SECTION I

1. (a) State and prove Invariance property of consistent estimators. Using this property, obtain consistent estimator of $\theta^2 + \theta - \sqrt{\theta}$, when X follows Bernoulli distribution with parameter θ.

- **Invariance Property of Consistent Estimators:** If T_n is a consistent estimator of a parameter θ, and $g(\theta)$ is a continuous function of θ, then $g(T_n)$ is a consistent estimator of $g(\theta)$.
- **Proof:** Since T_n is a consistent estimator of θ, it means that for any $\epsilon > 0$ and $\delta > 0$, there exists an integer N such that for all $n \geq N$, $P(|T_n - \theta| < \epsilon) > 1 - \delta$.

  Since $g(\theta)$ is a continuous function of θ, for any $\epsilon' > 0$, there exists an $\epsilon > 0$ such that if $|T_n - \theta| < \epsilon$, then $|g(T_n) - g(\theta)| < \epsilon'$.

  Therefore, $P(|g(T_n) - g(\theta)| < \epsilon') \geq P(|T_n - \theta| < \epsilon) > 1 - \delta$.

  This shows that as $n \to \infty$, $P(|g(T_n) - g(\theta)| < \epsilon') \to 1$, which is the definition of consistency for $g(T_n)$ as an estimator of $g(\theta)$.
- **Consistent Estimator of $\theta^2 + \theta - \sqrt{\theta}$ for Bernoulli Distribution:**
  - For a Bernoulli distribution with parameter θ, the probability mass function is $P(X = x) = \theta^x (1 - \theta)^{1-x}$ for $x = 0,1$.
  - The mean of a Bernoulli distribution is $E[X] = \theta$.
  - By the Weak Law of Large Numbers, the sample mean $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is a consistent estimator of the population mean θ. That is, $\overline{X} \xrightarrow{P} \theta$.
  - Let $g(\theta) = \theta^2 + \theta - \sqrt{\theta}$. This is a continuous function of θ for $\theta > 0$.
  - Using the invariance property of consistent estimators, since $\overline{X}$ is a consistent estimator of θ, then $g(\overline{X})$ is a consistent estimator of $g(\theta)$.

- Therefore, a consistent estimator of $\theta^2 + \theta - \sqrt{\theta}$ is $\overline{X}^2 + \overline{X} - \sqrt{\overline{X}}$.

1. (b) Let $X_1, X_2, \ldots, X_n$, be a random sample from population having p.d.f. $f(x, \theta) = (\theta + 1)x^\theta$, $0 < x < 1, \theta > -1$. Show that $\left[\frac{-(n-1)}{\sum_{i=1}^{n} \log x_i} - 1\right]$ is an unbiased estimator of θ.

- **Expected Value of** $\log X$**:** To show unbiasedness, we need to find $E\left[\frac{-(n-1)}{\sum_{i=1}^{n} \log x_i} - 1\right]$. This form is complex, so let's consider the expected value of $\log X$ first. $E[\log X] = \int_0^1 (\log x)(\theta + 1)x^\theta dx$ Let $u = -\log x$, so $x = e^{-u}$ and $dx = -e^{-u} du$. When $x = 0, u = \infty$. When $x = 1, u = 0$. $E[\log X] = \int_\infty^0 (-u)(\theta + 1)(e^{-u})^\theta (-e^{-u}) du$ $E[\log X] = \int_0^\infty u\,(\theta + 1)e^{-(\theta+1)u} du$ Let $v = (\theta + 1)u$. Then $u = \frac{v}{\theta+1}$ and $du = \frac{dv}{\theta+1}$. $E[\log X] = \int_0^\infty \frac{v}{\theta+1}(\theta + 1)e^{-v} \frac{dv}{\theta+1}$ $E[\log X] = \frac{1}{\theta+1}\int_0^\infty v\, e^{-v} dv$ The integral $\int_0^\infty v\, e^{-v} dv = \Gamma(2) = 1$. So, $E[\log X] = \frac{1}{\theta+1}$.

- **Distribution of** $\sum_{i=1}^{n} -\log X_i$**:** Let $Y_i = -\log X_i$. Then Y_i follows an exponential distribution with rate parameter $(\theta + 1)$. The p.d.f. of Y_i is $f(y_i) = (\theta + 1)e^{-(\theta+1)y_i}$ for $y_i > 0$. The sum of n independent and identically distributed exponential random variables, $Y = \sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} -\log X_i$, follows a Gamma distribution with shape parameter n and rate parameter $(\theta + 1)$. The p.d.f. of Y is $f_Y(y) = \frac{(\theta+1)^n}{\Gamma(n)} y^{n-1} e^{-(\theta+1)y}$ for $y > 0$.

- **Expected Value of** $\frac{1}{Y}$**:** We need to find $E\left[\frac{1}{\sum_{i=1}^{n} \log x_i}\right] = E\left[\frac{1}{-Y}\right]$. $E\left[\frac{1}{-Y}\right] = -\int_0^\infty \frac{1}{y} \frac{(\theta+1)^n}{\Gamma(n)} y^{n-1} e^{-(\theta+1)y} dy$ $E\left[\frac{1}{-Y}\right] = -\frac{(\theta+1)^n}{\Gamma(n)} \int_0^\infty y^{n-2}\, e^{-(\theta+1)y} dy$ The integral is of the form $\int_0^\infty y^{k-1}\, e^{-\lambda y} dy = \frac{\Gamma(k)}{\lambda^k}$. Here $k = n - 1$ and $\lambda = \theta + 1$. So, the integral is $\frac{\Gamma(n-1)}{(\theta+1)^{n-1}}$. $E\left[\frac{1}{-Y}\right] = -\frac{(\theta+1)^n}{\Gamma(n)} \frac{\Gamma(n-1)}{(\theta+1)^{n-1}}$ Since $\Gamma(n) = (n-1)\Gamma(n-1)$, $E\left[\frac{1}{-Y}\right] = -\frac{(\theta+1)^n}{(n-1)\Gamma(n-1)} \frac{\Gamma(n-1)}{(\theta+1)^{n-1}}$ $E\left[\frac{1}{-Y}\right] = -\frac{\theta+1}{n-1}$

- **Unbiasedness Check:** Now substitute this back into the estimator: $E\left[\frac{-(n-1)}{\sum_{i=1}^{n}\log x_i}-1\right]=E\left[-(n-1)\left(\frac{1}{\sum_{i=1}^{n}\log x_i}\right)-1\right]=-(n-1)E\left[\frac{1}{-Y}\right]-1=-(n-1)\left(-\frac{\theta+1}{n-1}\right)-1=(\theta+1)-1=\theta$

  Thus, the estimator is unbiased for θ.

2. (a) Show that the estimator of the form $a\overline{X}$ for in random sampling from N(0,σ^2), has the minimum mean-square error, when $a=\frac{\sigma^2}{\theta^2+\sigma^2/n}$, Which → 1 as $n\to\infty$ but < 1 when n is finite.

- **Mean and Variance of $\overline{X}$:** Given $X_1, X_2, \ldots, X_n$ is a random sample from $N(0,\sigma^2)$. This means the true mean is 0. However, the question states "N(0, σ^2)" and then refers to "θ". It's likely there's a typo and it should be $N(\theta,\sigma^2)$, with θ being the unknown mean we are estimating. Let's assume the distribution is $N(\theta,\sigma^2)$. For $X\sim N(\theta,\sigma^2)$, we have: $E[\overline{X}]=\theta$ $Var(\overline{X})=\frac{\sigma^2}{n}$

- **Mean Squared Error (MSE) of $T=a\overline{X}$:** The MSE of an estimator T of a parameter θ is given by $MSE(T)=Var(T)+[Bias(T)]^2$. First, let's find the bias of $T=a\overline{X}$: $E[T]=E[a\overline{X}]=aE[\overline{X}]=a\theta$ $Bias(T)=E[T]-\theta=a\theta-\theta=\theta(a-1)$

  Next, let's find the variance of T: $Var(T)=Var(a\overline{X})=a^2Var(\overline{X})=a^2\frac{\sigma^2}{n}$

  Now, substitute these into the MSE formula: $MSE(a\overline{X})=a^2\frac{\sigma^2}{n}+[\theta(a-1)]^2$ $MSE(a\overline{X})=a^2\frac{\sigma^2}{n}+\theta^2(a-1)^2$

- **Minimizing MSE with respect to a:** To find the value of a that minimizes the MSE, we take the derivative of $MSE(a\overline{X})$ with respect to a and set it to zero: $\frac{d}{da}MSE(a\overline{X})=\frac{d}{da}\left(a^2\frac{\sigma^2}{n}+\theta^2(a-1)^2\right)=0$ $2a\frac{\sigma^2}{n}+2\theta^2(a-1)=0$ $a\frac{\sigma^2}{n}+\theta^2a-\theta^2=0$ $a\left(\frac{\sigma^2}{n}+\theta^2\right)=\theta^2$ $a=\frac{\theta^2}{\frac{\sigma^2}{n}+\theta^2}$ $a=\frac{\theta^2}{\frac{\sigma^2+n\theta^2}{n}}$ $a=\frac{n\theta^2}{\sigma^2+n\theta^2}$

This expression can be rewritten by dividing the numerator and denominator by θ^2: $a = \frac{n}{n+\frac{\sigma^2}{\theta^2}}$ This can also be written as: $a = \frac{1}{1+\frac{\sigma^2}{n\theta^2}}$

The problem statement gives $a = \frac{\sigma^2}{\theta^2+\sigma^2/n}$. Let's recheck the derivative or the problem statement. If the question meant to minimize the MSE for an estimator of 0 (since $N(0, \sigma^2)$ is given), then $\theta = 0$. In that case, $a = 0$. However, the estimator form is $a\overline{X}$, which for $N(0, \sigma^2)$ would estimate 0. This seems contradictory.

Let's assume the question's provided minimum MSE 'a' is correct and work backward or confirm. Given $a = \frac{\sigma^2}{\theta^2+\sigma^2/n}$. This 'a' value suggests that θ is the parameter being estimated, and it implies a slightly different derivation of the MSE, perhaps for estimating a mean from a specific prior, or if the θ in the denominator is not the true mean.

Let's re-evaluate the problem statement carefully: "estimator of the form $a\overline{X}$ for in random sampling from N(0,σ^2)". If the true mean is 0, and we are estimating 0, then $\overline{X}$ is an unbiased estimator of 0. The estimator $a\overline{X}$ would also estimate 0. In this case, $E[\overline{X}] = 0$. $E[a\overline{X}] = a \cdot 0 = 0$. So, $a\overline{X}$ is always an unbiased estimator of 0. $Bias(a\overline{X}) = 0 - 0 = 0$. $MSE(a\overline{X}) = Var(a\overline{X}) + [Bias(a\overline{X})]^2 = a^2Var(\overline{X}) + 0 = a^2\frac{\sigma^2}{n}$. To minimize $a^2\frac{\sigma^2}{n}$, if 'a' is unrestricted, the minimum would be when $a = 0$. This doesn't make sense in the context of the given a.

