Statistics 1040 Final Exam, Fall 2005

Name or A-Number:

Directions: You have 110 minutes to complete the exam. The exam will be graded out of 100 points. Be sure to answer every question. You must show your work for full credit.

---

1. Read the following news article:

> ## "Wrap up" advice to stop colds
> Scientists say cold noses reduce ability to fight virus attacks
> Monday, November 14, 2005; Posted: 4:47 p.m. EST (21:47 GMT)
> **LONDON, England (CNN) – British researchers into the common cold say "catching a chill" really does help colds develop – and are advising to "wrap up warm" to keep viruses at bay.**
> Mothers and grandmothers have long warned that chilling the surface of the body, through wet clothes, feet and hair, causes common cold symptoms to develop. But much previous research has dismissed any link between chilling and viral infection as having no scientific basis.
> Now researchers in Cardiff, Wales, say they can prove drops in temperature to the body really can cause a cold to develop.
> Claire Johnson and Professor Ron Eccles, from Cardiff University's Common Cold Center, recruited 180 volunteers, half of whom they got to immerse their feet in ice and cold water for 20 minutes.
> The other 90 in tests during the common cold "season" sat with their feet in an empty bowl.
> During the next four or five days, almost a third (29 percent) of the chilled volunteers developed cold symptoms – compared to just 9 percent in the control group, the scientists said.

(a) (2 points) Is this an observational study or a designed experiment? Explain briefly.

(b) (2 points) What is the "treatment" in this study?

(c) (3 points) The article does not say how the researchers divided up the 180 people into two groups of 90. How *should* they do this? Explain clearly.

(d) (2 points) Is the study blind? Explain briefly.

(e) (3 points) Are there any possible confounding factors or problems with the study? Explain clearly.

(f) (10 points) Assume that the researchers followed your instructions in part (c) and ignore any other problems of which you might be aware. Using the numbers provided in the article, perform a statistical hypothesis test that the researchers could have used to test their belief that drops in temperature can cause a cold to develop.

You must clearly state the null and the alternative hypotheses, compute a test statistic and a P-value and clearly state your conclusions in terms of the language used in the newspaper article.

2. When the Tribbles invaded the spaceship Enterprise, suppose that crew member Spock decided to take the logical step of seeing what the crew was up against, and he wanted to graphically represent the sizes of the Tribbles. Suppose that the table below summarizes the heights of the 50 Tribbles he found on the bridge. (Class intervals include the left but not the right endpoints.)

Tribble Height (inches)	Number of Tribbles
3-5	19
5-7	10
7-9	12
9-11	7
11-13	2
	50

a. (9 points) Draw a histogram of these height data, with the vertical axis on the usual density scale, both axes labeled, and heights of bars clearly indicated.

b. (2 points) If a Tribble is in the 82^{nd} percentile for height, about how tall is it? (Note: Use the histogram, NOT the normal curve).

3. For college-aged men, the average height is 70 inches with a standard deviation of 3 inches, and the average weight is 162 pounds with a standard deviation of 30 pounds. The correlation between height and weight for college-aged men is 0.47.

a. (3 points) Find the equation of the line to predict weight from height for college-aged men.

b. (2 points) Predict the weight for a college-aged man who is 66 inches tall.

c. (2 points) If the man in part b is exactly on the SD line, will his true weight be larger or smaller than your answer to part b? Explain briefly. (Note that it is possible to answer this question even if you were not able to answer part b.)

4. In one Stat 1040 class, there are 48 students, of whom 13 are male and 35 are female. Of the male students, 6 are from the College of Education, and of the female students, 16 are from the College of Education. Two students are selected at random (like drawing without replacement) from the class.

a. (2 points) What is the chance that both students are women?

b. (2 points) What is the chance that neither of the students is male?

c. (2 points) What is the chance that the first student is male and the second student is female?

d. (2 points) What is the chance that at least one of the students is from the College of Education?

5. (4 points) A gambler loses ten times running at a game of chance. The gambler thinks he should keep playing because he is due for a win by the law of averages. A bystander advises him to quit, on the grounds that his luck is cold. What does statistics say?

6. Suppose researchers selected a simple random sample of 1,200 U.S. taxpayers and found that in 2004 these 1,200 received an average tax refund of $2,052, with a standard deviation of $431. According to the IRS, the average refund for all U.S. taxpayers that year was $2,063.

a. (2 points) What does it mean when we say that the researchers selected a "simple random sample" of 1,200 U.S. taxpayers?

b. (7 points) Using the researchers' results, find a 90 percent confidence interval for the average refund for all U.S. taxpayers in 2004.

c. (3 points) True or False, and explain briefly: The sizes of the tax returns of 90% of all U.S. taxpayers in 2004 are in this 90 percent confidence interval.

7. (6 points) Suppose that 13 percent of all people are left-handed. If a simple random sample of seven hundred people is considered, what are the chances that between 9.3 percent and 11.2 percent of them are left-handed?

