Suficiência

- Estatística suficiente;
- Teorema da fatorização;
- Suficiência conjunta;
- Suficiência mínima;

Um exemplo motivador

Suponha que os tempos de falha de um modelo de lâmpada podem ser modelados como $X_1, X_2, \ldots, X_n \sim \text{expo}(\theta)$.
Suponha que dois técnicos, Afonso e Bruna, medem cada um três lâmpadas, obtendo:

- $\boldsymbol{x}_\text{A} = \{1.64, 1.37, 0.13\}$ meses;
- $\boldsymbol{x}_\text{B} = \{0.48, 0.87, 1.79\}$ meses;

O chefe dos dois, Astolfo, suspeita que o tempo de falha seja, em média, 2 meses com desvio padrão de mais ou menos 1 mês. Para cada uma das amostras

- (i) Compute o estimador de Bayes θ sob perda quadrática;
- (ii) Estime θ por máxima verossimilhança.

Definição 23

*Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma distribuição indexada pelo parâmetro θ. Seja $T = r(X_1, X_2, \ldots, X_n)$ uma estatística. Dizemos que T é uma **estatística suficiente** para θ se e somente se*

$$f(X_1, X_2, \ldots, X_n \mid T = t, \theta) = f(X_1, X_2, \ldots, X_n \mid T = t, \theta'), \ \forall\, \theta, \theta' \in \Omega,$$

isto é, se a distribuição condicional da amostra dado o valor da estatística não depende de θ.

No exemplo anterior, tanto $\hat{\theta}_{\text{Bayes}}$ quanto $\hat{\theta}_{\text{EMV}}$ dependem de $X_1, X_2, \ldots, X_n$ apenas através de $r(X_1, X_2, \ldots, X_n) = T = \sum_{i=1}^{n} X_i$.

Uma observação importante

Definição 24 (Aleatorização auxiliar)

Suponha que T é suficiente para θ. O processo de simular $X'_1, \ldots, X'_n \mid T = r(X_1, X_2, \ldots, X_n)$ de modo que

$$f(X_1, X_2, \ldots, X_n \mid \theta) = f(X'_1, \ldots, X'_n \mid \theta), \forall \theta \in \Omega,$$

*é chamado de **aleatorização auxiliar** (em inglês, auxiliary randomisation).*

Observação 8 (A busca por bons estimadores)

Na busca por bons estimadores, estamos justificados em restringir a busca a funções de estatísticas suficientes.

Justificativa: Suponha que o estatístico A tem à sua disposição $X_1, X_2, \ldots, X_n$, enquanto B tem acesso somente a $T = r(X_1, X_2, \ldots, X_n)$. Se T é suficiente, B pode sempre fazer uma aleatorização auxiliar e gerar $X'_1, \ldots, X'_n$ com exatamente a mesma distribuição conjunta condicional a θ.

Teorema da fatorização (TF)

Teorema 14

Suponha que $X_1, X_2, \dots, X_n$ *perfazem uma amostra aleatória com f.d.p./f.m.p* $f(x \mid \theta)$, $\theta \in \Omega$. *Uma estatística* $T = r(X_1, X_2, \dots, X_n)$ *é suficiente para* θ *se, e somente se, para todo* $\mathbf{x} \in \mathcal{X}$ *e* $\theta \in \Omega$ *existem* u *e* v *não negativas tal que*

$$f_n(\mathbf{x} \mid \theta) = u(\mathbf{x})v[r(\mathbf{x}), \theta].$$

Prova: (Para v.a.s discretas). Para a "ida" notar que T é uma função determinística de X, ou seja, $\Pr(T = t \mid \boldsymbol{X} = \mathbf{x}, \theta) = 1$ e que só precisamos considerar $\mathbf{x} \in \{\mathbf{y} : r(\mathbf{y}) = t\}$. Para a "volta", mostrar que T suficiente implica que $\Pr(\boldsymbol{X} = \mathbf{x} \mid T = t, \theta)$ é função apenas de $\mathbf{x}$. Ver De Groot, Teorema 7.7.1 e Casella & Berger, Teorema 6.2.6.

