Statistics 1040, Section 004, Midterm 1 (200 Points)

Friday, February 13, 2004

Your Name: _____

Question 1: Observational Studies and Experiments **(50 Points)**

Forty men have agreed to be subjects in an experiment on the effectiveness of a new throat spray that is supposed to reduce snoring. These people were randomly divided into a treatment group and a control group.

1. **(10 Points)** Is the study described in this article an **observational study** or a **controlled experiment**? Circle your answer and explain!

2. **(15 Points)** When the groups were chosen, each of 40 men was given a spray bottle and told to use it every night for a week. Why are *all* the men given spray bottles? Are all the contents the same? Explain!

3. **(15 Points)** To find out whether the spray works, at the end of the study the men were asked the following question:

   > *Do you think that with the spray you snore more than before, less than before, or about the same as before?*

   It turns out that on average the men in *both groups* thought that they snored less than before. One explanation why the treatment group might answer this way is that the spray works. But why did the men in the control group answer this way? Provide two plausible reasons.

4. **(10 Points)** Suggest a better way to evaluate the effectiveness of the spray.

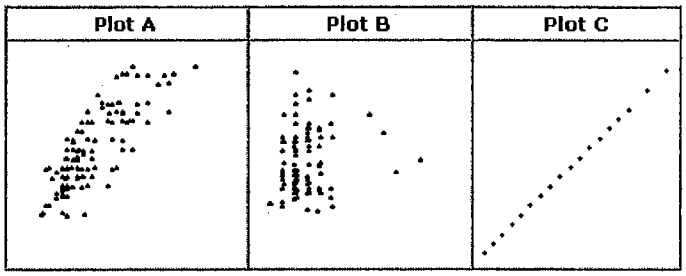
Question 2: Correlation **(30 Points)**

1. **(18 Points)** The following three plots are based on the heights and weights of 126 college students. Match each plot with the description:

   (a) X = Weight in pounds and Y = Height in inches for the 126 people;

   (b) X = Age in years and Y = Weight in pounds;

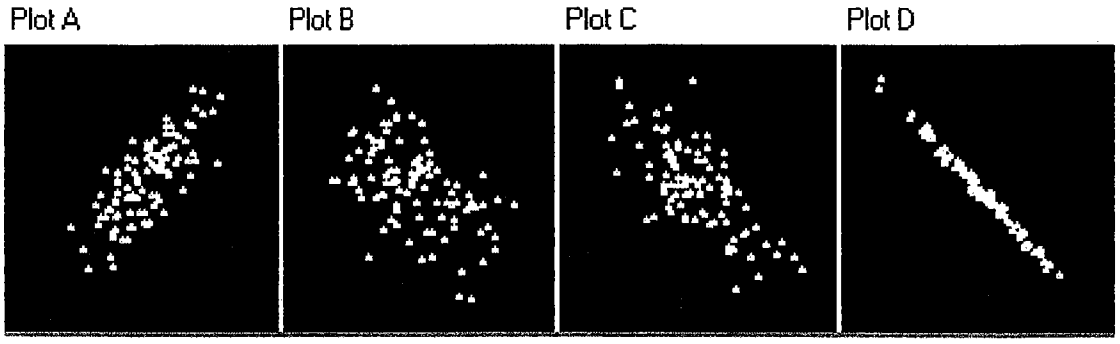
   (c) X = Weight in pounds and Y = Weight in kilograms for the 126 people.



   - Plot A: _____
   - Plot B: _____
   - Plot C: _____

2. **(12 Points)** For the following plots, match each plot with one of the following correlation coefficients:

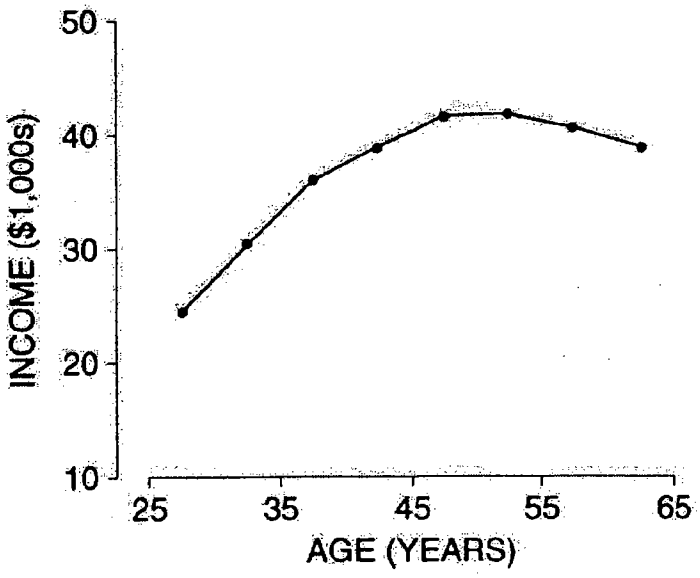
$$-1.03, \quad -0.99, \quad -0.70, \quad -0.50, \quad 0, \quad 0.20, \quad 0.76, \quad 0.99$$



   - Correlation for Plot A: r = _____
   - Correlation for Plot B: r = _____
   - Correlation for Plot C: r = _____
   - Correlation for Plot D: r = _____

Question 3: Representative Sample **(35 Points)**

The graph below summarizes the results from a study based on a representative sample of men age 25-64 in 1993, who were working full time that year; the graph shows average income for each age group.



