Termo-Estatística – Licenciatura: *3ª Aula* (06/03/2013)

Prof. Alvaro Vannucci

RELEMBRANDO

► Probabilidade condicional:

$$P(X/Y) = \frac{P(X \cap Y)}{P(Y)}$$

► Se dois eventos (**A** e **B**) forem independentes:

$$P(A \cap B) = P(A) \cdot P(B)$$

► Dados dois eventos independentes **A** e **B**:

(i) Probabilidade do evento **A** **ou** **B** ocorrer será: $P = P(A) + P(B)$

(ii) Probabilidade do evento **A** **e** **B** ocorrer será: $P = P(A) \cdot P(B)$

► Sendo *n* o número de objetos **indistinguíveis** (ou o número de medidas independentes de um sistema) que queremos alocar em *p* **posições disponíveis** (ou *p* medidas que forneceram um certo resultado "binário"), então as **combinações possíveis** entre estes elementos (ou a forma como as medidas realizadas e que dão resultados que satisfaçam uma certa exigência) será:

$$C_p^n = \frac{n!}{p!(n-p)!}$$

DISTRIBUIÇÕES DISCRETAS E CONTÍNUAS – FUNÇÕES DE DISTRIBUIÇÃO

► Como fazer para se determinar a largura de uma mesa?

► Certamente é sempre possível utilizar uma **régua** (ou trena) e efetuar a medida apropriada.

► Todos concordamos, porém, que um **valor mais confiável** se obteria **realizando algumas medidas** e, depois, **tirando a média**.

► Mas, mesmo tendo-se o valor médio de várias medidas, não se consegue saber, por exemplo, a **precisão** e os cuidados tomados pelo marceneiro, na confecção da mesa.

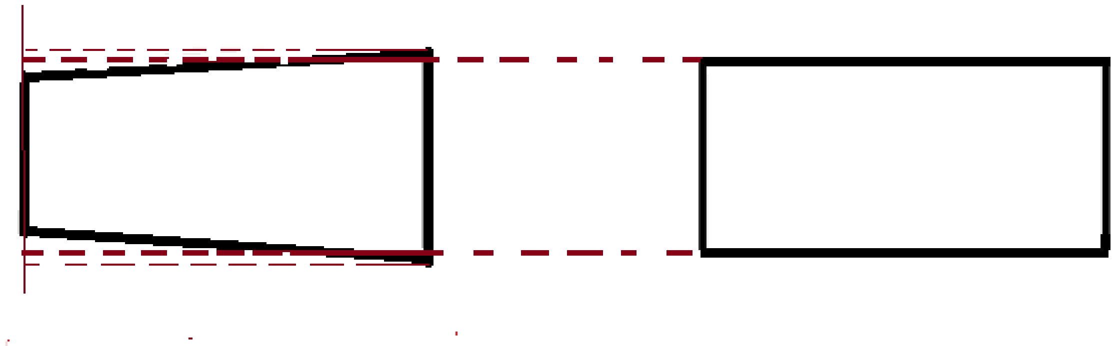

Figura 1: Duas mesas não idênticas.

► Note que as larguras das mesas acima podem ter os **mesmos valores** - na **média**. Só sabendo estes valores, porém, podemos nos **equivocar** e dizer que as **mesas são idênticas**!

► Daí que fazer um gráfico com as medidas realizadas, ou seja, construir uma **"curva de distribuição"** será uma forma **muito mais adequada** de se estudar um sistema, do que apenas saber os valores médios de suas grandezas.

► Analisemos, por exemplo, o jogo com 2 dados, em que é efetuada **a soma dos resultados** após cada jogada simultânea.

► Sabemos que a probabilidade de sair um número, em qualquer um deles (lembrando que são eventos independentes) é 1/6.

► Se pegarmos 2 dados e, antes de jogá-los, pedir aos alunos de uma classe, um a um, que tentem **adivinhar** o resultado da jogada que será feita, a **grande maioria** irá escolher um número entre **2 e 12**, sem atentar para o fato que alguns resultado têm maior probabilidade de sair do que outros!

