PhD Prelim Exam
THEORY

Summer 2004
(Given on 7/29/04)

1. a) Specify a probability space (Ω, F, P) and a random variable $X : \Omega \to \Re$ such that X has the Poisson distribution with mean 1.

b) Let

$$Y = \begin{cases} X & \text{if } X \text{ is even} \\ 0 & \text{otherwise} \end{cases}$$

Find $\sigma(Y)$, the σ-algebra generated by Y. Then show that if $f : \Omega \to \Re$ is $\sigma(Y)$-measurable, then $f(\omega_1) = f(\omega_2)$ whenever $X(\omega_1)$ and $X(\omega_2)$ are both odd.

2. Let $F_\lambda(\cdot)$ and $F_{\gamma,\sigma}(\cdot)$ denote respectively the Poisson(λ) and Normal(γ,σ) cumulative distribution functions (the normal mean is γ and the standard deviation is σ). Let

$$F(\cdot) = \alpha F_\lambda(\cdot) + (1-\alpha) F_{\gamma,\sigma}(\cdot)$$

for some $0 < \alpha < 1$. Let μ_F, μ_{F_λ}, and $\mu_{F_{\gamma,\sigma}}$ be the Lebesgue-Stieltjes measures induced by F, F_λ, and $F_{\gamma,\sigma}$ respectively.

a) Evaluate $\mu_F(A_1), \mu_F(A_2)$, and $\mu_F(A_3)$ for

$$A_1 = \{1,2,3\},$$

$$A_2 = Q = \{\text{rational numbers}\},$$

and

$$A_3 = Q^C \cap [-1,1]$$

b) Prove or disprove each of the following. "$\ll$" means "is dominated by." For those cases where domination holds, give the corresponding Radon-Nikodym derivative.

i) $\mu_{F_\lambda} \ll \mu_F$

ii) $\mu_F \ll \mu_{F_\lambda}$

iii) $\mu_{F_{\gamma,\sigma}} \ll \mu_F$

iv) $\mu_F \ll \mu_{F_{\gamma,\sigma}}$

c) Let $\{X_i\}_{i\ge1}$ be iid with cumulative distribution function F. Let

$$\delta_i = \begin{cases} 1 & \text{if } X_i \text{ is integer} \\ 0 & \text{otherwise} \end{cases}$$

Show that the following limits exist with probability 1 and identify those limits. (This will imply that all four parameters α, λ, γ, and σ can be estimated consistently.) You may use without proof any result proved in class.

i) $\displaystyle\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \delta_i$

ii) $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i$

iii) $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i\delta_i$

iv) $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \left(Z_i - \bar{Z}_n\right)^2$, where $Z_i = X_i\left(1-\delta_i\right)$ and $\bar{Z}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i$

v) $\lim_{n\to\infty} F_n\left(x\right)$, where $F_n\left(x\right) = \frac{1}{n}\sum_{i=1}^{n} I_{(-\infty,x]}\left(X_i\right)$ and $I_A\left(\cdot\right)$ denotes the indicator function of a set A

For the last case, v), can the convergence be strengthened from "point-wise in x" to "uniformly in x"? Explain.

Major Theory I

Solution to Stat 642 Prelim Qn Aug 04

(Krishna B. Athreya).

i) a) $\Omega = R$, $\mathcal{F} = \mathcal{B}(R)$, $P = \mu_F$, the <u>Lebesgue-Stieltjes</u> measure induced by F

$X(\omega) = \omega$

b) Let $A_j = \{\omega : X(\omega) = 2j\}$, $j = 0, 1, 2 \ldots$

Let $\mathcal{A} \equiv \{A_j : j = 0, 1, 2 \ldots\}$. Then

$\sigma(Y) = \sigma(\mathcal{A}) \equiv \{ \bigcup_{j \in I} A_j, \ I \subset \{0, 1, 2 \ldots\} \}$

Let $A^* = \left(\bigcup_{j \geq 0} A_j\right)^c$. Then $A^* \in \sigma(Y)$

Any f that is $\sigma(Y)$ measurable is constant on each A_j & also on A^*. This $\Rightarrow$ (b).

Since $A^* = \{\omega : X(\omega)$ is odd$\}$.

ii) a) $\mu_F(A_1) = \alpha \ \cancel{F_\lambda(1) \pm 1} \ \mu_{F_\lambda}\{1, 2, 3\} = \alpha e^{-\lambda}\left(\frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!}\right)$

$\mu_F(A_2) = \alpha \ \mu_{F_\lambda}(Q) = \alpha$

$\cancel{\mu_F(A_3) = (1-\alpha) \ \mu_{F_{\mu,\sigma}}(-1, +1) = (1-\alpha)} \ \Phi\left(\frac{1-\mu}{\sigma}\right) - \Phi\left(\frac{-1-\mu}{\sigma}\right)$

b) i) $\mu_{F_\lambda}(A) \leq \frac{1}{\alpha} \mu_F(A)$ for $\forall A \in \mathcal{B}(R)$

So $\mu_{F_\lambda} << \mu_F$ $\quad \frac{d\mu_{F_\lambda}}{d\mu_F} = \frac{1}{\alpha} I_N(x)$ a.e. μ_F

where $N = \{0, 1, 2, \ldots\}$

ii) $\mu_{F_\lambda}(Q^c) = 0$ but $\mu_F(Q^c) = 1 - \alpha$

So $\mu_F \not<< \mu_{F_\lambda}$

iii) $\mu_{F_{\mu,\sigma}} << \mu_F$ as in (i) & $\frac{d\mu_{F_{\mu,\sigma}}}{d\mu_F} = \frac{1}{(1-\alpha)} I_{N^c}(x)$

iv) $\mu_{F_{\mu,\sigma}}(Q) = 0$ but $\mu_F(Q) = 1$ So $\mu_F \not<< \mu_{F_{\mu,\sigma}}$

Major Theory I

c) By SLLN, w.p. 1

i) $\lim_{n\to\infty} \frac{1}{n} \sum_1^n \delta_i = E\delta_1 = \alpha$

ii) $\lim_n \frac{1}{n} \sum_1^n X_i = EX_1 = \alpha\lambda + (1-\alpha)\mu$

iii) $\lim_n \frac{1}{n} \sum_1^n X_i\delta_i = EX_1\delta_1 = \alpha\lambda$

iv) $\lim_n \sum_1^n (z_i - \bar{z}_n)^2 = \lim_n \left(\frac{1}{n}\sum_1^n z_i^2 - \bar{z}_n^2\right)$

$= Ez_1^2 - (Ez_1)^2$

$= (1-\alpha)(\mu^2+\sigma^2) - ((1-\alpha)\mu)^2$

v) $\forall$ $F_n(x) \to F(x) = P(X_1 \le x)$

By the Glivenko-Cantelli this convergence is uniform w-p-1.

