Biometria - EDAP

Descrição de amostras

(Leitura complementar ao capítulo 1)

Sumário:

Lembrete: Antes de ler esse texto já devem ter sido *lidos*:

Estatística, Hipótese, Método e Ciência,

População, Amostra, Variáveis, Dados e

Criação de Tabelas

Medidas de tendência central (ou Medidas de concentração ou Promédios)

Como o próprio nome já diz, *medidas de tendência central* são aquelas cujo valor tende a localizar-se no *centro* de uma série de dados.

Frequentemente, quando se analisa os valores de uma variável em uma amostra, constata-se que os dados não se distribuem uniformemente, havendo concentração em alguns pontos, notadamente próximos ao centro da distribuição.

Ou seja, é comum haver um grande número de elementos com valores próximos à média e poucos indivíduos apresentando valores extremos, isto é próximos aos valores mínimo e máximo..

Assim, de modo geral, se houver a necessidade ou interesse em apresentar informações de um conjunto de dados na forma resumida devemos apresentá-los em forma de medidas de tendência central.

Pode-se, portanto, estudar os valores numéricos que determinam a distribuição, procurando o ponto onde está a maior concentração de valores individuais, ou seja, as medidas de *tendência central*.

Delas, as mais importantes em estatística são: *Média*, *Mediana* e *Moda*.

Média

Há vários tipos de média (aritmética - simples ou ponderada, geométrica, harmônica, quadrática, cúbica, biquadrática).

A mais usada é a *média aritmética simples* ou, simplesmente, *média*, que é obtida dividindo-se a soma das observações pelo número delas. É um quociente geralmente representado pela letra M ou pelo símbolo $\bar{x}$ (lê-se "x barra").

Média de dados puros

Se tivermos uma série de N valores de uma variável x (x_1, x_2, x_3, x_4, ... x_n) a média será determinada pela expressão:

$\bar{x} = (x_1 + x_2 + x_3 + x_4 +...+ x_n) / N = \Sigma x / N$

Supondo os seguintes dados, já *ordenados*:

4	5	6	6	6	7	7	7	8	8
9	9	9	9	9	10	10	10	10	11
12	12	12	12	*12*	*13*	13	13	13	14
14	15	15	15	15	15	16	17	18	19
19	19	20	20	21	22	23	24	25	26

$\bar{x} = \Sigma / N = 664 / 50 = 13{,}28$

Para saber como efetuar os *cálculos* no BrOffice.org Calc, quando se tem todos os dados individuais, clicar aqui.

Lembrar que $\bar{x}$ refere-se à média da amostra (com n elementos) e deve ser distinguida da média da população, **μ** (com N elementos).

Média de dados agrupados

Como calcular a média *se os dados estiverem agrupados* em uma tabela de distribuição de frequências?

Evidentemente, a frequência tem que ser incluída no cálculo da média que passa a ser:

Média = $\bar{x}$ = Σfx / n

que pode ser calculada assim:

		dev
4	1	4
5	1	5
6	3	18
7	3	21
8	2	16
9	5	45
10	4	40
11	1	11
12	5. sub total: 25	60
13	4	52

14	2	28
15	5	75
16	1	16
17	1	17
18	1	18
19	3	57
20	2	40
21	1	21
22	1	22
23	1	23
24	1	24
25	1	25
26	1	26
Totais	50	664
	Σf = *N*	**Σfx**
Média:	***664/50***	**13.28**

Mediana

A mediana ocupa a posição central de uma série de dados ordenados.

- é o próprio valor central se a série for impar *ou*
- é a média aritmética dos dois valores centrais quando a sequência for par.

É absolutamente necessário que os valores estejam dispostos em *ordem* (crescente ou decrescente) de magnitude. Ou seja, a mediana divide os dados: 50% dos valores estão abaixo e 50% estão acima da mediana.

Portanto, mediana é o valor que divide uma série ordenada de modo que pelo menos a metade das observações sejam iguais ou maiores do que ela, e que haja pelo menos outra metade de observações maiores do que ela.

Chama-se de *EMd* o elemento mediano, aquele que indica a posição da mediana.

