Point Estimation and Sampling Distributions (1)

— Dr. Arti Singh, Assistant Professor, USAR, GGSIPU, EDC

— Reference: (Text Book)

Douglas and George, Applied statistics and probability for Engineers

Statistics

→ Descriptive statistics

→ Inferential statistics

Descriptive statistics

- Ungrouped data.
- Frequency distribution table
- Measures of Central tendency
  - Mean, Median, Mode
- Measures of dispersion
  - Variance, Mean deviation

Inferential statistics

- Sampling theory
- Parameter Estimation.
- Confidence Interval
- Hypothesis Testing
- Correlation & Regression

etc.

Population

Complete data set under study.

$X \rightarrow$ Random variable corresponding to population.

Generally, X is unknown. That is, its distribution and parameters are unknown.

Random sample (sample)

Small portion of population upon which statistical tests are conducted.

$X_1, X_2, \dots, X_n$ is random sample from population X.

sample size $= n$.

Properties of Random sample $-(X_1, X_2, \dots, X_n)$

- $X_1, X_2, \dots, X_n$ are i.i.d.s (independent identically distributed r.v.s)

$\downarrow$

Means, these are independent random variables.
And these have same distributions as that of population X.

For eg: If $X \sim N(\mu, \sigma^2)$, then each $X_i \sim N(\mu, \sigma^2)$ for $i = 1, 2, \dots, n$

- Statistical Inference (or Inferencial statistics)
statistical inference is concerned with making decisions about a population based on the information contained in a random sample from that population.

(~~Douglas & George~~

- Statistic: A statistic is any function of the ~~As observati~~ random sample, not ~~const~~ having unknown parameters.
- Sampling distribution: The probability distribution of a statistics is called a sampling distribution.

- Examples of statistics -

(a) sample mean $= \bar{X} = \dfrac{X_1+X_2+\dots+X_n}{n}$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i$$

(b) Sample variance $= S^2$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2 \quad \left[\text{or} = \frac{1}{n-1}\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right]\right]$$

Sample standard deviation $= S$

- Population mean $= \mu$

Population variance $= \sigma^2$

(Mean & variance of X. Generally unknown.)

Sample mean $= \bar{X}$

Sample variance $= S^2$

These are random variables.

Mean and Variance of Sample Mean ($\bar{X}$):

Note that $\bar{X}$ is a random variable. So we can talk about its mean & variance.

Also note that each X_i has mean μ & variance σ^2 (same that of population r.v. X).

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

$$= \frac{1}{n}X_1 + \frac{1}{n}X_2 + \dots + \frac{1}{n}X_n$$

$$E(\bar{X}) = \frac{1}{n}E(X_1) + \frac{1}{n}E(X_2) + \dots + \frac{1}{n}E(X_n)$$

$$= \frac{1}{n}\mu + \frac{1}{n}\mu + \dots + \frac{1}{n}\mu = \frac{n\mu}{n} = \mu$$

$$Var(\bar{X}) = \left(\frac{1}{n}\right)^2 Var(X_1) + \left(\frac{1}{n}\right)^2 Var(X_2) + \dots + \left(\frac{1}{n}\right)^2 Var(X_n)$$

(As X_i's are independent so no covariance terms.)

$$= \frac{1}{n^2}\sigma^2 + \frac{1}{n^2}\sigma^2 + \dots + \frac{1}{n^2}\sigma^2$$

$$= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

So $\bar{X}$ has mean μ and variance $\frac{\sigma^2}{n}$.

Mean and variance of sample variance (s^2)

Note that s^2 is a random variable. so we can talk about its mean and variance.

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}(x_i^2+\bar{x}^2-2x_i\bar{x})$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2+\sum_{i=1}^{n}\bar{x}^2-2\bar{x}\sum_{i=1}^{n}x_i\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2+n\bar{x}^2-2\bar{x}\,n\bar{x}\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2-n\bar{x}^2\right]$$

Taking Mean

$$E(s^2) = \frac{1}{n-1}\left[\sum_{i=1}^{n}E(x_i^2)-nE(\bar{x}^2)\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}(\sigma^2+\mu^2)-n\left(\frac{\sigma^2}{n}+\mu^2\right)\right]$$

$$= \frac{1}{n-1}\left[n(\sigma^2+\mu^2)-\sigma^2-n\mu^2\right]$$

$$= \frac{1}{n-1}\left[(n-1)\sigma^2\right] = \sigma^2$$

- Mean of sample variance is population variance.

- Taking variance

$$Var(s^2) = E(s^4)-(E(s^2))^2$$

after some calculations

$$= \frac{\mu_4}{n}-\frac{\sigma^4(n-3)}{n(n-1)} \quad \text{where } \mu_4 = E(X-\mu)^4$$

→ Proof not Much Important.

- If $X_1, X_2, \ldots, X_n$ are independent normal r.v.s.

Let population $X \sim N(\mu, \sigma^2)$.

