

REVISÃO DE ESTATÍSTICA GERAL

Planejamento e Análise Estatística de Experimentos Agronômicos

Décio Barbin (2003)

VARIAÇÃO DO ACASO: É toda variação devida a fatores não controláveis.

Exemplo: área experimental plana, de solo bem homogêneo e nela plantamos sementes selecionadas de milho. Fazemos a semeadura de maneira que as sementes sejam colocadas no solo, na mesma posição e mesma profundidade. Essas sementes irão germinar. As plantas crescerão e, no momento que emitirem o pendão vamos medir suas alturas do solo até a inserção da folha “bandeira”.

Fato: Observam-se plantas com diferentes alturas, mesmo com todo o cuidado em homogeneizar as condições experimentais e ter material biológico homogêneo.

Motivo: Fatores impossíveis de serem controlados (VARIAÇÃO DO ACASO)

COMO MEDIR A VARIAÇÃO DO ACASO?

- Média e Desvios de cada observação em relação a média
- Ex_1 : Altura de pés de milho (cm): Y

y_i : 203 208 198 200 202 192 197

$$\bar{y} = 200 \text{ cm}$$

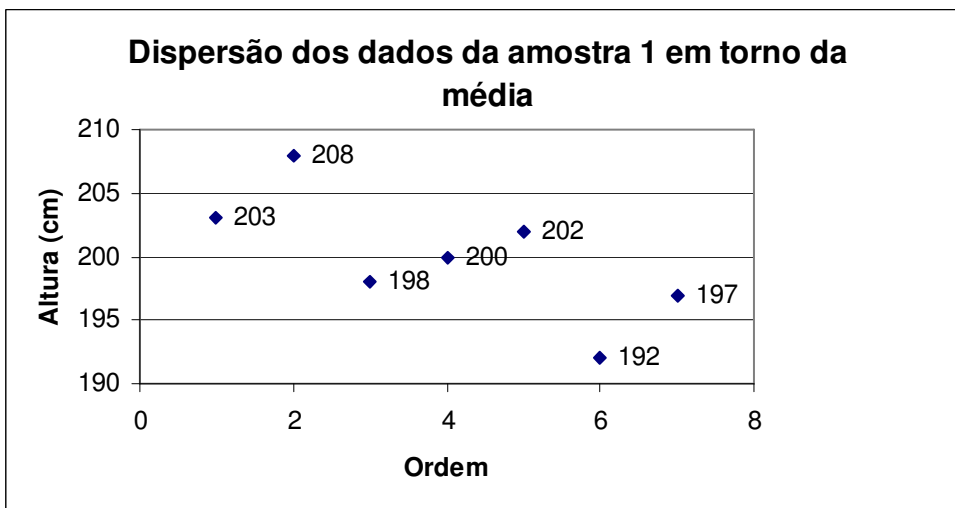
Desvios $\hat{e}_i$: 3 8 -2 0 2 -8 -3

Propriedade: $\sum_i \hat{e}_i = 0$

Modelo matemático:

$$y_i = \mu + e_i,$$

em que: μ é a média populacional, e e_i são os verdadeiros desvios populacionais.



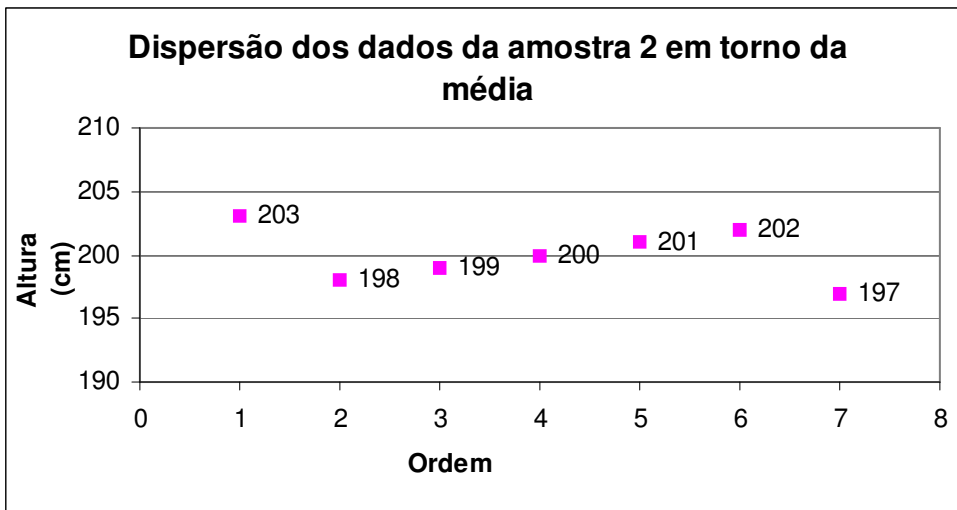
- **Ex₂: Altura de pés de milho (cm)**

y_i: 203 198 199 200 201 202 197

$$\bar{y} = 200 \text{ cm}$$

Desvios $\hat{e}_i$: 3 -2 -1 0 1 2 -3

Propriedade: $\sum_i \hat{e}_i = 0$



Menor dispersão dos dados (Menor VARIAÇÃO DO ACASO)

