

Part I

- 1.** Suppose λ follows a $\text{Gamma}(\alpha, \beta)$ distribution with the probability density function

$$f(\lambda) = \frac{\lambda^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha} e^{-\lambda/\beta}$$

for $\lambda > 0$ and some $\alpha, \beta > 0$. Conditional on λ , X follows a $\text{Poisson}(\lambda)$ distribution.

- a)** Derive the mean and variance for X .
 - b)** What is the marginal distribution for X ?
 - c)** Derive the conditional distribution of λ given $X = x$ for some $x > 0$.
 - d)** State $E(\lambda | X = x)$ for some $x > 0$.
- 2.** Suppose $\mathbf{X} = (X_1, X_2, X_3)^T$ follows a multivariate normal distribution with mean $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)^T$ and covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}.$$

- a)** What is the conditional distribution of (X_1, X_2) given $X_3 = x_3$? Provide expressions for the mean and covariance matrix for this conditional distribution.
- b)** Provide a general condition under which X_1 and X_2 are independent given that $X_3 = x_3$.

Part II

For problems **3** and **4**, let $\{X_n, n = 1, 2, \dots\}$ be a sequence of random variables.

- 3.** Show that, for a random variable X , if $E|X_n - X| \rightarrow 0$, then $X_n \rightarrow X$ in probability.
- 4.** Prove that if $X_n \rightarrow c$ in distribution for a constant c , then $X_n \rightarrow c$ in probability.

Part III

Let U_1, \dots, U_n be iid $\text{Uniform}(0,1)$ random variables. Let $U_{(1)}, \dots, U_{(n)}$ be the corresponding order statistics and suppose $1 \leq k < \ell \leq n$.

5. Find the joint pdf for $(U_{(k)}, U_{(\ell)})$.
6. What is the conditional pdf of $U_{(k)}$ given $U_{(\ell)} = u_\ell$?
7. Provide the conditional mean and variance of $U_{(k)}$ given $U_{(\ell)} = u_\ell$.
8. Compute $\text{cov}(U_{(k)}, U_{(\ell)})$.
9. Let $X_1 = U_{(k)}$ and $X_2 = U_{(\ell)} - U_{(k)}$. What is the joint pdf for (X_1, X_2) ?
10. Let $Y_n = (\prod_{i=1}^n U_i)^{-1/n}$. Show that $\sqrt{n}(Y_n - e)$ converges in distribution to a Normal random variable as $n \rightarrow \infty$, where e is the base of the natural logarithm. Provide the mean and variance for the asymptotic distribution.

1. a) $E(X) = E\{E(X|\lambda)\} = E(\lambda) = \alpha\beta$, and $\text{var}(X) = E\{\text{var}(X|\lambda)\} + \text{var}\{E(X|\lambda)\} = E(\lambda) + \text{var}(\lambda) = \alpha\beta + \alpha\beta^2$.

b) The marginal pmf of X is

$$\begin{aligned} f_X(x) &= \int_0^\infty \frac{\lambda^x e^{-\lambda}}{x!} \frac{1}{\Gamma(\alpha)\beta^\alpha} \lambda^{\alpha-1} e^{-\lambda/\beta} d\lambda \\ &= \frac{1}{\Gamma(\alpha)x!\beta^\alpha} \int_0^\infty \lambda^{x+\alpha-1} e^{-\lambda(1+\beta^{-1})} d\lambda \\ &= \frac{\Gamma(x+\alpha)}{\Gamma(\alpha)x!} \beta^{-\alpha} (1+\beta^{-1})^{-(x+\alpha)} \\ &= \frac{\Gamma(x+\alpha)}{\Gamma(\alpha)x!} \left(\frac{1}{1+\beta}\right)^\alpha \left(\frac{\beta}{1+\beta}\right)^x \end{aligned}$$

for $x = 0, 1, 2, \dots$

c) $f(\lambda \mid x) = \frac{1}{\Gamma(x+\alpha)} (1+\beta^{-1})^{x+\alpha} \lambda^{x+\alpha-1} e^{-\lambda(1+\beta^{-1})}$, and therefore $[\lambda \mid X]$ follows a Gamma($\alpha + X, \frac{\beta}{1+\beta}$).

d) $E(\lambda|X) = \frac{(\alpha+X)\beta}{1+\beta}$.

2. a) $[(X_1, X_2) \mid X_3 = x_3]$ follows a bivariate normal distribution with mean and covariance

$$\begin{aligned} \boldsymbol{\mu}_{12|3} &= \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} - \begin{pmatrix} \sigma_{13} \\ \sigma_{23} \end{pmatrix} \sigma_{33}^{-1} (x_3 - \mu_3) \\ \boldsymbol{\Sigma}_{12|3} &= \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} - \sigma_{33}^{-1} \begin{pmatrix} \sigma_{13} \\ \sigma_{23} \end{pmatrix} \begin{pmatrix} \sigma_{13} \\ \sigma_{23} \end{pmatrix}^T. \end{aligned}$$

b) Given $X_3 = x_3$, X_1 and X_2 are independent if and only if $\sigma_{12} - \sigma_{13}\sigma_{23}\sigma_{33}^{-1} = 0$.

3. By the Markov inequality, for any $\epsilon > 0$,

$$P(|X_n - X| > \epsilon) \leq \frac{E|X_n - X|}{\epsilon} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

hence $X_n \rightarrow X$ in probability.

4. Let $F_n(x)$ be the cdf for X_n . For any $\epsilon > 0$, $F_n(c-\epsilon) \rightarrow 0$ and $F_n(c+\epsilon) \rightarrow 1$, and therefore

$$P(|X_n - c| > \epsilon) = 1 - F_n(c+\epsilon) + F_n(c-\epsilon) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

5. The joint density function for $(U_{(k)}, U_{(\ell)})$ is

$$f_{U_{(k)}, U_{(\ell)}}(s, t) = \frac{n!}{(k-1)!(\ell-k-1)!(n-\ell)!} s^{k-1} (t-s)^{\ell-k-1} (1-t)^{n-\ell} \quad \text{for } 0 < s < t < 1.$$

6. The marginal distribution for $U_{(\ell)}$ is

$$f_{U_{(\ell)}}(t) = \frac{n!}{(\ell-1)!(n-\ell)!} t^{\ell-1} (1-t)^{n-\ell} \quad \text{for } 0 < t < 1,$$

therefore,

$$f_{U_{(k)}|U_{(\ell)}=t}(s | t) = \frac{(\ell-1)!}{(k-1)!(\ell-k-1)!} \left(\frac{s}{t}\right)^{k-1} \left(1 - \frac{s}{t}\right)^{\ell-k-1} t^{-1} \quad \text{for } 0 < s < t < 1.$$

This means $[U_{(k)} | U_{(\ell)} = t] \sim tBeta(k, \ell - k)$.

7. From the previous question $[U_{(k)} | U_{(\ell)} = t] \sim tBeta(k, \ell - k)$, therefore $E(U_{(k)} | U_{(\ell)} = t) = \frac{tk}{\ell}$ and $\text{var}(U_{(k)} | U_{(\ell)} = t) = \frac{t^2 k (\ell - k)}{\ell^2 (\ell + 1)}$.
8. $U_{(\ell)} \sim Beta(\ell, n - \ell + 1)$, $U_{(k)} \sim Beta(k, n - k + 1)$, therefore $E(U_{(\ell)}) = \frac{\ell}{n+1}$ and $E(U_{(k)}) = \frac{k}{n+1}$.

$$\begin{aligned} E(U_{(\ell)} U_{(k)}) &= E\left\{E(U_{(k)} | U_{(\ell)}) U_{(\ell)}\right\} \\ &= E\frac{k}{\ell} U_{(\ell)}^2 \\ &= \frac{k}{\ell} \left\{ \frac{\ell(n-\ell+1)}{(n+1)^2(n+2)} + \left(\frac{\ell}{n+1}\right)^2 \right\} \\ &= \frac{k}{\ell} \left\{ \frac{\ell(\ell+1)}{(n+1)(n+2)} \right\} \\ &= \frac{k(\ell+1)}{(n+1)(n+2)}, \end{aligned}$$

and therefore

$$\text{cov}\{U_{(\ell)}, U_{(k)}\} = \frac{k(\ell+1)}{(n+1)(n+2)} - \frac{k\ell}{(n+1)^2} = \frac{k(n-\ell+1)}{(n+1)^2(n+2)}.$$

9. The Jacobian is

$$J = \det \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = 1,$$

and therefore

$$f_{(X_1, X_2)}(x_1, x_2) = \frac{n!}{(k-1)!(\ell-k-1)!(n-\ell)!} x_1^{k-1} x_2^{\ell-k-1} (1-x_1-x_2)^{n-\ell}$$

for $0 < x_1, x_2 < 1$ and $x_1 + x_2 < 1$.

- 10.** Notice that $W_n = \log(Y_n) = \frac{1}{n} \sum_{i=1}^n -\log(U_i)$, $V_i = -\log(U_i)$ has a cdf $F_V(v) = P(-\log(U) < v) = P(U > e^{-v}) = 1 - e^{-v}$ for $v > 0$, and hence V_i are iid $\text{Exp}(1)$ random variables. By the central limit theorem,

$$\sqrt{n}(W_n - 1) \xrightarrow{d} \text{Normal}(0, 1).$$

Since $Y_n = \exp(W_n)$, by the delta method

$$\sqrt{n}(Y_n - e) \xrightarrow{d} \text{Normal}(0, e^2).$$

In Parts **I**, **II**, and **III** below, we consider a joint distribution for random pairs (X, Y) depending upon a parameter vector $\boldsymbol{\eta} = (\lambda, \beta) \in (0, \infty)^2$ specified by the marginal density for X

$$f_X(x; \boldsymbol{\eta}) = \lambda \exp(-\lambda x) I[x > 0]$$

and conditional density for $Y | X = x$

$$f_{Y|X}(y | x; \boldsymbol{\eta}) = \beta x \exp(-\beta xy) I[y > 0].$$

(That is, X has an exponential distribution with rate λ , and conditional on $X = x$ the random variable Y is exponential with rate βx .) We will write $f(x, y; \boldsymbol{\eta})$ for the joint pdf of (X, Y) .

Part I

Problems **1-5** concern a single pair (X, Y) with joint distribution described above.

