MATHEMATICAL PROBABILITY AND STATISTICS (II) QUALIFYING EXAM (MATH 541A AND 507A)

FALL 1994

(1) (a) Define "The family of random variables $\{X_1, X_2, \ldots\}$ is uniformly integrable."

(b) Give an example of a sequence of random variables X_n, $n = 1, 2, \ldots$ where $X_n \geq 0$, $\mathbb{E}X_n = 1$ and the family of random variables $\{X_1, X_2, \ldots\}$ is not uniformly integrable.

For parts (c) and (d) assume that $X_n \to X$ a.s. as $n \to \infty$ and that $X_n \geq 0$ for all n, ω.

(c) Assume that $\{X_1, X_2, \ldots\}$ is uniformly integrable. Prove that $\mathbb{E}X_n \to \mathbb{E}X$ as $n \to \infty$.

(d) ASsume that $\mathbb{E}X_n \to \mathbb{E}X$ as $n \to \infty$. Prove that $\{X_1, X_2, \ldots\}$ is uniformly integrable.

(e) Let $f : [0, \infty) \to [0, \infty)$ with

$$\lim_{x \to \infty} \frac{f(x)}{x} = \infty$$

If $X_n \geq 0$, $\mathbb{E}f(X_n) \geq c < \infty$, show that $\{X_1, X_2, \ldots\}$ is uniformly integrable.

(2) (a) Let Y_n, $n \geq 1$ be random variables and Y_n' independent of Y_n with the same distribution as Y_n. If Y_n converges in distribution show that $Y_n \to Y_n'$ also converges in distribution.

(b) Let $X_1, X_2, \ldots$ be independent and identically distributed with characteristic function $f(t) = \exp(-c|t|^\alpha)$ where $c > 0$, $\alpha > 0$ are constants, and let $S_n = X_1 + X_2 + \cdots + X_n$. Find constants a_n so $a_n S_n$ converges in distribution to some random variable Z. How is Z related to X_1?

(3) Let $\theta \in \Theta = (-\infty, \infty)$ and $X_1, X_2, \ldots, X_n$ be independent and identically distributed with density

$$f(x; \theta) = \boldsymbol{I}(|x - \theta| \leq 1/2)$$

as usual, denote the order statistics by $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$.

(a) Show that $(X_{(1)}, X_{(n)})$ is sufficient for θ.

(b) Show that $T_n = \frac{1}{2}(X_{(n)} + X_{(1)})$ is unbiased for θ.

(c) Compute the variance of T_n.

(d) Find a maximum likelihood estimate for θ based on $(X_{(1)}, X_{(n)})$. Is it unique?

(4) Let $\theta \in \Theta = (0, \infty)$ and $X_1, X_2, \ldots, X_n$ be independent and identically distributed with density

$$f(x; \theta) = \boldsymbol{I}(x \in [0, \theta])$$

Construct uniformly minimum variance unbiased estimators $q(\theta)$ for the following choices of $q(\theta)$, or prove they do not exist.

(a) $q(\theta) = \theta^k$ for $k \in \{1, 2, \ldots\}$.

(b) $q(\theta) = e^\theta$.

MATHEMATICAL PROBABILITY AND STATISTICS (II) QUALIFYING EXAM (MATH 541A AND 507A)

SPRING 1995

(1) Recall that we say the distribution functions F_n converge in distribution to the distribution function F, written $F_n \Rightarrow F$, if

$$\lim_{n\to\infty} F_n(x) = F(x)$$

at all continuity points of F. Show that $F_n \Rightarrow F$ if and only if for all bounded continuous functions h

$$\int h \, \mathrm{d}F_n \to \int h \, \mathrm{d}F \quad \text{as } n \to \infty$$

(2) Suppose $X, X_1, X_2, \dots$ are iid with values in $(1, \infty)$. State a condition on X which is necessary and sufficient for $\lim_{n\to\infty}(X_1 X_2 \cdots X_n)^{1/n}$ to exist almost surely and be finite. Demonstrate why your condition is necessary.

(3) (a) Let $X_1, \dots, X_n$ be independent normal variates with common variance σ^2. With $\mu_1, \dots, \mu_n$ not all zero, derive the level α most powerful test for the hypotheses

$$H_0 : \mathbb{E}X_1 = \cdots = \mathbb{E}X_n = 0 \quad \text{versus} \quad H_1 : \mathbb{E}X_1 = \mu_1, \dots, \mathbb{E}X_n = \mu_n$$

(b) Find the level α most powerful test for the above hypotheses when $\boldsymbol{X} = (X_1, \dots, X_n)$ is multivariate normal with known covariance matrix $\boldsymbol{\Sigma}$.

(4) (a) Complete the following statement of the factorization theorem. In a regular model, a statistic $T(X)$ is sufficient for θ if and only if there exists functions g and h such that:

(b) Let $X_1, X_2, \dots, X_n$ be independent Cauchy (θ) random variables each with density

$$p(x;\theta) = \frac{1}{\pi}\left(\frac{1}{1+(x-\theta)^2}\right)$$

Show that the order statistics $(X_{(1)}, \dots, X_{(n)})$ are *minimal* sufficient for θ.

4. Let $\mathbf{X} = (X_1, X_2)$ be bivariate normal, with commom unknown mean μ, and known variances σ_1^2, σ_2^2, and correlation ρ.

a) What is the Fisher information $\mathbf{I}(\mu)$ based on one observation of the pair (X_1, X_2)?

b) Express $\mathbf{I}(\mu)$ as a function of ρ in the special case $\sigma_1^2 = \sigma_2^2 = 1$.

c) Explain why the expression found is part b takes on the values it does, in the (further) special cases when $\rho = 0$ and $\rho = 1$.

d) What is the Fisher information $\mathbf{I}(\mu)$ based on one observation $\mathbf{X}$ of a multivariate p-dimensional normal vector whose components have common mean μ and known covariance matrix Σ?

5. Recall for $\mu > 0$ the exponential density with parameter μ is $\mu \exp(-\mu x)$ when x is positive. Let μ, ν be positive and suppose that $X_1, \ldots, X_n$ are exponential with parameter μ, and $Y_1, \ldots, Y_m$ are exponential with parameter ν and that all variables are independent.

a) Construct the generalized likelihood ratio test for the hypotheses $H_0 : \mu = \nu$ versus $H_1 : \mu \neq \nu$ and show that it can be based on the statistic

$$T(X_1, \ldots, X_n, Y_1, \ldots, Y_m) = \frac{X_1 + \cdots + X_n}{X_1 + \cdots + X_n + Y_1 + \cdots + Y_m}.$$

In particular, express the critical region of this test in terms of T.

b) What is the distribution of T under the null hypotheses?

c) For the case $n = m$, show that the rejection region can be written as $\{T : |T - a| > b\}$ and find a. Can the type I error probability in this case be written entirely in terms of the cumulative distribution function of T?

(b) Justify the claim: if $S_n/n \to a$ in probability, then $\phi'(0)$ exists and equals ia. You may assume without proof that for characteristic functions, if $\phi_n \to \phi$ pointwise , then the convergence is uniform on compact sets.

Q4 Assume $X, X_1, X_2, \ldots$ are i.i.d. with $X \geq 0$ always.
(a) Prove that if $\mathbb{E}X < \infty$ then $X_n/n \to 0$ almost surely.
(b) Prove that if $\mathbb{E}X = \infty$, then $\limsup X_n/n = \infty$ almost surely, and hence $(X_1 + \cdots + X_n)/n \to \infty$ almost surely.

Section II. Statistics

DO ANY TWO OF THE FOLLOWING THREE PROBLEMS.

