r = 0.53$.

Show your work!

1. **(7 Points)** Find the equation of the regression line for predicting number of mg sodium in an one–cup serving of cereals from calories.

2. **(4 Points)** Predict the number of mg sodium in an one–cup serving of cereals that has 200 calories per cup.

Please turn over!

3. **(5 Points)** Find the r.m.s. error for predicting mg sodium from calories.

4. **(4 Points)** Explain why it would not be a good idea to use the information in the question to estimate the amount of sodium for a cereal with 350 calories per cup.

Formulas:

$$\text{r.m.s. error} = \sqrt{1 - r^2} \times \text{SD}_y$$

$$\text{slope} = r \times \frac{\text{SD}_y}{\text{SD}_x} \qquad \text{intercept} = \text{avg}_y - \text{slope} \times \text{avg}_x$$

Statistics 1040, Sections 007 & 009, Quiz 7 (20 Points)

Monday, October 22, 2007

Your Name: _____

Question 1: Chance/Probability I (**6 Points**)

A deck of cards is shuffled and the top two cards are placed face down on a table.
Circle your answer:

1. There is a 1 chance in 52 for the first card to be the ace of clubs.
   True / False
2. There is a 1 chance in 52 for the second card to be the ace of diamonds.
   True / False
3. The chance of getting the ace of clubs and then the ace of diamonds is $1/52 \times 1/52$.
   True / False

Question 2: Chance/Probability II (**3 Points**)

A coin is tossed six times. Two possible sequences of results are

(i) H T T H T H (ii) H H H H H H

(The coin must land on H or T in the order given; H = heads, T = tails).

Which of the following is correct?

Circle your answer:

1. Sequence (i) is more likely.
2. Sequence (ii) is more likely.
3. Both sequences are equally likely.

Please turn over!

Question 3: Chance/Probability III (**3 Points**)

Two cards will be dealt off the top of a well–shuffled deck. You have a choice:

(i) To win \$1 if at least one of the two cards is a queen.

(ii) To win \$1 if the first card is a queen.

Circle your answer:

1. Option (i) is better for winning.
2. Option (ii) is better for winning.
3. Both options are equally good for winning.

Question 4: Chance/Probability IV (**4 Points**)

One event has chance 1/2, another has chance 1/3. Fill in the blanks, using one phrase from each pair below, to make **two true sentences**. Write out both sentences.

"If you want to find the chance that **at least one of the two events / both events** will happen, check to see if they are **independent / mutually exclusive**. If so, you can **add / multiply** the chances."

Correct Sentence 1: "If you want to find the chance that **at least one of the two events** will happen, check to see if they are _____. If so, you can _____ the chances."

Correct Sentence 2: "If you want to find the chance that **both events** will happen, check to see if they are _____. If so, you can _____ the chances."

Question 5: Chance/Probability V (**4 Points**)

A coin is tossed 10 times. **Circle your answer:**

1. The chance of getting 10 heads in a row is 1/1,024.
   True / False
2. Given that the first 9 tosses were heads, the chance of getting 10 heads in a row is 1/2.
   True / False

Statistics 1040, Sections 007 & 009, Quiz 8 (20 Points)

Friday, October 26, 2007

Your Name: _____

Question 1: Law of Averages I (**2 Points**)

A box contains 10,000 tickets: 4,000 $\boxed{0}$'s and 6,000 $\boxed{1}$'s. And 10,000 draws will be made at random with replacement from this box. Which of the following best describes the situation? **Circle your answer.** You don't have to provide any explanations.

1. The number of 1's will be 6,000 exactly.
2. The number of 1's is very likely to equal 6,000, but there is also some small chance that it will not be equal to 6,000.
3. The number of 1's is likely to be different from 6,000, but the difference is likely to be small compared to 10,000.

Question 2: Law of Averages II (**8 Points**)

Circle your answer for each of the following four parts. You don't have to provide any explanations.

1. (**2 Points**) A die will be rolled some number of times, and you win \$1 if it shows an ace ($\boxed{\bullet}$) more than 20% of the time.

   Which is better: **60 rolls** or **600 rolls**?

