



N.º 26 – A FREQUÊNCIA RELATIVA PARA ESTIMAR A PROBABILIDADE

Por: Maria Eugénia Graça Martins
Departamento de Estatística e Investigação Operacional da FCUL
memartins@fc.ul.pt

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

INTRODUÇÃO

O termo Probabilidade é utilizado diariamente de forma mais ou menos intuitiva, pois, nos mais variados aspetos da nossa vida, está presente a **Incerteza**.

Nesta ActivALEA pretende-se mostrar como se pode estimar a probabilidade de um acontecimento se realizar, utilizando a frequência relativa.

QUAL A PROBABILIDADE DE...

O Manuel frequenta o 9.º ano de escolaridade da escola básica 2,3 da *Minha Terra* e gosta muito da disciplina de Matemática. É mesmo a sua disciplina preferida! No entanto, nem todos os colegas têm o mesmo gosto. O Manuel gostaria de saber **qual a probabilidade de**, escolhido um colega da escola ao acaso, ele também **preferir a Matemática**.

Como ajudar o Manuel a resolver esta situação?

1. Introdução

Antes de respondermos à questão colocada pelo Manuel, vejamos outras situações mais vulgares em que usamos o termo **probabilidade**.

- Lançando uma moeda de um euro ao ar, a probabilidade de a face Nacional ficar voltada para cima é $\frac{1}{2}$. Significa que, lançando várias vezes a moeda ao ar, esperamos obter a face Nacional em 50% das vezes! Esta situação de lançar a moeda é utilizada com frequência quando temos de tomar uma decisão relativamente a duas alternativas A e B. Se não tivermos a mínima ideia sobre qual a alternativa escolher... lançamos a moeda e, se sair a face Nacional, escolhemos A; caso contrário, escolhemos B!
- Se em vez de uma moeda se lançar um dado de 4 faces (tetraedro) numeradas de 1 a 4, qual a probabilidade de sair a face com o número 3 quando se lança o dado? Se o dado tem 4 faces e se, à partida, não há nenhuma razão para qualquer uma das faces ser preferida, relativamente às outras, então a probabilidade de sair a face com o número 3 é $\frac{1}{4}$, ou seja, se lançarmos o dado muitas vezes espera-se que saia a face com o número 3 em 25% das vezes, pois há 4 faces todas com igual possibilidade de sair, como dissemos anteriormente. Uma situação em que este dado poderia ser útil é, por exemplo, numa questão com quatro respostas possíveis, numeradas de 1 a 4, das quais só uma é que está certa. Respondendo ao "acaso", para decidir qual a resposta a escolher, poderíamos lançar o dado e, de acordo com o número saído, decidíamos qual a resposta. A probabilidade de responder corretamente seria $\frac{1}{4}$.

Nos exemplos anteriores foi fácil atribuir um valor à probabilidade de sair a face Nacional ou sair a face 3, respetivamente no lançamento da moeda ou do dado de 4 faces. Mas nem sempre as situações são tão simples.

Voltemos então à questão colocada inicialmente e, só para simplificar, suponhamos que cada aluno tem apenas 6 disciplinas: Língua Portuguesa (LP), Língua Estrangeira (LE), História (H), Geografia (G), Matemática (M) e Ciências Naturais (CN). O Manuel pretende averiguar a probabilidade de a Matemática ser a disciplina preferida de um colega, escolhido ao acaso. Será que, como no exemplo anterior, se poderia pegar num dado de 6 faces, escrever o nome de cada disciplina em cada uma das faces e, seguidamente, em vez de se perguntar



ao aluno qual a sua disciplina preferida, lançar-se o dado e considerar-se como disciplina preferida a que estiver escrita na face que fica virada para cima?

