COLECTÂNEA DE EXERCÍCIOS RESOLVIDOS DE ESTATÍSTICA

Filipe Mahaluça

2019

ÍNDICE

ESTATÍSTICA DESCRITIVA

1. Das afirmações que se seguem assinale com F (Falsas) e com V (Verdadeiras):

    a) Um pesquisador que ordena uma lista de cidades segundo o ritmo de vida, do mais lento para o mais acelerado, está operando no nível de medida nominal (**Falso**).

    b) A frequência relativa é obtida adicionando-se a frequência absoluta ao somatório das frequências posteriores (**Falso**).

    c) No estudo de determinada característica associada a uma população, deve-se recorrer a uma amostra quando for impraticável (ou mesmo impossível) observar todo o grupo (**Falso**).

    d) Os quartis de uma distribuição são $Q_1 = 4,\ Q_2 = 6\ e\ Q_3 = \ 10.$ Essa distribuição é assimétrica positiva (**Verdade**).

    e) A Estatística Descritiva compreende as técnicas por meio das quais são tomadas decisões sobre uma população com base na observação de uma amostra (**Falso**).

    f) Uma população só pode ser caracterizada se forem observados todos os seus componentes (**Falso**).

    g) As distribuições de frequências que têm a maior concentração de dados à esquerda da média são denominadas assimétricas negativas (**Falso**).

    h) Um conjunto de 500 notas de Estatística, extraídas dos arquivos da secretaria de uma universidade, constitui uma relação de dados brutos (**Verdade**).

    i) A primeira etapa, de modo a se construir uma distribuição de frequência contínua, consiste em calcular a amplitude total da distribuição (**Falso**).

    j) Ao nascer, os bebés são pesados e medidos, para saber se estão dentro das tabelas padrão. Estas duas variáveis são contínuas (**Verdade**).

2. Em um aviário foi observada a distribuição dos frangos com relação ao peso apresentado conforme a tabela a seguir:

Peso (kg)	[0.96-0.98[	[0.98-1.00[	[1.00-1.02[	[1.02-1.04[	[1.04-1.06[	[1.06-1.08[
N° de frangos	60	160	280	260	160	80

    a) Determine o peso médio?

<u>Resolução</u>

A fórmula da média para dados agrupados é:

$$\bar{x} = \frac{\sum x_i * f_i}{\sum f_i}$$

Nota que é necessário determinar o valor de $\boldsymbol{x_i}$, recorrendo a tabela temos:

classe	x_i	f_i	$x_i * f_i$
[0.96-0.98[	0.97	60	58.2
[0.98-1.00[	0.99	160	158.4
[1.00-1.02[	1.01	280	282.8
[1.02-1.04[	1.03	260	267.8
[1.04-1.06[	1.05	160	168
[1.06-1.08[	1.07	80	85.6
Total		1000	1020.8

$$\bar{x} = \frac{\sum x_i * f_i}{n} = \frac{1020.8}{1000} = 1.0208$$

Interpretação: Em média cada frango pesa cerca de 1.02Kg.

b) Queremos dividir os frangos em 4 categorias, com relação ao peso, de modo que

- os 20% mais leves sejam da categoria D
- os 30% seguintes sejam da categoria C
- os 30% seguintes sejam da categoria B
- os 20% restantes sejam da categoria A

Apresente a tabela dos limites de peso entre as categorias A, B, C, D?

Resolução

Recorrendo a fórmula para cálculos de percentis temos:

$$P_i = lim_{inf} + \frac{\frac{in}{100} - f_{ac.ant}}{f_i} * a_i$$

Determinando as frequências acumuladas (F_i) *temos:*

classe	x_i	f_i	F_i
[0.96-0.98[	0.97	60	60
[0.98-1.00[	0.99	160	220
[1.00-1.02[	1.01	280	500
[1.02-1.04[	1.03	260	760
[1.04-1.06[	1.05	160	920
[1.06-1.08[	1.07	80	1000
Total		1000	

Categoria D: $P_{20} = 0.98 + \frac{200-60}{160} * 0.02 = 0.9975$

Categoria C: $P_{50} = 1.00 + \frac{500-220}{280} * 0.02 = 1.0200$

Categoria B: $P_{80} = 1.04 + \frac{800-760}{160} * 0.02 = 1.0450$

Distribuindo os dados em tabela temos:

Categoria	Limites do peso
D	[0.9600-0.9975[
C	[0.9975-1.0200[
B	[1.0200-1.0450[
A	[1.0450-1.0800[

c) Determine o peso modal.

Resolução

Pela fórmula de Czuber temos:

$$\hat{x} = lim_{inf\hat{x}} + \frac{f_{i\hat{x}} - f_{\hat{x}i-1}}{f_{i\hat{x}} - f_{\hat{x}i-1} + f_{i\hat{x}} - f_{\hat{x}i+1}} * h = 1.00 + \frac{280 - 160}{2 * 280 - (160 + 260)} * 0.02$$
$$= 1.0171$$

Interpretação: A maior parte dos frangos tem um peso de 1.0171Kg.

d) O proprietário decide separar deste lote os animais com peso inferior a um desvio padrão abaixo da média para receberem ração reforçada e também separar os animais com peso superior a 3/4 desvios padrões acima da média para usá-los como reprodutores. Qual a percentagem de animais que serão separados em cada caso?

Resolução

Primeiro vamos determinar o valor de desvio padrão

classe	x_i	f_i	$(x_i - \bar{x})^2 * f_i$
[0.96-0.98[	0.97	60	0.154838
[0.98-1.00[	0.99	160	0.151782
[1.00-1.02[	1.01	280	0.032659
[1.02-1.04[	1.03	260	0.022006
[1.04-1.06[	1.05	160	0.136422
[1.06-1.08[	1.07	80	0.193651
Total		1000	0.69136

$$S = \sqrt{\frac{\sum(x_i - \bar{x})^2 * f_i}{n-1}} = \sqrt{\frac{0.69136}{999}} = \sqrt{0.000691} = 0.0263$$

Peso inferior a um desvio padrão abaixo da média

$$\bar{x} - s = 1.0208 - 0.0263 = 0.9945$$

$$fr\ (p < \bar{x} - s\ ) = \frac{60+160}{1000} * 100\% = 22\%$$

Interpretação: Serão separados 22% dos animais com peso inferior a um desvio padrão abaixo da média.

Peso superior a 3/4 desvio padrão acima da média

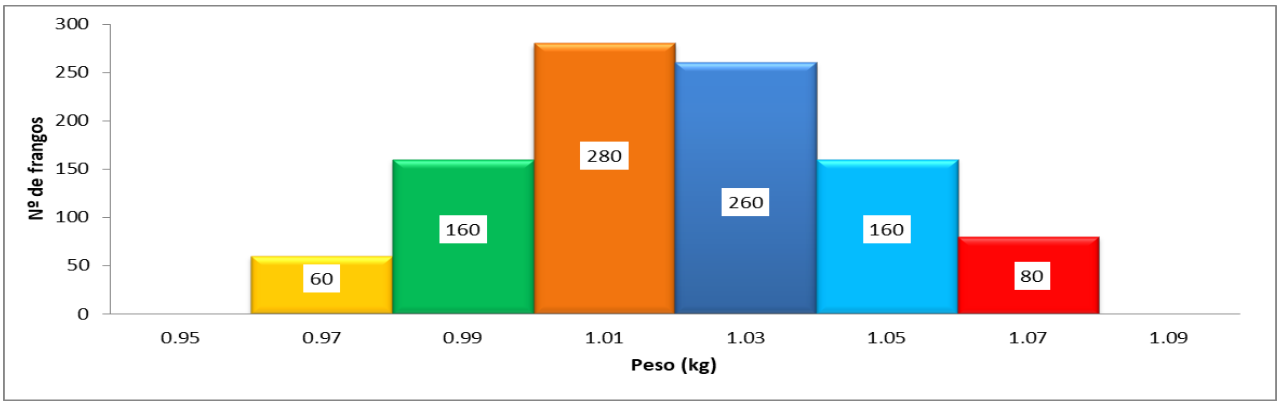
$$\bar{x} + \frac{3}{4}s = 1.0208 + \frac{3}{4} * 0.0263 = 1.04$$

$$fr\left(p > \bar{x} + \frac{3}{4}s\right) = \frac{160+80}{1000} * 100\% = 24\%$$

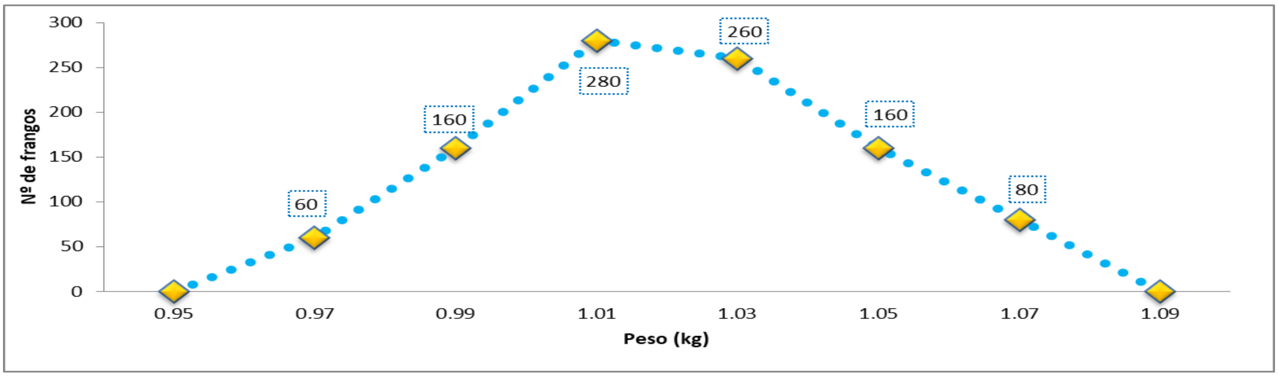
Interpretação: Serão separados 24% dos animais com peso superior a 3/4 desvios padrões acima da média.

e) Construa o histograma e polígono de frequências.

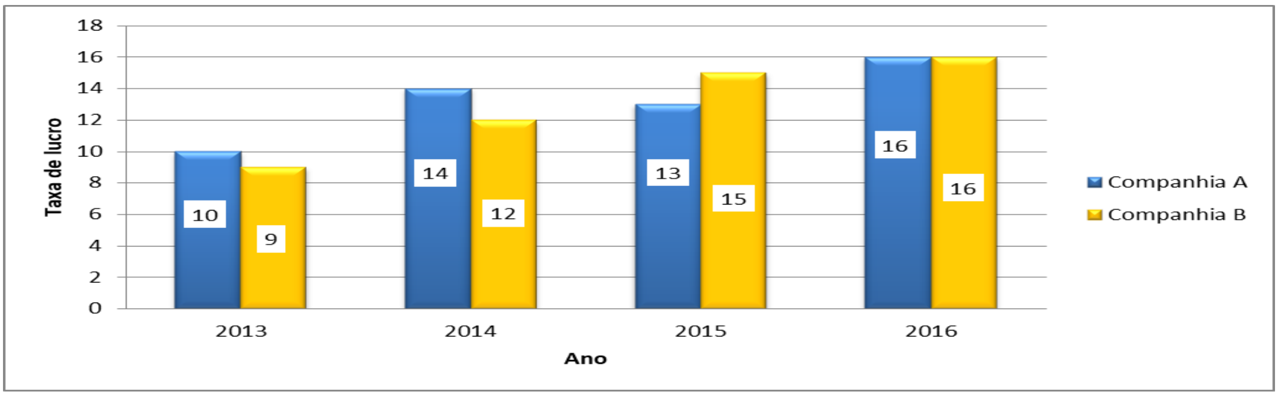
Histograma



Polígono de frequências



3. Para se estudar o desempenho de 2 companhias corretoras de acções, seleccionou-se de cada uma delas amostras das acções negociadas. Para cada acção seleccionada, computou-se a taxa de lucro apresentada durante um período de 2013 a 2016, obtendo-se o gráfico abaixo.



a) Que tipo de gráfico se trata?

Resposta: Trata-se de gráfico de barras ou colunas.

b) Indique e classifique a variável em estudo.

Resposta: A variável em estudo é a taxa de lucro e quanto a classificação é quantitativa discreta.

c) Apresente os dados em tabela.

Resolução

Ano	Taxa de lucro	
	Companhia A	**Companhia B**
2013	10	9
2014	14	12
2015	13	15
2016	16	16

d) Qual companhia teve melhor desempenho?

Resolução

Companhia A

$$\bar{x} = \frac{\sum x_i * f_i}{n} = \frac{10 + 14 + 13 + 16}{4} = 13.25$$

$$S = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{(10-13.25)^2 + (14-13.25)^2 * + (13-13.25)^2 + (16-13.25)^2}{3}} = 2.5$$

$$CV = \frac{2.5}{13.25} * 100\% = 18.9\%$$

Companhia B

$$\bar{x} = \frac{\sum x_i * f_i}{n} = \frac{9 + 12 + 15 + 16}{4} = 13$$

$$S = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{(9-13)^2 + (12-13)^2 + (15-13)^2 + (16-13)^2}{3}} = 3.16$$

$$CV = \frac{3.16}{13} * 100\% = 24.3\%$$

*Resposta: A **companhia A** foi a que teve melhor desempenho porque o seu coeficiente de variação é inferior a do **companhia B.***

e) Caracterize a distribuição do lucro para cada companhia quanto a achatamento (use a fórmula dos momentos).

Resolução

Fórmula dos momentos

Companhia A

$$K = \frac{m_4}{S^4} = \frac{\sum(x_i - \bar{x})^4}{n * S^4} = \frac{(10 - 13.25)^4 + (14 - 13.25)^4 * +(13 - 13.25)^4 + (16 - 13.25)^4}{4 * 2.5^4}$$

$$K = 1.08 < 3$$

Companhia B

$$K = \frac{m_4}{S^4} = \frac{\sum(x_i - \bar{x})^4}{n * S^4} = \frac{(9 - 13)^4 + (12 - 13)^4 * +(15 - 13)^4 + (16 - 13)^4}{4 * 3.16}$$

$$K = 0.89 < 3$$

Resposta: As duas companhias apresentam uma distribuição Leptocúrtica

4. Na Engenharia de Avaliações, diversas características de um imóvel são consideradas para o cálculo dos preços de venda e de aluguel de imóveis comerciais e residenciais. Classifique as seguintes variáveis:
   a) Área do imóvel; **quantitativa contínua**
   b) Localização do imóvel; **qualitativa nominal**
   c) Estado de conservação de imóvel; **qualitativa ordinal**
   d) Número de quartos do imóvel. **Quantitativa discreta**

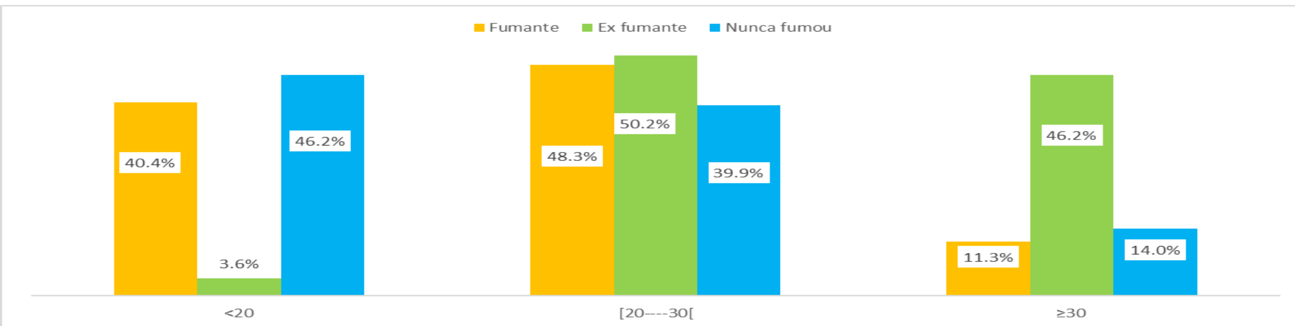
5. Construa o gráfico em frequências relativas apropriado para representar esses dados.

Hábitos de fumo	Idade (anos)		
	< 20	$[20 - 30[$	≥ 30
Fumante	143	171	40
Ex fumante	11	152	140
Nunca fumou	66	57	20

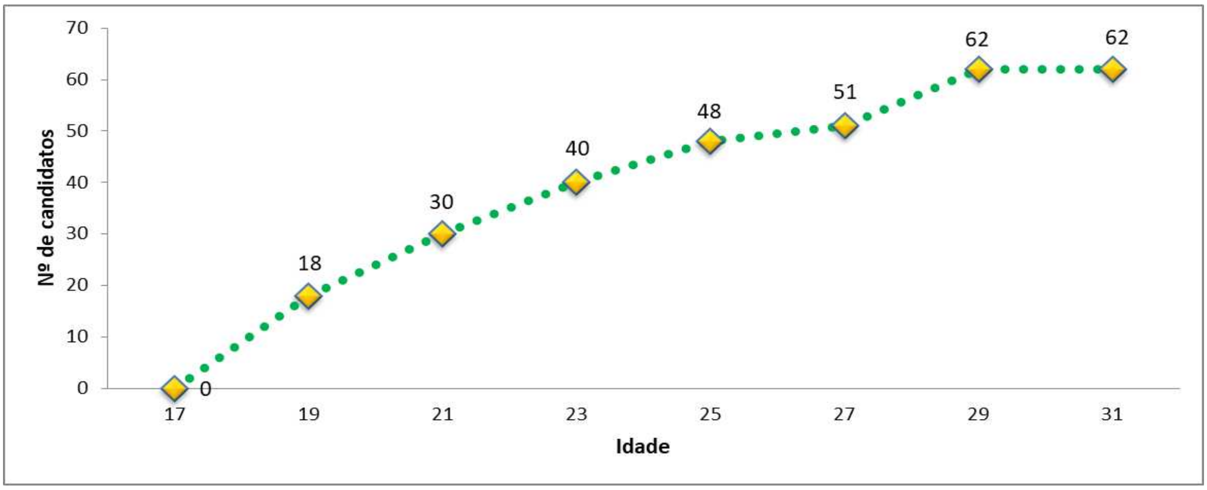
Resolução

Construíndo primeiramnte a tabela de frequências relativas, tendo em conta o denominador da soma das linhas temos:

Hábitos de fumo	Idade (anos)		
	<20	[20----30[	≥30
Fumante	40.4%	48.3%	11.3%
Ex fumante	3.6%	50.2%	46.2%
Nunca fumou	46.2%	39.9%	14.0%



6. A idade média dos candidatos a um determinado curso de aperfeiçoamento oferecido por uma empresa foi sempre baixa, da ordem de 24 anos. Como esse curso foi preparado para todas as idades, decidiu-se fazer uma campanha de divulgação. Para verificar se a campanha foi ou não eficiente, fez-se um levantamento da idade dos candidatos à última promoção, obtendo-se os resultados do gráfico a seguir:



a) Que tipo de gráfico se trata?

Resposta: Trata-se de Polígono de Frequências Acumuladas ou Ogiva

b) Apresente os dados em tabela (limites da classe, ponto médio da classe, frequência absoluta simples e acumulada).

Resolução

classe	x_i	f_i	F_i
[17-19[	18	18	18
[19-21[	20	12	30
[21-23[	22	10	40
[23-25[	24	8	48
[25-27[	26	3	51
[27-29[	28	11	62
Total		62	

c) Baseando-se nesses resultados, você diria que a campanha surtiu o efeito desejado?

Resolução

Primeiro temos que determinar o valor da média.

A fórmula da média para dados agrupados é:

$$\bar{x} = \frac{\sum x_i * f_i}{\sum f_i}$$

Nota que é necessário determinar o valor de x_i, recorrendo a tabela temos:

classe	x_i	f_i	$x_i * f_i$
[17-19[	18	18	324
[19-21[	20	12	240
[21-23[	22	10	220
[23-25[	24	8	192
[25-27[	26	3	78
[27-29[	28	11	308
Total		62	1362

$$\bar{x} = \frac{\sum x_i * f_i}{n} = \frac{1362}{62} = 21.97 \approx 22$$

Interpretação: Como a $\bar{x} < 24$, *pode-se concluir que a campanha não surtiu o efeito desejado*

d) Um outro pesquisador decidiu usar o seguinte critério: se a diferença $24 - \bar{x}$ fosse maior que o valor $\alpha = \frac{2*S}{\sqrt{n}}$, então a campanha teria sido efectiva. Qual a conclusão dele?

