Exercícios: Estimação intervalar e quantidades pivotais

Disciplina: Inferência Estatística (MSc)
Instrutor: Luiz Carvalho
Monitor: Isaque Pim

Novembro/2023

Motivação: Estimativas pontuais não são adequadas a todas as situações. Em várias aplicações queremos produzir estimativas que carreguem uma noção inerente de incerteza. Os intervalos de confiança são a ferramenta adequada para quantificar a incerteza sobre nossas inferências, ao mesmo tempo que vêm com garantias probabilísticas que fazem sentido dentro do paradigma clássico (frequentista). As quantidades pivotais ou pivôs são transformações (dependentes dos parâmetros) cuja distribuição não depende dos parâmetros de interesse, e são extremamente úteis na construção de intervalos de confiança, sejam eles exatos ou aproximados. Neste conjunto de exercícios vamos explorar as propriedades de pivôs e intervalos de confiança, tanto nas situações onde as quantidades são exatas quanto em situações de grandes amostras onde as propriedades se verificam aproximadamente.

Notação: Como convenção adotamos $\mathbb{R} = (-\infty, \infty)$, $\mathbb{R}_+ = (0, \infty)$ e $\mathbb{N} = \{1, 2, \ldots\}$.

Dos livros-texto:

a) KN, Ch9: 9.1, 9.2, 9.3, 9.8, 9.13, 9.21, 9.33(a), 9.34.

b) CB, Ch9: 14; **(PhD)**11, 30 e 31.

Extra:

1. Tome $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma distribuição normal com média $\mu \in \mathbb{R}$ e variância $\sigma^2 \in \mathbb{R}_+$, ambas desconhecidas. Defina $\bar{X}_n = \sum_{i=1}^n X_i/n$ e $S_n^2 = \sum_{i=1}^n (X_i - \bar{X}_n)/(n-1)$.

   Derive a distribuição (incluindo a densidade com respeito a Lebesgue) das seguintes quantidades:

   (a) Média padronizada:
   $$Z_n := \sqrt{n}\frac{\bar{X}_n - \mu}{\sigma};$$

   (b) Variância escalonada:
   $$V_n := \frac{(n-1)S_n^2}{\sigma^2};$$

(c) Quantidade t de Student:

$$T_n := \sqrt{n-1}\frac{Z_n}{\sqrt{V_n}}.$$

2. No exemplo anterior, mostre como construir um intervalo de confiança para σ^2 para (i) μ conhecida e (ii) μ desconhecida.

3. Tome $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma distribuição normal com média μ_1 e variância v_1 e $Y_1, Y_2, \ldots, Y_m$ uma amostra aleatória de uma distribuição normal com média μ_2 e variância v_2 com todos os parâmetros desconhecidos, i.e. $\theta = (\mu_1, v_1, \mu_2, v_2)$. Escreva

$$\boldsymbol{Z} = (X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_m)$$

e defina $\bar{X}_n$, $\bar{Y}_m$, S_x^2 e S_y^2 de forma análoga ao que é feito no item 1 desta lista.

Considere $\xi := g(\theta) = \mu_1 - \mu_2$ e assuma que $v_1 = v_2 = v$.

(a) Mostre que

$$t(Z, \xi) = \frac{(\bar{X}_n - \bar{Y}_m - \xi)/\sqrt{1/n + 1/m}}{\sqrt{\left[(n-1)S_x^2 + (m-1)S_y^2\right]/(n+m-2)}}$$

é quantidade pivotal e construa um intervalo de confiança (exato) de nível de confiança γ para ξ usando $t(\boldsymbol{Z}, \xi)$.

(b) Agora considere $\kappa = v_2/v_1$. Mostre que $R(\boldsymbol{Z}, \kappa) = S_y^2/(\kappa S_x^2)$ é quantidade pivotal e use-a para construir intervalo de confiança para κ.

4. (**Beta**) Considere uma amostra aleatória $X_1, X_2, \ldots, X_n$ de uma distribuição Beta com parâmetros θ (desconhecido) e $\beta = 1$, isto é com pdf comum

$$f_\theta(x) = \theta x^{\theta-1}\mathbb{I}(x \in (0,1)).$$

(a) Encontre a distribuição assintótica para o estimador de máxima verossimilhança de
   - $g_1(\theta) = \theta$;
   - $g_2(\theta) = \frac{\theta}{1+\theta}$.

(b) Encontre uma quantidade pivotal para θ e use-a para construir um intervalo de confiança de $(1-\alpha)\%$ para θ.

5. (**Beta escalonada**) Considere uma amostra aleatória $X_1, X_2, \ldots, X_n$ de uma distribuição com pdf comum

$$f_\theta(x) = \frac{\theta x^{\theta-1}}{\tau^\theta}\mathbb{I}(x \in (0,\tau)),$$

com $\tau > 0$ conhecida e $\theta > 0$ desconhecida.

(a) Encontre o estimador $\delta_{\text{EMV}}(\boldsymbol{X}_n)$ de máxima verossimilhança para θ;

(b) Mostre que

$$T(\boldsymbol{X}_n) := \frac{2n\theta}{\delta_{\text{EMV}}(\boldsymbol{X}_n)}$$

é uma quantidade pivotal e encontre sua distribuição. Depois use $T(\boldsymbol{X}_n)$ para construir um intervalo de confiança exato de $(1-\alpha)\times 100\%$ para θ;

(c) Construa um intervalo de confiança assintótico (aproximado) de $(1-\alpha)\times 100\%$ para θ.

(d) **[Computacional]** Compare os dois intervalos obtidos quanto à cobertura (o aproximado bate perto da cobertura nominal?) e largura dos intervalos obtidos.

6. **PhD**: Seja $\boldsymbol{X} \in \mathcal{X}$ uma amostra de uma distribuição na família paramétrica $\mathcal{P} = \{P_\theta : \theta \in \Omega\}$ e suponha que $T : \mathcal{X} \to \mathbb{R}$ é uma estatística com densidade $f_\theta(t)$. Tome $R(t,\theta)$ uma função monotônica de t para todo $\theta \in \Omega$. Mostre que se

$$f_\theta(t) = g\left(R(t,\theta)\right)\left|\frac{\partial}{\partial t}R(t,\theta)\right|$$

para alguma função mensurável g, então $R(T(\boldsymbol{X}),\theta)$ é quantidade pivotal. **Comentário: para uma aplicação, ver a questão 1 da P1 2022, que mostra como usar a cdf como pivô.**