1 Noções de Probabilidade

Já vimos que para se obter informações sobre alguma característica da população, podemos utilizar uma amostra. Estudaremos agora a probabilidade, que é uma ferramenta usada e necessária para se fazer ligações entre a amostra e a população, de modo que a partir de informações da amostra se possa fazer afirmações sobre características da população.

As probabilidades são utilizadas para exprimir a chance de ocorrência de determinado evento. O estudo das probabilidades é importante pois elas são a base para o estudo estatístico

A teoria de probabilidades tem por objetivo o estudo de fenômenos aleatórios. Um fenômeno é chamado de aleatório se ele tem a seguinte propriedade: quando observado repetidamente sob as mesmas condições ele produz resultados diferentes. Mesmo que a chance da ocorrência seja alta, os resultados não são conhecidos antes de ocorrer, mas de certa forma, mantém uma certa regularidade, o que permite determinar a chance de ocorrência; a Probabilidade.

Exemplos:

- Jogar uma moeda repetidamente e observar o resultado da face de cima;
- Jogar um dado e observar o número mostrado na face superior;
- Número de filhos de um casal;

Observação: quando a possibilidade de repetir o fenômeno está na mão do experimentador, este fenômeno aleatório é chamado de experimento aleatório.

1.1 Espaço Amostral e Eventos

Espaço amostral (Ω) - é o conjunto de todos os possíveis resultados de um experimento.

Um espaço amostral é

Exemplo:

- Lançamento de um dado não viciado. Neste caso o espaço amostral é

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

- Lançar uma moeda duas vezes e observar as faces obtidas

$$\Omega = \{(Ca, Co), (Ca, Ca), (Co, Ca), (Co, Co)\}$$

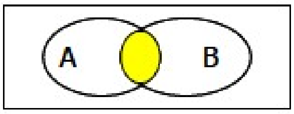
No lançamento de um dado pode-se interessar, por exemplo, somente na ocorrência de número ímpares. O subconjunto $A = \{1, 3, 5\}$ do espaço amostral Ω representa o evento A definido pela ocorrência de números ímpares.

Evento - é um subconjunto do espaço amostral que representa um resultado definido.

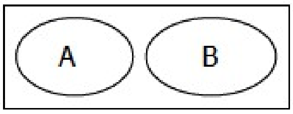
Ponto amostral - é apenas um elemento do espaço amostral.

1.1.1 Operação com eventos

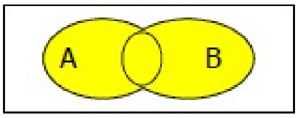
Sejam A e B dois eventos de um mesmo espaço amostral O evento intersecção de A e B, denotado $A \cap B$, e o evento em que A e B ocorrem simultaneamente.



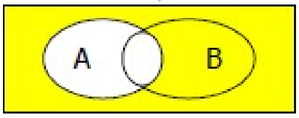
Dois eventos A e B são mutuamente exclusivos ou disjuntos se eles não podem ocorrer simultaneamente $A \cap B = \varnothing$.



O evento União de A e B, denotado $A \cup B$, e o evento em que A ocorre ou B ocorre (ou ambos).



O evento complementar de A, denotado A^c, é o evento em que A não ocorre.



Exemplo: Seja o espaço amostral $\Omega = \{1, 2, 3, 4, 5, 6\}$ e considere os eventos:

$$A = \{1, 3, 5\} \quad B = \{2, 4, 6\} \quad C = \{3, 4, 5, 6\}$$

Vamos fazer as seguintes operações:

$$\begin{aligned}
A \cap B &= \varnothing \quad \text{Conjuntos mutuamente exclusivos ou disjunto} \\
A \cap C &= \{3, 5\} \\
A \cup B &= \{1, 2, 3, 4, 5, 6\} = \Omega \\
A \cap B^c &= \{1, 3, 5\} = A \quad \text{os elementos de } \Omega \text{ que não estão no conjunto B} \Rightarrow B^c\{1, 3, 5\}
\end{aligned}$$

