

# Material Teórico - Módulo de INTRODUÇÃO À INFERÊNCIA ESTATÍSTICA

## A Curva Normal

Segundo Ano do Ensino Médio

Prof. Francisco Bruno Holanda  
Prof. Antonio Caminha Muniz Neto

12 de maio de 2020



PORTAL DA  
MATEMÁTICA  
OBMEP

# 1 Introdução

Na primeira aula deste módulo, aprendemos de maneira intuitiva que a distribuição (histograma) da média amostral “aproxima-se” de uma curva em formato de sino que é conhecida como curva normal. No presente material, definiremos essa curva de maneira explícita e teceremos comentários sobre algumas de suas propriedades. Porém, antes disso, utilizemos os exemplos a seguir para relembrar alguns fatos sobre probabilidade.

**Exemplo 1 (Caso discreto).** Quando jogamos um dado, a pontuação obtida é sempre um elemento do conjunto  $\{1, 2, 3, 4, 5, 6\}$ . Dessa forma, sempre podemos definir a probabilidade de uma determinada pontuação ser a sorteada. Por exemplo, no caso do dado ser justo, temos

$$P(1) = P(2) = P(3) = \dots = P(6) = \frac{1}{6}.$$

**Exemplo 2 (Caso contínuo).** Considere o segmento de reta formado pelos números reais do intervalo  $[0, 1]$ . Suponha que um ponto seja escolhido aleatoriamente (de maneira uniforme) neste intervalo. Qual é a probabilidade deste ponto ser  $\frac{1}{2}$ ?

A resposta para esta pergunta é 0, o que pode parecer estranho, já que existe a possibilidade do número escolhido ser exatamente  $\frac{1}{2}$ . Esta “estranheza” ocorre nesse exemplo por haver uma infinidade de valores possíveis, enquanto no exemplo anterior o número de possíveis valores é finito.

Por outro lado, caso a pergunta fosse: Qual é a probabilidade de escolhermos um ponto no intervalo  $[\frac{1}{3}, \frac{2}{3}]$ ?, a resposta seria mais naturalmente aceita. Esta probabilidade é  $\frac{1}{3}$  (que é igual à medida do intervalo), para o caso de aleatorização uniforme.

Assim como neste último exemplo, muitas situações aleatórias do nosso cotidiano estão “mais próximas” do caso contínuo do que do caso discreto. Considere, por exemplo, a possível altura que uma criança pode atingir ao chegar à vida adulta.

Assim, quando os possíveis cenários de um experimento aleatório são valores reais, necessitamos trabalhar com um objeto matemático conhecido como *distribuição de probabilidade* cuja definição<sup>1</sup> é apresentada a seguir:

**Definição 3.** Dado um experimento aleatório  $X$  cujos resultados são valores reais, sua **densidade** ou **distribuição** de probabilidade é uma função  $f : \mathbb{R} \rightarrow \mathbb{R}_+$  tal que a área abaixo do gráfico de  $f$ , acima do eixo das abscissas e compreendida entre as retas  $x = a$  e  $x = b$  é igual à probabilidade do resultado do experimento  $X$  ter resultado entre  $a$  e  $b$ .

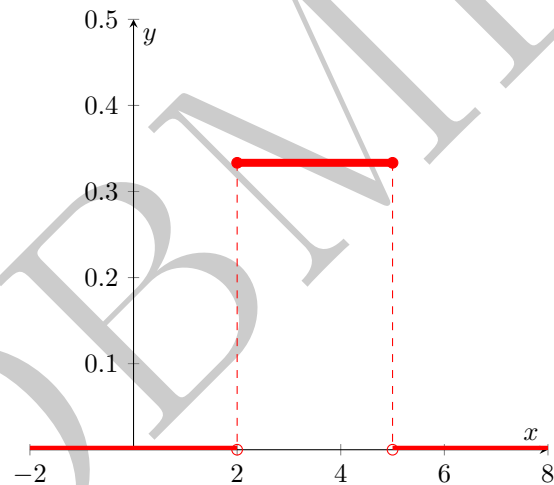
<sup>1</sup>Essa não é a definição mais geral para distribuição de probabilidade, porém é que mais facilmente se adequa ao Ensino Médio.

