1. (Previously Problem 5 of PS 4) Suppose that $\Theta$ is a random parameter with prior density

$$w(\theta) = \begin{cases} \alpha e^{-\alpha \theta} & \theta \geq 0 \\ 0 & \theta < 0 \end{cases}$$

where $\alpha > 0$ is known. Suppose our observation $Y$ is a Poisson random variable with rate $\Theta$. That is

$$p_{\theta}(y) = P(Y = y | \Theta = \theta) = \frac{\theta^y e^{-\theta}}{y!}$$

for $y = 0, 1, 2, \ldots$. Find the MMSE and MAP estimates of $\Theta$ based upon $Y$. How would you find the MMLAE estimate?

2. (Previously Problem 6 of PS 4) Suppose that $\theta > 0$ is a parameter of interest and that given $\theta$, $\{Y_k : 1 \leq k \leq n\}$ is a set of iid observations with marginal distribution function

$$F_\theta(y) = [F(y)]^{1/\theta}, \quad y \in \mathcal{R}$$

where $F$ is a known distribution function with pdf $f$.

(a) Show that

$$\hat{\theta}_{MV}(y) = -\frac{1}{n} \sum_{k=1}^{n} \log F(y_k)$$

is a minimum variance unbiased estimate of $\theta$.

(b) Suppose now that $\theta$ is replaced by a random variable $\Theta$ drawn at random using the prior density

$$w(\theta) = c^m e^{-c/\theta} \frac{1}{\Gamma(m) \theta^{m+1}} \quad \theta > 0$$

where $c$ and $m$ are positive constants. Use the fact that $E\{\Theta\} = c/(m+1)$ to show that the MMSE estimator of $\Theta$ from $Y$ is

$$\hat{\theta}_{MMSE}(y) = \frac{c - \sum_{k=1}^{n} \log F(y_k)}{m + n - 1}.$$ 

(c) Compare $\hat{\theta}_{MV}$ and $\hat{\theta}_{MMSE}$ with regard to the role of prior information.

3. Consider the observation model

$$Y_k = \sqrt{\theta} s_k R_k + N_k \quad 1 \leq k \leq n$$

where $\{s_k : 1 \leq k \leq n\}$ is a known signal, the $N_k$ and $R_k$ are iid $\mathcal{N}(0,1)$ random variables, and $\theta \geq 0$ is an unknown parameter.

(a) Find the likelihood equation for estimating $\theta$ from $\{Y_k : 1 \leq k \leq n\}$.

(b) Find the Cramer–Rao lower bound on the variance of unbiased estimates of $\theta$.

(c) Suppose that $s_k$ is a sequence of $\pm 1$'s. Find the MLE of $\theta$ explicitly.

(d) Compute the bias and variance of the estimate from (c) and compare with the CRB.

4. Consider the following questions concerning unbiased estimators in the classical setting, i.e., with a deterministic parameter.
(a) **Unbiased estimators don’t always exist.** Let $X$ be a binomial random variable with parameter $\theta$ ($0 \leq \theta \leq 1$), i.e., the pmf of $X$ assuming $\theta$ is the true parameter is

$$p_\theta(x) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \quad x = 0, 1, \ldots, n.$$  

Let $g(\theta) = \theta^p$. Show that there exists an unbiased estimator of $g(\theta)$ if $p$ is an integer such that $0 \leq p \leq n$ and there does not if $p$ is an integer such that $p > n$. In the first case, you do not need to solve for the unbiased estimator to prove its existence.

(b) **Unbiased estimators don’t always make sense.** Let $X$ be a discrete random variable taking values on the positive integers with the truncated Poisson distribution ($\theta > 0$)

$$p_\theta(x) = \frac{e^{-\theta}\theta^x}{(1-e^{-\theta})x!} \quad x = 1, 2, 3, \ldots.$$  

Show that an unbiased estimator of $g(\theta) = 1 - e^{-\theta}$ must be such that it equals 0 if $x$ is odd, and it equals 2 if $x$ is even.

5. Suppose that $Z_1, Z_2, \ldots, Z_m$ are independent, identically-distributed samples from a uniform distribution on the interval $(0, \theta)$, $0 < \theta < \infty$. Let $Y_1, Y_2, \ldots, Y_m$ be the order statistics for the sample. Prove that $Y_m$ is a sufficient statistic.

6. Consider the observation model

$$Z = \frac{1}{\theta} + V$$

where $V$ is a Gaussian random variable with zero mean and variance equal to one. Let $\psi = g(\theta) = 1/\theta$. In this problem we consider the estimation of the unknown parameters $\theta$ and $\psi$.

(a) Assume that $\theta$ is a nonrandom parameter.

(i) Find the maximum likelihood estimator $\hat{\theta}_{ML}$ of $\theta$.

(ii) Find the maximum likelihood estimator $\hat{\psi}_{ML}$ of $\psi$.

(b) Assume that $\theta$ is a realization of a random parameter $\Theta$ with probability density

$$p_\theta(\theta) = \frac{1}{\theta^2 \sqrt{2\pi}} \exp \left\{-\frac{1}{2\theta^2}\right\} \quad \theta \neq 0.$$  

Assume also that $\Theta$ and $V$ are independent.

(i) Find the maximum a posteriori estimator $\hat{\psi}_{MAP}$ of $\Psi = g(\Theta)$.

(ii) Find the maximum a posteriori estimator $\hat{\Theta}_{MAP}$ of $\Theta$.

7. Let $Y = [Y_1, Y_2, \ldots, Y_n]^T$ be a random vector where the individual components are i.i.d. Poisson random variables with parameter $\theta$.

(a) Show that $T(Y) = Y_1 + Y_2 + \cdots + Y_n$ is a complete sufficient statistic for $\theta$. You must explain why it is sufficient and why it is complete.

(b) Show that $T(Y)$ is also Poisson. Start with $n = 2$ and use induction. What is the parameter for the distribution of $T(Y)$?

(c) For any (fixed) integer $k \geq 0$ find a minimum variance unbiased estimate (MVUE) of the probability $P_\theta\{Y_1 = k\}$.

(d) For any (fixed) integer $k \geq 0$ find the maximum likelihood (ML) estimator of the probability $P_\theta\{Y_1 = k\}$.

Is this ML estimator biased?
8. This problem refers back to the previous problem. Recall that \( Y = [Y_1 \ Y_2 \ \cdots \ Y_n]^T \) was a random vector where the individual components were i.i.d. Poisson random variables with parameter \( \theta \). Previously, you showed that \( T(Y) = Y_1 + Y_2 + \cdots + Y_n \) is a complete sufficient statistic for \( \theta \), that \( T(Y) \) is Poisson with parameter \( n \theta \), and you found the MVUE and ML estimators for \( \phi = g(\theta) = P_{\theta}(Y_1 = 0) = e^{-\theta} \).

(actually, a slightly more general result was found before). These estimators were (for \( n > 1 \))

\[
\hat{\phi}_{\text{MVUE}}(y) = \left( \frac{n-1}{n} \right)^{T(y)} \\
\hat{\phi}_{\text{ML}}(y) = \exp \left( -\frac{T(y)}{n} \right)
\]

(a) Find the CRLB for estimating \( \theta \) based upon the \( n \) observations.

(b) The ML estimate for \( \theta \) is \( \hat{\theta}_{\text{ML}}(y) = T(y)/n \). Is \( \hat{\theta}_{\text{ML}}(y) \) unbiased? Is it efficient? Is it MVUE?

