Probabilidade

Prof. Igor Nascimento

22 de abril de 2021

1 Experimento aleatório

O experimento aleatório está diretamente associado ao tipo de pesquisa.
Os tipos de pesquisa científica, relacionada à análise de dados são:

- experimental: medicamento é bom?
- observacional: chove em novembro?
- amostral: qual a intenção de voto no candidato A?

Em algumas circunstâncias os tipos de pesquisa podem ser combinadas.
As variáveis podem ser:

- quantitativa: discreta ou contínua
- qualitativa: nominal ou ordinal

Diante disso, no módulo anterior respondemos a seguinte pergunta:

- Qual é o padrão da pressão sistólica do brasileiro.
- Qual é o padrão do tempo de vida da lâmpada ABC.

O experimento aleatório é formado pelo espaço amostral Ω que é formado por elementos w, isto é, $w \in \Omega$. Os elementos de Ω podem ser enumerável, finito ou infinito, ou não enumerável.

O conjunto enumerável está relacionado aos números naturais N Magalhães (2006). Um exemplo de um conjunto enumerável finito pode ser obtido com o experimento aleatório da observação do local de parada do ponteiro analógico de um relógio, portanto, $\Omega = \{1, 2, 3, 4, ..., 60\}$. Já para os conjunto de enumerável infinito temos, por exemplo, o número de casos de dengue em um ano em uma determinada região do país, e portanto . $\Omega = \{0, 1, 2, 3, 4, ..., \infty\}$. Apesar de ser muito pouco provável que haja 1 milhão de casos, essa seria uma representação adequada para esse experimento aleatório James (2011). O espaço amostral é completamente dependente do experimento. Isso é notado pela diferença entre os espaços amostrais dos experimentos (a) número de pessoas que nascem com Progeria (doença rara) no país e (b) número de pessoas diagnosticas com influenza no país. Veremos mais adiante que modelos probabilísticos são utilizadas de acordo com a distribuição de frequência observado na realidade, isto é, os modelos são utilizados para reproduzir a realidade Morettin e Bussab (2017).

O espaço amostral não enumerável de um experimento aleatório é, normalmente, associado a medidas físicas de comprimento, área ou volume, por exemplo. Dessa forma, os elementos w estão associadas à reta real R. A observação do local de parada de um ponteiro digital é um exemplo de experimento aleatório com espaço amostral não enumerável.

Um conjunto A, contido em $A \subset \Omega$, elementos $w_i \in A$ é denominado de evento e é comumente representado por letras maiúsculas.

Considerando um experimento aleatório que possui um espaço amostral Ω é possível determinar a probabilidade de ocorrência dos elementos $w_i \in \Omega$ por meio do cálculo de probabilidades, que mensuram a incerteza associado ao resultado do experimento.

Isso pode ser feito em uma abordagem clássica, empírica ou subjetiva.

A abordagem clássica é a mais simples, sendo possível adotá-la quando Ω é formado por elementos w_i equiprováveis, isto é, são igualmente prováveis. Nesse caso, a probabilidade do evento A é calculada por:

$$P(A) = \frac{\text{Número de elementos em } A}{\text{Número de elementos em } \Omega} \tag{1.1}$$

É importante destacar que para isso precisamos conhecer o espaço amostral para determinar o valor do numerador e do denominador. Métodos de inumerações de permutação, combinação e arranjos são amplamente utilizados para isso, e serão abordados em seções futuras.

