(1+11) = 12

I. Answer the following briefly. ~~72~~ **pts** 60

1. How many degrees of freedom do the following have? Show formulas. Find the critical value 8pts

20pts

a. a 3x4 contingency (3 rows, 4 columns)

$df = (r-1)(c-1)$ 1pt $= 2(3) = 6$ 1pt

crit $= 12.592$ 2pt

b. a regression in which there are 16 X values and 16 Y values

$n = 16$ 1pt

$df = n - 2 = 14$ 1pt

$t_{crit} = t_{.05(2)} = 2.145$

1pt 1pt

2. What is the relationship in words between the following pairs : 6pts

a. F and χ^2

χ^2 = special case of F with df denom $= \infty$ $\div \times 2)$

b. Poisson and binomial

Poisson special case of binomial with $p \to 0$, $n \to \infty$

c. t and Normal

Normal = special case of t with ∞ df

3. List the properties of the sampling distribution of the mean from the central limit theorem. 3pts

1) Normal

2) $\mu_{\bar{X}} = \mu$

3) $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$

1 each

4. Derive the test statistic formula for Anova intuitive method from the properties of the sampling distribution of the mean. 3pts

$F = \frac{n s_{\bar{X}}^2}{s_p^2}$ 1pt

s_p^2 = average of s^2 estimates σ^2 1pt

if Ho is true then:

3) $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$

$\therefore n\sigma_{\bar{X}}^2 = \sigma^2$ $\therefore n s_{\bar{X}}^2$ estimates σ^2 1pt

5. What is the problem with using model I regression to find the equation relating leaf length and width? Include a description of the method of finding the line of best fit for model I regression. 4pts

15

You cannot "fix" manipulate either variable, 2pts

Least squares regression minimizes the residuals between a Y point and the corresponding Y on the line implying X can't deviate from the line 2pts

6. a. What is the test statistic for factor A in a pure model I Anova? 2pts

$$F = \frac{MSA}{MSerror}$$

b. Explain why it is this formula in terms of expected mean squares. 2pts

A $\sigma^2 + \sigma^2_A$ 1½ to isolate the effect of A ÷ by error ½

error σ^2

c. Can you apply this test statistic to an Anova in which one factor is Model 1 and one Model 2? 1pt

no ½ interaction shows up in some factors for 2 way & (in Nested must use subgr as denominator)

7. For the following pairs state which is likely to be more powerful and explain why. 6 pts

a. 1 way or random block Anova

Random block 1pt 2pts { SSerror reduce by subtracting effect of blocks = $SS_{TOT} - SS_{Tr} - SS_B - SS_{error}$

1-way SSerror = $SS_{TOT} - SS_{Tr} - SS_{error}$

2pts SSerror is smaller in RB & MSerror is probably smaller $MSerror = \frac{SSerror}{df\,error}$ although df error is smaller as well $\therefore \uparrow F = \frac{MStr}{MSerror \downarrow}$

b. t or Mann Whitney U on normally distributed data

t on normally dist data 2 uses measurement data rather than ranks So dist. is narrower → less overlap 1pt

ex. mixed fixed A $\sigma^2 + \sigma^2_{A\times B} + \sigma^2_A$ $\sigma^2 + \sigma^2_{A\times B}$ 3 = error

you do not need to say this → random $\sigma^2 + \sigma^2_{blocks}$

$(1+11) = \cancel{72}$ ✓

I. Answer the following briefly. ~~72~~ pts 60

1. How many degrees of freedom do the following have? Show formulas. Find the critical value 8pts

a. a 3x4 contingency (3 rows, 4 columns)

20pts

$df = (r-1)(c-1)$ 1pt $= 2(3) = 6$ 1pt

crit $= 12.592$ 2pt

b. a regression in which there are 16 X values and 16 Y values

$n = 16$ 1pt

$df = n-2 = 14$ 1pt

$t_{crit} = t_{.05(2)} = 2.145$

↑1pt ↑1pt

2. What is the relationship in words between the following pairs : 6pts

a. F and χ^2

χ^2 = special case of F with df denom $= \infty$

$\div \times 2)$

b. Poisson and binomial

Poisson special case of binomial with $p \to 0$, $n \to \infty$

c. t and Normal

Normal = special case of t with ∞ df

3. List the properties of the sampling distribution of the mean from the central limit theorem. 3pts

1) Normal

1 each

2) $\mu_{\bar{X}} = \mu$

3) $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$

4. Derive the test statistic formula for Anova intuitive method from the properties of the sampling distribution of the mean. 3pts

