Part I

1. Let X and Y be real-valued random variables with finite second moments on the same probability space. Suppose $f(y) = \mathbf{E}[X|Y = y]$ is decreasing in y. Show that $\mathbf{Cov}(X, Y) \leq 0$. *Hint:* A decreasing function f has the property that $(y - z)(f(y) - f(z)) \leq 0$, for all $y, z \in \mathbb{R}$.

Part II

Let $X_1, \dots, X_n$ be an iid sample from the distribution on $\mathbb{R}$ with parameter $\theta > 0$ and pdf

$$f(x;\theta) = \begin{cases} \theta^2 x e^{-\theta x}, & \text{if } x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

2. Show that $\sum_{i=1}^n X_i$ is a sufficient statistic for θ.

3. Find the UMVUE for θ for $n \geq 2$.

Part III

Let $X_1, \dots, X_n$ be an iid sample from a (beta) distribution on $\mathbb{R}$ with parameters $\theta > 0$ and 1 and pdf

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

4. Show that $-2\theta \sum_{i=1}^n \log(X_i)$ is a pivotal quantity, i.e., its probability distribution does not depend on the unknown parameter.

5. Find a $100(1-\alpha)\%$ confidence interval for θ based on the pivotal quantity $-2\theta \sum_{i=1}^n \log(X_i)$.

Part IV

6. Let $X_1, \dots, X_n$ be an iid real-valued random variable with finite fourth moment.

(a) Find the limiting distribution of $\sqrt{n}(S_n^2 - \sigma^2)$, where

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i \text{ and } S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2.$$

(b) Find the limiting distribution of $\sqrt{n}(S_n - \sigma)$.

Part I

1. By the law of iterated expectations,

$$\mathbf{E}[XY] = \mathbf{E}\big[\mathbf{E}[XY|Y]\big] = \mathbf{E}\big[Y\,\mathbf{E}[X|Y]\big] = \mathbf{E}[Yf(Y)],$$

and similarly $\mathbf{E}[X] = \mathbf{E}\big[\mathbf{E}[X|Y]\big] = \mathbf{E}[f(Y)]$. Hence,

$$\mathbf{Cov}[X,Y] = \mathbf{E}[XY] - \mathbf{E}[X]\,\mathbf{E}[Y] = \mathbf{E}[Yf(Y)] - \mathbf{E}[f(Y)]\,\mathbf{E}[Y] = \mathbf{Cov}[Y, f(Y)].$$

Consequently, it is sufficient to show that $\mathbf{Cov}[Y, f(Y)] \leq 0$. Now, let Z be independently and identically distributed as Y. Since f is non-increasing, for each $y, z \in \mathbb{R}$, we have $(y-z)(f(y)-f(z)) \leq 0$. So, $\mathbf{E}[(Y-Z)(f(Y)-f(Z))] \leq 0$. Then by the independence of Y and Z,

$$\mathbf{E}[Yf(Y)] + \mathbf{E}[Zf(Z)] - \mathbf{E}[Y]\,\mathbf{E}[f(Z)] - \mathbf{E}[Z]\,\mathbf{E}[f(Y)] \leq 0.$$

Since Y is identically distributed as Z, $2\,\mathbf{E}[Yf(Y)] - 2\,\mathbf{E}[Y]\,\mathbf{E}[f(Y)] \leq 0$. Finally, it follows that $\mathbf{Cov}[Y, f(Y)] = \mathbf{E}[Yf(Y)] - \mathbf{E}[Y]\,\mathbf{E}[f(Y)] \leq 0$.

Part II

2. We have that

$$\begin{aligned}
\prod_{i=1}^{n} f(x_i, \theta) &= \prod_{i=1}^{n} \theta^2 x_i e^{-\theta x_i} \\
&= \theta^{2n} e^{-\theta \sum_{i=1}^{n} x_i} \prod_{i=1}^{n} x_i \\
&= g\Big(\sum_{i=1}^{n} x_i; \theta\Big) \cdot h(x_1, \ldots, x_n)
\end{aligned}$$

with $g\big(\sum_{i=1}^{n} x_i; \theta\big) = \theta^{2n} e^{-\theta \sum_{i=1}^{n} x_i}$ a nonnegative function (depending on $x_1, \ldots, x_n$ and θ) and $h(x_1, \ldots, x_n) = \prod_{i=1}^{n} x_i$ also a nonnegative function not depending on θ. Then according to the factorization theorem of Fisher and Neyman, we have that $\sum_{i=1}^{n} X_i$ is a sufficient statistic for θ.

3.

$$
\begin{aligned}
\mathbf{E}\left(\frac{1}{\sum_{i=1}^n X_i}\right) &= \int_0^\infty \frac{1}{x}\frac{\theta^{2n}}{\Gamma(2n)}x^{2n-1}e^{-\theta x}\,dx \\
&= \frac{\theta^{2n}}{\Gamma(2n)}\int_0^\infty x^{2n-2}e^{-\theta x}\,dx \\
&= \frac{\theta^{2n-1}}{\Gamma(2n)}\int_0^\infty \frac{t^{2n-2}}{\theta^{2n-2}}e^{-t}\,dt \\
&= \frac{\theta}{\Gamma(2n)}\Gamma(2n-1) = \frac{\theta}{2n-1}
\end{aligned}
$$

and hence $\dfrac{2n-1}{\sum_{i=1}^n X_i}$ is an unbiased estimator for θ. Also,

$$
\mathbf{E}\left\{\left(\frac{1}{\sum_{i=1}^n X_i}\right)^2\right\} = \int_0^\infty \frac{1}{x^2}\frac{\theta^{2n}}{\Gamma(2n)}x^{2n-1}e^{-\theta x}\,dx = \frac{\theta^2}{\Gamma(2n)}\Gamma(2n-2) \text{ if } n \geq 2,
$$

i.e., the variance is finite. According to the Lehmann-Scheffé theorem, $\dfrac{2n-1}{\sum_{i=1}^n X_i}$ is the UMVUE of θ as it is a function of a complete, sufficient statistic $\sum_{i=1}^n X_i$.

Part III

4. Let $Y = -\log(X)$. Then

$$
F_Y(y) = \mathbf{P}(Y \leq y) = \begin{cases} 1 - \mathbf{P}(X \leq e^{-y}), & \text{if } y \geq 0, \\ 0, & \text{if } y < 0. \end{cases}
$$

and

$$
f_Y(y) = \begin{cases} f_X(e^{-y})e^{-y} = \theta e^{-y(\theta-1)}e^{-y} = \theta e^{-\theta y}, & \text{if } y \geq 0, \\ 0, & \text{if } y < 0. \end{cases}
$$

Hence $Y = -\log(X) \sim \text{Exp}(\theta)$. The distribution of $-2\theta\sum_{i=1}^n \log(X_i) = 2\theta\sum_{i=1}^n Y_i$ can be found as follows. Consider the characteristic function of $2\theta\sum_{i=1}^n Y_i$:

$$
\mathbf{E}\left\{e^{it(2\theta\sum_{i=1}^n Y_i)}\right\} \overset{i.i.d.}{=} \{\mathbf{E}\left(e^{2it\theta Y}\right)\}^n = (\varphi_Y(2t\theta))^n = (1-2it)^{-n},
$$

where $\varphi_Y(t) = (1 - it\theta^{-1})^{-1}$. By the continuity theorem, it follows that $-2\theta\sum_{i=1}^n \log(X_i) = 2\theta\sum_{i=1}^n Y_i \sim \chi^2(2n)$, i.e., $-2\theta\sum_{i=1}^n \log(X_i)$ is a pivotal quantity.