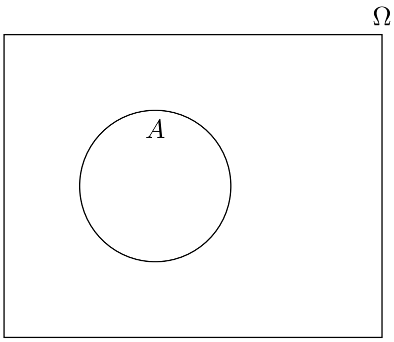
Quando não se conhece o espaço amostral, a abordagem empírica, também conhecida como frequentista, pode ser adotada. Nesse caso, a probabilidade do evento A é determinada pela frequência relativa da ocorrência de A (n_A) em n repetições do experimento aleatório associado.

$$P(A) \approx \frac{n_A}{n} \tag{1.2}$$

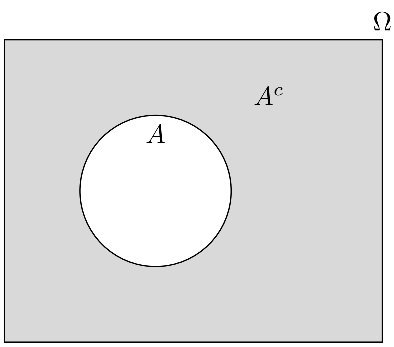
É importante que o resultado de cada repetição não interfira no resultado das demais. Esse é o conceito de independência de eventos, que detalharemos a seguir. A abordagem subjetiva determina a probabilidade de um evento A baseado no conhecimento prévio individual sobre o evento A.

2 Propriedades

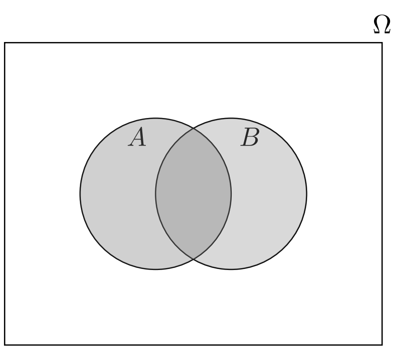
Para lidar com a medida de probabilidade é necessário definir algumas propriedades e características. Considere um evento A formado por um conjunto de elementos w de um espaço amostral Ω.



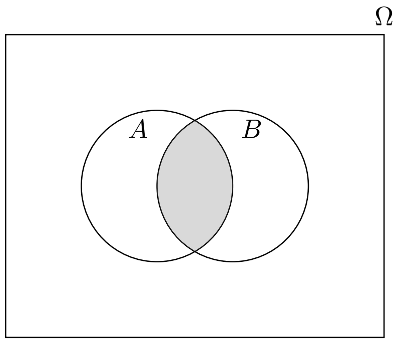
- $A =$, A é um conjunto vazio se não existe nenhum $w \in \Omega$ em A.
- A^c é o evento complementar de A, e representa tudo de Ω exceto aquilo que é A. Isto é, seja $w_i \in A^c$, então podemos garantir que w_i^c.



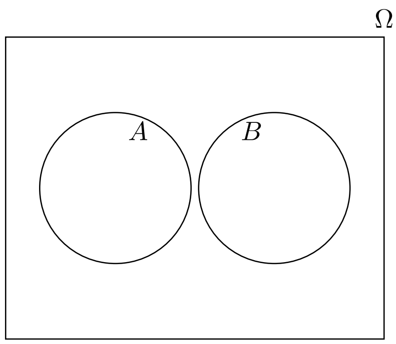
- $A \cup B$ é a união entre os eventos A e B. Dessa forma, a união desses eventos é compostar por elementos w, sendo $w \in A$ e/ou $w \in B$.



- $A \cap B$ é a intesercção entre os eventos A e B. Dessa forma, a intersecção desses eventos é compostar por elementos w, sendo $w \in A$ e $w \in B$.