O TF em ação

- Poisson;
- $f(x \mid \theta) = \theta x^{\theta-1}$, $x \in (0, 1)$ e $\theta > 0$;
- Normal;

O que acontece, por exemplo, no caso Normal com μ e σ^2 desconhecidos?

Definição 25 (Suficiência conjunta)

*Dizemos que um conjunto de estatísticas $\boldsymbol{T} = \{T_1, \ldots, T_k\}$ é **suficiente** (conjuntamente) se que a distribuição condicional conjunta de $X_1, X_2, \ldots, X_n$ dado $T_1 = t_1, \ldots, T_k = t_k$ não depende de θ.*

Observação 9 (TF para estatísticas suficientes conjuntas)

Para o caso de estatísticas suficientes conjuntas, vale um Teorema da fatorização:

$$f_n(\boldsymbol{x} \mid \theta) = u(\boldsymbol{x})v[r_1(\boldsymbol{x}), \ldots, r_k(\boldsymbol{x}), \theta].$$

Suficiência conjunta – exemplos

- Normal;
- Uniforme;

Observação 10 (Transformações biunívocas de estatísticas suficientes)

Se $\boldsymbol{T} = \{T_1, \ldots, T_k\}$ são estatísticas suficientes conjuntas, e $h : \mathcal{T} \to \mathbb{R}$ é um mapa inversível, então $\boldsymbol{T'} = h(\boldsymbol{T})$ também são suficientes conjuntas.

Suficiência mínima – motivação

Primeiro um exemplo motivador:

Definição 26 (Estatísticas de ordem)

Seja $\boldsymbol{X} = X_1, X_2, \ldots, X_n$ uma amostra aleatória. Dizemos que $Y_1, Y_2, \ldots, Y_n$ são ***estatísticas de ordem*** *se Y_1 é o menor valor de $\boldsymbol{X}$, Y_5 é o quinto menor valor e assim por diante.*

Teorema 15 (Estatísticas de ordem são suficientes conjuntas)

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória com f.d.p/f.m.p. $f(x \mid \theta)$. As estatísticas de ordem $Y_1, Y_2, \ldots, Y_n$ são suficientes conjuntas para θ.

Prova: Usar o fato de que a conjunta é o produto das marginais e a comutatividade da multiplicação em $\mathbb{R}$.

Definição 27

*Uma estatística T é dita **mínima suficiente** se T é suficiente e é função de qualquer outra estatística suficiente. Um vetor $\boldsymbol{T} = \{T_1, \ldots, T_k\}$ é dito **minimamente suficiente conjunto** se é função de qualquer outro vetor de estatísticas suficientes conjuntas.*

Observação 11 (Estatísticas de ordem são minimamente suficiente conjuntas no caso Cauchy)

$$f_n(\boldsymbol{x} \mid \theta) = \frac{1}{\pi^n \prod_{i=1}^{n} [1 + (x_i - \theta)^2]} \tag{18}$$

EMV e Bayes como estatísticas minimamente suficientes

Teorema 16

Se a função de verossimilhança admite fatorização como no Teorema 14, os estimadores de Bayes e de máxima verossimilhança são estatísticas minimamente suficientes.

Prova:

- EMV: notar que $f_x(\boldsymbol{x} \mid \theta) \propto v[r(\boldsymbol{x}), \theta]$;
- Bayes: escrever a perda esperada a posteriori explicitamente usando a verossimilhança na forma do TF.

Ver Teoremas 7.8.3 e 7.8.4 de De Groot.

O que aprendemos?

- Estatística suficiente;
  "Uma estatística T é suficiente para θ se $\Pr(\boldsymbol{X} = \boldsymbol{x} \mid T = x, \theta)$ não depende de θ."
- Teorema da fatorização;
  "Se T é suficiente para θ, podemos escrever a verossimilhança como o produto entre uma função que não depende de θ e uma função que só depende de $\boldsymbol{X}$ através de T."
- Os estimadores de Bayes e de máxima verossimilhança são minimamente suficientes.

Leitura recomendada

- De Groot seções 7.7 e 7.8;
- * Casella & Berger (2002), seção 6.2.
- **Exercícios recomendados**
  - De Groot.

    Seção 7.7: exercícios 4, 7, 13, 16;
    Seção 7.8: exercícios 3, 8, 12, 16.