(c) Now consider estimation of \( \phi = e^{-\theta} \). Why can we say that \( \hat{\phi}_{\text{ML}} \) is biased without calculation? Calculate the bias directly. From your result show that the ML estimator is asymptotically (as \( n \to \infty \)) unbiased.

(d) Directly calculate the variance of \( \hat{\theta}_{\text{ML}} \) as a function of \( \theta \) and \( n \) and argue that the estimator is consistent.

(e) Find the CRLB for estimating \( \phi = e^{-\theta} \) based upon the \( n \) observations and express it as a function of \( \theta \).

(f) Does the unbiased estimator \( \hat{\phi}_{\text{MVUE}} \) meet the CRLB for finite \( n \)? Answer the question without calculation.

(g) Now calculate the variance of \( \hat{\phi}_{\text{MVUE}} \) expressed as a function of \( \theta \) and \( n \). Form the ratio

\[
\frac{\text{Var}_{\theta}(\hat{\phi}_{\text{MVUE}}(Y))}{\text{the CRLB for } \phi}
\]

and directly explore the efficiency question of (f).
Suppose that $\Theta$ is a random parameter with prior density

$$w(\theta) = \begin{cases} \alpha e^{-\alpha \theta} & \theta \geq 0 \\ 0 & \theta < 0 \end{cases}$$

where $\alpha > 0$ is known. Suppose our observation $Y$ is a Poisson random variable with rate $\Theta$. That is

$$p_\theta(y) = P(Y = y|\Theta = \theta) = \frac{\theta^y e^{-\theta}}{y!}$$

for $y = 0, 1, 2, \ldots$. Find the MMSE and MAP estimates of $\Theta$ based upon $Y$. How would you find the MMAE estimate?
Prior density for $\Theta$ \( w(\theta) = \begin{cases} x e^{-x \theta} & \theta > 0 \\ 0 & \theta < 0 \end{cases} \) \( x > 0 \) is known

Conditional model \( p_\theta(y) = \frac{e^y e^{-\theta}}{y!} \) \( y = 0, 1, 2 \ldots \)

For all of the estimators mentioned we need to compute the cond. density \( w(\theta | y) \). From Bayes formula

\[
w(\theta | y) = \frac{p_\theta(y) w(\theta)}{\int_0^\infty p_\theta(y) w(\theta) \, d\theta}
\]

Need the denominator

\[
\int_0^\infty \frac{e^y e^{-\theta}}{y!} \, d\theta = \frac{x}{y!} \int_0^\infty \theta^{-x-1} e^{-\theta} \, d\theta
\]

\[
= \frac{x}{y!} \int_0^\infty \frac{\phi^y e^{-\phi}}{(1+x)^y} \frac{d\phi}{(1+x)} = \frac{x}{y!} \frac{1}{(1+x)^{y+1}}
\]

\[
= \frac{x}{(1+x)^{y+1}}
\]

\[
\therefore \ w(\theta | y) = \frac{\frac{e^y e^{-\theta}}{y!} \frac{x}{(1+x)^{y+1}}}{\frac{x}{(1+x)^{y+1}}} = \frac{e^y}{y!} \frac{e^{-(\theta+1)\theta}}{(1+x)^{y+1}}
\]

\[
\text{RMSE Compute the conditional mean}
\]

\[
\hat{\Theta}_{\text{RMSE}}(y) = \int_0^\infty \theta w(\theta | y) \, d\theta = \int_0^\infty \frac{\theta}{y!} \frac{e^{-(\theta+1)\theta}}{(1+x)^{y+1}} \, d\theta
\]

\[
= \frac{y+1}{x+1} \quad \text{(using same idea as above)}
\]
MAPP Compute the conditional mode
\[
\hat{\Theta}_{\text{MAP}}(y) = \arg\max_{\Theta > 0} \{ W(\Theta y) \} = \arg\max_{\Theta > 0} \left\{ \Theta^y e^{-x \Theta} \right\}
\]
\[
= \arg\max_{\Theta > 0} \left\{ y \log \Theta - (x+1) \Theta \right\} = \frac{y}{x+1}
\]

ABS By definition \( \hat{\Theta}_{\text{ABS}}(y) \) is the conditional median, the solution of
\[
\int_{\hat{\Theta}_{\text{ABS}}(y)} W(\Theta y) \, d\Theta = \frac{1}{2}
\]
Suppose that $\theta > 0$ is a parameter of interest and that given $\theta$, \{$Y_k : 1 \leq k \leq n$\} is a set of iid observations with marginal distribution function

$$F_\theta(y) = [F(y)]^{1/\theta}, \quad y \in \mathcal{R}$$

where $F$ is a known distribution function with pdf $f$.

(a) Show that

$$\hat{\theta}_{MV}(y) = -\frac{1}{n} \sum_{k=1}^{n} \log F(y_k)$$

is a minimum variance unbiased estimate of $\theta$.

(b) Suppose now that $\theta$ is replaced by a random variable $\Theta$ drawn at random using the prior density

$$w(\theta) = c^m \frac{e^{-c/\theta}}{\Gamma(m)\theta^{m+1}} \quad \theta > 0$$

where $c$ and $m$ are positive constants. Use the fact that $E(\Theta) = c/(m+1)$ to show that the MMSE estimator of $\Theta$ from $Y$ is

$$\hat{\theta}_{MMSE}(y) = \frac{c - \sum_{k=1}^{n} \log F(y_k)}{m + n - 1}.$$ 

(c) Compare $\hat{\theta}_{MV}$ and $\hat{\theta}_{MMSE}$ with regard to the role of prior information.

\(\bigcirc\) Typical: $E\{\Theta\} = \frac{c}{m-1}$

\(\bigstar\) where $c > 0$ and $m > 1$
\[ \theta > 0 \text{ a scalar parameter. } \{ y_k \} \text{ iid observations with marginal dist} \]
\[ F_0(y) = \left( F(y) \right)^{1/\theta} \text{ F a fixed known dist. with density } f. \]

From iid the distribution function of the vector observation \( Y = (Y_1, \ldots, Y_n)^T \) is
\[ P_{\theta}(y) = P_{\theta} \left\{ Y_1 \leq y_1, \ldots, Y_n \leq y_n \right\} = \prod_{k=1}^{n} \left[ F(y_k) \right]^{1/\theta} \]

The density is the derivative of the above. Since \( \frac{\partial}{\partial y} F_0(y) = \frac{\frac{F(y)}{\theta} f(y)}{f(y)} \)

it is easy to see that
\[ P_{\theta}(y) = \frac{1}{\theta^n} \prod_{k=1}^{n} \left[ F(y_k) \right]^{1/\theta} \prod_{k=1}^{n} \left( \frac{f(y_k)}{F(y_k)} \right) \]
\[ = \frac{1}{\theta^n} \exp \left\{ \frac{1}{\theta} \sum_{k=1}^{n} \log F(y_k) \right\} \prod_{k=1}^{n} \left( \frac{f(y_k)}{F(y_k)} \right) \]

This is a one parameter exponential family with statistic \( T(y) = \sum_{k=1}^{n} \log F(y_k) \). From completeness theorem for exponential families and fact that \((0, \infty)\) is open interval we have that \( T(y) \) is complete sufficient.

