

Nota sobre a Utilização Incorreta de Conceitos Estatísticos

A Sociedade Portuguesa de Estatística (SPE) congratula-se com o empenho da comunidade científica na análise dos dados relativos à pandemia do CoVid-19, que assim contribui para um maior conhecimento da forma de propagação da doença, propiciando condições para uma actuação informada das entidades responsáveis. Importa, no entanto, assegurar que as análises efectuadas têm em conta as características dos dados em análise, seleccionando métodos e procedimentos adequados. Infelizmente, nem sempre é este o caso nas análises que são publicadas. Chamou a atenção da SPE o artigo publicado no *Observador* na 5ª feira dia 9 de Abril, [“O pico já passou? Portugal estabilizou a curva e até pode já estar em planalto, mas a confirmação só chegará nos próximos dias”](#) e [relatório anexo](#), os quais reportam análises claramente não apropriadas, obtendo e discutindo valores que não têm qualquer significado.

Em particular, importa atentar nos seguintes aspectos:

1. A curva “epidémica” representada na figura da página 7 do referido relatório apresenta o número de novos casos em cada dia, trata-se assim de valores observados de uma variável aleatória indexada no tempo, i.e., a curva representa de facto uma série temporal.
2. Esta curva não constitui de forma alguma um gráfico de frequências de uma amostra aleatória, pois são valores correlacionados e não observações de variáveis identicamente distribuídas - basta ver que a média se altera no tempo.
3. Estatísticas tais como média, desvio padrão, coeficientes de assimetria e achatamento são medidas que permitem descrever as características de uma amostra aleatória. Não são por isso de todo adequadas para caracterizar a forma da “curva” representada na referida figura, não tendo assim os valores reportados qualquer significado.
4. O coeficiente de achatamento calculado numa amostra aleatória serve para comparar a distribuição dessa amostra com a distribuição Normal (curva de Gauss). Não pode ser usado para descrever as características da tendência de uma série temporal, como acontece quando se tem o número de novos casos em cada dia.

É certo que a palavra "achatamento" é usada nos dois contextos, o da curva epidémica (e seu achatamento) e os dados de uma amostra (e o correspondente coeficiente de achatamento) e isso pode induzir em erro quem tem poucos conhecimentos de Estatística. Mas uma curva epidémica e uma amostra são objetos bem diferentes.

Sublinhe-se que não se trata de uma questão de opinião sobre como avaliar a forma da curva. O que aqui está em causa é a utilização incorreta de conceitos de Estatística.

Estes e outros aspetos referentes aos cuidados a ter com a análise de dados foram também muito bem salientados no artigo publicado no *Observador* [“Não se deixe iludir pelas aparências. Não é tão simples quanto parece.”](#), no dia 8 de Abril, por especialistas de Estatística Prof. Lisete Sousa, Prof. Tiago Marques e Prof. Antónia Amaral Turkman. Uma leitura atenta deste artigo teria evitado a publicação repetida de notícias baseadas em estudos que usam incorretamente conceitos estatísticos, pouco contribuindo para uma opinião pública informada.

P’la Sociedade Portuguesa de Estatística



Maria Eduarda Silva
(Presidente)