

ESTATÍSTICAS

Na HP-12C

01/10/2009 Bertolo 1

O que é Estatística?

- A estatística pode ser entendida como um conjunto de ferramentas envolvidas no estudo de **métodos e procedimentos** usados para
 1. *coleccionar,*
 2. *classificar e*
 3. *analisar dados.*
- As ferramentas estatísticas oferecem também as medidas para se fazer *inferências científicas sumarizadas* resultantes de tais dados. Duas destas ferramentas são a **Média** e o **Desvio Padrão**.

01/10/2009 Bertolo 2

Conceito de Média

- Dado um conjunto de dados coletados, a média é definida como uma medida de tendência central e é a mais comumente usada.
- Seu valor é calculado como a soma de todos os pontos dados dividida pelo número de pontos dados incluídos.

01/10/2009

Bertolo

3

Conceito de Desvio Padrão

- O **desvio padrão** é um índice de variabilidade usado para caracterizar a dispersão entre os dados numa população dada ou uma amostra.
- Ele mede a dispersão ao redor da média.
- A propriedade do desvio padrão é tal que quando os dados subjacentes estão *normalmente distribuídos*, aproximadamente 68% de todos eles caem dentro de **um** desvio padrão em cada lado da média, e aproximadamente 95% de todos os valores caem dentro de **dois** desvios padrões de cada lado da média.
- Isto tem aplicação em muitos campos, particularmente quando se tenta decidir se um valor observado não é usual de ser significativamente diferente da média.

01/10/2009

Bertolo

4

Média e Desvio Padrão na HP-12C

- Na HP12C, os dados estatísticos são *armazenados* como um conjunto de somatórios resultantes dos dados coletados originalmente.
- O conjunto dos dados coletados originalmente *deve ser digitado antes* de se usar quaisquer características estatísticas disponíveis na HP12C, porque todos os valores produzidos por estas ferramentas estatísticas dependem deles.
- A organização da memória da HP12C permite o estudo dos dados estatísticos organizados como amostras de uma ou duas variáveis

01/10/2009

Bertolo

5

O que a HP-12C calcula?

- Como um procedimento geral, os dados são sempre coletados como um par de números, ou valores (x,y), por exemplo, *quantidade* e *preço* de várias mercadorias.
- HP-12C calcula as seguintes somas:
 - $\sum x_n$ $\sum y_n$ $\sum (x_n)^2$
 - $\sum (y_n)^2$ $\sum (x_n \times y_n)$

01/10/2009

Bertolo

6

Introduzindo os dados

- Para o caso de um par de dados:
- digita-se o dado y **ENTER** , o dado x e, depois, pressione a tecla $\Sigma+$
- A HP-12C calcula automaticamente as estatísticas e armazena nos registradores R1 a R6, como mostra a Tabela:

Registrador	Estatística
R1 e visor	N
R2	Σx
R3	Σx^2
R4	Σy
R5	Σy^2
R6	Σxy

01/10/2009

Bertolo

7

Exemplo

- Os preços de vendas das últimas 10 casas vendidas na comunidade Parkdale foram: \$198.000; \$185.000; \$205.200; \$225.300; \$206.700; \$201.850; \$200.000; \$189.000; \$192.100; \$200.400. Qual foi a média destes preços de venda e qual é o desvio padrão da amostra? Um preço de venda de \$240.000 será considerado não usual na mesma comunidade?
 - certifique-se de apagar as memórias somatório/estatística antes de iniciar o problema. Para isto, **f** Σ , antes de tudo
- | | | | | |
|----------|-----------|--------|-----------|--|
| • 198000 | $\Sigma+$ | 185000 | $\Sigma+$ | A cada valor digitado, seguido de um $\Sigma+$, o visor mostra o N, número de dados entrados. |
| • 205200 | $\Sigma+$ | 225300 | $\Sigma+$ | |
| • 206700 | $\Sigma+$ | 201850 | $\Sigma+$ | |
| • 200000 | $\Sigma+$ | 189000 | $\Sigma+$ | |
| • 192100 | $\Sigma+$ | 200400 | $\Sigma+$ | |
| | | | | |

01/10/2009

Bertolo

8

Exemplo - Continuação

- Para calcular a média aperte:
- **g** \bar{x} 200.355,00
- Para calcular o desvio padrão aperte:
- **g** **s** 11.189,04
- Baseado nestes números, aproximadamente 68% dos preços estão no intervalo de \$200.355,00 ± \$11.189,04.
- Aproximadamente 95% dos preços estão no intervalo \$200.355,00 ± 2×(\$11.189,04). A seqüência de teclas seguinte dá o limite inferior:
- **g** \bar{x} **ENTER** **g** **s** **2** **x** **x^y** **R↓** **-**
- 177.976,91 Análise passo a passo a pilha operacional.

