Let $X_0, X_1, X_2, \ldots$ be a <u>random</u> <u>autoregressive</u> <u>sequence</u> satisfying

$$X_j = \rho_j X_{j-1} + \epsilon_j, \quad j \geq 1$$

where (ρ_j, ϵ_j) $j=1,2,3,\ldots$ are independent, identically distributed random vectors and independent of X_0.

1. Express X_j in terms of X_0 and (ρ_i, ϵ_i) for $i=1,2,\ldots,j$.

2. Using 1 or otherwise show that the sequence $\{X_j\}$ has the <u>Markov property</u>, i.e., that the conditional distribution of X_j given $X_0 = x_0$, $X_1 = x_1, \ldots, X_{j-1} = x_{j-1}$ is the same as that of X_j given $X_{j-1} = x_{j-1}$ and is the same as that of $\rho_1 x_{j-1} + \epsilon_1$.

3. Assume that there exists a σ-finite measure m on $(R, B(R))$ (R is the set of real numbers, B(R) the σ-algebra of Borel sets on R) such that for each x, $\rho_1 x + \epsilon_1$ has a probability distribution that is dominated by m with a density $p(x, \cdot, \theta)$ where θ is a parameter with values in a set Θ. Assume further that the r.v. X_0 has a distribution that is also dominated by m with density $g(\cdot)$ that is independent of θ.

   a. Show that $\underline{X}_0^n = (X_0, X_1, \ldots, X_n)$ generates a dominated statistical experiment, clearly identifying the data space, the σ-algebra of events and the family of probability measures $P = \{P_\theta : \theta \in \Theta\}$ and the family of densities $f(\underline{x}_0^n, \theta)$.

   From now on assume that m is the Lebesgue measure, and ρ_1 and ϵ_1 are independent with distributions $N(\mu, \sigma^2)$ and $N(0, \tau^2)$ respectively. Let $\theta = (\mu, \sigma^2, \tau^2)$.

   b. Determine $p(x, \cdot, \theta)$ and $f(\underline{x}_0^n, \theta)$ explicitly.

   c. Assume σ^2 and τ^2 are known.

      i. Show that $f(\underline{x}_0^n, \theta)$ is an exponential family.

      ii. Determine a sufficient statistic for θ.

      iii. Find the m.l.e. of μ based on $\underline{X}_0^n$.

      iv. Let $\delta_{j+1} = \dfrac{(X_{j+1} - \mu X_j)\, X_j}{(\sigma^2 X_j^2 + \tau^2)}$.

Show that

$$E(\delta_{j+1} | X_0, X_1, \ldots, X_j) = 0$$

$$E(\delta_{j+1}^2 | X_0, X_1, \ldots, X_j) = \frac{X_j^2}{(\sigma^2 X_j^2 + \tau^2)}$$

v. Using (iv) or otherwise verify that the $\{\delta_j : j=1,2,\ldots\}$ are uncorrelated and compute the mean and variance of

$$Z_n = \sum_{j=0}^{n-1} \frac{X_j X_{j+1}}{(\sigma^2 X_j^2 + \tau^2)} - \mu \sum_{j=0}^{n-1} \frac{X_j^2}{(\sigma^2 X_j^2 + \tau^2)}$$

vi. Assume that $\left(\frac{1}{n} \sum_{j=0}^{n-1} \frac{X_j^2}{\sigma^2 X_j^2 + \tau^2}\right)$ converge w.p.1 to some constant $\gamma > 0$.

Show that the m.l.e. $\hat{\mu}_n$ of μ based on $\underline{X}_0^n$ converges to μ in probability and that the Fisher information $I_n(\theta)$ satisfies $n^{-1} I_n(\theta) \to \gamma$.

vii. Using the asymptotic theory of m.l.e., make a conjecture about the asymptotic distribution of $(\hat{\mu}_n - \mu)$.

Solution to Theory I Prelim Question March 99

1.
$$\begin{aligned} X_j &= \rho_j(\rho_{j-1} X_{j-2} + \epsilon_{j-1}) + \epsilon_j \\ &= \rho_j \rho_{j-1} X_{j-2} + \rho_j \cancel{\rho_j} \epsilon_{j-1} + \epsilon_j \\ &= \rho_j \rho_{j-1} (\rho_{j-2} X_{j-3} + \epsilon_{j-3}) + \rho_j \cancel{\rho_j} \epsilon_{j-1} + \epsilon_j \\ &= \cdots \\ &= \rho_j \rho_{j-1} \rho_{j-2} \cdots \rho_1 X_0 + \rho_j \rho_{j-1} \cdots \rho_2 \epsilon_1 + \rho_j \rho_{j-1} \cdots \rho_3 \epsilon_2 \\ &\quad + \cdots + \rho_j \epsilon_{j-1} + \epsilon_j \\ &= \psi(X_0, (\rho_1, \epsilon_1), (\rho_2, \epsilon_2), \ldots (\rho_j, \epsilon_j)) \end{aligned}$$

2. Since $X_j = \rho_j X_{j-1} + \epsilon_j$ and by 1, X_{j-1} is a function of X_0 and (ρ_i, ϵ_i), $i = 1, 2, \ldots j-1$ and (ρ_j, ϵ_j) is independent of X_0 and (ρ_i, ϵ_i) $i = 1, 2, \ldots j-1$, the conditional distribution of X_j given $X_{j-1} = x_{j-1}$, $X_{j-2} = x_{j-2}, \ldots X_1 = x_1, X_0 = x_0$ is the same as that of ~~$\rho_j X_{j-1} + \epsilon$~~ $\rho_j X_{j-1} + \epsilon_j$ given $X_{j-1} = x_{j-1}$ which is the same as $\rho_j x_{j-1} + \epsilon_j$.

