Chapter - 5

6th March'2024

Section 5.1 : Inverting tests to Derive Confidence Intervals.

Defn: Let, θ_0 be the true parameter value of θ.

$[A, B]$ is $100(1-\alpha)\%$ Confidence Interval (C.I)

if, $\mathbb{P}[A \le \theta_0 \le B \mid \theta = \theta_0] = \underbrace{1-\alpha}$

"Converge probability"

i.e. $1-\alpha$ is the proportion of C.I's, estimated from repeated random samples, that include the true parameter value θ_0.

You can invert a Hypothesis test (like, a LRT, for example) to obtain/derive a Confidence interval.

Now, Let's look at the Example 5.1.1 from the Complete Lecture Notes. $\longrightarrow$

(C) $n = 50$,

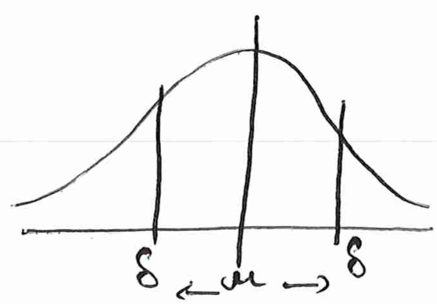
H_0 : No effect on profits.

X_i : Business i's change in profits, $i = 1, 2, \cdots, 50$

$\bar{X} = -1000$

$X_i \overset{iid}{\sim} N(\mu, \delta^2 = 600^2)$

$\rightarrow H_0 : \mu = 0$



$\boxed{\text{Basic Model}}$:-

$$L(\mu) = \prod_{i=1}^{n} \exp\left[-\frac{1}{2\delta^2}(x_i - \mu)^2\right]$$

★ See Ex 4.2.1 for Derivation

$$l(\mu) = -\frac{1}{2\delta^2} \sum_{i=1}^{n} (x_i - \mu)^2$$

Since, δ is known, μ is only unknown

$\therefore \boxed{k=1}$; $\hat{\mu} = \bar{x}$

Hypothesized Model :-

$$\delta = 600 \, ; \, \mu_0 = 0$$

So, $\boxed{q = 0}$

Perform LRT :-

$$D = 2\left[\ell(\hat{\mu}) - \ell(\mu_0)\right]$$

$$\Rightarrow D = 2\left[-\frac{1}{2\delta^2}\sum_{i=1}^{n}(x_i-\bar{x})^2 - \left(-\frac{1}{2\delta^2}\sum_{i=1}^{n}(x_i-\mu_0)^2\right)\right]$$

$$\vdots$$

$$\Rightarrow D = n(\bar{x}-\mu_0)^2/\delta^2$$

$$d_{obs} = \frac{50(-1000-0)^2}{600^2} = \cancel{83} \; 138.889$$

$$p\text{-value} = \mathbb{P}[D \geq d_{obs}] = \mathbb{P}[\chi^2_{(1)} \geq \cancel{83} \; 138.889] < \cancel{0.0004} \; 0.01 .$$

We have very strong evidence against H_0, with p-value < 0.01

The Data are not consistent with the Hypothesis that no effects on profits.

Question: What range of μ_0, would give us a p-value ≥ 0.05 ?

What range of μ_0 would be "reasonably consistent" with the data.

$\hookrightarrow$ these would be "~~re~~ reasonably consistent" with the data.

$$p\text{-value} = \mathbb{P}(\chi^2_{(1)} \geq d_{obs}) = 0.05$$

$$d_{obs} = 3.843 \text{, from tables.}$$

$$p\text{-value} \geq 0.05 \iff d_{obs} \leq 3.843$$

$$\frac{n(\bar{x} - \mu_0)^2}{s^2} \leq 3.843$$

$$\Rightarrow (\bar{x} - \mu_0)^2 \leq 3.843\, s^2/n$$

$$\Rightarrow 1.96\, s/\sqrt{n} \leq \bar{x} - \mu_0 \leq -1.96\, s/\sqrt{n}.$$

$$\Rightarrow \bar{x} - 1.96\, s/\sqrt{n} \leq \mu_0 \leq \bar{x} + 1.96\, s/\sqrt{n}$$

or, $\boxed{\bar{x} \pm 1.96\, s/\sqrt{n}}$ is our 95 % C.I for μ_0.

$95\% = 100(1-\alpha)\%$, $\alpha = 0.05$

Putting the values, we get,

$\{[-1166, -834]\}$ Any μ value in this interval is what we'd consider a "reasonable" estimate of μ_0, given the data.

Since, 0 is not in the 95% C.I, it is not a plausible value for μ_0, given the data. Hence, evidence against the null Hypothesis.