O exemplo mais simples de densidade de probabilidade é a **uniforme**, ilustrada no exemplo a seguir.

**Exemplo 4.** Se escolhermos ao acaso, mas de maneira uniforme, um ponto no intervalo  $[2, 5]$ , a densidade de probabilidade desse experimento é a função:

$$f(x) = \begin{cases} 0, & \text{se } x < 2, \\ \frac{1}{3}, & \text{se } 2 \leq x \leq 5, \\ 0, & \text{se } x > 5, \end{cases}$$

cujo gráfico é ilustrado a seguir em vermelho:



O exercício a seguir deixa claro porque a função acima é a correta.

**Exercício 5.** Considerando a distribuição de probabilidade do exemplo anterior, calcule a probabilidade do resultado

- (a) estar entre 3 e 4,5;
- (b) ser menor do que 4;
- (c) ser maior do que 7.

**Solução.**

- (a) A região do plano cuja área corresponde a essa probabilidade é um retângulo com dimensões  $\frac{1}{3}$  e  $4,5 - 3 = 1,5 = \frac{3}{2}$ , conforme apresentado na figura 1. Logo, a probabilidade (que é igual à área do retângulo) vale  $\frac{1}{3} \times \frac{3}{2} = \frac{1}{2}$ .
- (b) A região do plano cuja área corresponde a essa probabilidade é um retângulo com dimensões  $\frac{1}{3}$  e 2 conforme apresentado na figura 2. Logo, a probabilidade vale  $\frac{1}{3} \times 2 = \frac{2}{3}$ .
- (c) A região do plano cuja área corresponde a essa probabilidade é vazia. Logo, a probabilidade correspondente vale zero.

□

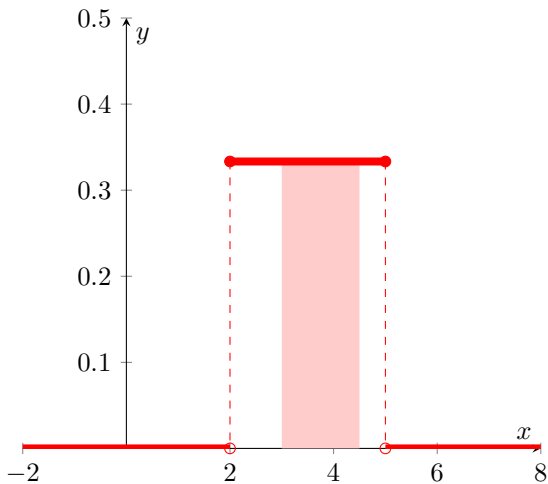


Figura 1: Figura relativa ao item (a).

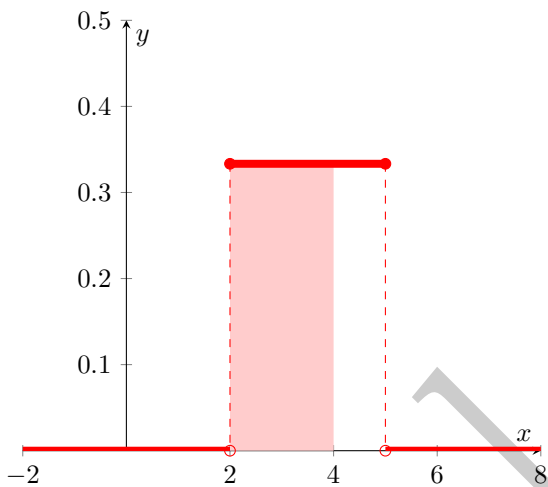


Figura 2: Figura relativa ao item (b).