5. Since $-2\theta\sum_{i=1}^n \log(X_i) = 2\theta\sum_{i=1}^n Y_i \sim \chi^2(2n)$, we have

$$\begin{aligned}1-\alpha &= \mathbf{P}\left(\chi^2_{2n;\frac{\alpha}{2}} \le -2\theta\sum_{i=1}^n \log(X_i) \le \chi^2_{2n;1-\frac{\alpha}{2}}\right)\\ &= \mathbf{P}\left(\frac{\chi^2_{2n;\frac{\alpha}{2}}}{-2\sum_{i=1}^n \log(X_i)} \le \theta \le \frac{\chi^2_{2n;1-\frac{\alpha}{2}}}{-2\sum_{i=1}^n \log(X_i)}\right).\end{aligned}$$

A $100(1-\alpha)\%$ confidence interval for θ is

$$\left[\frac{\chi^2_{2n;\frac{\alpha}{2}}}{-2\sum_{i=1}^n \log(X_i)}, \frac{\chi^2_{2n;1-\frac{\alpha}{2}}}{-2\sum_{i=1}^n \log(X_i)}\right].$$

Part IV

6. (a) We start from

$$\begin{aligned}\sqrt{n}(S_n^2-\sigma^2) &= \sqrt{n}\left(\frac{1}{n-1}\sum_{i=1}^n (X_i-\overline{X}_n)^2-\sigma^2\right)\\ &= \sqrt{n}\left(\frac{n}{n-1}\frac{1}{n}\sum_{i=1}^n \{(X_i-\mu)-(\overline{X}_n-\mu)\}^2-\sigma^2\right)\\ &= \frac{n\sqrt{n}}{n-1}\left(\frac{1}{n}\sum_{i=1}^n (X_i-\mu)^2-\frac{2}{n}(\overline{X}_n-\mu)\sum_{i=1}^n (X_i-\mu)+(\overline{X}_n-\mu)^2-\frac{n-1}{n}\sigma^2\right)\\ &= \frac{n\sqrt{n}}{n-1}\left(\frac{1}{n}\sum_{i=1}^n \{(X_i-\mu)^2-\sigma^2\}-(\overline{X}_n-\mu)^2+\frac{1}{n}\sigma^2\right)\\ &= \frac{n}{n-1}\frac{1}{\sqrt{n}}\sum_{i=1}^n \{(X_i-\mu)^2-\sigma^2\}-\frac{n\sqrt{n}}{n-1}(\overline{X}_n-\mu)^2+\frac{\sqrt{n}}{n-1}\sigma^2.\end{aligned}$$

Analyzing each term individually yields

1. $\frac{1}{\sqrt{n}}\sum_{i=1}^n\{(X_i-\mu)^2-\sigma^2\} = \frac{1}{\sqrt{n}}\sum_{i=1}^n Y_i$ with $Y_1,\dots,Y_n$ iid random variables. Also

$$\begin{aligned}\mathbf{E}\,Y_i &= \mathbf{E}(X_i-\mu)^2-\sigma^2 = \sigma^2-\sigma^2 = 0,\\ \mathbf{Var}\,Y_i &= \mathbf{Var}(X_i-\mu)^2 = \mathbf{E}(X_i-\mu)^4-[\mathbf{E}(X_i-\mu)^2]^2 = \mathbf{E}(X_i-\mu)^4-\sigma^4 = \tau^2 > 0.\end{aligned}$$

The latter is finite if $\mathbf{E}\,X^4 < \infty$. Then, according to the central limit theorem we have

$$\frac{\sum_{i=1}^n\{(X_i-\mu)^2-\sigma^2\}}{\sqrt{n\tau^2}} = \frac{\sum_{i=1}^n\{(X_i-\mu)^2-\sigma^2\}}{\tau\sqrt{n}} \xrightarrow{d} N(0,1),$$

or

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n\{(X_i-\mu)^2-\sigma^2\} \xrightarrow{d} N(0,\tau^2).$$

Finally, since $n/(n-1) \to 1$ and by Slutsky's theorem, we have

$$\frac{n}{n-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{(X_i-\mu)^2-\sigma^2\} \xrightarrow{d} N(0,\tau^2)$$

2. Again by Slutsky's theorem,

$$\frac{n\sqrt{n}}{n-1}(\overline{X}_n-\mu)^2 = \underbrace{\frac{n}{n-1}}_{\to 1}\underbrace{\sqrt{n}(\overline{X}_n-\mu)}_{\xrightarrow{d} N(0,\sigma^2)}\underbrace{(\overline{X}_n-\mu)}_{\xrightarrow{p} 0} \xrightarrow{d} 0,$$

and hence

$$\frac{n\sqrt{n}}{n-1}(\overline{X}_n-\mu)^2 \xrightarrow{p} 0.$$

3. For $\sigma^2 < \infty$, $\frac{\sqrt{n}}{n-1}\sigma^2 \to 0$ as $n \to \infty$.

Combining the above results and applying Slutsky's theorem yields the final result

$$\sqrt{n}(S_n^2-\sigma^2) \xrightarrow{d} N(0,\tau^2)$$

with $\tau^2 = \mathbf{E}(X-\mu)^4 - \sigma^4$ and $\mathbf{E}\,X^4 < \infty$.

(b) From Question 6a we know that

$$\sqrt{n}(S_n^2-\sigma^2) \xrightarrow{d} N(0,\tau^2)$$

with $\tau^2 = \mathbf{E}(X-\mu)^4 - \sigma^4$ and $\mathbf{E}\,X^4 < \infty$.
Applying the first-order delta method with $g(x) = \sqrt{x}$ yields

$$\sqrt{n}(S_n-\sigma) = \sqrt{n}(\sqrt{S_n^2}-\sqrt{\sigma^2}) \xrightarrow{d} N\left(0, \frac{\tau^2}{4\sigma^2}\right).$$

A random variable X has the exponential distribution, $X \sim E(a, \theta)$, if X has the pdf

$$f_X(x|a,\theta) = \theta^{-1} \exp[-(x-a)/\theta] \, \mathbb{I}[x \geq a],$$

where $\theta > 0$, $a \in \mathbb{R}$, and $\mathbb{I}[\cdot]$ denotes the indicator function (e.g., $\mathbb{I}[A] = 1$ if event A holds and 0 otherwise).
A random variable Y has the gamma distribution, $Y \sim \Gamma(\alpha, \beta)$, if Y has the pdf

$$f_Y(y|\alpha,\beta) = \frac{y^{\alpha-1} \, \exp[-\frac{y}{\beta}]}{\Gamma(\alpha)\beta^\alpha} \mathbb{I}[y > 0]$$

with $\alpha, \beta \in (0, \infty)$. You may use the fact that $\mathbb{E}(Y) = \alpha\beta$ and $\text{Var}(Y) = \alpha\beta^2$.

Part I

Assume $X_1, \cdots, X_n$ are iid $E(a, \theta)$. Let $\underset{\sim}{X} = (X_1, \cdots, X_n)$ be a random sample with $E(a, \theta)$.

