1 Revisão

1.a Introdução

Probabilidade: estuda experimentos ou fenômenos aleatórios e através dela é possível analisar as chances de um determinado evento ocorrer.

Quando calculamos a probabilidade, estamos associando um grau de confiança na ocorrência dos resultados possíveis de experimentos, cujos resultados não podem ser determinados antecipadamente. Desta forma, o cálculo da probabilidade associa a ocorrência de um resultado a um valor que varia de 0 a 1 e, quanto mais próximo de 1 estiver o resultado, maior é a certeza da sua ocorrência.

1.b Conceitos básicos

Experimento aleatório: é a observação sistemática de um fenômeno (evento aleatório) qualquer da população; pode ser repetido inúmeras vezes e nas mesmas condições e, mesmo assim, apresenta resultados diferentes.

Variável aleatória: é o valor assumido pelo fenômeno em um experimento qualquer. A variável é, portanto, o valor que pode assumir o evento dentro de um conjunto de valores possíveis chamado domínio da variável;

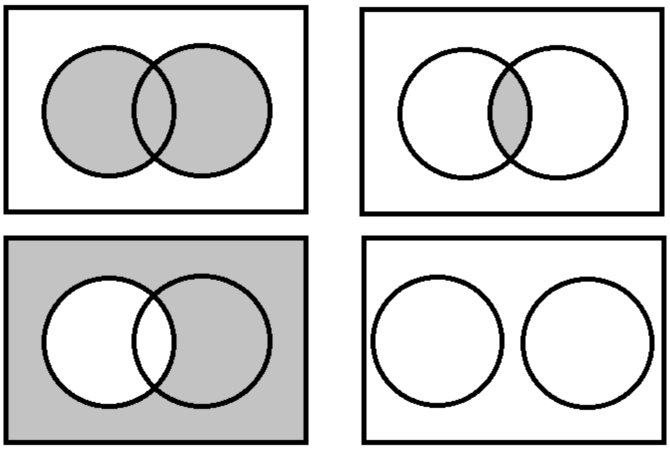
- variável aleatória contínua: pode assumir qualquer valor em uma escala de valores;
- variável aleatória discreta: os valores possíveis são números inteiros (contagens).

Espaço amostral: – conjunto de todos os resultados possíveis de um experimento (Ω).

Evento: subconjunto de um espaço amostral, que pode conter nenhum elemento (conjunto vazio $\varnothing$) ou todos os elementos de um espaço amostral. A probabilidade do evento ocorrer é a razão entre o número de resultados favoráveis e o número de resultados possíveis $\left(P(A) = \frac{\#A}{\#\Omega}\right)$;

- eventos independentes: quando a ocorrência de um evento não influência a ocorrência do outro evento ($P(A \cap B) = P(A)P(B)$, ou ainda $P(B|A) = P(B)$ e $P(B|A) = P(A)$);
- evento complementar: todos os resultados do espaço amostral que não fazem parte do evento desejado ($P(A^c) = 1 - P(A)$):
- evento soma (ou união): consiste na realização de pelo menos um dos eventos (um ou outro) e contém todos os elementos de ambos, incluindo os que são e os que não são comuns aos dois conjuntos ($P(A \cup B) = P(A) + P(B) - P(A \cap B)$);
- evento produto (ou intersecção): consiste na realização de ambos os eventos (um e outro — isto é, eles devem ocorrer simultaneamente) e contém os elementos comuns a ambos os conjuntos ($P(A \cap B) = P(B)P(A|B)$);
- eventos mutuamente exclusivos: não podem ocorrer simultaneamente ($P(A \cap B) = 0$).

Figura 1: Diagramas de Venn para eventos. *Da esq. para dir., de cima para baixo: união de eventos, intersecção de eventos, evento complementar, eventos mutuamente exclusivos.*



Fonte: Elaborada pelo autor

1.c Probabilidade condicional

Dois eventos são condicionados quando a ocorrência de um altera a probabilidade de ocorrência do outro. A probabilidade de ocorrer um evento A dado que ocorreu um evento B é dada por $P(A|B) = \frac{P(A \cap B)}{P(B)}$.

