# A Visão Frequencista

Rui Santos *rui.santos@ipleiria.pt*, Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Leiria, CEAUL — Centro de Estatística e Aplicações da Universidade de Lisboa

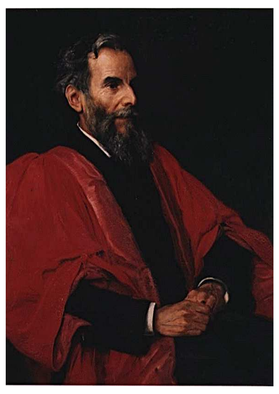
Trabalho financiado por Fundos Nacionais através da Fundação para a Ciência e a Tecnologia, no âmbito do projeto PEst-OE/MAT/UI0006/2011.

Na visão frequencista uma probabilidade só pode ser atribuída em fenómenos que podem ser repetidos um grande número de vezes nas mesmas condições, como tal, na estatística clássica (ou frequencista) é através da informação obtida na observação repetida de um fenómeno que podemos inferir acerca das suas características.

## A definição frequencista de probabilidade

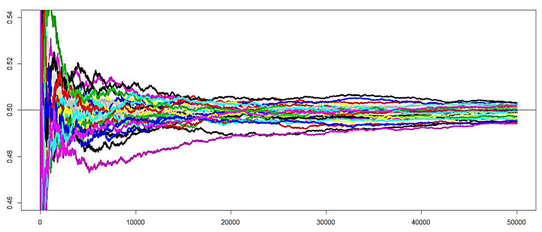
A utilização da frequência relativa como aproximação da probabilidade é uma consequência da aplicação da Lei dos Grandes Números, tendo sido usada, sem a preocupação de estabelecer fundamentos da inferência, desde os finais do século XVII, por exemplo nas tabelas de mortalidade de Edmond Halley (1656−1742) em 1693, ou no interessante trabalho de John Arbuthnot (1667−1735) em 1710, no qual conclui que regularmente nascem mais rapazes do que raparigas e que esta diferença não advém do acaso mas antes da vontade Divina.

Embora seja questionável a identificação do primeiro a introduzir a ideia de definir a probabilidade em termos de um limite das frequências relativas, muitos identificam a sua origem com os trabalhos de Robert Leslie Ellis (1817−1859) em 1843, nos quais considera que, numa longa série de testes, a frequência com que ocorre cada acontecimento tende a ser proporcional à sua probabilidade. Há também quem a identifique com os trabalhos de Bernhard Bolzano (1781−1841), Antoine Augustin Cournot (1801−1877) ou John Stuart Mill (1806−1873) todos com publicações em 1843 nas quais expõem interpretações frequencistas de probabilidade.



Venn

Todavia, o grande impulsionador desta interpretação foi John Venn (1834−1923), principalmente no seu livro *The Logic of Chance* publicado em 1866, no qual a probabilidade de um acontecimento é determinada observando a frequência relativa desse acontecimento numa sucessão numerável de experiências aleatórias idênticas e independentes, isto é, considerando que efetuamos uma experiência aleatória $n$ vezes onde o acontecimento $\mathbf{A}$ ocorre $n_{\mathbf{A}}$ vezes, a sua frequência relativa $f_n(\mathbf{A})$ será dada pelo rácio $\frac{n_{\mathbf{A}}}{n}$ e, caso haja limite deste rácio quando $n$ aumenta indefinidamente, então a probabilidade do acontecimento $\mathbf{A}$ é definida por:



$$\mathbb{P}(\mathbf{A}) = \lim_{n\to+\infty} \frac{n_{\mathbf{A}}}{n} = \lim_{n\to+\infty} f_n(\mathbf{A}).$$

**Figura**: Evolução da proporção de sucessos em 30 sequências de 50000 provas, inicialmente com alguma variabilidade, mas depois convergindo todas para o mesmo limite (para a probabilidade)

Este limite não tem o significado habitual utilizado em análise matemática, mas antes o da convergência em probabilidade, usualmente representada por $f_n(\mathbf{A}) \xrightarrow{\mathbb{P}} \mathbb{P}(\mathbf{A})$, que significa que a probabilidade da distância entre $f_n(A)$ e $\mathbb{P}(A)$ ser maior que um dado $\varepsilon \in \mathbb{R}^+$ converge para zero.

Entre os mais notáveis defensores desta interpretação destacam-se Richard von Mises (1883−1953), Ronald Fisher (1890−1962), Hans Reichenbach (1891−1953), Karl Raimund Popper (1902−1994) e Abraham Wald (1902−1950).

## A formalização de Richard von Mises

*"Since the first task of any scientific endeavor is a limitation of its scope, we limit our scope, roughly speaking, to the mathematical theory of repetitive events. We do not attempt to discuss in this book ideas and problems of subjective probability".* [von Mises, 1964]

Richard von Mises, em 1919, apoiando-se na interpretação de probabilidade desenvolvida por Venn, constrói uma axiomática para a Teoria da Probabilidade com base no conceito de *coletivo*. Von Mises representa cada resultado do espaço-amostra (conjunto constituído pelos possíveis resultados da experiência aleatória) por um número e considera a sequência formada pelos resultados, de uma experiência repetida indefinidamente, um *coletivo* no caso de se verificarem os seguintes postulados:

1. *Existência de limite* — as frequências relativas de um coletivo convergem para um limite.
2. *Aleatoriedade* — as frequências relativas de qualquer subsequência da sequência original têm o mesmo limite.

