01:XXX:XXX - Homework n

Pranav Tikkawar

November 5, 2024

1 Chapter 11: Confidence Intervals

Given a α such that $0 < \alpha < 1$, an Interval Estimation stratgey provides two statistics (r.v) L and R s.t $P(L < \theta < R) = 1 - \alpha$. The interval $[L, R]$ is called a $(1 - \alpha)\%$ confidence interval for θ.
We can say that if you repeat the expirment N times and gte N intervals, then $(1 - \alpha)\%$ of the intervals will contain the true value of θ.

Definition (Confidence Interval)**.** A confidence interval with (1-α) confidence level are two statistics L and R such that $P(L < \theta < R) = 1 - \alpha$.

Remark. If someone says after an expirment that $2 < \lambda < 2.1$ with 90% confidence, it means that on average that 90% of the intervals will contain the true value of λ.

We want CI to be symetric about $\bar{X}$

CI so far: $N(\mu, \sigma^2)$ population

1. σ^2 known: $\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$
2. σ^2 unknown: $\bar{X} \pm t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}$

We can think of $z_{\alpha/2}$ in the normal curve as the shaded area in the tails.

New Context: Two pops $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ and sample from both n_1 and n_2 from each

$$X_{11} \dots X_{1n_1} \sim N(\mu_1, \sigma_1^2)$$
$$X_{21} \dots X_{2n_2} \sim N(\mu_2, \sigma_2^2)$$

These are independent but not necessarily identically distributed.
Let $\bar{X}_1$ and $\bar{X}_2$ be the sample means and S_1^2 and S_2^2 be the sample variances.
Want CI for $\mu_1 - \mu_2$
Case 1: σ_1^2 and σ_2^2 are known.
We can use point estimators:

$$\bar{X}_1 - \bar{X}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

$$\bar{X}_1 - \bar{X}_2 \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Case 2: σ_1^2 and σ_2^2 are unknown. but $\sigma_1^2 = \sigma_2^2$
we can define a pooled sample variance:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

Remark. This is a weighted average of the sample variances. with weights $n_1 - 1$ and $n_2 - 1$

Remark.

$$S_p^2 := \frac{\sum^{n_1}(X_{1i} - \bar{X}_1)^2 + \sum^{n_2}(X_{2i} + \bar{X}_2)^2}{n_1 + n_2 - 2}$$
$$\sim \chi^2_{n_1+n_2-2}$$