Assumption based on typical problems: The problem likely intended to state that the population is $N(\theta, \sigma^2)$ and we are estimating θ. Let's proceed with this assumption, as it aligns with the given 'a' value. If we are estimating θ, and our estimator is $a\overline{X}$, we derived $a = \frac{n\theta^2}{\sigma^2+n\theta^2}$. This is not the $a = \frac{\sigma^2}{\theta^2+\sigma^2/n}$ given in the question.

Let's reconsider the formulation. Suppose we are estimating θ, but the mean of the population is θ_0. Or, perhaps the question implies that the estimator is for some quantity related to θ, where θ itself is not necessarily the mean.

Let's re-read: "Show that the estimator of the form $a\overline{X}$ for in random sampling from N(0,σ^2), has the minimum mean-square error, when $a = \frac{\sigma^2}{\theta^2+\sigma^2/n}$". This phrasing strongly implies that θ is **not** the population mean. It might be a fixed constant or another parameter. If the distribution is $N(0, \sigma^2)$, then the mean is 0. Let's assume we are estimating 0. Then $a\overline{X}$ is an estimator for 0. $E[a\overline{X}] = a \cdot 0 = 0$. So $a\overline{X}$ is unbiased for 0. $MSE(a\overline{X}) = Var(a\overline{X}) = a^2\frac{\sigma^2}{n}$. In this case, to minimize $a^2\frac{\sigma^2}{n}$, we would set $a = 0$. This still leads to a contradiction with the given a.

Crucial Re-interpretation: The phrase "for in random sampling from N(0,σ^2)" is slightly ambiguous. It is most likely a typo and should be "for θ in random sampling from N(θ,σ^2)". Let's proceed with this assumption, as it leads to the closest form of 'a'.

If we are estimating θ from $N(\theta, \sigma^2)$: We found $a = \frac{n\theta^2}{\sigma^2+n\theta^2}$. This is not the same as the given $a = \frac{\sigma^2}{\theta^2+\sigma^2/n}$.

Let's consider if the question intended for the estimator to be for θ, but it's not simply $a\overline{X}$ for θ. What if $\overline{X}$ itself is for 0, and we're looking for an estimator of θ in a different context. This is becoming too speculative.

Let's assume the question meant that the estimator is $a\overline{X}$ and the *true* parameter value for the expectation is θ and the variance is σ^2/n. This makes sense if the question means $N(\theta, \sigma^2)$.

Let's re-evaluate the derivative for the first 'a' derived: $a = \frac{n\theta^2}{\sigma^2+n\theta^2}$ This can be written as $a = \frac{1}{1+\frac{\sigma^2}{n\theta^2}}$.

The question states $a = \frac{\sigma^2}{\theta^2+\sigma^2/n}$. Let's test this form of 'a' in the MSE. This form is for the 'Shrinkage Estimator' or 'Bayesian Estimator'. Consider the estimator $\hat{\theta} = a\overline{X}$. We need to assume that θ is the parameter we are estimating and it's from $N(\theta, \sigma^2)$. $MSE(a\overline{X}) = a^2\frac{\sigma^2}{n} + \theta^2(a-1)^2$. If we substitute the given $a = \frac{\sigma^2}{\theta^2+\sigma^2/n}$: $MSE = \left(\frac{\sigma^2}{\theta^2+\sigma^2/n}\right)^2\frac{\sigma^2}{n} + \theta^2\left(\frac{\sigma^2}{\theta^2+\sigma^2/n} - 1\right)^2$ $MSE = \frac{\sigma^4}{(\theta^2+\sigma^2/n)^2}\frac{\sigma^2}{n} + \theta^2\left(\frac{\sigma^2-(\theta^2+\sigma^2/n)}{\theta^2+\sigma^2/n}\right)^2$ $MSE = \frac{\sigma^6}{n(\theta^2+\sigma^2/n)^2} + \theta^2\left(\frac{\sigma^2-\theta^2-\sigma^2/n}{\theta^2+\sigma^2/n}\right)^2$ $MSE = \frac{\sigma^6}{n(\theta^2+\sigma^2/n)^2} + \theta^2\left(\frac{(\sigma^2-\sigma^2/n)-\theta^2}{\theta^2+\sigma^2/n}\right)^2$ This looks computationally heavy to prove it's the minimum without setting the derivative.

Let's assume the question has a typo and the a value provided is derived from minimizing $MSE(a\overline{X})$ when estimating a parameter, let's call it μ, from $N(\mu, \sigma^2)$. Then $a = \frac{\mu^2}{\mu^2+\sigma^2/n}$. This is the result from the typical shrinkage estimator problem, where the parameter μ is treated as a random variable with some prior, or we are estimating $E[\mu]$. If the question *is* precisely stated with $N(0, \sigma^2)$, and $a\overline{X}$ is the estimator, and the MSE formula is general for an estimator of an arbitrary θ (not necessarily the mean), then we are looking for the minimum MSE for estimating θ. $MSE(a\overline{X};\theta) = E[(a\overline{X} - \theta)^2]$. $E[(a\overline{X}-\theta)^2] = E[a^2\overline{X}^2 - 2a\theta\overline{X} + \theta^2]$ Since $X \sim N(0, \sigma^2)$, $E[\overline{X}] = 0$ and $E[\overline{X}^2] = Var(\overline{X}) + (E[\overline{X}])^2 = \frac{\sigma^2}{n} + 0^2 = \frac{\sigma^2}{n}$. $MSE(a\overline{X};\theta) = a^2E[\overline{X}^2] - 2a\theta E[\overline{X}] + \theta^2$ $MSE(a\overline{X};\theta) = a^2\frac{\sigma^2}{n} - 2a\theta(0) + \theta^2$ $MSE(a\overline{X};\theta) = a^2\frac{\sigma^2}{n} + \theta^2$ To minimize this with respect to a, take the derivative: $\frac{d}{da}\left(a^2\frac{\sigma^2}{n} + \theta^2\right) = 2a\frac{\sigma^2}{n} = 0$ This implies $a = 0$. This again contradicts the given a.

Conclusion for 2. (a): There appears to be an inconsistency in the question's premise and the given 'a' value if the distribution is strictly $N(0, \sigma^2)$ and we are estimating θ. The given 'a' value is a standard result for a "shrinkage estimator" or a "Bayesian estimator" where the true mean is θ and θ itself is considered a random variable or a point estimate of it is sought in a specific context (e.g., estimating θ when θ is known). Let's assume the question implicitly refers to the classical problem where we estimate the mean θ from $N(\theta, \sigma^2)$ and the MSE minimizing constant a is sought for an estimator of the form $a\overline{X}$. In this case, as derived initially: $a = \frac{n\theta^2}{\sigma^2+n\theta^2}$. This is not the given $a = \frac{\sigma^2}{\theta^2+\sigma^2/n}$.

Given the exact wording and the provided 'a', let's assume it's for estimating **zero** from $N(0, \sigma^2)$, and θ in the formula for a is some other parameter related to a. This is the only way the question makes sense with the formula for a. If we are estimating the parameter 0 from $N(0, \sigma^2)$ using $a\overline{X}$. $MSE(a\overline{X}) = E[(a\overline{X} - 0)^2] = a^2E[\overline{X}^2]$. Since $E[\overline{X}] = 0$, $E[\overline{X}^2] = Var(\overline{X}) + (E[\overline{X}])^2 = \frac{\sigma^2}{n} + 0^2 = \frac{\sigma^2}{n}$. So, $MSE(a\overline{X}) = a^2\frac{\sigma^2}{n}$. To minimize this MSE, a should be 0. This contradicts the value of a given in the question.

Alternative interpretation based on "Shrinkage Estimator": If the problem is asking to show that the *given a is* the one that minimizes the MSE for some scenario. The form $a = \frac{\tau^2}{\tau^2+\sigma^2/n}$ often arises in a Bayesian context, where $\theta \sim N(0, \tau^2)$ and $X|\theta \sim N(\theta, \sigma^2)$. Then the Bayes estimator for θ is $E[\theta|X] = \frac{\tau^2}{\tau^2+\sigma^2/n}\overline{X}$. In this context, θ is the mean, and the estimator is $a\overline{X}$. Let's assume the θ in the given formula for a represents the standard deviation τ of the prior for θ. No, that wouldn't make sense since it's θ^2. Perhaps, θ is the actual mean, and $a = \frac{\theta^2}{\theta^2+\sigma^2/n}$. This is the form of the "James-Stein type" estimator or shrinkage estimator for the mean when estimating θ itself from $N(\theta, \sigma^2)$, but using a prior mean of 0. In this scenario, $a = \frac{\theta^2}{\theta^2+\sigma^2/n}$. Then, $MSE(a\overline{X}) = \left(\frac{\theta^2}{\theta^2+\sigma^2/n}\right)^2\frac{\sigma^2}{n} + \theta^2\left(\frac{\theta^2}{\theta^2+\sigma^2/n} - 1\right)^2$ $MSE = \frac{\theta^4\sigma^2}{n(\theta^2+\sigma^2/n)^2} + \theta^2\left(\frac{\theta^2-(\theta^2+\sigma^2/n)}{\theta^2+\sigma^2/n}\right)^2$ $MSE = \frac{\theta^4\sigma^2}{n(\theta^2+\sigma^2/n)^2} +$

$\theta^2\left(\frac{-\sigma^2/n}{\theta^2+\sigma^2/n}\right)^2$ $MSE = \frac{\theta^4\sigma^2}{n(\theta^2+\sigma^2/n)^2} + \theta^2\frac{\sigma^4/n^2}{(\theta^2+\sigma^2/n)^2}$ $MSE = \frac{1}{(\theta^2+\sigma^2/n)^2}\left(\frac{\theta^4\sigma^2}{n}+\frac{\theta^2\sigma^4}{n^2}\right)$ $MSE = \frac{1}{(\theta^2+\sigma^2/n)^2}\left(\frac{n\theta^4\sigma^2+\theta^2\sigma^4}{n^2}\right)$ $MSE = \frac{\theta^2\sigma^2}{n^2(\theta^2+\sigma^2/n)^2}(n\theta^2+\sigma^2)$ $MSE = \frac{\theta^2\sigma^2}{n^2(\frac{n\theta^2+\sigma^2}{n})^2}(n\theta^2+\sigma^2)$ $MSE = \frac{\theta^2\sigma^2}{n^2\frac{(n\theta^2+\sigma^2)^2}{n^2}}(n\theta^2+\sigma^2)$ $MSE = \frac{\theta^2\sigma^2}{n\theta^2+\sigma^2}$

Now, let's check the given $a = \frac{\sigma^2}{\theta^2+\sigma^2/n}$. This a is often encountered in the context of estimating variance when the mean is zero. But the estimator is $a\overline{X}$, which is for the mean.

Given the ambiguity, let's assume the question implicitly asks to verify the properties of the given 'a' value. Let $T = a\overline{X}$ where $a = \frac{\sigma^2}{\theta^2+\sigma^2/n}$. We are sampling from $N(0,\sigma^2)$. The parameter being estimated is not explicitly stated. If it's for estimating 0: $MSE = a^2\frac{\sigma^2}{n} = \left(\frac{\sigma^2}{\theta^2+\sigma^2/n}\right)^2\frac{\sigma^2}{n}$. This is just a value, not a minimization.

