

Analizando o Desvio Padrão

Esta tarefa de exploração enquadra-se no tema Estatística, trabalhado no 10.º Ano das Aprendizagens Essenciais de Matemática, concebida pelo grupo T3 da APM. Com a sua resolução comentada pretende-se mostrar algumas explorações possíveis. Na realização da tarefa foi utilizada a calculadora gráfica TI-Nspire.

Com o desenvolvimento da tarefa pretende-se que os alunos, através de uma situação real, apliquem conceitos estatísticos básicos, analisem e interpretem algumas propriedades desses

conceitos em contexto. Recorrendo à utilização da tecnologia, pretende-se que analisem o efeito que alterações nos dados iniciais produzem em medidas de localização e de dispersão, nomeadamente na média e no desvio padrão, contribuindo para a compreensão destes conceitos. Como extensão a esta proposta podem explorar-se outras situações de alteração dos dados iniciais, como, por exemplo, analisar a variação da média e do desvio padrão, quando se aumenta a frequência absoluta de cada dado, adicionando ou multiplicando por uma constante.

1. Na resolução da primeira questão, obtém-se uma tabela como a que se segue:

Tamanho	n _i	n _i ac	f _i (%)	f _i ac (%)
36	4	4	5,56	5,56
37	2	6	2,78	8,33
38	7	13	9,72	18,06
39	10	23	13,89	31,94
40	18	41	25,00	56,94
41	23	64	31,94	88,89
42	8	72	11,11	100,00

Com a análise da tabela podemos já identificar a mediana e os quartis desta distribuição.

Mínimo	36
Q1	39
Mediana	40
Q3	41
Máximo	42

Recorrendo à calculadora gráfica pode construir-se a tabela

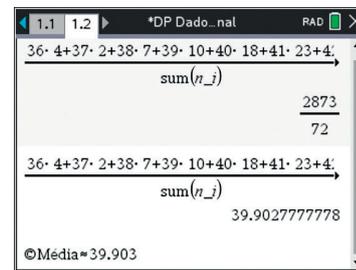
A tamanho	B n _i	C n _i ac	D f _i	E f _i ac	F
		=cumul	=n _i *1	=cumul	
1	36	4	4	5,55...	5,55...
2	37	2	6	2,77...	8,33...
3	38	7	13	9,72...	18,0...
4	39	10	23	13,8...	31,9...
5	40	18	41	25,56...	56,9...
E	f _i ac:=cumulativesum(f _i)				

Legenda:

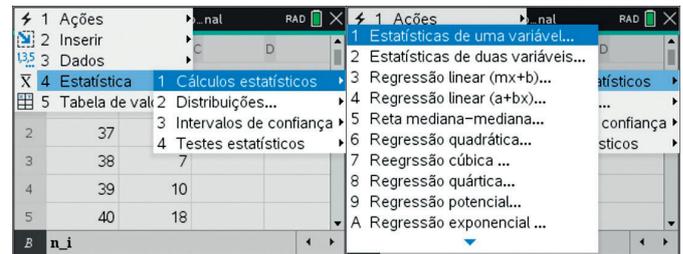
- Tamanho - valores assumidos pela variável em estudo;
- n_i – frequência absoluta da ocorrência de cada um dos valores da variável;
- n_iac – frequência acumulada das ocorrências (a última linha desta lista corresponde ao total de ocorrências)
- f_i – frequência relativa da ocorrência de cada um dos valores da variável;
- f_iac – frequência relativa acumulada da ocorrência de cada um dos valores da variável (a última linha desta lista tem que registar,100%)¹.

Na resolução da segunda questão pode calcular-se a média pela definição. Por observação dos dados, pode-se concluir que a moda é igual a 41 (valor com maior número de ocorrências).

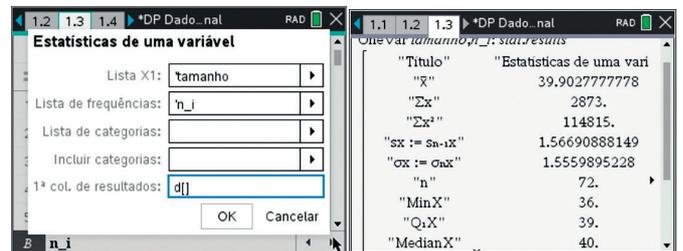
Cálculo da média pela definição



Como o objetivo é fazer variar dados e analisar as consequências, recorre-se aos comandos da calculadora gráfica, para calcular a média e o desvio padrão.



