Final for Statistics 301
Elementary Statistical Methods - Spring 2002
Material Covered: Chapters 1–11 of Workbook and text
29th April

This is a 2 hour final, worth 25% and marked out of 25 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 may also be used. No other aids are permitted.

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

---

1. Consider the following distribution table for the time length, in minutes, of rapid eye movements (REMs) a patient experiences during a night's sleep.

class interval	number of patients	relative number	proportion per 1 minute	percent (%)
0.5 to 11.5	5	$\frac{5}{20} = 0.25$	$\frac{.25}{11} \approx 0.023$	2.3%
11.5 to 20.5	7	$\frac{7}{20} = 0.35$	$\frac{.35}{9} \approx 0.039$	3.9%
20.5 to 27.5	4	$\frac{4}{20} = 0.20$	$\frac{.20}{7} \approx 0.029$	2.9%
27.5 to 40.5	4	$\frac{4}{20} = 0.20$	$\frac{.20}{13} \approx 0.015$	1.5%
total	20	1		

(a) [1 point] The percentage of patients who have between 0.5 and 11.5 minutes of REM is (circle closest one) **0.023** / **0.039** / **0.102** / **0.250** / **0.350**.

(b) [1 point] The number of patients who have between 3.5 and 18.5 minutes of REM is (circle closest one) **0.457** / **3.514** / **8.676** / **9.140** / **12.516**.

(c) [1 point] The number of patients who have at least 3.5 minutes of REM is (circle closest one) **1.264** / **6.818** / **10.353** / **14.452** / **18.636**.

2. Selected percentiles for the weight (in pounds) of males at Purdue University, North Central is given in the table below.

percentile	25	35	60	75	90	95
weight	150	165	180	190	210	245

(a) [1 point] About 25% of the males had weights above
(circle closest one) **25** / **150** / **165** / **190** / **245** pounds.

(b) [1 point] The percentage of males with weights between 165 and 180 pounds is
(circle closest one) **25** / **35** / **60** / **65** / **70** percent.

(c) [1 point] The percentage of males that weigh at least 180 pounds is
(circle closest one) **25** / **40** / **45** / **50** / **60** percent.

---

3. The predicted tomato yield is related to the amount of nitrogen fertilizer applied to a number of 10 square meter experimental plots in the following way:

predicted tomato yield = (15 kilograms tomato per kilogram nitrogen) × (nitrogen) + 300 kilograms

(a) [1 point] An unfertilized plot can be expected to yield
(circle closest one) **0** / **15** / **205** / **300** / **315** kilograms of tomatoes.

(b) [1 point] The following data set was used in determining the least squares line given above.

nitrogen, x	0.13	0.73	1.11	1.32	1.54	1.78	2.31	2.54	2.88
tomato yield, y	301.95	310.95	316.65	319.8	323.1	326.7	334.65	338.1	343.2

The standard error of the estimate, $S_e =$ _____.

---

4. For every five flips of a coin, a head comes up three (3) times and a tail comes up two (2) times.

(a) [1 point] The chance of flipping seven heads in a row is
(circle closest one) $\frac{\mathbf{2185}}{\mathbf{78125}}$ / $\frac{\mathbf{2186}}{\mathbf{78125}}$ / $\frac{\mathbf{2187}}{\mathbf{78125}}$ / $\frac{\mathbf{2188}}{\mathbf{78125}}$ / $\frac{\mathbf{2189}}{\mathbf{78125}}$.

(b) [1 point] The chance of flipping three heads, followed by four tails, is
(circle closest one) $\frac{\mathbf{432}}{\mathbf{78125}}$ / $\frac{\mathbf{433}}{\mathbf{78125}}$ / $\frac{\mathbf{434}}{\mathbf{78125}}$ / $\frac{\mathbf{435}}{\mathbf{78125}}$ / $\frac{\mathbf{436}}{\mathbf{78125}}$.

(c) [1 point] [1 point] the chance of flipping at least one head in seven flips is
(circle closest one) $\frac{\mathbf{77994}}{\mathbf{78125}}$ / $\frac{\mathbf{77995}}{\mathbf{78125}}$ / $\frac{\mathbf{77996}}{\mathbf{78125}}$ / $\frac{\mathbf{77997}}{\mathbf{78125}}$ / $\frac{\mathbf{77998}}{\mathbf{78125}}$.

5. The registrar keeps an alphabetical list of the 3,500 students at Purdue University, North Central (PU/NC). A number between 1 and 25 is chosen at random, say 12. The twelfth student on the alphabetical list and every twenty-fifth person after that is chosen for a sample of 140 students. We are interested in estimating the average GPA of students.

(a) [1 point] This (circle one) **is** / **is not** a simple random sample students although there is no selection bias in this method of sampling.

(b) [1 point] The chance error is measured by the (circle one) **parameter** / **sample** / **standard deviation** / **average** / **statistic**

(c) [1 point] Match the two columns below.

statistical terms	GPA example
(i) population	**(i)** all PU/NC students
(ii) sample	**(ii)** average GPA of all PU/NC students
(iii) statistic	**(iii)** GPAs of 140 PU/NC students
(iv) parameter	**(iv)** average GPA of 140 PU/NC students
	(v) 140 PU/NC students
	(vi) GPAs of 3,500 PU/NC students

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

---

6. [2 points] Of a simple random sample of five hundred (500) students taken from the students at Purdue University, North Central, 134 have a GPA greater than or equal to 3.3 (and so 366 students have a GPA less than 3.3). Circle true or false.