01/10/2009

Bertolo

9

Análise da Pilha Operacional do Exercício

g \bar{x} **ENTER** **g** **s** **2** **x** **x^y** **R↓** **-**

T	0,00	0,00	\bar{x}	\bar{x}	\bar{x}	\bar{x}	0,00	
Z	0,00	94,00	\bar{x}	0,00	\bar{x}	\bar{x}	\bar{x}	
Y	94,00	\bar{x}	0,00	s	0,00	2s	\bar{x}	\bar{x}
X	\bar{x}	\bar{x}	s	2	2s	0,00	2s	Resulta- do

01/10/2009

Bertolo

10

E o limite superior?

- O visor mostra o limite inferior.
- Para calcular o limite superior, se nenhuma operação foi realizada após as teclas acima, pressione:
- **$x \gt y$ g LSTx +**
- **222.733,09**
- O visor mostra o limite superior.
- Resposta: \$240.000,00 é um preço não usual para uma casa na comunidade *Parkdale* baseado nos últimos 10 preços de venda.

01/10/2009

Bertolo

11

Exercícios em sala

Uma máquina industrial (A) produz peças com desvio padrão de 1,2 gramas. Outra máquina (B) produziu 10 peças com os seguintes pesos (em gramas):
 100 101 99 98 101 102 100 97 100
 100

Qual das duas máquinas produz peças mais homogêneas?

Resp.: Média de B = 99,8 e desvio padrão de B = 1,476. A máquina A produz peças mais homogêneas.

Um estudo foi realizado por um professor em três turmas, obtendo a média e o desvio padrão das notas de sua disciplina, conforme abaixo. Qual a turma com menor variabilidade? O que isso significa?

TURMA	A	B	C
MÉDIA	6,5	8,0	8,0
DESVIO PADRÃO	1,2	2,7	2,5

Resp.: A turma A tem menor variabilidade. Esta turma tirou notas mais parecidas, entretanto isso não quer dizer que teve melhor desempenho.

01/10/2009

Bertolo

12

Desvio Padrão como uma Medida do Risco

O desvio padrão é freqüentemente usado pelos investidores para medirem o risco de uma ação ou um portfolio de ações. A idéia básica é que o desvio padrão é uma medida da volatilidade: quanto mais os retornos da ação variarem do valor do retorno médio daquela ação, mais volátil é a ação.

01/10/2009

Bertolo

13

Exemplo de Desvio Padrão como uma Medida do Risco

Considere os portfolios seguintes e seus respectivos retornos (em porcentagem) durante os últimos seis meses.

A			B		
ValorInicial	Retorno(%)	Valor Final	ValorInicial	Retorno(%)	Valor Final
1.000	0,75	1.008	1.000	1,50	1.015
1.008	1,00	1.018	1.015	5,00	1.066
1.018	3,00	1.048	1.066	12,00	1.194
1.048	-1,50	1.032	1.194	-9,00	1.086
1.032	0,50	1.038	1.086	-4,00	1.043
1.038	2,00	1.058	1.043	1,50	1.058

Ambos os portfolios terminam o período aumentando em valor de \$1.000 para \$1.058. Entretanto, eles diferem claramente na volatilidade. Os retornos mensais do Portfolio A variam de -1,5% a 3,0% enquanto os do Portfolio B variam de -9,0% a 12,0%.

O desvio padrão dos retornos é uma medida melhor da volatilidade daquele intervalo porque ele leva em conta todos os valores. Assim o desvio padrão dos seis retornos para o Portfolio A é 1,52; para o Portfolio B é 7,24

01/10/2009

Bertolo

14

Desvio Padrão é importante?

Um importante atributo do desvio padrão como uma medida de espalhamento é que se a média e desvio padrão de uma distribuição normal são conhecidos, é possível calcular o percentil associado com qualquer resultado dado. Numa distribuição normal, cerca de 68% dos resultados estão dentro de um desvio padrão da média e cerca de 95% dos resultados estão dentro de dois desvios padrões da média.

