Final for Statistics 113
Elements of Probability and Statistics - Spring 2000
Material Covered: Chapters 1–27 of Workbook and text
For: 3rd May

This is a 2 hour final, worth 28% and marked out of 28 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 final. A calculator and appropriate statistical tables may also be used. No other aids are permitted.

Name (please print): _____ . **ID Number:** _____.
last first

---

1. A search of past medical journals revealed the following information of the effect of level of education on the rate of dementia.

	education →	finished high school	finished college	subtotals
dementia?	none	450	102	552
	mild	51	24	75
	severe	10	7	17
	subtotals	511	133	644

(a) [1] This study seems to show level of education does influence the rate of dementia because (circle none, one or more)

(i) $\frac{450}{511} \neq \frac{102}{133}$
(ii) $\frac{51}{511} \neq \frac{51}{75}$
(iii) $\frac{450}{552}$, $\frac{51}{75}$ and $\frac{10}{17}$ are all unequal
(iv) $\frac{102}{552} \neq \frac{24}{75}$
(v) $\frac{51}{511} \neq \frac{75}{644}$

(b) [1] Possible confounders include (circle none, one or more)

(i) gender **(ii)** health **(iii)** happiness
(iv) musical tastes **(v)** genetic inheritance

(c) [1] It does not make sense to blind the patients in this observed study.

Explain. _____.

2. In a group of 6 disturbed males, the first episode of paranoia behavior occurred at the following ages.

22, 23, 23, 26, 26, 26

(a) [1] The histogram of this data is skewed (circle one) **left** / **right** and the mean is (circle one) **larger than** / **the same as** / **smaller than** the median.

(b) [1] The SD is (circle one) **larger than** / **the same as** / **smaller than** the range of the data.

(c) [1] In a group of disturbed females, the first episode of paranoia behavior occurred, on average, at age 29 with a SD of 1.1 years. The SD in the age at which paranoid behavior started for the *combined* group of males and females is (circle one) **larger than** / **the same as** / **smaller than** 1.1 years.

---

3. Consider the following incomplete table used to calculate the correlation coefficient, r, for artistic flair, x, and insanity, y. The standard deviation for x is 1.41 and the standard deviation for y is 1.85.

x	y	x in standard units	y in standard units	product
1	2	-1.41	-0.75	1.07
2	3	(i)	-0.22	0.15
3	1	0	(ii)	0
4	5	0.71	(iii)	(iv)
5	6	1.41	1.40	1.99

(a) [2] Complete this table by filling in the following table.

(i)	(ii)	(iii)	(iv)

(b) [1] If 3 was added to each value of y, this (circle one) **does** / **does not** change r. If $(x, y) = (5, 6)$ is changed to $(x, y) = (6, 5)$, this (circle one) **does** / **does not** change r.

4. Recall that there are two different errors, chance error and bias, and that we have the following relationship:

$$\text{individual measurement} = \text{exact value} + \text{chance error} + \text{bias}.$$

(a) [1] Of the four items listed in this relationship, (circle none, one or more)

(i) only the individual measurement is known for sure and all others (exact value, chance error and bias) are most likely unknown.

(ii) all are known; in fact, the individual measurement is determined by adding the exact value, chance error and bias.

(iii) the individual measurement is estimated by the average of all measurements.

(iv) only the exact value and chance error are estimated, by the average of the measurements and SD of the measurements, respectively.

(b) [2] It is known that 32% of all American writers are alcoholics. We find that 56% of a random sample of American writers are alcoholics. Match the statistical terms with this alcoholics example.

statistical terms	alcoholics example
(i) individual measurement	(i) 56%
(ii) exact value	(ii) 32%
(iii) chance error + bias	(iii) 24%
(iv) bias	(iv) unknown

statistical terms	(i)	(ii)	(iii)	(iv)
alcoholics example				

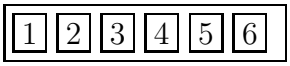
---

5. Suppose 2% of the 55,000 residences of Madeup town are criminally insane.

(a) [2] What is the chance, in a random sample of 115 individuals from Madeup, that more than 3% of these individuals are criminally insane? Circle closest one. **12%** / **23%** / **34%** / **39%** / **44%**.

