Practice Theory I Exam 2015

1). $X_1, \dots, X_n$ iid $f(x) = \alpha (x-\mu)^{\alpha-1} I(\mu \le x \le \mu+1)$

$0 < \alpha < \infty, \quad -\infty < \mu < \infty$

$X_{(1)} = \min(X), \quad X_{(n)} = \max(X)$

(a) Compute $E(X_1-\mu)^{-r}$ & show it is bounded for any $r < \alpha$

$$E[(X_1-\mu)^{-r}] = \int_{\mu}^{\mu+1} \alpha (x-\mu)^{-r}(x-\mu)^{\alpha-1}\, dx$$

$$= \int_{\mu}^{\mu+1} \alpha (x-\mu)^{(\alpha-r)-1}\, dx$$

let $z = x-\mu$

$dz = dx$

$$= \int_0^1 \alpha z^{(\alpha-r)-1}\, dz$$

$$= \frac{\alpha}{\alpha-r} z^{(\alpha-r)} \Big|_0^1$$

$$= \frac{\alpha}{\alpha-r} \quad \text{for } \alpha - r > 0$$

if $\alpha - r \le 0$ then $z^{(\alpha-r)}$ evaluated at 0 is undefined $\rightarrow \frac{1}{z^{(r-\alpha)}} = \frac{1}{0} = \infty$

(b) Assume that μ is known. Show that the MLE of α is

$$\tilde{\alpha}_n = \left[\frac{-1}{n}\sum_{i=1}^{n} \log(x_i - \mu)\right]^{-1}$$ and that

$$\sqrt{n}(\tilde{\alpha}_n - \alpha) \xrightarrow{d} N(0, \alpha^2)$$

$$f(\underset{\sim}{x}) = \prod_{i=1}^{n} f(x_i) = \alpha^n \left(\prod_{i=1}^{n} (x_i - \mu)^{\alpha-1}\right) I(\mu \le x_1, \dots, x_n \le \mu+1)$$

$$= \alpha^n \left(\prod_{i=1}^{n} (x_i - \mu)^{\alpha-1}\right) I(\mu \le x_{(1)}) I(x_{(n)} \le \mu+1)$$

μ considered known, assume

$$I(\mu \le x_{(1)}) = I(x_{(n)} \le \mu+1) = 1$$

$$\ell(\underset{\sim}{x}) = \log f(\underset{\sim}{x})$$
$$= n \log \alpha + \sum_{i=1}^{n} \log(x_i - \mu)^{\alpha-1}$$

$$= n \log \alpha + (\alpha-1) \sum_{i=1}^{n} \log(x_i - \mu)$$

$$\frac{\partial}{\partial \alpha} \ell(\underset{\sim}{x}) = \frac{n}{\alpha} + \sum_{i=1}^{n} \log(x_i - \mu) \overset{\text{set}}{=} 0$$

$$\Rightarrow \frac{n}{\alpha} = -\sum_{i=1}^{n} \log(x_i - \mu) \Rightarrow \frac{1}{\alpha} = \frac{-1}{n} \sum_{i=1}^{n} \log(x_i - \mu)$$

$$\Rightarrow \tilde{\alpha}_n = \left[\frac{-1}{n}\sum_{i=1}^{n} \log(x_i - \mu)\right]^{-1} \checkmark$$

Continued. →

According to a thm, if the log likelihood is continuously differentiable (twice continuously differentiable...) + (other conditions (?)), then

$$\sqrt{n}\,(\tilde{\alpha}_n - \alpha) \xrightarrow{d} N(0, I^{-1}(\alpha))$$

where $I(\alpha) = E[-\partial^2/\partial\alpha \, l(x_i)]$

$$f(x_i) = \alpha(x_i - \mu)^{\alpha-1} \underbrace{I(\mu \le x_i \le \mu+1)}_{\text{assume } = 1,\ \mu \text{ known.}}$$

$$l(x_i) = \log\alpha + (\alpha-1)\log(x_i - \mu)$$

$$\frac{\partial}{\partial\alpha} l(x_i) = \frac{1}{\alpha} + \log(x_i - \mu)$$

$$\frac{\partial^2}{\partial\alpha^2} l(x_i) = \frac{\partial}{\partial\alpha}\left[\frac{1}{\alpha} + \log(x_i - \mu)\right] = \frac{-1}{\alpha^2}$$

$$E[-\partial^2/\partial\alpha^2 \, l(x_i)] = E\{-(-1/\alpha^2)\} = \frac{1}{\alpha^2} = I(\alpha)$$

$$I^{-1}(\alpha) = \frac{1}{(1/\alpha^2)} = \alpha^2 \quad ✓$$

$$\Rightarrow \sqrt{n}\,(\hat{\alpha}_n - \alpha) \xrightarrow{d} N(0, \alpha^2) \quad ✓$$

(c) Assume both μ & α unknown.