Mi
0---------------Q1---------------Q2---------------Q3--------------Q4

A mediana é representada pelo símbolo Mi e, evidentemente, coincide com o segundo quartil (Q2).

Na amostra acima, em que há 50 valores a mediana é a média dos 2 valores centrais (12 e 13) portanto é 12,5.

Mediana de dados não agrupados

Se houver um número ímpar de valores ordenados é só verificar o valor que ocupa a *posição central*. Se houver um número par de valores ordenados toma-se a soma dos 2 valores que estão nas posições centrais e divide-se por 2.

Mediana de dados agrupados

Os mesmos dados anteriores foram agrupados nas seguintes classes:

Classes	Frequência	Frequência acumulada
4 - 7	8	8
8 - 11	12	20
*12 - 15 **	16	36
16 - 19	6	42
20 - 23	5	47
24 - 27	3	50

A *posição* da mediana deverá corresponder à observação de ordem *N/2*.

Neste exemplo a posição será 50/2 = 25. Ou seja, deverá estar na classe de 12-15, ocupando a 25a. posição, de acordo com a seguinte fórmula:

Mi = li + [(N/2 - fa) / fc . i]

em que:

Mi = mediana

li = limite inferior da classe que deve conter a Mi

N = total

fa = frequência acumulada anterior à classe que deve conter a Mi

i = intervalo de classe

fc = frequência da classe que deve conter a Mi

Portanto:

Mi = li + [(N / 2 - fa) / fc . i]

= 12 + [(50 / 2 - 20) / 16 . 4] = 12 + 0,3125 . 4 = 12 + 1,25 = 13,25

Para facilitar os cálculos utilize uma planilha especial: *Mediana* em amostras com até 25 classes com dados classificados *Copie* a planilha *comprimida* em formato ods

Média e Mediana

Como se pode interpretar a Mediana e a Média?

É preciso lembrar , primeiramente, que a Mediana pode ser usada tanto para variáveis quantitativas intervalares como para variáveis qualitativas ordinais, enquanto a Média só pode ser utilizada para variáveis intervalares.

Em segundo lugar, no caso das variáveis *quantitativas*, embora a Média seja um valor mais fácil de entender, tem o defeito de nos induzir em erro se a amostra contiver valores muito extremos.

Por *exemplo*, supondo uma amostra *A* em que foi estudada a idade de 7 indivíduos (1 a 7), em meses.

Mas, se o último elemento for 350, ao invés de 70, seria obtida a amostra *B*:

posição	1	2	3	4	5	6	7	*Total*	*Média*	*Mediana*
A	10	20	30	40	50	60	*70*	280	40	40
B	10	20	30	40	50	60	*150*	360	51,43	40

Deve-se notar que

- O valor da média é igual ao da mediana na amostra *A*, (40)
- As medianas são iguais em ambas as amostras (40)
- Mas, em *B,* a média saltou para 80, ou seja, ficou superior à maioria dos valores individuais.

Observando os 7 valores individuais da amostra, verifica-se que o número 40 é melhor representante da distribuição global da idade na amostra que o número 80.

Assim, no caso de *variáveis quantitativas,* quando o valor da mediana é muito diferente da média, é aconselhável considerar sempre a mediana como valor de referência mais importante, pois a média aritmética pode ser distorcida por valores discrepantes, o que se comprova ao observar na tabela que a idade do indivíduo 7 está bem distante da maioria dos outros.

Portanto, pode-se concluir que:

- A média reflete o valor de todas as observações e se a distribuição dos dados for aproximadamente simétrica a média tem valor próximo ao da mediana.
- A mediana é *mais robusta* do que a média como medida de localização, pois é menos sensível a alguns dados chamados de "outliers", ou seja, aos valores muito maiores ou muito menores do que os restantes.
- Quando a distribuição está enviesada para a esquerda (há alguns valores pequenos como "outliers"), a média tende a ser inferior à mediana. O oposto acontece quando a distribuição está enviesada para a direita, nesse caso a média tende a ser maior que a mediana, pois há alguns valores grandes como "outliers".

