

Introduction to Statistical Method

Estimation

- An **estimator** for a population parameter θ is a statistic and denoted by $\hat{\theta}$.
- Any given value of $\hat{\theta}$ is an **estimate**.

Desirable Properties of a Point Estimator

- The expected value of $\hat{\theta}$ should be θ .
- $\hat{\theta}$ should have **small variance** for **large sample size**.

Bias

- The difference $\theta - E[\hat{\theta}]$ is the bias of an estimator $\hat{\theta}$ for a population parameter θ .
- $E[\hat{\theta}] = \theta$ means $\hat{\theta}$ is unbiased.

Mean Square Error of Estimator

- The **mean square error** of $\hat{\theta}$ is defined as $MSE(\hat{\theta}) := E[(\hat{\theta} - \theta)^2]$.
- The mean square error measures the **overall quality of an estimator**.

$$\begin{aligned} MSE(\hat{\theta}) &= E[(\hat{\theta} - E[\hat{\theta}] + E[\hat{\theta}] - \theta)^2] = E[(\hat{\theta} - E[\hat{\theta}])^2 + (E[\hat{\theta}] - \theta)^2 + 2(\hat{\theta} - E[\hat{\theta}])(E[\hat{\theta}] - \theta)] \\ &= Var(\hat{\theta}) + (bias)^2 + 2E[(\hat{\theta} - E[\hat{\theta}])(E[\hat{\theta}] - \theta)] = Var(\hat{\theta}) + (bias)^2 + 2E[(\hat{\theta} - E[\hat{\theta}])(E[\hat{\theta}] - \theta)] \\ &= Var(\hat{\theta}) + (bias)^2 + 2(E[\hat{\theta}] - E[\hat{\theta}])(E[\hat{\theta}] - \theta) = Var(\hat{\theta}) + (bias)^2 \end{aligned}$$

Sample Mean Variance and Bias

Theorem: Unbiased Sample Mean

Let X_1, \dots, X_n be a random sample of size n from a distribution with mean μ . Then the sample mean \bar{X} is an unbiased estimator for μ .

Since $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$, then we see $\mu - E[\bar{X}] = \mu - \frac{1}{n} n\mu = 0$, thus unbiased.

Theorem: Sample Average Variance

Let \bar{X} be the sample of a random sample of size n from a distribution with mean μ and variance σ^2 . Then $Var \bar{X} = E[(\bar{X} - \mu)^2] = \frac{1}{n} \sigma^2$

Since $Var \bar{X} = Var(\frac{1}{n} \sum_{k=1}^n X_k) = \frac{1}{n^2} Var(\sum_{k=1}^n X_k)$, then check [ve401_note_4](#), we can see a fact that

$Var(X + Y) = Var X + Var Y + Cov(X, Y)$, where $Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$, then we get $Var \bar{X} = \frac{1}{n^2} n Var X = \frac{1}{n} \sigma^2$

The standard deviation of \bar{X} is given by $\sqrt{Var \bar{X}} = \sigma / \sqrt{n}$ and is called standard error of mean.

Theorem: Sample Variance

Sample variance $S^2 = \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X})^2$ is an unbiased estimator for σ^2 .

$$\begin{aligned} E\left[\frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X})^2\right] &= \frac{1}{n-1} E\left[\sum_{k=1}^n (X_k - \mu + \mu - \bar{X})^2\right] \\ &= \frac{1}{n-1} E\left[\sum_{k=1}^n (X_k - \mu)^2 - 2(\bar{X} - \mu)\left(\sum_{k=1}^n X_k - n\mu\right) + n(\mu - \bar{X})^2\right] \\ &= \frac{1}{n-1} E\left[\sum_{k=1}^n (X_k - \mu)^2 - n(\mu - \bar{X})^2\right] \\ &= \frac{1}{n-1} \left(\sum_{k=1}^n E[(X_k - \mu)^2] - nE[(\mu - \bar{X})^2]\right) = \frac{1}{n-1} \left(\sum_{k=1}^n \sigma^2 - n\frac{\sigma^2}{n}\right) = \sigma^2 \end{aligned}$$

Finding Estimator: Method of Moments

- We have an unbiased estimator $M_k = \frac{1}{n} \sum_{i=1}^n X_i^k$ for the k^{th} moments $E[X^k]$ given random samples X_1, \dots, X_n .
- The idea is then that population parameters θ_j can often be expressed in terms of moments of distribution. Replacing the moments in these expressions by their estimators for parameter θ_j .
- Estimators obtained in this way are not necessarily unbiased.