2. (**2 Points**) As in 1.), but you win the dollar if the percentage of aces is more than 15%.

   Which is better: **60 rolls** or **600 rolls**?

3. (**2 Points**) As in 1.), but you win the dollar if the percentage of aces is between 15% and 20%.

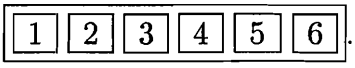
   Which is better: **60 rolls** or **600 rolls**?

4. (**2 Points**) As in 1.), but you win the dollar if the percentage of aces is exactly $16\frac{2}{3}\%$.

   Which is better: **60 rolls** or **600 rolls**?

Question 3: The Expected Value and Standard Error (**10 Points**)

One hundred draws are made at random with replacement from the box



1. (**1 Point**) If the sum of the draws is 321, what is their average?

2. (**1 Point**) If the average of the draws is 3.78, what is the sum?

3. (**8 Points**) Estimate the chance that the average of the draws is between 3 and 4. You have to make use of the normal curve to answer this part! **Show your work!**

Formulas:

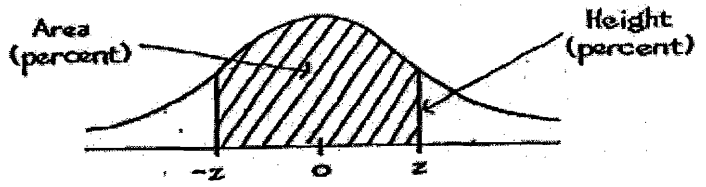
$$\text{box average} = \frac{\text{sum of all numbers in box}}{\text{how many numbers in box}}$$

$$\text{box SD} = \sqrt{\text{average of [(deviations from box average)}^2\text{]}}$$

$$\text{EV}_{sum} = \text{number of draws} \times \text{box average}$$

$$\text{SE}_{sum} = \sqrt{\text{number of draws}} \times \text{box SD}$$

Tables



A NORMAL TABLE

z	Area	z	Area	z	Area
0.00	0	1.50	86.64	3.00	99.730
0.05	3.99	1.55	87.89	3.05	99.771
0.10	7.97	1.60	89.04	3.10	99.806
0.15	11.92	1.65	90.11	3.15	99.837
0.20	15.85	1.70	91.09	3.20	99.863
0.25	19.74	1.75	91.99	3.25	99.885
0.30	23.58	1.80	92.81	3.30	99.903
0.35	27.37	1.85	93.57	3.35	99.919
0.40	31.08	1.90	94.26	3.40	99.933
0.45	34.73	1.95	94.88	3.45	99.944
0.50	38.29	2.00	95.45	3.50	99.953
0.55	41.77	2.05	95.96	3.55	99.961
0.60	45.15	2.10	96.43	3.60	99.968
0.65	48.43	2.15	96.84	3.65	99.974
0.70	51.61	2.20	97.22	3.70	99.978
0.75	54.67	2.25	97.56	3.75	99.982
0.80	57.63	2.30	97.86	3.80	99.986
0.85	60.47	2.35	98.12	3.85	99.988
0.90	63.19	2.40	98.36	3.90	99.990
0.95	65.79	2.45	98.57	3.95	99.992
1.00	68.27	2.50	98.76	4.00	99.9937
1.05	70.63	2.55	98.92	4.05	99.9949
1.10	72.87	2.60	99.07	4.10	99.9959
1.15	74.99	2.65	99.20	4.15	99.9967
1.20	76.99	2.70	99.31	4.20	99.9973
1.25	78.87	2.75	99.40	4.25	99.9979
1.30	80.64	2.80	99.49	4.30	99.9983
1.35	82.30	2.85	99.56	4.35	99.9986
1.40	83.85	2.90	99.63	4.40	99.9989
1.45	85.29	2.95	99.68	4.45	99.9991

Statistics 1040, Sections 007 & 009, Quiz 9 (20 Points)

Friday, November 2, 2007

Your Name: _____

Question 1: Normal Approximation for Probability Histograms I (**12 Points**)

A programmer is working on a new program, COIN, to simulate tossing a coin. As a preliminary test, he sets up the program to do one million tosses. The program returns with a count of 502,015 heads. The programmer looks at this and thinks: "*Hmmm. Two thousand and fifteen off. That's a lot. No, wait. Compare it to the million. Two thousand – forget the fifteen – out of a million is two out of a thousand. That's one in five hundred. One fifth of a percent. Very small. Good. COIN passes.*"

Do you agree that COIN passes? Answer **yes** or **no**, and **explain**. You should use box model calculations to support your answer.

