DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
January, 2011

Statistics (M. A. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let $(x_1, \ldots, x_n)$ be a sample from a population with pdf

$$f(x; \theta) = \frac{1}{\sqrt{2\pi}\theta} \exp \frac{-(x - \theta)^2}{2\theta^2}$$

with $\theta > 0$ as a parameter.

(i) Find the minimal sufficient statistic and justify its minimality and incompleteness.

(**Hint**: Suffice to construct two unbiased estimators of θ or of θ^2 that are functions of the minimal sufficient statistic.)

(ii) Calculate the Fisher information on θ in the sample and the efficiency of $\bar{x}$ as an estimator of θ.

(**Hint**: You need to recall the values of the third and fourth moments of the standard normal random variable.)

2. Let $X_1, \ldots, X_n$ be iid random variables uniformly distributed on $(\theta, 2\theta)$ with $\theta > 0$ as a parameter.
(i) Show that the family of pdf's of θ,

$$\pi(\theta; \alpha, a, b) = K(\alpha, a, b)\theta^\alpha, \ a < \theta < b$$

with $\alpha \in (0, \infty)$, $0 < a < b < \infty$ as parameters and $K(\alpha, a, b) = (\int_a^b \theta^\alpha d\theta)^{-1}$ the normalizing factor, is a conjugate family of prior pdf's.
(ii) If the parameters of the prior are $\alpha = 1$, $a = 2$, $b = 3$ and $n = 3$, calculate the Bayes estimators of θ for the loss function $L(\tilde{\theta}, \theta) = (\tilde{\theta} - \theta)^2$.

3. Let $(x_1, \ldots, x_m), (y_1, \ldots, y_n)$ be two independent samples from normal distributions $N(\mu_1, \sigma^2)$ and $N(\mu_2, 2\sigma^2)$, respectively, with (μ_1, μ_2, σ^2) as parameters.
(i) Find the minimal sufficient statistic for (μ_1, μ_2, σ^2).
(ii) If s_1^2, s_2^2 are the sample variances of the first and second sample, justify the existence of a better (i. e., with a smaller variance) than $(s_1^2 + s_2^2)/3$ unbiased estimator of σ^2.

4. Let $(x_1, \ldots, x_n)$ be a sample from a population with density

$$f(x; \theta) = \frac{\theta e^{\theta x}}{e^\theta - 1}, \ 0 < x < 1.$$

with $\theta > 0$ as a parameter.
(i) Show that the family of distributions of the sample has a monotone likelihood ratio (MLR) property.
(ii) Construct a uniformly most powerful (UMP) test of size α of

$$H_0 : \theta \le 1 \text{ vs } H_1 : \theta > 1.$$

5. Let $(x_1, \ldots, x_m), (y_1, \ldots, y_n)$ be two independent samples from populations with pdf's
$f_1(x; \theta) = \exp\{-(x - \theta)\}$, $x > \theta$ and $f_2(y; \theta) = 2\exp\{-2(y - \theta)\}$, $y > \theta$, respectively, with θ as a parameter.
(i) Find the maximum likelihood estimator (MLE) $\hat{\theta}_{m,n}$ of θ and calculate its variance.
(ii) Prove the consistency (as $m \to \infty$, $n \to \infty$) of the MLE.

6. Let $(x_1, \ldots, x_m)$, $(y_1, \ldots, y_n)$ be two independent samples of sizes m and n from normal populations $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, respectively, with all four parameters unknown

(i) Based on the sufficient statistics, construct a pivot for σ_1/σ_2

(ii) Express the distribution of the pivot in terms of an F-distribution and construct a confidence interval of level $1 - \alpha$ for σ_1/σ_2.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
January, 2011

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let $(x_1, \ldots, x_n)$ be a sample from a population with pdf

$$f(x; \theta) = \frac{1}{\sqrt{2\pi}\theta} \exp \frac{-(x - \theta)^2}{2\theta^2}$$

with $\theta > 0$ as a parameter.
(i) Find the minimal sufficient statistic and justify its minimality and incompleteness.
(**Hint:** Suffice to construct two unbiased estimators of θ or of θ^2 that are functions of the minimal sufficient statistic.)
(ii) Calculate the Fisher information on θ in the sample and the efficiency of $\bar{x}$ as an estimator of θ.
(**Hint:** You need to recall the values of the third and fourth moments of the standard normal random variable.)

2. Let $X_1, \ldots, X_n$ be iid random variables uniformly distributed on $(\theta, 2\theta)$ with $\theta > 0$ as a parameter.
(i) Show that the family of pdf's of θ,

$$\pi(\theta; \alpha, a, b) = K(\alpha, a, b)\theta^\alpha, \ a < \theta < b$$

with $\alpha \in (0, \infty)$, $0 < a < b < \infty$ as parameters and $K(\alpha, a, b) = (\int_a^b \theta^\alpha d\theta)^{-1}$ the normalizing factor, is a conjugate family of prior pdf's.
(ii) If the parameters of the prior are $\alpha = 1$, $a = 2$, $b = 3$ and $n = 4$, calculate the Bayes estimators of θ for the loss functions $L(\tilde{\theta}, \theta) = (\tilde{\theta} - \theta)^2$ and $L(\tilde{\theta}, \theta) = |\tilde{\theta} - \theta|$.

3. Let $(x_1, \ldots, x_m), (y_1, \ldots, y_n)$ be two independent samples from normal distributions $N(\mu_1, \sigma^2)$ and $N(\mu_2, 2\sigma^2)$, respectively, with (μ_1, μ_2, σ^2) as parameters.
(i) Find the minimal sufficient statistic for (μ_1, μ_2, σ^2).
(ii) Construct the uniformly minimum variance unbiased estimator (UMVUE) of σ^2.

4. Let $(x_1, \ldots, x_n)$ be a sample from a population with density

$$f(x; \theta) = \frac{\theta e^{\theta x}}{e^\theta - 1}, \ 0 < x < 1.$$

with $\theta > 0$ as a parameter.
(i) Construct a uniformly most powerful (UMP) test of size α of

$$H_0 : \theta \leq 1 \text{ vs } H_1 : \theta > 1.$$

(ii) Assuming n large and using the CLT, find (approximately) the threshold in the rejection region of the UMP test for $\alpha = .05$ and approximate its power function via the standard normal Φ

5. Let $(x_1, \ldots, x_m), (y_1, \ldots, y_n)$ be two independent samples from populations with pdf's
$f_1(x; \theta) = \exp\{-(x - \theta)\}, \ x > \theta$ and $f_2(y; \theta) = 2\exp\{-2(y - \theta)\}, \ y > \theta$, respectively, with θ as a parameter.
(i) Find the maximum likelihood estimator $\hat{\theta}_{m,n}$ of θ and calculate its mean.
(ii) Assuming $m = cn(1 + o(1))$, $c > 0$ and $n \to \infty$ find appropriate

$\{a_n, n = 1, 2, \ldots\}$ and the nondegenerate limiting distribution of $a_n(\hat{\theta}_{m,n} - \theta)$.

6. Let $(x_1, \ldots, x_m), (y_1, \ldots, y_n)$ be two independent samples of sizes m and n from normal populations $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, respectively, with all four parameters unknown.
(i) Based on the sufficient statistics, construct a pivot for σ_1/σ_2.
(ii) Express the distribution of the pivot in terms of an F-distribution and construct a confidence interval of level $1 - \alpha$ for σ_1/σ_2.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
AUGUST, 2010

Statistics (M.A. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use

---

1. For fixed λ, the random variables $X_1, \ldots, X_n$ are independent identically distributed with density $f(x|\lambda) = \lambda e^{-\lambda x} I_{[x>0]}$ Assume that the parameter λ is a random variable with prior distribution Gamma(α, β) for $\alpha, \beta > 0$ known, and that $\alpha + n > 3$

(a) Find the posterior distribution of λ and the Bayes estimator of λ corresponding to the quadratic loss function $L(a, \lambda) = (a - \lambda)^2$, where a is the estimated value or 'action' variable.

(b) Find the Bayes estimator of λ corresponding to the loss function $\bar{L}(a, \lambda) = (a - \lambda)^2/\lambda^2$.

2. Consider a sample $X_1, \ldots, X_n$ from the population with density

$$f(x) = \frac{1 + \lambda(x - \mu)^2}{1 + \lambda} \phi(x - \mu)$$

where $\phi(\cdot)$ is the standard normal density and $(\mu, \lambda) \in \mathbf{R} \times [0, \infty)$ is unknown. *Take as given that for a standard normal r.v. Z, $E(Z^2) = 1$, $E(Z^4) = 3$, and $E(Z^6) = 15$.*

(a) Find the method of moments estimator of (μ, λ), and show that it is consistent

(b) For the particular case $\mu = 0$, $\lambda = 1$, find the limiting nondegenerate (joint) distribution, as $n \to \infty$, of a suitably centered and scaled version of your estimator in part (a).

1

3. A dataset consists of independent triples (W_i, Y_i, Z_i) of independent random variables with distributions as follows,

$$W_i \sim \mathcal{N}(\lambda + \tau, 1), \quad Y_i \sim \text{Expon}(\lambda + \mu), \quad Z_i \sim \text{Poisson}(\mu + \tau), \quad i = 1, \ldots, n$$

(Here $\mathcal{N}$ denotes normal distribution parameterized by mean and variance, and $Expon(\theta)$ the exponential distribution with mean $1/\theta$.)

(a) With (λ, μ, τ) as unknown parameter, is this an exponential family ? Prove or disprove.

(b) Is the parameter (λ, μ, τ) identifiable based on this dataset and statistical model? Prove or disprove.

4. Suppose that you observe two independent samples of data, $X_1, \ldots, X_n$ with $\text{Expon}(\lambda)$ density and $Y_1, \ldots, Y_n$ with $\mathcal{N}(0, \sigma^2)$ density.

(a) Find a pivotal statistic for $\rho = \lambda/\sigma^2$ (as a function of the sufficient statistics for λ and σ^2), and find the form of a 95% confidence interval for ρ based on this pivotal quantity and indicate how you would find its endpoints if $n = m = 10$. What if $n = m = 100$?

(b) Find the form of the likelihood ratio test for $H_0 : \lambda/\sigma^2 = \rho_1$ (against the general alternative) for a fixed value ρ_1, and indicate how you could use such a test (with large-sample approximate cutoff) to define an approximate two-sided 95% confidence interval for ρ.

5. You observe a sample of 100 independent observations X_i from a population with the density

$$g(x) = C \sqrt{\lambda} \, \exp(-\lambda x^2 - \lambda^2 x^4), \quad -\infty < x < \infty$$

where $C = 0.73077$ is the normalizing constant, and $\lambda > 0$ is an unknown parameter.

(a) Find the minimal sufficient statistic for λ, and show that it is incomplete.

(b) Find an expression for the Maximum Likelihood Estimator $\hat{\lambda}$ of λ.

6. You observe a sample $X_1, \ldots, X_{20}$ with the density

$$f(x, \vartheta) = 2 (x/\vartheta^2) \, I_{[0 \leq x < \vartheta]}$$

with an unknown parameter $\vartheta > 0$, yielding

$$\min(X_i, i = 1, \ldots, 80) = 0.035, \quad \max(X_i, 1 \leq i \leq 80) = 3.66$$

Based on these data, give numerical expressions for

(a) the p-value of the most powerful hypothesis test for $H_0 : \vartheta = 4.0$ versus $H_A : \vartheta = 3.66$;

(b) the power of the most powerful hypothesis test of significance level 0.05 of $H_0 : \vartheta = 4.0$ versus $H_B : \vartheta \leq 3.75$, at the point alternative $\vartheta = 3.70$.

2

Statistics (Ph.D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. For fixed λ, the random variables $X_1, \ldots, X_n$ are independent identically distributed with density $f(x|\lambda) = \lambda e^{-\lambda x} I_{[x>0]}$. Assume that the parameter λ is a random variable with prior distribution Gamma(α, β) for $\alpha, \beta > 0$ known, and that $\alpha + n > 3$.

(a) Find the posterior distribution of λ and the Bayes estimator of λ corresponding to the quadratic loss function $L(a, \lambda) = (a - \lambda)^2$, where a is the estimated value or 'action' variable.

(b) Find the Bayes estimator of λ corresponding to the loss function $\bar{L}(a, \lambda) = (a - \lambda)^2/\lambda^2$.

2. Consider a sample $X_1, \ldots, X_n$ from the population with density

$$f(x) = \frac{1 + \lambda(x - \mu)^2}{1 + \lambda} \, \phi(x - \mu)$$

where $\phi(\cdot)$ is the standard normal density and $(\mu, \lambda) \in \mathbf{R} \times [0, \infty)$ is unknown. Take as given that for a standard normal r.v. Z, $E(Z^2) = 1$, $E(Z^4) = 3$, and $E(Z^6) = 15$.

(a) Find the method of moments estimator of (μ, λ), and show that it is consistent.

(b) For the particular case $\mu = 0$, $\lambda = 1$, find the limiting nondegenerate (joint) distribution, as $n \to \infty$, of a suitably centered and scaled version of your estimator in part (a).

3. A dataset consists of independent triples (W_i, Y_i, Z_i) of independent random variables with distributions as follows,

$$W_i \sim \mathcal{N}(\lambda + \tau, 1) , \quad Y_i \sim \mathrm{Expon}(\lambda + \mu) , \quad Z_i \sim \mathrm{Poisson}(\mu + \tau) , \quad i = 1, \dots, n$$

(Here $\mathcal{N}$ denotes normal distribution parameterized by mean and variance, and $Expon(\theta)$ the exponential distribution with mean $1/\theta$)

(a) With (λ, μ, τ) as unknown parameter, is this an exponential family ? Prove or disprove

(b) Is the parameter (λ, μ, τ) identifiable based on this dataset and statistical model? Prove or disprove.

(c) Is there a UMVU estimator for λ in this setting ? Prove or disprove.

4. Suppose that you observe two independent samples of data, $X_1, \dots, X_n$ with $\mathrm{Expon}(\lambda)$ density and $Y_1, \dots, Y_n$ with $\mathcal{N}(0, \sigma^2)$ density

(a) Find a pivotal statistic for $\rho = \lambda/\sigma^2$ (as a function of the sufficient statistics for λ and σ^2), and find the form of a 95% confidence interval for ρ based on this pivotal quantity and indicate how you would find its endpoints if $n = m = 10$. What if $n = m = 100$?

(b) Find the form of the likelihood ratio test for $H_0 : \lambda/\sigma^2 = \rho_1$ (against the general alternative) for a fixed value ρ_1, and indicate how you could use such a test (with large-sample approximate cutoff) to define an approximate two-sided 95% confidence interval for ρ.

5. You observe a sample of 100 independent observations X_i from a population with the density

$$g(x) = C \sqrt{\lambda} \, \exp(-\lambda x^2 - \lambda^2 x^4) , \quad -\infty < x < \infty$$

where $C = 0.73077$ is the normalizing constant, and $\lambda > 0$ is an unknown parameter.

(a) Find the minimal sufficient statistic for λ, and show that it is incomplete

(b) Find an expression for the Maximum Likelihood Estimator $\hat{\lambda}$ of λ.

6. You observe a sample $X_1, \dots, X_{20}$ with the density

$$f(x, \vartheta) = 2 \, (x/\vartheta^2) I_{[0 \le x < \vartheta]}$$

with an unknown parameter $\vartheta > 0$, yielding

$$\min(X_i, i = 1, \dots, 80) = 0.035 , \quad \max(X_i, 1 \le i \le 80) = 3.66$$

Based on these data, give numerical expressions for

(a) the p-value of the most powerful hypothesis test for $H_0 : \vartheta = 4.0$ versus $H_A : \vartheta = 3.66$;

(b) the power of the most powerful hypothesis test of significance level 0.05 of $H_0 : \vartheta = 4.0$ versus $H_B : \vartheta \le 3.75$, at the point alternative $\vartheta = 3.70$.

2

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
January, 2010

Statistics (M.A. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Based on an i.i.d. normal data sample $Z_1, \ldots, Z_n \sim \mathcal{N}(\mu, \sigma^2)$, with unknown parameters $\mu \in \mathbf{R}$ and $\sigma^2 \in (0, \infty)$, find and justify a UMVUE for σ. (The answer will be an explicit function of the data involving a constant expressed in terms of gamma function values depending on n.)

2. Let $T_1, \ldots, T_n$ be independent positive random variables with common Weibull$(3, \lambda)$ distribution function

$$F(t; \lambda) = \begin{cases} 1 - \exp(-\lambda t^3). & \text{if } t > 0 \\ 0 & \text{if } t \le 0. \end{cases}$$

Find 95% confidence intervals for the unknown median m of T_1 (where m is defined by $P[T_1 < m] = 1/2$), by two different methods:

(a) based on the maximum likelihood estimator of m and the δ-method,

(b) based on inverting the generalized likelihood ratio test to obtain a "test-based confidence interval"

3. Observed data $\mathbf{X} = (X_1, \ldots, X_n)$ are an i.i.d. sample drawn from one of two possible exponential distributions with unknown parameter $\theta \in \{1, 2\}$. That is, given $\theta = j$, the X_i have common density $f(x \mid j) = j \exp(-jx)I\{x > 0\}$. It is desired to identify the unknown θ when the prior distribution is $\pi(j) = P[\theta = j] = 1/2$, $j = 1, 2$. The loss function is $L(\theta, \hat{\theta}) = |\theta - \hat{\theta}|$

(a) Find the Bayes rule for identifying θ.

