Estimação

Parâmetro: funções de valores populacionais.

Estatística: funções de valores amostrais.

Exemplo 11.1. Uma amostra de n = 500 pessoas de uma cidade é escolhida, e a cada pessoa da amostra é feita uma pergunta a respeito de um problema municipal, para o qual foi apresentada uma solução pela prefeitura. A resposta à pergunta poderá ser SIM (favorável à solução) ou NÃO (contrária à solução). Deseja-se estimar a proporção de pessoas na cidade favoráveis à solução apresentada.

Se 300 pessoas responderam SIM à pergunta, então uma estimativa natural para essa proporção seria 300/500 ou 60%. Nossa resposta é baseada na suposição de que a amostra é representativa da população. Sabemos, também, que outra amostra poderia levar a outra estimativa. Conhecer as propriedades desses *estimadores* é um dos propósitos mais importantes da Inferência Estatística.

Definamos as v.a. $X_1, \ldots, X_n$, tais que:

$$X_i = \begin{cases} 1, \text{ se a i-ésima pessoa na amostra responder SIM}, \\ 0, \text{ se a i-ésima pessoa na amostra responder NÃO}, \end{cases}$$

e seja $p = P$ (sucesso), onde aqui *sucesso* significa resposta SIM à questão formulada.

$$\hat{p} = \frac{Y_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{\text{número de SIM}}{\text{número de indivíduos}} \,. \tag{11.1}$$

$$E(\hat{p}) = p, \tag{11.2}$$

$$\text{Var}(\hat{p}) = p(1-p)/n. \tag{11.3}$$

(11.2): estimador **não-viesado** (ou não-viciado) de p.

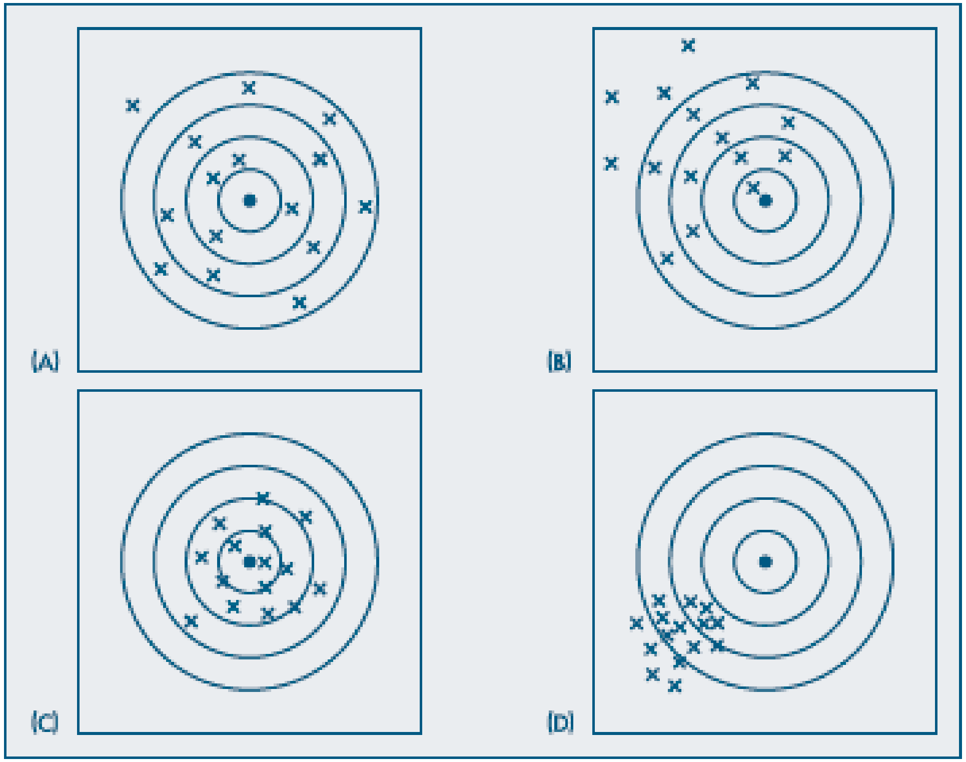
(11.3): estimador **consistente** de p.