3.a) By (2) for $A_0, A_1, \ldots, A_n \in \mathcal{B}(R)$

$$\begin{aligned} &P_\theta(X_0 \in A_0, X_1 \in A_1, \ldots, X_n \in A_n) \\ &= \int_{A_0} \int_{A_1} \cdots \left(\int_{A_n} p(x_{n-1}, x_n, \theta)\, m(dx_n) \right) p(x_{n-2}, x_{n-1}, \theta)\, m(dx_{n-1}) \\ &\qquad \cdots g(x_0)\, m(dx_0) \\ &= \int_{A_0 \times A_1 \times \cdots \times A_n} g(x_0)\, p(x_0, x_1, \theta)\, p(x_1, x_2, \theta) \cdots p(x_{n-1}, x_n, \theta) \\ &\qquad m(dx_0)\, m(dx_1) \cdots m(dx_n) \end{aligned}$$

Thus, the data space $\mathbb{X}$ is $R^{(n+1)}$, the event space is $\mathcal{X} = \mathcal{B}(R^{n+1})$ and $\lambda = m^{(n+1)}$ is the dominating measure and the family $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ where $P_\theta(A_0 \times A_1 \times \cdots \times A_n)$ is given above. Since the class $\{A_0 \times A_1 \times \cdots \times A_n : A_i \in \mathcal{B}(R)\}$ generates $\mathcal{B}(R^{n+1})$ this specifies P_θ completely. So

$$f_\theta(\underline{x}_0^n, \theta) = g(x_0)\, p(x_0, x_1, \theta) \cdots p(x_{n-1}, x_n, \theta)$$

Solution to Theory I Prelim Question March 99 (2)

3 b) Since P_1 and ϵ_1 are independent normal r.v. with distribution $N(\mu, \sigma^2)$ and $N(0, \tau^2)$ respectively $P_1 x + \epsilon_1 \sim N(\mu x, x^2\sigma^2 + \tau^2)$ and so

$$p(x, y, \theta) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{x^2\sigma^2+\tau^2}} \exp\left(-\frac{1}{2} \frac{(y-\mu x)^2}{x^2\sigma^2+\tau^2}\right)$$

and $$f(\underline{x}_0^n, \theta) = g(x_0) \prod_{j=0}^{n} \left(\frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{x_j^2\sigma^2+\tau^2}} \exp\left(-\frac{1}{2} \frac{(x_{j+1}-\mu x_j)^2}{x_j^2\sigma^2+\tau^2}\right)\right)$$

3c) i) If σ^2 and τ^2 are known Then

$$f(\underline{x}_0^n, \theta) = h(\underline{x}_0^n) \exp\left(\mu T_1(\underline{x}_0^n) + \mu^2 T_2(\underline{x}_0^n)\right)$$

where $$T_1(\underline{x}_0^n) = \left(\sum_{j=0}^{n-1} (x_j x_{j+1}) / (x_j^2\sigma^2+\tau^2)\right)$$

$$T_2(\underline{x}_0^n) = -\frac{1}{2} \sum_{j=0}^{n-1} x_j^2 / (x_j^2\sigma^2+\tau^2)$$

and $h(\cdot)$ is a function independent of μ

Thus it is an exponential family

ii) $(T_1(\underline{X}_0^n), T_2(\underline{X}_0^n))$ is a sufficient Statistic for μ or for $\theta = (\mu, \tau^2, \sigma^2)$ (since τ^2 & σ^2 are known)

iii) The m.l.e. of μ based on $\underline{X}_0^n$ is

$$\hat{\mu}_n = -\frac{T_1(\underline{X}_0^n)}{2 T_2(\underline{X}_0^n)} = + \frac{\sum_{j=0}^{n-1} (x_j x_{j+1}) / (x_j^2\sigma^2+\tau^2)}{\left(\sum_{j=0}^{n-1} x_j^2 / (x_j^2\sigma^2+\tau^2)\right)}$$

iv) By Markov property

$$E(\delta_{j+1} \mid X_0, \ldots, X_j) = X_j E\left(\left((P_{j+1}-\mu) X_j + \epsilon_{j+1}\right) \mid X_j\right)$$
$$= X_j (X_j \, 0 + 0) = 0$$

Since $(P_{j+1}, \epsilon_{j+1})$ is indep of X_j with mean $(\mu, 0)$.

By the same reasoning

$$E(\delta_{j+1}^2 \mid X_0, \ldots, X_j) = X_j^2 \frac{E\left(\left((P_{j+1}-\mu) X_j + \epsilon_{j+1}\right)^2 / X_j\right)}{(\sigma^2 X_j + \tau^2)^2}$$

Solution to Theory I Prelim Question March 99 (3)

$$= X_j^2 \left(X_j^2 \sigma^2 + \tau^2 \right)^{-1}$$

v) $$Z_n = \sum_{j=0}^{n-1} \frac{X_j \left(X_{j+1} - \mu X_j \right)}{\left(X_j^2 \sigma^2 + \tau^2 \right)} = \sum_{j=0}^{n-1} \delta_{j+1}$$