(b) What is the Bayes decision if $n = 5$ and it is observed that $\sum X_i = 5 \ln(2)$?

4. Let $X_1, \ldots, X_n$ be a random sample with density

$$f(x; \lambda, \vartheta) = \lambda e^{\vartheta} \exp(-\lambda e^{\vartheta} x)I\{x > 0\}$$

and let $(Y_1, \ldots, Y_n)$ be a random sample from

$$f(y; \lambda, \vartheta) = \lambda e^{-\vartheta} \exp(-\lambda e^{-\vartheta} y)I\{y > 0\}$$

(a) Find the shortest large sample confidence interval you can for ϑ based on these data when the parameters (λ, ϑ) are both unknown.

(b) Compute a similar confidence interval for ϑ using the additional knowledge that $\lambda = 2$. Compare the length of this confidence interval to the one in part (a)

5. Let $(Z_1, \ldots Z_n)$ be an i.i.d. sample with common probability mass function

$$p_Z(j) = \vartheta^2 I_{[j=1]} + 2\vartheta(1 - \vartheta) I_{[j=2]} + (1 - \vartheta)^2 I_{[j=3]}$$

(a) Find a UMP size α hypothesis test for $H_0 : \vartheta = 1/3$ versus $H_A : \vartheta > 1/3$. Give an explicit formula determining the constant threshold in the rejection region.

(b) Give the approximate power of the test, when $n = 100$, against the alternative $H_1 : \vartheta = 1/2$

6. Let $\mathbf{X} \equiv (X_1, \ldots, X_n)$ be a sample from a population with probability density function

$$f(x) = C(\beta, \vartheta) \, e^{-\beta x} \, I_{[x \geq \vartheta]}$$

where $\beta, \vartheta > 0$ are unknown parameters

(a) Find and justify a 2-dimensional sufficient statistic for the unknown parameter (β, ϑ) as a function of $\mathbf{X}$

(b) Are the parameters identifiable? Give the definition of identifiability and prove or disprove identifiability from first principles in this setting

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Based on an i.i.d. normal data sample $Z_1, \ldots, Z_n \sim \mathcal{N}(\mu, \sigma^2)$, with unknown parameters $\mu \in \mathbf{R}$ and $\sigma^2 \in (0, \infty)$, find and justify a UMVUE for σ. (The answer will be an explicit function of the data involving a constant expressed in terms of gamma function values depending on n.)

2. Let $T_1, \ldots, T_n$ be independent positive random variables with common Weibull$(3, \lambda)$ distribution function

$$F(t; \lambda) = \begin{cases} 1 - \exp(-\lambda t^3) & \text{if } t > 0 \\ 0 & \text{if } t \leq 0. \end{cases}$$

Find 95% confidence intervals for the unknown median m of T_1 (where m is defined by $P[T_1 < m] = 1/2$), by two different methods:

(a) based on the maximum likelihood estimator of m and the δ-method,

(b) based on inverting the generalized likelihood ratio test to obtain a "test-based confidence interval."

3. Observed data $\mathbf{X} = (X_1, \ldots, X_n)$ are an i.i.d. sample drawn from one of three possible exponential distributions with unknown parameter $\theta \in \{1, 2, 3\}$. That is, given $\theta = j$, the X_i have common density $f(x \mid j) = j \exp(-jx)I\{x > 0\}$. It is desired to identify the unknown θ when the prior distribution is $\pi(j) = P[\theta = j] = 1/3$, $j = 1, 2, 3$. The loss function is $L(\theta, \hat{\theta}) = |\theta - \hat{\theta}|$

(a) Find the Bayes rule for identifying θ.

(b) What is the Bayes decision if $n = 5$ and it is observed that $\sum X_i = 5\ln(2)$?

4. Let $X_1, \ldots, X_n$ be a random sample with density

$$f(x; \lambda, \vartheta) = \lambda e^{\vartheta} \exp(-\lambda e^{\vartheta} x)I\{x > 0\}$$

and let $(Y_1, \ldots, Y_n)$ be a random sample from

$$f(y; \lambda, \vartheta) = \lambda e^{-\vartheta} \exp(-\lambda e^{-\vartheta} y)I\{y > 0\}.$$

(a) Find the shortest large sample confidence interval you can for ϑ based on these data when the parameters (λ, ϑ) are both unknown.

(b) Compute a similar confidence interval for ϑ using the additional knowledge that $\lambda = 2$. Compare the length of this confidence interval to the one in part (a)

5. Let $(Z_1, \ldots Z_n)$ be an i.i.d. sample with common probability mass function

$$p_Z(j) = \vartheta^2 I_{[j=1]} + 2\vartheta(1 - \vartheta) I_{[j=2]} + (1 - \vartheta)^2 I_{[j=3]}$$

(a) Find a UMP size α hypothesis test for $H_0 : \vartheta = 1/3$ versus $H_A : \vartheta > 1/3$. Give an explicit formula determining the constant threshold in the rejection region.

(b) Give the approximate power of the test, when $n = 100$, against the alternative $H_1 : \vartheta = 1/2$

6. Let $\mathbf{X} \equiv (X_1, \ldots, X_n)$ be a sample from a population with probability density function

$$f(x) = C(\beta, \vartheta) \, e^{-\beta x} \, I_{[x \geq \vartheta]}$$

where $\beta, \vartheta > 0$ are unknown parameters.

(a) Find and justify a 2-dimensional sufficient statistic for the unknown parameter (β, ϑ) as a function of $\mathbf{X}$

(b) Is your sufficient statistic complete? Is it minimal?

(c) Are the parameters identifiable? Give the definition of identifiability and prove or disprove identifiability from first principles in this setting

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
AUGUST, 2009

Statistics (M.A. Version)

Instructions to the Student

a Answer all six questions. Each will be graded from 0 to 10.

b Use a different booklet for each question. Write the problem number and
   your code number (**NOT YOUR NAME**) on the outside cover

c Keep scratch work on separate pages in the same booklet.

d If you use a "well known" theorem in your solution to any problem, it
   is your responsibility to make clear which theorem you are using and to
   justify its use

---

1. Let the distribution P_ϑ of data X depend on a parameter $\vartheta \in \mathbf{R}$.

(a) An unbiased estimator $T = T(X)$ of a parameter $\tau = \tau(\vartheta)$ is said to
   be UMVUE, if it has variance less than or equal (for all ϑ) to any other
   finite-variance unbiased estimator of $\tau(\vartheta)$ Prove that T is UMVUE if
   and only if it is uncorrelated with every finite variance statistic S such
   that $E_\vartheta(S) = 0$. *Hint: consider the variance of $T + \epsilon S$ for real ϵ in a
   neighborhood of 0, and for S an unbiased finite-variance estimator of 0.*

(b) Prove that if $T_i(X)$ is the UMVUE of a parameter function $\tau_i(\vartheta)$, $i = 1, 2$,
   then $c_1 T_1(X) + c_2 T_2(X)$ is the UMVUE of $c_1 \tau_1(\vartheta) + c_2 \tau_2(\vartheta)$

2. Let $X_1, \ldots, X_m$ be i.i.d. exponential variables with mean μ_1 and let $Y_1, \ldots, Y_n$
be i i d. exponential variables with mean μ_2.

(a) Derive the likelihood ratio test for H_0: $\mu_1 = \mu_2$ and show that the test can
   be expressed in terms of $\bar{Y}/\bar{X}$

(b) Derive the power function of this test.

1

3. Suppose that $X_1, \ldots, X_n$ is a sample from a $\mathcal{N}(-\frac{1}{2}\sigma^2, \sigma^2)$ density

(a) Show that this distribution forms an exponential family, with complete sufficient statistic for σ^2, but that the statistic $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$ is **not** complete (i.e. that there are nontrivial functions of this statistic which are unbiased estimates of 0)

(b) Find an expression for the MLE $\hat{\sigma}^2$, and find its large-sample asymptotic variance

4. A sample of observations $Y_1, Y_2, \ldots, Y_n$ is observed with density $f_Y(y) = b^{-1} g((y-a)/b)$, where $g(x) \equiv (1 - |x|) I_{[|x|\le 1]}$, and where $a \in \mathbf{R}$, $b > 0$ are unknown parameters.

(a) Show that

$$\frac{\bar{Y} - a}{n^{-1}\sum_{i=1}^{n} |Y_i - \bar{Y}|} \qquad \text{and} \qquad \frac{1}{nb}\sum_{i=1}^{n} |Y_i - \bar{Y}|$$

are pivotal quantities for a and for b, respectively.

(b) Use the first of these pivotal quantities to construct a hypothesis test of significance level approximately 0.05 for large n, of the null hypothesis $H_0 : a \le 0$ versus $H_A : a > 0$. First describe the rejection region exactly in terms of distributions which could be calculated numerically, and then find a large sample approximation to the boundary point(s) of the rejection region

5. Consider independent observations $Y_i \sim \text{Poisson}(\alpha)$, where $\alpha > 0$ is an unknown parameter. Suppose that α has prior density $\alpha \sim \text{Gamma}(2, 3)$.

(a) Find the posterior density for α given $\{Y_i\}_{i=1}^{n}$.

(b) Find the Bayes estimator $\hat{\alpha}$ of α which minimizes the risk according to the loss function $L(\hat{\alpha}, \alpha) \equiv (\hat{\alpha} - \alpha)^2/\alpha$.

6. Suppose we observe two independent frequency counts, $X \sim \text{Binom}(n, \pi_1)$ and $Y \sim \text{Binom}(m, \pi_2)$, where n and m are large integers and $\pi_1, \pi_2 \in (0, 1)$ are unknown parameters

(a) If we let $\vartheta = \pi_1 - \pi_2$, then find the Rao Score and Wald test statistics for testing the hypothesis $H_0 : \vartheta = 0$ versus the general alternative $H_A : \vartheta \neq 0$ Are these test statistics equivalent (each a monotone function of the other) for finite n, m ? In what sense are these statistics asymptotically equivalent under H_0 when $n, m \to \infty$?

(b) When $X = 73$, $n = 120$, $Y = 115$, $m = 180$, find the large-sample approximate p-value of either the Wald or the Rao Score test of the hypothesis H_0 in (a) (*Your answer should be expressed as the normal distribution function value of an explicit numerical expression, which you need not reduce.*)

3

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
AUGUST, 2009

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10

b Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover

c. Keep scratch work on separate pages in the same booklet

d If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let the distribution P_ϑ of data X depend on a parameter $\vartheta \in \mathbf{R}$.

(a) An unbiased estimator $T = T(X)$ of a parameter $\tau = \tau(\vartheta)$ is said to be UMVUE, if it has variance less than or equal (for all ϑ) to any other finite-variance unbiased estimator of $\tau(\vartheta)$. Prove that T is UMVUE if and only if it is uncorrelated with every finite variance statistic S such that $E_\vartheta(S) = 0$. *Hint: consider the variance of $T + \epsilon S$ for real ϵ in a neighborhood of 0, and for S an unbiased finite-variance estimator of 0.*

(b) Prove that if $T_i(X)$ is the UMVUE of a parameter function $\tau_i(\vartheta)$, $i = 1, 2$, then $c_1 T_1(X) + c_2 T_2(X)$ is the UMVUE of $c_1 \tau_1(\vartheta) + c_2 \tau_2(\vartheta)$.

2. (a) State Basu's Theorem (that under suitable hypotheses, sufficient and ancillary statistics are independent), and prove it.

(b) Show how Basu's Theorem implies the independence of $n\bar{X}/(n\bar{X} + m\bar{Y})$ and $n\bar{X} + m\bar{Y}$ based on two independent samples $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ respectively from Gamma(a, λ) and Gamma(b, λ) densities

1

3. Suppose that $X_1, \ldots, X_n$ is a sample from a $\mathcal{N}(-\frac{1}{2}\sigma^2, \sigma^2)$ density

(a) Show that this distribution forms an exponential family, with complete sufficient statistic for σ^2, but that the statistic $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$ is **not** complete (i.e. that there are nontrivial functions of this statistic which are unbiased estimates of 0).

(b) Find an expression for the MLE $\hat{\sigma}^2$, and find its large-sample asymptotic variance

4. A sample of observations $Y_1, Y_2, \ldots, Y_n$ is observed with density $f_Y(y) = b^{-1} g((y-a)/b)$, where $g(x) \equiv (1 - |x|) I_{[|x| \leq 1]}$, and where $a \in \mathbf{R}$, $b > 0$ are unknown parameters.

(a) Show that

$$\frac{\bar{Y} - a}{n^{-1}\sum_{i=1}^{n} |Y_i - \bar{Y}|} \qquad \text{and} \qquad \frac{1}{nb}\sum_{i=1}^{n} |Y_i - \bar{Y}|$$

are pivotal quantities for a and for b, respectively

(b) Use the first of these pivotal quantities to construct a hypothesis test of significance level approximately 0.05 for large n, of the null hypothesis $H_0 : a \leq 0$ versus $H_A : a > 0$. First describe the rejection region exactly in terms of distributions which could be calculated numerically, and then find a large sample approximation to the boundary point(s) of the rejection region.

5. Consider independent observations $Y_i \sim \text{Poisson}(\alpha\, e^{\beta X_i})$, where $\alpha > 0$ and $\beta \in \{-b_0, b_0\}$ are unknown parameters but b_0 is known. Suppose that these unknown parameters have prior distribution which makes them independent, with $\alpha \sim \text{Gamma}(2, 3)$ and $\beta = \pm b_0$ with probabilities $1/2$ each. Assume that $X_i = \pm 1$ and that these values are fixed in advance with n even and $\sum_{i=1}^{n} X_i = 0$. (That is, half the observations have $X_i = 1$ and the other half have $X_i = -1$.)

(a) Find the posterior density for (α, β) given $\{Y_i\}_{i=1}^{n}$. (This distribution is continuous in the α variable but is a two-point discrete density in the β variable.) Show that α and β are independent according to this posterior density.

(b) Find the Bayes estimator $\hat{\alpha}$ of α (in the presence of unknown β) which minimizes the risk according to the loss function $L(\hat{\alpha}, \alpha) \equiv (\hat{\alpha} - \alpha)^2/\alpha$.

6. Suppose we observe two independent frequency counts, $X \sim \text{Binom}(n, \pi_1)$ and $Y \sim \text{Binom}(m, \pi_2)$, where n and m are large integers and $\pi_1, \pi_2 \in (0, 1)$ are unknown parameters.

(a) If we let $\vartheta = \pi_1 - \pi_2$, then find the Rao Score and Wald test statistics for testing the hypothesis $H_0 : \vartheta = 0$ versus the general alternative $H_A : \vartheta \neq 0$. Are these test statistics equivalent (each a monotone function of the other) for finite n, m ? In what sense are these statistics asymptotically equivalent under H_0 when $n, m \to \infty$?

(b) When $X = 73$, $n = 120$, $Y = 115$, $m = 180$, find the large-sample approximate p-value of either the Wald or the Rao Score test of the hypothesis H_0 in (a). (*Your answer should be expressed as the normal distribution function value of an explicit numerical expression, which you need not reduce.*)

3

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
JANUARY, 2009

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. A sample of observations X_i, $i = 1, \ldots, n$, follows a $\mathcal{N}(0, \sigma^2)$ distribution, where the parameter $\tau \equiv 1/\sigma^2$ has prior density

$$\pi_\tau(t) = \frac{\lambda^\alpha t^{\alpha - 1}}{\Gamma(\alpha)} \exp(-\lambda t)$$

(a) Show that this *inverse Gamma* family of prior densities is a conjugate prior family for the data $\mathbf{X} = \{X_i\}_{i=1}^n$, and find the parameters of the posterior density in terms of $\mathbf{X}$.

(b) Assume that $n \geq 6$, and find the Maximum Posterior Density estimator $\tilde{\sigma}^2$ of σ^2. That is, find the $\sigma^2 = s$ argument maximizing $\pi_{\sigma^2 \mid \mathbf{X}}(s \mid \mathbf{X})$, the posterior density of σ^2 given the data $X_1, \ldots, X_n$.

2. Suppose one has two independent data samples $\{X_i\}_{i=1}^n$ and $\{Y_i\}_{i=1}^n$ where $X_i \sim \mathcal{N}(1, \lambda)$ and $Y_i \sim \mathcal{N}(2, 1 + \lambda)$.

(a) Find (and justify that you have found) a minimal sufficient statistic for λ.

(b) Explain why the estimator $\hat{\lambda}$ obtained as the solution of the estimating equation

$$\sum_{i=1}^n \left\{ (X_i - 1)^2 - \frac{\lambda}{1 + \lambda} (Y_i - 2)^2 \right\} = 0$$

is asymptotically normally distributed with mean λ, and find its asymptotic variance.