Q1 Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables having exponential distribution with mean $1/\theta$.
(a) Let $X_{(1)} \leq \cdots \leq X_{(n)}$ denote the order statistics, and write $T_1 = X_{(1)}, T_2 = X_{(2)} - X_{(1)}, \ldots, T_k = X_{(k)} - X_{(k-1)}$ for $2 \leq k \leq n$. Find the joint distribution of $T_1, T_2, \ldots, T_k$.
(b) Light bulbs are known to have this exponential distribution. Suppose that a random sample of n bulbs is put under observation, and observation is stopped when the k^{th} bulb burns out. Find the maximum likelihood estimator of the mean lifetime of a bulb.
(c) Find a 95% confidence interval for this mean.

Q2 (a) Give the definition of a sufficient statistic.
(b) Let $X_1, \ldots, X_n$ be a random sample from a Poisson distribution with parameter λ. Find a sufficient statistic for λ.
(c) Show that $T = \frac{1}{n}\sum_{j=1}^{n} I(X_j = 0)$ is an unbiased estimator of $e^{-\lambda}$, and find its variance. Here, $I(A) = 1$ if A is true, $= 0$ if false.
(d) Use the Rao-Blackwell Theorem and (c) to find a better estimator of $e^{-\lambda}$.
(e) What optimality properties does the estimator in (d) have?

Q3 (a) State the Cramér-Rao Inequality.

(b) Let $X_1, \ldots, X_n$ be independent and identically distributed Bernoulli random variables with mean θ. Find the Cramér-Rao lower bound for the variance of unbiased estimators of $\tau(\theta) = \theta(1-\theta)$.

(c) State and prove the Neyman-Pearson Lemma.

(d) 1000 individuals were classified according to sex, and according to whether or not they were color-blind as follows:

	Male	Female
Normal	442	514
Color-blind	38	6

According to a genetic model, these numbers should have relative frequencies given by

$p/2$	$p^2/2 + pq$
$q/2$	$q^2/2$

where $q = 1 - p$ is the proportion of color-blind individuals in the population. Are the data consistent with this model?

MATH 507a/541b QUALIFYING EXAM - FALL 1998

To pass you must do well enough on both the Probability and the Statistics parts – high performance on one portion does not compensate for low performance on the other.

STATISTICS

1. There are g categories, $i = 1, 2, ..., g$ with probabilities $\pi_1, \pi_2, ..., \pi_g$. The random variable X is defined by

$$P(X = i) = \pi_i, \qquad 1 \le i \le g.$$

Let $X_1, X_2, ..., X_n$ be *iid* as X. Define Z_{ij} by

$$Z_{i,j} = \mathbf{I}(X_j = i).$$

The number of times the variables $X_1, \ldots, X_n$ fall in category i is given by

$$Y_i = \sum_{j=1}^{n} Z_{ij}.$$

a) Find the mean and variance of Y_1.

b) Find $P((Y_1, \ldots, Y_g) = (n_1, \ldots, n_g))$.

c) Find the mean vector μ and covariance matrix Σ of the vector $(Y_1, \ldots, Y_g)$.

d) Give a large sample test for the hypothesis $H_0 : \pi_1 = \pi_2$.

2. Let $X_1, \ldots, X_n$ be *iid* $\mathcal{N}(\mu, \sigma^2)$ with known mean μ.

a) Find the UMVE of the parameter σ^2 and prove it to be such.

b) Of *all* estimators of σ^2 of the form

$$\hat{\theta} = a \sum_{i=1}^{n} (X_i - \mu)^2, \quad a \in \mathbf{R},$$

find the one which achieves the smallest mean square error. (Hint: If Z is $\mathcal{N}(0,1)$ then $EZ^4 = 3$.)

Math 541b

1. Let a_n and μ_n be deterministic sequences tending to ∞ and μ respectively, and assume that the random variables X_n, properly scaled, converge in distribution to X; in particular, that

$$a_n(X_n - \mu_n) \xrightarrow{d} X.$$

a) Prove that if g is a function having a continuous derivative at μ, then

$$a_n(g(X_n) - g(\mu_n)) \xrightarrow{d} g'(\mu)X.$$

Now let $Y_1, \ldots, Y_n$ be a sample of independent exponential variables ('failure times') with density $f(t; \lambda) = \lambda e^{-\lambda t}$ for λ, t positive.
b) Calculate the Fisher information for λ in the sample.
c) Find, and justify, the limiting distribution of the maximum likelihood estimator for λ.
d) Suppose it is desired to estimate the probability that an exponential from the same distribution will not fail before time x; that is, we wish to estimate

$$q(\lambda) = P(Y > x) = e^{-\lambda x}.$$

What is the limiting distribution of the maximum likelihood estimator of $q(\lambda)$? (Hint: Use part a)

2. a) Prove the following form of the Neyman Pearson Lemma: If $\mathbf{X} \in \mathbf{R}^n$ is a random vector with density $f(\mathbf{x}; \theta)$, where $\theta \in \{\theta_0, \theta_1\}$, then the test for $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$ which rejects H_0 when $L(\mathbf{X}) = f(\mathbf{x}; \theta_1)/f(\mathbf{x}; \theta_0)$ exceeds a level k achieves the maximum power over all tests of size $P_0(L(\mathbf{X}) \geq k)$.
b) Let $X_1, \ldots, X_n$ be independent exponential variables with parameters either $\mu_1, \ldots, \mu_n$, or $\nu_1, \ldots, \nu_n$, known values. Find a simple test and test statistic for the Neyman Pearson tests that distinguish between the two hypotheses.

Math 541 Exam Portion

1.a) Let a_n and μ_n be deterministic sequences tending to ∞ and μ respectively, and assume that the random variables X_n, properly scaled, converge in distribution to X; in particular, that

$$a_n(X_n - \mu_n) \xrightarrow{d} X.$$

Prove that if g is a function having a continuous derivative at μ, then

$$a_n(g(X_n) - g(\mu_n)) \xrightarrow{d} g'(\mu)X.$$

b) State a multidimensional version of this fact.

Now let $X_1, \ldots, X_n$ be iid with mean μ and variance σ^2.

c) Find a method of moments estimator for the coefficient of variation

$$CV = \frac{\sigma}{\mu}.$$

d) Find the asymptotic distribution of the estimator in c). What moments of the X distribution need to exist?

2) Let $X_1, \ldots, X_n$ be iid normal with unknown mean μ and known variance σ^2.

a) Find the critical region for the Neyman Pearson test at level $\alpha \in (0, 1)$ for $H_0 : \mu = \mu_0$ versus $H_1 : \mu = \mu_1$ with $\mu_0 < \mu_1$.

b) Determine the power function $\beta(\mu)$ of this test.

MATH 507a/541 QUALIFYING EXAM. MAY 15, 2000.

PLEASE NOTE: To pass you must do well enough on both the Probability and the Statistics sections. High performance in one portion does not compensate for insufficient performance on the other.

STATISTICS SECTION

Q1. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables having density

$$f(x) = \begin{cases} e^{-(x-\theta)} & \text{if } x > \theta \\ 0 & \text{if } x \le \theta \end{cases}$$

for some $\theta \in (-\infty, \infty)$.

(a) State the Factorization Theorem for sufficient statistics.

(b) Find a one-dimensional sufficient statistic for θ.

(c) Find a 95% confidence interval for θ.

(d) Derive the likelihood ratio test of the null hypothesis that $\theta \ge 0$ against the alternative that $\theta < 0$.

Q2. Outputs $X_1, X_2, \ldots, X_n$ from a physical device are independent and identically distributed random variables having exponential distribution with (unknown) mean λ^{-1}. A measuring device records the values of the X_j as long as $X_j < c$, for some known threshold $c > 0$. If $X_j \ge c$ then the device returns the value c. Define

$$S_n = \sum_{j=1}^{n} X_j I(X_j < c), \quad T_n = \sum_{j=1}^{n} I(X_j \ge c),$$

where $I(A)$ denotes the indicator of the event A.

(a) Write down the likelihood function of the observed values in terms of S_n and T_n.

