Part I

Suppose X_i is $N(\mu_i, 1)$ for $i = 1, 2$, and X_1 and X_2 are independent. Define

$$Y = X_1 + X_2.$$

1. Find the joint distribution of $\begin{bmatrix} X_1 \\ X_2 \\ Y \end{bmatrix}$.

2. Find the conditional distribution of $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ given Y.

Part II

Suppose (Y_i, Z_i) for $i = 1, 2, \cdots, n$ are i.i.d. from a distribution with the probability density function $f_\theta(y, z) = e^{-\theta y - \theta^{-1} z}$, for $y > 0, z > 0$ and $\theta > 0$.

3. Find the MLE of θ based on $(Y_1, Z_1), \ldots, (Y_n, Z_n)$. Call it $\hat{\theta}_n$.

4. Show that $\hat{\theta}_n \longrightarrow \theta$ in probability as $n \to \infty$.

5. Under certain conditions $\hat{\theta}_n$ is asymptotically normal, i.e., $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, V(\theta))$ as $n \to \infty$. Find an expression for the asymptotic variance function $V(\theta)$.

Part III

Suppose that a random variable N follows a Poisson distribution with mean θ, i.e. $N \sim Poisson(\theta)$. Its pmf is $P(N = n) = \frac{\theta^n e^{-\theta}}{n!}, n = 0, 1, 2, \ldots$. Given N, suppose the conditional distribution of Y is χ^2_{2N}, i.e. $Y|N \sim \chi^2_{2N}$.

6. Show that $E[Y] = 2\theta$ and $V[Y] = 8\theta$. (Here $E[\cdot]$ and $V[\cdot]$ are expectation and variance with respect to the marginal distribution of Y.)

7. Show that $E[a^N] = e^{(a-1)\theta}$ for any arbitrary value a. (Here $E[\cdot]$ is expectation with respect to the distribution of N.)

Define $Z = \frac{Y-E[Y]}{\sqrt{V[Y]}}$.

8. Show that the moment generating function for Z is

$$M_Z(t) = e^{\frac{t^2}{2(1-t/\sqrt{2\theta})}}.$$

Hint: Use the fact that if a random variable $X \sim \chi^2_k$, the moment generating function of X is $M_X(t) = (1-2t)^{-k/2}$ for $t < 1/2$.

9. Show that as $\theta \to \infty$,

$$Z \xrightarrow{d} N(0,1).$$

Part IV

Let X and Y be independent random variables with $X \sim Exponential(\lambda)$ and $Y \sim Exponential(\mu)$, i.e., the pdf of X is $f(x) = \frac{1}{\lambda}e^{-\frac{x}{\lambda}}$ for $x > 0$ and the pdf of Y is $f(y) = \frac{1}{\mu}e^{-\frac{y}{\mu}}$ for $y > 0$. Let

$$Z = min(X,Y) \quad \text{and} \quad W = \begin{cases} 1 & \text{if } Z = X, \\ 0 & \text{if } Z = Y. \end{cases}$$

10. Find the joint distribution of Z and W.

11. Prove that Z and W are independent.

Part V

Let X and Y be independent $N(0,1)$ random variables and define a new random variable Z by

$$Z = \begin{cases} X & \text{if } XY > 0, \\ -X & \text{if } XY < 0. \end{cases}$$

12. Show that Z has a normal distribution.

13. Show that the joint distribution of Z and Y is NOT bivariate normal.

Key to Ph.D. Exam – Theory I – July 2018

Part I:

1. Define $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$, so that we can write

$$\begin{bmatrix} X_1 \\ X_2 \\ Y \end{bmatrix} = A \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}.$$

Since $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$,

$$\begin{bmatrix} X_1 \\ X_2 \\ Y \end{bmatrix} = A \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N\left(A \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, AA^T\right)$$

i.e. $\begin{bmatrix} X_1 \\ X_2 \\ Y \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_1+\mu_2 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}\right).$

2. $\begin{bmatrix} X_1 \\ X_2 \\ Y \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_1+\mu_2 \end{bmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)$, where $\Sigma_{11} = I_{2\times 2}$,

$\Sigma_{12} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\Sigma_{21} = \begin{bmatrix} 1 & 1 \end{bmatrix}$, $\Sigma_{22} = 2$.

So $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \Big| Y \sim N(\bar{\mu}, \bar{\Sigma})$

where $\bar{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \Sigma_{12}\Sigma_{22}^{-1}(Y - (\mu_1+\mu_2)) = \begin{bmatrix} \frac{Y+\mu_1-\mu_2}{2} \\ \frac{Y-\mu_1+\mu_2}{2} \end{bmatrix}$

$$\bar{\Sigma} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

So $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \Big| Y \sim N\left(\begin{bmatrix} \frac{Y+\mu_1-\mu_2}{2} \\ \frac{Y-\mu_1+\mu_2}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}\right).$

Part II:

3. $f(\underline{Y}, \underline{Z}; \theta) = \prod_{i=1}^{n} e^{-\theta y_i - \theta^{-1} z_i} = e^{-\theta \sum_{i=1}^{n} y_i - \theta^{-1} \sum_{i=1}^{n} z_i}$

$\log f(\underline{Y}, \underline{Z}; \theta) = -\theta \sum_{i=1}^{n} y_i - \theta^{-1} \sum_{i=1}^{n} z_i$

$\frac{\partial \log f(\underline{Y}, \underline{Z}; \theta)}{\partial \theta} = -\sum_{i=1}^{n} y_i + \frac{1}{\theta^2} \sum_{i=1}^{n} z_i$

Solve the equation $\frac{\partial \log f(\underline{Y}, \underline{Z}; \theta)}{\partial \theta} = 0$ for θ,

we obtain the MLE $\hat{\theta}_n$ for θ_0

$$\hat{\theta}_n = \sqrt{\frac{\sum_{i=1}^{n} z_i}{\sum_{i=1}^{n} y_i}} = \sqrt{\frac{\bar{z}}{\bar{y}}} \quad \text{because } \theta_0 > 0.$$

4. The marginal pdf of Y is:

$f(y) = \int_0^\infty f(y,z)\,dz = \int_0^\infty e^{-\theta_0 y - \theta_0^{-1} z}\,dz$

$= \theta_0 e^{-\theta_0 y} \underbrace{\frac{1}{\theta_0}\int_0^\infty e^{-\theta_0^{-1} z}\,dz}_{=1} = \theta_0 e^{-\theta_0 y}$

$\Rightarrow Y \sim \exp\left(\frac{1}{\theta_0}\right) \Rightarrow E[Y] = \frac{1}{\theta_0}$

The marginal pdf of Z is:

$f(z) = \int_0^\infty f(y,z)\,dy = \int_0^\infty e^{-\theta_0 y - \theta_0^{-1} z}\,dy$

$= \frac{1}{\theta_0} e^{-\theta_0^{-1} z} \underbrace{\theta_0 \int_0^\infty e^{-\theta_0 y}}_{=1} = \frac{1}{\theta_0} e^{-\theta_0^{-1} z}$

$\Rightarrow Z \sim \exp(\theta_0) \Rightarrow E[Z] = \theta_0.$

By the strong law of large numbers,

$\bar{z} \xrightarrow{a.s.} \theta_0, \quad \bar{y} \xrightarrow{a.s.} \frac{1}{\theta_0}$

By the mapping theorem,

$\hat{\theta}_n = \sqrt{\frac{\bar{z}}{\bar{y}}} \xrightarrow{a.s.} \sqrt{\frac{\theta_0}{\frac{1}{\theta_0}}} = \theta_0.$

5. Under certain conditions,

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, V(\theta_0))$$

where $V(\theta_0) = I^{-1}(\theta_0)$.

Here

$$I(\theta_0) = E\left(\left.\frac{\partial \log f(Y,Z)}{\partial \theta}\right|_{\theta=\theta_0}\right)^2$$

$$= -E\left(\left.\frac{\partial^2 \log f(Y,Z)}{\partial \theta}\right|_{\theta=\theta_0}\right)$$

$$= -E\left(-2\frac{1}{\theta_0^3} Z\right)$$

$$= \frac{2}{\theta_0^3} E(Z)$$

$$= \frac{2}{\theta_0^3} \times \theta_0 = \frac{2}{\theta_0^2}$$

So $V(\theta_0) = I^{-1}(\theta_0) = \frac{\theta_0^2}{2}$.

i.e.

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \rightarrow N(0, \frac{\theta_0^2}{2}).$$

Part III

6. $$E[Y] = E_N\left[E_{Y|N}(Y|N)\right] = E_N[2N] = 2E_N[N] = 2\theta.$$

$$V[Y] = E_N\left[V_{Y|N}(Y|N)\right] + V_N\left[E_{Y|N}(Y|N)\right]$$
$$= E_N[4N] + V_N[2N]$$
$$= 4\cdot E_N[N] + 4\cdot V_N[N]$$
$$= 4\theta + 4\theta = 8\theta.$$

7. $$E_N[a^N] = \sum_{N=0}^{\infty} a^N \cdot \frac{e^{-\theta}\theta^N}{N!} = e^{-\theta}\sum_{N=0}^{\infty}\frac{(a\theta)^N}{N!}$$
$$= e^{a\theta}\cdot e^{-\theta}\underbrace{\sum_{N=0}^{\infty}\frac{e^{-(a\theta)}(a\theta)^N}{N!}}_{=1}$$
$$= e^{(a-1)\theta}\cdot 1$$
$$= e^{(a-1)\theta}.$$

8. $$Z = \frac{Y-E[Y]}{\sqrt{V(Y)}} = \frac{Y-2\theta}{\sqrt{8\theta}}$$

$$M_Z(t) = E[e^{tZ}] = E\left[e^{\frac{t}{\sqrt{8\theta}}(Y-2\theta)}\right] = e^{-\frac{t}{\sqrt{2\theta}}\theta}E\left[e^{\frac{t}{\sqrt{8\theta}}Y}\right]$$
$$= e^{-\frac{t}{\sqrt{2\theta}}\theta}E_N\left[E_{Y|N}\left(e^{\frac{t}{\sqrt{8\theta}}Y}\right)\right]$$
$$= e^{-\frac{t}{\sqrt{2\theta}}\theta}\cdot E_N\left[\left(\frac{1}{1-\frac{2t}{\sqrt{8\theta}}}\right)^N\right] = e^{-\frac{t}{\sqrt{2\theta}}\theta}\cdot E_N\left[\left(\frac{1}{1-\frac{t}{\sqrt{2\theta}}}\right)^N\right]$$

Define $a = \dfrac{1}{1-\frac{t}{\sqrt{2\theta}}}$, then

$$M_Z(t) = e^{-\frac{t\theta}{\sqrt{2\theta}}}\cdot e^{(a-1)\theta}.$$

$$\text{or} = e^{\frac{t^2}{2\left(1-\frac{t}{\sqrt{2\theta}}\right)}}.$$

9\. According to the result in (8),

$$M_Z(t) = e^{\frac{t^2}{2(1-\frac{t}{\sqrt{2\theta}})}}.$$

When $\theta \to \infty$,

$$\Rightarrow -\frac{t}{\sqrt{2\theta}} \to 0 \quad \text{for any fixed } t$$

$$\Rightarrow 2(1-\frac{t}{\sqrt{2\theta}}) \to 2 \quad \text{for any fixed } t$$

$$\Rightarrow \frac{t^2}{2(1-\frac{t}{\sqrt{2\theta}})} \to \frac{t^2}{2} \quad \text{for any fixed } t.$$

Because the function $g(u) = e^u$ is a continuous function,

So $M_Z(t) \to e^{t^2/2}$ as $\theta \to \infty$.

And $e^{t^2/2}$ is the MGF of $N(0,1)$,

thus we have $Z \xrightarrow{d} N(0,1)$.

Part IV:

10.

$$P(Z \le z, W=0) = P(\min(X,Y) \le z,\ Y \le X)$$
$$= P(Y \le z,\ Y \le X)$$
$$= \int_0^z \int_y^\infty \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \frac{1}{\mu} e^{-\frac{y}{\mu}}\, dx\, dy$$
$$= \int_0^z \frac{1}{\mu} e^{-\frac{y}{\mu}} e^{-\frac{y}{\lambda}}\, dy$$
$$= \frac{\lambda}{\mu+\lambda}\left[1 - e^{-\left(\frac{1}{\mu}+\frac{1}{\lambda}\right) z}\right]$$

$$P(Z \le z, W=1) = P(\min(X,Y) \le z,\ X \le Y)$$
$$= P(X \le z,\ X \le Y)$$
$$= \int_0^z \int_x^\infty \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \frac{1}{\mu} e^{-\frac{y}{\mu}}\, dy\, dx$$
$$= \int_0^z \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \cdot e^{-\frac{x}{\mu}}\, dx$$
$$= \frac{\mu}{\mu+\lambda}\left[1 - e^{-\left(\frac{1}{\mu}+\frac{1}{\lambda}\right) z}\right]$$

11.

$$P(W=0) = P(Y \le X) = \int_0^{\infty}\int_y^{\infty} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \frac{1}{\mu} e^{-\frac{y}{\mu}} dx\,dy$$

$$= \int_0^{\infty} \frac{1}{\mu} e^{-\frac{y}{\mu}} e^{-\frac{y}{\lambda}} dy$$

$$= \frac{\lambda}{\mu+\lambda}$$

$$P(W=1) = P(X \le Y) = 1 - P(W=0) = \frac{\mu}{\mu+\lambda}.$$

$$P(Z \le z) = P(Z \le z, W=0) + P(Z \le z, W=1)$$

$$= \frac{\lambda}{\mu+\lambda}\left[1 - e^{-(\frac{1}{\mu}+\frac{1}{\lambda})z}\right] + \frac{\mu}{\mu+\lambda}\left[1 - e^{-(\frac{1}{\mu}+\frac{1}{\lambda})z}\right]$$

$$= 1 - e^{-(\frac{1}{\mu}+\frac{1}{\lambda})z}$$

Therefore, $P(Z \le z, W=0) = P(Z \le z)\,P(W=0)$

& $P(Z \le z, W=1) = P(Z \le z)\,P(W=1)$.

So, Z and W are independent.

part V:

12. For $z<0$,

$$
\begin{aligned}
P(Z\le z) &= P(X\le z \text{ and } XY>0) + P(-X\le z \text{ and } XY<0) \\
&= P(X\le z \text{ and } Y<0) + P(X\ge -z \text{ and } Y<0) \quad \text{since } z<0 \\
&= P(X\le z)P(Y<0) + P(X\ge -z)P(Y<0) \quad \text{symmetric of } X \,\&\, Y \\
&= P(X\le z)P(Y<0) + P(X\le z)P(Y>0) \\
&= P(X\le z)
\end{aligned}
$$

Similarly for $z>0$,

$$
\begin{aligned}
P(Z>z) &= P(X>z \text{ and } XY>0) + P(-X>z \text{ and } XY<0) \\
&= P(X>z \text{ and } Y>0) + P(X<-z \text{ and } Y>0) \\
&= P(X>z)P(Y>0) + P(X<-z)P(Y>0) \\
&= P(X>z)P(Y>0) + P(X>z)P(Y<0) \\
&= P(X>z)
\end{aligned}
$$

$$\Rightarrow P(Z\le z) = P(X\le z)$$

Thus, $Z \sim X \sim N(0,1)$

13. By definition of Z,

$$Z>0 \iff \begin{cases} X>0 \text{ and } Y>0 \\ \quad\text{or} \\ X<0 \text{ and } Y>0 \end{cases}$$

So Z and Y always have the same sign, hence they can not be bivariate normal.

A random variable X is said to have a Weibull distribution with a scale parameter θ, and a shape parameter κ, if its pdf is given by

$$f(x|\theta,\kappa) = \begin{cases} \frac{\kappa}{\theta} x^{\kappa-1} e^{-\left(\frac{x^\kappa}{\theta}\right)} & \text{for } x > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where $\kappa > 0$ and $\theta > 0$. We use $X \sim \mathrm{WEB}(\kappa,\theta)$ to indicate that the random variable X has a Weibull distribution with parameters κ and θ.

In this exam, we assume $X_1, X_2, \ldots, X_n$ are iid $\mathrm{WEB}(\kappa,\theta)$ and we assume κ is a known parameter and θ is an unknown parameter.

1. Show that $\mathrm{E}[\mathrm{X^r}] = \theta^{\mathrm{r}}\Gamma(1+\frac{\mathrm{r}}{\kappa})$, where $X \sim \mathrm{WEB}(\kappa,\theta)$.

2. Find $\tilde{\theta}_n$, the method of moments estimator (MME) of θ, based on $X_1, X_2, \ldots, X_n$.

3. Find $\hat{\theta}_n$, the maximum likelihood estimator (MLE) of θ, based on $X_1, X_2, \ldots, X_n$.

4. Show that $\hat{\theta}_n$ is an unbiased estimator for θ.

5. Show that $\tilde{\theta}_n$ and $\hat{\theta}_n$ are both consistent estimators for θ.

6. Define $\eta = \theta^\kappa$. prove that $\hat{\eta} = \frac{\Gamma(n)}{\Gamma(n+\kappa)}\left(\sum_{i=1}^n X_i^\kappa\right)^\kappa$ is the uniformly minimum variance unbiased estimator (UMVUE) of η.

7. Let $X_{(1)} = \min(X_1, \ldots, X_n)$ be the first order statistic based on the random sample $X_1, \ldots, X_n$. Prove that
$$\frac{d}{d\theta}\mathrm{E}_\theta\left[(\mathrm{X}_{(1)})^\kappa \middle| \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{X}_{\mathrm{i}}^\kappa\right] = 1$$
for $\theta > 0$.

8. Define a sequence of random variables
$$Q_n = Q_n(X_1, \ldots, X_n) = \frac{\kappa^{\gamma_1}\sqrt{n}(\hat{\theta}_n - \theta)}{(\hat{\theta}_n)^{\gamma_2}}$$
for $n \geq 1$. Find values of γ_1 and γ_2 so that Q_n is an asymptotical pivot.

9. Using the asymptotic distribution of Q_n in Problem **8**, find a confidence interval for θ with approximate confidence coefficient $1-\alpha$.

10. Consider testing $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$ based on $X_1, \ldots, X_n$. Show that the test that rejects H_0 when $\sum_{i=1}^n X_i^\kappa > c$ for some constant c is a uniformly most powerful (UMP) test. Determine c so that the test will have size α.

11. Construct a confidence interval with confidence coefficient $1 - \alpha$ for θ based on $T = \sum_{i=1}^{n} X_i^{\kappa}$.

For problems **12** to **14**, suppose θ has the Inverse-Gamma (IG) prior density

$$\pi(\theta) = \begin{cases} \frac{b^a}{\Gamma(a)\theta^{a+1}} e^{-\frac{b}{\theta}} & \text{for } \theta > 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $a > 0$ and $b > 0$.

