ActivALEA

active e actualize a sua literacia

N.º 32 – Independência de acontecimentos[1]
Tabelas de contingência
Probabilidade condicionada

Maria Eugénia Graça Martins
FCUL memartins@fc.ul.pt

Emília Oliveira
Escola Secundária de Tomaz Pelayo
ecmo.estp@gmail.com

Março 2025

Nesta ActivALEA aborda-se a independência de acontecimentos e a probabilidade condicionada, a partir de um exemplo/trabalho prático.

Uma turma de alunos do 12.º ano no âmbito do estudo do tema Probabilidade foi desafiada pela sua professora a desenvolver um trabalho de grupo que versasse este tópico.

O desafio foi aceite!

Um dos grupos de alunos lembrou-se que seria uma boa oportunidade para testar, junto dos colegas das outras turmas do 12.º ano, um jogo de computador que tinham desenvolvido na disciplina de Aplicações de Informática. É pressuposto que este jogo não envolva estratégia e só esteja dependente da "sorte" do aluno que joga.

Formalização da questão

O grupo de alunos de Aplicações de Informática desenvolveu um jogo de computador e pretende averiguar se o resultado do jogo, que pode ser **Ganha** ou **Perde**, está ou não relacionado com o facto de o aluno ser do sexo **Masculino** ou **Feminino**. Terão as raparigas mais habilidade do que os rapazes, para jogar o jogo? Será o contrário? Ou os resultados

[1] Esta tarefa foi adaptada de **Focus on Statistics** - *Investigations for the Integration of Statistics into Grades 9-12 Mathematics Classrooms*, Sara Brown**,** *Mathematics Institute of Wisconsin*, Patrick Hopfensperger
Retired, *University of Wisconsin-Milwaukee*, Henry Kranendonk, *Marquette University*, Copyright ©2020 by American Statistical Association, Alexandria, VA 22314-1943, página 139.

serão independentes do sexo? Para responder à questão, decidiram proceder à recolha da informação necessária junto dos colegas.

Planeamento da recolha de dados

O grupo de alunos que dirigia o trabalho organizou a recolha de dados junto dos 100 colegas do 12.º ano, combinando o dia em que lhes seria apresentado o jogo. A cada aluno foi pedido que, depois de jogar o jogo uma única vez, assinalasse na seguinte tabela (com um X) a situação verificada:

	Ganha	Perde
Masculino		
Feminino		

Apresentação dos dados recolhidos

Após todos os alunos terem assinalado a sua situação, o grupo encarregue do estudo resumiu os resultados das respostas dos colegas, obtendo a seguinte tabela:

	Ganha	Perde	Total
Masculino	11	29	40
Feminino	19	41	60
Total	30	70	100

Tabela de contingência

Uma **tabela de contingência**, de que a anterior é um exemplo, é uma tabela com **m** linhas e **n** colunas que se utiliza para organizar dados bivariados qualitativos, assim como para descrever a associação entre variáveis qualitativas ou quantitativas, cujos dados se apresentem categorizados (como, por exemplo, a variável Idade, quando se apresenta em categorias etárias).

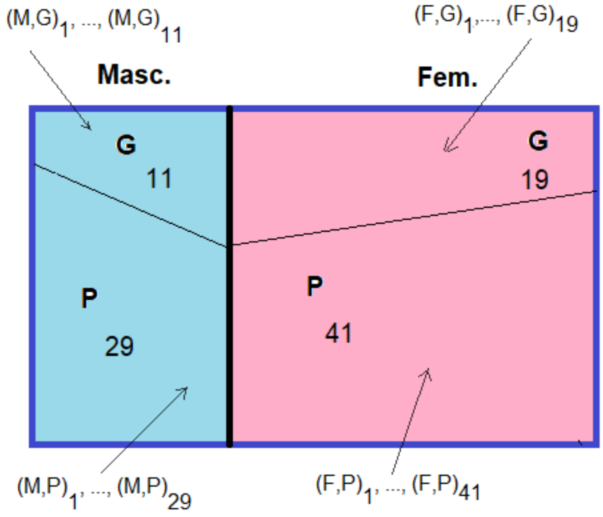
A tabela resume os elementos do **espaço de resultados** associado à experiência aleatória que consistiu em investigar, junto de cada um dos 100 alunos, o *Sexo* – M (Masculino) ou F (Feminino) e o Resultado do jogo – G (Ganha) ou P (Perde). O espaço de resultados, S, de dimensão 100, é dado pelo seguinte conjunto

$$S = \{(M,G)_1, (M,G)_2, \dots (M,G)_{11}, (M,P)_1, (M,P)_2, \dots, (M,P)_{29},$$