1

3. A data sample Y_i, $i = 1, \ldots, n$, is observed from the density $f(y, \vartheta) = \vartheta \, y^{-\vartheta - 1} I_{[y \geq 1]}$. Find the standard Method of Moments estimator $\bar{\vartheta}$ of ϑ, when $\vartheta > 2$, and find the efficiency of $\sqrt{\bar{\vartheta}}$ for estimating $\sqrt{\vartheta}$.

4. Let $\{X_i\}_{i=1}^n$ be i.i.d. with an $\text{Expon}(\lambda)$ distribution and let $\{Y_i\}_{i=1}^n$ be i.i.d. with an $\text{Expon}(1 + \lambda)$ distribution, where $\text{Expon}(\vartheta)$ has density $\vartheta \, e^{-\vartheta t} I_{[t > 0]}$. Note that $P(Y_1 > X_1) = \lambda/(1 + 2\lambda)$. (You need not show this.)

(a) Find an expression (which may be in the form of a single or double integral with explicitly given integrand) for an unbiased estimator of $\lambda/(1+2\lambda)$ which is a function of the sufficient statistic(s) for λ based on $\{X_i, Y_i\}_{i=1}^n$.

(b) Find a lower bound for the variance of all unbiased estimators of $\lambda/(1 + 2\lambda)$. Is there an estimator (based on the finite sample) that attains this bound?

5. Consider a data sample $\{Z_i\}_{i=1}^n$ from the density

$$f(z, \vartheta) \;=\; \vartheta \, I_{[1 \leq |z| < 2]} \;+\; \left(\frac{1}{2} - \vartheta\right) 3z^2 \, I_{[|z| < 1]}$$

with unknown parameter $\vartheta \in (0, \frac{1}{2})$. Find and justify a UMP level .05 test for the null hypothesis $H_0 : \vartheta = 1/4$ versus the alternative $H_1 : \vartheta \geq 1/3$. Does your test involve auxiliary randomization? Does there exist a nonrandomized test of size .05 which is UMP for H_0 versus H_1? Give (and justify) larger H_0' and H_1' parameter sets with respect to which your test is also a UMP level .05 test.

6. (a) Derive the Generalized Likelihood Ratio Test of level $\alpha = .05$ for $H_0 : \lambda_1 = \lambda_2$ versus $H_1 : \lambda_1 \neq \lambda_2$ based on independent data samples $\{X_i\}_{i=1}^n$ and $\{Y_i\}_{i=1}^n$ where $X_i \sim \text{Expon}(\lambda_1)$ and $Y_i \sim \text{Expon}(\lambda_2)$.

(b) Explain why you know for large n that a test based only on the data $\{X_i/(X_i + Y_i)\}_{i=1}^n$ will be less powerful than the test you found in part (a).

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let $(X_1, \ldots, X_n)$ be a sample from a population with probability density function

$$f(x; \theta) = \theta |x| / (1 + x^2)^{\theta + 1}, \quad -\infty < x < \infty$$

with $\theta > 0$ as a parameter.

(a) Construct the maximum likelihood estimator (MLE) of θ and decide whether it is unbiased.

(b) Under an appropriate normalization, find the limiting distribution of the MLE as $n \to \infty$.

2. Let $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ be independent samples of size n from populations with probability density functions

$$
\begin{aligned}
f_X(x; \sigma_1) &= (1/\sigma_1) \exp(-x/\sigma_1) I\{x > 0\}, \\
f_Y(y; \sigma_2) &= (1/\sigma_2) \exp(-y/\sigma_2) I\{y > 0\},
\end{aligned}
$$

where σ_1 and σ_2 are unknown positive parameters.

(a) Show that the likelihood ratio (LR) test of size α, $0 < \alpha < 1$, of the null hypothesis H_0: $\sigma_1 = \sigma_2$ (unspecified) versus the alternative H_1: $\sigma_1 \neq \sigma_2$ rejects H_0 if

$$
\text{either } 0 < \sqrt{\bar{X}/\bar{Y}} < c_1 \quad \text{or} \quad \sqrt{\bar{X}/\bar{Y}} > c_2
$$

and accepts H_0 otherwise, where $c_1 < c_2$ are the roots of the quadratic equation in z

$$
z^2 - 2b_n(\alpha)z + 1 = 0
$$

and the coefficient $b_n(\alpha) > 1$ is determined by n and α.

(b) Derive the power function of the LR test and prove that it is consistent; that is,

$$
\lim_{n \to \infty} P_{\sigma_1, \sigma_2}[\text{reject } H_0] = 1
$$

for any fixed σ_1, σ_2, $\sigma_1 \neq \sigma_2$.

3. Let $(x_1, \ldots, x_n)$ be a sample from a population with probability density function

$$
f(x; \theta) = \begin{cases} x/\theta^2 & \text{if } 0 < x < \theta \\ (2\theta - x)/\theta^2 & \text{if } \theta < x < 2\theta \\ 0 & \text{otherwise} \end{cases}
$$

with $\theta > 0$ as a parameter.

(a) Construct the method of moments estimator of θ and find, under proper normalization, its limiting distribution as $n \to \infty$.

(b) Based on the method of moments estimator, construct an asymptotic confidence interval of level $1 - \alpha$ for θ.

4. Assume that given θ, X has a binomial distribution with parameters (n, θ); that is, $X|\theta \sim \text{Bin}(n, \theta)$. Assume also that the prior distribution of θ is uniform on $(0, 1)$.

   (a) Prove that the family of posterior densities $\{\pi(\theta \,|\, X = x), x = 0, 1, \ldots, n\}$ has the monotone likelihood ratio (MLR) property.

   (b) Show that for all $u \in (0, 1)$, the posterior distributions satisfy

   $$x'' > x' \Rightarrow P(\theta > u \,|\, X = x'') > P(\theta > u \,|\, X = x').$$

5. Let the distribution of a random vector (X, Y, Z) be given by a positive density $p(x, y, z; \theta)$ depending on a parameter $\theta \in \Theta$. Denote by $p_1(x, y; \theta)$ the marginal density of (X, Y).

   (a) Show that if $T(X, Y)$ is sufficient for the family $\{p(x, y, z; \theta), \theta \in \Theta\}$, it is sufficient for the family $\{p_1(x, y; \theta), \theta \in \Theta\}$.

   (b) Prove that if both $T(X, Y)$ and (separately) $S(X, Z)$ are sufficient statistics for $\{p(x, y, z; \theta), \theta \in \Theta\}$, then X also is sufficient.

6. Let $X_1, \ldots, X_n$ be identically distributed and serially correlated normal random variables with $E[X_i] = \mu$, $\text{Var}\, X_i = \sigma^2$, $\text{Cov}(X_i, X_{i+1}) = \rho\sigma^2$, $i = 1, \ldots, n-1$, and all other covariances equal to zero.

   (a) Show that $E[\bar{X}] = \mu$,

   $$\text{Var}\,\bar{X} = \frac{\sigma^2}{n}\left[1 + 2\rho\left(1 - \frac{1}{n}\right)\right] \quad \text{and} \quad E[s^2] = \sigma^2\left[1 - \frac{2\rho}{n}\right]$$

   where $\bar{X}$ and s^2 are the usual sample mean and sample variance.

   (b) Show that s^2 is a consistent estimator of σ^2.

   (c) Under these assumptions, what is the limiting distribution of the Student t statistic $\sqrt{n}(\bar{X} - \mu)/s$? What is the effect of the serial correlation on large sample inference on μ?

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
August, 2008

Statistics (M.A. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let $(X_1, \ldots, X_n)$ be a sample from a population with probability density function
$$f(x; \theta) = \theta |x| / (1 + x^2)^{\theta+1}, \quad -\infty < x < \infty$$
with $\theta > 0$ as a parameter.

(a) Construct the maximum likelihood estimator (MLE) of θ and decide whether it is unbiased. *Hint:* Use Jensen's inequality.

(b) Under an appropriate normalization, find the limiting distribution of the MLE as $n \to \infty$.

2. Let $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ be independent samples of size n from populations with probability density functions

$$
\begin{aligned}
f_X(x; \sigma_1) &= (1/\sigma_1) \exp(-x/\sigma_1) I\{x > 0\}, \\
f_Y(y; \sigma_2) &= (1/\sigma_2) \exp(-y/\sigma_2) I\{y > 0\},
\end{aligned}
$$

where σ_1 and σ_2 are unknown positive parameters.

(a) Show that the likelihood ratio (LR) test of size α, $0 < \alpha < 1$, of the null hypothesis H_0: $\sigma_1 = \sigma_2$ (unspecified) versus the alternative H_1: $\sigma_1 \neq \sigma_2$ rejects H_0 if

$$
\text{either } 0 < \sqrt{\bar{X}/\bar{Y}} < c_1 \quad \text{or} \quad \sqrt{\bar{X}/\bar{Y}} > c_2
$$

and accepts H_0 otherwise, where $c_1 < c_2$ are the roots of the quadratic equation in z

$$
z^2 - 2b_n(\alpha)z + 1 = 0
$$

and the coefficient $b_n(\alpha) > 1$ is determined by n and α.

(b) Derive the power function of the LR test and prove that it is consistent; that is,

$$
\lim_{n \to \infty} P_{\sigma_1, \sigma_2}[\text{reject } H_0] = 1
$$

for any fixed σ_1, σ_2, $\sigma_1 \neq \sigma_2$.

3. Let $(x_1, \ldots, x_n)$ be a sample from a population with probability density function

$$
f(x; \theta) = \begin{cases} x/\theta^2 & \text{if } 0 < x < \theta \\ (2\theta - x)/\theta^2 & \text{if } \theta < x < 2\theta \\ 0 & \text{otherwise} \end{cases}
$$

with $\theta > 0$ as a parameter.

(a) Construct the method of moments estimator of θ and find, under proper normalization, its limiting distribution as $n \to \infty$.

(b) Based on the method of moments estimator, construct an asymptotic confidence interval of level $1 - \alpha$ for θ.

4. Assume that given θ, X has a binomial distribution with parameters (n, θ); that is, $X|\theta \sim \text{Bin}(n, \theta)$. Assume also that the prior distribution of θ is uniform on $(0, 1)$.

(a) Prove that the family of posterior densities $\{\pi(\theta|x), x = 0, 1, \ldots, n\}$ has the monotone likelihood ratio (MLR) property.

(b) Show that for all $u \in (0, 1)$, the posterior distributions satisfy

$$x'' > x' \Rightarrow P(\theta > u|X = x'') > P(\theta > u|X = x').$$

5. Let the distribution of a random vector (X, Y, Z) be given by a positive density $p(x, y, z; \theta)$ depending on a parameter $\theta \in \Theta$. Denote by $p_1(x, y; \theta)$ the marginal density of (X, Y). Show that if $T(X, Y)$ is sufficient for the family $\{p(x, y, z; \theta), \theta \in \Theta\}$, it is sufficient for the family $\{p_1(x, y; \theta), \theta \in \Theta\}$.

6. Let $Y_{ij} = \mu_i + e_{ij}$, $i = 1, \ldots, k$, $j = 1, \ldots, n$, where the e_{ij} are i.i.d. $N(0, \sigma^2)$ random variables..

(a) Prove that

$$\bar{Y}_{i\cdot} = \frac{1}{n}\sum_{j=1}^{n} Y_{ij}, \quad i = 1, \ldots, k,$$

and

$$s_p^2 = \frac{1}{k(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}(Y_{ij} - \bar{Y}_{i\cdot})^2$$

are mutually independent. What are the distributions of these statistics?

(b) Now assume $Y_{ij} = \mu + a_i + e_{ij}$ where the a_i are i.i.d. $N(0, \sigma_a^2)$ and the a_i and e_{ij} are mutually independent. Show that the Y_{ij} have the following correlation pattern:

$$\text{Cov}(Y_{ij}, Y_{kl}) = \begin{cases} \sigma_Y^2 & \text{if } i = k \text{ and } j = l \\ \rho\sigma_Y^2 & \text{if } i = k \text{ and } j \neq l \\ 0 & \text{if } i \neq k \end{cases}$$

where $\sigma_Y^2 = \sigma_a^2 + \sigma^2$ and $\rho = \sigma_a^2/(\sigma_a^2 + \sigma^2)$. How does your answer to (a) change?

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
January, 2008

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. (a) A statistician observes a pair $(\mathbf{X}, \mathbf{Y})$ of independent random vectors with positive density functions $f(\mathbf{x}; \theta)$ and $g(\mathbf{y}; \theta)$ depending on the parameter θ. Prove that $\mathbf{X}$ is sufficient for θ if and only if the distribution of $\mathbf{Y}$ does not depend on θ.

   (b) Now let the joint density of $(\mathbf{X}, \mathbf{Y})$ be $h(\mathbf{x} - \boldsymbol{\theta}, \mathbf{y})$, where the observation $\mathbf{X}$ and the parameter $\boldsymbol{\theta}$ are vectors of the same dimension. Prove that if $\mathbf{X}$ is sufficient for $\boldsymbol{\theta}$, then $\mathbf{X}$ and $\mathbf{Y}$ are independent. (*Hint*: Use the conditional density of $\mathbf{Y}$ given $\mathbf{X}$.)

2. Let $X_1, \ldots, X_n$ be a sample from a normal population with unknown mean μ and unknown variance σ^2.

 (a) Let $n \geq 4$. Prove that μ/σ has a unique minimum variance unbiased estimator of the form $c\bar{X}/s$. Evaluate the constant c explicitly.

 (b) What happens if $n \leq 3$?

3. Let $X_1, \ldots, X_n$ be a sample from the uniform distribution on $(-\theta, 2\theta)$ with $\theta > 0$ as a parameter.

(a) Calculate the MLE of θ, and find its density function and mean square error.

(b) Show that a UMP size α test of $H_0 : \theta \leq 1$ versus $H_A : \theta > 1$ can be based on the MLE from part (i), and explain why the power of this test against each alternative value of θ converges to 1 as $n \to \infty$.

4. Let $X_1, \ldots, X_n$ be a sample from a normal population with unknown mean μ and unknown variance σ^2.

(a) Prove that for any $c > 0$ the coverage probability of the confidence interval

$$\bar{X} - c \sum_1^n |X_i - \bar{X}| < \mu < \bar{X} + c \sum_1^n |X_i - \bar{X}| \qquad (*)$$

does not depend on μ or σ.

(b) Find c such that the mean length of the interval $(*)$ is equal to a given quantity Δ.

5. Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be independent samples from populations with densities

$$
\begin{aligned}
f(x; \theta_1) &= (x/\theta_1^2) \exp(-x/\theta_1), & x > 0, \\
f(y; \theta_2) &= (y/\theta_2^2) \exp(-y/\theta_2), & y > 0,
\end{aligned}
$$

respectively, where $\theta_1 > 0$ and $\theta_2 > 0$ are parameters.

(a) Find the critical region of the likelihood ratio test of size α of $H_0 : \theta_1 = \theta_2 = \theta$ (unspecified) versus $H_1 : \theta_1 \neq \theta_2$. (You do not need to find the critical constant of the test).

(b) Show that the above test is similar, that is, that the probability of rejecting H_0 when it is true is the same for all $\theta_1 = \theta_2 = \theta$.

6. Let $X_1, \ldots, X_n$ be a sample from a population with probability density

$$2|x - \theta|^3 \exp\{-(x - \theta)^4\}, \qquad -\infty < x < \infty,$$

where $\theta \in \mathbb{R}$ is an unknown parameter.

(a) Find the method of moments estimator of θ and, after an appropriate normalization, its nondegenerate limiting distribution. (*Hint:* Recall the identities $\Gamma(y + 1) = y\Gamma(y)$ for $y > 0$, and $\Gamma(1/2) = \sqrt{\pi}$).

(b) Calculate the Fisher information on θ and the asymptotic efficiency of the method of moments estimator.

(c) Find an integral formula for the Bayes estimator of θ with respect to the squared-error loss function if θ has a $N(0, 1)$ prior distribution.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
January, 2008

Statistics (M.A. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let $X_1, \ldots, X_n$ be a sample from a uniform distribution on $(-\theta, \theta)$ with $\theta > 0$ as a parameter.

(a) Calculate the MLE of θ, and find its density and mean squared error.

(b) Show that a UMP size α test of H_0: $\theta \leq 1$ versus H_A: $\theta > 1$ can be based on the MLE from part (i), and explain why the power of this test versus each alternative value of θ converges to 1 as $n \to \infty$.

2. Let $X_1, \ldots, X_n$ be i.i.d $N(\mu, \sigma^2)$.

(a) Show that $(\bar{X}, s)$ is a complete sufficient statistic.

(b) Assume $n \geq 4$. Find a constant c such that $c\bar{X}/s$ is an unbiased estimator of μ/σ.

(c) Prove that if $n \geq 4$, the statistic $c\bar{X}/s$ of part (b) is the unique minimum variance unbiased estimator of μ/σ.