1.2 Probabilidade

Probabilidade - freqüência relativa associada a um variável descritora de uma população. Num espaço amostral Ω, a probabilidade de ocorrer um evento A, representado por $P(A)$, é dado pela medida de A em Ω nas seguintes condições: Exemplo: A probabilidade de ocorrer face ímpar no lançamento de um dado não viciado é

$$P(A) = \frac{n}{N} = \frac{3}{6} = \frac{1}{2} = 0,5 = 50\%$$

Algumas propriedades de probabilidade:

- A probabilidade de ocorrência de Ω vale 1, ou seja, $P(\Omega) = 1$
- Probabilidade de em evento certo e de um evento impossível

$$P(\Omega) = 1; \quad P(\emptyset) = 0$$

- A probabilidade de ocorrência do evento A é não negativa, ou seja, $P(A) \geq 0$
- Domínio da Probabilidade

$$0 \leq P(A) \leq 1$$

- Regra da Adição de probabilidades de dois eventos A e B:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

No exemplo do lançamento de um dado seja os eventos $A = \{2, 4, 6\}$ e $B = \{3, 4, 5, 6\}$. A união entre os dois conjuntos daria $\{2, 3, 4, 5, 6\}$. Assim:

$$P(A \cup B) = \frac{5}{6} = 0,83 = 83\%$$

Utilizando a regra da adição teriamos:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{3}{6} + \frac{4}{6} - \frac{2}{6} = \frac{5}{6} = 0,83 = 83\%$$

em que $A \cap B = \{4, 6\}$

- Probabilidade complementar

$$P(A^c) = 1 - P(A)$$

No exemplo do lançamento de um dado seja o evento $A = \{3, 4, 5, 6\}$, então $A^c = \{1, 2\}$, logo

$$P(A) = \frac{4}{6} \quad e \quad P(A^c) = \frac{2}{6}$$

utilizando a regra da probabilidade complementar teriamos:

$$P(A^c) = 1 - P(A) = 1 - \frac{4}{6} = \frac{6-4}{6} = \frac{2}{6}$$

1.2.1 Probabilidade Condicional e Independência de Eventos

A probabilidade condicional surge, por exemplo, quando se deseja calcular a probabilidade de um evento A ocorrer sabendo que um evento B já ocorreu.

Sejam A e B dois eventos associados a um mesmo espaço amostral Ω. Denota-se por $P(A|B)$ a probabilidade condicionada do evento A, quando o evento B tiver ocorrido.

Sempre que calculamos $P(A|B)$, estamos essencialmente calculando $P(A)$ em relação ao espaço amostral reduzido devido a B ter ocorrido, em lugar de faze-lo em relação ao espaço amostral original Ω.

Dados dois eventos A e B , a probabilidade condicional de A dado que ocorreu B é representada por $P(A|B)$ e definida por

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0.$$

Isso significa que a probabilidade de A ocorrer, dado que B ocorreu, é igual à probabilidade de ocorrência simultânea de A e B dividida pela probabilidade de ocorrência de B.

Exemplo: Na tabela a seguir temos dados referentes a alunos matriculados em três cursos de uma universidade em dado ano.

Tabela 1: Dados referentes a alunos de uma dada universidade.

Cursos	Sexo		Total
	Feminino	Masculino	
Administração	70	40	110
Psicologia	10	20	30
Geologia	20	15	35
Total	100	75	175

Qual a probabilidade de escolhermos um aluno ao acaso e ele ser:

- Homem (H) e da Administração (Adm)?

$$P(H \cap Adm) = \frac{40}{175} = 0,2285$$

b) Homem (H) ou da Administração (Adm)?

$$\begin{aligned} P(H \cup Adm) &= P(H) + P(Adm) - P(H \cap Adm) \\ &= \frac{75}{175} + \frac{110}{175} - \frac{40}{175} = \frac{145}{175} = 0,8285 \end{aligned}$$

- Psicologia (Psi) ou Geologia (Geo)?