(a) **True** / **False** The standard error is given by 1.98%.

(b) **True** / **False** A 95% confidence interval of the percentage of students at PU/NC that have a GPA greater than or equal to 3.3 is 22.9% to 30.7%.

(c) **True** / **False** A 95% confidence interval of the percentage of 500 chosen students that have a GPA greater than or equal to 3.3 is 22.9% to 30.7%.

(d) **True** / **False** If we would like to be 95% confident that our estimate of the percentage of PU/NC students that have a GPA greater than or equal to 3.3, to have a margin of error of 0.05, then the sample size should be close to $n = 301$.

7. Professor Bumble compares the average number of late attendance assignments for the Spring 2001 internet students with the average number of late attendance assignments for internet students in previous semesters. He finds that, in previous semesters, an average of 2.3 attendance assignments per semester were late, whereas an average of 1.9 attendance assignments were late with a standard deviation of 1.1 for 1000 Spring 2001 internet students chosen at random. We would like to decide if this data supports the claim the population average number of late attendance assignments for the Spring 2001 internet students is different than the population average number of late attendance assignments of previous internet students at the 5% level of significance.

(a) [1 point] The p–value is
(circle closest one) **−11.50** / **−2.45** / **0** / **0.14** / **1.37**.

(b) [1 point] $E =$ (circle closest one) **0.034** / **0.068** / **0.126** / **0.222** / **0.341**.

(c) [1 point] The assumptions used (necessary) to carry out this test are (circle none, one or more)

(i) independent observations

(ii) normal distribution

(iii) small degrees of freedom

(iv) large sample size

(v) t distribution

---

8. [1 point] A random sample of 323 junior American hockey players showed that 23 eventually played at least one National Hockey League (NHL) game. A random sample of 1005 junior Canadian hockey players showed that 67 eventually played at least one NHL game. Do these data indicate that there is a higher proportion of American juniors than Canadian juniors who played at least one NHL game? Calculate the p–value and compare it to $\alpha = 0.05$ to help you decide.

9. Consider the following data of the hormone levels of fish subjected to different (habitat, population) treatments.

	habitats →	A	B	C	D	E
	low	68	68	66	68	67
population	medium	50	49	52	48	50
	high	37	41	40	41	40
	extreme	30	31	30	29	34

(a) [1 point] Complete the following ANOVA table, if we are interested in determining if the average hormone levels for the fish are different for different population levels at $\alpha = 0.01$.

Source	Sum Of Squares	Degrees of Freedom	Mean Squares
Treatments	_____	_____	_____
Residual (error)	_____	_____	_____
Total	_____	_____	

(b) [1 point] Complete the following ANOVA table, if we are interested in determining if the average hormone levels for the fish are different for different population levels at, notice, the different level of significance, $\alpha = 0.05$.

Source	Sum Of Squares	Degrees of Freedom	Mean Squares
Treatments	_____	_____	_____
Residual (error)	_____	_____	_____
Total	_____	_____	

(c) [1 point] In both ANOVA procedures, the SS_{BET} is (circle one) **smaller than** / **bigger than** the SS_W.

10. A study of 350 internet STAT 301 students is given in the following table. For example, although 20% of students were *expected* to be graduate students who had completed all of the necessary prerequisites (category (A)) for internet STAT 301, in fact, only 5% of the group of 350 were observed to be graduate students with the necessary prerequisites.

category	expected percentage	observed percentage
(A) graduate, with prerequisites	20%	5%
(B) graduate, without prerequisites	15%	10%
(C) undergraduate, with prerequisites	35%	70%
(D) undergraduate, without prerequisites	30%	15%

(a) [1 point] The observed number of undergraduate students without prerequisites is (circle closest one) **15** / **52.5** / **105** / **112.5** / **135**.

(b) [1 point] The observed χ^2 test statistic is (circle closest one) **1.9396** / **19.396** / **193.96** / **1939.6** / **19396**.

(1) (a) **0.250** (b) **9.140** (c) **18.636**

(2) (a) **190** (b) **25** (c) **40**

(3) (a) **300** (b) **0**

(4) (a) $\frac{2187}{78125}$ (b) $\frac{432}{78125}$ (c) $\frac{77997}{78125}$

(5) (a) **is not** (b) **SD** (c) (vi), (iii), (iv), (ii)

(6) (a) **True**, **True**, **False**, **True**

(7) (a) **0**, (b) **0.068**, (c) (i), (iv)

(8) **No**, since p–value = 0.389

(9) (a) ANOVA

Source	Sum Of Squares	Degrees of Freedom	Mean Squares
Treatments	3691.35	3	1230.45
Residual (error)	37.6	16	2.35
Total	3728.95	19	

(b) same answer as (a)
(c) **bigger than**

(10) (a) **52.5** (b) **193.96**