## 2 A curva normal

Além do caso uniforme, que observamos no exemplo anterior, a distribuição de probabilidade de um determinado experimento pode estar relacionada com outros tipos de curva. A seguir, definiremos uma importante função que possui uma ampla utilidade em Estatística.

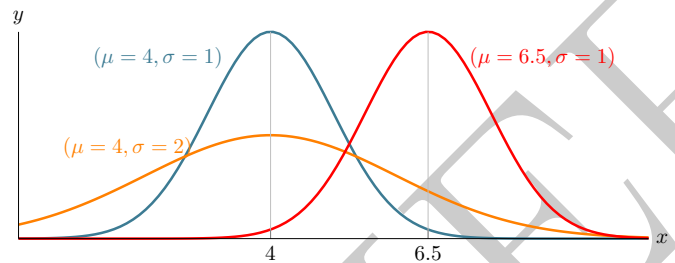
**Definição 6.** Dados dois valores reais  $\mu$  e  $\sigma > 0$ , a curva (ou função) normal de parâmetros  $(\mu, \sigma)$  é definida da seguinte forma:

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

onde  $\pi \cong 3,1415$  e  $e \cong 2,7182$ .

Os parâmetros  $\mu$  e  $\sigma$  são denominados de **média** e **desvio padrão**, respectivamente, da função normal.

A figura a seguir traz alguns gráficos para a função normal, para diferentes valores dos parâmetros  $\mu$  e  $\sigma$ , os quais deixam claro porque chamamos  $\mu$  de *média*<sup>2</sup>.



Veja que, quanto maior for o valor do parâmetro  $\mu$ , mais para a direita estará o gráfico. Por outro lado, quanto maior for o valor de  $\sigma$ , mais achatado será o gráfico.

Para construir os gráficos destas funções, basta utilizar as seguintes linhas de comando em R:

```
x <- seq(0,15,length=1000)
y <- dnorm(x,mean=4, sd=1)
plot(x,y, type="l", lwd=1)
```

- A primeira linha de comando serve para “guardar” os 1000 números que dividem o intervalo  $[0, 15]$  em subintervalos de mesma medida. Ou seja,  $x$  variará ao longo da sequência  $(x_1, x_2, \dots, x_{1000})$ , em que  $x_i = \frac{15(i-1)}{999}$ .
- A segunda linha de comando serve para “criar” outros 1000 valores e armazená-los na variável  $y$ , que percorrerá a sequência  $(y_1, y_2, \dots, y_{1000})$ , com  $y_i = f(x_i)$ , em que  $f$  é a curva normal com média 4 e desvio-padrão 1. (Em inglês, **mean** significa média e **sd** é a abreviação para *standard deviation* que significa desvio-padrão.)
- O terceiro comando serve para criar o gráfico gerado pelos pontos  $(x_i, y_i)$ .

**Exercício 7.** Use o site [www.tutorialspoint.com/execute\\_r\\_online.php](http://www.tutorialspoint.com/execute_r_online.php) para criar os gráficos de outras curvas normais, com:

(a)  $\mu = 5$  e  $\sigma = 2$ ;

(b)  $\mu = -3$  e  $\sigma = 8$ .

Para cada um destes exemplos, escolha um intervalo adequado.

Um caso especial importante é quando  $\mu = 0$  e  $\sigma = 1$ . Quando isto ocorrer, diremos que estamos trabalhando com a **curva normal padrão**, que estudaremos em maior profundidade nas próximas aulas.

<sup>2</sup>Devido ao *software* utilizado para gerar as figuras, o valor 6,5 aparece como 6.5. Isso ocorrerá algumas outras vezes, e não deve gerar confusão.