12. Derive the Bayes estimator of θ under the squared error loss.

13. Find a Bayes test for $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$.

14. Find a $(1 - \alpha)$ Bayes credible set for θ.

This problem requires background. The Weibull parameterization for X has cdf $P(X \le x) = 1 - e^{-x^\kappa/\theta}$ for $x > 0$ so that the cdf of $Y \equiv X^\kappa$ is then

$$P(Y \le y) = P(X \le y^{1/\kappa}) = 1 - e^{-y/\theta}, \quad \text{for } y > 0.$$

The key idea becomes that if $X \sim \text{WEB}(\kappa, \theta)$ in the Weibull parameterization given, then

$$Y \equiv X^\kappa \sim \text{Exp}(\theta) \quad (\textit{Exponential with mean } \theta > 0)$$

holds for given $\kappa > 0$, which gives the mean $\text{E}_\theta(Y) = \theta > 0$ and variance $\text{Var}_\theta(Y) = \theta^2$.

Letting

$$\Gamma(r) \equiv \int_0^\infty z^{r-1} e^{-z} dz, \quad \text{for real } r > 0,$$

denote the gamma function, it is also helpful that $Z \equiv Y/\theta \sim \text{Exp}(1)$ is distributed as a standard Exponential, which has the expected value property:

$$\text{E}(Z^r) = \int_0^\infty z^r e^{-z} dz = \Gamma(r+1), \quad \text{for real } r > 0.$$

With this background (i.e., moment details as above should perhaps be provided), the questions relate mostly to the Exponential distribution.

1. Question 1 uses incorrect moments, which affects Question 2. For $r > 0$, Question 1 becomes $\text{E}_\theta(X^r) = \text{E}_\theta([X^\kappa]^{r/\kappa}) = \text{E}_\theta(Y^c)$, using $Y = X^\kappa$ and setting $c \equiv r/\kappa > 0$ for simplicity. Then, by $Z \equiv Y/\theta \sim \text{Exp}(1)$, we have

$$\text{E}_\theta(X^r) = \text{E}_\theta(Y^c) = \text{E}_\theta([\theta \cdot Y/\theta]^c) = \text{E}_\theta(\theta^c Z^c) = \theta^c \text{E}(Z^c) = \theta^c \Gamma(c+1) = \theta^{r/\kappa} \Gamma\left(1 + \frac{r}{\kappa}\right).$$

2. In Question 2, using $r = 1$, the method of moments estimator becomes the θ-solution to

$$\bar{X}_n \equiv \frac{1}{n} \sum_{i=1}^n X_i = \text{E}_\theta(X_1) \equiv \theta^{1/\kappa} \Gamma\left(1 + \frac{1}{\kappa}\right);$$

this is $[\bar{X}_n / \Gamma(1 + \frac{1}{\kappa})]^\kappa$.

3. In Question 3, the MLE will be $\hat{\theta}_n = \bar{Y}_n \equiv \sum_{i=1}^n Y_i/n$, or the sample average of $Y_1 \equiv X_1^\kappa, \ldots, Y_n \equiv X_n^\kappa$ as a random sample of $\text{Exp}(\theta)$ variables.

4. For Question 4, the MLE (as an iid average) will be unbiased: $\text{E}_\theta(\bar{Y}_n) = \text{E}_\theta(Y_1) = \theta$.

5. In Question 5, $\hat{\theta}_n \xrightarrow{p} \text{E}_\theta(Y_1) = \theta$ holds by the WLLN applied to $\bar{Y}_n$. For the method of moments estimator, use $\bar{X}_n \xrightarrow{p} \text{E}_\theta(X_1) = \theta^{1/\kappa}\Gamma(1 + \frac{1}{\kappa})$ by the WLLN applied to $\bar{X}_n$; then, consistency follows by the continuous mapping theorem (i.e., $[\bar{X}_n/\Gamma(1 + \frac{1}{\kappa})]^\kappa$ is a continuous function of $\bar{X}_n$).

6. For Question 6, the main idea is that $\sum_{i=1}^n X_i^\kappa \equiv \sum_{i=1}^n Y_i$ is a complete & sufficient statistic for θ (using that the joint density of $X_1, \ldots, X_n$ is an exponential family). To find any UMVUE, it suffices to find an appropriate function of $\sum_{i=1}^n Y_i$ that is unbiased.

   This question would be easier if one had to find the UMVUE of $\eta = \theta^2$ (rather than $\eta = \theta^\kappa$). To find the UMVUE of $\eta = \theta^2$ (a parameter related to 2nd moments of Y_1), we would try 2nd moments of $\sum_{i=1}^n Y_i$:

$$\mathrm{E}_\theta\left(\sum_{i=1}^n Y_i\right)^2 = \mathrm{Var}_\theta\left(\sum_{i=1}^n Y_i\right) + \left(\mathrm{E}_\theta \sum_{i=1}^n Y_i\right)^2 = \sum_{i=1}^n \mathrm{Var}_\theta(Y_i) + (n\theta)^2 = n\theta^2 + n^2\theta^2.$$

   So then, $\left(\sum_{i=1}^n Y_i\right)^2/[n+n^2]$ would be the UMVUE of $\eta = \theta^2$.

   To find the UMVUE of $\eta = \theta^\kappa$, one would have to use that $Y_1, \ldots, Y_n$ are iid Exp(θ) so that $S \equiv \sum_{i=1}^n Y_i/\theta \sim Gamma(n,1)$. Moments of $\sum_{i=1}^n Y_i$ are then

$$\mathrm{E}_\theta\left(\sum_{i=1}^n Y_i\right)^r = \mathrm{E}_\theta \theta^r S^r = \theta^r \int_0^\infty s^r \cdot \frac{s^{n-1}e^{-s}}{\Gamma(n)} ds = \theta^r \frac{\Gamma(r+n)}{\Gamma(n)}, \quad r > 0;$$

   this requires having the $Gamma(n,1)$ density $(s^{n-1}e^{-s}/\Gamma(n))$ available for $S \equiv \sum_{i=1}^n Y_i/\theta$, which should be given here. So, the UMVUE of $\eta = \theta^\kappa$ is then $\left(\sum_{i=1}^n Y_i\right)^\kappa \Gamma(n)/\Gamma(\kappa+n)$.

7. Question 7 is incorrect. Because $\sum_{i=1}^n X_i^\kappa = \sum_{i=1}^n Y_i$ is a sufficient statistic for θ, the conditional distribution of $X_{(1)}^\kappa$ (or any statistic) given $\sum_{i=1}^n X_i^\kappa$ will not depend on θ, by definition. Consequently, the conditional expectation $\mathrm{E}_\theta(X_{(1)}^\kappa | \sum_{i=1}^n X_i^\kappa)$ cannot involve θ; that is, this is not a function of θ (or does not change with θ). Hence, the derivative of $\mathrm{E}_\theta(X_{(1)}^\kappa | \sum_{i=1}^n X_i^\kappa)$ with respect to θ must be 0 (not 1).

   The solution provided is incorrect because $X_{(1)}^\kappa/\theta$ is not a statistic; this is a pivotal quantity, but not an ancillary statistic (i.e., a statistic is computable solely from data while $X_{(1)}^\kappa/\theta$ involves the parameter θ). Also, $X_{(1)}^\kappa/\theta$ has the same distribution as the minimum of n iid standard Exponential variables so that $X_{(1)}^\kappa/\theta \sim \mathrm{Exp}(1/n)$ (not Exp(1)).

8. Question 8 uses the standard CLT for iid variables:

$$\sqrt{n}(\hat\theta_n - \theta) = \sqrt{n}(\bar Y_n - \mathrm{E}_\theta Y_1) \xrightarrow{d} N(0, \mathrm{Var}_\theta(Y_1) = \theta^2) \quad \text{as } n \to \infty.$$

   Because $\hat\theta_n = \bar Y_n \xrightarrow{p} \mathrm{E}_\theta(Y_1) = \theta > 0$ by Question 5, we apply Slutsky's theorem:

$$\frac{1}{\hat\theta_n}\sqrt{n}(\hat\theta_n - \theta) \xrightarrow{d} \frac{1}{\theta} \cdot N(0,\theta^2) \sim N(0,1) \quad \text{as } n \to \infty.$$

9. Question 9 should specify whether the interval intended is two-sided or not, but the interval is calibrated from $N(0,1)$ quantiles based on Question 8.

10. Questions 10 and 11 are largely the same. One uses that $Y_1, \ldots, Y_n$ are iid Exp(θ) so that $\sum_{i=1}^n Y_i/\theta \sim Gamma(n,1)$ is a pivot for purposes of testing (Question 10) or interval estimation (Question 11). Also, $2Y_i/\theta \sim \mathrm{Exp}(2) \sim \chi_2^2$ so that $2\sum_{i=1}^n Y_i/\theta \sim \chi_{2n}^2$ may be used, if one prefers.

11. Question 11 as above.

12. Question 12 is a bit ill-posed. The mean of the Inverse-Gamma Distribution ($IG(a,b)$) is required and should be given in the problem set-up as $b/(a-1)$ when $a > 1$.

1. In order to calculate the rth moments of the random variable X, we use the definition of Gamma function $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1}e^{-y}\,dy$ and change of variables as follows:

$$\begin{aligned}\mathbb{E}[X^r] &= \int_0^\infty \frac{\kappa}{\theta} x^{r+\kappa-1} e^{-(\frac{x^\kappa}{\theta})}\,dx \\ &= \theta^{\frac{r}{\kappa}} \int_0^\infty \kappa\, y^{\frac{r}{\kappa}} e^{-y}\,dy \quad \big(y := \frac{x^\kappa}{\theta}\big) \\ &= \theta^{\frac{r}{\kappa}}\,\Gamma\big(1+\frac{r}{\kappa}\big).\end{aligned}$$

2. Use problem **1** for $r=1$ to see $\mathbb{E}(X_1) = \theta^{\frac{1}{\kappa}}\,\Gamma(1+\frac{1}{\kappa})$ and obtain $\tilde{\theta}_n$ by solving $m_1 = \bar{X}_n = \theta^{\frac{1}{\kappa}}\,\Gamma(1+\frac{1}{\kappa})$. This yields $\tilde{\theta}_n = \left(\frac{\bar{X}_n}{\Gamma(1+\frac{1}{\kappa})}\right)^\kappa$.

3. The likelihood function is

$$L(\theta|x_1,\ldots,x_n) = \big(\frac{\kappa}{\theta}\big)^n \prod_{i=1}^n x_i^{(\kappa-1)}\ \exp\Big(-\frac{1}{\theta}\sum_{i=1}^n x_i^\kappa)\Big)$$

and hence the log-likelihood becomes

$$\log L(\theta|x_1,\ldots,x_n) = n\log\kappa - n\log\theta + (\kappa-1)\sum_{i=1}^n \log x_i - \frac{1}{\theta}\sum_{i=1}^n x_i^\kappa.$$

By taking the derivative with respect to θ and solving for $\theta = \hat{\theta}_n$, we have:

$$\frac{\partial}{\partial\theta}\log L(\theta|x_1,\ldots,x_n) = \frac{-n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^n x_i^\kappa\ |_{\theta=\hat{\theta}} = 0$$

and consequently we get $\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^n x_i^\kappa$. Therefore $\hat{\theta}_n$ maximize $\log L(\theta|x_1,\ldots,x_n)$ if $\frac{\partial^2}{\partial\theta^2}\log L(\theta|x_1,\ldots,x_n)|_{\theta=\hat{\theta}} < 0$. But $\frac{\partial^2}{\partial\theta^2}\log L(\theta|x_1,\ldots,x_n)|_{\theta=\hat{\theta}} = -\frac{n}{\hat{\theta}^2} < 0$ and this completes the proof.

4. We first note that $X_i^\kappa \sim \mathrm{Exp}(\theta)$ and consequently $\sum_{i=1}^n X_i^\kappa \sim \mathrm{Gamma}(n,\kappa)$. Therefore $\mathbb{E}[\hat{\theta}_n] = \mathbb{E}\Big(\frac{1}{n}\sum_{i=1}^n X_i^\kappa\Big) = \frac{n\theta}{n} = \theta$ which means $\hat{\theta}$ is an unbiased estimator for θ.

5. For $\tilde{\theta}_n$, the WLLN implies that $\bar{X}_n \xrightarrow{p} \mathbb{E}(X) = \theta^{\frac{1}{\kappa}}\,\Gamma(1+\frac{1}{\kappa})$. Define $g(x) = \big(\frac{x}{\Gamma(1+\frac{1}{\kappa})}\big)^\kappa$ and use the continuous mapping theorem to see $g(\bar{X}_n) \xrightarrow{p} g(\mathbb{E}(X)) = g(\theta^{\frac{1}{\kappa}}\,\Gamma(1+\frac{1}{\kappa}))$ as $n\to\infty$. Note that $g(\bar{X}_n) = \big(\frac{\bar{X}_n}{\Gamma(1+\frac{1}{\kappa})}\big)^\kappa = \tilde{\theta}_n$ and $g(\theta^{\frac{1}{\kappa}}\,\Gamma(1+\frac{1}{\kappa})) = \theta$. This means $\tilde{\theta}_n \xrightarrow{p} \theta$ and hence $\tilde{\theta}_n$ is consistent for θ. For $\hat{\theta}_n$, use the fact that $Y_i = X_i^\kappa \sim \mathrm{Exp}(\theta)$ for each $i = 1,\ldots,n$. Now, using WLLN, we have $\hat{\theta}_n = \bar{Y}_n \xrightarrow{p} \mathbb{E}(Y_1) = \theta$ which means $\hat{\theta}_n$ is consistent for θ.

6. We first note that

$$f(x_1,\ldots,x_n|\eta) = \left(\kappa\,\eta^{\frac{-1}{\kappa}}\right)^n \prod_{i=1}^{n} x_i{}^{\kappa-1} \exp\left(\frac{-1}{\eta^{\frac{1}{\kappa}}}\sum_{i=1}^{n} X_i^{\kappa}\right)$$

and $A \equiv \{\underset{\sim}{x} : f(\underset{\sim}{x}|\eta) > 0\}$ does not depend on η (here $\underset{\sim}{x} = (x_1,\ldots,x_n)$). Hence $f(\underset{\sim}{x}|\eta)$ belongs to the exponential family. Therefore, $T = \sum_{i=1}^{n} X_i^{\kappa}$ is complete and sufficient statistic for η. Next, we show that $\hat{\eta}$ is an unbiased estimator for η. In the following, we will use the fact that $Z := \sum_{i=1}^{n} X_i^{\kappa} \sim \text{Gamma}(n,\theta)$ and then we calculate $\mathbb{E}[Z^{\kappa}]$ to show that $\hat{\eta}$ is an unbiased estimator for η:

$$\begin{aligned}
\mathbb{E}[\hat{\eta}] &= \frac{\Gamma(n)}{\Gamma(n+k)}\mathbb{E}\Big[\Big(\sum_{i=1}^{n} X_i^{\kappa}\Big)^{\kappa}\Big] \\
&= \frac{\Gamma(n)}{\Gamma(n+k)}\mathbb{E}\big[Z^{\kappa}\big] \quad \big(Z \sim \text{Gamma}(n,\theta)\big) \\
&= \frac{\Gamma(n)}{\Gamma(n+k)}\int_0^{\infty} z^{\kappa}\,\frac{1}{\Gamma(n)\theta^n}\,z^n\,e^{-\frac{z}{\theta}}\,dz \\
&= \frac{\theta^{\kappa}}{\Gamma(n+k)}\int_0^{\infty} y^{n+\kappa-1}\,e^{-y}\,dy \quad (\frac{z}{\theta} = y) \\
&= \frac{\theta^{\kappa}\,\Gamma(n+k)}{\Gamma(n+k)} = \theta^{\kappa} = \eta
\end{aligned}$$

Finally, since $\hat{\eta}$ is an unbiased estimator of η and also it is a function of the complete and sufficient statistic $\sum_{i=1}^{n} X_i^{\kappa}$ then Lehmann- Scheffe Theorem implies that $\hat{\eta}$ is UMVUE for η.

7. The random variable $Y = \frac{(X_{(1)})^{\kappa}}{\theta} \sim \text{EXP}(1)$ and therefore Y is ancillary. On the other hand, $\sum_{i=1}^{n} X_i^{\kappa}$ is complete and sufficient statistic. Now, Basu's lemma implies that Y and $\sum_{i=1}^{n} X_i^{\kappa}$ are independent. That is $\mathbb{E}_{\theta}\Big[(X_{(1)})^{\kappa}\Big|\sum_{i=1}^{n} X_i^{\kappa}\Big] = \mathbb{E}_{\theta}\Big[(X_{(1)})^{\kappa}\Big] = \theta$. Now, taking the derivative of the both sides of the last equation with respect to θ gives the desired result.

8. By using the asymptotic properties of the MLEs, we have

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N\Big(0, \frac{1}{I_1(\theta)}\Big),$$

where

$$I_1(\theta) = -\mathbb{E}\Big[\frac{\partial^2}{\partial\theta^2}\log f(X|\theta)\Big] = \frac{1}{\theta^2}.$$

Define $g(x) = x^2$. According to WLLN, we have $\hat{\theta}_n \xrightarrow{p} \theta$ and then the continuous mapping theorem for function $g(x)$ implies that $g(\hat{\theta}_n) = \hat{\theta}_n^2 \xrightarrow{p} \theta^2$ as $n \to \infty$. Now by applying

Slutsky's theorem,

$$\frac{\sqrt{n}(\hat{\theta}_n - \theta)}{\hat{\theta}_n} \xrightarrow{d} \mathrm{N}\big(0, \frac{\theta^2}{\theta^2} = 1\big)$$

as $n \to \infty$. Therefore $\frac{\sqrt{n}(\hat{\theta}_n - \theta)}{\hat{\theta}_n^2}$ is an asymptotical pivot which means Q_n is also asymptotical pivot for $\gamma_1 = 0$ and $\gamma_2 = 1$.

9. From problem **8**, we know that $Q_n \to Q$ as $n \to \infty$ where the distribution of Q does not depend on θ. In fact, $Q \sim \mathrm{N}(0, 1)$ and we know that $\mathbb{P}\Big(z_{\frac{\alpha}{2}} \leq Z \leq z_{1-\frac{\alpha}{2}}\Big) = 1 - \alpha$. Therefore for large n,

$$\begin{aligned} C_{\underset{\sim}{X}} &= \Big\{\theta > 0, z_{\frac{\alpha}{2}} \leq Q_n(X_1, \dots, X_n, \theta) \leq z_{1-\frac{\alpha}{2}}\Big\} \\ &= \Big[\hat{\theta}_n - \frac{z_{1-\frac{\alpha}{2}}\hat{\theta}_n}{\sqrt{n}}, \hat{\theta}_n - \frac{z_{\frac{\alpha}{2}}\hat{\theta}_n}{\sqrt{n}}\Big] \end{aligned}$$

is a confidence interval for θ with approximate confidence coefficient $1 - \alpha$.