3. Let $X_1, \ldots, X_n$ be a sample from a normal population with unknown mean μ and unknown variance σ^2.

(a) Prove that for any $c > 0$ the coverage probability of the confidence interval

$$\bar{X} - c\sum_1^n |X_i - \bar{X}| < \mu < \bar{X} + c\sum_1^n |X_i - \bar{X}| \qquad (*)$$

does not depend on μ or σ.

(b) Find c such that the mean length of the interval $(*)$ is equal to a given quantity Δ.

4. Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be independent samples from populations with densities

$$
\begin{aligned}
f(x; \theta_1) &= (x/\theta_1^2)\exp(-x/\theta_1), & x > 0, \\
f(y; \theta_2) &= (y/\theta_2^2)\exp(-y/\theta_2), & y > 0,
\end{aligned}
$$

respectively, where $\theta_1 > 0$ and $\theta_2 > 0$ are parameters.

(a) Find the critical region of the likelihood ratio test of size α of $H_0 : \theta_1 = \theta_2 = \theta$ (unspecified) versus $H_1 : \theta_1 \neq \theta_2$. (You do not need to find the critical constant of the test).

(b) Show that the above test is similar, that is, that the probability of rejecting H_0 when it is true is the same for all $\theta_1 = \theta_2 = \theta$.

5. Let $X_1, \ldots, X_n$ be a sample from a population with probability density

$$f(x; \theta) = \begin{cases} \theta x^{\theta-1} & \text{if } 0 < x < 1, \\ 0 & \text{otherwise,} \end{cases}$$

where θ is a positive parameter.

(a) Find the method of moments estimator of θ and, after an appropriate normalization, its nondegenerate limiting distribution.

(b) Calculate the Fisher information on θ and the asymptotic efficiency of the method of moments estimator.

6. Given $\theta \in (0,1)$, let (X_0, X_1, X_2) have a trinomial distribution with n observations and cell probabilities

$$p_0(\theta) = (1-\theta)^2, \qquad p_1(\theta) = 2\theta(1-\theta), \qquad p_2(\theta) = \theta^2.$$

The parameter θ has a beta prior density:

$$\begin{aligned}\pi(\theta) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1} \qquad \text{if } 0 < \theta < 1, \\ &= 0 \qquad\qquad\qquad\qquad\qquad \text{otherwise,}\end{aligned}$$

where a and b are known constants.

(a) Assuming that the loss function is $L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$, find $\hat{\theta}_{ab}$, the Bayes estimator of θ.

(b) Show that $\hat{\theta}_{ab}$ is a function of the complete sufficient statistic for θ.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
AUGUST 2007

Statistics (Ph.D. Version)

Instructions to the Student

(a) Answer all six questions. Each will be graded from 0 to 10.

(b) Use a different booklet for each question. Write the problem number and your code
number (NOT YOUR NAME) on the first page.

(c) Keep scratch work on separate pages in the same booklet.

(d) If you use a "well known" theorem in your solution to any problem, it is your
responsibility to make clear which theorem you are using and to justify its use.

---

1. Let $X_1, ..., X_n$ be i.i.d. $N(\mu_1, 1)$ random variables.

    (a) Find the uniformly minimum variance estimators (UMVUE) of μ^2 and
        $\theta = P[X_1 \leq a]$, where a is a given constant. Justify the statement that your
        estimator is unique.

    (b) Is your estimator of μ^2 efficient?

2. Let $Y_1, ..., Y_k$ be independent Binomial (n, π_i) random variables, where
$\log(\pi_i / (1 - \pi_i)) = \beta x_i$ and $x_1, ..., x_k$ are known covariates.

    (a) Find a uniformly most powerful test of $H_0 : \beta \leq \beta_0$ vs. $H_1 : \beta > \beta_0$.

    (b) Let k be fixed and let $n \to \infty$. Under this large sample regime find
        approximate critical values for the test in part (a) and give an approximation
        to its power function.

3. Let $X_1, ..., X_n$ be i.i.d. $N(\mu_1, \sigma^2)$ and let $Y_1, ..., Y_n$ be $N(\mu_2, \sigma^2)$. Assume that the two
samples are independent and that σ^2 is known.

    (a) Show that $\overline{Y} / \overline{X}$ is a consistent and asymptotically normal estimator of
        $\theta = \mu_2 / \mu_1$. Derive the limiting distribution of $\overline{Y} / \overline{X}$ (suitably normalized).

    (b) Using the result of (a), how would you create a $1 - \alpha$ confidence interval for θ ?

    (c) Derive a pivot for θ based on $\overline{Y} - \theta\overline{X}$. Use this pivotal variable to derive an
        exact confidence set for θ. How do you reconcile the results of parts (b) and (c)?

4. Let $X_1, \ldots, X_n$ be i.i.d. with c.d.f. F and density f. Assume f is strictly positive and continuous for all x. Let $m = [np]$ denote the greatest integer less than or equal to np. The population pth quantile is ξ_p, the unique solution of the equation $F(\xi_p) = p$. The sample pth quantile is $\hat{\xi}_p = X_{(m)}$, the mth order statistic. Show that, as $n \to \infty$,

$$\sqrt{n}(\hat{\xi}_p - \xi_p) \to_d N(0, p(1-p)/f(\xi_p)^2.$$

5. Let the discrete random variable X have probability mass function $f(x)$. It is known a priori that

$$P[f = f_0] = \pi_0 \quad \text{and that} \quad P[f = f_{10}] = \pi_1 = 1 - \pi_0.$$

We want to choose a decision rule $\delta(X)$ to choose between f_0 and f_1. Any incorrect decision incurs a loss of 1, and a correct decision incurs a loss of 0. Find the Bayes solution to this problem.

6. Let $X_0, X_1, \ldots, X_n$ be a stationary Markov chain with state space $\{0,1\}$. That is,

$$p_{ij} = P[X_t = j \mid X_{t-1} = i] = P[X_t = j \mid X_{t-1} = i, X_{t-2} = i_{t-2}, \ldots, X_0 = i_0]$$

for $t = 1, \ldots, n$, and $i, j, i_{t-2}, \ldots, i_0 \in \{0,1\}$. Suppose further that

$$P[X_0 = 0] = P[X_0 = 1] = 1/2 \quad \text{and} \quad p_{00} = p_{11} = p \in (0,1).$$

   (a) Show that the joint distribution of $X_0, X_1, \ldots, X_n$ forms a one parameter exponential family with sufficient statistic $T = N_{00} + N_{11}$, where N_{ij} is the number of transitions from state i to state j. For example, 010011 has $n = 5$, $N_{00} = 1, N_{01} = 2, N_{10} = 1$ and $N_{11} = 1$.

   (b) Find the UMVUE of p.

Statistics (M.A. Version)

Instructions to the Student

(a) Answer all six questions. Each will be graded from 0 to 10.

(b) Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the first page.

(c) Keep scratch work on separate pages in the same booklet.

(d) If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let $X_1,...,X_n$ be i.i.d. $N(\mu,1)$.

   (a) Find the uniformly minimum variance estimators (UMVUE) of μ^2 and $\theta = P[X_1 \leq a]$, where a is a given constant. Justify the statement that your estimator is unique.

   (b) Is your estimator of μ^2 efficient? [**Hint:** If Z is $N(\mu,1)$, then $E[Z^3] = 0$ and $E[Z^4] = 3$.]

2. Let $Y_1,...,Y_k$ be independent Binomial (n, π_i) random variables, where
$$\log(\pi_i/(1 - \pi_i)) = \beta x_i$$
and $x_1,...,x_k$ are known covariates.

   (a) Find a uniformly most powerful test of $H_0 : \beta \leq \beta_0$ vs. $H_1 : \beta > \beta_0$.

   (b) Let k be fixed and let $n \to \infty$. Under this large sample regime find approximate critical values for the test in part (a) and give an approximation to its power function.

3. Let $X_1,...,X_n$ be i.i.d. exponential variables with mean μ_1 and let $Y_1,...,Y_n$ be exponential variables with mean μ_2. Assume that the two samples are independent.

   (a) Show that $\overline{Y}/\overline{X}$ is a consistent and asymptotically normal estimator of $\theta = \mu_2/\mu_1$. Derive the limiting distribution of $\overline{Y}/\overline{X}$ (suitably normalized).

   (b) Using the result of (a), how would you create a $1-\alpha$ confidence interval for θ?

   (c) Find an exact $1-\alpha$ confidence interval for θ. Explain how to find any necessary critical constants.

4. In a plant breeding experiment the offspring may be of any of four possible genotypes, with probabilities $\theta_1, \theta_2, \theta_3, \theta_4$, respectively. A genetic model imposes the following hypothesis:

$$H_0 : \theta_1 = \frac{2+\eta}{4}, \theta_2 = \theta_3 = \frac{1-\eta}{4}, \theta_4 = \frac{\eta}{4},$$

where η is an unknown parameter between 0 and 1. In n trials, genotype I is observed N_i times, $i = 1,2,3,4$.

   (a) Assuming H_0 is true, derive the maximum likelihood estimator of η.

   (b) If n is large, propose a test of H_0 and state the distribution of your test statistic. It is not necessary to derive an explicit formula for your test statistic.

5. Let $X_1,...,X_n$ be i.i.d Poisson variables with mean μ. Assume that μ has a gamma prior density

$$\pi(\mu) = \frac{r^a \mu^{a-1} e^{r\mu}}{\Gamma(a)} I\{\mu > 0\}$$

   (a) Find the Bayes estimator of μ with respect to squared error loss $L(\mu, \hat{\mu}) = (\hat{\mu} - \mu)^2$.

   (b) Show that as $n \to \infty$, $\sqrt{n}(\hat{\mu} - \mu) \to 0$, where $\hat{\mu}$ is the Bayes estimator found in part (a).

6. Based on a single observation X, find the most powerful test of $H_0 : X \sim N(0,1)$ vs. $H_0 : X \sim (1/2)e^{-|x|}$. Give the form of the critical region and provide a formula for the power of your test.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
January, 2007

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let $(X_1, \ldots, X_k)$ be a multinomial vector based on n trials and probability vector $\pi = (\pi_1, \ldots, \pi_k)$. Assume that n is large and that one wants to test the null hypothesis H_0: $\pi_1 = \pi_2$ against the general alternative.

(a) Show that the test statistic

$$Q = \frac{(X_1 - X_2)^2}{X_1 + X_2}$$

can be expressed in the form of a χ^2 test. That is,

$$Q = \sum_{i=1}^{k} \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

for an appropriate choice of "observed" and "expected" frequencies.

(b) Show that the test which rejects H_0 when $Q > \chi^2_{1,\alpha}$ has approximate level α.

2. Let X have the distribution

$$P_\theta[X = x] = \left(\begin{array}{c} x - 1 \\ r - 1 \end{array} \right) \theta^r (1 - \theta)^{x-r}$$

(a) Find $T(X)$, the uniformly minimum variance unbiased estimator of θ based on one observation of X.

(b) Prove that $T(X)$ is asymptotically efficient as $r \to \infty$.

3. Let $X_1, X_2, \ldots, X_n$ be i.i.d. Poisson random variables with mean μ.

(a) Find a uniformly minimum variance unbiased estimator (UMVUE) for μ and for $P_\mu[X_1 = 0]$.

(b) Show that no unbiased estimator of $1/\mu$ exists.

4. Let $\mathbf{X}$ be a data set whose distribution depends on a real-valued parameter θ and let $T(\mathbf{X}, \theta)$ be a pivotal quantity of the form $T(\mathbf{X}, \theta) = T_1(\mathbf{X}) - \theta$. Assume $T(\mathbf{X}, \theta)$ has a unimodal density function.

(a) Explain how one can construct a confidence interval for θ with fixed width d and maximum possible confidence coefficient.

(b) Given i.i.d. observations $X_1, \ldots, X_n$, construct a confidence interval for the mean θ of a $N(\theta, 1)$ distribution having fixed width d and maximum possible confidence probability.

(c) Repeat the above in sampling from the shifted exponential density $f(x; \theta) = \exp[-(x - \theta)]I\{x > \theta\}$.

5. The simple hypothesis H_0: $\theta = \theta_0$ is to be tested against the simple alternative $H_1 : \theta = \theta_1$ on the basis of data $\mathbf{X}$ which has density $f(\mathbf{x}; \theta_i)$ under H_i. Let ϕ be a most powerful test of H_0 vs. H_1 of level α. Show that ϕ has power at least α. When does ϕ have power exactly equal to α?

6. Let $X_1, \ldots, X_n$ be i.i.d. exponential random variables with density

$$f(x|\theta) = \begin{cases} \theta \exp(-\theta x) & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that θ has an exponential prior density

$$\pi(\theta) = \begin{cases} b \exp(-b\theta) & \text{if } \theta > 0, \\ 0 & \text{otherwise.} \end{cases}$$

where b is a known constant.

(a) Calculate the Bayes estimators of θ and $1/\theta$ under squared error loss, $L(\theta, \hat{\theta}) = (\theta - \hat{\theta})^2$.

(b) Write an equation for the Bayes estimator of θ if the loss function is changed to absolute error, $L(\theta, \hat{\theta}) = |\theta - \hat{\theta}|$. Describe how you could determine the estimator using standard statistical tables.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
August, 2006

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let X be an integer valued random variable with a power series distribution

$$p(k; \theta) = P_\theta[X = k] = \frac{a_k}{f(\theta)} \theta^k$$

where θ is a positive parameter and a_k, $k = m, m+1, \ldots$, is an infinite sequence of strictly positive numbers. Let $X_1, \ldots, X_n$ be i.i.d. with distribution $p(k; \theta)$ and let $T = X_1 + \cdots + X_n$.

(a) Show that T is a complete sufficient statistic with distribution

$$P_\theta[T = t] = \frac{b_t}{f(\theta)^n} \theta^t, \quad t = nm, nm+1, \ldots.$$

(b) Show that the uniformly minimum variance unbiased estimator (UMVUE) of θ^r is

$$u_r(T) = [b_{T-r}/b_T] I\{T \geq r\}$$

and that a UMVUE of its variance is $u_r(T)^2 - u_{2r}(T)$.

1

2. Let $(X_1, \ldots, X_n)$ be a sample from a population with density

$$f(x; \theta) = (\theta - 1)/x^\theta, \; x \geq 1$$

with $\theta > 1$ as a parameter.

(a) Prove that the MLE $\hat{\theta}_n = \hat{\theta}_n(X_1, \ldots, X_n)$ is a consistent estimator of θ.

(b) Prove that $E_\theta(\hat{\theta}_n) > \theta$ for all $\theta > 1$. (*Hint*: Apply Jensen's inequality.)

(c) Find the limiting distribution of $\sqrt{n}(\hat{\theta}_n - \theta)$.

3. Let $(X_1, \ldots, X_n)$ be a sample from a population with density

$$f(x; \theta) = \theta x^{\theta-1}, \; x > 0$$

with $\theta > 0$ as a parameter. Show that

$$T(X_1, \ldots, X_n; \theta) = -2\theta \log \left(\prod_{i=1}^{n} X_i \right)$$

is a pivot and, based on this, construct a $1 - \alpha$ confidence interval for θ. State explicitly how to construct the confidence limits using standard statistical tables.

4. Let $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n)$ be independent samples from populations with densities

$$f(x; \theta_1) = (x/\theta_1^2)e^{-x/\theta_1}, \; x > 0 \text{ and } f(y; \theta_2) = (y/\theta_2^2)e^{-y/\theta_2}, \; y > 0,$$

respectively, where $\theta_1 > 0$, $\theta_2 > 0$ are parameters.

(a) Develop the likelihood ratio test of size α of $H_0 : \theta_1 = \theta_2 = \theta$ (unspecified) versus $H_1 : \theta_1 \neq \theta_2$.

(b) Show how to find the critical value for this test using standard statistical tables, and show that $P(\text{reject } H_0 | H_0)$ equals α for all θ.

(c) Write an expression for the power of the test against alternatives of the form $\theta_1 = 2\theta_2$.

2

5. Let $(X_1, \ldots, X_n)$ be a sample from a population with density

$$f(x; \theta) = C \exp\{-|x - \theta|^3\}$$

where $\theta \in R$ is a parameter and C is the normalizing constant (you do not have to calculate it).

(a) Find the method of moments estimator $\tilde{\theta}_n$ of θ and the asymptotic distribution of $\sqrt{n}(\tilde{\theta}_n - \theta)$ as $n \to \infty$.

(b) Calculate the asymptotic efficiency of $\tilde{\theta}_n$.

6. Let $X_1, \ldots, X_n, X_{n+1}$ be conditionally i.i.d. $N(0, \sigma^2)$, given a parameter μ, let μ have a $N(a, b^2)$ prior distribution, and assume σ is known.

(a) Find the Bayes estimator of μ based on $X_1, \ldots, X_n$ only, with respect to the squared error loss function $L(\mu, \hat{\mu}) = (\hat{\mu} - \mu)^2$.