Fórmula da Probabilidade Total: Sejam $B_1, B_2, \ldots, B_n$ eventos que formam uma partição do espaço amostral (isto é, $\bigcup_{i=1}^{n} B_i = \Omega$) e $P(B_i) > 0$ para $i = 1, 2, \ldots, n$. Então, para qualquer evento $A \subseteq \Omega$:

$$P(A) = P(B_1 \cap A) + \cdots + P(B_n \cap A) = P(B_1)P(A|B_1) + \cdots + P(B_n)P(A|B_n) = \sum_i P(B_i)P(A|B_i).$$

Teorema de Bayes: Sejam $B_1, B_2, \ldots, B_n$ eventos que formam uma partição do espaço amostral (isto é, $\bigcup_{i=1}^{n} B_i = \Omega$) e $P(B_i) > 0$ para $\forall i$. Então, para qualquer evento $A \subseteq \Omega$, $P(A) > 0$:

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{P(A)} = \frac{P(B_i)P(A|B_i)}{P(B_1)P(A|B_1) + \ldots + P(B_n)P(A|B_n)} = \frac{P(B_i)P(A|B_i)}{\sum_i P(B_i)P(A|B_i)}.$$

1.d Distribuições de probabilidade

X é uma variável aleatória discreta:

- a cada possível resultado x_i associa-se um número $p(x_i) = P(X = x_i)$, denominado probabilidade de x_i. Os números $p(x_i)$, $i = 1, 2, \ldots$ devem satisfazer as seguintes condições: $p(x_i) \geq 0$ para todo i; $\sum_{i=1}^{\infty} p(x_i) = 1$. A função p é denominada distribuição (ou função massa) de probabilidade da variável aleatória X.

- se X é variável aleatória discreta que assume os valores $\{x_1, x_2, \ldots\}$ e seja $p(x)$ a função de probabilidade de X, isto é, $p(x_i) = \mathbb{P}(X = x_i)$. Então, o valor esperado de X, também chamado de esperança de X e denotado por $\mathbb{E}(X)$ ou μ_X, é definido por $\mathbb{E}(X) = \mu_X = \sum_{i=1}^{\infty} x_i p(x_i) = \sum_{i=1}^{\infty} x_i \mathbb{P}(X = x_i)$.
  - $\mathbb{E}[g(X)] = \sum_{x_i \in \mathcal{R}_X} g(x_i)p(x_i) = \sum_{x_i \in \mathcal{R}_X} g(x_i)\mathbb{P}(X = x_i)$.

X é uma variável aleatória contínua:

- a função é denominada função densidade de probabilidade (fdp), e satisfaz às seguintes propriedades: $f(x) \geq 0$, para todo $x \in \mathbb{R}_x$ e $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1$. Além disso, definimos para qualquer $c, d \in \mathbb{R}_x$, com $c < d$ que $P(c < X < d) = \int_c^d f(x)dx$. Ressalta-se que probabilidade de um ponto isolado é sempre zero, ou seja, $\mathbb{P}(X = c) = \displaystyle\int_c^c f(x)dx = 0$.
- $\frac{\partial}{\partial x} F_X(x) = f_X(x)$.
- se X é variável aleatória contínua com função densidade de probabilidade $f(x)$, define-se o valor esperado ou esperança matemática ou média de X por $\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x)dx$.

Função de distribuição acumulada: descreve como probabilidades são associadas aos valores ou aos intervalos de valores de uma variável aleatória. Ela representa a probabilidade de uma variável aleatória ser menor ou igual a um valor real x. Assim, a fda de uma variável aleatória X é uma função que a cada número real $x \in \mathbb{R}$ associa o valor $F(x) = P(X \leq x) \in [0, 1]$.

Relação entre a função de distribuição acumulada e a distribuição de probabilidade discreta: Seja X uma variável aleatória discreta cuja distribuição de probabilidade associa aos valores $x_1, x_2, \ldots$ as respectivas probabilidades $P(X = x_1), P(X = x_2), \ldots$. Como os valores de X são mutuamente exclusivos, temos que a função de distribuição acumulada é dada por $F(x) = \sum_{i \in A_x} P(X = x_i)$, com $\{i : x_i \leq x\}$. Assim, dada a distribuição de probabilidade de uma variável aleatória discreta, conseguimos determinar sua função de distribuição acumulada, ou ainda, dada a função de distribuição acumulada, podemos determinar a sua distribuição de probabilidade.

