14th Feb' 2024

Recap :→ LRTs for Simple Hypothesis
↳ 1-parameter Case
↳ Multi-parameter Case

Section 4.3 :- LRTs for Composite Hypothesis.

Composite Hypothesis : Reduces the # of unknown parameters, but not to zero.

i.e., atleast one parameter unspecified.

i.e, any hypothesis that is not simple.

Recall Ex 4.2.1 (Measurement Errors of Scale)

- Assume δ known, $H_0 : \mu = 226$ ← Simple ; $D \sim \chi^2_{(1)}$
- Assume δ unknown, $H_0 : \mu = 226, \delta = 1$ ← Simple ; $D \sim \chi^2_{(2)}$

What if δ is known and we want to test $H_0 : \mu = 226$?

$$\text{LRS} : D = 2\left[\underbrace{l(\hat{\mu}, \hat{\delta})}_{\substack{\text{Basic Model} \\ (\hat{\mu},\hat{\delta}) \text{ joint MLE}}} - \underbrace{l(\mu_0 = 226, \delta_0)}_{\text{"Hypothesized Model"}}\right]$$

Find the MLE of δ_0 which is the MLE of δ under H_0.

LRS for testing Composite Hypothesis H_0 is :

"tilde"

$$D = 2\left[\,l(\hat{\underline{\theta}}) - l(\tilde{\underline{\theta}})\,\right],$$

where, $\hat{\underline{\theta}}$ is the joint MLE of $\underline{\theta}$, and $\tilde{\underline{\theta}}$ is joint MLE of $\underline{\theta}$ under H_0

Under H_0 we have, $D \approx \chi^2_{(k-q)}$

where, K = # of functionally independent unknown parameters in basic model.

q = # " " " " " in hypothesized model (H_0).

In our Composite H_0 for Ex: 4.2.1, $K=2$, $q=1$ so,
$D \sim \chi^2_{(1)}$ still, but d_{obs} will be calculated differently.

L) Let's look at the Example 4.2.3 of the Complete Lecture Notes.

Here, $(X_1, X_2, \cdots, X_7) \sim$ Multinomial $(63, p_1, \cdots, p_7)$

Suppose,

H_0 : p_1 unspecified
$p_2 = \cdots = p_7 = p$, unknown

[H_0 does not specify numerical values for every parameter in the model, since p is unknown ∴ It is a Composite Hypo....]

Under Basic, we found $\hat{p}_i = x_i/n$, for $i=1(1)7$ (use this w/o proof)

$$L(p_1, p_2, \cdots, p_7) = \prod_{i=1}^{7} p_i^{x_i}$$

$$l(p_1, p_2, \cdots, p_7) = \sum_{i=1}^{7} x_i \ln(p_i) \leftarrow \text{Sub in } \hat{p}_i\text{'s for Basic model.}$$

Under H_0; $l(p_1, p) = x_1 \ln(p_1) + \sum_{i=2}^{7} x_i \ln(p_i)$

$\left[\sum_{i=1}^{7} p_i = 1 \overset{H_0}{\Longrightarrow} p_1 + 6p = 1 \text{ or, } p_1 = 1 - 6p \right]$

$$\Rightarrow l(p) = x_1 \ln(1-6p) + \sum_{i=2}^{7} x_i \ln(p).$$

at $\tilde{p}$, $l_H' = 0$

$$\frac{\partial l_H}{\partial p}\Big|_{\tilde{p}} = \frac{-6x_1}{1-6\tilde{p}} + \sum_{i=2}^{7} x_i/\tilde{p} = 0$$

$$\Rightarrow \frac{\sum_{i=2}^{7} x_i}{\tilde{p}} = \frac{6x_1}{1-6\tilde{p}}$$

$$\Rightarrow \sum_{i=2}^{7} x_i = 6n\tilde{p}$$

$$\Rightarrow \tilde{p} = \sum_{i=2}^{7} x_i/6n = \frac{41}{6(63)} \approx 0.1085$$

$$\Rightarrow \tilde{p}_1 = 1-6\tilde{p} \approx 0.349$$

$$d_{obs} = 2\left[\sum_{i=1}^{7} x_i \ln\left(\frac{x_i}{n}\right) - \left(x_1 \ln(\tilde{p}_1) + \sum_{i=2}^{7} x_i \ln(\tilde{p})\right)\right]$$

plug in x_i's, $\tilde{p}_1$, $\hat{p}$, and n, we get.

$$d_{obs} = 6.53$$

If, $D \approx \chi^2_{(k-q)}$, what is k? what is q?

$$\underline{\underline{k=6}},\ q=1$$

$$D \approx \chi^2_{(5)},\quad p\text{-value} \approx \mathbb{P}\big(\chi^2_{(5)} \geq \overset{d_{obs}}{6.53}\big) = 0.258$$

Conclusion: —

Expected frequencies under H_0: $E(\tilde{x}_1) = n\tilde{p}_1 \approx 21.987$
$= 22$

$i = 1, 2, \dots, 7$; $E(\tilde{x}_i) = n\tilde{p} \approx 6.835 \approx 6.8$.

We have no evidence against H_0, with a p-value of 0.258; The data support the hypothesis that fatal heart attacks were more likely to occur on Mondays for this cohort, and equally likely to occur on the other days.

☆ See Section 4.3.2 : Summary of LRT's.

A note about Degrees of freedom (d.o.f) :

↳ D.o.f represents the # of values in the calculation of a statistic that are free to vary. It is the number of components needed before a statistic is fully determind.

Because, we are assessing the "distance" between the hypothesized model from the basic model, we substract the d.o.f of the former from the latter. In a sense, d.o.f measure/reflect the complexity of a model.

When $k-q$ is big ⇒ Big reduction in complexity of model under H_0, i.e, we believe we can explain a phenomenon with a simpler model.

$k-q$ is small ⇒ Complexity isn't reduced by much (not necessarily a bad thing)

Of course, we want to use simple models as statisticians, but you can imagine that over simplifying can lead to a bad model fit; See fractured plastics Ex: 2.2.1, (v.s) when we re-parameterized it with λ in Sec 3.1.

o —— * —— * —— o