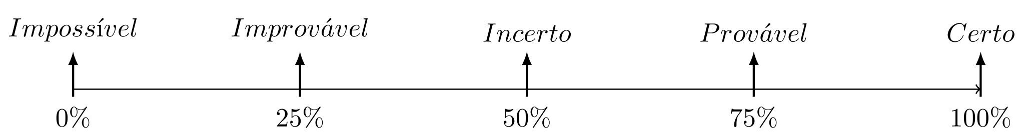


- A e B são disjuntos, ou mutualmente excludentes, se $w \in A$, então $w \notin B$ e se $w \in B$, então $w \notin A$. Dessa forma, $A \cap B =$



Por meio desses resultados e utilizando os axiomas de Kolmogorov, relacionados a probabilidade dos eventos de Ω Magalhães (2006), temos as seguintes propriedades:

- a probabilidade de qualquer evento $A \subset \Omega$ está limitada por $0 \leq P(A) \leq 1$.
- a probabilidade também pode ser apresentada em formato de percentual % e classificada da seguinte maneira:



- $P(\Omega) = 1$, a probabilidade de ocorrer algum evento em Ω é certa.
- A probabilidade de um evento impossível é 0: $p(\emptyset) = 0$
- A probabilidade da união de dois eventos A e B é (teoria dos conjuntos):

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

- Se dois eventos são mutualmente excludentes, ou seja, $A \cap B = \emptyset$, tem-se:

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

- Se dois eventos são independentes, então a probabilidade de ocorrer o evento A e B é:

$$P(A \cap B) = p(A) \times P(B)$$

- A probabilidade do evento A condicionada ao evento B é:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

2.1 Exemplos

Para compreender alguns desses axiomas considere os seguintes exempplos.

2.1.1 Conjuntos

Considere uma fábrica com 50 empregados. Um empregado não tem êxito em satisfazer os padrões de desempenho, se completa o (A) trabalho mais tarde e/ou monta produtos com (B) defeito. Foi observado que 5 dos 50 tinham completado o trabalho mais tarde, 6 dos 50 trabalhadores tinham montado peças defeituosas e 2 dos 50 tinham tanto completado mais tarde como montado produtos defeituosos. Quantos trabalhadores não satisfizeram os padrões de desempenho? Inicialmente, desejamos saber $A \cup B = A + B - A \cap B = 5 + 6 - 2 = 9$.

2.1.2 Espaço de probabilidade

Três cavalos A, B e C, estão em uma corrida; A: tem duas vezes mais probabilidades de ganhar que B, e B tem duas vezes mais probabilidade de ganhar que C. Quais são as probabilidades de vitória de cada um, isto é, P(A), P(B) e P(C)? Inicialmente, sabemos que $A + B + C = \Omega$ e, portanto $P(A) + P(B) + P(C) = P(\Omega) = 1$ e considerando que não é possível que dois cavalos sejam ganhadores ao mesmo tempo, $P(A \cap B \cap C) = P(A \cap B) = P(A \cap C) = P(B \cap C) = 0$. Além disso, $P(A) = 2P(B)$ e $P(B) = 2P(A)$, portanto, $P(A) = 2(2(P(C))) = 4P(C)$. Dessa forma, $4P(C) + 2P(C) + P(C) = P(C) = 1$. Com isso, $P(C) = 1/7$, $P(B) = 2/7$ e $P(A) = 4/7$.

2.1.3 Independência

A probabilidade de João (J) resolver uma questão de probabilidade e $P(J) = 2/3$ e a de Maria (M) $P(M) = 3/4$. Sabendo que ambos tentam de maneira independente, qual a probabilidade do problema ser resolvido? Inicialmente, desejamos saber $P(J \cup M) = P(J) + P(M) - P(J \cap M)$. Sabemos que $P(J \cap M) = P(J)P(M) = 2/3 \times 3/4 = 6/12$, pois J e M são independente. Logo, $P(J \cup M) = 2/3 + 3/4 - 1/2 = 11/12$.

3 Experimento de probabilidade

Imagine o apresentador do telejornal de sua preferência afirmar que, baseado nos especialistas, há uma previsão que a probabilidade de chuva para amanhã é de 90%. Considere que está marcado para amanhã um treino de corrida de rua com sua equipe. Você deveria planejar seu dia de amanhã contando com a corrida de rua? É razoável imaginar que, mesmo pouco conhecimento sobre probabilidade, que é muito provável que chova amanhã.

