



Data Science Academy

www.datascienceacademy.com.br

Microsoft Power BI Para Data Science

Probabilidade Marginal
x
Probabilidade Condicional

Suponha que todos os 100 empregados de uma empresa tenham sido entrevistados para afirmarem se são favoráveis ou contrários ao pagamento de salários elevados aos Presidentes de empresas brasileiras. A tabela abaixo fornece uma classificação cruzada dessas respostas para esses 100 empregados. Suponha que todo empregado responda que é a favor ou contra.

| | A Favor | Contra |
|---------------|----------------|---------------|
| Homem | 15 | 45 |
| Mulher | 4 | 36 |

A tabela mostra a distribuição de 100 empregados, com base em duas variáveis ou características: sexo do indivíduo (homem ou mulher) e opinião (a favor ou contra). Esse tipo de tabela é denominado tabela de contingência. Na tabela, cada uma das caixas que contém um número é chamada célula. Observe que existem quatro células. Cada célula fornece a frequência para duas características. Por exemplo, 15 empregados nesse grupo apresentam duas características: “homem” e “a favor de se pagar salários elevados para os Presidentes”. Podemos interpretar os números de outras células utilizando o mesmo método.

Adicionando os totais de linhas e os totais de colunas à tabela, temos:

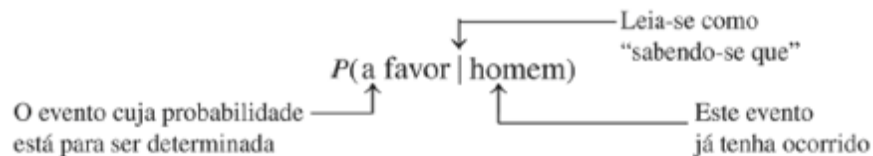
| | A Favor | Contra | Total |
|---------------|----------------|---------------|--------------|
| Homem | 15 | 45 | 60 |
| Mulher | 4 | 36 | 40 |
| Total | 19 | 81 | 100 |

Suponha que um empregado seja aleatoriamente selecionado, a partir desses 100 empregados. Esse empregado pode ser classificado simplesmente com base no sexo do indivíduo ou com base na opinião do indivíduo. Se apenas uma característica for considerada de cada vez, o empregado selecionado pode ser um homem, uma mulher, a favor ou contra. A probabilidade relativa a cada uma dessas quatro características, ou eventos, é denominada probabilidade marginal ou probabilidade simples. Essas probabilidades são conhecidas como probabilidades marginais uma vez que são calculadas dividindo-se as margens das linhas correspondentes (totais correspondentes às linhas) ou as margens das colunas (totais para as colunas) pelo total geral.

$$P(\text{homem}) = \frac{\text{Número de homens}}{\text{Número total de empregados}} = \frac{60}{100} = 0,60$$

Como podemos observar, a probabilidade de que um homem venha a ser selecionado é obtida dividindo-se o total da linha intitulada “Homem” (60) pelo total geral (100).

Agora, suponha que um empregado seja selecionado ao acaso entre esses 100 empregados. Suponha, também, que se saiba que esse empregado (selecionado) seja um homem. Em outras palavras, o evento de que o empregado selecionado seja um homem já ocorreu. Qual é a probabilidade de que o empregado selecionado seja a favor do pagamento de altos salários aos Presidentes? Essa probabilidade é escrita da seguinte maneira:



Essa probabilidade, $P(\text{a favor} \mid \text{homem})$ é chamada de probabilidade condicional de “a favor”, sabendo-se que o evento “homem” já tenha ocorrido. É lida como “a probabilidade de que o empregado selecionado seja a favor, sabendo-se que esse empregado é um homem”.