Recap :-> Intro to tests of Significance

8th Feb '2024

Section 4.2 :-> Likelihood Ratio Tests (LRTs) for Simple Hypothesis.

Determining test Statistics (and their Distributions under H_0) can be tricky, sometimes. The likelihood Ratio Statistic (LRS) in these Cases, has an initiative appeal.

Simple Hypothesis :-> Hypothesis that specifies numerical values for all model parameters

Exact Equality. (No inequalities here!)

$\boxed{\text{1-parameter Case}}$

LRS for testing $H_0 : \theta = \theta_0$ is

$$D \underset{\text{"defined"}}{\equiv} -2[r(\theta_0)] = 2[\ell(\hat{\theta}) - \ell(\theta_0)]$$

Where, $\hat{\theta}$ is the MLE of θ and $D \geq 0$ b/c $r(\theta) \leq 0$ for all θ.

"Intuition" behind D :

- D is big $\Rightarrow \ell(\theta_0)$ is "far" from $\ell(\hat{\theta}) \Rightarrow \theta_0$ is not very plausible.
- D is Small $\Rightarrow \ell(\theta_0)$ is pretty close to maximized likelihood $\Rightarrow \theta_0$ is more plausible.

Under the assumption that $H_0 : \theta = \theta_0$ is true,

$$D \approx \chi^2_{(1)}$$

"Approx. distributed", implies "Asymptotically Distributed"

Let, d_{obs} be our observed value of D.

$$D \approx \chi^2_{(1)} \Rightarrow \text{P-value} = \mathbb{P}(D \geq d_{obs})$$
$$\Rightarrow \text{P-value} \simeq \mathbb{P}(\chi^2_{(1)} \geq d_{obs})$$

"approx. equal to"

Now, let's look at an Example ;

<u>Example 4.2.1</u> from the Complete Lecture Notes

Here, X = Measurement error of a scale.
$X \sim N(\mu, \sigma = 1.3)$

Need LRS, $\Rightarrow$ Need $L(\mu)$ $\Rightarrow$ joint pdf.

$$f(x_1, x_2, \cdots, x_{10}; \mu) = \prod_{i=1}^{10} \boxed{\frac{1}{\sqrt{2\pi}\,\sigma}} \exp\left[-\frac{1}{2\sigma^2}(x_i - \mu)^2\right], \quad \mu \in \mathbb{R}, \ \sigma > 0$$

σ treated as constant,
So, we don't need this!

$$L(\mu) = \prod_{i=1}^{10} \exp\left[-\frac{1}{2\sigma^2}(x_i - \mu)^2\right]$$
$$= \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{10}(x_i - \mu)^2\right]$$

$$l(\mu) = -\frac{1}{2\sigma^2}\sum_{i=1}^{10}(x_i - \mu)^2 \ ; \quad \mu \in \mathbb{R}$$

$$\Rightarrow \hat{\mu} = \frac{\sum_{i=1}^{10} x_i}{10} = \bar{x}$$

(refer to 30th Jan, note)

$\hat{\mu} = \bar{x}$ $\qquad$ $\mu_0 = 226$

$$D = -2r(\mu_0) = 2[l(\hat{\mu}) - l(\mu_0)]$$
$$= 2\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{10}(x_i - \bar{x})^2 - \frac{-1}{2\sigma^2}\sum_{i=1}^{10}(x_i - \mu_0)^2\right]$$
$$= \frac{1}{\sigma^2}\left[\underline{\sum_{i=1}^{10}(x_i - \mu_0)^2} - \sum_{i=1}^{10}(x_i - \bar{x})^2\right]$$

$$\sum_{i=1}^{10}(x_i-\mu_0)^2 = \sum_{i=1}^{10}(x_i-\bar{x}+\bar{x}-\mu_0)^2$$

$$= \sum_{i=1}^{10}\left[(x_i-\bar{x})+(\bar{x}-\mu_0)\right]^2$$

$$= \underbrace{\sum_{i=1}^{10}(x_i-\bar{x})^2}+\underbrace{\sum_{i=1}^{10}2\left[(x_i-\bar{x})(\bar{x}-\mu_0)\right]}_{\text{Ex: Show this Cross-Product is equal to Zero.}}+\underbrace{\sum_{i=1}^{10}(\bar{x}-\mu_0)^2}$$

$$= \frac{1}{\sigma^2}\left[\sum_{i=1}^{10}(x_i-\mu_0)^2-\sum_{i=1}^{10}(x_i-\bar{x})^2\right]$$

$$= \frac{1}{\sigma^2}\left[\cancel{\sum_{i=1}^{10}(x_i-\bar{x})^2}+\sum_{i=1}^{10}(\bar{x}-\mu_0)^2-\cancel{\sum_{i=1}^{10}(x_i-\bar{x})^2}\right]$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{10}(\bar{x}-\mu_0)^2$$ No Subscripts here!

$$= \frac{10(\bar{x}-\mu_0)^2}{\sigma^2} = \underbrace{(\bar{x}-\mu_0)^2 \big/ \sigma^2/10}$$ Look familiar?

Recall from Stat 260:-

If, $X \sim N(\mu, \sigma^2)$, then $\bar{X} \sim N(\mu, \sigma^2/n)$; n is sample size.

then, $Z = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

So, $Z^2 = \dfrac{(\bar{X}-\mu)^2}{\sigma^2/n} \sim \chi^2_{(1)}$, from Section 3.2.

Hence, in this case, $D \sim \chi^2_{(1)}$. Exact equality!
even though in LRS def$^{\underline{n}}$ above, $D \approx \chi^2_{(1)}$

Then Calculating our p-value is also exact,

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

$$= 2\mathbb{P}[Z \geq \sqrt{d_{obs}}] \Big\}\ \text{2-tailed test.}$$

We use our data to calculate our d_{obs}:

$$d_{obs} = \frac{(227.49 - 226)^2}{1.3^2/10} = 13.14 \quad ; \bar{x} = 227.49$$

$$p\text{-value} = \mathbb{P}[\chi^2_{(1)} \geq 13.14] = 0.00029 < 0.01$$

Checklist for Conclusion :->

- Estimated parameter values & interval estimate (i)
- p-value (ii)
- In words, what is the result of the data. (iii)

Our estimated mean weight is 227.49 g, with a 10% L.I of 226.61 – 228.37 grams. (i) The p-value of our test was $2.9 * 10^{-4}$, showing very strong evidence against H_0. (ii) The data are not consistent with the hypothesis that the mean weight of the item is 226 grams. (iii)

— * —