$F = \frac{n s_{\bar{X}}^2}{sp^2}$ 1pt

sp^2 = average of s^2 estimates σ^2 1pt

if Ho is true then:

3) $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$

$\therefore n\sigma_{\bar{X}}^2 = \sigma^2 \therefore n s_{\bar{X}}^2$ estimates σ^2 1pt

8. On which scale (ordinal, interval, nominal, ratio) is each of the following variables (in bold)? Which statistic (or measure) is most appropriate (name and formula where possible) to indicate the following?

a. The proportion of variation in ***body size (log kg) in house sparrows*** associated with variation in ***latitude (degrees)*** 8pts

interval (log) or ratio OK 2

interval or ratio 2

coefficient of determination 2 r^2 2

2 5 / 5 / 0.

b. Sampling error associated with the mean ***number of plants per quadrat.*** 6pts

ratio 2

standard error or confidence limits 2

$s_{\bar{X}} = s/\sqrt{n}$ 2

$\bar{X} \pm t_{.05(2)(n-1)} \, s_{\bar{X}}$

9. Which non-parametric test(s) replaces each of the following parametric tests? 4pts

a. Pearson correlation Spearman rank 2pts rank 1pt

b. Paired t Wilcoxon or Sign

10. Which is more powerful a test – one with α = 0.05 or 0.01 assuming the same μ and σ? 2pts Illustrate your answer with a sketch(s) of power. (1pt). Label the reject and fail to reject regions (1pt), the level of significance(1pt), the probability of a type II error(1pt) and power(1pt). 7pts total

see notes

II. What error(s) has (have) been made in the following? The errors are included in the description. Some may have more than one. Be specific. What could you do with the data to correct the analysis? Include transformations, a change of test, and/or other measures as appropriate. 12pts

I should have specified the particular pop.

a. In a study of factors affecting the population growth of robins investigators performed a correlation between the number of eggs laid per female vs the number of females in a population to determine whether there are density dependent factors regulating reproduction.

2pts

1) # eggs / female counted low mean Poisson not normal
probably not homo var. or linear
$\sqrt{X}$ transform – or Spearman

females may be large enough to be normal

2) not mathematically independent 2pts
#/female vs female

do # eggs vs # females 2pts

b. A study was performed to measure the effect of the herbicide Atrazine on reproduction of the amphibian *Rana pipiens*. Sixty males received either no, low or high concentrations of Atrazine (20 per condition) as tadpoles. The degree of testes degeneration for each male was rated on a scale from 1= no degeneration to 4=much degeneration. Data were analyzed by 1 way Anova.

2pts

ranked data – likely not to be normal, homoscedastistic or additive

do Kruskal Wallis ANOVA 2pts
non-parametric (1pt)

1 1

III. For the following give the name of the sampling distribution, the null and alternate hypotheses for all sets of tests using symbols where appropriate, and the formula for the test statistic(s). 28pts

2 — 1 for o/c formula

1. To test whether two alleles for flower color segregate independently of one another.

$\chi 2$

Ho: random seg.
H1: not random

$$\chi^2 = \Sigma \frac{(O-E)^2}{E}$$

2. To determine whether a new drug relieves migraine pain more quickly than the old drug, 50 subjects were randomly assigned to either the new or the old drug and the time for the drug to take affect was noted.

t

Ho: $\mu_N \geq \mu_O$
H1: $\mu_N < \mu_O$

$$t = \frac{\bar{X}_O - \bar{X}_n}{S_{\bar{X}_O - \bar{X}_N}}$$

3. To compare each of 4 means with each other mean following a significant F

Tukey

Ho: $\mu_i = \mu_j$
H1: $\mu_i \neq \mu_j$

$$q = \frac{\bar{X}_i - \bar{X}_j}{SE}$$

4. To test for interaction in Anova.

F

Ho: no interaction between . . .
H1: interaction

$$F = \frac{MS_{A \times B}}{MS_{error}}$$

5. To determine whether variability in body length is higher than normal in a particular population of sparrows. The norm is known to be 1.34 mm with a variance of .86 while the population in question has a mean of 1.45 with a variance of 1.22 based on a sample of size 25.