Let's consider the problem as stated, and assume θ in the a formula is some unknown parameter, and the sampling is from $N(0,\sigma^2)$. And the quantity being estimated is θ. If the problem truly means that the observations are from $N(0,\sigma^2)$, but we are estimating a parameter θ, and our estimator is $a\overline{X}$, then this setup is problematic. This problem is a classic source of confusion due to inconsistent notation or a typo. Given the structure of such problems, the most probable intent is that we are sampling from $N(\theta,\sigma^2)$ and estimating θ. In that case, the 'a' that minimizes the MSE for $a\overline{X}$ is $a = \frac{n\theta^2}{\sigma^2+n\theta^2}$. This is not the given a.

Let's assume the question meant to say: "Show that the estimator of the form $a\overline{X}$ has the minimum mean-square error for estimating θ when $X_i \sim N(\theta,\sigma^2)$ if $a = \frac{\sigma^2}{\theta^2+\sigma^2/n}$". This still does not fit our derivation of $a = \frac{n\theta^2}{\sigma^2+n\theta^2}$.

What if the quantity being estimated is $\frac{1}{\theta}$? Or σ^2? No, the estimator is $a\overline{X}$.

The only way the given 'a' makes sense is if the question means: "Show that for an estimator of the form $a\overline{X}$ from $N(\theta, \sigma^2)$, if we are trying to estimate **zero** (as suggested by $N(0, \sigma^2)$ in the initial text), and θ itself is a parameter we are interested in. This still doesn't fit the 'a' form typically."

Re-reading again: "estimator of the form $a\overline{X}$ for in random sampling from N(0,σ^2)". The "for" part is missing the parameter being estimated. Given the MSE derivation for $a\overline{X}$ usually involves the parameter being estimated, let's assume it's θ. If $X \sim N(\theta, \sigma^2)$, then $MSE(a\overline{X}) = a^2\frac{\sigma^2}{n} + \theta^2(a-1)^2$. The a that minimizes this is $a = \frac{n\theta^2}{\sigma^2+n\theta^2}$.

The question presents $a = \frac{\sigma^2}{\theta^2+\sigma^2/n}$. This form is very similar to the posterior mean when θ has a prior distribution, often $N(0, \tau^2)$. If $\tau^2 = \sigma^2/\theta^2$ in some sense, this could be related.

Let's *assume* the given $a = \frac{\sigma^2}{\theta^2+\sigma^2/n}$ is correct and it is indeed the a that minimizes $MSE(a\overline{X})$ when estimating θ from $N(\theta, \sigma^2)$. We will show that this is the minimum MSE, by setting the derivative to zero and seeing if it matches. We need to solve $2a\frac{\sigma^2}{n} + 2\theta^2(a-1) = 0$ for a. $a\frac{\sigma^2}{n} + a\theta^2 - \theta^2 = 0$ $a(\frac{\sigma^2}{n} + \theta^2) = \theta^2$ $a = \frac{\theta^2}{\theta^2+\sigma^2/n}$. This is what we derived. The question provides $a = \frac{\sigma^2}{\theta^2+\sigma^2/n}$. These are different! There is a clear mismatch between the standard result for minimizing MSE for $a\overline{X}$ for estimating θ from $N(\theta, \sigma^2)$ and the a provided in the question.

Let's assume there's a typo in the provided a and the correct a for minimum MSE for estimating θ from $N(\theta, \sigma^2)$ is $a = \frac{\theta^2}{\theta^2+\sigma^2/n}$.

- **Derivation of minimum MSE** a: (as derived above) For $X_i \sim N(\theta, \sigma^2)$, $E[\overline{X}] = \theta$, $Var(\overline{X}) = \sigma^2/n$. The estimator is $T = a\overline{X}$. $MSE(T) = Var(T) + [Bias(T)]^2$ $E[T] = a\theta$, so $Bias(T) = a\theta - \theta = \theta(a-1)$. $Var(T) = a^2 Var(\overline{X}) = a^2 \frac{\sigma^2}{n}$. $MSE(a\overline{X}) = a^2 \frac{\sigma^2}{n} + \theta^2(a-1)^2$. To minimize MSE, differentiate with respect to a and set to zero: $\frac{d}{da} MSE(a\overline{X}) = 2a\frac{\sigma^2}{n} + 2\theta^2(a-1) = 0$ $a\frac{\sigma^2}{n} + a\theta^2 - \theta^2 = 0$ $a\left(\frac{\sigma^2}{n} + \theta^2\right) = \theta^2$ $a = \frac{\theta^2}{\theta^2 + \sigma^2/n}$.
- **Behavior of** a**:** As $n \to \infty$, $\sigma^2/n \to 0$. So, $a \to \frac{\theta^2}{\theta^2+0} = 1$. When n is finite: Since $\sigma^2/n > 0$ (assuming $\sigma^2 > 0$ and $n > 0$), we have $\theta^2 + \sigma^2/n > \theta^2$. Therefore, $\frac{\theta^2}{\theta^2+\sigma^2/n} < 1$. So, $a < 1$ when n is finite.

This derivation shows that if the problem intended for $a = \frac{\theta^2}{\theta^2+\sigma^2/n}$, then the conditions are met. However, the provided a in the question is different. This strongly suggests a typo in the question itself. I will assume the question intended to ask to prove for $a = \frac{\theta^2}{\theta^2+\sigma^2/n}$ for estimating θ from $N(\theta, \sigma^2)$.

2. (b) State Cramer-Rao Inequality. Let $X_1, X_2, \ldots, X_n$, be a random sample from Population having p.d.f. $f(x,\theta) = \frac{1}{\pi[1+(x-\theta)^2]}, -\infty < x < \infty, -\infty < \theta < \infty$. Find Cramer-Rao Lower bound for variance of an unbiased estimator of θ. Also, examine whether MVB estimator exists for θ.

- **Cramer-Rao Inequality:** Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with p.d.f. $f(x,\theta)$. Let $T = T(X_1, \ldots, X_n)$ be an unbiased estimator of a parameter θ. Under certain regularity conditions, the variance of T satisfies: $Var(T) \geq \frac{1}{nI(\theta)}$ where $I(\theta)$ is the Fisher Information of a single observation, given by: $I(\theta) = E\left[\left(\frac{\partial}{\partial\theta} \log f(X,\theta)\right)^2\right] =$

$-E\left[\frac{\partial^2}{\partial\theta^2}\log f(X,\theta)\right]$. If T is an unbiased estimator of $g(\theta)$, then $Var(T) \geq \frac{[g'(\theta)]^2}{nI(\theta)}$. For estimating θ, $g(\theta) = \theta$, so $g'(\theta) = 1$.

- **Cramer-Rao Lower Bound for the Cauchy Distribution:** The given p.d.f. is $f(x,\theta) = \frac{1}{\pi[1+(x-\theta)^2]}$. This is the p.d.f. of a Cauchy distribution.

  a. **Find** $\log f(X,\theta)$**:** $\log f(X,\theta) = \log\left(\frac{1}{\pi}\right) - \log[1+(X-\theta)^2]$
  $\log f(X,\theta) = -\log\pi - \log[1+(X-\theta)^2]$

  b. **Find the first derivative with respect to** θ**:** $\frac{\partial}{\partial\theta}\log f(X,\theta) = -\frac{1}{1+(X-\theta)^2}\cdot\frac{\partial}{\partial\theta}[1+(X-\theta)^2] = -\frac{1}{1+(X-\theta)^2}\cdot[2(X-\theta)(-1)] = \frac{2(X-\theta)}{1+(X-\theta)^2}$

  c. **Find the second derivative with respect to** θ**:** $\frac{\partial^2}{\partial\theta^2}\log f(X,\theta) = \frac{\partial}{\partial\theta}\left[\frac{2(X-\theta)}{1+(X-\theta)^2}\right]$ Using the quotient rule: $\frac{d}{du}\frac{u}{v} = \frac{u'v-uv'}{v^2}$ Numerator: $u = 2(X-\theta)$, $u' = -2$. Denominator: $v = 1+(X-\theta)^2$, $v' = 2(X-\theta)(-1) = -2(X-\theta)$. $\frac{\partial^2}{\partial\theta^2}\log f(X,\theta) = \frac{-2[1+(X-\theta)^2]-2(X-\theta)[-2(X-\theta)]}{[1+(X-\theta)^2]^2}$ $= \frac{-2-2(X-\theta)^2+4(X-\theta)^2}{[1+(X-\theta)^2]^2} = \frac{2(X-\theta)^2-2}{[1+(X-\theta)^2]^2} = 2\frac{(X-\theta)^2-1}{[1+(X-\theta)^2]^2}$

  d. **Calculate Fisher Information** $I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\log f(X,\theta)\right]$**:** $I(\theta) = -E\left[2\frac{(X-\theta)^2-1}{[1+(X-\theta)^2]^2}\right]$ $I(\theta) = -\int_{-\infty}^{\infty} 2\frac{(x-\theta)^2-1}{[1+(x-\theta)^2]^2}\frac{1}{\pi[1+(x-\theta)^2]}dx$ $I(\theta) = -\frac{2}{\pi}\int_{-\infty}^{\infty}\frac{(x-\theta)^2-1}{[1+(x-\theta)^2]^3}dx$ Let $y = x-\theta$. Then $dx = dy$. $I(\theta) = -\frac{2}{\pi}\int_{-\infty}^{\infty}\frac{y^2-1}{(1+y^2)^3}dy$ This integral can be solved using contour integration or by breaking it down. Consider the integral $J = \int_{-\infty}^{\infty}\frac{1}{(1+y^2)^k}dy$. For $k = 1$, $\int_{-\infty}^{\infty}\frac{1}{1+y^2}dy = [\arctan y]_{-\infty}^{\infty} = \pi$. We need to evaluate $\int_{-\infty}^{\infty}\frac{y^2}{(1+y^2)^3}dy$ and $\int_{-\infty}^{\infty}\frac{1}{(1+y^2)^3}dy$. A known result for Cauchy distribution: The Fisher Information for a Cauchy distribution is $I(\theta) = \frac{1}{2}$. Let's try to derive it. We can use integration by parts or complex analysis. Alternatively, $I(\theta) = E\left[\left(\frac{2(X-\theta)}{1+(X-\theta)^2}\right)^2\right] = \int_{-\infty}^{\infty}\left(\frac{2(x-\theta)}{1+(x-\theta)^2}\right)^2 \frac{1}{\pi(1+(x-\theta)^2)}dx$ $I(\theta) = \frac{4}{\pi}\int_{-\infty}^{\infty}\frac{(x-\theta)^2}{(1+(x-\theta)^2)^3}dx$ Let $y = x-\theta$. $I(\theta) = \frac{4}{\pi}\int_{-\infty}^{\infty}\frac{y^2}{(1+y^2)^3}dy$. Using integral tables or residues, $\int_{-\infty}^{\infty}\frac{y^2}{(1+y^2)^3}dy = \frac{\pi}{8}$. So, $I(\theta) = \frac{4}{\pi}\cdot\frac{\pi}{8} = \frac{1}{2}$.

e. **Cramer-Rao Lower Bound:** The CRLB for the variance of an unbiased estimator of θ is $\frac{1}{nI(\theta)}$. $CRLB = \frac{1}{n\cdot(1/2)} = \frac{2}{n}$.