Please turn over!

Question 2: Normal Approximation for Probability Histograms II (**8 Points**)

A coin is tossed 100 times. True or false? **Just circle your answer.** You don't have to give any explanation. **Answer each of the following questions separately!**

1. The expected value for the number of heads is 50.

   True / False

2. The expected value for the number of heads is 50, give or take 5 or so.

   True / False

3. The number of heads will be 50.

   True / False

4. The number of heads will be around 50, give or take 2 or so.

   True / False

Formulas:

$$\text{box average} = \frac{\text{sum of all numbers in box}}{\text{how many numbers in box}}$$

$$\text{box SD} = \sqrt{\text{average of } [(\text{deviations from box average})^2]}$$

$$\text{EV}_{sum} = \text{number of draws} \times \text{box average}$$

$$\text{SE}_{sum} = \sqrt{\text{number of draws}} \times \text{box SD}$$

Shortcut formulas for a box that contains only *two* different numbers:

$$\text{average} = \frac{(\text{smaller} \times \text{how many}) + (\text{bigger} \times \text{how many})}{\text{how many tickets in the box}}$$

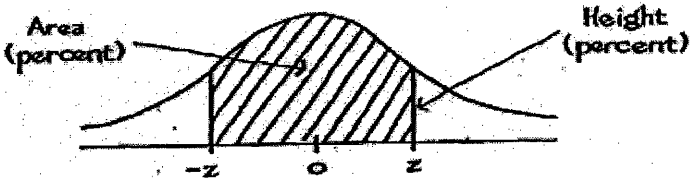
$$\text{SD} = (\text{bigger - smaller}) \times \sqrt{\frac{\text{fraction}}{\text{bigger}} \times \frac{\text{fraction}}{\text{smaller}}}$$

Shortcut formulas for a box that contains only $\boxed{0}$'s and $\boxed{1}$'s:

$$\text{average} = \frac{\text{number of } \boxed{1}\text{'s}}{\text{how many tickets in the box}}$$

$$\text{SD} = \sqrt{\frac{\text{fraction}}{\text{of } \boxed{1}\text{'s}} \times \frac{\text{fraction}}{\text{of } \boxed{0}\text{'s}}}$$

Tables



A NORMAL TABLE

z	*Area*	*z*	*Area*	*z*	*Area*
0.00	0	1.50	86.64	3.00	99.730
0.05	3.99	1.55	87.89	3.05	99.771
0.10	7.97	1.60	89.04	3.10	99.806
0.15	11.92	1.65	90.11	3.15	99.837
0.20	15.85	1.70	91.09	3.20	99.863
0.25	19.74	1.75	91.99	3.25	99.885
0.30	23.58	1.80	92.81	3.30	99.903
0.35	27.37	1.85	93.57	3.35	99.919
0.40	31.08	1.90	94.26	3.40	99.933
0.45	34.73	1.95	94.88	3.45	99.944
0.50	38.29	2.00	95.45	3.50	99.953
0.55	41.77	2.05	95.96	3.55	99.961
0.60	45.15	2.10	96.43	3.60	99.968
0.65	48.43	2.15	96.84	3.65	99.974
0.70	51.61	2.20	97.22	3.70	99.978
0.75	54.67	2.25	97.56	3.75	99.982
0.80	57.63	2.30	97.86	3.80	99.986
0.85	60.47	2.35	98.12	3.85	99.988
0.90	63.19	2.40	98.36	3.90	99.990
0.95	65.79	2.45	98.57	3.95	99.992
1.00	68.27	2.50	98.76	4.00	99.9937
1.05	70.63	2.55	98.92	4.05	99.9949
1.10	72.87	2.60	99.07	4.10	99.9959
1.15	74.99	2.65	99.20	4.15	99.9967
1.20	76.99	2.70	99.31	4.20	99.9973
1.25	78.87	2.75	99.40	4.25	99.9979
1.30	80.64	2.80	99.49	4.30	99.9983
1.35	82.30	2.85	99.56	4.35	99.9986
1.40	83.85	2.90	99.63	4.40	99.9989
1.45	85.29	2.95	99.68	4.45	99.9991