Since $E(\delta_{j+1} \mid X_0, \ldots, X_j) = 0$, the $\{\delta_j : j = 1, 2, \ldots\}$ are uncorrelated. So $EZ_n = 0$ and

$$V(Z_n) = \sum_{j=0}^{n-1} V(\delta_{j+1}) = \sum_{j=0}^{n-1} E(\delta_{j+1}^2) = \sum_{j=0}^{n-1} E\left(E(\delta_{j+1}^2 \mid X_j)\right)$$

$$= \sum_{0}^{n-1} E\left(\frac{X_j^2}{X_j^2 \sigma^2 + \tau^2} \right)$$

vi) By (iii) $\frac{\partial}{\partial \theta} \ln f(\underline{x}_0^n, \theta) = Z_n$ and

So $I_n(\theta) = E_\theta\left(\left(\frac{\partial}{\partial \theta} \ln f(\underline{x}_0^n, \theta) \right)^2 \right) = E_\theta Z_n^2 = V_\theta Z_n$

By hypothesis of (vi) $\frac{1}{n} \sum_{0}^{n-1} \frac{X_j^2}{X_j^2 \sigma^2 + \tau^2} \to \gamma$ w.p.1

and so by bounded convergence Theorem, $n^{-1} I_n(\theta)$ $= \frac{1}{n} V(Z_n) \to \gamma$.

Next, $$\left(\hat{\mu}_n - \mu \right) = \frac{Z_n}{\sum_{0}^{n-1} \frac{X_j^2}{(\sigma^2 X_j^2 + \tau^2)}}$$

Since $E\left(\frac{Z_n}{n} \right) = 0$, $V\left(\frac{Z_n}{n} \right) = \frac{1}{n^2} V(Z_n) \to 0$,

it follows by Chebychev that $\frac{Z_n}{n} \to 0$.

Also by hyp $\frac{1}{n} \sum_{0}^{n-1} \frac{X_j^2}{\sigma^2 X_j^2 + \tau^2} \to \gamma$, $0 < \gamma < \infty$

vii) By the asymptotic Theory of m.l.e. a natural conjecture is

$$\sqrt{n} \left(\hat{\mu}_n - \mu \right) \xrightarrow{d} N(0, \gamma^{-1}).$$

Let X be a random variable and $\phi: R \to R^+$ be convex. Let $\psi(a) = E\,\phi(X-a)$. Let α and β be the infimum and supremum of the set $A = \{a: \psi(a) < \infty\}$. (If $A = \phi$ then α and β are defined to be ∞).

a) Show that ψ is finite and convex on (α,β), if $\alpha < \beta$.

b) Show that if $\phi(x) = |x|^p$ for $p \ge 1$ and $E\,|X|^p < \infty$ then $\alpha = -\infty$ and $\beta = \infty$.

c) It is known that $\phi: R \to R^+$ convex implies that "chords turn counter clockwise", that is, for $a_1 < a_2 < b_1 < b_2$,

$$\frac{\phi(a_2) - \phi(a_1)}{a_2 - a_1} \le \frac{\phi(b_2) - \phi(b_1)}{b_2 - b_1}$$

This in turn implies that ϕ has right and left derivatives on all of R, say ϕ_+' and ϕ_-', that are both nondecreasing with $\phi_-'(\cdot) \le \phi_+'(\cdot)$. Using these facts or otherwise show that for $\alpha < a < \beta$ both $\phi_+'(X-a)$ and $\phi_-'(X-a)$ have finite expectations and

$$\psi_+'(a) = -E\,\phi_-'(X-a) \text{ and}$$
$$\psi_-'(a) = -E\,\phi_+'(X-a)$$

d) Show that $\psi(\cdot)$ is minimized at γ in (α,β) iff $\psi_+'(\gamma) \ge 0 \ge \psi_-'(\gamma)$.

e) Apply (d) to $\phi(x) = |x|$, $\phi(x) = |x|^2$, $\phi(x) = |x|^3$ and determine an equation for an optimal γ in each case (as explicitly as possible).

f) Let θ have a Beta (p,q), $p > 0$, $q > 0$, distribution on $[0,1]$. Given θ, define $X = \{\delta_1,\delta_2,\ldots,\delta_n\}$ where the δ_i's are i.i.d. Bernoulli (θ) random variables.

i) Find the posterior distribution of θ given X.

ii) Find an explicit equation for the Bayes estimate of θ based on X for the loss functions $L_1(\theta,a) = |\theta-a|$, $L_2(\theta,a) = (\theta-a)^2$ and $L_3(\theta,a) = |\theta-a|^3$.

h) Let X be a r.v. and $\phi: R \to R^+$ be convex and γ be any value that minimizes $\psi(a) = E\,\phi(X-a)$. Let $X_1,X_2,\ldots,X_n$ be i.i.d. with the same distribution as that of X. Let γ_n be any value that minimizes $\psi_n(\cdot)$ where

$$\psi_n(a) = \sum_{i=1}^{n} \phi(X_i-a)$$

For a ϕ of your choice from part (e) make a conjecture about the behavior of $(\gamma_n-\gamma)$ for large n and also about the asymptotic distribution of $(\gamma_n-\gamma)$.

a) For $\alpha < a < \beta$ $\exists$ $\alpha < a_1 < a < a_2 < \beta$ such that $\psi(a_1) < \infty$, $\psi(a_2) < \infty$.