(b) Show that the Maximum Likelihood Estimator of λ is

$$\hat{\lambda} = \frac{n - T_n}{S_n + cT_n}.$$

(c) Find the joint asymptotic distribution of (S_n, T_n).

Hint:

$$\int_0^c x\lambda e^{-\lambda x} dx = \lambda^{-1}\left(1 - (1 + c\lambda)e^{-c\lambda}\right)$$

and

$$\int_0^c x^2 \lambda e^{-\lambda x} dx = \lambda^{-1}\left(2 - (2 + 2c\lambda + c^2\lambda^2)e^{-c\lambda}\right).$$

(d) Using the result of the previous part, or otherwise, find the asymptotic distribution of $\hat{\lambda}$.

Math 541a Exam Portion. Spring 2001

Problem 1. a) Let $\mathbf{X} \sim \mathcal{N}(\lambda\mu, v^2\Sigma)$, where $\mu \in \mathbf{R}^n$ and $\Sigma \in \mathbf{R}^{n\times n}$ are a known vector and positive definite matrix respectively, and $\lambda \in \mathbf{R}$ and $v^2 > 0$ are unknown parameters in $\mathbf{R}$.

a) Find the maximum likelihood estimators $\hat{\lambda}$ and $\hat{v}^2$ of λ and v^2 on the basis of the observation $\mathbf{X}$.

b) Determine whether or not $\hat{\lambda}$ is unbiased for λ.

c) Calculate the variance of $\hat{\lambda}$.

d) *Demonstrate* what the estimators $\hat{\lambda}$ and $\hat{v}^2$ become when $\mu = (1, 1, \ldots, 1)$ and Σ is the identity matrix. Explain.

Problem 2. a) Let $\theta > 0$ be unknown and suppose that (X, Y) is uniform over the triangular region with vertices at $(0,0)$, $(\theta, 0)$ and $(0, \theta)$. Let (X_i, Y_i) be iid as (X, Y).

a) Find a one dimensional sufficient statistic T for θ, and prove it is sufficient.

b) Find an unbiased estimate of $\hat{\theta}$ which is a function of T.

c) Is $\hat{\theta}$ UMVU? Prove your claim.

507a Qualifying exam. September 12, 2001. Be sure to attempt the later parts of each problem even if you cannot do one of the earlier parts.

1.)

a) Prove that for any sequence X_n of random variables there exist positive constants c_n such that X_n/c_n converges to 0 almost surely.

b) Can you choose c_n so that this convergence is pointwise at every $\omega \in \Omega$?

c) Now suppose that $X_1, X_2, \ldots$ are iid, that $\mathbb{E}X_1$ exists and is finite, and that $c_n = n$. Prove that X_n/c_n converges to 0 almost surely.

2.) Assume $X, X_1, X_2, \ldots$ i.i.d. with characteristic function ϕ for X, i.e. $\phi(t) := \mathbb{E}e^{itX}$, and let $S_n := X_1 + \cdots + X_n$

a) For a random variable X, what special property of its characteristic function ϕ holds if and only if X and $-X$ have the same distribution? (Show both the implications.)

b) Express the characteristic function of the sample average, $\phi_{S_n/n}(t)$, in terms of ϕ.

c) If X has $\phi'(0) = 0$, show that $(X_1 + \cdots + X_n)/n$ converges to zero in probability. [HINTS: Since $\phi(0) = 1$, $\phi'(0) = 0$ if and only if $\phi(u) = 1 + o(u)$ as $u \to 0$. Also, for fixed t, as $n \to \infty$, $(1 + o(t/n))^n \to 1$ can be shown from $\log(1 + x) \sim x$ for small positive x.]

From now on assume that $X, X_1, X_2, \ldots$ are i.i.d. with the symmetric density

$$f(x) = c\frac{1}{x^2 \log |x|} \text{ for } |x| > 4; \quad f(x) = 0 \text{ otherwise,}$$

where c is an appropriate normalizing constant.

d) Show that $\mathbb{E}|X| = \infty$.

e) Show that the characteristic function ϕ for X has $\phi'(0) = 0$. [HINT: express $1 - \phi(t)$ as an integral over $x > 4$ and use the change of variables $y = tx$ to show that $|1 - \phi(t)|/t \to 0$ as $t \to 0$. You might use $|1 - \cos y| \le y^2 \; \forall y$, and dominated convergence.]

February 2002

Math 507a Exam

Problem 1.

Let $X_1, X_2, \ldots$ be a sequence of iid random variables so that $EX_1 = 0, E|X_1|^2 < \infty$. Show that

$$\lim_{n\to\infty} \max\left\{\frac{|X_1|}{\sqrt{n}}, \ldots, \frac{|X_n|}{\sqrt{n}}\right\} = 0$$

in probability.

Problem 2.

Let $X_1, X_2, \ldots$ be a sequence of random variables and $\varepsilon_1, \varepsilon_2, \ldots$, a sequence of real numbers so that $\varepsilon_n > 0$, $\sum_{n\geq 1} \varepsilon_n < \infty$, and $\sum_{n\geq 1} P(|X_n| > \varepsilon_n) < \infty$. Show that the series $\sum_{n\geq 1} |X_n|$ converges with probability one.

Math 541b Qualify Exam. Fall 2002

Problem 1. (EM) There are two possibly biased coins. The probability of heads for the first coin is $1/3$ and the probability of heads in the second coin is $p \in (0,1)$, an unknown parameter. An experiment consists of tossing the two coins together, which we do n times. Only X_i, the number of heads in the i^{th} experiment, is observable.

1. Let n_j, $j = 0, 1, 2$ be the number of experiments where j heads show up. Write the joint distribution of $(X_1, X_2, \cdots, X_n)$ in terms of n_0, n_1, n_2.

2. Write an equation for the maximum likelihood estimate (MLE) of p. Is it easy to solve this equation? If not, design an expectation-maximization (EM) algorithm for calculating this MLE.

3. Although we do not have a 'closed form' maximum likelihood estimator $\hat{p}$ for p, we can still study its approximate distribution. What is the approximate distribution of $\hat{p}$ when the sample size n tends to infinity?

4. It was suspected that the second coin is unbiased, that is, that $p = 1/2$. Outline a procedure for testing this hypothesis.

Problem 2. Let $\Theta = (0, \infty)$ and suppose that the density $f(x, y; \theta)$ of (X, Y) is uniform over region A, where $H_0 : A = [-\theta, \theta]^2$ (the square of side length 2θ centered at the origin, or $H_1 : A$ is the circle of radius θ centered at the origin. Let $(X_1, Y_1), \ldots, (X_n, Y_n)$, be i.i.d. with density $f(x, y; \theta)$.

a) For fixed known $\theta \in \Theta$, describe the (non-trivial) Neyman Pearson tests for the testing between the simple hypotheses H_0 vs. H_1?

b) A hypotheses test is said to be *consistent* if the probability of rejecting the null hypotheses when it is false tends to 1 as the sample size n tends to infinity. Prove that the test in part a) is consistent.

c) Describe a consistent test for the composite hypotheses H_0 vs. H_1 when θ is only known to lie in Θ.

