4. CLT, MLE

29 August 2024 08:04

Sampling Distributions

Central Limit Theorem

If we obtain a random sample and calculate a sample statistic (mean, SD, proportion) from that sample, then the sample statistic will be a random variable (since every sample is different).

The probability distribution of that sample statistic RV is called the sampling distribution of that statistic.

$$\boxed{\begin{array}{l} n < 30 \longrightarrow \text{small sample} \\ n \geq 30 \longrightarrow \text{large sample} \end{array}}$$

$$\boxed{\text{Ways to select } n\text{-sized sample} = {}^{N}C_{n}}$$

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

$X_1, X_2 \cdots X_n$ are random variables

$\bar{X} \rightarrow$ random variable denoting sample mean

Distribution of $\bar{X} \longrightarrow$ Sampling distribution of mean

Say $Y = aX$
↳ also random variable

$$\boxed{\begin{array}{l} E(Y) = a\,E(X) \\ V(Y) = a^2\,V(X) \end{array}}$$

$Z = aX + bY$

$E(Z) = a\,E(X) + b\,E(Y)$

$V(Y) = a^2\,V(X) + b^2\,V(Y)$

Now, say population has mean μ and variance σ^2.

$$E(\bar{X}) = E\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{\mu + \mu + \cdots + \mu}{n} = \frac{n\mu}{n} = \mu$$

$$V(\bar{X}) = V\left(\quad '' \quad \right) = \frac{1}{n^2}\left(V(X_1) + V(X_2) \cdots\right) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$$SD(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

Now, $\bar{X} \longrightarrow$ sample mean

$S_n = X_1 + X_2 + \cdots + X_n \longrightarrow$ sum of sample observations

If n is sufficiently large,

$$\boxed{\begin{aligned} \bar{X} &\sim \text{Normal}\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \\ S_n &\sim \text{Normal}\left(n\mu, n\sqrt{\sigma}\right) \end{aligned}}$$

Example 1

$\mu_{age} = 22.3$ years

$\sigma = 4$ years

$n = 64$

$P(\bar{X}_{age} > 23 \text{ years}) = ?$

$E(\bar{X}_{age}) = \mu_{age} = 22.3$

$SD(\bar{X}_{age}) = \dfrac{\sigma}{\sqrt{n}} = 0.5$

$$\begin{aligned} P(\bar{X}_{age} > 23 \text{ years}) &= 1 - P(\bar{X} < 23) \\ &= 1 - P\left(Z < \frac{23 - 22.3}{0.5}\right) \\ &= 1 - P(Z < 1.4) \\ &= 1 - 0.9192 \\ &= \underline{\underline{0.0808}} \end{aligned}$$

$P \rightarrow$ random variable of proportion $= \dfrac{X}{n}$

$E(P) = \dfrac{E(\bar{X})}{n} = \dfrac{\mu}{n}$

$V(P) = V\left(\dfrac{X}{n}\right) = \dfrac{1}{n^2} V(X) = \dfrac{\sigma^2}{n^2}$

$SD(p) = \frac{\sigma}{n}$

Example 2

$5\% \longrightarrow$ underfilled

$n = 200$ (sample)

$P(\hat{p} > 0.10) = ?$

$\hat{p} \longrightarrow$ proportion RV $= \frac{x}{n}$

$E(\hat{p}) = E\left(\frac{x}{n}\right) = \frac{1}{n} E(x) = \frac{\cancel{n}p}{\cancel{n}} = p$

Example 3

x	$P(x=x)$
0	0.48
1	0.39
2	0.12
3	0.01

$P(\bar{X} < 0.5) = ?$

$\mu = 1(0.39) + 2(0.12) + 3(0.01)$
$= 0.66$

$\sigma^2 = E(x^2) - [E(x)]^2$
$= 0.96 - 0.4356$
$= 0.5244$
$\sigma = 0.7241$

$E(\bar{x}) = \mu = 0.66$

$SD(\bar{x}) = \frac{\sigma}{\sqrt{n}} = 0.07241$

$$P(\bar{X} < 0.5) = P\left(Z < \frac{0.5 - 0.66}{0.07241}\right) = P(Z < -2.21) = 0.0136$$

MAXIMUM LIKELIHOOD ESTIMATION (MLE)

One way of estimating the parameters of the population through samples

$\bar{X}, S_n, \hat{p} \longrightarrow$ estimators

any value assumed by them $\longrightarrow$ estimate

Use of MLEs

If we somehow know the distribution (pdf/pmf) and the population parameters, we can form the likelihood function.

$$\boxed{\underbrace{L(\theta)}_{\text{joint pdf}} = f(x_1)f(x_2)\dots f(x_n)}$$

The parameter values which maximizes $L(\theta)$ is called the Maximum likelihood Estimate.

Step ①: Form likelihood function $L(\theta)$

Step ②: Differentiate L wrt θ & set it to 0 to find MLE.

Take $\log(L)$ to help differentiation

Example: Find MLE for λ in a population following the Poisson distribution

$$f(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$L(\lambda) = f(x_1)f(x_2)\dots f(x_n)$$

$$L(\lambda) = \frac{e^{-n\lambda}\lambda^{x_1+x_2+\dots+x_n}}{x_1!\,x_2!\dots x_n!}$$

$$\log L(\lambda) = \log e^{-n\lambda} + \log \lambda^{x_1+x_2+\dots+x_n} - \log(x_1!\,x_2!\dots x_n!)$$

$$\frac{d}{d\lambda}(\log L) = -n + (x_1+x_2+\dots+x_n)\frac{1}{\lambda}$$

Setting $\frac{d}{d\lambda}(\log L) = 0$,

$$-n + \frac{(x_1 + x_2 + \dots + x_n)}{\lambda} = 0$$

$$\Rightarrow \lambda = \frac{x_1 + x_2 + \dots + x_n}{n}$$

$$MLE = \lambda = \frac{\sum x_i}{n}$$

Example ②: Find the MLE for μ and σ of Normal(μ, σ)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

If $X \sim$ Binomial (n, p) with n known, find MLE for p

$$L(p) = \left[{}^nC_{x_1} p^{x_1} (1-p)^{n-x_1}\right]\left[{}^nC_{x_2} p^{x_2} (1-p)^{n-x_2}\right] \dots \left[{}^nC_{x_n} p^{x_n} (1-p)^{n-x_n}\right]$$

PROPERTIES OF A GOOD ESTIMATOR $(\hat{\theta})$

① Unbiasedness:

$$E(\hat{\theta}) - \theta = 0$$

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

NOTE:
$E(\hat{\theta}) - \theta \longrightarrow$ bias / amount of bias

② Consistency:

Even with increase in n, $\hat{\theta}$ should be close to true parameter θ

③ Efficiency:

③ Efficiency:

$Var(\hat{\theta})$ should be as small as possible

Note: Mean Squared Error (MSE) of estimates

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + (Bias)^2$$

$$MSE(\hat{\theta}) = E\left[(\hat{\theta}-\theta)^2\right]$$

Example ①: If $X \sim Binomial(n,p)$ with p unknown, find MSE of $\hat{p} = \frac{X}{n}$.

Example ②: If $X \sim Poisson(\lambda)$, find MSE of $\hat{\mu} = \bar{X}$

In all our estimators, bias = 0

$$E(\bar{X}-\mu) = E(\bar{X}) - \mu$$
$$= 0$$