Statistics 1040, Sections 003 & 004, Midterm 1 (200 Points)

February 14, 2003

Your Name: _____

Question 1: Observational Studies/Controlled Experiments (**40 Points**)

As part of a study on exercise and health, a group of 1,000 people was followed for 5 years. At the beginning of the the study, the researchers asked each person whether they exercised regularly or not. At the end of the study, the researchers measured several health–related variables, and in doing so, they noticed that the death rate for the exercise group was lower than for the no–exercise group.

1. **(15 Points)**Is the study described in this article an **observational study** or a **controlled experiment**? Circle your answer and **explain!**

2. **(5 Points)** When we follow people over a longer period of time, as in the study described in this article, we call such a study a _____ study.

3. **(20 Points)** Does the result of this study necessarily imply that if people who do not exercise start to exercise regularly, they will live longer, on average, than if they do not? **Explain!**

Question 2: Normal Distribution (**45 Points**)

A lawyer commutes daily from his suburban home to his midtown office. On the average, one way of his trip (to work) takes 24 minutes with a standard deviation of 4 minutes. Assume that the histogram of his trip times closely follows the normal curve.

Fill the blanks in the statements below and **show all the work needed to obtain the answer**:

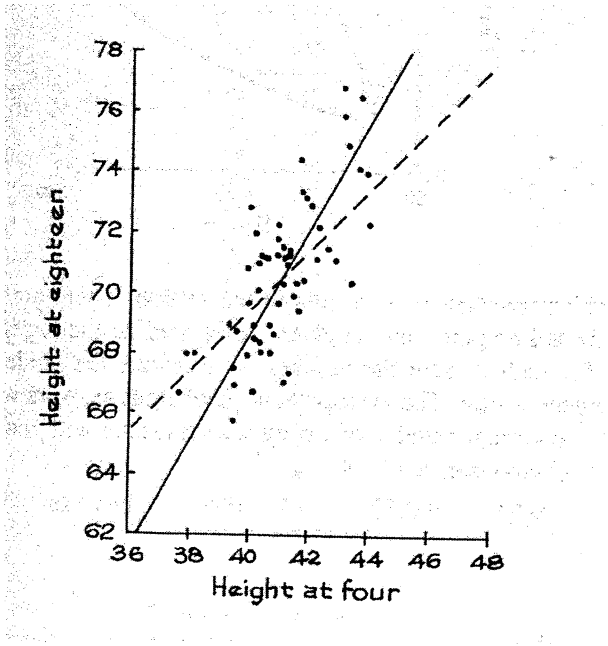
1. (**15 Points**) Assume the office opens at 9:00am and he leaves his house at 8:45am. The percentage of times he will be late for work (i.e., he needs more than 15 minutes to get to work) is _____%. This means 15 minutes is the _____th percentile of his trip times.

2. (**15 Points**) Not satisfied with the result above, he decides to leave his house at 8:30am already. Coffee is served at his office from 8:50am to 9:00am. The proportion of times he arrives at the office during the time coffee is served (i.e., he needs between 20 to 30 minutes to get to work) is _____%.

3. (**15 Points**) In 15% of his trips, he needs more than _____ minutes to get to work (i.e., the slowest 15% of his trip times are longer than this).

Question 3: Scatter Diagram & Correlation (**20 Points**)

A longitudinal study of human growth has been under way since 1929, at the Berkley Institute of Human Development. The scatter diagram below shows the heights of 64 boys, measured at ages 4 and 18.



1. (**5 Points**) The average height at age 4 is around

   38 inches 42 inches 44 inches 66 inches 68 inches 71 inches

2. (**5 Points**) The average height at age 18 is around

   38 inches 42 inches 44 inches 66 inches 68 inches 71 inches

3. (**5 Points**) The correlation coefficient is around

   0.50 0.80 0.95

4. (**5 Points**) Which is the SD line – solid or dashed?

Explain your answers!

Question 4: Regression (**45 Points**)

A statistical analysis was made of the midterm and final scores in a large course with following results:

average midterm score ≈ 50	SD ≈ 25
average final score ≈ 55	SD ≈ 15
correlation r ≈ 0.60	

1. (**15 Points**) Predict the final score for a student whose midterm score was 80.

2. (**15 Points**) Predict the final score for a student whose midterm score was 15.