(a) If an unbiased estimator based on \( T(y) \) can be found it must be uniformly minimum variance by RB Thm. Note
\[ E_{\theta} \left\{ \sum_{k=1}^{n} \log F(y_k) \right\} = n E_{\theta} \left\{ \log F(Y_1) \right\} = \frac{n}{\theta} \int_{-\infty}^{\infty} \log F(y_i) \left[ F(y_i) \right]^{1/\theta} f(y_i) \, dy \]

To evaluate the above integral note:
\[ d \log F(y_i) = \frac{f(y_i)}{F(y_i)} \, dy, \quad \left[ F(y_i) \right]^{1/\theta} = e^{\frac{1}{\theta} \log F(y_i)} \]

and substitute \( x = \log F(y_i) \)
\[ E_{\theta} \left\{ T(y) \right\} = \frac{n}{\theta} \int_{-\infty}^{\infty} x e^{x/\theta} \, dx = -n \theta \]
\[ \therefore \hat{\theta}_{MV}(y) = -\frac{1}{n} \sum_{k=1}^{n} \log F(y_k) \]
(b) $\phi$ is a realization of a random variable $\Theta$ drawn according to

$$W(\phi) = \frac{c^m e^{-c/\Theta}}{\Gamma(m) \Theta^{m+1}} \quad \Theta > 0$$

where $c,m$ pos. constants.

First check $E\{\Theta\}$ since instructor cannot be trusted.

$$E\{\Theta\} = \int_0^\infty \theta w(\theta) d\theta = \frac{c^m}{\Gamma(m)} \int_0^\infty \theta^m e^{-c/\Theta} d\theta$$

$$= \frac{c^m}{\Gamma(m)} \int_0^\infty x^{m-1} e^{-x} dx = \frac{c^m}{\Gamma(m)} \Gamma(m-1) = \frac{c}{m-1}$$

subst.

$x = c/\Theta$

To solve the problem we need to find $W(\phi | y)$. For simplicity let

$$h(y) = \prod_{k=1}^n \left( \frac{f(y_k)}{F(y_k)} \right) \quad T(y) = \sum_{k=1}^n \log F(y_k)$$

Then

$$p(\phi | y) w(\phi) = \frac{1}{\Theta^n} e^{T(\phi) / \Theta} h(y) \frac{c^m e^{-c/\Theta}}{\Gamma(m) \Theta^{m+1}}$$

$$= \frac{c^m h(y)}{\Gamma(m) \Theta^n \Theta^{m+1}} e^{-(c - T(y))/\Theta}$$

We could continue to evaluate $W(\phi | y)$ by computing $p(y)$ which is the integral of the term above. This turns out to be useless work because only the terms in $\Theta$ are of interest to us as far as computing the conditional mean:

$$W(\phi | y) = \frac{c^m h(y)}{\Gamma(m) p(y)} \frac{e^{-(c - T(y))/\Theta}}{\Theta^{n+m+1}}$$

Compare to

$$W(\phi) = \left( \frac{c^m}{\Gamma(m)} \right) \frac{e^{-c/\Theta}}{\Theta^{m+1}}$$

These only serve as normalizing constants.

Note that $c - T(y) > 0$ for all $y$ so the two densities above are of the same form. From inspection

$$E\{\Theta | y = y\} = \hat{\Theta}_{\text{MMSE}}(y) = \frac{C - T(y)}{n+m-1} = \frac{C - \sum_{k=1}^n \log F(y_k)}{n+m-1}$$
(c) In this problem the prior and posterior distributions have the same form, the only change is that the parameters of the dist. are updated as observations are made. A prior with this property is called a reproducing prior.

The prior parameters \( c \) and \( m \) can be thought of as coming from an earlier sample of size \( m \). As \( n \) becomes large relative to \( m \) the importance of the prior parameters in the estimate diminishes.

\[
\begin{align*}
\text{if } n \gg m & \quad \hat{\Theta}_{\text{MSE}} \approx \hat{\Theta}_{\text{HV}} \\
\text{if } m \gg n & \quad \hat{\Theta}_{\text{MSE}} \approx \frac{c}{m-1} \quad \text{(prior mean)}
\end{align*}
\]

Between the two extremes a balance is struck between prior information and observed information.
Consider the observation model

\[ Y_k = \sqrt{\theta} s_k R_k + N_k \quad 1 \leq k \leq n \]

where \( \{s_k : 1 \leq k \leq n\} \) is a known signal, the \( N_k \) and \( R_k \) are iid \( \mathcal{N}(0, 1) \) random variables, and \( \theta \geq 0 \) is an unknown parameter.

(a) Find the likelihood equation for estimating \( \theta \) from \( \{Y_k : 1 \leq k \leq n\} \).

(b) Find the Cramer–Rao lower bound on the variance of unbiased estimates of \( \theta \).

(c) Suppose that \( s_k \) is a sequence of \( \pm 1 \)'s. Find the MLE of \( \theta \) explicitly.

(d) Compute the bias and variance of the estimate from (c) and compare with the CRB.
Note that $Y_1, \ldots, Y_n$ are independent with $Y_k \sim \eta(\omega, 1 + \Theta S_k^2)$. The log likelihood is
\[
\log p_\Theta(y) = \sum_{k=1}^{n} \log \left\{ \frac{1}{\sqrt{2\pi \Theta S_k^2}} \frac{-y_k^2/2(1 + \Theta S_k^2)}{e} \right\} \\
= \sum_{k=1}^{n} \left\{ -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(1 + \Theta S_k^2) - \frac{y_k^2}{2(1 + \Theta S_k^2)} \right\}
\]

(a) To find the likelihood equation, we take the derivative of the above with respect to $\Theta$ and set equal to zero.
\[
\frac{\partial}{\partial \Theta} \log p_\Theta(y) = -\frac{1}{2} \sum_{k=1}^{n} \left\{ \frac{s_k^2}{1 + \Theta S_k^2} - \frac{y_k^2 s_k^2}{(1 + \Theta S_k^2)^2} \right\} = 0
\]

\[\hat{\Theta}_{ML}(y)\] is among the solutions of this.

(b) \[
\frac{\partial^2}{\partial \Theta^2} \log p_\Theta(y) = \sum_{k=1}^{n} \left\{ -\frac{s_k^4 y_k^2}{(1 + \Theta S_k^2)^3} + \frac{s_k^4}{2(1 + \Theta S_k^2)^2} \right\}
\]

\[I_\Theta = -E_\Theta \left\{ \frac{\partial^2}{\partial \Theta^2} \log p_\Theta(Y) \right\}\]
\[
= \sum_{k=1}^{n} \left\{ \frac{s_k^4 E_\Theta \{Y_k^2\}}{(1 + \Theta S_k^2)^3} - \frac{s_k^4}{2(1 + \Theta S_k^2)^2} \right\}\]
\[
= \frac{1}{2} \sum_{k=1}^{n} \frac{s_k^4}{(1 + \Theta S_k^2)^2}
\]

\[CRLB = \frac{1}{I_\Theta} = \frac{2}{\sum_{k=1}^{n} \frac{s_k^4}{(1 + \Theta S_k^2)^2}}
\]
(c) $S_k^2 = 1$ for all $k$. and the likelihood equation is

$$\frac{\sum_{k=1}^{n} y_k^2 - 1 - \theta}{(1 + \theta^2)} = 0 \Rightarrow \hat{\theta}_{ML}(\mathbf{y}) = \left(\frac{1}{n} \sum_{k=1}^{n} y_k^2 \right) - 1$$

(d) $E_\theta \left\{ \hat{\theta}_{ML}(\mathbf{y}) \right\} = \frac{1}{n} \sum_{k=1}^{n} E_\theta \left\{ Y_k^2 \right\} - 1$

$$= \frac{1}{n} \sum_{k=1}^{n} (1 + \theta) - 1 = \theta \text{ unbiased.}$$

Since the $Y_k$'s are independent.