Estimação

Em algumas situações, podemos ter mais de um estimador para um mesmo parâmetro, e desejamos saber qual deles é “melhor”. O julgamento pode ser feito analisando as propriedades desses estimadores.

Exemplo: Desejamos comprar um rifle e, após algumas seleções, restaram quatro alternativas, que chamaremos de rifles A, B, C e D. Foi feito um teste com cada rifle, que consistiu em fixá-lo num cavalete, mirar o centro de um alvo e disparar 15 tiros. Qual é a melhor arma?

Figura 11.1: Resultados de 15 tiros dados por 4 rifles.



Desse modo, podemos descrever cada arma da seguinte maneira:

Arma A: não-viesada, pouco acurada e baixa precisão.
Arma B: viesada, pouco acurada e baixa precisão.
Arma C: não-viesada, muito acurada e boa precisão.
Arma D: viesada, pouco acurada e alta precisão.

Acurácia: mede a proximidade de cada observação do **valor alvo** que se procura atingir.

Precisão: mede a proximidade de cada observação da **média** de todas as observações

Critérios:

1. Em média acertar o alvo: A e C;
2. Não ser muito dispersivo: C e D.

Estimação

Propriedades de estimadores

Definição: um **estimador** T do parâmetro θ é qualquer função das observações da amostra, ou seja,

$$T=f(X_1, ..., X_n).$$

Definição: o estimador T é **não viesado** para θ se

$$E(T) = \theta$$

Definição: **Estimativa** é o valor assumido pelo estimador em uma particular amostra.

Uma sequência $\{T_n\}$ de estimadores de θ é consistente se

$$\lim_{n \to \infty} E(T_n) = \theta \text{ e } \lim_{n \to \infty} Var(T_n) = 0$$

Definião: Se T e T´são dois estimadores **não-viesados** de um mesmo parâmetro θ, e ainda Var(T) < Var(T´), então T diz-se mais eficiente do que T´.

Exemplo 11.4. Considere uma população com N elementos e a variância populacional

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (X_i - \mu)^2, \tag{11.5}$$

onde $\mu = \frac{1}{N} \sum_{i=1}^{N} X_i$ é a média populacional. Um possível estimador para σ^2, baseado numa AAS de tamanho n extraída dessa população, é

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2. \tag{11.6}$$

Mostremos que esse estimador é viesado. Pela fórmula (3.11), temos que

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} X_i^2 - \bar{X}^2,$$

logo

$$E(\hat{\sigma}^2) = \frac{1}{n} \sum_{i=1}^{n} E(X_i^2) - E(\bar{X}^2).$$

Mas, pela definição de AAS e definição de variância de uma v.a., $E(X_i^2) = \mathrm{Var}(X_i) + [E(X_i)]^2 = \sigma^2 + \mu^2$. Também, usando o Teorema 10.1, temos que $E(\bar{X}^2) = \mathrm{Var}(\bar{X}) + [E(\bar{X})]^2 = \frac{\sigma^2}{n} + \mu^2$.

Seque-se que

$$E(\hat{\sigma}^2) = \frac{1}{n}\sum_{i=1}^{n}(\sigma^2 + \mu^2) - \left(\frac{\sigma^2}{n} + \mu^2\right),$$

ou seja,

$$E(\hat{\sigma}^2) = \frac{1}{n}(n(\sigma^2 + \mu^2)) - \frac{\sigma^2}{n} - \mu^2 = \sigma^2 - \frac{\sigma^2}{n} = \sigma^2\left(1 - \frac{1}{n}\right).$$

Finalmente,

$$E(\hat{\sigma}^2) = \left(\frac{n-1}{n}\right)\sigma^2. \tag{11.7}$$

De (11.7) vemos que $\hat{\sigma}^2$ é viesado para σ^2 e o viés é dado por

$$V = V(\hat{\sigma}^2) = E(\hat{\sigma}^2) - \sigma^2 = -\frac{\sigma^2}{n}. \tag{11.8}$$

