Homework 1

Pengkun Gu

9/16/2021

11.

If $T(X)$ is sufficient and $S(X) = G(T(X))$ where $\mathcal{T} \xrightarrow{G} \mathcal{S}$ is 1-1: The sufficiency of $T(X)$ tells that θ and X are independent providing $T(X)$, *i.e.* both $\theta \longrightarrow T(X) \longrightarrow X$ and $\theta \longrightarrow X \longrightarrow T(X)$ are Markov chains. Since $S = G(T(X))$ and that G is 1-1, there exist an invert $\mathcal{S} \xrightarrow{H} \mathcal{T}$ making $T = H(S)$, then $\theta \longrightarrow S(X) \longrightarrow T(X) \longrightarrow X$ is Markov, so do $\theta \longrightarrow S(X) \longrightarrow X$ and $\theta \longrightarrow X \longrightarrow S(X)$. So S is sufficient.

16.

For an n-dim sample space, the statistic $T_n = \sum_{i=1}^{n} |X_i|$, and the mass function is

$$\begin{aligned} f_\theta(x) &= \frac{1}{(2\theta)^n} exp(-\frac{\sum_{i=1}^{n} |X_i|}{\theta}) \\ &= \frac{1}{(2\theta)^n} exp(-\frac{T_n}{\theta}) \end{aligned}$$

By the factorization theorem, T is sufficient.

23.

~~Suppose σ here is an unknown parameter.~~

The likelihood functions of X-sample is

$$P_m(\mathbf{X}; \mu_1, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{m}{2}}} exp(-\frac{\sum_{i=1}^{m} (X_i - \mu_1)^2}{2\sigma^2})$$

And the likelihood functions of y-sample is

$$P_n(\mathbf{Y}; \mu_2, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} exp(-\frac{\sum_{i=1}^{n} (Y_i - \mu_2)^2}{2\sigma^2})$$

For hypotheesis $H_0 : \mu = \mu_1 = \mu_2$, the likelihood ratio statistic is

$$\begin{aligned} \log \boldsymbol{\lambda}(\mathbf{X}, \mathbf{Y}) &= \log \frac{\sup_{\mu_1, \mu 2, \sigma} [L_{\mathbf{X},\mathbf{Y}}(\mu_1, \mu 2, \sigma)]}{\sup_{\mu, \sigma} [L_{\mathbf{X},\mathbf{Y}}(\mu, \mu, \sigma)]} \\ &= -\frac{\sum_{i=1}^{m} (X_i - \bar{X})^2 + \sum_{i=1}^{n} (Y_i - \bar{Y})^2}{2\widetilde{\sigma}^2} + \frac{\sum_{i=1}^{m} (X_i - \hat{\mu})^2 + \sum_{i=1}^{n} (Y_i - \hat{\mu})^2}{2\widetilde{\sigma}_0^2} \\ &\quad - \frac{m+n}{2} \log \widetilde{\sigma}^2 + \frac{m+n}{2} \log \widetilde{\sigma}_0^2 \end{aligned}$$

where, the MLE take values at

$$\widetilde{\sigma}^2 = \frac{1}{m+n}(\sum_{i=1}^{m}(X_i - \bar{X})^2 + \sum_{i=1}^{n}(Y_i - \bar{Y})^2)$$

and the MLE under H_0

$$\hat{\mu} = \frac{1}{m+n}(\sum_{i=1}^{m} X_i + \sum_{i=1}^{n} Y_i)$$

and

$$\widetilde{\sigma}_0^2 = \frac{1}{m+n}(\sum_{i=1}^{m}(X_i - \hat{\mu})^2 + \sum_{i=1}^{n}(Y_i - \hat{\mu})^2)$$

We get

$$\begin{aligned}
\log \boldsymbol{\lambda}(\mathbf{X}, \mathbf{Y}) &= \frac{m+n}{2} \log \frac{\widetilde{\sigma}_0^2}{\widetilde{\sigma}^2} \\
&= \frac{m+n}{2} \log \frac{\sum_{i=1}^{m}(X_i - \bar{X})^2 + \sum_{i=1}^{n}(Y_i - \bar{Y})^2}{\sum_{i=1}^{m}(X_i - \hat{\mu})^2 + \sum_{i=1}^{n}(Y_i - \hat{\mu})^2} \\
&= \frac{m+n}{2} \log \frac{(m-1)s_X^2 + (n-1)s_Y^2}{(m-1)s_X^2 + (n-1)s_Y^2 + m(\bar{X} - \hat{\mu})^2 + n(\bar{Y} - \hat{\mu})^2}
\end{aligned}$$

We choose the statistic T where

$$\begin{aligned}
T(\mathbf{X}, \mathbf{Y}) &= \sqrt{\frac{m(\bar{X} - \hat{\mu})^2 + n(\bar{Y} - \hat{\mu})^2}{(m-1)s_X^2 + (n-1)s_Y^2}} \\
&= \sqrt{\frac{mn}{m+n}}(\frac{\bar{Y} - \bar{X}}{s})
\end{aligned}$$

and where

$$s^2 = \frac{1}{m+n-2}(\sum_{i=1}^{m}(X_i - \bar{X})^2 + \sum_{i=1}^{n}(Y_i - \bar{Y})^2)$$

Here T has the t-distribution $\mathcal{T}_{m+n-2}$, and for level $(1-\alpha)$, the criteria for rejecting H_0 is $T \geq t_{n-2}(1-\alpha)$.