$\text{Var}_\theta \left\{ \hat{\theta}_{ML}(\mathbf{y}) \right\} = \frac{1}{n^2} \sum_{k=1}^{n} \text{Var}_\theta \left\{ Y_k^2 \right\} = \frac{1}{n^2} \sum_{k=1}^{n} 2(1 + \theta)^2$

$$= \frac{2 \cdot (1 + \theta)^2}{n}$$

$= \text{CRLB for this Case } (S_k^2 = 1)$. 
Consider the following questions concerning unbiased estimators in the classical setting, i.e., with a deterministic parameter.

(a) **Unbiased estimators don’t always exist.** Let $X$ be a binomial random variable with parameter $\theta$ ($0 \leq \theta \leq 1$), i.e., the pmf of $X$ assuming $\theta$ is the true parameter is

$$p_\theta(x) = \binom{n}{x} \theta^x (1 - \theta)^{n-x} \quad x = 0, 1, \ldots, n.$$ 

Let $g(\theta) = \theta^p$. Show that there exists an unbiased estimator of $g(\theta)$ if $p$ is an integer such that $0 \leq p \leq n$ and there does not if $p$ is an integer such that $p > n$. In the first case, you do not need to solve for the unbiased estimator to prove its existence.

(b) **Unbiased estimators don’t always make sense.** Let $X$ be a discrete random variable taking values on the positive integers with the truncated Poisson distribution ($\theta > 0$)

$$p_\theta(x) = \frac{e^{-\theta} \theta^x}{(1 - e^{-\theta}) x!} \quad x = 1, 2, 3, \ldots$$

Show that an unbiased estimator of $g(\theta) = 1 - e^{-\theta}$ must be such that it equals 0 if $x$ is odd, and it equals 2 if $x$ is even.
(a) Unbiased Estimators don't always exist.

Suppose \( T(X) \) taking values to \( t_1, \ldots, t_n \) when \( X = \theta, \ldots, \rho \) respectively, is unbiased for \( g(\theta) = \theta^p \). Then must have

\[
E_\theta T(X) = \sum_{k=0}^{n} \binom{n}{k} t_k \theta^k (1-\theta)^{n-k} = \theta^p \quad \forall \theta \in [0, 1]
\]

Say \( p > n \)

Then unbiased condition implies

\[
\sum_{k=0}^{n} \binom{n}{k} t_k \theta^k (1-\theta)^{n-k} - \theta^p \equiv 0 \quad 0 \leq \theta \leq 1
\]

a polynomial of degree \( p \)

with "-1" (ie non-zero) leading coeff.

has an uncountable # of zeros,

hence must be identically zero,

i.e. contradiction

\( \implies \) can be no solution for the \( t_k \).

Say \( 0 \leq p \leq n \) (only \( p > 0 \) makes any sense).

Now the polynomial above is of degree \( \leq n \). Must show there is a solution for the \( t_k \) which make the poly.

coeffs. identically zero. Have

\[
\sum_{k=0}^{n} \binom{n}{k} t_k \theta^k \left[ \sum_{\ell=0}^{n-k} \binom{n-k}{\ell} (-1)^{n-k-\ell} \theta^\ell \right] - \theta^p \equiv 0
\]
$$L.H.S. = \sum_{k=0}^{n} \sum_{l=0}^{n-k} \binom{n}{k} \binom{n-k}{l} (-1)^{n-k-l} t_k \theta^k \frac{A_{k+l}}{A} - \theta^p$$

Let \( q = k + l \) and sum over \( q \) from 0 to \( n \).

$$L.H.S. = \sum_{q=0}^{n} \sum_{k=0}^{n-q} A_{k,q-k} - \theta^p$$

where

$$A_{k,q-k} = \binom{n}{k} \binom{n-k}{q-k} (-1)^{n-q} t_k \theta^q$$

Can now write in terms of the polynomial coefficients

$$L.H.S. = \sum_{q=0}^{n} \left[ \sum_{k=0}^{n-q} \binom{n}{k} \binom{n-k}{q-k} (-1)^{n-q} t_k \right] \theta^q - \theta^p$$

In order that \( L.H.S. \equiv 0 \) for \( 0 \leq \theta \leq 1 \) the coeff. must be zero. So,

for \( q \neq p \): \( \sum_{k=0}^{n-q} \binom{n}{k} \binom{n-k}{q-k} t_k = 0 \)

for \( q = p \): \( (-1)^p \sum_{k=0}^{n-p} \binom{n}{k} \binom{n-k}{p-k} t_k - 1 = 0 \)

\( \Rightarrow \) Must show that this system of linear equations can be solved for \( t_k \).
As an example say $0 < p < n$ and define $a_{jk}^p = \binom{n}{k}(q_{jk})^{n-k}$

Then

$q = 0:\quad a_{00}t_0 = 0 \quad (a_{00} = 1 \Rightarrow t_0 = 0)$

$q = 1:\quad a_{10}t_0 + a_{11}t_1 = 0 \quad (a_{11} = n \Rightarrow t_1 = 0)$

$\vdots$

$q = p - 1:\quad a_{p-1,0}t_0 + a_{p-1,1}t_1 + \cdots + a_{p-1,p-1}t_{p-1} = 0 \quad (a_{p-1,p-1} \neq 0 \Rightarrow t_{p-1} = 0)$

$q = p:\quad \text{(use fact that } t_k = 0 \text{ for } 0 \leq k \leq p - 1)$

\[-1 \frac{p}{p} a_{pp}t_p = 1 \quad \Rightarrow \quad t_p = (-1)^{n-p}/a_{pp}\]

$\vdots$

Can finish the solution in the same manner though from here on the $t_k$'s will most likely not be zero.
(b) **Unbiased estimators don't always make sense.**

Suppose $T(X)$ is unbiased est. of $g(\theta) = 1 - e^{-\theta}$. Then

$$\sum_{k=1}^{\infty} t_k \frac{e^{-\theta} \theta^k}{k! (1 - e^{-\theta})} = 1 - e^{-\theta} \quad \forall \theta > 0$$

(where $t_k$ are values taken by $T(X)$). Thus

$$\sum_{k=1}^{\infty} t_k \frac{\theta^k}{k!} = e^\theta (1 - e^{-\theta})^2 = e^\theta (1 - 2e^{-\theta} + e^{-2\theta})$$

$$= e^\theta - 2 + e^{-\theta}$$

$$= \sum_{\ell=0}^{\infty} \frac{\theta^\ell}{\ell!} - 2 + \sum_{\ell=0}^{\infty} \frac{(-1)^\ell \theta^\ell}{2!}$$

$$= \sum_{\ell=1}^{\infty} \frac{[1 + (-1)^\ell]}{2!} \theta^\ell$$

Since equality holds $\forall \theta > 1$, the two power series must have same coeffs. Hence

$$t_k = \left[1 + (-1)^k\right] = \begin{cases} 0 & k \text{ odd} \\ 2 & k \text{ even} \end{cases}$$
Suppose that $Z_1, Z_2, \ldots, Z_m$ are independent, identically-distributed samples from a uniform distribution on the interval $(0, \theta)$, $0 < \theta < \infty$. Let $Y_1, Y_2, \ldots, Y_m$ be the order statistics for the sample. Prove that $Y_m$ is a sufficient statistic.
\{Z_i : 1 \leq i \leq m\} iid with common density

\[ p_\theta(Z_i) = \begin{cases} \frac{1}{\theta} & 0 \leq Z_i \leq \theta \\ 0 & \text{else} \end{cases} \]