$\tilde{\mu}_n = X_{(1)} \qquad \hat{\mu}_n = X_{(n)} - 1$

$Y_n = n^{1/\alpha}(\tilde{\mu}_n - \mu)$

$Z_n = n(\mu - \hat{\mu}_n)$

Show for all $0 \le y, z < \infty$,
$P(Y_n > y, Z_n > z) \longrightarrow e^{-y^\alpha - \alpha z}$ as $n \to \infty$ &
that $Y_n, Z_n \ge 0$ a.s. for all $n \ge 1$

$$P(Y_n > y, Z_n > z) = P(n^{1/\alpha}(X_{(1)} - \mu) > y, n(\mu - X_{(n)} + 1) > z)$$
$$= P(X_{(1)} - \mu > yn^{-1/\alpha}, \mu - X_{(n)} + 1 > z/n)$$
$$= P(X_{(1)} > yn^{-1/\alpha} + \mu, -X_{(n)} > z/n - (\mu+1))$$
$$= P(X_{(1)} > yn^{-1/\alpha} + \mu^{-1/\alpha}, X_{(n)} < (\mu+1) - z/n)$$
$$= P(X_1, \dots, X_n > yn^{-1/\alpha} + \mu, X_1, \dots, X_n < (\mu+1) - z/n)$$
$$= [P(X_i > yn^{-1/\alpha} + \mu, X_i < (\mu+1) - z/n)]^n$$
$$= \left[\int_{yn^\alpha + \mu}^{(\mu+1) - z/n} f(x_i)\,dx\right]^n$$

Note: if $x_i > yn^{-1/\alpha} + \mu \Rightarrow x_i \ge \mu$ ✓ (for $y \ge 0$)
if $x_i < (\mu+1) - z/n \Rightarrow x_i \le \mu + 1$ ✓

$$= \left[\int_{yn^\alpha + \mu}^{(\mu+1) - z/n} \alpha(x-\mu)^{\alpha-1}\,dx\right]^n$$
$$= \left[\frac{\alpha}{\alpha}(x-\mu)^\alpha \Big|_{x = yn^\alpha + \mu}^{x = (\mu+1) - z/n}\right]^n$$
$$= \left[(1 - z/n)^\alpha - (yn^{-1/\alpha})^\alpha\right]^n \to \exp(-z\alpha - y^\alpha) \ (?)$$

$$\log(\text{above}) = n \log\left((1 - z/n)^\alpha - y^\alpha n^{-1}\right)$$
$$\cancel{= \frac{\log((1 - z/n)^\alpha - y^\alpha n^{2\alpha})}{1/n}}$$

$$\frac{\partial}{\partial n}(\text{numerator}) = \frac{1}{(1-z/n)^\alpha - y^\alpha n^{2\alpha}} \left(\alpha(1-z/n)^{\alpha-1}(z/n^2) - (2\alpha) y^\alpha n^{2\alpha-1}\right)$$

$$\frac{\partial}{\partial n}(\text{denominator}) = \frac{-1}{n^2}$$

$$\frac{\partial}{\partial n}(\text{total}) = -n^2 \left(\partial/\partial n \text{ numerator} \right)$$

$$= \frac{\alpha(1-z/n)^{\alpha-1}(z) - 2\alpha y^\alpha n^{2\alpha+1}}{(1-z/n)^\alpha - y^\alpha n^{2\alpha}}$$

$$P(x_i > yn^{-1/\alpha} + \mu) = \int_{yn^\alpha + \mu}^{\mu+1} f(x_i)\,dx$$

$$= \frac{\alpha}{\alpha}(x-\mu)^\alpha \Big|_{x = yn^{-1/\alpha}+\mu}^{\mu+1}$$

$$= (1) - (yn^{-1/\alpha})^\alpha$$

Note:

$X_{(1)}$ & $X_{(n)}$ indep

$\Rightarrow P(X_{(1)} > a,\ X_{(n)} < b)$

$= P(X_{(1)} > a)\, P(X_{(n)} < b)$

$= [P(x_i > a)]^n [P(x_i < b)]^n$

? X

$$P(x_i < (\mu+1) - z/n)$$

$$= \int_\mu^{\mu+1-z/n} \alpha(x-\mu)^{\alpha-1}\,dx$$

$$= \frac{\alpha}{\alpha}(x-\mu)^\alpha \Big|_\mu^{\mu+1-z/n}$$

$$= (1-z/n)^\alpha - 0$$

$$= (1-z/n)^\alpha$$

$$\cancel{\left(P(x_i > yn^{-1/\alpha}+\mu,\ x_i < (\mu+1) - z/n)\right)^n}$$

X $\Rightarrow P(x_i > yn^{-1/\alpha}+\mu)^n\, P(x_i < (\mu+1) - z/n)^n$

$$= (1 - y^\alpha n^{-1})^n (1 - z/n)^{n\alpha}$$

$$\lim_{n\to\infty}(\text{above}) = (e^{-y^\alpha})(e^{-z})^\alpha \quad (\text{product of limits})$$

$$= \exp(-y^\alpha - z\alpha)$$ ✓

(ii) Show $Y_n, Z_n \geq 0$ a.s. for all $n \geq 1$

In joint likelihood, we have the indicators
$I(\mu \leq X_{(1)})$ and $I(X_{(n)} \leq \mu+1)$

$\Rightarrow$ $X_{(1)}$ must be $\geq \mu$ and $X_{(n)}$ must be $\leq \mu+1$

$\Rightarrow$ $X_{(n)} = \mu + \delta$; $X_{(n)} = \mu + 1 - \gamma$ $\quad 0 \leq \delta, \gamma \leq 1$