O desvio padrão tem sido comprovado uma medida extremamente útil do espalhamento em parte porque ele é matematicamente tratável. Muitas fórmulas de estatística inferencial usam o desvio padrão.

01/10/2009

Bertolo

15

Média e Desvio Padrão de Duas Variáveis

- Um recenseador de terrenos quer calcular a relação entre a área construída e a área do terreno de oito casas localizadas na sua vizinhança. Inicialmente ele precisa saber a média e o desvio padrão para ambos parâmetros. Suas medidas permitiram-lhe construir o seguinte quadro:

Área do Terreno (m ²)	Área Construída (m ²)	Área do Terreno (m ²)	Área Construída(m ²)
12000	3120	9000	2080
10000	2560	10000	2700
11000	2920	13000	3280
14000	3300	12000	3080

- Certifique-se em apagar as memórias estatísticas/somatório antes de iniciar o problema. $f \Sigma$
- 3120 ENTER 12000 $\Sigma+$ Para calcular a média da de terreno: \bar{x} . Área de
- 2560 ENTER 10000 $\Sigma+$ terreno média: 11.375 m²
- 2920 ENTER 11000 $\Sigma+$ Agora pressionando **x><y** temos área média de
- 3300 ENTER 14000 $\Sigma+$ construção: 2.880 m².
- 2080 ENTER 9000 $\Sigma+$ Para calcular o desvio padrão: **g s** ...desvio padrão
- 2700 ENTER 10000 $\Sigma+$ para a área de terreno: 1.685,02 m².
- 3280 ENTER 13000 $\Sigma+$ Pressionando **x><y** temos o desvio padrão para a
- 3080 ENTER 12000 $\Sigma+$ área construída: 415,83 m².

01/10/2009

Bertolo

16

Exercício

- Uma pesquisa feita com sete vendedores de sua empresa revelou os dados da tabela dada a seguir. Quantas horas um vendedor trabalha, em média, por semana? Quanto ele vende, em média, por mês? Qual o desvio padrão das vendas e das horas trabalhadas por semana?

Vendedor	Horas por Semana	Vendas por Mês
1	32	R\$ 1.700.000,00
2	40	R\$ 2.500.000,00
3	45	R\$ 2.600.000,00
4	40	R\$ 2.000.000,00
5	38	R\$ 2.100.000,00
6	50	R\$ 2.800.000,00
7	35	R\$ 1.500.000,00

Resposta: Média das vendas é R\$ 2.171.428,57.
 Média das horas de trabalho por semana é 40 h.
 Desvio padrão das vendas: R\$ 482.059,08
 Desvio padrão das horas trabalhadas: 6,03 h.

01/10/2009

Bertolo

17

Média Ponderada

- Numa **média simples**, os valores individuais são adicionados e divididos pelo número de valores envolvidos. Com efeito, cada peso do valor ou contribuição à média é $1/n$, onde n é o número de valores na amostra.
- Comparativamente, uma **média ponderada** é uma média calculada dando diferentes pesos a alguns dos valores individuais.

Exemplos:

- uma média simples dos três números 5, 10 e 15 aplica-se um peso igual a $(1/3)$ para cada valor e a uma média simples dos três números 5, 10 e 15 aplica-se um peso igual a $(1/3)$ para cada valor e a média resultante é 10.
- Uma média ponderada ou média poderá aplicar um peso de 50% a 5 e 25% para cada um dos 10 e 15, resultando numa média ponderada de 8,75.
- Existem muitas situações onde um cálculo de média ponderada economiza uma grande porção de tempo do que usar uma abordagem de média simples.

01/10/2009

Bertolo

18

Média Ponderada - Fórmula

Dado um conjunto de dados coletados onde valores repetidos v_n ocorrem k_n vezes (peso), a média ponderada é calculada como:

$$\bar{X}_w = \frac{\sum (k_n \cdot v_n)}{\sum k_n}$$

Na HP-12C a média ponderada é calculada com o uso das teclas \boxed{g} $\boxed{X_w}$ e os conteúdos de dois somatórios são usados.

$$\bar{X}_w = \frac{\sum w X}{\sum w}$$

01/10/2009

Bertolo

19

Média Ponderada - Exemplo

- Um grande *shopping center* quer saber a média ponderada dos preços de venda de 2.000 unidades de um produto que tem o seu preço final ajustado de acordo com os primeiros dez dias de vendas. Calcule o preço médio e a média ponderada dos preços de vendas deste produto.