So α, λ, μ, σ^2 can be estimated consistently from the data $\{X_i\}_{i\ge1}$

This problem consists of 3 unrelated parts, labeled A, B, and C.

A. Let X be a random variable with $EX = 0$, $EX^2 = 1$, and with characteristic function, ϕ, such that $\int |\phi(t)|^p dt < \infty$ for some real number $p \in (1, \infty)$.

1. For each integer $r \geq 2$, find a random variable Z_r such that Z_r has characteristic function ϕ^r, i.e., $E\exp(itZ_r) = \phi(t)^r$ for all $t \in I\!R$.

2. Show that for $r \geq p$, the distribution of Z_r has a density with respect to the Lebesgue measure on $I\!R$.

3. It can be shown that
$$\lim_{r\to\infty} \int \left|\left[\phi(t/\sqrt{r})\right]^r - \exp(-t^2/2)\right| dt = 0. \tag{1}$$
Assuming (1), show that
$$\sup_{B\in\mathcal{B}(I\!R)} \left|P(Z_r \in B) - P(Z \in B)\right| \to 0 \quad \text{as} \quad r\to\infty,$$
where $\mathcal{B}(I\!R)$ is the Borel σ-field on $I\!R$ and where $Z \sim$ N(0,1).

B. Let X be a random variable having the Geometric distribution with parameter $\theta \in (0, 1)$, i.e.,
$$P_\theta(X \in A) = \sum_{x=0}^{\infty}[(1-\theta)\theta^x]\mathbb{1}_A(x), \quad A \in \mathcal{B}(I\!R),$$
where $\mathbb{1}_A$ denotes the indicator function of A. Let $\mathbf{A} = [0, 1]$ be the action space and $L(\theta, a) = (\theta - a)^2/(1-\theta)$ denote the loss function for a decision problem based on X.

1. Show that any nonrandomized decision rule, d, has risk function given by
$$R(\theta, d) = \sum_{k=0}^{\infty} c(k)\theta^k, \quad \theta \in (0, 1),$$
where $c(0) = d(0)$, $c(1) = d(1)^2 - 2d(0)$, and $c(k) = 1 - 2d(k-1) + d(k)^2$ for $k \geq 2$.

2. Let d_0 be the decsion rule given by $d_0(0) = 1/2$ and $d_0(i) = 1$ for all $i \geq 1$. Show that d_0 is an equalizer rule, i.e., $R(\theta, d_0)$ is constant in θ.

3. Let d be a nonrandomized equalizer rule.
   (a) Show that if $d(0) < 1/2$, then $d(i+1) < d(i) < 1$ for all $i \geq 1$.
   (b) Using (a) or otherwise, show that '$d(0) < 1/2$' is impossible and that $d(0)$ must be equal to 1/2. Hence, conclude that $d = d_0$.

This shows that d_0 is the only nonrandomized equalizer rule for this problem.

C. Let $X_1, \ldots, X_n$ be a collection of independent and identically distributed (iid) random variables with common distribution EXPONENTIAL (θ), $\theta \in (0, \infty)$, i.e.,

$$P_\theta(X_1 \in A) = \int_{A \cap (0,\infty)} \Big[\theta^{-1} \exp(-x/\theta)\Big] dx, \quad A \in \mathcal{B}(\mathbb{R}).$$

Find a uniformly most powerful unbiased (UMPU) size α test ($\alpha \in (0,1)$) for testing

$$H_0 : \theta = 1 \quad \text{against} \quad H_1 : \theta \neq 1.$$

(**NOTE:** *You need to write down equations for the constants that define the UMPU test, but do NOT need to solve them.*)

Major Theory II
STAT 643 SOLUTION SUMMER '2004

A.1 Let $X_1, \cdots, X_r$ be iid random variables with the same distribution as X. Take $Z_r = X_1 + \cdots + X_r$, $r \geq 2$. Then, for any $t \in \mathbb{R}$,

$$\Psi_r(t) \equiv E \exp(itZ_r) = E \exp(it[X_1 + \cdots + X_r])$$

$$= E \prod_{j=1}^{r} \{\exp(itX_j)\}$$

$$= \prod_{j=1}^{r} \underbrace{E \exp(itX_j)}_{= \phi(t)} \text{, by independence}$$

$$= [\phi(t)]^r$$

A.2. Note that $|\phi(t)| \leq 1$ for all $t \in \mathbb{R}$. Hence, for $r \geq p$,

$$\int_{-\infty}^{\infty} |E \exp(itZ_r)| \, dt = \int_{-\infty}^{\infty} |\phi(t)|^r \, dt$$

$$\leq \int_{-\infty}^{\infty} |\phi(t)|^p \, dt < \infty.$$

Since the characteristic function of Z_r is integrable, by the inversion formula, Z_r has a density, given by

$$f_r(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi_r(t) \exp(-itx)\, dt,$$

$x \in \mathbb{R}$.

A.3. Since Z_r has a density so does $Z_r/\sqrt{r}$. Let g_r denote the density of $Z_r/\sqrt{r}$.

Note that the characteristic function of $\frac{Z_r}{\sqrt{r}}$ is given by $E\exp\left(it\frac{Z_r}{\sqrt{r}}\right) = E\exp\left(i.[t/\sqrt{r}]Z_r\right)$

$= \psi_r(t/\sqrt{r}) = \left[\phi(t/\sqrt{r})\right]^r$, $t \in \mathbb{R}$. By the inversion formula, $g_r(x) = \int \psi_r(t/\sqrt{r})\, e^{-itx}\, dt/2\pi$, $x \in \mathbb{R}$, and

$f_Z(x) =$ density of $Z \sim N(0,1)$ $= \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) = \int e^{-t^2/2} \cdot e^{-itx} \frac{dt}{2\pi}$, $x \in \mathbb{R}$.