Como esse viés é negativo, o estimador $\hat{\sigma}^2$ em geral subestima o verdadeiro parâmetro σ^2. Por outro lado, por (11.8), o viés diminui com n, ou seja, formalmente, para $n \to \infty$, o viés de $\hat{\sigma}^2$ tende a zero. Note também que o viés de $\hat{\sigma}^2$ é uma função de σ^2. Uma estimativa do viés seria dada por

$$\hat{V} = -\frac{\hat{\sigma}^2}{n},$$

ou seja, substituímos o valor desconhecido de σ^2 por uma estimativa, como por exemplo $\hat{\sigma}^2$.

É fácil ver que para obter um estimador não-viesado de σ^2 basta considerar $(n/(n-1))\hat{\sigma}^2$, pois de (11.7) segue-se que

$$E\left(\frac{n}{n-1}\hat{\sigma}^2\right) = \sigma^2.$$

Logo, se definirmos

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2, \tag{11.9}$$

então $E(S^2) = \sigma^2$ e S^2 é um estimador não-viesado para σ^2. Essa é a razão para se usar $n - 1$, em vez de n, como denominador da variância da amostra.

Exemplo 11.5. Vimos que S^2, dado por (11.9), é não-viesado para σ^2. É possível demonstrar, no caso que $X_1,\ldots, X_n$ são observações de uma distribuição $N(\mu, \sigma^2)$, que

$$\mathrm{Var}(S^2) = \frac{2\sigma^4}{n-1}. \tag{11.13}$$

Como $E(S^2) = \sigma^2$, e $\lim_{n\to\infty} \mathrm{Var}(S^2) = 0$, segue-se que S^2 é um estimador consistente para σ^2. Dado o que foi dito acima, talvez fosse melhor escrever S_n^2.

Exemplo 11.6. Vimos que $E(\hat{\sigma}^2) = \sigma^2(1 - 1/n)$, de modo que $\lim_{n\to\infty} E(\hat{\sigma}^2) = \sigma^2$. Também, de (11.6) e (11.13) e supondo que as observações são de uma distribuição normal $N(\mu, \sigma^2)$, temos que

$$\mathrm{Var}(\hat{\sigma}^2) = \left(\frac{n-1}{n}\right)^2 \mathrm{Var}(S^2) = \frac{n-1}{n^2}(2\sigma^4), \tag{11.14}$$

o que mostra que $\mathrm{Var}(\hat{\sigma}^2) \to 0$, quando n $\to \infty$, logo $\hat{\sigma}^2 = \hat{\sigma}_n^2$ também é consistente para σ^2.

De (11.14) obtemos, também, que

$$\mathrm{Var}(\hat{\sigma}^2) < \frac{2\sigma^4}{n-1} = \mathrm{Var}(S^2). \tag{11.15}$$

Estimação

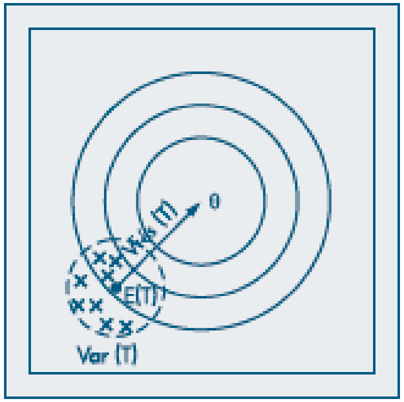
Propriedades de estimadores

$e = T - \theta$: o erro amostral que cometemos ao estimar o parâmetro θ da distribuição da v.a. X pelo estimador $T=f(X_1, ..., X_n)$, baseado na amostra $(X_1, ..., X_n)$.

Definição: Chama-se erro quadrático médio (EQM) do estimador T ao valor

$$EQM(T; \Theta) = E(e^2) = E(T-\theta)^2 = Var(T) + (\text{viés de } T)^2$$

Figura 11.2: Representação gráfica para o *EQM*.



Um estimador preciso tem variância pequena, mas pode ter EQM grande.