This density can also be written

\[ p_\theta(Z_i) = \frac{1}{\theta} \cdot 1_{[0,\theta]}(Z_i) = \frac{1}{\theta} \cdot 1_{[Z_i,\infty)}(\theta) \]

where

\[ 1_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases} \quad \text{(the indicator funct. of set A)} \]

\[ p_\theta(Z) = \prod_{i=1}^{m} p_\theta(Z_i) = \frac{1}{\theta^m} \prod_{i=1}^{m} 1_{[Z_i,\infty)}(\theta) \]

\[ z = [z_1 \ldots z_m]^T \]

\[ = \begin{cases} \frac{1}{\theta^m} & \text{if } \max\{z_i\} \leq \theta \\ 0 & \text{else} \end{cases} \]

The \{Y_i : 1 \leq i \leq m\} are the order statistics of the sample. In particular

\[ Y_m = \max\{Z_i\} \]

\[ p_\theta(Z) = \frac{1}{\theta^m} \cdot 1_{[y_m,\infty)}(\theta) = g_\theta(y_m) h(z) \]

By Fisher–Neyman Factorization Thm, \(Y_m\) is a suff. statistic.
Consider the observation model

\[ Z = \frac{1}{\theta} + V \]

where \( V \) is a Gaussian random variable with zero mean and variance equal to one. Let \( \psi = g(\theta) = 1/\theta \). In this problem we consider the estimation of the unknown parameters \( \theta \) and \( \psi \).

(a) Assume that \( \theta \) is a nonrandom parameter.
   (i) Find the maximum likelihood estimator \( \hat{\psi}_{ML} \) of \( \psi \).
   (ii) Find the maximum likelihood estimator \( \hat{\theta}_{ML} \) of \( \theta \).

(b) Assume that \( \theta \) is a realization of a random parameter \( \Theta \) with probability density

\[ p_{\Theta}(\theta) = \frac{1}{\sqrt{2\pi} \theta^2} \exp \left\{ -\frac{1}{2\theta^2} \right\} \quad \theta \neq 0. \]

Assume also that \( \Theta \) and \( V \) are independent.
   (i) Find the maximum a posteriori estimator \( \hat{\psi}_{MAP} \) of \( \psi = g(\Theta) \).
   (ii) Find the maximum a posteriori estimator \( \hat{\Theta}_{MAP} \) of \( \Theta \).
(a) \( \Theta \) and consequently \( \varphi = \frac{1}{\Theta} \) are non-random. First, consider ML estimation of \( \varphi \):

\[
Z = \varphi + V \quad \text{where} \quad V \sim N(0,1)
\]

\[
P_\varphi(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} (z - \varphi)^2 \right\}
\]

\[
\Rightarrow \log P_\varphi(z) = -\log\sqrt{2\pi} - \frac{1}{2} (z - \varphi)^2
\]

\[
\Rightarrow \varphi_{ML}(z) = \arg\min_{\varphi} (z - \varphi)^2 \quad \Rightarrow \quad \hat{\varphi}_{ML}(z) = z \quad (i)
\]

From the invariance principle for ML estimation have

\[
\hat{\Theta}_{ML}(z) = \frac{1}{\hat{\varphi}_{ML}(z)} = \frac{1}{z} \quad (ii)
\]

(b) Let \( \Psi = \frac{1}{\Theta} \) and write the observation equation as

\[
Z = \Psi + V
\]

where \( \Psi \perp V \). To continue we need to calculate the density of \( \Psi \) from that given for \( \Theta \). If you don't remember the formula for a transformation of random variables, it is still easy to compute. Let \( g(\Theta) = \frac{1}{\Theta} \).

**Case \( \varphi > 0 \)**

\[
P(g(\Theta) \leq \varphi) = 1 - P \left( 0 < \Theta \leq \frac{1}{\varphi} \right)
\]

\[
= 1 - \int_0^{1/\varphi} P_\Theta(x) \, dx
\]

Then

\[
P_\Psi(\varphi) = \frac{d}{d\varphi} \left\{ \text{above} \right\} = -P_\Theta \left( \frac{1}{\varphi} \right) \left[ -\varphi^{-2} \right]
\]

\[
= \frac{1}{\varphi^2} \frac{\varphi^2}{\sqrt{2\pi}} \exp\left\{ -\frac{\varphi^2}{2} \right\}
\]

**Case \( \varphi < 0 \)**

\[
P(g(\Theta) \leq \varphi) = \int_{1/\varphi}^{0} P_\Theta(x) \, dx
\]

\[
\Rightarrow \quad P_\Psi(\varphi) = \text{same formula as above.}
Therefore $\Phi$ is actually Gaussian!

$$P_\Phi(\Phi) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{\Phi^2}{2}\right\}$$

To compute the MAP estimate:

$$\hat{\Phi}_{\text{MAP}}(z) = \arg\max_\Phi P_{\Phi|z}(\Phi|z) = \arg\max_\Phi P_{z|\Phi}(z|\Phi)P_\Phi(\Phi)$$

Here

$$P_{z|\Phi}(z|\Phi)P_\Phi(\Phi) = \frac{1}{2\pi} \exp \left\{-\frac{1}{2}(z-\Phi)^2 \right\} \exp \left\{-\frac{\Phi^2}{2}\right\}$$

$$= \frac{1}{2\pi} \exp \left\{-\frac{1}{2} (z^2 - 2z\Phi + 2\Phi^2) \right\}$$

$$\Rightarrow \hat{\Phi}_{\text{MAP}}(z) = \arg\min_\Phi (z^2 - 2z\Phi + 2\Phi^2)$$

$$\frac{d}{d\Phi} \left\{ (i) \right\} = -2z + 4\Phi = 0 \Rightarrow \hat{\Phi}_{\text{MAP}}(z) = \frac{z}{2}$$

2nd deriv. shows its a minimum.

Now for $\hat{\Theta}_{\text{MAP}}(z)$ go from definition

$$\hat{\Theta}_{\text{MAP}}(z) = \arg\max_\Theta P_{z|\Theta}(z|\Theta)P_\Theta(\Theta)$$

Taking log:

$$\log P_{z|\Theta}(z|\Theta) + \log P_\Theta(\Theta)$$

$$= -\frac{1}{2} \log(2\pi) - \frac{1}{2}(z - \Theta)^2 - \frac{1}{2} \log(z^2) - 2 \log \Theta - \frac{1}{2} \Theta^2$$

$$= C(z) - \frac{1}{2}(z - \Theta)^2 - 2 \log \Theta - \frac{1}{2} \Theta^2$$

Taking $\frac{d}{\Theta}$ \{ above \} get MAP equation

$$-\left(z - \frac{1}{\Theta}\right) \frac{1}{\Theta} - \frac{2}{\Theta} + \frac{1}{\Theta^3} = -\frac{z}{\Theta^2} - \frac{z}{\Theta} + \frac{2}{\Theta^3} = 0$$

$$\Rightarrow -z\Theta - 2\Theta^2 + 2 = 0$$
The possibilities we must consider for MAP est. are:

1. roots of $\theta^2 + \frac{3}{2}\theta - 1 = 0$
2. $\theta = 0$
3. $\theta = \pm \infty$

Neither of these can be a global max:
For 2 $\lim_{\theta \to 0} P_{\theta|z}(z|\theta)p_\theta(\theta) = 0$
For 3 $\lim_{|\theta| \to \infty} P_{\theta|z}(z|\theta)p_\theta(\theta) = 0$

From quad. formula, $\theta = \frac{1}{4}\left[-z \pm \sqrt{z^2 + 16}\right]$ and $\hat{H}_{\text{MAP}}(z)$ must be chosen from these roots.

