This question contains two parts, labelled as A and B.

A. Let $(\mathbf{X}, \mathcal{X}, \mu)$ be a measure space, $f_i : \mathbf{X} \rightarrow \mathbb{R}$, $i = 1, \ldots, m+1$ $(m \geq 1)$ be μ-integrable functions, and $k_1, \ldots, k_m$ be nonnegative real numbers. Define the function $\phi_0 : \mathbf{X} \rightarrow [0,1]$ by

$$\phi_0(x) = \begin{cases} 1 & \text{if } f_{m+1}(x) > \sum_{i=1}^{m} k_i f_m(x), \\ 0 & \text{otherwise.} \end{cases}$$

and let $\alpha_i = \int \phi_0 f_i d\mu$, $i = 1, \ldots, m$. Show that ϕ_0 maximizes $\int \phi f_{m+1} d\mu$ over all measurable functions $\phi : \mathbf{X} \rightarrow [0,1]$ with $\int \phi f_i d\mu \leq \alpha_i$, $i = 1, \ldots, m$.

(*This is a generalization of the Neyman-Pearson Lemma, which corresponds to the case* $m = 1$.)

B. Consider the statistical experiment $(\mathbf{X}, \mathcal{X}, \{P_\theta : \theta \in \mathbb{R}\})$ with $\mathbf{X} = \mathbb{R}^n$, $\mathcal{X} = \mathcal{R}^n$, the Borel σ-field on $\mathbb{R}^n$, and

$$\frac{dP_\theta}{d\mu}(x) \equiv f(x;\theta) = (2\pi)^{-n/2} \exp(-\sum_{i=1}^{n}(x_i - \theta)^2/2), \quad x \in \mathbf{X},$$

where μ is the Lebesgue measure on $\mathbb{R}^n$. Thus, if $X = (X_1, \ldots, X_n)' : \mathbf{X} \rightarrow \mathbf{X}$ denotes the identitity mapping, then $X_1, \ldots, X_n$ are iid $N(\theta, 1)$ random variables under P_θ. We want to find a *Uniformly Most powerful* (UMP) size α test for the *two sided* hypotheses $H_0 : \theta \notin (\theta_0, \theta_1)$ *against* $H_1 : \theta \in (\theta_0, \theta_1)$ where $\theta_0, \theta_1 \in \mathbb{R}$ and $0 < \alpha < 1$ are given numbers. Complete the following steps.

i. For any positive numbers a, b, c, d, show that there exists a real number t_0 such that the function

$$g(t) = ae^{-bt} + ce^{dt}, \quad t \in \mathbb{R}$$

is strictly decreasing on $(-\infty, t_0)$ and is strictly increasing on (t_0, ∞), with a unique minimum at t_0.

ii. Using (i) above or otherwise, show that for any $K_1 > 0, K_2 > 0$ and any $\theta \in (\theta_0, \theta_1)$, there exists constants $C_1, C_2 \in \mathbb{R}$ such that $\{x \in \mathbf{X} : K_1 f(x;\theta_0) + K_2 f(x;\theta_1) < f(x;\theta)\} = \{x \in \mathbf{X} : C_1 < \sum_{i=1}^{n} x_i < C_2\}$.

iii. Let $\theta_2 \in (\theta_0, \theta_1)$ be given. Using A and (ii) above, show that a best test of size α for testing $H_0 : \theta = \theta_0$ *or* θ_1 *against* $H_1 : \theta = \theta_2$ is given by

$$\phi_0(x) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} x_i \in (C_1, C_2), \\ 0 & \text{otherwise,} \end{cases}$$

where C_1 and C_2 satisfy the condition

$$E_{\theta_i}\phi_0(X) = \alpha$$

for $i = 0, 1$.

iv. It can be shown that *for a* $N(0, \sigma^2)$ *random variable Z,* $P(a < Z < b) = P(a + c < Z < b + c)$ *for some real numbers* $a < b$ *and* $c \neq 0$ *implies a+b+c=0.* Using this or otherwise, show that the midpoint of the interval $(C_1 - n\theta_0, C_2 - n\theta_0)$ lies in $(0, \infty)$ and the midpoint of the interval $(C_1 - n\theta_1, C_2 - n\theta_1)$ is in $(-\infty, 0)$.

v. Using (iv) or otherwise, show that

$$\sup_{\theta \notin (\theta_0, \theta_1)} E_\theta \phi_0(X) = \alpha.$$

vi. Conclude that ϕ_0 is a UMP size α test for $H_0 : \theta \notin (\theta_0, \theta_1)$ *against* $H_1 : \theta \in (\theta_0, \theta_1)$

Ph.D. Prelim
March '98

Solution / STAT 642-643 (I).

A. Let $\phi : \mathcal{X} \xrightarrow{\text{mble}} [0,1]$, satisfy $\int \phi f_i \, d\mu \le \alpha_i$, $i = 1, \dots, m$. Then, from the definition of ϕ_0, we have

$$(\phi_0 - \phi)\left(f_{m+1} - \sum_{i=1}^{m} k_i f_i\right) \geq 0$$

$$\Rightarrow \quad 0 \le \int (\phi_0 - \phi)\left(f_{m+1} - \sum_{i=1}^{m} k_i f_i\right) d\mu$$

$$= \int \phi_0 f_{m+1} \, d\mu - \int \phi f_{m+1} \, d\mu$$

$$- \sum_{i=1}^{k} \underbrace{k_i}_{\in [0,\infty)} \underbrace{\left(\alpha_i - \int \phi f_i \, d\mu\right)}_{\geq 0}$$

$$\Rightarrow \quad \int \phi_0 f_{m+1} \, d\mu - \int \phi f_{m+1} \, d\mu$$

$$\geq \sum_{i=1}^{k} k_i \left(\alpha_i - \int \phi f_i \, d\mu\right) \geq 0.$$