By (1), for any $x \in \mathbb{R}$,

$$\left|g_r(x) - f_Z(x)\right| = \frac{1}{2\pi} \left|\int \left(\psi_r(t/\sqrt{r}) - e^{-t^2/2}\right) e^{-itx}\, dt\right|$$

$$\leq \frac{1}{2\pi} \int \left| \varphi_r(t/\sqrt{r}) - e^{-t^2/2} \right| dt \to 0 \text{ as } r \to \infty.$$

Hence, by Scheffe's Theorem,

$$\sup_{B \in \mathcal{B}} \left| P\left(\frac{Z_r}{\sqrt{r}} \in B \right) - P(Z \in B) \right|$$

$$= \sup_{B \in \mathcal{B}} \left| \int_B g_r(x)\, dx - \int_B f_Z(x)\, dx \right|$$

$$\leq \int \left| g_r(x) - f_Z(x) \right| dx$$

$$\to 0 \quad \text{as} \quad r \to \infty.$$

B(1)

$$R(\theta,d) = E_\theta L(\theta,d(X)) = (1-\theta)^{-1}\sum_{x=0}^{\infty}(\theta-d(x))^2\left[(1-\theta)\theta^x\right]$$

$$= \sum_{x=0}^{\infty}(\theta-d(x))^2\theta^x$$

$$= \sum_{x=0}^{\infty}\theta^{2+x} - 2\sum_{x=0}^{\infty}\theta^{x+1}d(x) + \sum_{x=0}^{\infty}(d(x))^2\theta^x$$

$$= \sum_{k=2}^{\infty}\theta^k - 2\sum_{k=1}^{\infty}\theta^k d(k-1) + \sum_{x=0}^{\infty}(d(x))^2\theta^x$$

$$= \sum_{k=0}^{\infty}c(k)\,\theta^k \quad , \quad 0<\theta<1$$

where $c(0) = d(0)^2$, $c(1) = d(1)^2 - 2d(0)$
and $c(k) = 1 - 2d(k-1) + d(k)^2$. $\Bigg\} \rightarrow (**)$

B.2.

Easy. verify that d_0 satisfies the equations $(**)$. $\left\{\begin{array}{ll} c(0)=0, & c(1)=0 \\ c(k)=0 & \forall k\geq 2 \end{array}\right\} \rightarrow (*)$

B.3

If possible, suppose that there exists a $d(x)$ satisfying $(*)$ such that ~~$d(x) \neq d_0(x)$ for some x.~~

$d(0) \neq \frac{1}{2}$. Note that $d(x) \in [0,1]$ $\forall x \in X = \{0,1,\dots\}$.

Hence, $C(1) = 0 \iff d(1)^2 - 2d(0) = 0 \iff$

$\boxed{d(1)^2 = 2d(0)} \implies 1 \geq d(1)^2 \geq 2(d(0)) \iff d(0) \leq \frac{1}{2}$.

Thus, $d(0) \neq \frac{1}{2} \implies d(0) < \frac{1}{2}$.

$$\implies d(1)^2 = 2d(0) < 1$$

$$\implies d(1) < 1 \longrightarrow ①.$$

Now, for $k \geq 2$, $c(k) = 0 \implies$

$$0 = d(k)^2 - 2d(k-1) + 1$$
$$= d(k)^2 - d(k-1)^2 + (d(k-1) - 1)^2$$

Thus, $d(k-1) < 1$ for some $k \geq 2$ $\implies d(k)^2 - d(k-1)^2 = -(d(k-1)-1)^2 < 0$

$$\implies d(k) < d(k-1) \longrightarrow ②$$

By induction and ① and ②, ~~d(k)~~

$1 > d(1) > d(2) > \cdots$; i.e. $\{d(k)\}_{k=1}^{\infty}$ is a decreasing sequence in $[0,1]$. ~~Now letting~~

Let $a = \lim_{k\to\infty} d(k) \in [0,1)$. But letting $k \to \infty$ in '$c(k) = 0$, $k \geq 2$', we have

$$0 = 1 - \lim_{k\to\infty} 2d(k-1) + \lim_{k\to\infty} d(k)^2 = 1 - 2a + a^2$$

~~$\implies a = 1$, which is a~~

$\Rightarrow a = 1$. This contradicts the fact that $a < 1$. Hence, $d(0) = 1/2$, and by (*), $d(1) = 1/2 = d(2) = \cdots$

C. The joint distribution of $X_1, \cdots, X_n$ is

$$f_\theta(x) = \prod_{j=1}^{n} \left\{ \tfrac{1}{\theta} \exp(-x_j/\theta) \, \mathbb{1}_{(0,\infty)}(x_j) \right\}$$

$$= \theta^{-n} \exp\left(-\textstyle\sum_{j=1}^{n} x_j/\theta\right) \mathbb{1}_{(0,\infty)}(x_{(1)}),$$

$x = (x_1, \cdots, x_n)^t \in \mathbb{R}^n$, where $x_{(1)} = \min\{x_j : 1 \le j \le n\}$.

$\Rightarrow$ $\underset{\text{joint distribution of } X_1,\cdots,X_n}{\{P_\theta : \theta \in \Theta\}}$ is an 1-parameter exponential family

with $Q(\theta) = -\frac{1}{\theta}$ increasing in $\theta \in (0,\infty) \equiv \Theta$.