χ^2

Ho: $\sigma^2 \leq .86$

H1: $\sigma^2 > .86$

$$\chi^2 = \frac{\nu s^2}{\sigma^2}$$

6. To test for a positive relation between two measurement variables in a mensurative experiment corr

r or t

Ho: $\rho \leq 0$

H1: $\rho > 0$

$$r \quad \text{or} \quad t = \frac{r - 0}{s_r}$$

7. To test for association between two variables when the data are frequencies

χ^2

Ho: the variables are not associated

H1: the variables are associated

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

IV. Test the following 3 hypotheses. Show your work. If you cannot do all the calculations assume some values and proceed so that you can get some credit. 72pts

1. Given below is the following data for the occurrence of a particular genetic abnormality. 100 areas (of equal population)were sampled and the number of areas with 0, 1, 2 mutants was determined. Does this abnormality occur at random? 26pts

# mutations per area	#areas
0	50
1	30
2	20
3	0

a. State the level of significance 1 pt

$\alpha = 0.05$

b. State the hypotheses in words. 2pts

Ho: it is random

H1: it is not random

c. Find the degrees of freedom 1 pt

Poisson

$4 - 1 - 1 = 2$

d. Calculate the mean, expected proportions, and expected numbers 4 pts

$\bar{X} = .7$ 1p

$p = \frac{e^{-\lambda}\lambda^x}{x!}$ 1p

P		E(AB)	$\frac{(O-E)^2}{E}$
.49	1p	49.0	1p
.347		34.70	
.1216		12.16	
3+ subtract		→ 3.24	

$\chi^2 = 9.11$ 2pts

math error

e. Calculate the test statistic. 4pts

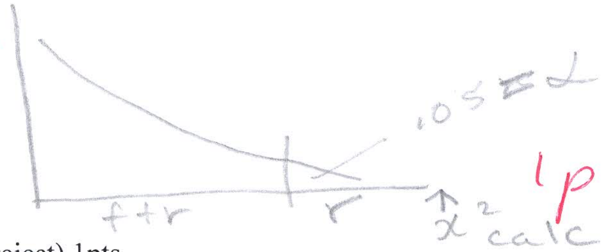
$\frac{(O-E)^2}{E}$ 2pts

f. Find p. 2pts

$0.025 > p > 0.01$

g. Find the critical and indicate it in a sketch showing the reject and fail to reject regions 3pts

$\chi^2_{crit} = 5.991$ 2pts



1pt

h. State the statistical conclusion (reject/fail to reject) 1pts

reject

i. Verbalize. 2pts

Differs significantly from random

j. Assuming you found the occurrence to be non-random, what measure would you use to determine whether it is regular or clumped? (Name, formula, and interpretation) 6pts

coefficient of dispersion $\frac{s^2}{\bar{x}}$

<1 regular (uniform)

1=0 random

>1 = contagious clumped

Random Block

2. An experiment was performed to measure the effect of crowding on testosterone levels in laboratory animals. 6 litters were chosen and one animal from each litter was randomly assigned as the focal animal to one of the three crowding conditions (extremely crowded, moderately crowded, and not crowded). (There are 18 focal animals in total-measured, 6 per treatment.) 26 pts

Say Nested - 2

Source of variation	SS	Df	MS SS/df	F calc $\frac{MS\ of}{MS_{error}}$	Fcrit 8 4.46	P
Crowding treatment	792	$3-1=2$	$792/2=396$	$396/87 = 4.55$	11 3.98 4.10 8 3.69	$.05>p>.025$
Litter	1305	$6-1=5$	$1305/5=261$	$261/87=3.00$	11 3.20 3.3	$.10>p>.05$
Error	870	10	$870/10=87$			
Total	2967	$18-1=17$				

a. Fill in the Anova table above. 13pts I have left room to show your work in the table.

b. State the level of significance 1pt

$\alpha = 0.05$

c. State all sets of hypotheses using symbols where appropriate. 4pts

1) $H_0: \mu_e = \mu_m = \mu_n$
H_1: not all means =

2) H_0: no variation among litters ∴ blocking not effective
H_1: variation among litters ∴ blocking effective

d. State the statistical conclusions and verbalize. 8pts

1) $F_{calc} = 4.55 > F_{crit} = 4.10 \rightarrow$ reject crowding significantly affects testosterone

2) $F_{calc} = 3.0 < F_{crit} = 3.33 \rightarrow$ fail to reject
no evidence of variation among litters

3. A study of factors affecting mortality from a certain type of cancer the correlation coefficient between mean annual temperature and mortality was found to be .62 based on 15 localities in Europe. Analyze. 20pts

a. State the hypotheses in words and symbols. 2pts

$H_0: \rho = 0$ no correlation betwen temp & mortaly

$H_1: \rho \neq 0$ correlation " " "

b. Find the test statistic. 4pts

$r = .62$

1-tail -1

or $t = \frac{r}{\sqrt{(1-r^2)/n-2}} = \frac{.62}{\sqrt{(1-.38)/13}} = \frac{.62}{\sqrt{\frac{.62}{13}}} \approx 2.85$

c. State the degrees of freedom 2pts

$15 - 2 = 13$

d. Find p 2pts

$0.592 < 0.62 < 0.641$

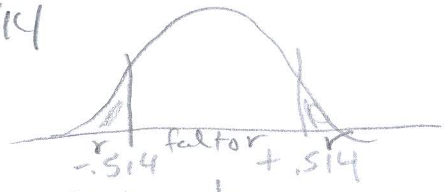
$0.02 > p\ 0.01$

e. Find the critical and indicate it in a sketch showing the reject and fail to reject regions 3pts

