Tratamento Estatístico de Dados em Física Experimental

Prof. Zwinglio Guimarães

2º semestre de 2016

Tópico 3 – Função de Propabilidade Binomial

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, que é obtido por:**

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

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

Outros parâmetros de uma f.d.p. (revisão)

- **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 uma f.d.p. (revisão)

O **momento de ordem n, $\boldsymbol{\mu_n}$**, é dado por:

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

O **momento central de ordem n, $\boldsymbol{\mu_n^0}$** é dado por:

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

Outros parâmetros úteis para caracterizar funções densidade de probabilidade: Momentos centrais normalizados (revisão)

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

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

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

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

Exercício 4 (correção)

1) Considere a função densidade de probabilidade, $f(x)$, definida como ($L > 0$):

$$f(x) = A\,exp\left(\frac{-|x|}{L}\right)$$

Faça uma rotina para gerar dados de acordo com essa função densidade de probabilidade. Em seguida, gere diversos conjuntos de experimentos numéricos com $N = 500$ dados cada, para o caso com $L = 1$, e determine (estimando todas as incertezas de forma numérica):

a) O desvio-padrão amostral, s_x, do primeiro conjunto gerado (com sua respectiva incerteza);

b) A média, x_m, do primeiro conjunto gerado (com sua respectiva incerteza);

c) A mediana, x_M, do primeiro conjunto gerado (com sua respectiva incerteza);

d) A curtose (*kurtosis*), $K_x = \frac{\mu_4^0}{\sigma^4}$, do primeiro conjunto gerado (com sua respectiva incerteza);

Exercício 4 (correção)

1) Considere a função densidade de probabilidade, $f(x)$, definida como ($L > 0$):

$$f(x) = A\, exp\left(\frac{-|x|}{L}\right)$$

- Desvio-padrão: $\sigma_0 = \sqrt{2} \cong 1{,}41$ $\qquad s_s \cong 0{,}071$
- Média: $x_0 = 0$ $\qquad s_{x_m} \cong 0{,}063$
- Mediana: $\langle x_M \rangle = 0$ $\qquad s_{x_M} \cong 0{,}046$
- Curtose: $K = 6$ $\qquad s_K \cong 1{,}3$

Características notáveis dessa função densidade de probabilidade:

- **A mediana é um estimador mais preciso que a media**
- **A curtose é bem maior que a da gaussiana ($K_{Gaussiana} = 3$)**

Exercício 4 (correção)

2) Responda os itens (**a**)-(**d**) acima para a função densidade de probabilidade definida abaixo com $n = 5$:

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

- Desvio-padrão: $\sigma_0 = \frac{1}{2}$ $\quad s_S \cong 0{,}011$
- Média: $x_0 = 0$ $\quad s_{x_m} \cong 0{,}022$
- Mediana: $\langle x_M \rangle = 0$ $\quad s_{x_M} \cong 0{,}037$
- Curtose: $K = \frac{48}{25} = 1{,}92$ $\quad s_K \cong 0{,}059$

Característica notável dessa função densidade de probabilidade:

- **A curtose é bem menor que a da gaussiana ($K_{Gaussiana} = 3$)**

Exercício 4 (correção)

3) Repita a questão acima para o caso em que $n = \frac{1}{2}$.

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

- Desvio-padrão: $\sigma_0 = \sqrt{7}/_7 \cong 0{,}378$ $\qquad s_s \cong 0{,}011$
- Média: $x_0 = 0$ $\qquad s_{x_m} \cong 0{,}017$
- Mediana: $\langle x_M \rangle = 0$ $\qquad s_{x_M} \cong 0{,}017$
- Curtose: $K = \frac{147}{55} \cong 2{,}67$ $\qquad s_K \cong 0{,}12$

Característica notável dessa função densidade de probabilidade:

- **A precisão da média e da mediana são equivalentes**

No caso de variáveis discretas

$$\int \phi(x) f(x) dx \longrightarrow \sum_i \phi(X_i)\ F(X_i)$$

onde $F(X)$ é a função de probabilidade de obter em uma medição o valor X.

- No caso de comparações com histogramas as expressões para variáveis discretas são usadas juntamente com a aproximação

$$F(X) \cong f(x)\Delta x$$

onde Δx é a largura de cada canal do histograma.