Algumas propriedades da esperança:

- se $X = c$ com c uma constante real, então $\mathbb{E}(X) = c$;
- sejam X e Y duas variáveis aleatórias quaisquer. Então, $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$;
- sejam X e Y variáveis aleatórias independentes. Então, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Variância:

$$\mathbb{V}(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right].$$

$$\mathbb{V}(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right] = \mathbb{E}\left\{X^2 - 2X\mathbb{E}(X) + [\mathbb{E}(X)]^2\right\},$$

ou seja,

$$\mathbb{V}(X) = \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + [\mathbb{E}(X)]^2 = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2.$$

Algumas propriedades da variância:

- se c é uma constante real, $\mathbb{V}(X + c) = \mathbb{V}(X)$;
- se c é uma constante real, $\mathbb{V}(cX) = c^2\mathbb{V}(X)$
- se X e Y forem variáveis aleatórias independentes, então $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.

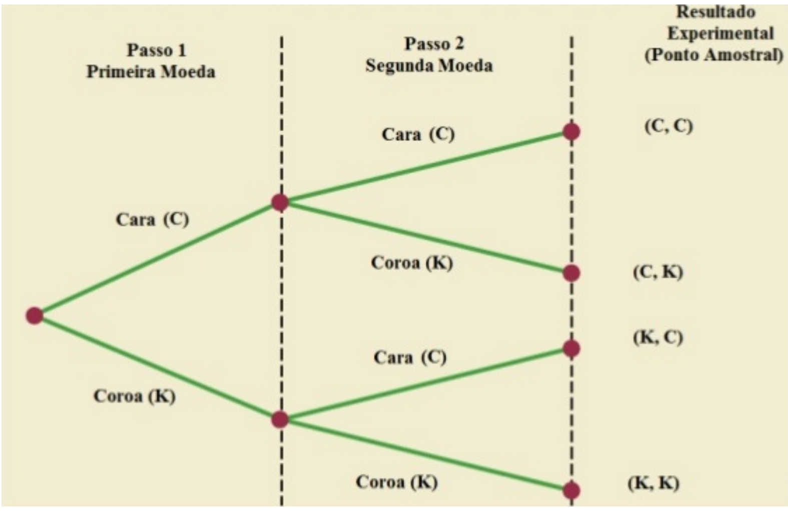
Árvore de probabilidades:

- usada para representar um espaço de probabilidade;
- cada nó no diagrama representa um evento e está associado com a probabilidade desse evento;

- a construção de um diagrama de árvores é feita simplesmente através da ramificação de todas as possibilidades de todos os experimentos;

- o somatório de todas as probabilidades das ramificações de um mesmo nó é sempre igual a 1.

Figura 2: Exemplo de árvore de probabilidades. *Experimento de lançar duas moedas, uma após a outra.*



Fonte: Adaptada de https://pt.slideshare.net/RicardoSantos11/probabilidade-estatstica-i

Covariância:
Se X e Y são variáveis aleatórias então a covariância entre X e Y é dada por:

$$Cov(X,Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))],$$
$$Cov(X,Y) = \mathbb{E}(XY - Y\mathbb{E}(X) - X\mathbb{E}(Y) + \mathbb{E}(X)\mathbb{E}(Y)) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

- se $Cov(X,Y) = 0$, dizemos que X e Y são não correlacionadas;

- se X e Y são independentes, então são não correlacionadas, pois neste caso $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$

- $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ não implica em independência, ou seja, a covariância zero não necessariamente implica independência.

Correlação:
Se X e Y são variáveis aleatórias então o coeficiente de correlação entre X e Y é dado por:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \mathbb{E}\left[\left(\frac{X - \mathbb{E}(X)}{\sigma_X}\right)\left(\frac{Y - \mathbb{E}(Y)}{\sigma_Y}\right)\right]$$

- $-1 \leq \rho(X,Y) \leq 1$;

- o coeficiente de correlação é independente de escala e de translação da variáveis, ou seja, $\rho(X,Y) = \rho(aX + b, cY + d)$, para $a < 0$, $c < 0$.