Finding Estimator: Method of Maximum Likelihood

- Assume we have a random sample x_1, \dots, x_n from the distribution of a random variable X with density f and parameter θ .
- Define likelihood function by $L(\theta) = \prod_{i=1}^n f(x_i)$.
- Find the θ by maximizing the $L(\theta)$.
- Replace θ with $\hat{\theta}$.

Distribution of Sample Mean

Theorem: MGF to distribution function

- X, Y are two random variables with moments generating functions m_X, m_Y .
- If $m_X(t) = m_Y(t)$ for all t in a neighborhood of zero, then $f_X(t) = f_Y(t)$ for all x .

Theorem: MGF of Random Variables' Sum

- X_1, X_2 are two random variables with moments generating functions m_{X_1}, m_{X_2} .
- Then if $Y = X_1 + X_2$, $m_Y = m_{X_1} + m_{X_2}$.

Theorem: MGF of Numerical Multiplied Random Variable

- X be a random variable with MGF $m_X = E[e^{Xt}]$.
- $Y = \alpha + \beta X$, then $m_Y(t) = E[e^{(\alpha + \beta X)t}] = e^{\alpha t} m_X(\beta t)$.

Two-Sided Confidence Intervals

Definition

$0 \leq \alpha \leq 1$, a $100(1 - \alpha)\%$ (two sided) confidence interval for a parameter θ is an interval $[L_1, L_2]$ such that $P[L_1 \leq \theta \leq L_2] = 1 - \alpha$.

Remark 1

The definition doesn't determine L_1, L_2 uniquely.

Remark 2

The population parameter θ is not random, but L_1, L_2 are random. Hence $[L_1, L_2]$ is random interval.

Interval Estimation for Mean

- random sample of size n from a normal population
- unknown mean μ and known variance σ^2
- Sample yields point estimate \bar{X} for μ .
- We are interested in finding $L = L(\alpha)$ that we can state with $100(1 - \alpha)\%$ confidence that $\mu = \bar{X} \pm L$.

$$z_{\alpha/2} \text{ defined for } \alpha \in [0, 1] \text{ with } \alpha/2 = P[Z \geq z_{\alpha/2}] = \frac{1}{\sqrt{2\pi}} \int_{z_{\alpha/2}}^{\infty} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-z_{\alpha/2}} e^{-x^2/2} dx$$

Thus $1 - \alpha = P[\bar{X} - L \leq \mu \leq \bar{X} + L] = P\left[\frac{\bar{X} - \mu - L}{\sigma/\sqrt{n}} \leq 0 \leq \frac{\bar{X} - \mu + L}{\sigma/\sqrt{n}}\right]$ (the later P is standard normal distribution)

$$\text{Then we let } Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

$$\begin{aligned} 1 - \alpha &= P\left[Z - \frac{L}{\sigma/\sqrt{n}} \leq 0 \leq Z + \frac{L}{\sigma/\sqrt{n}}\right] = P\left[-\frac{L}{\sigma/\sqrt{n}} \leq Z \leq \frac{L}{\sigma/\sqrt{n}}\right] \\ &= 2P\left[0 \leq Z \leq \frac{L}{\sigma/\sqrt{n}}\right] = 1 - 2P\left[\frac{L}{\sigma/\sqrt{n}} \leq Z \leq \infty\right] \end{aligned}$$

$$\text{This means } \frac{L}{\sigma/\sqrt{n}} = z_{\alpha/2} \Leftrightarrow L = \frac{\sigma \cdot z_{\alpha/2}}{\sqrt{n}}$$

Theorem

Let X_1, \dots, X_n be a random sample of size n from a normal distribution with mean μ and variance σ^2 .

A $100(1 - \alpha)\%$ confidence interval on μ is given by $\bar{X} \pm \frac{\sigma \cdot z_{\alpha/2}}{\sqrt{n}}$

Central Limit Theorem