Let $a = \lambda a_1 + (1-\lambda) a_2$ with $\lambda = \frac{a_2 - a}{a_2 - a_1}$

Then since ϕ is convex

$$\phi(X-a) = \phi(\lambda X + (1-\lambda)X - \lambda a_1 - (1-\lambda)a_2)$$
$$= \phi(\lambda(X-a_1) + (1-\lambda)(X-a_2))$$
$$\leq \lambda \phi(X-a_1) + (1-\lambda)\phi(X-a_2)$$

$\Longrightarrow$

$$\psi(a) \leq \lambda \psi(a_1) + (1-\lambda)\psi(a_2) < \infty$$

showing ψ is both finite and convex in (α, β)

b) $|X-a|^p \leq 2^{p-1}(|X|^p + |a|^p)$ by convexity of $\phi(x) = |x|^p$ for $p \geq 1$. So if $E|X|^p < \infty$ then $E|X-a|^p < \infty$ for all a.

c) Fix $\alpha < a_1 < a_2 < a < b < b_1 < b_2 < \beta$

Then $X-b_2 < X-b_1 < X-b < X-a < X-a_2 < X-a_1$

So

$$Z \equiv \frac{\phi(X-a) - \phi(X-b)}{(X-a) - (X-b)} \leq \frac{\phi(X-a_1) - \phi(X-a_2)}{(X-a_1) - (X-a_2)}$$

and $$\geq \frac{\phi(X-b_1) - \phi(X-b_2)}{(X-b_1) - (X-b_2)}$$

Let $$Z_1 = \frac{\phi(X-a_1) - \phi(X-a_2)}{(a_2 - a_1)}$$

$$Z_2 = \frac{\phi(X-b_1) - \phi(X-b_2)}{(b_2 - b_1)}$$

By hypothesis $E|Z_1| < \infty$ and $E|Z_2| < \infty$

Since $Z_2 \le Z \le Z_1$ it follows that

$$|Z| \le |Z_1| + |Z_2|.$$

Letting $a \uparrow b$ and applying LDCT yields

$$E\left|\phi_+'(X-b)\right| < \infty \text{ and } \psi_-'(b) = -E\phi_+'(X-b)$$

Similarly letting $b \downarrow a$ and applying LDCT yields

$$E\left|\phi_-'(X-a)\right| < \infty \text{ and } \psi_+'(a) = -E\phi_-'(X-a)$$

---

d) If $\psi(a)$ is minimised at γ then by convexity

$$\psi(b) \geq \psi(\gamma) \text{ for all } b \neq \gamma$$

and so $\psi_+'(\gamma) \geq 0 \geq \psi_-'(\gamma) \geq 0$.

Conversely $\psi_+'(\gamma) \geq 0 \Rightarrow \psi$ is increasing in (γ, ∞)

and $\psi_-'(\gamma) \le 0 \Rightarrow \psi$ is decreasing in $(-\infty, \gamma)$.

e) The equation for γ is

$$E\,\phi_-'(X-\gamma) \le 0 \le E\phi_+'(X-\gamma)$$

i) If $\phi(x) = |x|$ then $\phi_+'(x) = 1$ if $x \geq 0$ and -1 for $x < 0$ and $\phi_-'(x) = -1$ if $x \le 0$ and $+1$ for $x > 0$. So γ satisfies

$$-P(X \le \gamma) + P(X > \gamma) \le 0 \le P(X \geq \gamma) - P(X < \gamma)$$

$$\Leftrightarrow \quad 2P(X \le \gamma) \geq 1$$

$$2P(X \geq \gamma) \geq 1.$$

ie γ is a median of X.

ii) If $\phi(x) = |x|^p$ ~~for~~ $p > 1$ then $\phi'(x) = p|x|^{p-1}$ for $x > 0$

$= -p|x|^{p-1}$ for $x < 0$ and 0 for $x = 0$.

So γ satisfies

$$p\, E\left((X-\gamma)^{p-1} : X \geq \gamma\right) = p\, E\left((\gamma-X)^{p-1} : X \leq \gamma\right)$$

For $p=2$ this becomes

$$E(X-\gamma : X \geq \gamma) = E(\gamma - X : X \leq \gamma)$$

ie $\gamma = EX$

For $p=3$ this becomes

$$E((X-\gamma)^2 : X \geq \gamma) = E((X-\gamma)^2 : X \leq \gamma)$$

f)

i) The joint distribution of θ and X is given by

$$P(a < \theta < b,\ X = x)$$

$$= \frac{1}{B(p,q)} \int_a^b \theta^{p-1} (1-\theta)^{q-1}\ \theta^{\sum_1^n x_i} (1-\theta)^{n-\sum_1^n x_i}\, d\theta$$

where $X = (\delta_1, \delta_2, \ldots, \delta_n)$

$x = (x_1, x_2, \ldots, x_n)$

The posterior distribution of θ is given by

$$P(a < \theta < b \mid X = x)$$

$$= \frac{P(a < \theta < b,\ X = x)}{P(X = x)}$$

$$= \frac{1}{B(p,q)\, P(X=x)} \int_a^b \theta^{p+\sum_1^n x_i - 1} (1-\theta)^{q+n-\sum_1^n x_i - 1}\, d\theta$$

ie $\theta \mid X = x$ is $\text{Beta}\left(p + \sum_1^n x_i,\ q + \sum_1^n (1-x_i)\right)$

f) ii) Case 1 Bayes estimate γ when $L_1(\theta,a) = |\theta - a|$

It is the median of the Beta distribution with parameters $(p + \sum_1^n x_i, \; q + \sum_1^n (n - x_i))$

Thus γ satisfies

$$\int_0^\gamma x^{p+r-1}(1-x)^{q+n-r-1}\,dx$$

$$= \int_\gamma^1 x^{p+r-1}(1-x)^{q+n-r-1}\,dx$$

where $r = \sum_1^n x_i$

Both sides are polynomials in γ.