► Para vermos isto, vamos calcular as probabilidades pertinentes e preencher a tabela abaixo, lembrando a necessidade de calcular a probabilidade de sair certo resultado em um dado **E** no outro:

► valor **2**: (**1** no 1º dado e **1** no 2º dado).

$$\Rightarrow P_2 = \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{36} = 0,028 \text{ } ou \text{ } 2,8\%$$

► valor **3**: (**2** no 1º dado e **1** no 2º dado **ou** **1** no 1º dado e **2** no 2º dado)

$$\Rightarrow P_3 = \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} = \frac{2}{36} = 0,056 \text{ } ou \text{ } 5,6\%$$

► valor **4**: (**3** no 1º dado e **1** no 2º dado **ou** **1** no 1º dado e **3** no 2º dado **ou** **2** no 1º dado e **2** no 2º dado).

$$\Rightarrow P_4 = \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} = \frac{3}{36} = 0,083 \text{ } ou \text{ } 8,3\%$$

► valor **5**: (**1** no 1º dado e **4** no 2º dado **ou** **4** no 1º dado e **1** no 2º dado **ou** **2** no 1º dado e **3** no 2º dado **ou** **3** no 1º dado e **2** no 2º dado).

$$\Rightarrow P_5 = \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} = \frac{4}{36} = 0,111 \text{ } ou \text{ } 11,1\%$$

► valor **6**: (**5** no 1º dado e **1** no 2º dado **ou** **1** no 1º dado e **5** no 2º dado **ou** **3** no 1º dado e **3** no 2º dado **ou** **2** no 1º dado e **4** no 2º dado **ou** **4** no 1º dado e **2** no 2º dado).

$$\Rightarrow P_6 = \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} = \frac{5}{36} = 0,139 \text{ } ou \text{ } 13,9\%$$

► valor **7**: (**6** no 1º dado e **1** no 2º dado **ou** **1** no 1º dado e **6** no 2º dado **ou** **5** no 1º dado e **2** no 2º dado **ou** **2** no 1º dado e **5** no 2º dado **ou** **4** no 1º dado e **3** no 2º dado **ou** **3** no 1º dado e **4** no 2º dado).

$$\Rightarrow P_7 = \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} = \frac{6}{36} = 0,167 \text{ } ou \text{ } 16,7\%$$

► valor **8**: (**6** no 1º dado e **2** no 2º dado **ou** **2** no 1º dado e **6** no 2º dado **ou** **5** no 1º dado e **3** no 2º dado **ou** **3** no 1º dado e **5** no 2º dado **ou** **4** no 1º dado e **4** no 2º dado).

$$\Rightarrow P_8 = \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} = \frac{5}{36} = P_6$$, e os valores ***irão se repetir***!

$$\Rightarrow P_9 = P_5$$
$$\Rightarrow P_{10} = P_4$$
$$\Rightarrow P_{11} = P_3$$
$$\Rightarrow P_{12} = P_2$$

resultado (evento)	**2**	**3**	**4**	**5**	**6**	**7**
probabilidade	**2,8%**	**5,6%**	**8,3%**	**11,1%**	**13,9%**	**16,7%**
resultado (evento)	**8**	**9**	**10**	**11**	**12**	
probabilidade	**13,9%**	**11,1%**	**8,3%**	**5,6%**	**2,8%**	

Tabela 1: Probabilidades de sair certo resultado na soma dos dados.

► Note que, neste experimento, a construção do **espaço amostral** correspondente resultaria em **11 pontos** que **nada informariam** sobre as **diferentes probabilidades de cada evento**.

► Daí a conveniência da representação através de um gráfico, em casos como este.

► Antes, porém, vejamos o seguinte experimento: uma câmara com gás possui um orifício por onde escapam as moléculas que, após atravessarem uma região de vácuo, atingem diferentes partes de uma fita cintilante (situação em ***1D***).