Source: March 1993 Current Population Survey; CD-ROM supplied by the Bureau of the Census through the U.C. Survey Research Center.

True or **false**, and explain: the data show that on average, if a man keeps working, his income will increase until age 50 or so, then start decreasing. If false, how do you account for the pattern in the data?

Question 4: Normal Distribution **(45 Points)**

According to the U.S. Department of Agriculture, a 3–ounce serving of trimmed sirloin beef contains, on average, 7.4 grams of fat. Assume that the amount of fat for such servings closely follows the normal curve with a standard deviation of 0.4 gram. Find the percentage of servings that contain:

1. **(15 Points)** Between 6.9 and 7.2 grams of fat: _____ % of servings

2. **(15 Points)** More than 8.1 grams of fat: _____ % of servings

3. **(15 Points)** Only one percent of servings will contain more than _____ grams of fat.

Show your work!

Question 5: Regression **(35 Points)**

For women age 25 and over in the U.S. in 1993, the relationship between age and educational level (years of schooling completed) can be summarized as follows:

average age $\approx$ 48.7 years, SD $\approx$ 16.8 years
average ed. level $\approx$ 12.5 years, SD $\approx$ 3.1 years,
r $\approx$ –0.28

1. **(20 Points)** The predicted educational level (number of years of schooling) for a woman who is 45 years old is _____ years. **Show your work!**

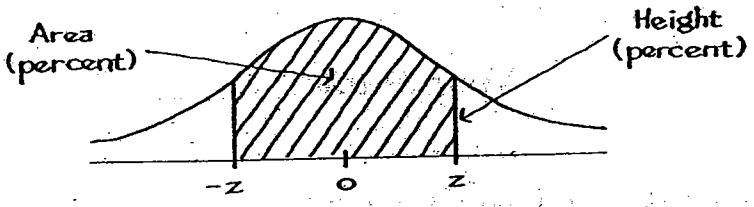
2. **(15 Points) True** or **false**, and explain: as you get older, you become less educated. If this statement is false, what accounts for the negative correlation?

Question 6: Valentine's Day Question **(5 Points)**

How many **free points** do you want to get for taking the Midterm Test just before the Valentine's Day (circle one)?

- 0 points
- 1 point
- 3 points
- 5 points

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 004, Midterm 2 (200 Points)

Friday, March 26, 2004

Your Name: _____

Question 1: Regression (**50 Points**)

From the subjects (all men) in a health survey, the following data were collected:
Average height = 68 inches SD = 2.5 inches
Average blood pressure = 120 mm SD = 15 mm
Correlation = -0.2.

The scatter diagram is football–shaped.
Show your work!

1. (**15 Points**) Find the regression equation for predicting blood pressure from height.

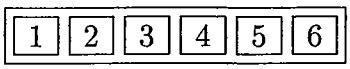
2. (**10 Points**) Using your regression equation, estimate the blood pressure of a man who is 73 inches tall.

3. (**15 Points**) Find the r.m.s. error for predicting the blood pressure from the height.

4. (**10 Points**) The correlation coefficient tells us that, on average, taller men have **higher / lower** blood pressures than shorter men, and that the relationship between blood pressure and height is quite **strong / weak**. (**Circle the correct word in each pair of choices.**)

Question 2: The Expected Value and the Standard Error (**40 Points**)

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



Show your work!

1. (**10 Points**) If the sum of the draws is 321, what is their average?
   It is: _____

2. (**10 Points**) If the average of the draws is 3.78, what is the sum?
   It is: _____

3. (**20 Points**) Using the normal curve, estimate the chance that the average of the draws is between 3 and 4.
   It is: _____

Question 3: Chances and Probabilities (**40 Points**)

I have a bag with 20 balls in it: 10 are red, 8 are blue, and 2 are green.
Show your work!

1. (**10 Points**) If I draw one ball at random from the bag, what is the chance that I get a red ball or a green ball?
   It is: _____%