Case 2 Bayes estimate γ when $L_2(\theta,a) = (\theta - a)^2$

It is the posterior mean of Beta $(p+r, q+n-r)$ and so

$$\gamma = \frac{p+r}{(p+q+n)} \quad \text{where } r = \sum_1^n x_i$$

Case 3 $L_3(\theta,a) = (\theta - a)^3$

Then γ satisfies

$$\int_0^\gamma (\gamma - x)^2 x^{p+r-1}(1-x)^{q+n-r-1}\,dx$$

$$= \int_\gamma^1 (x-\gamma)^2 x^{p+r-1}(1-x)^{q+n-r-1}\,dx$$

Both sides are polynomials in γ.

h) ~~Take $\phi(x) = x^2$~~

Conjecture: $\gamma_n \xrightarrow{p} \gamma$ (when γ is unique)

and $\gamma_n - \gamma$ is asymptotically normal

Let $\phi(x) = x^2$. Assume $EX^2 < \infty$. Then $\gamma = EX$, $\gamma_n = \bar{X}_n$

By ~~CLT~~ & WLLN + CLT $\bar{X}_n \xrightarrow{P} \gamma$ & $\sqrt{n}(\bar{X}_n - \gamma) \to N(0,\sigma^2)$.

1. Define
   (a) a best invariant decision rule
   (b) a minimax rule

2. Prove that if δ is a best invariant and admissible decision rule, then δ is minimax.

3. However, it is not necessary that a best invariant decision rule be admissible in order to be minimax. Show this by completing the details of the following example.

   Let X be a random variable with $P_\theta(X = \theta + 1) = 1/2 = P_\theta(X = \theta - 1)$, $\theta \in \mathbb{R}$, and let the loss function be given by $L(\theta, a) \equiv L_1(a - \theta) = |\theta - a| I(|\theta - a| \leq 1) + I(|\theta - a| > 1)$ where $I(\cdot)$ denotes the indicator function.

   (a) Show that the decision problem is invariant under the group of transformations $\mathcal{G} \equiv \{g_c : c \in \mathbb{R}\}$, where $g_c(x) = x + c$, $x \in \mathbb{R}$.

   (b) It follows from part (a) that an invariant rule under $\mathcal{G}$ is of the form $d(X) = X - b$, $b \in \mathbb{R}$. Show that the risk function of the invariant rule $d(X) = X - b$ at $\theta = 0$ is given by

   $$R(0, d) = \begin{cases} 1 - (|b|/2) & \text{if } |b| \leq 1 \\ |b|/2 & \text{if } 1 \leq |b| \leq 2 \\ 1 & \text{if } |b| > 2. \end{cases}$$

   (c) Conclude from (b) that the best invariant rules are given by $d_1(X) = X - 1$ and $d_2(X) = X + 1$, with $R(0, d_i) = 1/2$, $i = 1, 2$.

   (d) Let τ_n denote the uniform distribution on $(-n, n)$, $n \geq 1$. Then, show that for any nonrandomized decision rule $d(X)$ (not necessarily invariant), its Bayes risk $r(\tau_n, d)$ with respect to τ_n satisfies the following relations:

   $$\begin{aligned} r(\tau_n, d) &= (4n)^{-1} \left[\int_{-n+1}^{n+1} L_1(d(y) - y + 1) dy + \int_{-n-1}^{n-1} L_1(d(y) - y - 1) dy \right] \\ &\geq \frac{2n - 2}{4n}, \end{aligned}$$

   where $L_1(\cdot)$ is as defined above.

   (e) From (d), conclude that

   $$\lim_{n \to \infty} \inf_{\delta \in \mathcal{D}} r(\tau_n, \delta) = 1/2$$

   where $\mathcal{D}$ denotes the set of all decision rules.

   (f) From (c) and (e), conclude that $d_1(X)$ and $d_2(X)$ are minimax. Be sure to state any standard result that you are using.

   (g) Next define the rule d_0 by

   $$d_0(X) = \begin{cases} X + 1 & \text{if } X < 0 \\ X - 1 & \text{if } X \geq 0 \end{cases}$$

   It can be shown that $R(\theta, d_0) \leq 1/2$ for all θ, so that d_0 is at least as good as d_1 and d_2. Show that

   $$R(0, d_0) = 0.$$

   (Thus, in this example, the best invariant rules are minimax, but not admissible).

Solution - Theory-III / Ph.D. Prelim / Spring, 1999.

1. — [Class-Notes]

2. If possible, suppose that δ is not minimax. Then, there exists a decision rule δ_1 such that

$$\sup_{\theta \in \Theta} R(\theta, \delta) > \sup_{\theta \in \Theta} R(\theta, \delta_1). \quad (1)$$

Since δ is the best invariant rule, it has a constant risk function. Hence, from (1), we get for all $\theta_0 \in \Theta$,

$$R(\theta_0, \delta) = \sup_{\theta \in \Theta} R(\theta, \delta)$$

$$> \sup_{\theta \in \Theta} R(\theta, \delta_1) \geq R(\theta_0, \delta_1),$$

implying that δ_1 is better than δ. This contradicts the admissibility of δ.