B. II. (i) Easy. (Consider the sign of $g'(t)$).

II. (iii)

$$K_1 f(\underline{x};\theta_0) + K_2 f(\underline{x};\theta_1) \lesseqgtr f(x,\theta_2)$$

$$\Leftrightarrow \quad K_1 \exp\left(-\sum_{i=1}^{n}(x_i-\theta_0)^2/2\right) + K_2 \exp\left(-\sum_{i=1}^{n}(x_i-\theta_1)^2/2\right) \lesseqgtr \exp\left(-\sum_{i=1}^{n}(x_i-\theta_2)^2/2\right)$$

$$\Leftrightarrow \quad \left(K_1 e^{-n\theta_0^2/2}\right)\exp\left(\theta_0\left(\Sigma x_i\right)\right) + \left(K_2 e^{-n\theta_1^2/2}\right) e^{\theta_1 \Sigma x_i} \lesseqgtr \exp\left(\frac{-n\theta_2^2}{2}\right)\exp(\theta_2 \Sigma x_i)$$

$$\Leftrightarrow \quad a \exp\left(-(\theta_2-\theta_0)\Sigma x_i\right) + c\exp\left((\theta_1-\theta_2)\Sigma x_i\right) \lesseqgtr 1$$

where $a = K_1 \exp\left(n(\theta_2^2-\theta_0^2)/2\right)$ and

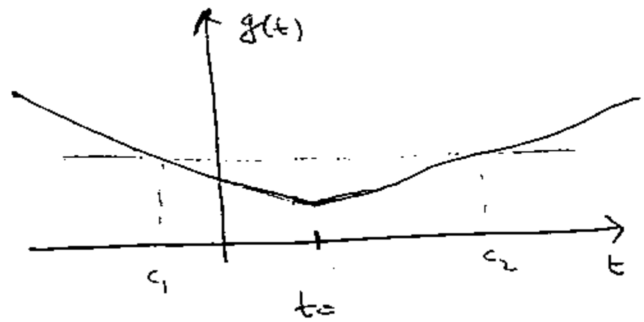
$c = K_2 \exp\left(n(\theta_2^2-\theta_1^2)/2\right)$

$$\Leftrightarrow \quad g(\Sigma x_i) \lesseqgtr 1$$

with $b = \theta_2 - \theta_0 > 0$

and $d = \theta_1 - \theta_2 > 0$.

Hence, there exist constants $c_1, c_2 \in \mathbb{R}$ such that



$$\left\{ \underline{x} : \ k_1 f(\underline{x}; \theta_0) + k_2 f(\underline{x}; \theta_1) \leq f(\underline{x}, \theta) \right\}$$

$$= \left\{ \underline{x} : \ c_1 \leq \sum x_i \leq c_2 \right\}.$$

B(iii) Follows from A and B(ii).

B(iv) [At the end]

B(v) Let $\Theta_0 = (\theta_1, \theta_2)^c$. For $\theta < \theta_1$,

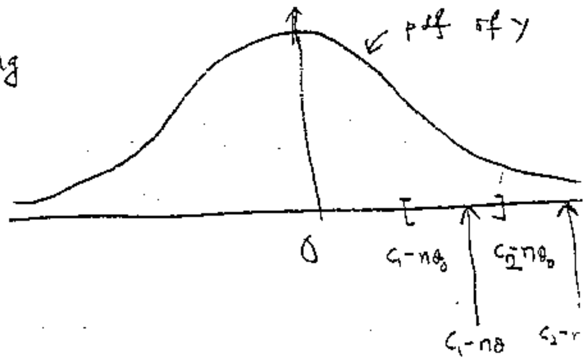
$$E_\theta \phi_0(\underline{x}) = P_\theta (c_1 < \Sigma X_i < c_2)$$

$$= P_\theta (c_1 - n\theta < \Sigma X_i - n\theta < c_2 - n\theta)$$

$$= P (c_1 - n\theta < Y < c_2 - n\theta)$$

where $Y \sim N(0, n)$. Since $\theta < \theta_0$,

$c_1 - n\theta > c_1 - n\theta_0$. Now, using II(iv) and the monotonicity of the $N(0, n)$ pdf on $(0, \infty)$, one can show that



$$E_\theta \phi_0(X)$$
$$= P\left(Y \in (c_1 - n\theta, c_2 - n\theta)\right)$$
$$< P\left(Y \in (c_1 - n\theta_0, c_2 - n\theta_0)\right) = \alpha \qquad \forall \theta < \theta_0$$

Similarly, it follows that $E_\theta \phi_0(X) < \alpha \quad \forall \theta > \theta_1$.

Thus, $\sup_{\theta \in \Theta_0} E_\theta \phi_0(X) = \alpha$.

B(vi') Let $\mathcal{C}_0 = \{ \phi : \sup_{\theta \in \Theta_0} E_\theta \phi(X) \le \alpha \}$

and $\mathcal{C} = \{ \phi ; \max_{i=0,1} E_{\theta_i} \phi(X) \le \alpha \}$.

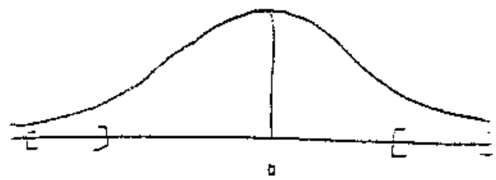
Then, clearly, $\mathcal{C}_0 \subset \mathcal{C}$. Also, $\phi_0 \in \mathcal{C}_0$ and by II(iii),

$$E_\theta\, \phi_0(x) \ge E_\theta\, \phi(x) \qquad \forall \theta \in (\theta_1, \theta_2) \longrightarrow (*)$$

for all $\phi \in \mathcal{C}$. Since $\mathcal{C}_0 \subset \mathcal{C}$, $(*)$ also holds for all ϕ in the smaller class $\mathcal{C}_0$.

This proves that ϕ_0 is the UMP size α test for $H_0: \theta \in \Theta_0$ vs $H_1: \theta \in (\theta_1, \theta_2)$.

B(iv) Note that



$$E_{\theta_0}\, \phi_0(x) = \alpha = E_{\theta_1}\, \phi_0(x)$$

$$\Leftrightarrow \quad P\left(c_1 - n\theta_0 < Y < c_2 - n\theta_0 \right) = \alpha$$

$$P\left(c_1 - n\theta_1 < Y < c_2 - n\theta_1 \right) = \alpha$$

Hence, with $a = c_1 - n\theta_0$, $b = c_2 - n\theta_0$ and $c = n\theta_0 - n\theta_1 < 0$,

$a + b + c = 0 \Rightarrow c_1 + c_2 = n\theta_0 + n\theta_1$. As a result, the midpoint of $(c_1 - n\theta_0,\ c_2 - n\theta_0) = \frac{1}{2}(c_1 + c_2 - 2n\theta_0)$
$> \frac{1}{2}(c_1 + c_2 - n\theta_0 - n\theta_1) = 0$ (since $\theta_0 < \theta_1$)

Similarly, the midpoint of $(c_1 - n\theta_1,\ c_2 - n\theta_1)$ is negative.

