Medidas de Posição e Dispersão

Gilberto Pereira Sassi

Universidade Federal Fluminense
Instituto de Matemática e Estatística
Departamento de Estatística

12 de maio de 2016

Alguns Conceitos Básicos

População todos elementos alvo do estudo;

Amostra Parte da população;

Parâmetro característica da população;

Estatística característica de amostra;

Estimativa valor da estatística para uma amostra específica.

Exemplo

No estudo de incidência de depressão em nível grave nos alunos no ciclo inicial no curso de medicina na UFF temos:

População todos os alunos regularmente matriculados no ciclo inicial no curso de medicina da UFF;

Amostra alunos selecionados para responderem o Inventário para Depressão de Beck;

Parâmetro porcentagem na populacional de alunos no ciclo inicial com depressão em nível grave;

Estatística Porcentagem de alunos com depressão em parte da população;

Estimativa Após coletada uma amostra, valor numérico da estatística.

Medidas de Posição e Dispersão

Dois tipos de medidas:

Medidas de posição	valor típico / valor representativo;
Medidas de Dispersão	Avalia como os valores da variável se distribuem.

Média

Seja $x_1, \ldots x_n$ valores de uma variável, então a média é dada por

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Exemplo

Na Tabela1, mostramos a nota final de três turmas de estatística.

Tabela 1: Notas dos alunos das três de estatística.

Turma A	4	5	5	6	6	7	7	8
Turma B	1	2	4	6	6	9	10	10
Turma C	0	6	7	7	7	7	8	

Turma A $\bar{x} = \dfrac{4+5+5+6+6+7+8}{8} = 6$

Turma B $\bar{x} = \dfrac{1+2+4+6+6+9+10+10}{8} = 6$

Turma C $\bar{x} = \dfrac{0+6+7+7+7+7+8}{8} = 6$

Dispersão unidimensional

Apesar de terem a mesma média, os dados se distribuem de forma distinta em torno da média.

Figura 1: Notas e média para a Turma A.

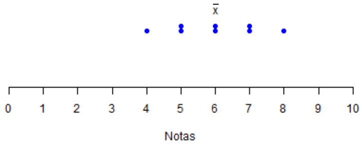


Figura 2: Notas e média para a Turma B.

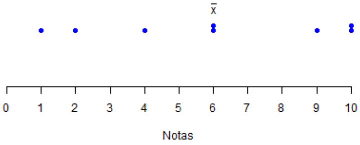
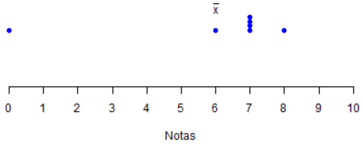


Figura 3: Notas e média para a Turma C.



Observações sobre Média.

1) A média indica o centro de um conjunto de valores;

2) A média resume um conjunto de valores em um único valor;

3) Notamos que as três turmas (A, B e C) têm a mesma média, mas os dados são distribuídos de forma diferente. A turma A tem notas mais homogêneas que as turmas A e C. Além disso, o aluno com nota 0 "puxa" a média para 6. Para avaliar essa distribuição podemos usar a variância e o desvio padrão.

Variância

Avalia como os dados estão distribuídos em torno da média em conjunto de valores.

Chamamos as distâncias $x_1 - \bar{x}$, $x_2 - \bar{x}$, $\cdots$, $x_n - \bar{x}$ de desvios. Poderíamos tomar a média dos desvios, mas considere o seguinte conjunto de dados $\{1, 2, 3\}$ com média $\bar{x} = 2$. Então os desvios são $-1, 0, 1$ cuja soma é 0 e teríamos a impressão que todos os valores são iguais a média. Como solução, podemos tirar a média dos desvios ao quadrado:

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n - 1}.$$

s^2 é chamado de variância.

O Desvio Padrão é a raiz quadrada da variância. Usamos o desvio padrão para manter a mesma unidade dos dados originais.

Variância – passos para calcular

i. Calculamos os desvios

$$(x_1 - \bar{x}), (x_2 - \bar{x}), \cdots, (x_n - \bar{x})$$

ii. Computamos os desvios ao quadrado

$$(x_1 - \bar{x})^2, (x_2 - \bar{x})^2, \cdots, (x_n - \bar{x})^2$$

iii. Calculamos a média dos desvios ao quadrado

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n - 1}$$

Exemplo

Considere as notas das três turmas de estatística apresentadas na Tabela 2.

Tabela 2: Notas e média para a as três turmas de estatística.