Resolução

Se $24 - \bar{x} > \frac{2*S}{\sqrt{n}}$ *pode-se concluir que a campanha surtiu efeito*

Primeiro vamos determinar o valor de desvio padrão, recorrendo a seguinte fórmula:

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2 * f_i}{n-1}}$$

Pela tabela temos:

classe	x_i	f_i	$(x_i - \bar{x})^2 * f_i$
[17-19[	18	18	288
[19-21[	20	12	48
[21-23[	22	10	0
[23-25[	24	8	32
[25-27[	26	3	48
[27-29[	28	11	396
Total		62	812

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2 * f_i}{n-1}} = \sqrt{\frac{812}{61}} = \sqrt{13.31148} = 3.648$$

$$24 - \bar{x} = 2 > \frac{2 * 3.648}{\sqrt{62}} = 0.927$$

Conclusão: A campanha foi efectiva.

e) Determine a idade mediana e modal dos candidatos.

Resolução

Mediana

Recorrendo a fórmula para cálculos de percentis temos:

$$P_i = lim_{inf} + \frac{\frac{in}{100} - f_{ac.ant}}{f_i} * a_i$$

$$Identificação\ da\ classe\ mediana: 50 * \frac{62}{100} = 31$$

$$P_{50} = 21 + \frac{31-30}{10} * 2 = 21.2$$

Interpretação: 50% dos candidatos tem uma idade máxima de até 21.2 anos.

<u>Moda</u>

Pela fórmula de Czuber temos:

$$\hat{x} = lim_{inf\hat{x}} + \frac{f_{i\hat{x}} - f_{\hat{x}i-1}}{f_{i\hat{x}} - f_{\hat{x}i-1} + f_{i\hat{x}} - f_{\hat{x}i+1}} * h = 17 + \frac{18-0}{2*18-(0+12)} * 2 = 18.5$$

Interpretação: A maior parte dos candidatos tem 18.5 anos de idade.

f) Se a efectividade do programa fosse baseada em distribuição assimétrica positiva, qual seria a sua conclusão?

<u>Resolução</u>

$$C_a = \frac{3*(\bar{x}-\tilde{x})}{s} = 3 * \frac{(22-21.2)}{3.648} = 0.657$$

Conclusão: Se a efectividade fosse baseada em distribuição assimétrica positiva, concluiria que foi efectiva, porque o coeficiente de assimetria é maior que zero.

g) Classifique a distribuição das idades dos candidatos quanto a achatamento.

<u>Resolução</u>

Recorrendo a fórmula dos percentis temos:

$$K = \frac{P_{75}-P_{25}}{2*(P_{90}-P_{10})}$$

$$P_i = lim_{inf} + \frac{\frac{in}{100} - f_{ac.ant}}{f_i} * a_i$$

$$P_{10} = 17 + \frac{6.2-0}{18} * 2 = 17.7$$

$$P_{25} = 17 + \frac{15.5-0}{18} * 2 = 18.7$$

$$P_{75} = 23 + \frac{46.5-40}{10} * 2 = 24.3$$

$$P_{90} = 27 + \frac{55.8-51}{11} * 2 = 27.9$$

Logo

$$K = \frac{P_{75}-P_{25}}{2*(P_{90}-P_{10})} = \frac{24.3-18.7}{2*(27.9-17.7)} = 0.275 > 0.263$$

Interpretação: Quanto a achatamento a distribuição das idades dos candidatos é Platicúrtica.

7. A empresa do ramo calçadista Pé nas Nuvens Calçados foi fundada em 1990. Desde então, sempre buscou trabalhar com uma abordagem de gestão inovadora, e tendo em vista o aumento de produtividade de seus vendedores, resolveu premiar com um aumento de 5% no salário, 36% de seus vendedores mais eficientes. Para isto, fez um levantamento de vendas semanais por vendedor. Os dados obtidos estão descritos na tabela abaixo.

Vendas (USD 1000)	[0--10[	[10--20[	[20--30[	[30--40[	[40--50[
N° de vendedores	10	25	9	7	6

a) A partir de qual volume de vendas o vendedor da Pé nas Nuvens Calçados será premiado?

Resolução

Para encontrar o valor mínimo dos 36% mais eficientes, deve-se determinar o Percentil 64, então:

Classe	f_i	F_i
[0--10[	10	10
[10--20[	25	35
[20--30[	9	44
[30--40[	7	51
[40--50[	6	57
Total	57	-

$$P_i = lim_{inf} + \frac{\frac{in}{100} - F_{i-1}}{f_i} * a_i$$

$$\frac{in}{100} = 64 * \frac{57}{100} = 36.48 \ \in [20 - 30[$$

$$P_{64} = 20 + \frac{36.48 - 35}{9} * 10 = 21.6$$

Resposta: O vendedor da Pé nas Nuvens Calçados será premiado a partir de USD 21600.00.

b) Determine o número modal do volume de vendas da empresa.

Resolução

$$\hat{x} = lim_{inf} + \frac{f_{i\hat{x}} - f_{\hat{x}i-1}}{2 * f_{i\hat{x}} - (f_{\hat{x}i-1} + f_{\hat{x}i+1})} * a_i \ = 10 + \frac{25 - 10}{2 * 25 - (10 + 9)} * 10 = 14.8$$

Resposta: O número modal do volume de vendas da empresa é de USD 14800.00.

8. Considere a tabela abaixo, sobre avaliações de nove lugares para se hospedar na Europa, responda:

Nome do Estabelecimento	País	Preço do Quarto	Número de Quartos	Pontuação por quarto
Graveteye Manor	Inglaterra	$$	18	83.6
Villa d´Este	Itália	$$$$	166	86.3
Hotel Prem	Alemanha	$	54	77.8
Hotel d´Europa	França	$$	47	76.8
Palace Luzern	Suíça	$$	326	80.9
Royal Crescent Hotel	Inglaterra	$$$	45	73.7
Hotel Sacher	Áustria	$$$	120	85.5
Duc de Bourgogne	Bélgica	$	10	76.9
Villa Gallici	França	$$	22	90.6

a) Classifique as variáveis.

Resolução

- *Nome do estabelecimento: Variável Qualitativa Nominal*
- *País: Variável Qualitativa Nominal*
- *Preço do quarto: Variável Quantitativa Discreta*
- *Número de quartos: Variável Quantitativa Discreta*
- *Pontuação por quarto: Variável Quantitativa Contínua*

b) Qual é a percentagem de hotéis com preços de quartos iguais à $$?

Resolução

$$fr\ (x = \$\$) = \frac{18 + 47 + 326 + 22}{808} * 100\% = 51.1\%$$

Resposta: A percentagem de hotéis com preços de quartos iguais à $$ é de 51.1%.

c) Calcule a pontuação global média dos quartos.

Resolução

x_i	f_i	$\sum x_i * f_i$
83.6	18	1504.8
86.3	166	14325.8
77.8	54	4201.2
76.8	47	3609.6
80.9	326	26373.4
73.7	45	3316.5
85.5	120	10260
76.9	10	769
90.6	22	1993.2
Somatório	808	66353.5

$$\bar{x} = \frac{\sum x_i * f_i}{\sum f_i} = \frac{83.6 * 18 + 86.3 * 166 + 77.8 * 54 + \cdots + 90.6 * 22}{808} = 82.1$$

d) Determine a mediana da pontuação dos quartos.

Resolução

x_i	f_i	F_i
83.6	18	18
86.3	166	184
77.8	54	238
76.8	47	285
80.9	326	611
73.7	45	656
85.5	120	776
76.9	10	786
90.6	22	808
Somatório	808	-

$$n = 2k\ ;\ k = \frac{808}{2} = 404$$

$$\tilde{x} = \frac{x_{404} + x_{405}}{2} = \frac{80.9 + 80.9}{2} = 80.9$$

9. Considere a seguinte distribuição de frequência correspondente aos diferentes preços em u.m de um determinado produto em vinte e duas lojas pesquisadas.

Preços	50,5	51	52	54.5	56	Total
Nº de lojas	2	5	8	6	1	22

a) Classifique as variáveis envolvidas.

Resolução

Preço: variável quantitativa continua

Nº de lojas: variável quantitativa discreta

b) Qual é o percentual das lojas que apresentam um preço de até 52,00 u.m?

Resolução

$$fr\ (x \leq 52) = \frac{2 + 5 + 8}{22} * 100\% = 68.18\%$$

Resposta: As lojas que apresentam um preço menor ou igual a 52 u.m perfazem 68.18%

c) Qual é o preço médio do produto em referência?

Resolução

$$\bar{x} = \frac{\sum x_i * f_i}{\sum f_i} = \frac{50.5 * 2 + 51 * 5 + 52 * 8 + 54.5 * 6 + 56 * 1}{22} = \frac{1155}{22} = 52.5$$

Resposta: o preço médio do produto em referência é de 52.2 u.m.

d) Calcule o desvio padrão e conclua quanto a variabilidade.

Resolução

$$s = \sqrt{\frac{\sum(x_{i-}\bar{x})^2 * f_i}{n-1}} = \sqrt{\frac{(50.5-52.5)+(51-52.5)+(52-52.5)+(54.5-52.5)+(56-52.5)}{21}}$$
$$= 1.05$$

Resposta: o desvio padrão é de 1.05 unidades monetárias

$$CV = \frac{s}{\bar{x}} * 100\% = \frac{1.05}{52.5} * 100\% = 2\%$$

Resposta: a distribuição possui uma variabilidade muito baixa.

10. Com o objectivo de medir o intervalo entre as transações de diferentes tipos de contas bancárias dos seus clientes, o banco ABC seleccionou aleatoriamente 28 contas bancárias em150 balcões espalhados pelo País, tendo obtido os seguintes dados:

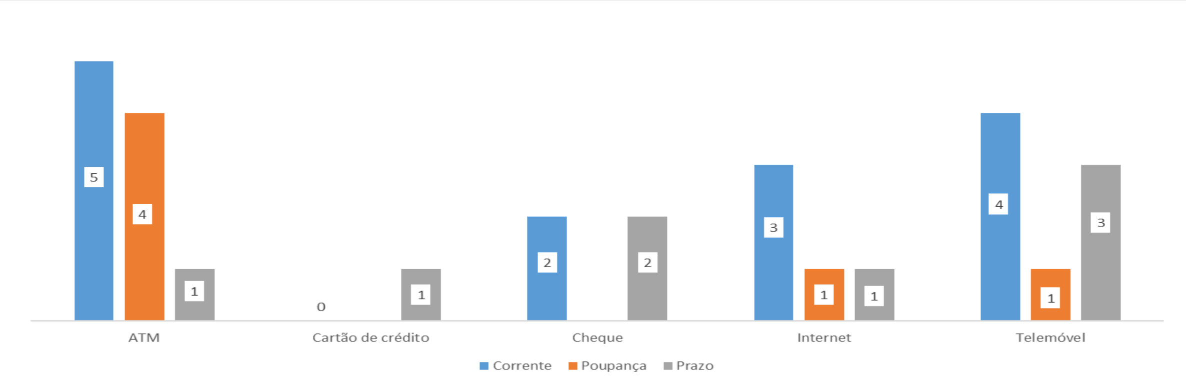
NIB	Tipo de Conta Bancaria	Meio de Movimentação	Rendimento mensal em u.m	NIB	Tipo de Conta Bancaria	Meio de Movimentação	Rendimento mensal em u.m
12A	Corrente	Cheque	120	12PH	Corrente	ATM	108
12B	Prazo	Cartão de crédito	190	12DS	Corrente	Telemóvel	192
12C	Poupança	ATM	65	12TZ	Corrente	Internet	64
12Z	Poupança	Internet	48	12UA	Prazo	Internet	36
12AB	Poupança	Telemóvel	90	12Og	Corrente	Telemóvel	74
12F	Corrente	Internet	130	12FS	Prazo	Telemóvel	72
12AN	Prazo	Telemóvel	200	12HC	Corrente	ATM	84
12CB	Prazo	Telemóvel	15	12UA	Prazo	ATM	92
12KL	Poupança	ATM	60	12QZ	Corrente	Internet	64
12BR	Corrente	ATM	46	12PF	Poupança	ATM	67
12M	Corrente	Telemóvel	92	12TZ	Corrente	Telemóvel	62
12KQ	Corrente	ATM	107	12OV	Corrente	ATM	64
12TY	Poupança	ATM	18	12IZ	Prazo	Cheque	89
12OA	Corrente	Cheque	97	12RM	Prazo	Cheque	36

a) Classifique as variáveis representadas na tabela acima em qualitativas (nominais ou ordinais) e quantitativas (discretas ou contínuas);

Resolução

- ✓ *NIB: Variável qualitativa ordinal*
- ✓ *Tipo de conta bancária: Variável qualitativa nominal*
- ✓ *Meio de movimentação: Variável qualitativa nominal*
- ✓ *Rendimento mensal: Variável quantitativa discreta*

b) Represente graficamente a distribuíção do Tipo de Conta Bancaria por Meio de Movimentação;



c) O banco adopta uma política diferenciada de concessão de crédito a dois tipos de clientes:

- **Grupo 1**: Aplicação da taxa de juro especial aos 14% dos clientes com o rendimento mais alto;
- **Grupo 2**: Aumento do tempo de reembolso aos 9% dos clientes com o rendimento mais baixo.

I. Qual é o menor rendimento do grupo 1?

Resolulão

Nota que para determinar o menor rendimento desse grupo, deve-se determinar o percentil 86. Pela fórmula dos percentis (Mendenhall e Sincich) para dados brutos temos:

$$P_i = X_m + (p - m) * (X_{m+1} - X_m)$$

$$p = \frac{i * (n+1)}{100} = \frac{86 * 29}{100} = 24.94$$

Logo

$$P_{86} = X_{24} + (24.94 - 0.94) * (X_{25} - X_{24})$$

Organizando os dados em Roll temos:

$$P_{86} = 120 + (24.94 - 24) * (130 - 120)$$

$$P_{86} = 129.4$$

Resposta: O menor rendimento mensal do grupo 1 é de 129.40 unidades monetárias.

II. Qual é o maior rendimento do grupo 2?

Resolução

Nota que para determinar o maior rendimento desse grupo, deve-se determinar o percentil 9. Pela fórmula dos percentis (Mendenhall e Sincich) para dados brutos temos:

$$P_i = X_m + (p - m) * (X_{m+1} - X_m)$$

$$p = \frac{i * (n+1)}{100} = \frac{9 * 29}{100} = 2.61$$

Logo

$$P_9 = X_2 + (2.61 - 2) * (X_3 - X_2)$$

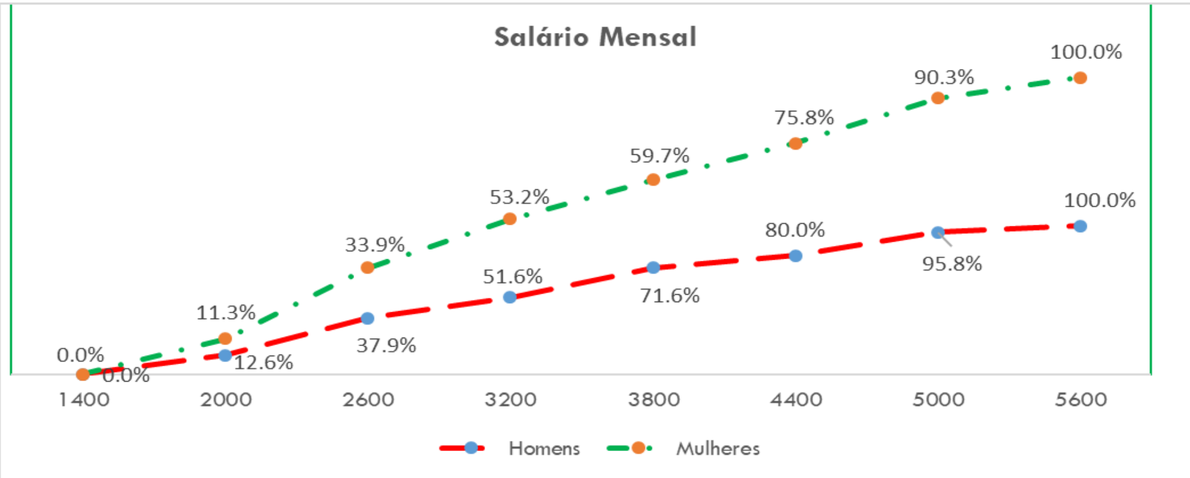
Organizando os dados em Roll temos:

$$P_9 = 18 + (2.61 - 2) * (36 - 18)$$

$$P_{86} = 28.98$$

Resposta: O maior rendimento mensal do grupo 2 é de 28.98 unidades monetárias.

11. Pretende-se investigar o nível de remuneração salarial dos homens e mulheres de certa categoria profissional. De duas amostras obtidas entre dois grupos (62 mulheres e 95 homens), destacam-se os seguintes resultados (em unidades monetárias):



a) Que tipo de grafico se trata?

Resposta:

Trata-se de polígono de frequências acumuladas ou Ogiva

b) Represente os dados em tabela de frequências.

Resolução

Nota que os dados estão representados em frequências relativas, conhecendo-se o tamanho de amostra de cada grupo temos:

		Frequência Absoluta Simples (f_i)		
Salário	x_i	*Homens*	*Mulheres*	*Total*
1400-2000	*1700*	*12*	*7*	*19*
2000-2600	*2300*	*24*	*14*	*38*
2600-3200	*2900*	*13*	*12*	*25*
3200-3800	*3500*	*19*	*4*	*23*
3800-4400	*4100*	*8*	*10*	*18*
4400-5000	*4700*	*15*	*9*	*24*
5000-5600	*5300*	*4*	*6*	*10*
Total		*95*	*62*	*157*

c) Usando a medida apropriada, diga qual dos grupos apresenta o salário mais homogéneo?

Resolução

Tratando-se de duas amostras, a melhor medida para medir a homogeneidade é o coeficiente de variação, cuja a fórmula é:

$$CV = \frac{s}{\bar{x}} * 100\%$$

Onde

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2 * f_i}{\sum f_i - 1}} \quad e \quad \bar{x} = \frac{\sum x_i * f_i}{n}$$

Recorrendo a tabela temos:

Salário	x_i	f_i		$x_i * f_i$		$(x_i - \bar{x})^2 * f_i$	
		Homens	Mulheres	Homens	Mulheres	Homens	Mulheres
1400-2000	1700	12	7	20400	11900	27113803.9	19169438.1
2000-2600	2300	24	14	55200	32200	19576660.4	15577585.8
2600-3200	2900	13	12	37700	34800	1194761.2	2482539.0
3200-3800	3500	19	4	66500	14000	1674189.5	84287.2
3800-4400	4100	8	10	32800	41000	6434606.1	5552653.5
4400-5000	4700	15	9	70500	42300	33608044.3	16285130.1
5000-5600	5300	4	6	21200	31800	17586987.3	22701914.7
Total		95	62	304300	208000	107189052.6	81853548.4

Logo:

$$\bar{x}_{Homens} = \frac{304300}{95} = 3203.2$$

$$\bar{x}_{Mulheres} = \frac{208000}{62} = 3354.8$$

$$s_{Homens} = \sqrt{\frac{107189052.6}{94}} = 1067.9$$

$$s_{Mulheres} = \sqrt{\frac{81853548.4}{63}} = 1158.4$$

$$CV_{Homens} = \frac{1067.9}{3203.2} * 100\% = 33.3\%$$

$$CV_{Mulheres} = \frac{1158.4}{3354.8} * 100\% = 34.5\%$$

Resposta: O grupo dos homens apresenta o salário mais homogéneo em relação ao das mulheres, porque tem o menor valor de coeficiente de variação.

d) Determine o salário mais frequente para cada grupo;

Resolução

Aqui pede se a moda e como as amplitudes são iguais pode-se recorrer a fórmula de Czuber, então:

$$\hat{x} = lim_{inf} + \frac{f_{i\hat{x}} - f_{\hat{x}i-1}}{2 * f_{i\hat{x}} - (f_{\hat{x}i-1} + f_{\hat{x}i+1})} * a_i$$

	f_i	
Salário	Homens	Mulheres
1400-2000	12	7
2000-2600	24	14
2600-3200	13	12
3200-3800	19	4
3800-4400	8	10
4400-5000	15	9
5000-5600	4	6
Total	95	62

→ **Classe modal**

$$\hat{x}_{Homens} = 2000 + \frac{24 - 12}{2 * 24 - (12 + 13)} * 600 = 2313.04$$

$$\hat{x}_{Mulheres} = 2000 + \frac{14 - 7}{2 * 14 - (7 + 12)} * 600 = 2466.7$$

Resposta: O salário mais frequente no grupo de homens é de 2,313.04 u.m e nas mulheres é de 2,466.70 u.m.

e) Classifique a distribuíção salarial (independentemente do sexo) desta categoria profissional quanto assimetria (Use a fórmula dos quartis).

