

Ciência dos dados

PROPRIEDADES DE ESPERANÇA E VARIÂNCIA

quando há combinação linear de variáveis aleatórias

Exemplo 1

Uma turma que cursa a disciplina Ciência dos dados, teve as seguintes estatísticas na avaliação intermediária:

X: nota parte papel → Média = 2,687 e DP = 1,070

Y: nota parte Python → Média = 2,663 e DP = 1,432

Ainda, a correlação entre essas notas foi -0,1484.

Calcule a nota média da AI e respectivo desvio padrão, ou seja, de $N = X + Y$ com as informações acima.

Propriedades da Esperança

Sejam a e b constantes e X e Y variáveis aleatórias:

a) $E(X+Y) = E(X) + E(Y)$

b) $E(aX+bY) = a E(X) + b E(Y)$

Se X e Y são **independentes** ou **dependentes**, os resultados acima são sempre válidos!!

Propriedades da Variância

Sejam X e Y variáveis aleatórias:

$$a) \text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y)$$

$$b) \text{Var}(X-Y) = \text{Var}(X) + \text{Var}(Y) - 2\text{Cov}(X,Y)$$

$$\text{sendo } \text{Cov}(X,Y) = E[(X-E(X))(Y-E(Y))]$$

Se X e Y são **independentes**, então

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$$

$$\text{Var}(X-Y) = \text{Var}(X) + \text{Var}(Y)$$

Propriedades da Variância

Sejam a e b constantes e X e Y variáveis aleatórias:

$$a) \text{Var}(aX+bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab \text{Cov}(X,Y)$$

$$b) \text{Var}(aX-bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) - 2ab\text{Cov}(X,Y)$$

$$\text{sendo } \text{Cov}(X,Y) = E[(X-E(X))(Y-E(Y))]$$

Se X e Y são **independentes**, então

$$\text{Var}(aX+bY) = a^2\text{Var}(X) + b^2\text{Var}(Y)$$

$$\text{Var}(aX-bY) = a^2\text{Var}(X) + b^2\text{Var}(Y)$$

Exemplo 2

Em uma determinada loja de roupas, o preço médio da calça jeans é de R\$ 118,00, com um desvio-padrão associado a essa variável de R\$ 22,00.

- a) Defina a variável aleatória X em termos do problema.
- b) Uma mãe de trigêmeas entra nessa loja e deseja comprar uma calça jeans para cada filha, mas para evitar brigas, comprará todas iguais. Qual o gasto total esperado dessa mãe e respectivo desvio padrão?
- c) Uma outra mãe de três filhas entra nessa loja e deseja comprar uma calça jeans para cada filha. Como suas filhas não têm gostos iguais, fará escolhas independentes e não necessariamente iguais. Qual o gasto total esperado dessa outra mãe e respectivo desvio padrão?

DEIXE TODOS OS RESULTADOS TEÓRICOS EXPLICITAMENTE DEMONSTRADOS NA SUA RESOLUÇÃO.

Propriedades da Esperança e Variância

Sejam X , Y e Z variáveis aleatórias quaisquer, então

$$E(X+Y+Z) = E(X) + E(Y) + E(Z)$$

$$\begin{aligned} \text{Var}(X+Y+Z) = & \text{Var}(X) + \text{Var}(Y) + \text{Var}(Z) + \\ & + 2\text{Cov}(X,Y) + 2\text{Cov}(X,Z) + 2\text{Cov}(Y,Z) \end{aligned}$$

Propriedades da Esperança e Variância

Sejam X_1, X_2, \dots, X_k variáveis aleatórias **independentes**, então

$$E(X_1 + \dots + X_k) = E(X_1) + \dots + E(X_k)$$

$$\text{Var}(X_1 + \dots + X_k) = \text{Var}(X_1) + \dots + \text{Var}(X_k)$$

Exercícios

Exercício 1

Um processo industrial pode ser executado em duas etapas independentes. O tempo gasto em cada etapa segue uma distribuição com média de 5 horas e desvio-padrão de 10 horas.

Um engenheiro resolveu modificar o processo de produção fazendo com que seja executado numa única fase, cujo tempo de execução é o dobro da primeira etapa do processo atual.

Vale a pena adotar o processo proposto pelo engenheiro?