- **Existence of MVB Estimator:** An MVB (Minimum Variance Bound) estimator exists if and only if the equality in the Cramer-Rao inequality holds. The equality holds if and only if $\frac{\partial}{\partial\theta}\log L(X_1,\ldots,X_n;\theta)$ can be written in the form $k(\theta)(T-\theta)$, where T is an unbiased estimator of θ and $k(\theta)$ is a function of θ only. For a single observation, $\frac{\partial}{\partial\theta}\log f(X,\theta) = \frac{2(X-\theta)}{1+(X-\theta)^2}$. This cannot be written in the form $k(\theta)(X-\theta)$. The term $\frac{1}{1+(X-\theta)^2}$ makes it impossible to separate X from θ in the required form. For the sample, $\frac{\partial}{\partial\theta}\log L = \sum_{i=1}^{n}\frac{2(X_i-\theta)}{1+(X_i-\theta)^2}$. This sum cannot be written as $k(\theta)(T-\theta)$. Therefore, an MVB estimator for θ does not exist for the Cauchy distribution. This is a known property of the Cauchy distribution; its mean does not exist, and standard estimation methods often fail.

3. (a) Let X and Y be two random variables having joint probability density function $f(x,y) = \frac{2}{\theta^2}\exp[-(x+y)]$, $0 < x < y < \infty$.

- **Marginal PDF of Y:** To find the marginal PDF of Y, we integrate $f(x,y)$ with respect to x. The limits for x are from 0 to y. $f_Y(y) = \int_0^y \frac{2}{\theta^2}e^{-(x+y)}dx = \frac{2}{\theta^2}e^{-y}\int_0^y e^{-x}\,dx$ $f_Y(y) = \frac{2}{\theta^2}e^{-y}[-e^{-x}]_0^y = \frac{2}{\theta^2}e^{-y}(-e^{-y}-(-e^0))$ $f_Y(y) = \frac{2}{\theta^2}e^{-y}(1-e^{-y}) = \frac{2}{\theta^2}(e^{-y}-e^{-2y})$, for $y > 0$.

- **(i) Show that the mean and variance of Y are respectively $\frac{3\theta}{2}$ and $\frac{5\theta^2}{4}$.**

  The problem statement has $\frac{2}{\theta^2}$ as a constant. This constant is not standard for a general exponential distribution. Let's recheck if the problem meant $e^{-(x+y)/\theta}$ or if the constant is fixed. Given $f(x,y) = \frac{2}{\theta^2}\exp[-(x+y)]$. This constant implies θ is a scale parameter. If the constant is $\frac{2}{\theta^2}$, then the integral over the sample space should be 1. $\int_0^\infty \int_0^y \frac{2}{\theta^2} e^{-(x+y)} dxdy = \int_0^\infty \frac{2}{\theta^2}(e^{-y} - e^{-2y})dy = \frac{2}{\theta^2}[-e^{-y} + \frac{1}{2}e^{-2y}]_0^\infty = \frac{2}{\theta^2}[-(0-1) + \frac{1}{2}(0-1)] = \frac{2}{\theta^2}[1-\frac{1}{2}] = \frac{2}{\theta^2}\cdot\frac{1}{2} = \frac{1}{\theta^2}$. For this to be a valid PDF, we must have $\frac{1}{\theta^2} = 1$, which implies $\theta^2 = 1$, so $\theta = 1$. This seems restrictive. There might be a typo in the PDF definition. A common joint PDF for transformed exponential variables involves scaling by θ. Let's assume the PDF is $f(x,y) = \frac{2}{\theta^2}e^{-(x+y)/\theta}$ for $0 < x < y < \infty$. In this case, the constant is correct. Let's work with this assumption. $f(x,y) = \frac{2}{\theta^2}e^{-x/\theta}e^{-y/\theta}$. Marginal PDF of Y: $f_Y(y) = \int_0^y \frac{2}{\theta^2}e^{-(x+y)/\theta}dx = \frac{2}{\theta^2}e^{-y/\theta}\int_0^y e^{-x/\theta}\,dx$ $f_Y(y) = \frac{2}{\theta^2}e^{-y/\theta}[-\theta e^{-x/\theta}]_0^y = \frac{2}{\theta^2}e^{-y/\theta}(-\theta e^{-y/\theta} - (-\theta e^0))$ $f_Y(y) = \frac{2}{\theta^2}e^{-y/\theta}(\theta - \theta e^{-y/\theta}) = \frac{2}{\theta}(e^{-y/\theta} - e^{-2y/\theta})$, for $y > 0$.

  Expected Value of Y: $E[Y] = \int_0^\infty y\frac{2}{\theta}(e^{-y/\theta} - e^{-2y/\theta})dy$ $E[Y] = \frac{2}{\theta}\left[\int_0^\infty y\, e^{-y/\theta}dy - \int_0^\infty y\, e^{-2y/\theta}dy\right]$ Recall $\int_0^\infty x\, e^{-ax}dx = \frac{1}{a^2}$. For the first integral, $a = 1/\theta$, so $\frac{1}{(1/\theta)^2} = \theta^2$. For the second integral, $a = 2/\theta$, so $\frac{1}{(2/\theta)^2} = \frac{\theta^2}{4}$. $E[Y] = \frac{2}{\theta}\left[\theta^2 - \frac{\theta^2}{4}\right] = \frac{2}{\theta}\left[\frac{3\theta^2}{4}\right] = \frac{3\theta}{2}$. This matches.

  Expected Value of Y squared: $E[Y^2] = \int_0^\infty y^2\frac{2}{\theta}(e^{-y/\theta} - e^{-2y/\theta})dy$ $E[Y^2] = \frac{2}{\theta}\left[\int_0^\infty y^2\, e^{-y/\theta}dy - \int_0^\infty y^2\, e^{-2y/\theta}dy\right]$ Recall $\int_0^\infty x^2\, e^{-ax}dx = \frac{2}{a^3}$. For the first integral, $a = 1/\theta$, so $\frac{2}{(1/\theta)^3} = 2\theta^3$. For the second integral, $a =$

$2/\theta$, so $\frac{2}{(2/\theta)^3} = \frac{2\theta^3}{8} = \frac{\theta^3}{4}$. $E[Y^2] = \frac{2}{\theta}\left[2\theta^3 - \frac{\theta^3}{4}\right] = \frac{2}{\theta}\left[\frac{8\theta^3 - \theta^3}{4}\right] = \frac{2}{\theta}\left[\frac{7\theta^3}{4}\right] = \frac{7\theta^2}{2}$.

Variance of Y: $Var(Y) = E[Y^2] - (E[Y])^2$ $Var(Y) = \frac{7\theta^2}{2} - \left(\frac{3\theta}{2}\right)^2 = \frac{7\theta^2}{2} - \frac{9\theta^2}{4} = \frac{14\theta^2 - 9\theta^2}{4} = \frac{5\theta^2}{4}$. This matches. So, the assumption that the PDF was $f(x, y) = \frac{2}{\theta^2} e^{-(x+y)/\theta}$ was correct. The original problem statement likely had a typo omitting the θ in the exponent.

- **(ii) Show that $E[Y|X = x] = \phi(x) = x + \theta$ and expected value of $X + \theta$ is that of Y.**
  - **Conditional PDF of Y given X=x:** $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$ First, find the marginal PDF of X: $f_X(x) = \int_x^\infty \frac{2}{\theta^2} e^{-(x+y)/\theta} dy = \frac{2}{\theta^2} e^{-x/\theta} \int_x^\infty e^{-y/\theta}\, dy$ $f_X(x) = \frac{2}{\theta^2} e^{-x/\theta}[-\theta e^{-y/\theta}]_x^\infty = \frac{2}{\theta^2} e^{-x/\theta}[0 - (-\theta e^{-x/\theta})]$ $f_X(x) = \frac{2}{\theta^2} e^{-x/\theta}(\theta e^{-x/\theta}) = \frac{2}{\theta} e^{-2x/\theta}$, for $x > 0$.

    Now, the conditional PDF: $f_{Y|X}(y|x) = \frac{\frac{2}{\theta^2} e^{-(x+y)/\theta}}{\frac{2}{\theta} e^{-2x/\theta}} = \frac{1}{\theta} e^{-(x+y)/\theta} e^{2x/\theta} = \frac{1}{\theta} e^{(-x-y+2x)/\theta} = \frac{1}{\theta} e^{-(y-x)/\theta}$, for $y > x$. This is an exponential distribution shifted by x. Let $Z = Y - x$. Then Z follows an exponential distribution with rate $1/\theta$ (mean θ).
  - **Conditional Expectation $E[Y|X = x]$:** $E[Y|X = x] = \int_x^\infty y \frac{1}{\theta} e^{-(y-x)/\theta} dy$ Let $u = y - x$, so $y = u + x$ and $dy = du$. When $y = x, u = 0$. When $y = \infty, u = \infty$. $E[Y|X = x] = \int_0^\infty (u + x) \frac{1}{\theta} e^{-u/\theta} du = \frac{1}{\theta}\left[\int_0^\infty u\, e^{-u/\theta} du + \int_0^\infty x\, e^{-u/\theta} du\right]$ The first integral is $E[U]$ for $U \sim Exp(1/\theta)$, which is θ. The second integral is $x \int_0^\infty e^{-u/\theta}\, du = x[-\theta e^{-u/\theta}]_0^\infty = x(0 - (-\theta)) = x\theta$. So, $E[Y|X = x] = \frac{1}{\theta}[\theta^2 + x\theta] = \theta + x$. Thus, $E[Y|X = x] = \phi(x) = x + \theta$. This matches.

- **Expected value of $X+\theta$ is that of Y:** $E[X+\theta]=E[X]+\theta$. We need to find $E[X]$. $E[X]=\int_0^\infty x\, f_X(x)dx=\int_0^\infty x\,\frac{2}{\theta}e^{-2x/\theta}dx$ Recall $\int_0^\infty x\,e^{-ax}dx=\frac{1}{a^2}$. Here $a=2/\theta$. $E[X]=\frac{2}{\theta}\frac{1}{(2/\theta)^2}=\frac{2}{\theta}\frac{\theta^2}{4}=\frac{\theta}{2}$. So, $E[X+\theta]=\frac{\theta}{2}+\theta=\frac{3\theta}{2}$. From part (i), we found $E[Y]=\frac{3\theta}{2}$. Thus, $E[X+\theta]=E[Y]$. This matches.

- **(iii) Show that the variance of $\phi(X)$ is less than that of Y.** We know $\phi(X)=X+\theta$. So, $Var(\phi(X))=Var(X+\theta)=Var(X)$. We need $Var(X)$. $E[X^2]=\int_0^\infty x^2\, f_X(x)dx=\int_0^\infty x^2\,\frac{2}{\theta}e^{-2x/\theta}dx$ Recall $\int_0^\infty x^2\,e^{-ax}dx=\frac{2}{a^3}$. Here $a=2/\theta$. $E[X^2]=\frac{2}{\theta}\frac{2}{(2/\theta)^3}=\frac{4}{\theta}\frac{\theta^3}{8}=\frac{\theta^2}{2}$. $Var(X)=E[X^2]-(E[X])^2=\frac{\theta^2}{2}-\left(\frac{\theta}{2}\right)^2=\frac{\theta^2}{2}-\frac{\theta^2}{4}=\frac{\theta^2}{4}$. So, $Var(\phi(X))=Var(X)=\frac{\theta^2}{4}$. From part (i), $Var(Y)=\frac{5\theta^2}{4}$. Clearly, $Var(\phi(X))=\frac{\theta^2}{4}<\frac{5\theta^2}{4}=Var(Y)$. This matches.