Since $z < 0 \implies \hat{H}_{\text{MAP}}(z) < 0$
$z > 0 \implies \hat{H}_{\text{MAP}}(z) > 0$

must have

$$\hat{H}_{\text{MAP}}(z) = \begin{cases} \frac{1}{4}\left[-z + \sqrt{z^2 + 16}\right] & \text{for } z > 0 \\ \frac{1}{4}\left[-z - \sqrt{z^2 + 16}\right] & \text{for } z < 0 \end{cases}$$
Let \( \mathbf{Y} = [ Y_1 \ Y_2 \ \cdots \ Y_n ]^T \) be a random vector where the individual components are i.i.d. Poisson random variables with parameter \( \theta \).

(a) Show that \( T(\mathbf{Y}) = Y_1 + Y_2 + \cdots + Y_n \) is a complete sufficient statistic for \( \theta \). You must explain why it is sufficient and why it is complete.

(b) Show that \( T(\mathbf{Y}) \) is also Poisson. Start with \( n = 2 \) and use induction. What is the parameter for the distribution of \( T(\mathbf{Y}) \)?

(c) For any (fixed) integer \( k \geq 0 \) find a minimum variance unbiased estimate (MVUE) of the probability

\[
P_\theta \{ Y_1 = k \}.
\]

(d) For any (fixed) integer \( k \geq 0 \) find the maximum likelihood (ML) estimator of the probability

\[
P_\theta \{ Y_1 = k \}.
\]

Is this ML estimator biased?

\[
P_\theta \{ Y_1 = k \}
= \frac{e^{-\theta} \theta^k}{k!}
= \text{pdf}(\theta).
\]
Thus
\[ p_\Theta(y) = p_\Theta^*(y_1) \cdots p_\Theta^*(y_n) = \frac{e^{-\Theta} \Theta^{y_i}}{y_i!} \quad y_i = 0, 1, 2, \ldots \]

(2) The NF factorization thm shows that \( T(y) = \sum_{i=1}^n y_i \) is sufficient for \( \Theta \). To see completeness note
\[ p_\Theta(y) = (e^{-\Theta}) (e^{\Theta T(y)} \log \Theta) \frac{1}{y_1! \cdots y_n!} \]

which is a one parameter exponential family. The param. space contains an interval and hence \( T(y) \) is a complete suff. stat.

(3) Say \( Y_1 \) and \( Y_2 \) are indep. Poisson random variables with parameters \( \Theta_1 \) and \( \Theta_2 \) respectively. Let \( T = Y_1 + Y_2 \) so if \( t \geq 0 \) is an integer
\[ \{T = t\} = \left\{ Y_1 + Y_2 = t \right\} = \bigcup_{s=0}^{t} \{Y_1 = s, Y_2 = t-s\} \]

The above is a disjoint union. With independence
\[ P\{T = t\} = \sum_{s=0}^{t} P\{Y_1 = s\} P\{Y_2 = t-s\} = \sum_{s=0}^{t} \frac{e^{-\Theta_1} \Theta_1^s}{s!} e^{-\Theta_2} \Theta_2^{t-s} \]
\[ = \frac{e^{-(\Theta_1+\Theta_2)}}{t!} \sum_{s=0}^{t} \frac{t!}{s! (t-s)!} \Theta_1^s \Theta_2^{t-s} \]
\[ = \frac{e^{-(\Theta_1+\Theta_2)}}{t!} (\Theta_1 + \Theta_2)^t \Rightarrow \text{T is Poisson} \]

This immediately implies that \( T(y) = Y_1 + \cdots + Y_n \) from (3) is Poisson with parameter \( n \Theta \).

(4) From (2) \( T \) is a complete suff. statistic. Therefore, from the Rao-Blackwell Theorem, if we can find any unbiased estimator \( \hat{g}(y) \) of \( g(\Theta) \) then the MVUE is given by
\[ \hat{g}(T(y)) = \mathbb{E}_\Theta \left[ \hat{g}(y) \mid T(y) = T(y) \right] \]

Of course
\[ \hat{g}(y) = \mathbb{1}_{\{Y_1 = k\}} \] (indicator function)
is unbiased.
Computing the cond. expectation

\[
E_{\theta} \left\{ I_{Y_i = k} \mid T = t \right\} = \frac{P_{\theta} \left\{ Y_i = k \mid T = t \right\}}{P_{\theta} \left\{ \sum_{i=1}^{n} Y_i = t \right\}}
\]

\[
= \frac{P_{\theta} \left\{ Y_i = k \sum_{i=1}^{n} Y_i = t-k \right\}}{P_{\theta} \left\{ \sum_{i=1}^{n} Y_i = t \right\}} \quad \text{(for } t \geq k) \]

\[
= \frac{P_{\theta} \left\{ Y_i = k \right\} P_{\theta} \left\{ \sum_{i=1}^{n} Y_i = t-k \right\}}{P_{\theta} \left\{ \sum_{i=1}^{n} Y_i = t \right\}} = \frac{e^{-\theta} \theta^k}{k!} \frac{e^{-n\theta} (n\theta)^{t-k}}{(t-k)!} \frac{e^{-n\theta}}{t!} \frac{(n\theta)^{t-k}}{(n\theta)^t}
\]

\[
= \frac{t!}{k! (t-k)!} \frac{\theta^k [(n-1) \theta]^{t-k}}{(n \theta)^t} = \frac{(t)}{\binom{k}{t}} (\frac{n-1}{n})^{t-k} = \binom{t}{k} (\frac{1}{n})^k (\frac{n-1}{n})^{t-k}
\]

\[
= \tilde{g}(t).
\]

1. First, find the ML est. of \( \theta \) and then use invariance principle.

\[
\log P_{\theta}(y) = C(y) - n\theta + T(y) \log \theta
\]

\[
\Rightarrow \frac{d}{d\theta} \log P_{\theta}(y) = -n + \frac{T(y)}{\theta} = 0 \Rightarrow \hat{\theta}_{ML}(y) = \frac{1}{n} \sum_{i=1}^{n} y_i
\]

From invariance the ML est. of \( P_{\theta} \{ Y_i = k \} \) is \( P_{\theta} \{ Y_i = k \mid \theta = \hat{\theta}_{ML} \} \)

\[
= e^{-\frac{1}{n} \sum_{i=1}^{n} y_i} \left( \frac{1}{n} \sum_{i=1}^{n} y_i \right)^{k} = \frac{e^{-t/\mu} \left( \frac{t}{\mu} \right)^k}{k!}
\]

This est. is a function of complete suff. stat. \( t \). Hence if unbiased it must be MVUE. But it is not equal to the MVUE from 1 so it must be biased.
This problem refers back to the previous problem. Recall that \( Y = \begin{bmatrix} Y_1 & Y_2 & \cdots & Y_n \end{bmatrix}^T \) was a random vector where the individual components were i.i.d. Poisson random variables with parameter \( \theta \). Previously, you showed that \( T(Y) = Y_1 + Y_2 + \cdots + Y_n \) is a complete sufficient statistic for \( \theta \), that \( T(Y) \) is Poisson with parameter \( n\theta \), and you found the MVUE and ML estimators for

\[ \phi = g(\theta) = P_\theta \{ Y_1 = 0 \} = e^{-\theta}. \]

(actually, a slightly more general result was found before). These estimators were (for \( n > 1 \))

\[
\hat{\phi}_{MVUE}(Y) = \left( \frac{n-1}{n} \right)^{T(Y)} \\
\hat{\phi}_{ML}(Y) = e^{T(Y)/n}.
\]

(a) Find the CRLB for estimating \( \theta \) based upon the \( n \) observations.