Moda

É o valor amostral que tem a maior frequência, ou seja, é o encontrado em maior número de vezes, portanto, é a observação mais "provável" da distribuição dos dados. É representada pela notação *Mo* e também é chamada de "*Modo*".

Portanto, numa amostra a moda pode não existir. E uma distribuição em que não há elementos repetidos é dita *amodal*.

Também deve-se considerar que a moda pode não ser única. Se dois valores aparecem em igual quantidade de vezes a distribuição é dita bimodal. Para três valores, trimodal, e assim, sucessivamente.

Importante é notar que se existe apenas *uma moda* em uma amostra, há apenas um grupo de indivíduos com suas variações, ou seja, a amostra é *homogênea*.

Mas, se houver *duas ou mais modas*, há *grupos* diferentes dentro daquela amostra. Diz-se, então, que a amostra é *heterogênea*.

Métodos para calcular a moda

- Simples inspeção

Verifica-se qual é a classe que tem a maior frequência. Essa classe se constitui na moda.

Se os dados estão agrupados a moda é o ponto médio da classe que tem a maior frequência.
Note-se que uma amostra pode ter uma moda ou mais (diz-se que a curva é unimodal, bimodal, trimodal...)

- Processo empírico
  Em distribuições moderadamente assimétricas pode ser usada a fórmula de Pearson, sendo que:

  $$Mo = 3\ Mi - 2M$$

  Como a média = Σx / N = 664 / 50 = 13,28.
  Portanto, Mo = 3 . 13,25 - 2 . 13,28 = 39,75 - 26,56 = 13,19

- Processo gráfico
  Usa-se o histograma gerado pelos dados, passando-se dois segmentos de reta entre o vértice esquerdo da maior coluna e o vértice direito da coluna seguinte e entre o vértice direito da maior coluna e o vértice esquerdo da coluna anterior. No ponto onde as retas se cruzam traça-se uma perpendicular à abcissa e o valor encontrado no eixo dos X é a moda.

Medidas Separatrizes: Mediana, Quartis, Decis e Percentis

Lembrando que se um conjunto de dados é *ordenado* em ordem de grandeza, o valor médio que divide o conjunto em duas partes iguais é a *mediana*, pode-se pensar em valores que dividam o conjunto em: quatro, dez ou cem partes iguais.

O *Quartil* é a medida que divide o conjunto em quatro partes iguais. Assim, para dividir uma reta em quatro partes, é necessário marcar três pontos. Então, haverá sempre três Quartis em um conjunto, que serão denominados por Q1 (primeiro quartil), Q2 (segundo quartil) e Q3 (terceiro quartil).

O *Decil*, por sua vez, que divide o conjunto em dez partes iguais (D1, D2, ..., D9). Daí se conclui que haverá, em um conjunto, nove decis.

Já o *Centil* (ou *Percentil*), é a medida separatriz que divide o conjunto em cem partes iguais. Portanto em um conjunto, haverá noventa e nove Centis (P1, P2, ..., P99).

```
                                                            Mi
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
0                             Q1                            Q2                            Q3
0           D1          D2          D3          D4          D5          D6          D7          D8          D9
0     P5    P10   P15   P20   P25   P30   P35   P40   P45   P50   P55   P60   P65   P70   P75   P80   P85   P90   P95
```

Portanto, *Q2*, *D5* e *P50* correspondem à mediana, da mesma forma que P25 e P75 correspondem a Q1 e Q3, respectivamente.

Medidas de dispersão

Variação ou dispersão é o grau com que os dados numéricos tendem a se espalhar em torno de um valor médio. Ou seja, medidas de dispersão são indicadores do grau de variabilidade demonstrada pelos indivíduos em torno das medidas de tendência central.

Para estudar a variação há várias medidas já definidas. Dentre elas destacam-se a variância e o desvio padrão.

Variância

A *variância*, representada por s^2, e é definida como o "*desvio quadrático médio da média*".

Note-se que como a variância mede os *desvios em relação à média* (ou seja, a diferença entre cada dado e a média) e avalia o *grau de dispersão* de um conjunto de dados.

