Lecture 2

Data Reduction in Statistical Inferences

from n-dimensional to 1 dimensional

$$X_i = \begin{cases} 1 & \text{w.p } \theta \\ 0 & 1-\theta \end{cases}$$

$x = 0, 1 \qquad f_{X_i}(x_i) = \theta^{x_i}(1-\theta)^{1-x_i}$

$X_1, \dots, X_n$ iid Bernoulli(θ)

$T = \sum_{i=1}^{n} X_i \qquad \frac{T}{n} \leftarrow$ Ratio of success

n ← Total number of event

$$\boxed{T = \sum_{i=1}^{n} X_i} \sim Bin(n, \theta)$$

$t = 0, 1, 2, \dots, n$

$$P(T=t) = \binom{n}{t} \theta^t (1-\theta)^{n-t}$$

$$P(X=x \mid T=t) = \frac{P(X=x \cap T=t)}{P(T=t)} = \begin{cases} 0 & \text{if } \sum_{i=1}^{n} x_i \neq t \\ \frac{P(X=x)}{P(T=t)} = \frac{\cancel{\theta^t}\cancel{(1-\theta)^{n-t}}}{\binom{n}{t}\cancel{\theta^t}\cancel{(1-\theta)^{n-t}}} \end{cases}$$

$$= \frac{1}{\binom{n}{t}} \text{ if } \sum_{i=1}^{n} x_i = t$$

Proof 5.4

Suppose $L(X, \theta) = g(T, \theta) h(x)$

$$P(X=x \mid T=t) = \frac{P(X=x \cap T=t)}{P(T=t)} = \begin{cases} 0 & \text{if } T(x) \neq t \\ = \frac{P(X=x)}{P(T=t)} = \frac{g(t,\theta)h(x)}{\sum g(t,\theta)h(\tilde{x})} \end{cases}$$

$\tilde{x} . T(\tilde{x}) = t$

$$= \frac{g(t,\theta)\cancel{h(x)}}{\cancel{g(t,\theta)}\sum h(\tilde{x})}$$

⇒ Assume sufficient

1. $P_\theta(X=x \mid T=t)\, P_\theta(T=t) \longrightarrow = h(x) g(t,\theta)$

Examples

$$f_{X_i}(x_i) : \theta^{x_i}(1-\theta)^{n-x_i}$$

$$L(X,\theta) = \underbrace{\theta^{T}(1-\theta)^{n-T}}_{g(\theta,T)}\, h(X)$$

$= 1$