(b) The ML estimate for \( \theta \) is \( \hat{\theta}_{ML}(Y) = T(Y)/n \). Is \( \hat{\theta}_{ML}(Y) \) unbiased? Is it efficient? Is it MVUE?

(c) Now consider estimation of \( \phi = e^{-\theta} \). Why can we say that \( \hat{\phi}_{ML} \) is biased without calculation? Calculate the bias directly. From your result show that the ML estimator is asymptotically (as \( n \to \infty \)) unbiased.

(d) Directly calculate the variance of \( \hat{\phi}_{ML} \) as a function of \( \theta \) and \( n \) and argue that the estimator is consistent.

(e) Find the CRLB for estimating \( \phi = e^{-\theta} \) based upon the \( n \) observations and express it as a function of \( \theta \).

(f) Does the unbiased estimator \( \hat{\phi}_{MVUE} \) meet the CRLB for finite \( n \)? Answer the question without calculation.

(g) Now calculate the variance of \( \hat{\phi}_{MVUE} \) expressed as a function of \( \theta \) and \( n \). Form the ratio

\[
\frac{\text{Var}_\theta(\hat{\phi}_{MVUE}(Y))}{\text{the CRLB for } \phi}
\]

and directly explore the efficiency question of (f).
Statement of the "previous problem"

Let $Y = \begin{bmatrix} Y_1 & Y_2 & \cdots & Y_n \end{bmatrix}^T$ be a random vector where the individual components are i.i.d. Poisson random variables with parameter $\theta$.

(a) Show that $T(Y) = Y_1 + Y_2 + \cdots + Y_n$ is a complete sufficient statistic for $\theta$. You must explain why it is sufficient and why it is complete.

(b) Show that $T(Y)$ is also Poisson. Start with $n = 2$ and use induction. What is the parameter for the distribution of $T(Y)$?

(c) For any (fixed) integer $k \geq 0$ find a minimum variance unbiased estimate (MVUE) of the probability $P_\theta\{Y_1 = k\}$.

(d) For any (fixed) integer $k \geq 0$ find the maximum likelihood (ML) estimator of the probability $P_\theta\{Y_1 = k\}$.

Is this ML estimator biased?
Define \( y = \left[ y_1, \ldots, y_n \right]^T \) and parameter \( \theta \) is > 0. Easy properties of the Poisson distribution tell us that

\[
E_\theta \{ Y_i \} = \theta, \quad \text{Var}_\theta \{ Y_i \} = \theta.
\]

Earlier we found \( \hat{\theta}_{ML}(y) = \frac{1}{n} \sum_{i=1}^{n} y_i \) so that

\[
E_\theta \left\{ \left( \hat{\theta}_{ML}(y) \right)^2 \right\} = \theta \quad \text{ie unbiased}
\]

\[
\text{Var}_\theta \left\{ \left( \hat{\theta}_{ML}(y) \right)^2 \right\} = \frac{1}{n^2} \theta = \frac{\theta}{n}
\]

Can compute the CRLB for unbiased estimation of parameter \( \theta \) from \( I_\theta = n i_\theta \)

\[
i_\theta = -E_\theta \left\{ \frac{\partial^2}{\partial \theta^2} \log P_\theta(y_i) \right\}
\]

To this end \( \log P_\theta(y_i) = -\theta + y_i \log \theta - \log y_i! \)

\[
\Rightarrow \frac{\partial}{\partial \theta} \log P_\theta(y_i) = -1 + y_i / \theta
\]

\[
\frac{\partial^2}{\partial \theta^2} \log P_\theta(y_i) = -y_i / \theta^2
\]

\[
i_\theta = \frac{1}{\theta^2} E_\theta \{ Y_i \}^2 = \frac{1}{\theta} \Rightarrow I_\theta = n / \theta
\]

\[\Rightarrow \text{CRLB} = \frac{\theta}{n}\]

See then that the ML estimator of \( \theta \) is efficient in that it meets the CRLB exactly. Hence it is also MVUE for \( \theta \).
Now consider again the estimation of $\phi = e^{-\Theta}$ (a function of theta — one-to-one in fact, $\Theta = -\log \phi$, $0 < \phi < 1$).

The two estimators considered before were

$$\hat{\phi}_{ML}(y) = e^{\frac{T(y)}{n}} = e^{-\hat{\phi}_{ML}(y)}$$

$$\hat{\phi}_{MVUE}(y) = \left(\frac{n-1}{n}\right)^{T(y)}$$

Is a function of a complete sufficient statistic and hence if it were unbiased would have to be MVUE since it's not MVUE it must be biased i.e

$$E_{\theta}\{\hat{\phi}_{ML}(y)\} \neq \phi = e^{-\Theta}$$

For this estimator would like to compute: actual bias, variance, and MSE. Similarly, would likely to have CRLB for estimating $\phi$.

Consider a direct computation

$$E_{\theta}\{\hat{\phi}_{ML}(y)\} = E_{\theta}\{e^{-T/n}\} \quad \text{where } T \text{ is Poisson with param. } \Phi \text{ under } \theta.$$ 

$$= \sum_{t=0}^{\infty} e^{-\frac{T}{n}} P_{\Theta}(T=t) = \sum_{t=0}^{\infty} e^{-\frac{T}{n}} \frac{e^{-\text{\(n\phi\)}} (n\phi)^t}{t!}$$
This is related to the moment generating function for a Poisson r.v. If $X \sim \text{Poisson}$ with param. $\lambda$

$$\n \Rightarrow \Psi_X(\alpha) = E\{e^{\alpha X}\} = \exp\{\lambda(e^\alpha - 1)\} \n$$

which is defined for all $\alpha$. Hence

$$\n E_\theta\{e^{-T/n}\} = \Psi_T(\alpha)\bigg|_{\alpha = -1/n}\n = \exp\{n\theta(e^{-1/n} - 1)\} \n$$

$\therefore$ Bias is

$$\n B_n(\theta) = \exp\{n\theta(e^{-1/n} - 1)\} - \exp\{-\theta\} \n$$

Now we can show that $n(e^{-1/n} - 1) \rightarrow -1$ as $n \rightarrow \infty$ implying that the ML estimator $\hat{\theta}_{ML}$ is at least asymptotically unbiased.