(b) [1] If, instead of a random sample of 115 individuals, 785 individuals were chosen, then the SE of the percentage of criminally insane would be (circle closest one) **0.0001** / **0.0050** / **0.0150** / **0.0230** / **0.0541**.

6. [3] Three tickets are sampled with replacement where order does not matter and at random from the following box of six tickets.



It is possible, for example, that the outcome "one 1" occurs. Two examples of the outcome "one 1" are given by the two groups of three ticket arrangements below.

Match the outcomes described in words in the left column with *all* the possible three ticket arrangements given in the right column. (It is possible that none, one or more of the three ticket arrangement on the right match(es) the outcomes described on the left!)

described outcomes	three ticket arrangements
(i) all 1s	(i) 1 1 1
(ii) at least one 1	(ii) 2 1 3
(iii) not all 1s	(iii) 3 5 2
(iv) no 1s	
(v) 1st ticket is 1 or 2nd ticket is 1 or 3rd ticket is 1	
(vi) 1st ticket is 1 and 2nd ticket is 1 and 3rd ticket is 1	

outcomes	(i)	(ii)	(iii)	(iv)	(v)	(vi)
three tickets						

---

7. Consider the following problems which involve flipping a fair coin a number of times.

(a) [1] You win \$1 if the number of heads equals the number of tails, no more and no less. It is better to flip the coin (circle none, one or more) **12** / **42** / **76** / **154** / **994** times.

(b) [1] You win \$1 the closer the *actual* number of heads flipped is to the *expected* number of heads. It is better to flip the coin (circle none, one or more) **12** / **42** / **76** / **154** / **994** times.

(c) [1] You win \$1 the smaller the fraction of the "(absolute) difference between the actual number of heads flipped and the expected number of heads" to the "total number of flips of the coin". It is better to flip the coin (circle none, one or more) **12** / **42** / **76** / **154** / **994** times.

8. Reconsider the data for artistic flair, x and level of insanity, y, given in question 3 above and repeated here for convenience.

x	1	2	3	4	5
y	2	3	1	5	6

(a) [2] The regression equation for predicting level of insanity from level of artistic flair is given by (circle one)

(i) $y = x + 3$

(ii) $y = 3x + 0.4$

(iii) $y = 0.4x + 1$

(iv) $y = x + 0.4$

(v) $y = 0.4x + 3$

(b) [1] The predicted value of insanity for a level of artistic flair of $x = 3.5$ is (circle closest one) **2.4** / **4.4** / **3.9** / **6.5** / **10.9**

---

9. [2] In a random sample of 32 Professors at PU/NC, 23 were found to have Type A personalities. A 95% confidence interval of the percentage of Professors with Type A personalities

is: _____.

---

10. Consider the following couple of questions which concern the t–tables.

(a) [1] $P(t_{10} > 1.23) \approx$ (circle one) $\mathbf{(0.75, 0.90)}$ / $\mathbf{(0.25, 0.50)}$ / $\mathbf{(0.10, 0.25)}$ / $\mathbf{(0.05, 0.10)}$ / $\mathbf{(0.025, 0.050)}$

(b) [1] The t table is used in test problems where (circle *best* answer)

(i) a small random sample size is taken, where it does not matter whether the distribution is normal or not (because of the central limit theorem).

(ii) a big sample is taken, where it does not matter whether the distribution is normal or not (because of the central limit theorem).

(iii) a small sample is taken, where it is known the distribution is normal.

(iv) a small random sample size is taken, where it is known the distribution is normal.

(v) a big random sample size is taken, where it is known the distribution is normal.

(1) (a) i, iii, iv and v (b) i, ii, iii and v (c) In this observed study, the data was most likely collected after the fact and so patients do not need to be blinded.

(2) (a) left, smaller; (b) smaller; (c) larger

(3) (a) -0.71, -1.30, 0.86, 0.61 (b) **does not**, **does**

(4) (a) (i) and (iv); (b) i, ii, iii, iv

(5) (a) **26%** (b) **0.0230**

(6) (a) (i); (i) and (ii); (ii) and (iii); (iii); (i) and (ii); (i)

(7) (a) **12** (b) **12** (c) **994**

(8) (a) (iv); (b) **3.9**

(9) $0.72 \pm 2(0.08) \approx 0.72 \pm 0.16 \approx (0.56, 0.88)$

(10) (a) $\mathbf{(0.10, 0.25)}$ (b) (iv)