- **(iv) Comment on the result.** The result $Var(E[Y|X])\leq Var(Y)$ is a fundamental property in probability theory, often stated as $Var(Y)=E[Var(Y|X)]+Var(E[Y|X])$. In our case, $Var(E[Y|X])=Var(\phi(X))=Var(X+\theta)=Var(X)$. And $Var(Y)$ is the total variance of Y. The inequality $Var(E[Y|X])\leq Var(Y)$ always holds. This is because conditioning on a random variable X generally reduces the variance, as some of the randomness in Y is explained by X. Here, $Var(X)=\frac{\theta^2}{4}$ and $Var(Y)=\frac{5\theta^2}{4}$. This shows that knowing X (or the expected value of Y given X) provides considerable information, as the variance of the conditional expectation is significantly smaller than the total variance of Y. The difference, $Var(Y)-Var(X)=\frac{5\theta^2}{4}-\frac{\theta^2}{4}=\frac{4\theta^2}{4}=\theta^2$, is $E[Var(Y|X)]$. Let's check $Var(Y|X=x)$. For $f_{Y|X}(y|x)=\frac{1}{\theta}e^{-(y-x)/\theta}$, this is an exponential distribution starting from x. $Y-x\sim Exp(1/\theta)$. So $Var(Y-x)=\theta^2$. $Var(Y|X=x)=Var(Y-x)=\theta^2$. Then $E[Var(Y|X)]=$

$E[\theta^2] = \theta^2$. And $Var(Y) = E[Var(Y|X)] + Var(E[Y|X]) = \theta^2 + \frac{\theta^2}{4} = \frac{5\theta^2}{4}$. This confirms the identity and the result. It highlights the reduction in variance due to conditioning.

3. (b) Define Minimum Variance Unbiased (MVU) Estimator. Let T_0 be an MVU estimator, while T_1 is an unbiased estimator with efficiency e_0. If ρ_{01}, be the correlation coefficient between T_0 and T_1, then prove that $\rho_{01} = \sqrt{e_0}$.

- **Minimum Variance Unbiased (MVU) Estimator:** An estimator T^* is called a Minimum Variance Unbiased (MVU) estimator (also known as a Uniformly Minimum Variance Unbiased Estimator or UMVUE) for a parameter θ if:
  - T^* is an unbiased estimator of θ, i.e., $E[T^*] = \theta$.
  - For any other unbiased estimator T of θ, $Var(T^*) \leq Var(T)$ for all possible values of θ.
- **Proof that $\rho_{01} = \sqrt{e_0}$:** Let T_0 be an MVU estimator of θ. So, $E[T_0] = \theta$. Let T_1 be another unbiased estimator of θ. So, $E[T_1] = \theta$. The efficiency of T_1 (relative to T_0) is defined as $e_0 = \frac{Var(T_0)}{Var(T_1)}$. Since T_0 is an MVU estimator, $Var(T_0) \leq Var(T_1)$, which implies $0 < e_0 \leq 1$.

  Consider the estimator $T_0 - T_1$. $E[T_0 - T_1] = E[T_0] - E[T_1] = \theta - \theta = 0$.

  Consider the variance of T_0: $Var(T_0) = Var(T_0 - T_1 + T_1)$ $Var(T_0) = Var((T_0 - T_1) + T_1)$ Using $Var(A + B) = Var(A) + Var(B) + 2Cov(A,B)$: $Var(T_0) = Var(T_0 - T_1) + Var(T_1) + 2Cov(T_0 - T_1, T_1)$.

  We know that T_0 is the MVU estimator. By the Rao-Blackwell theorem (or properties of sufficient statistics), if an MVU estimator exists, it is unique (up to sets of measure zero). Also, if an MVU estimator T_0 exists, then $Cov(T_0, T_1 - T_0) = 0$. This implies that T_0 is uncorrelated with any unbiased estimator of zero. Let $D = T_1 - T_0$. Then $E[D] = 0$. Since T_0 is MVU, it is uncorrelated with D. So, $Cov(T_0, T_1 - T_0) = 0$. This means

$Cov(T_0, T_1) - Cov(T_0, T_0) = 0$. $Cov(T_0, T_1) - Var(T_0) = 0$. Therefore, $Cov(T_0, T_1) = Var(T_0)$.

Now, let's use the definition of the correlation coefficient: $\rho_{01} = \frac{Cov(T_0,T_1)}{\sqrt{Var(T_0)Var(T_1)}}$ Substitute $Cov(T_0, T_1) = Var(T_0)$: $\rho_{01} = \frac{Var(T_0)}{\sqrt{Var(T_0)Var(T_1)}}$ $\rho_{01} = \frac{\sqrt{Var(T_0)}}{\sqrt{Var(T_1)}}$ $\rho_{01} = \sqrt{\frac{Var(T_0)}{Var(T_1)}}$ Since $e_0 = \frac{Var(T_0)}{Var(T_1)}$, we have: $\rho_{01} = \sqrt{e_0}$. This completes the proof.

4. (a) Let $X_1, X_2, \ldots, X_n$, be a random sample from U(0,θ) population. Show that the largest order statistic $X_{(n)}$ is a complete sufficient statistic for θ. Using Lehman Scheffe's theorem, find UMVUE of θ.

- **PDF of U(0, θ):** $f(x, \theta) = \frac{1}{\theta}$ for $0 < x < \theta$, and 0 otherwise. This can be written as $f(x, \theta) = \frac{1}{\theta} I(0 < x < \theta)$, where $I(\cdot)$ is the indicator function.
- **Sufficiency of $X_{(n)}$ using Factorization Theorem:** The likelihood function for a sample $X_1, \ldots, X_n$ is: $L(\theta | x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i, \theta) = \prod_{i=1}^{n} \frac{1}{\theta} I(0 < x_i < \theta)$ $L(\theta | x_1, \ldots, x_n) = \frac{1}{\theta^n} \prod_{i=1}^{n} I(0 < x_i < \theta)$ The product of indicator functions is $I(0 < x_1 < \theta$ and … and $0 < x_n < \theta)$. This is equivalent to $I(\max(x_i) < \theta$ and $\min(x_i) > 0)$. Since $x_i > 0$ is implicitly satisfied by the range, the condition simplifies to $I(x_{(n)} < \theta)$. So, $L(\theta | x_1, \ldots, x_n) = \frac{1}{\theta^n} I(x_{(n)} < \theta) I(x_{(1)} > 0)$. This can be factored as $g(T(\mathbf{x}), \theta) h(\mathbf{x})$, where $T(\mathbf{x}) = X_{(n)}$. $g(X_{(n)}, \theta) = \frac{1}{\theta^n} I(X_{(n)} < \theta)$ (This part depends on $X_{(n)}$ and θ). $h(\mathbf{x}) = I(X_{(1)} > 0)$ (This part does not depend on θ). According to the Fisher-Neyman Factorization Theorem, $X_{(n)}$ is a sufficient statistic for θ.
- **Completeness of $X_{(n)}$:** The PDF of $X_{(n)}$ for $U(0, \theta)$ is $f_{X_{(n)}}(y) = n\left(\frac{y}{\theta}\right)^{n-1} \frac{1}{\theta} = \frac{ny^{n-1}}{\theta^n}$ for $0 < y < \theta$. We need to show that if $E[k(X_{(n)})] = 0$ for all θ, then $k(y) = 0$ almost everywhere. $E[k(X_{(n)})] = \int_0^{\theta} k(y) \frac{ny^{n-1}}{\theta^n} dy = 0$ $\frac{n}{\theta^n} \int_0^{\theta} k(y) y^{n-1} dy = 0$ Since $\frac{n}{\theta^n} \neq 0$, we must have

$\int_0^\theta k\,(y)y^{n-1}dy = 0$ for all $\theta > 0$. Let $G(\theta) = \int_0^\theta k\,(y)y^{n-1}dy$. If $G(\theta) = 0$ for all θ, then its derivative with respect to θ must also be 0. By the Fundamental Theorem of Calculus, $\frac{d}{d\theta}G(\theta) = k(\theta)\theta^{n-1}$. Since $k(\theta)\theta^{n-1} = 0$ for all $\theta > 0$, and $\theta^{n-1} \neq 0$, it implies $k(\theta) = 0$ for all θ in $(0, \infty)$. Therefore, $X_{(n)}$ is a complete sufficient statistic for θ.

- **UMVUE of θ using Lehmann-Scheffé Theorem:** The Lehmann-Scheffé theorem states that if T is a complete sufficient statistic for θ, and $E[g(T)] = \theta$ for some function g, then $g(T)$ is the unique UMVUE of θ. We need to find an unbiased estimator of θ that is a function of $X_{(n)}$. Let's find $E[X_{(n)}]$: $E[X_{(n)}] = \int_0^\theta y\,\frac{ny^{n-1}}{\theta^n}dy = \frac{n}{\theta^n}\int_0^\theta y^n\,dy = \frac{n}{\theta^n}\left[\frac{y^{n+1}}{n+1}\right]_0^\theta = \frac{n}{\theta^n}\frac{\theta^{n+1}}{n+1} = \frac{n\theta}{n+1}$. So, $E[X_{(n)}] = \frac{n\theta}{n+1}$. To make it an unbiased estimator of θ, we need to multiply $X_{(n)}$ by a constant: Let $g(X_{(n)}) = cX_{(n)}$. $E[cX_{(n)}] = cE[X_{(n)}] = c\frac{n\theta}{n+1}$. For $E[cX_{(n)}] = \theta$, we must have $c\,\frac{n}{n+1} = 1$, so $c = \frac{n+1}{n}$. Therefore, $g(X_{(n)}) = \frac{n+1}{n}X_{(n)}$ is an unbiased estimator of θ. Since $X_{(n)}$ is a complete sufficient statistic for θ, by the Lehmann-Scheffé theorem, $\frac{n+1}{n}X_{(n)}$ is the UMVUE of θ.

4. (b) Let $X_1, X_2, \ldots, X_n$, be a random sample from the distribution having p.d.f. $f(x,\theta) = \frac{1}{\theta}e^{-x/\theta}$, $0 \le x < \infty, \theta > 0$. Show that $\frac{n-1}{n\overline{X}}$ is the only unbiased estimator of $\frac{1}{\theta}$ based on $\overline{X}$.

- **PDF of Exponential Distribution:** $f(x,\theta) = \frac{1}{\theta}e^{-x/\theta}$, $x \ge 0, \theta > 0$. This is an exponential distribution with mean θ.

