

Sequential probability ratio test, continued.

Stat 563

4-25-19

Remaining question was: How are $k_0 \neq k_1$ chosen?

(1)

$$\text{Let } R_n = \{ \vec{x} : k_0 < \Delta_j < k_1, \forall 1 \leq j \leq n-1 \text{ and } \Delta_n \leq k_0 \}$$

$$A_n = \{ \vec{x} : k_0 < \Delta_j < k_1, \forall 1 \leq j \leq n-1 \text{ and } \Delta_n \geq k_1 \}$$

$$\alpha = \text{Prob}(\text{Type I error})$$

$$= \text{Prob}(R_n | H_0) = P\left(\bigcup_{n=1}^{\infty} R_n | H_0\right)$$

$$\alpha = \sum_{n=1}^{\infty} P(R_n | H_0) = \sum_{n=1}^{\infty} \int_{R_n} L(\theta_0) d\vec{x}$$

(2)

Note: for $\vec{x} \in R_n$, $\Delta_n \leq k_0$

$$\frac{L(\theta_0)}{L(\theta_1)}$$

So, for $\vec{x} \in R_n$, $L(\theta_0) \leq k_0 L(\theta_1)$

$$\alpha \leq \sum_{n=1}^{\infty} \int_{R_n} k_0 L(\theta_1) d\vec{x} = \sum_{n=1}^{\infty} k_0 P(\vec{x} \in R_n | H_1)$$

$$\alpha \leq k_0 \sum_{n=1}^{\infty} P(\vec{r} \in R_n | H_1) \quad (3)$$

$$\begin{aligned} \text{let } t_n &= \text{Prob}(\text{procedure terminates at step } n | H_1) \\ &= P(R_n | H_1) + P(A_n | H_1) \end{aligned}$$

$$\begin{aligned} \alpha &\leq k_0 \sum_{n=1}^{\infty} (t_n - P(A_n | H_1)) \\ &= k_0 \left(\sum_{n=1}^{\infty} t_n - \sum_{n=1}^{\infty} P(A_n | H_1) \right) \end{aligned}$$

$$\leq k_0 (1 - P(\bigcup_{n=1}^{\infty} A_n | H_1)) \quad (4)$$

$$\alpha \leq k_0 (1 - \beta)$$

$$\begin{aligned} \text{Also, } 1 - \alpha &= P(\text{Acc } H_0 | H_0) \\ &= P(\bigcup_{n=1}^{\infty} A_n | H_0) \\ &= \sum_{n=1}^{\infty} P(A_n | H_0) \\ &= \sum_{n=1}^{\infty} \int_{A_n} L(\theta_0) d\vec{r} \end{aligned}$$

(5)

Note: For $\vec{\pi} \in A_n$, $\Delta_n \geq k_1$

$$\frac{L(\theta_1)}{L(\theta_0)}$$

$$L(\theta_1) \geq k_1 L(\theta_0)$$

$$1 - \alpha \geq \sum_{n=1}^{\infty} \int_{A_n} k_1 L(\theta_1) d\vec{\pi}$$

$$= k_1 \sum_{n=1}^{\infty} P(\vec{\pi} \in A_n | H_1)$$

$$= k_1 P\left(\bigcup_{n=1}^{\infty} A_n | H_1\right) = k_1 \beta$$

(6)

$$\text{So } \alpha \leq k_0(1-\beta) \text{ and } 1-\alpha \geq k_1\beta$$

$$\frac{\alpha}{1-\beta} \leq k_0 \text{ and } \frac{1-\alpha}{\beta} \geq k_1$$

Choose a desired α and β , called α^* and β^*

$$\text{Let } k_0 = \frac{\alpha^*}{1-\beta^*} \text{ and } k_1 = \frac{1-\alpha^*}{\beta^*}$$

$$\text{Then } \frac{\alpha}{1-\beta} \leq \frac{\alpha^*}{1-\beta^*} \text{ and } \frac{1-\alpha}{\beta} \geq \frac{1-\alpha^*}{\beta^*}$$

Does this imply that $\alpha \leq \alpha^*$ and $\beta \leq \beta^*$?

$$\alpha(1-\beta^*) \leq \alpha^*(1-\beta) \text{ and } (1-\alpha)\beta^* \geq \beta(1-\alpha^*) \quad (7)$$

$$\alpha - \alpha\beta^* \leq \alpha^* - \alpha^*\beta$$

$$\beta - \alpha^*\beta \leq \beta^* - \alpha\beta^*$$

$$\alpha + \beta - \cancel{\alpha\beta^*} - \cancel{\alpha^*\beta} \leq \alpha^* + \beta^* - \cancel{\alpha^*\beta} - \cancel{\alpha\beta^*}$$

$$\therefore \underline{\alpha + \beta \leq \alpha^* + \beta^*}$$

(8)

Look at simple linear regression as maximum likelihood problem.

