S&C

Sufficient Statistics

$S = s(X_1 \ldots X_n)$

① **Def of Joint Sufficiency:** $f_{\vec{x}|\vec{s}=s}(\vec{x}) = \dfrac{f_{\vec{x},\vec{s}}(\vec{x},\vec{s};\vec{\theta})}{f_{\vec{s}}(\vec{s};\theta)} = \dfrac{f_{\vec{x}}(\vec{x};\vec{\theta})}{f_{\vec{s}}(\vec{s};\vec{\theta})}$ is free of θ

Thm (Factorisation Criterion): $f(\vec{x};\theta) = g(\vec{s};\theta)\, h(\vec{x})$

Further Properties:

$l(\vec{\theta}) = \log g(s,\theta) + \log h(\vec{x})$

$\uparrow$ $\frac{dl\theta}{d\theta} = 0$ $\frac{\partial}{\partial\theta} g(s,\theta) = 0$ $\hat{\theta} = \hat{\theta}(s)$

① **MLEs & Order Statistics are jointly sufficient**

Thm If $S_1 \ldots S_k$ are joint sufficient for θ & $\hat{\theta}$ is a unique MLE for θ, then $\hat{\theta}$ is a function of S

Thm If $X_1 \ldots X_n$ is a random sample from a continuous distribution with pdf $f(x;\theta)$, then the order statistics: $X_{1:n}, X_{2:n}, \ldots X_{n:n}$ form a jointly sufficient set for θ

② **Improve Theorem (Rao-Blackwell)**

$T^* = E[T \mid \vec{S}] \rightarrow$ ① T is a function of S (not of θ); ② T^* is an unbiased estimator of $\tau(\theta)$; ③ $Var(T^*) \le Var(T)$

Completeness

① **Def:** family $\{f_T(t;\theta) \mid \theta \in \Omega\}$ is called complete if $E[u(T)] = 0$ for all $\theta \in \Omega \Rightarrow u(T) = 0$.

S&C Statistics: S is jointly sufficient for θ; pdf $f_S(s;\theta)$ of S is member of complete family of pdfs $\{f_S(s;\theta) \mid \theta \in \Omega\}$

② **Lehmann-Scheffé** $X_1 \ldots X_n$ have pdf $f(x_1 \ldots x_n;\theta)$ & let S be a vector of jointly C&S statistics for θ.

$\begin{cases} \text{unbiased for } \tau(\theta) \\ \text{a function of } S \Rightarrow T^* = T^*(\vec{S}) \end{cases} \Rightarrow T^*$ is an UMVUE of $\tau(\theta)$

REC

① **Def:** $f(x;\theta) = c(\theta)\, h(x) \exp\left[\sum_{j=1}^{k} q_j(\theta)\, t_j(x)\right] \quad x \in A$

② **Thm.** $S_1 = \sum_{i=1}^{n} t_1(X_i) \ldots S_k = \sum_{i=1}^{n} t_k(X_i)$ is a set of C&S for $\theta_1 \ldots \theta_k$

(range depend on θ cannot belong to REC)

Basu's Theorem:

① **Def: Stochastically independent** $P[a_1 \le X_1 \le b_1, \ldots, a_k \le X_k \le b_k] = \prod_{i=1}^{k} P[a_i \le X_i \le b_i]$

② **Basu's theorem.** $X_1 \ldots X_n$ have joint pdf $f(x_1 \ldots x_n;\theta)$, $S = (S_1 \ldots S_k)$ are C&S for θ.

T is another statistics.

If the distribution of T doesn't depend on θ, T & S are stochastically independent.

① $\dfrac{X_{k:n}}{X_{j:n}} \rightarrow \dfrac{Z_{k:n}}{Z_{j:n}} \rightarrow$ scale parameter free of $\theta \Rightarrow$ By Basu's theorem.

Example 10.2.4

$\vec{\theta} = (\mu, \sigma^2)$

Consider a sample from N(μ, σ^2), μ and σ^2 unknown.

$$\begin{aligned} f(x_1, \ldots, x_n; \mu, \sigma^2) &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu)^2\right] \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\left\{\sum_{i=1}^n x_i^2 - 2\mu\sum_{i=1}^n x_i + n\mu^2\right\}\right] \\ &= g(s_1, s_2; \mu, \sigma^2) h(x_1, \ldots, x_n) \end{aligned}$$

with

$$g(s_1, s_2; \mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\left\{s_2 - 2\mu s_1 + n\mu^2\right\}\right],$$

$$s_1 = \sum_{i=1}^n x_i \quad \text{and} \quad s_2 = \sum_{i=1}^n x_i^2; \qquad h(x_1, \ldots, x_n) = 1.$$

fact crit. $\Longrightarrow$ $S_1 = \sum_i X_i$ and $S_2 = \sum_i X_i^2$ are (jointly) suff. statistics for μ and σ^2.

① can't say S_1 & S_2 1 to 1 corresponds to μ & σ^2

② if μ known, S_1 & S_2 sufficient for σ^2

S_1, S_2 also complete for σ^2

$$\underbrace{\frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2\right)}_{u(S_1, S_2)}$$

③ if σ^2 known, S_1 sufficient for μ, no meaning for S_2