INSTITUTO SUPERIOR DE AGRONOMIA
Universidade de Lisboa

Departamento de Ciências e Engenharia de Biossistemas
Secção de Matemática

Matemática II

Conceitos básicos de Inferência Estatística

2014/2015

(F. Valente e M. Mesquita)

Inferência estatística

- Uma variável aleatória é descrita pela sua distribuição de probabilidade.
- A distribuição de probabilidade (modelo) é caracterizada por um ou mais parâmetros populacionais.

  Ex: $X \cap P(\lambda),\quad X \cap N(\mu, \sigma)$
- Em geral, a população é desconhecida logo a distribuição (o modelo) e/ou os parâmetros são desconhecidos.
- Um dos objectivos pode ser, assim, avaliar o valor real do(s) parâmetro(s) populacional(ais).

Inferência estatística (cont.)

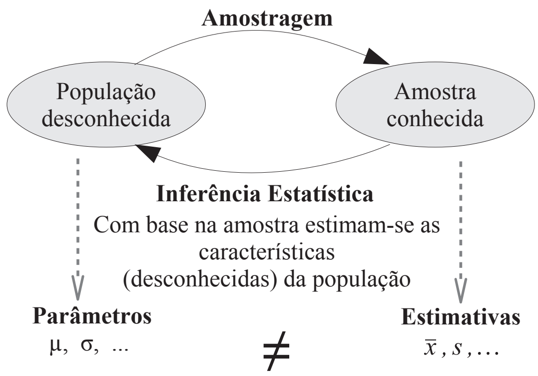
- As amostras observadas fornecem **estimativas** para o valor do(s) parâmetro(s) desconhecido(s).

  Exemplo: Amostra observada (1,5,6,2,3,3,2,4,1,2,4,4)

  $\bar{x} = 3.08(3)$ é uma estimativa para μ

  $s = 2.44697$ é uma estimativa para σ
- Diferentes amostras dão origem a diferentes estimativas para o valor do mesmo parâmetro.
- **Inferência estatística** – conjunto de métodos que tem por objectivo tirar conclusões (inferir) sobre "aspectos" desconhecidos da população (por exemplo, parâmetros) com base na análise de uma amostra.

Inferência estatística - motivação



Amostra aleatória e amostra observada

- Amostra deve dar uma "boa" representação da população.
- Considere-se:
  - X uma v.a. de interesse;
  - $X_1, \ldots, X_n$ v.a.s independentes e identicamente distribuídas (i.i.d.) a X.

  O vector $(X_1, \ldots, X_n)$ diz-se uma **amostra aleatória** (a.a.) de dimensão n respeitante à v.a. X.
- À observação particular da a.a. $(X_1, \ldots, X_n)$ dá-se o nome de **amostra observada** e representa-se por $(x_1, \ldots, x_n)$.

Estimação pontual

- **Estimador** é uma função da a.a., que não envolve parâmetros desconhecidos, cujo contradomínio é o espaço dos parâmetros.
- A cada amostra observada o **estimador** faz corresponder uma **estimativa**, isto é, uma avaliação do(s) parâmetro(s) desconhecido(s):

Parâmetro populacional (nº real desconhecido)	**Estimador** (variável aleatória)	**Estimativa** (nº real conhecido)
μ	$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
σ^2	$S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}$	$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$

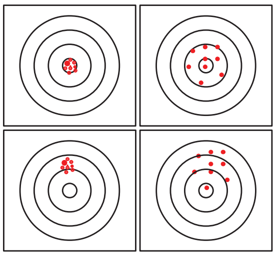
Estimação pontual

- Propriedades dos estimadores que poderão conduzir a "boas" estimativas.
- Exemplo:

preciso impreciso

centrado

enviezado



(Whitlock e Schluter, 2009)

Um "bom" estimador deve fornecer "em média" estimativas concentradas em torno do parâmetro a estimar.

Estimação pontual

- Seja $T = T(X_1, ..., X_n)$ um estimador de um parâmetro θ :
  - T é um **estimador centrado** de θ se $E[T] = \theta$
  - O **erro quadrático médio**, *EQM*, de T é dado por

$$EQM(T) = E\left[(T-\theta)^2\right]$$

- O *EQM* quantifica a dispersão esperada do estimador T em torno do parâmetro desconhecido θ. Um estimador será tanto "melhor" quanto menor for o seu *EQM*.

Bibliografia

- Neves, M. (2009) Introdução à Estatística e Probabilidade. Apontamentos de apoio às aulas:
  - Introdução à Inferência Estatística: *www.isa.utl.pt/dm/estat/estat/seb3.pdf*
- Pestana, D.D. e Velosa, S.F. (2002) Introdução à Probabilidade e à Estatística, vol.I, Fundação Caloustre Gulbenkian, Lisboa.
- Whitlock, M. e Schluter, D. (2009) The analysis of biological data. Roberts and Company Publishers, Colorado.