Statistics 1040, Sections 007 & 009, Quiz 10 (20 Points)

Friday, November 16, 2007

Your Name: _____

Question 1: The Accuracy of Percentages (**20 Points**)

The National Assessment of Educational Progress administers standardized achievement tests to nationwide samples of 17–year–olds in school. One year, the tests covered history and literature. You may assume that a simple random sample of size 6,000 was taken. Only 36.1% of the students in the sample knew that Chaucer wrote *The Canterbury Tales*, but 95.2% knew that Edison invented the light bulb.

1. (**10 Points**) Is it possible to find a 95% confidence interval for the percentage of all 17–year–olds in school who knew that Chaucer wrote *The Canterbury Tales*?
   Yes or **No**? – Circle your answer.
   If yes, calculate this CI (and show your work). If no, clearly indicate why this is not possible.

2. (**10 Points**) Is it possible to find a 95% confidence interval for the percentage of all 17–year–olds in school who knew that Edison invented the light bulb?
   Yes or **No**? – Circle your answer.
   If yes, calculate this CI (and show your work). If no, clearly indicate why this is not possible.

Please turn over!

Formulas:

$$\text{box average} = \frac{\text{sum of all numbers in box}}{\text{how many numbers in box}}$$

$$\text{box SD} = \sqrt{\text{average of [(deviations from box average)}^2\text{]}}$$

$$\text{EV}_{sum} = \text{number of draws} \times \text{box average}$$

$$\text{SE}_{sum} = \sqrt{\text{number of draws}} \times \text{box SD}$$

$$\text{EV}_{avg} = \text{box average} \qquad \text{SE}_{avg} = \frac{\text{SE}_{sum}}{\text{number of draws}}$$

Shortcut formulas for a box that contains only *two* different numbers:

$$\text{average} = \frac{(\text{smaller} \times \text{how many}) + (\text{bigger} \times \text{how many})}{\text{how many tickets in the box}}$$

$$\text{SD} = (\text{bigger - smaller}) \times \sqrt{\frac{\text{fraction}}{\text{bigger}} \times \frac{\text{fraction}}{\text{smaller}}}$$

Shortcut formulas for a box that contains only $\boxed{0}$'s and $\boxed{1}$'s:

$$\text{average} = \frac{\text{number of } \boxed{1}\text{'s}}{\text{how many tickets in the box}}$$

$$\text{SD} = \sqrt{\frac{\text{fraction}}{\text{of } \boxed{1}\text{'s}} \times \frac{\text{fraction}}{\text{of } \boxed{0}\text{'s}}}$$

$$\text{EV}_{\%} = \%\text{ of } \boxed{1}\text{'s in the box}$$

$$\text{SE}_{\%} = \frac{\text{SEsum}}{\text{number of draws}} \times 100\%$$

Statistics 1040, Sections 007 & 009, Quiz 11 (20 Points)

Wednesday, November 28, 2007

Your Name: _____

Question 1: The Accuracy of Averages **(20 Points)**

One year, there were about 3,000 institutions of higher learning in the U.S. (including junior colleges and community colleges). As part of a continuing study of higher education, the Carnegie Commission took a simple random sample of 400 of these institutions. The average enrollment in the 400 sample schools was 3,700, and the SD was 6,500. The Commission estimates the average enrollment at all 3,000 institutions to be around 3,700; they put a give–or–take number of 325 on this estimate.

Say whether each of the following statements is true or false, and explain. If you need more information to decide, say what you need and why.

1. **(4 Points)** An approximate 68%–confidence interval for the average enrollment of all 3,000 institutions runs from 3,375 to 4,025.

   Yes, **No**, or **Need more information**.
   Circle your answer and explain.