3. (a) Easy. (Here, θ is a location parameter)

(b)

$$
\begin{aligned}
R(0,d) &= E_0\, L(0, d(x)) \\
&= E_0 \left\{ |x-b| \cdot I(|x-b| \le 1) + I(|x-b|>1) \right\} \\
&= \frac{1}{2}\Big[\left\{ |1-b|\, I(|1-b| \le 1) + I(|1-b|>1) \right\} \\
&\qquad + \left\{ |-1-b| \cdot I(|-1-b| \le 1) + I(|-1-b|>1) \right\} \Big] \\
&= \frac{1}{2}\Big[\left\{ |1-b| \cdot I(0 \le b \le 2) + I(b \notin [0,2]) \right\} \\
&\qquad + \left\{ |1+b| \cdot I(-2 \le b \le 0) + I(b \notin [-2,0]) \right\} \Big]
\end{aligned}
$$

$$
= \begin{cases}
\frac{1}{2}[(1-b)+1] & \text{if } 0 \le b \le 1 \\
\frac{1}{2}[(b-1)+1] & \text{if } 1 < b \le 2 \\
\frac{1}{2}[1+1] & \text{if } |b| > 2) \text{ or } b = \\
\frac{1}{2}[1+(|b|-1)] & \text{if } -2 \le b < -1 \\
\frac{1}{2}[1+(1+b)] & \text{if } -1 \le b < 0
\end{cases}
$$

which is equivalent to the given expression.

(c) $\min_b R(0,d) = 1/2$, which is attained by $b = \pm 1$.

(d)

$$r(\tau_n, d) = \frac{1}{2n}\int_{-n}^{n} R(\theta, d)\, d\theta$$

$$= \frac{1}{2n}\int_{-n}^{n}\left[E_\theta L(\theta, d(x))\right] d\theta$$

$$= (2n)^{-1}\int_{-n}^{n}\left[\frac{1}{2}\left\{L_1\left(d(\theta+1)-\theta\right) + L_1\left(d(\theta-1)-\theta\right)\right\}\right] d\theta$$

$$= (4n)^{-1}\left\{\int_{-n+1}^{n+1} L_1\left(d(y)-(y-1)\right) dy \quad [\text{Put } y=\theta+1]\right.$$

$$\left. + \int_{-n-1}^{n-1} L_1\left(d(y)-(y+1)\right) dy\right\} \quad [\text{Put } y=\theta-1]$$

$$\geqslant (4n)^{-1}\int_{-n+1}^{n-1}\left[L_1\left(d(y)-y+1\right) + L_1\left(d(y)-y-1\right)\right] dy$$

$$= (2n)^{-1}\int_{-n+1}^{n-1}\left\{\underbrace{E_0 L_1\left(X-[y-d(y)]\right)}_{= R(0, d_y)}\right\} dy$$

, where $d_y = X - [y - d(y)]$ is an invariant decision rule

$$\geqslant (2n)^{-1}\int_{-n+1}^{n-1} R(0, d_1)\, dy$$

[A Best invariant rule]

$$= (2n-2)/4n.$$

(e) Clearly, $R(\theta, d_1) = R(0, d_1) = \frac{1}{2}$ for all $\theta \Rightarrow$
$\frac{1}{2} = r(\tau_n, d_1)$

$$\geq \inf_{\delta \in \mathcal{D}} r(\tau_n, \delta) = \inf_{d \in D} r(\tau_n, d) \quad \text{(why?)}$$

The set of all nonrandomized rules

$$\geq \frac{2n-2}{4n} \quad \text{(by (d))}$$

$$\Rightarrow \lim_{n \to \infty} \inf_{\delta \in \mathcal{D}} r(\tau_n, \delta) = \frac{1}{2}$$

(f) For $i = 1, 2$, d_i is an equalizer rule, which is also extended Bayes (by (e)). Hence, by Theorem 2.11.3 of [Ferguson] / [Class notes], d_i is minimax, $i = 1, 2$.

(g)
$$R(0, d_0) = E_0 L(0, d_0(x)) = E_0 L_1(d_0(x))$$
$$= \frac{1}{2}\left[L_1(d_0(1)) + L_1(d_0(-1))\right]$$
$$= \frac{1}{2}\left[L_1(0) + L_1(0)\right]$$
$$= 0$$
, QED.

NOTE: Indeed, it can be shown that $R(\theta, d_0) = 0$ for all $-1 \leq \theta < 1$.

1. (a) i. Define a nonrandomized test rule.

 ii. State the existence and the uniqueness parts of the Neyman-Pearson Lemma.

 (b) Let P_{θ_i} be a probability measure on $(\mathbb{R}^d, \mathcal{R}^d)$, $d \geq 1$, such that P_{θ_i} is dominated by the Lebesgue measure, $i = 0, 1$. Consider the testing problem $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$. The most powerful test of size $\alpha \in (0,1)$, is, in general, a randomized test. Show that for any $\alpha \in (0,1)$, there exists a nonrandomized most powerful test of size α for testing H_0 against H_1.

 [**Hint**: You may use the following result:

 Let $f : \mathbb{R}^d \to [0, \infty)$ be a Lebesgue integrable function such that for some $a \geq 0$ and some Borel set A, $\int_A f(x)dx = a$. Then, given any $0 \leq b \leq a$, there exists a Borel subset B of A such that $\int_B f(x)dx = b$.]

 (c) Does the randomized test you found in (b) contradict the uniqueness part of the Neyman-Pearson Lemma? Explain (briefly).

2. (a) Define the monotone likelihood ratio property.

 (b) Let $\mathcal{F} \equiv \{f(x,\theta) : \theta \in \Theta\}, \Theta \subset \mathbb{R}$, be a family of densities on the real line (with respect to some σ-finite measure μ). Suppose that $f(x,\theta)$ is strictly positive for all x and θ, and $\partial^2 f(x,\theta)/\partial x \partial \theta$ exists and is continuous on $\mathbb{R} \times \Theta$. Then, show that $\mathcal{F}$ has the monotone likelihood ratio property in x if and only if $\frac{\partial^2}{\partial\theta\partial x} \log f(x,\theta) \geq 0$ for all x, θ.