Hence, by a Theorem from the class-notes, a size α UMPU test for

$H_0 : \theta = 1$ vs $H_1 : \theta \neq 1$

is given by

$$\phi(x) = \begin{cases} 1 & \text{if } \quad \sum_{j=1}^n x_j < c_1 \text{ or } \sum_{j=1}^n x_j > c_2 \\ 0 & \text{otherwise} \end{cases}$$

where c_1 and c_2 are constants satisfying

$$E_{\theta=1}\,\phi(x) = \alpha$$

and

$$E_{\theta=1}\,\phi(x)\left(\sum_{j=1}^n x_j\right) = \alpha\; E_{\theta=1}\left(\sum_{j=1}^n x_j\right)$$

$$\Leftrightarrow \begin{cases} P_{\theta=1}\left(c_1 \le \sum_{j=1}^n x_j \le c_2\right) = 1-\alpha \\ E_{\theta=1}\left(\sum_{j=1}^n x_j\right) \mathbb{1}_{[c_1, c_2]}\left(\sum_{j=1}^n x_j\right) = (1-\alpha)\, E_{\theta=1}\left(\sum_{j=1}^n x_j\right) \end{cases}$$

$$\Leftrightarrow \begin{cases} \int_{c_1}^{c_2} \frac{1}{\Gamma(n)}\, y^{n-1} e^{-y}\, dy = 1-\alpha \\ \int_{c_1}^{c_2} \frac{1}{\Gamma(n)}\, y^{n} e^{-y}\, dy = (1-\alpha)\cdot n \end{cases}$$

1. Two independent random samples $X_1, X_2, \cdots, X_n$ and $Y_1, Y_2, \cdots, Y_m$ are respectively IID Poisson$(\theta\phi)$ and Poisson(θ/ϕ) for θ and ϕ both positive parameters.

(a) Derive $\hat{\theta}$ and $\hat{\phi}$, the joint maximum likelihood estimators (MLE) of θ and ϕ.

The Fisher information quantity defined for a univariate likelihood can be generalized to a multivariate likelihood like the one encountered in (a). Let $\alpha = (\alpha_1, \cdots, \alpha_k)'$ be a k-variate parameter taking value in $\Theta \subset R^k$ and $\ell(\alpha_1, \cdots, \alpha_k)$ be the log-likelihood function of α based on n random observations. For $i = 1, ..., k$, let $U_i(\alpha) = \frac{\partial \ell(\alpha)}{\partial \alpha_i}$ and $U(\alpha) = (U_1(\alpha), U_2(\alpha), \cdots, U_k(\alpha))'$. The Fisher information matrix is defined as

$$I(\alpha) = Cov(U) = \begin{pmatrix} Var(U_1(\alpha)), \ldots, Cov(U_1(\alpha), U_k(\alpha)) \\ \cdots\cdots \\ Cov(U_k(\alpha), U_1(\alpha)), \ldots, Var(U_k(\alpha)) \end{pmatrix}.$$

It can be shown that under certain regularity conditions the MLE of α is asymptotically normally distributed mean α and covariance $I^{-1}(\alpha)$.

(b) Derive the Fisher information matrix for (θ, ϕ).

(c) Identify the asymptotic distribution of $(\hat{\theta}, \hat{\phi})$ as $n \to \infty$.

(d) Show that the asymptotic correlation of $\hat{\theta}$ and $\hat{\phi}$ is positive if and only if $\phi < 1$.

2. An independent and identically distributed random sample $X_1, \cdots, X_n$ is taken from a distribution with probability density function

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

where $\theta > 0$. The maximum likelihood estimator of θ is

$$\hat{\theta}_{ML} = -\frac{n}{\sum_{i=1}^{n} log(X_i)}.$$

(You may take this as given).

(a) Find a pivotal quantity based on the entire random sample.

(b) Use the pivotal quantity obtained in (a) to construct an exact 90% confidence interval for θ.

3. One hundred independent and identically distributed exponential random variables are observed. Suppose the mean of the exponential distributions is $1/\alpha$. It is desired to test $\mathcal{H}_0 : \alpha = 2$ against various alternatives.

(The following information maybe useful: The 2.5%, 5%, 95%, 97.5% quantiles of the Gamma distribution with the shape parameter 100 and the scale parameter 1/2 are respectively 40.68, 42.07, 58.50, 60.27.)

(a) Derive the likelihood ratio test for $\mathcal{H}_0 : \alpha = 2$ against $\mathcal{H}_1 : \alpha = 3$ at significance level 5%. Express the test in terms of the maximum likelihood estimator of α.

(b) Derive the uniformly most powerful test of $\mathcal{H}_0 : \alpha = 2$ against $\mathcal{H}_1 : \alpha > 2$ at significance level 5%. Express the test in terms of the maximum likelihood estimator of α.

(a) Define $\mu_x = \theta\phi$ and $\mu_y = \theta/\phi$

$$L(\mu_x, \mu_y) = C\, e^{-n\mu_x} \mu_x^{\sum_1^n x_i} e^{-\mu_y} \mu_y^{\sum_1^m y_j}$$

The MLE's of (μ_x, μ_y) are $(\bar{X}_n, \bar{Y}_m)$.

$$\begin{cases} \theta\phi = \bar{X}_n \\ \theta/\phi = \bar{Y}_m \end{cases} \Rightarrow \hat{\theta} = \bar{Y}_m \sqrt{\bar{X}_n/\bar{Y}_m} \quad \hat{\phi} = \sqrt{\bar{X}_n/\bar{Y}_m}$$

(b) $(\bar{X}_n, \bar{Y}_m)^T$ is ~~AN~~ asymptotically $N_2\left(\begin{pmatrix}\mu_x \\ \mu_y\end{pmatrix}, \begin{pmatrix}\frac{\mu_x}{n} & 0 \\ 0 & \frac{\mu_y}{m}\end{pmatrix}\right)$ ~~normally distributed~~ distribution

$$(\hat{\theta}, \hat{\phi})^T = g(\bar{X}_n, \bar{Y}_m) \quad \text{where} \quad g(u,v) = \begin{pmatrix} v\sqrt{u/v} \\ \sqrt{u/v} \end{pmatrix}$$

is a smooth function over R_+^2.

The asymptotic normality of $(\hat{\theta}, \hat{\phi})$ is preserved under the smoothed transformation.

The delta method can be used to derive the asy. mean and asymptotic ~~var~~ covariance of ~~\theta~~ $(\hat{\theta}, \hat{\phi})$.