QUANTIFICANDO A VARIABILIDADE DOS DADOS EXPERIMENTAIS

- **VARIANCIA E DESVIO PADRÃO**

$$\sigma^2 = \frac{\sum_{i=1}^n e_i^2}{n} \quad (\text{variância populacional: Parâmetro})$$

$$S^2 = \frac{\sum_{i=1}^n \hat{e}_i^2}{n-1} \quad (\text{variância amostral: Estimador})$$

n: é o número de observações

n-1: graus de liberdade

Nos exemplos:

$$s_1^2 = 25,67 \text{ cm}^2 \quad \text{e} \quad s_2^2 = 4,67 \text{ cm}^2 \quad (\text{MENOR VARIAÇÃO DO ACASO})$$

$$\text{Desvios Padrão: } s_1 = \sqrt{25,67} = 5,07 \text{ cm} \quad \text{e} \quad s_2 = \sqrt{4,67} = 2,16 \text{ cm}$$

$$S^2 = \frac{\sum_{i=1}^n y_i^2 - \frac{\left(\sum_{i=1}^n y_i\right)^2}{n}}{n-1} \quad (\text{VARIÂNCIA AMOSTRAL BASEADA NOS VALORES OBSERVADOS})$$

$$S_1^2 = \frac{280154 - \frac{(1400)^2}{7}}{7-1} = 25,67 \text{ cm}^2 \quad \text{e} \quad S_2^2 = \frac{280028 - \frac{(1400)^2}{7}}{7-1} = 4,67 \text{ cm}^2$$

VARIÂNCIA DA MÉDIA E ERRO PADRÃO DA MÉDIA

$$\text{Var}(\bar{Y}) = \frac{S^2}{n} \quad (\text{VARIÂNCIA DA MÉDIA})$$

$$S(\bar{Y}) = \frac{S}{\sqrt{n}} \quad (\text{ERRO PADRÃO DA MÉDIA})$$

$$\text{var}_1(\bar{Y}) = 3,67 \text{ cm}^2, \quad \text{com} \quad s_1(\bar{Y}) = 1,91 \text{ cm}$$

$$\text{var}_2(\bar{Y}) = 0,67 \text{ cm}^2, \quad \text{com} \quad s_2(\bar{Y}) = 0,82 \text{ cm}$$

$$\text{Apresentação usual em artigos científicos: } \bar{y} \pm s(\bar{Y}) \begin{cases} 200 \pm 1,91 \text{ cm} \\ 200 \pm 0,82 \text{ cm} \end{cases}$$

COEFICIENTE DE VARIAÇÃO (VARIABILIDADE RELATIVA)

$$C.V = \frac{S}{\bar{Y}} \times 100 \quad (\text{É um número puro ou adimensional ou sem unidade})$$

$$CV_1 = \frac{5,07}{200} \times 100 = 2,53\%$$

$$CV_2 = \frac{2,16}{200} \times 100 = 1,08\%$$

PROBLEMA: Para dados relativos (positivos e negativos) o CV não tem sentido, pois a média poderá se aproximar de zero e com isso o CV tenderá a infinito.

SOLUÇÃO: Somar uma constante positiva a todos os dados tornando-os positivos.

EXERCÍCIOS:

- (1) Verificar o que ocorre com a média, variância e coeficiente de variação quando se soma uma constante k aos dados observados.
- (2) Idem, quando se se multiplicam os dados por uma constante k.

INTERVALO DE CONFIANÇA PARA A MÉDIA

Informa sobre a precisão da estimativa da média, em termos probabilísticos.

- **Obtido por:**

$$\bar{Y} \pm t \times S(\bar{Y})$$

t: é o valor da variável aleatória t-Student $t_{(\alpha, n-1)}$

Amostra₁: $200 \pm 2,45 \times 1,91$ cm
 $200 \pm 4,68$ cm
 $195,32 \text{ cm} < \mu < 204,68 \text{ cm}$

Amostra₂: $200 \pm 2,45 \times 0,82$ cm
 $200 \pm 2,01$ cm
 $197,99 \text{ cm} < \mu < 202,01 \text{ cm}$

- Qual a probabilidade dos intervalos anteriores conterem a verdadeira média?
- **Interpretação:** Antes de coletar a amostra, a confiança que se deposita no intervalo que será construído é de 95%.

Coletando-se 100 amostras nas mesmas condições, e para cada uma delas construir um intervalo de confiança, espera-se que 95% deles contenham o verdadeiro parâmetro.

