Tratamento Estatístico de Dados em Física Experimental

Prof. Zwinglio Guimarães

2º semestre de 2016

Tópico 5

A função de probabilidade binomial (re-revisão)

- A distribuição do número de ocorrências, n, em N **medições independentes com probabilidade individual fixa,** p, segue uma binomial:

$$P_{N,p}(n) = \frac{N!}{(N-n)!\,n!} p^n (1-p)^{N-n}$$

- A média do número de ocorrências é $n_0 = Np$

$$\langle n \rangle = \sum_{n=0}^{N} n \frac{N!}{(N-n)!\,n!} p^n (1-p)^{N-n} = Np$$

- E o desvio-padrão do número de ocorrências é:

$$\sigma_n = \sqrt{Np(1-p)}$$

A função de probabilidade de Poisson (revisão)

- Probabilidade de ocorrência de n sucessos quando o número de tentativas, N, é muito grande, mas a probabilidade de sucesso em cada tentativa, p, é muito baixa:

$$P_a(n) = \frac{a^n\, e^{-a}}{n!}$$

- Corresponde ao caso limite da Binomial com $N \to \infty$ e $p \to 0$, mas com o produto $a = Np$ mantido constante:

$$P_a(n) = \lim_{\substack{N\to\infty \\ Np=a}} \frac{N!}{(N-n)!\, n!} p^n (1-p)^{N-n}$$

Resultados importantes da Poisson (revisão)

- A Poisson é normalizada:

$$\sum_{n=0}^{\infty} \frac{a^n\, e^{-a}}{n!} = \left(\sum_{n=0}^{\infty} \frac{a^n}{n!}\right) e^{-a} = (e^a)e^{-a} = 1$$

- A média do número de ocorrências é:

$$\langle n \rangle = \sum_{n=0}^{\infty} n \frac{a^n\, e^{-a}}{n!} = a$$

- O desvio-padrão do número de ocorrências é:

$$\sigma_n = \sqrt{\langle n^2 \rangle - \langle n \rangle^2} = \sqrt{a}$$

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

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

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

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

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

O Teorema Central do Limite

- A soma, S, de variáveis aleatórias x_i, cada uma obedecendo à uma função densidade de probabilidade $f_i(x)$ com médias (verdadeiras) μ_i e variâncias σ_i^2 finitas, tende a uma gaussiana quando o número de variáveis N tente ao infinito.
- A variável S terá média (verdadeira) igual à soma das médias e a variância é igual à soma das variâncias:

$$S = \sum_{i=1}^{N} x_i$$

$$\mu_0 = \langle S \rangle = \sum_{i=1}^{N} \mu_i$$

$$\sigma_0^2 = \sum_{i=1}^{N} \sigma_i^2$$

O Teorema Central do Limite (outra forma de escrever)

- A variável soma normalizada:

$$S^* = \frac{S - \mu_0}{\sigma_0} = \frac{\sum_{i=1}^{N} x_i - \sum_{i=1}^{N} \mu_i}{\sqrt{\sum_{i=1}^{N} \sigma_i^2}}$$

$$S = \sum_{i=1}^{N} x_i$$

$$\mu_0 = \sum_{i=1}^{N} \mu_i$$

$$\sigma_0^2 = \sum_{i=1}^{N} \sigma_i^2$$

de N variáveis aleatórias x_i, que obedecem funções densidade de probabilidade $f_i(x)$ com médias (verdadeiras) μ_i e variâncias σ_i^2 finitas tem a seguinte propriedade:

$$\lim_{N \to \infty} P(S^* \leq a) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\, dz$$