Biometry mid-term exam Spring 2013

NAME: _____

(250 pts total):

Section 1 – Short answer

1. (10 pts) What is the likelihood function for *N* random variables $\{X_1, X_2, \ldots, X_N\}$ drawn from a Normal distribution with mean μ and variance σ^2?

2. (10 pts) What distribution has the following probability density function?

$$f(x|u) = \frac{\Gamma(\frac{u+1}{2})}{\Gamma(\frac{u}{2})} \frac{1}{(u\pi)^{1/2}} \frac{1}{(1+\frac{x^2}{u})^{(u+1)/2}}, \mathrm{u} > 0$$

3. (5 pts) What is E[X] for the distribution in Question 2? (Hint: Can you do this without integrating?) Show your work.

4. (15 pts) Complete the following equations:

$$Beta(\alpha = 1, \beta = 1) =$$

$$If\ X_1, X_2, \ldots, X_n\ \sim N(0,1), then\ \sum_{i=1}^{n} X_i^2 =$$

$$\lim_{\alpha \to \infty} Gamma(\alpha, \beta) \rightarrow$$

5. (10 pts) Fill in the blanks to construct the 1-α confidence interval for the ratio of population variances, assuming $X_A \sim N(\mu_A, \sigma_A^2)$ and $X_B \sim N(\mu_B, \sigma_B^2)$.

$$P(\qquad\qquad \leq \sigma_A^2/\sigma_B^2 \leq \qquad\qquad) = 1 - \alpha$$

6. (10 pts) Using an equation, state the Central Limit Theorem.

7. (1 pt each) Name the R functions used to do the following tests:

The t-test: _____

The F-test: _____

The Binomial (a.k.a. proportions) test: _____

The Kolmogorov-Smirnov test: _____

Biometry mid-term exam Spring 2013

8. (4 pts each) In a single-sample t-test, what effect does reducing the value of the significance level from 0.05 to 0.01 have on….

a. …the probability of committing a Type II error (i.e., $P(Fail\ to\ reject\ H_0|H_A)$)?

b. …the standard error of the mean?

c. …the power of the test?

d. …the sample size?

9. (10 pts each) What is the Wald test for the proportion of a binomial variable? (Hint: You need to state both the test statistic and its distribution under the null hypothesis.) Name two shortcomings of the confidence interval that can be obtained by inverting the Wald test for a binomial proportion.

Biometry mid-term exam Spring 2013

10. Fill in the three empty boxes.

Part I (15 pts)

Test	Hypothesis (assuming two-tailed tests)	Test statistic T	f(T|H_0) (Distribution of T under H_0)	Assumptions
Two sample paired t-test	$H_0: \mu_A = \mu_B$ $H_A: \mu_A \neq \mu_B$			

Part II (15 pts)

Under what conditions is the t-test sensitive to violations of its assumptions?

Section 2 – Long answer

11. (40 pts) The probability density function for the exponential distribution is given by

$$f(x|\lambda) = \lambda e^{-\lambda x} \; when \; x \geq 0, \quad f(x|\lambda) = 0 \; otherwise$$

a. What is the expected value E[X] of this distribution?

b. What is the variance Var[X] of this distribution?

c. Prove that the exponential distribution is memoryless, i.e. that

$$P(X > 40 | X > 30) = P(X > 10)$$

d. Are the events $X > 40$ and $X > 10$ independent? (Hint: What would be the definition of independence in this case?)

12. (25 pts)

a. (15 pts) What is the likelihood function for *N* random variables $\{X_1,X_2,\ldots,X_N\}$ if each random draw has a (binomial) probability π_1 of being drawn from a Normal distribution with mean μ and variance σ^2, and a $(1-\pi_1)$ probability of being drawn from a Poisson distribution with parameter λ?

b. (5 pts each) Briefly describe two methods for finding the MLE for the probability π_1 – one algebraic and one numerical. (Note: You do not actually have to calculate the MLE, just tell me how you would go about doing it.)

13.(15 pts) Define “p-value” and discuss at least two criticisms raised in the readings we discussed regarding the use of p-values.

14. (50 pts) In an experiment to see if moths can retain memories through metamorphosis, 5th instar *Manduca sexta* caterpillars were trained to associate a mild electrical shock with a specific odor (ethyl acetate; EA). They were then tested to see whether they preferred clean air or EA-doped air as larvae and again as adult moths, after metamorphosis (Blackiston et al. 2008. Retention of memory through metamorphosis: can a moth remember what it learned as a caterpillar? PLoS ONE 3: e1736). The results of the experiment were as follows:

	Caterpillar treatment	
Adult response	Received shock treatment	Control
Chose clean air	32	25
Chose EA-doped air	9	21
Total	41	46

We'll use the odds ratio to measure association between the caterpillar treatment and adult response:

$$O = \frac{p}{1-p}$$

where p=probability of choosing clean air (this is arbitrary, but let's define it this way for consistency), and the odds ratio (OR), which compares the odds of choosing clean air under the two treatments, is defined as

$$OR = \frac{O_1}{O_2}$$

Part I: What is the odds-ratio (OR) for choosing clean air in this experiment (10 pts)?

Part II: Describe in words how you could use the bootstrap to obtain the standard error and a 95th percentile confidence interval for the true (population) odds-ratio (OR). (10 pts).

Part III: Using the approach described in Part II, write the R code for actually calculating the standard error and the 95th percentile confidence interval for the true (population) odds-ratio (OR). (20 pts)

Part IV: How would you calculate the bias of the estimator for the odds-ratio (OR)? (10 pts)