PhD Prelim Exam
THEORY

Summer 2006
(Given on 7/20/06)

1. $[5+5]$ Let ν be a nonnegative r.v. Show that

$$E\nu = \int_0^\infty P(\nu \geq t)dt = \int_0^\infty P(\nu > t)dt.$$

Conclude that $E\nu < \infty$ iff $\sum_{k=1}^\infty P(\nu \geq k) < \infty$.

2. $[10+5]$ State the Borel-Cantelli lemmas and prove the first one.

3. $[5+8+7]$

(a) Let $\{Y_n\}_{n\geq 1}$ be independent r.v. such that for each n, $P(Y_n = 1) = \frac{1}{n} = 1 - P(Y_n = 0)$. Show that Y_n converges to zero in probability but not with probability one. (Hint: Use the second Borel Cantelli lemma.)

(b) Let $\{\nu_i\}_{i\geq 1}$ be a sequence of iidrv. Show that $\frac{\nu_k}{k} \to 0$ w.p.1 iff $E|\nu_1| < \infty$. (Hint: Use both Borel-Cantelli lemmas.)

(c) Show that if $Y_n \xrightarrow{p} Y$ then $Y_n \to Y$ in distribution. Show by an example that the converse is not true.

4. $[5+15]$ Let $\{\epsilon_i\}_{i\geq 1}$ be a sequence of iidrv with $E(\log|\epsilon_1|)^+ < \infty$. Let $|\rho| < 1$. Let $\hat{X}_n \equiv \sum_{j=1}^n \rho^{j-1}\epsilon_j, n \geq 1$.

(a) Show that $\lim_n \hat{X}_n \equiv \hat{X}_\infty$ exists w.p.1. (Hint: Use 3b to show $|\epsilon_j| = 0(\lambda^j)$ w.p.1 for any $\lambda > 1$ and the fact $|\rho| < 1$.)

(b) Now assume that $E\epsilon_1^2 < \infty$ and $E\epsilon_1 = 0$ and $|\rho| < 1$. Show that $\{\hat{X}_n\}_{n\geq 1}$ is Cauchy in mean square, that is,
$$E(\hat{X}_n - \hat{X}_m)^2 \to 0 \text{ as } n, m \to \infty.$$
Conclude that $E\hat{X}_\infty^2 < \infty$ and $E(\hat{X}_n - \hat{X}_\infty)^2 \to 0$ as $n \to \infty$.

5. $[5+5+5+5]$ Let $\{X_n\}_{n\geq 0}$ be a sequence of random variables defined by the stochastic recurrence equation
$$X_n = \rho X_{n-1} + \epsilon_n, n \geq 1$$
where $\{\epsilon_j\}_{j\geq 1}$ are iidrv independent of X_0 and ρ is a constant.

(a) Express X_n in terms of $\rho, X_0, \epsilon_1, \epsilon_2, \ldots, \epsilon_n$.

(b) Show that $X_n - \rho^n X_0$ and $\hat{X}_n \equiv \sum_1^n \rho^{j-1}\epsilon_j$ have the same distribution.

(c) Assume $|\rho| < 1$ and $E(\log|\epsilon_1|)^+ < \infty$. Show that X_n converges in distribution to $\hat{X}_\infty$ as defined in 4(a).

(d) Show that X_n does not converge in probability by considering $X_n - X_{n-1}$.

6. $[7+8]$

(a) Consider the problem of estimating ρ based on the data $\{X_j : 0 \leq j \leq n\}$ defined on a probability space (Ω, F, P). Derive the least squares estimator $\hat{\rho}_n$ of ρ defined as the minimizer of $\sum_{j=0}^{n-1}(X_{j+1} - \rho X_j)^2$. Show that $\hat{\rho}_n$ is F measurable.

(b) Show that if $\frac{\sum_{j=0}^{n-1} X_j\epsilon_{j+1}}{\left(\sum_0^{n-1} X_j^2\right)} \xrightarrow{p} 0$ then $\hat{\rho}_n$ is a <u>weakly consistent estimator</u> of ρ, i.e. $\hat{\rho}_n \xrightarrow{p} \rho$.

1. If γ is defined on a p. space (Ω, B, P) then

$$E\gamma = \int_\Omega \gamma \, dP = \int_\Omega \left(\int_{R^+} I_{[0, \gamma(\omega)]}(t) \, dt \right) dP$$

$$= \int_{R^+} \left(\int_\Omega I_{[0, \gamma(\omega)]}(t) \, dP \right) dt \qquad \text{by Tonelli}$$

$$= \int_{R^+} P(\gamma(\omega) \geq t) \, dt$$

Similarly $\quad E \int_\Omega \gamma(\omega) \, dP = \int_\Omega \left(\int_{R^+} I_{[0, \gamma(\omega))}(t) \right) dP$

$$= \int_{R^+} \left(\int_\Omega \, dP \right) dt = \int_{R^+} P(\gamma(\omega) > t) \, dt .$$

2. BW

3. a) Let $A_n \equiv (Y_n = 1)$. Then $\sum P(A_n) = \sum \frac{1}{n} = \infty$

Since $\{A_n\}$ are indep by Borel Cantelli 2nd lemma

$P(\text{infinitely many } A_n\text{'s happen}) = 1$

ie $P(\varlimsup X_n = 1) = 1$ Thus $X_n \not\to 0$ w.p.1

But $P(|X_n| > \epsilon) = \frac{1}{n} \qquad (\text{for any } \epsilon < 1)$

$\longrightarrow 0 \quad \text{as } n \to \infty. \quad \text{So} \quad X_n \xrightarrow{P} 0 .$

b) $E|V_i| < \infty \implies \forall \epsilon > 0, \quad E\frac{|V_i|}{\epsilon} < \infty$

$\implies \sum_k P(|V_i| > \epsilon k) < \infty$

$\implies \sum_k P\left(\left|\frac{V_k}{k}\right| > \epsilon \right) < \infty$

By Borel-Cantelli's first lemma

$\implies P\left(\varlimsup \left|\frac{V_k}{k}\right| \leq \epsilon \right) = 1 .$

This being true $\forall \epsilon > 0 \implies P\left(\lim \left|\frac{V_k}{k}\right| = 0 \right) = 1$

Conversely, if $\frac{Y_k}{k} \to 0$ w.p.1 then

$\forall \epsilon > 0$, the $P(\text{infinitely many } A_k\text{'s happen}) = 0$

where $A_k = \left(\left| \frac{Y_k}{k} \right| > \epsilon \right)$

By Borel-Cantelli second lemma $\sum P(A_k) < \infty$

$\implies \sum_k P(|Y_1| > \epsilon k) < \infty \implies E \frac{|Y_1|}{\epsilon} < \infty$

(by point 1)

$\implies E|Y_1| < \infty.$

A. a) Since $E(\log|\epsilon_1|)^+ < \infty$,

$\frac{(\log|\epsilon_k|)^+}{k} \to 0$ w.p.1.

