STAT 231 Midterm 2
Tuesday November 15, 2016, 4:40-6:10 p.m.

Seating is predetermined so please check your seat assignment at
https://odyssey.uwaterloo.ca/teaching/schedule
Bring your Watcard. Only Pink Tie or Blue Goggle calculators may be used.
You may bring one (1) double-sided, letter sized (8.5 x 11 inches), handwritten page of notes to the test (no photocopies).

Midterm Test 2 will focus on the material in Sections 4.1-4.7 and 5.1-5.2, however you must still know the material from Chapters 1-3.

You should know and understand the following:

Given observed data and a model, determine the maximum likelihood estimate of an unknown parameter (Sections 2.2-2.4)
Invariance property of maximum likelihood estimates. (Theorem 13, page 61)
Definition of a point estimate of a parameter θ (Definition 7, page 47)
Definition of a point estimator of a parameter θ (Definition 22, page 109)
Sampling distribution of an estimator (Section 4.2)

Relative likelihood function (Definition 23, page 113)
Log relative likelihood function (Definition 25, page 114)
$100p\%$ likelihood interval (Definition 24, p. 114)
How to obtain likelihood intervals **from a graph** of the relative likelihood function or the log relative likelihood function
Behaviour of likelihood functions, log relative likelihood functions, and likelihood intervals as the sample size increases (page 114)
Interpretation of likelihood intervals (Table 4.2, page 116)

In particular, you should know how to construct the relative likelihood functions for each of the following:
Binomial(n, θ); Geometric(θ); Negative Binomial(k, θ); Poisson(θ); Exponential(θ);
$G(\mu, \sigma)$, σ known; $G(\mu, \sigma)$, μ known

Definition of a confidence interval (Definition 27, page 117)
Interpretation of a $100p\%$ confidence interval (page 118)
Behaviour of the width of a confidence interval as the sample size increases (page 118)
Definition of a pivotal quantity (Definition 28, page 118)
How to find a confidence interval given a pivotal quantity or approximate pivotal quantity
$100p\%$ confidence interval for the mean of a Gaussian distribution with known standard deviation (Example 4.4.2, page 119)

$$\bar{y} \pm a\frac{\sigma}{\sqrt{n}} \text{ where } P(Z \le a) = \frac{1+p}{2}, \quad Z \backsim G(0,1)$$

Approximate $100p\%$ confidence interval for θ for Binomial(n, θ) data (Example 4.4.3, page 120)

$$\hat{\theta} \pm a\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}} \quad \text{where } \hat{\theta} = \frac{y}{n} = \text{sample proportion}$$
$$\text{and } P(Z \leq a) = \frac{1+p}{2}, \quad Z \backsim G(0,1)$$

For a Binomial experiment choose the sample size n such that

$$n \geq \left[\frac{a\,(0.5)}{d}\right]^2 \quad \text{where} \quad P(Z \leq a) = \frac{1+p}{2} \quad \text{and} \quad Z \backsim G(0,1)$$

to ensure approximate $100p\%$ confidence interval for θ has width $\leq 2d$ (Example 4.4.5, pages 121-122).

Approximate $100p\%$ confidence interval for θ for Poisson(θ) data (Chapter 4, Problem 24(c))

$$\bar{y} \pm a\sqrt{\frac{\bar{y}}{n}} \quad \text{where } P(Z \leq a) = \frac{1+p}{2}, \quad Z \backsim G(0,1)$$

Approximate $100p\%$ confidence interval for θ for Exponential(θ) data (Assignment 3, Problem 2)

$$\bar{y} \pm a\frac{\bar{y}}{\sqrt{n}} \quad \text{where } P(Z \leq a) = \frac{1+p}{2}, \quad Z \backsim G(0,1)$$

Section 4.5:
$\chi^2(k)$ distribution and its properties and how to read Chi-squared tables
$t(k)$ distribution and its properties and how to read the t tables

Section 4.6:
Likelihood ratio statistic and its approximate distribution for large n (Theorem 33, page 128)
Likelihood based confidence intervals, in particular, a 15% likelihood interval is an approximate 95% confidence interval and a 10% likelihood interval is an approximate 97% confidence interval.

Section 4.7:
$100p\%$ confidence interval for the mean μ of a Gaussian distribution with unknown standard deviation σ (page 133)

$$\bar{y} \pm a\frac{s}{\sqrt{n}} \quad \text{where } P(T \leq a) = \frac{1+p}{2}, \quad T \backsim t(n-1)$$

$100p\%$ confidence interval for the variance σ^2 of a Gaussian distribution with unknown mean μ

$$\left[\frac{(n-1)\,s^2}{b}, \frac{(n-1)\,s^2}{a}\right]$$
$$\text{where } P(W \leq a) = \frac{1-p}{2}, \quad P(W \leq b) = \frac{1+p}{2}, \quad W \backsim \chi^2(n-1)$$

and corresponding $100p\%$ confidence interval for the standard deviation σ

$$\left[\sqrt{\frac{(n-1)\,s^2}{b}}, \sqrt{\frac{(n-1)\,s^2}{a}}\right]$$

For a Gaussian experiment with known standard deviation σ, choose the sample size n such that

$$n \geq \left(\frac{a\sigma}{d}\right)^2 \quad \text{where} \quad P\left(Z \leq a\right) = \frac{1+p}{2} \quad \text{and} \quad Z \backsim G\left(0,1\right)$$

to ensure $100p\%$ confidence interval for μ has width $\leq 2d$ (page 135).

Section 5.1:
Null hypothesis, alternative hypothesis
Test statistic or discrepancy measure (Definition 38, page 158)
$p - value$ (Definition 39, page 159) and its interpretation (page 160)
You can also listen to the p-value bears:
http://www.youtube.com/watch?v=ax0tDcFkPic&feature=related
Guidelines to be used in STAT 231 for interpreting p-values (see Table 5.1, page 160)
Important points about tests of hypotheses including practical versus statistical significance (pages 162-163)
Relationship between confidence intervals and tests of hypotheses (page 167)

Section 5.2:
Hypothesis test for $H_0 : \mu = \mu_0$ for $G\left(\mu, \sigma\right)$ model, with unknown standard deviation σ (pages 164-165)

$$p - value = 2\left[1 - P\left(T \leq \frac{\left|\bar{y} - \mu_0\right|}{s/\sqrt{n}}\right)\right] \quad \text{where} \quad T \backsim t\left(n-1\right)$$

Hypothesis test for $H_0 : \sigma^2 = \sigma_0^2$ for $G\left(\mu, \sigma\right)$ model, with unknown mean μ (pages 168-169)

$$p - value = 2P\left(U \leq \frac{\left(n-1\right)s^2}{\sigma_0^2}\right) \quad \text{if } P\left(U \leq \frac{\left(n-1\right)s^2}{\sigma_0^2}\right) \leq 0.5$$

or

$$p - value = 2P\left(U \geq \frac{\left(n-1\right)s^2}{\sigma_0^2}\right) \quad \text{if } P\left(U \geq \frac{\left(n-1\right)s^2}{\sigma_0^2}\right) < 0.5$$

where $U \backsim \chi^2\left(n-1\right)$.

There will also be short answer and multiple choice questions on the R code used in Assignments 1, 2 and 3.

In preparation for the midterm test you should do the following end-of-chapter problems and sample midterm:

Chapter 4, Problems: 1-30

Chapter 5 Problems: 1-8

Sample Midterm 2 (pp. 399-402)