

Ciência dos Dados

Dúvidas e Revisão

Exercício 1

Considere que após a primeira prova de MatVar, Circuitos e Ciência dos dados, as notas dos alunos desta turma deram os seguintes resultados:

Medidas-resumo	MatVar	Ciência dos dados	Circuitos
<i>Média</i>	6,00	6,00	6,00
<i>Desvio padrão</i>	2,00	2,00	2,00
<i>Assimetria</i>	Esquerda	Simétrica	Direita

Se você tirou nota 6 na prova dessas três disciplinas, em qual estaria mais preocupado? E se sentindo mais confortável? Responda com base na tabela e não considerando sua habilidade pessoal em cada uma das disciplinas.

Exercício 2

Vendas (em bilhões de US\$)			Densidade		
0,00	--	0,20	0,6250		
0,20	--	0,30	2,5000		
0,30	--	0,70	0,9375		
0,70	--	1,10	?		
1,10	--	2,60	0,0463		
2,60	--	3,30	0,0595		

Das 1000 maiores empresas do país por vendas em 2009, conforme publicação da revista EXAME de julho de 2010, 72 empresas estão no setor de indústria da construção. As vendas (em bilhões de dólares) dessas empresas do setor de indústria da construção estão resumidas na tabela acima. Em média, essas empresas venderam 0,639 bilhões de dólares.

Deixe todas as contas indicadas, caso contrário, sua resposta será desconsiderada. Se utilizar a(s) coluna(s) em branco da tabela acima, explique em algum item desta questão como fez para obter os valores colocados nessa(s) coluna(s) (pode dar um exemplo para apenas uma classe).

Responda os itens a seguir.

- Qual o percentual de empresas do setor de indústria da construção com vendas entre 0,28 bilhões de dólares e 0,82 bilhões de dólares? **46,67%**
- Quais os valores de vendas que delimitam os 50% valores mais centrais das vendas das empresas do setor de indústria da construção? **0,25 e 0,70**
- Avalie a assimetria dos dados. Justifique sua resposta, caso contrário, sua resposta será desconsiderada. **Assimetria Positiva**

Exercício 3

Existe relação entre as alturas de um pai e seu filho (adulto)?

Sejam X : altura do pai (em cm) e Y : altura do seu filho adulto (em cm).

X	161	173	180	191
Y	170	173	172	185

- Com base numa medida de associação, descreva como essas variáveis se relacionam.
- Assumindo que Y é uma função linear de X , encontre os valores de a e b . Interprete-os.
- Para um filho que acaba de nascer, qual deve ser sua altura **prevista** quando for adulto sabendo que seu pai tem 185 cm?
- Um filho sempre será, em média, mais alto do que seu pai? Se não, encontre o ponto que muda esse comportamento.

Exercício 3 - Respostas

Respostas:

(A)

	PAI X	FILHO Y
Média	176,25	175,00
Desvio Padrão	10,89	5,87
Covariância	54,75	
Correlação	0,86	(A)

Correlação entre alturas do pai e filho é positiva e forte.

(B)

Coeficiente Linear	=b=	93,70
Coeficiente angular	=a=	0,46

$$\hat{y}_F = 0,46x + 93,70$$

Interpretação:

a: A cada aumento de 1 cm na altura do pai, o acréscimo esperado na altura do filho é de 0,46 cm.

b: não faz sentido fazer altura do pai (x) igual a zero.

Logo, AQUI o intercepto não tem interpretação.

Exercício 3 - Respostas

Respostas:

(C)

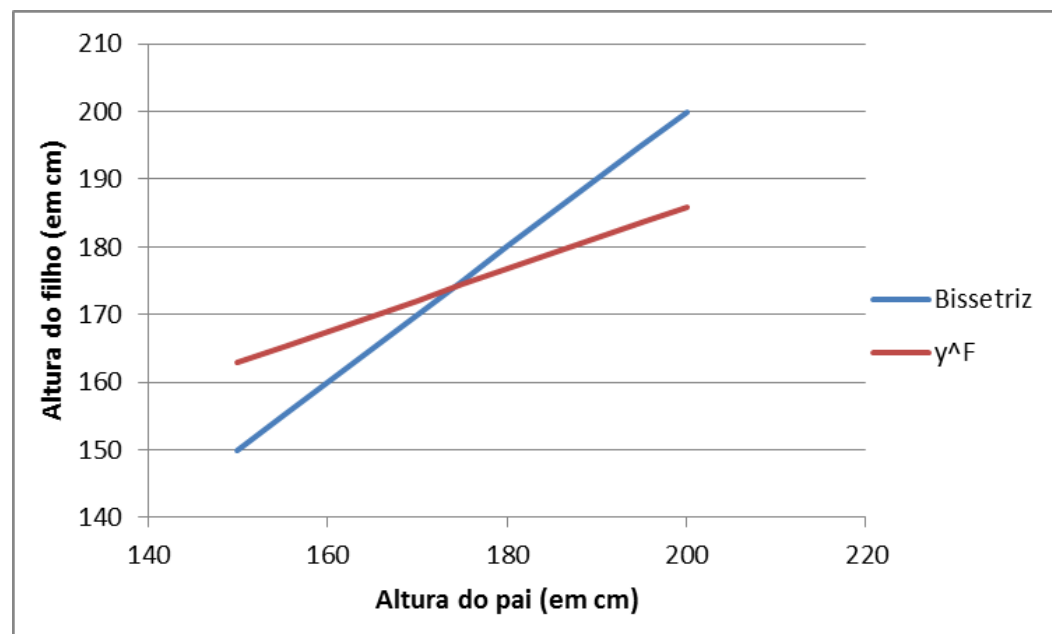
Altura do Pai =x=	185
(Altura do Filho)^ =y^=	178,80

$$\hat{y}_F = 0,46 * 185 + 93,70 = 178,80 \text{ cm}$$

A altura esperada para um filho, cujo pai tem 1,85 m, é 178,80 cm.