Considere 3 amostras, A, B e C, com médias iguais, em que o comprimento de um órgão (em mm) foi anotado.

											soma	média
A	*160*	162	165	168	172	*175*					1002	*167*
B	*160*	161	162	168	170	173	*175*				1169	*167*
C	*160*	162	163	164	165	167	170	171	173	*175*	1670	*167*

É importante notar que as amplitudes (175-160 = 15) e as médias (= 167) são iguais nas 3 amostras.

Para analisar a dispersão dos dados em torno da média, em cada amostra, obtém-se o desvio em relação à média ($x - \bar{x}$) e as suas somas.

É importante notar que não há média dos desvios ($x - \bar{x}$), porque a sua soma em cada amostra é sempre igual a zero.

Assim, também se obtém os quadrados dos desvios $(x - \bar{x})^2$ e as suas somas:

A	**desvio**	**desvio²**	***B***	**desvio**	**desvio²**	***C***	**desvio**	**desvio²**
	$(x - \bar{x})$	$(x - \bar{x})^2$		$(x - \bar{x})$	($(x - \bar{x})^2$		$(x - \bar{x})$	$(x - \bar{x})^2$
160	-7	49	*160*	-7	49	*160*	-7	49
162	-5	25	161	-6	36	162	-5	25
165	-2	4	162	-5	25	163	-4	16
168	1	1	168	1	1	164	-3	9
172	5	25	170	3	9	165	-2	4
175	8	64	173	6	36	167	0	0
			175	8	64	170	3	9
						171	4	16
						173	6	36
						175	8	64
1002	**0**	**168**	**1169**	**0**	**220**	**1670**	**0**	**228**

Ressalte-se que apesar da dispersão dos dados em torno da média ser a mesma nos 3 grupos, a *soma dos quadrados* dos desvios $(x - \bar{x})^2$ é maior no grupo ***C***, pois é o que possui maior número de dados.

Mas, para medir a dispersão dos dados em relação à média, deve-se usar a *variância*, (s^2), pois o valor obtido leva em consideração o *tamanho da amostra*.

A fórmula geral da variância é

$$(s^2) = \text{soma de quadrados dos desvios} / (n - 1).$$

Como a mostra *A* tem 6 elementos a variância é assim calculada: 168 / 5 = 33,6.

Do mesmo modo, para a amostra *B* = 220/ 6 = 36,67 e para a amostra *C* = 228/9 = 25,33.

Fórmulas

Considerando uma série de *N* valores de uma variável x ($x_1, x_2, x_3, x_4, \ldots x_n$), com média $\bar{x}$, a variância pode ser determinada por:

$$s^2 = \Sigma (x_i - \bar{x})^2 / (N - 1)$$

$$s^2 = \Sigma(x_i^2) - N \bar{x}^2 / (N - 1)$$

$$s^2 = \Sigma x^2 - [(\Sigma x)^2 / N] / (N - 1)$$

Assim, a variância é a medida que se obtém somando os quadrados dos desvios das observações da amostra, relativamente à sua média, e dividindo pelo número de observações da amostra menos um.

É importante notar que:

- a variância *nunca é negativa*, porque os quadrados são sempre positivos ou nulos. Assim, a unidade de variância é o quadrado da unidade de observação. (Exemplo: a variância de um conjunto de alturas medidas em centímetros será dada em centímetros quadrados).
- se todas as médias das amostras forem iguais, o valor da variância da média seria igual a zero.
- quanto maior for a variância menor é o grau de concentração dos indivíduos na amostra

Exemplo:

Apenas como exemplo, suponha que duas amostras apresentaram os seguintes valores de largura de um órgão, em cm:

A = 8, 10, 12, 14 e 16 e

B = 4, 8, 12, 16 e 20

Cálculo da variância de dados puros

Quando se tem todos os dados individuais, ainda sem nenhum tratamento, portanto sem agrupamento em *classes*, pode-se obter o quadrado dos valores individuais e das duas somatórias.