Resolução

O Coeficiente Quartílico de Assimetria tem a seguinte fórmula:

$$C_a = \frac{Q_1 + Q_3 - 2 * Q_2}{Q_3 - Q_1}$$

onde:

$$Q_i = lim_{inf} + \frac{\frac{in}{4} - F_{i-1}}{f_i} * a_i$$

Determinando as frequências acumuladas teremos:

Salário	f_i	F_i
1400-2000	19	19
2000-2600	38	57
2600-3200	25	82
3200-3800	23	105
3800-4400	18	123
4400-5000	24	147
5000-5600	10	157
Total	157	

Então:

$$Q_1 = 2000 + \frac{39.25 - 19}{38} * 600 = 2319.7$$

$$Q_2 = 2600 + \frac{78.5 - 57}{25} * 600 = 3116$$

$$Q_1 = 3800 + \frac{117.75 - 105}{18} * 600 = 4225$$

Logo:

$$C_a = \frac{2319.5 + 4225 - 2 * 3116}{4225 - 2319.5} = 0.164$$

Interpretação: $C_a = 0.164 > 0.000$, *A distribuíção é assimétrica positiva.*

12. Em 1600 fábricas diferentes de uma determinada indústria, obteve-se os seguintes dados em relação a produção dum certo produto:

- $Média\ =\ 18\ toneladas;$
- $1^{\underline{o}}\ Quartil\ =\ 9\ toneladas;$
- $70^{\underline{o}}\ Percentil\ =\ 25\ toneladas;$
- $Variança\ =\ 25\%\ da\ média.$

Pede-se:

a) Quantas fábricas produziram até 9 toneladas?

Resolução

Apartir do primeiro quartil temos:

$$25\%\ de\ 1600 = 400$$

Resposta: 400ou 25% de fábricas produzem até 9 toneladas.

b) Quantas fábricas produziram mais de 25 toneladas?

Resolução

Apartir do Percentil 70 temos:

$$(100\% - 70\%)\ de\ 1600 = 480$$

Resposta: 480 ou 30% de fábricas produzem até 9 toneladas.

c) Determine o coeficiente de variação.

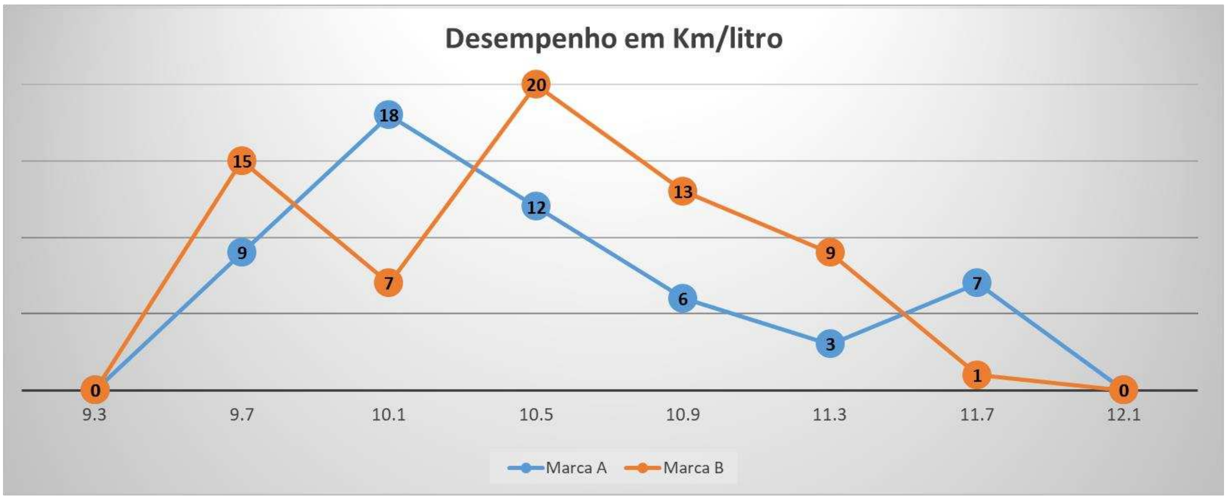
Resolução

Pela fórmula temos:

$$CV = \frac{s}{\bar{x}} * 100\% = \frac{\sqrt{0.25 * 18}}{18} * 100\% = 11.8\%$$

Interpretação: A dispersão da produção é de 11.8% medido pelo coeficiente de variação.

13. Duas marcas de veículos pretendem comparar o desempenho de seus modelos populares. Para isso, a marca A selecionou 55 veículos de sua produção e fez um teste de consumo. A marca B retirou uma amostra com 65 veículos e realizou o mesmo teste. Os dados em km/litro de cada marca estão representados no gráfico abaixo:



a) Que tipo de grafico se trata?

Resposta:

Trata-se de polígono de frequências acumuladas ou Ogiva

b) Represente os dados em tabela de frequências.

Resolução

Nota que os dados estão representados em frequências relativas, conhecendo-se o tamanho de amostra de cada grupo temos:

		Frequência Absoluta Simples (f_i)		
Desempenho	x_i	Marca A	Marca B	Total
9.5----9.9	9.7	9	15	24
9.9----10.3	10.1	18	7	25
10.3----10.7	10.5	12	20	32
10.7----11.1	10.9	6	13	19
11.1----11.5	11.3	3	9	12
11.5----11.9	11.7	7	1	8
Total		55	65	120

c) Usando a medida apropriada, diga qual das marcas apresenta o desempenho mais homogéneo?

Resolução

Tratando-se de duas amostras, a melhor medida para medir a homogeneidade é o coeficiente de variação, cuja a fórmula é:

$$CV = \frac{S}{\bar{x}} * 100\%$$

Onde

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2 * f_i}{\sum f_i - 1}} \quad e \quad \bar{x} = \frac{\sum x_i * f_i}{n}$$

Recorrendo a tabela temos:

Desempenho	x_i	f_i		$x_i * f_i$		$(x_i - \bar{x})^2 * f_i$	
		Marca A	*Marca B*	*Marca A*	*Marca B*	*Marca A*	*Marca B*
9.5----9.9	9.7	9	15	87.3	145.5	5.76	9.6
9.9----10.3	10.1	18	7	181.8	70.7	2.88	1.12
10.3----10.7	10.5	12	20	126	210	0	0
10.7----11.1	10.9	6	13	65.4	141.7	0.96	2.08
11.1----11.5	11.3	3	9	33.9	101.7	1.92	5.76
11.5----11.9	11.7	7	1	81.9	11.7	10.08	1.44
Total		55	65	576.3	681.3	21.6	20

Logo:

$$\bar{x}_A = \frac{576.3}{55} = 10.5 \qquad \bar{x}_B = \frac{681.3}{65} = 10.5$$

$$s_A = \sqrt{\frac{21.6}{54}} = 0.627 \qquad s_B = \sqrt{\frac{20}{64}} = 0.555$$

$$CV_A = \frac{0.627}{10.5} * 100\% = 5.97\% \qquad CV_B = \frac{0.55}{10.5} * 100\% = 5.24\%$$

Resposta: A marca B apresenta o desempenho mais homogéneo em relação a marca B, porque tem o menor valor de coeficiente de variação.

d) Determine o desempenho modal de cada marca;

<u>Resolução</u>

Aqui pede se a moda e como as amplitudes são iguais pode-se recorrer a fórmula de Czuber, então:

$$\hat{x} = lim_{inf} + \frac{f_{i\hat{x}} - f_{\hat{x}i-1}}{2 * f_{i\hat{x}} - (f_{\hat{x}i-1} + f_{\hat{x}i+1})} * a_i$$

Desempenho	f_i	
	Marca A	Marca B
9.5----9.9	9	15
9.9----10.3	18	7
10.3----10.7	12	20
10.7----11.1	6	13
11.1----11.5	3	9
11.5----11.9	7	1
Total	55	65

Classe modal A: 9.9----10.3

Classe modal B: 10.3----10.7

$$\hat{x}_A = 9.9 + \frac{18-9}{2*18-(12+9)}*0.4 = 10.14$$

$$\hat{x}_B = 10.3 + \frac{20-7}{2*20-(7+13)}*0.4 = 10.56$$

Resposta: A maior parte dos veículos da Marca A apresentam um desempenho de 10.14 km/litro e na Marca B o dempenho mais frequente é de 10.56km/litro.

e) Classifique a distribuíção do desempenho (independentemente da marca) quanto assimetria e achatamento (Use a fórmula dos separatizes).

Resolução

Assimetria:

O Coeficiente Quartílico de Assimetria tem a seguinte fórmula:

$$C_a = \frac{Q_1 + Q_3 - 2*Q_2}{Q_3 - Q_1}$$

onde:

$$Q_i = lim_{inf} + \frac{\frac{in}{4} - F_{i-1}}{f_i} * a_i$$

Determinando as frequências acumuladas teremos:

Desempenho	f_i	F_i
9.5----9.9	24	24
9.9----10.3	25	49
10.3----10.7	32	81
10.7----11.1	19	100
11.1----11.5	12	112
11.5----11.9	8	120
Total	120	

Então:

$$Q_1 = 9.9 + \frac{30-24}{25}*0.4 = 9.996$$

$$Q_2 = 10.3 + \frac{60-49}{32}*0.4 = 10.4375$$

$$Q_1 = 10.7 + \frac{90-81}{19}*0.4 = 10.889$$

Logo:

$$C_a = \frac{9.996 + 10.889 - 2*10.4375}{10.889 - 9.996} = 0.0112$$

Interpretação: $C_a = 0.112 > 0.000$, *A distribuíção é assimétrica positiva.*

Achatamento:

O Coeficiente Percentílico de Curtose tem a seguinte fórmula:

$$K = \frac{Q_3 - Q_1}{2 * (P_{90} - P_{10})}$$

onde:

$$P_i = lim_{inf} + \frac{\frac{in}{100} - F_{i-1}}{f_i} * a_i$$

Então:

$$P_{90} = 11.1 + \frac{108 - 100}{12} * 0.4 = 11.367$$

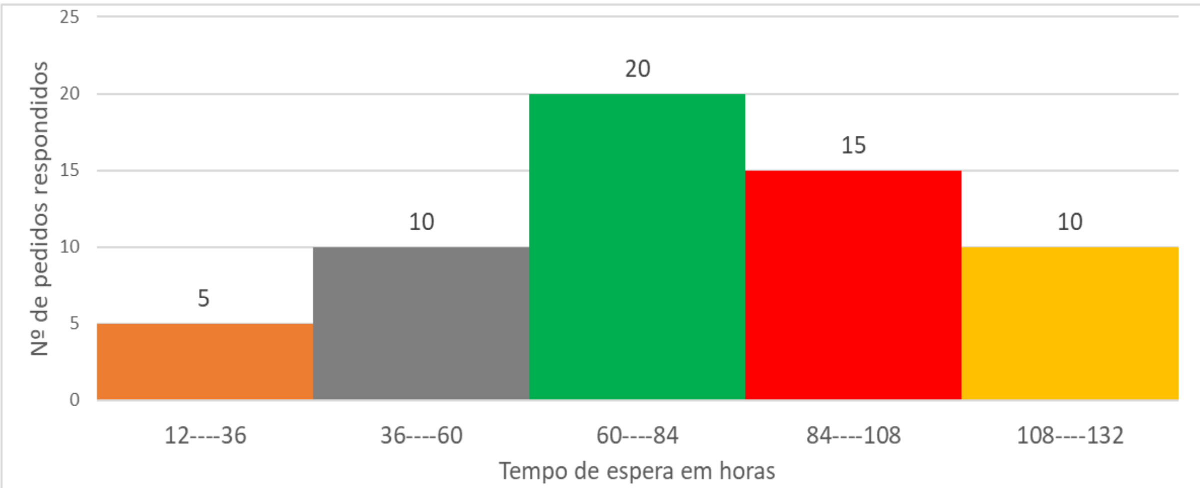
$$P_{10} = 9.5 + \frac{12 - 0}{24} * 0.4 = 9.7$$

Logo:

$$K = \frac{10.889 - 9.996}{2 * (11.367 - 9.7)} = 0.268$$

Interpretação: $K = 0.268 > 0.263$, *A distribuíção é platicúrtica.*

14. Considere o gráfico abaixo, que representa uma amostra de tempos (em horas) entre a formulação/apresentação de um pedido de financiamento e resposta num banco Comercial, em Maputo.



a) Identifique e classifique a variável em estudo;

Resposta: Trata-se do tempo de espera em horas. É uma variável quantitativa contínua.

b) Represente os dados em tabela de frequências absoluta e relativa (simples e acumulada);

Resolução

Como trata-se do histograma, a tabela de frequência será:

Classes	x_i	f_i	F_i	f_r	F_r
12----36	24	5	5	8.3%	8.3%
36----60	48	10	15	16.7%	25.0%
60----84	72	20	35	33.3%	58.3%
84----108	96	15	50	25.0%	83.3%
108----132	120	10	60	16.7%	100.0%
Total	-	60	-	100.0%	-

c) Determine a percentagem de pedidos respondidos em menos de 72 horas;

Resolução

$$fr(x < 72) = \frac{f(x = 24) + f(x = 48)}{n} = \frac{5 + 10}{60} * 100\% = 25\%$$

Resposta: A percentagem de pedidos respondidos em menos de 72 horas é de 25%.

d) Calcule o tempo de espera mais frequente;

Resolução

Aqui pede se a moda e como as amplitudes são iguais pode-se recorrer a fórmula de Czuber, então:

$$\hat{x} = lim_{inf} + \frac{f_{i\hat{x}} - f_{\hat{x}i-1}}{2 * f_{i\hat{x}} - (f_{\hat{x}i-1} + f_{\hat{x}i+1})} * a_i$$

Pela tabela de frequências, a classe modal é: 60----84

Logo:

$$\hat{x} = 60 + \frac{20 - 10}{2 * 20 - (10 + 15)} * 24 = 76$$

Interpretação: O tempo de espera mais frequente é de 76 horas.

e) Determine o tempo de espera máximo dos 15% de pedidos com resposta mais rápida.

Resolução

Para encontrar o tempo de espera máximo dos 15% de pedidos com resposta mais rápida, deve-se determinar o P_{15}.

Como os dados estão agrupados em classes, pode se usar a seguinte fórmula:

$$P_i = lim_{inf} + \frac{\frac{in}{100} - F_{i-1}}{f_i} * a_i$$

Pela tabela de frequências, a classe P_{15} é: 36----60

Então,

$$P_{15} = 36 + \frac{\frac{15 * 60}{100} - 5}{10} * 24 = 45.6$$

Interpretação: O tempo de espera máximo dos 15% de pedidos com resposta mais rápida é de 45.6 horas.

15. Um levantamento foi realizado com relação ao tempo com que os serviços de atendimento ao consumidor (SACs) de fabricantes de computadores solucionam chamados técnicos. Foram obtidos os seguintes resultados sobre o número de dias que os SACs de 14 fabricantes de computadores necessitaram para resolver certo problema.

Fabricante	1	2	3	4	5	6	7
Dias para resolver	13	27	11	14	14	17	16
Fabricante	8	9	10	11	12	13	14
Dias para resolver	21	27	12	14	20	40	17

a) Identifique e classifique a variável que está sendo estudada.

Resposta:

Tempo para solução de chamados técnicos nos serviços de atendimento ao consumidor (SACs) de fabricantes de computadores. Trata-se de uma variável quantitativa contínua.

b) Qual fabricante resolveu o problema mais rapidamente? Em quantos dias o problema foi resolvido?

Resposta:

Fabricante 3. Resolveu o problema em 11 dias.

c) Qual fabricante demorou mais para resolver o problema? Em quantos dias o problema foi resolvido?

Resposta:

Fabricante 13. Resolveu o problema em 40 dias.

d) Obtenha o número mediano de dias necessários para que o problema fosse resolvido.

Resolução

Trata-se de dados não agrupados. Primeiro deve-se organizar os dados em Roll:

Fabricante	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Dias	11	12	13	14	14	14	16	17	17	20	21	27	27	40

$$n = 2k\ (par), isto\ é, n = 14\ e\ k = 7, então:$$

$$\tilde{x} = \frac{x_k + x_{k+1}}{2} = \frac{x_7 + x_8}{2} = \frac{16 + 17}{2} = 16.5$$

Interpretação: 50% de fabricantes de computadores solucionam chamados técnicosno máximo até 16.5 dias.

e) Calcule os quartis Q1 e Q3 para o número de dias em questão.

Resolução

Pela definição de Mendenhall e Sincich para o i-ésimo percentil de n valores ordenado temos:

$$P_i = X_m + (p - m) * (X_{m+1} - X_m)$$

Onde

$$p = \frac{i * (n+1)}{100}$$

Para Q1

$$p = \frac{i * (n+1)}{100} = 0.25 * 15 = 3.75$$

$$P_{25} = X_3 + (3.75 - 3) * (X_{3+1} - X_3)$$

$$P_{25} = 13 + (3.75 - 3) * (14 - 13) = 13.75$$

Para Q3

$$p = \frac{i * (n+1)}{100} = 0.75 * 15 = 11.25$$

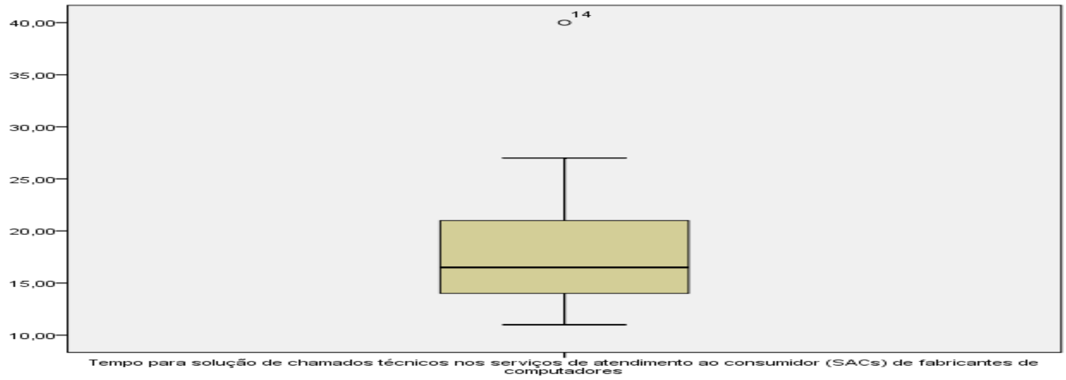
$$P_{75} = X_{11} + (11.25 - 11) * (X_{11+1} - X_{11})$$

$$P_{75} = 21 + (11.25 - 11) * (27 - 21) = 22.5$$

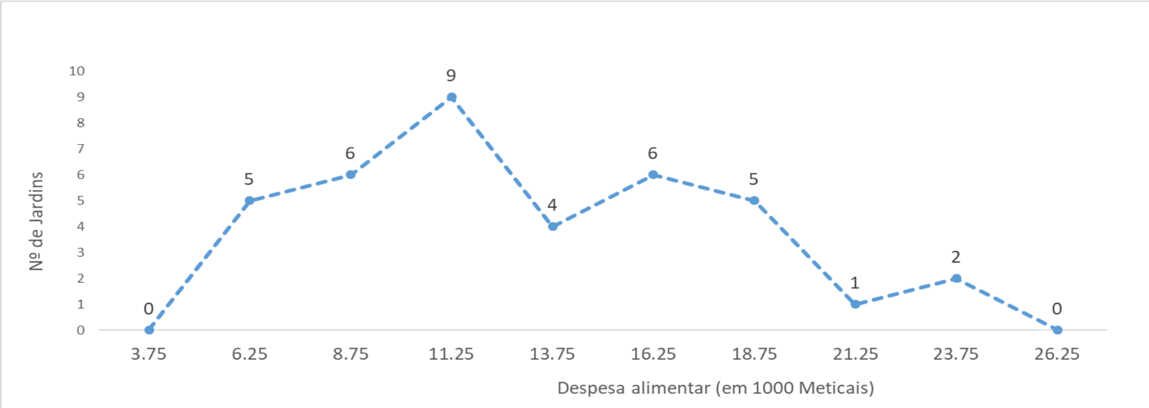
Interpretação: 25% e 75% de fabricantes de computadores solucionam chamados técnicos no máximo até 11.25 e 22.5 dias respectivamente.

f) Com base nos cinco valores calculados construa o gráfico *boxplot* para o número de dias necessários para que o problema fosse resolvido.

Resolução



16. Os gastos diários em compra de produtos alimentares em 1000 Meticais nos jardins infantis do distrito de Marracuene, estão apresentados no gráfico abaixo.



a) Identifique e classifique a variável em estudo.

