Fundamentos de Estatística

Last updated: July 16, 2025

Contents

※ Modelos probabilísticos

1.1 Distribuição contínuas e modelo normal

Dizemos que uma variável aleatória é contínua quando não conseguimos enumerar seus possíveis resultados, por esses formarem um conjunto infinito, num dado intervalo de números reais.

Por exemplo, a altura de um indivíduo, tomado ao acaso, é uma variáve aleatória continua, pois não é possível enumerar todos os valores possíveis de altura de indivíduos, mas podemos dizer, por exemplo, que o resultado será um número real do intervalo de zero a dois metros e meio, o qual contém infinitos números.

1.1.1 Distribuições contínuas

Em variáveis aleatórias contínuas, não existe interesse em atribuir probabilidade a cada particular valor, mas sim, para eventos formados por intervalos de valores.

A especificação da distribuição de probabilidades de uma variável aleatória continua é realizada por um modelo matemático que permite calcular probabilidades em qualquer intervalo de números reais.

Example 1.1. Considere um círculo, com medidas de ângulos, em graus, a partir da origem. Um ponteiro que é colocado a girar no sentido anti-horário.

Seja X a variável aleatória que indica o ponto em que o ponteiro para de girar. Como existem infinitos pontos no intervalor 0 a 360^o, esta variável aleatória é contínua. Primeiramente, vejamos a probabilidade de o ponteiro parar no quadrante 1, isto é, a probabilidade de X assumir um valor entre 0 e 90^o.

Peleo princípio da equiprobabilidade, as probabilidades de parada são iguais para os quatros quadrantes. Assim, a probabilidade de o ponteiro parar no primeiro quadrante deve ser igual a $\frac{1}{4}$.

Podemos representar o evento do ponteiro parar no quadrante 1 por $0 \leq X \leq 90$, e a probabilidade por $P(0 \leq X < 90)$.

A distribuição de probabilidades de uma variável aleatória contínua pode ser representada por uma função não negativa, com a área entre o eixo-X e a curva igual a 1 (um). Os eventos podem ser representados por intervalos no eixo-X, enquanto as probabilidades, pelas correspondentes áreas sob a curva.

A área de um retãngulo é dada por base x altura. Como a base e 360 e a área e 1, então a altura tem que ser $\frac{1}{360}$.

Example 1.2. Selecione, aleatoriamente , um aluno do sexo masculino de uma certa universidade. Seja X o valor de sua altura, em centímetros.

Temos uma variável aleatória contínua. É intuitivo que a probabilidade do estudante ter altura entre 165 cm e 175 cm seja bem maior do que entre 190 cm e 200 cm, mesmo que ambos os intervalos tenham a mesma amplitude.

Um modelo adequado para representar essa situação é o modelo conhecido como distribuição normal de probabilidades, onde existe um valor típico, ou valor médio. Intervalos em torno deste valor médio têm altas probabilidades de ocorrência, mas estas probabilidades diminuem na medida que nos afastamos deste valor médio, para a esquerda, ou para a direita.

1.1.2 Distribuição normais

A distribuição normal é caracterizada por uma função, cujo gráfico descreve uma curva em forma de sino. Esta distribuição depende dos dois parâmetros a seguir:

- μ (média ou valor esperado): especifica a posição central da distribuição de probabilidades.
- σ (desvio padrão): especifica a variabilidade da distribuição de probabilidades.

1.1.3 Valores padronizados e a distribuição normal padrão

Com o objetivo de facilitar a obtenção de determinadas áreas sob uma curva normal, podemos fazer uma transformação na variável, levando-a para uma distribuição normal com média 0 e desvio padrão 1.

Definition 1.3. A distribuição normal com média 0 e desvio padrão 1 é conhecida como distribuição normal padrão.

Para transformar um valor x, de uma distribuição normal com média μ e desvio padrão σ, em um valor z da distribuição normal padrão, basta fazer a seguinte operação:

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

Definition 1.4. O valor z conhecido como valor padronizado é uma medida relativa. Mede o quanto x se afasta da média (μ), em unidade de desvio padrão (σ).

Example 1.5. Suponha que a altura dos estudantes do sexo masculino de uma certa universidade tenha distribuição normal com média $\mu = 170$ cm e desvio padrão $\sigma = 10$ cm. Para o estudante com altura $x = 180$ cm, temos como valor padronizado:

$$z = \frac{180 - 170}{10} = 1$$

Podemos dizer que este estudante de altura 180 cm encontra-se a 1 desvio padrão acima da altura média dos estudantes do sexo masculino da universidade.

Seja X a altura, em centimentro, de um estudante do sexo masculino, selecionado ao acaso. Considere que temos interesse no evento $X > 180$. Podemos representar graficamente a equivalência da probabilidade deste evento, $P(X > 180)$, com área na distribuição normal padrão.