10. First, note that $f(\underset{\sim}{x}|\theta)$ have MLR in the real-valued statistic $T = t(\underset{\sim}{X}) = \sum_{i=1}^n X_i^\kappa$. In fact, for any $\theta_2 > \theta_1$, we have,

$$\begin{aligned} h(\underset{\sim}{x}) := \frac{f(\underset{\sim}{x}|\theta_2)}{f(\underset{\sim}{x}|\theta_1)} &= \frac{\theta_1}{\theta_2} \exp \sum_{i=1}^n x_i^\kappa \Big[\frac{1}{\theta_1} - \frac{1}{\theta_2}\Big] \\ &= g_{\theta_1,\theta_2}\big(t(\underset{\sim}{x})\big), \end{aligned}$$

where $g_{\theta_1,\theta_2}\big(t(\underset{\sim}{x})\big) = \frac{\theta_1}{\theta_2} \exp\Big\{t[\frac{1}{\theta_1} - \frac{1}{\theta_2}]\Big\}$ is non-decreasing as $t \to \infty$ (Note that this is true when κ is the known parameter). Now, apply Karrlin-Rubin Theorem to find the UMP test of $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$ based on $X_1, \dots, X_n$ as follows:

$$\phi(\underset{\sim}{X}) = \begin{cases} 1, & \text{if } \sum_{i=1}^n X_i^\kappa > c, \\ 0, & \text{otherwise.} \end{cases}$$

In order to have the test of size α, we need to find c. Write

$$\alpha = \mathbb{E}_{\theta_0}\big[\phi(\underset{\sim}{X})\big] = \mathbb{P}_{\theta_0}\Big(\sum_{i=1}^n X_i^\kappa > c\Big) = 1 - \mathbb{P}_{\theta_0}\Big(\frac{2\sum_{i=1}^n X_i^\kappa}{\theta_0} \leq \frac{2c}{\theta_0}\Big).$$

But, $\frac{2\sum_{i=1}^n X_i^\kappa}{\theta_0} \sim \chi^2_{2n}$ and hence $\frac{2c}{\theta_0} = \chi^2_{2n,1-\alpha}$. Therefore, the UMP test of size α has the form:

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

11. We know that $T = \sum_{i=1}^n X_i^\kappa \sim \text{Gamma}(n,\theta)$ and then $\frac{2T}{\theta} \sim \chi^2_{2n}$. Given t, $F(t|\theta) := \mathbb{P}(T \le t|\theta) = \mathbb{P}(\frac{2T}{\theta} \le \frac{2t}{\theta}|\theta) = \mathbb{P}(\chi^2_{2n} \le \frac{2t}{\theta}|\theta) = G(\frac{2t}{\theta})$, where $G(\cdot)$ is the cdf of χ^2_{2n}. We see that $F(t|\theta) = G(\frac{2t}{\theta}) \downarrow$ as $\theta \uparrow$ and hence by applying the tail theorem we have:

$$\frac{\alpha}{2} = \mathbb{P}(T \le t|\theta_U(t)) = G(\frac{2t}{\theta_U(t)})$$

and

$$1 - \frac{\alpha}{2} = \mathbb{P}(T \le t|\theta_L(t)) = G(\frac{2t}{\theta_L(t)})$$

and then $\big(\theta_L(t), \theta_U(t)\big) = \Big(\frac{2t}{\chi^2_{2n,\frac{1-\alpha}{2}}}, \frac{2t}{\chi^2_{2n,\frac{\alpha}{2}}}\Big)$ for any given value t of T. Hence

$$(\theta_L(T), \theta_U(T)) = \left(\frac{2T}{\chi^2_{2n,\frac{1-\alpha}{2}}}, \frac{2T}{\chi^2_{2n,\frac{\alpha}{2}}}\right) = \left(\frac{2\sum_{i=1}^n X_i^\kappa}{\chi^2_{2n,\frac{1-\alpha}{2}}}, \frac{2\sum_{i=1}^n X_i^\kappa}{\chi^2_{2n,\frac{\alpha}{2}}}\right)$$

is a confidence interval for θ with confidence coefficient $(1-\alpha)$.

12. The posterior distribution of θ is $\text{IG}(n+a, \sum_{i=1}^n x_i^\kappa + b)$. To see this write:

$$\begin{aligned} f(\theta|\underset{\sim}{x})(\theta) &= \pi(\theta) f(\underset{\sim}{x}|\theta) \\ &\propto \frac{1}{\theta^n}\, e^{\frac{-\sum_{i=1}^n x_i^\kappa}{\theta}}\, \theta^{-a-1}\, e^{-\frac{b}{\theta}} \\ &= \theta^{-(n+a)-1}\, e^{-\frac{1}{\theta}\left(\sum_{i=1}^n x_i^\kappa + b\right)}, \end{aligned}$$

where the last one is proportional to the density function of $\text{IG}(n+a, \sum_{i=1}^n x_i^\kappa + b)$. Now, the Bayes estimator of θ is given by

$$\mathbb{E}_{\theta|\underset{\sim}{x}}(\theta) = \frac{\sum_{i=1}^n x_i^\kappa + b}{n+a-1},$$

where we used the fact that $\mathbb{E}[X] = \frac{b}{a-1}$ if $X \sim \text{IG}(a,b)$ to obtain the Bayes estimator.

13. We reject H_0 if $\mathbb{P}(\theta \le \theta_0|\underset{\sim}{x}) < \frac{1}{2}$. Note that $\theta|\underset{\sim}{x} \sim \text{IG}(n+a, \sum_{i=1}^n X_i^\kappa + b)$ and then $\frac{1}{\theta} \sim \text{Gamma}(n+a, \frac{1}{\sum_{i=1}^n X_i^\kappa + b})$. Therefore, we have

$$\begin{aligned} \mathbb{P}(\theta \le \theta_0|\underset{\sim}{x}) &= \mathbb{P}\Big(\frac{1}{\theta} \ge \frac{1}{\theta_0}\Big|\underset{\sim}{x}\Big) \\ &= \mathbb{P}\Big(\frac{2\big(\sum_{i=1}^n X_i^\kappa + b\big)}{\theta} \ge \frac{2\big(\sum_{i=1}^n X_i^\kappa + b\big)}{\theta_0}\Big|\underset{\sim}{x}\Big) \\ &= \mathbb{P}\Big(\chi^2_{2(n+a)} \ge \frac{2\big(\sum_{i=1}^n X_i^\kappa + b\big)}{\theta_0}\Big|\underset{\sim}{x}\Big) < \frac{1}{2}. \end{aligned}$$

Consequently, we reject the null hypothesis if $\frac{2\left(\sum_{i=1}^{n} X_i^{\kappa}+b\right)}{\theta_0} > \chi^2_{2(n+a),0.5}$. Hence we have the following Bayesian test:

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

14. From Problems **12** and **13** we know that $\theta|\underset{\sim}{x} \sim \mathrm{IG}(n+a, \sum_{i=1}^{n} x_i^{\kappa}+b)$ and $\frac{1}{\theta} \sim \mathrm{Gamma}(n+a, \frac{1}{\sum_{i=1}^{n} X_i^{\kappa}+b})$. Then $\frac{2\left(\sum_{i=1}^{n} x_i^{\kappa}+b\right)}{\theta} \sim \chi^2_{2(n+a)}$. So,

$$1-\alpha = \mathbb{P}\big(L \leq \theta \leq U\big) = \mathbb{P}\big(\frac{1}{U} \leq \frac{1}{\theta} \leq \frac{1}{L}\big) = \mathbb{P}\Big(\frac{2\big(\sum_{i=1}^{n} x_i^{\kappa} + b\big)}{U} \leq \frac{1}{\theta} \leq \frac{2\big(\sum_{i=1}^{n} x_i^{\kappa} + b\big)}{L}\Big)$$

for any given value $\sum_{i=1}^{n} x_i^{\kappa}$. Thus

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

is a $(1-\alpha)$ credible set.

Part I

Let Ω be a non-empty set and $(\Omega, \mathcal{F})$ be a measurable space.

1. Let Ω' be a non-empty subset of Ω. Define

$$\mathcal{F}_{\Omega'} = \{A \cap \Omega' : A \in \mathcal{F}\}.$$

Show that $\mathcal{F}_{\Omega'}$ is a σ-algebra of subsets of Ω'.

2. Define a π-class with subsets of Ω.

3. Define a λ-system with subsets of Ω.

4. If a λ-system is also a π-class, show then it is a σ-algebra.

5. Let P_1 and P_2 be two probability measures on $(\Omega, \mathcal{F})$. Show that the class

$$\mathcal{L} \equiv \{A \in \mathcal{F} : P_1(A) = P_2(A)\}$$

is a λ-system.

Part II

6. Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\{X_n\}_{n \geq 1}$ be a monotone sequence of random variables and X be another random variable, all defined on $(\Omega, \mathcal{F}, P)$. Assume that $X_n \xrightarrow{P} X$ as $n \to \infty$. Show that $X_n \to X$ almost surely as $n \to \infty$.

Part III

Let $X_1 \sim$ Poisson (λ), $X_2 \sim$ Exponential (β), and $Y \sim$ Bernoulli (α), for some known $\beta, \lambda > 0$ and $0 < \alpha < 1$. Assume that X_1, X_2, and Y are independent. Define the random variable

$$Z = YX_1 + (1 - Y)X_2.$$

Denote the cdfs of X_1, X_2 and Z by F_1, F_2 and F respectively. Denote the Lebesgue-Stieltjes measures induced by F_1, F_2, F as μ_{F_1}, μ_{F_2} and μ_F respectively. For two measures μ and ν we write $\mu << \nu$ if μ is dominated by ν.

7. Write down F in terms of F_1 and F_2.

8. Prove whether or not the following dominations hold. In each case, if the domination holds then provide the corresponding Radon-Nikodym derivative.

a) $\mu_{F_1} << \mu_F$

b) $\mu_F << \mu_{F_1}$

c) $\mu_{F_2} << \mu_F$

d) $\mu_F << \mu_{F_2}$

Part IV

Let X_λ be a Poisson random variable with mean λ.

9. Show that the characteristic function of X_λ is given by $\phi_{X_\lambda}(t) = \exp[\lambda\{\exp(it) - 1\}]$.

10. State Lévy's continuity theorem.

11. Define

$$Y_\lambda = \frac{X_\lambda - \lambda}{\sqrt{\lambda}}.$$

Show that as $\lambda \to \infty$, Y_λ converges in distribution to $N(0, 1)$.

12. Consider a sequence of random variables $X_{\lambda_n}, n \geq 1$. Let X_{λ_n} be a Poisson random variable with mean λ_n and suppose $\lambda_n \to \lambda_0 < \infty$ as $n \to \infty$. Show that

$$\sum_{x=0}^{\infty} |P(X_{\lambda_n} = x) - P(X_0 = x)| \to 0, \text{ as } n \to \infty,$$

where X_0 is a Poisson random variable with mean λ_0.

Part V

Let $\{X_n, n \geq 1\}$ be independent random variables on a probability space $(\Omega, \mathcal{F}, P)$ with

$$P(X_n = n^{-\alpha}) = \frac{1}{2} = P(X_n = -n^{-\alpha})$$

for all $n \geq 1$.

13. Show that if $\alpha > 1/2$ then $S_n = \sum_{j=1}^{n} X_j$ converges almost surely.

14. Verify that $\alpha > 1/2$ is necessary for convergence (almost surely or in probability) of S_n.

Part VI

Let $\{X_{nj} : 1 \leq j \leq n\}_{n\geq 1}$ be a triangular array of random variables such that $X_{n1}, X_{n2}, \ldots, X_{nn}$ are independent, $E(X_{nj}) = 0$, and $E(X_{nj}^2) < \infty$ for all $j = 1, \ldots, n$.

15. State the Lindeberg condition for $\{X_{nj} : 1 \leq j \leq n\}_{n\geq 1}$.

16. In addition to the above assumptions for the triangular array $\{X_{nj} : 1 \leq j \leq n\}_{n\geq 1}$, assume that $\sum_{j=1}^{n} E(X_{nj}^2) = 1$ for all $n \geq 1$.

a) Show that

$$\lim_{n\to\infty} \max_{1\leq j\leq n} E(X_{nj}^2) = 0.$$

b) Show that

$$\max_{1\leq j\leq n} |X_{nj}| \xrightarrow{P} 0.$$

17. Let $s_n^2 = \text{Var}(\sum_{j=1}^{n} X_{nj})$. Assume that there exists a $\delta > 0$ such that as $n \to \infty$

$$\frac{\sum_{j=1}^{n} E|X_{nj}|^{2+\delta}}{s_n^{2+\delta}} \to 0.$$

Show that $\{X_{nj} : 1 \leq j \leq n\}_{n\geq 1}$ satisfies the Lindeberg condition.

1. a) Since $\Omega \in \mathcal{F}$, $\Omega' = \Omega \cap \Omega' \in \mathcal{F}_{\Omega'}$. b) Let $B \in \mathcal{F}_{\Omega'}$. To show $\Omega' \setminus B \in \mathcal{F}_{\Omega'}$. Since $B \in \mathcal{F}_{\Omega'}$, $B = A \cap \Omega'$ for some $A \in \mathcal{F}$. Thus $\Omega' \setminus B = A^c \cap \Omega' \in \mathcal{F}_{\Omega'}$ since $A^c \in \mathcal{F}$. c) Finally, let $B_n \in \mathcal{F}_{\Omega'}$ for all $n \geq 1$. To show $\cup_{n\geq 1} B_n \in \mathcal{F}_{\Omega'}$. Since $B_n \in \mathcal{F}_{\Omega'}$, $B_n = A_n \cap \Omega'$ for some $A_n \in \mathcal{F}$ and $\cup_{n\geq 1} B_n = \cup_{n\geq 1} A_n \cap \Omega'$. Thus $\cup_{n\geq 1} B_n \in \mathcal{F}_{\Omega'}$ as $\cup_{n\geq 1} A_n \in \mathcal{F}$.

2. A non empty class $\mathcal{C}$ of subsets of Ω is called a $\pi-$class if

$$A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}.$$

3. A class $\mathcal{L}$ of subsets of Ω is called a $\lambda-$ system if

$$\text{(i) } \Omega \in \mathcal{L} \qquad \text{(ii) } A \in \mathcal{L} \Rightarrow A^c \in \mathcal{L}$$

$$\text{\& (iii) for any sequence } A_1, A_2, \ldots \text{ of disjoint sets in } \mathcal{L},\ \bigcup_{n=1}^{\infty} A_n \in \mathcal{L}$$

4. Let $\mathcal{L}$ be a $\lambda-$ system. To show that it is also a σ-algebra, we need to show it is closed under countable union of not necessarily disjoint sets. Let $A_n \in \mathcal{L}$ for all $n \geq 1$. Define $B_1 = A_1, B_2 = A_2 \cap A_1^c, B_3 = A_3 \cap A_2^c \cap A_1^c$ so on. Since $\mathcal{L}$ is a π-class, $B_n \in \mathcal{L}$ for $n \geq 2$. Since B_n's are disjoint, $\cup_{n\geq 1} B_n \in \mathcal{L}$. So $\cup_{n\geq 1} A_n = \cup_{n\geq 1} B_n \in \mathcal{L}$.

5. Note that $\Omega \in \mathcal{L}$ since $P_1(\Omega) = 1 = P_2(\Omega)$. If $A \in \mathcal{L}$, $P_1(A^c) = 1 - P_1(A) = 1 - P_2(A) = P_2(A^c)$. Thus $A^c \in \mathcal{L}$. Finally, let $A_n, n \geq 1$ are disjoint and $A_n \in \mathcal{L}$ for $n \geq 1$. Then $P_1(\cup_n A_n) = \sum_n P_1(A_n) = \sum_n P_2(A_n) = P_2(\cup_n A_n)$. Thus $\cup_n A_n \in \mathcal{L}$. Thus $\mathcal{L}$ is a $\lambda-$ system.

6. Fix $\epsilon > 0$. Since $X_n \xrightarrow{P} X$ as $n \to \infty$, $\lim_{n\to\infty} P(|X_n - X| > \epsilon) = 0$. Since $\{X_n\}_{n\geq 1}$ is a monotone sequence of random variables,

$$\begin{aligned} \lim_{n\to\infty} P(|X_n - X| > \epsilon) &= \lim_{N\to\infty} P(\cup_{n\geq N} |X_n - X| > \epsilon) \\ &= P(\limsup_{N\to\infty} |X_n - X| > \epsilon) = P(|X_n - X| > \epsilon \ i.o.). \end{aligned}$$

Thus $X_n \to X$ almost surely.

7. Note that $F(z) = P(Z \leq z) = E[P(Z \leq z|Y)] = \alpha F_1(z) + (1-\alpha) F_2(z)$.

8.
- **a)** For any $A \in \mathcal{B}(\mathbb{R})$, $\mu_{F_1}(A) \leq \frac{1}{\alpha}\mu_F(A)$. Thus $\mu_{F_1} << \mu_F$. Denote $\mathbb{N} = \{0, 1, 2, \ldots\}$. Then $\frac{d\mu_{F_1}}{d\mu_F}(x) = \frac{1}{\alpha} I_N(x)$ a.e. (μ_F).
- **b)** Let Q denote the set of rationals. Since $\mu_F(Q^c) = 1 - \alpha, \mu_{F_1}(Q^c) = 0$, $\mu_F \not<< \mu_{F_1}$.
- **c)** For any $A \in \mathcal{B}(\mathbb{R})$, $\mu_{F_2}(A) \leq \frac{1}{1-\alpha}\mu_F(A)$. Thus $\mu_{F_2} << \mu_F$ and $\frac{d\mu_{F_2}}{d\mu_F} = \frac{1}{1-\alpha} I_{N^c}(x)$ a.e. (μ_F).
- **d)** Since $\mu_F(Q) = \alpha, \mu_{F_2}(Q) = 0$, $\mu_F \not<< \mu_{F_2}$.

9. The characteristic function of X_λ is given by

$$\begin{aligned}\phi_{X_\lambda}(t) = E(\exp(itX_\lambda)) &= \sum_{x=0}^{\infty} \exp(itx)\exp(-\lambda)\frac{\lambda^x}{x!} \\ &= \exp(-\lambda)\exp[\lambda\exp(it)] = \exp[\lambda\{\exp(it)-1\}].\end{aligned}$$

10. Suppose $X_n, n \geq 0$ is a sequence of random variables, each with characteristic function ϕ_{X_n}, $n \geq 0$.

(a) If $X_n \xrightarrow{d} X_0$, then for any $T > 0$,

$$\sup_{|t|\leq T} \left|\phi_{X_n}(t) - \phi_{X_0}(t)\right| \to 0 \quad \text{as } n \to \infty.$$

(b) If $\phi_{X_n}(t) \to g(t)$ as $n \to \infty$ for all $t \in \mathbb{R}$ and the function $g(\cdot)$ is continuous at zero, then $g(\cdot)$ is a characteristic function and $X_n \xrightarrow{d} X_0$, where X_0 is the random variable with characteristic function $g(\cdot)$.