### 3 Propriedades da curva normal

A seguir, listamos algumas propriedades notáveis das curvas normais:

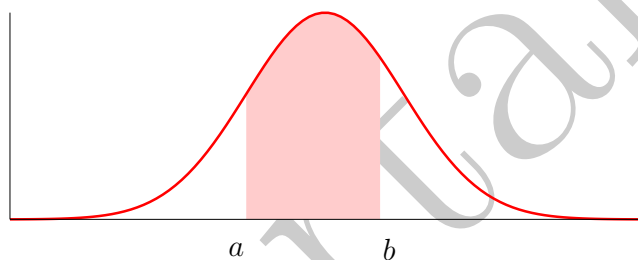
1. A curva é simétrica em torno da média, ou seja,  $f(\mu - x) = f(\mu + x)$ . Isto é facilmente verificável substituindo os valores  $\mu - x$  e  $\mu + x$  na fórmula que define a função.
2. Quanto mais à direita está a curva, maior será sua média.
3. Quanto menor for seu desvio-padrão, mais “alto” será o valor de  $f(\mu)$ . Assim, quanto maior for o valor de  $\sigma$ , mais “achatada” será a curva. Isso significa que será mais provável que valores mais distantes da média apereçam aleatoriamente.

Agora, suponha que uma determinada característica de uma população tenha densidade de probabilidade dada pela curva normal, com média  $\mu$  e desvio-padrão  $\sigma$ . Como já comentamos, a área da região do plano cartesiano delimitada pela curva normal, pelo eixo das abscissas e pelas retas verticais  $x = a$  e  $x = b$  representará a probabilidade de obtermos um valor de observação dessa característica compreendido entre  $a$  e  $b$ , quando selecionarmos ao acaso um dos elementos da população.

Doravante, utilizaremos a notação

$$P(a \leq z \leq b)$$

para o valor dessa probabilidade. A figura a seguir destaca a área correspondente.



Por se relacionar com uma medida de probabilidade, fazendo  $a$  diminuir mais e mais (escrevemos  $a \rightarrow -\infty$  e lemos  $a$  tende a  $-\infty$ ) e  $b$  aumentar mais e mais (escrevemos  $b \rightarrow +\infty$  e lemos  $b$  tende a  $+\infty$ ), temos que

$$P(-\infty < z < +\infty) = 1.$$

Em palavras, a probabilidade de obtermos um valor real (isto é, pertencente ao intervalo  $(-\infty, \infty)$ ) é igual a 1. Por outro lado, para qualquer intervalo menor, esse valor será estritamente menor do que 1.

**Definição 8.** Dada uma curva normal  $f(x)$ , com média  $\mu$  e desvio-padrão  $\sigma$ , a **função de distribuição acumulada** correspondente é  $F : \mathbb{R} \rightarrow \mathbb{R}$  tal que

$$F(k) = P(z \leq k).$$

No R, os valores de  $F$  podem ser facilmente obtidos através do comando `pnorm(k, mean=a, sd=b)`, onde o valor de  $k$  deve ser um número real tal que queremos calcular  $F(k)$ , ao passo que  $a$  é o valor da média e  $b$  o valor do desvio-padrão (que deve ser positivo).

Por exemplo, o comando

```
pnorm(0, mean=0, sd=1)
```

Nos dá o valor de  $P(z \leq 0)$  referente à probabilidade de se obter um valor menor ou igual a zero ao sortear aleatoriamente uma observação de um experimento que segue a distribuição normal padrão.

Após executar esse comando, obteremos como resultado

```
[1] 0.5,
```

o que já era esperado, uma vez que a curva normal padrão é simétrica em torno de zero.

**Exercício 9.** Dada a curva normal padrão, utilize o comando `pnorm` do R para calcular as seguintes probabilidades:

- (a)  $P(z \leq 2)$ ;
- (b)  $P(z \geq 1,3)$ ;
- (c)  $P(-2 \leq z \leq 1,3)$ .