$$(F,G)_1, (F,G)_2, \dots, (F,G)_{19}, (F,P)_1, (F,P)_2, \dots, (F,P)_{41}\}$$

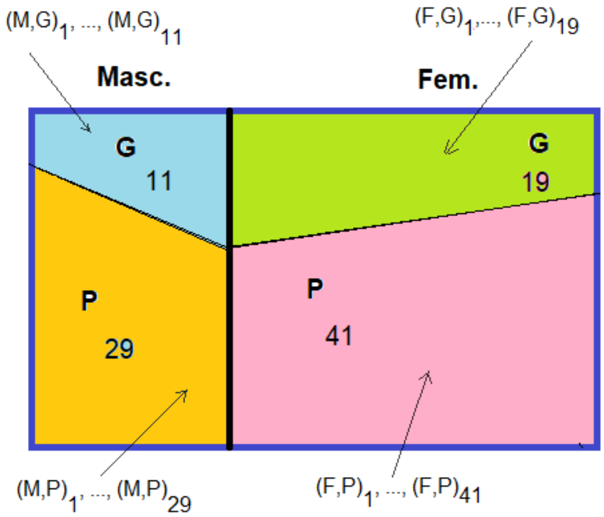
É usual utilizar **diagramas de Venn**[2] para visualizar o espaço de resultados e os **acontecimentos**, subconjuntos do espaço de resultados, como se apresenta a seguir:

[2] https://wikiciencias.casadasciencias.org/wiki/index.php/Diagrama_de_Venn



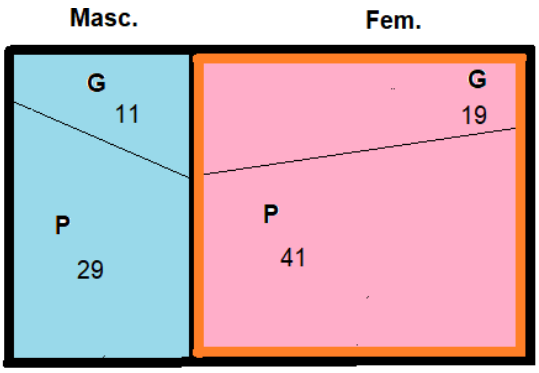
No diagrama anterior, o espaço de resultados é representado pelo retângulo maior e o acontecimento "sexo masculino" é constituído pelos 40 resultados $(M,G)_1,(M,G)_2,\ldots(M,G)_{11},(M,P)_1,(M,P)_2,\ldots,(M,P)_{29}$, correspondentes à parte pintada de azul. De forma análoga, o acontecimento "sexo feminino" é constituído pelos 60 resultados $(F,G)_1,(F,G)_2,\ldots(F,G)_{19},(F,P)_1,(F,P)_2,\ldots,(F,P)_{41}$, correspondentes à parte rosa.

Os acontecimentos "Sexo masculino e Perde" e "Sexo feminino e Ganha" são representados no diagrama seguinte, respetivamente na cor amarela e na cor verde:



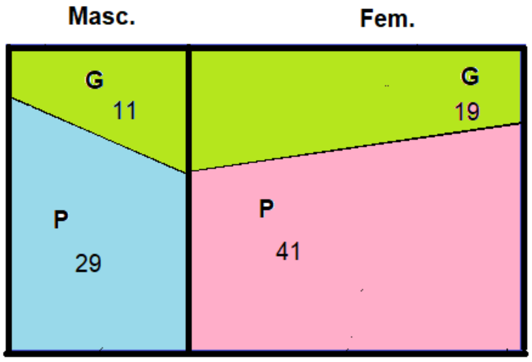
Algumas questões a serem respondidas pelos dados

1. Escolhendo, ao acaso, um aluno de entre os 100 que jogaram o jogo, qual a probabilidade de este ser do sexo feminino?



$P(\text{Feminino}) = \frac{60}{100} = 60\%$ (definição clássica ou de Laplace de Probabilidade)

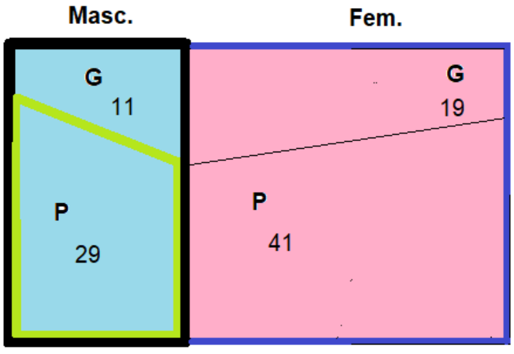
2. Escolhendo, ao acaso, um aluno de entre os 100 que jogaram o jogo, qual a probabilidade de este ganhar o jogo?



$P(\text{Ganha}) = \frac{11+19}{100} = 30\%$, (definição clássica ou de Laplace de Probabilidade)

3. Escolhendo, ao acaso, um aluno do sexo masculino, qual a probabilidade de este:

a) Ter perdido o jogo?



