Statistics 1040, Section 006, Quiz 1 (20 Points)

Friday, September 2, 2005

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
haphàzardly
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, Section 006, Quiz 2 (20 Points)

Friday, September 9, 2005

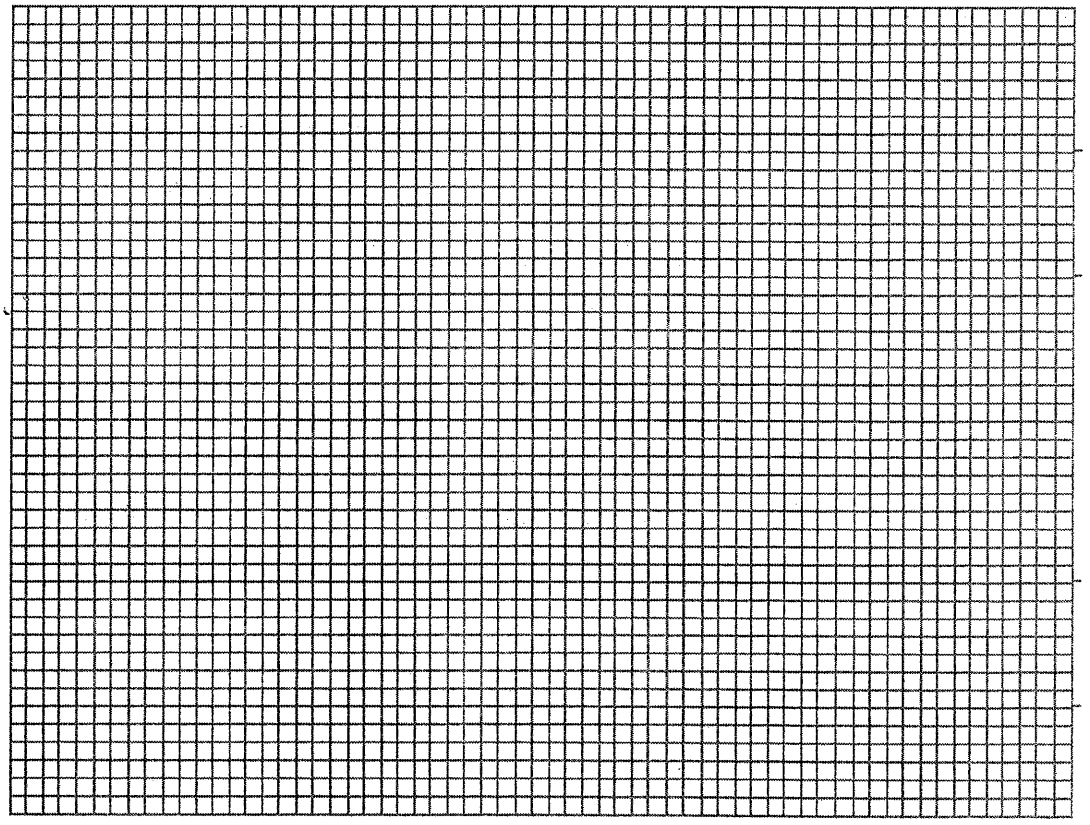
Your Name: _____

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, Section 006, Quiz 3 (20 Points)

Friday, September 16, 2005

Your Name: _____

Question 1: Measures of Center and Spread (**20 Points**)

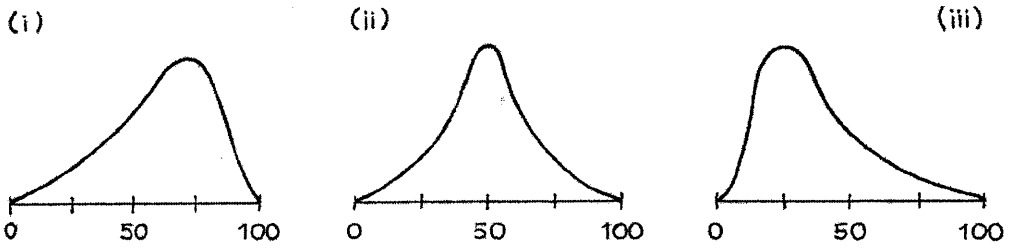
1. **(10 Points)** Find the average and the standard deviation of the following two lists of numbers:

Numbers	Average	Standard deviation
17, 17, 17, 17, 17	_____	_____
15, 16, 17, 18, 19	_____	_____

Show your work! Use formulas provided on the back where necessary.