You may use the following facts without proof:

- If $X \sim E(a, \theta)$, then $\mathbb{E}(X) = a + \theta$ and $\text{Var}(X) = \theta^2$.
- Define $Z_i = X_{(i)} - X_{(i-1)}$, where $X_{(j)}$ is the jth order statistic, and $X_{(0)} = 0$. Then $Z_1, \cdots, Z_n$ are independent and $2(n-i+1)Z_i/\theta$ has the chi-squared distribution with 2 degrees of freedom for $i = 1, \cdots, n$.
- Define $S(\underset{\sim}{X}) = \big(X_{(1)}, \sum_{i=1}^n (X_i - X_{(1)})\big)$. Then $S(\underset{\sim}{X})$ is jointly complete for (a, θ).

1. Show that $S(\underset{\sim}{X})$ is jointly sufficient for (a, θ).

2. Find the MLE of (a, θ) based on $\underset{\sim}{X}$, say $(\hat{a}_n, \hat{\theta}_n)$.

3. Show that $X_{(1)}$ and $\sum_{i=1}^n (X_i - X_{(1)})$ are independent.

4. Identify the distributions of $X_{(1)}$ and $\sum_{i=1}^n (X_i - X_{(1)})$.

5. Find the UMVUE's of a and θ, say $\tilde{a}_n$ and $\tilde{\theta}_n$.

6. Prove or disprove: $\tilde{\theta}_n$ is asymptotically more efficient than $\hat{\theta}_n$.

7. Determine the limiting distribution of $\sqrt{n}(\tilde{\theta}_n - \theta)$ as $n \to \infty$.

8. Find a variance-stabilizing transformation for $\tilde{\theta}_n$ and use this to determine a large-sample confidence interval for θ with approximate confidence coefficient $1 - \alpha$.

9. Find an LRT of size α for testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$, where θ_0 is a known constant. Recall that a is an unknown parameter.

Part II

For **Problems 10-13**, assume $X_1, \cdots, X_n$ are iid $E(0, \theta)$ and the prior distribution of θ^{-1} is $\Gamma(\alpha, \beta)$.

10. Find the posterior distribution of θ^{-1} and identify its parameters.

11. What is the Bayes estimator for θ^{-1} under the squared error loss function?

12. Show that for fixed $\theta_0^{-1} > 0$,

$$\phi(\underset{\sim}{X}) = \begin{cases} 1 & \text{if } \sum_{i=1}^n X_i > \frac{\theta_0}{2}\chi^2_{2(n+\alpha),0.5} - \frac{1}{\beta}, \\ 0 & \text{otherwise} \end{cases}$$

is the 0-1 loss Bayes test for $H_0 : \theta^{-1} \geq \theta_0^{-1}$ versus $H_1 : \theta^{-1} < \theta_0^{-1}$.

Here, $\chi^2_{2(n+\alpha),0.5}$ is a point such that $\mathbb{P}(\chi^2_{2(n+\alpha)} \leq \chi^2_{2(n+\alpha),0.5}) = \alpha$.

13. Find a $(1 - \alpha)$ Bayes credible set for θ^{-1}.

1. Let $\underset{\sim}{X} = (X_1, \cdots, X_n)$ be a random sample with the common pdf

$$f_X(x|a,\theta) = \theta^{-1} \exp[-(x-a)/\theta]\, \mathbb{I}[x \geq a],$$

where $\theta > 0$, $a \in \mathbb{R}$, and $\mathbb{I}[\cdot]$ denotes the indicator function (e.g., $\mathbb{I}[A] = 1$ if event A holds and 0 otherwise). The joint pdf of $\underset{\sim}{X}$ is

$$f_{\underset{\sim}{X}}(\underset{\sim}{x}|a,\theta) = \theta^{-n} \exp[-\sum_{i=1}^{n}(x_i - a)/\theta]\, \mathbb{I}[x_{(1)} > a].$$

By the factorization theorem, we see that $S^*(\underset{\sim}{X}) = \left(X_{(1)}, \sum_{i=1}^n X_i\right)$ is jointly sufficient for (a, θ). Now, consider a one-to-one function $h : \mathbb{R}^2 \to \mathbb{R}^2$ where $h(x, y) = (x, y+nx)$. Then, $S^*(\underset{\sim}{X}) = h(S(\underset{\sim}{X}))$ where $S(\underset{\sim}{X}) = \left(X_{(1)}, \sum_{i=1}^n X_i - nX_{(1)}\right)$. Since any one-to-one function of a sufficient statistics is sufficient we conclude that $S(\underset{\sim}{X}) = \left(X_{(1)}, \sum_{i=1}^n X_i - nX_{(1)}\right)$ is jointly sufficient for (a, θ).

2. The likelihood function

$$L(a,\theta|X_1,\ldots,X_n) = \theta^{-n} \exp[-\sum_{i=1}^{n}(X_i - a)/\theta]\, \mathbb{I}[X_{(1)} \geq a]$$

is zero if $a > X_{(1)}$ and increasing on $(0, X_{(1)}]$. Hence the MLE of a, $\hat{a}_n$, is $X_{(1)}$. Now, substitute $a = \hat{a}_n = X_{(1)}$ into $L(a, \theta|X_1, \ldots, X_n)$ and maximize the resulting likelihood function to see that the MLE of θ, $\hat{\theta}_n$, is $\frac{1}{n}\sum_{i=1}^n \left(X_i - X_{(1)}\right)$. More specifically, we have

$$L(X_{(1)},\theta|X_1,\ldots,X_n) = \theta^{-n} \exp\left[-\frac{\sum_{i=1}^n \left(X_i - X_{(1)}\right)}{\theta}\right]$$

and hence the log-likelihood becomes

$$\log L(X_{(1)},\theta|X_1,\ldots,X_n) = -n\log(\theta) - \frac{1}{\theta}\sum_{i=1}^{n}\left(X_i - X_{(1)}\right).$$

Taking the derivative of the log-likelihood with respect to θ and solving for $\theta = \hat{\theta}_n$, we have:

$$\frac{\partial}{\partial\theta}\log L(X_{(1)},\theta|X_1,\ldots,X_n) = \frac{-n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n}\left(X_i - X_{(1)}\right)|_{\theta=\hat{\theta}_n} = 0$$

and consequently we get $\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^n \left(X_i - X_{(1)}\right)$. Therefore, $\hat{\theta}_n$ maximizes the log-likelihood if we can show that

$$\frac{\partial^2}{\partial\theta^2}\log L(X_{(1)},\theta|X_1,\ldots,X_n)|_{\theta=\hat{\theta}_n} < 0.$$

Note that $\frac{\partial^2}{\partial \theta^2} \log L(X_{(1)}, \theta | X_1, \ldots, X_n)|_{\theta=\hat{\theta}_n} = -\frac{n}{\hat{\theta}_n^2} < 0$ and this completes the proof.