Here we use the approach of Fisher-Information.

$$\ell(\theta, \phi) = C - n\theta\phi - m\,\theta/\phi + \left(\sum x_i + \sum y_i\right)\log\theta + \left(\sum x_i - \sum y_i\right)\ell$$

$$\begin{cases} U_1 = \dfrac{\partial \ell}{\partial \theta} = -n\phi - m/\phi + \dfrac{\sum x_i + \sum y_i}{\theta} \\[2ex] U_2 = \dfrac{\partial \ell}{\partial \phi} = -n\theta + \dfrac{m\theta}{\phi^2} + \dfrac{\sum x_i - \sum y_i}{\phi} \end{cases}$$

$$Var(U_1) = \frac{n\cdot\theta\phi + m\,\theta/\phi}{\theta^2}, \quad Var(U_2) = \frac{n\theta\phi + m\,\theta/\phi}{\phi^2}$$

$$Cov(U_1, U_2) = \frac{n\theta\phi - m\,\theta/\phi}{\theta\phi}$$

Hence the Fisher Inf. matrix is

$$I(\theta, \phi) = \begin{pmatrix} \dfrac{n\phi + m/\phi}{\theta} & \dfrac{n\phi - m/\phi}{\phi} \\[2ex] \dfrac{n\phi - m/\phi}{\phi} & \dfrac{\theta(n\phi + m/\phi)}{\phi^2} \end{pmatrix}$$

$$I^{-1}(\theta, \phi) = \frac{1}{4mn}\begin{pmatrix} \theta(n\phi + m/\phi) & -n\phi^3 + m\phi \\ -n\phi^3 + m\phi & n\phi^3 + m\phi \end{pmatrix}$$

Hence

$$(\hat{\theta}, \hat{\phi}) \text{ is AN}\left(\begin{pmatrix} \theta \\ \phi \end{pmatrix}, \frac{1}{4mn}\begin{pmatrix} \theta(n\phi + m/\phi) & -n\phi^3 + m\phi \\ '' & n\phi^3 + m\phi \end{pmatrix}\right)$$

(c) $$\mathrm{Cov}(\hat{\partial}, \hat{\phi}) \approx \frac{-n\phi^3 + m\phi}{4mn} > 0$$

$$\text{iff} \quad -n\phi^3 + m\phi > 0 \quad \overset{n=m}{\Longleftrightarrow} \quad \text{iff} \quad -\phi^3 + \phi > 0$$

$$\text{As } \phi > 0 \quad \Rightarrow \quad \text{iff} \quad \phi < 1.$$

2

(a) Let $Y_i = -\alpha \log X_i \overset{d}{\sim} Exp(1) = Gamma(1,1)$

Here $\sum Y_i = -\alpha \sum \log X_i \overset{d}{\sim} Gamma(n,1)$

is a pivotal quantity

(b) Let $\gamma_{n,0.05}$ & $\gamma_{n,0.95}$ be the 0.05 and 0.95 quantiles of Gamma$(n,1)$

Then

$$Pr\left(\gamma_{n,0.05} < -\alpha \sum \log X_i < \gamma_{n,0.95} \right) = 0.90$$

Thus a 90% CI for α is

$$\cancel{\left(\frac{\gamma_{n,0.05}}{-\sum \log X_i} < \alpha < \right.}$$

$$\left(\frac{\gamma_{n,0.05}}{-\sum \log X_i} , \frac{\gamma_{n,0.95}}{-\sum \log X_i} \right)$$

3. (a) The LR is of the form

$$LR = \frac{L(\alpha_1)}{L(\alpha_0)} = \left(\frac{\alpha_1}{\alpha_0}\right)^{100} e^{(\alpha_0-\alpha_1)\sum_{i=1}^{100} T_i}$$

and for $\alpha_0 = 2, \alpha_1 = 3$ this is a monotonic decreasing function of $\sum T_i$. Hence the LR test is of the form "Reject $\mathcal{H}_0$ if $\sum T_i \leq c$". If we choose $c = 42.07$ then the size of the test is

$$\alpha = P\left(\sum T_i < 42.07\right) = P(\Gamma(100,2) < 42.07) = 0.05.$$

and so the test has size 5%. The ML estimator of α is $100/\sum T_i$. Hence we reject $\mathcal{H}_0$ if and only if

$$\sum T < 42.01 \Leftrightarrow \frac{100}{\sum T_i} > \frac{100}{42.07} \Leftrightarrow \hat{\alpha}_{ML} > 2.377.$$

(b)The likelihood is

$$L(\alpha) = \alpha^n e^{-\alpha \sum T_i}$$

which is of the form $f(\alpha)exp\{g(\alpha)S(T)\}$ where $S = \sum T_i$. The most powerful unbiased test therefore rejects $\mathcal{H}_0$ when $\sum T_i$ is either too small or too large. The distribution of T is $\Gamma(100,2)$ which is close to symmetric (because it is very close to normal!) and so the required test is Reject $\mathcal{H}_0$ if $\sum T_i < 40.28$ or $\sum T_i > 60.27$. In terms of the the ML estimator the test is

Reject $\mathcal{H}_0$ if $\hat{\alpha}_{ML} < 1.659$ or $\hat{\alpha}_{ML} > 2.458$.

the same test as in (b) is the MPT

1. Let Y have a Gamma(α, β) distribution with density given by

$$f_Y(y|\alpha,\beta) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha}\, y^{\alpha-1}\, e^{-y/\beta} & y > 0 \\ 0 & \text{otherwise} \end{cases}$$

where α and β are positive. Let μ ($\mu > 0$) be the mean of Y and σ^2 be the variance of Y.

(a) Derive the moment generating function (MGF) of Y.

(b) Derive the mean of Y, E(Y).

(c) Derive the variance of Y, var(Y).

2. Let $Y_1, Y_2, \ldots$ be independent and identically distributed as in part (1). Let $\overline{Y_n} = \frac{1}{n}\sum_{i=1}^n Y_i$.

(a) Argue (i.e., provide a short proof/demonstration) that $\overline{Y_n}$ is a consistent estimator of μ.

(b) Does E$(\log \overline{Y_n}) = \log$ E$(\overline{Y_n})$? Why or why not?

(c) Is $\log \overline{Y_n}$ a consistent estimator of $\log \mu$? Why or why not?

(d) What is the large sample distribution of $\sqrt{n}(\overline{Y} - \mu)$?

(e) What is the large sample distribution of $\sqrt{n}(\log \overline{Y} - \log \mu)$?

3. Let $Y_1, Y_2, \ldots$ be as in part (2), with $\alpha = 2$.

(a) Find the joint probability density of Y_1 and Y_2.