$Y_n = n^{1/\alpha}(X_{(1)} - \mu)$
$= n^{1/\alpha}(\mu + \delta - \mu)$
$= n^{1/\alpha}(\delta) \quad 0 \leq \delta \leq 1$

$Z_n = n(\mu - \hat{\mu}_n) = n(\mu - (\mu + 1 - \gamma - 1))$
$= n(\gamma) \quad 0 \leq \gamma \leq 1$

Since n, δ, γ are all ≥ 0 always,
$Y_n, Z_n \geq 0$ always ✓

(d) $\hat{\alpha}_n = \left[-\frac{1}{n} \sum_{i=1}^{n} \log(x_i - \hat{\mu}_n) \right]^{-1}$

Show for any $0 < s < \alpha$,

$\hat{\alpha}_n - \tilde{\alpha}_n = O_p(n^{-1 \wedge s})$

Note: $a \wedge b \equiv \min(a, b)$

Can use following facts:

(*) For any $0 < r \leq 1$, there exists a constant $0 < C_r < \infty$ such that $\log(1+\Delta) \leq C_r \Delta^r$ for all $0 \leq \Delta < \infty$

Use this fact to show

$$0 \leq \frac{1}{\tilde{\alpha}_n} - \frac{1}{\hat{\alpha}_n} \leq C_{s \wedge 1} |\hat{\mu}_n - \mu|^{s \wedge 1} \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^{-s \wedge 1}$$

& then complete the proof.

(i) Proving above statement

$$\frac{1}{\tilde{\alpha}_n} - \frac{1}{\hat{\alpha}_n} = \left[-\frac{1}{n} \sum_{i=1}^{n} \log(x_i - \mu) \right] - \left[-\frac{1}{n} \sum_{i=1}^{n} \log(x_i - \hat{\mu}_n) \right]$$

$$= -\frac{1}{n} \sum_{i=1}^{n} \log\left(\frac{x_i - \mu}{x_i - \hat{\mu}_n} \right) = +\frac{1}{n} \sum_{i=1}^{n} \log\left(\frac{x_i - \hat{\mu}_n}{x_i - \mu} \right)$$

$$= +\frac{1}{n} \sum_{i=1}^{n} \log\left(\frac{x_i - \mu + \mu - \hat{\mu}_n}{x_i - \mu} \right)$$

$$= +\frac{1}{n} \sum_{i=1}^{n} \log\left(1 + \left(\frac{\mu - \hat{\mu}_n}{x_i - \mu} \right) \right)$$

$\hat{\mu}_n = X_{(n)} - 1 \Rightarrow \mu - \hat{\mu}_n = \mu + 1 - X_{(n)}$

since $X_{(n)} \leq \mu + 1 \Rightarrow \mu - \hat{\mu}_n > 0$

Since $X_{(1)} \geq \mu \Rightarrow x_i - \mu > 0 \quad \forall i = 1, \dots, n$

$\Rightarrow \dfrac{\mu + \hat{\mu}_n}{x_i - \mu} > 0$ $\qquad \Delta = \dfrac{\mu - \hat{\mu}_n}{x_i - \mu}$

$\Rightarrow \log(1 + \Delta) > \log(1) = 0$ ✓

By the given stetement of fact,

$$\log(1+\Delta) \le C_r \Delta^r$$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} \log(1+\Delta_i) \le C_r \frac{1}{n}\sum_{i=1}^{n} \Delta_i^r$$

$$\Rightarrow \frac{1}{\tilde{\alpha}_n} - \frac{1}{\hat{\alpha}_n} = \frac{1}{n}\sum_{i=1}^{n} \log\left(1 + \frac{\mu - \hat{\mu}_n}{x_i - \mu}\right)$$

$$\le \frac{1}{n} C_r \sum_{i=1}^{n} \frac{(\mu - \hat{\mu}_n)^r}{(x_i - \mu)^r} \quad \text{where } r = 1 \wedge s$$

(note: $0 < (1 \wedge s) \le 1$ ✓)

$$= C_r |\mu - \hat{\mu}_n|^r \frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)^{-r}$$

(ii) Show $\hat{\alpha}_n - \tilde{\alpha}_n = O_p(n^{-(1 \wedge s)})$ Let $r = 1 \wedge s$

$$\frac{1}{\tilde{\alpha}_n} - \frac{1}{\hat{\alpha}_n} = \frac{\hat{\alpha}_n - \tilde{\alpha}_n}{\tilde{\alpha}_n(\hat{\alpha}_n)}$$

$$\Rightarrow \hat{\alpha}_n - \tilde{\alpha}_n \le \tilde{\alpha}_n \hat{\alpha}_n C_r |\mu - \hat{\mu}_n|^r \frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)^{-r}$$

$$O_p(n^{-r}) = n^{-r}(O_p(1))$$

$\Rightarrow$ Want to show $n^r(\hat{\alpha}_n - \tilde{\alpha}_n) = O_p(1)$

$$n^r(\hat{\alpha}_n - \tilde{\alpha}_n) = n^r \tilde{\alpha}_n \hat{\alpha}_n C_r |\mu - \hat{\mu}_n|^r \frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)^{-r}$$

