Capítulo 5

Distribuição Normal

Muitas variáveis aleatórias contínuas, tais como altura, comprimento, peso, entre outras, podem ser descritas pelo modelo Normal de probabilidades. Este modelo é, sem dúvida, o mais importante modelo de probabilidades devido ao seu papel central na inferência estatística (testes de hipóteses, intervalos de confiança, etc.).

Definição

Se a variável aleatória contínua X tem distribuição normal ou gaussiana com parâmetro de posição μ e de escala σ^2, denotada por $X \sim N(\mu, \sigma^2)$, então sua função densidade de probabilidade é dada por

$$f_X(x|\mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\} I_{(-\infty,\infty)}(x), \quad \mu \in \Re, \sigma^2 > 0.$$

A figura 5.1 ilustra a densidade da curva normal, que é simétrica em torno da média (μ).

Média e variância

$$\begin{aligned} E(X) &= \mu \\ Var(X) &= \sigma^2 \end{aligned}$$

A distribuição normal apresenta as seguintes características:

- A variável pode assumir qualquer valor real
- A distribuição é em forma de sino, simétrica em torno da média (μ), como se observa na figura 5.1

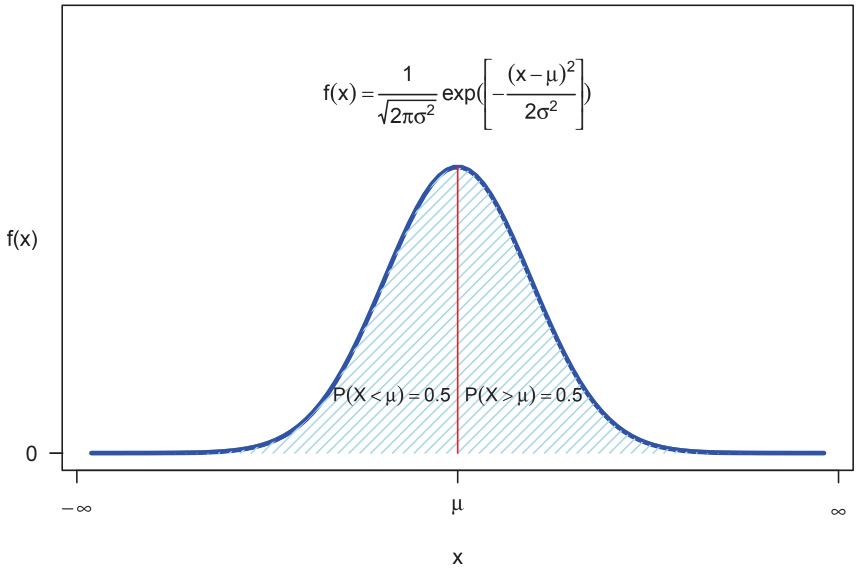

Figura 5.1: Densidade de probabilidade normal.

- A área total sob a curva vale 1, pois esta é a probabilidade de a variável assumir qualquer valor real
- O formato da curva é determinado unicamente pelos dois parâmetros da distribuição, μ e σ^2. Alterando a média, alteramos a posição da distribuição, como mostra a figura 5.2A. Alterando a variância, alteramos a dispersão da distribuição, como mostra a figura 5.2B.

Distribuição Normal Padrão

A integral da curva normal não tem solução analítica. Assim sendo, os cálculos são, em geral, baseados na *distribuição normal padrão*, de acordo com a transformação

$$Z = \frac{X - \mu}{\sigma}$$

de modo que $Z \sim N(0, 1)$, isto é, Z tem média zero e variância unitária. Esse resultado é válido para quaisquer valores de μ e σ^2. Isso permite a construção de tabelas contendo probabilidades em função dos valores de Z, do tipo

$$P(X \leq x) = P(Z \leq z)$$

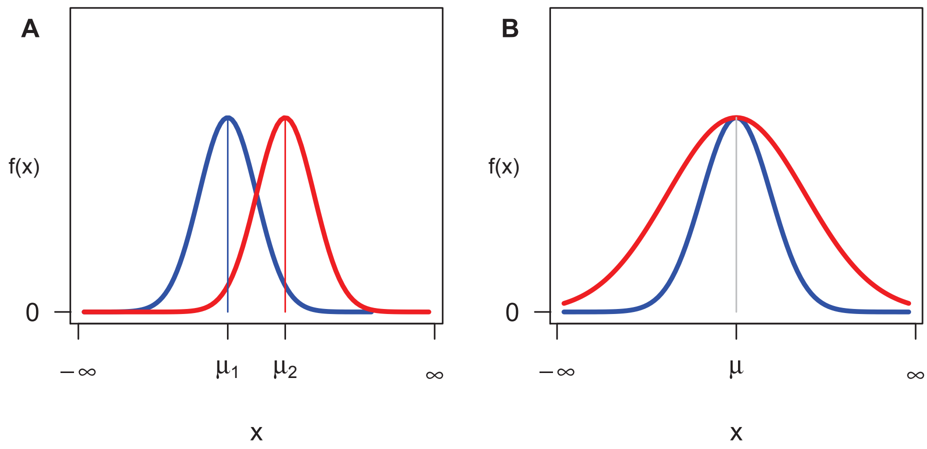

Figura 5.2: Densidades de probabilidade normal com (A) médias diferentes e mesma variância e (B) médias iguais e variâncias diferentes.