R: Atual: $E(T) = 10$ e $DP(T) = 14,14$

Novo: $E(T) = 10$ e $DP(T) = 20$

Exercício 2

Uma pessoa irá comprar uma camisa e uma calça. Os dois produtos serão adquiridos numa mesma loja. O preço da camisa segue uma distribuição aproximadamente normal de média \$45,00 e desvio-padrão \$3,00. O preço da calça segue uma distribuição aproximadamente normal de média \$60,00 e desvio-padrão \$5,00. Sabe-se ainda que a correlação entre os preços da camisa e da calça é da ordem de 0,25.

a) Uma pessoa possui apenas \$95,00 para comprar a camisa e a calça, indo a um único estabelecimento escolhido ao acaso, qual é a probabilidade de conseguir comprar os dois produtos? Considere que o gasto total com a compra de uma calça e uma camisa também segue uma distribuição aproximadamente normal. **6,06%**

b) Indo a um único estabelecimento, qual a probabilidade de se pagar mais caro na camisa do que na calça? **0,18%**

Exercício 3 - Atividade no Python...

Você foi designado pelo seu grupo de Acionamentos para comprar os componentes para o projeto final.

Depois de cotar em diversas lojas, apurou-se que um pacote com mil peças de resistor de carbono 1/8w tem o preço distribuído como uma normal com média 21 reais e desvio padrão de 2 reais, ou seja, $X \sim N(21; 4)$.

Já um pacote de jumpers/fios macho-fêmea com 40 unidades de 20 cm tem o preço distribuído como uma normal com média 18,90 reais e desvio padrão de 1,50 reais, ou seja, $Y \sim N(18,90; 2,25)$.

Dado que você vai precisar dos componentes citados acima, se todos os grupos (n) comprarem aleatoriamente em alguma loja, qual será a média e o desvio padrão do gasto dos grupos?

No fundo, desejamos saber: Qual é μ_G e σ_G , em que:

$$G = X + Y,$$

e X e Y são definidos pelos modelos citados acima.

Inicialmente, assuma que os preços dos produtos sejam independentes.

Assuma também que haja 100 grupos. Assim, simule $n = 100$ valores de cada uma das variáveis aleatórias (X e Y) respeitando as distribuições delas.

Exercício 4

O preço de compra, em reais, de cada unidade da matéria-prima M1 não é fixo (único), pois pode variar de região para região brasileira. Neste caso, assuma que o preço de M1 pode ser modelado por uma distribuição normal com média 15 reais e desvio padrão 3 reais.

a) Qual é o menor valor cobrado pela matéria-prima M1 referente aos 39% dos locais mais careiros? $P(M1 > x) = 0,39$. Assim, $x = 15,84$.

b) Um determinado produto, que será lançado no mercado, é composto por três unidades de mesmo valor da matéria-prima M1. Ainda, na fabricação desse produto, há um custo fixo de 30 reais por produto produzido. O valor comercializado de cada produto será definido para que se tenha um lucro de 35% sobre o valor da etiqueta. Encontre a esperança, variância e distribuição do valor comercializado de cada produto.

os parâmetros da distribuição Normal de V: valor de venda
→ média = 115,3833 e variância = 191,7089

Exercício 1

Dentre os clientes de uma seguradora, 25% são mulheres. Do histórico da empresa, sabe-se que 20% das mulheres acabam acionando o seguro para alguma indenização, enquanto que apenas 10% dos homens acionam o seguro para isso.

Para definir o preço do seguro a seguradora usa a informação de que a indenização de um acidente provocado por mulheres é da ordem de R\$800,00, enquanto que a indenização de acidentes causados por homens é da ordem de R\$2000,00. Obviamente se não ocorrer acidentes, não há indenização.

a) Qual é a distribuição de probabilidades do gasto com indenizações de uma apólice qualquer? Se há 10.000 pessoas na carteira, quanto se espera gastar com indenização **considerando independência entre as pessoas**? Com que desvio-padrão?

R: $E(X) = 1,9$ milhões e $DP(X) = 54,4$ mil

b) Admitindo que, em cada apólice, incida um custo fixo de R\$50, qual deve ser o preço médio de uma apólice para que a seguradora garanta um lucro médio de R\$100 por apólice? R: 340 reais