Let $(\mathbf{X}, \mathcal{X}, \{P_\theta : \theta > 0\})$ be a statistical experiment with $\mathbf{X} = (0, \infty)^n$ and $\mathcal{X}$ being the Borel σ-field on $\mathbf{X}$. Suppose that $X = (X_1 \ldots, X_n)'$ is the identity mapping from $\mathbf{X}$ onto $\mathbf{X}$. Thus, X is a random vector with joint distribution P_θ. Suppose that P_θ has a density f_θ with respect to the Lebesgue measure on $(0, \infty)^n$, given by

$$f_\theta(x_1, \ldots, x_n) = \theta^{-n} f(x_1/\theta, \ldots, x_n/\theta),$$

where f is a given probability density function on $(0, \infty)^n$. We want to find an *equivariant estimator* of the parameter θ using the loss function

$$L(\theta, a) = W(a/\theta), \ \theta > 0, a > 0, \tag{1}$$

where W is a nonnegative function on $(0, \infty)$. Let $G \equiv \{g_c : c > 0\}$ be a group of transformations on $\mathbf{X}$ defined by $g_c(x_1, \ldots, x_n) = (cx_1, \ldots, cx_n), \quad (x_1, \ldots, x_n)' \in \mathbf{X}$.

i. Find $\bar{G}$ and $\tilde{G}$, the induced groups of transformations on the parameter space $\Theta = (0, \infty)$ and the action space $\mathbf{A} = (0, \infty)$, respectively.

ii. Show that a nonrandomized decision rule $d(x)$ that is equivariant under G is of the form $d(x_1, \ldots, x_n) = x_1 d_1(y)$ for some function $d_1 : (0, \infty)^{n-1} \to \mathbf{A}$, where $y = (x_2/x_1, \ldots, x_n/x_1)'$.

iii. Suppose that $E_1 W(X_1 b) < \infty$ for some $b > 0$, where E_1 denotes expectation under P_θ when $\theta = 1$. Then, show that the best equivariant estimator (BEE) of θ with respect to the loss function $L(\theta, a)$ in (1) above is given by

$$d_0(X) = X_1 b_0(Y)$$

where $Y = (X_2/X_1, \ldots, X_n/X_1)'$ and $b_0(Y)$ is defined by $E_1(W(X_1 b_0(Y))|Y) = \inf_{b>0} E_1(W(X_1 b)|Y)$.

iv. Show that for the loss function $L_1(\theta, a) = \left(\frac{a}{\theta} - 1\right)^2$, the BEE of θ can be expressed as

$$d_0(X) = \frac{\int_0^\infty \theta^{-(n+2)} f(X_1/\theta, \ldots, X_n/\theta) d\theta}{\int_0^\infty \theta^{-(n+3)} f(X_1/\theta, \ldots, X_n/\theta) d\theta}.$$

(*This is an analog of Pitman's estimator for the scale parameter*).

v. Now suppose that $f(x_1, \ldots, x_n) = \prod_{i=1}^n x_i^{-2} I_{(1,\infty)}(x_i), \quad (x_1, \ldots, x_n)' \in \mathbf{X}$ and that $n \geq 3$, where $I_A(\cdot)$ denotes the indicator function of a set A. Find the BEE of θ under the following loss functions:

(a) $L_1(\theta, a) = \left(\frac{a}{\theta} - 1\right)^2$

(You need to find an exact expression for $d_0(X)$*).*

(b) $L_2(\theta, a) = |\log a - \log \theta|$

(Your answer may depend on the conditional distribution of a function of X_1 *given* Y*; You need to write down the form of* $d_0(X)$*, but need NOT find an exact expression.)*

SPRING'98 SOLUTION / Ph.D. Prelim STAT 642-643 (II)

(i) Suppose $X \sim P_\theta$. Then for any $c > 0$,

$$Y \equiv g_c(X) = (cx_1, \ldots, cx_n)'$$

has pdf $c^{-n} f_\theta(y_1/c, \ldots, y_n/c)$

$$= (\theta c)^{-n} f(y_1/\theta c, \ldots, y_n/\theta c), \quad y \in \mathcal{X}$$

$\Rightarrow \quad g_c(X) \sim P_{\theta c}$

$\Rightarrow \quad \bar{g}_c(\theta) = \theta c$.

Hence, $\bar{G} = \{\bar{g}_c : c > 0\}$. Similarly,

$$L(\bar{g}_c(\theta), \tilde{g}_c(a)) = L(\theta, a) \qquad \forall\ \theta, a$$

$\Longrightarrow \quad \tilde{g}_c(a) = ac$, and

$$\tilde{G} = \{\tilde{g}_c : c > 0\}.$$

(ii) A nonrandomized decision rule $d(x)$ is equivariant $\Longleftrightarrow$

$$\tilde{g}_c(d(x)) = d(g_c(x)) \qquad \forall x, \forall c$$

$$\Longleftrightarrow \quad c\,d(x) = d(cx_1, \ldots, cx_n) \quad \forall c > 0, \forall x$$

$\Rightarrow$ (with $c = 1/x_1$) $\quad d(x) = x_1\, d(1, x_2/x_1, \ldots, x_n/x_1)$, i.e. $$d(x) = x_1 \cdot d_1(y) \qquad \forall x \in \mathbb{X}$$

where $$d_1(y) = d(1, y_1, \cdots, y_{n-1}).$$

(iii) Note that for any $\theta \in \Theta$,

$$\{\bar{g}_c(\theta) : \bar{g}_c \in \bar{G}\} = \{c\theta : c > 0\} = \Theta,$$

so that the orbit of any θ is all of Θ. Since the risk function of an equivariant decision rule is constant on Θ, the BEE is obtained by minimizing

$R(1, d)$ over $d \in \mathbb{D}_I$, the set of all equivariant decision rules. By part (ii),

$$\inf_{d \in \mathbb{D}_I} R(1, d)$$

$$= \inf_{d_1} E_{\theta=1} W(X_1 d_1(Y))$$

$$= \inf_{d_1} E_{\theta=1}\left(E\{W(X_1 d_1(Y)) \mid Y\}\right)$$

$$\geq E_{\theta=1}\left(\inf_{b>0} E\{W(X_1 b) \mid Y\}\right) =$$

$$= E_{\theta=1}\left[E\left(W(X_1 b_0(Y))\mid Y\right)\right]$$

$$= E_{\theta=1} W(X_1 b_0(Y)) = R(1, d_0)$$

$$\geq \inf_{d\in \mathbb{D}_I} R(1, d)$$

Q.E.D.