3. Let $X_1, \ldots, X_n$ be iid random variables with the UNIFORM$(0,\theta)$ distribution, where $\theta > 0$.

 (a) Show that the joint distribution of $(X_1, \ldots, X_n)$ has the monotone likelihood ratio property.

 (b) For $\alpha \in (0,1)$ and $\theta_0 > 0$ fixed, consider the class of test rules

$$\mathcal{C}_\alpha = \{\phi : E_{\theta_0}\phi(\mathbf{X}) = \alpha, \quad \sup_{\theta \leq \theta_0} E_\theta \phi(\mathbf{X}) = \alpha \quad \text{and} \quad \phi(\mathbf{x}) = 1 \quad \text{if} \quad x_{(n)} > \theta_0\},$$

 where $\mathbf{X} = (X_1, \ldots, X_n)$, $\mathbf{x} = (x_1, \ldots, x_n)$ and $x_{(n)} = \max_{1 \leq i \leq n} x_i$. Show that any $\phi \in \mathcal{C}_\alpha$ is a UMP size α test for testing $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$.

 (c) In general, UMP tests for two-sided hypotheses do not exist. However, a size $\alpha \in (0,1)$ UMP test for the two-sided testing problem $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$ ($\theta_0 > 0$ fixed) exists in this case. Find such a UMP test.
 [**Hint**: First, find a size α UMP test for $H_0 : \theta = \theta_0$ against $H_1 : \theta < \theta_0$ and then, *combine* it with a suitable size α UMP test for $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$.]

1 (a) -

(b) By the Neyman - Pearson Lemma, a size α MP test for testing $H_0: \theta = \theta_0$ vs. $H_1: \theta = \theta_1$ is given by

$$\phi_\alpha(x) = \begin{cases} 1 & \text{if } f_1(x) > k f_0(x) \\ \gamma & \qquad = \\ 0 & \qquad < \end{cases}$$

where $f_i = dP_{\theta_i} / d(P_{\theta_0} + P_{\theta_1})$ and $k \in [0, \infty)$ is a constant satisfying $E_{\theta_0} \phi(X) = \alpha$.

Let $A_0 = \{x \in \mathscr{X} : f_1(x) > k f_0(x)\}$, and
$A = \{x \in \mathscr{X} : f_1(x) = k f_0(x)\}$, where $\mathscr{X}$ denotes the sample space. Then,

$$P_{\theta_0}(A_0) + \gamma P_{\theta_0}(A) = \alpha$$

$$\Rightarrow \quad \gamma \cdot P_{\theta_0}(A) = (\alpha - P_{\theta_0}(A_0)).$$

Next, set $b = \gamma \cdot P_{\theta_0}(A)$. Then, $0 \le b \le a \equiv P_{\theta_0}(A)$

Hence, by the result, there exists a subset $B \subset A$ such that $P_{\theta_0}(B) = b$.

Hence, $\tilde{\phi}_\alpha = I_{(A_0 \cup B)}$ is a nonrandomized MP test of size α.

(c) This does NOT contradict the uniqueness part of the NP Lemma, since $\tilde{\varphi}_\alpha(x) = \varphi_\alpha(x)$ for all $x \in A_0 \cup \{x : f_1(x) < k f_0(x)\}$.

2. (a) ——

(b) Suppose, $\mathcal{F}$ has MLR in x. Then, for any $\theta_1 > \theta_2$,

$$\frac{f(x, \theta_1)}{f(x; \theta_2)} \text{ is an increasing function of } x$$

$\Leftrightarrow$ $\log f(x;\theta_1) - \log f(x, \theta_2)$ is increasing in x

(1) $\Leftrightarrow$ $\frac{\partial}{\partial x}\left[\log f(x, \theta_1) - \log f(x, \theta_2)\right] \geq 0 \quad \forall x, \forall \theta_1 > \theta_2$

Hence, by the mean value theorem,

LHS of (1)

$= (\theta_1 - \theta_2) \frac{\partial^2}{\partial\theta\,\partial x} \log f(x, \theta^*) \geq 0 \quad \forall x, \forall \theta_1 > \theta_2$

where θ^* is a point between θ_2 and θ_1.

$\Leftrightarrow$ $\frac{\partial^2}{\partial\theta\,\partial x} \log (x, \theta^*) \geq 0 \quad \forall x, \forall \theta_1 > \theta_2$

Now, letting $\theta_1 \to \theta_2+$, we get $\frac{\partial^2}{\partial\theta\partial x} \log f(x, \theta) \geq 0 \ \forall x, \forall \theta$

Conversely, if $\frac{\partial^2}{\partial\theta\partial x} \log f(x, \theta) \geq 0 \ \forall x, \forall \theta$, then retracing the steps above, one can conclude that

$\frac{\partial}{\partial x}\left(\log f(x, \theta_1) - \log f(x, \theta_2)\right) \geq 0 \quad \forall x, \forall \theta_1 > \theta_2$, so that $\mathcal{F}$ has MLR in x.

3. (a)

$$f(\underset{\sim}{x}, \theta) = \text{the joint pdf of } X_1, \cdots, X_n$$
$$= \theta^{-n} \; I(0 < x_{(1)} \le x_{(n)} \le \theta)$$

where $x_{(1)} = \min_{1 \le i \le n} x_i$ and $x_{(n)} = \max_{1 \le i \le n} x_i$, and where $I(\cdot)$ denotes the indicator function.