By properties of MLE, $\tilde{\alpha}_n$ is consistent for α

$\Rightarrow \tilde{\alpha}_n \xrightarrow{P} \alpha$

By WLLN, $\frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)^{-r} \xrightarrow{P} E((x_i - \mu)^{-r})$

which we have shown to be bounded

in (a) when $r < \alpha$ and $1 \wedge s < \alpha$ ✓

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)^{-r} = O_p(1)$$

C_r is a constant $\Rightarrow C_r = O_p(1)$ ✓

$$\hat{\alpha}_n = \left[-\frac{1}{n}\sum_{i=1}^{n} \log(x_i - \hat{\mu}_n)\right]^{-1}$$

$$= \left[-\frac{1}{n}\sum_{i=1}^{n} \log(x_i - x_{(n)} + 1)\right]^{-1}$$

If we could show $\frac{1}{n}\sum_{i=1}^{n} \log(x_i - \hat{\mu}_n) \xrightarrow{P} \delta$

$\Rightarrow \hat{\alpha}_n \xrightarrow{P} (-\delta)^{-1}$ by the continuous mapping thm

$\Rightarrow \hat{\alpha}_n = O_p(1)$

?

$n^r |\mu - \hat{\mu}_n|^r$

$= n^r |\mu - x_{(n)} + 1|^r$

$F(n(\mu + 1 - x_{(n)}))(z)$

$= P(\mu + 1 - x_{(n)} \le z/n)$

$= P(x_{(n)} \le (\mu+1) - z/n)$

$= [P(x_i \le (\mu+1) - z/n)]^n$

$= [F_x(\mu + 1 - z/n)]^n$

$$F_x(t) = \int_{\mu}^{t} \alpha (x-\mu)^{\alpha-1} dx \qquad (\mu \le t \le \mu+1)$$

$$= (x-\mu)^{\alpha}\Big|_{x=\mu}^{t} = (t-\mu)^{\alpha}$$

$= (\cancel{\mu} + 1 - z/n - \cancel{\mu})^{n\alpha}$

$= (1 - z/n)^{n\alpha}$

$$\lim_{n\to\infty} \left[(1 - z/n)^n\right]^{\alpha} = e^{-\alpha z}$$

(7)

Therefore, $n|\mu - \hat{\mu}_n| = O_p(1)$

since $n|\mu - \hat{\mu}_n| \xrightarrow{d}$ some exponential dist

By the continuous mapping thm,

$[n|\mu - \hat{\mu}_n|]^r$ is also $O_p(1)$ ✓

Consequently,

$n^r(\hat{\alpha}_n - \tilde{\alpha}_n) = O_p(1) \cdot O_p(1) \cdot O_p(1) \cdot O_p(1) = O_p(1)$

$\Rightarrow \hat{\alpha}_n - \tilde{\alpha}_n = O_p(n^{-r})$

(e) Show for any $1/2 < s < \alpha < \infty$,
$\sqrt{n}(\hat{\alpha}_n - \alpha) \xrightarrow{d} N(0, \alpha^2)$

By work in (d),
$\hat{\alpha}_n - \tilde{\alpha}_n = O_p(n^{-r}) \quad r = 1 \wedge s$
$\Rightarrow \sqrt{n}(\hat{\alpha}_n - \tilde{\alpha}_n) = O_p(n^{-r+1/2})$
$\Rightarrow \sqrt{n}(\hat{\alpha}_n - \alpha) - \sqrt{n}(\tilde{\alpha}_n - \alpha) = O_p(n^{-r+1/2})$
$\Rightarrow \sqrt{n}(\hat{\alpha}_n - \alpha) = O_p(n^{-r+1/2}) + \underbrace{\sqrt{n}(\tilde{\alpha}_n - \alpha)}_{\xrightarrow{d} N(0,\alpha^2) \text{ from (b)}}$

$r = 1 \wedge s$
- if $r = 1 \Rightarrow -r + 1/2 = -1/2 \Rightarrow O_p(n^{-1/2}) = o_p(1)$
- if $r = s$ and $s > 1/2 \Rightarrow -s + 1/2 < 0$
$\Rightarrow O_p(n^{-s}) = o_p(1)$

$\Rightarrow \sqrt{n}(\hat{\alpha}_n - \alpha) = o_p(1) + \sqrt{n}(\tilde{\alpha}_n - \alpha)$

By Slutsky's thm, (or work in (b))
$o_p(1) + \sqrt{n}(\tilde{\alpha}_n - \alpha) \xrightarrow{d} 0 + N(0, \alpha^2)$
$\Rightarrow \sqrt{n}(\hat{\alpha}_n - \alpha) \xrightarrow{d} N(0, \alpha^2)$ ✓

2). Suppose the dist of a discrete RV X is as follows:

x	-2	-1	0	1	2
$p(x)$	$\theta_1(1-\theta_2)$	$(1/2-\alpha)\left(\frac{1-\theta_1}{1-\alpha}\right)$	$\alpha\left(\frac{1-\theta_1}{1-\alpha}\right)$	same as -1	$\theta_1\theta_2$

$0 < \theta_1 \le \alpha < 1/2$ (α known)

$0 < \theta_2 < 1$

Both (θ_1, θ_2) unknown.