Considerando que essa resposta nem sempre é simples como nesse caso, a estatística e a matemática nos ajudam a utilizar modelos probabilísticos para realizar o cálculo de probabilidades e nos ajudar a tomar decisões Faber e Larson (2010).

Veja alguns exemplos de situações em que se pode estudar probabilidade.

Considere o lançamento de um dado de 6 lados honesto. Temos que:

- Experimento: lançamento do dado
- Resultado: observar o número da face virada para cima
- Espaço amostra: $\Omega = \{1, 2, 3, 4, 5, 6\}$
- Probabilidade clássica: 1/6

Considere o lançamento de uma moeda honesta. temos que:

- Experimento: lançamento da moeda
- Resultado: observar a face virada para cima
- Espaço amostra: $\Omega = \{cara, coroa\}$
- Probabilidae clássica: 1/2

Uma empresa está investigando se o trânsito é um problema no bairro. Já foram pesquisadas 500 pessoas, e 200 responderam que "o trânsito é um problema". Qual a probabilidade de que a próxima pessoa do bairro questionada responda que o trânsito é um problema?

- Experimento: pergunta individual
- Resultado: resposta individual
- Espaço amostra: $\Omega = \{$não é problema,é problema$\}$
- Probabilidade empírica: $200/500 = 2/5$

4 Probabilidade condicional

Para compreender a probabilidade condicional, vamos analisar tabelas de contigência. Para isso, vamos considerar o tabela 1, apresentada no trabalho de Carvalho et al. (2001) que investigou a relação entre sexo e as categorias de Índice de Massa Corpórea (IMC).

	HOMEM	MULHER	Total
NORMAL	110	148	258
MAGRO	5	5	10
SOBREPESO	38	28	66
Total	153	181	334

Figura 1 – Tabela de contigência apresentada em Carvalho et al. (2001)

Qual a probabilidade se selecionarmos uma mulher (M) nesse grupo investigado? Pela abordagem clássica, temos que:

$$P(M) = \frac{181}{334} = 54.2\% \tag{4.1}$$

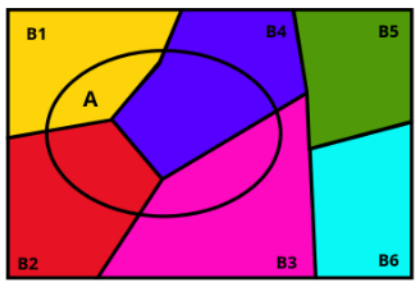
Qual a probabilidade se selecionarmos uma mulher, dentre os que estão em sobrepeso? Estamos restringindo o espaço amostral para aqueles que está em sobrepeso (S). Pela abordagem clássica e utilizando os axiomas de probabilidade, temos que:

$$P(M|S) = \frac{P(M \cap S)}{P(S)} = \frac{\frac{28}{334}}{\frac{66}{334}} = \frac{28}{66} \tag{4.2}$$

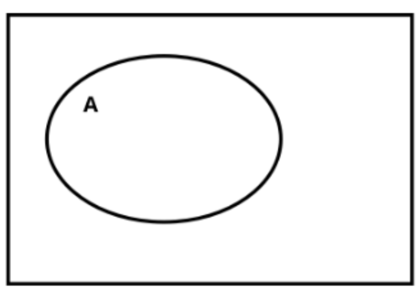
Dessa forma, a probabilidade condicional, como o nome sugere, condiciona o processo de cálculo de probabilidade.

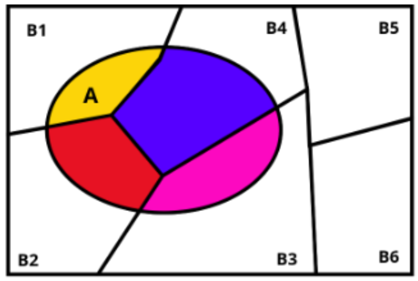
5 Regra da probabilidade total

A Lei da Probabilidade total é uma regra fundamental que relaciona as probabilidades conjuntas de um evento de interesse. Ela expressa a probabilidade total de um resultado que pode ser obtido por meio de vários eventos disjuntos (mutuamente exclusivos). Nos espaços amostrais Ω formados pela união das partes B_i disjuntas a probabilidade de qualquer evento em Ω.