**Solução.**

- (a) Basta computar `pnorm(2)` para encontrar 0,9772.
- (b) Primeiro computamos `pnorm(1.3)` e encontramos 0,9032; esta é a probabilidade do resultado do experimento ser um valor menor ou igual a 1,3. Para encontrarmos o valor pedido no item, subtraímos esse número de 1 (que é equivalente à probabilidade total), obtendo  $1 - 0,9032 = 0,0968$ .
- (c) Inicialmente, temos

$$P(-2 \leq z \leq 1,3) = P(z \leq 1,3) - P(z \leq -2).$$

Agora, a solução do item (b) dá  $P(z \leq 1,3) = 0,9032$ . Por outro lado, utilizando a simetria da curva normal e o resultado do item (a), segue que

$$P(z \leq -2) = P(z \geq 2) = 1 - P(z \leq 2) = 1 - 0,9772.$$

Então,

$$P(-2 \leq z \leq 1,3) = 0,9032 - (1 - 0,9772) = 0,8804.$$

□

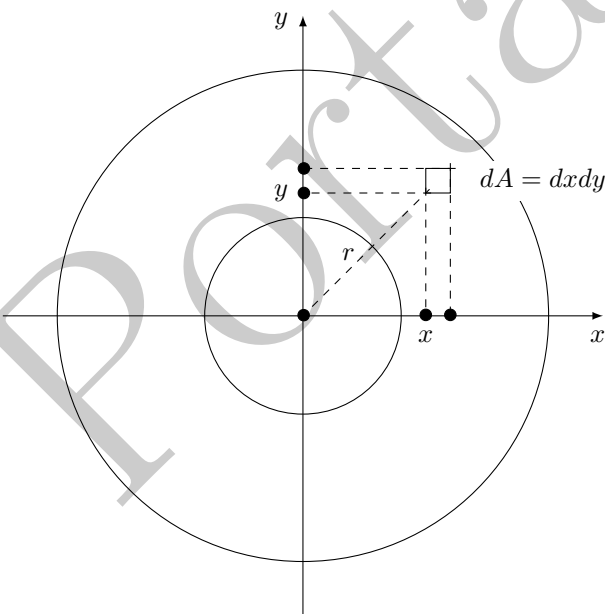
Ainda em relação ao exercício anterior, nem sempre temos à disposição um computador para calcular probabilidades relativas a uma distribuição normal com o auxílio do R. Na próxima aula, apresentaremos uma forma “menos tecnológica” de encontrar as probabilidades relativas a uma distribuição normal.

## 4 De onde vem a curva normal?

A esta altura, você pode estar se perguntando se realmente existem experimentos aleatórios tendo curvas normais como densidades de probabilidade. A verdade é que tais fenômenos abundam na Natureza, e alguns exemplos são: as alturas dos homens de uma mesma idade em uma grande população, as velocidades das moléculas de um gás perfeito, os resultados de um grande número de medições de uma grandeza física em uma experiência, etc.

Aqui, apresentaremos um exemplo de experimento mais *prosaico*, mas não menos interessante, para explicar como a curva normal aparece na descrição de fenômenos aleatórios. Seguimos a referência [3].

Considere o experimento, suposto aleatório, de lançar um dardo em um alvo circular e observar onde o dardo atinge o alvo. Escolha um sistema cartesiano de eixos com centro no centro do alvo, conforme mostrado na figura a seguir. Seja  $f(x)$  a densidade de probabilidade correspondente ao experimento de o dardo atingir o alvo num ponto de abscissa situada no intervalo  $[x, x + dx]$ , em que  $dx$  representa um pequeno incremento à abscissa  $x$ . Por simetria,  $f(y)$  é a densidade de probabilidade correspondente ao experimento de o dardo atingir o alvo num ponto de ordenada situada no intervalo  $[y, y + dy]$ , em que  $dy$  representa um pequeno incremento à ordenada  $y$ .