1. Give the marginal pdf of Y , say $f_Y(y; \boldsymbol{\eta})$. Does Y have a finite mean? Explain.

2. Identify a function of x and $\boldsymbol{\eta}$, say $\hat{y}(x; \boldsymbol{\eta})$, that minimizes

$$\mathbb{E}_{Y|X=x}^{\boldsymbol{\eta}} (Y - \hat{y}(x; \boldsymbol{\eta}))^2 \tag{*}$$

over choices of such a function. ($\mathbb{E}_{Y|X}^{\boldsymbol{\eta}}$ is standing here for expectation according to the $\boldsymbol{\eta}$ conditional distribution of $Y | X = x$.) What is the minimum value of (*) possible?

3. Is the entire family of joint distributions for (X, Y) (indexed by $\boldsymbol{\eta} \in (0, \infty)^2$ and described above) a regular exponential family? Explain.

4. Find the 2×2 Fisher Information matrix (about the parameter $\boldsymbol{\eta}$ at a point $\boldsymbol{\eta}_0 \in (0, \infty)^2$) for a single pair (X, Y) .

5. For 0-1 loss in a decision between $\boldsymbol{\eta}_0 = (1, 1)$ and $\boldsymbol{\eta}_1 = (2, 2)$ with prior probabilities $\pi_0 = \pi_1 = .5$, identify an optimal decision rule $d(x, y)$.

Part II

6. $\{P_\theta\}_{\theta \in \Theta}$ be a family of distributions for a random vector \mathbf{Z} . What does it mean for a statistic $S(\mathbf{Z})$ to be a minimal sufficient statistic for $\{P_\theta\}_{\theta \in \Theta}$?

For the rest of **Part II** (problems 7-12), let $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ be iid with joint pdf $f(x, y; \boldsymbol{\eta})$ defined on page 1.

7. For the full family of distributions (for all n observations (X_i, Y_i)) indexed by $\boldsymbol{\eta} \in (0, \infty)^2$ identify a low-dimensional minimal sufficient statistic, and argue that it is indeed minimal sufficient.

8. For the small sub-family of 2 distributions (for all n observations (X_i, Y_i)) indexed by $\boldsymbol{\eta} \in \{(1,1), (2,2)\}$ identify a 1-dimensional minimal sufficient statistic and say why you know it is minimal sufficient.

9. Find the maximum likelihood estimator for the vector $\boldsymbol{\eta}$. (Argue that your estimator does maximize the likelihood.)

10. Based on your answers to problems 4 and 9, specify the large sample (joint) distribution of the MLE vector. Use this and specify an approximate 95% confidence region for the vector $\boldsymbol{\eta}$.

For problems 11 and 12, let $g(\lambda, \beta) \geq 0$ specify a (possibly "improper" in the case that its integral over all $\boldsymbol{\eta}$ is not finite) prior distribution for $\boldsymbol{\eta}$ on $(0, \infty) \times \Re$.

11. Does the improper "uniform prior" specified by $g(\lambda, \beta) = 1$ lead to a "proper posterior" based on n iid pairs (X_i, Y_i) ? Explain.

12. For the proper prior with joint density

$$g(\lambda, \beta) \propto \exp(-\lambda - \beta),$$

identify the posterior distribution based on n iid pairs (X_i, Y_i) .

Part III

Now consider a hypothetical censoring mechanism that is independent of (X_i, Y_i) and makes (only) X_i available with probability p_x and makes (only) Y_i available with probability p_y , where $p_x + p_y < 1$. That is, the censoring mechanism produces data cases of the types indicated in the table below.

Data Type/Case	Probability
X	p_x
Y	p_y
(X, Y)	$1 - p_x - p_y$

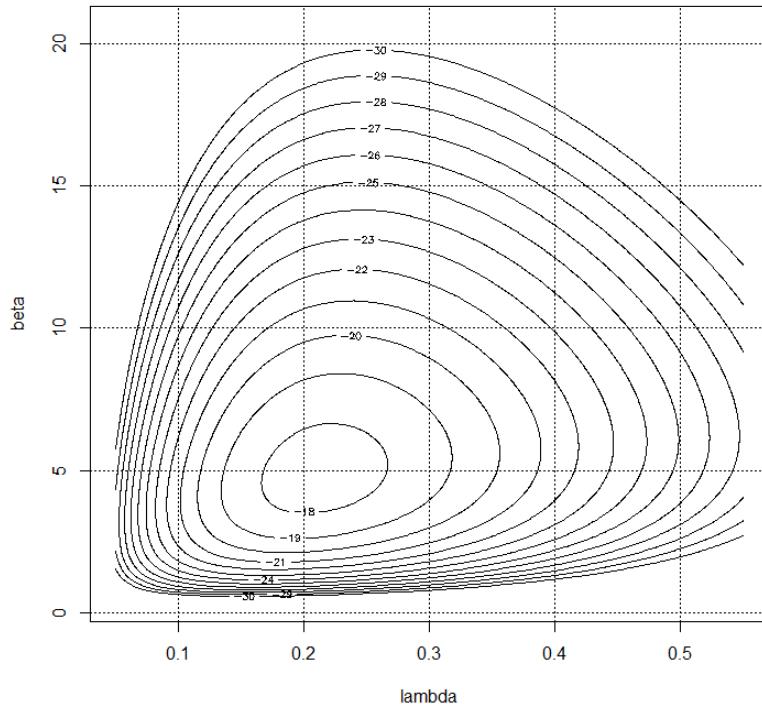
Assume now that one has n such iid data cases, a number n_x of which have available only values X_i , a number n_y of which have available only values Y_i , and a number $n_C = n - n_x - n_y$ of which are "complete," having available the pair (X_i, Y_i) .

13. Treating the values p_x and p_y as known, give a likelihood function for $\boldsymbol{\eta}$ in this scenario. You may use the notations $f, f_X, f_Y, f_{Y|X}$, and $f_{X|Y}$. That is, there is no need to use their explicit forms for the joint pdf, marginal pdfs, or conditional pdfs (but do show dependence upon parameters).

A simulated data set of $n = 20$ data cases generated using a particular set of p 's and a particular $\boldsymbol{\eta}$ is below.

x	y	x	y	x	y	x	y
---	.05	---	.10	12.12	---	9.40	.03
2.38	---	---	.07	---	.01	---	.01
3.73	.12	4.65	---	.02	---	3.22	.02
---	.07	5.68	---	5.74	.03	.89	.05
10.23	---	3.99	---	2.29	---	.14	2.02

Below is a contour plot for (and some other information about) a log-likelihood function $l(\lambda, \beta)$ based on the dataset on the previous page.



$$l(.212, 4.885) = -17.567$$

$$\frac{\partial}{\partial \lambda} l(\lambda, \beta) \Big|_{(.212, 4.885)} = 0$$

$$\frac{\partial}{\partial \beta} l(\lambda, \beta) \Big|_{(.212, 4.885)} = 0$$

$$\frac{\partial^2}{\partial \lambda^2} l(\lambda, \beta) \Big|_{(.212, 4.885)} = -385.83$$

$$\frac{\partial^2}{\partial \beta^2} l(\lambda, \beta) \Big|_{(.212, 4.885)} = -.368$$

$$\frac{\partial^2}{\partial \lambda \partial \beta} l(\lambda, \beta) \Big|_{(.212, 4.885)} = 2.386$$

14. Using the information concerning the plot, give Wald approximate 95% two-sided confidence limits

- a) for λ , and
- b) for β .

15. What is an approximate p -value for a likelihood ratio test of $H_0 : (\lambda, \beta) = (.3, 10)$?

Part IV

This part concerns a *context different from the earlier parts*. Here consider iid observations X_1, X_2, \dots, X_n with marginal pdf

$$f(x; \lambda, c) = \lambda \exp(-\lambda(x-c)) I[x > c]$$

(This is a non-regular family of shifted exponential distributions on \mathfrak{R} , with shift parameter c and rate parameter $\lambda > 0$.)

For \bar{X}_n the sample mean and $m_n = \min(X_1, X_2, \dots, X_n)$, it is a fact that you may use without proof that

$$\begin{pmatrix} \sqrt{n} \left(\bar{X}_n - c - \frac{1}{\lambda} \right) \\ n(m_n - c) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} U \\ V \end{pmatrix}$$

for U normal with mean 0 and standard deviation $\frac{1}{\lambda}$ *independent of* V that is (un-shifted) exponential with rate parameter λ .

A second order Taylor approximation of a function $g(u, v)$ at the point (u_0, v_0) is

$$\begin{aligned} g(u, v) \approx & g(u_0, v_0) + (u - u_0) g_u(u_0, v_0) + (v - v_0) g_v(u_0, v_0) + \frac{1}{2} (u - u_0)^2 g_{uu}(u_0, v_0) \\ & + \frac{1}{2} (v - v_0)^2 g_{vv}(u_0, v_0) + (u - u_0)(v - v_0) g_{uv}(u_0, v_0) \end{aligned}$$

where the subscripts on g indicate various partial derivatives.

16. Let $l_n(\lambda, c)$ be the log likelihood function, and $\begin{pmatrix} \hat{\lambda}_n \\ \hat{c}_n \end{pmatrix}$ the maximum likelihood estimator for the vector $\begin{pmatrix} \lambda \\ c \end{pmatrix}$. Apply a second Taylor order approximation of the log likelihood function at the MLE vector to show that for large n ,

$$2(l_n(\hat{\lambda}_n, \hat{c}_n) - l_n(1, 0)) \approx 2n\hat{\lambda}_n \hat{c}_n + \frac{n}{\hat{\lambda}_n^2} (1 - \hat{\lambda}_n)^2 + 2n\hat{c}_n (1 - \hat{\lambda}_n) = \frac{n}{\hat{\lambda}_n^2} (1 - \hat{\lambda}_n)^2 + 2n\hat{c}_n .$$

17. Based on your answer to problem 16, determine the asymptotic distribution of $2(l_n(\hat{\lambda}_n, \hat{c}_n) - l_n(1, 0))$ as $n \rightarrow \infty$ if $\lambda = 1$ and $c = 0$. (This distribution is NOT χ^2_2 .)