2. **(4 Points)** If a statistician takes a simple random sample of 400 institutions out of 3,000, and goes one SE either way from the average enrollment of the 400 sample schools, there is about a 68% chance that this interval will cover the average enrollment of all 3,000 schools.

   Yes, **No**, or **Need more information**.
   Circle your answer and explain.

Please turn over!

3. (**4 Points**) About 68% of the schools in the sample had enrollments in the range 3,700 $\pm$ 6,500.

   Yes, **No**, or **Need more information**.
   Circle your answer and explain.

4. (**4 Points**) It is estimated that 68% of the 3,000 institutions of higher learning in the U.S. enrolled between 3,700 - 325 = 3,375 and 3,700 + 325 = 4,025 students.

   Yes, **No**, or **Need more information**.
   Circle your answer and explain.

5. (**4 Points**) The normal curve can't be used to figure confidence levels here at all, because the data don't follow the normal curve.

   Yes, **No**, or **Need more information**.
   Circle your answer and explain.

Formulas:

$$\text{box average} = \frac{\text{sum of all numbers in box}}{\text{how many numbers in box}}$$

$$\text{box SD} = \sqrt{\text{average of } [(\text{deviations from box average})^2]}$$

$$\text{EV}_{sum} = \text{number of draws} \times \text{box average}$$

$$\text{SE}_{sum} = \sqrt{\text{number of draws}} \times \text{box SD}$$

$$\text{EV}_{avg} = \text{box average} \qquad \text{SE}_{avg} = \frac{\text{SE}_{sum}}{\text{number of draws}}$$

Shortcut formulas for a box that contains only *two* different numbers:

$$\text{average} = \frac{(\text{smaller} \times \text{how many}) + (\text{bigger} \times \text{how many})}{\text{how many tickets in the box}}$$

$$\text{SD} = (\text{bigger - smaller}) \times \sqrt{\frac{\text{fraction}}{\text{bigger}} \times \frac{\text{fraction}}{\text{smaller}}}$$

Shortcut formulas for a box that contains only $\boxed{0}$'s and $\boxed{1}$'s:

$$\text{average} = \frac{\text{number of } \boxed{1} \text{ 's}}{\text{how many tickets in the box}}$$

$$\text{SD} = \sqrt{\frac{\text{fraction}}{\text{of } \boxed{1} \text{ 's}} \times \frac{\text{fraction}}{\text{of } \boxed{0} \text{ 's}}}$$

$$\text{EV}_{\%} = \% \text{ of } \boxed{1} \text{ 's in the box}$$

$$\text{SE}_{\%} = \frac{\text{SE}_{\text{sum}}}{\text{number of draws}} \times 100\%$$

Statistics 1040, Sections 007 & 009, Quiz 12 (20 Points)

Wednesday, December 5, 2007

Your Name: _____

Question 1: Tests of Significance I (**20 Points**)

A random sample of 18 Penn State students was selected. The nose length of each student was measured in millimeters, and was recorded in the following data set:

41 57 43 42 55 35 36 40 45 55 55 36 45 44 45 47 37 48

A student thinks that the average nose length of Penn State students is fairly large. He thinks that the average nose length could be more than 44 millimeters. Does this sample of 18 students provide evidence to suggest that the average nose length is more than 44 millimeters? State the null and the alternative hypothesis, calculate the test statistic (after finding the average and SD of the sample), obtain the P–value, and clearly state your conclusions. Assume that the nose lengths follow the normal curve.

Indicate whether this is a **z–test, t–test, or 2–sample z–test**. Circle your answer and explain why you have chosen that test. (If we **cannot** conduct a test, state so and explain why not!)

Show your work!