Spring 2003 Math 541b Exam

1. Let $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$ be a vector of given probabilities, and $\mathbf{e}_i$ the i^{th} unit vector in $\mathbf{R}^n$ with a 1 in position i and zeros elsewhere. Let $\mathbf{Y}, \mathbf{Y}_1, \dots, \mathbf{Y}_n$ be i.i.d. with distribution

$$P(\mathbf{Y} = \mathbf{e}_i) = \theta_i, \quad \mathbf{X}_n = \sum_{j=1}^{n} \mathbf{Y}_j \quad \text{and} \quad \bar{\mathbf{X}}_n = \frac{1}{n}\sum_{j=1}^{n} \mathbf{Y}_j$$

so that $\mathbf{X}_n = (n_1, \dots, n_k)$ has the multinomial distribution $M(n, \boldsymbol{\theta})$.

a) Find the mean vector $\boldsymbol{\mu}$ and covariance matrix Σ of $\mathbf{Y}$.

b) Write the usual chi-squared statistic

$$V_n = \sum_{i=1}^{k} \frac{(n_j - n\theta_j)^2}{n\theta_j}$$

as

$$V_n = n(\bar{\mathbf{X}}_n - \boldsymbol{\theta})' P^{-1} (\bar{\mathbf{X}}_n - \boldsymbol{\theta})$$

for some diagonal matrix P.

c) Find the asymptotic distribution of $\sqrt{n}(\bar{\mathbf{X}}_n - \boldsymbol{\theta})$.

d) Show that as $n \to \infty$, $V_n \to_d \chi_{k-1}$, a chi squared random variable on $k-1$ degrees of freedom. Hint: Write the asymptotic distribution in terms of a vector with covariance matrix Γ which satisfies $\Gamma' = \Gamma$ and $\Gamma^2 = \Gamma$.

2. Suppose data $X_1, X_2, \cdots, X_n$ are independent identically distributed normal random variables with mean μ and variance σ_0^2. Suppose that μ is random with (prior) normal distribution $\mathcal{N}(\mu_0, \sigma_0^2)$. What is the conditional distribution (posterior) of μ given the data? Give the mean and variance of the posterior distribution of μ in terms of $\overline{X}, \mu_0, \sigma_0^2$.

Math 541b, Fall 2003

1. Consider a test with critical region of the form $\{T \geq c\}$ for testing $H : \theta = 0$ versus $K : \theta > 0$. Suppose that T has a continuous distribution F_θ. Define the p-value as

$$U = 1 - F_0(T).$$

a) Show that if the test has level α, the power is

$$\beta(\theta) = P\{U \leq \alpha\} = 1 - F_\theta(F_0^{-1}(1 - \alpha)),$$

where $F_0^{-1}(\mu) = \inf\{t : F_0(t) \geq u\}$.

b) Define the expected p-value as $EPV(\theta) = E_\theta U$. Let T_0 denote a random variable with distribution F_0, which is independent of T. Show that $EPV(\theta) = P(T_0 \geq T)$.

c) Suppose that for each $\alpha \in (0, 1)$, the uniformly most powerful test is of the form $I(T \geq c)$. Let $EPV_T(\theta)$ be the expected p-value of $I(T \geq c)$ and $EPV_{T^*}(\theta)$ be the expected p-value for another test T^*. Show that for any $\theta > 0$, $EPV_T(\theta) \leq EPV_{T^*}(\theta)$.

d) Consider the problem of testing $H_0 : \mu = 0$ versus $H_1 : \mu > 0$ on the basis of $N(\mu, 1)$ sample $X_1, X_2, \cdots, X_n$. Let $T = \overline{X}$. Show that $EPV(\theta) = \Phi(-\sqrt{n}\mu/\sqrt{2})$, where Φ denotes the standard normal distribution.

2. Let k and densities $f_1, \ldots, f_k$ be known, and consider an i.i.d. sample from the mixture distribution

$$f(x; \theta) = \sum_{j=1}^{k} \theta_j f_j(x)$$

where

$$\Theta = \{\theta \in \mathbf{R}^k : \theta_j \geq 0, \sum_{j=1}^{k} \theta_j = 1\}.$$

a) Write down the equations for which the maximum likelihood estimate of θ is the solution.

b) Describe the EM procedure for finding the MLE.

c) Calculate the information and determine the asymptotic distribution of the MLE for the (single) parameter θ_1 when $k = 2$ and the densities f_1 and f_2 are variance 1 normals with unequal means.

Spring 2004 Math 541b Exam

1. Ratio Estimation.

a) (Midzuno's Procedure) Let $0 < n < N$ and $(x_1, y_1), \ldots, (x_N, y_N)$ be a fixed set of pairs of numbers with $x_i > 0$, and let

$$\theta = \frac{\bar{y}_N}{\bar{x}_N} = \frac{\sum_{i=1}^N y_i}{\sum_{i=1}^N x_i}.$$

Let I be a random index with distribution

$$P(I = i) = \frac{x_i}{\sum_{i=1}^N x_i}, i = 1, \ldots, N,$$

and let a sample S of size n consist of (x_I, y_I) and a simple random sample of $n - 1$ of the remaining pairs. Let

$$T = \frac{\sum_{j \in S} y_j}{\sum_{j \in S} x_j}.$$

Find ET. Hint: For a simple random sample of size n, let I_i be the indicator that pair i is included and $\bar{x}_S$ the average of the x values in that sample. With I_i^* the indicator that pair i is included using Midzuno's scheme, show

$$E\left(\bar{x}_s f(I_1, \ldots, I_N)\right) = \bar{x}_N E f(I_1^*, \ldots, I_N^*).$$

b) Let $X_i \sim \mathcal{N}(\mu_X, 1), Y_i \sim \mathcal{N}(\mu_Y, 1), i = 1, \ldots, n$ be independent normal variables. Find a confidence interval for the ratio of means

$$\theta = \frac{\mu_Y}{\mu_X}$$

Hint: First consider

$$U = \bar{Y} - \theta \bar{X}.$$

2. An individual has two coins; one is unbiased and the other one is biased with head (H) probability p. The person chooses the first coin with probability $1-\alpha$ and the second coin with probability α. Both p and α are unknown parameters. He then tosses the chosen coin three times. Let N be the number of times "H" appears.

a). What is the distribution of N?

b). The person does n such experiments, where in each experiment, he chooses a coin and tosses it three times. Let n_i, $i = 0, 1, 2, 3$ be the number of experiments in which i heads appear, $n_0 + n_1 + n_2 + n_3 = n$. What is the likelihood function of the observed data? What is the set of equations for the maximum likelihood estimates of α and p?

c). Design an EM algorithm for estimating α and p.

d). How would you, in principle, use Wald's statistic to construct a $1-\beta$ confidence region for (α, p)? (Recall that for testing the hypothesis $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$ Wald's test statistic

$$W_n(\theta_0) = n(\hat{\theta}_n - \theta_0)^t I(\theta_0)(\hat{\theta}_n - \theta_0),$$

has an approximate χ^2-distribution under the null hypothesis, where $\hat{\theta}_n$ is the maximum likelihood estimate of θ and $I(\theta)$ is the information matrix.)

Math 541b Qualifying Exam (One-hour)

1. Let $X = (X_1, \cdots, X_n)$ be a sample from the uniform distribution on $(0, \theta)$. Show that

   (a) For testing H: $\theta \le \theta_0$ against K: $\theta > \theta_0$, any test is UMP at level α for which $E_{\theta_0} \phi(X) = \alpha$, $E_\theta \phi(X) \le \alpha$ for $\theta \le \theta_0$, and $\phi(x) = 1$ when $x_{(n)} > \theta_0$, where we denote the order statistics by $X_{(1)} \le X_{(2)} \le \cdots, \le X_{(n)}$.

   (b) For testing H: $\theta \le \theta_0$ against K: $\theta \ne \theta_0$, a UMP test exists, and is given by $\phi(x) = 1$ when $x_{(n)} > \theta_0$ or $x_{(n)} \le \theta_0 \alpha^{1/n}$, and by $\phi(x) = 0$ otherwise.

2. Suppose that $Y = (Y_1, Y_2)$, where Y_1 takes values from $\{1, 2\}$, and Y_2 takes values from $\{1, 2, 3\}$. We assume that $a_{ij} = \Pr(Y_1 = i, Y_2 = j) > 0$ for all (i, j). We want to use Gibbs sampler to obtain the joint distribution of (Y_1, Y_2).

   (a) Consider the following systematic version of Gibbs sampling. In each round, we first update the value of Y_1 and then the value of Y_2. Please write down the transition probability matrix for each update.