Theory II Key 2017 Statistics Prelim 1/6

1. $f(x, y; \lambda, \beta) = \lambda \exp(-\lambda x) x \beta \exp(-x \beta y) I[(x, y) \in (0, \infty)^2]$

$$f_Y(y; \lambda, \beta) = \int_0^\infty x \lambda \beta \exp(-x(\lambda + \beta y)) dx I[y > 0]$$

$$= \frac{\lambda \beta}{(\lambda + \beta y)} \int_0^\infty x(\lambda + \beta y) \exp(-x(\lambda + \beta y)) dx I[y > 0]$$

$$= \frac{\lambda \beta}{(\lambda + \beta y)^2} I[y > 0]$$

$EY = \int_0^\infty y \frac{\lambda \beta}{(\lambda + \beta y)^2} dy$ and $\frac{y}{(\lambda + \beta y)^2} / \frac{1}{y} \rightarrow \frac{1}{\beta^2}$ so that
since $\int_0^\infty \frac{1}{y} dy = \infty$ so also is $EY = \infty$.

2. $\hat{y}(x, \eta) = E_{\eta}[y | X=x] = \frac{1}{\beta x}$

The minimum is the conditional variance of
 $Y | X=x$ i.e. $(\frac{1}{\beta x})^2$

3. $f(x, y; \eta) = x \lambda \beta \exp(-\lambda x - \beta xy) I[(x, y) \in (0, \infty)^2]$

Since the natural parameter space is $(\lambda, \beta) \in (0, \infty)^2$
which is an open rectangle, this is indeed a
regular exponential family with natural sufficient
statistic (X, XY) .

4. $\ln f(x, y; \eta) = \ln x + \ln \lambda + \ln \beta - \lambda x - \beta xy$

$$\frac{\partial}{\partial \lambda} \ln f(x, y; \eta) = \frac{1}{\lambda} - x \quad \frac{\partial}{\partial \beta} \ln f(x, y; \eta) = \frac{1}{\beta} - xy$$

$$\frac{\partial^2}{\partial \lambda^2} \ln f(x, y; \eta) = -\frac{1}{\lambda^2} \quad \frac{\partial^2}{\partial \beta^2} \ln f(x, y; \eta) = -\frac{1}{\beta^2}$$

$$\frac{\partial^2}{\partial \lambda \partial \beta} \ln f(x, y; \eta) = 0$$

$$I(\lambda, \beta) \Big|_{(\lambda_0, \beta_0)} = E_{\eta} \left[\begin{pmatrix} \frac{\partial^2}{\partial \lambda^2} \ln f & \frac{\partial^2}{\partial \lambda \partial \beta} \ln f \\ \frac{\partial^2}{\partial \lambda \partial \beta} \ln f & \frac{\partial^2}{\partial \beta^2} \ln f \end{pmatrix} \right] \Bigg|_{\eta=(\lambda_0, \beta_0)} = \begin{pmatrix} \frac{1}{\lambda_0^2} & 0 \\ 0 & \frac{1}{\beta_0^2} \end{pmatrix}$$

5. This is $I[.5f(x,y;2,2) > .5f(x,y;1,1)]$ i.e.

$$\begin{aligned} I\left[\frac{f(x,y;2,2)}{f(x,y;1,1)} > 1\right] &= I\left[\frac{4x \exp(-2x-2xy)}{x \exp(-x-xy)} > 1\right] \\ &= I[4 \exp(-x-xy) > 1] \\ &= I[x(1+y) < \ln(4)] \end{aligned}$$

6. $S(\underline{z})$ is first sufficient, meaning that the conditional dsn for \underline{z} given $S(\underline{z})$ is the same for every $\theta \in \Theta$. And it is minimal meaning that for any other sufficient statistic $T(\underline{z})$, $\exists q(s)$ such that $T(\underline{z}) = q(S(\underline{z}))$

7. This is a regular exponential family with

$$f(x,y; \lambda, \beta) = \underbrace{\lambda^n}_{C(\eta)} \underbrace{\beta^n}_{h(\underline{x},y)} \prod_{i=1}^n x_i \exp\left(-\lambda \sum_{i=1}^n x_i - \beta \sum_{i=1}^n x_i y_i\right)$$

and in such a family $(T_1(\underline{x},y), T_2(\underline{x},y))$ is minimal sufficient.

8. In a two-dsn family the likelihood ratio is minimal sufficient. Here this is

$$\begin{aligned} \frac{4 \sum x_i \exp(-2 \sum x_i - 2 \sum x_i y_i)}{\sum x_i \exp(-\sum x_i - \sum x_i y_i)} &= 4 \exp(-\sum x_i - \sum x_i y_i) \\ &= 4 \exp\left(-\sum_{i=1}^n x_i (1+y_i)\right) \end{aligned}$$

9. The log likelihood is

$$\ell_n(\lambda, \beta) = \ln(f(x_i, y_i; \lambda, \beta)) = n \ln \lambda + n \ln \beta - \lambda \sum x_i - \beta \sum x_i y_i$$

Setting $\frac{\partial}{\partial \lambda} \ell_n(\lambda, \beta) = 0$ and solving gives $\hat{\lambda}_n = \frac{n}{\sum x_i}$.

and setting $\frac{\partial}{\partial \beta} \ell_n(\lambda, \beta) = 0$ and solving gives $\hat{\beta}_n = \frac{n}{\sum x_i y_i}$.

Note that $\frac{\partial^2}{\partial \lambda^2} \ell_n(\lambda, \beta) = -\frac{n}{\lambda^2}$ $\frac{\partial^2}{\partial \beta^2} \ell_n(\lambda, \beta) = -\frac{n}{\beta^2}$

and $\frac{\partial^2}{\partial \lambda \partial \beta} \ell_n(\lambda, \beta) = 0$ so that the Hessian

for the log-likelihood at $(\hat{\lambda}_n, \hat{\beta}_n)$ is $\text{diag}\left(-\frac{n}{\hat{\lambda}_n^2}, -\frac{n}{\hat{\beta}_n^2}\right)$
 which is negative definite, so that $(\hat{\lambda}_n, \hat{\beta}_n)$ produces
 a unique local maximum and \therefore a maximum of
 $\ell_n(\lambda, \beta)$ on $(0, \infty)^2$. $(\hat{\lambda}_n, \hat{\beta}_n)$ is thus the MLE.

$$10. \quad \sqrt{n} \left(\begin{pmatrix} \hat{\lambda}_n \\ \hat{\beta}_n \end{pmatrix} - \begin{pmatrix} \lambda \\ \beta \end{pmatrix} \right) \xrightarrow{d} MVN(0, I_1(\beta)) \quad 4/6$$

$$I_1^{-1}(\beta) = \begin{pmatrix} \lambda^2 & 0 \\ 0 & \beta^2 \end{pmatrix} \text{ and then}$$

$$\sqrt{n} \left(\begin{pmatrix} \hat{\lambda}_n \\ \hat{\beta}_n \end{pmatrix} - \begin{pmatrix} \lambda \\ \beta \end{pmatrix} \right)' I_1(\beta) \left(\sqrt{n} \left(\begin{pmatrix} \hat{\lambda}_n \\ \hat{\beta}_n \end{pmatrix} - \begin{pmatrix} \lambda \\ \beta \end{pmatrix} \right) \right) \xrightarrow{d} \chi^2_2$$

and further

$$n \left(\begin{pmatrix} \hat{\lambda}_n \\ \hat{\beta}_n \end{pmatrix} - \begin{pmatrix} \lambda \\ \beta \end{pmatrix} \right)' I_1(\beta) \left(\begin{pmatrix} \hat{\lambda}_n \\ \hat{\beta}_n \end{pmatrix} - \begin{pmatrix} \lambda \\ \beta \end{pmatrix} \right) \xrightarrow{d} \chi^2_2$$

So an elliptical large n confidence set for (λ, β) is

$$\left\{ \begin{pmatrix} \lambda \\ \beta \end{pmatrix} \mid \left(\begin{pmatrix} \hat{\lambda}_n \\ \hat{\beta}_n \end{pmatrix} - \begin{pmatrix} \lambda \\ \beta \end{pmatrix} \right)' \text{diag} \left(\frac{1}{\hat{\lambda}_n^2}, \frac{1}{\hat{\beta}_n^2} \right) \left(\begin{pmatrix} \hat{\lambda}_n \\ \hat{\beta}_n \end{pmatrix} - \begin{pmatrix} \lambda \\ \beta \end{pmatrix} \right) \leq \frac{1}{n} \chi^2_n \right\}$$

for χ^2_n the upper 5% pt of the χ^2_2 dist (2.92).

11.

$$\text{Again } f(\tilde{x}, \tilde{y}; \lambda, \beta) = \lambda^n \beta^n \prod_{i=1}^n x_i \exp \left(-\lambda \sum_{i=1}^n x_i - \beta \sum_{i=1}^n x_i y_i \right)$$

so that

$$\begin{aligned} \int_0^\infty \int_0^\infty 1 \cdot f(\tilde{x}, \tilde{y}; \lambda, \beta) d\lambda d\beta &= \prod_{i=1}^n x_i \int_0^\infty \lambda^n \exp(-T_1 \lambda) d\lambda \int_0^\infty \beta^n \exp(-T_2 \beta) d\beta \\ &= \prod_{i=1}^n x_i \frac{1}{(T_1)^{n+1}} \int_0^\infty u^n \exp(-u) du \frac{1}{(T_2)^{n+1}} \int_0^\infty u^n \exp(-u) du \\ &= \prod_{i=1}^n x_i \frac{1}{(T_1)^{n+1}} \frac{1}{(T_2)^{n+1}} \Gamma(n+1) \Gamma(n+1) < \infty \end{aligned}$$

and the posterior for $g(\lambda, \beta) \propto 1$ is indeed "proper."