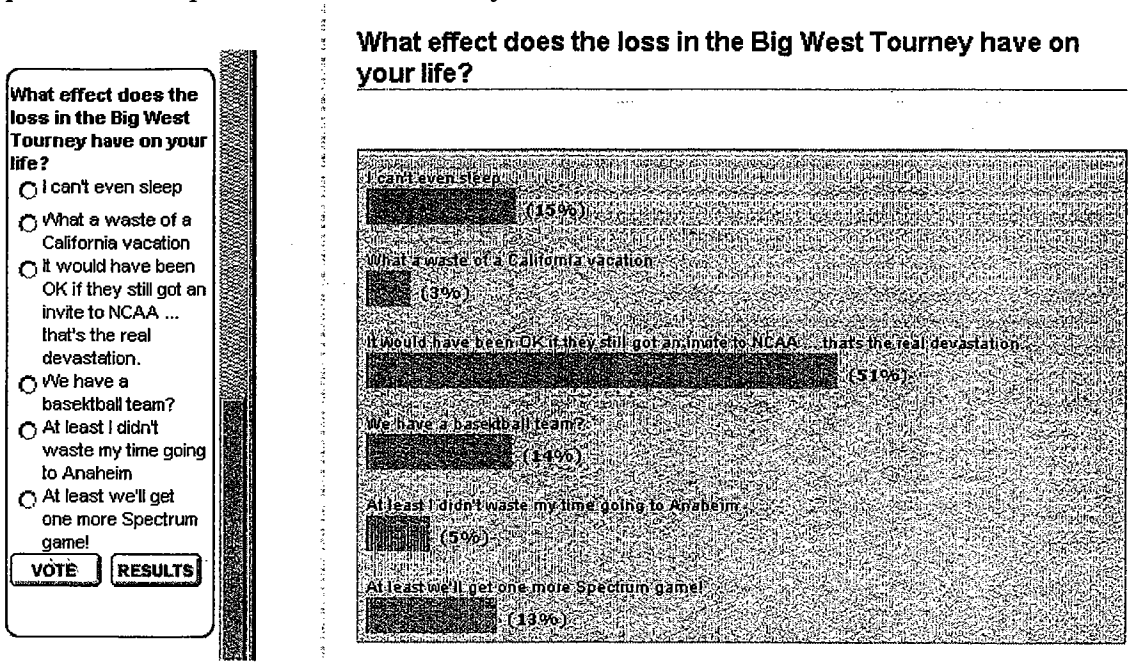
2. (**10 Points**) If I draw two balls at random **without replacement**, what is the chance that I get a red ball, followed by a green ball?
   It is: _____%

3. (**10 Points**) If I draw three balls at random **with replacement**, what is the chance that I get all three red balls?
   It is: _____%

4. (**10 Points**) If I draw three balls at random **with replacement**, what is the chance that I get at least one red ball?
   It is: _____%

Question 4: Sampling (**40 Points**)

What effect does the loss in the Big West Tourney have on your life? This was the question asked in a recent survey posted on the Utah Statesman Online Web page. A snapshot of the question and the survey results are shown below:



1. (**20 Points**) Whereas more than 50% of the students think that not being invited to the NCAA was the real devastation, 14% seem to be surprised to hear that USU has a basketball team. Interesting... But can we really take the results of this survey seriously? **Yes – seriously / no – not seriously.** Circle your answer and explain why or why not!

2. (**20 Points**) Explain how **you** would organize a survey to obtain students' opinion on the effect of the loss in the Big West Tourney on their life.

Question 5: Chance Errors in Sampling (**30 Points**)

Five hundred draws are made at random from the box

$$\boxed{60{,}000 \times \boxed{0} \qquad 20{,}000 \times \boxed{1}}.$$

True or false, and explain:

1. (**5 Points**) True / false: The expected value for the percentage of 1's among the draws is exactly 25%.

2. (**5 Points**) True / false: The expected value for the percentage of 1's among the draws is around 25%, give or take 2% or so.

3. (**5 Points**) True / false: The percentage of 1's among the draws will be around 25%, give or take 2% or so.

4. (**5 Points**) True / false: The percentage of 1's among the draws will be exactly 25%.

5. (**5 Points**) True / false: The percentage of 1's in the box is exactly 25%.

6. (**5 Points**) True / false: The percentage of 1's in the box is around 25%, give or take 2% or so.

Formulas:

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

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

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

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

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

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

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

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

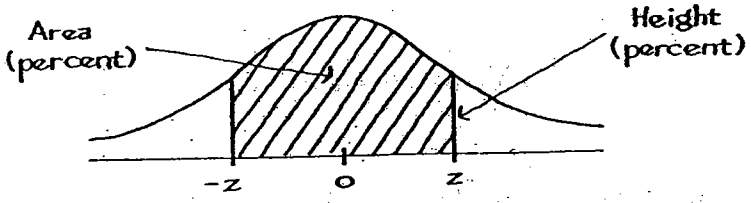
$$\text{SD} = (\text{bigger - smaller}) \times \sqrt{\frac{\text{fraction}}{\text{bigger}} \times \frac{\text{fraction}}{\text{smaller}}}$$

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

$$\text{average} = \frac{\text{number of } \boxed{1} \text{ 's}}{\text{how many tickets in the box}} \qquad \text{SD} = \sqrt{\frac{\text{fraction}}{\text{of } \boxed{1} \text{ 's}} \times \frac{\text{fraction}}{\text{of } \boxed{0} \text{ 's}}}$$

$$\text{EV}_{\%} = \%\text{ of } \boxed{1} \text{ 's in the box} \qquad \text{SE}_{\%} = \frac{\text{SE}_{\text{sum}}}{\#\text{draws}} \times 100\%$$

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