   (b) Consider the following random-scan version of Gibbs sampling. In each step, we flip a coin with chance λ of obtaining a head, where $0 < \lambda < 1$. If it is a head, we update the value of Y_1. Otherwise, we update the value of Y_2. Please show that the associated Markov chain is in detailed balance. Show this scheme indeed converges to the joint distribution of (Y_1, Y_2).

Math 541b Qualifying Exam, Spring 2005 (One-hour)

1. Let $X_1, \cdots, X_n$ be independent identically distributed samples from the uniform distribution $(\theta, \theta+1)$, $\theta \in R$. Suppose that $n \geq 2$. Let $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ be the order statistics from the smallest to the largest.

   (a) Show that the uniformly most powerful (UMP) test for testing $H_0 : \theta \leq 0$ versus $H_1 : \theta > 0$ is of the form

   $$T_*(X_{(1)}, X_{(n)}) = \begin{cases} 0 & X_{(1)} < 1 - \alpha^{1/n}, X_{(n)} < 1 \\ 1 & \text{otherwise} \end{cases}$$

   (b) Find a level $100(1-\alpha)\%$ confidence interval for θ.

2. Suppose that the length of life X of a light bulb manufactured by a certain process has an exponential distribution with unknown mean $1/\theta$, that is, the probability density function for $X|\theta$ is

   $$f(x|\theta) = \theta e^{-\theta x}.$$

   Let $X_1, X_2, \cdots, X_n$ be a random sample from the population.

   (a) Prove that the gamma prior distribution for θ with density function

   $$g(\theta|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha}\theta^{\alpha-1}e^{-x/\beta}$$

   is a conjugate prior.

   (b) Find the Bayesian estimate of θ corresponding to the quadratic loss function.

3. Let $(I_i, Y_i), 1 \leq i \leq n$ be independent identically distributed according to P_θ, $\theta = (\lambda, \mu) \in (0,1) \times R$ where

   $$P_\theta[I_1 = 1] = \lambda = 1 - P_\theta[I_1 = 0],$$

   and given $I_1 = j$, $Y_1 \sim N(\mu, \sigma_j^2), j = 0, 1$ and $\sigma_0 \neq \sigma_1$ known.

   (a) Find the maximum likelihood estimate of $\theta = (\lambda, \mu)$, when they exist.

   (b) Suppose that $I_i, i = 1, 2, \cdots, n$ are not observed. Give as explicitly as possible the E-Step and the M-step of the EM algorithm for this problem.

541b Qualifying Exam

Fall, 2005

Name:

1	
2	
total	

1. Suppose $X_1, X_2, \cdots, X_n$ are independent observations from the location model with density $f(x-\theta)$, $-\infty < \theta < \infty$, where f is differentiable and the Fisher information for θ is finite.

a) Show that the Fisher information $I(f)$ for θ is constant, and compute $I(f)$.

We consider the test of level α for $H_n : \theta = \theta_0$ versus $K_n : \theta = \theta_0 + h/\sqrt{n}$, where $h > 0$. Under the null $P^n_{\theta_0}$, the following expansion is valid:

$$\log \frac{f(X_1, X_2, \cdots, X_n; \theta = \theta_0 + h/\sqrt{n})}{f(X_1, X_2, \cdots, X_n; \theta = \theta_0)} = \frac{h}{\sqrt{n}} \sum_{i=1}^{n} \frac{-f'(X_i - \theta_0)}{f(X_i - \theta_0)} - \frac{1}{2} h^2 I(f) + o_{P^n_{\theta_0}}(1).$$

b) Show that the log-likelihood-ratio tends to $N(-\frac{1}{2}h^2 I(f), h^2 I(f))$ in distribution.

c) Show that the rejection region of the asymptotically most powerful test of level α is of the form $\sum_{i=1}^{n} \frac{-f'(X_i - \theta_0)}{f(X_i - \theta_0)} > c_n(\alpha)$, for some $c_n(\alpha)$. Find $c_n(\alpha)$.

d) When f is double exponential, namely,

$$f(x-\theta) = \frac{1}{2} \exp\{-|x-\theta|\}.$$

find the asymptotically most powerful test of level α

2. Consider an aperiodic and irreducible Markov Chain on a finite state space S with transition matrix $P = (p_{ij})_{i,j \in S}$.

a) Show that if the probabilities $\pi_i, i \in S$, satisfy the detail balance equation

$$\pi_i p_{ij} = \pi_j p_{ji} \quad i, j \in S,$$

then they give the unique stationary distribution of the chain.

b) Let q_{ij} be a 'proposal' transition rule on S. Given $\pi_i, i \in S$, show how to construct transitions probabilities p_{ij}, depending on q and the quantity

$$r_{ij} = \min\{1, \frac{\pi_j q_{ji}}{\pi_i q_{ij}}\},$$

which satisfy the detail balance equation. What conditions, if any, should the proposal q satisfy in order that π be the unique stationary distribution?

c) Let S_n be the collection of rooted binary trees on n vertices, where each vertex has either 0 or 2 descendants. Construct, in general, a Markov Chain on S that has the uniform stationary distribution, and calculate one transition probability for a simple small example. If it adds clarity, you may illustrate your proposal distribution and subsequent calculation with figures.

Spring 2006 Math 541b Exam

1. Let $X_1, \cdots, X_n$ be i.i.d. samples from a Weibull distribution with density $f(x, \lambda) = \lambda\, cx^{c-1}e^{-\lambda x^c}$, where $x > 0$, and c is a known positive constant and $\lambda > 0$ is the scale parameter of interest. Let $\mu = 1/\lambda$.

   (a) Show that $\sum_{i=1}^n X_i^c$ is an optimal test statistic or testing H: $\mu = \mu_0$ versus K: $\mu = \mu_1 > \mu_0$. That is, the most powerful test takes the form:
   $$\begin{cases} \text{reject } H & \text{if } \sum_{i=1}^n X_i^c > \text{critical value} \\ \text{accept } H & \text{if } \sum_{i=1}^n X_i^c \le \text{critical value}\,. \end{cases}$$

   (b) Show that λX_i^c follows the standard exponential distribution Exp(1).

   (c) Find the critical value for the size α most powerful test.

   (d) Show that the power of the most powerful test of size α is given by
   $$\beta(\mu_1) = 1 - G_n(\frac{\mu_0}{\mu_1} g_n(1-\alpha))\,.$$
   where G_n is the distribution function of $\Gamma(n, 1)$, $g_n(1-\alpha)$ is the $(1-\alpha)$th quantile of $\Gamma(n, 1)$, and prove that $\beta(\mu)$ is increasing in μ.

   (e) Show that the most powerful test of size α for the simple hypotheses in (a) is uniformly most powerful, at size α, for testing the composite hypotheses H: $\mu \le \mu_0$ versus K: $\mu > \mu_0$.

   (f) When n is large, please use normal approximation to find the critical value and power.

2. Let $X_i, B_i, i = 1, \ldots, n$ be independent Bernoulli variables where X_i has unknown success probability $p \in (0, 1)$, and B_i has success probability 1/3. Suppose we observes
   $$Y_i = B_i X_i + (1 - B_i)(1 - X_i), \quad i = 1, \ldots, n$$
   that is, we see the original X_i with probability 1/3, and $1 - X_i$ with probability 2/3.

   (a) Write the log likelihood in term so the sum $S_n = \sum_{i=1}^n Y_i$, and the equation one would solve for finding the maximum likelihood estimator.

(b) Introduce appropriate missing data for the implementation of the EM algorithm and write out the full likelihood, and the maximum likelihood estimator using this data.

(c) Detail the steps of the EM algorithm.