Moving on, let's try to calculate the MSE and variance of $\hat{\theta}_{ML}$.

$$\n E_\theta\{(e^{-T/n})^2\} = E_\theta\{e^{-2T/n}\} = \Psi_T(\alpha)\bigg|_{\alpha = -2/n}\n = \exp\{n\theta(e^{-2/n} - 1)\} \n$$

Now

$$\n \text{Var}_\theta\{e^{-T/n}\} = E_\theta\{(e^{-T/n})^2\} - (E_\theta\{e^{-T/n}\})^2 \n$$
\[ \hat{\phi}_{ML} \]

\[ \text{Var}_{\theta} \{ e^{-\frac{\gamma}{n}} \} = \exp \{ n\theta(e^{-\frac{2}{n}} - 1) \} - \exp \{ 2n\theta(e^{-\frac{1}{n}} - 1) \} \]

Using power series expansion \( e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots \), have:

\[ n\left(e^{-\frac{1}{n}} - 1\right) = n\left(-\frac{1}{n} + \frac{1}{2!}(-\frac{1}{n})^2 + \frac{1}{3!}(-\frac{1}{n})^3 + \cdots\right) \]

\[ \rightarrow -1 \text{ as } n \rightarrow \infty \]

\[ n\left(e^{-\frac{2}{n}} - 1\right) = n\left(-\frac{2}{n} + \frac{1}{2!}(-\frac{2}{n})^2 + \frac{1}{3!}(-\frac{2}{n})^3 + \cdots\right) \]

\[ \rightarrow -2 \text{ as } n \rightarrow \infty \]

which shows that \( \text{Var}_{\theta} \{ e^{-\frac{\gamma}{n}} \} \rightarrow 0 \text{ as } n \rightarrow \infty \).

Finally,

\[ M_{SE_{\theta}} \{ \hat{\phi}_{ML} \} = E_{\theta}\{ (\hat{\phi}_{ML}(\gamma) - \phi(\theta))^2 \} \]

\[ \quad = \text{Var}_{\theta} \{ \hat{\phi}_{ML}(\gamma) \} + (B(\theta))^2 \]

\[ = \exp \{ n\theta(e^{-\frac{2}{n}} - 1) \} - \exp \{ 2n\theta(e^{-\frac{1}{n}} - 1) \} \]

\[ \quad + \left[ \exp \{ n\theta(e^{-\frac{1}{n}} - 1) \} - e^{-\theta} \right]^2 \]

\[ = \exp \{ n\theta(e^{-\frac{2}{n}} - 1) \} - 2e^{-\theta}\exp \{ n\theta(e^{-\frac{1}{n}} - 1) \} + e^{-2\theta} \]
Re: CRLB for estimating $\phi = e^{-\theta}$.

Let $\phi = g(\theta) = e^{-\theta}$ defined for $\theta > 0$. The inverse function is also defined

$$\theta = -\log \phi = \log \left( \frac{1}{\phi} \right) \equiv g^{-1}(\phi)$$

for $0 < \phi \leq 1$. Also define

$$f_{g(\theta)}(y_i) = P_{g(\theta)}(y_i) \iff f_{\phi}(y_i) = P_{g(\phi)}(y_i)$$

and

$$P(y_i, \theta) = \log P_{g(\theta)}(y_i), \quad F(y_i, \phi) = \log f_{\phi}(y_i)$$

whence

$$F(y_i, \phi) = P(y_i, g(\phi)) \Rightarrow F(y_i, g(\theta)) = P(y_i, \theta)$$

To calculate the CRLB for estimating $\phi$ we want

$$\frac{d^2}{d\phi^2} F(y_i, \phi) \approx F''(y_i, \phi)$$

and would like to express it in terms of the CRLB for estimating $\theta$ (which we have already calculated).

$$F'(y_i, \phi) = P'(y_i, g(\phi)) \frac{d}{d\phi} g(\phi)$$

$$F''(y_i, \phi) = P''(y_i, g(\phi)) \left[ \frac{d}{d\phi} g(\phi) \right]^2 + P'(y_i, g(\phi)) \frac{d^2}{d\phi^2} g(\phi)$$

or written as a function of $\theta$ and using the fact that

$$g'(g(\theta)) = \frac{1}{g'(\theta)}$$
we can say that

\[ F''(Y_1, g(\theta)) = \frac{P''(Y_1, \theta)}{[g'(\theta)]^2} + P'(Y_1, \theta) \cdot \{ \text{a nonrand. funct. of } \theta \} \]

expectation wrt \( \theta \)

of this term is zero.

\[ \therefore J_g(\theta) = \frac{i_\theta}{[g'(\theta)]^2} \]

\[ \Rightarrow J_g(\theta) = \frac{1}{[g'(\theta)]^2} \]

Fisher's inf. for estimating \( \theta \).

Fisher's inf. for estimating \( g(\theta) \).

Since \( g'(\theta) = -e^{-\theta} \) get

\[ J_g(\theta) = \frac{n}{\theta e^{-2\theta}} \Rightarrow \text{CRLB for est. } \phi = g(\theta) = \frac{\theta e^{-2\theta}}{n} \]
Direct Computation of $\text{Var}_\theta \{ \hat{\phi}_{\text{MVUE}}(r) \}$

Know that $E_\theta \{ \hat{\phi}_{\text{MVUE}}(r) \} = E_\theta \{ (\frac{n-1}{n})^T \} = e^{-\theta}$ by unbiasedness. ... can be directly verified as well. Similarly

$$E_\theta \{ \frac{(n-1)^2}{n} \} = E_\theta \{ \hat{\phi}_{\text{MVUE}}(r) \}$$

$$= \sum_{t=0}^{\infty} \frac{(n-1)^2 e^{-\theta} (n\theta)^t}{t!} = e^{-\theta} \sum_{t=0}^{\infty} \frac{(n-1)^2 \Theta^{t}}{t!}$$

$$= e^{-\theta} \sum_{t=0}^{\infty} \frac{(n-1)^2 \Theta^{t}}{t!} = e^{-\theta} e^{(\frac{1}{n^2} - 2 + \frac{1}{n} - 1)^2}$$

$$= e^{(\frac{1}{n^2} - 2 + \frac{1}{n} - 1)^2}$$

$$= e^{(\frac{1}{n^2} - 2 + \frac{1}{n} - 1)^2}$$

$$\text{Var}_\theta \{ \hat{\phi}_{\text{MVUE}}(r) \} = e^{(\frac{1}{n^2} - 2 + \frac{1}{n} - 1)^2} - e^{(\frac{1}{n^2} - 2 + \frac{1}{n} - 1)^2}$$

$$= e^{(\frac{1}{n^2} - 2 + \frac{1}{n} - 1)^2} - e^{(\frac{1}{n^2} - 2 + \frac{1}{n} - 1)^2}$$

$$= e^{\frac{\Theta}{e} n - 1}$$

CRLB

$$\text{Var}_\theta \{ \hat{\phi}_{\text{MVUE}}(r) \} = n \left( \frac{e^{\Theta n} - 1}{\Theta} \right)$$

$$= n \left( \frac{e^{\Theta n} - 1}{\Theta} \right)$$

$$= 1 + \frac{1}{2} \left( \frac{\Theta}{n} \right)^2 + \frac{1}{3!} \left( \frac{\Theta}{n} \right)^3 + \cdots > 1$$
We actually can tell in advance that $\hat{\phi}_{\text{MVUE}}(Y)$ cannot meet the CRLB because if it did it would of necc. be equal to $\hat{\phi}_{\text{ML}}(Y)$ and this is not the case.