HW6

① Seja $X_1, \dots, X_n$ uma amostra aleatória para uma população com pdf

$$f(x|\theta) = \frac{1}{2\theta}, \quad -\theta < x < \theta, \quad \theta > 0$$

Encontre, se existir um, um melhor estimador não viesado de θ.

Para encontrar um estimador não-viesado ótimo de θ, primeiro encontramos uma estatística suficiente completa. A pdf conjunta é

$$L(\theta X_1, \dots, X_n|\theta) = \prod_{i=1}^{n} \frac{1}{2\theta} \mathbb{1}(-\theta < X_i < \theta)$$

$$= \left(\frac{1}{2\theta}\right)^n \prod_{i=1}^{n} \mathbb{1}(0 < |X_i| < \theta)$$

$$= \frac{1}{(2\theta)^n} \mathbb{1}(\max|X_i| < \theta)$$

Fazendo,

$$\frac{L(X_1, \dots, X_n|\theta)}{L(Y_1, \dots, Y_n|\theta)} = \frac{\mathbb{1}(\max|X_i| < \theta)}{\mathbb{1}(\max|Y_i| < \theta)}$$

Logo, $\max|X_i|$ é uma estatística suficiente completa.
Agora seja $Y = \max|X_i|$, note que, $f_Y(y) = \frac{ny^{n-1}}{\theta^n}$, $0 < y < \theta$
~~Suponha $g(Y)$ é uma função tal que~~ Agora

$$EY = \int_0^\theta y \frac{ny^{n-1}}{\theta^n} dy = \frac{n}{n+1}\theta \Rightarrow E\left(\frac{n+1}{n} Y\right) = \theta.$$

Assim, $\frac{n+1}{n}\max|X_i|$ é o melhor estimador não viesado para θ pois é uma função de uma estatística fo suficiente completa.

/ /

(2) Seja $X_1, \ldots, X_n$ uma amostra aleatória para uma população com média μ e variância σ^2.

(a) Mostre que o estimador $\sum a_i X_i$ é não viesado para μ se $\sum a_i = 1$.

(b) Dentre todos os estimadores lineares não viesados encontre o que tem variância mínima e calcule ela.

(a) $E(\sum a_i X_i) = \sum a_i EX_i = \mu \sum a_i = \mu$. Então é não-viesado

(b) $Var(\sum a_i X_i) = \sum a_i^2 Var X_i = \sigma^2 \sum a_i^2$. Então queremos minimizar $\sum a_i^2$ sujeito a $\sum a_i = 1$.

$$\sum a_i^2 = \sum \left[\left(a_i - \frac{1}{n}\right) + \frac{1}{n}\right]^2 = \sum \left(a_i - \frac{1}{n}\right)^2 + \frac{1}{n}$$

é mínimo quando $a_i = \frac{1}{n}$ $\forall i$. Então $\sum \frac{1}{n} X_i = \bar{X}$ é o estimador linear não-viesado com variância mínima.

/ /

③ Seja $X_1, \dots, X_n \overset{iid}{\sim}$ iid de uma distribuição com média μ e variância σ^2, e seja S^2 o estimador não viesado ~~usual~~ usual para σ^2 ($S^2 = \frac{1}{n-1}\sum(X_i - \bar{X})^2$).

ⓐ Mostre que, para qualquer estimador da forma aS^2, onde a é constante

$$MSE(aS^2) = E[aS^2 - \sigma^2]^2 = a^2 Var\, S^2 + (a-1)^2\sigma^4.$$

ⓑ Mostre que

$$Var(S^2) = \frac{1}{n}\left(k - \frac{n-3}{n-1}\sigma^4\right),$$

onde $k = E[X-\mu]^4/\sigma^4$ é a kurtosis.

ⓒ Mostre que, sob normalidade, o kurtosis é $3\sigma^4$ e estabeleça, neste caso, o estimador ~~que~~ da forma aS^2 que tem MSE mínimo é $\frac{n-1}{n+1}S^2$.

ⓓ Se não assumirmos normalidade, mostre que o $MSE(aS^2)$ é minimizado com

$$a = \frac{n-1}{(n+1) + \frac{(k-1)(n-1)}{n}}$$

o que é inútil, pois depende de um parametro.

ⓔ Mostre que

(i) para distribuições com $k>3$, o ótimo a satisfaz $a < \frac{n-1}{n+1}$

(ii) para distribuições com $k<3$, o ótimo a satisfaz $\frac{n-1}{n+1} < a < 1$.

/ /

(a)

$$E[aS^2-\sigma^2]^2 = E[a^2S^4 - 2a\,S^2\sigma^2 + \sigma^4]$$

$$= a^2ES^4 - 2a\sigma^2ES^2 + \sigma^4$$

$$= a^2(\sigma^4 + \text{Var}\,S^2) - 2a\sigma^4 + \sigma^4$$

$$= a^2\,\text{Var}\,S^2 + (a-1)^2\sigma^4.$$

(b) Fora do ~~escopo~~ alcance do curso.