Resposta: Trata-se da despesa alimentar em 1000 Meticais. É uma variável quantitativa discreta.

b) Organize os dados em tabela de frequências (classe, ponto médio, frequências absoluta e relativa simples e acumulada).

Resolução

Como trata-se do histograma, a tabela de frequência será:

Classes	x_i	f_i	F_i	f_r	F_r
5----7.5	6.25	5	5	13.2%	13.2%
7.5----10	8.75	6	11	15.8%	28.9%
10----12.5	11.25	9	20	23.7%	52.6%
12.5----15	13.75	4	24	10.5%	63.2%
15----17.5	16.25	6	30	15.8%	78.9%
17.5----20	18.75	5	35	13.2%	92.1%
20----22.5	21.25	1	36	2.6%	94.7%
22.5----25	23.75	2	38	5.3%	100.0%
Total	-	38	-	100.0%	-

c) Qual é a percentagem dos jardins infantis que gastam pelo menos 10 mil meticais em despesas alimentares por dia.

Resolução

$$fr(x \geq 10) = \frac{38-11}{38} * 100\% = 71.1\%$$

Resposta: A percentagem dos jardins infantis que gastam pelo menos 10 mil meticais em despesas alimentares por dia é de 71.1%.

d) Calcule o valor das despesas alimentares mais observadas.

Resolução

Aqui pede se a moda e como as amplitudes são iguais pode-se recorrer a fórmula de Czuber, então:

$$\hat{x} = lim_{inf} + \frac{f_{i\hat{x}} - f_{\hat{x}i-1}}{2 * f_{i\hat{x}} - (f_{\hat{x}i-1} + f_{\hat{x}i+1})} * a_i$$

Pela tabela de frequências, a classe modal é: 10----12.5

Logo:

$$\hat{x} = 10 + \frac{9-6}{2*9-(6+4)} * 2.5 = 10.9375$$

Interpretação: O valor das despesas alimentares mais observadas é de 10.937,50 Meticais.

e) Calcule a despesa máxima dos 75% dos jardins infantis.

Resolução

Para encontrar as despesas máximas dos 75% dos jardins infantis, deve-se determinar o P_{75}.

Como os dados estão agrupados em classes, pode se usar a seguinte fórmula:

$$P_i = lim_{inf} + \frac{\frac{in}{100} - F_{i-1}}{f_i} * a_i$$

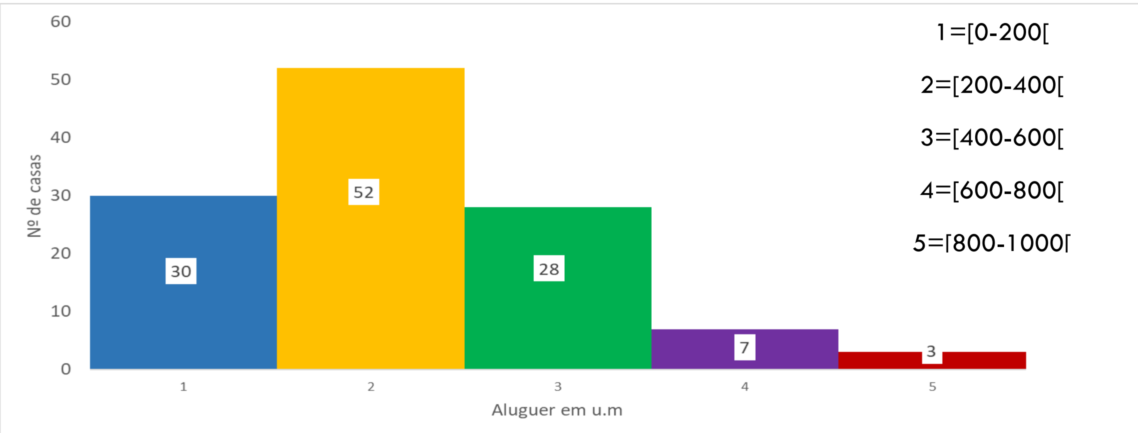
Pela tabela de frequências, a classe P_{75} é: 15----17.5

Então,

$$P_{75} = 15 + \frac{\frac{75*38}{100} - 24}{6} * 2.5 = 16.875$$

Interpretação: A despesa máxima dos 75% dos jardins infantis é de 16.875,00 Meticais.

17. Uma imobiliária gere o aluguer em u.m de residências particulares, segundo o gráfico abaixo:



a) Identifique e classifique a variável em estudo;

Resposta: Trata-se do preço de aluguer de cadas em unidades monetárias. É uma variável quantitativa discreta.

b) Organize os dados em tabela de frequências (classe, ponto médio, frequências absoluta e relativa simples e acumulada).

Resolução

Como trata-se do histograma, a tabela de frequência será:

Classes	x_i	f_i	F_i	f_r	F_r
0----200	100	30	30	25%	25%
200----400	300	52	82	43.3%	68.3%
400----600	500	28	110	23.3%	91.6%
600----800	700	7	117	5.8%	97.4%
800----1000	900	3	120	2.5%	100%
Total	-	120	-	100.0%	-

c) Qual é a percentagem de residências com um aluguer superior a 500 u.m?

Resolução

$$fr(x > 500) = \frac{7+3}{120} * 100\% = 8.3\%$$

Resposta: A de residências com um aluguer superior a 500 u.m é de 8.31%.

d) Quanto se esperaria receber por mês se todas as unidades estivessem ocupadas?

Resolução

$$\sum x_i f_i = 40200$$

Resposta: Se todas as unidades estivessem ocupadas eperaria receber por mês cerca de 40200,00 u.m.

e) Calcule o aluguer médio para estas residências.

Resolução

Pela fórmula para dados agrupados em classes temos:

$$\bar{x} = \frac{\sum x_i f_i}{n} = \frac{40200}{120} = 335$$

Resposta: O aluguer médio para estas residências é de 335,00 u.m

f) Qual é o preço mínimo dos 10% de casas mais caras?

Resolução

Para encontrar o o preço mínimo dos 10% de casas mais caras, deve-se determinar o P_{90}.

Como os dados estão agrupados em classes, pode se usar a seguinte fórmula:

$$P_i = lim_{inf} + \frac{\frac{in}{100} - F_{i-1}}{f_i} * a_i$$

Pela tabela de frequências, a classe P_{90} é: 400----600

$$Então,$$

$$P_{90} = 400 + \frac{\frac{90 * 120}{100} - 82}{28} * 200 = 585.71$$

Interpretação: O preço mínimo dos 10% de casas mais caras é de 585,71 u.m.

18. Foram avaliadas, por cirurgiões-dentistas com especialização em Ortodontia, crianças no estágio de dentadura decídua, entre 3 e 6 anos de idade. Não tinham hábitos de sucção 615. Das demais, 190 tinham o hábito de sucção do polegar, 588 usavam chupeta, 618 usavam mamadeira.

a) Indique e classifique a variável em estudo;

Resposta:

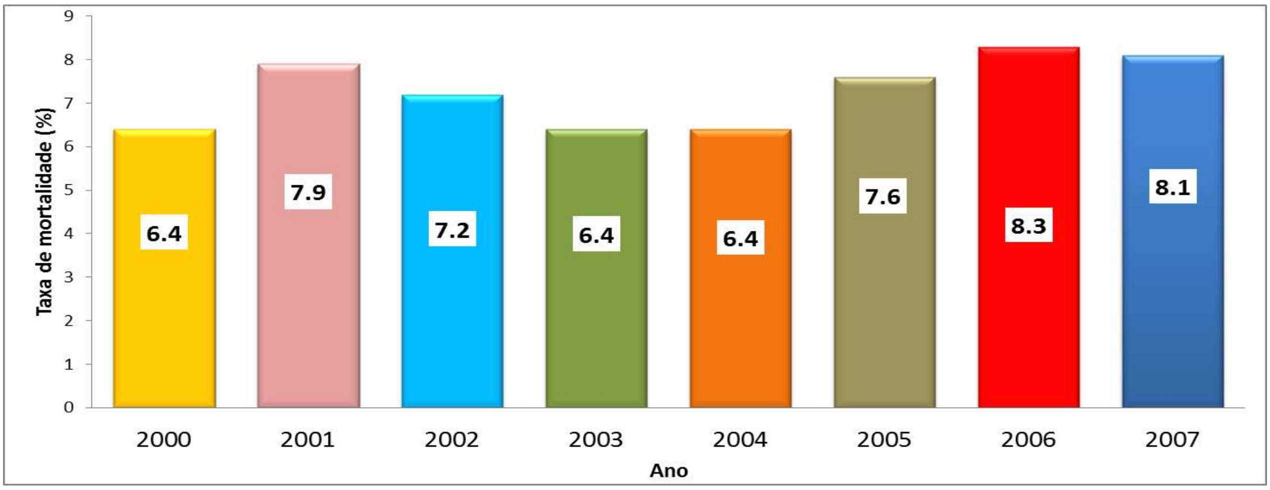
A variável em estudo é o tipo de hábito de sucção e quanto a classificação trata-se de uma variável qualitativa nominal.

b) Apresente os dados em tabela frequência (frequência absoluta simples e relativa simples).

Resolução

Tipo de hábito de sucção	f_i	f_r
Sem hábitos de sucção	615	31%
Sucção do polegar	190	9%
Sucção Chupeta	588	29%
Sucção Mamadeira	618	31%
Total	2011	100%

19. O gráfico a seguir representa as taxas de mortalidade em % de ano 2000 a 2007 duma certa unidade sanitária.



a) Que tipo de gráfico se trata?

Resposta:

Trata se de gráfico de barras

b) Determine a taxa média de mortalidade;

Resolução

$$\bar{x} = \frac{\sum xi}{n} = \frac{6.4 + 7.9 + 7.2 + 6.4 + 6.4 + 7.6 + 8.3 + 8.1}{8} = 7.3$$

Resposta: A taxa média de mortalidade é de 7.3%.

c) Calcule a taxa mediana de mortalidade;

Resolução

Organização de dados em Rol:

6.4; 6.4; 6.4; 7.2; 7.6; 7.9; 8.1; 8.3

Logo o n é par

$$n = 2k \quad \rightarrow \quad k = \frac{n}{2} = 4$$

Então:

$$\tilde{x} = \frac{x_k + x_{k+1}}{2} = \frac{x_4 + x_5}{2} = \frac{7.2 + 7.6}{2} = 7.4$$

Resposta: A taxa mediana de mortalidade é de 7.4%.

d) Determine a taxa modal de mortalidade.

Resposta:

A taxa modal de mortalidade é de 6.4%.

20. A média aritmética das idades de 15 pacientes internados na enfermaria de medicina no mês de Janeiro dum hospital era de 40 anos. Um paciente perdeu a vida e a média das idades passou a ser 41 anos. Qual era a idade desse paciente?

Resolução

Pela fórmula da média para dados não agrupados temos:

$$\bar{x} = \frac{\sum x_i}{n}$$

Determinando o total de idades dos 15 pacientes temos:

$$\sum x_i = \bar{x} * n = 15 * 40 = 600$$

Logo com a morte dum paciente o número de pacientes reduz para 14, assim:

$$\bar{y} = \frac{\sum y_i}{n - 1}$$

$$\sum y_i = \sum x_i - k$$

$$41 * 14 = 600 - k$$

$$k = 26$$

Resposta: *A idade desse paciente era de 26 anos.*

21. A tabela a seguir apresenta a taxa de desemprego em % da população economicamente activa no período de 1989 a 1997:

Ano	1989	1990	1991	1992	1993	1994	1995	1996	1997
Taxa de desemprego (%)	2,30	3,90	4,10	4,50	4,40	3,40	4,40	3,80	4,80

a) Qual a moda da variável?

Resolução

$$\hat{X} = 14.4$$

Resposta: A taxa modal de desemprego da população economicamente activa no período de 1989 a 1997 é de 4.4%.

b) Determine e interprete a média.

Resolução

Recorrendo a fórmula para cálculo da média temos:

$$\bar{X} = \frac{\sum x_i f_i}{n} = \frac{2.3 + 3.9 + 4.1 + 4.5 + 2 * 4.4 + 3.4 + 3.8 + 4.8}{8} = 4$$

Resposta: A taxa média de desemprego da população economicamente activa no período de 1989 a 1997é de 4%.

c) Determine e interprete a mediana.

1º ROll:

Ano	1	2	3	4	5	6	7	8	9
Taxa de desemprego	2,3	3,4	3,8	3,9	4,1	4,4	4,4	4,5	4,8

$$K = 2m + 1 \qquad m = 4$$

$$Logo, \tilde{x} = x_{m+1} = x_5 = 4.1$$

Interpretação: 50% da população economicamente activa no período de 1989 a 1997 atingiu uma taxa de até 4.1%.

d) Classifique a distribuição quanto a assimetria e achatamento.

Resolução

Recorrendo a fórmula dos momentos teremos:

Taxa	di	M2	M3	M4
2,3	-1,7	2,89	-4,913	8,3521
3,4	-0,6	0,36	-0,216	0,1296
3,8	-0,2	0,04	-0,008	0,0016
3,9	-0,1	0,01	-0,001	0,0001
4,1	0,1	0,01	0,001	1E-04
4,4	0,4	0,16	0,064	0,0256
4,4	0,4	0,16	0,064	0,0256
4,5	0,5	0,25	0,125	0,0625
4,8	0,8	0,64	0,512	0,4096
		0,502222	-0,48578	1,000756

Lembre- se que:

$$d_i = x_i - \bar{x}$$

$$S = \sqrt{\frac{(x_i - \bar{x})^2}{\sum f_i}} = 0.502$$

$$m_3 = \frac{(x_i - \bar{x})^3}{\sum f_i} = -0.48578$$

$$m_4 = \frac{(x_i - \bar{x})^4}{\sum f_i} = 1.000756$$

Então:

$$C_a = \frac{m_3}{S^3} = \frac{-0.48578}{0.502^3} = -1.432$$

$$K = \frac{m_4}{S^4} = \frac{1.000756}{0.502^4} = 2.433$$

Logo a distribuição quanto a assimetria é assimétrica negativa; e quanto ao achatamento é Platicúrtica.

22. No quadro seguinte apresentam-se o número de transacções efectuadas em cada uma das lojas dos Supermercados XXX em U.M, classificadas por níveis de despesa, e o número de empregados existentes em cada uma delas.

Escalão das despesas	[0-----10[	[10-----20[	[20-----30[	[40-----50[	Nº de Empregados
Loja A	74	78	30	18	30
Loja B	29	44	26	9	20

a) Será possível afirmar que "em ambas as lojas, mais de 70% das transacções têm um valor inferior a 20 U.M."?

Resolução

Recorrendo a tabela de frequência teremos:

	Loja A			Loja B		
Despesas	fi	fr	fr(ac)	fi	fr	fr(ac)
[0----10[	74	0,370	0,370	29	0,269	0,269
[10----20[	78	0,390	0,760	44	0,407	0,676
[20----30[	30	0,150	0,910	26	0,241	0,917
[30----40[	18	0,090	1,000	9	0,083	1,000
Somatório	200	1		108	1,000	

Resposta: A análise das frequências relativas acumuladas mostra que só há 67,6% de transacções abaixo de 20 u.m. na loja B: A afirmação não é verdadeira.

b) Determine o valor médio por transacção e o valor médio das transacções por empregado, em cada uma das lojas.

Despesas	Loja A			Loja B		
	xi	fi	xi*fi	xi	fi	xi*fi
[0----10[	5	74	370	5	29	145
[10----20[	15	78	1170	15	44	660
[20----30[	25	30	750	25	26	650
[30----40[	30	18	540	30	9	270
Somatório		200	2830		108	1725

Loja A (Valor médio por transacção):

$$\bar{X} = \frac{\sum x_i f_i}{n} = \frac{2830}{200} = 14.2$$

Loja B (Valor médio por transacção):

$$\bar{X} = \frac{\sum x_i f_i}{n} = \frac{1725}{108} = 16$$

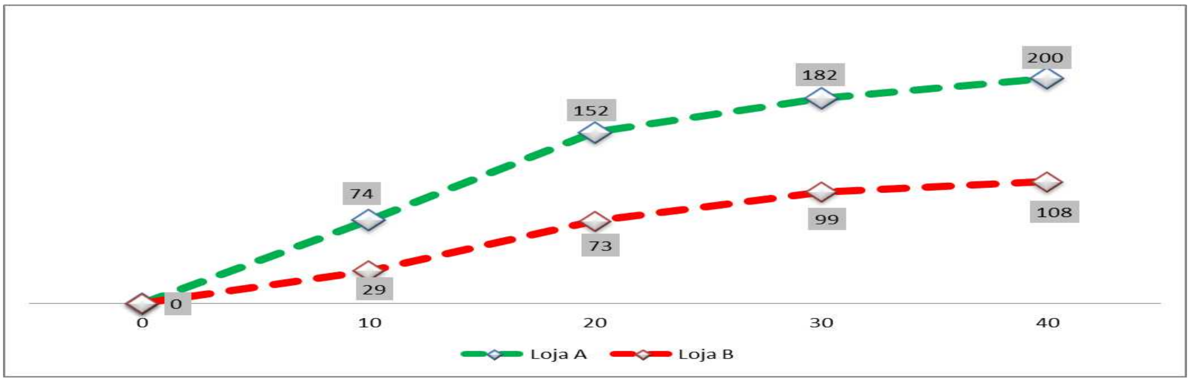
Loja A (Valor médio das transacções por empregado):

$$\bar{X} = \frac{\sum x_i f_i}{n} = \frac{2830}{30} = 94.3$$

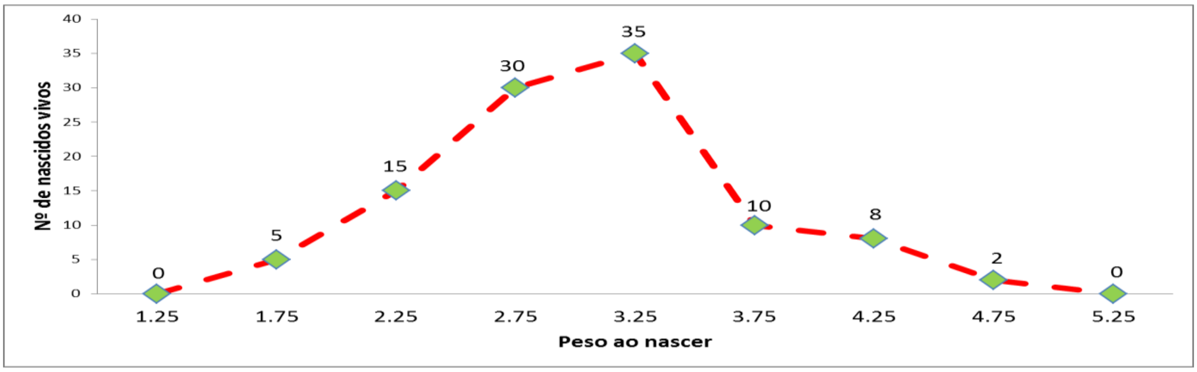
Loja B (Valor médio das transacções por empregado):

$$\bar{X} = \frac{\sum x_i f_i}{n} = \frac{1725}{20} = 86.3$$

c) Represente graficamente a ogiva de cada uma das distribuições.



23. O gráfico a seguir representa o peso ao nascer de nascidos vivos, em quilogramas no ano 2016 dum centro de saúde em Moçambique.



a) Que tipo de gráfico se trata?

Resposta:

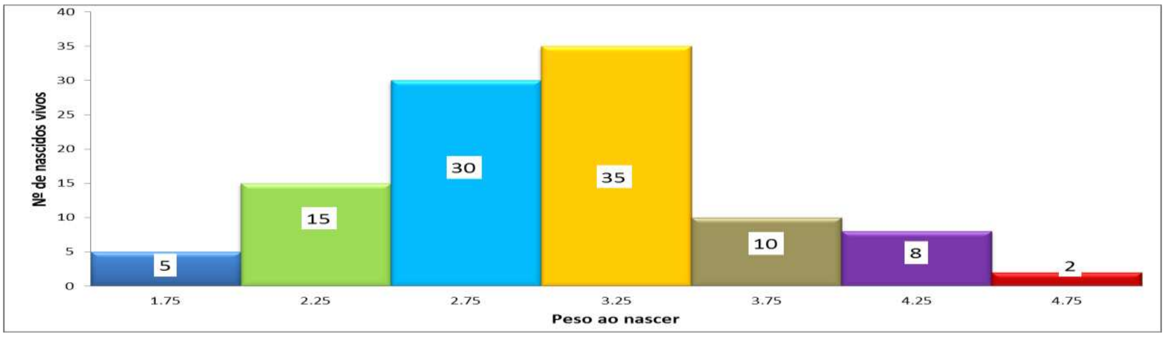
Trata-se de polígono de frequências

b) Represente os dados em tabela de frequência.