(b) Find a predictor of X_{n+1} based on $X_1, \ldots, X_n$ which minimizes the mean squared error of prediction $E[(\hat{X}_{n+1} - X_{n+1})^2]$.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
JANUARY, 2006

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let $(X_1, \ldots, X_n)$ be a sample of an odd size $n = 2m+1$ from a population with pdf $f(x; \theta) = (1/2) \exp\{-|x - \theta|\}$ with $\theta \in R$ as a parameter.

(i) Find the maximum likelihood estimator (MLE) of θ and show that its distribution is symmetric around θ.

(ii) Prove that the variance of the MLE is finite and does not depend on θ. (You do not have to calculate the variance.)

1

2. Let $Y_1, \ldots, Y_n$ be independent gamma variables with means $\mu_i = \exp(\alpha + \beta x_i)$ where the x_i are nonrandom observed real numbers. The Y_i have a common known shape parameter r.

(a) Using the parameterization

$$f(y; \mu) = \frac{1}{\Gamma(r)} \left(\frac{ry}{\mu} \right)^r \exp\left(\frac{ry}{\mu} \right) \frac{1}{y}$$

show that

$$E[\log(Y_i)] = \Psi(r) - \log r + \log \mu_i = \Psi(r) - \log r + \alpha + \beta x_i,$$

where
$$\Psi(r) = (d/dr) \log \Gamma(r) = \Gamma'(r)/\Gamma(r).$$

Show also that $\mathrm{Var}[\log Y_i]$ does not depend on μ_i. [The actual variance is $\Psi'(r)$, but you do not have to prove this.]

(b) Suppose that β is estimated by ordinary least squares regression of $\log Y_i$ on x_i, resulting in the estimated regression coefficient $\hat{\beta}_{\mathrm{LS}}$. Assume that, as $n \to \infty$, $\sum (x_i - \bar{x})^2 \to \infty$. Show that $\hat{\beta}_{\mathrm{LS}}$ is consistent and has asymptotic efficiency $1/[r\Psi'(r)]$ with respect to the MLE $\hat{\beta}_{\mathrm{ML}}$.

3. Given θ, let $X_1, \ldots, X_n$ be i.i.d. with density $f(x|\theta)$ and suppose that θ is a real parameter with prior density $\pi(\theta)$. It is desired to estimate θ with respect to squared error loss: $L(a, \theta) = (a - \theta)^2$.

(a) Show that if an estimator $\tilde{\theta}(X_1, \ldots, X_n)$ satisfies

$$E(\tilde{\theta}|\theta) = \theta + b(\theta) \text{ with } b(\theta) > 0 \text{ for all } \theta,$$

then $\tilde{\theta}$ cannot be a Bayes estimator with respect to any prior distribution of θ.

(b) Prove that if the X_i are i.i.d. with a $N(0, 1/\theta)$ density, then the MLE of θ can not be a Bayes estimator.

4. Let X_i, $i = 1, \ldots, n$, be i.i.d. with $P_\theta[X_i = 0] = (1 - \theta)/2$, $P_\theta[X_i = 1] = 1/2$, and $P_\theta[X_i = 2] = \theta/2$.

   (a) Find a minimal sufficient statistic for θ. Is this statistic also complete?

   (b) Find $\hat{\theta}$, the maximum likelihood estimator of θ.

   (c) Show that $\hat{\theta}$ is consistent and derive the limiting distribution of a suitably normalized version of $\hat{\theta}$.

5. Let $X_1, \ldots, X_m$ be i.i.d. $N(\mu_1, \sigma_1^2)$ random variables and let $Y_1, \ldots, Y_n$ be i.i.d. $N(\mu_2, \sigma_2^2)$ random variables. Let $\bar{X}$ and $\bar{Y}$ denote the sample means and let s_X^2 and s_Y^2 denote the (unbiased) sample variances. Define the pooled variance by
$$s_p^2 = \frac{(m-1)s_X^2 + (n-1)s_Y^2}{m+n-2}.$$
It is well known that the interval
$$I = \left[\bar{X} - \bar{Y} - ts_p\sqrt{1/m + 1/n}, \bar{X} - \bar{Y} - ts_p\sqrt{1/m + 1/n} \right]$$
is a $1 - \alpha$ confidence interval for $\mu_1 - \mu_2$ if $t = t_{m+n-2, \alpha/2}$ and $\sigma_1 = \sigma_2$.

   (a) Let $m \to \infty$, $n \to \infty$ in such a way that $m/(m+n) \to \lambda \in (0, 1)$. Without assuming $\sigma_1 = \sigma_2$, derive the limiting distribution of
$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{s_p\sqrt{1/m + 1/n}}.$$

   (b) In the asymptotic setting of (b), show that generally I does not have asymptotic coverage probability $1 - \alpha$. However, if either $\sigma_1 = \sigma_2$ or if $\sigma_1 \neq \sigma_2$ and $\lambda = 1/2$, then I does indeed have asymptotic coverage probability $1 - \alpha$.

   (c) If $s_p\sqrt{1/m + 1/n}$ is replaced by $\sqrt{s_X^2/m + s_Y^2/n}$ in the definition of I, the resulting confidence interval has asymptotic coverage probability $1 - \alpha$ for any values of λ, σ_1 and σ_2.

3

6. Let $\{N(t), t \geq 0\}$ be a Poisson process with rate λ. Consider the following two schemes for testing $H_0 : \lambda \leq a$ with significance level α.

(i) Observe $N(t)$ over the time interval $[(0, \tau]$ and base a test on $X = N(\tau)$.

(ii) Observe $N(t)$ until the time T when the rth event occurs, and base a test on the random variable T.

Find most powerful tests of H_0 for both schemes and explain how the critical constants are to be computed. Show that the power of either test can be made greater than or equal to β by choosing either τ or r sufficiently large.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
AUGUST, 2005

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let $X_1, \ldots, X_n$ be i.i.d. random variables with density

$$f(x; \lambda) = \begin{cases} \lambda \exp(-\lambda x) & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Show that the likelihood ratio test of the null hypothesis H_0: $\lambda = 1$ rejects H_0 when either

$$\bar{X} < c_1 \quad \text{or} \quad \bar{X} > c_2.$$

What is the exact distribution of the test statistic under H_0?

(b) For a given significance level α, show that the critical constants c_1 and c_2 satisfy the system of equations

$$\begin{aligned} P[\bar{X} < c_1] + P[\bar{X} > c_2] &= \alpha \\ c_1 \exp(c_2) &= c_2 \exp(c_1). \end{aligned}$$

[It is not necessary to solve the system.]

2. Suppose that S is a sufficient statistic for a family of distributions $\mathcal{P} = \{P_\theta \mid \theta \in \Theta\}$.

(a) Let T be another statistic which is independent of S for all $\theta \in \Theta$. Prove that $P_\theta[T \in B]$ does not depend on θ. That is, for any $\theta_1 \neq \theta_2$,

$$P_{\theta_1}[T \in B] = P_{\theta_2}[T \in B].$$

(b) Suppose that both S and T are (separately) sufficient for $\mathcal{P}$. If S and T are independent, then $\mathcal{P}$ consists of a single distribution P_{θ_0}.

3. Let $X_1, \ldots, X_n$ be i.i.d. with density

$$f(x; \theta) = \frac{3\theta^3}{(\theta + x)^4}$$

for $0 < x < \infty$ and $\theta > 1$. Calculate the efficiency of the best linear unbiased estimator $c\bar{X}$.

4. Let $X_1, \ldots, X_n$ be a random sample with a common $N(\mu, \sigma^2)$ distribution.

(a) Find c_1 such that

$$\hat{\sigma}_1 = \frac{c_1}{n} \sum_{i=1}^{n} |X_i - \bar{X}|$$

is an unbiased estimator of σ.

(b) Let

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

and let c_2 be chosen so that $\hat{\sigma}_2 = c_2 s$ is also an unbiased estimator of σ. Compare the performances of $\hat{\sigma}_1$ and $\hat{\sigma}_2$.

5. Let X be the number of successes in n Bernoulli trials with success probability θ. It is desired to estimate p under the quadratic loss function

$$L(\hat{p}, p) = \frac{(\hat{\theta} - \theta)^2}{\theta(1 - \theta)}.$$

(a) Find $\hat{\theta}_\pi$, the Bayes estimator with respect to the prior density

$$\pi(\theta) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} \theta^{a-1}(1 - \theta)^{b-1}$$

where $a > 1$ and $b > 1$ are known constants.

(b) Let $\hat{\theta} = X/n$ be the usual maximum likelihood estimator of θ. Prove that

$$\sqrt{n}(\hat{\theta}_\pi - \hat{\theta}) \to 0$$

in probability as $n \to \infty$.

6. Let $X_1, \ldots, X_n$ be i.i.d. with a common uniform distribution on $[-\theta, \theta]$.

(a) Prove that $\hat{\theta}_n$, the maximum likelihood estimator of θ, is consistent.

(b) Find a sequence of normalizing constants a_n, $n = 1, 2, \ldots$, such that $a_n(\hat{\theta}_n - \theta)$ has a nondegenerate limiting distribution.

3

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
JANUARY, 2005

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let $\mathcal{P}_1 = \{P_\theta, \ \theta \in \Theta_1\}$ and $\mathcal{P}_2 = \{P_\theta, \ \theta \in \Theta_2\}$ be two families of probability distributions on the same measurable space, and let a statistic T be sufficient for $\mathcal{P}_1$ and (separately) for $\mathcal{P}_2$. Set $\mathcal{P} = \{P_\theta, \ \theta \in \Theta_1 \cup \Theta_2\}$.

(a) Prove that if $\Theta_1 \cap \Theta_2 \neq \emptyset$, then T is sufficient for $\mathcal{P}$.

(b) Prove or disprove that T is sufficient for $\mathcal{P}$ when $\Theta_1 \cap \Theta_2 = \emptyset$.

2. Let $(x_1, x_2, \ldots, x_n)$ be a sample from a population with density

$$f(x; \theta) = 2\theta^2/x^3, \quad x > \theta,$$

where $\theta \in (0, \infty)$ is an unknown parameter.

(a) Find $\tilde{\theta}_n$, the uniformly minimum variance unbiased estimator (UMVUE) of θ.

(b) Find the asymptotic distribution of $2n(\tilde{\theta}_n - \theta)$ as $n \to \infty$.

1

3. Given $\Lambda = \lambda$, the random variables $X_1, X_2, \ldots, X_n$ are independent and identically distributed with density

$$f(x; \lambda) = \lambda e^{-\lambda x}, \ x \geq 0.$$

The prior density of Λ is Gamma(a, b):

$$\pi(\lambda) = \frac{1}{b^a \Gamma(a)} \lambda^{a-1} e^{-\lambda/b}$$

where $a > 0$ and $b > 0$ are known constants.

(a) For $X_1 = x_1, \ldots, X_n = x_n$, find the posterior distribution of Λ.

(b) If λ_0, λ_1 are the prior mean and mode of Λ and λ^*, $\tilde{\lambda}$ are the posterior mean and mode of Λ, show that with probability 1,

$$|\lambda^* - \tilde{\lambda}| < |\lambda_0 - \lambda_1|.$$

4. Let $X_1, \ldots, X_n$ be a random sample from a population with uniform distribution on (θ_1, θ_2), where $\theta_1 < 0 < \theta_2$.

(a) Write the critical region of the likelihood ratio test for testing $H_0 : \theta_2 = -\theta_1$.

(b) Show that when all the observations are of the same sign, the likelihood ratio does not exceed $(1/2)^n$.

5. Let X_{ij}, $i = 1, \ldots, k$, $j = 1, \ldots, n$, be independent random variables with $X_{ij} \sim N(\mu_i, \sigma^2)$. Find the maximum likelihood estimator of the parameter vector $(\mu_1, \ldots, \mu_k, \sigma^2)$. Investigate the consistency of the MLE's as $n \to \infty$.

6. Let X and Y be independent exponential variables with densities

$$f_X(x) = \begin{cases} \lambda_X \exp(-x/\lambda_X) & \text{if } x > 0 \\ 0 & \text{if } x \leq 0, \end{cases}$$

$$f_Y(y) = \begin{cases} \lambda_Y \exp(-y/\lambda_Y) & \text{if } y > 0 \\ 0 & \text{if } y \leq 0. \end{cases}$$

(a) Define $\theta = \lambda_X/\lambda_Y$. Show that the distribution of $U = \theta X/Y$ does not depend on any unknown parameters; that is, U is a pivotal quantity. Use U to construct a confidence interval for θ.

(b) Show that $V = \lambda_X X + \lambda_Y Y$ is also a pivotal quantity. Use V to construct a two-dimensional confidence set for (λ_X, λ_Y).

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
AUGUST, 2004

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let (X, Y) be a random vector whose distribution is given by a density $f(x, y; \theta)$ depending on a parameter θ.
(i) Assuming X and Y independent, find a necessary and sufficient condition, on the marginal density of Y, for X to be sufficient for θ.
(ii) For general X, Y (not necessarily independent), find a necessary and sufficient condition on the marginal distributions of X and Y under which X is sufficient for θ and so is Y.

2. Let $(x_1, x_2, \ldots, x_n)$ be a sample from a population X with density $\lambda e^{-\lambda x}$, $x \geq 0$.
(i) Find the uniformly minimum variance unbiased estimator $\hat{g}_n = \hat{g}(x_1, \ldots, x_n)$ of the parameter function $g(\lambda) = 1/\lambda^2$.
(ii) For an appropriate sequence $\{c_n,\ n = 1, 2, \ldots\}$, find the nondegenerate limiting distribution of $c_n(\hat{g}_n - g(\lambda))$ as $n \to \infty$.

3. Given a parameter θ, $X_1, X_2, \ldots, X_n$ are independent normally distributed random variables with $E(X_i|\theta) = \alpha_i\theta$, $\text{var}(X_i|\theta) = 1$, $i = 1, \ldots, n$; $\alpha_1, \ldots, \alpha_n$ known. The prior distribution of θ is normal with mean θ_0 and variance σ_0^2. Find the Bayesian estimator (i. e., the one that minimizes the Bayesian risk) of θ with respect to the loss function $L(\tilde{\theta}, \theta) = |\tilde{\theta} - \theta|$.

4. Let $(x_1, x_2, \ldots, x_n)$ be a sample from a population with uniform distribution on $(\theta, 2\theta)$ with $\theta > 0$ as a parameter.
(i) Find the maximum likelihood estimator of θ and its distribution function.
(ii) Based on the MLE, construct a confidence interval for θ of level $1 - \alpha$.

5. A trial may result in three possible outcomes, A with probability p_1, B with probability p_2, and C with probability $1 - p_1 - p_2$, with (p_1, p_2) as a parameter. In n independent trials A was observed n_1 times and B was observed n_2 times.
(i) Show that the likelihood ratio test of the null hypothesis $H_0 : p_1 = p_2$ versus the alternative $H_1 : p_1 \neq p_2$ accepts H_0 if $n_1 = n_2 > 0$.
(ii) Find an approximate critical region for testing H_0 of asymptotic level 0.05 for large n.

6. Let $(x_1, y_1), \ldots, (x_n, y_n)$ be a sample from a bivariate normal population with

$$E(x_i) = \theta, \; E(y_i) = 0, \; \text{var}(x_i) = \text{var}(y_i) = 1, \; \text{corr}(x_i, y_i) = \rho, \; i = 1, \ldots, n$$

with θ as a parameter (and ρ known).
(i) Find the maximum likelihood estimator $\hat{\theta}_n$ of θ and calculate its efficiency.
(ii) For an appropriate sequence $\{a_n, \; n = 1, 2, \ldots\}$, find the nondegenerate limiting distribution of $a_n(\hat{\theta}_n - \theta)$ as $n \to \infty$.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
JANUARY, 2004

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let $\mathcal{P} = \{P_\theta\}$ be a family of distributions of a random element $X \in \mathcal{X} \subset R^n$ parametrized by a discrete parameter $\theta \in \{\theta_0, \theta_1, \theta_2\}$. Assuming P_θ is given by a density $p(x; \theta)$ which is everywhere positive on $\mathcal{X}$, develop the minimal sufficient statistic for θ.

2. Let $(x_1, x_2, \ldots, x_n)$ be a sample from a population X with density $e^{-(x-\theta)}$, $x \geq \theta$. Find the uniformly minimum variance unbiased estimator of the value of the parameter function $g(\theta) = P_\theta(1 < X < 2)$.

3. Given a parameter θ, $X_1, X_2, \ldots, X_n$ are independent identically distributed random variables with $E(X_i|\theta) = \theta$, $\text{var}(X_i|\theta) = \sigma^2$. The prior distribution of θ has known mean a and variance b^2. The linear Bayesian estimator of θ, by definition, minimizes the overall risk within the class of estimators of the form $c_0 + \sum_{i=1}^n c_i X_i$ where the constant coefficients $c_0, c_1, \ldots, c_n$

1

do not depend upon the data. Find the linear Bayesian estimator of θ with respect to the squared-error loss function.

