Final for Statistics 503
Statistical Methods In Biology - Fall 1999
Material Covered: Chapters 1–9 of Workbook and text
15th December

This is a 2 hour final, worth 22% and marked out of 22 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. Assume the number of pups, X, born to a Collie dog, Tess, is a random variable with the following distribution,

x	1	2	3	4
$P(X = x)$	0.4	0.4	0.1	0.1

(a) [1] The chance the average number of pups born to Tess on *two* separate random independent occasions is 2.5 is (circle one) **0.04** / **0.08** / **0.12** / **0.16** / **0.20**. [Hint: For example, 2 pups are born to Tess this year and 3 pups are born to Tess next year.]

(b) [1] The expected average number of pups born to Tess on twelve separate random independent occasions is (circle one) **1.9** / **2.0** / **2.1** / **2.2** / **2.3**.

---

2. Doggy Chow Incorporated claims dog owners spend \$105 on dog food a month. A dog owners club claims the average is lower than \$105. Suppose, in a random sample of 9 dog owners, the sample average spent on dog food is \$102 and the sample standard deviation is \$3.50.

(a) [1] The observed value of the test statistic

is equal to _____.

(b) [1] The chance that the average cost of dog food for 9 dog owners is less than \$102, assuming an average of \$105,

is _____.

3. Consider the following random sample of data which is the result of an investigation of the relationship between the level of anxiety a person experiences and whether or not the person is a dog owner. We are interested in calculating a 95% confidence interval of the difference in the proportion of dog owners who have normal anxiety levels and non dog owners who have normal anxiety levels.

	dog owner	not a dog owner	subtotals
normal anxiety levels	100	101	201
reduced anxiety levels	103	53	156
subtotals	203	154	357

(a) [1] The standard error is given by
(circle closest one) **0.05** / **0.06** / **0.07** / **0.08** / **0.09**.

(b) [1] Of the four conditions that make it possible to calculate a 95% confidence interval (where a normal distribution approximation to a binomial distribution is used), (circle one) **none** / **one** / **two** / **three** / **four** of these conditions are satisfied in this question.

---

4. One ticket is drawn at random from *each* of the two boxes shown below (two tickets, in total, are drawn):

(A) $\boxed{3}\ \boxed{6}\ \boxed{9}$

(B) $\boxed{2}\ \boxed{4}\ \boxed{6}\ \boxed{8}\ \boxed{10}$

(a) [1] The chance the number drawn from A is

larger than the number drawn from B is _____.

(b) [1] The chance the
"number drawn from A is *smaller* than the number drawn from B"
or "the sum of the numbers is 10"

is _____.

5. Try the following questions.

(a) [1] A list of numbers is given. Match the two columns below.

column I	column II
(i) s of list is 0; this means	**(i)** there are no numbers on the list
(ii) $\bar{x}$ of list is 0; this means	**(ii)** the numbers on the list must be zero
	(iii) all numbers on the list must be the same
	(iv) the median of the list must be 0
	(v) the product of the list must be 0
	(vi) the sum of the list must be 0

column I	(i)	(ii)
column II		

[Hint: For example, the list, $\{0, 2, 5, 9\}$, has an average of 4.]

(b) [1] A personality test is administered to a large group of Scotch Terriers. Five scores are given below, in original units and as z scores.

original units	79	64	52	(ii)
z scores	1.8	0.8	(i)	-1.4

Complete the following table.

(i)	(ii)

6. In a particular town with 25,000 households, a simple random sample of 500 households revealed 398 households had dogs.

(a) [1] Match the two columns below.

statistical terms	dog example
(i) population	**(i)** all households
(ii) sample	**(ii)** percentage of all households with dogs
(iii) statistic	**(iii)** dog or not of 500 households
(iv) parameter	**(iv)** percentage of 500 households with dogs
	(v) 500 households
	(vi) dog or not of 25,000 households

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

(b) [1] An 88% confidence interval

is given by _____,

(c) [1] An 88% confidence interval means (circle none, one or more)

(i) there is an 88% chance that the particular confidence interval calculated above captures the true proportion of households with dogs.