(iv) For $L_1(\theta, a) = \left(\frac{a}{\theta}-1\right)^2$, $W(x) = (x-1)^2$

and hence, $b_0(y)$ minimizes $E_{\theta=1}\{(X_1 b - 1)^2 \mid Y\}$ over $b \in (0,\infty)$ $\Rightarrow$ $b_0(y)$ minimizes

$$h(b) = b^2 E_1(X_1^2|y) - 2b\, E_1(X_1|y) + 1$$

$$\Rightarrow \quad b_0(Y) = E_1(X_1|Y) / E_1(X_1^2|Y)$$

Hence,

$$d_0(X) = X_1\, E_1(X_1|Y)/E_1(X_1^2/Y).$$

Note that under $\theta = 1$, the conditional pdf of X_1 given Y is given by

$$f_{X_1|Y=y}(x_1) = \frac{f(x_1, y_1 x_1, \dots, y_{n-1} x_1)\; x_1^{n-1}}{\int_0^\infty f(x_1, y_1 x_1, \dots, y_{n-1} x_1)\, x_1^{n-1}\, dx_1}$$

Hence,

$$d_0(x) = X_1 \cdot \frac{\int_0^\infty x_1 \cdot f(x_1, Y_1 x_1, \cdots, Y_{n-1} x_1)\, x_1^{n-1}\, dx_1 / [\cdot]}{\int_0^\infty x_1^2 f(x_1, \underline{Y_1 x_1}, \cdots, \underline{Y_{n-1} x_1})\, x_1^{n-1}\, dx_1 / [\cdot]}$$

r.v.

The denomin... of $f_{x_1|Y}$ ...

Put $x_1 = X_1/\theta$ $\Rightarrow dx_1 = \frac{X_1}{\theta^2}\, d\theta$

$$\Rightarrow = X_1 \cdot \frac{\int_0^\infty (X_1/\theta)\, f(X_1/\theta, Y_1 X_1/\theta, \cdots, Y_{n-1} X_1/\theta)\, (X_1/\theta)^{n-1} (X_1/\theta^2) \ldots}{\int_0^\infty (X_1/\theta)^2 f(\frac{X_1}{\theta}, \frac{Y_1 X_1}{\theta}, \cdots, Y_{n-1} X_1/\theta)\, (X_1/\theta)^{n-1} \left(\frac{X_1}{\theta^2}\right) d\ldots}$$

$$= \frac{\cancel{X_1^{n+2}} \int_0^\infty \theta^{-(n+2)} f(X_1/\theta, \cdots, X_n/\theta)\, d\theta}{\cancel{X_1^{n+2}} \int_0^\infty \theta^{-(n+3)} f(X_1/\theta, \cdots, X_n/\theta)\, d\theta}$$

V (a).

$$d_0(X) = \frac{\int_0^\infty \theta^{-(n+2)} \left\{\prod_{i=1}^n (X_i/\theta)^{-2}\right\} I(X_{(1)}/\theta > 1)\, d\theta}{\int_0^\infty \theta^{-(n+3)} \left\{\prod_{i=1}^n (X_i/\theta)^{-2}\right\} I(X_{(1)}/\theta > 1)\, d\theta}$$

$$= \frac{\int_0^{X_{(1)}} \theta^{2n-(n+2)}\, d\theta \cdot \cancel{\left(\prod_{i=1}^n X_i^{-2}\right)}}{\int_0^{X_{(1)}} \theta^{2n-(n+3)}\, d\theta \cdot \cancel{\left(\prod_{i=1}^n X_i^{-2}\right)}}$$

$$n \geq 3 \quad = \frac{X_{(1)}^{n-1}/(n-1)}{X_{(1)}^{n-2}/(n-2)} = \frac{n-2}{n-1} \cdot X_{(1)}$$

1/(b)

$d_0(x) = X_1 \, b_0(Y)$, where $b_0(Y)$ minimizes

$$g(b) \equiv E_{\theta=1} W(X_1 b \mid Y) , \quad b > 0$$

$$= E_1\left(\left|\log X_1 b\right| \mid Y\right)$$

$$= E_1\left(\left|\log X_1 - \log b\right| \mid Y\right) , \quad b > 0$$

Note that $E_1\left(\left|\log X_1 - c\right| \mid Y\right)$ is minimized over $c \in \mathbb{R}$ at $c = \text{median}_{\theta=1}\left(\log X_1 \mid Y\right)$.

Hence,

$$\log b_0(Y) = \text{median}_{\theta=1}\left(\log X_1 \mid Y\right)$$

$$\Rightarrow b_0(Y) = \exp\left(\text{median}_{\theta=1}\left(\log X_1 \mid Y\right)\right)$$

a) For a sequence of random vectors $\{X_n\}$, a random vector X and a vector of constants **c**, define:

i) $X_n \xrightarrow{\mathcal{L}} X$ (convergence in distribution)

ii) $X_n \xrightarrow{P} \mathbf{c}$ (convergence in probability)

b) Prove that if $\{X_n\}$ is a sequence of (1-dimensional) random variables such that $X_n \xrightarrow{\mathcal{L}} c$, then $X_n \xrightarrow{P} c$.