4. Let $(x_1, x_2, \ldots, x_n)$ be a sample from a population with density

$$f(x; \theta) = \frac{1}{B(3,3)} (x - \theta)^2 (1 - x + \theta)^2, \ \theta \leq x \leq 1 + \theta.$$

with θ as a parameter. Here $B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1 - x)^{\beta-1} dx$ is the beta function, $B(\alpha, \beta) = (\alpha - 1)!(\beta - 1)!/(\alpha + \beta - 1)!$ for integer $\alpha > 0$. $\beta > 0$. Develop the method of moments estimator of θ and calculate its efficiency

5. Let $(x_1, x_2, \ldots, x_m)$ and $(y_1, y_2, \ldots, y_n)$ be two independent samples from normal populations with parameters (μ_1, σ_1^2) and (μ_2, σ_2^2), respectively. Assuming $\sigma_1^2 = 4\sigma^2$, $\sigma_2^2 = \sigma^2$, σ^2 unknown. develop a test of a constant level α for all $\sigma^2 > 0$ of the null hypothesis $H_0 : \mu_1 = \mu_2$ versus the alternative $H_1 : \mu_1 \neq \mu_2$.

6. Let $(x_1, y_1), \ldots, (x_n, y_n)$ be a sample from a bivariate normal population with mean vector (θ_1, θ_2). regarded as a parameter and covariance matrix I, the identity matrix. For the maximum likelihood estimator $\hat{g}_n$ of the value of the parameter function $g(\theta_1, \theta_2) = (\theta_1 - \theta_2)^2$, find the nondegenerate limiting distribution of $a_n(\hat{g}_n - g(\theta_1, \theta_2))$ as $n \to \infty$ for a properly chosen a_n.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
AUGUST, 2003

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. $X_1, ..., X_n$ are iid from the density $f(x, \theta)$, θ real valued, and consider the estimator $\hat{\theta}$ that minimizes $\sum_i \rho(x_i - \theta)$ for some fixed function ρ. Define the derivative $\psi = \rho'$, and assume that ρ is symmetric and ψ is strictly increasing and sufficiently smooth. Thus, the estimator is the solution to

$$\sum_{i=1}^{n} \psi(x_i - \theta) = 0.$$

If θ_0 is the true parameter and ρ is at least three times continuously differentiable in a neighborhood of θ_0, and if all of the random variables X_i are uniformly bounded (say with values in $[-1, 1]$), obtain the asymptotic distribution of $\sqrt{n}(\hat{\theta} - \theta_0)$. Assume $E_{\theta_0} \psi(X - \theta_0) = 0$.

1

2. $X_1, ..., X_n$ is a random sample from a density $f(x, \theta)$ with scalar θ.

a. Give the definition of a minimal sufficient statistic.

b. Suppose $X_1, ..., X_n$ are iid from $N(\theta, \theta)$. Find a minimal sufficient statistic.

c. Suppose T_1 is sufficient, T_2 is minimal sufficient, and U is unbiased for θ. Define $U_i = E(U|T_i)$, $i = 1, 2$. Show that $\text{Var}(U_2) \leq \text{Var}(U_1)$.

3. a. Mr. Z has a coin he suspects may not be balanced. Thus, flipping the coin independently n times and getting the 0-1 data $X_1, ..., X_n$ (1 being success) and the average $Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, he decided to study the long term behavior of $n[1/4 - Y_n(1 - Y_n)]$. To help Mr. Z decide whether the coin is fair, find whether $n[1/4 - Y_n(1 - Y_n)]$ as $n \to \infty$ has a proper limiting distribution and if so tell what it is.

b. Show that if $\sqrt{n}(W_n - \mu) \to N(0, \sigma^2)$ in distribution as $n \to \infty$, then $W_n \to \mu$ in probability.

4. Consider a multinomial experiment with m categories and cell probabilities $p_1, ..., p_m$ where $\sum_{i=1}^{m} p_i = 1$. Let $X_1, ..., X_m$ be the observed frequencies such that $\sum_{i=1}^{m} X_i = n$. Suppose the statistician is interested in testing the hypothesis H_0 that the cell probabilities depend on a k-dimensional parameter θ versus the alternative H_1 that the cell probabilities are free subject to the fact they are nonnegative and sum to 1. Denote by Λ the likelihood ratio test statistic.

a. Show that for suitable O_i and E_i we have

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

b. What is the asymptotic distribution of $-2 \log \Lambda$?

5. Let $X_1, ..., X_n$ be a random sample from the exponential distribution with density $f(x, \lambda) = \frac{1}{\lambda} \exp(-x/\lambda)$, $x > 0$.

a. Find the UMP test for testing $H_0 : \lambda = 7$ versus $H_1 : \lambda < 7$ at level α.

b. Find a $1 - \alpha$ confidence interval for λ by inverting the UMP test.

c. Find the expected length of the interval in part (b).

6. Let X be a Poisson random variable with parameter θ. Given a loss function $L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2/\theta$ and a prior density $\pi(\theta) = (\theta^2/2) \exp(-\theta)$, $\theta > 0$, find the Bayes estimator of θ.

2

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
JANUARY, 2003

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Suppose both $X_1, ..., X_m$ and $Y_1, ..., Y_n$ are random samples from the Exponential(1) distribution. With sample means $\overline{X}, \overline{Y}$, let

$$B_{m,n} = \frac{m\overline{X}}{m\overline{X} + n\overline{Y}}$$

where $m/(m+n) \to \alpha$ as $m, n \to \infty$.

a. What is the distribution of $B_{m,n}$?

b. For $0 < \alpha < 1$, derive the asymptotic distribution as $m, n \to \infty$ of

$$\frac{\sqrt{m+n}\left(B_{m,n} - \frac{m}{m+n}\right)}{\sqrt{\alpha(1-\alpha)}}$$

Explain your derivation carefully.

2. Two independent and identically distributed measurements X_1, X_2 are made from a parametric density with parameter $\theta > 0$

$$f(x, \vartheta) = \begin{cases} \frac{1}{2} \vartheta \, e^{-\vartheta x} & if \quad x \geq 0 \\ \frac{1}{2} \vartheta^{-1} e^{x/\vartheta} & if \quad x < 0 \end{cases}$$

(a) Find the MLE of ϑ in terms of X_1, X_2.

(b) Find the most powerful hypothesis test of $H_0 : \vartheta = 1$ versus the alternative $H_1 : \vartheta = 2$ of size 0.10, giving an equation to determine the rejection cutoff but not solving it explicitly. Justify that your test either is or is not UMP versus $H_A : \vartheta > 1$.

3. Let $X_1, ..., X_n$ be a random sample from a $N(\mu, 1)$ population, and let the prior pdf of μ be $N(0, 1)$. Assuming quadratic loss, obtain the Bayes estimator and the corresponding minimum Bayes risk.

4. Let $X_1, ..., X_n$ be iid Poisson(λ), and let $\overline{X}$ and S^2 be the sample mean and sample variance, respectively.

a. Show that $\overline{X}$ is the best unbiased estimator of λ.

b. Show that $E(S^2 \mid \overline{X}) = \overline{X}$.

c. Use (b) to demonstrate explicitly that $\text{Var}(S^2) > \text{Var}(\overline{X})$.

5. Let $X_1, ..., X_n$ be a random sample from a pdf $f(x, \theta)$ where θ is k dimensional, and suppose $\hat{\theta}$ is the maximum likelihood estimator of θ.

a. Under appropriate regularity conditions, show that

$$\sqrt{n}(\hat{\theta} - \theta_0) \to N(0, I_1^{-1}(\theta_0))$$

where θ_0 is the true parameter and $I_1(\theta_0)$ is the Fisher information matrix evaluated at θ_0.

b. Illustrate the result in terms of $X_1, ..., X_n$ from Poisson(λ) where λ is a scalar parameter.

6. Let $X_1, ..., X_n$ denote the times to failure of n pieces of equipment, where $X_1, ..., X_n$ are iid Exponential(λ). Consider the hypothesis

$$H_0 : 1/\lambda = \mu \leq \mu_0$$

a. Show that the test $\overline{X} \geq \mu_0 x(1-\alpha)/2n$, where $x(1-\alpha)$ is the $(1-\alpha)$th quantile of the χ^2_{2n} distribution, is a size α test.

b. Derive the power in terms of the χ^2_{2n} distribution.

c. Approximate the power by appealing to the central limit theorem, and draw the graph of the approximate power function.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
AUGUST, 2002

Statistics (MA Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

Standard notations: iid = independent identically distributed; I_A denotes the indicator of event or property A, equal to 1 if A holds and equal to 0 otherwise; for $\lambda > 0$, $Expon(\lambda)$ refers to the Exponential distribution with density $f(x) = \lambda e^{-\lambda x} I_{[x \geq 0]}$; $\sim$ means "is distributed as"; for a random sample of variables $X_1, \ldots, X_n$, the notation $\overline{X}$ denotes $n^{-1} \sum_{i=1}^{n} X_i$; and $\arg\max_a h(a)$ denotes the (or any) value of a which maximizes the value $h(a)$. Moreover, as needed, use the normal percentage-point notation $z_\alpha \equiv \Phi^{-1}(1 - \alpha)$ for the $(1 - \alpha)$ quantile of the standard normal distribution.

---

(1). Suppose that the numbers of sightings X_i of each of 100 species of birds in a single season in a certain locale are independent random variables with probability distribution $Poisson(\lambda_i)$, where the unknown parameters λ_i are unobserved independent random variables distributed $Expon(\vartheta)$ (with density $g(\lambda) = \vartheta^{-1} e^{-\lambda/\vartheta}$) for some unknown positive parameter ϑ.

(a). Find the likelihood for the unknown parameter ϑ in terms only of the sample $X_1, X_2, \ldots, X_{100}$ (and *not* in terms of unobserved random variables).

(b). Given that $X_1 = 7$, $\vartheta = 6$, and $\overline{X} = 6.5$, how would you estimate λ_1 ? Will the estimator you provide, when expressed as a function of the sample $(X_1, \ldots, X_n)$, be consistent for λ_1 when n is large ?

(2). Suppose that the random variables in a sample $Y_1, Y_2, \ldots, Y_n$ are *iid* with values in $[0, 1]$, and that an investigator knows that the underlying probability density $f_Y(y)$ has the form

$$f_Y(y) = f_Y(y, a, b) \equiv \begin{cases} a & \text{for } 0 \leq y < 1 \\ b & \text{for } 1 \leq y < 2 \\ 1 - a - b & \text{for } 2 < y \leq 3 \end{cases}$$

for some $a, b > 0$, $a + b < 1$, but thinks that $a = 2b$.

(a). Find the exact likelihood ratio statistic for testing the hypothesis $H_0 : a = 2b$ versus the general alternative, and give the approximate cutoff for large n , for a size-α test.

(b). Find a good size-0.05 test for the hypothesis $H_0' : a = 0.4$ versus $H_0' : a \neq 0.4$ based upon the maximum likelihood estimator of a. Use the limiting distribution for large n of your estimator to find the approximate rejection region of your test.

(3). Suppose that the random pairs (X_i, Y_i) are *iid* for $i = 1, \ldots, n$, with
$$X_i, \sim \mathcal{N}(\mu, \sigma^2) \quad , \qquad Y_i \sim Expon(\lambda e^{X_i}) \quad \text{given } X_i$$

(a). Show that these variables form an exponential family, and find a sufficient statistic.

(b). Find the asymptotic variance for large n of the maximum likelihood estimator for $E(Y_1)$ (regarded as a function of the three unknown parameters μ, σ^2, λ). **But do not find the MLE's themselves.**

(4) Let the sample $X_1, \ldots, X_n$ be distributed as $\mathcal{N}(\mu, 1)$, where the parameter μ has prior probability mass function $\pi_k = P(\mu = k)$ respectively equal to $0.2, 0.5, 0.3$ for $k = 3, 4, 5$.

(a). Find the posterior probabilities of the three possible values $\mu = 3,\ 4,\ 5$ given that $\overline{X} = 3.7$ for $n = 10$.

(b). Let a decision rule δ be defined in any way whatever with values $\delta(z) = 3,\ 4,$ or 5 as a function of $\overline{X} = z$, and let the loss-function value $L(a, \mu)$ for the selected value a when the correct value is μ be defined as $L(a, \mu) = I_{[\mu \neq a]}$. Find a univariate integral expression, in terms of the function $\delta(\cdot)$, for the expected loss (Bayesian risk) of the procedure which selects value $a = \delta(\overline{X})$.

(5) Suppose that *iid* observed data values Y_i for $i = 1,\ 2,\ \ldots,\ n$ have the density
$$f(y) \ = \ \lambda\, e^{-\lambda(y-\mu)}\, I_{[y > \mu]}, \quad y > 0$$
where $\lambda > 0$, $\mu > 0$ are unknown parameters.

(a). Find minimal sufficient statistics for (μ, λ). These are complete for $n \geq 2$. Prove this for $n = 2$.

(b). Suppose that you observe a sample of $n = 2$ variables Y_i, and that you are told that $\mu < 10$. Find and justify a UMVUE for $e^{-\lambda(10-\mu)}$.

(6) Suppose that *iid* random pairs $(X_i,\ Y_i)$ are such that
$$X_i \sim \mathcal{N}(\mu, \sigma_x^2) \quad , \qquad Y_i \sim \mathcal{N}(a + bX_i,\ \sigma_e^2)$$

(a). Find the joint asymptotic distribution as $n \to \infty$ of
$$\sqrt{n}\left(\overline{X} - \mu,\ \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)\,Y_i\right)$$

(b). Conditionally given $\overline{X} = \mu + \sigma_x$, find the upper quartile (which means the same as the 75th percentile or 0.75 quantile) of $\overline{Y}$.

3

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
AUGUST, 2002

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

Standard notations: *iid* = independent identically distributed; I_A denotes the indicator of event or property A, equal to 1 if A holds and equal to 0 otherwise; for $\lambda > 0$, *Expon*(λ) refers to the Exponential distribution with density $f(x) = \lambda e^{-\lambda x} I_{[x \geq 0]}$; $\sim$ means "is distributed as"; for a random sample of variables $X_1, \ldots, X_n$, the notation $\overline{X}$ denotes $n^{-1} \sum_{i=1}^{n} X_i$; and $\arg\max_a h(a)$ denotes the (or any) value of a which maximizes the value $h(a)$. Moreover, as needed, use the normal percentage-point notation $z_\alpha \equiv \Phi^{-1}(1 - \alpha)$ for the $(1 - \alpha)$ quantile of the standard normal distribution.

---

(1). Suppose that the numbers of sightings X_i of each of 100 species of birds in a single season in a certain locale are independent random variables with probability distribution *Poisson*(λ_i), where the unknown parameters λ_i are unobserved independent random variables distributed *Expon*(ϑ) (with density $g(\lambda) = \vartheta^{-1} e^{-\lambda/\vartheta}$) for some unknown positive parameter ϑ.

1

(a). Find the likelihood for the unknown parameter ϑ in terms only of the sample $X_1, X_2, \ldots, X_{100}$ (and *not* in terms of unobserved random variables).

(b). Find a 95% one-sided confidence interval (e.g., the test-based interval related to the relevant optimal one-sided test) of the form $[0, U]$ for ϑ based on the sample $X_1, X_2, \ldots, X_n$ in terms of a sufficient statistic. Express U in terms of the quantiles of a known distribution depending on no unkown parameters.

(2). Suppose that the random variables in a sample $Y_1, Y_2, \ldots, Y_n$ are *iid* with values in $[0, 1]$, and that an investigator knows that the underlying probability density $f_Y(y)$ has the form

$$ f_Y(y) \;=\; f_Y(y, a, b) \;\equiv\; \begin{cases} a & \text{for } 0 \le y < 1 \\ b & \text{for } 1 \le y < 2 \\ 1 - a - b & \text{for } 2 < y \le 3 \end{cases} $$

for some $a, b > 0$, $a + b < 1$, but thinks that $a = 2b$.

(a). Find the exact likelihood ratio statistic for testing the hypothesis $H_0 : a = 2b$ versus the general alternative, and give the approximate cutoff for large n, for a size-α test.

(b). Find a good size-0.05 test for the hypothesis $H_0' : a = 0.4$ versus $H_0' : a \ne 0.4$ based upon the maximum likelihood estimator of a. Use the limiting distribution for large n of your estimator to find the approximate rejection region of your test.

(3). Suppose that the random pairs (X_i, Y_i) are *iid* for $i = 1, \ldots, n$, with

$$ X_i \sim \mathcal{N}(\mu, \sigma^2) \;, \qquad Y_i \sim Expon(\lambda\, e^{X_i}) \;\; \text{given } X_i $$

(a). Show that these variables form an exponential family, and find a minimal sufficient statistic.

(b). Find the asymptotic variance for large n of the maximum likelihood estimator for $E(Y_1)$ (regarded as a function of the three unknown parameters μ, σ^2, λ). **But do not find the MLE's themselves.**

(4) Let the sample $X_1, \ldots, X_n$ be distributed as $\mathcal{N}(\mu, 1)$, where the parameter μ has prior probability mass function $\pi_k = P(\mu = k)$ respectively equal to $0.2, 0.5, 0.3$ for $k = 3, 4, 5$.

(a). Find the posterior probabilities of the three possible values $\mu = 3, 4, 5$ given that $\overline{X} = 3.7$ for $n = 10$.

