FINAL EXAMINATION

Statistics 305

Term 1, 2005-2006

Monday, December 12, 2005

Time: 3:30pm – 6:00pm

Student Name (**Please print in caps**): ____

Student Number: ____

Notes:

- This examination has 9 problems on the 11 following pages, plus 3 pages of tables. Check to ensure that you have a complete paper.
- The amount each part of each question is worth is shown in [ ] on the left-hand side of the page.
- Where appropriate, record your answers in the blanks provided on the right-hand side of the page.
- Your solutions must be justified; show the work and state the reason(s) leading to your answer for each question in the space provided immediately under the question.
- Clear and complete solutions are essential; little partial credit will be given.
- This is a closed book exam.
- A single **two-sided** 8.5 x 11 page of notes is allowed.
- Calculators are allowed (but not for symbolic differentiation or integration).
- No devices (including calculators) that can store text or send/receive messages are allowed.

Problem	Total Available	Score
1.	3	
2.	10	
3	7	
4.	12	
5.	17	
6.	8	
7.	22	
8.	13	
9.	8	
Total	100	

1. Consider $\boldsymbol{T}$, any estimator of the parameter θ. The **mean squared error** and **bias** of $\boldsymbol{T}$ are defined as $\boldsymbol{MSE(T)} = \boldsymbol{E[(T - \theta)^2]}$ and $\boldsymbol{Bias(T) = E(T) - \theta}$. Show that:

[3]

$$\boldsymbol{MSE(T)} = \boldsymbol{Var(T)} + \boldsymbol{Bias^2(T)}.$$

2. If $\boldsymbol{X}$ has a gamma distribution with shape parameter α and scale parameter λ, then $\boldsymbol{E(X)} = \alpha/\lambda$, $\boldsymbol{Var(X)} = \alpha/\lambda^2$ and $\boldsymbol{X}$ has a moment generating function given by:

$$M_{\boldsymbol{X}}(\boldsymbol{t}) = (1 - \boldsymbol{t}/\lambda)^{-\alpha}.$$

[8] a) Show that as $\alpha \to \infty$, the distribution of $\boldsymbol{W} = [\boldsymbol{X} - \boldsymbol{E}(\boldsymbol{X})]/\sqrt{\boldsymbol{Var}(\boldsymbol{X})}$ tends to the standard normal distribution.

[2] b) Explain how the above result yields an approximation to the $\chi^2_{\boldsymbol{f}}$ distribution for large values of the degrees of freedom $\boldsymbol{f}$.

3. Suppose $\boldsymbol{X_1}, \boldsymbol{X_2}, \ldots, \boldsymbol{X_n}$ is a simple random sample from a continuous distribution with density function $\boldsymbol{f(x)}$. Consider the new random variable $\boldsymbol{T}$, defined as:

$$\boldsymbol{T = max(X_1, X_2, \ldots, X_n)}.$$

[3] a) Find an expression for $\boldsymbol{g(t)}$, the density function of the random variable $\boldsymbol{T}$. __________________

[4] b) Suppose $\boldsymbol{f(x)}$ is the uniform density on $(0, \theta)$. Use the above result to obtain an exact $1 - \alpha$ one-sided confidence interval for θ of the form $(\boldsymbol{A}, \infty)$ based on the statistic $\boldsymbol{T}$. Give an explicit expression for the lower bound $\boldsymbol{A}$.
Hint: Consider $\boldsymbol{P}(\boldsymbol{T} < \boldsymbol{c}\,\theta) = 1 - \alpha$. __________________

4. Suppose $\boldsymbol{X_1}, \boldsymbol{X_2}, \ldots, \boldsymbol{X_n}$ is a simple random sample of size $\boldsymbol{n}$ from a Poisson distribution with parameter λ. You want to estimate $\theta = \exp(-\lambda) = \boldsymbol{P(X_i = 0)}$. Note that $\lambda > 0$, so $0 < \theta < 1$. Express your answers to b) – d) below entirely in terms of θ (not λ).

[3] a) Find $\hat{\theta}_{\boldsymbol{MM}}$, the method of moments estimator (MME) of θ. __________________

[5] b) Find a second order approximation to the expected value of the MME, $\hat{\theta}_{\boldsymbol{MM}}$.
Is $\hat{\theta}_{\boldsymbol{MM}}$ asymptotically unbiased? __________________

[2] c) Find the asymptotic variance of the MME, $\hat{\theta}_{\boldsymbol{MM}}$. __________________

[2] d) What is the asymptotic distribution of the MME, $\hat{\theta}_{MM}$? __________________