8. (10 points) On the M&M web page, it claims that they produce 13% brown, 14% yellow, 13% red, 24% blue, 20% orange, and 16% green milk chocolate M&M's. Suppose we buy a bag of milk chocolate M&M's and come up with the following numbers of each color:

Color	Number
brown	50
yellow	47
red	41
blue	94
orange	102
green	47
	381

Test the hypothesis that our bag of M&M's is like a simple random sample of M&M's from a population with the specified percentages of each color. You must state a null and an alternative hypothesis, find a test statistic and a P-value and clearly state your conclusions in terms of what you have learned about the color of milk chocolate M&M's.

9. A Stat 1040 instructor recently tried to answer the following question:

> "Does it help to improve Stat 1040 students' quiz scores when the answers to a quiz are handed out the lecture *before* the actual quiz?"

On Monday, students were given a handout with some worked examples. The students were strongly advised to look through this handout to prepare for the Wednesday quiz. The quiz consisted of one of the problems from the handout. The 38 participating students averaged 13.5 points (out of 20) with an SD of 5.8 points. In the past, students of this instructor averaged 11.4 points (out of 20) on a similar quiz. Do his Stat 1040 students score higher if they have access to the solutions before the quiz, or not? For the purpose of this question, you should assume that the 38 students who took this particular quiz are like a simple random sample of all Stat 1040 students taught by this instructor.

(a) (2 points) Clearly state the null and the alternative hypotheses.

(b) (5 points) Compute a test statistic.

(c) (1 point) Find the P-value.

(d) (1 point) Do you reject the null hypothesis? Explain why or why not.

(e) (1 point) Clearly state your conclusions.

10. In a large city, there are 5 electoral precincts. There are two mayoral candidates, A and B. A political science student takes a simple random sample of 1870 voters from this city and asks them which precinct they live in and whether they voted for candidate A or B. She makes the following table:

	Candidate		
Precinct	A	B	Total
1	400	229	629
2	184	154	338
3	101	200	301
4	182	186	368
5	136	98	234
Total	1003	867	1870

We want to test the hypothesis that voting for mayoral candidates A and B is independent of precinct for voters from this city.

(a) (2 points) Clearly state the null and the alternative hypotheses.

(b) (4 points) Compute a test statistic.

(c) (1 point) Find the degrees of freedom.

(d) (1 point) What can you say about the P-value?

(e) (1 point) Do you reject the null hypothesis? Explain why or why not.

(f) (1 point) Clearly state your conclusions.

Memory Aids

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

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

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

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

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

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

$EV_{\text{sum}} =$ number of draws $\times$ ave_{box}

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

$EV_{\text{ave}} = \text{ave}_{\text{box}}$

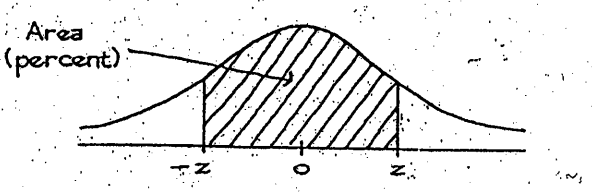
$SE_{\text{ave}} = \dfrac{SE_{\text{sum}}}{\text{number of draws}}$

$EV_{\%} =$ % of 1's in the box

$SE_{\%} = \left(\dfrac{SE_{\text{sum}}}{\text{number of draws}}\right) \times 100\%$

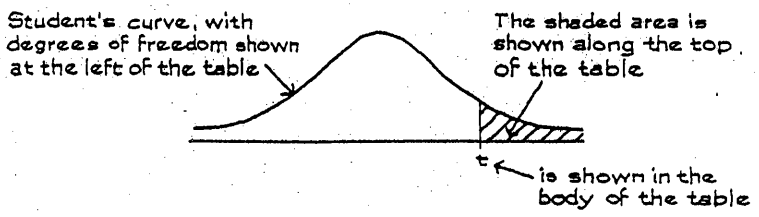
$SE_{\text{diff}} = \sqrt{a^2 + b^2}$ where a is the SE for the first quantity, b is the SE for the second quantity, and the two quantities are independent

A NORMAL TABLE



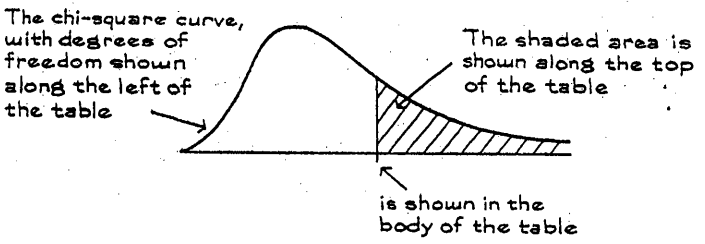
z	Area	z	Area	z	Area
0.00	0	1.50	86.64	3.00	99.730
0.05	3.99	1.55	87.89	3.05	99.771
0.10	7.97	1.60	89.04	3.10	99.806
0.15	11.92	1.65	90.11	3.15	99.837
0.20	15.85	1.70	91.09	3.20	99.863
0.25	19.74	1.75	91.99	3.25	99.885
0.30	23.58	1.80	92.81	3.30	99.903
0.35	27.37	1.85	93.57	3.35	99.919
0.40	31.08	1.90	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).