(b) Find the joint probability density of $|Y_1 - Y_2|$ and Y_1.
Note that $|Y_1 - Y_2| = Y_1 - Y_2$ when $Y_1 \geq Y_2$ and $|Y_1 - Y_2| = Y_2 - Y_1$ when $Y_1 < Y_2$.

(c) Find the probability density of $|Y_1 - Y_2|$.

(d) Find the maximum likelihood estimator of β when you have observed $Y_1 = y_1, \ldots, Y_n = y_n$.

(e) Find the large sample distribution of the maximum likelihood estimator of β.

4. Let $Y_1, Y_2, \ldots$ be distributed as in part (3), (i.e., iid Gamma(2,β)). Suppose further that *a priori* β has an Inverse-Gamma(α_0, β_0) distibution with probability density given by

$$f_\beta(\beta) = \begin{cases} \frac{1}{\Gamma(\alpha_0)\,\beta_0^{\alpha_0}} \left(\frac{1}{\beta}\right)^{\alpha_0+1} e^{-1/(\beta\beta_0)} & y > 0 \\ 0 & \text{otherwise} \end{cases}$$

When $\alpha_0 > 2$, the prior mean and variance of β are respectively $(\beta_0\,(\alpha_0 - 1))^{-1}$ and $(\beta_0^2\,(\alpha_0 - 1)^2\,(\alpha_0 - 2))^{-1}$.

(a) What are the (marginal) mean and variance of Y_1?

(b) What is the posterior distribution of β given $Y_1 = y_1, \ldots, Y_n = y_n$, the conditional distribution of β given $Y_1 = y_1, \ldots, Y_n = y_n$?

1. a) $M_Y(t) = \int_0^\infty e^{ty} \frac{1}{\Gamma(\alpha)\beta^\alpha} y^{\alpha-1} e^{-y/\beta} dy$

$= \frac{1}{\Gamma(\alpha)\beta^2} \int_0^\infty y^{\alpha-1} e^{-y\left(\frac{1}{\beta}-t\right)} dy$ $\qquad \frac{1}{\beta} - t = \frac{1-\beta t}{\beta}$

$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \Gamma(\alpha) \left(\frac{\beta}{1-\beta t}\right)^\alpha = \left(\frac{1}{1-\beta t}\right)^\alpha = (1-\beta t)^{-\alpha}$

b) $\frac{dM_Y(t)}{dt} = -\alpha(1-\beta t)^{-\alpha-1}(-\beta)$ $\qquad \frac{dM_Y(0)}{dt} = \alpha\beta = EY$

c) $\frac{d^2 M_Y(t)}{dt^2} = -(\alpha+1)\alpha(1-\beta t)^{-\alpha-2}(-\beta^2)$ $\qquad \frac{d^2 M_Y(0)}{dt^2} = (\alpha+1)\alpha\beta^2 = EY^2$

$\operatorname{var} Y = EY^2 - (EY)^2 = (\alpha+1)\alpha\beta^2 - \alpha^2\beta^2 = \alpha\beta^2.$

2. a)

$P(|\bar{Y} - \alpha\beta| < \varepsilon) = P((\bar{Y}-\alpha\beta)^2 < \varepsilon^2) \geq 1 - \frac{E(\bar{Y}-\alpha\beta)^2}{\varepsilon^2}$

$= 1 - \frac{\operatorname{var} \bar{Y}}{\varepsilon^2} = 1 - \frac{\alpha\beta^2}{n \cdot \varepsilon^2}$

$\longrightarrow 1$ as $n \to \infty$.

So $\bar{Y}$ is consistent for $EY = \alpha\beta$.

b) No log is concave. By Jensen's Inequality, $E\log\bar{Y} < \log E\bar{Y}$.

c) Yes log is a continuous function and $\bar{Y}$ is consistent for μ.

2d) CLT: $\bar{Y} \sim N\left(\alpha\beta, \frac{\alpha\beta^2}{n}\right)$. [Moment generating function MGF exists near zero]

e) Delta method & CLT

$g(\bar{Y}) = \log \bar{Y} \qquad g'(\bar{y}) = \frac{1}{\bar{y}} \qquad E\bar{Y} = \alpha\beta$

$\log \bar{Y} \sim N\left(\log(\alpha\beta), \frac{\alpha\beta^2}{n}\left(\frac{1}{\alpha\beta}\right)^2\right)$.

3. a) $f(y_1, y_2) = \left[\frac{1}{\Gamma(\alpha)\beta^\alpha}\right]^2 (y_1 y_2)^{\alpha-1} e^{-(y_1+y_2)/\beta}$ $\quad \alpha = 2$

b) $V = |Y_1 - Y_2| \qquad W = Y_1 \qquad Y_1 \geq Y_2 \quad V = Y_1 - Y_2 \quad Y_2 = W - V$

$Y_1 = W \qquad Y_1 < Y_2 \quad V = Y_2 - Y_1 \quad Y_2 = W + V$

For either transformation $|J| = 1$.

$f_{V,W}(v,w) = f(w, w-v) + f(w, w+v)$

c) Let $\alpha = 2$

$f_{V,W}(v,w) = \frac{1}{\beta^4} w(w-v) e^{-(2w-v)/\beta} + \frac{1}{\beta^4} w(w+v) e^{-\frac{2w+v}{\beta}}$

$f_V(v) = \int_0^\infty f_{V,W}(v,w)\, dw$ which is not too bad.