Tecla [Menu]



Associamos os valores da variável (tamanho) e as respetivas frequências (n_i)

Obtemos um resumo das estatísticas da distribuição entre as quais o desvio padrão da amostra (sx) e desvio padrão da população (σx)

¹ A função cumulativesum encontra-se no menu Listas e Folha de Cálculo3 ▶ : Dados ▶ 7:Operações da lista ▶ 1: Lista de somas cumulativas (Versão 6.0)

Nota: para melhor compreender o comportamento do desvio padrão quando se altera a frequência de um determinado valor dos dados,

seria preferível recorrer à definição de variância, e a partir desse valor obter o desvio padrão, como veremos na alínea 4.

A	ta...	B	n _i	C	ssx _i	D	ni _i ssx _i	
=							=n _i *ssx _i	
1	36	4	15.23184...			60.927391...		
2	37	2	8.426247...			16.852495...		
3	38	7	3.620647...			25.344534...		
4	39	10	0.815047...			8.1504784		
5	40	18	0.009447...			0.17006112		
C	ssx _i =(tamanho-39,9028) ²							

Legenda:

- Tamanho - valores assumidos pela variável em estudo;
- n_i - frequência da ocorrência de cada um dos valores da variável;
- ssx_i - quadrado dos desvios de cada um dos valores relativamente à média;
- ni_issx_i - quadrado dos desvios multiplicado pela frequência de cada uma das ocorrências.

Na questão 3, a construção do diagrama de extremos e quartis da distribuição é realizada recorrendo à calculadora gráfica, podendo ler-se no diagrama o valor da mediana e dos quartis.

Legenda: f_j

- Adicionar variável X
- Adicionar variável Y
- Adicionar variável X com lista de resumo
- Adicionar variável Y com lista de resumo
- Definir legenda

Gráfico de resumo

Lista X: tamanho

Lista de resumo: n_i

OK Cancelar

[CTRL]+[Page]



[Menu] 1Tipo de gráfico

Deslocando o cursor no diagrama, obtemos cada um dos quartis pretendidos

Na questão 4 faz-se variar um dos dados. Alterando a frequência do número de pares de sapatos de tamanho 42 vendidos, podemos verificar, no resumo das estatísticas, o que acontece à média e ao desvio padrão em cada uma das situações referidas nas alíneas a) a c), concretamente se o desvio padrão estaria sempre a aumentar.

8 pares de sapatos, tamanho 42

18 pares de sapatos, tamanho 42

A	tamanhc	n _i	C	D
=				=OneVar(ta
2	37	2	\bar{x}	39.902777...
3	38	7	Σx	2873.
4	39	10	Σx^2	114815.
5	40	18	sx := Sn...	1.5559895227995
6	41	23	σx := σn...	1.5559895...

28 pares de sapatos, tamanho 42

38 pares de sapatos, tamanho 42

A	tamanhc	n _i	C	D
=				=OneVar(ta
2	37	2	\bar{x}	40.358695...
3	38	7	Σx	3713.
4	39	10	Σx^2	150095.
5	40	18	sx := Sn...	1.6346624...
6	41	23	σx := σn...	1.6257541...

50 pares de sapatos, tamanho 42

(e indo mais além ...) 100 pares de sapatos, tamanho 42

inhc	n _i	C	D	
=			=OneVar(ta	
2	37	2	\bar{x}	40.675438...
3	38	7	Σx	4637.
4	39	10	Σx^2	188903.
5	40	18	sx := Sn-iX	1.6047254...
6	41	23	σx := σnX	1.5976716...

Continuando ...

200 pares de sapatos, tamanho 42

500 pares de sapatos, tamanho 42

A	tamanhc	n _i	C	D
=				=OneVar(ta
2	37	2	\bar{x}	41.428030...
3	38	7	Σx	10937.
4	39	10	Σx^2	453503.
5	40	18	sx := Sn...	1.2403736...
6	41	23	σx := σn...	1.2380222...