Then random sample $X_1, X_2, \ldots, X_n \equiv X_i \sim N(\mu, \sigma^2)$

- Sampling distribution of $\bar{X}$:

$$\bar{X} = \frac{1}{n}X_1 + \frac{1}{n}X_2 + \ldots + \frac{1}{n}X_n \quad (\text{linear combination})$$

Linear combination of normal is again normally distributed.

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

Also $$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim Z \quad (N(0,1) \text{ distribution}).$$

- Here question arises that what is the distribution of $\bar{X}$ if X_i's are not normal. That is, if underlying population X is not normal.

Answer to this question lies in the next discussed central limit theorem.

Central Limit theorem: (CLT)

If $x_1, x_2, \ldots, x_n$ is a random sample of size n taken from a population (either finite or infinite) with mean μ & variance σ^2. Then

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \Rightarrow Z \text{ as } n \to \infty.$$

- CLT is the underlying reason why many of the r.v.s encountered in engineering and science are normally distributed.
- In general, if $n > 30$, the ~~CLT~~ CLT can be applied.
- For two independent populations with means μ_1 and μ_2, variances σ_1^2 and σ_2^2, and sample means $\bar{X}_1$ & $\bar{X}_2$ with sample sizes n_1 & n_2 we get

$$\frac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} \approx Z \quad \left(\text{Approximately equal to } Z\right).$$

General concepts of Point estimation

let $\theta \rightarrow$ unknown parameter. (A constant)

$\hat{\theta} \rightarrow$ An estimator of θ (A random Variable)

Unbiased estimator: An estimator $\hat{\theta}$ is said to be an unbiased estimator of θ if

$$E(\hat{\theta}) = \theta.$$

MVUE: If we consider all unbiased estimators of θ, the one with the smallest variance is called the minimum variance unbiased estimator (MVUE).

- **Some Notations:** (Population Moments)

1st order population moment about origin $= E(X)$

2nd order population moment about origin $= E(X^2)$

In general,
k^{th} order population moment about origin $= E(X^k)$

- Note that 1st order population moment about origin is nothing but mean.

- 2nd order population moment about mean (μ)
$= E(X-\mu)^2 =$ Variance.

Sample moments:-

M_1 = 1st order sample moment about origin $= \frac{1}{n}\sum_{i=1}^{n} x_i$

M_2 = 2nd order sample moment about origin $= \frac{1}{n}\sum_{i=1}^{n} x_i^2$

$\vdots$

M_k = kth order sample moment about origin $= \frac{1}{n}\sum_{i=1}^{n} x_i^k$

Method of ~~moments~~ parameter estimation (10)

Two methods

- ↳ Method of moments.
- ↳ Method of Maximum Likeli-hood estimator. (MLE)

Method of moments: Consider population with m unknown parameters, $\theta_1, \theta_2, \ldots, \theta_m$. We equate first m population moment to corresponding sample moment.

$$E(X) = \frac{1}{n} \Sigma X_i = M_1$$

$$E(X^2) = \frac{1}{n} \Sigma X_i^2 = M_2$$

$$\vdots$$

$$E(X^m) = \frac{1}{n} \Sigma X_i^m = M_m$$

We solve above equations to find esti-mators $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m$.

- For example, For exponential (λ).

  one parameter, $m=1$. $\theta_1 = \lambda$.

  For Binomial (n, p)

  Two parameters, $m=2$, $\theta_1 = n$, $\theta_2 = p$.

- Exponential (λ)

Let λ be unknown. Using method of moments we find an estimator of λ.

One equation

Population Moment = Sample moment.

$$E(X) = \frac{1}{n} \Sigma X_i$$

$$\Rightarrow \frac{1}{\lambda} = \bar{X}$$

$$\text{So} \quad \lambda = \frac{1}{\bar{X}}$$

Estimator of λ = $\boxed{\hat{\lambda} = \frac{1}{\bar{X}}}$

Eg (Douglas & George - Chapter 7 - eg 7.7)

Suppose that the time to failure of an electronic module used in an automobile engine controller is tested at an elevated temperature to accelerate the failure mechanism.

Eight units are randomly selected and tested, with following failure time (in hours):

$x_1 = 11.96$, $x_2 = 5.03$, $x_3 = 67.40$, $x_4 = 16.07$, ..., $x_8 = 22.38$

Suppose time to failure is exponentially distributed with parameter λ. Estimate λ.

$$\hat{\lambda} = \frac{1}{\bar{x}} = \frac{1}{\frac{1}{n}\Sigma x_i} = \frac{1}{\frac{1}{8}(11.96 + 5.03 + \ldots + 22.38)} = 0.0462$$

- Binomial (n, p) (Method of moments)

$m = 2$ (Two unknown parameters). (Sample size $= N$)