$$\begin{aligned} P(Psi \cup Geo) &= P(Psi) + P(Geo) - P(Psi \cap Geo) \\ &= \frac{30}{175} + \frac{35}{175} - 0 = \frac{65}{175} = 0,3714 \end{aligned}$$

- De ser um aluno da psicologia dado que é mulher.

$$\begin{aligned} P(Psi|M) &= \frac{P(Psi \cap M)}{P(M)} \\ &= \frac{\frac{10}{175}}{\frac{100}{175}} = \frac{10}{175}\frac{175}{100} = \frac{10}{100} = 0,10 \end{aligned}$$

Das expressões acima resulta a **regra do produto**, que se refere ao cálculo da probabilidade do evento interseção,

$$P(A \cap B) = P(A|B).P(B)$$

A ordem do condicionamento pode ser invertida. Para três eventos, por exemplo, pode-se escrever:

$$P(A \cap B \cap C) = P(A).P(B|A).P(C|A \cap B) \tag{1}$$

Dois eventos A e B são **independentes** se a ocorrência de um não altera a probabilidade de ocorrência do outro, isto é, $P(A|B) = P(A)$ ou $P(B|A) = P(B)$, ou ainda, a seguinte forma equivalente:

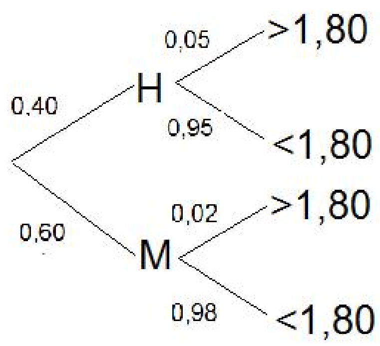
$$P(A \cap B) = P(A).P(B)$$

1.3 Árvores de probabilidade

A contrução de uma **árvore de probabilidade** fornece uma ferramenta muito útil para a solução de problemas envolvendo duas ou mais etapas. A árvore consiste em uma representação gráfica na qual diversas possibilidades são representadas, juntamente com as respectivas probabilidades condicionadas a cada situação. Isso permite, pela utilização direta da regra do produto das probabilidades, associar a cada nó terminal da árvore a respectiva probabilidade.

O uso das árvores de probabilidade ajudam e simplificam o entendimento da aplicação de dois teoremas que serão apresentados a seguir, conforme será visto no exemplo.

Exemplo: Em certo colégio, 5% dos homens e 2% das mulheres têm mais de 1,80m de altura. Por outro lado, 40% dos estudantes são homens. Sorteando-se um estudante aleatoriamente, qual a probabilidade de:



- Ser mulher (M) e ter mais de 1,80m?

$$P(M \cap > 1,80) = 0,60 \times 0,02 = 0,012$$

- Ter mais de 1,80m?

$$\begin{aligned} P(>1,80) &= P(M\cap >1,80) + P(H\cap >1,80) \\ P(H\cap >1,80) &= 0,40 \times 0,05 = 0,02 \\ P(>1,80) &= 0,012 + 0,02 = 0,032 \end{aligned}$$

- Um estudante é escolhido ao acaso e tem mais de 1,80m. Qual a probabilidade de que o estudante seja mulher?

$$P(M| >1,80) = \frac{P(M\cap >1,80)}{P(>1,80)} = \frac{0,012}{0,032} = 0,375$$

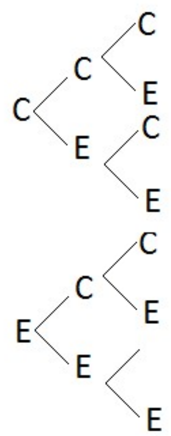
1.4 Variável Aleatória

Variável Aleatória - variável descritora de populações, cujos valores são associados a probabilidades de ocorrência.