3. First, we show that $X_{(1)}$ is complete and sufficient for a when θ is fixed. The sufficiency comes from the factorization theorem. To see $X_{(1)}$ is complete for a, let $\mathbb{E}_{(a,\theta)}\big[g(X_{(1)})\big] = 0$ for any value of a and fixed θ. This means

$$\frac{n}{\theta} \int_a^\infty g(x) \exp\Big[-\frac{n(x-a)}{\theta}\Big] dx = 0$$

for any value of a and fixed θ. Therefore $\int_a^\infty g(x) \exp[-\frac{nx}{\theta}] dx = 0$, which implies $g(x) = 0$ almost everywhere. Hence for fixed θ, $X_{(1)}$ is complete and sufficient for a. Next, note that for any $i = 1, \cdots, n$, $X_i - a \sim E(0, \theta)$. Therefore, for any fixed θ, $\sum_{i=1}^n (X_i - X_{(1)}) = \sum_{i=1}^n (X_i - a) - (X_{(1)} - a)$ is ancillary. Now, Basu's theorem implies that $X_{(1)}$ and $\sum_{i=1}^n (X_i - X_{(1)})$ are independent for any fixed θ. Since θ is arbitrary, $X_{(1)}$ and $\sum_{i=1}^n (X_i - X_{(1)})$ are independent for any (a, θ).

4. By the pdf of $X_{(1)}$ we can see that $X_{(1)} \sim E(a, \frac{\theta}{n})$. To find the distribution of $\sum_{i=1}^n (X_i - X_{(1)})$ we write

$$\sum_{i=1}^n (X_i - X_{(1)}) = \sum_{i=2}^n (X_{(i)} - X_{(1)}) = \sum_{i=2}^n (n - i + 1) Z_i,$$

where $Z_i = X_{(i)} - X_{(i-1)}$ for $i = 1, \cdots, n$ and $X_{(0)} = 0$. Based on the fact provided in the exam, $\frac{2(n-i+1)Z_i}{\theta} \sim \chi_2^2$. Therefore, $\sum_{i=1}^n (X_i - X_{(1)}) \sim \frac{\theta \chi_{2(n-1)}^2}{2}$.

5. From Problem **1** and the fact given in the exam, we know $S(\underset{\sim}{X}) = \big(X_{(1)}, \sum_{i=1}^n (X_i - X_{(1)})\big)$ is jointly complete and sufficient for (a, θ). Let $\tilde{\theta}_n = \frac{1}{n-1} \sum_{i=1}^n (X_i - X_{(1)})$. From Problem **4**, we know that $\sum_{i=1}^n (X_i - X_{(1)}) \sim \frac{\theta \chi_{2(n-1)}^2}{2}$ and hence $\mathbb{E}\big[\tilde{\theta}_n\big] = \theta$, which means $\tilde{\theta}_n$ is an unbiased estimator for θ. Also, $\tilde{\theta}_n$ is a function of the complete and sufficient statistic $S(\underset{\sim}{X})$. Therefore, the Lehmann-Scheffe (L-S) theorem implies that $\tilde{\theta}_n$ is the UMVUE for θ. For parameter a, let $\tilde{a}_n = X_{(1)} - \frac{1}{n(n-1)} \sum_{i=1}^n (X_i - X_{(1)})$. Then $\tilde{a}_n$ is an unbiased estimator for a and also a function of the complete and sufficient statistic $S(\underset{\sim}{X})$. Again, the L-S theorem implies that $\tilde{a}_n$ is the UMVUE for a.

6. From Problem **2**, we know that $\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n \big(X_i - X_{(1)}\big)$ and hence $\hat{\theta}_n \sim \frac{\theta}{2n} \chi_{2n}^2$. From Problem **5**, we see that $\tilde{\theta}_n = \frac{1}{n-1} \sum_{i=1}^n \big(X_i - X_{(1)}\big) \sim \frac{\theta}{2(n-1)} \chi_{2n}^2$. Now, by the definition of asymptotic relative efficiency, we have $\mathrm{ARE}(\hat{\theta}_n, \tilde{\theta}_n, \theta) = \lim_{n\to\infty} \frac{\mathrm{Var}(\tilde{\theta}_n)}{\mathrm{Var}(\hat{\theta}_n)} = 1$. Therefore, both $\hat{\theta}_n$ and $\tilde{\theta}_n$ have the same asymptotic efficiency.

7. From Problem **5**, we see that $\frac{2(n-1)\tilde{\theta}_n}{\theta} \sim \chi^2_{2(n-1)}$. Therefore, the central limit theorem implies that $\sqrt{n}(\tilde{\theta}_n - \theta) \xrightarrow{d} N(0, 2\theta^2)$ as $n \to \infty$.

8. From Problem **7**, we know $\sqrt{n}(\tilde{\theta}_n - \theta) \xrightarrow{d} N(0, 2\theta^2)$ as $n \to \infty$. A function $g : \mathbb{R} \to \mathbb{R}$ is called a variance-stabilizing transformation for $\tilde{\theta}_n$ if $g'(\theta)\sigma(\theta) = 1$ where $\sigma^2(\theta) = 2\theta^2$. A simple calculus shows that $g(\theta) = \frac{1}{\sqrt{2}} \log\theta$. Therefore, $Q_n = \sqrt{\frac{n}{2}}(\log(\tilde{\theta}_n) - \log(\theta)) \xrightarrow{d} N(0, 1)$ as $n \to \infty$ and this shows that Q_n is asymptotically pivotal. Now, a large-sample confidence interval for θ with approximate confidence coefficient $1 - \alpha$ is given by

$$C_{\underset{\sim}{X}} = \{(a, \theta) : z_{\alpha/2} \le \sqrt{\frac{n}{2}}(\log(\tilde{\theta}_n) - \log(\theta)) \le z_{1-\alpha/2}\}$$

which is the same as

$$\Big[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - X_{(1)}) \exp\big(-\sqrt{\frac{n}{2}}z_{1-\alpha/2}\big), \frac{1}{n-1}\sum_{i=1}^{n}(X_i - X_{(1)}) \exp\big(-\sqrt{\frac{n}{2}}z_{\alpha/2}\big)\Big].$$

9. Let $\underset{\sim}{\eta} = (a, \theta)$. Under the null hypothesis H_0, $(X_{(1)}, \theta_0)$ is the MLE of $\underset{\sim}{\eta}$. Therefore, the likelihood ratio $\lambda(\underset{\sim}{X})$ is

$$\begin{aligned}
\lambda(\underset{\sim}{X}) &= \Big(\frac{\hat{\theta}_n}{\theta_0}\Big)^n \exp\left[\sum_{i=1}^{n}(X_i - X_{(1)})\Big(\frac{1}{\hat{\theta}_n} - \frac{1}{\theta_0}\Big)\right] \\
&= \Big(\frac{\hat{\theta}_n}{\theta_0}\Big)^n \exp(n) \exp\Big[-n\Big(\frac{\hat{\theta}_n}{\theta_0}\Big)\Big] \\
&= h\Big(\frac{\hat{\theta}_n}{\theta_0}\Big),
\end{aligned}$$

where $h(t) = t^n \exp(n)\exp(-nt)$ for $t = \frac{\hat{\theta}_n}{\theta_0}$. For $k \in (0, 1)$, we can see that

$$\begin{aligned}
\lambda(\underset{\sim}{X}) > k &\iff \frac{\hat{\theta}_n}{\theta_0} \in (k_1, k_2) \\
\lambda(\underset{\sim}{X}) = k &\iff \frac{\hat{\theta}_n}{\theta_0} \in \{k_1, k_2\} \\
\lambda(\underset{\sim}{X}) < k &\iff \frac{\hat{\theta}_n}{\theta_0} \notin [k_1, k_2],
\end{aligned}$$