So $\forall \eta > 0$, $\log|\epsilon_k| < \eta k$ for all large k w.p.1

$\implies |\epsilon_k| = O((e^\eta)^k)$ w.p.1

b) $\implies |\epsilon_k| < C(\omega)(e^\eta)^k$ for large k w.p.1

$\implies |\rho^k \epsilon_k | < C(\omega)(\rho e^\eta)^k$

Now since $|\rho| < 1$, choose $\eta > 0 \implies \rho e^\eta < 1$.

b) $E(\hat{X}_n - \hat{X}_m)^2 = \sum V(\hat{X}_n - \hat{X}_m) / \text{since } E\epsilon_i = 0)$

$= \sum_{n+1}^{m} (\rho^2)^i (E\epsilon_i^2) \to 0$ as $n, m \to \infty$, $n < m$,

Let $X_0, (\epsilon_j)_{j \geq 1}$ be defined on (Ω, B, P).

Then $\{\hat{X}_n\}_{n \geq 1} \subset L^2(\Omega, B, P)$ which is complete

So $\hat{X}_n$ converges in L^2 to some Y say.

$E \epsilon_1^2 < \infty$

But Since $\Rightarrow E(\log|\epsilon_1|)^+ < \infty$ & $|\rho| < 1$ by (4)(a)

$\hat{X}_n \longrightarrow \hat{X}_\omega$ wp 1.   So $Y = \hat{X}_\omega$

So $\hat{X}_n \to \hat{X}_\omega$ in L^2.

5. a) $X_n = \rho X_{n-1} + \epsilon_n$

$= \rho(\rho X_{n-2} + \epsilon_{n-1}) + \epsilon_n$

$= \rho^2 (\rho X_{n-3} + \epsilon_{n-3}) + \rho \epsilon_{n-1} + \epsilon_n$

$= \rho^n X_0 + \rho^{n-1} \epsilon_1 + \rho^{n-2} \epsilon_2 + \cdots + \rho \epsilon_{n-1} + \epsilon_n$

$\Rightarrow$ b) $X_n - \rho^n X_0 = \sum_{j=1}^{n} \rho^{n-j} \epsilon_j \sim \hat{X}_n = \sum_{0}^{n-1} \rho^j \epsilon_j$

Since $(\epsilon_j)_{j \geq 1}$ are iid.

c) Since X_n & $\hat{X}_n$ have the same distn

$+ \hat{X}_n \longrightarrow \hat{X}_\omega$ wp1 $\Rightarrow \hat{X}_n \longrightarrow \hat{X}_\omega$ in distribution

It follows that $X_n \longrightarrow \hat{X}_\omega$ in distribution.

d) $X_n - X_{n-1} = (\rho-1) X_{n-1} + \epsilon_n$

Also X_{n-1} & ϵ_n are indep. Since $X_{n-1} \xrightarrow{d} \hat{X}_\omega$

$X_n - X_{n-1} \xrightarrow{d} (\rho-1) \hat{X}_\omega + \epsilon_1$ where $\hat{X}_\omega$ & ϵ_1 are

indep.   Thus $X_n - X_{n-1} \not\to 0$ in distribution

So $X_n - X_{n-1} \not\xrightarrow{p} 0 \Rightarrow \{X_n\}$ cannot converge

in   probability

6 a) $\dfrac{d}{dp}\left(\sum_0^{n-1}(X_{j+1}-PX_j)^2\right) = 2\sum_0^{n-1}(X_{j+1}-PX_j)X_j = 0$

$\dfrac{d^2}{dp^2} = -2\beta\sum_0^{n-1}x_i^2 < 0$

So $P_n \equiv \dfrac{\sum_0^{n-1} X_{j+1}X_j}{\sum_0^{n-1}X_j^2}$ is the minimizer.

Since it is a rational fn of $X_0 X_1 X_2 \ldots X_n$

it is a r.v.

b)   Since $X_{j+1} = P X_j + \epsilon_{j+1}$

$P_n = \dfrac{P\sum_0^{n-1}x_j^2 + \sum_0^{n-1}X_j\,\epsilon_{j+1}}{\sum_0^{n-1}x_j^2}$

$= P + \dfrac{\sum_0^{n-1}X_j\,\epsilon_{j+1}}{\sum_0^{n-1}x_j^2}$

$\xrightarrow{p} P$ since $\dfrac{\sum_0^{n-1}X_j\,\epsilon_{j+1}}{\sum_0^{n-1}x_j^2} \xrightarrow{p} 0$ by hyp.

ST643   Prelim Question,   2006

The notion of conditional independence.

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{G}_i$, for $i = 1, 2, 3$, be three σ-algebras of events. Then $\mathcal{G}_1$ and $\mathcal{G}_2$ are said to be **conditionally independent given $\mathcal{G}_3$** if **for any $A_i \in \mathcal{G}_i$, $i = 1$ and 2,**

$$P(A_1 A_2 | \mathcal{G}_3) = P(A_1 | \mathcal{G}_3) P(A_2 | \mathcal{G}_3) \quad \text{w.p.1}$$

Let Z_i, $i = 1, 2, 3$, be three random variables defined on the probability space and $\sigma(Z_i)$ be the σ-algebra generated by Z_i. Then Z_1 and Z_2 are said to be conditionally independent given Z_3 if $\sigma(Z_1)$ and $\sigma(Z_2)$ are conditionally independent given $\sigma(Z_3)$.

1. (10 Marks) Prove that if $\mathcal{G}_1$ and $\mathcal{G}_2$ are conditionally independent given $\mathcal{G}_3$ then

$$P(A_1 | \sigma(\mathcal{G}_2 \cup \mathcal{G}_3)) = P(A_1 | \mathcal{G}_3) \quad \text{w.p.1} \tag{1}$$

where $\sigma(\mathcal{G}_2 \cup \mathcal{G}_3)$ is the smallest σ-algebra containing both $\mathcal{G}_1$ and $\mathcal{G}_2$.

Missing at Random

Let $Y \in R$ be the outcome random variable and $X \in R^p$ be a covariate random variable of a statistical experiment. It is common that the outcome variable Y can be missing for the experiment. Let δ be the missing indicator variable such that

$$\delta = \begin{cases} 1, & \text{if } Y \text{ is observed;} \\ 0, & \text{if } Y \text{ is missing} \end{cases}$$

An important notion in analyzing data with missing values is the following notion of Strongly Ignorable Missing at Random or Missing at Random (MAR):

$$Y \text{ and } \delta \text{ are conditionally independent given } X \tag{2}$$

2. (5 marks) Show that the MAR implies that $E(\delta | Y, X) = E(\delta | X)$.

Let $w(X) = E(\delta | X)$ which is the so-called propensity score function.

3. (5 marks) Show that MAR implies $P(\delta = 1 | X, w(X)) = P(\delta = 1 | w(X))$.

4. (10 marks) Let $b(\cdot)$ be a measurable function from $R^p \to R$. Prove that $P(\delta = 1 | X, b(X)) = P(\delta = 1 | b(X))$ if and only if $b(x)$ is finer than $w(x)$ in the sense that $w(x) = f(b(x))$ for some measurable function f.

Suppose that the above said statistical experiment is repeated independently for $m + n$ times. The observed data consist of two parts:

$$(\delta_1 = 1, Y_1, X_1), ..., (\delta_m = 1, Y_m, X_m)$$

and

$$(\delta_{m+1} = 0, ?, X_{m+1}), ..., (\delta_{m+n} = 0, ?, X_{m+n}).$$

where ? stands for missing values. Here m is a positive integer. Let $N = n + m$.