A function of two variables $g(x, y)$ is **differentiable** at (x_0, y_0) provided g has first partials g_1 and g_2 at (x_0, y_0) and

$$g(x,y) = g(x_0,y_0) + g_1(x_0,y_0)(x - x_0) + g_2(x_0,y_0)(y - y_0)$$
$$+ E_1(x,y)(x - x_0) + E_2(x,y)(y - y_0)$$

for functions E_1 and E_2 that have limits 0 as $(x, y) \to (x_0, y_0)$.

It is a fact that if the first partials of g exist and are continuous in a neighborhood of (x_0, y_0), then g is differentiable at (x_0, y_0).

---

c) Suppose the sequence of 2-dimensional random vectors $\left\{\begin{pmatrix} X_n \\ Y_n \end{pmatrix}\right\}$ and sequence of constants $\{a_n\}$ are such that $a_n \to \infty$ and

$$a_n\left(\begin{pmatrix} X_n \\ Y_n \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}\right) \xrightarrow{\mathcal{L}} W$$

for $W \sim \mathrm{N}_2(\mathbf{0},\boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is nonsingular. Suppose g is differentiable at (x_0, y_0) and prove that $g(X_n, Y_n)$ is asymptotically normal and identify the mean and variance of the limit distribution.

d) Suppose the sequence of 2-dimensional random vectors $\left\{\begin{pmatrix} X_n \\ Y_n \end{pmatrix}\right\}$ and sequence of constants $\{a_n\}$ are such that $a_n \to \infty$,

$$a_n(X_n - x_0) \xrightarrow{\mathcal{L}} W$$

for $W \sim \mathrm{N}(0,\sigma^2)$, and

$$a_n(Y_n - y_0) \xrightarrow{P} 0 \ .$$

Suppose further that g is differentiable at (x_0, y_0) and prove that $g(X_n, Y_n)$ is asymptotically normal and identify the mean and variance of the limit distribution.

Henceforth assume that $X_1, X_2, ..., X_n$ are iid with marginal density (with respect to Lebesgue measure on $\mathcal{R}^1$)

$$f(x|\alpha,\theta) = I[0 \le x \le \theta]c(\alpha,\theta)x^\alpha \ \text{ for } \theta > 0 \text{ and } \alpha > -1 \ .$$

Denote moments of the marginal distribution as

$$\mathrm{E}_{\alpha,\theta}X_1^j = \mu_j(\alpha,\theta) \ .$$

And at some point it may be helpful to know that

$$\mathrm{E}_{\alpha,\theta}\ln X_1 = \ln\theta - \frac{1}{\alpha+1} \ \text{ and that } \ \mathrm{E}_{\alpha,\theta}(\ln X_1)^2 < \infty \ .$$

e) Using the first two moments, find method of moments estimators of α and θ, $\tilde{\alpha}_n$ and $\tilde{\theta}_n$. Show that these are consistent.

f) Using any of the results in parts **a)-d)** that you find helpful, show that $\tilde{\alpha}_n$ from part **e)** is asymptotically normal, and carefully describe how you would find the variance of the limit distribution. (You need not carry out all of the details of your program.)

g) Find maximum likelihood estimators of α and θ, $\widehat{\alpha}_n$ and $\widehat{\theta}_n$. Show directly that

$$\sqrt{n}\left(\widehat{\theta}_n - \theta\right) \xrightarrow{P_{\alpha,\theta}} 0 \ ,$$

and that $\widehat{\alpha}_n$ is consistent for α.

h) Using any of the results in parts **a)-d)** and **g)** that you find helpful, show that $\widehat{\alpha}_n$ is asymptotically normal, and carefully describe how you would find the variance of the limit distribution. (You need not carry out all of the details of your program.)

i) Briefly describe how you would create a large sample confidence interval for α based on either **f)** or **h)**.

Theory III / Key / 1998 Ph.D. Prelim

c) $a_n\left(g(X_n,Y_n)-g(x_0,y_0)\right)=g_1(x_0,y_0)\,a_n(X_n-x_0)+g_2(x_0,y_0)\,a_n(Y_n-y_0)$
$+E_1(X_n,Y_n)\,a_n(X_n-x_0)+E_2(X_n,Y_n)\,a_n(Y_n-y_0)$

$$= \left(g_1(x_0,y_0),g_2(x_0,y_0)\right)a_n\left(\begin{pmatrix}X_n\\Y_n\end{pmatrix}-\begin{pmatrix}x_0\\y_0\end{pmatrix}\right)$$
$$+E_1(X_n,Y_n)\,a_n(X_n-x_0)+E_2(X_n,Y_n)\,a_n(Y_n-y_0)$$

Since $(X_n,Y_n)\xrightarrow{P}(x_0,y_0)$ $E_1(X_n,Y_n)\xrightarrow{P}0$ and $E_2(X_n,Y_n)\xrightarrow{P}0$.
Since $a_n(X_n-x_0)$ and $a_n(Y_n-x_0)$ have normal limit dsns, the Slutsky theorem says that the last 2 terms above converge to 0 in probability. The first is a linear combination of coordinates of a vector that has a limiting normal dsn. So it has limiting dsn that is univariate normal with mean 0 and variance

$$\left(g_1(x_0,y_0),g_2(x_0,y_0)\right)\Sigma\begin{pmatrix}g_1(x_0,y_0)\\g_2(x_0,y_0)\end{pmatrix}$$

d) As above
$a_n\left(g(X_n,Y_n)-g(x_0,y_0)\right)=g_1(x_0,y_0)\,a_n(X_n-x_0)+g_2(x_0,y_0)\,a_n(Y_n-y_0)$
$+E_1(X_n,Y_n)\,a_n(X_n-x_0)+E_2(X_n,Y_n)\,a_n(Y_n-y_0)$

But now, the fact that $a_n(Y_n-y_0)\xrightarrow{P}0$ means that only the term $g_1(x_0,y_0)\,a_n(X_n-x_0)$ fails to converge to 0 in probability. The normal limit for $a_n(X_n-x_0)$ then implies that this has limiting dsn that is univariate normal with mean 0 and variance

$$\left(g_1(x_0,y_0)\right)^2\sigma^2$$

e) $c(\alpha,\theta) = \dfrac{1}{\int_0^\theta z^\alpha dz} = \dfrac{\alpha+1}{\theta^{\alpha+1}}$ so $E_{\alpha,\theta} X^j = c(\alpha,\theta)\int_0^\theta z^{\alpha+j} dz$

$$= \frac{c(\alpha,\theta)}{c(\alpha+j,\theta)}$$

$$= \theta^j\left(\frac{\alpha+1}{\alpha+j+1}\right)$$

(*)