where $k_1 < k_2$ and $h(k_1) = h(k_2) = k$. Now, an LRT $\phi(\underset{\sim}{X})$ has the form

$$\phi(\underset{\sim}{X}) = \begin{cases} 1, & \text{if } \frac{\hat{\theta}_n}{\theta_0} \notin [k_1, k_2], \\ 0, & \text{otherwise} \end{cases}$$

with $\mathbb{E}_{(X_{(1)},\theta_0)}\phi(\underset{\sim}{X}) = \alpha$. Next, we need to find k_1 and k_2. Note that under H_0, $\frac{\hat{\theta}_n}{\theta_0} = \frac{1}{n}\sum_{i=1}^{n}\frac{(X_i - X_{(1)})}{\theta_0}$ and hence $\frac{2n\hat{\theta}_n}{\theta_0} \sim \chi^2_{2(n-1)}$. This implies that

$$\mathbb{P}\Big(2nk_1 \leq \frac{2\hat{\theta}_n}{\theta_0} \leq 2nk_2\Big) = 1 - \alpha.$$

Define $\chi^2_{n,\alpha}$ as a critical point such that $\mathbb{P}(\chi^2_n \leq \chi^2_{n,\alpha}) = \alpha$. Using the definition of $\chi^2_{n,\alpha}$, we can see that $k_1 = \frac{\chi^2_{2(n-1),\alpha_1}}{2n}$ and $k_2 = \frac{\chi^2_{2(n-1),1-\alpha+\alpha_1}}{2n}$ where $0 < \alpha_1 < \alpha$. Finally, the LRT can be written as

$$\phi(\underset{\sim}{X}) = \begin{cases} 1, & \text{if } \frac{\hat{\theta}}{\theta_0} < \frac{\chi^2_{2(n-1),\alpha_1}}{2n} \text{ or } \frac{\hat{\theta}}{\theta_0} > \frac{\chi^2_{2(n-1),1-\alpha+\alpha_1}}{2n}, \\ 0, & \text{otherwise} \end{cases}$$

and this gives the desire result.

10. The posterior distribution of θ^{-1} is $\Gamma\Big(n+\alpha, (\sum_{i=1}^{n} X_i + \beta^{-1})^{-1}\Big)$. To see this write:

$$\begin{aligned} f_{(\theta|\underset{\sim}{x})}(\theta) &= \pi(\theta) f(\underset{\sim}{x}|\theta) \\ &\propto \big(\theta^{-1}\big)^{\alpha-1} \exp\Big[-\frac{\theta^{-1}}{\beta}\Big] \big(\theta^{-1}\big)^{n} \exp\Big[-\theta^{-1}\sum_{i=1}^{n} x_i\Big] \\ &= \big(\theta^{-1}\big)^{n+\alpha-1} \exp\Big[-\theta^{-1}\sum_{i=1}^{n}\big(x_i + \frac{1}{\beta}\big)\Big], \end{aligned}$$

where the last line is proportional to the pdf of $\Gamma\Big(n+\alpha, \frac{1}{\sum_{i=1}^{n} x_i + \frac{1}{\beta}}\Big)$.

11. The Bayes estimator of θ under the squared error loss function is given by

$$\mathbb{E}_{\theta^{-1}|\underset{\sim}{X}}(\theta^{-1}) = \frac{n+\alpha}{\sum_{i=1}^{n} X_i + \frac{1}{\beta}}$$

and this gives the desired result.

12. We reject H_0 if $\mathbb{P}(\theta^{-1} \geq \theta_0^{-1}|\underset{\sim}{X}) < \frac{1}{2}$. Note that

$$\begin{aligned} \mathbb{P}\Big(\theta^{-1} \geq \theta_0^{-1}|\underset{\sim}{X}\Big) &= \mathbb{P}\Big(2\theta^{-1}\big(\sum_{i=1}^{n} X_i + \frac{1}{\beta}\big) \geq 2\theta_0^{-1}\big(\sum_{i=1}^{n} X_i + \frac{1}{\beta}\big)\big|\underset{\sim}{X}\Big) \\ &= \mathbb{P}\Big(\chi^2_{2(n+\alpha)} \geq 2\theta_0^{-1}\big(\sum_{i=1}^{n} X_i + \frac{1}{\beta}\big)\big|\underset{\sim}{X}\Big) < \frac{1}{2} \end{aligned}$$

and hence we reject the null hypothesis if $2\theta_0^{-1}\big(\sum_{i=1}^n X_i + \frac{1}{\beta}\big) < \chi^2_{2(n+\alpha),0.5}$. Therefore

$$\phi(\underset{\sim}{X}) = \begin{cases} 1, & \text{if } \sum_{i=1}^n X_i > \frac{\theta_0 \chi^2_{2(n+a),0.5}}{2} - \frac{1}{\beta}, \\ 0, & \text{otherwise.} \end{cases}$$

13. We know that $\theta^{-1}\big|\underset{\sim}{X} \sim \Gamma\big(n+a, \frac{1}{\sum_{i=1}^n X_i + \frac{1}{\beta}}\big)$ and $2\theta^{-1}\big(\sum_{i=1}^n X_i + \frac{1}{\beta}\big)|\underset{\sim}{X} \sim \chi^2_{2(n+\alpha)}$ (see the solution of Problem **10**). Now, set $L(\underset{\sim}{X})$ and $U(\underset{\sim}{X})$ such that

$$1-\alpha = \mathbb{P}\Big(2L(\underset{\sim}{X})(\sum_{i=1}^n X_i + \frac{1}{\beta}) \leq 2\theta^{-1}(\sum_{i=1}^n X_i + \frac{1}{\beta}) \leq 2U(\underset{\sim}{X})(\sum_{i=1}^n X_i + \frac{1}{\beta})\Big)$$

for any given values of $\sum_{i=1}^n X_i$ and β. Thus

$$\left[\frac{\chi^2_{2(n+\alpha),\alpha/2}}{2(\sum_{i=1}^n X_i + \frac{1}{\beta})}, \frac{\chi^2_{2(n+a),1-\alpha/2}}{2(\sum_{i=1}^n X_i + \frac{1}{\beta})}\right]$$

is a $(1-\alpha)$ credible set for θ^{-1}.

This question set is grouped into three separate parts: Parts I-II involve properties of probabilities/measures, sometimes in connection to independence. **Part III** relates to modes of convergence of random variables (e.g., almost surely or in distribution). Throughout this series of questions, you may use results from any previous questions, regardless of your solutions to these questions.

Part I

For a sequence of events A_n, $n \geq 1$, on a probability space $(\Omega, \mathcal{F}, P)$, let $B = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n$.

1. Give $P(B)$ (under the Borel-Cantelli lemma) when $\sum_{n=1}^{\infty} P(A_n) < \infty$.

2. Explain why the event B is often referred to as $B \equiv$"A_n infinitely often (i.o.)."

Let X_n, $n \geq 1$, be a sequence of iid standard exponential random variables on a common probability space $(\Omega, \mathcal{F}, P)$ (where $P(X_n \leq x) = 1 - e^{-x}$ for $x \geq 0$).