Preço por unidade	# de unidades vendidas	Preço por unidade	# de unidades vendidas
R\$ 24,20	354	R\$ 24,14	288
R\$ 24,10	258	R\$ 24,06	240
R\$ 24,00	209	R\$23,95	186
R\$ 23,90	133	R\$ 23,84	121
R\$ 23,82	110	R\$ 23,75	101

Certifique-se em apagar as memórias estatísticas/somatório antes de iniciar o problema.

\boxed{f} $\boxed{\Sigma}$
Médias regulares e médias ponderadas podem ser calculadas dos mesmos dados acumulados na HP12C, desde que a ordem dos valores seja entrada corretamente: valor **ENTER** peso.

Para calcular a média ponderada dos preços de venda: \boxed{g} $\boxed{X_w}$ 24,03

Para calcular o preço médio:

\boxed{g} \boxed{X} $\boxed{\downarrow}$ 23,98
Note que a tecla $\boxed{\downarrow}$ é pressionada porque o valor que aparece no visor após \boxed{g} ser pressionados é a média dos pesos e não será de nenhuma utilidade neste exemplo.

24.20 ENTER 354 $\Sigma+$ 24.14 ENTER 288 $\Sigma+$
24.10 ENTER 258 $\Sigma+$ 24.06 ENTER 240 $\Sigma+$
24.00 ENTER 209 $\Sigma+$ 23.95 ENTER 186 $\Sigma+$
23.90 ENTER 133 $\Sigma+$ 23.84 ENTER 121 $\Sigma+$
23.82 ENTER 110 $\Sigma+$ 23.75 ENTER 101 $\Sigma+$

01/10/2009

Bertolo

20

Exercício

- Estimar os custos do combustível numa viagem de férias permite planejar melhor as próximas viagens. A média ponderada é uma referência melhor para se calcular a média atual quando se compra álcool em postos com diferentes preços por litro. Calcule a média ponderada e o custo médio por litro de álcool comprado.
- A tabela abaixo se refere a uma viagem de férias regular e relaciona o álcool comprado (peso) em litros pelo preço do litro (valor).

Litros de Álcool	Preço por litro	Litros de Álcool	Preço por litro
12	R\$ 1,26	9	R\$ 1,32
13	R\$ 1,20	29	R\$ 1,12
31	R\$ 1,18	13	R\$ 1,25

Resposta: Embora o preço médio do álcool seja \$1,22 por litro, a média ponderada para esta viagem foi de \$1.20 por litro. Note que a tecla **R↓** é pressionada porque o valor que aparece no visor após **g X** ser pressionados é a média dos pesos e não será de nenhuma utilidade neste exemplo.

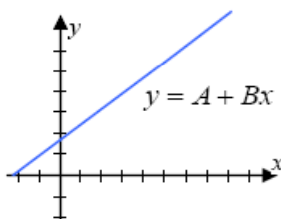
01/10/2009

Bertolo

21

Regressão Linear

- A regressão linear é um método estatístico para se encontrar uma linha reta suave que melhor se ajusta a dois ou mais pares de dados de uma amostra que está sendo analisada. Qualquer linha reta como aquela uma mostrada na Figura 1 tem dois coeficientes específicos que a localizam precisamente num sistema de coordenadas planas: um intercepto em y que denominamos de **A** e uma inclinação **B**. Estes coeficientes compõem a equação da linha reta $y = A + Bx$. É importante mencionar também que a correlação $|r|$ é sempre 1 quando somente dois pontos forem entrados.



01/10/2009

Bertolo

22

E como fazer na HP-12C?

- Na HP12C, somatórios resultantes de dados estatísticos são apropriados para cálculos de regressão linear. Dadas as coordenadas y e x de quaisquer dois ou mais pontos pertencentes a uma curva, os coeficientes de regressão linear podem ser facilmente encontrados.

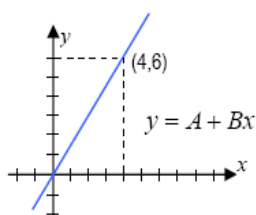
01/10/2009

Bertolo

23

Exemplo #01

- Baseado na informação apresentada no gráfico da Figura abaixo encontre o *intercepto* y e a *inclinação* para caracterizar a linha reta. Note que a linha cruza o eixo x na origem $(0,0)$.