Suppose that a parametric model for the propensity function $w(x)$ is available, i.e.

$$w(x) = w(x, \theta_0),$$

where the form of w is known up to a unknown r-dimensional parameter $\theta \in R^r$ whose true value is θ_0. The conditional binomial log likelihood of θ is

$$\ell_B(\theta) = \sum_{i=1}^{N} [\delta_i \log w(X_i, \theta) + (1 - \delta_i) \log\{1 - w(X_i, \theta)\}]. \tag{3}$$

Let $\hat{\theta}_N$ be a solution of the conditional likelihood score equation:

$$\frac{\partial \ell_B(\theta)}{\partial \theta} = 0. \tag{4}$$

Let us assume in additional to MAR the following conditions:

C1: There exists an open neighbourhood of the true parameter θ_0, say Θ_{θ_0}, such that for any x, and $\theta \in \Theta_{\theta_0}$, $0 < C_1 \leq w(x; \theta) \leq C_2 < 1$ for two fixed constants C_1 and C_2.

C2: The propensity score $w(x, \theta)$ is twice continuously differentiable for each θ in Θ_θ

C3: Both $E\left(\frac{\partial w(X_i, \theta)}{\partial \theta} \frac{\partial w(X_i, \theta)^T}{\partial \theta}\right)$ and $E\left(\frac{\partial^2 w(X_i, \theta)}{\partial \theta^2}\right)$ are finite matrices of full ranks for all $\theta \in \Theta_\theta$.

C4: $\hat{\theta}_N$ converges to θ_0 in probability as $N \to \infty$.

Let

$$q_N(\theta) = N^{-1} \sum_{i=1}^{N} \frac{\delta_i - w(X_i, \theta)}{w(X_i, \theta)\{1 - w(X_i, \theta)\}} \frac{\partial w(X_i, \theta)}{\partial \theta}.$$

2

5. (5 marks) Show that as $N \to \infty$ $q_N(\theta_0)$ is asymptotically normally distributed and derive the mean and variance of the asymptotic distribution.

6. (5 marks) Let $w_i(\theta) = w(X_i, \theta)$ and

$$
\begin{aligned}
B_N(\theta) &= N^{-1} \sum_{i=1}^{N} \left[\frac{\delta_i - w_i(\theta)}{w_i(\theta)\{1 - w_i(\theta)\}} \right] \left[\frac{\partial^2 w_i(\theta)}{\partial \theta^2} - \frac{\{1 - 2w_i(\theta)\}}{w_i(\theta)(1 - w_i(\theta))} \frac{\partial w_i(\theta)}{\partial \theta} \frac{\partial w_i^T(\theta)}{\partial \theta} \right] \\
&\quad + N^{-1} \sum_{i=1}^{N} \left[\frac{1}{1 - w_i(\theta)} \frac{\partial w_i(\theta)}{\partial \theta} \frac{\partial w_i^T(\theta)}{\partial \theta} \right].
\end{aligned}
$$

Derive the limit to which $B_N(\theta_0)$ converges in probability as $N \to \infty$.

7. (10 Marks) Establish the asymptotic normality of $\hat{\theta}_N$ and derive the mean and variance of the asymptotic distribution.

1. $\forall \; A_i \in \mathcal{G}_i \; , \; i=1,2,3 \; ,$ $P(A_1 A_2 | \mathcal{G}_3)$

$$\int_{A_2 A_3} I_{A_1} \, dp = \int_{A_3} I_{A_1 A_2} \, dp \overset{\text{Def}}{=} \int_{A_3} E\left(I_{A_1 A_2} | \mathcal{G}_3\right) \, dp$$

$$\overset{\text{Cond.}}{\underset{\text{indpt}}{=\!=}} \int_{A_3} P(A_1 | \mathcal{G}_3) \, P(A_2 | \mathcal{G}_3) \, dp = \int_{A_3} E\left\{ P(A_1|\mathcal{G}_3) \, I_{A_2} \right\}^{|\mathcal{G}_3} \, dp$$

$$= \int_{A_3} I_{A_2} \, P(A_1 | \mathcal{G}_3) \, dp = \int_{A_2 A_3} P(A_1 | \mathcal{G}_3) \, dp \qquad (1)$$

Let $\quad D = \left\{ A_2 A_3 \; | \; A_2 \in \mathcal{G}_2, \; A_3 \in \mathcal{G}_3 \right\}$ which forms

a $\;\Pi-$ Class $\quad$ and $\quad \sigma(D) = \sigma(\mathcal{G}_2 \cup \mathcal{G}_2)$.

Thus from a theorem given in 643

$$E\left\{ I_{A_1} | \sigma(\mathcal{G}_2 \cup \mathcal{G}_3) \right\} = E\left\{ I_{A_1} | \mathcal{G}_3 \right\}$$

$$\Rightarrow \quad P\left\{ A_1 | \sigma(\mathcal{G}_2 \cup \mathcal{G}_3) \right\} = P(A_1 | \mathcal{G}_3)$$

2. Let $\mathcal{G}_1 = \sigma(\bar{\sigma})$, $\;\mathcal{G}_2 = \sigma(Y)$, $\;\mathcal{G}_3 = \sigma(X)$

$$A_1 = \{\bar{\sigma} = 1\} \qquad \text{and Apply to Conclusion of 1}$$

3. As $\sigma(x, w(x)) = \sigma(x)$, $P(\delta=1 | x, w(x)) = P(\delta=1|x) = E(\delta|x)$,

<div align="right">(3.1)</div>

Since $\sigma(w(x)) \subset \sigma(x)$,

$$E(\delta | w(x)) = E\{ E(\delta | x) | w(x) \}$$

$$= E(w(x) | w(x)) = w(x)$$

$$= E(\delta | x) \qquad (3.2)$$

$$\Rightarrow \quad P(\delta=1 | x, w(x)) = P(\delta=1 | w(x))$$

Hence, given $w(x)$, δ & X are conditionally indpt.

4. "$\Leftarrow$" If $b(x)$ is finer than $w(x)$, i.e

$$\sigma(b(x)) \supset \sigma(w(x)), \quad \text{repeat above } (3.1) \text{ & } (3.2)$$

$$\Rightarrow \quad P(\delta=1 | x, b(x)) = E(\delta|x) = w(x) \qquad (3.1)'$$

and

$$E(\delta | b(x)) = E(w(x) | b(x)) = w(x)$$

$$\text{as } \sigma(w(x)) \subset \sigma(b(x)) \qquad \Rightarrow \text{ required}.$$

"⟹" Suppose that $b(x)$ is not finer than $w(x)$, i.e

$$\exists \; x_1, x_2 \quad s.t \quad w(x_1) \neq w(x_2) \quad but \quad b(x_1) = b(x_2)$$

Hence
$$P(\delta=1 \mid X=x_1) \neq P(\delta=1 \mid X=x_2)$$

However, as $b(x_1) = b(x_2)$, from hypothesis

$$w(x_1) = P(\delta=1 \mid x_1, b(x_1)) = P(\delta=1 \mid b(x_1)) = P(\delta=1 \mid b(x_2))$$

$$= P(\delta=1 \mid x_2, b(x_2)) = w(x_2), \quad \text{contradiction}.$$

5. As q_n is a sum of IID r. vectors,

$$E\left(\frac{\delta_i - W(x_i;\theta)}{W(x_i,\theta)\{1-W(x_i,\theta)\}}\frac{\partial W}{\partial\theta}\right) = E\left(\frac{\frac{\partial W}{\partial\theta}}{W(x_i,\theta)(1-W(x_i,\theta))}E(\delta_i - W_i | x_i)\right)$$

$$= 0 \qquad \text{Since} \qquad E(\delta_i | x_i) = W(x_i,\theta) \stackrel{\wedge}{=} W_i$$

and

$$Var\left(\frac{\delta_i - W(x_i,\theta)}{W_i(1-W_i)}\frac{\partial W}{\partial\theta}\right) = E\left(\frac{(\delta_i - W_i)^2}{W_i^2(1-W_i)^2}\frac{\partial W(x_i)}{\partial\theta}\frac{\partial^T W(x)}{\partial\theta}\right)$$

$$= E\left(\frac{\frac{\partial W(x_i,\theta)}{\partial\theta}\frac{\partial^T W(x_i,\theta)}{\partial\theta}}{W_i(1-W_i)}\right)$$

which is bounded away from infinite and

above from zero due to C1 & C3.