(b). Let a decision rule δ be defined in any way whatever with values $\delta(z) = 3, 4,$ or 5 as a function of $\overline{X} = z$, and let the loss-function value $L(a, \mu)$ for the selected value a when the correct value is μ be defined as $L(a, \mu) = I_{[\mu \neq a]}$. Find a univariate integral expression, in terms of the function $\delta(\cdot)$, for the expected loss (Bayesian risk) of the procedure which selects value $a = \delta(\overline{X})$.

(c). Use (b) to show that the essentially unique Bayes (minimum-risk) decision rule is defined by

$$\delta(\overline{X}) = \arg\max_a \left(\pi_a \exp(n\overline{X} a - na^2/2) \right)$$

(5) Suppose that *iid* observed data values Y_i for $i = 1, 2, \ldots, n$ have the density

$$f(y) = \lambda e^{-\lambda(y-\mu)} I_{[y > \mu]}, \quad y > 0$$

where $\lambda > 0$, $\mu > 0$ are unknown parameters.

(a). Find minimal sufficient statistics for (μ, λ). These are complete for $n \geq 2$. Prove this for $n = 2$.

(b). Suppose that you observe a sample of $n = 2$ variables Y_i, and that you are told that $\mu < 10$. Find and justify a UMVUE for $e^{-\lambda(10-\mu)}$.

(6) Suppose that *iid* random pairs (X_i, Y_i) are such that

$$X_i \sim \mathcal{N}(\mu, \sigma_x^2) \quad , \qquad Y_i \sim \mathcal{N}(a + bX_i, \sigma_e^2)$$

(a). Find the joint asymptotic distributions as $n \to \infty$ of each of

$$\sqrt{n}\left(\overline{X} - \mu, \ \overline{Y} - a - b\mu\right) \qquad \text{and} \qquad \sqrt{n}\left(\overline{X} - \mu, \ \frac{1}{n}\sum_{i=1}^n (X_i - \mu) Y_i\right)$$

(b). Conditionally given $\overline{X} = \mu + \sigma_x$, find the upper quartile (which means the same as the 75th percentile or 0.75 quantile) of $\overline{Y}$.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
JANUARY, 2002

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

Standard notations: r.v. = random variable; iid = independent identically distributed; $[x]$ = greatest integer less than or equal to x ; $\sim$ means "is distributed as"; and I_A denotes the indicator of event or property A, equal to 1 if A holds and equal to 0 otherwise.

---

(1). Suppose that for $i = 1, 2, \ldots, N$ and $j = 1, \ldots, n$, the r.v.'s $Y_{ij} \sim \mathcal{N}(\vartheta_i, \sigma^2)$ are mutually independent., where the parameters $\{\vartheta_i\}_{i=1}^{N}$ and σ^2 are unknown.

(a). For arbitrary $n \geq 2$ and $N \geq 2$, find the maximum likelihood estimators of ϑ_i, $i = 1, 2, \ldots, N$, and σ^2.

(b) If $n = 2$ and N gets large, show that the MLE $\hat{\sigma}^2$ converges in probability but *is not consistent*.

(c). Is the MLE $\hat{\sigma}^2$ for σ^2 consistent if $n = 1 + [\log(N)]$ and $N \to \infty$?

1

(2). Assume that an unknown positive scalar parameter ϑ follows the density

$$f(t) \;=\; c^{-a-1}\,(a+1)\,t^{a}\,I_{[0\le t\le c]}$$

for positive (known) parameters a, c. Assume also that given ϑ, a random sample X_i, $i = 1, 2, \ldots, n$, is drawn from a $Uniform[0, 1/\vartheta]$ density. Find the Bayes estimator for ϑ based on the observed sample $\{X_i\}_{i=1}^{n}$:

(a) under the squared-error loss function $L(\vartheta, a) = (\vartheta - a)^2$; and

(b) under the absolute-error loss function $L(\vartheta, a) = |\vartheta - a|$.

(3). A sample of random iid normally distributed 2-dimensional vectors (X_i, Y_i), $i = 1, 2, \ldots, n$, is assumed to follow the $\mathcal{N}\left(\begin{pmatrix} \mu \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 1/3 \\ 1/3 & 1 \end{pmatrix} \right)$ distribution. Find the UMVUE for μ based on these data, and justify this property for the estimator you give.

(4). The data-sample $V_1, V_2, \ldots, V_n$ is assumed to follow the "truncated exponential" density $f(v, a, b) = b\,e^{-b(v-a)}\,I_{[v\ge a]}$. Find the best critical region of size $\alpha \in (0.1)$ for testing

$$\mathbf{H_0} : \;\; a = \mu,\, b = \lambda \quad \text{versus} \quad \mathbf{H_1} : \;\; a = \mu',\, b = \lambda'$$

where $\mu' < \mu$ and $0 < \lambda < \lambda'$ are constants. Explain, with justification, whether or not this critical region is Uniformly Most Powerful for testing $\mathbf{H_0}$ versus the composite alternative $\mathbf{H_A} : a < \mu,\, b > \lambda$.

(5). Let $X_1, X_2, \ldots, X_n$ be a sample of iid r.v.'s with the common density $f(x, \lambda) = 2\lambda x\,e^{-\lambda x^2}\,I_{[x>0]}$.

(a) Find the Maximum Likelihood Estimator $\hat{\lambda}$ for λ, as a function of the sample, and find the limiting distribution for large n of $\sqrt{n}\,(\hat{\lambda} - \lambda)$.

(b) Find the method-of-moments estimator $\tilde{\lambda}$ for λ, and find its asymptotic relative efficiency (ARE) with respect to the maximum Likelihood Estimator.

(6). Independent exponential r.v.'s X_i, Y_i for $i = 1, 2, \ldots, n$ are defined so that $X_i \sim Expon(2\lambda)$, $Y_i \sim Expon(3\lambda^2)$, where $\lambda > 0$ is an unknown parameter.

(a) Show that the parametric joint density for $\{(X_i, Y_i)\}_{i=1}^{n}$ is of exponential-family form, and give (with justification) a minimal sufficient statistic for it.

(b) Show that $S = Y_1/X_1^2$ has distribution not depending upon λ, but that S is not independent of the minimal sufficient statistic for λ.

(c) Use either the facts proved in (b) *or another method* to establish that there is no complete sufficient statistic in (a).

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
AUGUST, 2001

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let $X_1, \ldots, X_n$ be a random sample from a population with probability density function (pdf)

$$f(x; \theta) = \begin{cases} \theta f_1(x), & \text{if} \quad x < 0 \\ (1 - \theta) f_2(x), & \text{if} \quad x \geq 0 \end{cases}$$

where $f_1 \geq 0$, $f_2 \geq 0$, θ is a parameter such that $0 < \theta < 1$, and

$$\int_{-\infty}^{0} f_1(x) dx = \int_{0}^{\infty} f_2(x) dx = 1.$$

Prove or disprove that there exists a complete sufficient statistic for θ.

2. Let $X_1, \ldots, X_n$ be a random sample of size n from a normal population $N(\mu, \sigma^2)$, $\mu \in \mathcal{R}$, $\sigma^2 \in \mathcal{R}_+$.

(i) Prove that if $n > 5$, the statistic

$$a_n \frac{\bar{X}^2}{S^2} - \frac{1}{n}$$

for some a_n (you do not need to compute a_n explicitly) is a minimum variance unbiased estimator of μ^2/σ^2.

(ii) What happens if $2 \leq n \leq 5$?

3. Let $X_1, \ldots, X_n$ be independent observations with

$$P_\theta(X_i < x) = \begin{cases} 1 - e^{-\lambda_i(x-\theta)}, & \text{if} \quad x \geq \theta \\ 0, & \text{if} \quad x < \theta \end{cases}$$

where $\lambda_1, \lambda_2, \ldots$ are given positive numbers and $\theta \in \mathcal{R}$ is a parameter. Find necessary and sufficient conditions on $\lambda_1, \lambda_2, \ldots$ for the consistency of $\hat{\theta}(X_1, \ldots, X_n)$, the minimum variance unbiased estimator of θ based on the first n observations.

4. Given θ, $X_1, \ldots, X_n$ are independent identically distributed random variables uniformly distributed in $(0, \theta)$. The prior distribution of θ is Pareto with density

$$\pi(\theta) = \frac{\alpha - 1}{\theta^\alpha}, \quad \alpha, \theta > 1$$

(i) Prove that the posterior density $\pi(\theta \mid X_1, \ldots, X_n)$ depends only on $X_{(n)} = \max(X_1, \ldots, X_n)$. That is, $\pi(\theta \mid X_1, \ldots, X_n) = \pi(\theta \mid X_{(n)})$.

(ii) Prove that $\pi(\theta \mid X_{(n)} = u)$ does not depend on u for $u < 1$.

5. Let $X_1, \ldots, X_n$ be independent copies of a random variable X, and let $\phi(u), \psi(v)$ be functions such that $E[\phi(X)] = E[\psi(X)] = 0$, $E|\phi(X)|^2 = a_{11} < \infty$, $E|\psi(X)|^2 = a_{22} < \infty$, $E[\phi(X)\psi(X)] = a_{12} = a_{21}$. Assume the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

is positive definite. Define $\bar{\phi}_n = \frac{1}{n} \sum_{i=1}^n \phi(X_i)$, $\bar{\psi}_n = \frac{1}{n} \sum_{i=1}^n \psi(X_i)$,

$$A^{-1} = \begin{pmatrix} a^{11} & a^{12} \\ a^{21} & a^{22} \end{pmatrix}.$$

Prove that as $n \to \infty$ we have the convergence in distribution,

$$n[a^{11}\bar{\phi}_n^2 + 2a^{12}\bar{\phi}_n\bar{\psi}_n + a^{22}\bar{\psi}_n^2] \longrightarrow \chi_2^2$$

6. Let $(X_1', ..., X_{n_1}')$, $(X_1'', ..., X_{n_2}'')$ be two independent random samples from normal populations $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$, respectively, with unknown σ^2. Develop the likelihood ratio test (i.e. obtain its critical region) for testing the null hypothesis $H_0 : \mu_2 = 2\mu_1$ (equivalently, $\mu_1 = \mu$, $\mu_2 = 2\mu$, μ unspecified.)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
JANUARY, 2001

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ be a family of distributions on a measurable space $(\mathcal{X}, \mathcal{A})$, and let $T : (\mathcal{X}, \mathcal{A}) \to (\mathcal{T}, \mathcal{B})$ be a *statistic*. Let $Q_\theta(B) = P_\theta(T^{-1}B)$, $B \in \mathcal{B}$, $\mathcal{Q} = \{Q_\theta, \theta \in \Theta\}$.

Prove that if T is sufficient for $\mathcal{P}$ and $S : (\mathcal{T}, \mathcal{B}) \to (\mathcal{S}, \mathcal{C})$ is sufficient for $\mathcal{Q}$ then S is sufficient for $\mathcal{P}$.

2. Let $X_1, \ldots, X_n$ be a random sample from a population with pdf

$$f(x - \theta) = \frac{1}{3\sqrt{2}}(x - \theta)^4 \exp\{-(x - \theta)^2/2\}, \quad \theta \in R$$

Compute the Fisher information on θ contained in the sample and find the efficiency of $\bar{X} = (X_1 + \cdots + X_n)/n$ as an estimator for θ.

3. Two identical coins with $P(H) = 1 - P(T) = p$ are tossed independently n times. Let n_0, n_1, n_2 denote the number of times when both coins show T, one shows T and the other shows H, both coins show H, respectively.

Find whether n_2/n is an admissible estimator for p^2 assuming quadratic loss.

4. Let $X_1, \ldots, X_n$ be of the form

$$X_i = \theta a_i + \epsilon_i, \quad i = 1, \ldots, n$$

where $a_1, \ldots, a_n$ are known constants, θ is a parameter to be estimated, and $\epsilon_1, \ldots, \epsilon_n$ are independent random variables with mean 0 and variance $\text{Var}(\epsilon_i) = \sigma_i^2 < \infty$.

Find the minimum variance linear unbiased estimator of θ and compute its variance.

Assuming $\epsilon_1, \epsilon_2, \ldots$ are normally distributed, study the consistency of the aforementioned estimator as $n \to \infty$.

5. Let the prior distribution of a parameter θ be Beta(α, β) with pdf,

$$\pi(\theta; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1}(1 - \theta)^{\beta-1}, \quad \theta \in (0.1)$$

where $\alpha > 0, \beta > 0$ and Γ is the gamma function. Given θ, the observations $X_1, \ldots, X_n$ are independent binary random variables with

$$P(X_i = 1|\theta) = 1 - P(X_i = 0|\theta) = \theta$$

(i) Find the Bayes estimator of θ assuming quadratic loss. (ii) Are $X_1, \ldots, X_n$ independent ?

6. Let $(X'_1, \ldots, X'_{n_1})$, $(X''_1, \ldots, X''_{n_2})$ be two independent random samples from normal populations $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$, respectively, with unknown σ^2. Develop a t-test at significance level α for testing the null hypothesis $H_0 : \mu_2 = 2\mu_1$.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
AUGUST 2000

Statistics (Ph.D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

---

1. Let X_j, $j = 1, 2, \ldots, n$, be a sample of n independent normal random variables where X_j has a $N(\mu, \sigma_j^2)$ distribution and the σ_j^2 are known but not necessarily equal. Find the best linear unbiased estimate of μ.

2. Let X_j, $j = 1, 2, \ldots, n$, be i.i.d. random variables having density

$$f(x; \theta) = \exp[-(x - \theta)] \quad \text{for } x > \theta.$$

Let $T = \min_{1 \le j \le n}(X_j)$ and denote the density of T by $g(t, \theta)$.

(a) Show that T is complete and sufficient for the family of of densities $\{g(t; \theta) \mid \theta \in (-\infty, \infty)\}$.

(b) Find the UMVU estimate for θ.

1

3. Suppose X has a binomial distribution with n trials and probability θ of success at each trial, where $0 < \theta < 1$.

(a) Find the Bayes estimator for θ for the uniform prior density with respect to the loss function the loss function

$$W(\theta, \delta) = \frac{(\theta - \delta)^2}{\theta(1 - \theta)}.$$

(b) Find the minimax estimate of θ with respect to the loss function $W(\theta, \delta)$ given in (a).

4. Let X be a single observation from the Weibull density

$$f(x; \theta) = \begin{cases} \theta x^{\theta-1} \exp(-x^\theta) & \text{if } x > 0 \\ 0 & \text{if } x \leq 0. \end{cases}$$

The parameter θ is positive.

(a) Show that the most powerful size α test of $H_0 : \theta = 1$ vs. $H_1 : \theta = 2$ rejects iff $k_1 < X < k_2$. Indicate how k_1 and k_2 are computed.

(b) Generalize the result of (a) to find the most powerful test of H_0 against a simple alternative of the form $H_1 : \theta = c_1$, where $c_1 > 1$, and show that no uniformly most powerful test of $H_0 : \theta = 1$ vs. $H_1 : \theta > 1$ exists.

(c) Show that the family of Weibull densities $\{f(x; \theta) \mid \theta > 0\}$ does not have the monotone likelihood ratio property with respect to X.

5. Let $U_1, \ldots, U_n$ be i.i.d. uniform random variables on the interval (a, b). Let $U_{(1)} = \min\{U_1, \ldots, U_n\}$ and let $U_{(n)} = \max\{U_1, \ldots, U_n\}$.

(a) Derive the joint distribution of $U_{(1)}$ and $U_{(n)}$.

(b) Show that the random variables $n(U_{(1)} - a)$ and $n(b - U_{(n)})$ have a limiting joint distribution and conclude that they are asymptotically independent.

6. Let X_i, $i = 1, \ldots, n$, be i.i.d. exponential random variables with parameter λ and let Y_i, $i = 1, \ldots, n$, be i.i.d. exponential random variables with parameter $\rho\lambda$. The X_i and Y_i are mutually independent, λ and ρ are unknown positive parameters, and the exponential parameterization is such that $E(X_1) = 1/\lambda$.

(a) Find a simplified form for the rejection region in the likelihood ratio test of the null hypothesis $H_0 : \rho = 1$ versus the general (two-sided) alternative.

(b) Give the exact distribution of the test statistic used to define the rejection region in (a).

(c) Give an approximate rejection region for the hypothesis test in (a) at level $\alpha = 0.10$ when n is large, where the cutoff point(s) are specified in terms of percentage points for the standard normal distribution.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
JANUARY, 2000

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number
    and your code number (NOT YOUR NAME) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it
    is your responsibility to make clear which theorem you are using and
    to justify its use.

---

1. Let $X_1, \ldots, X_m, Y_1, \ldots, Y_n$ be independent Poisson variables. The X_i
have common mean μ_x and the Y_j have common mean μ_y. It is proposed to
test $H_0 : \mu_x = \mu_y$ vs. $H_1 : \mu_x \neq \mu_y$ by rejecting H_0 when the statistic

$$ T = \frac{(\bar{X} - \bar{Y})^2}{(1/m + 1/n)(m\bar{X} + n\bar{Y})/(m + n)} $$

is large.