1 000 pares de sapatos, tamanho 42

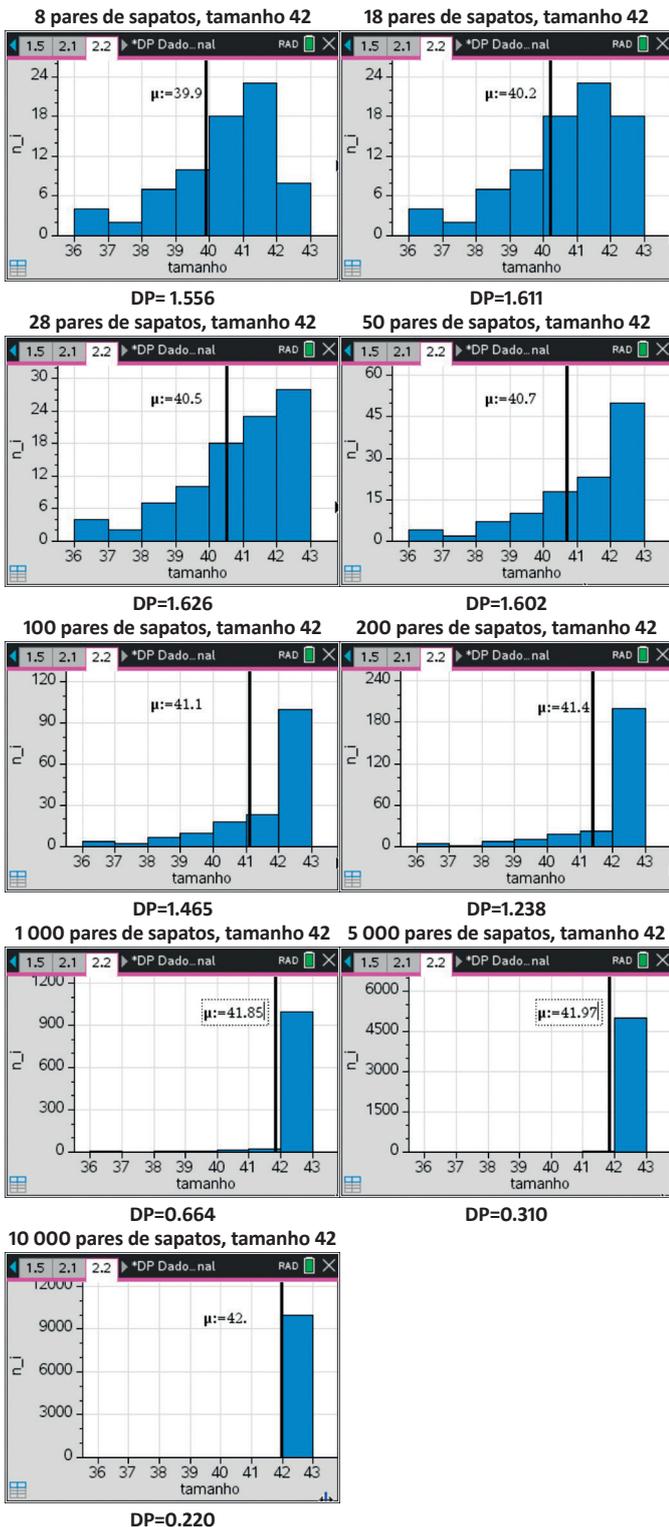
5 000 pares de sapatos, tamanho 42

A	tamanhc	n _i	C	D
=				=OneVar(ta
2	37	2	\bar{x}	41.858082...
3	38	7	Σx	44537.
4	39	10	Σx^2	1864703.
5	40	18	sx := Sn...	0.6646358...
6	41	23	σx := σn...	0.6643234...

10 000 pares de sapatos, tamanho 42 50 000 pares de sapatos, tamanho 42

A	tamanhc	n _i	C	D
=				=OneVar(ta
2	37	2	\bar{x}	41.984996...
3	38	7	Σx	422537.
4	39	10	Σx^2	17740703.
5	40	18	sx := Sn...	0.2203802...
6	41	23	σx := σn...	0.2203693...

Para explicar a variação do desvio padrão, representa-se, através de um conjunto de histogramas, variando apenas o número de ocorrências dos pares de sapatos de tamanho 42 e registando o valor da média e do desvio padrão (DP) em cada uma das situações indicadas.