Repare-se que agora, como sabemos que o aluno é do sexo masculino, o conjunto dos resultados possíveis ficou restrito a esses alunos, sublinhado a negro no diagrama anterior, pelo que,

$$P(\text{Perde } \textbf{sabendo que} \text{ é Masculino}) = \frac{\text{Nº resultados) da forma } (M,P)}{\text{Nº de resultados da forma } (M,G) \text{ ou } (M,P)} = \frac{29}{40} = 72{,}5\% \quad (1)$$

Nota: Repare-se que a partir de (1), podemos escrever

$$P(\text{Perde } \textbf{sabendo que} \text{ é Masculino}) = \frac{\text{Nº resultados da forma } (M,P)}{\text{Nº de resultados da forma } (M,G) \text{ ou } (M,P)}$$

$$= \frac{\dfrac{\text{Nº de resultados da forma } (M,P)}{\text{Nº de resultados do espaço de resultados}}}{\dfrac{\text{Nº de resultados da forma } (M,G) \text{ ou } (M,P)}{\text{Nº de resultados do espaço de resultados}}}$$

$$= \frac{\frac{29}{100}}{\frac{40}{100}}$$

$$= \frac{\text{P(Masculino} \cap \text{Perde)}}{\text{P(Masculino)}}$$

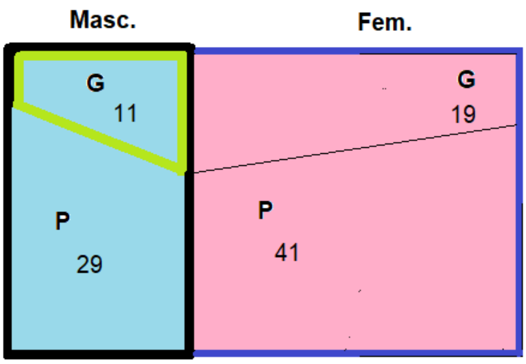
ou seja,

$$P(\text{Perde } \textbf{sabendo que} \text{ é Masculino}) = \frac{P(\text{Masculino } \cap \text{ Perde})}{P(\text{Masculino})}$$

De um modo geral, dados os acontecimentos A e B, com $P(B) > 0$, tem-se

$$P(\text{A } \textbf{sabendo que} \text{ B}) = \frac{P(A \cap B)}{P(B)}$$

b) Ter ganho o jogo?



$$P(\text{Ganha } \textbf{sabendo que} \text{ é Masculino}) = \frac{11}{40} = 27{,}5\%$$

Nota – Repare-se que em 2. se obteve, para a probabilidade de um aluno escolhido ao acaso,

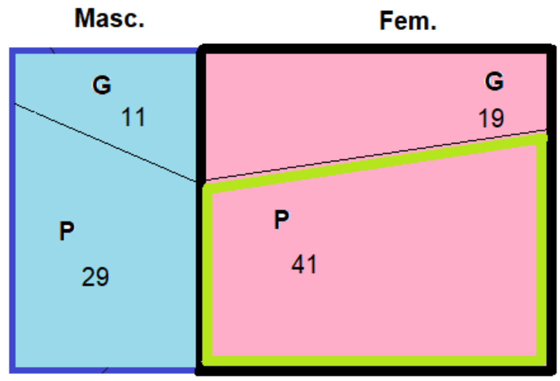
$$P(\text{ganha}) = 30\%$$

Com a informação adicional de que o aluno é do sexo masculino, tem-se

$$P(\text{Ganha } \textbf{sabendo que} \text{ é Masculino}) = 27{,}5\%$$

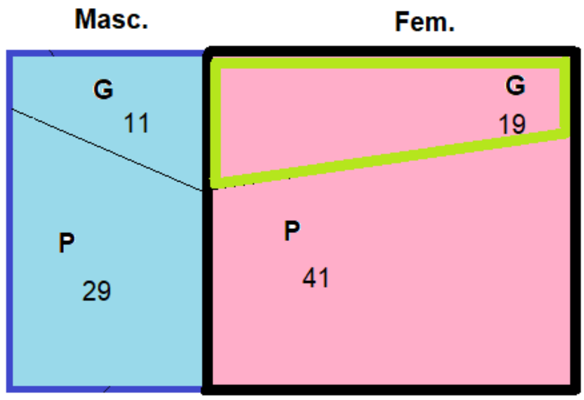
4. Escolhendo, ao acaso, um aluno do sexo feminino, qual a probabilidade de este:

a) Ter perdido o jogo?



$$P(\text{Perde } \textbf{sabendo que} \text{ é Feminino}) = \frac{41}{60} \approx 68{,}3\%$$

b) Ter ganho o jogo?