we have

$$\sqrt{n}\, q_n \stackrel{d}{\longrightarrow} N\left(0, E\frac{\frac{\partial W}{\partial\theta}\frac{\partial W^T}{\partial\theta}}{W(1-W)}\right)$$

as $\min(m,t) \to \infty$.

6   B_n is a sum of IID r. vectors

$$E(B_n) = E\left(\frac{1}{W(x_i,\theta)}\frac{\partial W(x_i,\theta)}{\partial\theta}\frac{\partial W^T(x_i,\theta)}{\partial\theta}\right) \stackrel{\wedge}{=} \Lambda_\theta$$

Hence $B_n \stackrel{P}{\longrightarrow} \Lambda_\theta$ as $\min\{m,t\} \to \infty$

7. $\dfrac{\partial l_B|\theta}{\partial \theta} = \sum_1^n \left\{ \dfrac{\delta_i \frac{\partial w(x_i,\theta)}{\partial \theta}}{w(x_i,\theta)} - \dfrac{(1-\delta_i)}{1-w(x_i,\theta)} \dfrac{\partial w(x_i,\theta)}{\partial \theta} \right\}$

$= \sum_1^n \dfrac{\delta_i - w(x_i,\theta)}{w(x_i,\theta)\{1-w_d(x_i,\theta)\}} \dfrac{\partial w(x_i,\theta)}{\partial \theta} = n\, g_n(\theta)$

Note that

As

$\dfrac{\partial l_B(\theta)}{\partial \theta}\bigg|_{\theta=\hat\theta_n} = 0$

$\dfrac{\partial^2 l_B(\theta)}{\partial \theta^2} = n\, \dfrac{\partial g_n(\theta)}{\partial \theta} = n\, B_n(\theta)$

From Taylor expansion

$0 = g_n(\theta_0) + B_n(\tilde\theta_n)(\hat\theta_n - \theta)$ for some $\tilde\theta_n$ between

$\hat\theta_n$ & θ. As $\hat\theta_n \xrightarrow{P} \theta \Rightarrow \tilde\theta_n \xrightarrow{P} \theta$ too

C2 $\Rightarrow$ $B_n(\theta')$ is Continuous in (H).

Apply (Continuous) mapping Theorem and Part 6

$B_n(\tilde\theta_n) \xrightarrow{P} \Lambda_\theta > 0$ positive definite

Part 5 $\Rightarrow$ $\sqrt{n}\, g_n(\theta) \xrightarrow{d} N\left(0, E\dfrac{\frac{\partial w}{\partial \theta}\frac{\partial w^T}{\partial \theta}}{w(1-w)}\right)$

Apply Slutsky Th $\sqrt{n}(\hat\theta_n - \theta) \xrightarrow{d} N\left(0, \Lambda_\theta^{-1} E\dfrac{\frac{\partial w}{\partial \theta}\frac{\partial w^T}{\partial \theta}}{w(1-w)} \Lambda_\theta^{-1}\right)$

a) Prove the following simple lemma. (You may use the lemma in what follows even if you can not prove it.).

Lemma Suppose that F is a continuous distribution with probability density function on $(-\infty,\infty)$

$$f(x) = C\exp\left(ax^2 + bx\right)$$

(for real numbers $a < 0, C$, and b). Then F is Normal with mean $\mu = -b/2a$ and variance $\sigma^2 = -1/2a$.

---

b) Suppose that (T,W) is a random vector such that $T \sim N\left(\delta,\gamma^2\right)$ and that conditioned on $T = t$, $U \sim N\left(t,\eta^2\right)$. Find the conditional distribution of T given that $U = u$.

c) Now suppose that $\mu \sim N\left(0,\gamma^2\right)$ and that conditioned on μ, variables $W_1, W_2,\ldots,W_n$ are iid $N(\mu,1)$. Let $\overline{W_n} = \frac{1}{n}\sum_{i=1}^{n} W_i$. What is the conditional distribution of μ given that $\overline{W_n} = \overline{w}_n$? (Hint: what is the conditional distribution of $\overline{W_n}$ given μ?) Evaluate the function of w,

$$m_n(w) \equiv E\left[\mu \,|\, \overline{W_n} = w\right]$$

d) Now suppose that μ is a fixed unknown quantity and that the variables $W_1, W_2,\ldots,W_n$ are iid $N(\mu,1)$. With $m_n(w)$ as defined in c) consider the random quantity $m_n\left(\overline{W_n}\right)$. Show that this converges to a constant in probability and identify the limit.

Now suppose that n values $0 = x_1 < x_2 < \cdots < x_n = 1$ are known, and that for two real numbers μ_1 and μ_2 and a $c \in (0,1)$ we define the function

$$\mu(x) \equiv \mu_1 I[x < c] + \mu_2 I[x \geq c]$$

Suppose further that (given the parameters μ_1, μ_2, and c) variables $Y_1, Y_2,\ldots,Y_n$ are independent Normal random variables with variance 1, and means

$$EY_i = \mu_i = \mu(x_i)$$

(Y_i has mean μ_1 if $x_i < c$, and otherwise has mean μ_2). Consider the statistical problem with parameter vector (μ_1,μ_2,c).

e) Write out a likelihood for this problem, $L_n(\mu_1, \mu_2, c)$. For fixed c, what values of μ_1 and μ_2 maximize $L_n(\cdot, \cdot, c)$? Call these $\hat{\mu}_1(c)$ and $\hat{\mu}_2(c)$ and use the notations

$$f(y|\mu) \text{ for the } N(\mu, 1) \text{ pdf} \quad \text{and} \quad n_1(c) = \sum_{i=1}^{n} I[x_i < c]$$

f) Is there a unique maximum likelihood estimator for the parameter vector (μ_1, μ_2, c)? Explain carefully.

g) As explicitly as is possible, give a likelihood ratio test statistic for testing the hypothesis $H_0 : c = .5$ versus $H_a : c \neq .5$.