5. Total precipitation (in mm) has been recorded on a daily basis at the Vancouver airport for many years. The largest of these daily totals for a given year is called the **annual maximum**; it describes the amount of precipitation on the "wettest" day of the year. You have a data file listing the annual maxima for the last $\boldsymbol{n}$ years. These $\boldsymbol{n}$ annual maxima do not show any trend over time, so a statistical model in which these are modelled as independent and identically distributed random variables seems reasonable. Assume that the exponential distribution provides an adequate model. So, if $\boldsymbol{X_i}$ denotes the annual maxima for year $\boldsymbol{i}$, our statistical model is: $\boldsymbol{X_1}, \boldsymbol{X_2}, \ldots, \boldsymbol{X_n}$ is a simple random sample from the population with density function $\boldsymbol{f(x)}$, given by:

$$\boldsymbol{f}(\boldsymbol{x}) = \lambda \exp(-\lambda \boldsymbol{x}) \qquad \text{for } \boldsymbol{x} > 0.$$

[2] a) Our primary interest is in θ, the probability that next year's annual maximum will exceed 100 mm. Express θ in terms of λ. __________________

[4] b) Find $\hat{\lambda}_{ML}$, the maximum likelihood estimator (***MLE***) of λ ? __________________

[3] c) What is the asymptotic distribution of the ***MLE***, $\hat{\lambda}_{ML}$? __________________

[2] d) What is $\hat{\theta}_{ML}$, the ***MLE*** of θ ? __________________

[6] e) What is the asymptotic distribution of $\hat{\theta}_{ML}$?
Note: Express your answer entirely in terms of θ. __________________

6. A single observation is drawn from an exponential distribution with rate = θ; that is, from the density function $\boldsymbol{f(x)}$, given by:

$$\boldsymbol{f}(\boldsymbol{x}) = \theta \exp(-\theta \boldsymbol{x}) \qquad \text{for } \boldsymbol{x} > 0.$$

You have decided to test $\boldsymbol{H}_0$: $\theta = 1$ versus $\boldsymbol{H}_1$: $\theta \neq 1$ by rejecting $\boldsymbol{H}_0$ if either $\boldsymbol{x} < 0.04$ or $\boldsymbol{x} > 4.60$, where $\boldsymbol{x}$ is the observed value of the single observation.

[4] a) What is the probability of Type I error of this test? __________________

[4] b) What is its probability of Type II error when the true value of θ is 10? __________________

7. Suppose $\boldsymbol{X_1}, \boldsymbol{X_2}, \ldots, \boldsymbol{X_n}$ is a random sample of size $\boldsymbol{n}$ from a $\boldsymbol{N}(\mu, 1)$ population.

[5] a) Show that $\overline{\boldsymbol{X}}$ is a sufficient statistic for μ.

[6] b) Use the Neyman-Pearson Lemma to find the form of the most powerful test for testing $\boldsymbol{H}_0$: $\mu = 0$ versus $\boldsymbol{H}_1$: $\mu = 1$. Write down the explicit form of this test, giving the critical value required to achieve a significance level of α.

[2] c) Show that this test is uniformly most powerful for $\boldsymbol{H}_0$: $\mu = 0$ versus $\boldsymbol{H}_1$: $\mu > 0$.

[5] d) Suppose we use this test for testing $\boldsymbol{H}_0$: $\mu \leq 0$ versus $\boldsymbol{H}_1$: $\mu > 0$. Evaluate the power function of the test and show that it is monotonically increasing. Sketch the power as a function of μ (roughly).

[4] e) If this test is carried out at a significance level of $\alpha = 0.05$, what is the smallest sample size $\boldsymbol{n}$ so that the power of this test exceeds 0.90 when $\mu = 1$?

8. The true average wingspan of a certain species of insects is known to be 12.6 mm. A biologist is studying a similar species. A random sample of $\boldsymbol{n} = 5$ of these insects had wingspans of 12.7, 13.1, 13.3, 12.9 and 12.5 mm. (Note that the sample average is $\bar{\boldsymbol{x}} = 12.90$ mm and the sample standard deviation is $\boldsymbol{s} = \sqrt{0.10}$ mm.) The biologist wishes to check whether the species under study has the same true average wingspan. Based on experience with wingspan data from many species of insects, the biologist is willing to use the normal distribution as a statistical model for this wingspan data.

[2] a) What are the null and alternative hypotheses?

[5] b) What is the generalized likelihood ratio test for this problem? Write down its explicit form, giving the critical value required to achieve a significance level of α. **Note:** You don't need to derive it; you can just state the result.

[4] c) Using this test, what is the $\boldsymbol{p}$-value for the observed data?

[2] d) What does the biologist conclude if he wishes to carry out the test at the 5% significance level?

9. In 30 sets of gill nets in a lake, the following counts of fish were obtained:

Number of fish in set =	0	1	2	3	4
Frequency of this number of fish =	12	10	6	0	2

[8] Do the data provide convincing evidence against the null hypothesis that the number of fish per set can be modeled as following a Poisson distribution? Be sure to explain clearly the steps in the work leading to your conclusion. ____________________

THE END