Um dos pontos que pertence à curva é $(0,0)$ e o outro é $(4,6)$. Ambos devem ser entrados para se calcular a equação da linha. Certifique-se de limpar as memórias estatísticas/somatório antes de início do problema.

`f Σ 0 ENTER 0 Σ+ 6 ENTER 4 Σ+`

Agora calcule a inclinação (B) entrando com: (Desde que A é zero)

`1 g y,r 1,50`

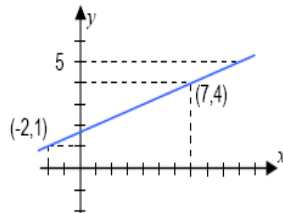
01/10/2009

Bertolo

24

Exemplo #02

- Baseado na informação apresentada no gráfico da Figura 5, compute o intercepto y e a inclinação para caracterizar a linha reta. Daí, então, use o x -previsto para computar a coordenada x relacionada à $y=5$.



Os pares de dados devem ser entrados antes de se computar os coeficientes.

1 ENTER 2 CHS $\Sigma+$

Você está certo de que as memórias estatísticas estavam limpas?

4 ENTER 7 $\Sigma+$

Como a linha não cruza o eixo x na origem, estimamos y quando $x = 0$ para achar o **A**, intercepto- y :

0 g y,r A = 1,67

Para calcular a inclinação, pressione agora:

1 g y,r x>y R↓ x>y - B = 0,33

Agora é necessário estimar x para $y=5$.

5 x,r

x = 10

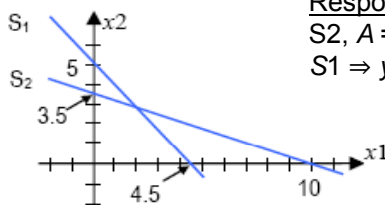
01/10/2009

Bertolo

25

Exercício

- A programação linear é uma técnica comum usada para resolver problemas de pesquisa operacional por inspeção gráfica. Baseado na informação apresentada no gráfico da Figura 10, compute o intercepto- y e a inclinação para ambas as linhas S_1 e S_2 .



Resposta: Para S_1 , $A = 5$ e $B = -1.11$.

Para S_2 , $A = 3.5$ e $B = -0.35$.

$S_1 \Rightarrow y = 5 - 1.11x$ $S_2 \Rightarrow y = 3.5 - 0.35x$

01/10/2009

Bertolo

26

Onde se aplica a Regressão Linear Simples?

A regressão linear simples é um modelo estatístico usado em várias áreas:

Variável dependente Y	Variável independente X
Renda	Consumo
Gasto com controle de qualidade (R\$)	Número de defeitos nos produtos
Memória RAM de computador	Tempo de resposta do sistema
Área construída do imóvel (m ²)	Preço do imóvel

01/10/2009

Bertolo

27

Em Economia

Em Economia, a demanda por x unidades de um produto ao preço unitário de p unidades monetárias (u.m.) é dada por uma equação envolvendo essas variáveis, chamada **equação de demanda**. Também a oferta de x unidades de um produto ao preço unitário de p (u.m.) é dada por uma equação, chamada **equação de oferta**.

Considere as tabelas abaixo:

unidades	Preço	unidades	Preço
0	6,00	0	1,00
1	5,50	1	3,00
2	5,00	2	5,00
3	4,50	3	7,00
4	4,00	4	9,00
5	3,50	5	11,00
6	3,00	6	13,00
7	2,50	7	15,00
8	2,00	8	17,00
9	1,50	9	19,00

- Encontre as equações de demanda e de oferta.
- Construa os gráficos de demanda e de oferta.
- O ponto de equilíbrio é atingido quando forem vendidas quantas unidades? A que preço?