- **Sufficiency and Completeness of $\overline{X}$ (or $\sum X_i$):** The sum $T = \sum_{i=1}^n X_i$ from an exponential distribution follows a Gamma distribution with shape parameter n and rate parameter $1/\theta$. The PDF of T is $f_T(t) = \frac{(1/\theta)^n}{\Gamma(n)}t^{n-1}e^{-t/\theta}$ for $t > 0$. By the Factorization Theorem, $T = \sum X_i$ (or equivalently $\overline{X} = T/n$) is a sufficient statistic for θ. To show completeness, we need to show that if $E[k(T)] = 0$ for all θ, then $k(t) = 0$ a.e. $E[k(T)] = \int_0^\infty k(t) \frac{(1/\theta)^n}{\Gamma(n)} t^{n-1} e^{-t/\theta} dt = 0$. $\frac{(1/\theta)^n}{\Gamma(n)} \int_0^\infty k(t) t^{n-1} e^{-t/\theta} dt = 0$. Since $\frac{(1/\theta)^n}{\Gamma(n)} \neq 0$, we have $\int_0^\infty k(t) t^{n-1} e^{-t/\theta} dt = 0$. This is the Laplace transform of $k(t)t^{n-1}$ at $s = 1/\theta$. Since the Laplace transform is unique, if it's zero for all $s > 0$, then $k(t)t^{n-1} = 0$ almost everywhere. Since $t^{n-1} \neq 0$ for $t > 0$, it implies $k(t) = 0$ almost everywhere. Thus, $T = \sum X_i$ (and hence $\bar{X}$) is a complete sufficient statistic for θ.

- **Finding UMVUE of $\frac{1}{\theta}$:** We need an unbiased estimator of $\frac{1}{\theta}$ based on $\bar{X}$ (or $T = n\bar{X}$). Consider $E[1/T]$. We know $T \sim Gamma(n, 1/\theta)$. $E[1/T] = \int_0^\infty \frac{1}{t} \frac{(1/\theta)^n}{\Gamma(n)} t^{n-1} e^{-t/\theta} dt$ $E[1/T] = \frac{(1/\theta)^n}{\Gamma(n)} \int_0^\infty t^{n-2} e^{-t/\theta} dt$ The integral is of the form $\int_0^\infty x^{k-1} e^{-ax} dx = \frac{\Gamma(k)}{a^k}$. Here $k = n - 1$ and $a = 1/\theta$. $E[1/T] = \frac{(1/\theta)^n}{\Gamma(n)} \frac{\Gamma(n-1)}{(1/\theta)^{n-1}}$ $E[1/T] = \frac{(1/\theta)^n}{(n-1)\Gamma(n-1)} \frac{\Gamma(n-1)}{(1/\theta)^{n-1}}$ $E[1/T] = \frac{1/\theta}{n-1} = \frac{1}{(n-1)\theta}$. So, $E\left[\frac{1}{T}\right] = \frac{1}{(n-1)\theta}$. We want an unbiased estimator of $\frac{1}{\theta}$. Multiply $1/T$ by $(n-1)$: $E\left[\frac{n-1}{T}\right] = \frac{n-1}{(n-1)\theta} = \frac{1}{\theta}$. Since $T = n\bar{X}$, substitute this into the expression: $\frac{n-1}{T} = \frac{n-1}{n\bar{X}}$. Thus, $\frac{n-1}{n\bar{X}}$ is an unbiased estimator of $\frac{1}{\theta}$. Since $\bar{X}$ is a complete sufficient statistic for θ, and $\frac{n-1}{n\bar{X}}$ is an unbiased estimator of $\frac{1}{\theta}$ based on $\bar{X}$, by the Lehmann-Scheffé theorem, it is the unique UMVUE of $\frac{1}{\theta}$. Therefore, it is the only unbiased estimator of $\frac{1}{\theta}$ based on $\bar{X}$.

SECTION II

5. (a) Describe the Method of Moments. Let $X_1, X_2, \ldots, X_n$, be a random sample from the distribution with p.d.f. $f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma\alpha} x^{\alpha-1} e^{-\beta x}, x \geq 0, \alpha > 0, \beta > 0$. Estimate the parameters α and β by the Method of Moments.

- **Method of Moments (MOM):** The Method of Moments is a technique for estimating parameters of a probability distribution. The principle is to equate the theoretical moments of the distribution (which are functions of the parameters) to the corresponding sample moments (which are calculated from the observed data). If there are k parameters to be estimated, then the first k theoretical moments are equated to the first k sample moments.

- The r-th theoretical moment (about the origin) is $\mu'_r = E[X^r]$.
- The r-th sample moment (about the origin) is $m'_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r$. The method involves solving the system of equations: $\mu'_1 = m'_1$ $\mu'_2 = m'_2$ ... $\mu'_k = m'_k$ The solutions for the parameters are the Method of Moments estimators.

- **Estimation for Gamma Distribution:** The given p.d.f. is that of a Gamma distribution, $X \sim Gamma(\alpha, \beta)$, where α is the shape parameter and β is the rate parameter (or $1/\beta$ is the scale parameter). We need to estimate two parameters, α and β, so we will use the first two moments.

f. **First Theoretical Moment (μ'_1):** For a Gamma distribution, $E[X] = \frac{\alpha}{\beta}$. The first sample moment is $m'_1 = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Equating them: $\frac{\alpha}{\beta} = \overline{X}$ (Equation 1)

g. **Second Theoretical Moment (μ'_2):** For a Gamma distribution, $Var(X) = \frac{\alpha}{\beta^2}$. Also, $Var(X) = E[X^2] - (E[X])^2 = \mu'_2 - (\mu'_1)^2$. So, $\mu'_2 = Var(X) + (\mu'_1)^2 = \frac{\alpha}{\beta^2} + \left(\frac{\alpha}{\beta}\right)^2 = \frac{\alpha}{\beta^2} + \frac{\alpha^2}{\beta^2} = \frac{\alpha(\alpha+1)}{\beta^2}$. The second sample moment is $m'_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$. Equating them: $\frac{\alpha(\alpha+1)}{\beta^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ (Equation 2)

h. **Solving for α and β:** From Equation 1, $\alpha = \beta\overline{X}$. Substitute this into Equation 2: $\frac{\beta\overline{X}(\beta\overline{X}+1)}{\beta^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ $\frac{\overline{X}(\beta\overline{X}+1)}{\beta} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ $\overline{X}^2 + \frac{\overline{X}}{\beta} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ $\frac{\overline{X}}{\beta} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2$ We know that the sample variance $S^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2$. So, $\frac{\overline{X}}{\beta} = S^2$. Therefore, $\hat{\beta}_{MOM} = \frac{\overline{X}}{S^2}$.

Now substitute $\hat{\beta}_{MOM}$ back into $\alpha = \beta\overline{X}$: $\hat{\alpha}_{MOM} = \frac{\overline{X}}{S^2}\overline{X} = \frac{\overline{X}^2}{S^2}$.

Thus, the Method of Moments estimators for α and β are: $\hat{\alpha} = \frac{\overline{X}^2}{S^2}$ $\hat{\beta} = \frac{\overline{X}}{S^2}$ where $\overline{X}$ is the sample mean and $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ is the sample variance.

5. (b) Let $X_1, X_2, \ldots, X_n$, be a random sample from a distribution with p.d.f. $f(x;\theta) = \frac{1}{\theta}x^{\frac{1-\theta}{\theta}}$, $0 < x < 1, 0 < \theta < \infty$. Find the Maximum likelihood estimator of θ. Also, examine the estimator for unbiasedness property.

- **Likelihood Function:** $L(\theta|x_1,\ldots,x_n) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n}\frac{1}{\theta}x_i^{\frac{1-\theta}{\theta}}$ $L(\theta|x_1,\ldots,x_n) = \left(\frac{1}{\theta}\right)^n \prod_{i=1}^{n} x_i^{\frac{1}{\theta}-1}$ $L(\theta|x_1,\ldots,x_n) = \theta^{-n}\exp\left[\left(\frac{1}{\theta} - 1\right)\sum_{i=1}^{n}\log x_i\right]$
- **Log-Likelihood Function:** $\log L(\theta) = -n\log\theta + \left(\frac{1}{\theta} - 1\right)\sum_{i=1}^{n}\log x_i$
- **First Derivative of Log-Likelihood:** $\frac{\partial}{\partial\theta}\log L(\theta) = -n\frac{1}{\theta} - \frac{1}{\theta^2}\sum_{i=1}^{n}\log x_i$
- **Setting the Derivative to Zero for MLE:** $\frac{\partial}{\partial\theta}\log L(\theta) = 0$ $-n\frac{1}{\theta} - \frac{1}{\theta^2}\sum_{i=1}^{n}\log x_i = 0$ Multiply by $-\theta^2$: $n\theta + \sum_{i=1}^{n}\log x_i = 0$ $n\theta = -\sum_{i=1}^{n}\log x_i$ $\hat{\theta}_{MLE} = -\frac{1}{n}\sum_{i=1}^{n}\log x_i$
- **Examining Unbiasedness of $\hat{\theta}_{MLE}$:** We need to find $E[\hat{\theta}_{MLE}] = E\left[-\frac{1}{n}\sum_{i=1}^{n}\log X_i\right] = -\frac{1}{n}\sum_{i=1}^{n}E[\log X_i]$. Since X_i are i.i.d., $E[\log X_i] = E[\log X]$. $E[\log X] = \int_0^1(\log x)\frac{1}{\theta}x^{\frac{1-\theta}{\theta}}dx$ Let $k = \frac{1-\theta}{\theta}$. So the p.d.f. is $\frac{1}{\theta}x^k$. $E[\log X] = \frac{1}{\theta}\int_0^1(\log x)x^k dx$. We use integration by parts: $\int u dv = uv - \int v du$. Let $u = \log x$, $dv = x^k dx$. Then $du = \frac{1}{x}dx$, $v = \frac{x^{k+1}}{k+1}$. $E[\log X] = \frac{1}{\theta}\left[\left[(\log x)\frac{x^{k+1}}{k+1}\right]_0^1 - \int_0^1\frac{x^{k+1}}{k+1}\frac{1}{x}dx\right]$ The first term: $[\log x \cdot x^{k+1}]_{x=1} = \log(1) \cdot$

$1^{k+1} = 0$. As $x \to 0$, $\lim_{x\to 0}(\log x)x^{k+1}$. Since $k+1 = \frac{1-\theta}{\theta} + 1 = \frac{1-\theta+\theta}{\theta} = \frac{1}{\theta}$. Since $\theta > 0$, $k+1 > 0$. So $\lim_{x\to 0}(\log x)x^{1/\theta} = 0$. (This is a standard limit, $x^a \log x \to 0$ as $x \to 0$ for $a > 0$). So the first term is 0. $E[\log X] = \frac{1}{\theta}\left[-\int_0^1 \frac{x^k}{k+1}dx\right] = -\frac{1}{\theta(k+1)}\int_0^1 x^k\, dx = -\frac{1}{\theta(k+1)}\left[\frac{x^{k+1}}{k+1}\right]_0^1 = -\frac{1}{\theta(k+1)}\left(\frac{1}{k+1} - 0\right) = -\frac{1}{\theta(k+1)^2}$. Substitute $k+1 = \frac{1}{\theta}$: $E[\log X] = -\frac{1}{\theta(1/\theta)^2} = -\frac{1}{\theta(1/\theta^2)} = -\theta$.