11. The characteristic function of Y_λ is

$$\phi_{Y_\lambda}(t) = E(\exp(itY_\lambda)) = \exp[\lambda\{\exp(i\frac{t}{\sqrt{\lambda}}) - 1\} - it\sqrt{\lambda}]$$

Consider the Taylor series expansion of $\log\phi_{Y_\lambda}(t)$:

$$\begin{aligned}\log\phi_{Y_\lambda}(t) &= \lambda\sum_{k=0}^{\infty}\frac{(it)^k}{(\sqrt{\lambda})^k k!} - \lambda - it\sqrt{\lambda} \\ &= -\frac{t^2}{2} - \frac{it^3}{\sqrt{\lambda}6} + \ldots.\end{aligned}$$

Thus $\log\phi_{Y_\lambda}(t) \to -t^2/2$ as $\lambda \to \infty$ for all $t \in \mathbb{R}$. Thus by the Lévy's Continuity theorem as $\lambda \to \infty$, Y_λ converges in distribution to $N(0,1)$.

12. Since $f_{X_{\lambda_n}}(x) = \exp(-\lambda_n)\frac{\lambda_n^x}{x!} \to \exp(-\lambda_0)\frac{\lambda_0^x}{x!} = f_{X_{\lambda_0}}(x)$ for all $x = 0, 1, 2, \ldots$, by Scheffé's Theorem we have the result.

13. Note that $E(X_n) = 0$ and $\sum_{n=1}^{\infty} E(X_n^2) = \sum_{n=1}^{\infty}\frac{1}{n^{2\alpha}} < \infty$ for $\alpha > 1/2$. By the Khintchine-Kolmogorov convergence theorem, $S_n = \sum_{j=1}^{n} X_j$ converges almost surely.

14. Let $X_n^{(1)} \equiv X_n\mathbb{I}(|X_n| \leq 1)$. Since $\sum_{n=1}^{\infty} P(|X_n| > 1) = 0$, $E(X_n^{(1)}) = 0$ for all $n \geq 1$, $\sum_{n=1}^{\infty}\text{var}(X_n^{(1)}) = \sum_{n=1}^{\infty}\frac{1}{n^{2\alpha}} < \infty$ for $\alpha > 1/2$, by Kolmogorov's 3-Series Theorem, $\alpha > 1/2$ is necessary for convergence (almost surely or in probability) of S_n.

15. Let $\{X_{nj} : 1 \le j \le n\}_{n=1}^{\infty}$ be a triangular array with

$$\mathrm{EX_{nj}} = 0, \qquad 0 < \mathrm{EX_{nj}^2} = \sigma_{\mathrm{nj}}^2 < \infty, \qquad \mathrm{v_n^2} = \sum_{\mathrm{j=1}}^{\mathrm{n}} \sigma_{\mathrm{nj}}^2.$$

Then, we say $\{X_{nj}\}$ satisfies the Lindeberg Condition if, for all $\epsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{v_n^2} \sum_{j=1}^{n} \mathrm{EX_{nj}^2} \mathbb{I}(|\mathrm{X_{nj}}| > \epsilon \mathrm{v_n}) = 0.$$

16. **a**) Fix $\epsilon > 0$. Since $v_n = 1$

$$\begin{aligned} E(X_{nj}^2) &= E(X_{nj}^2 I(|X_{n,j}| \le \epsilon)) + E(X_{nj}^2 I(|X_{n,j}| > \epsilon)) \\ &\le \epsilon^2 + \frac{1}{v_n^2} \sum_{j=1}^{n} E(X_{nj}^2 I(|X_{n,j}| > \epsilon v_n)) \\ &\to \epsilon^2 \text{ as } n \to \infty. \end{aligned}$$

Thus for all $\epsilon > 0$

$$\lim_{n\to\infty} \max_{1\le j\le n} E(X_{nj}^2) \le \epsilon^2.$$

The proof follows by letting $\epsilon \downarrow 0$.

b) Since $v_n = 1$,

$$\begin{aligned} P(\max_{1\le j\le n} |X_{nj}| > \epsilon) &\le P(\cup_{j=1}^{n} |X_{nj}| > \epsilon) \\ &\le \sum_{j=1}^{n} P(|X_{nj}| > \epsilon) \\ &\le \sum_{j=1}^{n} E(\frac{|X_{nj}|^2}{\epsilon^2} I(|X_{nj}| > \epsilon)) \\ &= \frac{1}{\epsilon^2 v_n^2} \sum_{j=1}^{n} E(|X_{nj}|^2 I(|X_{nj}| > \epsilon v_n)) \to 0 \text{ as } n \to \infty. \end{aligned}$$

17.

$$\begin{aligned} \frac{1}{v_n^2} \sum_{j=1}^{n} \mathrm{E}[\mathrm{X_{nj}^2} \mathbb{I}(|\mathrm{X_{nj}}| > \epsilon \mathrm{v_n})] &\le \frac{1}{v_n^2} \sum_{j=1}^{n} [\mathrm{EX_{nj}^2} \frac{|\mathrm{X_{nj}}|^\delta}{\epsilon^\delta \mathrm{v_n^\delta}} \mathbb{I}(|\mathrm{X_{nj}}| > \epsilon \mathrm{v_n})] \\ &\le \frac{1}{v_n^{2+\delta} \epsilon^\delta} \sum_{j=1}^{n} [\mathrm{E}|\mathrm{X_{nj}}|^{2+\delta} \mathbb{I}(|\mathrm{X_{nj}}| > \epsilon \mathrm{v_n})] \\ &\to 0 \text{ as } n \to \infty. \end{aligned}$$

Thus the Lindeberg condition holds.

Part I

In what follows, you will need the mean and variance of the truncated normal random variable. Let $X \sim \mathcal{N}(\mu, \sigma^2)$, $\phi(\cdot)$ be the standard normal pdf and $\Phi(\cdot)$ be the standard normal cdf. If we condition on X being larger than some constant a, then the variance of the truncated normal random variable is

$$\mathrm{Var}(X \mid X > a) = \sigma^2 \left[1 + \left(\frac{a-\mu}{\sigma}\right) \frac{\phi\left(\frac{a-\mu}{\sigma}\right)}{1-\Phi\left(\frac{a-\mu}{\sigma}\right)} - \left(\frac{\phi\left(\frac{a-\mu}{\sigma}\right)}{1-\Phi\left(\frac{a-\mu}{\sigma}\right)} \right)^2 \right].$$

1. Show that the mean of the truncated normal random variable is

$$E[X \mid X > a] = \mu + \frac{\sigma\phi\left(\frac{a-\mu}{\sigma}\right)}{1-\Phi\left(\frac{a-\mu}{\sigma}\right)}.$$

Part II

Consider independent flips of a coin and focus on the possibility, ϵ, that "the coin lands heads up." Define

$$u_n = \Pr(\epsilon \text{ occurs at trial } n)$$

and

$$f_n = \Pr(\epsilon \textit{ first} \text{ occurs at trial } n),$$

and take $u_0 \equiv 1$ and $f_0 \equiv 0$.

2. If $p \in [0, 1]$ is the probability of heads, what is u_n?

3. What is f_n?

4. Define the probability generating function (pgf) $F(s) = \sum_{k=0}^{\infty} f_k s^k$, $|s| \le 1$ for the wait time, T, until the first occurrence of ϵ. Show that

$$F(s) = \frac{ps}{1-(1-p)s}.$$

5. Use $F(s)$ to show that $E[T] = \frac{1}{p}$.

Suppose you are able to collect data on the waiting time for the first head in independent coin flipping trials. Unfortunately, your data recording device is broken so it only registers a value when the first head occurs on the first or second trial. Thus $T_1, T_2, \ldots, T_n \in \{1, 2\}$ are iid random variables with the distribution of **Problem 3** truncated above at 2. (Truncation means you observe *no result* when the first head occurs at trial 3 or beyond. All recorded observations satisfy $T_i \le 2$.) **Problems 6–10** are based on the data $T_1, T_2, \ldots, T_n$.

6. Find the MLE $\hat{p}$ based on data $T_1, T_2, \ldots, T_n$.

7. Verify that the method-of-moments estimator (MME) based on data $T_1, T_2, \ldots, T_n$, say $\tilde{p}$, is the same as the MLE $\hat{p}$ when $\tilde{p} \ge 0$.

8. What is the Cramér-Rao lower bound (call it $v^*(p)$) on the variance for unbiased estimators of p based on $T_1, T_2, \ldots, T_n$?

9. Use the delta method and the above results to obtain approximations for $E[\hat{p}]$ and $\text{Var}(\hat{p})$.

10. Verify that $\sqrt{n}\,(\hat{p}-p)\,/\sqrt{v^*(p)} \xrightarrow{\text{d}} \mathcal{N}(0,1)$. (In this sense, the MLE achieves the Cramér-Rao lower bound in the limit.)

Part III

Now consider the possibility, ϵ_k, of k consecutive heads. Disallow overlaps, so in repeated trials, a subsequent occurrence of ϵ_k cannot build on a previous occurrence of ϵ_k. Letting E_i denote the outcome of the ith coin flip, if $E_{m-1}=0, E_m=1, E_{m+1}=1$, and $E_{m+2}=1$, with 1 indicating heads, then ϵ_2 occurs at $m+1$ *but not* at $m+2$. Assume, as usual, that trials are independent and the same coin is used for all flips.

11. Let V_k be the time of the first occurrence of ϵ_k. Prove that

$$E[V_k] = 1 + E[V_{k-1}] + (1-p)E[V_k].$$

and use the result to show that $E[V_2] = \frac{p+1}{p^2}$.

12. Redefine u_n and f_n from Part II now for the possibility, ϵ_k, of k consecutive heads. From these probabilities, define $U(s) = \sum_{k=0}^{\infty} u_k s^k$ (which is *not* a pgf) and the pgf $F(s)$ as in **Problem 4**. Show that

$$U(s) = \frac{1}{1-F(s)}.$$

13. Show that if ϵ_2 is the possibility of two heads in a row, then

$$F(s) = \frac{p^2 s^2}{p^2 s^2 + (1-s)(1+sp)}.$$

14. Find the variance of V_2.

15. Provide an asymptotic formula for an approximate probability that the 38th non-overlapping occurrence of two heads in independent flips of a fair coin occurs after the 178th flip.

Part I

In what follows, you will need the mean and variance of the truncated normal random variable. Let $X \sim \mathcal{N}(\mu, \sigma^2)$, $\phi(\cdot)$ be the standard normal pdf and $\Phi(\cdot)$ be the standard normal cdf. If we condition on X being larger than some constant a, then the variance of the truncated normal random variable is

$$\mathrm{Var}(X \mid X > a) = \sigma^2 \left[1 + \left(\frac{a-\mu}{\sigma}\right)\frac{\phi\left(\frac{a-\mu}{\sigma}\right)}{1-\Phi\left(\frac{a-\mu}{\sigma}\right)} - \left(\frac{\phi\left(\frac{a-\mu}{\sigma}\right)}{1-\Phi\left(\frac{a-\mu}{\sigma}\right)}\right)^2\right].$$

1. Show that the mean of the truncated normal random variable is

$$E[X \mid X > a] = \mu + \frac{\sigma\phi\left(\frac{a-\mu}{\sigma}\right)}{1-\Phi\left(\frac{a-\mu}{\sigma}\right)}.$$

Solution: In the next two solutions, I use $\phi(\cdot;\mu,\sigma)$ and $\Phi(\cdot;\mu,\sigma)$ to represent the pdf and cdf of $\mathcal{N}(\mu,\sigma^2)$ random variables. Let $C = \frac{1}{1-\Phi(a;\mu,\sigma)}$.

$$\begin{aligned}
E[X \mid X > a] &= \frac{1}{C}\int_a^\infty y\phi(y;\mu,\sigma)dy \\
&= \frac{1}{C}\left[\sigma\int_a^\infty \frac{y-\mu}{\sigma}\phi(y;\mu,\sigma)dy + \mu\int_a^\infty \phi(y;\mu,\sigma)dy\right] \\
&= \frac{1}{C}\left[\sigma\int_{\frac{a-\mu}{\sigma}}^\infty \frac{ze^{-z^2/2}}{\sqrt{2\pi}}dz + \mu C\right] \\
&= \frac{\sigma}{C\sqrt{2\pi}}\int_{\frac{(a-\mu)^2}{2\sigma^2}}^\infty e^{-u}du + \mu \\
&= \frac{\sigma e^{-\frac{(a-\mu)^2}{2\sigma^2}}}{C\sqrt{2\pi}} + \mu = \frac{\sigma\phi\left(\frac{a-\mu}{\sigma}\right)}{1-\Phi\left(\frac{a-\mu}{\sigma}\right)} + \mu.
\end{aligned}$$

Part II

Consider independent flips of a coin and focus on the possibility, ϵ, that "the coin lands heads up." Define

$$u_n = \Pr(\epsilon \text{ occurs at trial } n)$$

and

$$f_n = \Pr(\epsilon \textit{ first} \text{ occurs at trial } n),$$

and take $u_0 \equiv 1$ and $f_0 \equiv 0$.

2. If $p \in [0,1]$ is the probability of heads, what is u_n?

Solution:

$$u_n = p$$

3. What is f_n?

Solution:

$$f_n = (1-p)^{n-1}p$$

4. Define the probability generating function (pgf) $F(s) = \sum_{k=0}^{\infty} f_k s^k$, $|s| \leq 1$ for the wait time, T, until the first occurrence of ϵ. Show that

$$F(s) = \frac{ps}{1-(1-p)s}.$$

Solution:

$$\begin{array}{rcll} F(s) & = & \sum_{k=0}^{\infty} f_k s^k & \text{definition of pgf} \\ & = & \sum_{k=1}^{\infty} f_k s^k & f_0 = 0 \\ & = & \sum_{k=1}^{\infty} (1-p)^{k-1} p s^k & \text{definition of } f_k \\ & = & ps \sum_{k=1}^{\infty} [(1-p)s]^{k-1} & \text{constant factor} \\ & = & ps \sum_{k=0}^{\infty} [(1-p)s]^{k} & \text{change of variable } k \to k-1 \\ & = & \frac{ps}{1-(1-p)s} & \text{geometric series for } (1-p)|s| < 1. \end{array}$$

5. Use $F(s)$ to show that $E[T] = \frac{1}{p}$.

Solution:

$$\begin{array}{rcl} E[T] & = & \left. \dfrac{dF(s)}{ds} \right|_{s=1} \\ & = & \left. \left(\dfrac{p}{1-(1-p)s} + \dfrac{ps(1-p)}{[1-(1-p)s]^2} \right) \right|_{s=1} \\ & = & \dfrac{p}{p} + \dfrac{p-p^2}{p^2} = \dfrac{1}{p}. \end{array}$$

Suppose you are able to collect data on the waiting time for the first head in independent coin flipping trials. Unfortunately, your data recording device is broken so it only registers a value when the first head occurs on the first or second trial. Thus $T_1, T_2, \ldots, T_n \in \{1, 2\}$ are iid random variables with the distribution of **Problem 3** truncated above at 2. (Truncation means you observe *no result* when the first head occurs at trial 3 or beyond. All recorded observations satisfy $T_i \le 2$.) **Problems 6–10** are based on the data $T_1, T_2, \ldots, T_n$.

6. Find the MLE $\hat{p}$ based on data $T_1, T_2, \ldots, T_n$.

Solution: For $p \in (0, 1)$ the truncated probability mass function is

$$P(T_i = t \mid T_i \le 2) = \frac{\mathbb{1}\{t \le 2\}(1-p)^{t-1}p}{1-(1-p)^2} = \frac{\mathbb{1}\{t \le 2\}(1-p)^{t-1}}{2-p},$$

and the likelihood and log likelihood are

$$\begin{aligned} L(p \mid t_1, t_2, \ldots, t_n) &= \prod_{i=1}^{n} \left[\frac{(1-p)^{t_i-1}}{2-p}\right] \\ l(p \mid \boldsymbol{t}) &= n\left[(\bar{t}-1)\ln(1-p) - \ln(2-p)\right]. \end{aligned}$$

Taking derivatives and solving the score function for p', we find

$$\begin{aligned} \frac{dl(p \mid \boldsymbol{t})}{dp} &= -n\left[\frac{\bar{t}-1}{1-p} - \frac{1}{2-p}\right] \\ 0 &= (\bar{t}-1)(2-p') - (1-p') \\ p' &= \frac{3-2\bar{t}}{2-\bar{t}}, \end{aligned} \tag{1}$$

but there is no solution when $\bar{t} = 2$ and the solution can be negative for $\bar{t} > 1.5$. First we check that this solution yields a maximum when $\bar{t} < 2$,

$$\begin{aligned} \frac{d^2 l(p \mid \boldsymbol{t})}{dt^2} &= -n\left[\frac{\bar{t}-1}{(1-p)^2} - \frac{1}{(2-p)^2}\right] \\ &= \frac{-n}{(1-p)^2(2-p)^2}\left[(\bar{t}-1)(2-p)^2 - (1-p)^2\right], \end{aligned}$$

which will be negative at p' if

$$\begin{aligned} (\bar{t}-1)(2-p)^2 - (1-p)^2\big|_{p=p'} &> 0 \\ (\bar{t}-1)\left(2 - \frac{3-2\bar{t}}{2-\bar{t}}\right)^2 - \left(1 - \frac{3-2\bar{t}}{2-\bar{t}}\right)^2 &> 0 \\ (\bar{t}-1)\left(\frac{1}{2-\bar{t}}\right)^2 - \left(\frac{\bar{t}-1}{2-\bar{t}}\right)^2 &> 0 \end{aligned}$$

$$\frac{\bar{t}-1}{2-\bar{t}} \quad > \quad 0,$$

which is satisfied for all $1 < \bar{t} < 2$. When $\bar{t} = 1$, $l(p \mid \boldsymbol{t})$ is increasing on $[0, 1]$, so $\hat{p} = 1$, which is consistent with the estimator (1). When $\bar{t} > 1.5$, $\hat{p}' < 0$, but $l(p \mid \boldsymbol{t})$ is decreasing on $[0, 1]$, and the maximum is at $\hat{p} = 0$. Overall, we conclude the MLE is

$$\hat{p} = \max\left\{0, \frac{3-2\bar{t}}{2-\bar{t}}\right\}.$$

7. Verify that the method-of-moments estimator (MME) based on data $T_1, T_2, \ldots, T_n$, say $\tilde{p}$, is the same as the MLE $\hat{p}$ when $\tilde{p} \geq 0$.

Solution: We need the expectation of our truncated random variable.

$$E[T] \quad = \quad \frac{1}{2-p} + \frac{2(1-p)}{2-p} = \frac{3-2p}{2-p}.$$

Set this equal to the sample mean $\bar{t}$ to find the MME is

$$\tilde{p} = \frac{3-2\bar{t}}{2-\bar{t}}.$$

Note, this estimator may be negative.