(a) Suggest an intuitive justification for the use of T.

(b) Find the approximate distribution of T under H_0.

(c) Assume $m = n$. As $n \to \infty$, give an approximation to the power of the
    test against the simple alternative $H_1^* : \mu_x = \mu_{x0}, \quad \mu_y = \mu_{y0}$, where
    $\mu_{x0} > \mu_{y0}$.

1

2. Let $X_1, \ldots, X_n$ have common mean μ and common variance σ^2, and let $\text{Cov}(X_i, X_j) = \sigma^2 \rho(|i - j|)$. As usual, $\bar{X}_n = (1/n) \sum_{i=1}^{n} X_i$.

(a) If $\rho(k) = \rho > 0$ for all k, then $\bar{X}_n$ is not a consistent estimator of μ.

(b) If $|\rho(k)| \leq M\gamma^k$, where $0 < \gamma < 1$, then $\bar{X}_n$ is a consistent estimator of μ.

3. Let $Y_1, \ldots, Y_n$ be jointly normal with $E[Y_j] = \theta$ and $\text{Cov}(Y_i, Y_j) = v_{ij}$. The parameter θ is unknown and the covariance matrix V is known and positive definite. Find the MLE of θ and decide whether it is unbiased and efficient.

4. Let $X_1, \ldots, X_n$ be i.i.d. with density $f(x)$. Show that, conditional on $X_{(j)} = a$, the $j - 1$ observations to the left of a and the $n - j$ observations to the right of a are distributed as $j - 1$ i.i.d. random variables with density $f(x)/F(a)$ and $n - j$ i.i.d. random variables with density $f(x)/[1 - F(a)]$, respectively, with the two sets being (conditionally) independent of each other.

5. Given the parameter θ, an observable random variable X has the negative binomial distribution:

$$P[X = k|\theta] = \frac{\Gamma(r + k)}{k!\Gamma(r)} \theta^r (1 - \theta)^k$$

where $k = 0, 1, 2, \ldots$, $0 < \theta < 1$ and r is a known constant. In addition, θ has a beta prior density given by:

$$p(\theta) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1 - \theta)^{b-1}.$$

Assuming squared error loss, find the Bayes estimators of each of the following parameters: θ, $\eta = 1/\theta$ and $\mu = E_\theta[X]$.

6. Let the discrete random vector X belong to an exponential family with joint probability function

$$P(\mathbf{X} = \mathbf{x}; \theta) = C(\theta)h(\mathbf{x}) \exp[\theta_j \sum_{j=1}^{k} T_j(\mathbf{x})]$$

depending on the k-dimensional parameter $\theta \in \Theta$, the natural parameter space.

(a) Find the distribution of

$$\mathbf{T} = [T_1(\mathbf{X}), \dots, T_k(\mathbf{X})]$$

and verify that it also belongs to an exponential family.

(b) Show that the marginal distribution of $\mathbf{T}_r = [T_1(\mathbf{X}), \dots, T_r(\mathbf{X})]$ belongs to an exponential family and find the natural parameter space and the normalization constant.

(c) Show that the conditional density of $\mathbf{T}_r$ given $[T_{r+1}(\mathbf{X}), \dots, T_k(\mathbf{X})] = [t_{r+1}, \dots, t_k]$ also belongs to an exponential family.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
AUGUST, 1999

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and justify its use.

---

1. Let $X_1, \ldots, X_n$ be a random sample from the uniform distribution on $(-\theta, \theta)$, where θ is an unknown positive parameter. Find the maximum likelihood estimator of θ and prove that it is consistent.

2. Let $X_1, \ldots, X_m, Y_1, \ldots, Y_n$ be independent normal random variables, let $E(X_i) = \mu_1, \operatorname{Var} X_i = \sigma_1^2, \quad i = 1, \ldots, m$, let $E(Y_j) = \mu_2, \operatorname{Var} Y_j = \sigma_2^2, \quad j = 1, \ldots, n$, and assume that both m and n are large.

(a) Find a confidence interval for $\delta = \mu_1 - \mu_2$, with approximate coverage probability $1 - \alpha$, assuming that the variances are unknown but equal.

(b) Find an approximate $1 - \alpha$ confidence interval for $\mu_1 - \mu_2$, assuming that the variances are unknown and possibly unequal.

(c) In general, does the interval of part (a) still contain δ with approximate probability $1 - \alpha$ if $m = n$ but $\sigma_1 \neq \sigma_2$?

(d) What is the approximate coverage probability of the interval of (a) if $\sigma_1^2/\sigma_2^2 = \rho$ and $m/n = R$?

3. Let X and Y be independent random variables with distributions $F(x|\xi)$ and $G(y|\eta)$, respectively, where ξ and η are real-valued parameters. Suppose that the prior distribution of (ξ, η) is such that ξ and η are independent with distributions P and Q. With respect to squared error loss, let $T_1(X)$ be the Bayes estimator of ξ based on X and let $T_2(Y)$ be the Bayes estimator of η based on Y.

(a) Show that, with respect to squared error loss, $T_1(X) - T_2(Y)$ is the Bayes estimator of $\xi - \eta$ based on (X, Y).

(b) Suppose that $\eta > 0$ and that with respect to squared error loss, $T_3(Y)$ is the Bayes estimator of $1/\eta$ based on Y. Show that, with respect to squared error loss, $T_1(X)T_3(Y)$ is the Bayes estimator of ξ/η based on (X, Y).

4. Let $X_1, \ldots, X_n$ be i.i.d exponential random variables with mean θ.

(a) Find a test of $H_0 : \theta = \theta_0$ vs. $H_0 : \theta > \theta_0$. Show how to calculate its critical value(s) using common statistical tables.

(b) What optimality properties, if any, does your test possess?

(c) How would you determine the required sample size if it is desired to test $H_0 : \theta = 1$ with $P[\text{Type I error}] = \alpha$ and power β against the alternative $\theta = 2$?

5. Let $X_1, \ldots, X_n$ be i.i.d. Poisson random variables with mean θ.

(a) Compute $\hat{\eta}$, the maximum likelihood estimator (MLE) of $\eta = \theta^2$, and $\bar{\eta}$, the uniformly minimum variance unbiased estimator (UMVUE) of η.

(b) Does $\text{Var}\,\bar{\eta}$ attain the Cramér-Rao lower bound?

(c) Determine the limiting distributions of both $\sqrt{n}(\hat{\eta}-\eta)$ and $\sqrt{n}(\overset{\sim}{\eta}-\eta)$.

6. Let $Y_1 = \beta_1 + e_1$, $Y_2 = \beta_2 + e_2$, and $Y_3 = \beta_1 + \beta_2 + e_3$, where the e_i, $i = 1, 2, 3$ are i.i.d. with common mean 0 and common variance σ^2.

(a) Find the best linear unbiased estimators of β_1 and β_2.

(b) Find the variance-covariance matrix of the estimators obtained in (a).

(c) Find an unbiased estimator of σ^2. What is its distribution under the additional assumption that the e_i, $i = 1, 2, 3$ are normally distributed?

(d) Find the best linear unbiased estimator of $2\beta_1 + 3\beta_2$ and the variance of this estimator.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
JANUARY, 1999

Statistics (M.A. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and justify its use.

---

1. Let $X_1, ..., X_n$ be independent and identically distributed (i.i.d.) random variables from the Poisson distribution with parameter λ.

(i) Find the UMVUE for λ.

(ii) Find the UMVUE for $\exp(-\lambda)$.

2. Let X be continuous with density $f(x;\theta)$, where θ is a real parameter with prior density $\pi(\theta)$. We wish to estimate θ, subject to the loss function $L(\theta;d) = W(\theta)(\theta - d)^2$, where $W(\theta)$ is a function of θ.

(i) Show that the Bayes estimate for θ is

$$\hat{\theta}_\pi = \frac{E[\theta W(\theta)|X]}{E[W(\theta)|X]}.$$

(ii) Let X be a binomial random variable with parameters n and θ. Suppose that the prior density of θ is $\pi(\theta) = 6\theta(1 - \theta)$ on (0,1) and the loss function is

$$L(\theta;d) = \frac{(\theta - d)^2}{\theta(1 - \theta)}.$$

Find the Bayes estimate for θ.

3. Suppose that X is a discrete variable with probability mass function (pmf) $f(x;\theta)$. Let $T = T(X)$ be a function of X with pmf $g(t;\theta)$.

(i) Show that

$$E\left\{\frac{\partial \log f(X;\theta)}{\partial \theta}\right\} = 0$$

and that

$$E\left\{\frac{\partial \log f(X;\theta)}{\partial \theta} \,\middle|\, T\right\} = \frac{\partial \log g(T;\theta)}{\partial \theta}.$$

(ii) Starting from

$$0 \leq \mathrm{Var}\left\{\frac{\partial \log f(X;\theta)}{\partial \theta} - \frac{\partial \log g(T;\theta)}{\partial \theta}\right\},$$

show that

$$\mathrm{Var}\left\{\frac{\partial \log f(X;\theta)}{\partial \theta}\right\} \geq \mathrm{Var}\left\{\frac{\partial \log g(T;\theta)}{\partial \theta}\right\}.$$

(iii) In what situation does the above inequality become an equality?

(iv) Give a statistical explanation of the results in parts (ii) and (iii).

Note: You may assume any needed mathematical regularity conditions without proof.

4. Suppose that Y_{ij}, $(i = 1, 2; \; j = 1, 2, 3)$ are independent random variables and that Y_{ij} has the normal distribution $N(\mu_i, \sigma^2)$.

(i) Find the least squares estimators for μ_1 and μ_2 and an unbiased estimator for σ^2.

(ii) To test $H_0 : (\mu_1, \mu_2) = (a_1\mu, a_2\mu)$, where a_1 and a_2 are known constants and μ is an unknown parameter, we can use the statistic

$$W = \frac{4(a_1 Y_{2\cdot} - a_2 Y_{1\cdot})^2}{(a_1^2 + a_2^2)[3\sum_{i=1}^{2}\sum_{j=1}^{3} Y_{ij}^2 - (Y_{1\cdot}^2 + Y_{2\cdot}^2)]},$$

where $Y_{i\cdot} = \sum_{j=1}^{3} Y_{ij}$, $i = 1, 2$. Find the sampling distribution of W under H_0.

5. Suppose that random variables X_i, $i = 1, 2, ..., n$, are independently and identically distributed with density

$$f_X(x) = f(x, \beta, \lambda) = 2\lambda\beta x \exp(-\lambda\beta x^2)$$

and that Y_i, $i = 1, 2, ..., n$, are independently and identically distributed with density

$$f_Y(y) = f(y, \lambda) = 2\lambda y \exp(-\lambda y^2),$$

where β and λ are unknown positive parameters.

(i) Find a two-dimensional sufficient statistic for the unknown parameter vector (β, λ) in terms of the samples X_i, $i = 1, 2, ..., n$ and Y_i, $i = 1, 2, ..., n$.

(ii) Find the form of the likelihood ratio test of $H_0 : \beta = 1$ versus the one-sided alternative $H_1 : \beta > 0$ with significance level 0.05, and show that it coincides with the one-sided test based upon the maximum likelihood estimator $\hat{\beta}$ of β.

3

6. Let $X_1, ..., X_n$ be an i.i.d sample from the gamma density

$$f(x; \theta) = \frac{\theta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\theta x}$$

where θ is an unknown positive parameter and α is a known positive constant.

(i) Find $\hat{\theta}$, the maximum likelihood estimator of θ.

(ii) What is the limiting distribution of $\sqrt{n}(\hat{\theta} - \theta)$?

.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
GRADUATE WRITTEN EXAMINATION
JANUARY, 1999

Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (**NOT YOUR NAME**) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and justify its use.

---

1. Let $X_1, ..., X_n$ be independent and identically distributed (i.i.d.) random variables from the Poisson distribution with parameter λ.

(i) Find the UMVUE for λ.

(ii) Find the UMVUE for $\exp(-\lambda)$.

2. Let X be continuous with density $f(x; \theta)$, where θ is a real parameter with prior density $\pi(\theta)$. We wish to estimate θ, subject to the loss function $L(\theta; d) = W(\theta)(\theta - d)^2$, where $W(\theta)$ is a function of θ.

(i) Show that the Bayes estimate for θ is

$$\hat{\theta}_\pi = \frac{E[\theta W(\theta)|X]}{E[W(\theta)|X]}.$$

(ii) Let X be a binomial random variable with parameters n and θ. Suppose that the prior density of θ is $\pi(\theta) = 6\theta(1 - \theta)$ on (0,1) and the loss function is

$$L(\theta; d) = \frac{(\theta - d)^2}{\theta(1 - \theta)}.$$

Find the Bayes estimate for θ.

3. Suppose that X is a discrete variable with probability mass function (pmf) $f(x; \theta)$. Let $T = T(X)$ be a function of X with pmf $g(t; \theta)$.

(i) Show that

$$E\left\{ \frac{\partial \log f(X; \theta)}{\partial \theta} \mid T \right\} = \frac{\partial \log g(T; \theta)}{\partial \theta}.$$

(ii) Starting from

$$0 \le \mathrm{Var}\left\{ \frac{\partial \log f(X; \theta)}{\partial \theta} - \frac{\partial \log g(T; \theta)}{\partial \theta} \right\},$$

show that

$$\mathrm{Var}\left\{ \frac{\partial \log f(X; \theta)}{\partial \theta} \right\} \ge \mathrm{Var}\left\{ \frac{\partial \log g(T; \theta)}{\partial \theta} \right\}.$$

(iii) In what situation does the above inequality become an equality?

(iv) Give a statistical explanation of the results in parts (ii) and (iii).

Note: You may assume any needed mathematical regularity conditions without proof.

4. Suppose that Y_{ij}, $(i = 1, 2; \ j = 1, 2, 3)$ are independent random variables and that Y_{ij} has the normal distribution $N(\mu_i, \sigma^2)$.

(i) Find the least squares estimators for μ_1 and μ_2 and an unbiased estimator for σ^2.

(ii) To test $H_0 : (\mu_1, \mu_2) = (a_1\mu, a_2\mu)$, where a_1 and a_2 are known constants and μ is an unknown parameter, we can use the statistic

$$W = \frac{4(a_1 Y_2. - a_2 Y_1.)^2}{(a_1^2 + a_2^2)[3 \sum_{i=1}^{2} \sum_{j=1}^{3} Y_{ij}^2 - (Y_1^2 + Y_2^2.)]},$$

where $Y_{i.} = \sum_{j=1}^{3} Y_{ij}$, $i = 1, 2$. Find the sampling distribution of W under H_0.

5. Suppose that random variables X_i, $i = 1, 2, ..., n$, are independent and identically distributed with density

$$f_X(x) = f(x, \beta, \lambda) = 2\lambda\beta x \exp(-\lambda\beta x^2)$$

and that Y_i, $i = 1, 2, ..., n$, are independent and identically distributed with density

$$f_Y(y) = f(y, \lambda) = 2\lambda y \exp(-\lambda y^2),$$

where β and λ are unknown positive parameters.

(i) Find a two-dimensional sufficient statistic for the unknown parameter (β, λ) in terms of the samples X_i, $i = 1, 2, ..., n$ and Y_i, $i = 1, 2, ..., n$.

(ii) Find the form of the likelihood ratio test of $H_0 : \beta = 1$ versus the one-sided alternative $H_1 : \beta > 1$ with significance level 0.05, and show that it coincides with the one-sided test based upon the maximum likelihood estimator $\hat{\beta}$ of β.

(iii) By any method you know, give the approximate distribution of $\hat{\beta}$ under the null hypothesis when n is large.

6. Let $X_1, ..., X_n$ be an i.i.d sample from a density $f(x; \theta)$, where θ is an unknown scalar parameter. Define

$$W_\alpha(\theta) = \int_{-\infty}^{\infty} f^{1+\alpha}(z; \theta)dz - (1 + 1/\alpha) \sum_{i=1}^{n} f^\alpha(X_i; \theta)/n,$$

where $\alpha > 0$ is a known constant. We estimate θ by $\hat{\theta}$, the quantity that minimizes $W_\alpha(\theta)$.

(i) Show that $\hat{\theta}$ satisfies the equation $U_n(\hat{\theta}) = 0$, where

$$U_n(\theta) = n^{-1} \sum_{i=1}^{n} u(X_i; \theta) f^\alpha(X_i; \theta) - \int_{-\infty}^{\infty} u(z; \theta) f^{1+\alpha}(z; \theta)dz$$

and $u(x; \theta) = (\partial/\partial\theta) \log f(x; \theta)$.

(ii) Find the limiting distribution of $\sqrt{n} U_n(\theta_0)$, where θ_0 is the true value of θ.

(iii) Find the asymptotic distribution of $\sqrt{n}(\hat{\theta} - \theta_0)$.

(iv) Can you see any connection between above estimate and the maximum likelihood estimate? (Hint: Consider the value of α).

Note: You may assume any needed mathematical regularity conditions without proof.

4