Spring 2007 Math 541b Exam

1. Let $\mathbf{p} = (p_1, \ldots, p_c)$ be a vector of positive numbers summing to one, and $\mathbf{X} \sim \mathcal{M}(n, \mathbf{p})$, the multinomial distribution given by

$$P(\mathbf{X} = \mathbf{k}) = \binom{n}{\mathbf{k}} \mathbf{p}^{\mathbf{k}},$$

where $\mathbf{k} = (k_1, \ldots, k_c)$ are non-negative integers summing to n,

$$\binom{n}{\mathbf{k}} = \frac{n!}{\prod_{i=1}^{n} k_i!} \quad \text{and} \quad \mathbf{p}^{\mathbf{k}} = \prod_{i=1}^{n} p_i^{k_i}.$$

For a given probability vector $\mathbf{p}_0$ we test $H_0 : p = p_0$ versus $H_1 : p \neq p_0$ using the chi-squared test statistic

$$V^2 = \sum_{i=1}^{c} \frac{(X_i - np_{i,0})^2}{np_{i,0}}.$$

   (a) Calculate the mean vector and the covariance matrix of $\mathbf{X}$.

   (b) Define a matrix $\mathbf{P}$ such that

$$V^2 = n^{-1}(\mathbf{X} - n\mathbf{p})'\mathbf{P}^{-1}(\mathbf{X} - n\mathbf{p}).$$

   (c) Show that

$$n^{-1/2}(\mathbf{X} - n\mathbf{p}) \rightarrow_p Y \sim \mathcal{N}_c(0, \Sigma)$$

   (d) Find the distribution of $U = P^{-1/2}Y$, and show that the covariance matrix of U is a projection. (Recall that Q is a projection matrix if $Q' = Q^2 = Q$.) Hint: show

$$P^{-1/2}\Sigma P^{-1/2} = I - P^{-1/2}\mathbf{p}\mathbf{p}'P^{-1/2}.$$

   (e) Show that

$$V^2 \rightarrow_d \chi^2_{c-1},$$

   that is, that V^2 converges in distribution to a chi squared distribution with $c - 1$ degrees of freedom.

2. Suppose $X_1, \cdots, X_n$ are independently and identically distributed with variance σ^2.

(a) Show that the estimate of variance $\hat{\theta} = \sum_{i=1}^{n}(x_i - \bar{x})^2/n$ has bias equal to $-\sigma^2/n$ as an estimator of σ^2.

(b) Show that the bias of the jackknife estimate is $-s^2/n$, where $s^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2/n$.

Fall 2007 Math 541b Exam

1. Let $p_0(\mathbf{x})$ and $p_1(\mathbf{x})$ be two distinct density functions on $\mathbf{R}^d$.

   (a) Based on $\mathbf{X}_1, \ldots, \mathbf{X}_n$ independent and identically distributed with density p, it is desired to test $H_0 : p = p_0$ versus $H_1 : p = p_1$. Prove the Neymann Pearson Lemma, that the test which rejects H_0 when the likelihood ratio

   $$L_n = \prod_{i=1}^{n} \frac{p_1(\mathbf{X}_i)}{p_0(\mathbf{X}_i}$$

   exceeds a threshold has maximum power over all tests having the same Type I error.

   (b) Let $E_i, i \in \{0, 1\}$ be the expectation when the density of $\mathbf{X}$ is $p_i(\mathbf{x})$. Prove that

   $$K(1,0) < 0 \quad \text{where} \quad K(1,0) = E_0 \log \frac{p_1(\mathbf{X})}{p_0(\mathbf{X})}.$$

   (c) Let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be an infinite sequence of observations independent and identically distributed with density p, which are observed one at a time. Let $a < 0 < b$ and consider the test which, after n observations, rejects H_0 if $\log L_n > b$, accepts H_0 if $\log L_n < a$, and takes an additional observation otherwise. By considering $n^{-1} \log L_n$, show that this test always terminates, whether H_0 or H_1 is true.

2. Let $X_1, \ldots, X_n$ be i.i.d. with mean μ and variance σ^2, and g a function with continuous derivative.

   Show that the jackknife estimate of variance of

   $$\hat{\theta} = g(\overline{X})$$

   is asymptotically equivalent to what is produced using the delta method.

Spring 2008 Math 541b Exam

1. Suppose that $X_1, \cdots, X_n$ are i.i.d. samples from the uniform distribution on $(0, \theta)$.

   (a) Show that the MLE of θ is
   $$\hat{\theta} = \max(X_1, \cdots, X_n)$$

   (b) Show that $n(\theta - \hat{\theta})$ converges in distribution to an exponential. Please specify the parameter.

   (c) Let $\lambda_n = \sup_\theta(L_n(\theta)/L_n(\theta_0))$, where $L_n(\theta)$ is the likelihood of $X_1, \cdots, X_n$. Show that
   $$2\log\lambda_n = \begin{cases} 2n\log\frac{\theta_0}{\hat{\theta}} & \hat{\theta} \le \theta_0 \\ \infty & \hat{\theta} > \theta_0 \end{cases}$$

   (d) Show that $2\log\lambda_n \longrightarrow \chi_2^2$ in distribution. Please notice that the degree of freedom is 2.

   (e) What is the asymptotic distribution of the likelihood ratio test under the general regularity conditions? Is the result in the last part consistent with the general result?

2. Consider the following formulation of the EM algorithm for the estimation of the parameter $\psi \in \Omega$ upon observing incomplete data y which is obtained through a (many to one) function as $y = y(x)$. The incomplete data is distributed according to $g(y;\psi)$; the full data x according to $g_c(x;\psi)$. These two densities are are related through the mapping $y = y(x)$ by
   $$g(y;\psi) = \int_{x:y=y(x)} g_c(x,\psi)dx.$$

   Begin with any initial value $\psi^{(0)} \in \Omega$, then iterate the following E and M steps.

   E step. Calculate
   $$Q(\psi, \psi^{(k)}) = E_{\psi^{(k)}}\left(\log L_c(\psi)|y\right).$$

   M step. Let $\psi^{(k+1)}$ be any value in Ω that maximizes $Q(\psi, \psi^{(k)})$, that is
   $$Q(\psi^{(k+1)}, \psi^{(k)}) \ge Q(\psi, \psi^{(k)}) \quad \text{for all } \psi \in \Omega.$$

(a) Argue that the conditional density of x given y when $y = y(x)$ is

$$k(x|y, \psi) = \frac{g_c(x; \psi)}{g(y; \psi)}, \tag{1}$$

and zero otherwise. Letting

$$H(\psi, \psi^{(k)}) = E_{\psi^{(k)}}(\log k(x|y, \psi)|y),$$

prove that

$$H(\psi, \psi^{(k)}) \leq H(\psi^{(k)}, \psi^{(k)}) \quad \text{for all } \psi \in \Omega.$$

(b) Use (1) to write the log likelihood of the incomplete data $\log L(\psi) = \log g(y; \psi)$ as a difference involving the functions Q and H.

(c) Prove that the log likelihood sequence $\log L(\psi^{(k)})$ is monotone nondecreasing (Hint: consider differences).

Fall 2008 Math 541b Exam

1. Let $X_1, \ldots, X_n$ be i.i.d. from a normal distribution with unknown mean μ and known variance 1. Suppose that negative values of X_i are truncated at 0, so that instead of X_i we actually observe

$$Y_i = \max\{0, X_i\}, \quad i = 1, \ldots, n,$$

from which we would like to estimate μ.

   (a) Explain how to use the EM algorithm to estimate μ from $Y_1, \ldots, Y_n$. Specifically, give the complete log-likelihood function $\log L_c(\mu)$ (i.e., the log of the joint density of $X_1, \ldots, X_n$) and a recursive formula for the successive EM estimates $\mu^{(k+1)}$. Write these in terms of the density ϕ and c.d.f. Φ of the standard normal distribution. *Hint:* To simplify things, assume that $X_1, \ldots, X_m$ are not truncated, and $X_{m+1}, \ldots, X_n$ are.

   (b) Find the partial log-likelihood function $\log L(\mu)$ (i.e., the log of the joint density of $Y_1, \ldots, Y_n$) and use it to write down a (nonlinear) equation which the MLE $\widehat{\mu}$ satisfies. Use this equation to manually verify that $\widehat{\mu}$ is indeed a fixed point of the recursion found in (a).

