Math Problems

YILONG LIU

May 2025

1 Question 1: An unbiased estimator of the variance of i.i.d random variables

Let $(Y_i)_{1\leq i\leq n}$ be n i.i.d random variables. Let $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ denote the average of these variables. Let Y be a random variable with the same distribution as the Y_is. The goal of the exercise is to show that $\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ is an unbiased estimator of $V(Y)$, the variance of the Y_is.

(1) Show that $\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = \frac{1}{n-1}\sum_{i=1}^{n} Y_i^2 - \frac{n}{n-1}(\bar{Y})^2$.

(2) Use the result in question 1) to prove that $E\left(\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2\right) = V(Y)$.

1.1 Question (1) Solution

We can expand the left hand side equation and have:

$$\begin{aligned}
\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(Y_i^2 - 2Y_i\bar{Y} + \bar{Y}^2) \\
&= \frac{1}{n-1}(\sum_{i=1}^{n} Y_i^2) - \frac{2}{n-1}\bar{Y}(\sum_{i=1}^{n} Y_i) + \frac{n}{n-1}\bar{Y}^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n} Y_i^2 - \frac{2}{n-1}\bar{Y}n\bar{Y} + \frac{n}{n-1}\bar{Y}^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n} Y_i^2 - \frac{n}{n-1}\bar{Y}^2
\end{aligned}$$

1.2 Quesiton (2) Solution

According to the result in question (1), we have the following:

$$
\begin{aligned}
E[\frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\bar{Y})^2] &= E[\frac{1}{n-1}\sum_{i=1}^{n}Y_i^2-\frac{n}{n-1}(\bar{Y})^2] \\
&= \frac{1}{n-1}E[\sum_{i=1}^{n}Y_i^2]-\frac{n}{n-1}E[\bar{Y}^2] \\
&= E[\frac{1}{n-1}\sum_{i=1}^{n}Y_i^2]-E[\frac{n}{n-1}\bar{Y}^2] \\
&= E[Y^2]-E^2[Y]
\end{aligned}
$$

Note: As $\frac{1}{n-1}\sum_{i=1}^{n}Y_i^2$ is the unbiased estimator of $E[Y^2]$.

2 Exercise 1, Question 2: A super consistent estimator

Assume you observe an iid sample of n random variables $(Y_i)_{1\leq i\leq n}$ following the uniform distribution on $[0,\theta]$, where θ is an unknown strictly positive real number we would like to estimate. Let Y be a random variable with the same distribution as the Y_is.

(1) Compute $E(Y)$. Write θ as a function of $E(Y)$.

(2) Use question (1) to propose an estimator $\hat{\theta}_{MM}$ for θ using the method of moments (reminder: that method amounts to replacing expectations by sample means).

(3) Show that $\hat{\theta}_{MM}$ is an asymptotically normal estimator of θ, and show that its asymptotic variance is $4V(Y)$.

Consider the following alternative estimator for θ: $\hat{\theta}_{ML}=\max\limits_{1\leq i\leq n}\{Y_i\}$.

(4) Why does using $\hat{\theta}_{ML}$ to estimate θ sounds like a natural idea?

(5) Show that

$$
P(\hat{\theta}_{ML}\leq x)=\begin{cases} 0 & \text{if } x<0 \\ \left(\frac{x}{\theta}\right)^n & \text{if } x\in[0,\theta] \\ 1 & \text{if } x>\theta \end{cases}
$$

2.1 Question (1) Solution

By definition:

$$
E[Y]=\int_0^\theta y\cdot f(y)dy
$$

As it is uniform distribution: $f(y) = \frac{1}{\theta}$ Then we have the follows:

$$\begin{aligned} E[Y] &= \int_0^\theta \frac{y}{\theta} dy \\ &= \frac{1}{2\theta} \cdot y^2|_0^\theta \\ &= \frac{\theta}{2} \end{aligned}$$

Therefore, $E[Y] = \frac{\theta}{2}$, $\theta = 2E[Y]$.

2.2 Question (2) Solution

From (1), $\theta = 2E[Y]$, we propose that the estimator $\hat{\theta}_{MM}$ for θ to be:

$$\hat{\theta}_{MM} = \frac{2}{n} \sum_{i=1}^{n} Y_i$$

based on n random variables $(Y_i)_{1 \leq i \leq n}$.

2.3 Question (3) Solution

From (1), $\theta = 2E[Y]$, we propose that the estimator $\hat{\theta}_{MM}$ for θ to be:

$$\hat{\theta}_{MM} = \frac{2}{n} \sum_{i=1}^{n} Y_i$$

based on n random variables $(Y_i)_{1 \leq i \leq n}$.

By the Central Limit Theorem, we have:

$$\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^{n} Y_i - E[Y] \right) \xrightarrow{d} \mathcal{N}(0, V(Y))$$

which implies that

$$\sqrt{n}(\hat{\theta}_{MM} - \theta) = \sqrt{n} \left(2\bar{Y} - 2E[Y] \right) = 2\sqrt{n}(\bar{Y} - E[Y]) \xrightarrow{d} \mathcal{N}(0, 4V(Y)).$$

Therefore, $\hat{\theta}_{MM}$ is an asymptotically normal estimator of θ, and its asymptotic variance is $4V(Y)$.

2.4 Question (4) Solution

It is because that the random variables $(Y_i)_{1 \leq i \leq n}$ follows a uniform distribution on $[0, \theta]$. Therefore for any i such that $1 \leq i \leq n$, we have $Y_i \in [0, \theta]$, and it is then a natural idea to assume the key parameter $\hat{\theta}_{ML} = max_{1 \leq i \leq n} Y_i$.

2.5 Question (5) Solution

By definition, $P(\hat{\theta}_{ML} \leq x) = P(max_{1 \leq i \leq n}\{Y_i\} \leq x)$, then:

if $x < 0$, then the probility is 0 as $Y_i \in [0, \theta]$ for any i.

if $x \in [0, \theta]$, then:

$$\begin{aligned} P(max_{1 \leq i \leq n}\{Y_i\} \leq x) &= \Pi_{i=1}^{n} P(Y_i \leq x) \\ &= \Pi_{i=1}^{n} F(X) \\ &= (\frac{x - 0}{\theta - 0})^n \\ &= (\frac{x}{\theta})^n \end{aligned}$$

if $x > \theta$, then the probability is 1 as $Y_i \in [0, \theta]$ for any i.