Exemplo:

Considere 2 amostras, A e B, com 5 dados cada:

Amostras	A			B	
	x	x^2		x	x^2
	8	64		4	16
	10	100		8	64
	12	144		12	144
	14	196		16	256
	16	256		20	400
Total	*60*	*760*		*60*	*880*

Pode-se calcular a variância do seguinte modo:

Variância A

$s^2_A = \Sigma x^2 - [(\Sigma x)^2 / N] / (N - 1) = [760 - (60^2 / 5)] / 4 =$

$= (760 - 720) / 4 = 10$

Variância B

$s^2B = \Sigma x^2 - [(\Sigma x)^2/ N] / (N - 1)$

$= [880 - (60^2 / 5)] / 4 = (880 - 720) / 4 = 40$

Notar que na amostra A os indivíduos estão mais *concentrados*, distribuindo-se entre o valor mínimo = 8 e o máximo = 16

E, na amostra B estão mais *dispersos* (distribuindo-se ente 4 e 20).

Assim, na amostra ***A*** a variância ($s^2_A = 10$) é menor que a da *B* ($s^2_B = 40$).

Exercício

Acessar a lista de exercícios 1g. Responder a questão 6.

Desvio Padrão

O *desvio padrão* é obtido simplesmente encontrando-se a raiz quadrada do valor obtido para a variância. É representado por *s*.

Utilizando os dados do exemplo anterior:

s_A = *raiz* s^2 A = *raiz* 10 = 3,16

s_B = *raiz* s^2 B = *raiz* 40 = 6,32

No caso da variância lida-se com *unidades ao quadrado*, representando uma grandeza em duas dimensões.

Para ter uma medida da variabilidade ou dispersão usa-se o *desvio padrão* que tem a mesma *unidade* dos dados originais. É um desvio médio em relação à média do conjunto de dados.

É importante notar, portanto, que quando se fala de desvio padrão ou de variância, fala-se de *medidas de dispersão*, que diferem apenas por uma transformação matemática, ou seja, diferem somente em um ajuste de escala.

Pela sua própria definição nota-se que o desvio padrão é uma medida que só pode assumir valores não negativos e que, quanto maior for, indicará mais variabilidade nos dados e que maior será a dispersão deles.

Média, variância e desvio padrão em dados classificados

Como calcular a variância e o desvio padrão se não tivermos todos os dados individuais, ou seja, quando a mostra está dividida em classes?

Por *exemplo*, supondo que na literatura tivéssemos obtido os dados sobre as frequências dos intervalos de classes apresentados para a característica idade:

Limites	*f*
4 a 7	0
8 a 11	2
12 a 15	1
16 a 19	3
20 a 23	8
24 a 27	11

Como seria possível calcular, a partir apenas desses dados:

a. média
b. variância
c. desvio padrão

Há um método fácil. Primeiramente, calcula-se o *valor central* de cada intervalo de classe (x).

Limites	***Centro x***	***f***	***fx***	***x^2***	***fx^2***
4 a 7	5,5	0	0,00	30,25	0,00
8 a 11	9,5	2	19,00	90,25	180,50
12 a 15	13,5	1	13,50	182,25	182,25
16 a 19	17,5	3	52,50	306,25	918,75
20 a 23	21,5	8	172,00	462,25	3698,00
24 a 27	25,5	11	280,50	650,25	7152,75
Total		25	537,50	1721,50	12132,25

E calcula-se:

Média = M	$\Sigma fx / n$	= 537,50 / 25 = 21,50
Variância = s^2 =	$\Sigma fx^2 - [(\Sigma fx)^2 / N]\} / (N - 1)$	= $[12132,25 - (537,50^2 / 25)] / 24$ = 24,00
Desvio padrão = s =	*raiz* s^2	= *raiz* 24,00 = 4,8990

Para facilitar os cálculos utilize uma planilha especial:

Média, Variância e Desvio Padrão em amostras com até 25 classes
com dados classificados
Copie a planilha comprimida em formato xls *ou* em ods
http://www.cultura.ufpa.br/dicas/biome/biozip/amostra.zip

Outras medidas de dispersão

A *amplitude* de um conjunto de números é a diferença entre o maior e o menor dos números do conjunto. É a medida mais simples que já mostra a dispersão dos dados.