H_0: $\theta_1 = \alpha$; $\theta_2 = 1/2$ (level α)

H_A: $\theta_1 < \alpha$; $\theta_2 \neq 1/2$ based on one obs X.

(a) Derive the α level likelihood test (LRT) for H_0 vs. H_1 & obtain its power function.

$$\Lambda = \frac{\sup_{\theta \in \Theta_0} L(\underset{\sim}{x} | \theta)}{\sup_{\theta \in \Theta} L(\underset{\sim}{x} | \theta)} = \prod_{i=-2}^{2} \Lambda_i^{I(x=i)}$$

$$\Lambda(-2) = \frac{p(\theta_1 = \alpha, \theta_2 = 1/2 \mid x = -2)}{\sup_{\theta \in \Theta} p(\theta_1, \theta_2 \mid x = -2)}$$

$$= \frac{\alpha(1/2)}{\alpha(1-0)} = \frac{1}{2}$$

Note: under general case:

$0 < \theta_1 \le \alpha < 1/2$ (α known)

$\Rightarrow$ most θ_1 can be is α ~~$(\alpha < 1/2)$~~ & least θ_1 can be is 0

$0 < \theta_2 < 1$

$\Rightarrow$ most θ_2 can be is 1 & least is 0

To maximize $\theta_1(1-\theta_2)$, plug in max value of θ_1 & min value of θ_2

$$\Lambda(-1) = \frac{p(\theta_1 = \alpha,\ \theta_2 = 1/2 \mid x = -1)}{\sup_{\theta \in \Theta} p(\theta_1, \theta_2 \mid x = -1)}$$

$$= \frac{(\cancel{1/2 - \alpha})\left(\frac{1-2\alpha}{1-\alpha}\right)}{(\cancel{1/2 - \alpha})\left(\frac{1-0}{1-\alpha}\right)}$$

$$= 1 - \alpha$$

$$\Lambda(0) = \frac{\alpha\left(\frac{1-2\alpha}{1-\alpha}\right)}{\cancel{\alpha}\left(\frac{1-0}{1-\alpha}\right)} = 1 - \alpha$$

$$\Lambda(1) = \Lambda(-1) = 1 - \alpha$$

$$\Lambda(2) = \frac{\alpha(1/2)}{\alpha(1)}$$

$$= \frac{1}{2}$$

$$\Lambda(-2) = \Lambda(2) = \frac{1}{2}$$

$$\Lambda(-1) = \Lambda(0) = \Lambda(-1) = 1 - \alpha$$

$$\Lambda = \left(\frac{1}{2}\right)^{I(|x|=2)} (1-\alpha)^{I(|x|=1)} (1-\alpha)^{I(x=0)} < K$$

$$\Rightarrow \left(\frac{1}{2}\right)^{I(|x|=2)} (1-\alpha)^{I(|x|<2)} < K$$

$$\Rightarrow \left(\frac{1}{2}\right)^{I(|x|=2)} (1-\alpha)^{(1 - I(|x|=2))} < K$$

$$\Rightarrow \left(\frac{1/2}{1-\alpha}\right)^{I(|x|=2)} (1-\alpha) < K$$

$$\Rightarrow \left(\frac{1/2}{1-\alpha}\right)^{I(|x|=2)} < K_1 = \frac{K}{1-\alpha}$$

$\rightarrow$

$$\Rightarrow -I(|x|=2)\log(2(1-\alpha)) < \overbrace{\log K_1}^{K_2}$$

$$\Rightarrow I(|x|=2)(-1) < K_3$$

Note: $\alpha < 1/2 \Rightarrow 2(1-\alpha) > 2(1/2) = 1$
$\Rightarrow \log(2(1-\alpha)) > 0$

$$\Rightarrow I(|x|=2) > K_4$$ → Always reject if $|x|=2$
$\Rightarrow I(|x|=2)=1$

$$\Rightarrow \phi(x) = \begin{cases} \gamma & |x|<2 \\ 0 & |x|=2 \end{cases}$$

α level:

$$E[\phi(x)] = B\phi(x)$$
$$= \gamma P(|x|<2)$$

$P(x=\{-1,0,1\} \mid H_0 \text{ true})$

At H_0, $\gamma P(|x|<2) = \alpha$

$$\Rightarrow \gamma\left[(1/2-\alpha)\left(\frac{1-\cancel{\alpha}}{1-\cancel{\alpha}}\right) + \alpha\left(\frac{1-\cancel{\alpha}}{1-\cancel{\alpha}}\right) + (1/2-\alpha)\left(\frac{1-\cancel{\alpha}}{1-\cancel{\alpha}}\right)\right] = \alpha$$

$$\Rightarrow \gamma\left[1/2 - \cancel{\alpha} + \cancel{\alpha} + 1/2 - \alpha\right] = \alpha$$

$$\Rightarrow \gamma[1-\alpha] = \alpha$$

$$\Rightarrow \gamma = \frac{\alpha}{1-\alpha}$$

Power $B\phi(x) = E_{\Theta}[\phi(x)]$
= prob of rejecting H_0 for various possible Θ

