

Descrição do serviço

A Estatística é, hoje em dia, uma ciência reconhecida como essencial e incontornável em Medicina.

A análise de um conjunto de dados deve começar por uma análise descritiva e exploratória, cujo objetivo principal é resumir e organizar a informação. Os diversos procedimentos descritivos, numéricos e gráficos, possibilitam a caracterização da amostra e visam expor regularidades/padrões e evidenciar contrastes. Embora a maioria dos estudos envolva múltiplas variáveis é comum iniciar-se a análise dos dados com uma descrição univariada (variável a variável). Uma forma prática de iniciar este processo passa pela exploração gráfica dos dados. A análise gráfica não só facilita a identificação de eventuais erros de introdução de dados, como também, revela a forma da população subjacente e evidencia eventuais observações discrepantes.

Uma das formas mais comuns no resumo das características amostrais consiste no cálculo de medidas de localização, de escala e de forma. A decisão de quais as medidas adequadas à descrição dos dados depende do tipo de dados e de algumas características dos dados, designadamente, da forma da distribuição e da existência de *outliers*. Em variáveis quantitativas, bem comportadas, os valores centrais são bem descritos pela média amostral, podendo complementar-se essa informação com o cálculo da mediana e da moda. Ainda neste tipo de variáveis a variabilidade das observações é, normalmente, resumida pelo desvio-padrão. Em distribuições muito assimétricas e/ou na presença de outliers, as medidas mais comuns na descrição dos dados, como sejam a média e o desvio-padrão, podem ser muito influenciadas por estes valores, induzindo a leituras distorcidas, sendo aconselhável, nestes casos, recorrer a medidas mais resistentes, como seja a mediana e a amplitude interquartílica.

Embora a análise estatística se inicie, de facto, com a descrição univariada dos dados, muito poucas serão as circunstâncias em que tal será suficiente. O passo seguinte será a descrição de associações entre variáveis. Na análise da associação entre variáveis quantitativas é útil começar com a representação gráfica da relação entre as duas variáveis, o que pode ser feito através de um gráfico de pontos (*Scatterplot*). Em Medicina, são muitas as variáveis ordinais referidas na bibliografia deste domínio científico. A título de exemplo, mencionam-se a Escala de Apgar, as Escalas de Avaliação Neurológica e as Escalas de Classificação Nutricional. Quando se pretende quantificar associações entre escalas ordinais ou entre estas e escalas quantitativas o coeficiente adequado será o Coeficiente de Correlação de Spearman. Outras duas medidas de associação muito usadas em estudos de investigação médica são o Risco Relativo (*Relative Risk*) e a Razão de Chances (*Odds Ratio*), respetivamente, em estudos prospetivos e retrospectivos.

Frequentemente a análise inicial de dados é complementada com a apresentação de estimativas intervalares dos parâmetros, ou seja, de intervalos de confiança. A inferência de resultados de uma amostra para a população é, em regra, o objetivo último da análise estatística de dados. Neste âmbito, sob determinadas condições e risco, é possível testar hipóteses e prever acontecimentos futuros. Frequentemente muitas das questões motivadoras de estudos de investigação envolvem e implicam a modelação de variáveis resposta bem como análises espaço-temporais. Metodologias como a análise de sobrevivência (e.g., Kaplan Meier curves, regressão de Cox, modelação paramétrica), a análise de dados categorizados (modelos log-lineares) e os modelos lineares generalizados (e.g., ANOVA, ANCOVA, regressão linear múltipla, regressão logística) são exemplos de métodos analíticos frequentemente usados neste contexto.

