STAT 512 - Sample Final Exam

Instructions: Answer *any six* of these eight equally weighted questions. You may quote, without proof, results which were established in class (unless you are explicitly asked to prove them) but <u>state what they are</u>. This is important and will affect your score.

Time: 3 hours.

1. (a) State what it means for a sequence to be a Cauchy sequence.

   (b) Show that the sequence $\{s_n\}_{n=1}^{\infty}$, where $s_n = \sum_{i=1}^{n} (1/i)$, is not a Cauchy sequence and is therefore divergent.

2. (a) Obtain a series expansion of $(1-x)^{-3}$, valid for $|x| < 1$.

   (b) Recall the definition of the hyperbolic cosine: $\cosh(z) = \left(e^z + e^{-z}\right)/2$. A random variable X has probability generating function $\phi(z) = \cosh(z)/\cosh(1)$. What is the probability distribution of X?

3. Suppose that $X_1, ..., X_n$ are independent, identically distributed random variables with density $f(x) = \lambda e^{-\lambda x}$ $(x > 0)$ for some positive parameter λ. Show that the density of $Y = \sum_{i=1}^{n} X_i$ is $g(y) = (\lambda y)^{n-1} \lambda e^{-\lambda y} / (n-1)!$.

4. Let $X_1, ..., X_n$ be independent, identically distributed random variables with mean μ, variance σ^2 $(0 < \sigma^2 < \infty)$ and finite third moment. Show that, as $n \to \infty$,

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_k - \mu) \xrightarrow{\mathcal{L}} N\left(0, \sigma^2\right).$$

5. (a) Suppose we make n observations of a random variable X, obtaining the numerical values $x_1 < \cdots < x_n$. Show that their average $\bar{x}$ can be represented as a Riemann-Stieltjes integral: $\bar{x} = \int_{-\infty}^{\infty} x dF_n(x)$ for some distribution function $F_n(x)$. Plot this distribution function.

   (b) Let X be a nonnegative, integer valued random variable with a finite mean. Show that

$$E[X] = \sum_{n=0}^{\infty} P(X > n).$$

...over

6. Consider the problem of estimating a parameter vector $\boldsymbol{\theta}$, by Least Squares, in the linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \textit{random error}$. Here $\mathbf{y} : n \times 1$ and $\mathbf{X} : n \times p$ are constants and $\mathbf{X}$ has rank $p < n$. Suppose that the parameters are required to satisfy q independent linear constraints of the form $\mathbf{A}\boldsymbol{\theta} = \mathbf{0}_{q\times 1}$, where $\mathbf{A}_{q\times p}$ has rank q. Thus the mathematical problem is

$$\textit{Minimize } S(\boldsymbol{\theta}) = ||\mathbf{y} - \mathbf{X}\boldsymbol{\theta}||^2 \textit{over } \boldsymbol{\theta} \in R^p, \textit{ subject to constraints } \mathbf{A}\boldsymbol{\theta} = \mathbf{0}_{q\times 1}.$$

Show that any solution to this problem is of the form

$$\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}_0 - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'\mathbf{t},$$

where $\hat{\boldsymbol{\theta}}_0 = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ and $\mathbf{t}$ is a $q \times 1$ vector chosen so that $\hat{\boldsymbol{\theta}}$ satisfies the given constraints.

7. Consider the nonlinear regression problem $y_i = f(\mathbf{x}_i, \boldsymbol{\theta}) + \varepsilon_i$, $i = 1, ..., n$. Outline the Gauss-Newton algorithm for least squares estimation in this model. Show that if the iterative procedure converges, then its limit is a stationary point of $S(\boldsymbol{\theta}) = \sum [y_i - f(\mathbf{x}_i, \boldsymbol{\theta})]^2$.

8. (a) Let $u(t)$ and $v(t)$ be linear functions of $t \in [0, 1]$, with $v(t) > 0$. Define $w(t) = u^2(t)/v(t)$ and show that $w(t)$ is convex.

   (b) Let $\mathcal{F}$ be a convex class of distributions functions F on the real line, with densities f. Recall that the asymptotic variance of an M-estimate of location, using scores $\psi_0(x) = -f_0'(x)/f_0(x)$ for a member F_0 of $\mathcal{F}$, is

$$V(\psi_0, F) = \frac{E_F\left[\psi_0^2(X)\right]}{\left\{E_F\left[\psi_0'(X)\right]\right\}^2}.$$

   Show that, in order that $V(\psi_0, F)$ be maximized by F_0, it is necessary and sufficient that

$$\int \left(2\psi_0'(x) - \psi_0^2(x)\right)(f_1(x) - f_0(x))\, dx \geq 0,$$

   for all $F_1 \in \mathcal{F}$.