Consider a Bayes version of the inference problem for (μ_1, μ_2, c). In particular, suppose that we give (μ_1, μ_2, c) a (prior) distribution G under which the parameters are independent with

$$\mu_1 \sim N(0, \gamma^2)$$
$$\mu_2 \sim N(0, \gamma^2)$$
$$c \sim U(0, 1)$$

Let $g(\mu | 0, \gamma^2)$ stand for the $N(0, \gamma^2)$ probability density and use the notation

$$h_1(c, Y_1, Y_2, \ldots, Y_n) = \int_{-\infty}^{\infty} \left(\prod_{i=1}^{n_1(c)} f(Y_i | \mu) \right) g(\mu | 0, \gamma^2) d\mu \quad \text{and}$$

$$h_2(c, Y_1, Y_2, \ldots, Y_n) = \int_{-\infty}^{\infty} \left(\prod_{i=n_1(c)+1}^{n} f(Y_i | \mu) \right) g(\mu | 0, \gamma^2) d\mu$$

h) Evaluate $E[\mu_1 | c, Y_1, Y_2, \ldots, Y_n]$

i) Write (in terms of the functions h_1 and h_2) a conditional pdf for c given $Y_1, Y_2, \ldots, Y_n$. (Notice that this pdf is constant on each interval (x_{i-1}, x_i).)

j) Use your answers to **h)** and **i)** to evaluate $E[\mu_1 | Y_1, Y_2, \ldots, Y_n]$

k) For an integer $1 < i < n$ evaluate $E[\mu(x_i) | Y_1, Y_2, \ldots, Y_n]$

a)
$$f(x) = c \exp\left(a\left[\left(x+\frac{b}{2a}\right)^2 - \left(\frac{b}{2a}\right)^2\right]\right)$$

$$= \underbrace{c \exp\left(-\left(\frac{b}{2a}\right)^2\right)}_{\text{a constant}} \underbrace{\exp\left(a\left(x - \left(-\frac{b}{2a}\right)\right)^2\right)}_{\substack{\text{proportional to a} \\ \text{normal pdf with mean}}}$$

$$-\frac{b}{2a} \text{ and } \frac{-1}{2\,\text{variance}} = a$$

i.e. variance $= -\frac{1}{2a}$

b)   The joint pdf for (T, U) is

$$f_{T,U}(t,u) = f_T(t)\, f_{U|T}(u|t)$$

$$= \frac{1}{\sqrt{2\pi\gamma^2}} \exp{-\frac{1}{2\delta^2}(t-\delta)^2} \; \frac{1}{\sqrt{2\pi\eta^2}} \exp{-\frac{1}{2\eta^2}(u-t)^2}$$

For fixed u, this is a function of t proportional to

$$f_{T|U}(t|u) \propto \exp\left(-\frac{1}{2}\left(\frac{1}{\gamma^2}+\frac{1}{\eta^2}\right)t^2 + \left(\frac{\delta}{\gamma^2}+\frac{u}{\eta^2}\right)t\right)$$

Now apply the lemma. It implies that conditioned on $U=u$, T is normal with

$$\text{mean } \mu_{T|U=u} \quad \frac{\frac{1}{\eta^2}u + \frac{1}{\gamma^2}\delta}{\frac{1}{\eta^2} + \frac{1}{\gamma^2}}$$

and   variance $\sigma^2_{T|U=u} = \left(\frac{1}{\eta^2} + \frac{1}{\gamma^2}\right)^{-1}$

c) Note that conditioned on μ, $\overline{W}_n \sim N(\mu, \frac{1}{n})$.
Apply part b) to conclude that conditioned on $\overline{W}_n = w$, μ is normal with mean

$$m_n(w) = \frac{nw + \frac{0}{\gamma^2}}{n + \frac{1}{\gamma^2}} = \left(\frac{1}{1 + \frac{1}{n\gamma^2}}\right) w = \left(\frac{n\gamma^2}{n\gamma^2 + 1}\right) w$$

and variance

$$V_n = \left(n + \frac{1}{\gamma^2}\right)^{-1} = \left(\frac{n\gamma^2}{n\gamma^2 + 1}\right) \frac{1}{n}$$

d) For a fixed μ, $\overline{W}_n \xrightarrow{P} \mu$ by LLN. Since for any fixed γ^2,

$$\frac{n\gamma^2}{n\gamma^2 + 1} \longrightarrow 1$$

The continuity of the product implies that

$$m_n(\overline{W}_n) = \frac{n\gamma^2}{n\gamma^2 + 1} (\overline{W}_n) \xrightarrow{P} \mu$$

(and we have consistency of $m_n(\overline{W}_n)$ for μ).

e)
$$L_n(\mu_1, \mu_2, c) = \prod_{i \le n_1(c)} f(y_i | \mu_1) \prod_{i > n_1(c)} f(y_i | \mu_2)$$

For fixed c, This is clearly maximized by choosing μ_1 and μ_2 to minimize respectively

$$\sum_{i=1}^{n_1(c)} (y_i - \mu_1)^2 \quad \text{and} \quad \sum_{n_1(c)+1}^{n} (y_i - \mu_2)^2$$

Standard arguments then imply that

$$\hat{\mu}_1(c) = \bar{y}_1(c) \equiv \frac{1}{n_1(c)} \sum_{i=1}^{n_1(c)} y_i \quad \text{and} \quad \hat{\mu}_2(c) = \bar{y}_2(c) = \frac{1}{n - n_1(c)} \sum_{i=n_1(c)+1}^{n} y_i$$

f) Consider maximizing The profile likelihood

$$L_n^*(c) \equiv L_n(\hat{\mu}_1(c), \hat{\mu}_2(c), c)$$

as a function of c. As c runs from 0 to 1, This changes values only as one crosses an x_i, i.e. this is constant on all intervals (x_{i-1}, x_i) and is, incidentally, right-continuous on $(0,1)$. So if x_{i_0} is any element of $(0,1)$ where $L_n^*(\cdot)$ is maximized, $L_n^*(\cdot)$ is constant on $[x_{i_0}, x_{i_0+1})$ and There is no unique MLE of (μ_1, μ_2, c).

g) A likelihood ratio statistic for testing $H_0 : c = .5$ vs $H_a : c \neq .5$ is

$$\lambda = \frac{\sup_c L_n(\hat{\mu}_1(c), \hat{\mu}_2(c), c)}{L_n(\hat{\mu}_1(.5), \hat{\mu}_2(.5), .5)}$$

Note That maximization over c can be done by searching over The set of discrete values, $x_1, x_2, \ldots, x_n$. Thus

$$\lambda = \frac{\max_{j=1,\ldots,n-1} \exp\left(-\frac{1}{2} \sum_{i=1}^{j} (y_i - \bar{y}_1(x_j))^2 - \frac{1}{2} \sum_{i=j+1}^{n} (y_i - \bar{y}_2(x_j))^2\right)}{\exp\left(-\frac{1}{2} \sum_{\substack{i \text{ s.t.} \\ x_i < .5}} (y_i - \bar{y}_1(.5))^2 - \frac{1}{2} \sum_{\substack{i \text{ s.t.} \\ x_i \geq .5}} (y_i - \bar{y}_2(.5))^2\right)}$$