A probabilidade do evento $P(A)$ pode ser calculada com ajuda de eventos disjuntos de Ω.





$$P(A) = P(A \cap B_1) \cup P(A \cap B_2) \cup ... \cup P(A \cap B_N)$$

Utilizando a probabilidadae condicional e o fato dos elementos B_i serem disjuntos:

$$P(A) = P(A|B_1)P(B_1) \cup P(A|B_2)P(B_2) \cup ... \cup P(A|B_N)P(B_N)$$

$$P(A) = \sum_{i=1}^{N} P(A|B_i)P(B_i)$$

Considere o seguinte exemplo. Suponha que duas fábricas sejam responsáveis pelo fornecimento de todas as lâmpadas do mercado. As lâmpadas da fábrica X trabalham por mais de 5000 horas em 99% dos casos, enquanto as lâmpadas de Y trabalham por mais de 5000 horas em 95% dos casos. Sabe-se que a fábrica X é reponsável por 60% das lâmpadas do mercado. Qual é a probabilidade de que uma lâmpada comprada dure mais de 5000 horas?

As lampadas só podem ser X ou Y, porem cada um dos tipos possuem durabilidades distintas. Sabendo que esse mercado de lâmpadas é disjunto, podemos calcular a probabilidade de duramar mais de 5000 horas P(D) por meio da regra da probabilidade total:

$$P(D) = P(D|X)P(X) + P(D|Y)P(Y) \tag{5.1}$$

$$P(D) = 99\% \times 60\% + 95\% \times 40\% = 97.4\% \tag{5.2}$$

Dessa forma, de maneira geral, as lâmpadas da região têm probabilidade 97.4% de durar mais de 5000 horas. Perceba que esse valor está entre o nível de X e o nível de Y.

6 Teorema de Bayes

O Teorema de Bayes utiliza probabilidades condicional e a Regra da probabilidade total para o cálculo da probabilidade de eventos.

A principal característica é atualizar o seu conhecimento sobre um evento baseado nas informações atuais.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B|A)}{P(B)} = P(A)\frac{P(B|A)}{P(B)}$$

A informação à *priori* P(A) é atualizada após o conhecimento sobre B $P(A|B)$. No teorema de Bayes temos que o valor da probabilidade a posteriori é o produto da probabilidade a priori e de uma razão de probabilidades da informação B.

Considere o seguinte exemplo. A probabilidade a priori dos eventos A e B são $P(A) = 0.4$ E $P(B) = 0.6$. Sabe-se que $P(A \cap B) = 0)$ e que que $P(R|A) = 0.2$ e $P(R|B) = 0.05$, R sendo um evento de interesse. Sabendo essa informação sobre R, qual é a probabilidade P(A/R) e P(B/R).

Nesse caso, temos a seguinte representação para o conhecimento a priori de A e B.

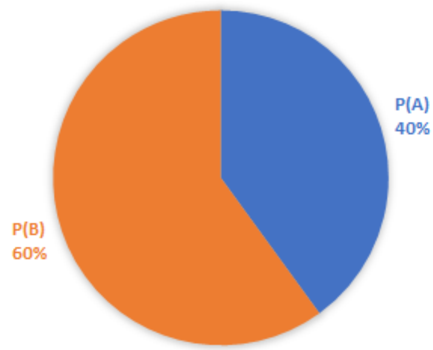


Figura 2 – Conhecimento sobre A e B antes de R.