12. For $g(\lambda, \beta) \propto \exp(-\lambda - \beta)$ the posterior is proportional to

$$\lambda^n \beta^n \exp \left(-\lambda \left(\sum_{i=1}^n x_i + 1 \right) - \beta \left(\sum_{i=1}^n x_i y_i + 1 \right) \right)$$

This is a product of Γ densities. $\lambda \sim \Gamma(n+1, \sum_{i=1}^n x_i + 1)$

$$\beta \sim \Gamma(n+1, \sum_{i=1}^n x_i y_i + 1)$$

13. A likelihood function is

$$\prod_{\substack{i \text{ s.t. } x_i \\ \text{case is} \\ \text{complete}}} f(x_i, y_i; \lambda, \beta) \prod_{\substack{i \text{ s.t. } y_i \\ \text{is missing}}} f(x_i; \lambda) \prod_{\substack{i \text{ s.t. } x_i \\ \text{is missing}}} f(y_i; \lambda, \beta)$$

14. The MLE is $\begin{pmatrix} \hat{\lambda} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} .212 \\ 4.885 \end{pmatrix}$

The negative Hessian at the MLE is $-H = \begin{pmatrix} 385.83 & -2.386 \\ -2.386 & .368 \end{pmatrix}$

Its inverse (an estimated covariance matrix) is $\begin{pmatrix} 2.7 \times 10^{-3} & -2.386 \\ -2.386 & 2.831 \end{pmatrix}$

So approximate 95% confidence limits are

$$\text{For } \lambda: .212 \pm 1.96 \sqrt{2.7 \times 10^{-3}} \quad \text{i.e. } .212 \pm .101$$

$$\text{For } \beta: 4.885 \pm 1.96 \sqrt{2.831} \quad \text{i.e. } 4.885 \pm 3.297$$

15. The loglikelihood at (.3, 10) is approximately -20.7.

The maximum loglikelihood is -17.6. Thus

$$Z(\lambda(.212, 4.885) - \lambda(.3, 10)) \approx Z(3.1) = 6.2$$

The large n reference dist is χ^2_2 . The p-value is

$$P[\text{a } \chi^2_2 \text{ r.v.} > 6.2] \text{ which is between .025 and .01.}$$

16. The loglikelihood is

$$l_n(\lambda, c) = n \ln(\lambda) - \lambda \sum_{i=1}^n (x_i - c)$$

For any λ this is increasing in c up to $c = \min(x_1, \dots, x_n)$
 $\hat{c}_n = \min(x_1, \dots, x_n)$. Then

$$\frac{\partial}{\partial \lambda} l_n(\lambda, \hat{c}_n) = \frac{n}{\lambda} - \sum_{i=1}^n (x_i - \hat{c}_n)$$

$$\Rightarrow \hat{\lambda}_n = \frac{n}{\sum_{i=1}^n (x_i - \hat{c}_n)} = \frac{1}{\bar{x}_n - \frac{\hat{c}_n}{n}}$$

$$\frac{\partial}{\partial c} l_n(\lambda, c) = n \lambda \quad \frac{\partial}{\partial \lambda} l_n(\lambda, c) = \frac{n}{\lambda} - \sum_{i=1}^n (x_i - c) = \frac{n}{\lambda} - \sum x_i + nc$$

$$\frac{\partial^2}{\partial c^2} l_n(\lambda, c) = 0 \quad \frac{\partial^2}{\partial \lambda^2} l_n(\lambda, c) = -\frac{n}{\lambda^2} \quad \frac{\partial^2}{\partial \lambda \partial c} l_n(\lambda, c) = n$$

$$\ell_n(1,0) \approx \ell_n(\hat{\lambda}_n, \hat{c}_n) + (1-\hat{\lambda}_n) \left(\frac{n}{\hat{\lambda}_n} - \sum_{i=1}^n (x_i - \hat{c}_n) \right) = 0 (0 - \hat{c}_n) (n \hat{\lambda}_n) + \frac{1}{2} (1-\hat{\lambda}_n)^2 \left(-\frac{n}{\hat{\lambda}_n^2} \right) + \frac{1}{2} (0 - \hat{c}_n)^2 (0) + (1-\hat{\lambda}_n)(0 - \hat{c}_n)n$$

$$\Rightarrow 2(\ell_n(\hat{\lambda}_n, \hat{c}_n) - \ell_n(1,0)) \approx 2\hat{\lambda}_n n \hat{c}_n + \frac{1}{\hat{\lambda}_n^2} n (1-\hat{\lambda}_n)^2 + 2n \hat{c}_n (1-\hat{\lambda}_n) = 2n \hat{c}_n + \frac{1}{\hat{\lambda}_n^2} n (1-\hat{\lambda}_n)^2$$

17. Under $(\lambda, c) = (1, 0)$ the first term above converges in dsh to $2V$ for $V \sim \text{Exp}(1)$. This is χ^2_2 .

The 2nd term is

$$\begin{aligned} \frac{1}{(\hat{\lambda}_n)} &= n \left(1 - \frac{1}{\bar{x}_n - \frac{\hat{c}_n}{n}} \right)^2 = \frac{1}{(\hat{\lambda}_n)^2} \left(\sqrt{n} \left(1 - \frac{1}{(\bar{x}_n - \frac{\hat{c}_n}{n})} \right) \right)^2 \\ &= \left(\bar{x}_n - \frac{\hat{c}_n}{n} \right)^2 n \left(\frac{\bar{x}_n - \frac{\hat{c}_n}{n} - 1}{\bar{x}_n - \frac{\hat{c}_n}{n}} \right)^2 \\ &= n \left((\bar{x}_n - 1) - \frac{\hat{c}_n}{n} \right)^2 \\ &= n (\bar{x}_n - 1)^2 - 2\hat{c}_n (\bar{x}_n - 1) + \frac{1}{n} \hat{c}_n^2 \end{aligned}$$

Now $2\hat{c}_n (\bar{x}_n - 1)$ and $\frac{1}{n} \hat{c}_n^2$ both converge to 0 in probability under $(\lambda, c) = (1, 0)$. By the CLT and cont \leq mapping theorem $n(\bar{x}_n - 1)^2$ converges in dsh to χ^2_1 independent of $2n\hat{c}_n$. So ultimately the whole converges in dsh to χ^2_3 under $(\lambda, c) = (1, 0)$

Part I

Let Ω be a non-empty set. Suppose \mathcal{A} is a class of subsets of Ω such that

- $\Omega \in \mathcal{A}$
- $A \in \mathcal{A}$ implies that $A^c \in \mathcal{A}$
- \mathcal{A} is closed under finite *disjoint* unions.

1. Show by an example that \mathcal{A} need *not* be an algebra.

Part II

Let $\mathcal{F}_1, \mathcal{F}_2, \dots$ be classes of subsets of a common non-empty set Ω .

2. Suppose that the \mathcal{F}_n 's are algebras satisfying $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ for $n = 1, 2, \dots$. Show that $\cup_{n=1}^{\infty} \mathcal{F}_n$ is an algebra.
3. Suppose that the \mathcal{F}_n 's are σ -algebras satisfying $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ for $n = 1, 2, \dots$. Show by an example that $\cup_{n=1}^{\infty} \mathcal{F}_n$ need *not* be a σ -algebra.

Part III

Let (Ω, \mathcal{F}, P) be a probability space. Suppose $\{A_n \in \mathcal{F}, n \geq 1\}$ are events with $P(A_n) = 1$ for all n .

4. Show that $P(\cap_{n=1}^{\infty} A_n) = 1$.

Part IV

Let $\{X_n\}_{n \geq 1}$ be a sequence of iid random variables with common distribution function F . Assume that $E[|X_1|] < \infty$.

5. Show that $X_n/n \xrightarrow{a.s.} 0$ as $n \rightarrow \infty$.
6. Assume that $F(x) = 1 - \exp(-x)$, $x > 0$, that is, X_n 's are iid exponential random variables with mean 1. Show that

$$P\left(\limsup_{n \rightarrow \infty} \frac{X_n}{\log n} = 1\right) = 1.$$

Hint: You may use the result in problem 4 and the fact that for positive ϵ_k 's with $\epsilon_k \downarrow 0$

$$\left\{ \limsup_{n \rightarrow \infty} \frac{X_n}{\log n} = 1 \right\} = \left[\cap_k \left\{ \liminf_{n \rightarrow \infty} \left\{ \frac{X_n}{\log n} \leq 1 + \epsilon_k \right\} \right\} \right] \cap \left[\cap_k \left\{ \frac{X_n}{\log n} > 1 - \epsilon_k \text{ i.o.} \right\} \right].$$

Here i.o. refers to infinitely often (in n).

Part V

7. Let $\{X_n, n \geq 1\}$ be a sequence of random variables on a probability space (Ω, \mathcal{F}, P) and suppose that

$$\sup_{n \geq 1} E(|X_n|g(|X_n|)) < \infty$$

for some nondecreasing function $g : [0, \infty] \rightarrow [0, \infty]$ with $g(x) < \infty$ for $0 \leq x < \infty$ and $\lim_{x \rightarrow \infty} g(x) = \infty$. Prove that $\{X_n, n \geq 1\}$ is uniformly integrable.

Part VI

Let $\{X_n, n \geq 1\}$ be a sequence random variables on a probability space (Ω, \mathcal{F}, P) . Let $S_n = \sum_{j=1}^n X_j$ for $n \geq 1$.

8. Show that $X_n \xrightarrow{a.s.} 0$ as $n \rightarrow \infty$ implies $S_n/n \xrightarrow{a.s.} 0$ as $n \rightarrow \infty$.
9. Show by an example that $S_n/n \xrightarrow{a.s.} 0$ as $n \rightarrow \infty$ does *not* imply $X_n \xrightarrow{a.s.} 0$ as $n \rightarrow \infty$.
10. Let $p \geq 1$. Show that $X_n \xrightarrow{L_p} 0$ as $n \rightarrow \infty$ implies $S_n/n \xrightarrow{L_p} 0$ as $n \rightarrow \infty$.
11. Show that $S_n/n \xrightarrow{p} 0$ as $n \rightarrow \infty$ implies $X_n/n \xrightarrow{p} 0$ as $n \rightarrow \infty$.

Part VII

Let $\{X_n, n \geq 1\}$ be a sequence of independent random variables with

$$\begin{aligned} P(X_n = 0) &= \frac{1}{2} \left(1 - \frac{1}{n^2}\right), \\ P(X_n = 1) &= \frac{1}{4} = P(X_n = -1), \\ P(X_n = n) &= \frac{1}{4n^2} = P(X_n = -n), \end{aligned}$$

for all $n \geq 1$.

12. Show that the triangular array $\{X_{nj} : 1 \leq j \leq n\}_{n \geq 1}$ with $X_{nj} \equiv X_j/\sqrt{n}$, $1 \leq j \leq n$, $n \geq 1$ does not satisfy the Lindeberg condition.
13. Let $S_n = \sum_{j=1}^n X_{nj}$, $n \geq 1$. Show that there exists $\sigma \in (0, \infty)$, such that $S_n \xrightarrow{d} N(0, \sigma^2)$ as $n \rightarrow \infty$.
14. Find σ^2 in problem 13.

