Final[1] for Statistics 301
Elementary Statistical Methods - Spring 1998
Material Covered: entire course
4th May

This is a 2 hour final, worth 27% and marked out of 27 points. The total possible points awarded for each question is given in square brackets at the beginning of each question. Anything that can fit on two sides of an $8\frac{1}{2}$ by 11 inch piece of paper may be used as a reference during this quiz. A calculator and appropriate statistical tables may also be used. No other aids are permitted.

Name (please print): ______________________________ . **ID Number:**
.

last   first

---

1. [2] At Purdue University North Central, 120 students are randomly selected and asked the distance of their commute to campus. From this group, an average of 9.8 miles is computed. Match the items in Column II with the statistical terms in Column I. (*All* of the items in Column I will be used up in the matching procedure; however, two items in Column II will be left unmatched.)

Column I	Column II
(a) data point	**(a)** Process used to select 120 students and determine their distance
(b) experiment	**(b)** 9.8 miles
(c) parameter	**(c)** All students at PU/NC
(d) population	**(d)** Commute distance for one student
(e) sample	**(e)** Average commute distance for all students
(f) statistic	**(f)** The 120 students
	(g) The 120 commute distances
	(h) 8 mile commute distance for a particular student

---

2. [2] From a sample of 80 scores on a chemistry exam, the following results were obtained:

mean = 68	midrange = 72	third quartile = 84
median = 70	standard deviation = 11	range = 52
mode = 74	first quartile = 58	

[1] All questions are based on questions taken from Zarcone, A.J. and Johnson, R. "Test Bank For Johnson's Elementary Statistics", Duxbury Press, 1996.

(a) The score earned by more students than any other: ____________________.

(b) The number of students who scored between 58 and 84 on the exam: ________.

(c) The highest score earned on the exam: ____________________________.

(d) The number that scored within three standard deviations of the mean: ______.

3. Consider the following three questions on the correlation coefficient, r. (Each question, (a), (b) and (c), is *separate* from the rest; in other words, they are *not* related to one another.)

(a) [1] Select the most likely value for the linear correlation coefficient, r, for the two variables from among those given.

- x = age (in years) of an individual
- y = number of times a year the individual visits a doctor

(i) 0.98 (ii) 0.82 (iii) 0.07 (iv) -0.65 (v) -1.28

(b) [1] Suppose we used the equation $y = -3x + 4$ to generate eight ordered pairs (x, y) and then used these ordered pairs to determine r. The value of r will be,

(i) -3 (ii) -1 (iii) 0 (iv) 1 (v) 2

(c) [1] **True** / **False** Two variables are not (or, at most, very minimally) related if $r = -0.02$.

---

4. Consider the following two probability questions. (Probability question (a) is separate from, is not related to, question (b).)

(a) [2] Suppose one individual is chosen at random from 500 individuals who are classified in the following way.

	huffer puffer	even breather
smoker	97	53
nonsmoker	99	251

(i) Given that the individual is a smoker, the chance s/he is a "huffer puffer" is: ________________________.

(ii) The chance the individual is a smoker *or* is an "even breather" is: ________.

(b) [2] Three marbles are randomly selected from a box containing 11 red marbles and 7 blue marbles. The probability all three are red if they are selected

(i) with replacement, is: ________________________.

(ii) without replacement, is: ________________________.

5. [3] Suppose that samples of size 2 are picked, with a sampling with replacement where order matters procedure, from a small population of five numbers, 10, 20, 30, 40 and 50, where each number has an equal chance of being chosen.

(a) The mean of the population is given by: ______________________.

(b) The standard deviation of population is given by: ______________________

(c) The mean of the sample mean is given by: ______________________.

(d) The standard error of the sample mean is given by: ______________________.

(e) Determine the sampling distribution by filling in the following table:

$\bar{x}$									
$P(\bar{X} = \bar{x})$									

(f) Suggest an appropriate simulation procedure by completing the following table,

if random number					9,0
then "choose" number	10				

---

6. Consider the following questions about the normal distribution. (Parts (a) and (b) are unrelated to one another.)

(a) [1] The random variable X has a normal distribution with a mean of 47.5 and a standard deviation of 4.0. If $P(X < a) = 0.05$,

then $a =$ ______________________.

(b) [2] In which of the following binomial distributions is the normal approximation appropriate?