8. What is the Cramér-Rao lower bound (call it $v^*(p)$) on the variance for unbiased estimators of p based on $T_1, T_2, \ldots, T_n$?

Solution:

$$\begin{aligned}
-\frac{d^2 l(p \mid T)}{dt^2} &= \frac{T-1}{(1-p)^2} - \frac{1}{(2-p)^2} \\
E\left[-\frac{d^2 l(p \mid T)}{dt^2}\right] &= \frac{E[T]-1}{(1-p)^2} - \frac{1}{(2-p)^2} \\
&= \frac{1}{(2-p)(1-p)} - \frac{1}{(2-p)^2} = \frac{1}{(2-p)^2(1-p)}
\end{aligned}$$

The Cramér-Rao lower bound on the variance for unbiased estimators of p is hence

$$v^*(p) = (2-p)^2(1-p).$$

9. Use the delta method and the above results to obtain approximations for $E[\hat{p}]$ and

Var($\hat{p}$).

10. Verify that $\sqrt{n}\,(\hat{p}-p)\,/\sqrt{v^*(p)} \xrightarrow{\text{d}} \mathcal{N}(0,1)$. (In this sense, the MLE achieves the Cramér-Rao lower bound in the limit.)

Solution:

There is a singularity at $x = 2$, and I am not quite sure of the consequences even though in our application we stay well away from it.

By the Central Limit Theorem, we have

$$\sqrt{n}\left(\overline{T}-\mu_T\right) \xrightarrow{d} \mathcal{N}\left(0,\sigma_T^2\right),$$

where $E[T_1] = \mu_T = \frac{3-2p}{2-p}$ from **Problem 7** and $\text{Var}(T_1) = \sigma_T^2$ is to be computed. The result follows so long as σ_T^2 is finite. Since

$$E[T_1^2] = \frac{1}{2-p} + \frac{4(1-p)}{2-p} = \frac{5-4p}{2-p},$$

we have

$$\sigma_T^2 = E[T_1^2] - (E[T_t])^2 = \frac{5-4p}{2-p} - \left(\frac{3-2p}{2-p}\right)^2 = \frac{1-p}{(2-p)^2},$$

which is finite for $p \in (0,1)$.
Consider transform $g(x) = \frac{3-2x}{2-x}$, whose derivative is

$$g'(x) = \frac{-2}{2-x} + \frac{(3-2x)}{(2-x)^2} = \frac{-1}{(2-x)^2}.$$

By the delta method, the asymptotic variance of $g(\overline{T})$ is

$$\text{Var}\left[g(\overline{T})\right] = \frac{1-p}{n(2-p)^2}\left|g'(\mu)\right|^2 = \frac{1-p}{n(2-p)^2\left(2-\frac{3-2p}{2-p}\right)^4} = \frac{1}{n}(1-p)(2-p)^2,$$

and since $g(\cdot)$ is an involution ($g^{-1}(x) = g(x)$), we have $g(\mu_T) = p$ and

$$\sqrt{n}\left(\frac{3-2\overline{T}}{2-\overline{T}} - p\right) \xrightarrow{d} \mathcal{N}\left(0,(1-p)(2-p)^2\right).$$

Therefore, our MLE $\hat{p} = \max\left\{0, \frac{3-2\overline{T}}{2-\overline{T}}\right\}$ approximately follows a left 0-truncated normal with (pre-truncation) mean and variance

$$\mu = p$$

$$\sigma^2 = \frac{(1-p)(2-p)^2}{n}.$$

Let $\alpha = \frac{-\mu}{\sigma} = \frac{-\sqrt{n}p}{(2-p)\sqrt{2-p}}$. The mean and variance after truncation are

$$\begin{aligned} E[\hat{p}] &= p + \frac{\phi(\alpha)(2-p)\sqrt{1-p}}{\sqrt{n}\,[1-\Phi(\alpha)]} \\ \mathrm{Var}(\hat{p}) &= \frac{(1-p)(2-p)^2}{n}\left[1 - \frac{\sqrt{n}p\phi(\alpha)}{[1-\Phi(\alpha)]\,(2-p)\sqrt{1-p}} - \left(\frac{\phi(\alpha)}{1-\Phi(\alpha)}\right)^2\right] \end{aligned}$$

The mean squared error is

$$MSE(\hat{p}) = \mathrm{Var}(\hat{p}) + \mathrm{Bias}(\hat{p})^2 = \frac{(1-p)(2-p)^2}{n}.$$

In the limit $n \to \infty$, the MLE is unbiased because the second term vanishes and so the variance approaches the Cramér-Rao lower bound.

Part III

Now consider the possibility, ϵ_k, of k consecutive heads. Disallow overlaps, so in repeated trials, a subsequent occurrence of ϵ_k cannot build on a previous occurrence of ϵ_k. Letting E_i denote the outcome of the ith coin flip, if $E_{m-1} = 0, E_m = 1, E_{m+1} = 1$, and $E_{m+2} = 1$, with 1 indicating heads, then ϵ_2 occurs at $m+1$ *but not* at $m+2$. Assume, as usual, that trials are independent and the same coin is used for all flips.

11. Let V_k be the time of the first occurrence of ϵ_k. Prove that

$$E[V_k] = 1 + E[V_{k-1}] + (1-p)E[V_k].$$

and use the result to show that $E[V_2] = \frac{p+1}{p^2}$.

Solution: There are two mutually exclusive, exhaustive events: (1) either the first run of k heads occurs one time unit after the first run of $k-1$ heads with probability p, or (2) it doesn't, one time step is spent on a tail with probability $1-p$, and the wait starts all over again. Application of the law of total expectation produces

$$E[V_k] = p\,(1 + E[V_{k-1}]) + (1-p)\,(1 + E[V_{k-1}] + E[V_k])\,,$$

which trivially simplifies to the recurrence given.

To solve the equation, we have $E[V_0] = 0$, $E[V_1] = p$, and by the recurrence

$$E[V_2] = 1 + p^2 + (1-p)(p + E[V_2]),$$

which is solved as

$$\begin{aligned} [1-(1-p)]E[V_2] &= 1+p^2+p-p^2 \\ E[V_2] &= \frac{1+p}{p}. \end{aligned}$$

One might also find the solution by hypothesizing generating function $h(s) = \sum_{k=1}^{\infty} E[V_k]s^k$ and solving for the coefficients consistent with the recurrence.

12. Redefine u_n and f_n from Part II now for the possibility, ϵ_k, of k consecutive heads. From these probabilities, define $U(s) = \sum_{k=0}^{\infty} u_k s^k$ (which is *not* a pgf) and the pgf $F(s)$ as in **Problem 4**. Show that

$$U(s) = \frac{1}{1-F(s)}.$$

Solution: By assumptions of the question,

$$\Pr(\epsilon_k \text{ occurs at trial } n+m \mid \epsilon_k \text{ occurs at trial } n) = \Pr(\epsilon_k \text{ occurs at trial } m). \quad (2)$$

Therefore, if ϵ_k occurs *first* at trial n, then the probability it occurs again at trial $n+m$ is given by u_m since the sequence $E_{n+1}, E_{n+2}, \ldots, E_{n+m}$ has the exact same properties as $E_1, E_2, \ldots, E_m$ under the assumptions (of independence and homogeneity – same coin). The joint probability of a first occurrence at trial n and recurrence at $n+m$ is thus $f_n u_m$. By the Law of Total Probability, the probability of an occurrence at trial $n > 0$ is

$$u_n = f_1 u_{n-1} + f_2 u_{n-2} + \cdots + f_n.$$

Since $u_0 \equiv 1$, we can multiply the last term by u_0. The right-hand-side are the coefficients of s^n in $F(s)U(s)$, $n > 0$. Therefore,

$$U(s) - 1 = F(s)U(s).$$

Since $1 - F(s) \neq 0$ for $|s| < 1$, we can solve for $U(s)$ to get the desired result.

13. Show that if ϵ_2 is the possibility of two heads in a row, then

$$F(s) = \frac{p^2 s^2}{p^2 s^2 + (1-s)(1+sp)}.$$

Solution: We have $u_0 = 1$ and $u_1 = 0$. For $n > 1$, we have $E_{n-1} = E_n = 1$ with probability p^2. There are two mutually exclusive possibilities: (1) either ϵ occurred at trial $n-1$ or (2) ϵ occurred at trial n. Thus,

$$u_{n-1}p + u_n = p^2$$

Multiply both sides by s^n and sum over $n > 1$.

$$\begin{array}{rcll}
\sum_{n=2}^{\infty} u_{n-1}ps^n + \sum_{n=2}^{\infty} u_n s^n & = & \sum_{n=2}^{\infty} p^2 s^n & \\
sp\sum_{n=2}^{\infty} u_{n-1}s^{n-1} + \sum_{n=2}^{\infty} u_n s^n & = & p^2 s^2 \sum_{n=2}^{\infty} s^{n-2} & \text{constant term, both sides} \\
sp\sum_{n=1}^{\infty} u_n s^n + \sum_{n=1}^{\infty} u_n s^n & = & p^2 s^2 \sum_{n=0}^{\infty} s^n & \text{change of variable and } u_1 = 0 \\
(sp+1)\left[U(s) - 1\right] & = & \frac{p^2 s^2}{1-s} & U(s) = 1 + \sum_{n=1}^{\infty} u_n s^n \\
& & & \text{and geometric series}
\end{array}$$

Thus,

$$U(s) = \frac{p^2 s^2}{(1-s)(1+sp)} + 1 = \frac{p^2 s^2 + (1-s)(1+sp)}{(1-s)(1+sp)}.$$

Then,

$$\begin{array}{rcll}
F(s) & = & \frac{U(s)-1}{U(s)} & \text{from (??)} \\
& = & \frac{p^2 s^2}{p^2 s^2 + (1-s)(1+sp)} & \text{from above}
\end{array}$$

14. Find the variance of V_2.

Solution: The following confirms are previous work in **Problem 11**.

$$\begin{array}{rcll}
F(s)\left[p^2 s^2 + (1-s)(1+sp)\right] & = & p^2 s^2 & \text{definition of } F(s) \\
F'(s)\left[p^2 s^2 + (1-s)(1+sp)\right] + F(s)\left[2p^2 s - 1 + p - 2sp\right] & = & 2p^2 s & \text{derivative } \frac{d}{ds} \\
E[T]p^2 + 2p^2 - 1 - p & = & 2p^2 & F'(1) = E[T] \\
\mu := E[T] & = & \frac{1+p}{p^2}. &
\end{array}$$

$$\begin{array}{rcll}
F''(s)\left[p^2 s^2 + (1-s)(1+sp)\right] & & & \\
+ F'(s)\left[2p^2 s - (1+sp) + p(1-s)\right] & & & \text{derivative } \frac{d^2}{ds^2} \\
+ F'(s)\left[2p^2 s - 1 + p - 2sp\right] + F(s)\left[2p^2 - 2p\right] & = & 2p^2 & \\
E[T(T-1)]p^2 + E[T]\left[4p^2 - 2p - 2\right] - 2p(1-p) & = & 2p^2 & F''(1) = E[T(T-1)] \\
E[T^2]p^2 + \frac{1+p}{p^2}\left[3p^2 - 2p - 2\right] & = & 2p & \text{simplify \& } \textbf{Problem 11}
\end{array}$$

yields

$$E[T^2] = \frac{-p^3 - p^2 + 4p + 2}{p^4}.$$

The variance is

$$\sigma^2 := E[T^2] - \frac{(1+p)^2}{p^4} = -\frac{p^3 + 2p^2 - 2p - 1}{p^4}.$$

15. Provide an asymptotic formula for an approximate probability that the 38th non-overlapping occurrence of two heads in independent flips of a fair coin occurs after the 178th flip.

Solution: We have observed 38 realizations of T. We know that

$$\sqrt{n}\left(\overline{T} - \frac{1+p}{p^2}\right) \xrightarrow{d} \mathcal{N}\left(0, -\frac{p^3 + 2p^2 - 2p - 1}{p^4}\right)$$

We seek the probability

$$\Pr\left(\overline{T} > \frac{178}{38}\right) \approx 1 - \Phi\left(\frac{178}{38}; \frac{(1+p)}{p^2}, -\frac{p^3 + 2p^2 - 2p - 1}{38p^4}\right),$$

which is 0.9581211.

A random variable X is said to have a Weibull distribution with a scale parameter θ, and a shape parameter κ, if its pdf is given by

$$f(x|\theta,\kappa) = \begin{cases} \frac{\kappa}{\theta} x^{\kappa-1} e^{-\left(\frac{x^\kappa}{\theta}\right)} & \text{for } x > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where $\kappa > 0$ and $\theta > 0$. We use $X \sim \mathrm{WEB}(\kappa, \theta)$ to indicate that the random variable X has a Weibull distribution with parameters κ and θ.

In this exam, we assume $X_1, X_2, \ldots, X_n$ are iid $\mathrm{WEB}(\kappa, \theta)$ and we assume κ is a known parameter and θ is an unknown parameter.

1. Show that $\mathrm{E[X^r]} = \theta^{\mathrm{r}}\Gamma(1 + \frac{\mathrm{r}}{\kappa})$, where $X \sim \mathrm{WEB}(\kappa, \theta)$.

2. Find $\tilde{\theta}_n$, the method of moments estimator (MME) of θ, based on $X_1, X_2, \ldots, X_n$.

3. Find $\hat{\theta}_n$, the maximum likelihood estimator (MLE) of θ, based on $X_1, X_2, \ldots, X_n$.

4. Show that $\hat{\theta}_n$ is an unbiased estimator for θ.

5. Show that $\tilde{\theta}_n$ and $\hat{\theta}_n$ are both consistent estimators for θ.

6. Define $\eta = \theta^\kappa$. prove that $\hat{\eta} = \frac{\Gamma(n)}{\Gamma(n+\kappa)}\left(\sum_{i=1}^n X_i^\kappa\right)^\kappa$ is the uniformly minimum variance unbiased estimator (UMVUE) of η.

7. Let $X_{(1)} = \min(X_1, \ldots, X_n)$ be the first order statistic based on the random sample $X_1, \ldots, X_n$. Prove that

$$\frac{d}{d\theta}\mathrm{E}_\theta\left[(\mathrm{X}_{(1)})^\kappa \Big| \sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{X_i^\kappa}\right] = 1$$

for $\theta > 0$.

8. Define a sequence of random variables

$$Q_n = Q_n(X_1, \ldots, X_n) = \frac{\kappa^{\gamma_1}\sqrt{n}(\hat{\theta}_n - \theta)}{(\hat{\theta}_n)^{\gamma_2}}$$

for $n \geq 1$. Find values of γ_1 and γ_2 so that Q_n is an asymptotical pivot.

9. Using the asymptotic distribution of Q_n in Problem **8**, find a confidence interval for θ with approximate confidence coefficient $1 - \alpha$.

10. Consider testing $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$ based on $X_1, \ldots, X_n$. Show that the test that rejects H_0 when $\sum_{i=1}^n X_i^\kappa > c$ for some constant c is a uniformly most powerful (UMP) test. Determine c so that the test will have size α.

11. Construct a confidence interval with confidence coefficient $1-\alpha$ for θ based on $T = \sum_{i=1}^{n} X_i^{\kappa}$.

For problems **12** to **14**, suppose θ has the Inverse-Gamma (IG) prior density

$$\pi(\theta) = \begin{cases} \frac{b^a}{\Gamma(a)\theta^{a+1}} e^{-\frac{b}{\theta}} & \text{for } \theta > 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $a > 0$ and $b > 0$.

12. Derive the Bayes estimator of θ under the squared error loss.

13. Find a Bayes test for $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$.

14. Find a $(1-\alpha)$ Bayes credible set for θ.

1. In order to calculate the rth moments of the random variable X, we use the definition of Gamma function $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y}\, dy$ and change of variables as follows:

$$\begin{aligned}\mathbb{E}[X^r] &= \int_0^\infty \frac{\kappa}{\theta} x^{r+\kappa-1} e^{-(\frac{x^\kappa}{\theta})}\, dx \\ &= \theta^{\frac{r}{\kappa}} \int_0^\infty \kappa\, y^{\frac{r}{\kappa}} e^{-y}\, dy \quad \big(y := \frac{x^\kappa}{\theta}\big) \\ &= \theta^{\frac{r}{\kappa}}\, \Gamma\big(1 + \frac{r}{\kappa}\big).\end{aligned}$$

2. Use problem **1** for $r = 1$ to see $\mathbb{E}(X_1) = \theta^{\frac{1}{\kappa}}\, \Gamma(1 + \frac{1}{\kappa})$ and obtain $\tilde{\theta}_n$ by solving $m_1 = \bar{X}_n = \theta^{\frac{1}{\kappa}}\, \Gamma(1 + \frac{1}{\kappa})$. This yields $\tilde{\theta}_n = \Big(\frac{\bar{X}_n}{\Gamma(1+\frac{1}{\kappa})}\Big)^\kappa$.

3. The likelihood function is

$$L(\theta|x_1, \ldots, x_n) = \big(\frac{\kappa}{\theta}\big)^n \prod_{i=1}^n x_i^{(\kappa-1)} \exp\Big(-\frac{1}{\theta}\sum_{i=1}^n x_i^\kappa)\Big)$$

and hence the log-likelihood becomes

$$\log L(\theta|x_1, \ldots, x_n) = n \log\kappa - n\log\theta + (\kappa - 1)\sum_{i=1}^n \log x_i - \frac{1}{\theta}\sum_{i=1}^n x_i^\kappa.$$

By taking the derivative with respect to θ and solving for $\theta = \hat{\theta}_n$, we have:

$$\frac{\partial}{\partial\theta}\log L(\theta|x_1, \ldots, x_n) = \frac{-n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^n x_i^\kappa\, |_{\theta=\hat{\theta}} = 0$$

and consequently we get $\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^n x_i^\kappa$. Therefore $\hat{\theta}_n$ maximize $\log L(\theta|x_1, \ldots, x_n)$ if $\frac{\partial^2}{\partial\theta^2}\log L(\theta|x_1, \ldots, x_n)|_{\theta=\hat{\theta}} < 0$. But $\frac{\partial^2}{\partial\theta^2}\log L(\theta|x_1, \ldots, x_n)|_{\theta=\hat{\theta}} = -\frac{n}{\hat{\theta}^2} < 0$ and this completes the proof.

4. We first note that $X_i^\kappa \sim \mathrm{Exp}(\theta)$ and consequently $\sum_{i=1}^n X_i^\kappa \sim \mathrm{Gamma}(n, \kappa)$. Therefore $\mathbb{E}[\hat{\theta}_n] = \mathbb{E}\Big(\frac{1}{n}\sum_{i=1}^n X_i^\kappa\Big) = \frac{n\theta}{n} = \theta$ which means $\hat{\theta}$ is an unbiased estimator for θ.