3. (**15 Points**) Apparently, students who had very good midterm scores did not study enough later on and, consequently, did not do so well on the final. On the other hand, students with poor midterm scores worked hard during the rest of the semester and considerably improved their performance on the final. Do you agree with this statement? Why or why not? Explain!

Question 5: Change of Scale (**40 Points**)

From the subjects 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 r = -0.2	

You want to provide a summary of these results to a friend in Europe and report heights in centimers (instead of inches) to make it easier for your friend to interpret the results. Recall that 1 inch = 2.54 centimeters.

1. (**10 Points**) The average height (in centimeters) is: _____

2. (**10 Points**) The SD [for the height] (in centimeters) is: _____

3. (**10 Points**) The correlation r is: _____

4. (**10 Points**) Circle the correct answer: r is measured in

   - centimeters
   - inches
   - mm
   - inches · mm
   - centimeters · mm
   - inches · centimeters
   - none of these – r is a unitless number

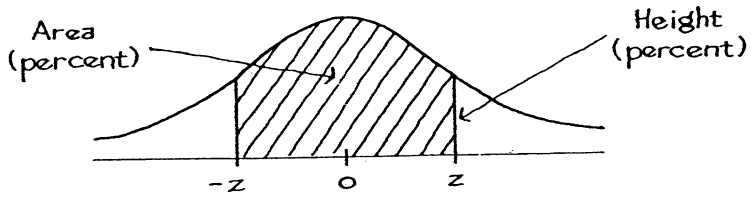
Question 6: Valentine's Day Question (**10 Points**)

Mark the correct box:

(**10 Points**):
☐ **YES**, I would like to get 10 points for free for taking the Valentine's Day Midterm.

(**0 Points**):
☐ **No, thank you, I don't need any additional 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

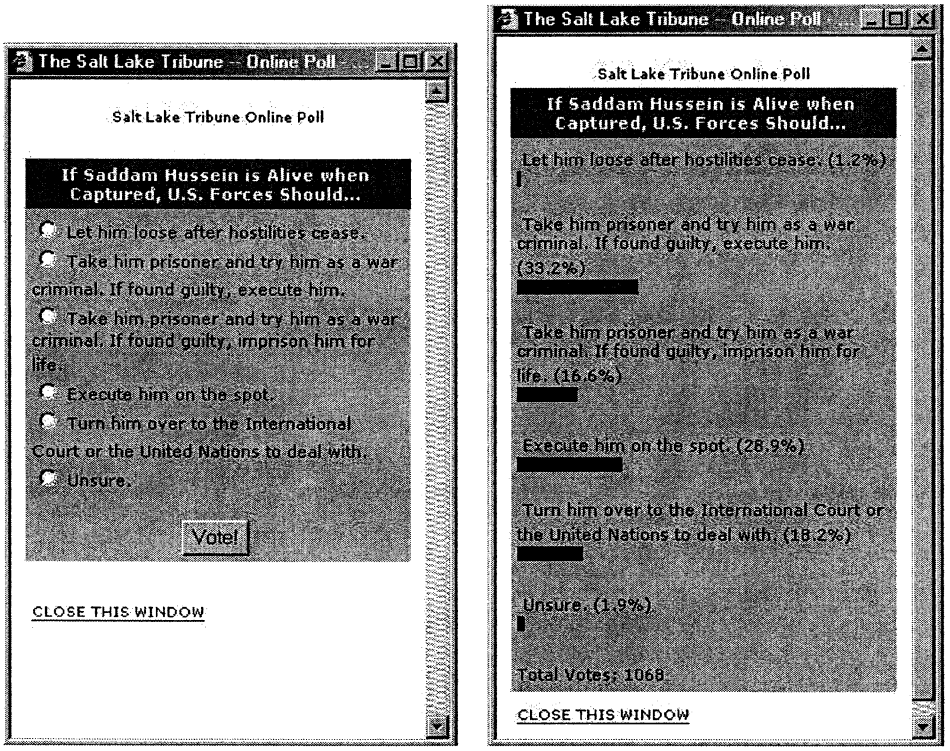
Statistics 1040, Sections 003 & 004, Midterm 2 (200 Points)

March 28, 2003

Your Name: _____

Question 1: Sampling (**30 Points**)

The pictures below show the *The Salt Lake Tribune* Online Poll of 03/25/2003:



(**30 Points**) So, based on this poll, it seems that about 1/3 of the nation would take S.H. prisoner and try him as a war criminal, and a slightly smaller proportion of the nation (28.9%) would execute him on the spot. Do you think the results of this poll objectively represent public opinion in the US? Answer **yes** or **no** and give **three arguments** to justify your answer, based on the facts given, but *NOT* on your personal opinion.