Exemplo: Um estudante é submetido a três questões de múltipla escolha, em cada questão tinha cinco alternativas. Logo a chance de acerta uma questão no chute é 20$\%$

- Correto (C) - $P(C) = 20\% = \frac{1}{5}$
- Errado (E) - $P(E) = 80\% = \frac{4}{5}$

A questões e resultados possíveis são:



$$\Omega = \{CCC, CCE, CEC, CEE, ECC, ECE, EEC, EEE\}$$

Supondo que sua variável aleatória é acertar a questão, temos que o ocorrência no espaço amostral pode ser:

$$\Omega = \left\{\frac{CCC}{3}, \frac{CCE}{2}, \frac{CEC}{2}, \frac{CEE}{1}, \frac{ECC}{2}, \frac{ECE}{1}, \frac{EEC}{1}, \frac{EEE}{0}\right\}$$

As probabilidade dos pontos amostrais são:

$$\begin{aligned}
P(CCC) &= \frac{1}{5}\frac{1}{5}\frac{1}{5} = \frac{1}{125} \\
P(CCE) &= \frac{1}{5}\frac{1}{5}\frac{4}{5} = \frac{4}{125} \\
P(CEC) &= \frac{1}{5}\frac{4}{5}\frac{1}{5} = \frac{4}{125} \\
P(CEE) &= \frac{1}{5}\frac{4}{5}\frac{4}{5} = \frac{16}{125} \\
P(ECC) &= \frac{4}{5}\frac{1}{5}\frac{1}{5} = \frac{4}{125} \\
P(ECE) &= \frac{4}{5}\frac{1}{5}\frac{4}{5} = \frac{16}{125} \\
P(EEC) &= \frac{4}{5}\frac{4}{5}\frac{1}{5} = \frac{16}{125} \\
P(EEE) &= \frac{4}{5}\frac{4}{5}\frac{4}{5} = \frac{64}{125}
\end{aligned}$$

Pode-se construir uma tabela, em que X é o número de questões corretas e f(x) é a probabilidade de ocorrer o resultado X.

x	0	1	2	3
f(x)	64/125	48/125	12/125	1/125

Nesta tabela X assume os valores $(X = 0, 1, 2, 3)$ que são valores numéricos que descrevem os resultados da experiência, logo os valores de X são de uma variável aleatória.

Uma função que transforma em resultados de um espaço amostral em números reais, chama-se variável aleatória.

- X é o nome da variável aleatória definida. Ex. número de questões corretas;
- x são os valores assumidos pela variável. Ex. $x = 0, 1, 2, 3$.

1.5 Função de Probabilidade Discreta

É uma função $f(x)$ que associa a cada valor x da variável aleatória a sua respectiva probabilidade. Esta função deve atender duas condições:

1. $f(x) \geq 0$;
2. $\sum f(x) = 1$

Ex.: Para a três questões, considerando X número de acertos e x=(0,1,2,3)

x	0	1	2	3
f(x)	64/125	48/125	12/125	1/125

Verificação da duas condições:

1. $f(x) \geq 0$;

- Para $x < 0 \rightarrow f(x) = 0$
- Para $0 \leq x \leq 2 \rightarrow f(x) > 0$
- Para $x > 2 \rightarrow f(x) = 0$

2. $\sum f(x) = \dfrac{64}{125} + \dfrac{48}{125} + \dfrac{12}{125} + \dfrac{1}{125} = \dfrac{125}{125} = 1$

   Uma função de probabilidade discreta pode ser representada por

$$f(x) \text{ ou } P(x) \text{ ou } P(X = x)$$

Outra forma de representar uma distribuição de probabilidade de uma variável aleaória é por meio de sua função de distribuição acumulado, que é definida por