5. For $\tilde{\theta}_n$, the WLLN implies that $\bar{X}_n \xrightarrow{p} \mathbb{E}(X) = \theta^{\frac{1}{\kappa}}\, \Gamma(1 + \frac{1}{\kappa})$. Define $g(x) = \big(\frac{x}{\Gamma(1+\frac{1}{\kappa})}\big)^\kappa$ and use the continuous mapping theorem to see $g(\bar{X}_n) \xrightarrow{p} g(\mathbb{E}(X)) = g(\theta^{\frac{1}{\kappa}}\, \Gamma(1+\frac{1}{\kappa}))$ as $n \to \infty$. Note that $g(\bar{X}_n) = \big(\frac{\bar{X}_n}{\Gamma(1+\frac{1}{\kappa})}\big)^\kappa = \tilde{\theta}_n$ and $g(\theta^{\frac{1}{\kappa}}\, \Gamma(1 + \frac{1}{\kappa})) = \theta$. This means $\tilde{\theta}_n \xrightarrow{p} \theta$ and hence $\tilde{\theta}_n$ is consistent for θ. For $\hat{\theta}_n$, use the fact that $Y_i = X_i^\kappa \sim \mathrm{Exp}(\theta)$ for each $i = 1, \ldots, n$. Now, using WLLN, we have $\hat{\theta}_n = \bar{Y}_n \xrightarrow{p} \mathbb{E}(Y_1) = \theta$ which means $\hat{\theta}_n$ is consistent for θ.

6. We first note that

$$f(x_1,\dots,x_n|\eta) = \left(\kappa\,\eta^{\frac{-1}{\kappa}}\right)^n \prod_{i=1}^{n} x_i^{\kappa-1} \exp\Big(\frac{-1}{\eta^{\frac{1}{\kappa}}}\sum_{i=1}^{n} X_i^{\kappa}\Big)$$

and $A \equiv \{\underset{\sim}{x} : f(\underset{\sim}{x}|\eta) > 0\}$ does not depend on η (here $\underset{\sim}{x} = (x_1,\dots,x_n)$). Hence $f(\underset{\sim}{x}|\eta)$ belongs to the exponential family. Therefore, $T = \sum_{i=1}^{n} X_i^{\kappa}$ is complete and sufficient statistic for η. Next, we show that $\hat{\eta}$ is an unbiased estimator for η. In the following, we will use the fact that $Z := \sum_{i=1}^{n} X_i^{\kappa} \sim \text{Gamma}(n,\theta)$ and then we calculate $\mathbb{E}[Z^{\kappa}]$ to show that $\hat{\eta}$ is an unbiased estimator for η:

$$\begin{aligned}
\mathbb{E}[\hat{\eta}] &= \frac{\Gamma(n)}{\Gamma(n+k)}\mathbb{E}\Big[\Big(\sum_{i=1}^{n} X_i^{\kappa}\Big)^{\kappa}\Big] \\
&= \frac{\Gamma(n)}{\Gamma(n+k)}\mathbb{E}\big[Z^{\kappa}\big] \quad \big(Z \sim \text{Gamma}(n,\theta)\big) \\
&= \frac{\Gamma(n)}{\Gamma(n+k)}\int_0^{\infty} z^{\kappa}\,\frac{1}{\Gamma(n)\theta^n}\,z^n\,e^{-\frac{z}{\theta}}\,dz \\
&= \frac{\theta^{\kappa}}{\Gamma(n+k)}\int_0^{\infty} y^{n+\kappa-1}\,e^{-y}\,dy \quad (\frac{z}{\theta} = y) \\
&= \frac{\theta^{\kappa}\,\Gamma(n+k)}{\Gamma(n+k)} = \theta^{\kappa} = \eta
\end{aligned}$$

Finally, since $\hat{\eta}$ is an unbiased estimator of η and also it is a function of the complete and sufficient statistic $\sum_{i=1}^{n} X_i^{\kappa}$ then Lehmann- Scheffe Theorem implies that $\hat{\eta}$ is UMVUE for η.

7. The random variable $Y = \frac{(X_{(1)})^{\kappa}}{\theta} \sim \text{EXP}(1)$ and therefore Y is ancillary. On the other hand, $\sum_{i=1}^{n} X_i^{\kappa}$ is complete and sufficient statistic. Now, Basu's lemma implies that Y and $\sum_{i=1}^{n} X_i^{\kappa}$ are independent. That is $\mathbb{E}_{\theta}\Big[(X_{(1)})^{\kappa}\Big|\sum_{i=1}^{n} X_i^{\kappa}\Big] = \mathbb{E}_{\theta}\Big[(X_{(1)})^{\kappa}\Big] = \theta$. Now, taking the derivative of the both sides of the last equation with respect to θ gives the desired result.

8. By using the asymptotic properties of the MLEs, we have

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N\Big(0, \frac{1}{I_1(\theta)}\Big),$$

where

$$I_1(\theta) = -\mathbb{E}\Big[\frac{\partial^2}{\partial\theta^2}\log f(X|\theta)\Big] = \frac{1}{\theta^2}.$$

Define $g(x) = x^2$. According to WLLN, we have $\hat{\theta}_n \xrightarrow{p} \theta$ and then the continuous mapping theorem for function $g(x)$ implies that $g(\hat{\theta}_n) = \hat{\theta}_n^2 \xrightarrow{p} \theta^2$ as $n \to \infty$. Now by applying

Slutsky's theorem,

$$\frac{\sqrt{n}(\hat{\theta}_n - \theta)}{\hat{\theta}_n} \xrightarrow{d} \mathrm{N}\big(0, \frac{\theta^2}{\theta^2} = 1\big)$$

as $n \to \infty$. Therefore $\frac{\sqrt{n}(\hat{\theta}_n - \theta)}{\hat{\theta}_n^2}$ is an asymptotical pivot which means Q_n is also asymptotical pivot for $\gamma_1 = 0$ and $\gamma_2 = 1$.

9. From problem **8**, we know that $Q_n \to Q$ as $n \to \infty$ where the distribution of Q does not depend on θ. In fact, $Q \sim \mathrm{N}(0,1)$ and we know that $\mathbb{P}\Big(z_{\frac{\alpha}{2}} \le Z \le z_{1-\frac{\alpha}{2}}\Big) = 1 - \alpha$. Therefore for large n,

$$C_{\underset{\sim}{X}} = \Big\{\theta > 0, z_{\frac{\alpha}{2}} \le Q_n(X_1, \dots, X_n, \theta) \le z_{1-\frac{\alpha}{2}}\Big\}$$

$$= \Big[\hat{\theta}_n - \frac{z_{1-\frac{\alpha}{2}}\hat{\theta}_n}{\sqrt{n}}, \hat{\theta}_n - \frac{z_{\frac{\alpha}{2}}\hat{\theta}_n}{\sqrt{n}}\Big]$$

is a confidence interval for θ with approximate confidence coefficient $1 - \alpha$.

10. First, note that $f(\underset{\sim}{x}|\theta)$ have MLR in the real-valued statistic $T = t(\underset{\sim}{X}) = \sum_{i=1}^n X_i^\kappa$. In fact, for any $\theta_2 > \theta_1$, we have,

$$h(\underset{\sim}{x}) := \frac{f(\underset{\sim}{x}|\theta_2)}{f(\underset{\sim}{x}|\theta_1)} = \frac{\theta_1}{\theta_2} \exp \sum_{i=1}^n x_i^\kappa \Big[\frac{1}{\theta_1} - \frac{1}{\theta_2}\Big]$$

$$= g_{\theta_1,\theta_2}\big(t(\underset{\sim}{x})\big),$$

where $g_{\theta_1,\theta_2}\big(t(\underset{\sim}{x})\big) = \frac{\theta_1}{\theta_2} \exp\Big\{t[\frac{1}{\theta_1} - \frac{1}{\theta_2}]\Big\}$ is non-decreasing as $t \to \infty$ (Note that this is true when $\underline{\kappa \text{ is the known parameter}}$). Now, apply Karrlin-Rubin Theorem to find the UMP test of $H_0 : \theta \le \theta_0$ vs $H_1 : \theta > \theta_0$ based on $X_1, \dots, X_n$ as follows:

$$\phi(\underset{\sim}{X}) = \begin{cases} 1, & \text{if } \sum_{i=1}^n X_i^\kappa > c, \\ 0, & \text{otherwise.} \end{cases}$$

In order to have the test of size α, we need to find c. Write

$$\alpha = \mathbb{E}_{\theta_0}\big[\phi(\underset{\sim}{X})\big] = \mathbb{P}_{\theta_0}\Big(\sum_{i=1}^n X_i^\kappa > c\Big) = 1 - \mathbb{P}_{\theta_0}\Big(\frac{2\sum_{i=1}^n X_i^\kappa}{\theta_0} \le \frac{2c}{\theta_0}\Big).$$

But, $\frac{2\sum_{i=1}^n X_i^\kappa}{\theta_0} \sim \chi^2_{2n}$ and hence $\frac{2c}{\theta_0} = \chi^2_{2n,1-\alpha}$. Therefore, the UMP test of size α has the form:

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

11. We know that $T = \sum_{i=1}^n X_i^\kappa \sim \text{Gamma}(n, \theta)$ and then $\frac{2T}{\theta} \sim \chi^2_{2n}$. Given t, $F(t|\theta) := \mathbb{P}(T \le t|\theta) = \mathbb{P}(\frac{2T}{\theta} \le \frac{2t}{\theta}|\theta) = \mathbb{P}(\chi^2_{2n} \le \frac{2t}{\theta}|\theta) = G(\frac{2t}{\theta})$, where $G(\cdot)$ is the cdf of χ^2_{2n}. We see that $F(t|\theta) = G(\frac{2t}{\theta}) \downarrow$ as $\theta \uparrow$ and hence by applying the tail theorem we have:

$$\frac{\alpha}{2} = \mathbb{P}(T \le t|\theta_U(t)) = G(\frac{2t}{\theta_U(t)})$$

and

$$1 - \frac{\alpha}{2} = \mathbb{P}(T \le t|\theta_L(t)) = G(\frac{2t}{\theta_L(t)})$$

and then $\big(\theta_L(t), \theta_U(t)\big) = \Big(\frac{2t}{\chi^2_{2n,\frac{1-\alpha}{2}}}, \frac{2t}{\chi^2_{2n,\frac{\alpha}{2}}}\Big)$ for any given value t of T. Hence

$$(\theta_L(T), \theta_U(T)) = \left(\frac{2T}{\chi^2_{2n,\frac{1-\alpha}{2}}}, \frac{2T}{\chi^2_{2n,\frac{\alpha}{2}}}\right) = \left(\frac{2\sum_{i=1}^n X_i^\kappa}{\chi^2_{2n,\frac{1-\alpha}{2}}}, \frac{2\sum_{i=1}^n X_i^\kappa}{\chi^2_{2n,\frac{\alpha}{2}}}\right)$$

is a confidence interval for θ with confidence coefficient $(1 - \alpha)$.

12. The posterior distribution of θ is $\text{IG}(n + a, \sum_{i=1}^n x_i^\kappa + b)$. To see this write:

$$\begin{aligned} f(\theta|\underset{\sim}{x})(\theta) &= \pi(\theta) f(\underset{\sim}{x}|\theta) \\ &\propto \frac{1}{\theta^n}\, e^{\frac{-\sum_{i=1}^n x_i^\kappa}{\theta}}\, \theta^{-a-1}\, e^{-\frac{b}{\theta}} \\ &= \theta^{-(n+a)-1}\, e^{-\frac{1}{\theta}\left(\sum_{i=1}^n x_i^\kappa + b\right)}, \end{aligned}$$

where the last one is proportional to the density function of $\text{IG}(n + a, \sum_{i=1}^n x_i^\kappa + b)$. Now, the Bayes estimator of θ is given by

$$\mathbb{E}_{\theta|\underset{\sim}{x}}(\theta) = \frac{\sum_{i=1}^n x_i^\kappa + b}{n + a - 1},$$

where we used the fact that $\mathbb{E}[X] = \frac{b}{a-1}$ if $X \sim \text{IG}(a, b)$ to obtain the Bayes estimator.

13. We reject H_0 if $\mathbb{P}(\theta \le \theta_0|\underset{\sim}{x}) < \frac{1}{2}$. Note that $\theta|\underset{\sim}{x} \sim \text{IG}(n + a, \sum_{i=1}^n X_i^\kappa + b)$ and then $\frac{1}{\theta} \sim \text{Gamma}(n + a, \frac{1}{\sum_{i=1}^n X_i^\kappa + b})$. Therefore, we have

$$\begin{aligned} \mathbb{P}(\theta \le \theta_0|\underset{\sim}{x}) &= \mathbb{P}\Big(\frac{1}{\theta} \ge \frac{1}{\theta_0}\Big|\underset{\sim}{x}\Big) \\ &= \mathbb{P}\Big(\frac{2\big(\sum_{i=1}^n X_i^\kappa + b\big)}{\theta} \ge \frac{2\big(\sum_{i=1}^n X_i^\kappa + b\big)}{\theta_0}\Big|\underset{\sim}{x}\Big) \\ &= \mathbb{P}\Big(\chi^2_{2(n+a)} \ge \frac{2\big(\sum_{i=1}^n X_i^\kappa + b\big)}{\theta_0}\Big|\underset{\sim}{x}\Big) < \frac{1}{2}. \end{aligned}$$

Consequently, we reject the null hypothesis if $\frac{2\left(\sum_{i=1}^n X_i^\kappa + b\right)}{\theta_0} > \chi^2_{2(n+a),0.5}$. Hence we have the following Bayesian test:

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

14. From Problems **12** and **13** we know that $\theta|\underset{\sim}{x} \sim \mathrm{IG}(n+a, \sum_{i=1}^n x_i^\kappa + b)$ and $\frac{1}{\theta} \sim \mathrm{Gamma}(n+a, \frac{1}{\sum_{i=1}^n X_i^\kappa + b})$. Then $\frac{2\left(\sum_{i=1}^n x_i^\kappa + b\right)}{\theta} \sim \chi^2_{2(n+a)}$. So,

$$1-\alpha = \mathbb{P}\big(L \leq \theta \leq U\big) = \mathbb{P}\Big(\frac{1}{U} \leq \frac{1}{\theta} \leq \frac{1}{L}\Big) = \mathbb{P}\Big(\frac{2\big(\sum_{i=1}^n x_i^\kappa + b\big)}{U} \leq \frac{1}{\theta} \leq \frac{2\big(\sum_{i=1}^n x_i^\kappa + b\big)}{L}\Big)$$

for any given value $\sum_{i=1}^n x_i^\kappa$. Thus

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

is a $(1-\alpha)$ credible set.

Part I

Let Ω be a non-empty set and $(\Omega, \mathcal{F})$ be a measurable space.

1. Let Ω' be a non-empty subset of Ω. Define

$$\mathcal{F}_{\Omega'} = \{A \cap \Omega' : A \in \mathcal{F}\}.$$

Show that $\mathcal{F}_{\Omega'}$ is a σ-algebra of subsets of Ω'.

2. Define a π-class with subsets of Ω.

3. Define a λ-system with subsets of Ω.

4. If a λ-system is also a π-class, show then it is a σ-algebra.

5. Let P_1 and P_2 be two probability measures on $(\Omega, \mathcal{F})$. Show that the class

$$\mathcal{L} \equiv \{A \in \mathcal{F} : P_1(A) = P_2(A)\}$$

is a λ-system.

Part II

6. Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\{X_n\}_{n\geq 1}$ be a monotone sequence of random variables and X be another random variable, all defined on $(\Omega, \mathcal{F}, P)$. Assume that $X_n \stackrel{P}{\to} X$ as $n \to \infty$. Show that $X_n \to X$ almost surely as $n \to \infty$.

Part III

Let $X_1 \sim$ Poisson (λ), $X_2 \sim$ Exponential (β), and $Y \sim$ Bernoulli (α), for some known $\beta, \lambda > 0$ and $0 < \alpha < 1$. Assume that X_1, X_2, and Y are independent. Define the random variable

$$Z = YX_1 + (1 - Y)X_2.$$

Denote the cdfs of X_1, X_2 and Z by F_1, F_2 and F respectively. Denote the Lebesgue-Stieltjes measures induced by F_1, F_2, F as μ_{F_1}, μ_{F_2} and μ_F respectively. For two measures μ and ν we write $\mu << \nu$ if μ is dominated by ν.

7. Write down F in terms of F_1 and F_2.

8. Prove whether or not the following dominations hold. In each case, if the domination holds then provide the corresponding Radon-Nikodym derivative.

a) $\mu_{F_1} << \mu_F$

b) $\mu_F << \mu_{F_1}$

c) $\mu_{F_2} << \mu_F$

d) $\mu_F << \mu_{F_2}$

Part IV

Let X_λ be a Poisson random variable with mean λ.

9. Show that the characteristic function of X_λ is given by $\phi_{X_\lambda}(t) = \exp[\lambda\{\exp(it) - 1\}]$.

10. State Lévy's continuity theorem.

11. Define

$$Y_\lambda = \frac{X_\lambda - \lambda}{\sqrt{\lambda}}.$$

Show that as $\lambda \to \infty$, Y_λ converges in distribution to $N(0, 1)$.

12. Consider a sequence of random variables $X_{\lambda_n}, n \geq 1$. Let X_{λ_n} be a Poisson random variable with mean λ_n and suppose $\lambda_n \to \lambda_0 < \infty$ as $n \to \infty$. Show that

$$\sum_{x=0}^{\infty} |P(X_{\lambda_n} = x) - P(X_0 = x)| \to 0, \text{ as } n \to \infty,$$

where X_0 is a Poisson random variable with mean λ_0.

Part V

Let $\{X_n, n \geq 1\}$ be independent random variables on a probability space $(\Omega, \mathcal{F}, P)$ with

$$P(X_n = n^{-\alpha}) = \frac{1}{2} = P(X_n = -n^{-\alpha})$$

for all $n \geq 1$.

13. Show that if $\alpha > 1/2$ then $S_n = \sum_{j=1}^{n} X_j$ converges almost surely.

14. Verify that $\alpha > 1/2$ is necessary for convergence (almost surely or in probability) of S_n.

Part VI

Let $\{X_{nj} : 1 \le j \le n\}_{n\ge1}$ be a triangular array of random variables such that $X_{n1}, X_{n2}, \ldots, X_{nn}$ are independent, $E(X_{nj}) = 0$, and $E(X_{nj}^2) < \infty$ for all $j = 1, \ldots, n$.