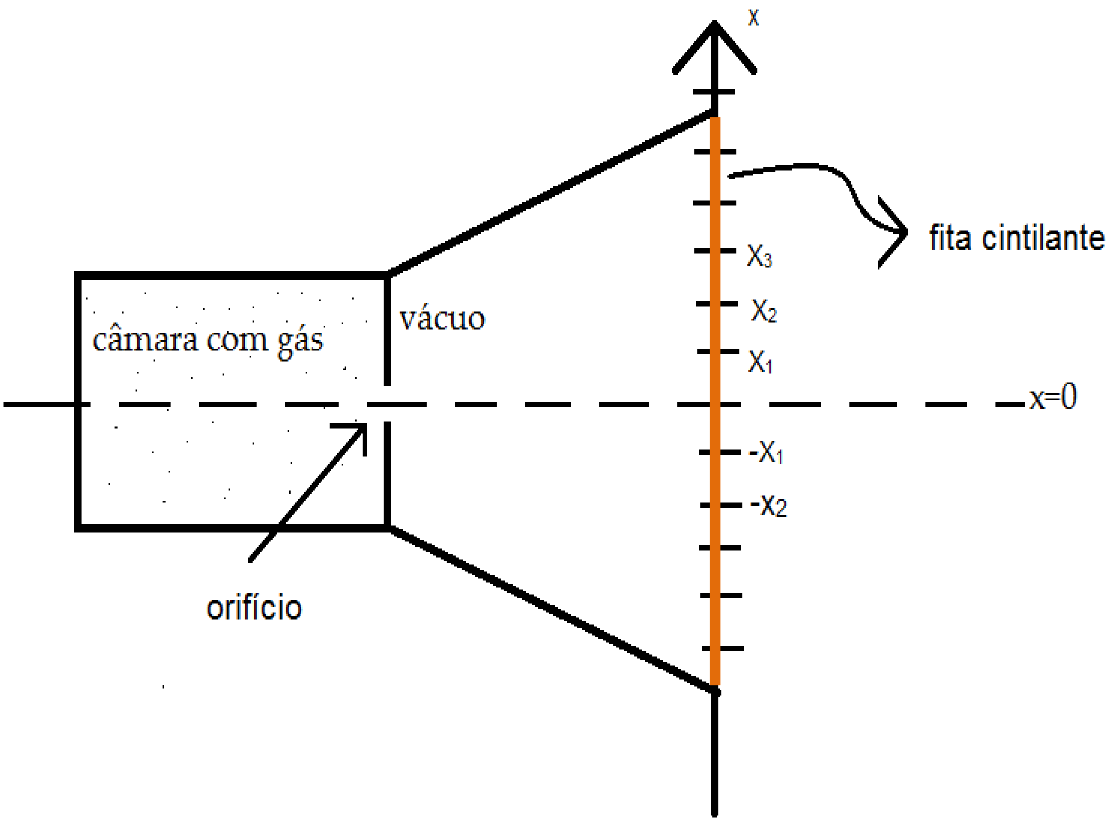

Figura 2: Representação do experimento.

► Para avaliar como as moléculas se distribuem espacialmente ao longo de *x*, ao colidirem com a fita, vamos dividi-la em intervalos Δx = 1cm, por exemplo, e contar quantas interações em um intervalo de tempo Δt ocorreram em cada trecho, e dividir pelo número total de moléculas que saíram do orifício naquele intervalo Δt (para termos a probabilidade de cada evento).

► O resultado pode ser apresentado através de um **histograma**:

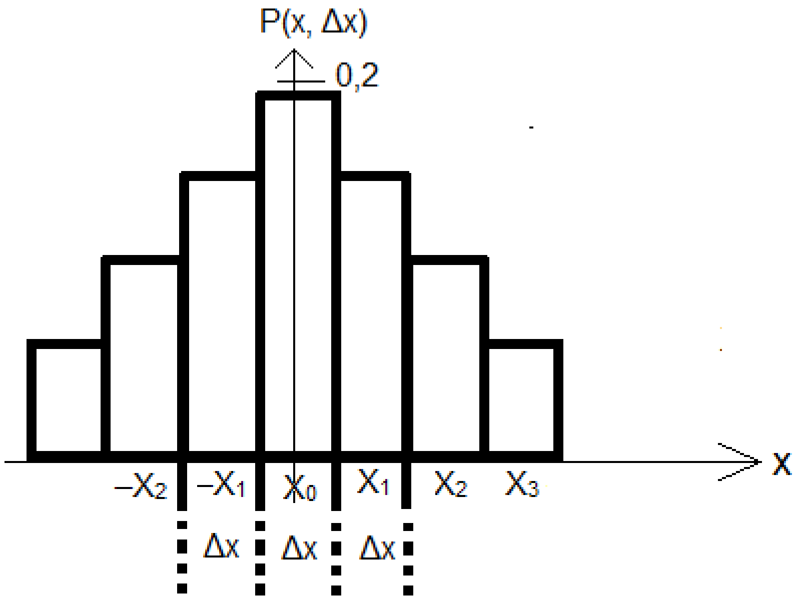

Figura 3: Representação dos dados (discretos) através de um histograma.