(c) Sob normalidade $K=3$. Temos que

$$K = \frac{E[X-\mu]^4}{\sigma^4} = E\left[\frac{X-\mu}{\sigma}\right]^4 = EZ^4$$

onde $Z \sim N(0,1)$. Agora, sendo $g(z) = z^3$ temos

$$K = EZ^4 = Eg(Z)Z \overset{(*)}{=} 1.\,E(3Z^2) = 3EZ^2 = 3.$$

(*) pelo Lemma de Stein que diz que sendo $X \sim N(\theta, \sigma^2)$ e g diferenciável com $E|g'(X)| < \infty$. Então

$$E[g(X)(X-\theta)] = \sigma^2 Eg'(X).$$

Para minimizar $MSE(S^2)$ em geral, escrevendo $\text{Var}\,S^2 = B\sigma^4$. Então minimizar ~~Var S^2~~ $MSE\,S^2$ é equivalente a minimizar $a^2B + (a-1)^2$. E Igualando a zero para obter que o valor que minimiza a é $1/(B+1)$. Usando a expressão da parte (b), com a premissa de normalidade a é minimizado com

$$\frac{1}{B+1} = \frac{1}{\frac{1}{n}\left(3-\frac{n-3}{n-1}\right)+1} = \frac{n-1}{n+1}.$$

(d) Pela parte (c) avaliamos $1/(B+1)$ com

$$B = \frac{1}{n}\left(k - \frac{n-3}{n-1}\right)$$

e daí chegamos em

$$a = \frac{n-1}{(n+1) + \frac{(k-3)(n-1)}{n}}.$$

/ /

(9) Suponha que quando o raio de um circulo é medido, um erro é cometido que tem distribuição $N(0, \sigma^2)$. Se n independentes medidas são feitas, encontre um estimador não viesado para a area do circulo. É o melhor estimador não viesado?

Sendo $X_i \sim N(r, \sigma^2)$, então $\bar{X} \sim N(r, \sigma^2/n)$ e $E\bar{X}^2 = r^2 + \sigma^2/n$. Daí $E[\pi \bar{X}^2 - \pi \sigma^2/n] = \pi r^2$ é o melhor estimador não-viesado porque $\bar{X}$ é uma estatística suficiente completa.

/ /

⑤ Para cada pdf, seja $X_1, \dots, X_n$ uma amostra desta distribuição. Em cada caso, encontre o melhor estimador não-viesado de θ^r.

ⓐ $f(x|\theta) = \frac{1}{\theta}$, $0 < x < \theta$, $r < n$

ⓑ $f(x|\theta) = e^{-(x-\theta)}$, $x > \theta$

ⓒ $f(x|\theta) = \frac{e^{-x}}{e^{-\theta} - e^{-b}}$, $\theta < x < b$ com b conhecido.

De forma geral. Supondo $X \sim f(x|\theta) = c(\theta)m(x)$, $a < x < \theta$. Então $1/c(\theta) = \int_a^\theta m(x)\,dx$, e a cdf de X é $F(x) = c(\theta)/c(x)$, $a < x < \theta$. Seja $Y = \max X$ uma estatística. Como $x < \theta$ e já avaliamos antes Y é uma estatística suficiente completa. Daí, qualquer função $T(Y)$ que seja um estimador não viesado de $h(\theta)$ é o melhor. Pelo Teorema 5.4.4 [Casela e Berger] a pdf de Y é $g(y|\theta) = n\,m(y)\,c(\theta)^n / c(y)^{n-1}$, $a < y < \theta$. Considerando

$$\int_a^\theta f(x|\theta)\,dx = 1 \quad e \quad \int_a^\theta T(y)\,g(y|\theta)\,dy = h(\theta)$$

Reescrevendo

$$\int_a^\theta m(x)\,dx = \frac{1}{c(\theta)} \quad e \quad \int_a^\theta \frac{T(y)\,n\,m(y)}{c(y)^{n-1}}\,dy = \frac{h(\theta)}{c(\theta)^n}$$

Diferenciando os 2 lados das 2 equações temos

$$m(\theta) = \frac{c'(\theta)}{c(\theta)^2} \quad e \quad \frac{T(\theta)\,n\,m(\theta)}{c(\theta)^{n-1}} = \frac{c(\theta)^n h'(\theta) - h(\theta)\,n\,c(\theta)^{n-1}\,c'(\theta)}{c(\theta)^{2n}}$$

Digitalizado com CamScanner

/ /

Trocando θ por Y e resolvendo as duas equações para $T(Y)$ temos que o melhor estimador não-viesado de $h(\theta)$ é

$$T(Y) = h(Y) + \frac{h'(Y)}{n\,m(Y)\,c(Y)}$$

Para $h(\theta) = \theta^r$, $h'(\theta) = r\theta^{r-1}$.

(a) Aqui $m(x) = 1$ e $c(\theta) = \frac{1}{\theta}$, daí

$$T(Y) = Y^r + \frac{rY^{r-1}}{n(1/Y)} = \frac{n+r}{n} Y^r$$

(b) Se θ é o menor ponto final do suporte, o menor valor da estatística de ordem $Y = \min X_i$ é uma estatística suficiente completa. Pelo argumento acima o melhor estimador não-viesado de $h(\theta)$ é

$$T(Y) = h(Y) - \frac{h'(Y)}{n\,m(Y)\,c(Y)}.$$

Para esta pdf, $m(x) = e^{-x}$ e $c(\theta) = e^{\theta}$. Daí

$$T(Y) = Y^r - \frac{rY^{r-1}}{ne^{-Y}e^{Y}} = Y^r - \frac{rY^{r-1}}{n}.$$

(c) Para esta pdf, $m(x) = e^{-x}$ e $c(\theta) = 1/(e^{-\theta} - e^{-b})$ Daí

$$T(Y) = Y^r - \frac{rY^{r-1}}{ne^{-Y}}(e^{-Y} - e^{-b})$$

$$= Y^r - \frac{rY^{r-1}(1 - e^{-(b-Y)})}{n}.$$