3. Determine the probability $P(\text{"}X_n > \alpha \log n \text{ i.o."})$ for each real $\alpha > 0$, stating any results used.

Hint: This is 1 if and only if $0 < \alpha \leq 1$.

4. Using your result from **Question 3**, show that $\limsup_{n\to\infty} \dfrac{X_n}{\log n} = 1$ almost surely.

Part II

5. For any two events A and B, show that $|P(A \cap B) - P(A)P(B)| \leq \frac{1}{4}$.

Hint: Apply the Cauchy-Schwarz inequality to random variables $X = I_A$ and $Y = I_B$, which are Bernoulli variables defined by indicator functions.

6. Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be two σ-algebras on a common probability space $(\Omega, \mathcal{F}, P)$ and define $\alpha \equiv \sup\{|P(A \cap B) - P(A)P(B)| : A \in \mathcal{G}_1, B \in \mathcal{G}_2\}$. Show that $\mathcal{G}_1$ and $\mathcal{G}_2$ are independent if and only if $\alpha = 0$.

7. If X and Y are both positive random variables on a common probability space $(\Omega, \mathcal{F}, P)$ with finite variances, show that

$$\mathrm{Cov}(X, Y) = \int_0^{\infty} \int_0^{\infty} \big[P(X > u, Y > v) - P(X > u)P(Y > v)\big] du dv.$$

Hint: Note $P(X > u, Y > v) = \mathrm{E}[I(X > u)I(Y > v)]$ for indicators $I(X > u)$ and $I(Y > v)$. In addition, $\int_0^{\infty} I(X > u) du = \int_0^X 1 du = X$ and $\int_0^{\infty} I(Y > v) dv = \int_0^Y 1 dv = Y$.

8. Supposing the positive random variables X, Y in **Question 7** are both bounded by some $c > 0$ with probability 1 (i.e., $P(X > c) = 0 = P(Y > c)$), show that

$$|\mathrm{Cov}(X, Y)| \leq \alpha c^2,$$

where $\alpha \equiv \sup\{|P(A \cap B) - P(A)P(B)| : A \in \sigma\langle X \rangle, B \in \sigma\langle Y \rangle\}$ and where $\sigma\langle X \rangle$ and $\sigma\langle Y \rangle$ denote the σ-algebras generated by X and Y, respectively.

Part III

Let $\{Y_n\}_{n\geq 0}$ and $\{Z_n\}_{n\geq 1}$ denote generic sequences of (real-valued) random variables.

9. Define the meaning of a sequence $\{Y_n\}_{n\geq 1}$ being tight (or stochastically bounded).

10. If $Y_n \xrightarrow{d} Y_0$ as $n \to \infty$ for some random variable Y_0, briefly explain why $\{Y_n\}_{n\geq 1}$ is then tight.

11. If $\{Y_n\}_{n\geq 1}$ is tight and if $Z_n \xrightarrow{p} 0$ as $n \to \infty$, show that $Y_n Z_n \xrightarrow{d} 0$.

12. Let $g : \mathbb{R} \to \mathbb{R}$ be a Borel measurable function with discontinuity points $D_g \equiv \{y \in \mathbb{R} : g(\cdot)$ is not continuous at $y\}$. If $Y_n \xrightarrow{d} Y_0$ as $n \to \infty$ for a random variable Y_0 where $P(Y_0 \in D_g) = 0$, show that $g(Y_n) \xrightarrow{d} g(Y_0)$.

Hint: Use the embedding theorem.

Suppose that $X_1, X_2, \ldots, X_n$ are iid Bernoulli$(1/2)$ random variables on a common probability space. Using the sample mean $\bar{X}_n \equiv \sum_{i=1}^n X_i/n$ along with $Q_n \equiv \sqrt{n}(2\bar{X}_n - 1)$ and an indicator function $I(\cdot)$, define a non-negative statistic λ_n as

$$\lambda_n \equiv Q_n^2 I(Q_n < 0) + R_n,$$

where R_n denotes a random variable bounded by

$$|R_n| \leq Q_n^2 S_n + nI(\bar{X}_n = 0), \quad \text{with } S_n \equiv |2\bar{X}_n - 1| \left(\frac{1}{(2\bar{X}_n)^3} + 1 - \bar{X}_n \right).$$

This form of $\lambda_n \geq 0$ above arises from an expansion of a log-likelihood ratio statistic (for testing $H_0 : \theta = 1/2$ with independent Bernoulli(θ), $\theta \in (0, 1/2]$, random variables).

In the following, let Z denote a standard normal variable and let $g(x) = x^2 I(x < 0)$, $x \in \mathbb{R}$.

13. Carefully explain why $g(Q_n) = Q_n^2 I(Q_n < 0)$ converges in distribution to $g(Z) = Z^2 I(Z < 0)$ as $n \to \infty$.

14. Show that the conditional distribution of $g(Z)$, given that $Z < 0$, is chi-square with 1 df.

Hint: Consider $P(Z^2 \leq x) = P(-\sqrt{x} \leq Z \leq \sqrt{x})$ and $P(Z^2 \leq x | Z < 0)$ for $x > 0$.

15. Prove that $S_n \to 0$ almost surely as $n \to \infty$, stating any standard results used.

16. Using the definition of convergence in probability, show $nI(\bar{X}_n = 0) \xrightarrow{p} 0$ as $n \to \infty$.

17. Based on the results in **Questions 9-16**, carefully show that, as $n \to \infty$, $\lambda_n \xrightarrow{d} B \cdot C$ holds, where B, C denote independent random variables with $B \sim$Bernoulli(1/2) and $C \sim \chi_1^2$ (chi-square with 1 df).

18. Briefly explain why $\lim_{n\to\infty} \mathrm{E}\lambda_n I(\lambda_n \leq 1)$ exists.

1. $P(B) = 0$ by the BC Lemma when $\sum_{n=1}^{\infty} P(A_n) < \infty$. ($P(B) = 1$ by the BC Lemma when $\{A_n\}_{n\geq 1}$ are independent and $\sum_{n=1}^{\infty} P(A_n) = \infty$.)

2. For $\omega \in \Omega$, $\omega \in B \equiv \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n$ if and only if, for each integer $N \geq 1$, there exists an integer $n \equiv n(\omega, N) \geq N$ (depending possibly on ω and N) where $\omega \in A_n$. For this reason, the event B is equivalent to "A_n i.o.".

3. Note that, for a given $\alpha > 0$,

$$\sum_{n=1}^{\infty} P(X_n > \alpha \log n) = \sum_{n=1}^{\infty} e^{-\alpha \log n} = \sum_{n=1}^{\infty} e^{\log n^{-\alpha}} = \sum_{n=1}^{\infty} \frac{1}{n^{\alpha}},$$

which is finite iff $\alpha > 1$. By the BC lemma, $P(\text{"}X_n > \alpha \log n \text{ i.o."}) = 0$ if $\alpha > 1$; also $P(\text{"}X_n > \alpha \log n \text{ i.o."}) = 1$ holds if $0 < \alpha \leq 1$ by the BC lemma and the independence of events $\{X_n > \alpha \log n\}_{n \geq 1}$.