► Como esperado, a probabilidade de impacto das moléculas com a fita **vai diminuindo para regiões mais afastadas** do ponto $x = 0$.

► Como a altura de cada barra representa o número de moléculas que atingem cada faixa da tira (dividido pelo número total de partículas N que saíram do orifício em Δt), podemos então perceber que cada barra do histograma, que possui certa área, indica a **probabilidade** (maior ou menor) de uma molécula **atingir certa faixa** Δx da fita.

► Observe que se desejarmos **maior precisão** para o experimento, em termos de posição, podemos dividir a fita em um **maior número de faixas** de largura Δx **menor** (definindo uma área menor para cada barra).

► Por exemplo, escolhendo Δx duas vezes menor, temos o histograma:

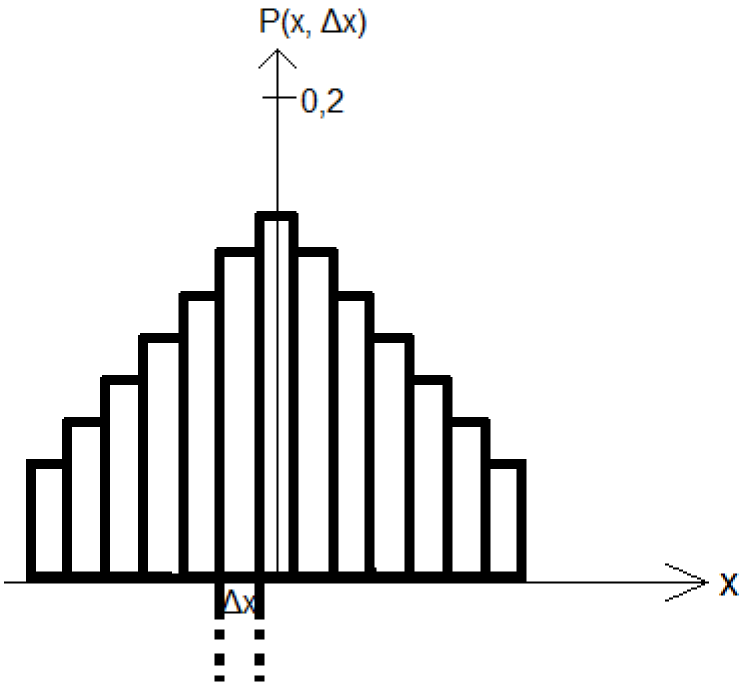

Figura 4: Representação do histograma escolhendo intervalo Δx menores.

► Note que **<u>as alturas das barras</u>** (<u>e suas áreas</u>) **<u>diminuem</u>**! Ou seja, o número de átomos que atingem cada faixa é menor e, portanto, a **<u>probabilidade</u>** de atingir determinada faixa é **<u>menor</u>**.

► Assim, à medida que **<u>diminuímos</u>** Δx, o histograma vai tornando-se **<u>cada vez mais *'liso'*</u>** de forma que cada probabilidade P(x, Δx), consequentemente, vai se tornando **<u>cada vez menor</u>**.

► No **<u>limite</u>** em que $\Delta x \to dx$ (infinitesimal) o **<u>histograma toma a forma de uma curva contínua</u>**, que designaremos por ***f(x)***:

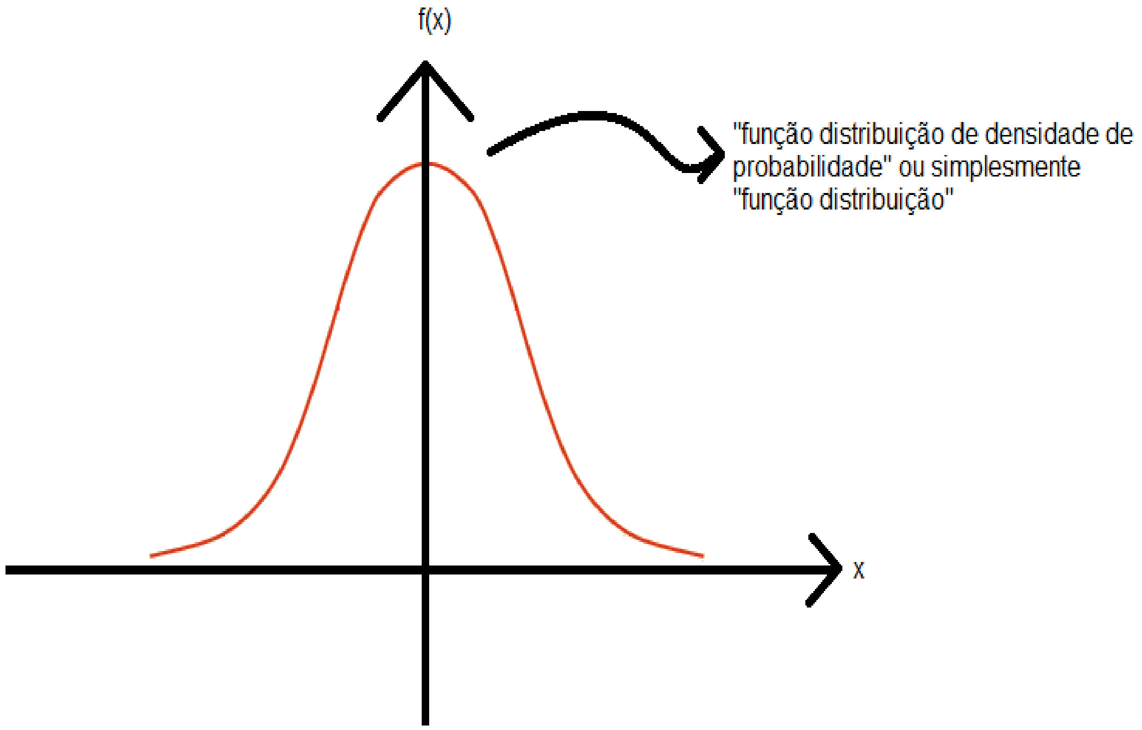

Figura 5: Representação da função distribuição.

► Vejamos agora um ponto importante a respeito da distribuição contínua de probabilidade: supondo a distribuição do gráfico acima sendo **simétrica** em relação a x = 0, qual é a probabilidade da molécula atingir a tira cintilante em valores de *x* **positivos**? E para valores de *x* **quaisquer**?

► Intuitivamente podemos afirmar que será **1/2 = 0,5** no primeiro caso e **1,0** no segundo.

► Ou seja, a **probabilidade** está relacionada com a **área sob f(x)**, no intervalo Δx de interesse:

$$P_{\Delta x}(x) = \int_{x_1}^{x_2} f(x)dx$$

► De forma que:

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

► Observe que, então, o **<u>cálculo da probabilidade</u>** a partir de uma **<u>“função de distribuição de densidade de probabilidade”</u>** envolve um **<u>cálculo de área</u>**.

► **<u>Quanto maior</u>** Δx, **<u>maior a área</u>** e **<u>maior a probabilidade</u>** (impondo que $P \leq 1$, sempre).

► Quando $\Delta x \to 0 \Rightarrow P \to 0$

► Por esta linha de raciocínio é que nós percebemos a razão da curva chamar-se **“<u>distribuição de densidade de probabilidade</u>”**.

► O termo **“<u>densidade</u>”** deve-se ao fato que, para se obter a probabilidade, **<u>deve-se somar</u>**, para **<u>cada intervalo dx</u>**, o **<u>produto de f(x) por dx</u>** (de forma semelhante ao cálculo da **<u>massa</u>**, por exemplo, no qual multiplicamos a densidade volumétrica de massa ρ pelo volume: $m = \rho\,\Delta V$).

►Resumindo:

(i) $P(x, dx) = f(x)dx \equiv$ probabilidade infinitesimal

(ii) $P_{\Delta x}(x) = \int_{x_1}^{x_2} f(x)dx$

(iii) $\int_{\substack{espaço \\ todo}} f(x)\,dx = 1$; para f(x) ser normalizada