Desta forma, a probabilidade da disciplina preferida ser a Matemática (ou qualquer outra das 6 disciplinas) seria $\frac{1}{6}$. Será que este procedimento está correto? **NÃO!** Só estaria correto se o aluno respondesse ao acaso, indicando uma qualquer das 6 disciplinas! Mas, na realidade, as coisas não se passam assim... As disciplinas não estão todas em igualdade de circunstâncias, pois cada aluno tem alguma ou algumas disciplinas preferidas. Vamos partir do princípio de que cada aluno tem apenas uma disciplina preferida.

Então, como proceder?



2. A frequência relativa para estimar a probabilidade

Uma vez que o procedimento anterior não está correto, uma solução será averiguar junto de alguns alunos qual a sua disciplina preferida e depois estimar a probabilidade de a disciplina favorita ser a Matemática, através da **frequência relativa** ou percentagem de respostas que a indicam como preferida.

Por exemplo: se, numa amostra de 15 alunos, 6 deles disserem que preferem a Matemática, a frequência relativa atribuída à Matemática como disciplina preferida é de $6/15 = 0,4$, ou seja, 40%.

Este valor obtido para a frequência relativa é considerado uma **estimativa** para a **probabilidade** de um aluno, escolhido ao acaso, preferir a Matemática.

De um modo geral, estimamos a probabilidade de um determinado acontecimento se realizar pela frequência relativa com que esse acontecimento se verifica numa série de repetições da experiência, que conduza a que o acontecimento se possa realizar ou não.

Na situação anterior, a realização da experiência que conduzia a que o acontecimento se realizasse ou não consistia em perguntar a um aluno escolhido ao acaso qual a disciplina favorita. O acontecimento realizava-se sempre que a resposta fosse "Matemática". Esta experiência diz-se aleatória, pois, antes de fazermos a pergunta, não sabemos a resposta, embora saibamos que será uma de 6 possíveis. As respostas dos alunos inquiridos constituem a amostra utilizada para estimar a probabilidade do acontecimento "a disciplina preferida é a Matemática". Ao número de respostas dadas, igual ao número de alunos inquiridos, chamamos **dimensão da amostra**.

Se dissermos que temos um sucesso sempre que o acontecimento se realiza, então uma estimativa para a probabilidade do acontecimento se realizar é dada pela frequência relativa

$$\frac{\text{numero de sucessos}}{\text{dimensao da amostra}}$$

A estimativa será tanto melhor quanto maior for a dimensão da amostra.

No caso presente, quanto mais alunos forem inquiridos, melhor será a estimativa para a probabilidade da disciplina preferida ser a Matemática.

3. Como ajudar o Manuel?

Para ajudar o Manuel, o professor de Matemática decidiu planificar na turma a recolha da informação necessária para conseguir dar uma resposta à questão colocada pelo Manuel. Assim, como na escola havia 15 turmas do 9.º ano, sem contar com a turma do Manuel, o professor organizou 15 grupos (de 1 ou 2 alunos), em que cada grupo iria perguntar a 10 alunos escolhidos ao acaso, de cada uma das 15 turmas, qual a sua disciplina favorita. Para facilitar a recolha de informação, cada grupo tinha uma folha com 10 quadrados – tantos

quantos o número de alunos a inquirir - e em cada quadrado escrevia 1 ou 0, conforme a disciplina preferida fosse a Matemática ou outra qualquer:

