HW 9

① Se $L(x)$ e $U(x)$ satisfazem $P_\theta(L(x) \leq \theta) = 1-\alpha_1$ e $P_\theta(U(x) \geq \theta) = 1-\alpha_2$, e $L(x) \leq U(x)$ $\forall x$, mostre que $P_\theta(L(x) \leq \theta \leq U(x)) = 1-\alpha_1-\alpha_2$.

Sendo $A = \{x : L(x) \leq \theta\}$ e $B = \{x : U(x) \geq \theta\}$. Então $A \cap B = \{x : L(x) \leq \theta \leq U(x)\}$ e $1 \geq P(A \cup B) = P\{L(x) \leq \theta$ ou $\theta \leq U(x)\} \geq P\{L(x) \leq \theta$ ou $\theta < L(x)\} = 1$, pois $L(x) \leq U(x)$. Além disso, $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 1-\alpha_1 + 1-\alpha_2 - 1 = 1-\alpha_1-\alpha_2$.

/ /

② Seja $X_1, \dots, X_n$ independentes com distribuição

$$P(X_i \le x) = \begin{cases} 0 & x \le 0 \\ (x/\beta)^{\alpha} & 0 < x < \beta \\ 1 & x \ge \beta \end{cases}$$

ⓐ Se α é conhecido, α_0, encontre um limite superior de confiança para β com ~~o~~ coeficiente de confiança 0.95.

ⓑ Usando os dados

22.0, 23.9, 20.9, 23.8, 25.0, 24.0
21.7, 23.8, 22.8, 23.1, 23.1, 23.5
23.0, 23.0

para construir um intervalo de confiança para β. Suponha que α é conhecido e igual ao MLE.

ⓐ O MLE de β é $X_{(n)} = \max X_i$. Desde que β é um parametro $X_{(n)}/\beta$ é um pivo, e

$$0.05 = P_{\beta}(X_{(n)}/\beta \le c) = P(X_i \ \forall i \le c\beta)$$

$$= \left(\frac{c\beta}{\beta}\right)^{\alpha_0 n} = c^{\alpha_0 n}$$

então $c = 0.05^{1/\alpha_0 n}$. Daí, $0.95 = P_{\beta}(X_{(n)}/\beta > c) = P(X_{(n)}/c > \beta)$, e $\{\beta : \beta < X_{(n)}/(0.05^{1/\alpha_0 n})\}$ é um limite superior de confiança 95% para β.

ⓑ $\hat{\alpha} = 12.59$ e $X_{(n)} = 25$. Então o intervalo de confiança é $(0,\ 25/[0.05^{1/12.59 \cdot 14}]) = (0,\ 24\cancel{5}.43)$

(3) Observando $X_1, ..., X_n$ onde $X_i \sim Bernoulli(p)$, queremos um intervalo da forma $[0, U(X_1, ..., X_n)]$, onde $P_p(p \in [0, U(X_1, ..., X_n)]) \geq 1-\alpha$.

(a) Mostre que a inversão da região de aceitação do teste

$$H_0: p = p_0 \quad vs \quad H_1: p < p_0$$

da um intervalo de confiança e nível de confiança dessa forma.

(b) Encontre as equações para construção do intervalo de confiança.

(a) Podemos basear o teste em $T = \sum X_i = binomial(n, p)$, desde que T seja suficiente para p. Como a distribuição binomial tem razão de verossimilhança monótona pelo Teorema de Karlin-Rubin o teste que rejeita H_0 se $T > k(p_0)$ é o teste UMP deste tamanho. Para cada p_0, podemos escolher $k(p_0)$ para ter um teste de nível α. $k(p_0)$ é definido para ser um inteiro entre 0 e n que satisfaz as inequações

$$\sum_{y=0}^{k(p_0)} \binom{n}{y} p_0^y (1-p_0)^{n-y} \geq 1-\alpha$$

e $$\sum_{y=0}^{k(p_0)} \binom{n}{y} p_0^y (1-p_0)^{n-y} < 1-\alpha.$$

(b) $k(p)$ é o inteiro que satisfaz simultaneamente.

$$\sum_{y=k(p)}^{n} \binom{n}{y} p^y (1-p)^{n-y} \geq 1-\alpha$$

e $$\sum_{y=k(p)+1}^{n} \binom{n}{y} p^y (1-p)^{n-y} < 1-\alpha.$$

/ /

(4) (a) Encontre o ~~t.c.~~ conjunto de confiança $1-\alpha$ para a obtido invertendo o LRT de $H_0: a = a_0$ vs. $H_1: a \neq a_0$ baseado na amostra $X_1, \dots, X_n \sim N(\theta, a\theta)$, onde θ é desconhecido.

