1. (a) (5 marks) Let $Y_n \sim \chi^2_n$. State with reasons why

$$Z_n = \frac{Y_n - n}{\sqrt{2n}} \xrightarrow{\mathcal{D}} Z \sim N(0,1) \text{ as } n \to \infty.$$

Note: $E(Y_n) = n$ and $\text{Var}(Y_n) = 2n$.

(b) (5 marks) Let s_n^2 denote the sample variance of n i.i.d. $N(\mu, \sigma^2)$ random variables. Prove that

$$\frac{\sqrt{n-1}[s_n^2 - \sigma^2]}{\sigma^2\sqrt{2}} \xrightarrow{\mathcal{D}} Z \sim N(0,1) \text{ as } n \to \infty.$$

(c) (5 marks) Suppose that $X_i \sim N(\mu, \sigma^2)$, $i = 1, 2, \cdots, n$ and $Z_i \sim N(0,1)$ and all r.v.'s are independent. State, with reasons, the distribution of $\dfrac{\sqrt{n}(\overline{X} - \mu)}{\sigma s_Z}$, where s_Z^2 is the sample variance of $Z_1, Z_2, \cdots, Z_n$.

2. Let X be a Bernoulli r.v. with parameter p.

(a) (5 marks) Show that there exists no unbiased estimator $T = T(X)$ of $\sqrt{p}$.

(b) (10 marks) Let $X_1, X_2, \cdots, X_n$ be i.i.d. Bernoulli r.v.'s with parameter p. Show that for <u>any</u> estimator $T = T(X_1, \cdots, X_n)$, $E(T)$ must be a polynomial in p of degree no greater than n. Hence prove the generalization of (a): There exists no unbiased estimator of $\sqrt{p}$ for any $n \geq 1$.
[Hint: Divide through by $\sqrt{p}(\neq 0)$ and find $\lim\limits_{p \downarrow 0}$.]

3. (a) (10 marks) Let $X_1, X_2, \cdots, X_n$ be i.i.d. r.v.'s with density $f(x;\theta)$. Prove that

$$E\left[\frac{\partial}{\partial\theta} \ln f(\mathbf{X};\theta)\right]^2 = -E\left[\frac{\partial^2}{\partial\theta^2} \ln f(\mathbf{X};\theta)\right].$$

(b) (5 marks) Let $X_1, X_2, \cdots, X_n$ be n i.i.d. Poisson r.v.'s with probability function $p_X(x) = \dfrac{\lambda^x e^{-\lambda}}{x!}$, $x = 0, 1, 2, \cdots$.
Show that $\hat{\lambda} = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i$ is a uniform minimum variance unbiased estimator of λ, by using the Cramér-Rao inequality.

4. (a) (5 marks) Prove the following theorem.

   Let T be a sufficient statistic based on a set of n random variables $X_1, X_2, \cdots, X_n$, with joint density $f(\mathbf{x};\theta)$. Then, for any $\mathbf{x}$ and $\mathbf{y}$, with $T(\mathbf{x}) = T(\mathbf{y})$, $f(\mathbf{x};\theta)/f(\mathbf{y};\theta)$ is independent of θ, provided both numerator and denominator are nonzero.

   (b) (5 marks) Suppose that $\mathbf{X}$ is sampled from $f(\mathbf{x};\theta)$ and $T = T(\mathbf{X})$ is calculated, where T is sufficient for θ. If $\mathbf{X}$ is then discarded and T retained, explain how one may simulate an $\mathbf{X}'$, without knowing θ, such that $\mathbf{X}'$ has the same distribution as $\mathbf{X}$.

   (c) (10 marks) Let $X_1, X_2, \cdots, X_n$ be n i.i.d. $N(\mu, \sigma^2)$ r.v.'s where μ and σ^2 are unknown. Find a minimal sufficient statistic for (μ, σ^2).

5. (15 marks) Let $X_1, X_2, \cdots, Y_n$ be n i.i.d. $U(0,\theta)$ r.v.'s. Assuming that $Y_n = \max(X_1, X_2, \cdots, X_n)$ is minimal sufficient for θ, find the unique uniform minimum variance unbiased estimator for θ. State, by name, any results you use.

   [Note: $f_{Y_n}(y) = n\left(\frac{y}{\theta}\right)^{n-1}\frac{1}{\theta}$ for $0 < y < \theta$, and 0 elsewhere.] Prove that your estimator is a consistent estimator of θ.

6. (a) (5 marks) Let $(X_1, X_2, \cdots, X_n)$ be n i.i.d. $N(\mu, 1)$ r.v.'s where μ is unknown. Derive the maximum likelihood estimator (MLE) of μ.

   (b) (5 marks) Find an expression for the MLE of the 95th percentile of X_i. [Note: $x_{.95}$ is the 95th percentile of X_i if $P(X_i \le x_{.95}) = .95$.] Do not evaluate.

   (c) (5 marks) Suppose instead of observing the X_i's themselves, you observe only, K, the number of X_i's that are ≥ 0. Based on <u>this</u> information, write down the likelihood equation for μ.

7. (a) (5 marks) Let X be a single exponential observation with density $f(x;\theta) = \frac{1}{\theta}e^{-x/\theta}$, for $x > 0$, and 0 elsewhere. Find a $100(1-\alpha)\%$ confidence interval for θ. [Hint: $\frac{2}{\theta}X \sim \chi^2_2$.]

   (b) (5 marks) You plan to conduct an opinion poll to estimate the proportion of McGill undergraduate students that are in favour of allowing Coka Cola to have sole selling rights on campus. You intend to find a 95% confidence interval for this proportion, that will be no wider than .06. How many students should be polled in order to meet these goals?

8. (a) (10 marks) Let $X_1, X_2, \cdots, X_n$ be n i.i.d. $N(\mu, 4)$ r.v.'s. Find the most powerful α-level of

$$\begin{array}{lll} & H_0: & \mu = \mu_0 \\ \text{vs} & H_a: & \mu = \mu_1, \quad \mu_1 < \mu_0 \end{array}$$

Is this test uniformly most powerful for testing the following hypothesis?

$$\begin{array}{lll} & H_0: & \mu = \mu_0 \\ \text{vs} & H_a: & \mu < \mu_0. \end{array}$$

Justify your answer.

(b) You conduct an experiment to test $H_0: \mu = 3$ vs $H_a: \mu > 3$, where μ is the mean reaction time of certain brain cells. You find $p = .01$. Please mark each of the statements below as 'true' or 'false'.

i. You have disproved the null hypothesis.
ii. The probability that the null hypothesis is true is .01.
iii. You have proved that the alternative hypothesis is true.
iv. You can deduce the probability of the alternative hypothesis.
v. You know, if you decided to reject the null hypothesis, the probability that you are making the wrong decision is $< .05$.

Some formulae:

$$\overline{X} \pm t_{n-1}(\alpha/2)\frac{s}{\sqrt{n}}, \qquad \hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \qquad (\hat{p} - p_0)\Big/\sqrt{\frac{p_0(1-p_0)}{n}}.$$

McGILL UNIVERSITY

FACULTY OF SCIENCE

FINAL EXAMINATION

MATHEMATICS 189-357B

STATISTICS

Examiner: Professor D.B. Wolfson | Date: Wednesday, April 12, 2000
Associate Examiner: Professor K. Worsley | Time: 9:00 A.M. - 12:00 Noon.

INSTRUCTIONS

Answer all questions.
Calculators are permitted.
Tables have been provided.
Some formulae are provided on the last page.

This exam comprises the cover, three pages of questions and one page of tables.