Antes de tudo, vamos interpretar tais probabilidade. Perceba que o espaço amostral é formado por A e B de forma mutuamente exclusiva, sendo a maior parte, 60% relacionada a B. Além disso, vemos que quando estamos condicionados ao envento A, a probabilidade de R é 20%. Quando estamos em B, a probabilidade de R é 5%. Isso quer dizer, por exemplo, que é mais provável que R aconteça quando estamos condicionadas ao envento A.

A apesar de não sabermos a probabilidade P(R) sabemos qual é a probabilidade de R caso seja A ou B, e esses dois eventos representam Ω.

Pelas propriedades da probabilidade condicional, temos que $P(A \cap R)$:

$$P(R|A) = \frac{P(A \cap R)}{P(A)} \tag{6.1}$$

Assim, temos:

$$P(A \cap R) = P(R|A) \times P(A) = 0.2 \times 0.4 = 0.08 \tag{6.2}$$

Pelas propriedades da probabilidade condicional, temos que $P(B \cap R)$:

$$P(R|B) = \frac{P(B \cap R)}{P(B)} \tag{6.3}$$

Assim, temos:

$$P(B \cap R) = P(R|B) \times P(B) = 0.05 \times 0.6 = 0.03 \tag{6.4}$$

Pelas regras da probabilidade total, temos que:

$$P(R) = P(R|A)P(A) + P(R|B)P(B) = 0.08 + 0.03 = 0.11$$

Dessa forma, para encontrar a probabilidade de P(A—R), podemos utilizar o teorema de Bayes.

$$P(A/R) = P(A)\frac{P(R|A)}{P(R)} = 0.4\frac{0.2}{0.11} = 0.4 \times 1.81 = 0.7272$$

Dessa forma, a probabilidade inicial de P(A) era de 40%, no entanto, após saber que R aconteceu, a probabilidade associada ao evento A aumenta em 1.81 vezes. Isto é, a probabilidade de A é atualizada pelo o que conheçemos sobre R.

De forma análoga, pelas regras da probabilidade total, temos que:

$$P(R) = P(R|A)P(A) + P(R|B)P(B) = 0.08 + 0.03 = 0.11$$

Para encontrar a probabilidade de P(B—R), podemos utilizar o teorema de Bayes.

$$P(B|R) = P(B)\frac{P(R|B)}{P(R)} = 0.6\frac{0.05}{0.11} = 0.6 \times 0.4545 = 0.2727$$

Dessa forma, a probabilidade inicial de P(B) era de 60%, no entanto, após saber que R aconteceu, a probabilidade associada ao evento B reduz em 0.45 vezes. Isto é, a probabilidade de B é atualizada pelo o que conheçemos sobre R.

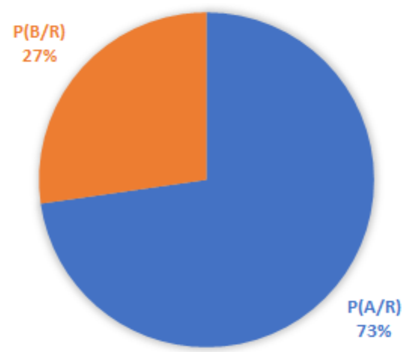


Figura 3 – Conhecimento sobre A e B após R.

De maneira geral, o evento B tinha maior probabilidade do que o de A, antes de sabermos sobre R. No entanto, como a probabilidade do evento R é maior quando relacionada ao evento A, ocorre um aumento da probabilidade de A.

Referências

CARVALHO, C. et al. Consumo alimentar de adolescentes matriculados em um colégio particular de teresina, piauí, brasil. 2001.

FABER, B.; LARSON, R. *Estatística Aplicada*. 4. ed. São Paulo: Pearson Prentice Hall, 2010.

JAMES, B. R. *Probabilidade: um curso em nível intermediário*. 3. ed. Rio de Janeiro: IMPA, 2011.

MAGALHÃES, M. N. *Probabilidade e variáveis aleatórias*. São Paulo: Edusp, 2006.

MORETTIN, P. A.; BUSSAB, W. O. *Estatística básica*. São Paulo: Editora Saraiva, 2017.