(b) Uma questão similar pode ser quanto feita sobre a família $N(\theta, a\theta^2)$. Se $X_1, \dots, X_n \overset{iid}{\sim} N(\theta, a\theta^2)$, onde θ é desconhecido, encontre o ~~c.~~ conjunto de confiança $1-\alpha$ baseado no ~~t.~~ LRT de $H_0: a = a_0$ vs. $H_1: a \neq a_0$.

(a) Sendo $X_1, \dots, X_n \overset{iid}{\sim} N(\theta, a\theta)$

$$f(x \mid \theta, a) = \frac{1}{\sqrt{2\pi a\theta}} \exp\left\{-\frac{(x_i - \theta)^2}{2a\theta}\right\}$$

Sobre H_0 a razão da verossimilhança

$$\lambda(\underset{a_0}{x}) = \frac{\sup_{\theta} L(\theta, a_0)}{\sup_{\theta, a} L(\theta, a)}.$$

O ~~E~~ MLE restrito é

$$\hat{\theta}_0 = \frac{-a_0 + \sqrt{a_0^2 + 4\sum x_i^2}}{2} \quad (\text{onde } a = a_0)$$

e irrestrito

$$\hat{\theta} = \bar{X} \quad e \quad \hat{a} = \frac{\sum(X_i - \bar{X})^2}{n\bar{X}}.$$

Então,

$$\lambda(x \mid a_0) = \frac{\left(\frac{\hat{a}\hat{\theta}}{a_0\hat{\theta}_0}\right)^{n/2} \exp\left\{-\frac{1}{2a_0\hat{\theta}_0}\sum(X_i - \hat{\theta}_0)^2\right\}}{\exp\left\{-\frac{1}{2\hat{a}\hat{\theta}}\sum(X_i - \hat{\theta})^2\right\}}.$$

$$= \left(\frac{1}{2\pi a_0 \hat{\theta}_0}\right)^{\frac{n}{2}} e^{\frac{n}{2}} e^{-\frac{1}{2a_0\theta_0}\sum(x_i-\hat{\theta}_0)^2}.$$

A região de rejeição com tamanho α é $\{x : \lambda(x) \leq c_\alpha\}$, e ~~te~~ um conjunto de confiança $1-\alpha$ é $\{\alpha_0 : \lambda(x) \geq c_\alpha\}$.

(b) ~~O MLE restrito é~~

$$\cancel{\hat{\theta}_0 = \bar{x} + \sqrt{\bar{x} + 4a_0 (\theta}$$

$$f(x|\theta, a) = \frac{1}{\sqrt{2\pi a \theta^2}} \exp\left\{-\frac{(x_i-\theta)^2}{2a\theta^2}\right\}$$

Daí, o MLE restrito $(a = a_0)$ é (verifique o exercício 8.8 de [CB])

$$\hat{\theta}_0 = \bar{x} + \sqrt{\bar{x} + 4a_0(\hat{\sigma}^2 + \bar{x}^2)} / 2a_0.$$

E o MLE irrestrito com

$$\hat{\theta} = \bar{x} \text{ e } \hat{a} = \frac{1}{n\bar{x}^2} \sum (x_i - \bar{x})^2 = \frac{\hat{\sigma}^2}{\bar{x}^2}.$$

onde $\hat{\sigma} = \frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2$. Portanto

$$\lambda(x) = \left(\frac{\hat{\sigma}}{\theta_0}\right)^n \exp\left\{\frac{n}{2} - \frac{\sum(x_i-\hat{\theta}_0)^2}{2\hat{\theta}_0^2}\right\}.$$

A região de rejeição com tamanho α é $\{x : \lambda(x) \leq c_\alpha\}$, e um conjunto de confiança $1-\alpha$ é $\{\alpha_0 : \lambda(x) \geq c_\alpha\}$.