1) The joint pdf for $Y_1, Y_2, \ldots, Y_n, \mu_1, \mu_2, c$ is

$$\left(\prod_{i=1}^{n} f(y_i \mid \mu(z_i)) \right) g_1(\mu_1 \mid 0, \delta^2) \, g(\mu_2 \mid 0, \delta^2) \, I[0 < c < 1] \quad \circledast$$

depends on μ_1, μ_2, c

For fixed c, if I integrate out μ_2 I'm left with

$$\left(\prod_{i=1}^{n_1(c)} f(y_i \mid \mu_1) \right) g(\mu_1 \mid 0, \sigma^2) \, h_2(c, y)$$

which, as a function of μ_1, is proportional to

$$\exp\left(-\frac{1}{2}\left(n_1(c)\mu_1^2 - 2\mu_1 \sum_{i=1}^{n_1(c)} y_i \right) \right) \exp - \frac{1}{2\delta^2} \mu^2$$

i.e. by the Lemma is Normal with mean

$$\left(-\sum_{i=1}^{n_1(c)} y_i \right) \bigg/ 2\left(-\frac{n_1(c)}{2} - \frac{1}{2\delta^2} \right) = \bar{y}_1(c) \left(\frac{n_1(c)}{n_1(c) + \frac{1}{\delta^2}} \right)$$

$$= \bar{y}_1(c) \left(\frac{n_1(c)\delta^2}{n_1(c)\delta^2 + 1} \right)$$

So $\quad E[\mu_1 \mid c, Y_1, \ldots, Y_n] = \bar{Y}_1(c) \left(\dfrac{n_1(c)\delta^2}{n_1(c)\delta^2 + 1} \right)$

(note that this is formally correct even in the special case where $n_1(c) = 0$)

2). The joint pdf for $c, Y_1, Y_2, \ldots, Y_n$ is obtained by integrating μ_1, μ_2 out at $\circledast$. This produces

(note $h_1(c, y) = 1$ for $c < x_1$ and $h_2(c, y) = 1$ for $x_n < c$)

$$h_1(c, y) \, h_2(c, y) \, I[0 < c < 1]$$

This function is constant on the intervals (x_{i-1}, x_i). So the marginal of $Y_1, \ldots, Y_n$ is

$$f(y) \equiv \int_0^1 h_1(c, y) h_2(c, y) \, dc$$

$$= x_1 h_2(x_1, \underline{y}) + \sum_{i=1}^{n} (x_i - x_{i-1}) h_1(x_i, \underline{y}) h_2(x_i, \underline{y})$$

$$+ (1 - x_n) h_1(1, \underline{y})$$

and the conditional pdf of c given $Y_1, Y_2, \ldots, Y_n$ is thus

$$h_1(c, \underline{y}) h_2(c, \underline{y}) / f(\underline{y}) \qquad \text{on } (0,1)$$

j)
$$E[\mu_1 | Y_1, Y_2, \ldots, Y_n] = E\left[\underbrace{E[\mu_1 | c, Y_1, \ldots, Y_n]}_{\text{computed in g)}} \Big| Y_1, Y_2, \ldots, Y_n\right]$$

$$= \frac{1}{f(\underline{Y})}\left[\sum_{i=2}^{n} (x_i - x_{i-1}) h_1(x_i, \underline{Y}) h_2(x_i, \underline{Y}) \overline{Y}_1(x_i) \left(\frac{n_1(x_i) \gamma^2}{n_1(x_i)\gamma^2 + 1}\right)\right.$$

$$\left. + (1 - x_n) h_1(1, \underline{Y}) \overline{Y}_1(1) \left(\frac{n\gamma^2}{n\gamma^2 + 1}\right)\right]$$

k) There is (obviously) a formula for $E[\mu_2 | c, Y_1, Y_2, \ldots, Y_n]$ parallel to the one developed in g). So then

$$E[\mu(x_i) | c, Y_1, \ldots, Y_n] = I[x_i < c]\, \overline{Y}_1(c) \left(\frac{n_1(c) \gamma^2}{n_1(c)\gamma^2 + 1}\right)$$

$$+ I[x_i \geq c]\, \overline{Y}_2(c) \left(\frac{(n - n_1(c)) \gamma^2}{(n - n_1(c))\gamma^2 + 1}\right)$$

Notice that $\overline{Y}_1(c)$ and $\overline{Y}_2(c)$ and $n_1(c)$ are constant on each (x_{i-1}, x_i), so

$$E[\mu(x_i) | Y_1, \ldots, Y_n] = E\left[E[\mu(x_i) | Y_1, \ldots, Y_n] \Big| Y_1, \ldots, Y_n\right]$$

$$= \frac{1}{f(\underline{Y})}\left[x_1 h_2(x_1, \underline{Y}) \overline{Y}_2 \left(\frac{n\gamma^2}{n\gamma^2 + 1}\right) + \sum_{j=2}^{i} (x_j - x_{j-1}) h_1(x_j, \underline{Y}) h_2(x_j, \underline{Y}) \overline{Y}_2(x_j) \left(\frac{(n - n_1(x_j)) \gamma^2}{(n - n_1(x_j))\gamma^2 + 1}\right)\right.$$

$$+ \sum_{j=i+1}^{n} (x_j - x_{j-1}) h_1(x_j, \underline{Y}) h_2(x_j, \underline{Y}) \overline{Y}_1(x_j) \left(\frac{n_1(x_j) \gamma^2}{n_1(x_j)\gamma^2 + 1}\right)$$

$$\left. + (1 - x_n) h_1(1, \underline{Y}) \overline{Y}_1(1) \left(\frac{n\gamma^2}{n\gamma^2 + 1}\right)\right]$$

Let W have a Weibull(δ, τ) distribution with probability density function (pdf) given by

$$f_W(w|\delta,\tau) = \begin{cases} \tau\delta^{-\tau}w^{\tau-1}e^{-(w/\delta)^\tau} & 0 \le w \\ \\ 0 & \text{otherwise} \end{cases}$$

where δ and τ are positive.

1. Derive the cumulative distribution function of W.

2. Explain how you could generate a random value from a Weibull(δ, τ) distribution assuming that you are able to generate a random value from a Uniform(0,1) distribution.

3. Let W_1 and W_2 be independent Weibull(δ, τ) random variables. Let $S = W_1 + 3W_2$. Find the joint probability density function (pdf) of S and W_1. What is the probability density function of S given that $W_1 = w_1$, where w_1 is a nonnegative value?

4. Let V_1 and V_2 be independent Exponential(β) random variables with mean β ($\beta > 0$). Let $T = V_1 + V_2$. What is the probability density function of V_1 given that $T = t$, where t is a nonnegative value?

5. Let W have a Weibull(δ, τ) distribution. Show that random variable $X = W^\tau$ has an Exponential(δ^τ) distribution.

6. Derive the moment generating function of X, where X has an Exponential(δ^τ) distribution. The answer is $(1 - \delta^\tau)^{-1}$.

7. Show that the mean and variance of X are δ^τ and $\delta^{2\tau}$, respectively.

   Let $X_1, X_2, \ldots, X_n$ be independent Exponential(δ^τ) random variables.

8. Argue that $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ is a consistent estimator of δ^τ.

   Suppose that the value of τ is known.

9. Propose a consistent estimator of the δ. Argue that it is consistent.

10. Find the limit of the distribution of $\sqrt{n}(\bar{X}_n - \delta^\tau)$ as $n \to \infty$.

11. Find the limit of the distribution of $\sqrt{n}(\bar{X}_n^{1/\tau} - \delta)$ as $n \to \infty$.

1

1. Derive the cumulative distribution function of W. Let $y = (w/\delta)^\tau$. $dy = \tau(w/\delta)^{\tau-1}(1/\delta)dw$. Let $u = (v/\delta)^\tau$. $du = \tau(v/\delta)^{\tau-1}(1/\delta)dv$.

$$
\begin{aligned}
F(w) &= \int_0^w \tau\delta^{-\tau}v^{\tau-1}e^{-(v/\delta)^\tau}\,dv \\
&= \int_0^y e^{-u}\,du \\
&= 1 - e^{-y} \\
&= 1 - e^{-(w/\delta)^\tau}
\end{aligned}
$$

for $w \geq 0$ and $F(w) = 0$ for $w < 0$.