(D)

Altura do Pai = Altura do Filho?	$y^{\wedge}=x$	173,52
-------------------------------------	----------------	--------



Exercício 4

Uma revendedora de veículos trabalha com duas marcas de automóveis: A e B. 70% dos que adquirem carros populares, escolhem a marca A. Dentre os que adquirem carros não populares, 80% compram A. Sabe-se que 60% das vendas são de carros populares.

a. Qual é a probabilidade de um consumidor comprar um carro da marca A? 0,74

b. Sabendo que uma pessoa comprou um carro da marca A, qual é a probabilidade de ter sido um carro popular? 0,568

Exercício 5

80% dos funcionários de uma agência financeira são homens.

Nos últimos 2 anos, dos funcionários promovidos, 89% eram homens. Das mulheres, 15% foram promovidas.

Qual é a probabilidade de alguém ser promovido num prazo de 2 anos? 0,273

Se uma pessoa é do sexo masculino, qual é a probabilidade de ser promovida? E se for do sexo feminino? 0,30 e 0,15

Um comitê sindical levantou a questão de que estava havendo discriminação.

A acusação procede?

Exercício 6

Uma das responsabilidades da área de Tecnologia da Informação (TI) de uma empresa é controlar o fluxo de mensagens eletrônicas e detectar a presença de SPAMs (mensagens não desejadas). A empresa Mauro, & Falchetti tem, atualmente, 850 funcionários e, por consequência, 850 endereços eletrônicos (e-mails) individuais cadastrados. Os responsáveis pela área de TI dessa empresa detectaram que a probabilidade de um funcionário receber SPAMs, diariamente, no seu e-mail individual é de 8% e que estes são enviados de forma independente para cada funcionário da empresa.

Num dia qualquer, em que pelo menos um funcionário recebeu SPAMs no seu e-mail individual entre 40 funcionários selecionados aleatoriamente, qual é a probabilidade de cinco deles terem recebido SPAMs no seu e-mail individual?

Seja R o número de funcionários que receberam spam. Assim, $R \sim \text{Binomial}(40, 0,08)$.
Assim, $P(R = 5 | R > 0) = 0,121$.

Exercício 7

Uma das grandes preocupações do comércio varejista é com relação ao custo gerado pela devolução de mercadorias. A empresa *TemAqui* possui 260 lojas espalhadas na Grande São Paulo, tendo cada loja a taxa de 20 devoluções por quinzena.

Observação: Para cada item, descreva a variável aleatória e sua respectiva distribuição.

a) Se uma loja tiver devolução de mercadorias em um determinado dia, qual a probabilidade de serem até duas devoluções?

Considere D o número de devoluções. Do enunciado, temos que a taxa de devolução é de 1,33 mercadorias/dia. Assim, $P(D \leq 2 | D > 0) = 0,7954$

b) Para minimizar as devoluções de mercadorias feitas pelos clientes, a empresa *TemAqui* dará um bônus de R\$ 15,00 para cada loja que não tiver devolução de mercadorias em um dia; dará R\$ 9,00 se cada loja tiver uma ou duas devoluções em um mesmo dia e deixará sem bônus nos demais número de devoluções de mercadorias. Qual o bônus diário esperado para cada loja?

Defina uma nova variável B (bônus) e calcule a $P(B = 15)$, $P(B = 9)$ e $P(B = 0)$. Assim, $E(B) = 9,2262$.

c) Considere que uma amostra de 10 lojas da *TemAqui* foi selecionada aleatoriamente. Qual o número médio de lojas com até duas devoluções em um único dia entre as 10 ($=n$) selecionadas?

Defina Y : número de lojas com até duas devoluções em um único dia entre as 10 e $p = P(D \leq 2) = 85,01\%$. Assim, $E(Y) = 10 * 0,8501 = 8,5$ lojas...

Exercício 8

Em um processo industrial, o diâmetro de um rolamento é uma parte importante do processo.

Sabe-se que a probabilidade de um rolamento ter diâmetro maior do 2,98 cm é de 80%.

Sabe-se, também que a probabilidade de que um rolamento tenha diâmetro abaixo de 2,97 cm é de 10%.

Admitindo que o diâmetro de um rolamento segue uma distribuição normal, determine a média e o desvio-padrão dos diâmetros dos rolamentos que saem da linha de produção. **2,9993 e 0,0227**

Determine a especificação que represente a maior distância da média, para mais ou para menos, contendo 95% dos rolamentos produzidos.

$$1,96 * 0,0227 = 0,04449$$

Exercício 9

Um determinado calçado é vendido em lojas populares e em lojas sofisticadas. De todas as lojas, 70% são populares e 30% são sofisticadas. Nas lojas populares seu preço segue uma distribuição normal com média 8 e desvio-padrão 1,2. Já em lojas sofisticadas, o preço também segue uma distribuição normal de média 16 e desvio-padrão 3.

- a) Determine o primeiro quartil da distribuição de preços de uma loja popular. **7,196**
- b) Gastou-se mais de \$10,00 para comprar o calçado. Qual é a probabilidade da compra ter sido feita numa loja popular? **10,2%**

Distribuição Normal : Valores de $P(Z \leq z) = A(z)$

	Segunda decimal de z									
	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000