Name_____

Stat 1040, Spring 2004
Final Test, Friday April 23rd, 7:30-9:20 am

Show your work. The test is worth 100 points and you have 110 minutes.

1. (9 pts) In an experiment, 216 stamped, addressed letters are lost, and the rate of return is recorded. Some of the letters were addressed to Mr. M. J. Davis; some to Dandee Davis, c/o Hooters Club; some to M.J. Davis, c/o Friends of the Communist Party. Test to see if there is a relationship between the addressee and the rate of return. Clearly state the null and alternative hypotheses, calculate the appropriate test statistic, find the p-value, and state your conclusion.

	Davis	Hooters	Communist	Total
Returned	32	17	29	78
Not Returned	40	55	43	138
Totals	72	72	72	216

2.(6 pts) In a random sample of 200 homes in Jefferson, Illinois (population 25,000), it was found that the average number of televisions sets per home was 3.2 with a standard deviation of 0.9.

a. Construct a 95% confidence interval for the average number of television set per home in Jefferson, Illinois.

b. T F 95% of the homes in Jefferson have between (your lower limit) and (your upper limit) televisions. (2 pts)

c. T F Your interval in part (a) represents a 95% confidence interval for the average number of television set in the sample. (2 pts)

d. (3 pts) Suppose that you find out that televisions sets in Jefferson, Illinois do not follow the normal curve, but have a long right tail. Can you still rely on your results in part (a)? Explain

3. Roger has tossed a fair coin six times and got heads every time. His friend tells him that his chance of getting a tail on the next toss is almost certain. Do you agree? (3 pts)

4. A bookstore has a shelf with 15 math books, 10 stat books, and 25 biology books. (2 pts each). Suppose that three books are drawn at random **(without replacement)**:

a. What is the probability that all three books are stat books?

b. What is the probability the one of each type of book is selected?

c. What is the probability that none of the books selected are stat books?

d. What is the probability that at least one of the books selected is a stat book?

e. Given that the first book was a biology book, what is the probability that the second book is a stat book?

f. What is the probability that the first book is math book or that the first book is a stat book?

5. (6 pts) A Logan radio station wishes to determine the reaction of students to President Hall's application to the University of Tennessee. They set up a booth on the quad in front of Widtsoe Hall and pass out surveys from 10 to 12 Thursday morning. Of 700 surveys that were distributed, 206 are returned and of these 169 are glad President Hall is returning. The radio station reports that 82% of USU students are in glad President Hall is returning. Do you agree with the radio station's conclusion? Clearly justify your conclusion.

6. A proponent of Dr. Atkins claims that this diet is better than the South Beach diet. In a simple random sample of 400 Atkins diet participants, 300 lost weight and in a simple random sample of 500 South Beach dieters, 360 lost weight. Can we conclude from these data that the Atkins diet is better than the South Beach diet?

a. (10 pts) Clearly state the null and alternative hypotheses, calculate the appropriate test statistic, find the p-value, and state your conclusion.

b. (3 pts) Is this an observational study or controlled experiment? Justify your choice.

7. (4 pts) A marketing firm wants to determine whether Barbie dolls are still a good market item. To estimate the percentage of people who want to buy one, the firm takes a simple random sample of 500 people in Utah and a simple random sample of 500 people in California. Other things being equal,

a. The accuracy in Utah will be about the same as the accuracy in California.
b. The accuracy in Utah will be less than the accuracy in California.
c. The accuracy in Utah will be more than the accuracy in California.
d. Cannot be determined from the information given.

8. In Pearson's data about fathers and sons, fathers had an average height of 68 inches with a standard deviation of 2.7 inches. (4 points each)

a. Suppose one father is 64 inches tall, express his height as a percentile.

b. Another father is in the 78^{th} percentile, express his height in inches.

9. For men 18 – 24 in the HANES sample the following is recorded:
average height = 70 inches SD = 2.7 inches
average weight = 162 pounds SD = 30 pounds r = 0.6

a. (6 pts) Calculate the regression equation for predicting weight from height.

b. (3 pts) Predict the weight of a man who is 66 inches tall.

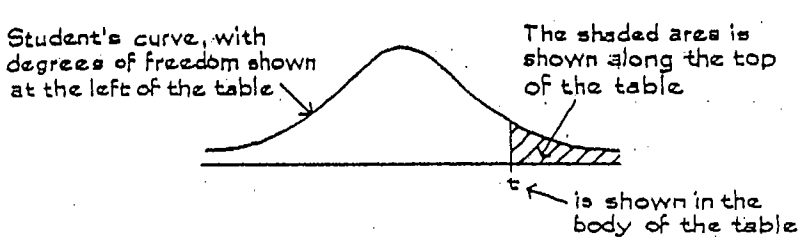
c. (3 pts) About how far off do you expect your prediction to be?

d. (4 pts) Men who were 73 inches tall averaged about 182 pounds. True or false and explain, the ones who weighed 182 pounds averaged about 73 inches tall.

10. (8 pts) Suppose that the ages at which children first walk alone are normally distributed with mean 11.5 months. A simple of sample of 12 children is selected from children having low birth weights. The average age these children first walked alone is 12.2 months with a standard deviation of 1.1. Can we conclude from these data that children with low birth weights walk later than the average? Clearly state the null and alternative hypotheses, calculate the appropriate test statistic, find the p-value, and state your conclusion.

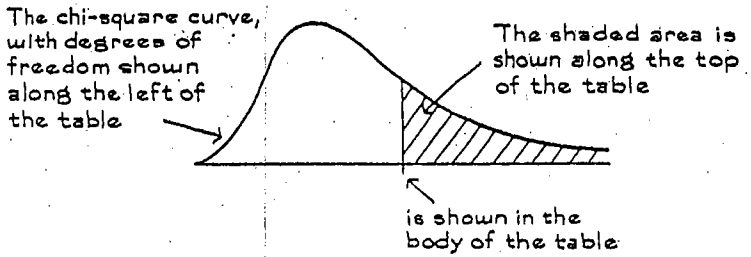
11. (8 pts) Many companies are experimenting with "flex-time," which is supposed to reduce absenteeism. One company employees have averaged 6.3 days off work in the past. The company introduces "flex-time" and a year later a simple random sample of 100 employees is selected. They average 5.5 days off 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.

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.

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

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

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

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

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

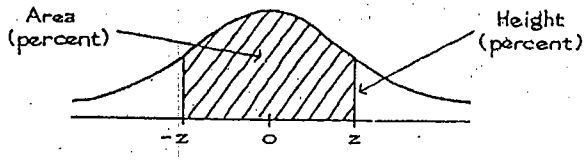
$$EV_{\text{sum}} = \text{number of draws} \times \text{ave}_{\text{box}}$$
$$SE_{\text{sum}} = \sqrt{\text{number of draws}} \times SD_{\text{box}}$$

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

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

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