2. Explain how you could generate a random value from a Weibull distribution assuming that you are able to generate a random value from a Uniform(0,1) distribution.

   - Generate u from Uniform(0,1).
   - Let $u = F(w) = 1 - e^{-(w/\delta)^\tau}$.
   - Solve for w.
   - $w = \delta(-\ln(1-u))^{1/\tau}$

3. Let W_1 and W_2 be independent Weibull(δ, τ) random variables. Let $S = W_1 + 3W_2$. Find the joint probability density function (pdf) of S and W_1. What is the probability density function of S given that $W_1 = w_1$, where w_1 is a nonnegative value?

   - $f(s|w_1) = f(w_1, s)/f(w_1)$ for $s \geq w_1$; $f(s|w_1) = 0$ otherwise.
   - $S = W_1 + 3W_2$, so $W_1 = W_1$, $W_2 = (1/3)(S - W_1)$, $|J| = 1/3$.
   - $f(w_1, s) = \tau^2\delta^{-2\tau}w_1^{\tau-1}((1/3)(s - w_1))^{\tau-1}e^{-(w_1/\delta)^\tau}e^{-((s-w_1)/(3\delta))^\tau}(1/3)$
   - $f(s|w_1) = \tau\delta^{-\tau}((1/3)(s - w_1))^{\tau-1}e^{-((s-w_1)/(3\delta))^\tau}(1/3)$ for $s \geq w_1$ and $f(s|w_1) = 0$ otherwise.

4. Let V_1 and V_2 be independent Exponential(β) random variables with mean β ($\beta > 0$). Let $T = V_1 + V_2$. What is the probability density function of V_1 given that $T = t$, where t is a nonnegative value?

   - $f(v_1|t) = f(v_1, t)/f(v_1)$ for $t \geq v_1$; $f(t|v_1) = 0$ otherwise.
   - $f(v_1, t) = \beta^{-2}e^{-(v_1/\beta)}e^{-((t-v_1)/\beta))} = \beta^{-2}e^{-(t/\beta)}$
   - $f(t) = \beta^{-2}te^{-(t/\beta)}$; Gamma(2,$\beta$).
   - $f(v_1|t) = 1/t$ for $0 \leq v_1 \leq t$, Uniform(0,t).

5. Let W have a Weibull(δ, τ) distribution. Show that random variable $X = W^\tau$ has an Exponential distribution.

- $X = W^\tau; W = X^{1/\tau}$
- $dW = (1/\tau)X^{1/\tau-1}$
- $f(x) = \tau\delta^{-\tau}(x^{1/\tau})^{\tau-1}e^{-x/\delta^\tau}(1/\tau)x^{1/\tau-1}$
- $f(x) = \delta^{-\tau}e^{-x/\delta^\tau}$, $x \geq 0$; $f(x) = 0$ otherwise.

6. Derive the moment generating function of X. The answer is $(1 - \delta^\tau)^{-1}$.

$$
\begin{aligned}
M_X(t) &= \int_0^\infty e^{tx}\alpha^{-\beta}e^{-x/\alpha^\beta}dx \\
&= \alpha^{-\beta}\int_0^\infty e^{-x(1/\alpha^\beta-t)}dx \\
&= \alpha^{-\beta}\int_0^\infty e^{-x(1-\alpha^\beta t)/\alpha^\beta}dx \\
&= \alpha^{-\beta}\alpha^\beta/(1 - t\alpha^\beta) \\
&= (1 - t\alpha^\beta)^{-1}
\end{aligned}
$$

7. Show that the mean and variance of X are δ^τ and $\delta^{2\tau}$, respectively.

   - Differentiate MGF once, evaluate at $t = 0$, produces EX. Differentiate MDF twice, evaluate at $t = 0$, produces EX^2. $\text{Var}(X) = E(X^2) - (EX)^2$.
   - Intergrate $\int_0^\infty xf(x)dx$, use kernel of Gamma random variable, produces EX. Intergrate $\int_0^\infty x^2f(x)dx$, use kernel of Gamma random variable, produces EX^2. $\text{Var}(X) = E(X^2) - (EX)^2$.

   Let $X_1, X_2, \ldots, X_n$ be independent Exponential(δ^τ) random variables.

8. Argue that $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ is a consistent estimator of δ^τ.

   - The mean of X's is δ^τ.
   - The variance of X's is finite.
   - So $\bar{X}_n$ is consistent.

   Suppose that the value of τ is known.

9. Propose a consistent estimator of δ. Argue that it is consistent.

   - $\bar{X}_n$ is consistent.
   - $g(y) = y^{1/\tau}$ is a smooth function.
   - $g(\bar{X}_n) = \bar{X}_n^{1/\tau}$ is consistent for $E(X)^{1/\tau} = \delta$.

10. Find the limit of the distribution of $\sqrt{n}(\bar{X}_n - \delta^\tau)$ as $n \to \infty$. Answer: $N(0, \delta^{2\tau})$

11. Find the limit of the distribution of $\sqrt{n}(\bar{X}_n^{1/\tau} - \delta)$ as $n \to \infty$.

   - $g(y) = y^{1/\tau}$, $g'(y) = (1/\tau)y^{1/\tau-1}$, $g'(y)^2 = (1/\tau^2)y^{2(1/\tau-1)}$
   - Large sample: $\sqrt{n}(\bar{X}_n^{1/\tau} - \delta)$ is $N(0, \delta^{2\tau}(1/\tau^2)(\delta^\tau)^{2(1/\tau-1)} = N(0, \delta^2/\tau^2)$

2

A. Let $X_1, \ldots, X_n$ $(n \geq 2)$ be a collection of independent and identically distributed (iid) random variables with Uniform $(-\theta, \theta)$, $\theta > 0$, distribution, i.e., the probability density function (pdf) of X_1 is

$$f_{X,\theta}(x) = \begin{cases} (2\theta)^{-1} & \text{for} \quad x \in (-\theta, \theta) \\ 0 & \text{otherwise.} \end{cases}$$

Let $X_{(1)} = \min_{1 \leq i \leq n} X_i$ and $X_{(n)} = \max_{1 \leq i \leq n} X_i$, and

A.1 Write down the joint pdf of $(X_1, \ldots, X_n)$.

A.2 Show that $(X_{(1)}, X_{(n)})$ is sufficient for θ.

A.3 Show, by an example, that $(X_{(1)}, X_{(n)})$ is not complete for θ.

B. Let $Y_i = |X_i|^{1/2}$, $i = 1, \ldots, n$, where $X_1, \ldots, X_n$ are as in part A. Let $Y_{(n)} = \max_{1 \leq i \leq n} Y_i$. Note that the pdf of Y_1 is given by

$$f_\theta(y) = \begin{cases} 2y/\theta & \text{for} \quad y \in (0, \theta^2) \\ 0 & \text{otherwise} \end{cases}$$

and that $Y_1^2 \sim$ Uniform $(0, \theta)$.

B.1 Write down the joint pdf of $(Y_1, \ldots, Y_n)$.

B.2 Show that $Y_{(n)}$ is complete and sufficient for θ.

B.3 Find the uniformly minimum variance unbiased estimator (UMVUE) of θ^3.
(**Hint:** You may use the fact that $EW^3 = n/(n+3)$ where $W \sim \text{BETA}(n, 1)$.)