Two equations

$$E(X) = \frac{1}{N}\sum x_i$$

$$E(X^2) = \frac{1}{N}\sum x_i^2$$

$\Rightarrow$

$$np = M_1 \quad \text{--- (1)}$$

$$np(1-p) + (np)^2 = M_2 \quad \text{--- (2)}$$

$$\left(\begin{aligned} &Var(X) = E(X^2) - (E(X))^2 \\ &\Rightarrow E(X^2) = Var(X) + (E(X))^2 \\ &\qquad = np(1-p) + n^2p^2. \end{aligned} \right.$$

From equation (1), put in (2).

$$M_1(1-p) + M_1^2 = M_2$$

Solving for p.

$$p = \frac{M_1^2 - M_2}{M_1} + 1$$

So $\boxed{\hat{p} = M_1 - \frac{M_2}{M_1} + 1}$

Using equation (1)

$$\hat{n} = \frac{M_1}{\hat{p}}$$

$$\Rightarrow \boxed{\hat{n} = \frac{M_1^2}{M_1^2 - M_2 + M_1}}$$

where

$$M_1 = \frac{1}{N}\sum x_i$$

$$M_2 = \frac{1}{N}\sum x_i^2$$

- For Normal & Gamma refer book Douglas & George. Chapter 7 (Eg 7.8 and Eg 7.9)

Method of Maximum Likelihood Estimator (MLE)

- Consider population X with PDF/PMF as $f(x, \theta)$. (Here θ is unknown parameter.).
[It could be $f(x, \theta_1, \theta_2)$ → If there are two unknown parameters.]

- $X_1, X_2, \ldots, X_n$ Random sample.

let $x_1, x_2, \ldots, x_n$ → Corresponding observed values of random sample.

- Consider

<u>Likelihood function.</u>

$$L(\theta) = f(x_1, \theta) f(x_2, \theta) \ldots f(x_n, \theta)$$

or

$$L(\theta_1, \theta_2) = f(x_1, \theta_1, \theta_2) f(x_2, \theta_1, \theta_2) \ldots f(x_n, \theta_1, \theta_2)$$

- The MLE of θ is that value of θ which <u>maximizes</u> function $L(\theta)$. (Maximizing $L(\theta)$ ⟺ Maximizing $\ln L(\theta)$).

- So from calculus to find point of maxima or minima we need.

In case of θ_1, θ_2.

$$\boxed{\frac{d \ln L(\theta)}{d\theta} = 0} \quad \text{OR} \quad \boxed{\begin{aligned} \frac{\partial \ln L(\theta)}{\partial \theta_1} &= 0 \\ \frac{\partial \ln L(\theta)}{\partial \theta_2} &= 0. \end{aligned}}$$

↳ Solve for θ.

$\ln L(\theta)$ ↓ $\log_e L(\theta)$

→ Solve for θ_1 & θ_2

Example Consider population distributed as $N(\mu, \sigma^2)$ where μ & σ^2 are unknown. (14)

Obtain MLEs (Maximum Likelihood Estimators) of μ & σ^2.

$$f(x, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{(PDF)}.$$

Likelihood function

$$L(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x_1-\mu)^2} \times \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x_2-\mu)^2} \times \cdots \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x_n-\mu)^2}$$

$$= \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

Taking logarithm

$$\ln L = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\frac{\partial \ln L}{\partial \mu} = 0 \Rightarrow \boxed{+\frac{2}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = 0} \quad \text{— (1)}$$

$$\frac{\partial \ln L}{\partial \sigma^2} = 0 \Rightarrow \boxed{-\frac{n}{2} \times \frac{2\pi}{2\pi\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0} \quad \text{— (2)}$$

Considering eqn ①.

$\Rightarrow \sum_{i=1}^{n}(x_i - \mu) = 0 \Rightarrow \sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\mu = 0$

$\Rightarrow \sum x_i - n\mu = 0$

$\Rightarrow \hat{\mu} = \frac{1}{n}\sum x_i$

$\Rightarrow \boxed{\hat{\mu} = \bar{x}}$

Putting $\hat{\mu} = \bar{x}$ in eqn ② and simplifying :-

$-\cancel{\frac{n}{2}} - n + \frac{1}{\cancel{2}\sigma^2}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 0$

$\Rightarrow \boxed{\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$

• Note that • $\hat{\mu}$ is an unbiased estimator of μ.

$\boxed{\begin{aligned} E(\hat{\mu}) &= E(\bar{x}) \\ &= \mu \end{aligned}}$

• $\hat{\sigma}^2$ is not an unbiased estimator of σ^2.

$E(\hat{\sigma}^2)$

$= E\left(\frac{1}{n}\sum(x_i - \bar{x})^2\right)$

$= \frac{n-1}{n}E(S^2)$

$= \frac{n-1}{n}\sigma^2$

So $E(\hat{\sigma}^2) \neq \sigma^2$

Biased estimator.

- See ~~also~~ Douglas & George

  Chapter 7, examples 7.10
  example 7.11
  Exam 7.12

  For MLEs of other distributions.