4. Define events $B_k =$"$\frac{X}{\log n} > 1 - \frac{1}{k}$ i.o." and $C_k =$"$\frac{X}{\log n} > 1 + \frac{1}{k}$ i.o." for $k \geq 2$. By Question 3, $P(B_k) = 1$ and $P(C_k) = 0$ for each $k \geq 2$. Define the event $A = \cap_{k=1}^{\infty}(B_k \cap C_k^c)$ where $P(A) = 1$ holds (e.g., $P(A^c) \leq \sum_{k=1}^{\infty}[P(B_k^c) + P(C_k)] = 0$). Pick $k \geq 2$. For $\omega \in A$, "$X_n(\omega)/\log n > 1 - 1/k$ holds i.o.," implying $\limsup_{n\to\infty} X_n(\omega)/\log n \geq 1 - 1/k$. Likewise, for $\omega \in A$, "$X_n(\omega)/\log n \leq 1 + 1/k$ holds eventually" (or "$X_n(\omega)/\log n > 1 + 1/k$ i.o." does not hold), implying $\limsup_{n\to\infty} X_n(\omega)/\log n \leq 1 + 1/k$. Hence, for $\omega \in A$,

$$1 - 1/k \leq \limsup_{n\to\infty} \frac{X_n(\omega)}{\log n} \leq 1 + 1/k$$

holds for any $k \geq 2$, implying that $\limsup_{n\to\infty} X_n(\omega)/\log n = 1$. Because $P(A) = 1$, then $\limsup_{n\to\infty} X_n/\log n = 1$ almost surely.

5. By the Cauchy-Schwarz inequality for $X = I_A$, $Y = I_B$,

$$|P(A \cap B) - P(A)P(B)| = |\text{Cov}(X, Y)| \leq \sqrt{\text{Var}(X)\text{Var}(Y)} = \sqrt{P(A)[1 - P(A)]P(B)[1 - P(B)]}.$$

Because the function $p(1-p)$, $p \in [0, 1]$, has a maximum of $1/4$ at $p = 1/2$, we have $0 \leq P(A)[1 - P(A)], P(B)[1 - P(B)] \leq 1/4$ and the bound follows.

6. $\mathcal{G}_1$ and $\mathcal{G}_2$ are independent if and only if $P(A \cap B) = P(A)P(B)$ holds for any $A \in \mathcal{G}_1$ and $B \in \mathcal{G}_2$. Hence, if $\mathcal{G}_1$ and $\mathcal{G}_2$ are independent, then $\alpha = 0$ holds for $\alpha \equiv \sup\{|P(A \cap B) - P(A)P(B)| : A \in \mathcal{G}_1, B \in \mathcal{G}_2\}$. Conversely, if $\alpha = 0$ holds, then $P(A \cap B) = P(A)P(B)$ follows for any $A \in \mathcal{G}_1$ and $B \in \mathcal{G}_2$, implying $\mathcal{G}_1$ and $\mathcal{G}_2$ are independent.

7. Fubini's theorem (non-negative integrands) allows the integration orders to be switched in

$$\begin{aligned}
\int_0^{\infty} \int_0^{\infty} P(X > u, Y > v) du dv &= \int_0^{\infty} \int_0^{\infty} \mathrm{E}\big[I(X > u)I(Y > v)\big] du dv \\
&= \mathrm{E}\left[\int_0^{\infty} \int_0^{\infty} I(X > u)I(Y > v) du dv\right] \\
&= \mathrm{E}\left[\int_0^{\infty} I(X > u) du \cdot \int_0^{\infty} I(Y > v) dv\right] \\
&= \mathrm{E}\left[\int_0^{X} 1 du \cdot \int_0^{Y} 1 dv\right] = \mathrm{E}[XY];
\end{aligned}$$

likewise, it holds that

$$\int_0^\infty P(X > u)du = \int_0^\infty \mathrm{E}I(X > u)du = \mathrm{E}\left(\int_0^\infty I(X > u)du\right) = \mathrm{E}(X)$$

and $\int_0^\infty P(Y > v)dv = \mathrm{E}(Y)$. Since $\mathrm{E}(X), \mathrm{E}(Y)$ and $\mathrm{E}(XY)$ are all finite by assumption, the result follows.

8. By Question 7, we have

$$|\mathrm{Cov}(X,Y)| \leq \int_0^\infty \int_0^\infty |P(X > u, Y > v) - P(X > u)P(Y > v)|dudv.$$

By assumption, $P(X > u, Y > v)$ and $P(X > u)P(Y > v)$ are both zero if either $u > c$ or $v > c$ hold; and $|P(X > u, Y > v) - P(X > u)P(Y > v)|$ is bounded by $\alpha \equiv \sup\{|P(A \cap B) - P(A)P(B)| : A \in \sigma\langle X\rangle, B \in \sigma\langle Y\rangle\}$ for any $u, v > 0$. Consequently, for any $(u, v) \in (0, \infty)^2$, we have

$$|P(X > u, Y > v) - P(X > u)P(Y > v)| \leq \begin{cases} 0 & \text{if } u > c \text{ or if } v > c \\ \alpha & \text{if } (u, v) \in (0, c] \times (0, c], \end{cases}$$

and, hence,

$$|\mathrm{Cov}(X,Y)| \leq \int_0^c \int_0^c \alpha dudv = \alpha c^2.$$

9. A sequence $\{Y_n\}_{n\geq 1}$ of random variables is tight (or bounded in probability) if the marginal probability of any single random variable being large can be uniformly controlled: for any ϵ, there exists some $M \equiv M(\epsilon)$ where $\sup_{n\geq 1} P(|Y_n| > M) < \epsilon$.

10. If $Y_n \xrightarrow{d} Y_0$ as $n \to \infty$, then the probability structure of Y_n converges to that of Y_0 and, consequently, a bound on the probability of "large Y_0" translates to a bound on the (marginal) probability of "large Y_n," where the same bound for Y_0 applies to Y_n for all sufficiently large n. This suffices to ensure tightness of $\{Y_n\}_{n\geq 1}$.

More technically, fix $\epsilon > 0$ and pick a constant $M > 0$ where $P(Y_0 = \pm M) = 0$ and $P(|Y_0| > M) < \epsilon$. Then, since $Y_n \xrightarrow{d} Y_0$ implies $\lim_{n\to\infty} P(|Y_n| > M) = P(|Y_0| > M)$, there exists some N where $\sup_{n\geq N} P(|Y_n| > M) < \epsilon$. For $i = 1, \ldots, N$, pick M_i so that $P(|Y_i| > M_i) < \epsilon$. Then, $\tilde{M} \equiv \max\{M, M_1, \ldots, M_N\}$ gives $\sup_{n\geq 1} P(|Y_n| > \tilde{M}) < \epsilon$.

11. One way to show $Y_n Z_n \xrightarrow{d} 0$ (equivalent to $Y_n Z_n \xrightarrow{p} 0$) is to show that: given an arbitrary subsequence $\{n_j\}$ of $\{n\}$, there exists a further subsequence $\{n_k\} \subset \{n_j\}$ such that $Y_{n_k} Z_{n_k} \xrightarrow{d} 0$. Take some $\{n_j\}$. Because $\{Y_n\}$ is tight by assumption, there exists a subsequence $\{n_k\} \subset \{n_j\}$ and a random variable $\tilde{Y}_0$ (possibly depending on $\{n_k\}$) such that $Y_{n_k} \xrightarrow{d} \tilde{Y}_0$. But then $Z_{n_k} \xrightarrow{p} 0$ (by assumption) and Slutsky's theorem gives $Y_{n_k} Z_{n_k} \xrightarrow{d} 0$, as needed.