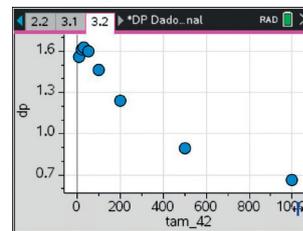
Graficamente, compreendemos que aumentando significativamente o número de pares de sapatos de tamanho 42, a média vai-se aproximando cada vez mais desse mesmo valor, e, portanto, a diferença entre esse valor e o valor médio populacional, vai diminuindo cada vez mais.

A exploração realizada favorece a compreensão do conceito de desvio padrão, como medida da dispersão dos valores da

variável em estudo em torno da média populacional, o que nem sempre é facilmente compreendido pelos alunos. Nestas experiências é visível que um baixo desvio padrão indica que os valores tendem a estar concentrados próximos da média populacional e um alto desvio padrão indica que os valores estão amplamente espalhados.

Podemos concluir que, aumentando significativamente as ocorrências de um determinado valor da variável, o desvio padrão tenderá a diminuir, embora essa diminuição não seja necessariamente monótona como podemos confirmar no gráfico abaixo.

Representação gráfica do valor do DP em função da variação do número de pares de sapatos de tamanho 42



Também algebricamente, é fácil compreender o comportamento do desvio padrão. Relembrando a definição de desvio padrão como a raiz quadrada da variância e recordando as fórmulas:

$$\text{Variância} = \frac{\sum (x_i - \mu)^2}{\sum n_i}$$

$$\text{Desvio padrão} = \sqrt{\text{variância}}$$

8 pares de sapatos, tamanho 42				100 pares de sapatos, tamanho 42			
B	n _i	C	D	B	n _i	C	D
=			=OneVar('ta	=			=OneVar('ta
9	Q ₁ X		39.	9	Q ₁ X		40.5
10	MedianX		40.	10	MedianX		42.
11	Q ₃ X		41.	11	Q ₃ X		42.
12	MaxX		42.	12	MaxX		42.
13	SSX := Σ(x- \bar{x}) ²		174.31944...	13	SSX := Σ(x- \bar{x}) ²		351.96951...
B1	=4			B7	=100		

$$\sqrt{\frac{174.32}{72}} \approx 1.56$$

$$\sqrt{\frac{351.97}{164}} \approx 1.47$$

1000 pares de sapatos, tamanho 42			
B	n _i	C	D
=			=OneVar('ta
9	Q ₁ X		42.
10	MedianX		42.
11	Q ₃ X		42.
12	MaxX		42.
13	SSX := Σ(x- \bar{x}) ²		469.57048...
B7	=1000		

$$\sqrt{\frac{469.57}{1064}} \approx 0.66$$

ANTÓNIO JOSÉ MENDES

AGRUPAMENTOS DE ESCOLAS DE VALBOM

MANUEL LAGIDO

ESCOLA SECUNDÁRIA JOSÉ RÉGIO, VILA DO CONDE

Analisando o Desvio Padrão

Uma sapataria especializada em sapatos para jovens realizou, num determinado mês, as vendas que estão registadas no quadro que se segue:

Tamanho dos sapatos	Nº de pares de sapatos vendidos
36	4
37	2
38	7
39	10
40	18
41	23
42	8

1. Constrói uma tabela de frequências absolutas e relativas, simples e acumuladas, referentes a estes dados.
2. Determina a média, a moda e o desvio padrão destes dados.
3. Constrói o diagrama de extremos e quartis desta distribuição e indica o valor da mediana e dos quartis.
4. Imagina, agora, que tinha havido um engano no registo das vendas e que tinham sido vendidos 18 pares de sapatos (em vez de 8) do tamanho 42.
 - a) Qual a alteração dos valores da média e do desvio padrão com estes novos dados, relativamente aos valores já calculados anteriormente?
 - b) E se tivessem sido vendidos 28 pares de sapatos do tamanho 42? Quais seriam, agora, os valores da média e do desvio padrão?
 - c) E se tivessem sido vendidos 38 pares de sapatos do tamanho 42? E se fossem 50 pares de sapatos?
 - d) Com o aumento das vendas do tamanho 42, como variaram os valores da média e do desvio padrão? Como explicas essas variações?