Please turn over!

2. **(10 Points)** Below are sketches of histograms for three lists.



(a) In a scrambled order, the averages are 40, 50, 60. Match the histograms with averages:

Histogram (i): average =

Histogram (ii): average =

Histogram (iii): average =

(b) Match the histograms with the description (circle your answer):

- The median is less than the average. Histogram (i), (ii), or (iii).
- The median is about equal to the average. Histogram (i), (ii), or (iii).
- The median is bigger than the average. Histogram (i), (ii), or (iii).

(c) The SD for histogram (i) is a lot smaller than that for histogram (iii). **True** or **false**? Circle your answer and explain:

Formulas:

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

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

Statistics 1040, Section 006, Quiz 4 (20 Points)

Friday, September 23, 2005

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 (http://ftp.ets.org/pub/gre/994950.pdf). **Show your work!**

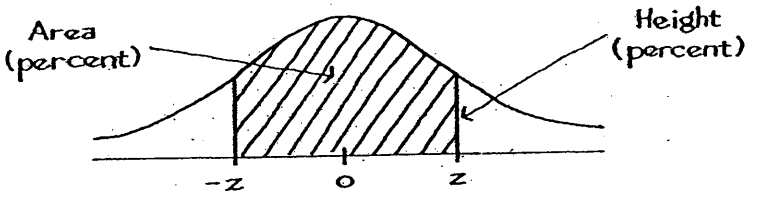
- (**7 Points**) The percentage of examinees who scored **less than 570** on the GRE test is roughly _____ %.

- (**7 Points**) The percentage of examinees who scored **between 270 and 620** is about _____ %.

- (**6 Points**) In order to be among the top 5%, 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, Section 006, Quiz 5 (20 Points)

Friday, September 30, 2005

Your Name: _____

Question 1: Change Of Scale (**12 Points**)

Conversion of temperature from Celsius to Fahrenheit is another example of what statisticians call **change of scale**. The formula for conversion is

$$F^\circ = \frac{9}{5}C^\circ + 32^\circ.$$

A group of people have an average body temperature of 38.0° Celsius, with a standard deviation of 0.3° Celsius.

1. (**8 Points**) If we translate these results into degrees Fahrenheit, the average temperature would be _____ degrees Fahrenheit, with a standard deviation of _____ degrees Fahrenheit.

2. (**4 Points**) Someone's temperature is 1.8 standard deviations above average on the Celsius scale. When converting this temperature to standard units for an investigator who is using the Fahrenheit scale, we have to report _____ standard units to this investigator.

Question 2: Correlation (**8 Points**)

1. (**4 Points**) If women always marry men who were five years older, the correlation between ages of husbands and wives would be _____. **Choose one of the options below, and explain.**

2. (**4 Points**) In reality, the correlation between ages of husbands and wives in the US is _____. **Choose one of the options below, and explain.**

Options: (a) exactly −1 (b) close to −1 (c) close to 0
(d) close to 1 (e) exactly 1

Statistics 1040, Section 006, Quiz 6 (20 Points)

Friday, October 14, 2005

Your Name: _____

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

A researcher is interested in the extent to which lead particles emitted from automobiles are absorbed by competitive cyclists. For a large group of cyclists they found the following:

Hours of training:	average = 16.2,	SD = 5.9	
Blood lead (μmol/L):	average = .42,	SD = .19,	r=0.6.

The scatter plot of the data is football–shaped.

Show your work!

1. (**7 Points**) Find the equation of the regression line for predicting blood lead from training time.