15. State the Lindeberg condition for $\{X_{nj} : 1 \le j \le n\}_{n\ge1}$.

16. In addition to the above assumptions for the triangular array $\{X_{nj} : 1 \le j \le n\}_{n\ge1}$, assume that $\sum_{j=1}^n E(X_{nj}^2) = 1$ for all $n \ge 1$.

a) Show that
$$\lim_{n\to\infty} \max_{1\le j\le n} E(X_{nj}^2) = 0.$$

b) Show that
$$\max_{1\le j\le n} |X_{nj}| \xrightarrow{P} 0.$$

17. Let $s_n^2 = \text{Var}(\sum_{j=1}^n X_{nj})$. Assume that there exists a $\delta > 0$ such that as $n \to \infty$
$$\frac{\sum_{j=1}^n E|X_{nj}|^{2+\delta}}{s_n^{2+\delta}} \to 0.$$

Show that $\{X_{nj} : 1 \le j \le n\}_{n\ge1}$ satisfies the Lindeberg condition.

1. a) Since $\Omega \in \mathcal{F}$, $\Omega' = \Omega \cap \Omega' \in \mathcal{F}_{\Omega'}$. b) Let $B \in \mathcal{F}_{\Omega'}$. To show $\Omega' \setminus B \in \mathcal{F}_{\Omega'}$. Since $B \in \mathcal{F}_{\Omega'}$, $B = A \cap \Omega'$ for some $A \in \mathcal{F}$. Thus $\Omega' \setminus B = A^c \cap \Omega' \in \mathcal{F}_{\Omega'}$ since $A^c \in \mathcal{F}$. c) Finally, let $B_n \in \mathcal{F}_{\Omega'}$ for all $n \geq 1$. To show $\cup_{n\geq 1} B_n \in \mathcal{F}_{\Omega'}$. Since $B_n \in \mathcal{F}_{\Omega'}$, $B_n = A_n \cap \Omega'$ for some $A_n \in \mathcal{F}$ and $\cup_{n\geq 1} B_n = \cup_{n\geq 1} A_n \cap \Omega'$. Thus $\cup_{n\geq 1} B_n \in \mathcal{F}_{\Omega'}$ as $\cup_{n\geq 1} A_n \in \mathcal{F}$.

2. A non empty class $\mathcal{C}$ of subsets of Ω is called a $\pi-$class if

$$A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}.$$

3. A class $\mathcal{L}$ of subsets of Ω is called a $\lambda-$ system if

$$\text{(i) } \Omega \in \mathcal{L} \qquad \text{(ii) } A \in \mathcal{L} \Rightarrow A^c \in \mathcal{L}$$

$$\text{\& (iii) for any sequence } A_1, A_2, \ldots \text{ of disjoint sets in } \mathcal{L}, \ \bigcup_{n=1}^{\infty} A_n \in \mathcal{L}$$

4. Let $\mathcal{L}$ be a $\lambda-$ system. To show that it is also a σ-algebra, we need to show it is closed under countable union of not necessarily disjoint sets. Let $A_n \in \mathcal{L}$ for all $n \geq 1$. Define $B_1 = A_1, B_2 = A_2 \cap A_1^c, B_3 = A_3 \cap A_2^c \cap A_1^c$ so on. Since $\mathcal{L}$ is a π-class, $B_n \in \mathcal{L}$ for $n \geq 2$. Since B_n's are disjoint, $\cup_{n\geq 1} B_n \in \mathcal{L}$. So $\cup_{n\geq 1} A_n = \cup_{n\geq 1} B_n \in \mathcal{L}$.

5. Note that $\Omega \in \mathcal{L}$ since $P_1(\Omega) = 1 = P_2(\Omega)$. If $A \in \mathcal{L}$, $P_1(A^c) = 1 - P_1(A) = 1 - P_2(A) = P_2(A^c)$. Thus $A^c \in \mathcal{L}$. Finally, let $A_n, n \geq 1$ are disjoint and $A_n \in \mathcal{L}$ for $n \geq 1$. Then $P_1(\cup_n A_n) = \sum_n P_1(A_n) = \sum_n P_2(A_n) = P_2(\cup_n A_n)$. Thus $\cup_n A_n \in \mathcal{L}$. Thus $\mathcal{L}$ is a $\lambda-$ system.

6. Fix $\epsilon > 0$. Since $X_n \xrightarrow{P} X$ as $n \to \infty$, $\lim_{n\to\infty} P(|X_n - X| > \epsilon) = 0$. Since $\{X_n\}_{n\geq 1}$ is a monotone sequence of random variables,

$$\begin{aligned}\lim_{n\to\infty} P(|X_n - X| > \epsilon) &= \lim_{N\to\infty} P(\cup_{n\geq N} |X_n - X| > \epsilon) \\ &= P(\limsup_{N\to\infty} |X_n - X| > \epsilon) = P(|X_n - X| > \epsilon \ i.o.).\end{aligned}$$

Thus $X_n \to X$ almost surely.

7. Note that $F(z) = P(Z \leq z) = E[P(Z \leq z|Y)] = \alpha F_1(z) + (1-\alpha) F_2(z)$.

8. **a)** For any $A \in \mathcal{B}(\mathbb{R})$, $\mu_{F_1}(A) \leq \frac{1}{\alpha}\mu_F(A)$. Thus $\mu_{F_1} << \mu_F$. Denote $\mathbb{N} = \{0, 1, 2, \ldots\}$. Then $\frac{d\mu_{F_1}}{d\mu_F}(x) = \frac{1}{\alpha} I_N(x)$ a.e. (μ_F).

b) Let Q denote the set of rationals. Since $\mu_F(Q^c) = 1 - \alpha, \mu_{F_1}(Q^c) = 0$, $\mu_F \not<< \mu_{F_1}$.

c) For any $A \in \mathcal{B}(\mathbb{R})$, $\mu_{F_2}(A) \leq \frac{1}{1-\alpha}\mu_F(A)$. Thus $\mu_{F_2} << \mu_F$ and $\frac{d\mu_{F_2}}{d\mu_F} = \frac{1}{1-\alpha} I_{N^c}(x)$ a.e. (μ_F).

d) Since $\mu_F(Q) = \alpha, \mu_{F_2}(Q) = 0$, $\mu_F \not<< \mu_{F_2}$.

9. The characteristic function of X_λ is given by

$$\begin{aligned}\phi_{X_\lambda}(t) = E(\exp(itX_\lambda)) &= \sum_{x=0}^{\infty} \exp(itx)\exp(-\lambda)\frac{\lambda^x}{x!}\\ &= \exp(-\lambda)\exp[\lambda\exp(it)] = \exp[\lambda\{\exp(it)-1\}].\end{aligned}$$

10. Suppose $X_n, n \geq 0$ is a sequence of random variables, each with characteristic function ϕ_{X_n}, $n \geq 0$.

(a) If $X_n \xrightarrow{d} X_0$, then for any $T > 0$,

$$\sup_{|t|\leq T} \Big|\phi_{X_n}(t) - \phi_{X_0}(t)\Big| \to 0 \quad \text{as } n \to \infty.$$

(b) If $\phi_{X_n}(t) \to g(t)$ as $n \to \infty$ for all $t \in \mathbb{R}$ and the function $g(\cdot)$ is continuous at zero, then $g(\cdot)$ is a characteristic function and $X_n \xrightarrow{d} X_0$, where X_0 is the random variable with characteristic function $g(\cdot)$.

11. The characteristic function of Y_λ is

$$\phi_{Y_\lambda}(t) = E(\exp(itY_\lambda)) = \exp[\lambda\{\exp(i\frac{t}{\sqrt{\lambda}}) - 1\} - it\sqrt{\lambda}]$$

Consider the Taylor series expansion of $\log \phi_{Y_\lambda}(t)$:

$$\begin{aligned}\log \phi_{Y_\lambda}(t) &= \lambda \sum_{k=0}^{\infty} \frac{(it)^k}{(\sqrt{\lambda})^k k!} - \lambda - it\sqrt{\lambda}\\ &= -\frac{t^2}{2} - \frac{it^3}{\sqrt{\lambda}6} + \dots .\end{aligned}$$

Thus $\log \phi_{Y_\lambda}(t) \to -t^2/2$ as $\lambda \to \infty$ for all $t \in \mathbb{R}$. Thus by the Lévy's Continuity theorem as $\lambda \to \infty$, Y_λ converges in distribution to $N(0,1)$.

12. Since $f_{X_{\lambda_n}}(x) = \exp(-\lambda_n)\frac{\lambda_n^x}{x!} \to \exp(-\lambda_0)\frac{\lambda_0^x}{x!} = f_{X_{\lambda_0}}(x)$ for all $x = 0, 1, 2, \dots$, by Scheffé's Theorem we have the result.

13. Note that $E(X_n) = 0$ and $\sum_{n=1}^{\infty} E(X_n^2) = \sum_{n=1}^{\infty} \frac{1}{n^{2\alpha}} < \infty$ for $\alpha > 1/2$. By the Khintchine-Kolmogorov convergence theorem, $S_n = \sum_{j=1}^{n} X_j$ converges almost surely.

14. Let $X_n^{(1)} \equiv X_n \mathbb{I}(|X_n| \leq 1)$. Since $\sum_{n=1}^{\infty} P(|X_n| > 1) = 0$, $E(X_n^{(1)}) = 0$ for all $n \geq 1$, $\sum_{n=1}^{\infty} \mathrm{var}(X_n^{(1)}) = \sum_{n=1}^{\infty} \frac{1}{n^{2\alpha}} < \infty$ for $\alpha > 1/2$, by Kolmogorov's 3-Series Theorem, $\alpha > 1/2$ is necessary for convergence (almost surely or in probability) of S_n.

15. Let $\{X_{nj} : 1 \leq j \leq n\}_{n=1}^{\infty}$ be a triangular array with

$$\mathrm{EX_{nj}} = 0, \qquad 0 < \mathrm{EX_{nj}^2} = \sigma_{\mathrm{nj}}^2 < \infty, \qquad \mathrm{v_n^2} = \sum_{\mathrm{j=1}}^{\mathrm{n}} \sigma_{\mathrm{nj}}^2.$$

Then, we say $\{X_{nj}\}$ satisfies the Lindeberg Condition if, for all $\epsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{v_n^2} \sum_{j=1}^{n} \mathrm{EX_{nj}^2} \mathbb{I}(|\mathrm{X_{nj}}| > \epsilon \mathrm{v_n}) = 0.$$

16. **a**) Fix $\epsilon > 0$. Since $v_n = 1$

$$\begin{aligned} E(X_{nj}^2) &= E(X_{nj}^2 I(|X_{n,j}| \leq \epsilon)) + E(X_{nj}^2 I(|X_{n,j}| > \epsilon)) \\ &\leq \epsilon^2 + \frac{1}{v_n^2} \sum_{j=1}^{n} E(X_{nj}^2 I(|X_{n,j}| > \epsilon v_n)) \\ &\to \epsilon^2 \text{ as } n \to \infty. \end{aligned}$$

Thus for all $\epsilon > 0$

$$\lim_{n\to\infty} \max_{1\leq j\leq n} E(X_{nj}^2) \leq \epsilon^2.$$

The proof follows by letting $\epsilon \downarrow 0$.

b) Since $v_n = 1$,

$$\begin{aligned} P(\max_{1\leq j\leq n} |X_{nj}| > \epsilon) &\leq P(\cup_{j=1}^{n} |X_{nj}| > \epsilon) \\ &\leq \sum_{j=1}^{n} P(|X_{nj}| > \epsilon) \\ &\leq \sum_{j=1}^{n} E(\frac{|X_{nj}|^2}{\epsilon^2} I(|X_{nj}| > \epsilon)) \\ &= \frac{1}{\epsilon^2 v_n^2} \sum_{j=1}^{n} E(|X_{nj}|^2 I(|X_{nj}| > \epsilon v_n)) \to 0 \text{ as } n \to \infty. \end{aligned}$$

17.

$$\begin{aligned} \frac{1}{v_n^2} \sum_{j=1}^{n} \mathrm{E}[\mathrm{X_{nj}^2} \mathbb{I}(|\mathrm{X_{nj}}| > \epsilon \mathrm{v_n})] &\leq \frac{1}{v_n^2} \sum_{j=1}^{n} [\mathrm{EX_{nj}^2} \frac{|\mathrm{X_{nj}}|^\delta}{\epsilon^\delta \mathrm{v_n^\delta}} \mathbb{I}(|\mathrm{X_{nj}}| > \epsilon \mathrm{v_n})] \\ &\leq \frac{1}{v_n^{2+\delta} \epsilon^\delta} \sum_{j=1}^{n} [\mathrm{E}|\mathrm{X_{nj}}|^{2+\delta} \mathbb{I}(|\mathrm{X_{nj}}| > \epsilon \mathrm{v_n})] \\ &\to 0 \text{ as } n \to \infty. \end{aligned}$$

Thus the Lindeberg condition holds.

Part I

In what follows, you will need the mean and variance of the truncated normal random variable. Let $X \sim \mathcal{N}(\mu, \sigma^2)$, $\phi(\cdot)$ be the standard normal pdf and $\Phi(\cdot)$ be the standard normal cdf. If we condition on X being larger than some constant a, then the variance of the truncated normal random variable is

$$\mathrm{Var}(X \mid X > a) = \sigma^2 \left[1 + \left(\frac{a-\mu}{\sigma} \right) \frac{\phi\left(\frac{a-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{a-\mu}{\sigma}\right)} - \left(\frac{\phi\left(\frac{a-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{a-\mu}{\sigma}\right)} \right)^2 \right].$$

1. Show that the mean of the truncated normal random variable is

$$E[X \mid X > a] = \mu + \frac{\sigma \phi\left(\frac{a-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{a-\mu}{\sigma}\right)}.$$

Part II

Consider independent flips of a coin and focus on the possibility, ϵ, that "the coin lands heads up." Define

$$u_n = \Pr(\epsilon \text{ occurs at trial } n)$$

and

$$f_n = \Pr(\epsilon \textit{ first} \text{ occurs at trial } n),$$

and take $u_0 \equiv 1$ and $f_0 \equiv 0$.

2. If $p \in [0, 1]$ is the probability of heads, what is u_n?

3. What is f_n?

4. Define the probability generating function (pgf) $F(s) = \sum_{k=0}^{\infty} f_k s^k$, $|s| \le 1$ for the wait time, T, until the first occurrence of ϵ. Show that

$$F(s) = \frac{ps}{1 - (1-p)s}.$$

5. Use $F(s)$ to show that $E[T] = \frac{1}{p}$.

Suppose you are able to collect data on the waiting time for the first head in independent coin flipping trials. Unfortunately, your data recording device is broken so it only registers a value when the first head occurs on the first or second trial. Thus $T_1, T_2, \ldots, T_n \in \{1, 2\}$ are iid random variables with the distribution of **Problem 3** truncated above at 2. (Truncation means you observe *no result* when the first head occurs at trial 3 or beyond. All recorded observations satisfy $T_i \le 2$.) **Problems 6–10** are based on the data $T_1, T_2, \ldots, T_n$.

6. Find the MLE $\hat{p}$ based on data $T_1, T_2, \ldots, T_n$.

7. Verify that the method-of-moments estimator (MME) based on data $T_1, T_2, \ldots, T_n$, say $\tilde{p}$, is the same as the MLE $\hat{p}$ when $\tilde{p} \ge 0$.

8. What is the Cramér-Rao lower bound (call it $v^*(p)$) on the variance for unbiased estimators of p based on $T_1, T_2, \ldots, T_n$?

9. Use the delta method and the above results to obtain approximations for $E[\hat{p}]$ and $\mathrm{Var}(\hat{p})$.

10. Verify that $\sqrt{n}\,(\hat{p} - p)\,/\sqrt{v^*(p)} \xrightarrow{\mathrm{d}} \mathcal{N}(0,1)$. (In this sense, the MLE achieves the Cramér-Rao lower bound in the limit.)

Part III

Now consider the possibility, ϵ_k, of k consecutive heads. Disallow overlaps, so in repeated trials, a subsequent occurrence of ϵ_k cannot build on a previous occurrence of ϵ_k. Letting E_i denote the outcome of the ith coin flip, if $E_{m-1} = 0, E_m = 1, E_{m+1} = 1$, and $E_{m+2} = 1$, with 1 indicating heads, then ϵ_2 occurs at $m+1$ *but not* at $m+2$. Assume, as usual, that trials are independent and the same coin is used for all flips.

11. Let V_k be the time of the first occurrence of ϵ_k. Prove that

$$E[V_k] = 1 + E[V_{k-1}] + (1-p)E[V_k].$$

and use the result to show that $E[V_2] = \frac{p+1}{p^2}$.

12. Redefine u_n and f_n from Part II now for the possibility, ϵ_k, of k consecutive heads. From these probabilities, define $U(s) = \sum_{k=0}^{\infty} u_k s^k$ (which is *not* a pgf) and the pgf $F(s)$ as in **Problem 4**. Show that

$$U(s) = \frac{1}{1 - F(s)}.$$

13. Show that if ϵ_2 is the possibility of two heads in a row, then

$$F(s) = \frac{p^2 s^2}{p^2 s^2 + (1-s)(1+sp)}.$$

14. Find the variance of V_2.

15. Provide an asymptotic formula for an approximate probability that the 38th non-overlapping occurrence of two heads in independent flips of a fair coin occurs after the 178th flip.

Part I

In what follows, you will need the mean and variance of the truncated normal random variable. Let $X \sim \mathcal{N}(\mu, \sigma^2)$, $\phi(\cdot)$ be the standard normal pdf and $\Phi(\cdot)$ be the standard normal cdf. If we condition on X being larger than some constant a, then the variance of the truncated normal random variable is

$$\operatorname{Var}(X \mid X > a) = \sigma^2 \left[1 + \left(\frac{a-\mu}{\sigma}\right) \frac{\phi\left(\frac{a-\mu}{\sigma}\right)}{1-\Phi\left(\frac{a-\mu}{\sigma}\right)} - \left(\frac{\phi\left(\frac{a-\mu}{\sigma}\right)}{1-\Phi\left(\frac{a-\mu}{\sigma}\right)} \right)^2 \right].$$

1. Show that the mean of the truncated normal random variable is

$$E[X \mid X > a] = \mu + \frac{\sigma\phi\left(\frac{a-\mu}{\sigma}\right)}{1-\Phi\left(\frac{a-\mu}{\sigma}\right)}.$$

Solution: In the next two solutions, I use $\phi(\cdot;\mu,\sigma)$ and $\Phi(\cdot;\mu,\sigma)$ to represent the pdf and cdf of $\mathcal{N}(\mu,\sigma^2)$ random variables. Let $C = \frac{1}{1-\Phi(a;\mu,\sigma)}$.

$$\begin{aligned}
E[X \mid X > a] &= \frac{1}{C}\int_a^\infty y\phi(y;\mu,\sigma)dy \\
&= \frac{1}{C}\left[\sigma \int_a^\infty \frac{y-\mu}{\sigma}\phi(y;\mu,\sigma)dy + \mu\int_a^\infty \phi(y;\mu,\sigma)dy\right] \\
&= \frac{1}{C}\left[\sigma\int_{\frac{a-\mu}{\sigma}}^\infty \frac{ze^{-z^2/2}}{\sqrt{2\pi}}dz + \mu C\right] \\
&= \frac{\sigma}{C\sqrt{2\pi}}\int_{\frac{(a-\mu)^2}{2\sigma^2}}^\infty e^{-u}du + \mu \\
&= \frac{\sigma e^{-\frac{(a-\mu)^2}{2\sigma^2}}}{C\sqrt{2\pi}} + \mu = \frac{\sigma\phi\left(\frac{a-\mu}{\sigma}\right)}{1-\Phi\left(\frac{a-\mu}{\sigma}\right)} + \mu.
\end{aligned}$$

Part II

Consider independent flips of a coin and focus on the possibility, ϵ, that "the coin lands heads up." Define

$$u_n = \Pr(\epsilon \text{ occurs at trial } n)$$

and

$$f_n = \Pr(\epsilon \textit{ first} \text{ occurs at trial } n),$$

and take $u_0 \equiv 1$ and $f_0 \equiv 0$.

