

1. Uniformly cool.

Seja $\mathbf{X}_n = (X_1, X_2, \dots, X_n)$ uma amostra aleatória de uma distribuição uniforme em $(0, \theta]$ com densidade comum com respeito a Lebesgue

$$f_\theta(x) = \frac{1}{\theta} \mathbb{I}(0 < x \leq \theta).$$

Defina $m := \min(\mathbf{X}_n)$ e $M := \max(\mathbf{X}_n)$.

- a) (20 pontos) Encontre o estimador de máxima verossimilhança para θ e mostre que ele é suficiente e completo. Discuta se este estimador é viesado e se é consistente.

- b) (10 pontos) Mostre que $U = \frac{m}{M}$ é ancilar.

Dica: Para uma amostra aleatória de tamanho n com cdf comum F e $a < b$, vale

$$\Pr(m \leq a, M \leq b) = [F(a)]^n - [F(b) - F(a)]^n.$$

- c) (10 pontos) Calcule $E_\theta[U]$.

Dica: Compute $E_\theta[m]$ e lembre-se de que se X e Y são v.a.s independentes, $E[XY] = E[X]E[Y]$.

2. So many dice...

Seja $\mathbf{X}_n = (X_1, X_2, \dots, X_n)$ uma amostra aleatória de uma distribuição binomial com número de tentativas $N \in \mathbb{N}$ conhecido e probabilidade de sucesso $\theta \in (0, 1)$ desconhecida. Considere estimar a probabilidade de exatamente um sucesso:

$$g(\theta) := P_\theta(X = 1) = N\theta(1 - \theta)^{N-1}.$$

- a) (10 pontos) Mostre que $\delta_1(\mathbf{X}_n) = \mathbb{I}(X_1 = 1)$ é um estimador não-viesado de $g(\theta)$;
- b) (10 pontos) Encontre uma estatística suficiente e completa, $T(\mathbf{X}_n)$.
- c) (10 pontos) Argumente que δ_1 não é suficiente e que portanto pode ser melhorado usando T . Mostre como fazer isso e obter um estimador $\delta_2(\mathbf{X}_n)$.
- d) (10 pontos) Exiba a fórmula para δ_2 e mostre que ele é ENVVUM.

1. Miss Independent

Seja X_1, \dots, X_n variáveis aleatórias i.i.d absolutamente contínuas com densidade comum

$$f_\theta(x) = \theta e^{-\theta x} \mathbb{I}\{x > 0\},$$

onde $\theta > 0$ é desconhecido.

- a) (5 pontos) Encontre a densidade de θX_i ;
- b) (15 pontos) Seja, $X_{(1)} \leq \dots \leq X_{(n)}$ as estatísticas de ordem e $\bar{X} = (X_1 + \dots + X_n)/n$ a média amostral. Mostre que \bar{X} e $X_{(1)}/X_{(n)}$ são independentes.

2. You can't truncate sufficiency

Tome X_1, X_2, \dots, X_n uma amostra aleatória de uma distribuição normal com média $\mu \in \mathbb{R}$ e variância $\sigma^2 \in \mathbb{R}_+$. Sabemos que média amostral $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$ e a variância amostral $\bar{V}_n = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ são estatísticas suficientes para a distribuição normal com $\theta = (\mu, \sigma^2)$.

Em algumas situações, essas observações só podem acontecer dentro de uma região $A \in \mathcal{B}$. Por regularidade, assuma que $k(\theta) = P_\theta(X \in A) > 0$. Nesse caso, a distribuição apropriada para a nova variável aleatória restrita X^* é a distribuição condicional

$$P_\theta(X^* \in B) = P_\theta(X \in B | X \in A), \quad B \in \mathcal{B}.$$

Esta distribuição é chamada de *truncamento* de X em um conjunto A .

- a) (10 pontos) Mostre que a média e a variância amostrais continuam sendo estatísticas suficientes para uma distribuição normal truncada em um intervalo $A = (a, b)$ com $-\infty \leq a < b \leq \infty$.
- b) Seja $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ uma família de distribuições dominada por uma medida σ -finita μ . Seja A um boreliano. Considere P_θ^* o truncamento de P_θ em A e $\mathcal{P}^* = \{P_\theta^* : \theta \in \Omega\}$. Mostre que:
- (10 pontos)* Se T é suficiente para \mathcal{P} , então é suficiente para \mathcal{P}^* ;
 - (10 pontos)* Se T é completa para \mathcal{P} , então é completa para \mathcal{P}^* .

4. Pfizer Inc.

Modelos de *resposta à dosagem* são de grande interesse da indústria farmacêutica. Uma possível aplicação seria na área de avaliação de risco de substâncias tóxicas, onde o realizador do experimento está interessado em estimar a probabilidade de alguma resposta (e.g, aparição de um tumor) em função da dosagem de toxina⁶ aplicada. Nesta questão vamos estudar as propriedades de suficiência mínima destes modelos.

Para isso, considere que vamos aplicar diferentes volumes de toxinas em ratos. $0 < d_1 < d_2 < \dots < d_m$ foram as dosagens escolhidas para este estudo. Para cada dosagem, foi escolhida uma população de n_i ratos. Os n_i 's são fixos. Seja $x = (x_1, \dots, x_m)$ o número de respostas à aplicação das doses, respectivamente. Isto é, para o i -ésimo grupo, de n_i ratos, x_i tiveram a reação desejada. Seja $p_\theta(d)$ a probabilidade de resposta a uma dose $d > 0$, onde θ é p -dimensional com $(p \leq m)$ em um espaço de parâmetros Θ . A distribuição de probabilidade de x é dada por

$$f_M(x|\theta) = \prod_{i=1}^m \binom{n_i}{x_i} p_\theta(d_i)^{x_i} [1 - p_\theta(d_i)]^{n_i - x_i},$$

onde M representa uma escolha particular de modelo de resposta à dose (podemos escolher diferentes funções para modelar p_θ). Seja \mathcal{F}_M a família de distribuições induzida por M .

- a) (5 pontos) Mostre que $T(x) = x$ é suficiente para \mathcal{F}_M para todo M .
- b) (15 pontos) Mostre o seguinte resultado preliminar. Seja \mathcal{P} uma família finita de distribuições com densidades p_i , $i = 0, 1, \dots, k$, todas tendo o mesmo suporte. Então, a estatística

$$T(X) = \left(\frac{p_1(X)}{p_0(X)}, \frac{p_2(X)}{p_0(X)}, \dots, \frac{p_k(X)}{p_0(X)} \right)$$

é suficiente mínima.

Dica: Mostre primeiro que se U é uma estatística suficiente para \mathcal{P} , então a razão $\frac{p_i(x)}{p_j(x)}$ é uma função apenas de $U(x)$.

- c) (10 pontos) Seja A uma matriz $m \times m$ com seus elementos definidos por

$$A_{i,j} = \ln \left(\frac{p_{\theta_j}(d_i)[1 - p_{\theta_0}(d_i)]}{p_{\theta_0}(d_i)[1 - p_{\theta_j}(d_i)]} \right).$$

Mostre que se existirem vetores $(\theta_0, \theta_1, \dots, \theta_m)$ tais que a matriz A é invertível, então $T(x) = x$ é suficiente mínima.