So $\mu_1(\alpha,\theta) = \theta\,\dfrac{\alpha+1}{\alpha+2}$ and $\mu_2(\alpha,\theta) = \theta^2\left(\dfrac{\alpha+1}{\alpha+3}\right)$ and thus

$$\frac{\mu_2(\alpha,\theta)}{\mu_1^2(\alpha,\theta)} = \frac{(\alpha+2)^2}{(\alpha+1)(\alpha+3)} = \frac{\alpha^2+4\alpha+4}{\alpha^2+4\alpha+3} \equiv r(\alpha)$$

$r(\alpha)$ is differential and monotone on $(-1,\infty)$ and $\therefore$ has nice differentiable inverse on $(1,\infty)$, call it $s(r)$. $(s(r(\alpha))=\alpha)$

For $M_{jn} = \frac{1}{n}\sum_{i=1}^n X_i^j$ the jth sample moment, we might then estimate α with

$$\tilde{\alpha}_n = s\left(\frac{M_{2n}}{M_{1n}^2}\right)$$

and remembering (*) set $\tilde{\theta}_n = \dfrac{\tilde{\alpha}_n+2}{\tilde{\alpha}_n+1} M_{1n}$.

The WLLN shows that for any α,θ $M_{jn} \xrightarrow{P_{\alpha,\theta}} \mu_j(\alpha,\theta)$. So since s is cont$^{\underline{s}}$, $g(m_1,m_2) = s\left(\frac{m_2}{m_1^2}\right)$ is cont$^{\underline{s}}$ at $(\mu_1(\alpha,\theta),\mu_2(\alpha,\theta))$ and so $\tilde{\alpha}_n = g(M_{1n},M_{2n})$ is consistent for α. Then since

$$h(a,m) = \frac{a+2}{a+1}\,m \text{ is cont}^{\underline{s}} \text{ provided } a>-1$$

and $\tilde{\theta}_n = h(\hat{\alpha}_n, M_{1n})$ we have that for any α,θ

$$\hat{\theta}_n \xrightarrow{P_{\alpha,\theta}} \frac{\alpha+2}{\alpha+1}\left(\theta\,\frac{\alpha+1}{\alpha+2}\right) = \theta$$

f) Apply part c). Under (α, θ)

$$\sqrt{n}\left(\begin{pmatrix} M_{1n} \\ M_{2n} \end{pmatrix} - \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \right) \xrightarrow{\mathcal{L}} N_2\Big(0, \overbrace{\left(\mu_{i+j}(\alpha,\theta) - \mu_i(\alpha,\theta)\mu_j(\alpha,\theta) \right)}^{V(\alpha,\theta)}\Big)$$

Use $g(x,y) = s\left(\frac{y}{x^2}\right)$ so $g_1(x,y) = s'\left(\frac{y}{x^2}\right)\left(-\frac{y}{x^3}\right)$
$g_2(x,y) = s'\left(\frac{y}{x^2}\right)\left(\frac{1}{x^2}\right)$

and the derivative of s at r is $\frac{1}{r'(s(r))}$. The limiting variance is (as in c))

$$g_1(\mu_1(\alpha,\theta), \mu_2(\alpha,\theta)) \cdot V(\alpha,\theta) \begin{pmatrix} g_1(\mu_1(\alpha,\theta), \mu_2(\alpha,\theta)) \\ g_2(\mu_1(\alpha,\theta), \mu_2(\alpha,\theta)) \end{pmatrix}$$

g) The likelihood is

$$L(\alpha,\theta) = f(\underset{\sim}{x} \mid \alpha,\theta) = I\left[\max x_i \le \theta\right] \frac{(\alpha+1)^n}{\theta^{n(\alpha+1)}} \prod_{i=1}^{n} x_i^{\alpha}$$

Clearly, this is maximized as a function of θ (for any α) by $\theta = \max x_i$. Substituting $\hat{\theta}_n = \max x_i$ into $L(\alpha,\theta)$ one gets the profile log likelihood

$$\log L(\alpha, \hat{\theta}_n) = n\left[\log(\alpha+1) - (\alpha+1)\log\hat{\theta}_n\right] + \alpha \sum \log x_i$$

$$\frac{\partial}{\partial\alpha}(\text{above}) = \frac{n}{\alpha+1} - n\log\hat{\theta}_n + \sum \log x_i$$

Setting this $=0$ gives $\quad -\frac{n}{\alpha+1} = \sum \log x_i - n\log\hat{\theta}_n$

$$-\frac{1}{\alpha+1} = \frac{1}{n}\sum \log x_i - \log\hat{\theta}_n$$

$$-(\alpha+1) = \frac{1}{\frac{1}{n}\sum \log x_i - \log\hat{\theta}_n}$$

$$-\alpha = 1 + \frac{1}{\frac{1}{n}\sum \log x_i - \log \hat{\theta}_n}$$

$$\alpha = -\left(1 + \frac{1}{\frac{1}{n}\sum \log x_i - \log \hat{\theta}_n}\right)$$

So use $\hat{\theta}_n = \max X_i$ and $\hat{\alpha}_n = -\left(1 + \frac{1}{\frac{1}{n}\sum \log X_i - \log \hat{\theta}_n}\right)$

Note that $P_{\alpha,\theta}[\hat{\theta}_n \le \theta] = 1$ and for $\epsilon > 0$

$$P\left[\sqrt{n}(\theta - \hat{\theta}_n) > \epsilon\right] = P_{\alpha,\theta}\left[\theta - \frac{\epsilon}{\sqrt{n}} > \max X_i\right]$$

$$= \left(P_{\alpha,\theta}\left[X_1 < \theta - \frac{\epsilon}{\sqrt{n}}\right]\right)^n$$