Now, substitute $E[\log X]$ back into $E[\hat{\theta}_{MLE}]$: $E[\hat{\theta}_{MLE}] = -\frac{1}{n}\sum_{i=1}^{n}(-\theta) = -\frac{1}{n}(n(-\theta)) = \theta$. Since $E[\hat{\theta}_{MLE}] = \theta$, the Maximum Likelihood Estimator $\hat{\theta}_{MLE} = -\frac{1}{n}\sum_{i=1}^{n}\log X_i$ is an unbiased estimator of θ.

6. (a) What is failure-censored sample in life testing experiment? Obtain the Maximum likelihood estimator of the expected life time and reliability function in case of failure censored sample from the life time distribution: $f(x,\sigma) = \frac{1}{\sigma}e^{-x/\sigma}$, $x > 0, \sigma > 0$.

- **Failure-Censored Sample (Type II Censoring):** In life testing experiments, it is often impractical or too time-consuming to observe all units until they fail, especially if the products have long lifetimes. Censoring occurs when the exact lifetime of a unit is not observed. A failure-censored sample (specifically, Type II censoring) occurs when an experiment is terminated as soon as a predetermined number of failures (r) have occurred, out of n units put on test. The observed lifetimes are the first r ordered failure times, $X_{(1)} \le X_{(2)} \le \ldots \le X_{(r)}$. The remaining $n-r$ units are still functioning at time $X_{(r)}$ (the r-th failure time), and their exact lifetimes are unknown, but they are known to be at least $X_{(r)}$.

- **MLE for Exponential Lifetime Distribution with Type II Censoring:** The lifetime distribution is $f(x,\sigma) = \frac{1}{\sigma}e^{-x/\sigma}$, which is an exponential distribution with mean σ. So, the expected lifetime is σ. The reliability function is $R(t) = P(X > t) = \int_t^\infty \frac{1}{\sigma} e^{-x/\sigma} dx = [-e^{-x/\sigma}]_t^\infty = 0 - (-e^{-t/\sigma}) = e^{-t/\sigma}$.

Let $X_{(1)}, X_{(2)}, \dots, X_{(r)}$ be the first r ordered failure times from a sample of n units. The likelihood function for Type II censoring is given by: $L(\sigma|x_{(1)}, \dots, x_{(r)}) = \frac{n!}{(n-r)!}\left(\prod_{i=1}^r f\ (x_{(i)}, \sigma)\right)[R(x_{(r)}, \sigma)]^{n-r}$ $L(\sigma) = \frac{n!}{(n-r)!}\left(\prod_{i=1}^r \frac{1}{\sigma} e^{-x_{(i)}/\sigma}\right)[e^{-x_{(r)}/\sigma}]^{n-r}$ $L(\sigma) = \frac{n!}{(n-r)!}\sigma^{-r} \exp\left(-\frac{1}{\sigma}\sum_{i=1}^r x_{(i)}\right) e^{-(n-r)x_{(r)}/\sigma}$ $L(\sigma) = \frac{n!}{(n-r)!}\sigma^{-r} \exp\left(-\frac{1}{\sigma}\left[\sum_{i=1}^r x_{(i)} + (n-r)x_{(r)}\right]\right)$

Let $S_r = \sum_{i=1}^r x_{(i)} + (n-r)x_{(r)}$. This is the total observed time on test. The log-likelihood function is: $\log L(\sigma) = \log\left(\frac{n!}{(n-r)!}\right) - r\log\sigma - \frac{S_r}{\sigma}$

MLE of Expected Lifetime (σ): Differentiate $\log L(\sigma)$ with respect to σ and set to zero: $\frac{\partial}{\partial\sigma}\log L(\sigma) = -\frac{r}{\sigma} + \frac{S_r}{\sigma^2} = 0$ Multiply by σ^2: $-r\sigma + S_r = 0$ $r\sigma = S_r$ $\hat{\sigma}_{MLE} = \frac{S_r}{r} = \frac{\sum_{i=1}^r X_{(i)} + (n-r)X_{(r)}}{r}$. This is the MLE of the expected lifetime σ.

MLE of Reliability Function ($R(t)$): The reliability function is $R(t) = e^{-t/\sigma}$. By the invariance property of MLEs, the MLE of $R(t)$ is obtained by substituting the MLE of σ into the expression for $R(t)$. $\hat{R}_{MLE}(t) = e^{-t/\hat{\sigma}_{MLE}} = e^{-t/\left(\frac{\sum_{i=1}^r X_{(i)} + (n-r)X_{(r)}}{r}\right)}$.

6. (b) Show by means of an example that: (i) MLE may not be unbiased. (ii) MLE may not be unique.

- **(i) MLE may not be unbiased (Example: Variance of Normal Distribution):** Let $X_1, X_2, \dots, X_n$ be a random sample from a Normal distribution $N(\mu, \sigma^2)$, where μ is known (e.g., $\mu = 0$). The PDF is $f(x, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/(2\sigma^2)}$. The likelihood function is $L(\sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-X_i^2/(2\sigma^2)}$ $L(\sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n X_i^2\right)$. The log-likelihood function is $\log L(\sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n} X_i^2$. Differentiating with respect to σ^2: $\frac{\partial}{\partial\sigma^2}\log L(\sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n} X_i^2$. Setting to zero: $-\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n} X_i^2 = 0$ $\frac{n}{2\sigma^2} = \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n} X_i^2$ $n\sigma^2 = \sum_{i=1}^{n} X_i^2$ $\widehat{\sigma^2}_{MLE} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$. Now, let's check for unbiasedness: $E[\widehat{\sigma^2}_{MLE}] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2\right] = \frac{1}{n}\sum_{i=1}^{n} E\,[X_i^2]$. Since $\mu = 0$, $E[X_i^2] = Var(X_i) + (E[X_i])^2 = \sigma^2 + 0^2 = \sigma^2$. $E[\widehat{\sigma^2}_{MLE}] = \frac{1}{n}\sum_{i=1}^{n} \sigma^2 = \frac{1}{n}(n\sigma^2) = \sigma^2$. In this case (where μ is known to be 0), the MLE is unbiased.

Let's use the more common example where μ is unknown. Let $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$, both μ and σ^2 are unknown. The MLEs are $\hat{\mu}_{MLE} = \overline{X}$ and $\widehat{\sigma^2}_{MLE} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$. We know that $E[\overline{X}] = \mu$, so $\hat{\mu}_{MLE}$ is unbiased. However, for $\widehat{\sigma^2}_{MLE}$: $E[\widehat{\sigma^2}_{MLE}] = E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2\right] = E\left[\frac{(n-1)S^2}{n}\right]$ where $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ is the unbiased sample variance, with $E[S^2] = \sigma^2$. So, $E[\widehat{\sigma^2}_{MLE}] = \frac{n-1}{n}E[S^2] = \frac{n-1}{n}\sigma^2$. Since $\frac{n-1}{n} \neq 1$, $\widehat{\sigma^2}_{MLE}$ is a biased estimator of σ^2. It is asymptotically unbiased, as $\frac{n-1}{n} \to 1$ as $n \to \infty$.

- **(ii) MLE may not be unique (Example: Uniform Distribution):** Let $X_1, X_2, \ldots, X_n$ be a random sample from $U(\theta, \theta+1)$. The PDF is $f(x,\theta) = 1$ for $\theta < x < \theta + 1$, and 0 otherwise. The likelihood function is $L(\theta|x_1,\ldots,x_n) = \prod_{i=1}^{n} 1 \cdot I(\theta < x_i < \theta + 1)$ $L(\theta|x_1,\ldots,x_n) = I(\theta < x_1 < \theta+1, \ldots, \theta < x_n < \theta + 1)$ This is 1 if $\theta < x_i < \theta + 1$ for all i, and 0 otherwise. These inequalities can be rewritten as: $x_i > \theta$ for all $i \Rightarrow \theta < \min(x_i) = X_{(1)}$. $x_i < \theta + 1$ for all $i \Rightarrow \theta + 1 > \max(x_i) = X_{(n)} \Rightarrow \theta > X_{(n)} - 1$. So, the likelihood function is 1 if $X_{(n)} - 1 < \theta < X_{(1)}$, and 0 otherwise. The likelihood function is constant (equal to 1) for any θ in the interval $(X_{(n)} - 1, X_{(1)})$. Any value of θ in this interval maximizes the likelihood function. For example, if $X_{(n)} - 1 < X_{(1)}$, then there is an interval of possible θ values. For instance, if $X_{(n)} = 5$ and $X_{(1)} = 4.5$, then $4 < \theta <$

4.5. Any θ in (4,4.5) is an MLE. If $X_{(n)} = 5$ and $X_{(1)} = 5$, then $4 < \theta < 5$. Any θ in (4,5) is an MLE. This shows that the MLE is not unique.

7. (a) Let $X_1, X_2, \ldots, X_n$, be a random sample from a distribution having p.d.f. $f(x, \theta) = e^{-(x-\theta)}$, $\theta \leq x < \infty, -\infty > \theta < \infty$. Obtain 100 (1-$\alpha$)% confidence interval for θ.

- **PDF of Shifted Exponential Distribution:** $f(x, \theta) = e^{-(x-\theta)} = e^{\theta} e^{-x}$ for $x \geq \theta$. This is an exponential distribution with rate 1, shifted by θ.

- **Finding a Pivotal Quantity:** Consider $X_{(1)} = \min(X_1, \ldots, X_n)$. The CDF of X is $F(x, \theta) = \int_{\theta}^{x} e^{-(u-\theta)}\, du = e^{\theta} \int_{\theta}^{x} e^{-u}\, du = e^{\theta}[-e^{-u}]_{\theta}^{x} = e^{\theta}(-e^{-x} - (-e^{-\theta})) = e^{\theta}(e^{-\theta} - e^{-x}) = 1 - e^{-(x-\theta)}$ for $x \geq \theta$. The CDF of $X_{(1)}$ is $F_{X_{(1)}}(y) = P(X_{(1)} \leq y) = 1 - P(X_{(1)} > y)$. $P(X_{(1)} > y) = P(X_1 > y, \ldots, X_n > y) = \prod_{i=1}^{n} P(X_i > y) = [P(X > y)]^n$. $P(X > y) = 1 - F(y, \theta) = 1 - (1 - e^{-(y-\theta)}) = e^{-(y-\theta)}$ for $y \geq \theta$. So, $P(X_{(1)} > y) = [e^{-(y-\theta)}]^n = e^{-n(y-\theta)}$ for $y \geq \theta$. $F_{X_{(1)}}(y) = 1 - e^{-n(y-\theta)}$ for $y \geq \theta$. This is the CDF of a shifted exponential distribution with rate n and shift parameter θ. Let $Y = n(X_{(1)} - \theta)$. $P(Y \leq y_0) = P(n(X_{(1)} - \theta) \leq y_0) = P(X_{(1)} - \theta \leq y_0/n) = P(X_{(1)} \leq \theta + y_0/n)$. Using the CDF of $X_{(1)}$: $F_{X_{(1)}}(\theta + y_0/n) = 1 - e^{-n((\theta+y_0/n)-\theta)} = 1 - e^{-n(y_0/n)} = 1 - e^{-y_0}$. So, $Y = n(X_{(1)} - \theta)$ follows an exponential distribution with rate 1 (mean 1), i.e., $Y \sim Exp(1)$. Y is a pivotal quantity because its distribution does not depend on θ.

