Central Limit Theorem and Normal Approximation

Tsering Sherpa

2025

Central Limit Theorem (CLT)

The Central Limit Theorem states that the sampling distribution of the sample mean $\bar{X}$ approaches a normal distribution as the sample size n increases, regardless of the shape of the population distribution.

Conditions for CLT to apply:

- If $n \geq 30$: Sampling distribution of $\bar{X}$ is approximately normal, even if the population distribution is not.
- If $n < 30$: The population distribution should be approximately normal, and the sample should not have clear outliers.

CLT for Population Mean

If X is a random variable with mean μ and standard deviation σ, and n is the sample size:

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

Normal Approximation to the Binomial

- The binomial distribution $B(n, p)$ can be approximated by a normal distribution when n is large and p is not too close to 0 or 1.
- Rule of thumb: Use the normal approximation if

$$np \geq 10 \quad \text{and} \quad n(1-p) \geq 10$$

- The approximate normal distribution is:

$$X \sim N(np, \sqrt{np(1-p)})$$

- A **continuity correction** of ± 0.5 is applied when approximating discrete binomial probabilities with a continuous normal distribution.

Example:

If $X \sim B(n, p)$, then to approximate $P(X \leq k)$ using the normal distribution:

$$P(X \leq k) \approx P\left(Z \leq \frac{k + 0.5 - np}{\sqrt{np(1-p)}}\right)$$

Using R for Normal Probabilities

- To find the z-score corresponding to a cumulative probability:

```
qnorm(p)
```

- To compute the probability to the left of a z-value:

```
pnorm(z)
```