2. Let f denote the true density function of X, and consider testing the simple hypotheses

$$H_0 : f = f_0 \quad \text{vs.} \quad H_1 : f = f_1$$

for given densities f_0, f_1. For a fixed value $\pi \in (0, 1)$, suppose that the probabilities $\pi_0 = \pi$ and $\pi_1 = 1 - \pi$ can be assigned to H_0 and H_1 prior to the experiment. We will describe tests of H_0 vs. H_1 by their indicator functions

$$\psi(X) = \begin{cases} 1, & \text{the test rejects } H_0 \\ 0, & \text{the test accepts } H_0. \end{cases}$$

   (a) Show that the overall probability of an error resulting from using a test ψ is

$$\pi E_0 \psi(X) + (1 - \pi) E_1 [1 - \psi(X)]. \tag{1}$$

   (b) Call the test ψ^* minimizing (1) the *Bayes optimal* test. By writing (1) as a single E_0 expectation using the "change of measure" technique

$$E_1(\cdot) = E_0 \left[(\cdot) \frac{f_1(X)}{f_0(X)} \right],$$

   show that the Bayes optimal test is equivalent to a simple likelihood ratio test. Also, give the value of the likelihood ratio test's critical value.

   (c) Argue that the Bayes optimal test is hence most powerful for detecting f_1 at a certain significance level. Write down an expression for this significance level, and also give an upper bound for it as a function of π.

   (d) The *posterior probability* of H_i is the conditional probability that H_i is true, given $X = x$. Show that the posterior probability of H_i is

$$\frac{\pi_i f_i(x)}{\pi_0 f_0(x) + \pi_1 f_1(x)}. \tag{2}$$

   Show that the Bayes optimal test is also equivalent to choosing which hypothesis has the larger posterior probability.

Spring 2009 Math 541b Exam

1. Let $X_1, \ldots, X_{n_1}$ be i.i.d. $N(\mu_1, \sigma_1^2)$ and $Y_1, \ldots, Y_{n_2}$ i.i.d. $N(\mu_2, \sigma_2^2)$, where $\mu_1, \mu_2, \sigma_1, \sigma_2$ are all unknown and σ_1, σ_2 are not necessarily assumed to be equal. This problem concerns the generalized likelihood ratio (GLR) test of

$$H_0 : \mu_1 = \mu_2 \quad \text{vs.} \quad H_1 : \mu_1 \neq \mu_2.$$

   (a) Letting $L(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$ be the log of the likelihood function of $(X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2})$, find the (unrestricted) maximum likelihood estimates (MLEs) $\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}_1^2, \widehat{\sigma}_2^2$ and write down

$$L(\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}_1^2, \widehat{\sigma}_2^2)$$

   in as simple form as possible.

   (b) For the H_0-restricted MLEs $\widehat{\widehat{\mu}}, \widehat{\widehat{\sigma}}_1^2, \widehat{\widehat{\sigma}}_2^2$,

      i. Write down formulas for $\widehat{\widehat{\sigma}}_i^2$ as functions of $\widehat{\sigma}_i^2$ and $\widehat{\widehat{\mu}}$,

      ii. Find a cubic equation, not depending explicitly on $\widehat{\widehat{\sigma}}_i$, that $\widehat{\widehat{\mu}}$ satisfies. You do not need to solve this equation.

   (c) Give the asymptotic distribution of the GLR statistic under H_0, as $n_1, n_2 \to \infty$.

2. Let $\widehat{\theta}_n$ be a parameter estimate computed from the random sample $X_1, \ldots, X_n$, let $\widehat{\theta}_{(i)}$ be the estimate computed from

$$X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n,$$

   and let

$$\widehat{\theta}_{(\cdot)} = \frac{1}{n} \sum_{i=1}^{n} \widehat{\theta}_{(i)}.$$

   Recall that the Jackknife estimate of the bias of $\widehat{\theta}$ is given by

$$b_{JACK} = (n-1)(\widehat{\theta}_{(\cdot)} - \widehat{\theta}_n).$$

   Now let $X_{(1)}, \ldots, X_{(n)}$ be the order statistics of the sample, which we assume are from a continuous distribution. The sample median $\widehat{m}_n$ is

$$\widehat{m}_n = \begin{cases} X_{((n+1)/2)}, & n \text{ odd} \\ (X_{(n/2)} + X_{((n/2)+1)})/2, & n \text{ even.} \end{cases}$$

   (a) Compute the jackknife estimate of bias b_{JACK} for the sample median $\widehat{m}_n$ in both cases.

   (b) Let b denote the true bias of $\widehat{m}_n$. Is b_{JACK} always unbiased for b in the even case?

Fall 2009 Math 541b Exam

1. With $\mathbf{X}_1, \ldots, \mathbf{X}_n$ i.i.d. with density $p(\mathbf{x}; \theta), \theta \in \mathbb{R}^d$, consider the usual likelihood ratio test for $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1$ based on

$$\Lambda_n = \frac{\sup_{\theta \in \Theta_0} \prod_{i=1}^n p(\mathbf{x}_i; \theta)}{\sup_{\theta \in \Theta_0 \cup \Theta_1} \prod_{i=1}^n p(\mathbf{x}_i; \theta)}.$$

Let

$$\lambda_n = -2 \log \Lambda_n.$$

   (a) Under sufficient regularity there exists Y such that,

$$\lambda_n \rightarrow_d Y \quad \text{as } n \rightarrow \infty.$$

   State the distribution of Y.

   (b) In parts b) and c) assume $\theta = (\theta_1, \theta_2)$, and let $p(\mathbf{x}; \theta)$ be the density of the bivariate normal $\mathbf{X}$ whose components have unknown mean (θ_1, θ_2), known variances σ_1^2, σ_2^2 and known correlation coefficient $\rho \in (-1, 1)$. For testing with $\Theta_0 = \{\theta_1 = 0, \theta_2 = 0\}$ and $\Theta_1 \cup \Theta_2 = \mathbb{R}^2$, verify that the distribution of λ is as specified in part a).

   (c) For testing with $\Theta_0 = \{\theta_1 = 0, \theta_2 = 0\}$ and $\Theta_1 \cup \Theta_2 = \{\theta_1 > 0, \theta_2 \in \mathbb{R}\}$, determine the distribution of λ. Compare this distribution to the one in part b), and explain.

   (d) State whether we can generalize the distributional result in c) asymptotically under the regularity assumed for the convergence in a), and justify your answer.

2. Let $X_1, \ldots, X_n$ be i.i.d. from the exponential distribution $\mathcal{E}(a, b)$ with density

$$b^{-1} \exp\{-(x - a)/b\} \cdot \mathbf{1}\{x \geq a\}, \quad -\infty < a < \infty, \quad b > 0,$$

and let $X_{(1)} = \min\{X_1, \ldots, X_n\}$.

   (a) Determine the uniformly most powerful (UMP) test for testing $H_0 : a = a_0$ vs. $H_1 : a \neq a_0$ when b is assumed known.

(b) Show that the most powerful level-α test of H_0 vs. $H_1' : a = a_1$, for some given $a_1 < a_0$, has power equal to

$$\beta(a_1) = 1 - (1 - \alpha)e^{-n(a_0 - a_1)/b}. \tag{1}$$

(c) Show that, for the problem in part (a) but with b unknown, any level-α test which rejects when

$$\frac{X_{(1)} - a_0}{\sum(X_i - X_{(1)})} \notin (C_1, C_2) \tag{2}$$

is most powerful at all alternatives (a_1, b) with $a_1 < a_0$ (independent of the particular choice of C_1, C_2).