$$\phi(x) = \begin{cases} \alpha/(1-\alpha) & |x|<2 \\ 0 & |x|=2 \end{cases}$$

from generic $p(x)$

$$B\phi(x) = \left(\frac{\alpha}{1-\alpha}\right)\left[(1/2-\alpha)\left(\frac{1-\theta_1}{1-\alpha}\right) + \alpha\left(\frac{1-\theta_1}{1-\alpha}\right) + (1/2-\alpha)\left(\frac{1-\theta_1}{1-\alpha}\right)\right]$$

$$= \left(\frac{\alpha}{1-\cancel{\alpha}}\right)\left[\left(\frac{1-\theta_1}{1-\alpha}\right)(1/2-\cancel{\alpha}+\alpha+1/2-\alpha)\right]$$

$$= \frac{\alpha(1-\theta_1)}{1-\alpha}$$

(b) Consider the transformation $g(x) = -x$
Show the dist of X under Ho, the hypotheses, & the LRT are invariant under this transformation.

(i)

x	-2	-1	0	1	2
$p(x \mid H_0)$	$\alpha(1/2)$	$(1/2 - \alpha)$	α	$(1/2 - \alpha)$	$\alpha(1/2)$

since $P(x = -2 \mid H_0) = P(x = 2 \mid H_0)$
and $P(x = -1 \mid H_0) = P(x = 1 \mid H_0)$

$\Rightarrow$ Can multiply X by (-1) & dist will be the same

Also:

$\lvert x \rvert$	2	1	0
$p(\lvert x \rvert \mid H_0)$	α	$1 - 2\alpha$	α

- Dist of $\lvert x \rvert$ remains unchanged ✓
  - $\Rightarrow$ Transformation of $g(x) = -x$ will not change anything ✓

(ii) Ho & H1 do not vary wrt X ✓

(iii) LRT invariant:
The LRT is based on $\lvert x \rvert$, so $g(x) = -x$ transformation will not change anything

X

(c) Derive a uniformly most powerful (UMP) α level invariant test for the above hypothesis & compare its powerfunction w/ that of the LRT

$$|X| \sim \text{MultiBin}(1, \underbrace{\alpha, (1-2\alpha), \alpha}_{\text{Should leave in } (\theta_1, \theta_2, \alpha) \text{ form}})$$

to find sufficient statistics of θ_1, θ_2

$P(|X|=0) = \alpha$

$P(|X|=1) = 1-2\alpha$

$P(|X|=2) = \alpha$

Check: $\alpha + 1 - 2\alpha + \alpha = 1$ ✓

$$f(|x|) = \alpha^{I(|x|=0)} (1-2\alpha)^{I(|x|=1)} \alpha^{I(|x|=2)}$$

$$= \exp\left((I(|x|=0) + I(|x|=2)) \log \alpha + I(|x|=1) \log(1-2\alpha) \right)$$

$$I(|x|=0) = 1 - I(|x|=1) - I(|x|=2)$$

$$= \exp\left((1 - I(|x|=1) - \cancel{I(|x|=2)} + \cancel{I(|x|=2)}) \log \alpha + I(|x|=1) \log(1-2\alpha) \right)$$

$$= \exp\left(\underbrace{I(|x|=1)}_{g(x)} \underbrace{\log\left(\frac{1-2\alpha}{\alpha}\right)}_{\theta} + \underbrace{\log \alpha}_{b(\theta)} \right)$$

Sufficient statistic for $\alpha = |x|$

$\longrightarrow$

$$\frac{1-2\alpha}{\alpha} = \frac{1}{\alpha} - 2 \qquad 0 < \alpha < 1/2$$

$\Rightarrow \; 2 < 1/\alpha < \infty$

$\Rightarrow \; 0 < 1/\alpha - 2 < \infty$

by thm, UMP test is

$$\phi(x) = \begin{cases} 1 & |x| > k \\ \gamma & |x| = k \\ 0 & |x| < k \end{cases}$$ Should signs be flipped?

Want:

$\alpha = P_{\Theta_0}(|x| > k) + \gamma P_{\Theta_0}(|x| = k)$

$P_{\Theta_0}(|x| > 1) = P_{\Theta_0}(x = \{-2, 2\})$

$= \alpha(1/2) + \alpha(1/2)$

$= \alpha$ ✓ let $\gamma = 0$

$$\phi(x) = \begin{cases} 1 & |x| = 2 \\ 0 & |x| < 2 \end{cases}$$

$B_\phi(x) = E_\Theta[\phi(x)]$

$= P_\Theta(|x| = 2)$

3). $(x_1, y_1), \ldots, (x_n, y_n)$ iid RVs.

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \underbrace{\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}}_{\Sigma\text{, all elements} > 0} \right)$$

Goal: estimate (μ_x, μ_y) + Σ

Actual data collected may have Y missing

R_i indicates in Y_i observed.

Observed data:

$\{ R_i, R_i(x_i, y_i)^T + (1-R_i)(x_i, 0)^T \}$ for $i = 1, \ldots, n$.