Coeficiente de variação

Representando-se a dispersão absoluta pelo desvio-padrão (s), define-se a seguinte dispersão relativa, chamada de coeficiente de variação, como a razão entre o desvio padrão e a média amostral:

$$C = s / \bar{x}$$

O C (ou CV ou V) é geralmente expresso em porcentagem.

É importante notar que esse coeficiente é independente das unidades de medida usadas, por isso é útil para comparar distribuições, mesmo que as unidades das variáveis sejam diferentes.

Simetria e Assimetria

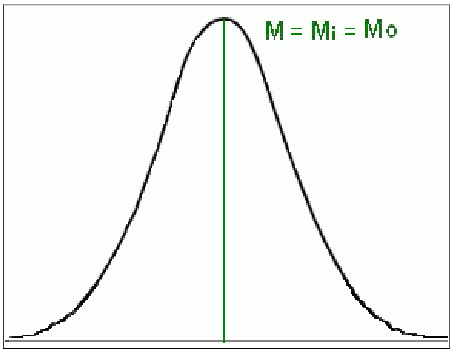
As distribuições de frequências não diferem apenas quanto ao valor médio e à variabilidade. Deve-se considerar também a sua *forma*, que pode ser simétrica ou assimétrica. Assim, uma das características mais importantes de uma distribuição de frequências é a *simetria* ou a falta dela.

Simetria: diz-se que uma distribuição de frequências é *simétrica* quando a média, mediana e moda são iguais, ou seja, coincidem num mesmo ponto, apresentando o mesmo valor.

Assimetria: Já, quando a média, mediana e a moda apresentam valores diversos, caindo em pontos diferentes da distribuição, diz-se que a distribuição de frequências é *assimétrica*.

O deslocamento desses pontos pode acontecer para a direita ou para a esquerda. Portanto, quanto ao grau de deformação, uma curva de frequência de uma distribuição unimodal pode ser:

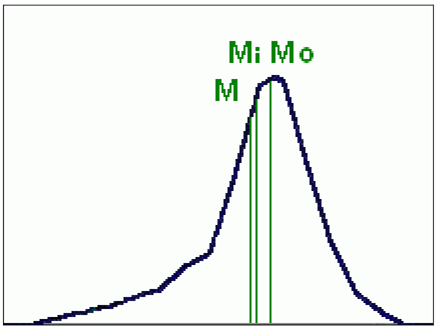
- Simétrica
- Assimétrica Positiva
- Assimétrica Negativa



Simétrica - Tem um só "pico" e apresenta o máximo de frequência no centro, diminuindo gradativamente em ambos os lados, até atingir valores extremos da escala.

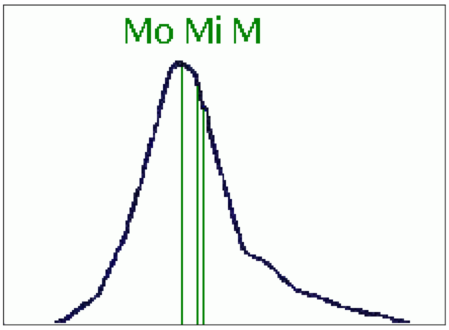
Nesse caso, *Média* = *Mediana* = *Moda* (M = Mi = Mo).

Também é chamada de campanular ou em forma de sino ou, ainda, por curva normal ou curva de Gauss.



Assimétrica Negativa - Tem um só "pico". A moda apresenta-se no máximo de frequência, sendo maior que a mediana e a média.

Nesse caso, M < Mi < Mo, ou seja a moda é o maior valor dos 3 promédios.



Assimétrica Positiva - Tem um só "pico". A moda apresenta-se no máximo de frequência, sendo menor que a mediana e a média.

Nesse caso, Mo < Mi < M, ou seja a média é o maior valor dos 3 promédios.

Curva normal - uma curva unimodal simétrica

Se desenharmos uma curva com os valores amostrados e obtivermos uma curva unimodal simétrica, em forma de sino, sabe-se que média, moda e mediana estão no ápice da curva, sendo que a distribuição de valores maiores que a média e a dos valores menores que a média é especular.