1. Let $\Omega = \{1, 2, 3, 4\}$ and let

$$\mathcal{A} = \{\phi, \Omega, \{1, 2\}, \{3, 4\}, \{1, 3\}, \{2, 4\}, \{1, 4\}, \{2, 3\}\}$$

It is easy to see that \mathcal{A} satisfies all three conditions in the question, but \mathcal{A} is not an algebra as it is not closed under finite unions.

2. Note that $\Omega \in \bigcup_{n=1}^{\infty} \mathcal{F}_n$ as $\Omega \in \mathcal{F}_n$ for all n . Let $A \in \bigcup_{n=1}^{\infty} \mathcal{F}_n$. Thus $A \in \mathcal{F}_{n_1}$ for some n_1 implying $A^c \in \mathcal{F}_{n_1}$ as \mathcal{F}_{n_1} is an algebra. Hence $A^c \in \bigcup_{n=1}^{\infty} \mathcal{F}_n$. Finally let $A, B \in \bigcup_{n=1}^{\infty} \mathcal{F}_n$. Thus $A \in \mathcal{F}_{n_1}$ and $B \in \mathcal{F}_{n_2}$ for some n_1, n_2 . Without loss of generality assume $n_1 \leq n_2$. Since $\mathcal{F}_n \uparrow$ as $n \uparrow$, $\mathcal{F}_{n_1} \subset \mathcal{F}_{n_2}$. Thus $A, B \in \mathcal{F}_{n_2}$ implying $A \cup B \in \mathcal{F}_{n_2} \subset \bigcup_n \mathcal{F}_n$ since \mathcal{F}_n 's are algebras.
3. Let $\Omega = \mathbb{N}$, the set of natural numbers. Let \mathcal{F}_n be the σ -algebra that consists of all subsets of $\{1, 2, \dots, n\}$ and their complements in Ω . Thus $\mathcal{F}_1 = \{\phi, \{1\}, \{2, 3, \dots\}, \mathbb{N}\}$, $\mathcal{F}_2 = \{\phi, \{1\}, \{2\}, \{1, 2\}, \{2, 3, \dots\}, \{1, 3, 4, \dots\}, \{3, 4, \dots\}, \mathbb{N}\}$, etc. Note that $\{2i\} \in \bigcup_n \mathcal{F}_n$ for each i , but $\bigcup_i \{2i\}$ does not lie in $\bigcup_n \mathcal{F}_n$ as $\bigcup_i \{2i\} \notin \mathcal{F}_n$ for any n . Thus $\bigcup_n \mathcal{F}_n$ is not a σ -algebra.

4. Note that

$$P\{(\bigcap_n A_n)^c\} = P\{\bigcup_n A_n^c\} \leq \sum_{n=1}^{\infty} P(A_n^c) = \sum_{n=1}^{\infty} 0 = 0.$$

Thus $P\{(\bigcap_n A_n)^c\} = 0$ implying $P\{\bigcap_n A_n\} = 1$.

5. Since $E|X_1| < \infty$, for all $\epsilon > 0$, $\sum_n P(|X_1| > \epsilon n) < \infty$. By Borel-Cantelli lemma, $\sum_n P(|X_n| > \epsilon n) < \infty$ implies $P(|X_n|/n > \epsilon \text{ i.o.}) = 0$. Thus $\limsup_{n \rightarrow \infty} |X_n|/n \leq \epsilon$ with probability one for all $\epsilon > 0$. This in turn implies $X_n/n \xrightarrow{a.s.} 0$.

6. Letting $\epsilon_k \downarrow 0$, we know that

$$\left\{ \limsup_{n \rightarrow \infty} \frac{X_n}{\log n} = 1 \right\} = \left[\bigcap_k \left\{ \liminf_{n \rightarrow \infty} \left\{ \frac{X_n}{\log n} \leq 1 + \epsilon_k \right\} \right\} \right] \cap \left[\bigcap_k \left\{ \frac{X_n}{\log n} > 1 - \epsilon_k \text{ infinitely often}(n) \right\} \right].$$

Now

$$\sum_n P\left(\frac{X_n}{\log n} > 1 - \epsilon_k\right) = \sum_n P(X_n > (1 - \epsilon_k) \log n) = \sum_n \exp\{-(1 - \epsilon_k) \log n\} = \sum_n \frac{1}{n^{1-\epsilon_k}} = \infty.$$

So by Borel-Cantelli results we have

$$P\left(\frac{X_n}{\log n} > 1 - \epsilon_k \text{ infinitely often}(n)\right) = 1.$$

Similarly

$$\sum_n P\left(\frac{X_n}{\log n} > 1 + \epsilon_k\right) = \sum_n P(X_n > (1 + \epsilon_k) \log n) = \sum_n \exp\{-(1 + \epsilon_k) \log n\} = \sum_n \frac{1}{n^{1+\epsilon_k}} < \infty.$$

So by Borel-Cantelli lemma we have

$$P\left(\limsup_{n \rightarrow \infty} \left\{ \frac{X_n}{\log n} > 1 + \epsilon_k \right\}\right) = P\left(\frac{X_n}{\log n} > 1 + \epsilon_k \text{ infinitely often}(n)\right) = 0$$

implying

$$P\left(\liminf_{n \rightarrow \infty} \left\{ \frac{X_n}{\log n} \leq 1 + \epsilon_k \right\}\right) = P\left(\limsup_{n \rightarrow \infty} \left\{ \frac{X_n}{\log n} > 1 + \epsilon_k \right\}\right)^c = 1.$$

Thus 6 follows from 4.

7. Since g is nondecreasing, for $a > 0$

$$\sup_{n \geq 1} E(|X_n| I(|X_n| > a)) \leq \frac{1}{g(a)} \sup_{n \geq 1} E(|X_n| g(|X_n|) I(|X_n| > a)).$$

Let $M \equiv \sup_{n \geq 1} E(|X_n| g(|X_n|))$. Thus

$$\sup_{n \geq 1} E(|X_n| I(|X_n| > a)) \leq \frac{M}{g(a)} \rightarrow 0$$

as $a \rightarrow \infty$. Hence $\{X_n, n \geq 1\}$ is uniformly integrable.

8. This follows since if a sequence converges to zero, its Cesaro averages also converge to zero.

9. Take $X_n(\omega) = (-1)^n$ for all ω .

10. Since $\|X_n\|_p \rightarrow 0$, by Minkowski inequality

$$\left\| \frac{\sum_{i=1}^n X_i}{n} \right\|_p \leq \frac{1}{n} \sum_{i=1}^n \|X_i\|_p \rightarrow 0.$$

11. Note that

$$\frac{X_n}{n} = \frac{S_n}{n} - \frac{n-1}{n} \frac{S_{n-1}}{n-1} \xrightarrow{P} 0 - (1)0 = 0.$$

12. We have $EX_{nj} = 0$, $EX_{nj}^2 = 1/n$ and $v_n^2 = \sum_{j=1}^n EX_{nj}^2 = 1$. Fix $\epsilon > 0$. For all large n , we have $\epsilon\sqrt{n} > 1$. Thus

$$\frac{1}{v_n^2} \sum_{j=1}^n EX_{nj}^2 I(|X_{nj}| > \epsilon v_n) = \frac{1}{n} \sum_{j=1}^n EX_j^2 I(|X_j| > \epsilon\sqrt{n}) = \frac{1}{n} \sum_{j=1}^n \frac{1}{2} I(|j| > \epsilon\sqrt{n}) = \frac{n - j_n + 1}{2n} \rightarrow \frac{1}{2},$$

as $n \rightarrow \infty$, where j_n is the smallest integer such that $j_n \geq \epsilon\sqrt{n}$.

13. Let $Y_n \equiv X_n I(|X_n| \leq 1)$. Note that $EY_n = 0$, $EY_n^2 = \frac{1}{2}$, and $s_n^2 \equiv \sum_{j=1}^n EY_j^2 = \frac{n}{2}$. Fix $\epsilon > 0$. For all large n , we have $\epsilon\sqrt{n/2} > 1$. Thus

$$\frac{1}{s_n^2} \sum_{j=1}^n EY_j^2 I(|Y_j| > \epsilon s_n) = \frac{2}{n} \sum_{j=1}^n EY_j^2 I(|Y_j| > \epsilon\sqrt{n/2}) \rightarrow 0,$$

as $n \rightarrow \infty$. Thus the Lindeberg condition is satisfied. So

$$\frac{\sum_{j=1}^n Y_j}{s_n} \xrightarrow{d} N(0, 1)$$

as $n \rightarrow \infty$. That is, $\sum_{j=1}^n Y_j / \sqrt{n} \xrightarrow{d} N(0, 1/2)$.

Let $Z_n \equiv X_n I(|X_n| > 1)$. Thus $X_n = Y_n + Z_n$. Now $Z_n^{(1)} \equiv Z_n I(|Z_n| \leq 1) \equiv 0$. So $\sum_{n=1}^{\infty} EZ_n^{(1)} = 0$ and $\sum_{n=1}^{\infty} \text{var}(Z_n^{(1)}) = 0$. Also

$$\sum_{n=1}^{\infty} P(|Z_n| > 1) = \sum_{n=1}^{\infty} P(X_n = \pm n) = \sum_{n=1}^{\infty} \frac{1}{2n^2} < \infty.$$

Since Z_n 's are independent by Kolmogorov's 3-series theorem, $\sum_{j=1}^n Z_j \xrightarrow{a.s.} \sum_{j=1}^{\infty} Z_j < \infty$. So $\sum_{j=1}^n Z_j / \sqrt{n} \xrightarrow{a.s.} 0$. Hence, by Slutsky's theorem $S_n = \sum_{j=1}^n X_j / \sqrt{n} = \sum_{j=1}^n Y_j / \sqrt{n} + \sum_{j=1}^n Z_j / \sqrt{n} \xrightarrow{d} N(0, 1/2)$.

14. From the above discussion it follows that $\sigma^2 = \frac{1}{2}$.

Part I

Let X_1, \dots, X_n be an iid sample from $N(\mu, 1)$ and Y_1, \dots, Y_m be an iid sample from $N(\theta, 1)$. Assume X_1, \dots, X_n and Y_1, \dots, Y_m are independent. Consider hypothesis

$$H_0 : \mu - \theta = 1, \quad vs. \quad H_1 : \mu - \theta \neq 1.$$

1. Find the likelihood ratio test (identify a test statistic and rejection region).
2. Show that the rejection region in Problem 1 can be represented in terms of $|\bar{x} - \bar{y} - 1|$.

