

557: MATHEMATICAL STATISTICS II

P-VALUE CALCULATIONS

Example 1 Suppose that $X_1, \dots, X_n \sim N(\theta, 1)$. To test

$$\begin{aligned} H_0 &: \theta = 0 \\ H_1 &: \theta = 1 \end{aligned}$$

the most powerful test at level α is based on the critical region of the form $\bar{X} > c_n$. Now suppose that $n = 8$ and $\bar{x} = 1.240$. To compute the p -value, $p(\underline{x})$, we evaluate

$$p(\underline{x}) = \sup_{\theta \in \Theta_0} \Pr[T(\underline{X}) \geq T(\underline{x}) \mid \theta = 0] = \Pr[\bar{X} \geq \bar{x} \mid \theta = 0] = 1 - \Phi(\sqrt{n}\bar{x}).$$

In this case,

$$p(\underline{x}) = 1 - \Phi(\sqrt{8} \cdot 1.240) = 1 - \Phi(3.507) = 2.264 \times 10^{-4}$$

Example 2 Suppose that $X_1, \dots, X_{n_1} \sim N(\theta_1, \sigma^2)$ and $Y_1, \dots, Y_{n_2} \sim N(\theta_2, \sigma^2)$ are independent random samples. To test

$$\begin{aligned} H_0 &: \theta_1 = \theta_2 = \theta, \sigma^2 \text{ unspecified} \\ H_1 &: \theta_1 \neq \theta_2, \sigma^2 \text{ unspecified} \end{aligned}$$

using a likelihood ratio test, we use the statistic

$$T(\underline{X}, \underline{Y}) = \frac{\bar{X} - \bar{Y}}{\sqrt{s_P^2 (1/n_1 + 1/n_2)}}$$

where

$$s_P^2 = \frac{1}{n_1 + n_2 - 2} \left[\sum_{i=1}^{n_1} (X_i - \bar{X})^2 + \sum_{i=1}^{n_2} (Y_i - \bar{Y})^2 \right].$$

and the rejection region

$$T(\underline{X}, \underline{Y})^2 \geq c.$$

Under H_0 , $T(\underline{X}, \underline{Y}) \sim \text{Student}(n_1 + n_2 - 2)$. Now suppose that once the data are observed, we have $n_1 = 12$, $n_2 = 7$, $\bar{x} = 3.214$, $\bar{y} = 2.572$ and $s_P^2 = 1.710$. Then

$$T(\underline{x}, \underline{y}) = \frac{3.214 - 2.572}{\sqrt{1.710 (1/12 + 1/7)}} = 1.032$$

so that

$$\begin{aligned} p(\underline{x}) &= \sup_{\theta \in \Theta_0} \Pr[T(\underline{X}, \underline{Y})^2 \geq T(\underline{x}, \underline{y})^2 \mid H_0] = \Pr[T(\underline{X}, \underline{Y}) \leq -T(\underline{x}, \underline{y}) \mid H_0] + \Pr[T(\underline{X}, \underline{Y}) \geq T(\underline{x}, \underline{y}) \mid H_0] \\ &= \Pr[T(\underline{X}, \underline{Y}) \leq -1.032 \mid H_0] + \Pr[T(\underline{X}, \underline{Y}) \geq 1.032 \mid H_0] \\ &= \text{St}_{n_1+n_2-2}(-1.032) + 1 - \text{St}_{n_1+n_2-2}(1.032) \\ &= 0.158 + 0.158 = 0.316. \end{aligned}$$

where $\text{St}_\nu(x)$ is the inverse cdf for the Student-t distribution with ν degrees of freedom evaluated x . Note that the supremum is removed as the distribution of $T(\underline{X}, \underline{Y})$ given H_0 is independent of the parameters.

Example 3 Suppose that $X \sim \text{Binomial}(n, \theta)$. To test

$$\begin{aligned} H_0 &: \theta = \theta_0 \\ H_1 &: \theta > \theta_0 \end{aligned}$$

consider using the statistic $T(X) = X$ and rejection region $\mathcal{R} \equiv \{X \geq c_n\}$. Suppose that $\theta_0 = 0.2$, $n = 20$ and $x = 14$. Then

$$p(\underline{x}) = \sup_{\theta \in \Theta_0} \Pr[T(X) \geq T(\underline{x})|\theta] = \Pr[X \geq 14|\theta = 0.2] = \sum_{x=14}^{20} \binom{20}{x} 0.2^x 0.8^{20-x} = 1.845 \times 10^{-6}$$

Example 4 Suppose that $X_1, \dots, X_n \sim \text{Exponential}(\theta)$. To test

$$\begin{aligned} H_0 &: \theta = \theta_0 \\ H_1 &: \theta < \theta_0 \end{aligned}$$

consider using the statistic $T(\underline{X}) = \sum_{i=1}^n X_i$ and rejection region $\mathcal{R} \equiv \{T(\underline{X}) > c_n\}$. Suppose that $\theta_0 = 3$, $n = 15$ and $T(\underline{x}) = 4.197$. Then

$$p(\underline{x}) = \sup_{\theta \in \Theta_0} \Pr[T(\underline{X}) \geq T(\underline{x})|\theta] = \Pr[T(\underline{X}) \geq 4.197|\theta_0 = 3].$$

Now for any θ , $T(\underline{X}) \sim \text{Gamma}(n, \theta)$ so, using the function `pgamma` in R.

$$p(\underline{x}) = \int_{4.197}^{\infty} \frac{\theta_0^n}{\Gamma(n)} x^{n-1} e^{-\theta_0 x} dx = 0.716$$

Example 5 Suppose that $X_1, \dots, X_n \sim \text{Uniform}(0, \theta)$. To test

$$\begin{aligned} H_0 &: \theta = \theta_0 \\ H_1 &: \theta > \theta_0 \end{aligned}$$

consider using the statistic $T(\underline{X}) = X_{(n)} = \max\{X_1, \dots, X_n\}$ and rejection region $\mathcal{R} \equiv \{T(\underline{X}) > c_n\}$. Suppose that $\theta_0 = 2$, $n = 8$ and $T(\underline{x}) = 1.842$. Then

$$p(\underline{x}) = \sup_{\theta \in \Theta_0} \Pr[T(\underline{X}) > T(\underline{x})|\theta] = \Pr[X_{(n)} > 1.842|\theta_0 = 2]$$

Now for any θ , by results for maximum order statistics the pdf of $T(\underline{X})$ is

$$F_{T|\theta}(t|\theta) = \left(\frac{t}{\theta}\right)^n \quad 0 < t < \theta$$

so, for a level α test we have

$$1 - \left(\frac{c_n}{\theta_0}\right)^n = \alpha \quad \therefore \quad c_n = \theta_0(1 - \alpha)^{1/n}$$

If $n = 8$, and $\alpha = 0.05$, then $c_n = 1.987$. Finally, we have

$$p(\underline{x}) = \Pr[X_{(n)} > 1.842|\theta_0 = 2] = 1 - \left(\frac{1.842}{2}\right)^8 = 0.482.$$

If $T(\underline{x}) = 1.971$, then

$$p(\underline{x}) = \Pr[X_{(n)} > 1.971|\theta_0 = 2] = 1 - \left(\frac{1.971}{2}\right)^8 = 0.110$$