Resolução

Classe	x_i	f_i
[1.5----2[	1.75	5
[2----2.5[	2.25	15
[2.5----3[	2.75	30
[3----3.5[	3.25	35
[3.5----4[	3.75	10
[4----4.5[	4.25	8
[4.5----5[	4.75	2
Total		105

c) Esboce o histograma correspondente.



d) Classifique a distribuição quanto a assimetria.

Resolução

Pela fórmula das separatrizes temos:

$$C_a = \frac{Q_1 + Q_3 - 2 * Q_2}{Q_3 - Q_1}$$

$$P_i = lim_{inf} + \frac{\frac{in}{100} - F_{i-1}}{f_i} * a_i$$

Classe	x_i	f_i	F_i
[1.5----2[	1.75	5	5
[2----2.5[	2.25	15	20
[2.5----3[	2.75	30	50
[3----3.5[	3.25	35	85
[3.5----4[	3.75	10	95
[4----4.5[	4.25	8	103
[4.5----5[	4.75	2	105
Total		105	

$$Q_1 : \frac{25}{100} * 105 = 26.25;\ Q_2 : \frac{50}{100} * 105 = 52.5;\ Q_3 : \frac{75}{100} * 105 = 78.75$$

$$Q_1 = 2.5 + \frac{26.25 - 20}{30} * 0.5 = 2.6$$

$$Q_2 = 3 + \frac{52.5 - 50}{35} * 0.5 = 3.04$$

$$Q_3 = 3 + \frac{78.75 - 50}{35} * 0.5 = 3.4$$

Logo:

$$C_a = \frac{2.6 + 3.4 - 2 * 3.04}{3.4 - 2.6} = -0.1$$

Conclusão: A distribuição das idades quanto a assimetria é assimétrica negativa

$$(C_a < 0)$$

e) Tire as conclusões em relação a achatamento.

Resolução

Pela fórmula dos separatrizes temos:

$$K = \frac{P_{75} - P_{25}}{2 * (P_{90} - P_{10})}$$

$$P_{90} : \frac{90}{100} * 105 = 94.5;\ P_{10} : \frac{10}{100} * 105 = 10.5$$

Classe	x_i	f_i	F_i
[1.5----2[	1.75	5	5
[2----2.5[	2.25	15	20
[2.5----3[	2.75	30	50
[3----3.5[	3.25	35	85
[3.5----4[	3.75	10	95
[4----4.5[	4.25	8	103
[4.5----5[	4.75	2	105
Total		105	

$$P_{10} = 2 + \frac{10.5 - 5}{15} * 0.5 = 2.183$$

$$P_{90} = 3.5 + \frac{94.5 - 85}{10} * 0.5 = 3.975$$

$$K = \frac{3.4 - 2.6}{2 * (3.975 - 2.183)} = 0.223$$

Conclusão: A distribuição das idades quanto a achatamento é leptocúrtica

$$(K < 0.263)$$

24. Considere os dados a seguir representando os resultados de 50 exames de sangue, referentes á fracção de colesterol de muito baixa densidade (Very Low Density Lipo-protein, VLDL) em miligramas por decilitro (mg/dl), em indivíduos do sexo feminino:

    30 35 32 28 25 26 28 30 35 40 26 27 45 36 30 30 26 34 28 29 22 30 28 36 30 28 35 40 39 29 30 28 34 39 26 28 30 34 35 34 28 29 34 35 37 48 30 22 26 30

    a) Determinar a amplitude total dos dados.

Resolução:

$$At = x_{max} - x_{min} = 48 - 22 = 26$$

Resposta: A amplitude total é de 26mg/dl.

    b) Escolher um número de classes conveniente e construir o histograma.

Resolução:

Usando a regra de Sturges para dados agrupados em intervalos de classe teremos:

$$k = int[1 + 3.22 * log_{10} n] = int[1 + 3.22 * log_{10} 50] \approx 7$$

Logo

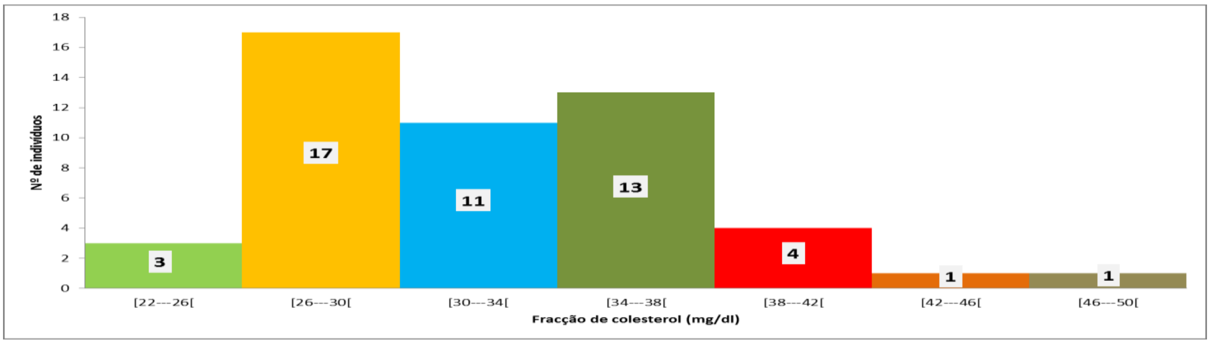
$$a_i = \frac{x_{max} - x_{min}}{k} = \frac{26}{7} \approx 4$$

Então:

A tabela de frequências será:

Fracção de colesterol	x_i	f_i
[22----26[	24	3
[26----30[	28	17
[30----34[	32	11
[34----38[	36	13
[38----42[	40	4
[42----46[	44	1
[46----50[	48	1
Total		50

O histograma correspondente é:



    c) Calcule as medidas de tendência central e compare.

Resolução

Média

Recorrendo a tabela teremos:

Fracção de colesterol	x_i	f_i	$x_i * f_i$
[22----26[	24	3	72
[26----30[	28	17	476
[30----34[	32	11	352
[34----38[	36	13	468
[38----42[	40	4	160
[42----46[	44	1	44
[46----50[	48	1	48
Total		50	1620

$$\bar{x} = \frac{\sum x_i * f_i}{n} = \frac{1620}{50} = 32.4$$

Mediana

Recorrendo a tabela das frequências acumuladas temos:

Fracção de colesterol	x_i	f_i	F_i
[22----26[	24	3	3
[26----30[	28	17	20
[30----34[	32	11	31
[34----38[	36	13	44
[38----42[	40	4	48
[42----46[	44	1	49
[46----50[	48	1	50
Total		50	

$$\tilde{x} = lim_{inf} + \frac{\frac{1}{2} * n - F_{i-1}}{f_i} * a_i$$

$$\tilde{x} = \frac{1}{2} * n = \frac{1}{2} * 50 = 25$$

pelas frequências acumuladas a classe mediana será [30---34[, Logo:

$$\tilde{x} = 30 + \frac{25 - 20}{11} * 4 = 31.8$$

Moda

Pela fórmula de Czuber temos:

$$\hat{x} = lim_{inf} + \frac{f_{i\hat{x}} - f_{\hat{x}i-1}}{2 * f_{i\hat{x}} - (f_{\hat{x}i-1} + f_{\hat{x}i+1})} * a_i$$

1° Identificação da classe modal:

A maior frequência absoluta simples é 17 e pertence a classe [26---30 [, logo:

$$\hat{x} = 26 + \frac{17 - 3}{2 * 17 - (3 + 1)} * 4 = 28.8$$

Resposta:

- ✓ *$\bar{x}$= 32.4: A fracção média de colesterol de muito baixa densidade (Very Low Density Lipo-protein, VLDL) em miligramas por decilitro (mg/dl), em indivíduos do sexo feminino é de 32.4.*
- ✓ *50% dos indivíduos do sexo feminino contem uma fracção de colesterol de muito baixa densidade (Very Low Density Lipo-protein, VLDL) em miligramas por decilitro (mg/dl) de até 31.8.*
- ✓ *A maior parte dos indivíduos do sexo feminino contem uma fracção de colesterol de muito baixa densidade (Very Low Density Lipo-protein, VLDL) em miligramas por decilitro (mg/dl) de 28.8*

25. A anemia por deficiência de ferro é a forma mais comum de subnutrição nos países em desenvolvimento, afectando cerca de 50% das crianças e mulheres e 25% dos homens. Em um estudo planeado para se investigar esse assunto, comparou-se o conteúdo de ferro de determinados alimentos da Etiópia. Foi analisado o conteúdo de ferro na carne de vaca preparada com diversos temperos locais, nos legumes cozidos e em um assado de verduras condimentado e mediu-se a quantidade de ferro na comida, em miligramas de ferro por 100 gramas de alimento preparado.

Conteúdo de ferro/100g para cada tipo de alimento

Carne	1.84	1.96	1.97	2.06	2.14	2.17	2.22	2.27	2.28	2.36
Legumes	2.17	2.32	2.34	2.40	2.41	2.41	2.43	2.43	2.48	2.54
Verduras	0.76	0.92	0.99	1.03	1.07	1.30	1.45	1.53	1.55	1.68

Qual dos alimentos apresenta mais heterogeneidade na quantidade de ferro?

Resolução:

Usando o coeficiente de variação teremos:

$$CV = \frac{S}{\bar{x}} * 100\%$$

Carne

$$\bar{x} = \frac{\sum x_i * f_i}{n} = \frac{1.84 + 1.96 + 1.97 + 2.06 + \cdots + 2.36}{10} = 2.127$$

$$S = \sqrt{\frac{\sum(x_i - \bar{x})^2}{\sum f_i - 1}} = \sqrt{\frac{(1.84 - 2.127)^2 + (1.96 - 2.127)^2 + \cdots + (2.36 - 2.127)^2}{9}} = 0.027$$

$$CV = \frac{S}{\bar{x}} * 100\% = \frac{0.027}{2.127} * 100\% = 1.3\%$$

Legumes

$$\bar{x} = \frac{\sum x_i * f_i}{n} = \frac{2.17 + 2.32 + 2.34 + 2.40 + \cdots + 2.54}{10} = 2.393$$

$$S=\sqrt{\frac{\sum(x_i-\bar{x})^2}{\sum f_i-1}}=\sqrt{\frac{(2.17-2.393)^2+(2.32-2.393)^2+\cdots+(2.54-2.393)^2}{9}}=0.01$$

$$CV=\frac{s}{\bar{x}}*100\%=\frac{0.01}{2.393}*100\%=0.4\%$$

Verduras

$$\bar{x}=\frac{\sum x_i*f_i}{n}=\frac{0.76+0.92+0.99+1.03+\cdots+1.68}{10}=1.231$$

$$S=\sqrt{\frac{\sum(x_i-\bar{x})^2}{\sum f_i-1}}=\sqrt{\frac{(0.76-1.231)^2+(0.92-1.231)^2+\cdots+(1.68-1.231)^2}{9}}=0.096$$

$$CV=\frac{s}{\bar{x}}*100\%=\frac{0.096}{1.231}*100\%=7.8\%$$

Resposta: O alimento que apresenta mais heterogeneidade na quantidade de ferro são as verduras.

26. A média aritmética dos salários de quatro funcionários de uma empresa era 1.300,00 USD. Ao ser contratado um novo funcionário, a média passou a ser 1.340,00 USD. Qual o salário desse funcionário?

Resolução

Recorrendo a fórmula para cálculo da média temos:

$$\bar{X}=\frac{\sum x_i f_i}{n}$$

$$\sum x_i f_i=1300*4=5200$$

Sabendo que com a entrada dum novo funcionário a média passou a ser 1340, então:

$$1340=\frac{5200+x}{5}$$

$$x=5*1340-5200$$

$$x=1500$$

Logo: O salário desse funcionário é de 1500,00USD.

27. Considere a seguinte distribuição de frequência correspondente aos diferentes preços em u.m de um determinado produto em vinte e duas lojas pesquisadas.

Preços	50.5	51	52	54.5	56	Total
Nº de lojas	2	5	8	6	1	22

a) Qual é o percentual das lojas que apresentam um preço de até 52,00 u.m?

Resolução

$$fr\ (x\leq 52)=\frac{2+5+8}{22}*100\%=68.18\%$$

Resposta: As lojas que apresentam um preço menor ou igual a 52 u.m perfazem 68.18%

b) Qual é o preço médio do produto em referência?

Resolução

$$\bar{x} = \frac{\sum x_i * f_i}{\sum f_i} = \frac{50.5 * 2 + 51 * 5 + 52 * 8 + 54.5 * 6 + 56 * 1}{22} = \frac{1155}{22} = 52.5$$

Resposta: o preço médio do produto em referência é de 52.2 u.m.

c) Calcule a o desvio padrão e classifique a variabilidade.

Resolução

$$s^2 = \frac{\sum(x_{i-}\bar{x})^2 * f_i}{n-1} = \frac{(50.5 - 52.5) + (51 - 52.5) + (52 - 52.5) + (54.5 - 52.5) + (56 - 52.5)}{21} = 1.05$$

O desvio padrão é de 1.05 unidades monetárias

$$CV = \frac{s}{\bar{x}} * 100\% = \frac{1.05}{52.5} * 100\% = 2\%$$

Resposta: a distribuição possui uma variabilidade muito baixa.

TEORIA DE PROBABILIDADES

28. Dois equipamentos, A e B, para processamento de dosagens bioquímicas são colocados para teste de controlo de qualidade por 120 horas. A probabilidade de que um erro de cálculo aconteça em um equipamento do tipo A é de 1/30, no tipo B, 1/80 e em ambos, 1/1000. Qual a probabilidade de que:

a) Pelo menos um dos equipamentos tenha apresentado erro?

Resolução

Pelos dados temos:

$$P(A) = \frac{1}{30}; P(B) = \frac{1}{80}; P(A \cap B) = \frac{1}{1000}$$

Então:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{30} + \frac{1}{80} - \frac{1}{1000} = 0.0448$$

Resolução: A probabilidade de pelo menos um dos equipamentos tenha apresentado erro é de 4.5%.

b) Apenas o equipamento A tenha apresentado erro?

Resolução

$$P(A \cap \bar{B}) = P(A) - P(A \cap B) = \frac{1}{30} - \frac{1}{1000} = 0.032$$

Resolução: A probabilidade de apenas o equipamento A tenha apresentado erro é de 3.2%.

c) Verifique se os erros de cálculos dos equipamentos A e B são ou não independentes.

Resolução

Dois eventos são independentes se:

- $P(A \backslash B) = P(A)$
- $P(B \backslash A) = P(B)$
- $P(A \cap B) = P(A) * P(B)$

Então:

$$P(A \backslash B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{1000}}{\frac{1}{80}} = 0.08 \neq \frac{1}{30}$$

Logo os erros de cálculos dos equipamentos A e B não são independentes.

29. Um empreiteiro apresentou orçamentos separados para a execução da parte eléctrica e da parte de encanamento de um edifício. Ele acha que a probabilidade de ganhar a concorrência da parte eléctrica é de 50%. Caso ele ganhe a parte eléctrica, a chance de ganhar a parte de encanamento é de 3/4, caso contrário esta probabilidade é de 1/3.

    a) Qual a probabilidade do empreiteiro ganhar os dois contractos?

Resolução

Sejam definidos os seguintes eventos:

$$EL = O\ empreiteiro\ ganhar\ a\ concorrência\ da\ parte\ eléctrica$$

$$EN =\ O\ empreiteiro\ ganhar\ a\ concorrência\ da\ parte\ de\ encanamento$$

Pelos dados temos:

$$P(EL) = 1/2; P(EN\backslash EL) = 3/4; P(EN\backslash \overline{EL}) = 1/3$$

Então:

$$P(EL \cap EN) = P(EL) * P(EN\backslash EL) = \frac{1}{2} * \frac{3}{4} = \frac{3}{8}$$

Resposta: A probabilidade do empreiteiro ganhar os dois contractos é de 37.5%.

    b) Qual a probabilidade do empreiteiro ganhar apenas um?

Resolução

$$P(EL \cap \overline{EN}) + P(\overline{EL} \cap EN) = \frac{1}{2} * \frac{1}{4} + \frac{1}{2} * \frac{1}{3} = 0.292$$

Resposta: A probabilidade do empreiteiro ganhar apenas um contrato é de 29.2%.

    c) Qual a probabilidade do empreiteiro perder a parte eléctrica e perder a parte de encanamento?

Resolução

$$P(\overline{EL} \cap \overline{EN}) = \frac{1}{2} * \frac{2}{3} = \frac{1}{3}$$

Resposta: A probabilidade do empreiteiro perder a parte eléctrica e perder a parte de encanamento é de 33.3%

30. Para seleccionar seus funcionários, uma empresa oferece aos candidatos um curso de treinamento durante uma semana. No final do curso, eles são submetidos a uma prova e 25% são classificados como bons (B), 50% como médios (M) e os restantes 25% como fracos (F). Para facilitar a selecção, a empresa pretende substituir o treinamento por um teste contendo questões referentes a conhecimentos gerais e específicos. Para isso, gostaria de conhecer qual a probabilidade de um indivíduo aprovado no teste ser considerado fraco, caso fizesse o curso. Assim, neste ano, antes do início do curso, os candidatos foram submetidos ao teste e receberam o conceito aprovado (A) ou reprovado (R). No final do curso 80% dos funcionários bons foram aprovados, metade dos médios

reprovados e 20% dos fracos aprovados. Um funcionário acaba de ser aprovado, qual é a probabilidade de ser fraco?

Resolução

Pelos dados temos:

$$P(B) = 0.25; P(M) = 0.50; P(F) = 0.25; P(A\backslash B) = 0.80; P(R\backslash M) = 0.50; P(A\backslash F) = 0.20$$

Recorrendo ao teorema de Bayes temos:

$$P(F\backslash A) = \frac{P(F \cap A)}{P(A)} = \frac{0.25 * 0.20}{0.25 * 0.80 + 0.50 * 0.50 + 0.25 * 0.20} = 0.10$$

Resposta: A probabilidade de o funcionário ser fraco dado que acaba de ser aprovado é de 10%.

31. Um software que detecta fraudes em cartões telefónicos detecta o número de áreas metropolitanas onde as chamadas são originadas a cada dia. São obtidos os seguintes dados:
    - 1.6% dos usuários legítimos chamam de duas ou mais áreas metropolitanas em um mesmo dia.
    - 25% dos usuários fraudulentos chamam de duas ou mais áreas metropolitanas em um mesmo dia.
    - A proporção de usuários fraudulentos é de 0.84%.

    Se um mesmo usuário faz chamadas de duas ou mais áreas metropolitanas em um mesmo dia, qual é a probabilidade de que o usuário seja fraudulento?

Resolução

Sejam definidos os seguintes eventos:

$$A = Usuários\ \ chamam\ de\ duas\ ou\ mais\ áreas\ metropolitan$$

$$F = Usuários\ fraudulentos$$

Pelos dados temos:

$$P(F) = 0.84\%; P(A\backslash \bar{F}) = 1.6\%; P(A\backslash F) = 0.25\%$$

Então:

$$P(F\backslash A) = \frac{P(F \cap A)}{P(A)} = \frac{0.0084 * 0.25}{0.0084 * 0.25 + 0.9916 * 0.016} = 0.11689$$

Resposta: A probabilidade de que o usuário seja fraudulento dado que faz chamadas de duas ou mais áreas metropolitanas em um mesmo dia é de 11.7%.

32. No design preliminar de produtos são utilizadas avaliações de clientes. No passado, 95% dos produtos de alto sucesso receberam boas avaliações, 60% dos produtos de sucesso moderado receberam boas avaliações, e 10% dos produtos de pobre desempenho receberam boas avaliações. Além disso, 40% dos produtos tiveram alto sucesso, 35% tiveram sucesso moderado e

25% tiveram desempenho pobre. Qual é a probabilidade de que o produto consiga uma boa avaliação?

Resolução

Sejam definidos os seguintes eventos:

$$A = Produtos\ de\ alto\ sucesso$$

$$M = Produtos\ de\ sucesso\ moderado$$

$$P = Produtos\ de\ desempenho\ pobre$$

$$B = Produtos\ receberam\ uma\ boa\ avaliação$$

Pelos dados temos:

$$P(A) = 0.40; P(M) = 0.35; P(P) = 0.25; P(B\backslash A) = 0.95; P(B\backslash M) = 0.60; P(B\backslash P) = 0.10$$

Então:

$$P(B) = P(A \cap B) + P(M \cap B) + P(P \cap B) = 0.40 * 0.95 + 0.35 * 0.60 + 0.25 * 0.10$$
$$= 0.615$$

Resposta: A probabilidade de que que o produto consiga uma boa avaliação é de 61.5%.