A tabela da distribuição normal padrão que usaremos neste curso nos fornece a probabilidade da variável Z assumir um valor entre zero e um particular valor z_0, isto é,

$$P(0 \leq Z \leq z_0)$$

como mostrado na figura 5.3.

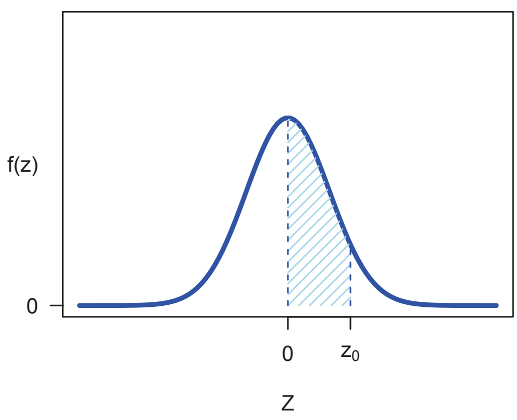

Figura 5.3: Densidade de probabilidade normal padrão e área sombreada indicando $P(0 \leq Z \leq z_0)$.

Aplicações

A distribuição normal é, certamente, a mais importante das distribuições probabilísticas, devido a quatro principais motivos:

1. Devido ao papel central que desempenha nas técnicas de inferência.

2. Outras importantes distribuições amostrais são derivadas da distribuição normal, tais como: qui-quadrado, F e t-Student.

3. Há ainda o *teorema central do limite* (TCL), que permite aproximações de vários outros modelos, contínuos e discretos, em grandes amostras. São exemplos clássicos as aproximações dos modelos binomial e Poisson.

5.1 Alguns resultados úteis

Corolário 1. *Seja $X \sim Binomial(n, \pi)$. Pelo teorema central do limite, para n grande e π não muito extremo (próximo de 0 ou 1), tem-se que*

$$X \sim Normal(\mu = n\pi, \sigma^2 = n\pi(1-\pi))$$

Corolário 2. *Se $X \sim Normal(0, 1)$, então*

$$\mu + \sigma X \sim Normal(\mu, \sigma^2)$$

5.2 Exercícios

1) A observação dos pesos (X) de um grande número de espigas de milho apresentou distribuição normal com média $\mu = 120$g e desvio padrão $\sigma = 10$g. Num programa de melhoramento genético, entre outras características, uma linhagem deve satisfazer a condição: $112 < X < 140$g. Num programa envolvendo 450 linhagens, qual deve ser o número provável de linhagens que atende à essa condição?

2) Sabe-se que o peso médio, em arrobas, de abate de bovinos é normalmente distribuído com média 18 e variância 2,25. Um lote de 5000 cabeças, com essa característica, foi destinado ao frigorífico que abate só a partir de um peso mínimo w. Sabendo-se que foram abatidas 4200 cabeças, pede-se:

a. O número esperado de bovinos com peso entre 17 e 19 arrobas.

b. O valor de w.

3) As notas de uma prova são normalmente distribuídas com média 73 e variância 225. Os 15% melhores alunos recebem o conceito A e os 11,9% piores alunos recebem conceito R. Determine:

a. A nota mínima para receber conceito A.

b. A nota mínima para ser aprovado.

c. $P(X \geq 55,3)$.

4) Se $X \sim Normal(3,4)$, calcule o valor de k tal que $P(X \geq k) = 2P(X \leq k)$.

5) O diâmetro de um cabo elétrico é normalmente distribuído com média 0,8 mm e variância 0,0004 mm^2. Dentre uma amostra de 1000 cabos, quantos esperamos que tenha diâmetro...

a. ... maior ou igual a 0,81 mm?

b. ... entre 0,73 e 0,86 mm?

c. ... menor que 0,78 mm?

6) Um pesquisador decidiu que, para facilitar a classificação das aves em experimentos de nutrição, deve-se dividir as poedeiras no início da postura em três grupos **equiprováveis** de peso, a saber: poedeiras pesadas, poedeiras médias, poedeiras leves. Encontre os pesos correspondentes à cada classe, sabendo-se que o peso médio das aves nessa idade é 1,5 kg, e o desvio padrão é de 0,17 kg. Asmita a distribuição normal para a variável peso.

Respostas

1) 345
2) **a**. 2486 **b**. 16,52
3) **a**. 88,66 **b**. 55,3 **c**. 0,881
4) 2,14
5) **a**. 308,5 $\approx$ 309 cabos **b**. 998,7 $\approx$ 999 cabos **c**. 158,7 $\approx$ 159 cabos
6) Leves: peso $<$ 1,43 kg; Médias: $1{,}43 \leq \text{peso} \leq 1,57$; Pesadas: peso $>$ 1,57