Final for Statistics 301
Elementary Statistical Methods–Fall 2000
Material Covered: Chapters 1–11 of Workbook and text
11th December

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 quiz. A calculator may also be used. No other aids are permitted.

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

---

1. [2 points] One hundred and twenty (120) pea plants are selected at random and the number of pea pods produced per plant is measured (observed). From this group, an average number of pea pods per plant is computed. Match the columns: *All* of the items in the first column will be used up in the matching procedure; however, one item in the second column will be left unmatched.

statistical terms	pea pods example
(a) observation	**(a)** average number of pea pods per plant for 120 pea plants
(b) variable	**(b)** all pea plants
(c) parameter	**(c)** number of pea pods per plant for all pea plants
(d) statistical population	**(d)** number of pea pods for a pea plant
(e) sample	**(e)** average number of pods per plant for all pea plants
(f) statistic	**(f)** 120
(g) sample size	**(g)** number of pea pods per plant for 120 pea plants
	(h) number of pea pods for a particular pea plant

terms	(a)	(b)	(c)	(d)	(e)	(f)	(g)
pea pod example							

2. A tumor is cancerous 16% of the time; seven tumors are inspected. Assume this problem obeys the conditions of a binomial experiment.

(a) [1 point] The chance that at least 5 of the tumors are cancerous is (circle one) **0.0017** / **0.1721** / **0.3334** / **0.6734** / **0.9983**

(b) [1 point] The variance in the number of tumors is (circle one) **0.27** / **0.94** / **1.12** / **2.67** / **2.88**

(c) [1 point] A Poisson approximation in this example is *not* advisable because (circle one)

(i) $n \neq 100$ and $np \leq 20$

(ii) $n > 100$ and $np \leq 10$

(iii) $n < 100$ and $np \leq 20$

(iv) $n \geq 100$ and $np \leq 10$

(v) $n \leq 100$ and $np \leq 10$

---

3. Consider the χ^2 distribution.

(a) [1 point] For this distribution with 12 degrees of freedom,

$P(\chi^2 \leq 3.24) =$ _____.

(b) [1 point] The 37th percentile for this distribution with 9 degrees of freedom

is _____.

(c) [1 point] A χ^2 distribution with 5 degrees of freedom is (circle one)

(i) more skewed left than

(ii) as skewed left as

(iii) as symmetric as

(iv) as right skewed as

(v) more right skewed than

a χ^2 distribution with 25 degrees of freedom.

4. The average survival of leukemia patients in the midwest is assumed to be 17 months from time of diagnosis. The Cancer Research Society (CRS), however, claims the average survival time to be longer than this. The average survival time of a random sample of size $n = 15$ patients is $\bar{x} = 18.5$ months and the standard deviation in survival time is $s = 5.5$ months. Does this data support the CRS's claim at $\alpha = 0.05$? Assume normality.

(a) [1 point] A test of the CRS claim involves using the (circle best one)

(i) normal distribution.

(ii) t distribution with 14 degrees of freedom.

(iii) χ^2 distribution with 14 degrees of freedom.

(iv) F distribution with (14,15) degrees of freedom.

(v) binomial distribution.

(b) [1 point] The p–value

is _____.

(c) [1 point] A 95% confidence interval for μ

is _____.

---

5. The observed data of the incidence of colon cancer in parents and their children from a random sample of 329 families in a midwestern city is given in the table below.

	children have colon cancer	children do not colon cancer	
parents have colon cancer	18	12	30
parents do not have colon cancer	22	277	299
	40	289	329

(a) [1 point] The test of whether or not those children who have colon cancer is dependent on whether or not the parents have colon cancer at $\alpha = 0.05$ has an observed χ^2 test statistic value of (circle closest one)
56.7 / **63.3** / **68.2** / **70.0** / **70.7**

(b) [1 point] The test of whether there is a larger proportion of children with colon cancer of parents who had colon cancer than of children with colon cancer of parents who did not have colon cancer at $\alpha = 0.05$ has an observed z test statistic value of (circle closest one)
2.38 / **3.33** / **5.80** / **6.79** / **8.41**

6. In a controlled randomized experiment, the effect of different levels of (bad) fat and salt on the average volume (in mm^3) of cancer tumor nodules in mice is investigated. For instance, the average volume of cancerous tumor nodules found in the first three mice, fed a diet with a low dose of (bad) fat and a low dose of salt, are 4.2, 3.1 and 2.9 mm^3, respectively.