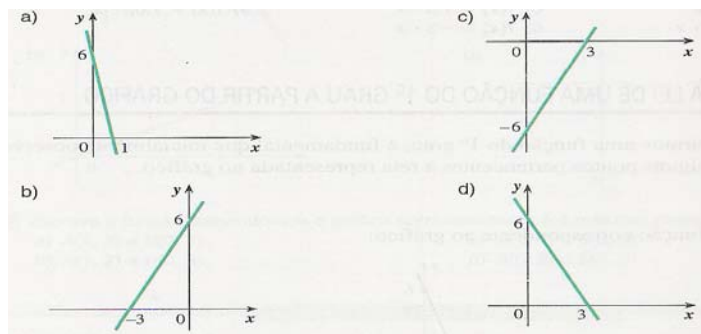
01/10/2009

Bertolo

28

Exercício de Reconhecimento

O gráfico da função $y = -2x + 6$ para $x \in \mathbb{R}$ é:



01/10/2009

Bertolo

29

Exercício de Vestibulares

- Na produção de peças, uma indústria tem um custo fixo de R\$ 8,00 mais custo variável de R\$ 0,50 por unidade produzida . Sendo x o número de unidades produzidas:
 - escreva a lei da função que fornece o custo total de peças;
 - calcule o custo de 100 unidades;
- (FGVSP) Os gastos de consumo (C) de uma família e sua renda (x) são tais que $C = 200 + 0,8x$. Podemos então afirmar que :
 - se a renda aumenta em 500, o consumo aumenta em 500,
 - se a renda diminui em 500, o consumo diminui em 500.
 - se a renda aumenta em 1000, o consumo aumenta em 800.
 - se a renda diminui em 1000, o consumo diminui em 2800.
- (VUNESP) Por uma mensagem dos Estados Unidos para o Brasil, via fax, a Empresa de Correios e Telégrafos (ECT) cobra R\$ 1,37 pela primeira pagina e R\$ 0,67 por pagina que segue, completa ou não. Qual o número mínimo de mensagens para que o preço ultrapasse o valor de R\$ 10,00 é de:

(A) 8 (B) 10 (C) 12 (D) 14 (E) 16

01/10/2009

Bertolo

30

Praticando

Construir a equação de regressão, com ela fazer um gráfico e determinar a altura do filho de um pai com 164 cm, para a distribuição de altura dos pais (X) e a altura dos filhos (y):

x	y
164	166
166	166
169	171
169	166
171	171
173	171
173	178
176	173
178	178

Resp: $y = 22 + 0,872 x$

01/10/2009

Bertolo

31

R² =

Praticando

Chama-se coeficiente de determinação R² (R-quadrado) a razão:

$$R^2 = \frac{\sum (\hat{y} - \bar{y})^2}{\sum (y - \bar{y})^2}$$

$$R^2 = \frac{125,454}{176,886} = 0,709 \text{ ou } 70,9\%$$

O coeficiente de determinação é uma medida descritiva da proporção da variação de Y que pode ser explicada por x. Segundo o modelo, temos R² = 70% dentre os 9 indivíduos estudados, i.é., 70% da variação das alturas é determinada pelos pais e 30% por outros fatores.

x	y	(y-y _{méd}) ²	(\hat{y} -y _{méd}) ²
164	166	26,122	37,255
166	166	26,122	19,011
169	171	0,0	3,42
169	166	26,122	3,42
171	171	0,012	0
173	171	0,012	3,041
173	178	47,458	3,041
176	173	3,568	19,009
178	178	47,458	37,257

01/10/2009

Bertolo

32

Coeficiente de Variação

Os retornos mensais dos investimentos em ações A e B durante os últimos 6 meses estão apresentados na tabela seguinte:

Qual dos dois apresenta maior dispersão?

Solução

Seria muito bom ressaltar que quando comparamos distribuições pode acontecer delas terem unidades e/ou valores de média bem diferentes.

	A	B
$x_{\text{médio}}$	9,33%	7,17%
s	3,72%	1,17%
CV	39,9%	16,9%

	A	B
	5%	6%
	9%	7%
	15%	9%
	12%	7%
	9%	6%
	6%	8%

É importante introduzir o coeficiente de variação, definido por: Trata-se de uma medida relativa, enquanto o s é uma medida absoluta.

O coeficiente da variação mede a variabilidade (em Finanças = RISCO do investimento).

Aqui neste exemplo temos CV da ação A é MAIOR que o CV da ação B. A ação A oferece maior risco.

$$CV = \frac{s}{x_{\text{médio}}}$$

01/10/2009

Bertolo

33

Mais Exercícios e Correlação.

1. As taxas de juros cobradas nos empréstimos para compra de eletrodomésticos em oito das maiores lojas da cidade são:

6,00% 4,80% 5,30% 4,75% 4,10% 5,40% 3,90% 5,20%.

