Exercícios de Revisão: Teoria de probabilidade.

Disciplina: Inferência Estatística
Professor: Luiz Max de Carvalho

1 Desigualdades probabilísticas

As desigualdades probabilísticas são ferramentas de grande utilidade na prática estatística. São úteis, por exemplo, na demonstração de teoremas de convergência que veremos mais à frente no curso.

(a)

Teorema 1 (Desigualdade de Markov). *Seja X uma variável aleatória contínua não-negativa e $t > 0$. Então*

$$\Pr(X \geq t) \leq \frac{E[X^n]}{t^n}. \tag{1}$$

Demonstre o Teorema 1. *Dica*: use a linearidade e a monotonicidade da integral.

(b)

Teorema 2 (Desigualdade de Chebychev). *Seja Y uma variável aleatória com média $E[Y] =: \mu$ e variância $\operatorname{Var}(Y) =: \sigma^2$, ambas finitas. Mais uma vez, $t > 0$. Então*

$$\Pr(|Y - \mu| \geq t) \leq \frac{\operatorname{Var}(Y)}{t^2}. \tag{2}$$

Demonstre o Teorema 2.

2 Distribuições da média e variância amostrais.

Considere uma **amostra aleatória** $X_1, X_2, \ldots, X_n$, $n \in \mathbb{N}$ de variáveis aleatórias de uma mesma distribuição com média $E[X_i] = \mu$ e variância $\mathrm{Var}(X_i) = \sigma^2$.

Definição 1 (Média amostral)**.** *A média amostral de $X_1, X_2, \ldots, X_n$ é*

$$\bar{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i. \tag{3}$$

(a) Demonstre o seguinte resultado:

Teorema 3 (Média e variância em uma amostra i.i.d.)**.** *Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias independentes e identicamente distribuídas, com média μ e variância σ^2. Temos que (i) $E[\bar{X}_n] = \mu$ e (ii) $\mathrm{Var}(\bar{X}_n) = \frac{\sigma^2}{n}$.*

(b) Comente sobre como as premissas de identidade de distribuição e independência são utilizadas em sua demonstração do item anterior.

3 Lei (fraca) dos grandes números e Teorema central do limite

As leis dos grandes números são resultados fundamentais da teoria de probabilidade, nos permitindo fazer afirmações sobre o comportamento de processos estocásticos à medida que o número de observações aumenta. Da mesma forma, os teoremas centrais do limite[1] tratam da distribuição **assintótica**[2] de certas variáveis aleatórias.

Primeiro, uma definição.

Definição 2 (Convergência em probabilidade)**.** *Dizemos que uma sequência de variáveis aleatórias converge em probabilidade para b se, para todo $\epsilon > 0$, temos*

$$\lim_{n\to\infty} \Pr\left(|Z_n - b| < \epsilon\right) = 1.$$

Neste caso, escrevemos $Z_n \xrightarrow{p} b$.

(a) Mostre que o seguinte teorema vale:

[1] Sim, existem vários.

[2] Isto é, à medida que o número de observações $n \to \infty$.

Teorema 4 (Lei Fraca dos Grandes Números)**.** *Sejam* $X_1, X_2, \ldots, X_n$ *variáveis aleatórias independentes e identicamente distribuídas, com média* μ *e variância* σ^2*. Então*

$$\bar{X}_n \xrightarrow{p} \mu.$$

(b)

Teorema 5 (Teorema Central do Limite (Lindeberg e Lévy))**.** *Sejam* $X_1, X_2, \ldots, X_n$ *variáveis aleatórias independentes e identicamente distribuídas, com média* μ *e variância* σ^2*. Então, para cada* x*, temos*

$$\lim_{n\to\infty} \Pr\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq x\right) = \Phi(x),$$

onde

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_0^x \exp\left(-\frac{t^2}{2}\right) dt,$$

é a função de distribuição (cumulativa) normal padrão.

Mostre que o Teorema 5 vale. *Dica*: Ver Casella & Berger (2002), página 237.

4 Aplicações.

Para fixar o conteúdo, agora veremos algumas aplicações dos conceitos trabalhados e dos resultados demonstrados.

(a) Suponha que uma moeda justa é lançada n vezes. Seja X_i a variável aleatória que é 1 se o i-ésimo lançamento dá cara e 0 caso contrário. Quantos lançamentos devemos fazer para que

$$\Pr(0.4 \leq \bar{X}_n \leq 0.6) \geq 0.7\,?$$

Responda à pergunta utilizando (i) a desigualdade de Chebychev e (ii) probabilidades binomiais, obtidas atráves de uma tabela ou programa de computador.

(b) Compare os resultados do item anterior e discuta se a aproximação por Chebychev é boa. Que consequências práticas (em termos de custo de amostragem, por exemplo) haveria em usar o resultado aproximado? O que isso diz sobre a desigualdade de Chebychev?

(c) Suponha que $X_1, \ldots, X_{12}$ são variáveis aleatórias independentes com distribuição uniforme em $(0,1)$. Defina

$$p := \Pr\left(\left|\bar{X}_n - \frac{1}{2}\right| \leq 0.1\right).$$

Determine quanto vale p utilizando: (i) o teorema central do limite (TCL) e (ii) a expressão exata.

(d) Compare os resultados obtidos no item anterior e discuta se a aproximação utilizando o TCL é boa.