Now

$$P_{\alpha,\theta}\left[X_1 < \theta - \frac{\epsilon}{\sqrt{n}}\right] = \int_{\theta - \frac{\epsilon}{\sqrt{n}}}^{\theta} \frac{\alpha+1}{\theta^{\alpha+1}} x^\alpha \, dx$$

$$\le \frac{\alpha+1}{\theta^{\alpha+1}} \frac{\epsilon}{\sqrt{n}} \max\left(\theta^\alpha, \left(\theta - \frac{\epsilon}{\sqrt{n}}\right)^\alpha\right)$$

$$\text{for } n \text{ large enough } \le 2\theta^\alpha \frac{\alpha+1}{\theta^{\alpha+1}} \epsilon \frac{1}{\sqrt{n}}$$

So $n \log P_{\alpha,\theta}\left[X_1 < \theta - \frac{\epsilon}{\sqrt{n}}\right] \le n \log\left(\frac{2(\alpha+1)\epsilon}{\theta}\right) - \frac{n}{2}\log n$

which $\to -\infty$. Thus $P\left[\sqrt{n}(\theta - \hat{\theta}_n) > \epsilon\right] \to 0$. Then

$\frac{1}{n}\sum \log X_i \xrightarrow{P_{\alpha,\theta}} \log\theta - \frac{1}{\alpha+1}$. So that with

with $g(x,y) = -\left(1 + \frac{1}{x - \log y}\right)$, the continuity of

g at $\left(\log\theta - \frac{1}{\alpha+1}, \theta\right)$ implies the consistency of $\hat{\alpha}_n$.

h) Apply part d). Under (α, θ)

$$\sqrt{n}\left(\frac{1}{n}\sum \log X_i - \left(\log\theta - \frac{1}{\alpha+1}\right)\right) \xrightarrow{\mathcal{L}} N(0, \sigma^2(\alpha,\theta))$$

The (finite) variance of $\log X_1$

Use $g(x,y) = -\left(1 + \frac{1}{x - \log y}\right)$ so that

$$g_1(x,y) = -\frac{1}{(x-\log y)^2}$$

and the limiting variance is (as in d))

$$\left(g_1\left(\log\theta - \frac{1}{\alpha+1}, \theta\right)\right)^2 \sigma^2(\alpha,\theta)$$

i) One might, for example, use Slutsky and ∴ employ

$$\hat{\alpha}_n \pm z\sqrt{\frac{1}{n}\left(g_1\left(\log\hat{\theta}_n - \frac{1}{\hat{\alpha}_n+1}, \hat{\theta}_n\right)\right)^2 \sigma^2(\hat{\alpha}_n, \hat{\theta}_n)}$$

a) Consider a statistical model $\{P_\theta\}_{\theta\in\Theta}$ for the observable X. Define what it means for a statistic $T(X)$ to be:

i) sufficient for θ
ii) minimal sufficient
iii) complete

Henceforth, let $X_1, X_2, \ldots$ be a sequence of iid Bernoulli (p) random variables for $p \in [0,1]$. Define

$$S_n = \sum_{i=1}^{n} X_i \text{ and } \boldsymbol{X}_n = (X_1, X_2, \ldots, X_n) \quad .$$

b) For n a fixed positive integer, first consider a statistical problem based on the observable $\boldsymbol{X}_n$. In this problem:

i) What is the conditional distribution of S_{n-1} given S_n?

ii) Let d be a nonconstant function mapping $\{0, 1, 2, \ldots, n-1\}$ to $(0,1)$. Find an estimator of p that is better than $d(S_{n-1})$ under squared error loss and argue carefully that it is indeed better.

c) Consider now inference for p based on a randomly stopped portion of the $X_1, X_2, \ldots$ sequence formulated as follows. On the "wedge-shaped" 2-dimensional grid of points

$$\Gamma = \left\{(n, s) \in \{1, 2, 3, \ldots\}^2 \,|\, s \leq n\right\} \;,$$

let $\Delta \subset \Gamma$ be a "stop sampling boundary" such that $\forall p$

$$P_p[\text{for some } n,\ (n, S_n) \in \Delta] = 1 \quad .$$

(For all p there is probability 1 that the random sequence $\{(n, S_n)\}$ eventually reaches an element of Δ.) Let

$$N^* = \min\{n | (n, S_n) \in \Delta\} \quad \text{and} \quad \boldsymbol{Y} = \boldsymbol{X}_{N^*} = (X_1, X_2, \ldots, X_{N^*}) \quad .$$

$\boldsymbol{Y}$ is clearly equivalent to

$$\boldsymbol{S} = (S_1, S_2, \ldots, S_{N^*}) \quad ,$$

which can be thought of as specifying a random path through Γ terminating at a point (N^*, S_{N^*}) of Δ. We will consider inference based entirely on $\boldsymbol{S}$ (or $\boldsymbol{Y}$).

i) Argue carefully that (N^*, S_{N^*}) is sufficient for p in this model for S.

For $(n, s) \in \Delta$, let $k_{(n,s)}$ be the number of possible paths through Γ that end at (n, s). That is, let

$$k_{(n,s)} = \#\{s = (s_1, s_2, ..., s_{n^*}) | \, n^* = n \text{ and } s_{n^*} = s\}$$

ii) For $(n, s) \in \Delta$, what is $P_p[(N^*, S_{N^*}) = (n, s)]$ in terms of p and $k_{(n,s)}$?

iii) Show that $I[S_1 = 1]$ is unbiased for p and provided $(1,1) \notin \Delta$ that $I[S_2 = 2]$ is unbiased for p^2.

iv) Provided $(1,1) \notin \Delta$, find estimators of p and p^2 that are functions of the sufficient statistic (N^*, S_{N^*}).

In general (N^*, S_{N^*}) need not be a complete statistic.

v) Show that with $\Delta = \{(2,1), (3,0), (3,1), (3,2), (3,3)\}$, (N^*, S_{N^*}) is not complete.

A sufficient condition for the completeness of (N^*, S_{N^*}) turns out to be that Δ is finite and contains exactly

$$m(\Delta) = \max\{n | (n, s_n) \in \Delta\} + 1$$

elements. (Don't try to prove this.)

vi) In cases where Δ is finite and contains $m(\Delta)$ elements, identify a minimal sufficient statistic and argue carefully that it is minimal.

vii) Argue carefully that in cases where Δ is finite and contains $m(\Delta)$ elements, it will always be possible to construct a UMVUE of any polynomial in p of degree less than $m(\Delta)$. (Outline the construction.)