Spring 2010 Math 541b Exam

1. Consider the multinomial model with cell probabilities $p = (p_1, p_2, \cdots, p_m) \in \Omega$, in which $\sum_{i=1}^m p_i = 1$, $p_i > 0$, for $i = 1, \cdots, m$. And the count data generated from a multinominal is denoted by $x = (x_1, x_2, \cdots, x_m)$, where $\sum_{i=1}^m x_i = n$. Under the null hypothesis H_0, the vector of cell probabilities p is specified by $p = p(\theta) \in \omega_0$, where $\theta \in \Theta$. One example is the parameter specification in the Pearson's chi-square test of independence. In the alternative model H_A, the parameter vector $p \in \Omega - \omega_0$.

   (a) Denote $\hat{p}_i = x_i/n$, and the maximum likelihood estimate of θ under H_0 by $\hat{\theta}$. Show that the log-likelihood ratio test statistic is

   $$-2\log\Lambda = 2\sum_{i=1}^{m} O_i \log(\frac{O_i}{E_i}),$$

   where $O_i = n\hat{p}_i$ and $E_i = np(\hat{\theta})$.

   (b) Assume that $\dim\omega_0 = k$, what is the asymptotic distribution of the likelihood ratio test statistic as $n \longrightarrow +\infty$.

   (c) What is the Pearson test statistic in this general scenario?

   (d) Show that the likelihood ratio test statistic and the Pearson test statistic are approximately equivalent. (Hint: You may use a Taylor expansion.)

2. Consider the Bernoulli-Laplace model in which there are two urns, each containing M balls, and of these $2M$ total balls M are black and M are white. Suppose at each time step, one ball is chosen from each urn at random and they are interchanged. Let $X_n \in \{0, \ldots, M\}$ be the number of black balls in the first urn just after the nth step.

   (a) Find the transition probabilities of the Markov Chain X_n.

   (b) Find the stationary distribution of X_n. Prove stationarity.

Fall 2010 Math 541b Exam

1. Let $p(x)$ and $q(x)$ be two distinct density functions that are positive on $\mathbb{R}$.

   (a) We define the Kullack Leibler divergence as

   $$D(p||q) = E_p \log \frac{p(x)}{q(x)},$$

   where E_p means taking the expectation under density $p(x)$. Prove that $D(p||q)$ is strictly positive.

   (b) Let $X_1, \ldots, X_n$ be i.i.d. with density

   $$p(x;\theta) = \begin{cases} p(x) & \theta = 0 \\ q(x) & \theta = 1. \end{cases}$$

   For some fixed $\delta > 0$, consider the test $H_0 : \theta = 0$ versus $H_1 : \theta = 1$ which has rejection region A_n given by

   $$A_n = \{(X_1, \ldots, X_n) : e^{n(D(p||q)-\delta)} \leq \prod_{i=1}^{n} \frac{p(X_i)}{q(X_i)} \leq e^{n(D(p||q)+\delta)}\}^c.$$

   Prove that the sequence of Type I errors $P(A_n)$ tend to zero as $n \to \infty$, where $P(\cdot)$ is the probability under $p(x)$.

   (c) With $Q(\cdot)$ the probability under $q(x)$, prove that the Type II errors $Q(A_n^c)$ satisfy

   $$\lim_{\delta \to 0} \lim_{n \to \infty} \frac{1}{n} \log Q(A_n^c) = -D(p||q).$$

2. Suppose that X follows a Poisson distribution $\mathcal{P}(\theta)$, $\theta > 0$. Let θ have a Gamma $\Gamma(p, \lambda)$ distribution.

   (a) Show that the posterior distribution is $\Gamma(p + x, 1 + \lambda)$

   (b) Show that if we take the loss function as $l(\theta, a) = (a - \theta)^2$, then the Bayes estimate of θ is $(p + x)/(1 + \lambda)$.

   (c) Find the Bayes estimate if the loss function is $l(\theta, a) = (a - \theta)^2/\theta$

Fall 2010 Math 541b Exam

1. Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be independent pairs of independent random variables X_i and Y_i, where X and Y have continuous distribution functions $F(x)$ and $G(x)$, respectively. We want to test the hypothesis $H_0 : F = G$ versus the alternative $H_1 : F \neq G$. Let $Z_i = X_i - Y_i$ and rank the numbers $|Z_1|, |Z_2|, \cdots, |Z_n|$ in increasing order, and let r_i be the rank of $|Z_i|$, so that the smallest absolute value receives the rank of 1.

   Define the Wilcoxon signed-rank statistic W as

   $$W = \sum_{i=1}^{n} \text{sign}(Z_i) r_i,$$

   where $\text{sign}(z) = 1$ if $z > 0$, $\text{sign}(z) = -1$ if $z < 0$, and $\text{sign}(z) = 0$ if $z = 0$.

   (a) Calculate $P(\text{sign}(Z_i)) = 1)$ and $P(\text{sign}(Z_i)) = 0)$ under the null hypotheses H_0.

   (b) Calculate the mean and the variance of W under the null hypothesis H_0.

   (c) Propose a test for H_0 versus H_1 that has approximate Type I error level α when the sample size is large.

2. Let $X_1, X_2, \cdots, X_n$ be i.i.d observations drawn from a mixture of two normal densities $\mathcal{N}(\mu_1, 1)$ and $\mathcal{N}(\mu_2, 1)$, where α and $1 - \alpha$ are the probabilities that a given observation is taken from the first and second normal distribution, respectively. We suppose that (μ_1, μ_2, α) are unknown.

   (a) Write down the likelihood function of the observed data $(X_1, X_2, \cdots, X_n)$ as a function of $\theta = (\mu_1, \mu_2, \alpha)$.

   (b) In order to design an EM algorithm to estimate θ, define missing data $(Z_1, Z_2, \cdots, Z_n)$ where $Z_i = 1$ if X_i is drawn from the first population $\mathcal{N}(\mu_1, 1)$, and $Z_i = 0$ if it comes from the second population. Write the likelihood function of the complete data $((X_1, Z_1), (X_2, Z_2), \cdots, (X_n, Z_n))$.

   (c) Design an EM algorithm to estimate the parameter vector θ.

1. Let $Y_1, \ldots, Y_n$ be a random sample from the uniform density on $[0, \theta]$, where θ is an unknown parameter.

   (a) Calculate the Likelihood function $L(\theta|Y_1, \ldots, Y_n) = p(Y_1, \ldots, Y_n|\theta)$ for θ.

   (b) Let $\Omega = \{\theta : 0 < \theta \leq \theta_0\}$. Show that $\max\{L(\theta|Y_1, \ldots, Y_n) : \theta \in \Omega\} = (Y_{\max})^{-n}$, where $Y_{\max} = \max\{Y_1, \ldots, Y_n\}$.

   (c) Write out the form of the Generalized Likelihood Ratio λ for the test of $H_0 : \theta = \theta_0$ versus $H_1 : \theta < \theta_0$.

   (d) Show that the Generalized Likelihood Ratio Test calls for H_0 to be rejected at level α if $Y_{\max} \leq \theta_0 \sqrt[n]{\alpha}$.

2. Let $(I_1, Y_1), \ldots, (I_n, Y_n)$ be i.i.d. from distribution P_θ, where $\theta = (\lambda, \mu) \in (0, 1) \times \mathbb{R}$,

$$P_\theta(I_i = 1) = \lambda = 1 - P_\theta(I_i = 0),$$

   and, given $I_i = j$, $Y_i \sim N(\mu, \sigma_j^2)$, where $\sigma_0 \neq \sigma_1$ are known positive values.

   (a) Write down the complete likelihood function $L_c(\lambda, \mu)$ assuming that all of $(I_1, Y_1), \ldots, (I_n, Y_n)$ are observed.

   (b) Give explicitly the maximum likelihood estimates of λ and μ.

   (c) Now suppose that the I_i are not observed. Give as explicitly as possible the E- and M-steps of the EM algorithm, including recursive formulae for the EM iterates $\lambda^{(k)}$ and $\mu^{(k)}$.