3c) continued

$$f_V(v) = \int_0^\infty \left[\frac{1}{\beta^4} w(w-v) e^{-\frac{2w-v}{\beta}} + \frac{1}{\beta^4} w(w+v) e^{-\frac{2w+v}{\beta}} \right] dw$$

$$= \frac{e^{v/\beta}}{\beta^4} \int_0^\infty w^2 e^{-w/(\frac{\beta}{2})} dw - \frac{v}{\beta^4} e^{v/\beta} \int_0^\infty w e^{-w/(\frac{\beta}{2})} dw$$

$$+ \frac{e^{-v/\beta}}{\beta^4} \int_0^\infty w^2 e^{-w/(\frac{\beta}{2})} dw + \frac{v}{\beta^4} e^{-v/\beta} \int_0^\infty v e^{-w/(\frac{\beta}{2})} dw$$

Kernels: $\text{Gamma}(3, \frac{\beta}{2})$ $\qquad$ $\text{Gamma}(2, \frac{\beta}{2})$

$$= \frac{1}{\beta^4} \left(\Gamma(3) \left(\frac{\beta}{2}\right)^3 \left[e^{v/\beta} + e^{-v/\beta} \right] - \Gamma(2) \left(\frac{\beta}{2}\right)^2 v \left[e^{v/\beta} - e^{-v/\beta} \right] \right)$$

d) $f(\underset{\sim}{y} | \beta) = \left(\frac{1}{\beta^2}\right)^n (\prod y_i) e^{-\Sigma y_i/\beta}$

$\ell(\beta | \underset{\sim}{y}) = -2n \log \beta + \Sigma \log y_i - \Sigma y_i / \beta$

$\ell'(\beta | \underset{\sim}{y}) = \frac{-2n}{\beta} + \frac{\Sigma y_i}{\beta^2} = 0 \qquad \Sigma y_i = 2n\beta$

$\hat{\beta} = \frac{1}{2} \bar{y}$

3, e) $\ell''(\beta|\underset{\sim}{y}) = \dfrac{2n}{\beta^2} - \dfrac{2\Sigma y_i}{\beta^3} = \dfrac{2(n\beta - \Sigma y_i)}{\beta^3}$

$$\hat{\beta} \sim N\left(\beta, \frac{\beta^3}{2(n2\beta - n\beta)}\right)$$

variance approximated by $\dfrac{\frac{1}{8}\bar{y}^3}{2\left(n\bar{Y} - \frac{n\bar{Y}}{2}\right)} = \dfrac{\bar{y}^2}{8n}$

4.

a) $EY = E[E(Y|\beta)] = E\alpha\beta = \alpha/(\beta_0(\alpha_0-1))$

$$\operatorname{var} Y = E \operatorname{var} Y|\beta + \operatorname{var} E Y|\beta$$
$$= E\alpha\beta^2 + \operatorname{var}\alpha\beta$$
$$= \alpha[\operatorname{var}\beta + (E\beta)^2] + \alpha^2 \operatorname{var}\beta$$
$$= \frac{\alpha(\alpha+1)}{\beta_0^2(\alpha_0-1)^2(\alpha_0-2)} + \frac{\alpha}{\beta_0^2(\alpha_0-1)^2}$$
$$= \frac{\alpha(\alpha+1+\alpha_0-2)}{\beta_0^2(\alpha_0-1)^2(\alpha_0-2)}$$

$\underline{\underline{\alpha = 2}}$

b) $$f(\beta|\underset{\sim}{y}) \propto \left(\frac{1}{\beta}\right)^{\alpha_0+1} e^{-\frac{1}{\beta\beta_0}} \left(\frac{1}{\beta}\right)^{2n} e^{-\Sigma y_i/\beta}$$

$$\beta|\underset{\sim}{y} \sim IG\left(\alpha_0 + 2n, \left(\frac{1+\beta_0\Sigma y_i}{\beta_0}\right)^{-1}\right).$$

Suppose that (X_1, Y_1), ..., (X_n, Y_n) are independent and have the same distribution as (X, Y), where the probability density function of Y is

$$f_Y(y) = \begin{cases} \theta e^{-\theta y} & \text{if } y > 0; \\ 0 & \text{otherwise,} \end{cases}$$

and for $y > 0$, the conditional probability density function of X given $Y = y$ is

$$f_{X|Y}(x|y) = \begin{cases} \theta y e^{-\theta y x} & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Here $\theta > 0$ is an unknown parameter.

(a) Let $Z = XY + Y$. Show that Z and X are independent.

(b) Show that

$$E(Y|X) = \frac{2}{\theta(1+X)}$$

and

$$Var(Y|X) = \frac{2}{\theta^2(1+X)^2}.$$

(c) Suppose we observe $Y_1, \ldots, Y_n$ only. Let $\hat{\theta}$ be the maximum likelihood estimator of θ. Show that

$$\hat{\theta} = \frac{n}{\sum_{i=1}^n Y_i}.$$

(d) Show that the $\hat{\theta}$ in part (c) is consistent for θ.

(e) Suppose we observe X_1, ..., X_n, Y_1, ..., Y_n. Find the maximum likelihood estimator of θ and denote it by $\tilde{\theta}$. Show that $\tilde{\theta}$ has a smaller asymptotic variance than the $\hat{\theta}$ in part (c).

(f) Suppose we observe Y_1^*, ..., Y_n^*, where

$$Y_i^* = \begin{cases} 0.5 & \text{if } 0 < Y_i \leq 1, \\ 1.5 & \text{if } 1 < Y_i \leq 2, \\ 2 & \text{if } Y_i > 2; \\ 0 & \text{otherwise.} \end{cases}$$

Find a consistent estimator for θ based on Y_1^*, ..., Y_n^*.

(g) Construct an approximate $(1 - \alpha)$ confidence interval for $\log \theta$ based on the $\tilde{\theta}$ in part (e). Express your answer using n, α and Φ^{-1}, where Φ is the cumulative distribution function of the standard normal distribution.

(a) We will first find the joint PDF of (X, Z) from the joint PDF of (X, Y). Since $Y = Z/(1+X)$,

$$
\begin{aligned}
f_{X,Z}(x,z) &= f_{X,Y}\left(x, \frac{z}{1+x}\right)\left|det\begin{pmatrix} \frac{\partial}{\partial x}x & \frac{\partial}{\partial z}x \\ \frac{\partial}{\partial x}\frac{z}{1+x} & \frac{\partial}{\partial z}\frac{z}{1+x} \end{pmatrix}\right| \\
&= \begin{cases} \frac{\theta^2 z e^{-\theta z}}{(1+x)^2} & \text{if } x > 0, z > 0; \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}
$$

The marginal probability density function of X is then

$$
\begin{aligned}
f_X(x) &= \int f_{X,Z}(x,z)dz \\
&= \begin{cases} \int_0^\infty \frac{\theta^2 z e^{-\theta z}}{(1+x)^2}dz & \text{if } x > 0; \\ 0 & \text{otherwise,} \end{cases} \\
&= \begin{cases} \frac{1}{(1+x)^2} & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}
$$

Similarly, the marginal PDF of Z is

$$
\begin{aligned}
f_Z(z) &= \int f_{X,Z}(x,z)dx \\
&= \begin{cases} \int_0^\infty \frac{\theta^2 z e^{-\theta z}}{(1+x)^2}dx & \text{if } z > 0; \\ 0 & \text{otherwise,} \end{cases} \\
&= \begin{cases} \theta^2 z e^{-\theta z} & \text{if } z > 0, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}
$$

From the above calculation, we have $f_{X,Z}(x,z) = f_X(x)f_Z(z)$, so X and Z are independent.