Question 2: The Expected Value and Standard Error (**60 Points**)

Three hundred draws will be made at random with replacement from the box $\boxed{1}\boxed{3}\boxed{5}\boxed{7}$.

Show your work!

1. (**25 Points**) Estimate the chance that the sum of draws will be more than 1,100.

2. (**35 Points**) Estimate the chance that there will be fewer than 80 $\boxed{5}$'s.

Question 3: Regression (**50 Points**)

For a random sample of 25 car models, the average weight (in pounds) was 3240, with an SD of 530. The average gas mileage (in miles per gallon) was 22.3 with an SD of 4.3. The correlation between weight and gas mileage was -0.79. The scatter diagram was football shaped. **Show your work!**

1. (**15 Points**) Find the regression equation for predicting gas mileage from weight.

2. (**10 Points**) Using your regression equation, predict the gas mileage of a car that weights 3200 pounds.

3. (**10 Points**) Find the r.m.s. error for predicting gas mileage from weight.

4. (**15 Points**) Would you be surprised if someone told you that one of these cars weighting 3600 pounds got 24 miles per gallon? Why or why not? Explain your reasoning, using the r.m.s. error.

Question 4: Probability and Chance (**40 Points**)

A bookshelf contains 8 novels, 7 books of poems, 1 dictionary, and 2 copies of Freedman, Pisani, and Purves's "Statistics" textbook. I pick two books at random without replacement. **Answer each of the following questions separately. Show your work!**

1. (**5 Points**) What is the chance that the first book is a novel or a dictionary?

2. (**13 Points**) If I want to study statistics, what is the chance that I pick at least one copy of the Stats textbook?

3. (**10 Points**) What is the chance that the first two books are both novels?

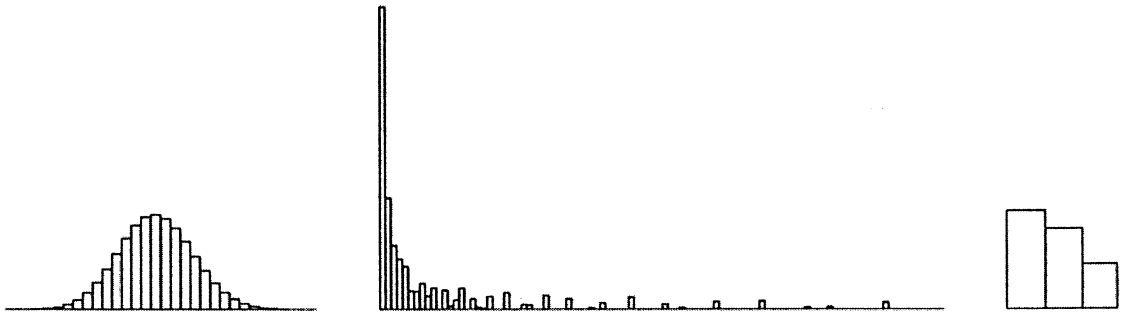
4. (**12 Points**) What is the chance that I pick one book of poems and a dictionary (in any order)?

Question 5: Normal Approximation for Probability Histograms (**20 Points**)

Twenty-five draws are made at random with replacement from the box $\boxed{1}\boxed{1}\boxed{2}\boxed{2}\boxed{3}$.

(**20 Points**) One of the graphs below is an (empirical) histogram for the numbers drawn. One is the probability histogram for the sum. And one is the probability histogram for the product. Which is which? **Explain!**

(i) (ii) (iii)



- An (empirical) histogram for the numbers drawn is _____.
  Explanation:

- The probability histogram for the sum is _____.
  Explanation:

- The probability histogram for the product is _____.
  Explanation:

Formulas:

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

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

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

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

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

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

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

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:

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

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

Formulas:

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

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

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

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

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

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

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

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

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

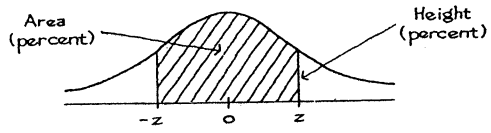
$$\text{SD} = (\text{bigger - smaller}) \times \sqrt{\frac{\text{fraction}}{\text{bigger}} \times \frac{\text{fraction}}{\text{smaller}}}$$

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

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

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

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