Turma	Notas								Média
A	4	5	5	6	6	7	7	8	6
B	1	2	4	6	6	9	10	10	6
C	0	6	7	7	7	7	8		6

Turma A $s^2 = \frac{(4-6)^2+(5-6)^2+(5-6)^2+(6-6)^2+(6-6)^2+(7-6)^2+(7-6)^2+(8-6)^2}{8-1} = 1,71$

Turma B $s^2 = \frac{(1-6)^2+(2-6)^2+(4-6)^2+(6-6)^2+(6-6)^2+(9-6)^2+(10-6)^2+(10-6)^2}{8-1} = 12,29$

Turma C $s^2 = \frac{(0-6)^2+(6-6)^2+(7-6)^2+(7-6)^2+(7-6)^2+(7-6)^2+(8-6)^2}{7-1} = 7,33$

Podemos notas que B tem a maior variância e tem os dados mais dispersos.

Forma Alternativa para Variância

Podemos calcular a variância usando:

$$s^2 = \frac{\sum x_i^2 - n \cdot \bar{x}^2}{n - 1}$$

em que

$$\sum x_i^2 = x_1^2 + x_2^2 + \cdots + x_3^2$$
$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$
$$n = \text{número de elementos na amostra}$$

Exemplo

Tabela 3: Notas, média e tamanha amostral n para as três turmas.

Turma	Notas								$\bar{x}$	n
A	4	5	5	6	6	7	7	8	6	8
B	1	2	4	6	6	9	10	10	6	8
C	0	6	7	7	7	7	8		6	7

Turma A $\sum x_i^2 = 300$, $\bar{x}^2 = 36$ e $s^2 = \frac{\sum x_i^2 - 8 \cdot 36}{7} = 1,71$;

Turma B $\sum x_i^2 = 374$, $\bar{x}^2 = 36$ e $s^2 = \frac{\sum x_i^2 - 8 \cdot 36}{7} = 12,29$;

Turma C $\sum x_i^2 = 296$, $\bar{x}^2 = 36$ e $s^2 = \frac{\sum x_i^2 - 7 \cdot 36}{6} = 6,28$;

Distribuição de Frequência – Média e Variância

Considere a variância quantitativa discreta com a seguinte distribuição de frequência:

Tabela 4: Distribuição de Frequência para uma variável quantitativa discreta.

Variável	Frequência
l	f_l
l+1	f_{l+1}
⋮	⋮
k-1	f_{k-1}
k	f_k
Total	n

A média e variância podem ser calculados por

$$\bar{x} = \frac{l \cdot f_l + (l+1) f_{l+1} + \cdots + (k-1) \cdot f_{k-1} + k \cdot f_k}{n}$$

$$s^2 = \frac{\sum x_i^2 - n \cdot \bar{x}^2}{n-1}$$

$$s^2 = \frac{l \cdot f_l + (l+1) \cdot f_{l+1} + \cdots + (k-1) \cdot f_{k-1} + k \cdot f_k}{n-1}$$

em que $\sum x_i^2 = l \cdot f_l + (l+1) \cdot f_{l+1} + \cdots + (k-1) \cdot f_{k-1} + k \cdot f_k$.

Distribuição de Frequência – Média e Variância

Considere a distribuição de frequência para a turma A.

Notas	Frequência
4	1
5	2
6	2
7	2
8	1
Total	8

$$\bar{x} = \frac{4 + 5 \cdot 2 + 6 \cdot 2 + 7 \cdot 2 + 8}{8} = 6$$

$$\sum x_i^2 = 4^2 + 2 \cdot 5^2 + 2 \cdot 6^2 + 2 \cdot 7^2 = 300$$

$$s^2 = \frac{300 - 8 \cdot 36}{7} = 1,71$$

Distribuição de Frequência – Média e Variância Variância Quantitativa Contínua

Considere a distribuição de frequência da taxa da mortalidade infantil dos municipios da microrregião do oeste catarinense em 1982:

Taxa de Mortalidade Infantil	Frequência	Ponto Médio
0 ⊢ 10	1	5
10 ⊢ 20	10	15
20 ⊢ 30	15	25
30 ⊢ 40	7	35
40 ⊢ 50	0	45
50 ⊢ 60	0	55
60 ⊢ 70	1	65
Total	34	–

Então,

$$\bar{x} = \frac{1 \cdot 5 + 10 \cdot 15 + 15 \cdot 25 + 7 \cdot 35 + 1 \cdot 65}{34} = 24,71$$

$$\sum x_i^2 = 1 \cdot 5^2 + 10 \cdot 15^2 + 15 \cdot 25^2 + 7 \cdot 35^2 + 1 \cdot 65^2 = 24450$$

$$s^2 = \frac{24450 - 34 \cdot (24,71)^2}{33} = 111,82$$