	(bad) fat dosage →	low (L)	medium (M)	high (H)
salt dosage ↓	low (L)	4.2, 3.1, 2.9	4.4, 3.2, 3.3	5.9, 6.2, 5.5
	medium (M)	4.1, 3.7, 3.9	5.2, 4.5, 4.2	6.2, 6.7, 5.4
	high (H)	5.2, 4.1, 3.4	5.3, 4.1, 4.9	6.7, 7.5, 6.9

(a) [1 point] Complete the following ANOVA table using the data above using the "salt" factor as the treatment variable *and ignoring the "(bad) fat" factor.*

Source	Degrees of Freedom	Sum Of Squares	Mean Squares
Treatment (Salt)	_____	_____	_____
Error	_____	_____	_____
Total	_____	_____	

(b) [1 point] Complete the following ANOVA table using the data above using the "(bad) fat" factor as the treatment variable *and ignoring the "salt" factor.*

Source	Degrees of Freedom	Sum Of Squares	Mean Squares
Treatment (Fat)	_____	_____	_____
Error	_____	_____	_____
Total	_____	_____	

(c) [1 point] The observed F statistic calculated for the ANOVA analysis using only the "salt" factor as the treatment is (circle one) **more significant than / as significant as / less significant than** the observed F statistic calculated for the ANOVA analysis using only the "(bad) fat" factor as the treatment.

---

7. Some permutations and combinations questions.

(a) [1 point] The number of different permutations that can be made from the letters in the word *arrghh* (a sound often made by students disgusted with permutations and combinations questions) is (circle one)
$\mathbf{6!}$ / $\boldsymbol{P_{6,6}}$ / $\frac{\mathbf{6!}}{\mathbf{2!2!}}$ / $\boldsymbol{P_{6,2}}$ / $\boldsymbol{P_{2,6}}$

(b) [1 point] The number of ways of dealing a five–card hand from a deck of 52 playing cards such that this hand includes all four aces is (circle one)
$\mathbf{52!}$ / $\boldsymbol{C_{52,5}}$ / $\boldsymbol{C_{5,4}}$ / $\boldsymbol{C_{48,1}}$ / $\boldsymbol{C_{52,48}}$

8. Consider the following table of wheat yield (kilograms) versus amount of water (liters per square meter).

amount of water, x	0.13	0.54	0.73	1.11	1.32	1.54	1.78	2.31	2.54	2.88
wheat yield, y	7.1	7.0	7.5	8.8	9.1	9.4	10.0	9.2	9.0	8.5

(a) [1 point] The residual at $x = 1.11$ is _____.

(b) [1 point] S_e = (circle closest one) **0.747** / **0.818** / **1.112** / **1.272** / **1.379**

(c) [1 point] A 95% CI for y_p at $x = 1.22$

is _____.

---

9. [2 points] Circle true or false.

(a) **True** / **False**. The t distribution is a "flatter" version of the standard normal distribution. The larger the sample size, n, the less flat the t distribution becomes and the more like the standard normal it becomes.

(b) **True** / **False**. A sampling distribution is a probability distribution for a statistic.

(c) **True** / **False**. Even though the sampling distribution for the average becomes more normal–shaped as the random sample size increases, the mean of the average, $\mu_{\bar{X}}$, remains the same and is equal to μ_X.

(d) **True** / **False**. Even though the sampling distribution becomes more normal–shaped as the random sample size increases, the standard deviation of the average, $\sigma_{\bar{X}}$ decreases and is equal to $\frac{\sigma_X}{\sqrt{n}}$ (not $\frac{\sigma_X}{n}$).

---

10. Consider the following table of the sweetness measurements taken from a sample of Yummy chocolate bars.

sweetness measurements	13	54	73	111	132	154	178	231	254	288

(a) [1 point] Chebyshev's rule says there

should be a proportion of at least _____ of the data within 2.6 standard deviations of the mean.

(b) [1 point] A box and whiskers plot indicates the data is (circle one) **badly skewed left** / **more or less symmetric** / **badly skewed right**.

(1) h, d, e, c, g, a, f

(2) (a) **0.0017**; (b) **0.94** (c) (v) $n \leq 100$ and $np \leq 10$

(3) (a) 0.0064; (b) 7.068 (c) **more right skewed than**.

(4) (a) (ii) t distribution with 14 degrees of freedom; (b) 0.154; (c) (15.454, 21.546).

(5) (a) **70.7**; (b) **8.41**.

(6) (a) 2, 4.927, 2.464; 24, 38.238, 1.593; 27, 43.165
(b) 2, 31.201, 15.600; 24, 11.964, 0.499; 26, 43.165
(c) **less significant than**

(7) (a) $\frac{\mathbf{6!}}{\mathbf{2!2!}}$; (b) $\boldsymbol{C_{48,1}}$.

(8) (a) 0.52; (b) **0.818**; (c) $\mathbf{(6.37, 10.35)}$

(9) **False** (should be random sample), **True**, **True**, **True**

(10) (a) **0.85**; (b) **more or less symmetric**.