2. If $p \in [0,1]$ is the probability of heads, what is u_n?

Solution:

$$u_n = p$$

3. What is f_n?

Solution:

$$f_n = (1-p)^{n-1}p$$

4. Define the probability generating function (pgf) $F(s) = \sum_{k=0}^{\infty} f_k s^k$, $|s| \leq 1$ for the wait time, T, until the first occurrence of ϵ. Show that

$$F(s) = \frac{ps}{1-(1-p)s}.$$

Solution:

$$\begin{array}{rcll} F(s) & = & \sum_{k=0}^{\infty} f_k s^k & \text{definition of pgf} \\ & = & \sum_{k=1}^{\infty} f_k s^k & f_0 = 0 \\ & = & \sum_{k=1}^{\infty} (1-p)^{k-1} p s^k & \text{definition of } f_k \\ & = & ps \sum_{k=1}^{\infty} [(1-p)s]^{k-1} & \text{constant factor} \\ & = & ps \sum_{k=0}^{\infty} [(1-p)s]^{k} & \text{change of variable } k \to k-1 \\ & = & \frac{ps}{1-(1-p)s} & \text{geometric series for } (1-p)|s| < 1. \end{array}$$

5. Use $F(s)$ to show that $E[T] = \frac{1}{p}$.

Solution:

$$\begin{aligned} E[T] &= \left.\frac{dF(s)}{ds}\right|_{s=1} \\ &= \left.\left(\frac{p}{1-(1-p)s} + \frac{ps(1-p)}{[1-(1-p)s]^2}\right)\right|_{s=1} \\ &= \frac{p}{p} + \frac{p-p^2}{p^2} = \frac{1}{p}. \end{aligned}$$

Suppose you are able to collect data on the waiting time for the first head in independent coin flipping trials. Unfortunately, your data recording device is broken so it only registers a value when the first head occurs on the first or second trial. Thus $T_1, T_2, \ldots, T_n \in \{1, 2\}$ are iid random variables with the distribution of **Problem 3** truncated above at 2. (Truncation means you observe *no result* when the first head occurs at trial 3 or beyond. All recorded observations satisfy $T_i \leq 2$.) **Problems 6–10** are based on the data $T_1, T_2, \ldots, T_n$.

6. Find the MLE $\hat{p}$ based on data $T_1, T_2, \ldots, T_n$.

Solution: For $p \in (0, 1)$ the truncated probability mass function is

$$P(T_i = t \mid T_i \leq 2) = \frac{\mathbb{1}\{t \leq 2\}(1-p)^{t-1}p}{1-(1-p)^2} = \frac{\mathbb{1}\{t \leq 2\}(1-p)^{t-1}}{2-p},$$

and the likelihood and log likelihood are

$$\begin{aligned} L(p \mid t_1, t_2, \ldots, t_n) &= \prod_{i=1}^{n} \left[\frac{(1-p)^{t_i-1}}{2-p}\right] \\ l(p \mid \boldsymbol{t}) &= n\left[(\bar{t}-1)\ln(1-p) - \ln(2-p)\right]. \end{aligned}$$

Taking derivatives and solving the score function for p', we find

$$\begin{aligned} \frac{dl(p \mid \boldsymbol{t})}{dp} &= -n\left[\frac{\bar{t}-1}{1-p} - \frac{1}{2-p}\right] \\ 0 &= (\bar{t}-1)(2-p') - (1-p') \\ p' &= \frac{3-2\bar{t}}{2-\bar{t}}, \end{aligned} \tag{1}$$

but there is no solution when $\bar{t} = 2$ and the solution can be negative for $\bar{t} > 1.5$. First we check that this solution yields a maximum when $\bar{t} < 2$,

$$\begin{aligned} \frac{d^2 l(p \mid \boldsymbol{t})}{dt^2} &= -n\left[\frac{\bar{t}-1}{(1-p)^2} - \frac{1}{(2-p)^2}\right] \\ &= \frac{-n}{(1-p)^2(2-p)^2}\left[(\bar{t}-1)(2-p)^2 - (1-p)^2\right], \end{aligned}$$

which will be negative at p' if

$$\begin{aligned} (\bar{t}-1)(2-p)^2 - (1-p)^2\big|_{p=p'} &> 0 \\ (\bar{t}-1)\left(2 - \frac{3-2\bar{t}}{2-\bar{t}}\right)^2 - \left(1 - \frac{3-2\bar{t}}{2-\bar{t}}\right)^2 &> 0 \\ (\bar{t}-1)\left(\frac{1}{2-\bar{t}}\right)^2 - \left(\frac{\bar{t}-1}{2-\bar{t}}\right)^2 &> 0 \end{aligned}$$

$$\frac{\bar{t}-1}{2-\bar{t}} \quad > \quad 0,$$

which is satisfied for all $1 < \bar{t} < 2$. When $\bar{t} = 1$, $l(p \mid \boldsymbol{t})$ is increasing on $[0, 1]$, so $\hat{p} = 1$, which is consistent with the estimator (1). When $\bar{t} > 1.5$, $\hat{p}' < 0$, but $l(p \mid \boldsymbol{t})$ is decreasing on $[0, 1]$, and the maximum is at $\hat{p} = 0$. Overall, we conclude the MLE is

$$\hat{p} = \max\left\{0, \frac{3-2\bar{t}}{2-\bar{t}}\right\}.$$

7. Verify that the method-of-moments estimator (MME) based on data $T_1, T_2, \ldots, T_n$, say $\tilde{p}$, is the same as the MLE $\hat{p}$ when $\tilde{p} \geq 0$.

Solution: We need the expectation of our truncated random variable.

$$E[T] \quad = \quad \frac{1}{2-p} + \frac{2(1-p)}{2-p} = \frac{3-2p}{2-p}.$$

Set this equal to the sample mean $\bar{t}$ to find the MME is

$$\tilde{p} = \frac{3-2\bar{t}}{2-\bar{t}}.$$

Note, this estimator may be negative.

8. What is the Cramér-Rao lower bound (call it $v^*(p)$) on the variance for unbiased estimators of p based on $T_1, T_2, \ldots, T_n$?

Solution:

$$\begin{aligned}
-\frac{d^2 l(p \mid T)}{dt^2} &= \frac{T-1}{(1-p)^2} - \frac{1}{(2-p)^2} \\
E\left[-\frac{d^2 l(p \mid T)}{dt^2}\right] &= \frac{E[T]-1}{(1-p)^2} - \frac{1}{(2-p)^2} \\
&= \frac{1}{(2-p)(1-p)} - \frac{1}{(2-p)^2} = \frac{1}{(2-p)^2(1-p)}
\end{aligned}$$

The Cramér-Rao lower bound on the variance for unbiased estimators of p is hence

$$v^*(p) = (2-p)^2(1-p).$$

9. Use the delta method and the above results to obtain approximations for $E[\hat{p}]$ and

$\text{Var}(\hat{p})$.

10. Verify that $\sqrt{n}\,(\hat{p}-p)\,/\sqrt{v^*(p)} \xrightarrow{\text{d}} \mathcal{N}(0,1)$. (In this sense, the MLE achieves the Cramér-Rao lower bound in the limit.)

Solution:

There is a singularity at $x = 2$, and I am not quite sure of the consequences even though in our application we stay well away from it.

By the Central Limit Theorem, we have

$$\sqrt{n}\left(\overline{T} - \mu_T\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma_T^2\right),$$

where $E[T_1] = \mu_T = \frac{3-2p}{2-p}$ from **Problem 7** and $\text{Var}(T_1) = \sigma_T^2$ is to be computed. The result follows so long as σ_T^2 is finite. Since

$$E[T_1^2] \quad = \quad \frac{1}{2-p} + \frac{4(1-p)}{2-p} = \frac{5-4p}{2-p},$$

we have

$$\sigma_T^2 \quad = \quad E[T_1^2] - \left(E[T_t]\right)^2 = \frac{5-4p}{2-p} - \left(\frac{3-2p}{2-p}\right)^2 = \frac{1-p}{(2-p)^2},$$

which is finite for $p \in (0,1)$.
Consider transform $g(x) = \frac{3-2x}{2-x}$, whose derivative is

$$g'(x) \quad = \quad \frac{-2}{2-x} + \frac{(3-2x)}{(2-x)^2} = \frac{-1}{(2-x)^2}.$$

By the delta method, the asymptotic variance of $g(\overline{T})$ is

$$\text{Var}\left[g(\overline{T})\right] \quad = \quad \frac{1-p}{n(2-p)^2}\left|g'(\mu)\right|^2 = \frac{1-p}{n(2-p)^2\left(2-\frac{3-2p}{2-p}\right)^4} = \frac{1}{n}(1-p)(2-p)^2,$$

and since $g(\cdot)$ is an involution ($g^{-1}(x) = g(x)$), we have $g(\mu_T) = p$ and

$$\sqrt{n}\left(\frac{3-2\overline{T}}{2-\overline{T}} - p\right) \xrightarrow{d} \mathcal{N}\left(0, (1-p)(2-p)^2\right).$$

Therefore, our MLE $\hat{p} = \max\left\{0, \frac{3-2\overline{T}}{2-\overline{T}}\right\}$ approximately follows a left 0-truncated normal with (pre-truncation) mean and variance

$$\mu \quad = \quad p$$

$$\sigma^2 = \frac{(1-p)(2-p)^2}{n}.$$

Let $\alpha = \frac{-\mu}{\sigma} = \frac{-\sqrt{n}p}{(2-p)\sqrt{2-p}}$. The mean and variance after truncation are

$$\begin{aligned} E[\hat{p}] &= p + \frac{\phi(\alpha)(2-p)\sqrt{1-p}}{\sqrt{n}\,[1-\Phi(\alpha)]} \\ \mathrm{Var}(\hat{p}) &= \frac{(1-p)(2-p)^2}{n}\left[1 - \frac{\sqrt{n}p\phi(\alpha)}{[1-\Phi(\alpha)]\,(2-p)\sqrt{1-p}} - \left(\frac{\phi(\alpha)}{1-\Phi(\alpha)}\right)^2\right] \end{aligned}$$

The mean squared error is

$$MSE(\hat{p}) = \mathrm{Var}(\hat{p}) + \mathrm{Bias}(\hat{p})^2 = \frac{(1-p)(2-p)^2}{n}.$$

In the limit $n \to \infty$, the MLE is unbiased because the second term vanishes and so the variance approaches the Cramér-Rao lower bound.

Part III

Now consider the possibility, ϵ_k, of k consecutive heads. Disallow overlaps, so in repeated trials, a subsequent occurrence of ϵ_k cannot build on a previous occurrence of ϵ_k. Letting E_i denote the outcome of the ith coin flip, if $E_{m-1} = 0, E_m = 1, E_{m+1} = 1$, and $E_{m+2} = 1$, with 1 indicating heads, then ϵ_2 occurs at $m+1$ *but not* at $m+2$. Assume, as usual, that trials are independent and the same coin is used for all flips.

11. Let V_k be the time of the first occurrence of ϵ_k. Prove that

$$E[V_k] = 1 + E[V_{k-1}] + (1-p)E[V_k].$$

and use the result to show that $E[V_2] = \frac{p+1}{p^2}$.

Solution: There are two mutually exclusive, exhaustive events: (1) either the first run of k heads occurs one time unit after the first run of $k-1$ heads with probability p, or (2) it doesn't, one time step is spent on a tail with probability $1-p$, and the wait starts all over again. Application of the law of total expectation produces

$$E[V_k] = p\,(1 + E[V_{k-1}]) + (1-p)\,(1 + E[V_{k-1}] + E[V_k])\,,$$

which trivially simplifies to the recurrence given.

To solve the equation, we have $E[V_0] = 0$, $E[V_1] = p$, and by the recurrence

$$E[V_2] = 1 + p^2 + (1-p)(p + E[V_2]),$$

which is solved as

$$\begin{aligned}[1-(1-p)]E[V_2] &= 1+p^2+p-p^2 \\ E[V_2] &= \frac{1+p}{p}.\end{aligned}$$

One might also find the solution by hypothesizing generating function $h(s) = \sum_{k=1}^{\infty} E[V_k]s^k$ and solving for the coefficients consistent with the recurrence.

12. Redefine u_n and f_n from Part II now for the possibility, ϵ_k, of k consecutive heads. From these probabilities, define $U(s) = \sum_{k=0}^{\infty} u_k s^k$ (which is *not* a pgf) and the pgf $F(s)$ as in **Problem 4**. Show that

$$U(s) = \frac{1}{1-F(s)}.$$

Solution: By assumptions of the question,

$$\Pr(\epsilon_k \text{ occurs at trial } n+m \mid \epsilon_k \text{ occurs at trial } n) = \Pr(\epsilon_k \text{ occurs at trial } m). \quad (2)$$

Therefore, if ϵ_k occurs *first* at trial n, then the probability it occurs again at trial $n+m$ is given by u_m since the sequence $E_{n+1}, E_{n+2}, \ldots, E_{n+m}$ has the exact same properties as $E_1, E_2, \ldots, E_m$ under the assumptions (of independence and homogeneity – same coin). The joint probability of a first occurrence at trial n and recurrence at $n+m$ is thus $f_n u_m$. By the Law of Total Probability, the probability of an occurrence at trial $n > 0$ is

$$u_n = f_1 u_{n-1} + f_2 u_{n-2} + \cdots + f_n.$$

Since $u_0 \equiv 1$, we can multiply the last term by u_0. The right-hand-side are the coefficients of s^n in $F(s)U(s)$, $n > 0$. Therefore,

$$U(s) - 1 = F(s)U(s).$$

Since $1 - F(s) \neq 0$ for $|s| < 1$, we can solve for $U(s)$ to get the desired result.

13. Show that if ϵ_2 is the possibility of two heads in a row, then

$$F(s) = \frac{p^2 s^2}{p^2 s^2 + (1-s)(1+sp)}.$$

Solution: We have $u_0 = 1$ and $u_1 = 0$. For $n > 1$, we have $E_{n-1} = E_n = 1$ with probability p^2. There are two mutually exclusive possibilities: (1) either ϵ occurred at trial $n-1$ or (2) ϵ occurred at trial n. Thus,

$$u_{n-1}p + u_n = p^2$$

Multiply both sides by s^n and sum over $n > 1$.

$$\begin{array}{rcll}
\sum_{n=2}^{\infty} u_{n-1}ps^n + \sum_{n=2}^{\infty} u_n s^n &=& \sum_{n=2}^{\infty} p^2 s^n & \\
sp\sum_{n=2}^{\infty} u_{n-1}s^{n-1} + \sum_{n=2}^{\infty} u_n s^n &=& p^2s^2\sum_{n=2}^{\infty} s^{n-2} & \text{constant term, both sides} \\
sp\sum_{n=1}^{\infty} u_n s^n + \sum_{n=1}^{\infty} u_n s^n &=& p^2s^2\sum_{n=0}^{\infty} s^n & \text{change of variable and } u_1 = 0 \\
(sp+1)\left[U(s)-1\right] &=& \frac{p^2s^2}{1-s} & U(s) = 1 + \sum_{n=1}^{\infty} u_n s^n \\
&&& \quad \text{and geometric series}
\end{array}$$

Thus,

$$U(s) = \frac{p^2s^2}{(1-s)(1+sp)} + 1 = \frac{p^2s^2 + (1-s)(1+sp)}{(1-s)(1+sp)}.$$

Then,

$$\begin{array}{rcll}
F(s) &=& \frac{U(s)-1}{U(s)} & \text{from (??)} \\
&=& \frac{p^2s^2}{p^2s^2+(1-s)(1+sp)} & \text{from above}
\end{array}$$

14. Find the variance of V_2.

Solution: The following confirms are previous work in **Problem 11**.

$$\begin{array}{rcll}
F(s)\left[p^2s^2 + (1-s)(1+sp)\right] &=& p^2s^2 & \text{definition of } F(s) \\
F'(s)\left[p^2s^2 + (1-s)(1+sp)\right] + F(s)\left[2p^2s - 1 + p - 2sp\right] &=& 2p^2s & \text{derivative } \frac{d}{ds} \\
E[T]p^2 + 2p^2 - 1 - p &=& 2p^2 & F'(1) = E[T] \\
\mu := E[T] &=& \frac{1+p}{p^2}. &
\end{array}$$

$$\begin{array}{rcll}
F''(s)\left[p^2s^2 + (1-s)(1+sp)\right] &&& \\
+\, F'(s)\left[2p^2s - (1+sp) + p(1-s)\right] &&& \text{derivative } \frac{d^2}{ds^2} \\
+\, F'(s)\left[2p^2s - 1 + p - 2sp\right] + F(s)\left[2p^2 - 2p\right] &=& 2p^2 & \\
E[T(T-1)]p^2 + E[T]\left[4p^2 - 2p - 2\right] - 2p(1-p) &=& 2p^2 & F''(1) = E[T(T-1)] \\
E[T^2]p^2 + \frac{1+p}{p^2}\left[3p^2 - 2p - 2\right] &=& 2p & \text{simplify \& } \textbf{Problem 11}
\end{array}$$

yields

$$E[T^2] = \frac{-p^3 - p^2 + 4p + 2}{p^4}.$$

The variance is

$$\sigma^2 := E[T^2] - \frac{(1+p)^2}{p^4} = -\frac{p^3 + 2p^2 - 2p - 1}{p^4}.$$

15. Provide an asymptotic formula for an approximate probability that the 38th non-overlapping occurrence of two heads in independent flips of a fair coin occurs after the 178th flip.

Solution: We have observed 38 realizations of T. We know that

$$\sqrt{n}\left(\overline{T} - \frac{1+p}{p^2}\right) \xrightarrow{d} \mathcal{N}\left(0, -\frac{p^3 + 2p^2 - 2p - 1}{p^4}\right)$$

We seek the probability

$$\Pr\left(\overline{T} > \frac{178}{38}\right) \approx 1 - \Phi\left(\frac{178}{38}; \frac{(1+p)}{p^2}, -\frac{p^3 + 2p^2 - 2p - 1}{38p^4}\right),$$

which is 0.9581211.