Memory Aids

Please note that these are provided for your convenience, but it is your responsibility to know how and when to use them.

$$\text{rms error} = \sqrt{1 - r^2} \times SD_Y$$

$$\text{slope} = r \times \frac{SD_Y}{SD_X}$$

$$\text{intercept} = \text{ave}_Y - \text{slope} \times \text{ave}_X$$

$$SD^+ = \sqrt{\frac{\text{number of draws}}{\text{number of draws} - 1}} \times SD$$

$$SD_{\text{box}} = \sqrt{\text{fraction of 0's} \times \text{fraction of 1's}}$$

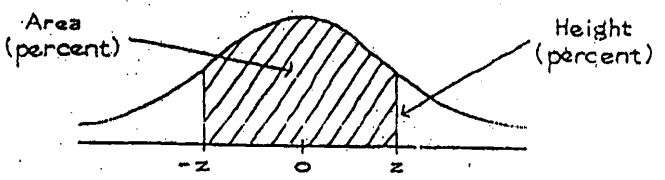
$EV_{\text{sum}} = \text{number of draws} \times \text{ave}_{\text{box}}$
$SE_{\text{sum}} = \sqrt{\text{number of draws}} \times SD_{\text{box}}$

$EV_{\text{ave}} = \text{ave}_{\text{box}}$
$SE_{\text{ave}} = \dfrac{SE_{\text{sum}}}{\text{number of draws}}$

$EV_{\%} =$ % of 1's in the box
$SE_{\%} = \left(\dfrac{SE_{\text{sum}}}{\text{number of draws}}\right) \times 100\%$

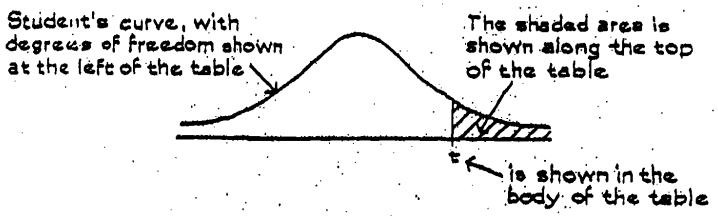
$SE_{\text{diff}} = \sqrt{a^2 + b^2}$ where a is the SE for the first quantity, b is the SE for the second quantity, and the two quantities are independent

A NORMAL TABLE



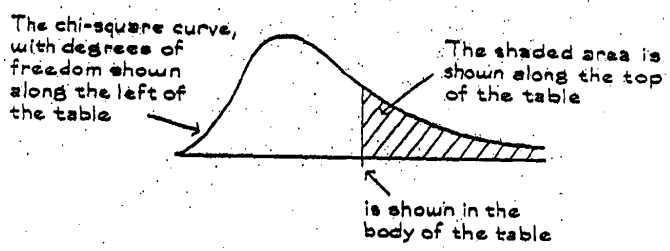
z	Area	z	Area	z	Area
0.00	0	1.50	86.64	3.00	99.730
0.05	3.99	1.55	87.89	3.05	99.771
0.10	7.97	1.60	89.04	3.10	99.806
0.15	11.92	1.65	90.11	3.15	99.837
0.20	15.85	1.70	91.09	3.20	99.863
0.25	19.74	1.75	91.99	3.25	99.885
0.30	23.58	1.80	92.81	3.30	99.903
0.35	27.37	1.85	93.57	3.35	99.919
0.40	31.08	1.90	94.26	3.40	99.933
0.45	34.73	1.95	94.88	3.45	99.944
0.50	38.29	2.00	95.45	3.50	99.953
0.55	41.77	2.05	95.96	3.55	99.961
0.60	45.15	2.10	96.43	3.60	99.968
0.65	48.43	2.15	96.84	3.65	99.974
0.70	51.61	2.20	97.22	3.70	99.978
0.75	54.67	2.25	97.56	3.75	99.982
0.80	57.63	2.30	97.86	3.80	99.986
0.85	60.47	2.35	98.12	3.85	99.988
0.90	63.19	2.40	98.36	3.90	99.990
0.95	65.79	2.45	98.57	3.95	99.992
1.00	68.27	2.50	98.76	4.00	99.9937
1.05	70.63	2.55	98.92	4.05	99.9949
1.10	72.87	2.60	99.07	4.10	99.9959
1.15	74.99	2.65	99.20	4.15	99.9967
1.20	76.99	2.70	99.31	4.20	99.9973
1.25	78.87	2.75	99.40	4.25	99.9979
1.30	80.64	2.80	99.49	4.30	99.9983
1.35	82.30	2.85	99.56	4.35	99.9986
1.40	83.85	2.90	99.63	4.40	99.9989
1.45	85.29	2.95	99.68	4.45	99.9991