Calcule a média e o desvio padrão. **Resp: $x_{\text{médio}} = 4,93\%$ e $s = 0,69\%$**

2. Os retornos anuais das ações X e Y durante os últimos 5 anos foram:

X	Y
12%	12%
15%	16%
12%	15%
11%	9%
14%	13%

- Qual as médias dos retornos das ações X e Y? **Resp: $x_{\text{médio}} = 12,80$ e $Y_{\text{médio}} = 13\%$.**
- Quais os desvios padrões dos retornos das ações X e Y? **$S_A = 1,64$ e $S_B = 2,74$.**
- Quais os coeficientes de variações das ações X e Y? **$CV_A = 0,13$ e $CV_B = 0,21$**
- Qual ação apresenta maior risco? **A ação Y**
- Calcule a COVARIÂNCIA entre as ações X e Y.

A covariância resume num único número a tendência e a força da relação linear entre 2 variáveis, e é dado por :

$$s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

Inserir na HP-12C os dados ao lado .

RCL 6. Iremos obter **13**. Após dividir este por **n-1**. O resultado será **3,25**

$X_i - X_{\text{Médio}}$	$Y_i - Y_{\text{Médio}}$
-0,80	-1,00
2,20	3,00
-0,80	2,00
-1,80	-4,00
1,20	0,00

01/10/2009

Bertolo

34

O Coeficiente de Correlação r_{XY} .

É definido como:

$$r_{XY} = \frac{S_{XY}}{S_X S_Y}$$

No exercício anterior, temos $13 / (1,64 \times 2,74) = 0,72$.

$$-1 \leq r_{XY} \leq 1.$$

- Se $r = +1$ As duas séries de valores estão PERFEITAMENTE CORRELACIONADAS DE FORMA POSITIVA.
- Se $r = -1$ As duas séries de valores estão PERFEITAMENTE CORRELACIONADAS DE FORMA NEGATIVA.
- Se $r = 0$ Não há relação.

O coeficiente de determinação r^2 mede a explicação da reta de regressão e pode ser obtido como elevando o coeficiente de correlação ao quadrado. O r^2 dá a porcentagem das variações de Y que podem ser explicadas pela variação de X.

O r_{XY} é encontrado na HP-12C toda vez que pressionamos $\mathbf{g y^{\wedge}}$, \mathbf{r} . Basta trocar a pilha x com a y. Aliás, é aquele valor que até agora jogamos fora.

01/10/2009

Bertolo

35

Exemplo Completo.

O diretor de vendas de uma rede de varejo nacional necessita analisar a relação entre o investimento em propaganda e as vendas da empresa. O objetivo é dispor de uma equação matemática que permita realizar projeções de vendas a partir de investimentos em propaganda. O departamento de vendas preparou a tabela abaixo com as vendas em milhões e os investimentos em propaganda em milhões dos últimos dez anos. Definir um modelo que represente a relação entre as duas variáveis ou amostras.

Propaganda	30	21	35	42	37	20	8	17	35	25
Vendas	430	335	520	490	470	210	195	270	400	480

Primeiramente inserir todos as despesas de propaganda e as vendas na HP-12C utilizando a tecla $\Sigma+$, após limpar os registros estatísticos com $f \Sigma$.

01/10/2009

Bertolo

36

Exemplo Completo – cont....

Propaganda	30	21	35	42	37	20	8	17	35	25
Vendas	430	335	520	490	470	210	195	270	400	480

Qual a equação de regressão de Vendas (Y) versus Propaganda (X)?

Resp: $Y = 117,07 + 9,74 X$

Qual o coeficiente de correlação entre as vendas e a propaganda?

Resp: $r_{XY} = 0,859366$

Qual as médias de vendas e de propaganda?

Vendas $Y_{\text{médio}} = 380,00$ Propaganda $X_{\text{médio}} = 27$

Qual os desvios padrões de vendas e de propaganda?

Vendas $s_Y = 120,16$ Propaganda $s_X = 10,60$

Qual o coeficiente de determinação? O que ele significa

Resp: $r^2 = 0,738510$. Significa que 74% das variações das vendas são explicadas pelas variações em propaganda.

Fazer um gráfico da reta de regressão.

Projetar as vendas para investimentos de 20, 30 e 45 milhões em propaganda.

Resp: 311,83; 409,21; 555,29 milhões respectivamente.

01/10/2009

Bertolo

37