Then, for any $\theta_1 > \theta_2$,

$$\frac{f(\underset{\sim}{x}, \theta_1)}{f(\underset{\sim}{x}, \theta_2)} = \begin{cases} (\theta_2/\theta_1)^n & \text{if } \quad 0 < x_{(n)} \le \theta_2 \\ +\infty & \text{if } \quad \theta_2 < x_{(n)} \le \theta_1 \end{cases}$$

for all $\underset{\sim}{x} \in A_{\theta_1,\theta_2} \equiv \{\underset{\sim}{x} \in \mathbb{R}^n : f(\underset{\sim}{x}, \theta_1) + f(\underset{\sim}{x}; \theta_2) > 0\}$

$= \{\underset{\sim}{x} : 0 < x_{(1)} \le x_{(n)} \le \theta_1\}$.

$\Rightarrow \dfrac{f(\underset{\sim}{x}; \theta_1)}{f(\underset{\sim}{x}; \theta_2)}$ is a nondecreasing function of $x_{(n)}$ $\Rightarrow$ $\{f(\underset{\sim}{x}; \theta) : \theta > 0\}$ has MLR in $x_{(n)}$.

(b) By the MLR property, a size α UMP test for $H_0: \theta \le \theta_0$ vs $H_1: \theta > \theta_0$ is given by

2 $\leftarrow$ $$\varphi_1(\underset{\sim}{x}) = \begin{cases} 1 & \text{if } x_{(n)} > c \\ 0 & \text{if } x_{(n)} \le c \end{cases}$$

where $E_{\theta_0}\phi_1(\underset{\sim}{X}) = \alpha \iff P_{\theta_0}(X_{(n)} > c) = \alpha$

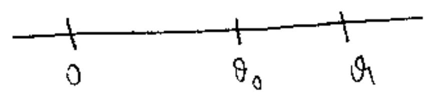
$\iff P_{\theta_0}(X_{(n)} \leq c) = 1-\alpha \iff \left(\frac{c}{\theta_0}\right)^n = 1-\alpha$

$\Rightarrow c = \left[(1-\alpha)^{1/n}\right]\theta_0.$

Hence, the power function of $\phi_1(\cdot)$ for $\theta > \theta_0$ is given by

$$\gamma_{\phi_1}(\theta) = E_\theta\, \phi_1(\underset{\sim}{X}) = P_\theta(X_{(n)} > c)$$

$$= 1 - \left[P_\theta(X_1 \leq c)\right]^n$$



$$= 1 - \left(\frac{c}{\theta}\right)^n$$

$$= 1-(1-\alpha)\left(\theta_0/\theta\right)^n$$

Next, note that for any $\phi \in \mathcal{C}_\alpha$, its power function, for $\theta > \theta_0$, is given by

$$E_\theta\, \phi(\underset{\sim}{X}) = \theta^{-n}\int_{0<x_{(1)}<x_{(n)}<\theta} \phi(\underset{\sim}{x})\, dx$$

$$= \theta^{-n}\left[\int_{0<x_{(1)}<x_{(n)}<\theta_0} \phi(\underset{\sim}{x})\, d\underset{\sim}{x} + \int_{\theta_0<x_{(n)}<\theta} \phi(\underset{\sim}{x})\, d\underset{\sim}{x}\right]$$

$$= (\theta_0/\theta)^n\left[E_{\theta_0}\phi(\underset{\sim}{X})\right] + P_\theta(\theta_0 < X_{(n)} < \theta)$$

$$\left(\text{since } \phi(\underset{\sim}{x}) \equiv 1 \ \ \forall\, x_{(n)} > \theta_0\right)$$

$$= \alpha \left(\theta_0/\theta\right)^n + 1 - \left(\theta_0/\theta\right)^n$$

$$= 1 - (1-\alpha)\left(\theta_0/\theta\right)^n = \gamma_{\phi_1}(\theta) \qquad \forall \theta > \theta_0.$$

$\Rightarrow$ Any $\phi \in \mathcal{C}_\alpha$ is a UMP size α test.

(c) By the MLR property, a UMP size α test for testing $H_0: \theta = \theta_0$ against $H_1: \theta < \theta_0$ is given by

$$\phi_2(x) = \begin{cases} 1 & \text{if } x_{(n)} < c_2 \\ 0 & \text{if } x_{(n)} \geq c_2 \end{cases}$$

where $E_{\theta_0} \phi_2(\underset{\sim}{X}) = \alpha \iff P_{\theta_0}(X_{(n)} < c_2) = \alpha$

$$\iff \left(\frac{c_2}{\theta_0}\right)^n = \alpha \iff c_2 = \theta_0 \alpha^{1/n}.$$

Next define the test ϕ^* by

$$\phi^*(x) = \begin{cases} 1 & \text{if } x_{(n)} < \theta_0 \alpha^{1/n} \ \underline{\text{or}} \ x_{(n)} > \theta_0 \\ 0 & \text{otherwise.} \end{cases}$$

Then, check that

[1] $E_{\theta_0} \phi^*(\underset{\sim}{X}) = \alpha$,

[2] $E_\theta \phi^*(\underset{\sim}{X}) = \gamma_{\phi_1}(\theta)$ for all $\theta > \theta_0$

[3] $E_\theta \phi^*(X) = E_\theta \phi_2(X)$ for all $\theta < \theta_0$

Hence, conclude that ϕ^* is a size α UMP test for $H_0: \theta = \theta_0$ vs. $H_1: \theta \neq \theta_0$.