Theory IV / Key / 1998 Ph.D. Prelim

b) i) $$P_p\left[S_{n-1}=u \text{ and } S_n=t\right] = \begin{cases} p\,P_p\left[S_{n-1}=u\right] & \text{if } t=u+1 \\ (1-p)\,P_p\left[S_{n-1}=t\right] & \text{if } u=t \\ 0 & \text{otherwise} \end{cases}$$

So $$P_p\left[S_{n-1}=u \mid S_n=t\right] = \begin{cases} \dfrac{p\binom{n-1}{t-1}p^{t-1}(1-p)^{n-t}}{\binom{n}{t}p^t(1-p)^{n-t}} & \text{if } u=t-1 \\ 1-\text{above} & \text{if } u=t \\ 0 & \text{otherwise} \end{cases}$$

for $t=0,1,2,\ldots,n$. That is, for such t

$$P_p\left[S_{n-1}=u \mid S_n=t\right] = \begin{cases} \dfrac{t}{n} & \text{if } u=t-1 \\ \dfrac{n-t}{n} & \text{if } u=t \\ 0 & \text{otherwise} \end{cases}$$

ii) $\delta = E\left[d(S_{n-1}) \mid S_n\right] = \dfrac{S_n}{n}\, d(S_n-1) + \dfrac{n-S_n}{n}\, d(S_n)$ is better than $d(S_n)$. (Use any arbitrary definitions for "$d(-1)$" and "$d(n+1)$".) This follows from the facts that squared error loss is strictly convex in the action and for any $p\in(0,1)$ there is positive p probability that δ differs from $d(S_{n-1})$.

c) i) $$P_p\left[S=(s_1,s_2,\ldots,s_n)\right] = p^{s_n}(1-p)^{n-s_n}\, I\left[\begin{array}{l}(s_1,s_2,\ldots,s_n) \\ \text{specifies a path} \\ \text{through } \Gamma \text{ with} \\ \text{each } s_j \le s_{j+1} \\ \text{first hitting } \Delta \\ \text{at } (n,s_n)\end{array}\right]$$

which for

$$g((a,b),p) = p^a(1-p)^{b-a} \quad \text{and} \quad h(s) = I[\qquad]$$

is clearly of the form

$$g((n, s_n), p)\, h(s).$$

So by the factorization Theorem, (N^*, S_{N^*}) is sufficient for p.

ii) Clearly for $(n,s) \in \Delta$ $\quad P_p[(N^*, S_{N^*}) = (n,s)] = k_{(n,s)}\, p^s (1-p)^{n-s}$

iii) $E_p I[S_1 = 1] = P_p[X_1 = 1] = p$

$E_p I[S_2 = 2] = P_p[X_1 = 1 \text{ and } X_2 = 1] = p^2$

iv) We can do this by finding the conditional means of $I[S_1 = 1]$ and $I[S_2 = 2]$ given the sufficient statistic. For the first, for $(n,s) \in \Delta$

$$E\left[I[S_1 = 1] \,\middle|\, (N^*, S_{N^*}) = (n,s) \right] = P\left[X_1 = 1 \,\middle|\, (N^*, S_{N^*}) = (n,s) \right]$$

$$= \frac{k'_{(n,s)}\, p^s (1-p)^{n-s}}{k_{(n,s)}\, p^s (1-p)^{n-s}} = \frac{k'_{(n,s)}}{k_{(n,s)}}$$

for $k'_{(n,s)} = \#\{ s = (1, s_2, s_3, \dots, s_{n^*}) \mid n^* = n \text{ and } s_{n^*} = s \}$

Similarly, $\quad E\left[I[S_2 = 2] \,\middle|\, (N^*, S_{N^*}) = (n,s) \right] = \dfrac{k''_{(n,s)}}{k_{(n,s)}}$

for $k''_{(n,s)} = \#\{ s = (1, 1, s_3, s_4, \dots, s_{n^*}) \mid n^* = n \text{ and } s_{n^*} = s \}$

v) Note that $I[(N^*, S_{N^*}) = (2,1)]$ has mean $2p(1-p) = 2p - 2p^2$

The constructions from iv) show that

$$\delta_1 = \begin{cases} 0 & \text{if } (N^*, S_{N^*}) = (3,0) \text{ or } (3,1) \\ \frac{1}{2} & \text{if } (N^*, S_{N^*}) = (2,1) \\ 1 & \text{if } (N^*, S_{N^*}) = (3,2) \text{ or } (3,3) \end{cases}$$

has $E_p \delta_1 = p$ and

$$\delta_2 = \begin{cases} 0 & \text{if } (N^*, S_{N^*}) = (2,1), (3,0), (3,1) \\ 1 & \text{if } (N^*, S_{N^*}) = (3,2), (3,3) \end{cases}$$

has $E_p \delta_2 = p^2$. So then

$$(2\delta_1 - 2\delta_2) - I[(N^*, S_{N^*}) = (2,1)]$$

is a nontrivial function of (N^*, S_{N^*}) with mean $0 \ \forall p$ and (N^*, S_{N^*}) is not complete.

vi) Bahadur's Theorem says that under these conditions, the (complete sufficient) statistic (N^*, S_{N^*}) is minimal.

vii) For each $n < m(\Delta)$ there is at least one point in Γ that is "reachable" with a possible path — call that point (n, r_n). Now $E_p I[S_n = r_n]$ is a polynomial of degree n in p. Working by induction, one may thus find a set of constants $c_{n0}, c_{n1}, \ldots, c_{nn}$ that $c_{n0} + \sum_{i=1}^{n} c_{ni} E_p I[S_i = r_i] = p^n$.

Then for a polynomial $a_0 + a_1 p + a_2 p^2 + \cdots + a_\ell p^\ell = g(p)$ with $\ell < m(\Delta)$

$$a_0 + \sum_{k=1}^{\ell} a_k \left(c_{k0} + \sum_{i=1}^{k} c_{ki} I[S_i = r_i] \right)$$

is a linear combination of $I[S_i = r_i]$ $i = 1, 2, \ldots, \ell$ with mean $g(p)$. Replacing the indicators with their conditional means given (N^*, S_{N^*}) produces the UMVUE of $g(p)$.