Statistics 2 Final Exam
Fall 2002

Printed Name ___________________
(Please also print your name at the top of each page)

Signature _____________________

Student ID # _________________

Circle your TA: Ellen Chien Chris Furgiuele Megumi Harada

Pawel Lasieki John Stambaugh

Circle the time of your section: 9:00 10:00 11:00 12:00 1:00

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Total

There are 19 problems. A Normal table and a Chi-square table are at the end of the exam. You can leave your answers in unreduced forms, for example as fractions or factorials

---

1. Three cards are dealt off the top of a well-shuffled deck. The deck has 52 cards, of which there are 13 of each suit—spades, clubs, diamonds, and hearts. Find the chance that:

    a. [2] You get three hearts.

    b. [2] You get at least one heart.

    c. [2] You get no diamonds.

    d. [2] You get a heart or a diamond

2. [2] A certain species of fish weighs an average of 1 pound. In the population of these fish, the SD is 0.2 pounds. If a fisherman catches 100 of these fish and their weights are independent, the chance that their total weight is more than 103 pounds is about equal to ______.

3. [2] Fisherman A computes the correlation between the length and weight of a collection of fish. Using the same collection, Fisherman B first sorts the fish into 12 distinct species and finds the average length and weight for each species. He then computes the correlation of the resulting 12 pairs of numbers. Circle one:

    a. The correlation coefficient found by Fisherman A will be larger than that found by Fisherman B.
    b. The correlation coefficient found by Fisherman B will be larger than that found by Fisherman A.
    c. The two correlation coefficients will be about the same.

4. [2] In a survey of rents in San Francisco, the median monthly rent was $1200 and the average monthly rent was $1800. Circle one:

    a. The difference between the 90^{th} and 50^{th} percentiles is greater than the difference between the 50^{th} and 10^{th} percentiles.
    b. The difference between the 90^{th} and 50^{th} percentiles is less than the difference between the 50^{th} and 10^{th} percentiles.
    c. The difference between the 90^{th} and 50^{th} percentiles is about equal to the difference between the 50^{th} and 10^{th} percentiles.

5. [2] A certain genetic trait has a chance of occurring equal to 25%. Among 10 individuals with independent genetic make-ups, what is the chance that exactly 9 have this genetic trait?

6. Data on the heights of fathers and sons can be summarized as follows:

    average height of fathers = 68 inches SD = 2.7 inches
    average height of sons = 69 inches SD = 2.7 inches
    correlation = 0.8

    a. [2] On average the sons of 70 inch fathers are about ____ inches taller than sons of 66 inch fathers.

    b. [2] If a son's height is guessed from his father's height, there is about a 68% chance that the guess is right within ____ points.

7. [2] The regression equation for predicting the Math SAT from the Verbal SAT is

```
MathSAT = 0.6 x VerbalSAT + 245
```

The RMS error of the regression line is 80 points.

   a. [2] If a student scores 500 on the Verbal SAT, his predicted Math SAT would equal ______.

   b. [2] Of 10,000 students who scored 500 on the Verbal SAT, the number who scored more than 500 on the Math SAT was about equal to ______.

8. The National Institute for Standards and Technology has a small weight called NB10 that is supposed to weigh 10 grams. The average of 100 weighings of NB10 on an unbiased scale was 406 micrograms below 10 grams, and the SD of the weighings was 6 micrograms.

   a. [2] The SE of the average equals _______.

   b. [2] A 95% confidence interval for the number of micrograms below 10 grams is _____ to ______.

9. [2] In a simple random sample of 625 voters in Berkeley, 60% said they planned to vote in favor of a particular state ballot initiative. An approximate 95% confidence interval for the percentage of voters in Berkeley who planned to vote for the initiative is _______ to _______.

10. [2] If a simple random sample of the same size as in the previous problem is taken in Los Angeles, which has a much larger population, the results will be (circle one):

    a. Substantially more accurate than those for Berkeley.
    b. Substantially less accurate.
    c. About the same accuracy.
    d. Can't tell without further information

11. [2] If a newly minted penny is spun on its edge, it does not have exactly a 50-50 chance of coming to rest with heads or tails up. In order to estimate the chance that it does land head up with a standard error of 5%, the coin would have to be spun ______ times.

12. A company developed a new treatment for dandruff and performed a double-blind randomized experiment to see if it worked better than the old treatment. The new treatment performed better in the experiment. The p-value was 18%. Label the following statements true or false by circling T or F. (Each question is worth 1 point).
    a. [T F] The new drug is significantly better than the old.
    b. [T F] We can reject the null hypothesis.
    c. [T F] It is plausible that the superior performance of the new drug was due to chance.
    d. [T F] There is an 18% chance that the null hypothesis is true.
    e. [T F] The alternative hypothesis is 18 times more likely the null.

13. A simple random sample of 225 people over the age of 18 was taken in a certain city, and they were asked how many miles they had traveled on vacation the previous year. A 95% confidence interval of 488 to 592 miles was calculated from these results. Label each of the following statements True or False. (Each question is worth 1 point)

    a. [T F] The average of the 225 distances is about 540 miles
    b. [T F] The SD of the distances is about 390 miles
    c. [T F] The histogram of the 225 distances is close to the normal curve
    d. [T F] The probability histogram for the sample average is close to the normal curve.
    e. [T F] About 95% of the vacationers in that city traveled between 488 and 592 miles the previous year.