33. Em um escritório de contabilidade, o contador-chefe tem três auxiliares, um que trabalha em tempo integral e os outros dois que trabalham em tempo parcial. O funcionário de tempo integral é responsável por 50% dos balancetes, enquanto cada um dos funcionários de tempo parcial responde pela metade dos balancetes restantes. Nos últimos 2 meses, a proporção de balancetes com erros oriundos do funcionário de tempo integral foi de 5%, enquanto para os funcionários de tempo parcial essas proporções foram de 6% e 8%. O chefe resolve, então, fazer um novo treinamento, discutindo os principais erros encontrados. No mês seguinte ao treinamento, a proporção de balancetes com erro cai pela metade, com cada funcionário de tempo parcial produzindo a mesma proporção de balancetes com erro, igual á metade da proporção de erros do funcionário de tempo integral. Quais são as novas proporções de balancetes com erro de cada funcionário?

Resolução

Consideremos os eventos

$$I\ =\ auxiliar\ a\ tempo\ integral$$

$$P_1\ =\ auxiliar\ 1\ a\ tempo\ parcial$$

$$P_2\ =\ auxiliar\ 2\ a\ tempo\ parcial$$

$$E = balancetes\ com\ erros$$

Pelo problema temos:

$$P(I) = 0.50;\ P(P_1) = 0{,}25;\ P(P_2) = 0{,}25$$

$$P(E/I) = 0.05; P(E/P_1) = P(E/P_2) = 0.25$$

Temos dois casos (antes e depois de treinamento)

Pelo teorema de probabilidade total pode se determinar a probabilidade dos balancetes com erros ***antes de treinamento***, *assim teremos:*

$$P(E) = P(I) * P(E/I) + P(P_1) * \ P(E/P_1) + \ P(P_2) * P(E/P_2)$$

$$P(E) = 0.50 * 0.05 + 0.25 * 0.06 + 0.25 * 0.08$$

$$P(E) = 0.06$$

Após o treinamento

a proporção dos erros caiu pela metade, isto é passou a ser $\frac{1}{2} * P(E) = \frac{0.06}{2} = 0.03$, *e*

$$P(E/I) = 1/2 * \{P(E/P_1) = P(E/P_2)\}\,; P(E/P_1) = P(E/P_2)$$

Então:

$$\frac{1}{2} * P(E) = \ 0.50 * P(E/I) + 0.25 * 2 * P(E/I) + \ 0.25 * 2 * P(E/I)$$

$$0.03 = 1.5 * P(E/I)$$

$$Logo:$$

$$P(E/I) = 0.02; P(E/P_1) = 0.04; \ P(E/P_2) = 0.04$$

Resposta: Após o treinamento, as novas proporções de balancetes com erro de cada funcionário passou a ser de 2%, 4% e 4% respectivamente.

34. Um empresário é conhecido por ser muito cauteloso com relação a suas informações. Ele tem um registro minucioso da composição de cada área de sua empresa e sabe que:

Área/Departamento	Executivo Sénior	Executivo Pleno	Executivo júnior
Financeiro	2	3	4
Advocacia	3	2	2
Contabilidade	4	1	1

Sabendo que ele realiza uma visita surpresa dum dos departamentos, escolhendo aleatoriamente um deles e consegue identificar de modo imediato um executivo sénior. Qual é a probabilidade de ser do departamento de contabilidade? Assuma que os três departamentos são igualmente prováveis de serem visitados (as portas das salas são idênticas e equiprováveis).

Resolução

Seja os seguintes eventos:

C=*departamento de contabilidade;*

A=departamento de advocacia

F=departamento financeiro

P=Executivo pleno

J=Executivo Júnior

S=Executivo sénior

Pede-se P(C/J):

Recorrendo a teorema de Bayes temos:

$$P(C/J) = \frac{P(C \cap J)}{P(J)} = \frac{P(C \cap J)}{P(C \cap J) + P(F \cap J) + P(A \cap J)}$$

Pelos dados temos:

$$P(F) = P(A) = P(C) = \frac{1}{3}$$

$$P(J/C) = \frac{1}{6}; P(J/F) = \frac{4}{9}; P(J/A) = \frac{2}{7}$$

$$P(C/J) = \frac{\frac{1}{3} * \frac{1}{6}}{\frac{1}{3} * \frac{1}{6} + \frac{1}{3} * \frac{4}{9} + \frac{1}{3} * \frac{2}{7}} = 0.186$$

Resposta: A probabilidade de ser do departamento de contabilidade, dado que na sua visita identificou um executivo júnior é de 18.6%.

35. Um certo tipo de motor eléctrico falha se ocorrer uma das seguintes situações: emperramento dos mancais, queima dos rolamentos ou desgaste das escovas. Suponha que o emperramento seja duas vezes mais provável do que a queima, esta sendo quatro vezes mais provável do que o desgaste das escovas. Qual é a probabilidade de que a falha seja devida a cada uma dessas circunstâncias?

<u>Resolução</u>

Sejam definidos os seguintes eventos:

$$M = Emperramento\ dos\ mancais$$

$$R = \ Queima\ dos\ rolamentos$$

$$E = Desgaste\ das\ escovas$$

Pelos dados temos:

$$P(M) = 2 * P(R); P(R) = 4 * P(E)$$

E sabe-se que pela definição:

$$P(M) + P(R) + P(E) = 1$$

$$2 * P(R) + P(R) + \frac{1}{4} * P(R) = 1$$

$$P(R) = \frac{4}{13}; P(M) = 2 * \frac{4}{13} = \frac{8}{13}; P(E) = \frac{\frac{4}{13}}{4} = \frac{1}{13}$$

Resposta: A probabilidade de que a falha seja devida a cada uma dessas circunstâncias é de 4/13, 8/13 e 1/13 respectivamente.

36. Um equipamento electrónico é formado por 2 componentes, I e II. Suponha que:

- A chance do componente I falhar é 0,20;
- A chance de apenas o componente II falhar é 0,15 e
- A chance de I e II falharem simultaneamente é 0,15.

a) Calcule a probabilidade de apenas o componente I falhar.

Resolução

$$P(I \cap \overline{II}) = P(I) - P(I \cap II) = 0.20 - 0.15 = 0.05$$

Resposta: A probabilidade de apenas o componente I falhar é de 5%.

b) Calcule a probabilidade do componente I falhar dado que o componente II falhou.

Resolução

$$P(I \backslash II) = \frac{P(I \cap II)}{P(II)} = \frac{0.15}{0.15 + 0.15} = 0.5$$

Resposta: A probabilidade do componente I falhar dado que o componente II falhou é de 50%

37. O Ministério da Economia da Espanha acredita que a probabilidade da inflação ficar abaixo de 3% este ano é de 0,20; entre 3% e 4% é de 0,45 e acima de 4% é de 0,35. O Ministério acredita que, com inflação abaixo de 3%, a probabilidade de se criar mais 200.000 empregos é de 0.6, diminuindo essa probabilidade para 0,3 caso a inflação fique entre 3% e 4%; no entanto, com inflação acima de 4%, isso é totalmente impossível. No ano seguinte, um economista estrangeiro constata que foram criados 200.000 empregos novos. Qual é a probabilidade de a inflação ter ficado entre 3% e 4%.

Resolução

Sejam definidos os seguintes eventos:

$$A = Inflação\ \ abaixo\ de\ 3\%$$

$$B = \ Inflação\ entre\ 3\%\ e\ 4\%$$

$$C = Inflação\ acima\ de\ 4\%$$

$$E = Criar\ mais\ 200.000\ empregos$$

Pelos dados temos:

$$P(A) = 0.20; P(B) = 0.45; P(C) = 0.35;\ P(E \backslash A) = 0.6; P(E \backslash B) = 0.3; P(E \backslash C) = 0$$

Então:

$$P(B \backslash E) = \frac{P(B \cap E)}{P(E)} = \frac{0.45 * 0.3}{0.20 * 0.6 + 0.45 * 0.3 + 0.35 * 0} = 0.529$$

Resposta: A probabilidade de a inflação ter ficado entre 3% e 4% dado que foram criados 200.000 empregos novos é de 52.9%.

38. Um inspector trabalhando para uma companhia de manufactura tem uma probabilidade de 99% de identificar correctamente um item com defeito e 0.5% de probabilidade de classificar incorrectamente um produto bom como defeituoso. A companhia tem evidências de que sua linha produz 0.9% de itens defeituosos.

    a) Qual é a probabilidade de um item seleccionado para inspecção ser classificado como defeituoso?

Resolução

Sejam definidos os seguintes eventos:

$$D = Itens\ defeituosos$$

$$B = \ Itens\ bons$$

$$C = Classificar\ correctamente$$

$$CD = Item\ classificado\ como\ defeituoso$$

Pelos dados temos:

$$P(D) = 0.9\%; P(B) = 90.1\%;\ P(C\backslash D) = 99\%;\ P(\bar{C}\backslash D) = 1\%; P(\bar{C}\backslash B) = 0.9\%;$$

$$P(C\backslash B) = 99.5\%$$

Então:

$$P(CD) = P(B \cap \bar{C}) + P(D \cap C) = 0.901 * 0.005 + 0.009 * 0.99 = 0.0134$$

Resposta: A probabilidade de um item seleccionado para inspecção ser classificado como defeituoso é de 1.34%.

    b) Se um item seleccionado aleatoriamente é classificado como não-defeituoso, qual é a probabilidade dele ser realmente bom?

Resolução

$$P(B\backslash\overline{CD}) = \frac{P(B \cap C)}{P(\overline{CD})} = \frac{0.901 * 0.995}{1 - 0.0134} = 0.90867$$

Resposta: A probabilidade de um item seleccionado para inspecção ser realmente bom dado que foi classificado como não defeituoso é de 1.34%.

39. Um empresário é conhecido por ser muito cauteloso com relação a suas informações. Ele tem um registro minucioso da composição de cada área de sua empresa e sabe que:

Área/Departamento	Executivo Sénior	Executivo Pleno	Executivo júnior
Financeiro	2	3	4
Advocacia	3	2	2
Contabilidade	4	1	1

Ele realiza visita surpresa dum dos departamentos, escolhendo aleatoriamente um deles e consegue identificar de modo imediato dois executivos. Um deles é sénior e o outro, pleno. Assuma que os três departamentos são igualmente prováveis de serem visitados (as portas das salas são idênticas e equiprováveis). Qual é a probabilidade de ser o departamento de contabilidade o visitado?

Resolução

Sejam definidos os seguintes eventos:

$$F = Deprtamento\ Financeiro$$

$$A = Departamento\ de\ Advocacia$$

$$C = Departamento\ de\ Contabilidade$$

$$S = Executivo\ Sénior$$

$$P = Executivo\ Pleno$$

$$S = Executivo\ Júnior$$

Pelos dados temos:

$$P\big(C\backslash(S\cap P)\big) = \frac{P(C\cap S\cap P)}{P(S\cap P)}$$

$$P\big(C\backslash(S\cap P)\big) = \frac{\frac{1}{3}*\frac{4}{6}*\frac{1}{5}+\frac{1}{3}*\frac{1}{6}*\frac{4}{5}}{\frac{1}{3}*\frac{4}{6}*\frac{1}{5}+\frac{1}{3}*\frac{1}{6}*\frac{4}{5}+\frac{1}{3}*\frac{3}{7}*\frac{2}{6}+\frac{1}{3}*\frac{2}{7}*\frac{3}{6}+\frac{1}{3}*\frac{2}{9}*\frac{3}{8}+\frac{1}{3}*\frac{3}{9}*\frac{2}{8}} = 0.371$$

Resposta: A probabilidade de ser o departamento de contabilidade o visitado, dado que foram escolhidos aleatoriamente um executivo pleno e outro sénior é de 37.1%.

40. Em um levantamento dum bairro de 1000 moradores, verifica-se que:

- 220 têm curso superior;
- 160 são casados;
- 100 estão empregados;
- 20 têm curso superior, são casados e estão empregados;
- 60 têm curso superior e estão empregados;
- 90 têm curso superior e são casados

Escolhe-se ao acaso um morador desse bairro. Qual é a probabilidade de que ele tenha apenas curso superior?

Resolução

Seja: $S = o\ morador\ tem\ curso\ superior; C = o\ morador\ é\ casado;$

$$E = o\ morador\ é\ empregado; D = o\ morador\ é\ desempregado$$

Pelos dados temos:

$$n = 1000;\ \#\ S = 220;\ \#\ C = 160;\ \#\ E = 100;\ \#\ S\cap C\cap E = 20$$

$$\#\ S\cap E = 60;\ \#\ S\cap C = 90;\ ; Pede - se\ P(S\cap \bar{C}\cap \bar{E})$$

Logo:

$$P(S \cap \bar{C} \cap \bar{E}) = P(S) - P(S \cap E) - P(S \cap C) + P(S \cap C \cap E)$$

$$P(S \cap \bar{C} \cap \bar{E}) = \frac{220 - 60 - 90 + 20}{1000}$$

$$P(S \cap \bar{C} \cap \bar{E}) = 0.09$$

Resposta: A probabilidade de que um morador desse bairro tenha apenas curso superior é de 9%.

41. Um gerente de banco tem que decidir se concede ou não empréstimo aos clientes que o solicitam. Ele analisa diversos dados para estudar a possibilidade de o cliente vir a inadimplente. Com base em dados passados, ele estima em 12% a taxa de inadimplência. Dentre os inadimplentes, ele tem 85% de chance de tomar a decisão certa, enquanto que essa chance aumenta para 95% entre os clientes adimplentes. Esse gerente acaba de recusar um empréstimo. Qual é a probabilidade de ele ter tomado a decisão incorrecta?

Resolução

Seja: $I = o\ cliente\ é\ inadiplente; A = o\ cliente\ é\ adiplente;$

$$R = o\ gerente\ recusa\ o\ empréstimo$$

Pelos dados temos:

$$P(I) = 0.12; P(A) = 0.88;\ P(R \backslash I) = 0.85; P(\bar{R} \backslash A) = 0.95;$$

$$P(R \backslash A) = 0.05\ Aqui\ pede - se\ P(A \backslash R)?$$

Nota que se ele acaba de recusar um empréstimo, só pode ter tomado a decisão incorrecta se recusou conceder o empréstimo a um adiplente

Logo, pelo teorema de Bayes temos:

$$P(A \backslash R) = \frac{P(A \cap R)}{P(R)} = \frac{P(A) * P(R \backslash A)}{P(A) * P(R \backslash A) + P(I) * P(R \backslash I)}$$

$$P(A \backslash R) = \frac{0.88 * 0.05}{0.88 * 0.05 + 0.12 * 0.85}$$

$$P(A \backslash R) = 0.301$$

Resposta: A probabilidade do gerente ter tomado a decisão incorrecta dado que acaba de recusar um empréstimo é de 30.1%.

42. Uma pesquisa realizada sobre a preferência dos consumidores por três categorias de veículos A, B e C de uma indústria automobilística revelou que dos 100 entrevistados, 65 preferiam o veículo A; 68 preferiam o veículo B; 50 preferiam o veículo C; 40 preferiam os veículos A e B; 35 preferiam os veículos A e C; 30 preferiam os veículos B e C e 4 não preferem nenhuma das três categorias de veículos. Um consumidor é selecionado ao acaso entre os entrevistados. Calcule a probabilidade de que ele prefira simultaneamente as três categorias.

Resolução

Pelos dados temos:

$$n = 100;\ \#\, A = 65;\ \#\, B = 68;\ \#\, C = 50;\ \#\, A \cap B = 40$$

$$\#\, A \cap C = 60;\ \#\, B \cap C = 30;\ Pede - se\ P(A \cap B \cap C)$$

Logo pelo teorema de união de três eventos temos:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C\ ) - P(B \cap C) + P(A \cap B \cap C)$$

$$1 - 0.04 = 0.65 + 0.68 + 0.50 - 0.40 - 0.35 - 0.30 + P(A \cap B \cap C)$$

$$P(A \cap B \cap C) = 0.18$$

Resposta: A probabilidade de que ele prefira simultaneamente as três categorias é de 18%.

43. O gerente de Recursos Humanos de uma empresa escolhe estagiários oriundos de dois cursos de Administração. No curso 1, a proporção de alunos com boa formação em informática é de 60%, enquanto no outro curso, essa proporção cai para 45%. Um estagiário acaba de ser contratado. A probabilidade de que tenha boa formação em informática é 0,50. Qual é a preferência (probabilidade) do gerente pelo curso 1?

<u>Resolução</u>

Seja: $I = curso\ 1; II = curso\ 2; B = boa\ formação\ em\ informática$

Pelos dados temos:

$$P(B\backslash I) = 0.60; P(B\backslash II) = 0.45; P(B) = 0.50. Aqui\ pede - se\ P(I)$$

Pelo teorema de probabilidade total temos:

$$P(B) = P(I \cap B) + P(II \cap B)$$

$$P(B) = P(I) * P(B\backslash I) + P(II) * P(B\backslash II)$$

$$0.50 = P(I) * 0.60 + 0.45 * \left(1 - P(B\backslash I)\right)$$

$$0.50 = 0.60 * P(I) + 0.45 - 0.45 * P(I)$$

$$P(I) = \frac{0.05}{0.15} = 0.333$$

Resposta: A preferência (probabilidade) do gerente pelo curso 1 é 33.3%.

44. O chefe do Sector de Compras de uma empresa trabalha com 3 grandes distribuidores de material de escritório. O distribuidor 1 é responsável por 70% dos pedidos, enquanto cada um dos outros 2 distribuidores responde por 15% dos pedidos. A proporção de pedidos com atraso dos três distribuídores é de 4%, 5% e 6% respectivamente. Determine a proporção dos pedidos que chegam com atraso.

<u>Resolução</u>

Seja: $I = pedidos\ do\ distribuídor\ I; II = pedidos\ do\ distribuídor\ II;$

$$A = pedidos\ com\ atraso$$

Pelos dados temos:

$$P(I) = 0.70; P(II) = P(III) = 0.15; P(A\backslash I) = 0.04; P(A\backslash II) = 0.05; P(A\backslash III) = 0.06$$

$$Aqui\ pede\ se\ P(A)$$

Pelo teorema de probabilidade total temos:

$$P(B) = P(I \cap A) + P(II \cap A) + P(III \cap A)$$

$$P(B) = P(I) * P(A\backslash I) + P(II) * P(A\backslash II) + P(III) * P(A\backslash II)$$

$$P(B) = 0.70 * 0.04 + 0.15 * 0.05 + 0.15 * 0.06$$

$$P(B) = 0.0445$$

Resposta: A proporção dos pedidos que chegam com atraso é de 4.45%.

45. Para seleccionar seus funcionários, uma empresa oferece aos candidatos um curso de treinamento durante uma semana. No final do curso, eles são submetidos a uma prova e 30% são classificados como bons, 42% como médios e os restantes 28% como fracos. Para facilitar a selecção, a empresa pretende substituir o treinamento por um teste contendo questões referentes a conhecimentos gerais e específicos. Para isso, gostaria de conhecer qual a probabilidade de um indivíduo aprovado no teste ser considerado fraco, caso fizesse o curso. Assim, neste ano, antes do início do curso, os candidatos foram submetidos ao teste e receberam o conceito aprovado ou reprovado. No final do curso 80% dos funcionários bons foram aprovados, metade dos médios reprovados e 20% dos fracos aprovados. Um funcionário acaba de ser aprovado, qual é a probabilidade de ser fraco?

<u>Resolução</u>

Seja: $B = funcionários\ bons; M = funcionários\ médios; F = funcionários\ fracos;$

$$A = funcionários\ aprovados$$

Pelo dados temos:

$$P(B) = 0.30; P(M) = 0.42; P(F) = 0.28; P(A\backslash B) = 0.80; P(A\backslash M) = 0.50;$$

$$P(A\backslash F) = 0.20$$

Pede-se: $P(F\backslash A)$

Então:

$$P(F\backslash A) = \frac{P(F \cap A)}{P(A)}$$

Pelo teorema de Bayes temos:

$$P(F\backslash A) = \frac{P(F \cap A)}{P(B \cap A) + P(M \cap B) + P(F \cap B)}$$

$$P(F\backslash A) = \frac{P(F \cap A)}{P(B) * (A\backslash B) + P(M) * P(A\backslash M) + P(F) * P(A\backslash F)}$$

$$P(F\backslash A) = \frac{0.28 * 0.20}{0.30 * 0.80 + 0.42 * 0.50 + 0.28 * 0.20}$$

$$P(F\backslash A) = 0.1107$$

Resposta: A proporção do funcionário ser fraco dado que acaba de ser aprovado é de 11.1%.