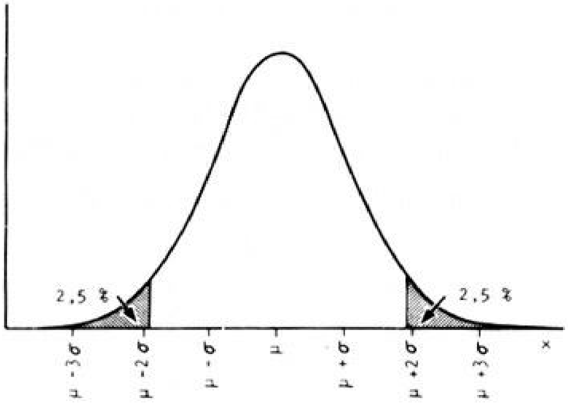
É importante lembrar que:

Símbolos usados	***Amostra***	***População***
Média	M ou $\bar{x}$	**µ**
Desvio padrão	***s***	**σ**

Assim, se passarmos uma reta exatamente no meio da curva e chamarmos de zero a intersecção entre elas (ponto onde está a média) e se for utilizada uma escala com o número de desvios padrão no eixo dos X, ou seja valores positivos do desvio padrão à direita: +1**σ**, + 2**σ**, + 3**σ** e valores negativos para a esquerda: -1**σ**, -2**σ**, -3**σ** serão encontrados os seguintes valores de probabilidade:

Sub-área da curva

População	**Amostra**	**Valor**
$\mu \pm 1\sigma$	$\bar{x} \pm 1s$	*68,26 %*
$\mu \pm 2\sigma$	$\bar{x} \pm 2s$	*95,44 %*
$M \pm 3\sigma$	$\bar{x} \pm 3s$	*99,74%*



(Se desejar saber como esses valores de *probabilidade* foram obtidos, clicar aqui).

Importante é notar que ao estudar uma variável com distribuição normal em duas ou mais amostras em geral é necessário saber se uma amostra difere significativamente das outras, ou seja, se elas podem ser consideradas como extraídas da mesma população.

Como a distribuição normal é determinada pela média e desvio padrão (ou variância) é óbvio que se as médias e variâncias de 2 ou + amostras não diferirem significativamente pode-se aceitar que elas foram extraídas da mesma população.

Assimetria

A assimetria, representada pela notação "*As*" é característica das distribuições *deformadas*.

Pearson propôs a seguinte fórmula de cálculo:

$$As = (\bar{x} - Mo) / s$$

Entretanto, a assimetria pode dar-se na cauda esquerda ou na direita da curva de distribuição de frequências, pois, se

As > 0: a distribuição é assimétrica positiva (à direita)
As < 0: a distribuição é assimétrica negativa (à esquerda)

- Curva ou Distribuição de Frequências com *Assimétrica Positiva*
  Neste caso a cauda é mais alongada à direita da ordenada máxima (ordenada correspondente à moda) e a média aritmética apresenta um valor maior do que a mediana, e esta, por sua vez, tem valor maior do que a moda. Há uma predominância de valores superiores a moda.

$$\bar{x} > Mi > Mo$$

- Curva ou Distribuição de Frequências com *Assimétrica Negativa*
  Neste caso a cauda é mais alongada à esquerda da ordenada máxima e predominam valores inferiores à moda.

De forma geral, quanto mais o valor se afastar do zero tanto maior será o grau de assimetria da curva.

Exemplo:

Usando os dados numéricos anteriores e sabendo-se que o desvio padrão é 5,58, calcule a simetria

As = (13,28 -13,19) / 5,58 = 0,0161

Este "site", destinado prioritariamente aos alunos de Fátima Conti, pretende auxiliar quem esteja começando a se interessar por Bioestatística, computadores e programas, estando em permanente construção. Sugestões e comentários são bem vindos. Agradeço antecipadamente.

——

Endereço dessa página:

HTML: http://www.cultura.ufpa.br/dicas/biome/bioamos.htm

PDF: http://www.cultura.ufpa.br/dicas/biopdf/bioamos.pdf

Última alteração: 24 mar 2011 (Solicito conferir datas. Pode haver atualização só em HTML).