Another option for showing $Y_n Z_n \xrightarrow{p} 0$ is brute force. Pick $\epsilon > 0$ and $M > 0$. Then,

$$\begin{aligned} P(|Y_n Z_n| > \epsilon) &\leq P(|Y_n| > M) + P(|Y_n Z_n| > \epsilon, |Y_n| \leq M) \\ &\leq P(|Y_n| > M) + P(|Z_n| > \epsilon/M) + P(|Y_n Z_n| > \epsilon, |Y_n| \leq M, |Z_n| \leq \epsilon/M), \end{aligned}$$

where $P(|Y_n Z_n| > \epsilon, |Y_n| \leq M, |Z_n| \leq \epsilon/M) = 0$. Because $\lim_{n\to\infty} P(|Z_n| > \epsilon/M) = 0$ by $Z_n \xrightarrow{p} 0$, we may bound

$$\limsup_{n\to\infty} P(|Y_n Z_n| > \epsilon) \leq \limsup_{n\to\infty} P(|Z_n| > \epsilon/M) + \limsup_{n\to\infty} P(|Y_n| > M) \leq \sup_{n\geq 1} P(|Y_n| > M);$$

since $M > 0$ was arbitrary and since $\sup_{n\geq 1} P(|Y_n| > M) \to 0$ as $M \to \infty$ by tightness, it follows that $\limsup_{n\to\infty} P(|Y_n Z_n| > \epsilon) = 0$. In other words, $P(|Y_n Z_n| > \epsilon) \to 0$ holds for any ϵ and so $Y_n Z_n \xrightarrow{p} 0$ follows; the latter is equivalent to $Y_n Z_n \xrightarrow{d} 0$.

12. Since $Y_n \xrightarrow{d} Y_0$, the embedding theorem gives that there exist random variables X_n, $n \geq 0$, on a common probability space $(\Omega, \mathcal{F}, P)$ where $X_n \to X_0$ almost surely and where X_n has the same distribution as Y_n for each $n \geq 0$. The event $A \equiv$"$\lim_{n\to\infty} X_n = X_0$" and event $B \equiv$"$X_0 \notin D_g$" both have probability 1 (since $P(X_0 \in D_g) = P(Y_0 \in D_g) = 0$). For $\omega \in A \cap B$, $\lim_{n\to\infty} X_n(\omega) = X_0(\omega)$ holds and $g(\cdot)$ is continuous at $X_0(\omega)$; consequently, $\lim_{n\to\infty} g(X_n(\omega)) = g(X_0(\omega))$ holds too for $\omega \in A \cap B$ where $P(A \cap B) = 1$. Hence, $g(X_n) \to g(X_0)$ almost surely, implying that $g(X_n) \xrightarrow{d} g(X_0)$. Due to matching distributions, this gives $g(Y_n) \xrightarrow{d} g(Y_0)$.

13. By the CLT for iid Bernoulli$(1/2)$ variables, we have $\sqrt{n}(\bar{X}_n - 1/2) \xrightarrow{d} N(0, \mathrm{Var}(X_1) = 1/4)$ as $n \to \infty$; consequently, $Q_n \equiv 2\sqrt{n}(\bar{X}_n - 1/2) \xrightarrow{d} Z \sim N(0,1)$. The function $g(x) = x^2 I(x < 0)$ is discontinuous only at $x = 0$, where $P(Z = 0) = 0$. By Question 12, it follows that $g(Q_n) \xrightarrow{d} g(Z) = Z^2 I(Z < 0)$.

14. For $x > 0$, the conditional cdf of $g(Z)$ is given by $P(g(Z) \leq x | Z < 0) = P(Z^2 \leq x, Z < 0)/P(Z < 0) = 2P(-\sqrt{x} \leq Z \leq \sqrt{x}, Z < 0) = 2P(-\sqrt{x} \leq Z < 0) = P(-\sqrt{x} \leq Z < 0) + P(0 < Z \leq \sqrt{x}) = P(-\sqrt{x} \leq Z \leq \sqrt{x}) = P(Z^2 \leq x)$, using symmetry of the standard normal distribution. Hence, the variable $g(Z)$, conditionally on $Z < 0$, has the same distribution as Z^2 (i.e., χ^2_1).

15. By the SLLN, $\bar{X}_n \to 1/2$ as $n \to \infty$ almost surely. The mapping $h(x) = |2x - 1|((2x)^{-3} + 1 - x)$, $x \in \mathbb{R}$, is continuous at $x = 1/2$ so that $S_n = h(\bar{X}_n) \to h(1/2) = 0$ almost surely.

16. Fix $\epsilon > 0$. Then, $P(nI(\bar{X}_n = 0) > \epsilon) \leq P(\bar{X}_n = 0) = \prod_{i=1}^n P(X_i = 0) = 2^{-n} \to 0$ as $n \to \infty$.

17. Note that $\lambda_n = g(Q_n) + R_n$, where $g(Q_n) \xrightarrow{d} g(Z) = Z^2 I(Z < 0)$ by Question 13. Hence, $\lambda_n \xrightarrow{d} Z^2 I(Z < 0)$ will follow by Slutsky's theorem if $|R_n| \xrightarrow{p} 0$ is shown. Note that $|R_n| \leq Q_n^2 S_n + nI(\bar{X}_n = 0)$. Since $Q_n^2 \xrightarrow{d} Z^2$ holds by $Q_n \xrightarrow{d} Z$ and the continuous mapping theorem, we have Q_n^2 is tight (Question 10); taking this along with $S_n \xrightarrow{p} 0$ from Question 15, we have $Q_n^2 S_n \xrightarrow{p} 0$ by Question 11. Then, from $nI(\bar{X}_n = 0) \xrightarrow{p} 0$ in Question 16, we also have $Q_n^2 S_n + nI(\bar{X}_n = 0) \xrightarrow{p} 0$ by Slutsky's theorem; this implies $|R_n| \xrightarrow{p} 0$. Finally, the distribution of $Z^2 I(Z < 0)$ is equivalent to that of $B \cdot C$, for independent variables B, C where $C \sim \chi^2_1$ and $B \sim$Bernoulli(1/2); Question 14 entails this part when taking $B = I(Z < 0)$ and $V = Z^2$.

18. Because $\lambda_n \xrightarrow{d} B \cdot C$ by Question 17 and because the function $xI(x \leq 1)$, $x \in \mathbb{R}$, is discontinuous only at $x = 1$ (where $P(B \cdot C = 1) = P(B = 1, C = 1) = 0$), we have $\lambda_n I(\lambda_n \leq 1) \xrightarrow{d} (B \cdot C)I(B \cdot C \leq 1)$ by Question 12. Then, convergence in distribution combined with uniform integrability (i.e., $\lambda_n I(\lambda_n \leq 1) \in [0,1]$ is bounded) give that $\lim_{n\to\infty} \mathrm{E}\lambda_n I(\lambda_n \leq 1) = \mathrm{E}(B \cdot C)I(B \cdot C \leq 1)$. The latter equals $(2)^{-3/2} \int_0^1 x^{0.5} e^{-x/2} dx / [\Gamma(0.5)]$.