$$P(\text{Ganha } \textbf{sabendo que} \text{ é Feminino}) = \frac{19}{60} \approx 31{,}7\%$$

Probabilidade condicionada

A probabilidade de um acontecimento A, **sabendo que** se verificou outro acontecimento B, $P(B) > 0$, chama-se **probabilidade condicionada** ou **condicional**, representa-se por $P(A|B)$ e calcula-se a partir da expressão

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

Assim, utilizando esta notação, as probabilidades calculadas anteriormente podem ser escritas assim:

$P(\textbf{Perde}|\textbf{Masculino}) = 72{,}5\%$ $\quad$ $P(\textbf{Ganha}|\textbf{Masculino}) = 27{,}5\%$

$P(\textbf{Perde}|\textbf{Feminino}) = 68{,}3\%$ $\quad$ $P(\textbf{Ganha}|\textbf{Feminino}) = 31{,}7\%$

5. Tendo em conta os resultados obtidos nos pontos anteriores, pode-se construir a seguinte tabela de probabilidades condicionadas ou condicionais:

	Probabilidade condicionada de Ganhar tendo em conta o sexo	**Probabilidade condicionada de Perder tendo em conta o sexo**	**Total**
Masculino	27,5%	72,5%	100%
Feminino	31,7%	68,3%	100%
Total	30%	70%	100%

6. Interpretação da tabela de probabilidades condicionadas:

6.1 Se o facto de ganhar o jogo for totalmente devido ao acaso e não estiver relacionado com o sexo, qual a probabilidade de um aluno, escolhido ao acaso, ganhar o jogo?

Como se calculou em 2., tem-se

$$P(\text{Ganha}) = 30\%$$

Repare-se que esta probabilidade aparece na última linha da tabela, já que não está condicionada pelo sexo (Masculino ou Feminino). De forma análoga, tem-se

$$P(\text{Perde}) = 70\%$$

6.2 Se a probabilidade de ganhar o jogo for totalmente devida ao acaso, qual a probabilidade de um aluno ganhar, **sabendo-se que** é do sexo masculino?

Da tabela das probabilidades condicionais construída em 5., vem

$$P(\text{Ganha}|\text{Masculino}) = 27{,}5\%$$

6.3 Se a probabilidade de ganhar o jogo for totalmente devida ao acaso, qual a probabilidade de um aluno ganhar, **sabendo-se que** é do sexo feminino?

Da tabela das probabilidades condicionais construída em 5., vem

$$P(\text{Ganha}|\text{Feminino}) = 31{,}7\%$$

Acontecimentos independentes

Dois acontecimentos A e B, com $P(A) > 0$ e $P(B) > 0$, dizem-se independentes quando a ocorrência de um deles não altera a probabilidade de ocorrência do outro, ou seja,

$$P(A|B) = P(A) \;\text{ ou }\; P(B|A) = P(B).$$

6.4 Tendo em consideração a definição de independência de acontecimentos dada anteriormente, serão os acontecimentos Ganha e Masculino **independentes**?

Pelas alíneas anteriores, verificou-se que $P(\text{Ganha}|\text{Masculino}) = 27{,}5\%$ e $P(\text{Ganha}) = 30\%$, pelo que $P(\text{Ganha}|\text{Masculino}) \neq P(\text{Ganha})$, concluindo-se que os acontecimentos Ganha e Masculino **não são independentes**.

O que significa os acontecimentos não serem independentes? Significa que, quando se seleciona um aluno ao acaso e se calcula a probabilidade de que ele **G**anha, essa probabilidade é 30%, mas, se tivermos o conhecimento de que ele é do sexo **M**asculino, essa probabilidade **alterou-se** e diminuiu para $27{,}5\%$, significando que o facto de ganhar não é **independente** do sexo.

E agora...?

O grupo encarregue do trabalho pretendeu aprofundar um pouco mais este tema e, com a ajuda da professora, arranjou uns dados fictícios, admitindo que seriam o resultado das respostas dos 100 colegas. Esses dados são apresentados na seguinte tabela:

	Ganha	Perde	Total
Masculino	12	28	40
Feminino	18	42	60
Total	**30**	**70**	**100**

Será que se as respostas fossem as da tabela anterior, os acontecimentos

Ganha e Masculino seriam independentes?

$$P(\text{Ganha}) = \frac{30}{100} = 30\%$$

$$P(\text{Ganha}|\text{Masculino}) = \frac{12}{40} = 30\%$$

Como se verifica, o conhecimento de que um aluno, selecionado ao acaso, era do sexo Masculino, não alterou a probabilidade de que Ganha, pelo que os acontecimentos são **independentes**, ou seja, o facto de que um aluno Ganha é independente de ser Masculino ou Feminino.