14. In a simple random sample of 1500 people, Gallup poll were asked to rate the ethical standards of several occupations. Only 18% gave lawyers a rating of "very high or high," whereas 20% gave building contractors that rating.

    a. [1] True or False: a two sample z test can be used to test whether this difference is statistically significant.

    b. [1] True or False: a two sample z test could be used to compare the percentages of men and women who rated building contractors "very high or high."

15. [2] IQ tests were given to a group of men who had newly enlisted and were in training for the Army, and various other information was gathered as well. True or False: a two sample z test could be validly used to compare the average IQ scores of those whose parents had been divorced to those whose parents had not been divorced.

16. [2] In a simple random sample of families with at least two children, the genders of the first and second borns were compared. True or False: a two sample z test could be validly used to test the significance of the difference between the percentage of first borns who were male and the percentage of second borns who were male.

17. [2] The nurses in a certain hospital were randomly divided into two groups. One group was given vitamin C daily and the other group was given a daily placebo. After six months the number of days of sick leave taken by the two groups were compared. The experiment was doubly blinded. True or False: the two sample z test could be validly used to compare the averages of the two groups.

18. 50 rats were randomly divided into two groups of 25 each. The rats in one group were given steroids. They were then timed running a maze. The rats in the non-steroid group had an average time of 10 seconds with an SD equal to 2 seconds and the rats in the steroid group had an average time of 9 seconds with an SD equal to 3 seconds.

    a. [2] What is the SE of the difference of the two averages?

    b. [2] What is the value of the test statistic for testing the null hypothesis that steroids made no difference?

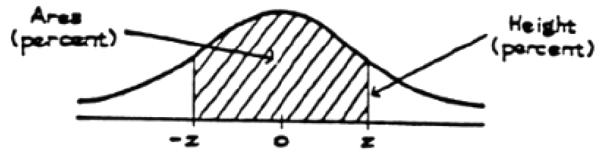
    c. [2] What is the P-value of the test?

19. A machine has a record of producing 80% excellent, 15% good, and 5% unacceptable items. After extensive repairs, in a sample of 200 items, 150 were excellent, 40 were good, and 10 were unacceptable. Test the hypothesis that after the repairs, the machine is performing in accordance with the historical record.

    a. [2] The value of the test statistic for testing this null hypothesis is ______.

    b. [2] The P-value of the test is _______.

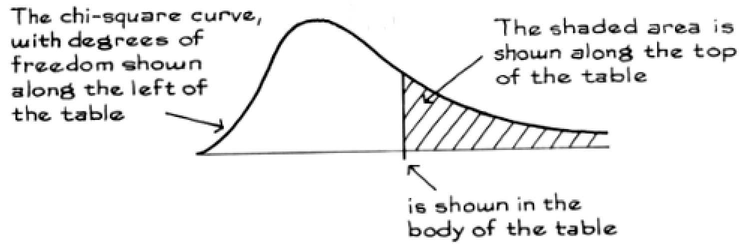
Table



A NORMAL TABLE

z	Height	Area	z	Height	Area	z	Height	Area
0.00	39.89	0	1.50	12.95	86.64	3.00	0.443	99.730
0.05	39.84	3.99	1.55	12.00	87.89	3.05	0.381	99.771
0.10	39.69	7.97	1.60	11.09	89.04	3.10	0.327	99.806
0.15	39.45	11.92	1.65	10.23	90.11	3.15	0.279	99.837
0.20	39.10	15.85	1.70	9.40	91.09	3.20	0.238	99.863
0.25	38.67	19.74	1.75	8.63	91.99	3.25	0.203	99.885
0.30	38.14	23.58	1.80	7.90	92.81	3.30	0.172	99.903
0.35	37.52	27.37	1.85	7.21	93.57	3.35	0.146	99.919
0.40	36.83	31.08	1.90	6.56	94.26	3.40	0.123	99.933
0.45	36.05	34.73	1.95	5.96	94.88	3.45	0.104	99.944
0.50	35.21	38.29	2.00	5.40	95.45	3.50	0.087	99.953
0.55	34.29	41.77	2.05	4.88	95.96	3.55	0.073	99.961
0.60	33.32	45.15	2.10	4.40	96.43	3.60	0.061	99.968
0.65	32.30	48.43	2.15	3.96	96.84	3.65	0.051	99.974
0.70	31.23	51.61	2.20	3.55	97.22	3.70	0.042	99.978
0.75	30.11	54.67	2.25	3.17	97.56	3.75	0.035	99.982
0.80	28.97	57.63	2.30	2.83	97.86	3.80	0.029	99.986
0.85	27.80	60.47	2.35	2.52	98.12	3.85	0.024	99.988
0.90	26.61	63.19	2.40	2.24	98.36	3.90	0.020	99.990
0.95	25.41	65.79	2.45	1.98	98.57	3.95	0.016	99.992
1.00	24.20	68.27	2.50	1.75	98.76	4.00	0.013	99.9937
1.05	22.99	70.63	2.55	1.54	98.92	4.05	0.011	99.9949
1.10	21.79	72.87	2.60	1.36	99.07	4.10	0.009	99.9959
1.15	20.59	74.99	2.65	1.19	99.20	4.15	0.007	99.9967
1.20	19.42	76.99	2.70	1.04	99.31	4.20	0.006	99.9973
1.25	18.26	78.87	2.75	0.91	99.40	4.25	0.005	99.9979
1.30	17.14	80.64	2.80	0.79	99.49	4.30	0.004	99.9983
1.35	16.04	82.30	2.85	0.69	99.56	4.35	0.003	99.9986
1.40	14.97	83.85	2.90	0.60	99.63	4.40	0.002	99.9989
1.45	13.94	85.29	2.95	0.51	99.68	4.45	0.002	99.9991

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).