Part II

Suppose $X_1, \dots, X_n, Y_1, \dots, Y_n$ ($n \geq 3$) are independent random variables, with X_i 's iid exponential with mean θ (> 0), and Y_i 's iid exponential with mean $\frac{1}{\theta}$. Define

$$T_1 = n^{-1} \sum_{i=1}^n X_i, \quad T_2 = n^{-1} \sum_{i=1}^n Y_i, \quad T_3 = (n-1) \left(\sum_{i=1}^n Y_i \right)^{-1}.$$

3. Find the Cramer-Rao lower bound for the variance of an unbiased estimator of θ based on $X_1, \dots, X_n, Y_1, \dots, Y_n$.
4. Is there an unbiased estimator of θ for which the bound in Problem 3 is attained? Explain your answer.
5. Find $\hat{\theta}_{MLE}$, the MLE of θ .
6. Show that the MLE in Problem 5 is consistent for θ .
7. Find the limiting distribution of $n^{1/2}(\hat{\theta}_{MLE} - \theta)$. Does the variance of the limiting distribution attain the Cramer-Rao lower bound in Problem 3?
8. Find a suitable α that minimizes the variance of $U = \alpha T_1 + (1 - \alpha) T_3$. Denote this by α_0 . Does the variance of the estimator $\alpha_0 T_1 + (1 - \alpha_0) T_3$ attain the Cramer-Rao lower bound in Problem 3?

Part III

A teacher has her class participate in an experiment to estimate a parameter in the following way. One group of students obtains a sample of $Uniform(0, \theta)$ random variables, X_1, \dots, X_n . These values are recorded on a sheet of paper. A second team uses the data from the first team to compute the areas, X_1^2, \dots, X_n^2 . These results are recorded on a sheet of paper not the same as the sheet with the original data. A third team is then provided with the results from the first two teams to do a statistical analysis to estimate θ , but one of the sheets got lost and

the third team received only a sheet of paper which has either X_1, \dots, X_n or X_1^2, \dots, X_n^2 and does not know which. Due to this complication, the teacher asks you for statistical advice on how to proceed.

You begin by setting up the problem parametrically. Let $\Theta = \{(\theta, i) : \theta > 0, i = 1, 2\}$. If $\zeta = (\theta, i) \in \Theta$, then the density having this parameter is given by

$$f_\zeta(x) = \begin{cases} \frac{1}{\theta} & 0 \leq x \leq \theta \\ \frac{1}{2\theta\sqrt{x}} & 0 \leq x \leq \theta^2 \end{cases} \quad \begin{array}{ll} \text{if } \zeta = (\theta, 1); \\ \text{if } \zeta = (\theta, 2). \end{array}$$

Let $M_n = \max_{1 \leq i \leq n} X_i$.

9. Find $P_{(\theta,1)}(X_1 \leq x)$ and $P_{(\theta,2)}(\sqrt{X_1} \leq x)$, for any $x \in (0, \theta)$.

10. For any $0 < \epsilon < \theta$, show that

$$\lim_{n \rightarrow \infty} P_{(\theta,1)}(M_n < \theta - \epsilon) = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} P_{(\theta,2)}(\sqrt{M_n} < \theta - \epsilon) = 0.$$

Hence, infer that

$$\lim_{n \rightarrow \infty} P_{(\theta,1)}(|M_n - \theta| > \epsilon) = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} P_{(\theta,2)}(|\sqrt{M_n} - \theta| > \epsilon) = 0.$$

Note that the latter results say that M_n is consistent for θ under the model $P_{(\theta,1)}$ and $\sqrt{M_n}$ is consistent for θ under the model $P_{(\theta,2)}$. That is, the consistent estimator of θ changes according to which data $(X_1, \dots, X_n$ or $X_1^2, \dots, X_n^2)$ the third team receives.

11. Suppose we define a random variable

$$Y_n = \frac{\left(\frac{1}{n} \sum_{i=1}^n X_i\right)^2}{\frac{1}{n} \sum_{i=1}^n X_i^2}.$$

Then, show using the Law of Large Numbers that Y_n converges in probability to

$$\begin{cases} \frac{3}{4} & \text{if } \zeta = (\theta, 1), \text{ or} \\ \frac{5}{9} & \text{if } \zeta = (\theta, 2) \end{cases}.$$

12. Suppose the third team wants to construct one consistent estimator of θ regardless of the data $(X_1, \dots, X_n$ or $X_1^2, \dots, X_n^2)$ the team receives. Then, we may define

$$\hat{\theta}_n = \begin{cases} M_n & \text{if } Y_n > \frac{3}{5} \\ \sqrt{M_n} & \text{if } Y_n < \frac{3}{5} \end{cases}$$

Note that $\frac{3}{4} > \frac{3}{5} > \frac{5}{9}$. Now use Problems 10, 11 and the fact that $\frac{3}{4} > \frac{3}{5} > \frac{5}{9}$, and give a heuristic reasoning (or a mathematical reasoning) as to why

$$\lim_{n \rightarrow \infty} P_\zeta(|\hat{\theta}_n - \theta| > \epsilon) = 0 \quad \text{for all } \zeta = (\theta, i).$$

Part I

Let X_1, \dots, X_n be an iid sample from $N(\mu, 1)$ and Y_1, \dots, Y_m be an iid sample from $N(\theta, 1)$. Assume X_1, \dots, X_n and Y_1, \dots, Y_m are independent. Consider hypothesis

$$H_0 : \mu - \theta = 1, \quad v.s. \quad H_1 : \mu - \theta \neq 1.$$

- Find the likelihood ratio test (identify a test statistic and rejection region).

Solution: The likelihood function is

$$\begin{aligned} L(\mu, \theta) &= (2\pi)^{-n/2} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n (x_i - \mu)^2 \right\} (2\pi)^{-m/2} \exp \left\{ -\frac{1}{2} \sum_{i=1}^m (y_i - \theta)^2 \right\} \\ &= (2\pi)^{-(n+m)/2} \exp \left\{ -\frac{(n-1)s_x^2 + (m-1)s_y^2}{2} - \frac{n(\bar{x} - \mu)^2 + m(\bar{y} - \theta)^2}{2} \right\} \end{aligned}$$

where \bar{x} and \bar{y} are sample averages and s_x^2 and s_y^2 are sample variance. The unrestricted MLEs of μ and θ are $\hat{\mu} = \bar{x}$ and $\hat{\theta} = \bar{y}$ respectively. Under H_0 , we have $\mu = 1 + \theta$, and the log-likelihood function (ignoring some constants) is

$$\log L(\theta) = -\frac{n}{2}(\bar{x} - \theta - 1)^2 - \frac{m}{2}(\bar{y} - \theta)^2$$

Set $\partial \log L / \partial \theta$ to zero, and solve it to obtain

$$\hat{\theta}_0 = \frac{n\bar{x} + m\bar{y} - n}{m + n}$$

The LRT statistic is

$$\lambda(\bar{x}, \bar{y}) = \frac{L(\hat{\theta}_0 + 1, \hat{\theta}_0)}{L(\bar{x}, \bar{y})} = \exp \left\{ -\frac{n}{2}(\bar{x} - \hat{\theta}_0 - 1)^2 - \frac{m}{2}(\bar{y} - \hat{\theta}_0)^2 \right\}.$$

The rejection region is $\lambda(\bar{x}, \bar{y}) < c$ for some constant c .

- Show that the rejection region in Problem 1 can be represented in terms of $|\bar{x} - \bar{y} - 1|$.

Solution: The second derivative of $\log L(\theta)$ is $-n - m < 0$. So $\hat{\theta}_0$ is indeed the MLE of θ . The LRT statistic is

$$\lambda(\bar{x}, \bar{y}) = \exp \left\{ -\frac{n}{2}(\bar{x} - \hat{\theta}_0 - 1)^2 - \frac{m}{2}(\bar{y} - \hat{\theta}_0)^2 \right\} = \exp \left(-\frac{mn}{2(m+n)}(\bar{x} - \bar{y} - 1)^2 \right).$$

The rejection region is $\lambda(\bar{x}, \bar{y}) < c$ for some constant c , or equivalently,

$$|\bar{x} - \bar{y} - 1| > k$$

where k is a constant such that under H_0 ,

$$P(|\bar{x} - \bar{y} - 1| > k) = \alpha$$

The distribution of $\bar{X} - \bar{Y}$ is normal with mean $\mu - \theta$ and variance $1/n + 1/m$. Under H_0 , the distribution of $\bar{X} - \bar{Y} - 1$ is normal with mean 0 and variance $1/n + 1/m$. So we can choose $k = z_{\alpha/2}\sqrt{1/n + 1/m}$. The rejection region is

$$\left\{x_1 \dots x_n, y_1, \dots y_m : |\bar{x} - \bar{y} - 1| > z_{\alpha/2}\sqrt{1/n + 1/m}\right\}.$$

Part II

Suppose $X_1, \dots, X_n, Y_1, \dots, Y_n$ ($n \geq 3$) are independent random variables, with X_i 's iid exponential with mean θ (> 0), and Y_i 's iid exponential with mean $\frac{1}{\theta}$. Define

$$T_1 = n^{-1} \sum_{i=1}^n X_i, \quad T_2 = n^{-1} \sum_{i=1}^n Y_i, \quad T_3 = (n-1) \left(\sum_{i=1}^n Y_i \right)^{-1}.$$

- 3.** Find the Cramer-Rao lower bound for the variance of an unbiased estimator of θ based on $X_1, \dots, X_n, Y_1, \dots, Y_n$.

Solution: Joint pdf of $X_1, \dots, X_n, Y_1, \dots, Y_n$ is

$$f(x_1, \dots, x_n, y_1, \dots, y_n) = \frac{1}{\theta^n} \exp \left\{ -\frac{\sum_{i=1}^n x_i}{\theta} \right\} \theta^n \exp \left\{ -\theta \sum_{j=1}^n y_j \right\} = \exp \left\{ -\frac{\sum_{i=1}^n x_i}{\theta} - \theta \sum_{j=1}^n y_j \right\}$$

$$\begin{aligned} \frac{d \log(f)}{d\theta} &= \frac{\sum_{i=1}^n x_i}{\theta^2} - \sum_{j=1}^n y_j, \\ \frac{d^2 \log(f)}{d\theta^2} &= -\frac{2 \sum_{i=1}^n x_i}{\theta^3}. \end{aligned}$$

Therefore,

$$I(\theta) = E_\theta \left(-\frac{d^2 \log(f)}{d\theta^2} \right) = \frac{2n\theta}{\theta^3} = \frac{2n}{\theta^2}.$$

So the Cramer-Rao lower bound (CRLB) to the variance of an unbiased estimator of θ is $\frac{\theta^2}{2n}$.