B.4 Show that the family of joint pdfs of $(Y_1, \ldots, Y_n)$ has monotone likelihood ratio in $Y_{(n)}$.

B.5 Using part B.4, find a size α $(\alpha \in (0, 1))$ uniformly most powerful (UMP) test for testing the hypotheses $H_0 : \theta \geq 1$ against $H_1 : \theta < 1$.
(**Note:** You need to find the constants appearing in the UMP test *explicitly*).

A.1

$$f_{X_1, \cdots, X_n, \theta}(x_1, \cdots, x_n) = \prod_{i=1}^n f_{X, \theta}(x_i)$$

$$= \begin{cases} (2\theta)^{-n} & \text{if } -\theta < x_{(1)} < x_{(n)} < \theta \\ 0 & \text{o.w.} \end{cases}$$

A-2.

Follows from (A.1) + the Factorization Theorem

A-3.

It is easy to check that θ is a scale parameter for the joint distribution of $(x_1, \cdots, x_n)$. Let $Z_i = X_i/\theta$, $i=1, \cdots, n$. And define $Z_{(1)} = \min_{1 \le i \le n} Z_i$, $Z_{(n)} = \max_{1 \le i \le n} Z_i$. Then, the (joint) distribution of $Z_1, \cdots, Z_n$ is free of θ. Let

$$a = E\left[\tan^{-1}\left(Z_{(n)} / Z_{(1)} \right) \right].$$

bounded $\Rightarrow$ Expectation exists.

Since θ is a scale parameter, it follows that

$$\tan^{-1}\left(X_{(n)}/X_{(1)} \right) \stackrel{d}{=} \tan^{-1} Z_{(n)}/Z_{(1)}, \text{ which}$$

has a density ("s.t. may use the transformation technique to derive it!") and hence, is non-degenerate.

$\Rightarrow$
$$\begin{cases} E_\theta\left[\tan^{-1}\frac{X_{(n)}}{X_{(1)}} - a\right] = 0 & \forall \theta \\[2em] \text{but} \quad P_\theta\left(\left[\tan^{-1}\frac{X_{(n)}}{X_{(1)}} - a\right] = 0\right) = 0 & \forall \theta. \end{cases}$$

$\Rightarrow$ $(X_{(1)}, X_{(n)})$ is not $\underset{\text{(boundedly)}}{\text{complete}}$ for θ.

B.1.

$$f_{Y_1, \dots, X_n; \theta}(y_1, \dots, y_n) = \prod_{i=1}^{n} f_\theta(y_i)$$

$$= \begin{cases} 2^n \prod_{i=1}^{n} y_i / \theta^n & \text{if } 0 < y_{(1)} < y_{(n)} < \theta^{\frac{1}{2}} \\[1em] 0 & \text{otherwise.} \end{cases}$$

B.2.

Factorization Theorem $\Rightarrow$ Sufficiency!

Completeness Pf. — May skip —

Let $h(\cdot)$ be a function with $E_\theta h(Y_{(n)}) = 0 \ \forall \theta$
(and $E_\theta |h(Y_{(n)})| < \infty \ \forall \theta$).

$(\Rightarrow)$
$$\int_0^\infty h(y) \cdot f_{Y_{(n)}, \theta}(y) \, dy = 0 \qquad \forall \theta$$

$(=)$ $\displaystyle\int_0^{\sqrt{\theta}^{\,a}} h(y) \cdot n \cdot \left[\frac{y^{2(n-1)}}{\theta^{n-1}}\right] \cdot \frac{2y}{\theta}\, dy = 0 \quad \forall \theta > 0$

$(=)$ $\displaystyle\int_0^{\sqrt{\theta}} h(y)\, y^{2n-1}\, dy = 0 \quad \forall \theta > 0$

$(=)$ $\displaystyle\int_0^\theta h^+(y)\, y^{n-1}\, dy = \int_0^\theta h^-(y)\, y^{2n-1}\, dy \quad \forall \theta > 0$

$\Longrightarrow$ $h^+(y) = h^-(y)$ a.e. (Leb)

$\Longrightarrow$ $h = 0$ a.e. (Leb).

B-3.   Enough to find a function g such that

$$E_\theta\, g(Y_{(n)}) = \theta^3 \quad \forall \theta > 0.\quad \text{Try } g(y) = y^6 \ (\text{as}$$

$\sqrt{\theta}$ is a scale - parameter for $Y_{(n)}$). By the Hint, and the fact that $Y_1^2 \sim UNIF(0,\theta)$ one has

$$E_\theta\, Y_{(n)}^6 = (E\, W^3)\, \theta^3 = \frac{n}{n+3} \cdot \theta^3$$

$\Longrightarrow$ $Y_{(n)}^6 \cdot \left(\frac{n+3}{n}\right)$ is a the UMVUE of θ^3.

B.4   For $0 < \theta_1 < \theta_2 < \infty$,

$$\frac{f_{\underline{Y}_{(n)}, \theta_2}(\underline{y})}{f_{\underline{Y}_{(n)}, \theta_1}(\underline{y})} = \frac{2^n \prod_{i=1}^{n} y_i \; \theta_2^{-n} \; I\left(0 < y_{(1)} < y_{(n)} < \sqrt{\theta_2}\right)}{2^n \left(\prod_{i=1}^{n} y_i\right) \theta_1^{-n} \; I\left(0 < y_{(1)} < y_{(n)} < \sqrt{\theta_1}\right)}$$

$$= \begin{cases} \dfrac{\theta_1^n}{\theta_2^n} & \text{if} \quad 0 < y_{(n)} < \sqrt{\theta_1} \\[2em] +\infty & \text{if} \quad y_{(n)} \in [\sqrt{\theta_1}, \sqrt{\theta_2}), \end{cases}$$

on the set $\mathcal{S}_{\theta_1, \theta_2} \equiv \{\underline{y}: \quad f_{\underline{y}, \theta_1}(\underline{y}) + f_{\underline{y}, \theta_2}(\underline{y}) > 0\}$

$$= \{\underline{y}: \quad 0 < y_{(1)} < y_{(n)} < \sqrt{\theta_2}\}.$$

$\Rightarrow$ $\dfrac{f_{\underline{y}, \theta_2}(\underline{y})}{f_{\underline{y}, \theta_1}(\underline{y})}$ is a nondecreasing function of $y_{(n)}$ on $\mathcal{S}_{\theta_1, \theta_2}$ $\Rightarrow$ $(Y_1, \cdots, Y_n)$ has MLR in $Y_{(n)}$.

B5.   A   size α UMP test for $H_0: \theta \geq 1$ vs $H_1: \theta < 1$

is given by

$$\phi(\underline{y}) = \begin{cases} 1 & \text{if } y_{(n)} < c \\ 0 & \text{if } y_n \geq c \end{cases}$$

where $c \in \mathbb{R}$ is such that $P_{\theta=1}(Y_{(n)} < c) = \alpha$.

Now -

$$\alpha = P_{\theta=1}(Y_{(n)} < c) = \left[P_{\theta=1}(Y_1 < c) \right]^n$$

$$= \left[\int_0^c 2y \, dy \right]^n \qquad \text{(assuming } c \in (0,1))$$

$$= c^{2n}$$

$$\implies c = \alpha^{\frac{1}{2n}}.$$