A *t*-TABLE



Degrees of freedom	*25%*	*10%*	*5%*	*2.5%*	*1%*	*0.5%*
1	1.00	3.08	6.31	12.71	31.82	63.66
2	0.82	1.89	2.92	4.30	6.96	9.92
3	0.76	1.64	2.35	3.18	4.54	5.84
4	0.74	1.53	2.13	2.78	3.75	4.60
5	0.73	1.48	2.02	2.57	3.36	4.03
6	0.72	1.44	1.94	2.45	3.14	3.71
7	0.71	1.41	1.89	2.36	3.00	3.50
8	0.71	1.40	1.86	2.31	2.90	3.36
9	0.70	1.38	1.83	2.26	2.82	3.25
10	0.70	1.37	1.81	2.23	2.76	3.17
11	0.70	1.36	1.80	2.20	2.72	3.11
12	0.70	1.36	1.78	2.18	2.68	3.05
13	0.69	1.35	1.77	2.16	2.65	3.01
14	0.69	1.35	1.76	2.14	2.62	2.98
15	0.69	1.34	1.75	2.13	2.60	2.95
16	0.69	1.34	1.75	2.12	2.58	2.92
17	0.69	1.33	1.74	2.11	2.57	2.90
18	0.69	1.33	1.73	2.10	2.55	2.88
19	0.69	1.33	1.73	2.09	2.54	2.86
20	0.69	1.33	1.72	2.09	2.53	2.85
21	0.69	1.32	1.72	2.08	2.52	2.83
22	0.69	1.32	1.72	2.07	2.51	2.82
23	0.69	1.32	1.71	2.07	2.50	2.81
24	0.68	1.32	1.71	2.06	2.49	2.80
25	0.68	1.32	1.71	2.06	2.49	2.79

A CHI-SQUARE TABLE



Degrees of freedom	*99%*	*95%*	*90%*	*70%*	*50%*	*30%*	*10%*	*5%*	*1%*
1	0.00016	0.0039	0.016	0.15	0.46	1.07	2.71	3.84	6.64
2	0.020	0.10	0.21	0.71	1.39	2.41	4.60	5.99	9.21
3	0.12	0.35	0.58	1.42	2.37	3.67	6.25	7.82	11.34
4	0.30	0.71	1.06	2.20	3.36	4.88	7.78	9.49	13.28
5	0.55	1.14	1.61	3.00	4.35	6.06	9.24	11.07	15.09
6	0.87	1.64	2.20	3.83	5.35	7.23	10.65	12.59	16.81
7	1.24	2.17	2.83	4.67	6.35	8.38	12.02	14.07	18.48
8	1.65	2.73	3.49	5.53	7.34	9.52	13.36	15.51	20.09
9	2.09	3.33	4.17	6.39	8.34	10.66	14.68	16.92	21.67
10	2.56	3.94	4.86	7.27	9.34	11.78	15.99	18.31	23.21
11	3.05	4.58	5.58	8.15	10.34	12.90	17.28	19.68	24.73
12	3.57	5.23	6.30	9.03	11.34	14.01	18.55	21.03	26.22
13	4.11	5.89	7.04	9.93	12.34	15.12	19.81	22.36	27.69
14	4.66	6.57	7.79	10.82	13.34	16.22	21.06	23.69	29.14
15	5.23	7.26	8.55	11.72	14.34	17.32	22.31	25.00	30.58
16	5.81	7.96	9.31	12.62	15.34	18.42	23.54	26.30	32.00
17	6.41	8.67	10.09	13.53	16.34	19.51	24.77	27.59	33.41
18	7.00	9.39	10.87	14.44	17.34	20.60	25.99	28.87	34.81
19	7.63	10.12	11.65	15.35	18.34	21.69	27.20	30.14	36.19
20	8.26	10.85	12.44	16.27	19.34	22.78	28.41	31.41	37.57

Source: Adapted from p. 112 of Sir R. A. Fisher, *Statistical Methods for Research Workers* (Edinburgh: Oliver & Boyd, 1958).