- 4.** Is there an unbiased estimator of θ for which the bound in Problem 3 is attained? Explain your answer.

Solution: An unbiased estimator $W(X, Y)$ of a parametric function $\tau(\theta)$ will attain the CRLB to the variance provided

$$a(\theta) \{W(X, Y) - \tau(\theta)\} = \frac{d}{d\theta} \log\{f(x, y|\theta)\} \tag{1}$$

for some $a(\theta)$. Thus, for $\tau(\theta) = \theta$, we need that (1) is satisfied for a suitable $a(\theta)$. This will not be true. Note that in this example $(\sum_{i=1}^n X_i, \sum_{j=1}^n Y_j)$ is minimal sufficient statistic. That is (T_1, T_3) is a minimal sufficient statistic. However,

$$\begin{aligned} E_\theta(T_1) &= \theta = E_\theta(T_3) \\ \Rightarrow E_\theta(T_1 - T_3) &= 0 \quad \text{for all } \theta \\ \Rightarrow (T_1, T_3) &\text{ is not a complete sufficient statistic.} \end{aligned}$$

5. Find $\hat{\theta}_{MLE}$, the MLE of θ .

Solution:

$$\begin{aligned} L(\theta) &= \exp \left\{ - \left(\frac{\sum_{i=1}^n x_i}{\theta} + \theta \sum_{j=1}^n y_j \right) \right\} \\ &= \exp \left\{ - \left(\frac{\sqrt{\sum_{i=1}^n x_i}}{\sqrt{\theta}} - \sqrt{\theta} \sqrt{\sum_{j=1}^n y_j} \right)^2 \right\} \exp \left(-2 \sqrt{\sum_{i=1}^n x_i} \sqrt{\sum_{j=1}^n y_j} \right) \\ &\leq \exp \left(-2 \sqrt{\sum_{i=1}^n x_i} \sqrt{\sum_{j=1}^n y_j} \right). \end{aligned}$$

The equality holds when

$$\frac{\sqrt{\sum_{i=1}^n x_i}}{\sqrt{\theta}} - \sqrt{\theta} \sqrt{\sum_{j=1}^n y_j} = 0 \Rightarrow \hat{\theta}_{MLE} = \sqrt{\frac{\sum_{i=1}^n x_i}{\sum_{j=1}^n y_j}}$$

6. Show that the MLE in Problem 5 is consistent for θ .

Solution: By WLLN, $T_1 \xrightarrow{p} \theta$, $T_2 \xrightarrow{p} \theta^{-1}$. Then

$$\hat{\theta}_{MLE} = \sqrt{\frac{\sum_{i=1}^n x_i}{\sum_{j=1}^n y_j}} \xrightarrow{p} \sqrt{\frac{\theta}{\theta^{-1}}} = \theta.$$

This proves the consistency of $\hat{\theta}_{MLE}$. Note that by CLT

$$\begin{aligned} \sqrt{n}(T_1 - \theta) &\xrightarrow{d} N(0, \theta^2), \\ \sqrt{n}(T_2 - \theta^{-1}) &\xrightarrow{d} N(0, \theta^{-2}). \end{aligned}$$

Since T_1 and T_2 are independently distributed, we get that

$$\sqrt{n} \begin{pmatrix} T_1 - \theta \\ T_2 - \theta^{-1} \end{pmatrix} \xrightarrow{d} N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \theta^2 & 0 \\ 0 & \theta^{-2} \end{pmatrix} \right).$$

Let $\Sigma = \begin{pmatrix} \theta^2 & 0 \\ 0 & \theta^{-2} \end{pmatrix}$, $g(x_1, x_2) = \sqrt{\frac{x_1}{x_2}}$, then

$$\frac{\partial g}{\partial x_1} = \frac{1}{2\sqrt{x_1 x_2}}, \quad \frac{\partial g}{\partial x_2} = -\frac{1}{2} \sqrt{\frac{x_1}{x_2^3}}.$$

Then $\nabla_g(x_1, x_2)|_{x_1=\theta, x_2=\theta^{-1}} = \frac{1}{2}(1, -\theta^2)^\top = \nabla_g(\theta)$. By the delta method

$$\sqrt{n} (\hat{\theta}_{MLE} - \theta) = \sqrt{n} \{g(T_1, T_2) - g(\theta, \theta^{-1})\} \xrightarrow{d} N(0, \nabla_g^\top(\theta) \Sigma \nabla_g(\theta)).$$

7. Find the limiting distribution of $n^{1/2}(\hat{\theta}_{MLE} - \theta)$. Does the variance of the limiting distribution attain the Cramer-Rao lower bound in Problem 3?

Solution:

$$\nabla_g^\top(\theta) \Sigma \nabla_g(\theta) = \frac{1}{4} (1, -\theta^2) \begin{pmatrix} \theta^2 & 0 \\ 0 & \theta^{-2} \end{pmatrix} \begin{pmatrix} 1 \\ -\theta^2 \end{pmatrix} = \frac{1}{4} (\theta^2 + \theta^2) = \frac{\theta^2}{2}.$$

Thus,

$$\sqrt{n} (\hat{\theta}_{MLE} - \theta) \xrightarrow{d} N\left(0, \frac{\theta^2}{2}\right) \Rightarrow V_\theta(\hat{\theta}_{MLE}) = \frac{\theta^2}{2n}.$$

By Problem 3, the asymptotic variance of the MLE of θ attains the cramer-rao bound.

8. Find a suitable α that minimizes the variance of $U = \alpha T_1 + (1 - \alpha)T_3$. Denote this by α_0 . Does the variance of the estimator $\alpha_0 T_1 + (1 - \alpha_0)T_3$ attain the Cramer-Rao lower bound in Problem 3?

Solution:

$$\begin{aligned} V_\theta(T_1) &= \frac{\theta^2}{n}, \quad V_\theta(T_3) = (n-1)^2 \left\{ E\left(\frac{\theta^2}{G^2}\right) - \frac{\theta^2}{(n-1)^2} \right\}, \quad \text{where } G \sim \text{Gamma}(n). \\ \Rightarrow V_\theta(T_3) &= (n-1)^2 \theta^2 \left\{ \frac{1}{(n-1)(n-2)} - \frac{1}{(n-1)^2} \right\} = \frac{\theta^2}{n-2}. \end{aligned}$$

Then,

$$V_\theta(U) = \frac{\alpha^2 \theta^2}{n} + (1-\alpha)^2 \frac{\theta^2}{n-2} = \theta^2 g(\alpha),$$

where

$$\begin{aligned} g(\alpha) &= \frac{\alpha^2}{n} + \frac{(1-\alpha)^2}{n-2} \\ g'(\alpha) &= \frac{2\alpha}{n} - \frac{2(1-\alpha)}{n-2} \\ g''(\alpha) &= \frac{2}{n} + \frac{2}{n-2} > 0 \\ g'(\alpha) = 0 &\Rightarrow \frac{\alpha}{1-\alpha} = \frac{n}{n-2} \Rightarrow \alpha = \frac{n}{2(n-1)} \end{aligned}$$

Thus,

$$V_\theta\{U(\alpha_0)\} = \theta^2 \left\{ \frac{\alpha_0^2}{n} + \frac{(1-\alpha_0)^2}{n-2} \right\} > \frac{\theta^2}{2n},$$

CRLB is not attained.

Part III

A teacher has her class participate in an experiment to estimate a parameter in the following way. One group of students obtains a sample of $\text{Uniform}(0, \theta)$ random variables, X_1, \dots, X_n . These values are recorded on a sheet of paper. A second team uses the data from the first team to compute the areas, X_1^2, \dots, X_n^2 . These results are recorded on a sheet of paper not the same as the sheet with the original data. A third team is then provided with the results from the first two teams to do a statistical analysis to estimate θ , but one of the sheets got lost and the third team received only a sheet of paper which has either X_1, \dots, X_n or X_1^2, \dots, X_n^2 and does not know which. Due to this complication, the teacher asks you for statistical advice on how to proceed.

You begin by setting up the problem parametrically. Let $\Theta = \{(\theta, i) : \theta > 0, i = 1, 2\}$. If $\zeta = (\theta, i) \in \Theta$, then the density having this parameter is given by

$$f_\zeta(x) = \begin{cases} \frac{1}{\theta} & 0 \leq x \leq \theta \quad \text{if } \zeta = (\theta, 1) \\ \frac{1}{2\theta\sqrt{x}} & 0 \leq x \leq \theta^2 \quad \text{if } \zeta = (\theta, 2) \end{cases}$$

Let $M_n = \max_{1 \leq i \leq n} X_i$.

9. Find $P_{(\theta,1)}(X_1 \leq x)$ and $P_{(\theta,2)}(\sqrt{X_1} \leq x)$, for any $x \in (0, \theta)$.

Solution: For $x \in (0, \theta)$

$$\begin{aligned} P_{(\theta,1)}(X_1 \leq x) &= \int_0^x f_{(\theta,1)}(x)dx = \int_0^x \frac{1}{\theta} dx = \frac{x}{\theta} \\ P_{(\theta,2)}(\sqrt{X_1} \leq x) &= P_{(\theta,2)}(0 \leq X_1 < x^2) = \int_0^{x^2} f_{(\theta,2)}(x)dx = \int_0^{x^2} \frac{1}{2\theta\sqrt{x}} dx = \frac{x}{\theta}. \end{aligned}$$

10. For any $0 < \epsilon < \theta$, show that

$$\lim_{n \rightarrow \infty} P_{(\theta,1)}(M_n < \theta - \epsilon) = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} P_{(\theta,2)}(\sqrt{M_n} < \theta - \epsilon) = 0.$$

Hence, infer that

$$\lim_{n \rightarrow \infty} P_{(\theta,1)}(|M_n - \theta| > \epsilon) = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} P_{(\theta,2)}(|\sqrt{M_n} - \theta| > \epsilon) = 0.$$

Note that the latter results say that M_n is consistent for θ under the model $P_{(\theta,1)}$ and $\sqrt{M_n}$ is consistent for θ under the model $P_{(\theta,2)}$. That is, the consistent estimator of θ changes according to which data $(X_1, \dots, X_n$ or $X_1^2, \dots, X_n^2)$ the third team receives.