$R_1, \ldots, R_n$ iid + indep of $(x_i, y_i)^T$

$\pi = P(R_i = 1)$ = known positive constant

Joint Likelihood of the observed data:

$$\prod_{i=1}^{n} \left[f(x_i, y_i)^{R_i} \left\{ \int f(x_i, y)\, dy \right\}^{1-R_i} \pi^{R_i} (1-\pi)^{1-R_i} \right]$$

$f(x_i, y_i)$ = joint density of (x_i, y_i)

(a) Show all model parameters are identifiable

Show if $f_1(x_i, y_i) = f_2(x_i, y_i)$

$\Rightarrow \underset{\sim}{\theta}_1 = \underset{\sim}{\theta}_2$ where $\underset{\sim}{\theta}_k = [\mu_{xk}, \mu_{yk}, \sigma_{11k}, \sigma_{12k}, \sigma_{22k}]^T$

(5) Write down the detail of the EM algorithm for calculating the maximum likelihood estimators for the parameters.

$E[\log \text{likelihood} \mid \text{observed } x_i, \underset{\sim}{\mu}^{(k)}, \Sigma^{(k)}, \text{obs. } y_i]$

– given obs. data + current iteration

$$= \sum_{i=1}^{n} R_i \log f(x_i, y_i) + R_i \log \pi + (1-R_i)\log(1-\pi)$$
$$+ E\left[\sum_{i=1}^{n} (1-R_i) \log f(x_i, y_i) \mid x_i, \underset{\sim}{\mu}^{(k)}, \Sigma^{(k)}\right]$$

Note: For $R_i = 1$, (x_i, y_i) fully observed

$\Rightarrow$ Expectation given obs (x_i, y_i) = constant

Since R_i indep of (x_i, y_i), can factor out of expectation

$$f(x_i, y_i) = \frac{1}{\sqrt{2\pi}} \frac{1}{|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}[x_i - \mu_x, y_i - \mu_y] \Sigma^{-1} \begin{bmatrix} x_i - \mu_x \\ y_i - \mu_y \end{bmatrix}\right)$$

$$|\Sigma| = \sigma_{11}\sigma_{22} - \sigma_{12}^2$$
$$|\Sigma|^{1/2} = \sqrt{\sigma_{11}\sigma_{22} - \sigma_{12}^2}$$

$$\Sigma^{-1} = \frac{1}{(\sigma_{11}\sigma_{22} - \sigma_{12}^2)} \begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{12} & \sigma_{11} \end{bmatrix}$$

$$\log f(x_i, y_i) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_{11}\sigma_{22} - \sigma_{12}^2) +$$
$$- \frac{1}{2}[x_i - \mu_x, y_i - \mu_y] \Sigma^{-1} \begin{bmatrix} x_i - \mu_x \\ y_i - \mu_y \end{bmatrix}$$

To perform the EM algorithm, take the derivative of the eqn listed (take derivative inside the expectation) wrt the desired paramete & set this derived eqn = 0.

$$\frac{\partial}{\partial \mu_x} \log f(x_i, y_i) = \frac{\partial}{\partial \mu_x}\left[\frac{-1}{2}\frac{1}{(\sigma_{11}\sigma_{22}-\sigma_{12}^2)}\begin{pmatrix}(x_i-\mu_x)^2(\sigma_{22}) + \\ (x_i-\mu_x)(y_i-\mu_y)(-2\sigma_{12})\end{pmatrix}\right.$$

$$[x \quad y]\begin{bmatrix} a & b \\ c & d\end{bmatrix}\begin{bmatrix} x \\ y\end{bmatrix} = [ax+cy, \ bx+dy]\begin{bmatrix} x \\ y\end{bmatrix}$$

$$= ax^2 + cxy + bxy + dy^2$$

$$= ax^2 + (c+b)xy + dy^2$$

$$= \frac{-1}{2}\frac{1}{(\sigma_{11}\sigma_{22}-\sigma_{12}^2)}\left(2(x_i-\mu_x)(-1)\sigma_{22} + (-1)(y_i-\mu_y)(-2\sigma_{12})\right)$$

$$= \frac{1}{\sigma_{11}\sigma_{22}-\sigma_{12}^2}\left(\sigma_{22}(x_i-\mu_x) - \sigma_{12}(y_i-\mu_y)\right)$$

Similarly, find $\frac{\partial}{\partial \Sigma}\log f(x_i, y_i)$ for $\Sigma = (\mu_x, \mu_y, \sigma_{11}, \sigma_{12}, \sigma_{22})$

Then, perform the following steps, using $\partial/\partial\mu_x$ as an illustration:

$$E\left[\partial/\partial\mu_x \text{ log likelihood} \mid obs(x_i, y_i), \mu^{(k)}, \Sigma^{(k)}\right]$$

$$= \frac{\partial}{\partial \mu_x}\left[\sum_{i=1}^{n} R_i \log f(x_i, y_i)\right] + \sum_{i=1}^{n}(1-R_i)\, E\left[\frac{\partial}{\partial \mu_x}\log f(x_i, y_i) \mid obs(x_i, y_i), \mu^{(k)}, \Sigma^{(k)}\right]$$