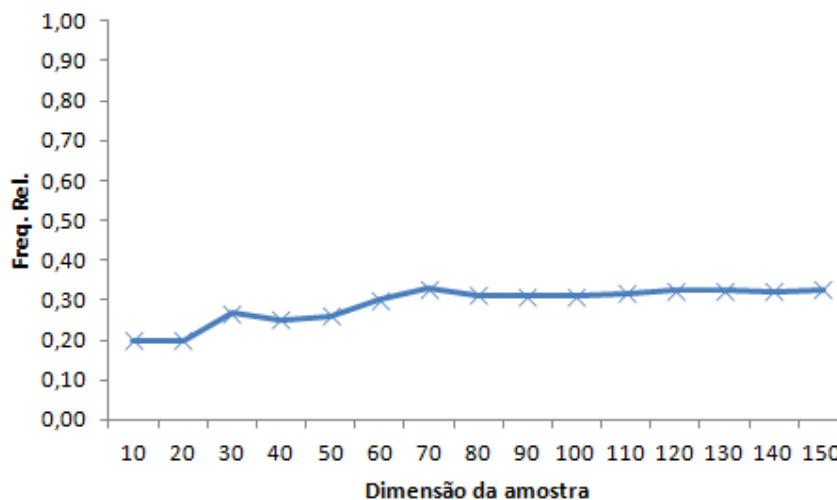
1.º	2.º	3.º	4.º	5.º	6.º	7.º	8.º	9.º	10.º	Total

Depois de todos terem feito a recolha da informação necessária, na aula iriam resumir os dados recolhidos numa tabela que o professor tinha construído para o efeito (na tabela, representa-se por sucesso a resposta "Matemática"):

Grupo	Dim. Amostra	N.º sucessos	Dim amostra acumulada	N.º sucessos acumulados	Freq. rel. (N.º sucessos acum./Dim. amostra acum)
1	10	2	10	2	0,20
2	10	2	20	4	0,20
3	10	4	30	8	0,27
4	10	2	40	10	0,25
5	10	3	50	13	0,26
6	10	5	60	18	0,30
7	10	5	70	23	0,33
8	10	2	80	25	0,31
9	10	3	90	28	0,31
10	10	3	100	31	0,31
11	10	4	110	35	0,32
12	10	4	120	39	0,33
13	10	3	130	42	0,32
14	10	3	140	45	0,32
15	10	4	150	49	0,33

Se anteriormente se disse que a frequência relativa com que o acontecimento se realiza nos dá uma estimativa para a probabilidade do acontecimento se realizar, então neste momento temos 15 estimativas e não sabemos qual considerar! **Sabemos sim...**, pois também se disse que, quanto maior for a dimensão da amostra, melhor será a estimativa da probabilidade. Assim, a estimativa que vamos considerar para a probabilidade de um colega do Manuel preferir a Matemática é, aproximadamente, 33%.

Para melhor visualizar a evolução da frequência relativa, à medida que a dimensão da amostra aumenta, podemos construir um gráfico de linhas como o que se apresenta a seguir:





4. Conclusão

Do que se disse anteriormente, é importante fixar o seguinte: o procedimento utilizado não permite dizer que 33% é a probabilidade de um colega do Manuel, escolhido ao acaso, preferir a Matemática. O que se obteve foi um valor aproximado para essa probabilidade.

Mas, então, nunca se consegue obter um valor que possamos dizer que é a probabilidade do acontecimento se realizar? Consegue-se, com algum esforço... Neste caso que estamos a analisar, teria de se perguntar a "todos" os colegas do Manuel qual a sua disciplina preferida e no final calcular a frequência relativa

$$\frac{\textit{número de sucessos}}{\textit{total de alunos}}$$

Admitindo que o total de alunos era 560 e que o número de sucessos era 196 (respostas indicando a Matemática como preferida), então a probabilidade de um aluno da escola, escolhido ao acaso, preferir a Matemática seria 0,35 (196/560) ou 35%.

A frequência relativa, afinal, utiliza-se com mais frequência do que se pensa!

Como, de um modo geral, a *população* de onde se recolhe a *amostra* para obter dados que nos permitam estimar a probabilidade de certo acontecimento se realizar tem dimensão grande, contentamo-nos com uma estimativa dessa probabilidade. O que temos de ter presente é que, quanto maior for a dimensão da amostra, melhor será a estimativa.

Este procedimento que acabámos de descrever, em que se recolhe uma amostra e se utiliza a frequência relativa para estimar probabilidades, é o processo utilizado pelas empresas que fazem sondagens.