46. Imaginas que a sua empresa foi adjudicada para produzir material de votação de um processo eleitoral e possui três impressoras gráficas offset que podes designares de I1, I2, I3. Sabe-se que a impressora 1 tem uma capacidade de fazer 15000 impressões por hora, a segunda com capacidade de 13000 impressões por hora e a terceira com capacidade de 12000 impressões por hora. Os registos mostram algumas imperfeiçoes das impressões de tal forma que 5%, 4% e 3% das impressões saem com defeito da I1, I2, I3 respectivamente. Sabendo que uma impressão contem defeito, calcule a probabilidade de ser proveniente da primeira impressora.

<u>Resolução</u>

Pelos dados temos:

$$P(I_1) = \frac{15000}{40000} = 0.375;\ P(I_2) = \frac{13000}{40000} = 0.325;\ P(I_3) = \frac{12000}{40000} = 0.3$$

$$P(D\backslash I_1) = 0.05; P(D\backslash I_2) = 0.04; P(D\backslash I_3) = 0.03;$$

$$Aqui\ pede - se\ P(I_1\backslash D)?$$

Logo, pelo teorema de Bayes temos:

$$P(I_1\backslash D) = \frac{P(I_1 \cap D)}{P(D)} = \frac{P(I_1) * (D\backslash I_1)}{\sum(I_i\backslash D) * P(I_i)} =$$

$$P(I_1\backslash D) = \frac{P(I_1) * (D\backslash I_1)}{(D\backslash I_1) * P(I_1) + (D\backslash I_2) * P(I_2) + (D\backslash I_3) * P(I_3)}$$

$$P(I_1\backslash D) = \frac{0.375 * 0.05}{0.05 * 0.375 + 0.04 * 0.325 + 0.03 * 0.3}$$

$$P(A\backslash R) = 0.4596$$

Resposta: A probabilidade da impressão ser proveniente da primeira impressora, dado que contem defeito é de 45.96%.

47. Os estudos epidemiológicos indicam que 20% dos idosos sofrem de uma deterioração neuropsicológica. Sabe-se que a tomografia axial computadorizada (TAC) é capaz de detectar esse transtorno em 80% dos que sofrem disso, mas que também resulta 3% de falso positivo entre pessoas com boa saúde. Se for escolhido um idoso ao acaso, sendo o resultado do seu TAC positivo, qual é a probabilidade de que ele realmente esteja enfermo?

<u>Resolução</u>

Sejam os eventos:

$$E = Doente\ está\ enfermo; P = Exame\ deu\ positivo; N = Exame\ deu\ negativo$$

Pelos dados temos:

$P(E) = 20\%; P(\bar{E}) = 80\%; P(P\backslash E) = 80\%; P(N\backslash E) = 20\%; P(P\backslash\bar{E}) = 3\%;$

$P(P\backslash\bar{E}) = 97\%$ Pede-se $P(E\backslash P)$

Recorrendo ao teorema de Bayes temos:

$$P(E\backslash P) = \frac{P(E) * P(P\backslash E)}{P(E) * P(P\backslash E) + P(\bar{E}) * P(P\backslash\bar{E})} = \frac{0.20 * 0.80}{0.20 * 0.80 + 0.80 * 0.03} = 0.869$$

Resposta: Se for escolhido um idoso ao acaso, sendo o resultado do seu TAC positivo, a probabilidade de que ele realmente esteja enfermo é de 86.9%

48. Um grupo de pessoas foi classificado quanto a peso e pressão arterial de acordo com as proporções do quadro a seguir:

Pressão	Peso			Total
	Excesso	Normal	Deficiente	
Elevada	0.10	0.08	0.02	0.20
Normal	0.15	0.45	0.20	0.80
Total	0.25	0.53	0.22	1.00

a) Qual a probabilidade de uma pessoa deste grupo, escolhida ao acaso, ter pressão elevada?

<u>Resolução</u>

Seja os eventos:

$PE = Pressão\ elevada; PN = Pressão\ normal; EP = Excesso\ de\ peso;$

$PN = Peso\ normal; PD = Peso\ deficiente$

Pede-se $P(PE)$

Então:

$$P(PE) = P(PE \cap EP) + P(PE \cap PN) + P(PE \cap PD)$$

$$P(PE) = 0.10 + 0.08 + 0.02 = 0.20$$

Resposta: A probabilidade de uma pessoa deste grupo, escolhida ao acaso, ter pressão elevada é de 20%.

b) Verifica-se que a pessoa escolhida tem excesso de peso, qual a probabilidade de ter também pressão elevada?

<u>Resolução:</u>

$$P(PE\backslash EP) = \frac{P(PE \cap EP)}{P(EP)} = \frac{0.10}{0.25} = 0.40$$

Resposta: A probabilidade da pessoa deste grupo ter pressão elevada dado que tem excesso de peso é de 40%.

c) Os eventos "excesso de peso" e "pressão elevada" são independentes?

Resolução

Para que esses eventos sejam independentes é necessário que:

1) $P(EP \backslash PE) = P(EP)$

2) $P(PE \backslash EP) = P(PE)$

3) $P(EP \cap PE) = P(EP) * P(PE)$

Então

$$P(EP \backslash PE) = \frac{P(PE \cap EP)}{P(PE)} = \frac{0.10}{0.20} = 0.50 \neq P(EP) = 0.25$$

$$P(PE \backslash EP) = \frac{P(PE \cap EP)}{P(EP)} = \frac{0.10}{0.25} = 0.40 \neq P(PE) = 0.20$$

$$P(PE \cap EP) = 0.10 \neq 0.20 * 0.25 = 0.05$$

Conclusão: Nenhum dos pressupostos foi verificado, logo os eventos "excesso de peso" e "pressão elevada" não são independentes

49. O quadro abaixo apresenta a situação de estudantes de uma certa universidade por sexo e bolsa de estudos. Sorteado um estudante ao acaso, qual a probabilidade de que ele possua as seguintes características:

	Bolseiro (B)	Não Bolseiro (NB)	Total
Mulheres (M)	20	15	35
Homens (H)	30	45	75
Total	50	60	110

a) Ser bolseiro;

Reolução:

$$P(B) = \frac{\# B}{\# n} = \frac{50}{110} = 0.4545$$

Resposta: A probabilidade de que ele seja bolseiro é de 45.45%

b) Ser homem sabendo que foi seleccionado um bolseiro;

Reolução:

$$P(H \backslash B) = \frac{P(H \cap B)}{P(B)} = \frac{30}{50} = 0.60$$

Resposta: A probabilidade de que ele seja homem dado que é bolseiro é de 60%

c) Seleccionadas duas pessoas, uma após outra, sem reposição, qual é a probabilidade de que ambas sejam mulheres sabendo que foram selecionadas pessoas do mesmo sexo.

Resolução

Pede-se:

$$P\left[(M1 \cap M2) \backslash \left((M1 \cap M2) \cup (H1 \cap H2)\right)\right]?$$

$$P\left[(M1 \cap M2) \backslash \left((M1 \cap M2) \cup (H1 \cap MH)\right)\right] = \frac{\frac{35}{110} * \frac{34}{109}}{\frac{35}{110} * \frac{34}{109} + \frac{75}{110} * \frac{74}{109}} = 0.18$$

50. Uma pesquisa realizada entre 1000 consumidores, registrou que 650 deles trabalham com cartões de crédito da bandeira MasterCard, que 550 trabalham com cartões de crédito da bandeira VISA e que 200 trabalham com cartões de crédito de ambas as bandeiras. Qual a probabilidade de ao escolhermos deste grupo uma pessoa que utiliza a bandeira VISA, ser também um dos consumidores que utilizam cartões de crédito da bandeira MasterCard?

Resolução:

Consideremos os eventos

A= Consumidores que trabalham com cartões de crédito da bandeira MasterCard

B= Consumidores que trabalham com cartões de crédito da bandeira VISA

Pelos dados temos:

$$P(A \backslash B) = \frac{P(A \cap B)}{P(B)} = \frac{200}{550} = 0.364$$

51. Suponha que uma loja dispõe se de dois produtos substitutos, a preferência do produto A pelos Clientes é de 46.67%, do produto B é de 55% e de ambos os produtos é de 21.67. Se um cliente é escolhido ao acaso, determine a probabilidade de:

a) Preferir produto A ou B;

Resolução

$$P(AUB) = P(A) + P(B) - P(A \cap B)$$

$$P(AUB) = 0.466 + 0.55 - 0.2167 = 0.7993$$

b) Preferir somente o produto A;

Resolução

$$P(A \cap \bar{B})A = P(AUB) - P(B) = 0.7993 - 0.55 = 0.2493$$

c) Preferir o produto A na condição de ser usuário do produto B.

Resolução

$$P(A \backslash B) = \frac{P(A \cap B)}{P(B)} = \frac{0.2167}{0.55} = 0.394$$

VARIÁVEL ALEATÓRIA DISCRETA

52. Um vendedor de equipamentos pesados pode visitar, num dia, um ou dois clientes, com mesma probabilidade. De cada contacto, pode resultar a venda de um equipamento por 50.000 Mt com probabilidade 1/10 ou nenhuma venda, com probabilidade 9/10. Indicando por Y o valor total de vendas diárias deste vendedor:

a) Escreva a função de probabilidade de Y.

Resolução

Considere os seguintes eventos:

$$A = vendedor\ visita\ um\ cliente$$

$$B = vendendor\ visita\ dois\ clientes$$

$$V = a\ visita\ resulta\ na\ venda\ de\ \ equipamento\ do\ primeiro\ cliente$$

$$V_1 = a\ visita\ resulta\ na\ venda\ de\ \ equipamento\ do\ segundo\ cliente$$

$$V_2 = a\ visita\ resulta\ na\ venda\ de\ \ equipamento\ do\ terceiro\ cliente$$

Então:

$$P(A) = \frac{1}{3}, P(B) = \frac{2}{3}, P(V/A) = \frac{1}{10}, P(\bar{V}/A) = \frac{9}{10}, P(V_1 \cap V_2/B) = \frac{2}{300},$$

$$P(V_1 \cap \bar{V}_2/B) = \frac{18}{300}, P(\bar{V}_1 \cap V_2/B) = \frac{18}{300}, P(\bar{V}_1 \cap \bar{V}_2/B) = \frac{162}{300}$$

$$A\ f.d.p\ das\ vendas\ \text{é}:$$

Nº de vendas	0	1	2
P (Vendas)	9/10+162/300=252/300	1/10+18/300+18/300=46/300	2/300

Seja $v.a: Y = o\ valor\ total\ de\ vendas\ diárias\ deste\ vendedor$. *Então a f.d.p será:*

Y (USD)	0	50000	100000
P(Y)	252/300	46/300	2/300

b) Calcule o valor total esperado de vendas diárias.

Resolução

Recorrendo a fórmula para cálculo de esperança matemática em v.a discretas temos:

$$E(Y) = \sum Y_i * P(Y_i)$$

$$E(Y) = 0 * \frac{252}{300} + 50000 * \frac{46}{300} + 100000 * \frac{2}{300} = 8333.33$$

Resposta: O valor total esperado de vendas diárias é de 8333 Mt.

53. Um gestor de uma empresa esta a considerar a hipótese de substituir uma das maquinas, para tal, realizou um estudo sobre a situação da maquina durante uma semana. A tabela seguinte mostra o numero de avarias e a correspondente probabilidade.

Nº de avarias	3	4	5	6	7
Probabilidade	0.10	k	0.42	0.10	0.20

a) Determine o valor de k

Resolução:

Nota que a tabela acima indica uma f.d.p de uma variavel aleatória discreta.

Logo:

$$\sum P(x) = 1$$

Então:

$$0.10 + k + 0.42 + 0.10 + 0.20 = 1$$

$$K = 0.18$$

b) Determine o número médio de avarias na empresa.

Resolução

$$E(x) = \sum x_i * P(x_i) = 3 * 0.1 + 4 * 0.18 + 5 * 0.42 + 6 * 0.1 + 7 * 0.2 = 5.12$$

c) Sabe se que cada avaria custa 7000Mt em perdas. Determine o prejuízo médio da empresa.

Resolução

$$Prejuízo\ Médio\ = \ Custo\ Unitário * E(x) = 7000 * 5.12 = 35840{,}00\ Mts$$

54. A procura de bolos de eventos da dona Zulfat no meio de semana é uma variável aleatória discreta com a seguinte distribuição de probabilidade

$$f(x) = k\frac{2^x}{x!}\ , x = 0,1,2,3,4$$

a) Determine o valor de K.

Resolução

Para que seja uma f.d.p o $\sum P(x_i) = 1$

Logo:

$$k * \left(\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!} + \frac{2^4}{4!}\right) = 1$$

$$k = \frac{1}{7}$$

Resposta: O valor de k é de 1/7.

b) Suponha que cada bolo é vendido por 500.00 MT. A dona Zulfate produz diariamente 3 bolos. Qualquer bolo que não tenha sido vendido ao fim do dia deve ser jogado fora provocando um prejuízo de 300.00 MT. Quanto espera a dona Zulfat ganhar em cada dia?

<u>Resolução</u>

$$Seja\ v.a: G = valor\ ganho\ por\ dia; Aqui\ pede - se\ E(G)$$

$$Pela\ fórmula\ temos:$$

$$E(x) = \sum_{i=1}^{n} x_i * P(x_i)$$

Determinando a f.d.p da v.a G, teremos:

x	0	1	2	3	4	Somatório
Prejuízo	-900	-600	-300	0	0	-
$Ganho\ (G)$	-900	-100	700	1500	1500	-
$P(G)$	1/7	2/7	2/7	4/21	2/21	1
$\sum G_1 * P(G_i)$	-128.57	-28.57	200	285.71	142.86	471.43

Logo:

$$E(x) = \sum G_1 * P(G_i) = 471.43$$

Resposta: A dona Zulfat espera gamha 471.43 Mt por dia

55. Na produção de uma peça são empregadas 2 máquinas. A primeira é utilizada para efectivamente produzir as peças e o custo de produção é de 50,00 u.m por peça. Das peças produzidas nessa máquina, 20% são defeituosas. As peças defeituosas são colocadas na segunda máquina para a tentativa de recuperação. Nessa segunda máquina o custo de produção é de 35,00 u.m mas apenas 55% das peças são de facto recuperadas. Cada peça perfeita é vendida por 120,00 u.m e cada peça defeituosa é vendida por 30,00 u.m. Determine o lucro esperado por peça.

<u>Resolução</u>

Seja:

$$1 = Máquina 1; 2 = Máquina\ 2;\ D = peças\ defeituosas; C = custo\ de\ produção;$$

$$V = valor\ da\ venda\ por\ peça;\ L = Lucro;$$

$$Nota\ que\ aqui\ pede - se\ E(L),$$

Pelos dados temos:

$$P(\overline{D}_1) = 0.80;\ C_1 = 50\ u.m;\ V_{\overline{D}_1} = 120\ u.m;$$

$$L = V_{\bar{D}_1} - C_1 = 120 - 50 = 70$$

$$P(D_1 \cap \bar{D}_2) = 0.20 * 0.55 = 0.11;\ C_{1\cap 2} = (50 + 35) = 85\ u.m;$$

$$V_{D_1 \cap \bar{D}_2} = 120\ u.m;\ L = V_{D_1 \cap \bar{D}_2} - C_{1\cap 2} = 120 - 85 = 35$$

$$P(D_1 \cap D_2) = 0.20 * 0.45 = 0.09;\ C_{1\cap 2} = (50 + 35) = 85\ u.m;$$

$$V_{D_1 \cap D_2} = 30\ u.m;\ L = V_{D_1 \cap D_2} - C_{1\cap 2} = 30 - 85 = 35$$

Sabe-se que

$$E(x) = \sum_{i=1}^{n} x_i * P(x_i)$$

Determinando a f.d.p temos com a v.a L:=Lucro, teremos:

L (u.m)	-55	35	70	Somatório
P(L)	0.09	0.11	0.80	1
$\sum L_i * P(L_i)$	-4.95	3.85	56	54.9

Logo:

$$E(L) = \sum L_i * P(L_i)$$

$$E(L) = 54.9$$

Resposta: O lucro esperado por peça é de 54.9 u.m.

56. Os rendimentos previsionais dos projectos A e B seguem uma variável aleatória segundo a tabela abaixo, dentro de um ambiente macroeconómico classificado em baixa, moderada e alta. Imagines que possuis um budget e pretende investir no projecto de menor risco (menor variabilidade).

Situações da económicas	Probabilidades	Retorno em % do projecto A	Retorno em % do projecto B
Baixa	0,25	13	7
Moderada	0,60	14	19
Alta	0.15	17	23

Qual seria o projecto de eleição e porquê?

<u>Resolução</u>

Nota que o projecto de eleição será aquele que contem a menor variabilidade.

Tratando-se de variável aleatória discreta temos:

$$V(X) = E(x^2) - [E(x)]^2$$

Onde

$$E(X^2) = \sum_{i=1}^{n} x_i^2.P(x_i) \quad e \quad E(x) = \sum_{i=1}^{n} x_i * P(x_i)$$

Logo:

Projecto A

$$V(X) = (13^2 * 0.25 + 14^2 * 0.6 + 17^2 * 0.15) - (0.25 * 13 + 0.60 * 14 + 0.15 * 17) = 1.56\%$$

Projecto B

$$V(X) = (7^2 * 0.25 + 17^2 * 0.6 + 23^2 * 0.15) - (0.25 * 7 + 0.60 * 17 + 0.15 * 23) = 32.64\%$$

Resposta: Elegeria o projecto A porque apresenta menor risco.

57. Em determinado sector de uma loja de departamentos, o número de produtos vendidos em um dia pelos funcionários é uma variável aleatória P com a seguinte distribuição de probabilidades (esses números foram obtidos dos resultados de vários anos de estudo):

$Nº\ de\ produtos$	0	1	2	3	4	5	6
$Probabilidade\ de\ venda$	0.1	0.4	0.2	0.1	0.1	0.05	0.05

Cada vendedor recebe comissões de venda, distribuídas da seguinte forma: se ele vende até 2 produtos em um dia, ele ganha uma comissão de 25,00 u.m por produto vendido. A partir da terceira venda, a comissão passa para 30,00 u.m. Qual a comissão média de cada um deles?

Resolução

Seja $v.a$: $C = comissão\ de\ venda\ recebida\ \ por\ cada\ vendedor.$

Aqui pede-se $E(C)$

Nota que primeiro deve-se determinar a f.d.p da comissão de venda

Logo:

$Nº\ de\ produtos$	*0*	*1*	*2*	*3*	*4*	*5*	*6*
$Comissão$	*0*	*25*	50	*80*	*110*	*140*	*170*
$Probabilidade\ de\ venda$	*0.1*	*0.4*	*0.2*	*0.1*	*0.1*	*0.05*	*0.05*

Então:

$$E(C) = \sum C_i * P(C_i)$$

$$E(C) = 0 * 0.1 + 25 * 0.4 + 50 * 0.2 + 80 * 0.1 + 110 * 0.1 + 140 * 0.05 + 170 * 0.05 = 54.5$$

Resposta: A comissão média de cada um deles é de 54.5 u.m.

58. Um revendedor de produtos informáticos recebe de vários fornecedores certo tipo de mouses, que tem custo diferenciado. Levando-se em conta a proporção fornecida e o preço apresentado por cada fornecedor, pode-se considerar que o custo unitário de cada mouse em u.m, escolhida ao acaso, é uma variável aleatória C. Admitindo a seguinte distribuição de probabilidade para C:

ci	1,00	1,10	1,20	1,30	1,40
$P(ci)$	0,2	0,3	0,2	0,2	0,1

Supondo que o revendedor venda cada um desses mouses acrescentando 50% sobre o custo, além de um adicional de 0,10 u.m pelo frete, calcular a média da nova variável aleatória preço de revenda R.

Resolução

Seja $v.a$: $R = preço\ de\ revenda$

Aqui pede-se $E(C)$

Nota que primeiro deve-se determinar a f.d.p do preço de revenda R,

Logo:

ci	1,00	1,10	1,20	1,30	1,40
Ri	1.6	1.75	1.9	2.05	1.08
$P(Ri)$	0,2	0,3	0,2	0,2	0,1

Então:

$$E(C) = \sum R_i * P(R_i)$$

$$E(C) = 1.6 * 0.2 + 1.75 * 0.3 + 1.9 * 0.2 + 2.05 * 0.2 + 1.08 * 0.1 = 1.743$$

Resposta: A média da nova variável aleatória preço de revenda R é de 1.743 u.m.