   Equation: _____

2. (**3 Points**) Use the regression equation from part 1. to predict the blood lead for a cyclist who trained for 21 hours.

   Answer: _____

Please turn over!

3. **(5 Points)** Find the r.m.s. error for predicting blood lead from training time of cyclist.

   Answer: _____

4. **(5 Points)** Would you be surprised to learn that a cyclist who trained for 3 hours had a blood lead of .8 μmol/L? Support your answer with a brief explanation and calculation.

   Answer: **Yes, surprised / No, not surprised**

   Explanation:

Formulas:

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

$$\text{slope} = r \times \frac{\text{SD}_y}{\text{SD}_x}$$

$$\text{intercept} = \text{avg}_y - \text{slope} \times \text{avg}_x$$

Statistics 1040, Section 006, Quiz 7 (20 Points)

Friday, October 21, 2005

Your Name: _____

Question 1: Chance/Probability (**20 Points**)

1. A deck of 52 cards is shuffled and two cards are drawn without replacement.

   (a) (**3 Points**) What is the chance that the first card is a ♡ or a ♢?

   (b) (**4 Points**) What is the chance that the first card is a ♡ and the second card is a ♢ ?

   (c) (**4 Points**) What is the chance that both cards are ♡ ?

   (d) (**4 Points**) What is the chance that neither card is a ♡?

2. (**5 Points**) There are two options:

   (a) You toss a coin 100 times; on each toss, if it lands heads you win \$1, if it lands tails you lose \$1.

   (b) You draw 100 times at random with replacement from the box $\boxed{1}\boxed{0}$. On each draw, you are paid (in dollars) the number on the ticket.

   Which option is better? Or are they the same? Explain briefly.

Statistics 1040, Section 006, Quiz 8 (20 Points)

Friday, October 28, 2005

Your Name: _____

Question 1: EV, SE, and Normal Curve (**15 Points**)

In a certain town, there are 40,000 registered voters, of whom 15,000 are Democrats. A survey organization is about to take a simple random sample of 1,000 registered voters. **Show your work!**

1. (**5 Points**) Find the box model.

2. (**5 Points**) The expected number of Democrats in this sample of 1,000 is _____ with an SE of _____.

3. (**5 Points**) The chance that **at least** 500 of the voters in the sample are Democrats is about _____ %.

Please turn over!

Question 2: Law of Averages (**5 Points**)

A box contains red and green marbles; there are more red marbles than green ones. Marbles are drawn one at a time from the box, at random with replacement. You win a dollar if a red marble is drawn more often than a green one. There are two choices:

- A: 50 draws are made from the box.
- B: 500 draws are made from the box.

Choose (i.e., circle) one of the four options below. **Explain your answer.**

1. A gives a better chance of winning.
2. B gives a better chance of winning.
3. A and B give the same chance of winning.
4. Can't tell without more information.

Formulas:

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

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

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

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

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

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

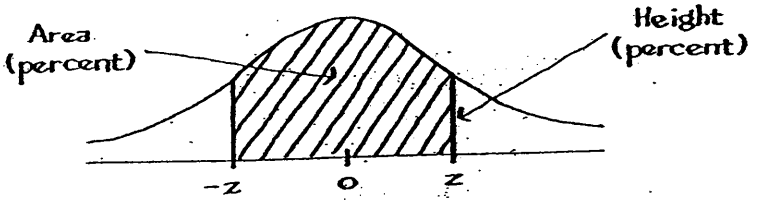
$$\text{SD} = (\text{bigger - smaller}) \times \sqrt{\frac{\text{fraction}}{\text{bigger}} \times \frac{\text{fraction}}{\text{smaller}}}$$

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, Section 006, Quiz 9 (20 Points)

Friday, November 4, 2005

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:

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

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

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

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

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

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

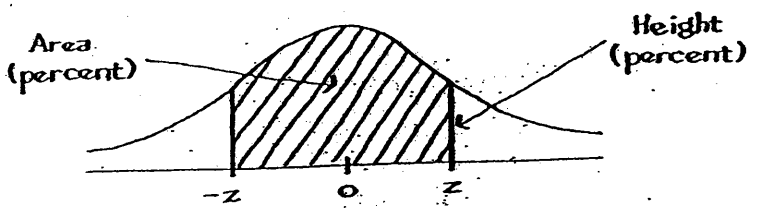
$$\text{SD} = (\text{bigger - smaller}) \times \sqrt{\frac{\text{fraction}}{\text{bigger}} \times \frac{\text{fraction}}{\text{smaller}}}$$

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, Section 006, Quiz 10 (20 Points)

Friday, November 18, 2005

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:

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

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

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

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

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

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:

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

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

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

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

Statistics 1040, Section 006, Quiz 11 (20 Points)

Wednesday, November 30, 2005

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:

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

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

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

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

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

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, Section 006, Quiz 12 (20 Points)

Wednesday, December 7, 2005

Your Name: _____

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

Many companies are experimenting with "flex–time", which is supposed to reduce absenteeism. One company employees have averaged 6.3 days off work in the past. The company introduces "flex–time" and a year later a simple random sample of 100 employees is selected. They average 5.5 days off work with a standard deviation of 2.9. Test to determine if "flex–time" reduces absenteeism. Clearly state the null and alternative hypotheses, calculate the appropriate test statistic, find the P–value, and state your conclusion.

Show your work!

1. **(3 Points)** State the null and the alternative hypotheses for this problem, in words and in terms of the box model.

2. **(5 Points)** Calculate the appropriate test statistic.

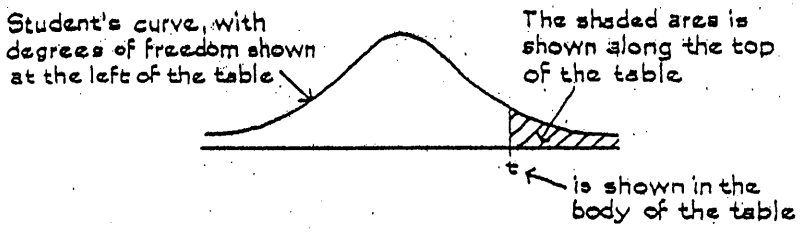
Please turn over!

3. **(4 Points)** Obtain the (approximate) P–value (use the appropriate table!).

4. **(6 Points)** State your conclusions in terms of rejecting (or not rejecting) the null hypothesis and in your own words. (If approriate, also speak of statistically significant or highly statistically significant.)

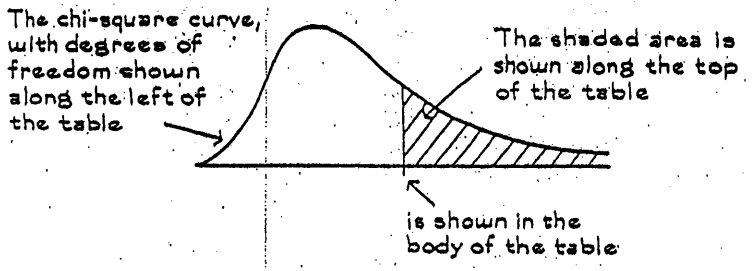
5. **(2 Points)** The test that has to be used in this question is a
   z–test / t–test / 2–sample z–test.
   Circle your answer and explain briefly why you chose this particular test to answer the question.

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).

Memory Aids

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

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

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

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

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

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

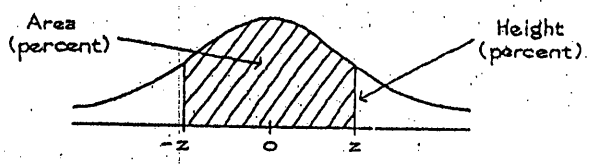
$$EV_{\text{sum}} = \text{number of draws} \times \text{ave}_{\text{box}}$$
$$SE_{\text{sum}} = \sqrt{\text{number of draws}} \times SD_{\text{box}}$$

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

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

$SE_{\text{diff}} = \sqrt{a^2 + b^2}$ 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