Assumindo que  $f(t)$  varia *continuamente* com  $t$ , temos que  $f(x)$  é uma aproximação razoável para  $f$  em todo o intervalo  $[x, x + dx]$ , contanto que o incremento  $dx$  seja pequeno. Dessa forma,  $f(x)dx$  é uma aproximação razoável para a probabilidade de que o dardo atinja o alvo em um ponto de abscissa situada no intervalo  $[x, x + dx]$ . Da mesma forma,  $f(y)dy$  é uma aproximação razoável para a probabilidade de que o dardo atinja o alvo em um ponto de ordenada situada no intervalo  $[y, y + dy]$ .

Agora, fazendo a suposição (razoável!) de que os desvios do dardo em relação ao alvo nas direções  $x$  e  $y$  são independentes um do outro, temos que o produto

$$f(x)dx \cdot f(y)dy = f(x)f(y)dxdy = f(x)f(y)dA$$

é uma aproximação razoável para a probabilidade de que o dardo atinja o alvo em um ponto do retângulo de área  $dA$  mostrado na figura.

Por fim, supondo adicionalmente que o experimento possui *simetria circular* (o que também é razoável), concluímos que a última probabilidade acima será a mesma para todo retângulo de área  $dA = dxdy$ , situado a uma mesma distância  $r$  do centro do alvo. Matematicamente, isto se resume a assumir que  $f(x)f(y)$  é uma função de  $r^2$ , que, pelo Teorema de Pitágoras, vale  $x^2 + y^2$ . Em resumo,

$$f(x)f(y) = g(x^2 + y^2). \quad (1)$$

A partir daí, admitindo que os gráficos das funções  $f$  e  $g$  são *suaves* (isto é, que são curvas contínuas e sem *quinas*), um argumento simples envolvendo um pouco de Cálculo garante que as únicas funções  $f$  que satisfazem (1) são as da forma

$$f(t) = Ae^{Bt^2}.$$

Como  $f$  é uma densidade de probabilidade, a área sob o gráfico de  $t \mapsto f(t)$ , quando  $t$  varia de  $-\infty$  a  $+\infty$ , vale 1; isso garante que  $B$  deve ser negativo. Então, escrevendo  $B = -\frac{1}{2\sigma^2}$  e calculando a área sob o gráfico de  $f(t) = Ae^{-\frac{t^2}{2\sigma^2}}$  (também com o auxílio dos métodos do Cálculo), obtemos  $A = \frac{1}{\sigma\sqrt{2\pi}}$ , de sorte que

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

## 5 Sugestões aos professores

Separe pelo menos dois encontros de 50 minutos para ensinar o conteúdo deste material. Se possível, ministre a aula em um único encontro de 140 minutos. Este assunto é um pouco mais abstrato que os anteriores, porém, será muito importante aos alunos nas próximas aulas, quando abordaremos alguns métodos de inferência estatística. Por isso, é fundamental a paciência do professor em deixar claros os assuntos deste material.

Durante esta aula, realizamos uma série de exemplos utilizando o R. Este é um programa livre que está disponível

para *download* no site [www.r-project.org](http://www.r-project.org) e é compatível com os principais sistemas operacionais. Além disso, vários sites na internet simulam este *software* de maneira online e gratuita. Um deles é [https://www.tutorialspoint.com/execute\\_r\\_online.php](https://www.tutorialspoint.com/execute_r_online.php). Copie os códigos apresentados no exemplo 3 e veja o que acontece! Tente também modificar os códigos para obter histogramas para outros exercícios.

## Referências

- [1] João Ismael Pinheiro et al. *Estatística Básica: a arte de trabalhar com dados*. Campus, 2009.
- [2] Pedro A. Morettin and Wilton de O. Bussab. *Estatística Básica*. Saraiva, 2010.
- [3] G.F. Simmons. *Differential Equations with Applications and Historical Notes*. Textbooks in Mathematics. CRC Press, 2016.