$$F(x) = P(X \leq x) = \sum_{i=1}^{n} P(X = x_i)$$

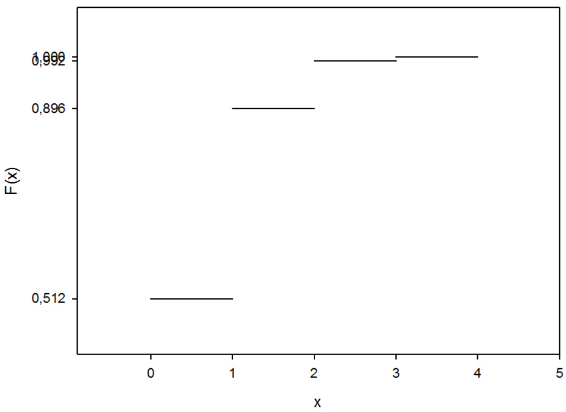
Utilizando o exemplo das questões, temos que a função de distribuição é

x	0	1	2	3
f(x)	64/125	48/125	12/125	1/125

Assim a função de distribuição acumulado é dada por

x	0	1	2	3
F(x)	64/125	112/125	124/125	125/125

E sua representação gráfica:



1.5.1 Esperança Matemática e Variância de uma VAD

Definição: Seja X uma V.A.D., com valores possíveis $x1, x2, ..., xn$; Seja $P(x_i) = P(X = x_i), i = 1, 2, ..., n$. Então, o valor esperado de X (ou Esperança Matemática de X), denotado por E(X) é definido como

$$E(X) = \sum_{i=1}^{\infty} x_i P(x_i)$$

esta expressão é também denominado o valor médio de X.

Definição: Seja X uma V.A.D. . Define-se a variância de X, denotada por $V(X)$ ou σ_X^2, da seguinte maneira:

$$V(X) = \sum_{i=1}^{\infty} (x_i - E(X))^2 \, P(x_i) = \text{ ou } V(X) = E(X^2) - (E(X))^2$$

e a raiz quadrada positiva de V(X) é denominada o desvio-padrão de X, e denotado por σ_X.

No exemplo das questões

$$\begin{aligned}
E(X) &= \sum_{i=1}^{4} x_i P(x_i) = 0\frac{64}{125} + 1\frac{48}{125} + 2\frac{12}{125} + 3\frac{1}{125} = 0 + \frac{48}{125} + \frac{24}{125} + \frac{3}{125} = 0,60 \\
V(x) &= \sum_{i=1}^{4} (x_i - E(X))^2 \, P(x_i) = (0-0,60)^2\frac{64}{125} + (1-0,60)^2\frac{48}{125} + (2-0,60)^2\frac{12}{125} + (3-0,60)^2\frac{1}{125} \\
&= 0,36\frac{64}{125} + 0,16\frac{48}{125} + 1,96\frac{12}{125} + 5,76\frac{1}{125} \\
&= \frac{23,04}{125} + \frac{7,68}{125} + \frac{23,52}{125} + \frac{5,76}{125} = \frac{60}{125} = 0,48 \\
V(X) &= E(X^2) - (E(X))^2 \\
E(X^2) &= \sum_{i=1}^{4} x_i^2 P(x_i) = 0^2\frac{64}{125} + 1^2\frac{48}{125} + 2^2\frac{12}{125} + 3^2\frac{1}{125} = 0\frac{64}{125} + 1\frac{48}{125} + 4\frac{12}{125} + 9\frac{1}{125} \\
&= 0 + \frac{48}{125} + \frac{48}{125} + \frac{9}{125} = \frac{105}{125} = 0,84 \\
V(X) &= 0,84 - (0,60)^2 = 0,84 - 0,36 = 0,48
\end{aligned}$$

1.6 Função de probabilidade contínua ou função de densidade de probabilidade (fdp).

Se a variável aleatória é contínua a sua função de probabilidade é uma função contínua conhecida por função de densidade de probabilidade (fdp). Esta função atende duas condições:

1. $f(x) \geq 0 \quad \forall x \in R$
2. $\displaystyle\int_R f(x)dx = 1$

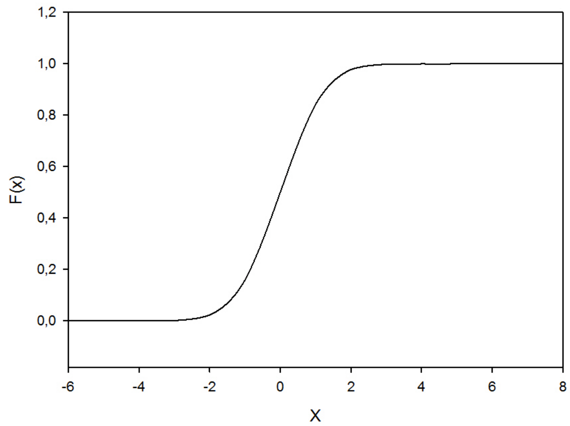
Das duas condições verifica-se que

$$P(a < x < b) = \int_a^b f(x)dx$$

No casa das variáveis contínuas a função de distribuição acumulada, que é definida por

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)dx$$

E sua representação gráfica:



Ex.: O tempo gasto, em minutos, por um estudante para responder a uma questão de um teste é uma variável aleatória contínua com função dada por

$$f(x) = \begin{cases} \frac{x}{4} & \text{para } 1 \leq x \leq 3 \\ 0 & \text{para outros valores} \end{cases}$$

Pela notação verifica-se que o estudante gasta um tempo entre 1 e 3 minutos.

Verificar as duas condições

1. $f(x) \geq 0 \quad \forall x \in R$
   - Para $x < 1 \rightarrow f(x) = 0$
   - Para $1 \leq x \leq 3 \rightarrow f(x) > 0$
   - Para $x > 3 \rightarrow f(x) > 0$
2. $\displaystyle\int_R f(x)dx = 1$

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} \frac{x}{4}dx = \int_1^3 \frac{x}{4}dx = \frac{1}{4}\int_1^3 xdx = \frac{1}{4}\frac{x^2}{2}\Big]_1^3 = \frac{1}{4}\left(\frac{3^2}{2} - \frac{1^2}{2}\right) = \frac{1}{4}\left(\frac{9}{2} - \frac{1}{2}\right) = \frac{1}{4}\frac{8}{2} = 1$$

Para obter a probabilidade utiliza-se a integral, por exemplo,

$$\begin{aligned} P(2 < x < 3) &= \int_2^3 \frac{x}{4}dx \\ &= \frac{1}{4}\int_2^3 xdx \\ &= \frac{1}{4}\frac{x^2}{2}\Big]_2^3 \\ &= \frac{1}{4}\left(\frac{3^2}{2} - \frac{2^2}{2}\right) \\ &= \frac{1}{4}\left(\frac{9}{2} - \frac{4}{2}\right) \\ &= \frac{1}{4}\frac{5}{2} = \frac{5}{8} = 0,625 \end{aligned}$$

1.6.1 Esperança Matemática e Variância de uma fdp

Definição: Seja X uma V.A. continua, com fdp $f(x)$. Então, o valor esperado de X (ou Esperança Matemática de X), denotado por E(X) é definido como

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

esta expressão é também denominado o valor médio de X.

Definição: Seja X uma V.A.D. . Define-se a variância de X, denotada por $V(X)$ ou σ_X^2, da seguinte maneira:

$$V(X) = \int_{-\infty}^{\infty} (x - E(X))^2 f(x) dx \text{ ou } V(X) = E(X^2) - (E(X))^2$$

em que

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$$

e a raiz quadrada positiva de V(X) é denominada o desvio-padrão de X, e denotado por σ_X.

No exemplo da o tempo gasto, em minutos, por um estudante para responder a uma questão de um teste, temos que:

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x f(x) dx = \int_{1}^{3} x \frac{x}{4} dx = 2,17 \\
V(X) &= \int_{-\infty}^{\infty} (x - E(X))^2 f(x) dx = \int_{1}^{3} (x - 2,17)^2 \frac{x}{4} dx = 9,70 \\
E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx = \int_{1}^{3} x^2 \frac{x}{4} dx = 0,30 \\
V(X) &= E(X^2) - (E(X))^2 = 5 - (2,17)^2 = 0,30
\end{aligned}$$