(b) $E(Y|X)$ and $Var(Y|X)$ can be obtained from the conditional PDF of Y given X or from the following calculation. Let $Z = XY + Y$ as in part (a). Then

$$
\begin{aligned}
E(Y|X) &= E\left(\frac{Z}{1+X}|X\right) \\
&= \frac{E(Z)}{1+X}
\end{aligned}
$$

and

$$
\begin{aligned}
Var(Y|X) &= E(Y^2|X) - (E(Y|X))^2 \\
&= E\left(\frac{Z^2}{(1+X)^2}|X\right) - \frac{(E(Z))^2}{(1+X)^2} \\
&= \frac{E(Z^2) - (E(Z))^2}{(1+X)^2},
\end{aligned}
$$

where $E(Z)$ and $E(Z^2)$ can be obtained from

$$E(Z^k) = \int z^k f_Z(z)dz,$$

or from the following:

$$\begin{aligned} E(Z) &= E(E(XY|Y)) + E(Y) \\ &= E(YE(X|Y)) + \frac{1}{\theta} \\ &= \frac{1}{\theta} + E\left(\frac{Y}{\theta Y}\right) = \frac{2}{\theta} \end{aligned}$$

and

$$\begin{aligned} E(Z^2) &= E(E(Y^2(X+1)^2|Y)) \\ &= E(Y^2(E(X^2|Y) + 2E(X|Y) + 1)) \\ &= E\left(Y^2\left(\frac{2}{\theta^2 Y^2} + \frac{2}{\theta Y} + 1\right)\right) \\ &= \frac{6}{\theta^2}. \end{aligned}$$

Thus

$$E(Y|X) = \frac{2}{\theta(1+X)}$$

and

$$Var(Y|X) = \frac{6/\theta^2 - 4/\theta^2}{(1+X)^2} = \frac{2}{\theta^2(1+X)^2}.$$

(c) Since $Y_1, \ldots, Y_n$ are independent, the joint PDF of $(Y_1, \ldots, Y_n)$ is the product of the marginal PDFs. The log likelihood function based on $Y_1, \ldots, Y_n$ is then

$$\begin{aligned} \log L(\theta) &= \log \prod_{i=1}^n f_Y(Y_i) \\ &= n \log\theta - \theta(\sum_{i=1}^n Y_i). \end{aligned}$$

The first derivative of $\log L(\theta)$ with respect to θ is

$$\frac{n}{\theta} - \sum_{i=1}^n Y_i,$$

which is positive if $\theta \in (0, n/\sum_{i=1}^n Y_i)$, negative if $\theta \in (n/\sum_{i=1}^n Y_i, \infty)$, and zero if $\theta = n/\sum_{i=1}^n Y_i$. So the log likelihood function is maximized at $\theta = n/\sum_{i=1}^n Y_i$, which is the MLE $\hat{\theta}$.

(d) Note that

$$1/\hat{\theta} = \frac{\sum_{i=1}^{n} Y_i}{n}$$

which converges to $E(Y) = 1/\theta$ by the strong law of large numbers, so $1/\hat{\theta}$ is consistent for $1/\theta$. By the continuous mapping theorem, $\hat{\theta}$ is consistent for θ.

(e) Let $\tilde{\theta}$ be the MLE of θ based on $X_1, Y_1, \ldots, X_n, Y_n$. Then it can be shown that

$$\frac{1}{\tilde{\theta}} = \frac{\sum_{i=1}^{n} Y_i(X_i + 1)}{2n}.$$

Let $Z = Y(X+1)$, then from previous calculation,

$$E(Y(X+1)/2) = E(Z)/2 = 1/\theta$$

and

$$Var(Y(X+1))/4 = (E(Z^2) - (E(Z))^2)/4 = 1/(2\theta^2).$$

By central limit theorem, $\sqrt{n}(1/\tilde{\theta} - 1/\theta)$ is aymptotically normal with mean zero and variance $1/(2\theta^2)$. By delta method, $\sqrt{n}(\tilde{\theta} - \theta)$ is aymptotically normal with mean zero and variance $\theta^4/(2\theta^2) = \theta^2/2$, which implies that $\tilde{\theta}$ is consistent for θ.

Similarly, $\sqrt{n}(1/\hat{\theta} - 1/\theta)$ is aymptotically normal with mean zero and variance $Var(Y) = 1/\theta^2$, which implies that $\sqrt{n}(\hat{\theta} - \theta)$ is aymptotically normal with mean zero and variance $\theta^4/\theta^2 = \theta^2 < \theta^2/2$. Therefore, $\tilde{\theta}$ has a smaller asymptotic variance than $\hat{\theta}$.

(f) Note that $P(Y_i^* = 2) = P(Y_i > 2) = e^{-2\theta}$ for $i = 1, \ldots, n$. Let $T_n = \sum_{i=1}^{n} I(Y_i^* = 2)/n$, then T_n is consistent for $e^{-2\theta}$ by the law of large numbers and $(\log T_n)/(-2)$ is consistent for θ by the continuous mapping theorem.

(g) By delta method, $\sqrt{n}(\log\tilde{\theta} - \log\theta)$ is aymptotically normal with mean zero and variance $(\theta^4 Var(Y(X+1))/4)(1/\theta)^2 = 1/2$. An approximate $(1-\alpha)$ confidence interval for θ^2 is $(\log\tilde{\theta} - c, \log\tilde{\theta} + c)$, where

$$c = \sqrt{1/2}\,\Phi^{-1}(\alpha/2)/\sqrt{n}.$$