(ii) there is an 22% chance that the particular confidence interval calculated above does *not* capture the true proportion of households with dogs.

(iii) we are 88% confident that the sample mean falls inside the confidence interval.

(iv) for all of the possible random samples from the 25,000 households, 88% of the confidence intervals calculated for these random samples will capture the sample proportion of households with dogs.

(v) for all of the possible random samples from the 25,000 households, 88% of the confidence intervals calculated for these random samples will capture the true proportion of households with dogs.

7. For dog owners in Indiana in 1999, data on the relationship between education (years of schooling completed), x, and earned income, y, can be summarized as follows.

$$\sum_{i=1}^{n} x_i = 145, \quad \sum_{i=1}^{n} y_i = 212, \quad \sum_{i=1}^{n} x_i^2 = 2785,$$

$$\sum_{i=1}^{n} y_i^2 = 5320, \quad \sum_{i=1}^{n} x_i y_i = 3764, \quad n = 10$$

(a) [1] The correlation coefficient is given by
(circle closest one) **−0.91** / **−0.52** / **0.06** / **0.52** / **0.92**.

(b) [1] The slope is given by
(circle closest one) **0.23** / **0.63** / **0.73** / **0.87** / **1.01**.

(c) [1] The linear regression equation is given by (circle one)

(i) $\hat{y} = 0.23x - 0.34$

(ii) $\hat{y} = 0.63x + 1.54$

(iii) $\hat{y} = 0.73x + 2.34$

(iv) $\hat{y} = 0.87x + 5.45$

(v) $\hat{y} = 1.01x + 6.54$

---

8. [1] In a study of annual family expenditures for dog health care, two populations were surveyed with the following results:

Population 1: $n_1 = 40$, $\bar{x}_1 = \$351$
Population 2: $n_2 = 35$, $\bar{x}_2 = \$321$

If it known that the population variances are $\sigma_1^2 = 2800$ and $\sigma_2^2 = 3250$, what is the probability of obtaining sample results $(\bar{x}_1 - \bar{x}_2)$ as large as those shown if there is no difference in the means of the two populations (do *not* pool)? (Circle closest one).
P–value = **0.67** / **0.79** / **0.81** / **0.93** / **0.99**.

---

9. Try the following questions.

(a) [1] **True** / **False** A randomized complete block design is a generalization of a two sample test with dependent or paired samples.

(b) [1] **True** / **False** Really, a randomized complete block design is the same as a repeated measures design, where the blocking is taken, in particular, over the subjects or experimental units used in the experiment.

10. Does the following data indicate that, as a consequence of using the five different *re*motivation methods (A, B, C, D and E), the average levels of remotivation for dogs are different at $\alpha = 0.01$?

	remotivation methods →	A	B	C	D	E
initial	nil	58	68	60	68	64
motivation	very low	62	70	65	80	69
level	low	67	78	68	81	70
	average	70	81	70	89	74

(a) [1] Complete the following ANOVA table.

Source	Sum Of Squares	Degrees of Freedom	Mean Squares
Treatments	(i)	4	(iv)
Residual (error)	(ii)	(iii)	35.6
Total	1166.8		

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

(b) [1] Based on the ANOVA table, the data (circle one) **supports** / **does not support** the claim the average levels of remotivation are different.

(c) [1] If a completely randomized design is used for this data, *initial* motivation level would probably be considered a (circle none, one or more) **treatment** / **confounding variable** / **external variable** / **response** / **type of bias**.

(1) (a) 0.16; (b) 1.9.

(2) (a) -2.57; (b) 0.016

(3) (a) 0.05; (b) four.

(4) (a) $\frac{7}{15}$; (b) $\frac{8}{15}$.

(5) (a) iii,vi; (b) 0,31.

(6) (a) vi,iii,iv,ii; (b) (0.768,0.824); (c) v.

(7) (a) 0.92; (b) 1.01; (c) v.

(8) (a) 0.99

(9) (a) true; (b) true

(10) (a) 632.8,532,15,158.2; (b) does not support; (c) confounding variable, external variable.