Von Mises

Para von Mises a probabilidade só existe quando aplicada a *coletivos*, sendo definida pelo limite das frequências relativas. Assim sendo, a probabilidade só está definida em experiências que podem ser repetidas indefinidamente e nunca em acontecimentos únicos. O segundo postulado (também denominado *princípio da exclusão do sistema de jogo*) garante a aleatoriedade dos resultados, pois, no caso de existência de regularidades, poder-se-iam prever alguns resultados — por exemplo, se num jogo de *cara* (0) ou *coroa* (1) os resultados fossem sempre alternados, teríamos uma sequência $0101010\cdots$ e as frequências relativas do acontecimento *cara* teriam limite $\frac{1}{2}$, porém os limites das subsequências de lançamentos pares (ou ímpares) não têm o mesmo limite, logo não pode ser considerado um *coletivo* e consequentemente a probabilidade não está definida. Estes axiomas foram muito criticados por diversos autores, no entanto, a Teoria Frequencista de Probabilidade foi desenvolvida de forma a ser, na atualidade, uma das mais influentes.

## Estatística clássica/frequencista

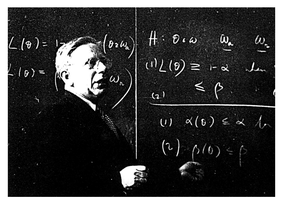
Se considerarmos que uma amostra é uma sequência retirada de um *coletivo*, a primeira condição garante a homogeneidade distribucional (que as variáveis aleatórias que compõem a amostra tenham a mesma distribuição marginal) e a segunda condição garante a aleatoriedade dos resultados (independência entre as variáveis). Desta forma, nesta visão de von Mises, numa amostra as variáveis aleatórias são independentes e identicamente distribuídas, o que torna o tratamento matemático da função densidade conjunta mais elementar, pois as hipóteses adjacentes, independência e homogeneidade, simplificam bastante a sua análise.

Na estatística clássica a base são as distribuições amostrais, distribuições de uma função da amostra aleatória — $h(X_1, X_2, \cdots, X_n)$. Os estimadores, estatísticas de testes ou preditores são funções $h(X_1, X_2, \cdots, X_n)$ da amostra aleatória $(X_1, X_2, \cdots, X_n)$ sendo, como tal, também uma variável aleatória. É fundamental, em toda a inferência na estatística clássica, determinar estas distribuições amostrais. Para tal, a hipótese de que as variáveis aleatórias $(X_1, X_2, \cdots, X_n)$ são independentes e identicamente distribuídas, consequência dos dois axiomas de Richard von Mises, é primordial.

Contudo, nem sempre é analiticamente possível deduzir as distribuições das estatísticas (dependendo da distribuição inerente a cada observação e da função $h$ utilizada). Deste modo, no determinar da distribuição de cada estatística, tornou-se usual o recurso à simulação de amostras (nos caso em que a distribuição de cada observação seja conhecida). Por outro lado, considerando que se pretende analisar todas as possíveis amostras de dimensão $n$, que é computacionalmente impossível para amostras e populações de dimensão razoável, a metodologia de simulação *Bootstrap* desempenha um papel central na investigação atual desta visão.

## A teoria da decisão de Wald

Abraham Wald, em 1939, considera que as duas principais aplicações da Estatística (estimação e testes de hipóteses) podem ser considerados casos particulares de problemas de decisão. Deste modo, Wald reconstrui a Teoria da Estatística como uma Teoria da Decisão (nome que só viria a ser utilizado a partir de 1950, com os trabalhos de Erich Leo Lehmann (1917−2009)).



Wald

Desde a origem do Cálculo de Probabilidades, logo nos trabalhos de Pascal, é utilizado o conceito de valor esperado do ganho para tomar decisões (escolher a melhor estratégia de jogo). Em 1738 Daniel Bernoulli (1700−1782), na tentativa de clarificar o célebre paradoxo de São Petersburgo, criou a função de utilidade e usou o valor esperado da utilidade (onde a utilidade de ganhar o montante M é dado por log(M)), considerando que os valores monetários tem utilidade marginal decrescente. Deste modo, baseado na Teoria dos Jogos que John von Neumann (1903−1957) e Oskar Morgenstern (1902−1977) criaram por volta de 1944, Wald concebeu a teoria da decisão de forma a obter o melhor processo de decisão quando há incerteza sobre o que pode ocorrer. Neste sentido, criou conceitos fundamentais na Estatística, tais como a função perda, a função risco, regras de decisão admissíveis, regras de decisão minimax, entre outras. Outro contributo importante de Wald foi a introdução da análise estatística sequencial.