A função de probabilidade binomial

- Distribuição do número de ocorrências em N medições independentes com probabilidade individual p

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

- Essa distribuição é chamada de binomial por causa da semelhança com o binômio de Newton

$$(a+b)^N = \sum_{n=0}^{N} \frac{N!}{(N-n)!\,n!} a^n b^{N-n}$$

Resultados importantes da binomial

- A binomial é normalizada:

$$\sum_{n=0}^{N} \frac{N!}{(N-n)!\,n!} p^n (1-p)^{N-n} = [p + (1-p)]^N = 1$$

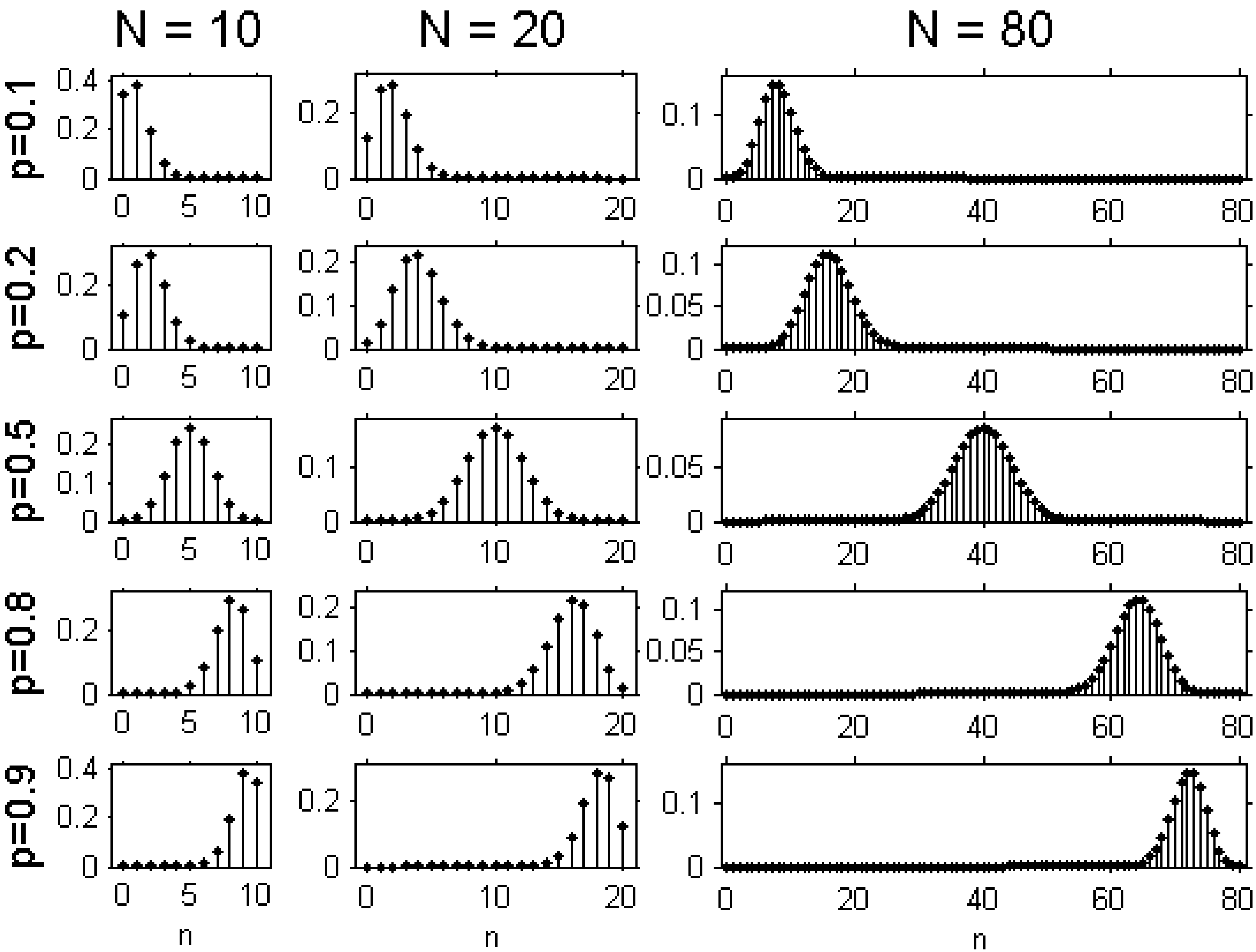
- A média do número de ocorrências é $n_0 = Np$:

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

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

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

Exemplos de binomial



Descreva, de forma qualitativa, como o desvio-padrão e a Assimetria dependem do número de tentativas, N, e da probabilidade de sucesso individual, p.