6 5 Se T é uma v.a. contínua com cdf $F_T(t|\theta)$ e $\alpha_1+\alpha_2=\alpha$, mostre que uma região de aceitação a da hipótese $H_0: \theta=\theta_0$ é $\{t: \alpha_1 \leq F_T(t|\theta_0) \leq 1-\alpha_2\}$, com confiança associada $1-\alpha$ ao conjunto $\{\theta: \alpha_1 \leq F_T(t|\theta) \leq 1-\alpha_2\}$

Supondo $F_T(t|\theta) \sim Uniform(0,1)$ (me parece que não é possível resolver sem essa hipótese). Daí

$$P_{\theta_0}(\{T: \alpha_1 \leq F_T(T|\theta_0) \leq 1-\alpha_2\}) = P_{\theta}(\alpha_1 \leq U \leq 1-\alpha_2) = 1-\alpha_2-\alpha_1$$

onde $U \sim Uniform(0,1)$. Desde que

$$t \in \{t: \alpha_1 \leq F_T(t|\theta) \leq 1-\alpha_2\} \Leftrightarrow \theta \in \{\theta: \alpha_1 \leq F_T(t|\theta) \leq 1-\alpha_2\}$$

o mesmo cálculo chega que o intervalo tem confiança $1-\alpha_2-\alpha_1$.

/ /

⑥ Encontre um intervalo de confiança $1-\alpha$ para θ, dado $X_1, \ldots, X_n$ iid de um pdf

ⓐ $f(x|\theta) = 1, \quad \theta - \frac{1}{2} < x < \theta + \frac{1}{2}$

ⓑ $f(x|\theta) = \frac{2x}{\theta^2}, \quad 0 < x < \theta, \; \theta > 0$

ⓐ A distribuição é uma $Uniform(\theta - 1/2, \theta + \frac{1}{2})$ e $X - \theta \sim Uniform(-\frac{1}{2}, \frac{1}{2})$.

$$P(a \leq X \leq b) = b - a$$

Escolhemos a e b de forma que $b - a = 1 - \alpha$. Uma escolha é $a = -\frac{1}{2} + \frac{\alpha}{2}$ e $b = \frac{1}{2} - \frac{\alpha}{2}$.

ⓑ Tomando $T = X/\theta$, então T tem pdf $f(t) = 2t, \; 0 \leq t \leq 1$,

$$P(a \leq X/\theta \leq b) = \int_a^b 2t\,dt = b^2 - a^2$$

Escolhemos a e b de forma que $b^2 - a^2 = 1 - \alpha$. Uma escolha é $a = \sqrt{\alpha/2}$ e $b = \sqrt{1 - \alpha/2}$.

/ /

7 (a) Seja $X_1, \dots, X_m \sim Poisson(\lambda)$ e $Y = \sum X_i$. Construa um intervalo para λ invertendo ~~a~~ o LRT.

(b) Use o seguinte dado

155, 104, 66, 50, 36, 40, 30, 35, 42.

e construa um intervalo de confiança 90%.

(a) A estatística LRT para $H_0: \lambda = \lambda_0$ vs $H_1: \lambda \neq \lambda_0$ é

$$g(y) = \frac{e^{-n\lambda_0}(n\lambda_0)^y}{e^{-n\hat{\lambda}}(n\hat{\lambda})^y}$$

onde $Y = \sum X_i \sim Poisson(n\lambda)$ e $\hat{\lambda} = \frac{Y}{n}$ que é o MLE de λ.
Para λ_0, a região de aceitação $A(\lambda_0)$ é definida por

$$A(\lambda_0) = \{y : g(y) > c(\lambda_0)\}$$

$c(\lambda_0)$ é escolhido de forma que a probabilidade de Y cair na região de aceitação é pelo menos $1-\alpha$, onde α é o nível de significância.

(b) Pelos dados temos $Y = 558$ e $n = 9$, então do item (a) o intervalo de confiança é (57.78, 66.45). ~~Para~~