$=$

$\longrightarrow$

$$= \sum_{i=1}^{n} R_i \left[\frac{1}{\sigma_{11}\sigma_{12} - \sigma_{12}^2} \left(\sigma_{22}(x_i - \mu_x) - \sigma_{12}(y_i - \mu_y) \right) \right] +$$

$$+ \sum_{i=1}^{n} (1-R_i) E\left[\frac{\sigma_{22}(x_i - \mu_x) - \sigma_{12}(y_i - \mu_y)}{\sigma_{11}\sigma_{22} - \sigma_{12}^2} \,\middle|\, \text{obs}(x_i, y_i), \underset{\sim}{\mu}^{(k)}, \Sigma^{(k)} \right]$$

$$\overset{set}{=} 0$$

$$\Rightarrow \sum_{i=1}^{n} R_i \left(\sigma_{22}\cancel{(x_i - \mu_x)} - \sigma_{12}(y_i - \mu_y) \right) +$$

$$+ \sum_{i=1}^{n} (1-R_i) \left[\sigma_{22}\cancel{(x_i - \mu_x)} - \sigma_{12} \left[E(y_i \mid \text{obs } x_i, \underset{\sim}{\mu}^{(k)}, \Sigma^{(k)}) - \mu_y \right] \right] \overset{set}{=} 0$$

$$\Rightarrow \sum_{i=1}^{n} (-\sigma_{12})(y_i - \mu_y) R_i + \sum_{i=1}^{n} (1-R_i)(-\sigma_{12}) \left[E[y_i \mid x_i] - \mu_y \right] = 0$$

$$\Rightarrow \sum_{i=1}^{n} R_i (y_i - \mu_y) + \sum_{i=1}^{n} (1-R_i) \left[E[y_i \mid x_i] - \mu_y \right] = 0$$

$$\sum_{i=1}^{n} R_i \mu_y + \sum_{i=1}^{n} (1-R_i) \mu_y = \sum_{i=1}^{n} \mu_y = n\mu_y$$

$$\Rightarrow \sum_{i=1}^{n} y_i R_i + \sum_{i=1}^{n} (1-R_i) E[y_i \mid x_i] - n\mu_y \overset{set}{=} 0$$

$$\Rightarrow \mu_y^{(k+1)} = \frac{1}{n} \left(\sum_{i=1}^{n} y_i R_i + \sum_{i=1}^{n} (1-R_i) E[y_i \mid x_i] \Big|_{\mu^{(k)}, \Sigma^{(k)}} \right)$$

Continue for other derivatives to find othe MLEs by EM algorithm.

$$E[y_i \mid x_i, R_i = 0]$$
$$= E[y_i \mid x_i] \quad \text{since } R_i, y_i \text{ indep}$$
$$= \mu_y + \sigma_{12}\sigma_{22}^{-1}(x_i - \mu_x)$$

$$\mu_y^{(k+1)} = \frac{1}{n} \left(\sum_{i=1}^{n} y_i R_i + \sum_{i=1}^{n} (1-R_i) \left(\mu_y^{(k)} + \sigma_{12}^{(k)} \sigma_{22}^{-1(k)} (x_i - \mu_x^{(k)}) \right) \right)$$

(d) To estimate μ_Y, we can impute missing Y_i values as follows:

- Fit linear reg. model $Y = \alpha + \beta X + \varepsilon$ using only the complete data $(R_i = 1)$ & assuming ε & X indep.
- For subject i w/ missing Y_i, then impute
  $\hat{Y}_i = \hat{\alpha} + \hat{\beta} x_i$
  - $(\hat{\alpha}, \hat{\beta})$ MLE's obtained under model using complete data,

$$\hat{\mu}_Y = \frac{1}{n}\left(\sum_{i=1}^{n} R_i Y_i + (1-R_i)\hat{Y}_i\right)$$

Identify the true values for α & β in terms of μ & Σ.

- Dist of ε?

$$E[Y_i \mid x_i] = \mu_y + \sigma_{12}\sigma_{22}^{-1}(x_i - \mu_x)$$
$$= (\mu_y - \sigma_{12}\sigma_{22}^{-1}\mu_x) + \sigma_{12}\sigma_{22}^{-1} x_i$$

Based on model,

$$E[Y_i \mid x_i] = \alpha + \beta x_i$$

$$\alpha = \mu_y - \sigma_{12}\sigma_{22}^{-1}\mu_x$$
$$\beta = \sigma_{12}\sigma_{22}^{-1}$$

$$\varepsilon = Y - (\alpha + \beta x)$$

$$E[Y_i - (\alpha + \beta x_i) \mid x_i] = 0$$

$$\mathrm{Cov}(Y_i - (\alpha + \beta x_i) \mid x_i) = \mathrm{Cov}(Y_i \mid x_i)$$
$$= \sigma_{11} - \sigma_{12}\sigma_{22}^{-1}\sigma_{12}$$
$$= \sigma_{11} - \sigma_{12}^2/\sigma_{22} \longrightarrow$$

Y_i indep, X_i indep, ε_i indep

$$\Rightarrow \varepsilon \sim N(\underset{\sim}{0}, (\sigma_{11} - \sigma_{12}^2/\sigma_{22}) I)$$