Solution:

$$P_{(\theta,1)}(M_n < \theta - \epsilon) = P_{(\theta,1)}\left(\max_{i \leq i \leq n} X_i < \theta - \epsilon\right) = P_{(\theta,1)}^n(X_1 < \theta - \epsilon) = \left(\frac{\theta - \epsilon}{\theta}\right)^n.$$

Since $0 < \epsilon < \theta$, $0 < \frac{\theta - \epsilon}{\theta} < 1$. Thus,

$$\lim_{n \rightarrow \infty} P_{(\theta,1)}(M_n < \theta - \epsilon) = \lim_{n \rightarrow \infty} \left(\frac{\theta - \epsilon}{\theta}\right)^n = 0.$$

Similarly,

$$\begin{aligned} P_{(\theta,2)}(\sqrt{M_n} < \theta - \epsilon) &= P_{(\theta,2)}\left(\sqrt{\max_{i \leq i \leq n} X_i} < \theta - \epsilon\right) = P_{(\theta,2)}\left(\max_{i \leq i \leq n} \sqrt{X_i} < \theta - \epsilon\right) \\ &= P_{(\theta,2)}^n(\sqrt{X_1} < \theta - \epsilon) = \left(\frac{\theta - \epsilon}{\theta}\right)^n. \end{aligned}$$

Thus,

$$\lim_{n \rightarrow \infty} P_{(\theta,2)}(\sqrt{M_n} < \theta - \epsilon) = \lim_{n \rightarrow \infty} \left(\frac{\theta - \epsilon}{\theta}\right)^n = 0.$$

Furthermore, under $P_{(\theta,1)}$

$$\{|M_n - \theta| > \epsilon\} = \{M_n - \theta > \epsilon\} \cup \{M_n - \theta < -\epsilon\} = \{M_n > \theta + \epsilon\} \cup \{M_n < \theta - \epsilon\}.$$

Since for $1 \leq i \leq n$, $0 \leq X_i \leq \theta$, $0 \leq M_n \leq \theta$, thus, $\{M_n > \theta + \epsilon\} = \emptyset$. So

$$\lim_{n \rightarrow \infty} P_{(\theta,1)}(|M_n - \theta| > \epsilon) = \lim_{n \rightarrow \infty} P_{(\theta,1)}(M_n < \theta - \epsilon) = 0.$$

Similarly for the condition under $P_{(\theta,2)}$. For every $1 \leq i \leq n$, $0 \leq X_i \leq \theta^2$, $0 \leq \sqrt{M_n} \leq \theta$. So

$$\begin{aligned} \{|\sqrt{M_n} - \theta| > \epsilon\} &= \{\sqrt{M_n} - \theta > \epsilon\} \cup \{\sqrt{M_n} - \theta < -\epsilon\} \\ &= \{\sqrt{M_n} > \theta + \epsilon\} \cup \{\sqrt{M_n} < \theta - \epsilon\} = \{\sqrt{M_n} < \theta - \epsilon\}. \end{aligned}$$

Hence,

$$\lim_{n \rightarrow \infty} P_{(\theta,2)}(|\sqrt{M_n} - \theta| > \epsilon) = \lim_{n \rightarrow \infty} P_{(\theta,2)}(\sqrt{M_n} < \theta - \epsilon) = 0.$$

11. Suppose we define a random variable

$$Y_n = \frac{\left(\frac{1}{n} \sum_{i=1}^n X_i\right)^2}{\frac{1}{n} \sum_{i=1}^n X_i^2}.$$

Then, show using the Law of Large Numbers that Y_n converges in probability to

$$\begin{cases} \frac{3}{4} & \text{if } \zeta = (\theta, 1), \text{ or} \\ \frac{5}{9} & \text{if } \zeta = (\theta, 2) \end{cases}.$$

Solution: Under $P_{(\theta,1)}$, based on the information provided, we find that

$$E(X_i) = \frac{\theta}{2}, \quad Var(X_i) = \frac{\theta^2}{12}, \quad E(X_i^2) = Var(X_i) + \{E(X_i)\}^2 = \frac{\theta^2}{3}.$$

Since $X_i \sim U(0, \theta)$, using the Law of Large Numbers,

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n X_i &\rightarrow E(X_i) = \frac{\theta}{2}, \quad \frac{1}{n} \sum_{i=1}^n X_i^2 \rightarrow E(X_i^2) = \frac{\theta^2}{3} \quad \text{in probability} \\ \Rightarrow Y_n &= \frac{\left(\frac{1}{n} \sum_{i=1}^n X_i\right)^2}{\frac{1}{n} \sum_{i=1}^n X_i^2} \rightarrow \frac{\left(\frac{\theta}{2}\right)^2}{\frac{\theta^2}{3}} = \frac{3}{4} \quad \text{in probability}. \end{aligned}$$

Under $P_{(\theta,2)}$, based on the information provided, we find that

$$\begin{aligned} E(X_i) &= \int_0^{\theta^2} x \frac{1}{2\theta\sqrt{x}} = \frac{1}{2\theta} \int_0^{\theta^2} x^{\frac{1}{2}} dx = \frac{\theta^2}{3} \\ E(X_i^2) &= \int_0^{\theta^2} x^2 \frac{1}{2\theta\sqrt{x}} = \frac{1}{2\theta} \int_0^{\theta^2} x^{\frac{3}{2}} dx = \frac{\theta^4}{5}. \end{aligned}$$

Since $X_i \sim U(0, \theta)$, using the Law of Large Numbers,

$$\Rightarrow Y_n = \frac{\left(\frac{1}{n} \sum_{i=1}^n X_i\right)^2}{\frac{1}{n} \sum_{i=1}^n X_i^2} \rightarrow \frac{\left(\frac{\theta^2}{3}\right)^2}{\frac{\theta^4}{5}} = \frac{5}{9} \quad \text{in probability.}$$

So, in summary, we obtain that Y_n converges in probability to

$$\begin{cases} \frac{3}{4} & \text{if } \zeta = (\theta, 1), \text{ or} \\ \frac{5}{9} & \text{if } \zeta = (\theta, 2) \end{cases}.$$

- 12.** Suppose the third team wants to construct one consistent estimator of θ regardless of the data $(X_1, \dots, X_n$ or $X_1^2, \dots, X_n^2)$ the team receives. Then, we may define

$$\hat{\theta}_n = \begin{cases} M_n & \text{if } Y_n > \frac{3}{5} \\ \sqrt{M_n} & \text{if } Y_n < \frac{3}{5} \end{cases}$$

Note that $\frac{3}{4} > \frac{3}{5} > \frac{5}{9}$. Now use Problems 10, 11 and the fact that $\frac{3}{4} > \frac{3}{5} > \frac{5}{9}$, and give a heuristic reasoning (or a mathematical reasoning) as to why

$$\lim_{n \rightarrow \infty} P_\zeta(|\hat{\theta}_n - \theta| > \epsilon) = 0 \quad \text{for all } \zeta = (\theta, i).$$

Solution: If $\zeta = (\theta, 1)$,

$$\begin{aligned} 0 &\leq \lim_{n \rightarrow \infty} P_\zeta(|\theta - \hat{\theta}_n| > \epsilon) \\ &= \lim_{n \rightarrow \infty} P_\zeta\left(|\theta - \hat{\theta}_n| > \epsilon, Y_n > \frac{3}{5}\right) + \lim_{n \rightarrow \infty} P_\zeta\left(|\theta - \hat{\theta}_n| > \epsilon, Y_n < \frac{3}{5}\right) \\ &= \lim_{n \rightarrow \infty} P_{(\theta,1)}\left(|M_n - \theta| > \epsilon, Y_n > \frac{3}{5}\right) + \lim_{n \rightarrow \infty} P_{(\theta,1)}\left(|M_n - \theta| > \epsilon, Y_n < \frac{3}{5}\right) \\ &\leq \lim_{n \rightarrow \infty} P_{(\theta,1)}(|M_n - \theta| > \epsilon) + \lim_{n \rightarrow \infty} P_{(\theta,1)}\left(Y_n - \frac{3}{4} < \frac{3}{5} - \frac{3}{4}\right) \\ &\leq 0 + \lim_{n \rightarrow \infty} P_{(\theta,1)}\left(\left|Y_n - \frac{3}{4}\right| > \frac{3}{20}\right) = 0 \end{aligned}$$

So,

$$\lim_{n \rightarrow \infty} P_{(\theta,1)}\left(|\theta - \hat{\theta}_n| > \epsilon\right) = 0.$$

if $\zeta = (\theta, 2)$,

$$\begin{aligned}
 0 &\leq \lim_{n \rightarrow \infty} P_\zeta \left(|\theta - \hat{\theta}_n| > \epsilon \right) \\
 &= \lim_{n \rightarrow \infty} P_\zeta \left(|\theta - \hat{\theta}_n| > \epsilon, Y_n > \frac{3}{5} \right) + \lim_{n \rightarrow \infty} P_\zeta \left(|\theta - \hat{\theta}_n| > \epsilon, Y_n < \frac{3}{5} \right) \\
 &\leq \lim_{n \rightarrow \infty} P_{(\theta, 2)} \left(|M_n - \theta| > \epsilon \right) + \lim_{n \rightarrow \infty} P_{(\theta, 2)} \left(Y_n - \frac{5}{9} < \frac{3}{5} - \frac{5}{9} \right) \\
 &\leq 0 + \lim_{n \rightarrow \infty} P_{(\theta, 2)} \left(\left| Y_n - \frac{5}{9} \right| > \frac{2}{45} \right) = 0
 \end{aligned}$$

So,

$$\lim_{n \rightarrow \infty} P_{(\theta, 2)} \left(|\theta - \hat{\theta}_n| > \epsilon \right) = 0.$$

In summary, $\lim_{n \rightarrow \infty} P_\zeta \left(|\theta - \hat{\theta}_n| > \epsilon \right) = 0$ for $\zeta = (\theta, 1)$ and $(\theta, 2)$.