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

↑
known
constants

$$\varepsilon_i \sim \text{iid } N(0, \sigma^2)$$

$$i = 1, \dots, n$$

$$Y_i \sim N(\beta_0 + \beta_1 X_i, \sigma^2), \text{ all independent}$$

$$L(\beta_0, \beta_1, \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_i - \beta_0 - \beta_1 x_i)^2}$$

$$L = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2} \quad (9)$$

$$l(\beta_0, \beta_1, \sigma^2) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$$

$$\frac{\partial l}{\partial \beta_0} = -\frac{1}{2\sigma^2} \sum_{i=1}^n 2(y_i - \beta_0 - \beta_1 x_i)(-1) \stackrel{\text{set}}{=} 0$$

$$\sum y_i - n\beta_0 - \beta_1 \sum x_i = 0$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\frac{\partial l}{\partial \beta_1} = -\frac{1}{2\sigma^2} \sum_{i=1}^n 2(y_i - \beta_0 - \beta_1 x_i)(-x_i) \stackrel{\text{set}}{=} 0 \quad (10)$$

$$\sum x_i y_i - \beta_0 \sum x_i - \beta_1 \sum x_i^2 = 0$$

$$\sum x_i y_i - (\bar{y} - \beta_1 \bar{x}) \sum x_i - \beta_1 \sum x_i^2 = 0$$

$$\sum x_i y_i - \bar{y} \sum x_i = \beta_1 (\sum x_i^2 - \bar{x} \sum x_i)$$

$$\hat{\beta}_1 = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{\sum x_i^2 - n \bar{x}^2} = \frac{S_{xy}}{S_{xx}}$$

Note: These are the same $\hat{\beta}_0$ and $\hat{\beta}_1$

(11)

that minimize the error sum of squares.

$$\frac{\partial L}{\partial \sigma^2} = -\frac{n}{2} \frac{2\pi}{2\pi\sigma^2} - \frac{1}{2} \frac{-1}{\sigma^4} \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 \stackrel{!}{=} 0$$

$$-n\sigma^2 + \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 = 0$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (y_i - (\bar{y} - \hat{\beta}_1 \bar{x}) - \hat{\beta}_1 x_i)^2$$

$$= \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y} - \hat{\beta}_1 (x_i - \bar{x}))^2$$

$$\hat{\sigma}^2 = \frac{1}{n} \left[\underbrace{\sum_{i=1}^n (y_i - \bar{y})^2}_{S_{yy}} + \hat{\beta}_1^2 \underbrace{\sum_{i=1}^n (x_i - \bar{x})^2}_{S_{xx}} - 2\hat{\beta}_1 \underbrace{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}_{S_{xy}} \right] \quad (12)$$

$$\hat{\sigma}^2 = \frac{1}{n} [S_{yy} + \hat{\beta}_1^2 S_{xx} - 2\hat{\beta}_1 S_{xy}]$$

$$= \frac{1}{n} \left[S_{yy} + \left(\frac{S_{xy}}{S_{xx}} \right)^2 S_{xx} - 2 \left(\frac{S_{xy}}{S_{xx}} \right) S_{xy} \right]$$

$$\hat{\sigma}^2 = \frac{1}{n} \left[S_{yy} - \frac{S_{xy}^2}{S_{xx}} \right] \quad \left(\text{This disagrees with} \right. \\ \left. \text{MSE} = \frac{SSE}{n-2} \right)$$

⑬

HW #4 (hp9 (p.459))

#4, 7, 17

due 5/2

9.4 Let X_1, \dots, X_n be a random sample from a $n(0, \sigma_X^2)$, and let Y_1, \dots, Y_m be a random sample from a $n(0, \sigma_Y^2)$, independent of the X s. Define $\lambda = \sigma_Y^2 / \sigma_X^2$.

- (a) Find the level α LRT of $H_0: \lambda = \lambda_0$ versus $H_1: \lambda \neq \lambda_0$.
- (b) Express the rejection region of the LRT of part (a) in terms of an F random variable.
- (c) Find a $1 - \alpha$ confidence interval for λ .

9.7 (a) Find the $1 - \alpha$ confidence set for a that is obtained by inverting the LRT of $H_0: a = a_0$ versus $H_1: a \neq a_0$ based on a sample X_1, \dots, X_n from a $n(\theta, a\theta)$ family, where θ is unknown.

- (b) A similar question can be asked about the related family, the $n(\theta, a\theta^2)$ family. If X_1, \dots, X_n are iid $n(\theta, a\theta^2)$, where θ is unknown, find the $1 - \alpha$ confidence set based on inverting the LRT of $H_0: a = a_0$ versus $H_1: a \neq a_0$.

9.17 Find a $1 - \alpha$ confidence interval for θ , given X_1, \dots, X_n iid with pdf

- (a) $f(x|\theta) = 1, \theta - \frac{1}{2} < x < \theta + \frac{1}{2}$.
- (b) $f(x|\theta) = 2x/\theta^2, 0 < x < \theta, \theta > 0$.