1-tail .441 -1.771

r critical $= r_{.05(2)(13)} = 0.514$

t crit $= 2.160$



f. State the statistical conclusions 1pts

r calc $= .62 > r$ crit reject

$p < 0.05$

g. Verbalize 2pts

there is a sign correlation between temp & mortality

h. Is the relationship biologically significant? Explain using a calculation. 4pts

coefficent of determination $r^2 = 0.62^2 = 0.38$

38% of the variation in both in this data set is explained by their correlation –

Fairly biol. significant

Statement only +1

IV. For the following state the type of test. Some types may occur more than once and others may not occur at all. Assume that one of the following tests is appropriate for the data or that the data will be transformed to meet assumptions. 33 pts

Unpaired t	Single classification Anova	Regression (I)
Paired t	Random block Anova (2way w.o. replication)	Correlation
	Factoral Anova (2way with replication)	χ^2 Goodness of fit to the (specify)
	Nested (hierarchical) Anova	χ^2 Contingency

1. In a study of the competitive abilities of *C.maculosa* investigators compared the size of the rosette of European and N. American plants. They chose 9 European and 9 American populations at random and measured 15 plants in each population.

   Nested ANOVA

2. Investigators wished to determine the relationship between the dose of a hypnotic drug and length of sleep. They used 10 doses (2, 4…20mM/kg) with 2 subjects/ dose and recorded the length of sleep in minutes for each subject.

   regression

3. A study was performed to determine which parts of the body are most likely to develop melanoma (a skin cancer). 300 women with melanoma were classified as to location of the lesion. 45 had it on the head/neck, 75 on the trunk, 34 on the arms, and 146 on the legs. (Assume the skin areas are equal.)

   χ^2 G of fit   1:1:1:1

   -2 for Corl
   -1 for wrong G of fit

4. A study investigated the hypothesis that infants who cry more have higher IQs. They measured the crying intensity of 50 infants aged 4-10 days and then measured the IQ of each when they reached 3 years.

correlation

5. A study was performed to look at the effect of photoperiod and genotype on the growth of barley plants infected by mildew isolate. Three photoperiods were used (8D/16L, 10D/14L, and 12D/12L. These were tested on 4 genotypes of barley. 10 plants were used in each of the 12 combinations

Factorial

6. In a study of geographic variation in ABO blood groups investigators compared the distribution of blood types in African Canadians from 3 different locales in Canada to determine whether there is a difference in the occurrence of the groups between locales. In each locale they typed 50 randomly chosen African Canadians.

χ^2 contingency -1 for G of fit

7. A study of the effect of sleep deprivation measured the ability of subjects to form new memories and whether the formation of new memories is dependent on the context in which they were learned. Sixty males were divided into 6 groups (10 per group): sleep deprived with positive, negative, or neutral learning environment and not deprived with the same environments. The investigators then measured the number of correct responses in a 10 minute period for each subject.

-2 for G of fit

Factorial -1 for χ^2 contingency

8. A study wished to compare the concentration of particulate pollutants in 5 areas in Montreal. Thirty days were selected at random and on each day the concentration of particulates was measured in each of the five areas.

random block ANOVA

9. In a study of red tides investigators wished to determine whether the major source of the toxin domoic acid in a particular episode is the diatom *Pseudo- nischia*. They measured both the concentration of the alga and the toxin in ten areas.

correlation

10. A study was performed to investigate the effect of disease on survival in the gray tree frog. Nine artificial ponds were used with the following treatments: Three ponds with no infection, three ponds with light infection, three ponds with heavy infection. The percentage survival of tadpoles was recorded for each of the 9 ponds.

1-way ANOVA

11. A study was performed to investigate the effectiveness of the drug Fluvoxamine in treating binge-eating disorder. 80 patients with the disorder were randomly assigned to either the drug or a placebo and at the end of an 8 week period they were classified into 3 categories with respect to the effect of the drug: no improvement, moderate improvement, total remission.

χ^2 contingency

-1 for G of fit