(i) Appropriate / Inappropriate $n = 40$, $p = 0.1$

(ii) Appropriate / Inappropriate $n = 100$, $p = 0.01$

(iii) Appropriate / Inappropriate $n = 400$, $p = 0.05$

(iv) Appropriate / Inappropriate $n = 50$, $p = 0.02$

7. Consider the following questions on confidence intervals and tests.

(a) [1] What value is always located at the center of a confidence interval for μ?

(i) E (maximum error of estimate) (ii) μ (iii) $\bar{X}$ (iv) σ (v) s

(b) [1] **True** / **False** If our decision in a hypothesis test is to reject the null hypothesis, then we are certain that the null hypothesis is false.

(c) [1] **True** / **False** If a hypothesis test concerning a population mean is conducted at a level of significance equal to 0.05, then the probability of a Type II error equals 0.95 for any value of μ associated with the alternative hypothesis.

---

8. Consider more questions on confidence intervals and tests. (The test in (a) is separate from the test in (b).)

(a) [1] The program director for an accounting program wishes to test, at 5%, the hypothesis that her students score higher than the national average of 615 on the final exam. She randomly selects 11 recent graduates of the two–year program and discovers that $\bar{x} = 630$, and $s = 23$.

statement of test of hypotheses: ______________________________.

test statistic: ______________________________.

p–value, p: ______________________________.
conclusion: (circle one) **accept null** / **reject null** because (circle one) $p > \alpha$ / $p < \alpha$

interpretation of conclusion: ______________________________.

(b) [2] Twenty–three males with cholesterol readings in the range from 250 to 275 were randomly divided into two groups of 14 each. The two groups were put on two different diets and after 5 months, the change in cholesterol was determined for each person. The results are given below.

	diet 1	diet 2
$\bar{x}$	20.5	14.8
s	5.5	6.5
n	14	14

Does this data support the claim the variance in cholesterol readings in the first diet is less than the variance in cholesterol readings in the second diet at $\alpha = 0.05$?

statement of test of hypotheses: ______________________________.

p–value, p: ______________________________.

level of significance, α: ______________________________.
conclusion: (circle one) **accept null** / **reject null** because (circle one) $p > \alpha$ / $p < \alpha$

interpretation of conclusion: ______________________________.

9. [2] Use the following data to test if the mean number of magazines sold is the same or different for three different store locations at $\alpha = 0.05$.

front window	24	26	25	25	30
back of store	26	30	35	40	45
check out	24	24	32	33	43

Source	Sum Of Squares	Degrees of Freedom	Mean Squares
Between Groups			
Within Groups			
Total			

statement of the test is given by, __________________________________,

p–value, $p =$ __________________________________,

level of significance, α: __________________________________.
conclusion: (circle one) **accept null** / **reject null** because (circle one) $p > \alpha$ / $p < \alpha$

interpretation of conclusion: __________________________________.
____

10. [2] Consider the following table.

x	1	2	3	3
$P(X = x)$	0.60	0.20	0.15	0.05

(a) **True** / **False** This is a probability distribution

(b) **True** / **False** The histogram of this distribution is skewed to the right.

(c) **True** / **False** $P(X < 3) = 0.80$.

(d) **True** / **False** This is a binomial distribution.

1. h, a, e, c, f, b, g, d
2. (a) 74; (b) 40; (c) 98; (d) everyone
3. (a) ii; (b) ii; (c) False
4. (a) 0.65, 0.802; (b) 0.202, 0.228
5. (a) 30; (b) 14.1; (c) 30; (d) 10;
   (e) 10,15,20,25,30,35,40,45,50 and $\frac{1}{25}, \frac{2}{25}, \frac{3}{25}, \frac{4}{25}, \frac{5}{25}, \frac{4}{25}, \frac{3}{25}, \frac{2}{25}, \frac{1}{25}$
   (f) 1,2–10;3,4–20;5,6–30;7,8–40;9,0–50
6. (a) 40.9; (b) (a) inappropriate, (b) inappropriate, (c) appropriate, (d) inappropriate
7. (a) iii; (b) False; (c) False
8. (a) $H_o : \mu = 615$ vs $H_a : \mu > 615$, $t = 2.16$, $p = 0.0279$, reject null, scored higher; (b) $H_o : \frac{\sigma_1}{\sigma_2} = 1$ vs $H_a : \frac{\sigma_1}{\sigma_2} < 1$, $p = 0..716$, accept null, variances same
9. 212.8, 2, 106.4; 499.6, 12, 41.67 $H_o : \mu_1 = \mu_2 = \mu_3$ vs $H_a : \mu_i \neq \mu_j$, $p = 0.119$, accept null, average same
10. (a) True; (b) True; (c) True; (d) False