- **Constructing the Confidence Interval:** We need to find y_L and y_U such that $P(y_L < Y < y_U) = 1 - \alpha$. Since $Y \sim Exp(1)$, its PDF is $f_Y(y) = e^{-y}$ for $y > 0$. We usually choose y_L and y_U such that $P(Y < y_L) = \alpha/2$ and $P(Y > y_U) = \alpha/2$. $P(Y < y_L) = \int_0^{y_L} e^{-y}\, dy = [-e^{-y}]_0^{y_L} = 1 - e^{-y_L} = \alpha/2$. $e^{-y_L} = 1 - \alpha/2$ $-y_L = \log(1 - \alpha/2) \Rightarrow y_L = -\log(1 - \alpha/2)$.

  $P(Y > y_U) = \int_{y_U}^{\infty} e^{-y}\, dy = [-e^{-y}]_{y_U}^{\infty} = e^{-y_U} = \alpha/2$. $-y_U = \log(\alpha/2) \Rightarrow y_U = -\log(\alpha/2)$.

So, we have $y_L = -\log(1-\alpha/2)$ and $y_U = -\log(\alpha/2)$. $P(-\log(1-\alpha/2) < n(X_{(1)}-\theta) < -\log(\alpha/2)) = 1-\alpha$. Now, isolate θ: $-\log(1-\alpha/2) < n(X_{(1)}-\theta)$ $\frac{-\log(1-\alpha/2)}{n} < X_{(1)}-\theta$ $\theta < X_{(1)} - \frac{-\log(1-\alpha/2)}{n} = X_{(1)} + \frac{\log(1-\alpha/2)}{n}$ (Upper bound for θ)

$n(X_{(1)}-\theta) < -\log(\alpha/2)$ $X_{(1)}-\theta < \frac{-\log(\alpha/2)}{n}$ $\theta > X_{(1)} - \frac{-\log(\alpha/2)}{n} = X_{(1)} + \frac{\log(\alpha/2)}{n}$ (Lower bound for θ)

Combining these, the $100(1-\alpha)\%$ confidence interval for θ is: $\left(X_{(1)} + \frac{\log(\alpha/2)}{n}, X_{(1)} + \frac{\log(1-\alpha/2)}{n}\right)$. Since $\log(\alpha/2)$ is negative (as $\alpha/2 < 1$), and $\log(1-\alpha/2)$ is negative (as $1-\alpha/2 < 1$), the lower bound is actually the larger one if we just consider the absolute values of the log terms. For example, if $\alpha = 0.05$, $\alpha/2 = 0.025$, $1-\alpha/2 = 0.975$. $\log(0.025) \approx -3.689$ $\log(0.975) \approx -0.025$ So, $X_{(1)} - 3.689/n < \theta < X_{(1)} - 0.025/n$. The interval is $(X_{(1)} + \frac{\log(\alpha/2)}{n}, X_{(1)} + \frac{\log(1-\alpha/2)}{n})$.

7. (b) Describe a method of constructing confidence intervals. Find 100 (1-α)% confidence interval for binomial proportion based on a random sample of size n for large samples.

- **Method of Constructing Confidence Intervals (Pivotal Quantity Method):** A common method for constructing confidence intervals is the Pivotal Quantity Method.

   i. **Identify a parameter of interest:** Let θ be the unknown parameter for which we want to construct a confidence interval.

   j. **Find a pivotal quantity:** A pivotal quantity $Q(X_1, \ldots, X_n; \theta)$ is a function of the random sample $X_1, \ldots, X_n$ and the parameter θ, whose sampling distribution does not depend on θ or any other unknown parameters (nuisance parameters).

   k. **Determine the distribution of the pivotal quantity:** Find the probability distribution of Q.

l. **Find two constants q_L and q_U:** Choose q_L and q_U such that $P(q_L < Q < q_U) = 1 - \alpha$, where $1 - \alpha$ is the desired confidence level. These constants are typically chosen from the tables of the pivotal quantity's distribution (e.g., Z-table, t-table, chi-square table) such that $P(Q < q_L) = \alpha/2$ and $P(Q > q_U) = \alpha/2$ for a symmetric interval.

m. **Invert the inequality:** Manipulate the inequality $q_L < Q(X_1, \dots, X_n; \theta) < q_U$ algebraically to isolate the parameter θ. The resulting inequality will be of the form $L(X_1, \dots, X_n) < \theta < U(X_1, \dots, X_n)$, where L and U are the lower and upper bounds of the confidence interval, respectively. These bounds are functions of the sample observations and are random variables.

- **Confidence Interval for Binomial Proportion (Large Samples - Wald Interval):** Let $X \sim Bin(n, p)$, where p is the unknown binomial proportion. The point estimator for p is the sample proportion $\hat{p} = X/n$. For large samples, by the Central Limit Theorem, the sampling distribution of $\hat{p}$ is approximately normal with mean $E[\hat{p}] = p$ and variance $Var(\hat{p}) = \frac{p(1-p)}{n}$. Therefore, a suitable pivotal quantity is: $Z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$. However, this Z depends on p in the denominator. For large samples, we can substitute $\hat{p}$ for p in the standard error, leading to the estimated standard error $SE(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. So, the approximate pivotal quantity for large samples is: $Z_{approx} = \frac{\hat{p}-p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$, which approximately follows a standard normal distribution $N(0,1)$.

  For a $100(1-\alpha)\%$ confidence interval, we find $z_{\alpha/2}$ such that $P(-z_{\alpha/2} < Z_{approx} < z_{\alpha/2}) = 1 - \alpha$. $-z_{\alpha/2} < \frac{\hat{p}-p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} < z_{\alpha/2}$ Multiply by $SE(\hat{p})$: $-z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < \hat{p} - p < z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ Subtract $\hat{p}$ from all parts: $-\hat{p} -$

$z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < -p < -\hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ Multiply by -1 and reverse inequalities: $\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Thus, the $100(1-\alpha)\%$ confidence interval for the binomial proportion p for large samples (Wald interval) is: $\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$.

8. Describe any three of the following: (a) Minimum Chi-square method of estimation. (b) Rao-Blackwell theorem. (c) Fisher-Neyman criterion. (d) Sufficient conditions for consistency.

- **(a) Minimum Chi-square method of estimation:** The Minimum Chi-square method is a method of estimation primarily used for grouped data or when working with multinomial distributions. It involves choosing parameter estimates that minimize a chi-square goodness-of-fit statistic, which measures the discrepancy between observed frequencies and expected frequencies (calculated based on the estimated parameters). Let $N_1, N_2, \ldots, N_k$ be the observed frequencies in k categories, with $\sum N_j = n$. Let $p_1(\theta), p_2(\theta), \ldots, p_k(\theta)$ be the theoretical probabilities for these categories, which depend on the parameter(s) θ. The expected frequencies are $E_j(\theta) = np_j(\theta)$. The Pearson chi-square statistic is given by: $\chi^2(\theta) = \sum_{j=1}^{k} \frac{(N_j - np_j(\theta))^2}{np_j(\theta)}$ The minimum chi-square estimator $\hat{\theta}_{MCS}$ is the value of θ that minimizes this $\chi^2(\theta)$ statistic. This method is asymptotically equivalent to Maximum Likelihood Estimation under certain conditions. It is computationally more involved than Method of Moments but provides good asymptotic properties.

- **(b) Rao-Blackwell theorem:** The Rao-Blackwell theorem provides a method for improving an unbiased estimator by conditioning it on a sufficient statistic. It states that if T is an unbiased estimator of a parameter θ, and S is a sufficient statistic for θ, then the conditional expectation $T^* = E[T|S]$ is also an unbiased estimator of θ, and its variance is less than or equal to the variance of T. That is, $Var(T^*) \leq Var(T)$. If S is also a complete sufficient statistic, then T^* is the unique Minimum Variance Unbiased Estimator (MVUE) of θ. In essence, the theorem says that if you have an unbiased estimator, you can "Rao-Blackwellize" it by conditioning on a sufficient statistic to obtain a new estimator that is no worse (and often better) in terms of variance, while maintaining unbiasedness. This improved estimator is often the UMVUE if the sufficient statistic is also complete.

- **(c) Fisher-Neyman criterion (Factorization Theorem for Sufficiency):** The Fisher-Neyman criterion, also known as the Factorization Theorem, provides a convenient way to check if a statistic is sufficient. It states that a statistic $T(\mathbf{X}) = T(X_1, \dots, X_n)$ is a sufficient statistic for a parameter θ (or a vector of parameters $\boldsymbol{\theta}$) if and only if the joint probability density function (or probability mass function) of the sample, $f(\mathbf{x}; \boldsymbol{\theta})$, can be factored into two non-negative functions: $f(\mathbf{x}; \boldsymbol{\theta}) = g(T(\mathbf{x}), \boldsymbol{\theta})h(\mathbf{x})$ where:
  - $g(T(\mathbf{x}), \boldsymbol{\theta})$ depends on the sample $\mathbf{x}$ only through the statistic $T(\mathbf{x})$ and on the parameter(s) $\boldsymbol{\theta}$.
  - $h(\mathbf{x})$ does not depend on the parameter(s) $\boldsymbol{\theta}$ at all. This theorem simplifies the identification of sufficient statistics, as it avoids the direct calculation of conditional distributions, which can be complex.
- **(d) Sufficient conditions for consistency:** Consistency is a desirable property of an estimator, meaning that as the sample size n increases, the estimator converges in probability to the true value of the parameter. That is, $\hat{\theta}_n \xrightarrow{P} \theta$. Two common sufficient conditions for the consistency of an estimator $\hat{\theta}_n$ (often used for Maximum Likelihood Estimators, but applicable more broadly) are:
  n. **Asymptotic Unbiasedness and Variance tending to Zero:** If an estimator $\hat{\theta}_n$ is asymptotically unbiased, i.e., $E[\hat{\theta}_n] \to \theta$ as $n \to \infty$, AND its variance tends to zero as $n \to \infty$, i.e., $Var(\hat{\theta}_n) \to 0$ as $n \to \infty$, then $\hat{\theta}_n$ is a consistent estimator of θ. This is a direct consequence of Chebyshev's inequality.

o. **Regularity Conditions for MLEs:** For Maximum Likelihood Estimators (MLEs), consistency is often guaranteed under a set of "regularity conditions" on the probability distribution function and the parameter space. These conditions generally include:

- The parameter space is an open interval.
- The support of the distribution does not depend on the parameter.
- The first and second derivatives of the log-likelihood function with respect to the parameter exist.
- The likelihood function is identifiable (different parameter values lead to different distributions).
- The integral/summation and differentiation operations can be interchanged.
- The Fisher information exists and is positive. Under these conditions, the MLE is consistent, asymptotically efficient, and asymptotically normally distributed.