Tratamento Estatístico de Dados em Física Experimental

Prof. Zwinglio Guimarães

2º semestre de 2017

Tópico 2 – 2ª semana – 22 e 24/08

A função densidade de probabilidade (revisão)

- A função densidade de probabilidade rege a probabilidade de se obter um dado experimental no intervalo $[x_a, x_b]$:

$$P(x \in [x_a, x_b]\,) = \int_{x_a}^{x_b} f(x)dx$$

onde $f(x)$ é a função densidade de probabilidade.

- $f(x)$ tem dimensão $[u.\,x]^{-1}$
- Se $f(x)$ é uma função densidade de probabilidade, então:

$$f(x) \geq 0$$

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

Principais parâmetros de uma função densidade de probabilidade (revisão)

- **O valor médio (verdadeiro), x_0, que é obtido por:**

$$x_0 = \langle\, x \,\rangle = \int_{-\infty}^{+\infty} x\, f(x)\, dx$$

- **E o desvio-padrão (verdadeiro), σ_0, que é obtido a partir de:**

$$\sigma_0^2 = \langle\, (x - x_0)^2 \,\rangle = \int_{-\infty}^{+\infty} (x - x_0)^2\, f(x)\, dx$$

➢ **Consideração prática:** $\langle\, (x - x_0)^2 \,\rangle = \langle\, x^2 \,\rangle - x_0^2$

Exemplo 1

- **Função densidade de probabilidade do erro devido a arredondamentos de amplitude *L***

$$f(x) = \begin{cases} A & \text{se} \quad |x| \leq \frac{L}{2} \\ 0 & \text{caso contrário} \end{cases}$$

- $A = {}^{1}/_{L}$
- $x_0 = 0$
- $\sigma_0 = {}^{L}/_{\sqrt{12}}$
- $P(x \in [x_0 - \sigma_0, x_0 + \sigma_0]) = {}^{\sqrt{3}}/_{3} \cong 0{,}577$

Exemplo 3

- **Função densidade de probabilidade do intervalo de tempo entre dois eventos aleatórios independentes**

$$f(x) = \begin{cases} A\, e^{\left(\frac{-x}{L}\right)} & \text{se} \quad x \geq 0 \\ 0 & \text{caso contrário} \end{cases}$$

- $A = {}^{1}/_{L}$
- $x_0 = L$
- $\sigma_0 = L$

Exercício da última aula

- **Calcular** $\boldsymbol{A}$**,** $\boldsymbol{\sigma_0}$ ***e*** $\boldsymbol{P(x \in [x_0 - \sigma_0, x_0 + \sigma_0])}$ ***para:***

$$f(x) = \begin{cases} A\left(1 - \left|\frac{x}{L}\right|^n\right) & \text{se} \quad |x| \leq L \\ 0 & \text{caso contrário} \end{cases}$$

- n=1 (triangular): $A = \frac{1}{L}$ $\sigma = \frac{L}{\sqrt{6}}$ $P \cong 0{,}650$
- n=2 (parabólica): $A = \frac{3}{4\,L}$ $\sigma = \frac{L}{\sqrt{5}}$ $P \cong 0{,}626$
- n=3 (cúbica): $A = \frac{2}{3\,L}$ $\sigma = \frac{L\sqrt{2}}{3}$ $P \cong 0{,}612$
- n=4 (4º grau): $A = \frac{5}{8\,L}$ $\sigma = \frac{L\sqrt{5}}{\sqrt{21}}$ $P \cong 0{,}603$

Parâmetros de localização de funções densidade de probabilidade

- **Moda, $\boldsymbol{x_{mp}}$**: valor de x em que $f(x)$ é máximo

- **Mediana, $\boldsymbol{x_M}$**: valor de x tal que a probabilidade de se obter um dado com $x \leq x_M$ é igual ao de $x \geq x_M$. Ou seja:

$$\int_{-\infty}^{x_M} f(x)dx = \int_{x_M}^{+\infty} f(x)dx = 0.5$$

Momentos de funções densidade de probabilidade

Os **momentos de ordem n**:

$$\mu_n = \langle x^n \rangle = \int_{-\infty}^{+\infty} x^n \, f(x) \, dx$$

Os **momentos centrais de ordem n**:

$$\mu_n^0 = \langle (x - x_0)^n \rangle = \int_{-\infty}^{+\infty} (x - x_0)^n \, f(x) \, dx$$

Parâmetros úteis para caracterizar a forma de funções densidade de probabilidade: Momentos centrais normalizados

- Obliquidade ou Assimetria ("***skewness***"):

$$S = \frac{\mu_3^0}{\sigma^3} = \frac{\langle\, (x - x_0)^3 \,\rangle}{\sigma^3}$$

- Curtose ("***kurtosis***"):

$$K = \frac{\mu_4^0}{\sigma^4} = \frac{\langle\, (x - x_0)^4 \,\rangle}{\sigma^4}$$

A função de distribuição acumulada

- A função de distribuição acumulada, $g(x)$, corresponde à integral da função densidade de probabilidade, $f(x)$, desde $-\infty$ até x:

$$g(x) = \int_{-\infty}^{x} f(x')dx'$$

Obs: os pacotes estatísticos costumam conter funções de distribuição acumulada para as funções densidade de probabilidade mais comuns. Usadas em testes estatísticos para avaliar se o valor de x não é muito pequeno ($g(x)\sim 0$) nem muito grande ($g(x)\sim 1$).

A função densidade de probabilidade gaussiana

- A função densidade de probabilidade gaussiana é dada por:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-x_0}{\sigma}\right)^2}$$

"...lei em que todos creem. Os experimentais pensam que é um teorema matemático e os matemáticos, que é um fato experimental.“

(citado em J.-P. Benzécri, *Histoire et Prehistoire de l'Analyse des Données*. Paris, Bordas, 1982)

Resultados importantes sobre a gaussiana (1)

- Uma mudança de variáveis muito comum consiste em escrever a gaussiana em termos do erro normalizado:

$$z = \left(\frac{x - x_0}{\sigma}\right)$$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

A função $f(z)$ é conhecida como **Gaussiana Padrão**.

Resultados importantes sobre a gaussiana (2) - Intervalos de confiança

- Alguns intervalos de confiança (I.C.) correspondentes à integral da gaussiana padrão são muito conhecidos:

   a) $P(|z| \leq 1) = \int_{-1}^{1} f(z)dz = 0{,}683 \rightarrow$ I.C. de 68,3 %

   b) $P(|z| \leq 2) = \int_{-2}^{2} f(z)dz = 0{,}954 \rightarrow$ I.C. de 95,4 %

   c) $P(|z| \leq 3) = \int_{-3}^{3} f(z)dz = 0{,}997 \rightarrow$ I.C. de 99,7 %

Usualmente esses resultados são apresentados em tabelas de integrais da gaussiana padrão entre 0 e z: $I(z) = \int_{0}^{z} f(z')dz'$

Obs: É comum haver uma função que corresponde a integral da gaussiana padrão entre $-\infty$ e z: $C(z) = \int_{-\infty}^{z} f(z')dz'$