59. A organização financeira Betha Ltda. verificou que o lucro unitário L, obtido numa operação financeira é dado pela seguinte expressão:

$$L = 1{,}9\,V - 0{,}9\,C - 4{,}5$$

Sabendo-se que o preço de venda unitário V tem uma distribuição de média 50,00 u.m e desvio padrão de 2,00 u.m, e que o preço de custo unitário C tem uma distribuição de média 45,00 u.m e desvio padrão 1,50 u.m, qual é a média e o desvio padrão do lucro unitário?

Resolução

Recorrendo as propriedades do valor esperado e variança para variáveis aleatórias temos:

Média

$$E(L) = E(1{,}9\,V - 0{,}9\,C - 4{,}5)$$
$$E(L) = 1.9 * 50 - 0.9 * 45 - 4.5 = 50$$

Desvio Padrão

$$Dp(L) = Dp(1{,}9\,V - 0{,}9\,C - 4{,}5)$$
$$Dp(L) = 1.9 * 2 - 0.9 * 1.5 = 2.45$$

Resposta: A média e o desvio padrão do lucro unitário é de 50 e 2.45 u.m.

60. Uma empresa comercializa um produto que possui um determinado prazo de validade. Para cada unidade vendida do produto, a empresa recebe 800,00 Mt e tem um custo de 200,00 Mt. Nestas condições a probabilidade de um produto ser vendido antes de vencer seu prazo de validade é de 80%. Quanto a empresa espera lucrar em uma unidade do produto?

Resolução

Seja $v.a$: $L = Lucro\ em\ unidade\ do\ produto$

Aqui pede-se $E(L)$

Nota que primeiro deve-se determinar a f.d.p do lucro

Logo:

$Nº\ de\ produtos$	1	1
$Custo$	200	200
$Valor\ recebido\ da\ venda$	0	800
$Lucro$	-200	600
$Probabilidade\ de\ venda$	0.20	0.80

Então:

$$E(L) = \sum L_i * P(L_i)$$

$$E(L) = -200 * 0.20 + 600 * 0.80 = 440$$

Resposta: Em uma unidade do produto a empresa espera lucrar 440 Meticais.

61. Uma transportadora possui uma frota de quatro caminhões de aluguer. Sabe-se que o aluguer é feito por dia e que a distribuição diária do número x de camiões alugados é a seguinte:

x	0	1	2	3	4
$P(x)$	0.1	0.2	0.3	0.3	0.1

Determine a média e a variança do lucro diário, sabendo-se que:

- O valor do aluguer por dia é de 300,00 USD;
- A despesa total diária com manutenção de cada veículo é de 140,00 USD quando este é alugado e de 15,00 USD quando o veículo não é alugado.

Resolução

Seja v.a L=Lucro diário por veículo alugado

Vamos construir uma tabela ilustrativa do lucro

N° de dias	Despesa de manutenção	Lucro
0	15	-15
1	140	300-140=160
2	2*140	2*300-2*140=320
3	3*140	3*300-3*140=480
4	4*140	4*300-4*140=640

Então a f.d.p será dada por:

L	-15	160	320	480	640
$P(L)$	0.1	0.2	0.3	0.3	0.1

Média

$$E(L) = \sum L_i * P(L_i)$$

$$E(L) = -15 * 0.1 + 160 * 0.2 + 320 * 0.3 + 480 * 0.3 + 640 * 0.1 = 334.5$$

Variança

$$V(X) = E(x^2) - [E(x)]^2$$

$$V(L) = (-15)^2 * 0.1 + 160^2 * 0.2 + 320^2 * 0.3 + 640^2 * 0.1 - 334.5^2 = 340.52.25$$

62. O tempo t, em minutos, necessário para um operário processar certa peça é uma v.a. com fdp dada na tabela abaixo.

t	2	3	4	5	6	7
$P(t)$	0.1	0.1	0.3	0.2	0.2	0.1

Para cada peça processada, o operário ganha um fixo de 2 u.m (unidade monetária) mas, se ele processa a peça em menos de 6 minutos, ganha 0,50 u.m por cada minuto poupado.

a) Encontre a função de distribuição da v.a. G = quantia (em u.m.) ganha por peça.

Resolução

t	2	3	4	5	6	7
G	2+0.5*4=4	2+0.5*3=3.5	2+0.5*2=3.0	2+0.5*1=2.5	2+0.5*0=2	2+0.5*0=2
$P(t)$	0.1	0.1	0.3	0.2	0.2	0.1

Logo

G	4	3.5	3	2.5	2
$P(t)$	0.1	0.1	0.3	0.2	0.3

b) Determine a quantia média ganha por peça.

Resolução

$$E(G) = \sum G_i * P(G_i) = 4 * 0.1 + 3.5 * 0.1 + 3 * 0.3 + 2.5 * 0.2 + 2 * 0.3 = 2.75$$

VARIÁVEL ALEATÓRIA CONTÍNUA

63. A demanda diária de arroz num supermercado, em centenas de quilos, é uma variável aleatória com função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{2x}{3} & se\ 0 \leq x < 1 \\ 1 - \frac{x}{3} & Se\ 1 \leq x \leq 3 \\ 0 & Caso\ contrário \end{cases}$$

a) Qual é a probabilidade de se vender mais de 150 kg num dia escolhido ao acaso?

Resolução

$$P(> 1.5) = \int_{1.5}^{3} \left(1 - \frac{x}{3}\right) dx = (3 - 1.5) - \frac{1}{6} * (3^2 - 1.5^2) = 0.375$$

A probabilidade de se vender mais de 150 kg num dia escolhido ao acaso é de 37.5%.

b) Qual a quantidade de arroz que deve ser deixada à disposição dos clientes diariamente para que não falte arroz em 95% dos dias?

Resolução

$$\int_{0}^{k} f(x)\, dx = 0.95$$

$$\int_{k}^{3} f(x)\, dx = 0.05$$

$$\int_{k}^{3} \left(1 - \frac{x}{3}\right) dx = 0.05$$

$$(3 - k) - \frac{1}{6} * (3^2 - k^2) = 0.05$$

$$k = 1.21$$

Resposta: A quantidade de arroz que deve ser deixada à disposição dos clientes diariamente para que não falte arroz em 95% dos dias é de 121Kg.

64. Considere a função:

$$f(x) = \begin{cases} \frac{x}{4} & se\ 0 \leq x \leq 2 \\ 1 - \frac{x}{4} & Se\ 2 < x \leq 4 \\ 0 & Caso\ contrário \end{cases}$$

a) Determine o valor de k tal que $P(x \leq k) = 0.60$.

Resolução

Tratando-se de v.a contínua temos:

$$\int_0^k f(x)dx = 0.60$$

$$\int_0^2 \frac{x}{4}dx + \int_2^k \left(1 - \frac{x}{4}\right) = 0.60$$

$$k^2 - 8k + 12.8 = 0$$

$$k_1 = 2.2 \; e \; k_2 = 5.8$$

Sendo $k = P_{60}$

Então:

$$2 < k < 4$$

Resposta: O valor de k tal que $P(x \leq k) = 0.60$ é de 2.2.

b) Calcule $P\left(x \leq \frac{5}{4} \,/\, \frac{3}{4} \leq x \leq \frac{9}{4}\right)$

<u>Resolução</u>

Pelo teorema de probabilidade condicional temos:

$$P\left(x \leq \frac{5}{4} \,/\, \frac{3}{4} \leq x \leq \frac{9}{4}\right) = \frac{P\left(x \leq \frac{5}{4} \cap \frac{3}{4} \leq x \leq \frac{9}{4}\right)}{P\left(\frac{3}{4} \leq x \leq \frac{9}{4}\right)}$$

$$P\left(x \leq \frac{5}{4} \,/\, \frac{3}{4} \leq x \leq \frac{9}{4}\right) == \frac{\int_{3/4}^{5/4} \frac{x}{4}dx}{\int_{3/4}^{2} \frac{x}{4}dx + \int_2^{9/4}\left(1 - \frac{x}{4}\right)dx} = 0.2285$$

65. Considere a função:

$$f(x) = \begin{cases} 0.06 + 0.04x & se\ 1 \leq x \leq 6 \\ 0 & Caso\ contrário \end{cases}$$

a) Calcule $P(2 \leq x \leq 3)$

Solução

$$P(2 \leq X \leq 3) = \int_2^3 (0.06 + 0.04x)dx = (3-2) * 0.06 + \frac{9-4}{2} * 0.04 = 0.16$$

b) Encontre a função de distribuição de X.

$$F(x) = \begin{cases} 0 & se\ t < 1 \\ \int_1^x (0.06 + 0.04t)dt & se\ 1 \leq t < 6 \\ 1 & se\ t \geq 6 \end{cases}$$

$$\rightarrow F(x) = \begin{cases} 0 & se\ x < 1 \\ 0.02x^2 + 0.06x - 0.08 & se\ 1 \leq x < 6 \\ 1 & se\ x \geq 6 \end{cases}$$

c) Determine o valor de k tal que $\mathrm{P}(\mathrm{x} \leq \mathrm{k}) = 0.6$

$$\int_{1}^{k} (0.06 + 0.04x)dx = 0.6$$

$$0.02 * (k^2 - 1) + 0.06 * (k - 1) = 0.6$$

$$k = 5.8$$

Nota: O valor de k aqui determinado corresponde ao percentil 60

d) Calcule a esperança e a variância de X.

$$E(x) = \int_{-\infty}^{+\infty} x * f_x(x)dx$$

$$E(x) = \int_{1}^{6} x * (0.06 + 0.04x)dx = \frac{0.06}{2} * (6^2 - 1^2) + \frac{0.04}{3} * (6^3 - 1^3) = 3.92$$

e

$$Var(x) = \int_{-\infty}^{+\infty} x^2 * f_x(x)dx - \left[\int_{-\infty}^{+\infty} x * f_x(x)dx\right]^2$$

$$\frac{0.06}{3} * (6^3 - 1^3) + \frac{0.04}{4} * (6^4 - 1^4) - 3.92^2 = 1.88$$

DISTRIBUÍÇÃO DE PROBABILIDADES DE VARIÁVEL ALEATÓRIA DISCRETA

66. Suponha que a probabilidade de que o Metical se desvalorize em relação ao Dólar numa determinada semana seja 0.30, e que o que acontece numa determinada semana seja independente do que acontece numa outra semana. Nota: 1 mês equivale a 4 semanas.

   a) Qual é a probabilidade de que o Metical se desvalorize pelo menos uma semana num período de 2 meses?

Resolução

$$Seja\ v.a: x = periodo\ de\ desvalorização\ do\ metical$$

$$x \sim Bin(n = 8; p = 0.30)$$

$$k = pelo\ menos\ uma\ semana\ num\ período\ de\ 2\ meses$$

$$P(x = k) = \binom{n}{k} * p^k * q^{n-k}$$

Pede-se $P(x \geq 1)$

$$P(x \geq 1) = 1 - P(x < 1) = 1 - P(x = 0)$$

$$P(x \geq 1) = 1 - \binom{8}{0} * 0.30^0 * 0.70^{8-0} = 0.942$$

Resposta: A probabilidade de que o Metical se desvalorize pelo menos uma semana num período de 2 meses é de 94.2%.

   b) Calcule os parâmetros da distribuição.

Resolução

Esperança matemática:

$$E(x) = n.p = 8 * 0.3 = 2.4$$

Variança:

$$V(x) = n.p.q = 8 * 0.3 * 0.7 = 1.68$$

Resposta: Os valores de Esperança e Variança são 2.4 e 1.68 respectivamente

67. Estudos anteriores mostraram que há 73% de chance de consumidores do sexo feminino apresentar uma reação positiva a anúncios publicitários com crianças. Uma agência está conduzindo um estudo, apresentando um novo anúncio para 5 consumidoras.

   a) Qual é a probabilidade de que pelo menos 4 das 5 consumidoras apresentem reação positiva?

Resolução

$$Seja\ v.a: x = nº\ de\ \text{consumidores do sexo feminino apresentar uma reação positiva}$$

$$x \sim Bin(n = 5; p = 0.73)$$

$$k = pelo\ menos\ uma\ semana\ num\ período\ de\ 2\ meses$$

$$P(x = k) = \binom{n}{k} * p^k * q^{n-k}$$

Pede-se $P(x \geq 4)$

$$P(x \geq 4) = P(x = 4) + P(x = 5)$$

$$P(x \geq 4) = \binom{5}{4} * 0.73^4 * 0.27^{5-4} + \binom{5}{5} * 0.73^5 * 0.27^{5-5} = 0.591$$

Resposta: A probabilidade de que pelo menos 4 das 5 consumidoras apresentem reação positiva é de 59.1%.

b) Determine o valor esperado e variança do número de consumidoras que apresentam reação positiva.

Resolução

Esperança matemática:

$$E(x) = 5 * 0.73 = 8 * 0.3 = 3.65$$

Variança:

$$V(x) = n.p.q = 5 * 0.73 * 0.27 = 0.9855$$

Resposta: Os valores de Esperança e Variança são 3.65 e 0.9855 respectivamente

68. Uma empresa especializada em fabrico e distribuíção de garrafas plásticas para a conservação de água, sabe que 85% das garrafas fabricadas não apresentam nenhum defeito. Os compradores I, II e III classificam as quantidades adquiridas em categorias A e B, pagando 1300 u.m. e 900 u.m respectivamente, do seguinte modo:

- Comprador I: retira uma amostra de 5 garrafas; se encontrar pelo menos uma defeituosa, classifica como B;
- Comprador II: retira uma amostra de 10 garrafas; se encontrar pelo menos 2 defeituosas, classifica como B.
- Comprador III: retira uma amostra de 6 garrafas; se encontrar mais que 1 defeituosas, classifica como B.

Em média, qual comprador é menos vantajoso para a empresa?

Resolução

Nota que para identificar o comprador que fornece o maior lucro para a empresa, deve-se determinar quantia média paga por comprador.

$$Seja: Q_t = quantia\ paga$$

Sabendo que temos duas catregorias, então:

$$E(Q_t) = Q_A * P(A) + Q_B * P(B)$$

Comprador I

Primeiro deve-se determinar a probabilidade de pagar uma das quantias.

Pelos dados temos:

$$p = 0.15; 1 - p = 0.85;\ n = 5; P(Q = 900) = P(x \geq 1)$$

$$onde\ x = nº\ de\ garrafas\ defeituosas$$

Logo:

$$x \sim Bin(n, p)\ \ onde\ \ P(x = k) = \binom{n}{k} * p^k * q^{n-k}$$

Então:

$$P(Q = 900) = P(x \geq 1) = 1 - P(x < 1) = 1 - P(x = 0) = 1 - \binom{5}{0} * 0.15^0 * 0.85^5 = 0.5563$$

Assim:

$$E(Q_t) = Q_A * P(A) + Q_B * P(B) = 1300 * (1 - 0.5563) + 900 * 0.5563$$
$$E(Q_t) = 1077.48$$

<u>Comprador II</u>

Primeiro deve-se determinar a probabilidade de pagar uma das quantias.

Pelos dados temos:

$$p = 0.15; n = 10; P(Q = 900) = P(x \geq 2)$$

$$onde\ x = nº\ de\ garrafas\ defeituosas$$

Logo:

$$x \sim Bin(n, p)\ \ onde\ \ P(x = k) = \binom{n}{k} * p^k * q^{n-k}$$

Então:

$$P(Q = 900) = P(x \geq 2) = 1 - P(x < 2) = 1 - \{P(x = 0) + P(x = 1)\}$$
$$= 1 - \left\{\binom{10}{0} * 0.15^0 * 0.85^{10} + \binom{10}{1} * 0.15^1 * 0.85^9\right\} = 0.4557$$

Assim:

$$E(Q_t) = Q_A * P(A) + Q_B * P(B) = 1300 * (1 - 0.4557) + 900 * 0.4557$$
$$E(Q_t) = 1117.72$$

<u>Comprador III</u>

Primeiro deve-se determinar a probabilidade de pagar uma das quantias.

Pelos dados temos:

$$p = 0.15; n = 6; P(Q = 900) = P(x > 1)$$

$$onde\ x = nº\ de\ garrafas\ defeituosas$$

Logo:

$$x \sim Bin(n, p)\ \ onde\ \ P(x = k) = \binom{n}{k} * p^k * q^{n-k}$$

Então:

$$P(Q = 900) = P(x > 1) = 1 - P(x \leq 1) = 1 - \{P(x = 0) + P(x = 1)\}$$
$$= 1 - \left\{\binom{6}{0} * 0.15^0 * 0.85^6 + \binom{6}{1} * 0.15^1 * 0.85^5\right\} = 0.2235$$

Assim:

$$E(Q_t) = Q_A * P(A) + Q_B * P(B) = 1300 * (1 - 0.2235) + 900 * 0.7765$$
$$E(Q_t) = 1708.3$$

Resposta: $E(I) < E(II) < E(III), Logo\ o$ *comprador I é menos vantajoso para a empresa.*

69. Determinada empresa tem três eventuais compradores de seu produto, que pagam preços em função da qualidade:

- O comprador A paga 1.300 dólares por artigo (Grupo A), se em uma amostra de seis artigos não encontrar nenhum defeituoso e 650 dólares pelo restante (Grupo B);
- O comprador B paga 900 dólares por artigo (Grupo A), desde que encontre no máximo um artigo defeituoso em seis artigos, pagando pelo restante 700 dólares (Grupo B);
- O comprador C paga 620 dólares por artigo (Grupo A), aceitando até dois defeituosos em uma amostra de seis, e paga pelo restante 430 dólares (Grupo B);

Qual dos compradores não deveria ser escolhido pelo empresário, se ele sabe que na produção 90% são totalmente perfeitos?

Resolução

Nota que para identificar o comprador que não deveria ser escolhido pelo empresário, deve-se determinar quantia média paga por comprador.

$$Seja: Q_t = quantia\ paga$$

Sabendo que temos duas catregorias, então:

$$E(Q_t) = Q_A * P(A) + Q_B * P(B)$$

Comprador A

Primeiro deve-se determinar a probabilidade de pagar uma das quantias.

Pelos dados temos:

$$p = 0.10; 1 - p = 0.90;\ n = 6; P(Q = 1300) = P(x = 0)$$

$$onde\ x = nº\ de\ artigos\ defeituosas$$

Logo:

$$x \sim Bin(n, p)\ \ onde\ \ P(x = k) = \binom{n}{k} * p^k * q^{n-k}$$

Então:

$$P(Q = 1300) = P(x = 0) = \binom{6}{0} * 0.10^0 * 0.90^6 = 0.5314$$

Assim:

$$E(Q_t) = Q_A * P(A) + Q_B * P(B) = 1300 * (0.5314) + 650 * (1 - 0.5314)$$

$$E(Q_t) = 995.41$$

Comprador B

Primeiro deve-se determinar a probabilidade de pagar uma das quantias.

Pelos dados temos:

$$p = 0.1; n = 10; P(Q = 900) = P(x \leq 1)$$

$$onde\ x = nº\ de\ garrafas\ defeituosas$$

Logo:

$$x \sim Bin(n,p) \ \ onde \ \ P(x=k) = \binom{n}{k} * p^k * q^{n-k}$$

Então:

$$P(Q=900) = P(x \le 1) = P(x=0) + P(x=1) = \binom{6}{0} * 0.10^0 * 0.90^6 + \binom{6}{1} * 0.10^1 * 0.90^5$$
$$= 0.8857$$

Assim:

$$E(Q_t) = Q_A * P(A) + Q_B * P(B) = 900 * 0.8857 + 700 * (1 - 0.8857)$$
$$E(Q_t) = 877.14$$

Comprador C

Primeiro deve-se determinar a probabilidade de pagar uma das quantias.

Pelos dados temos:

$$p = 0.10; n = 6; P(Q = 620) = P(x \le 2)$$

$$onde\ x = nº\ de\ garrafas\ defeituosas$$

Logo:

$$x \sim Bin(n,p) \ \ onde \ \ P(x=k) = \binom{n}{k} * p^k * q^{n-k}$$

Então:

$$P(Q=900) = P(x \le 2) = P(x=0) + P(x=1) + P(x=2)$$
$$= \binom{6}{0} * 0.10^0 * 0.90^6 + \binom{6}{1} * 0.10^1 * 0.90^5 + \binom{6}{2} * 0.10^2 * 0.90^4 = 0.9842$$

Assim:

$$E(Q_t) = Q_A * P(A) + Q_B * P(B) = 620 * 0.9842 + 430 * (1 - 0.9842)$$

$$E(Q_t) = 616.998$$

Resposta: $E(A) > E(B) > E(C), Logo\ o$ comprador C não deveria ser escolhido pelo empresário.

70. Tendo faltado às aulas num determinado dia, um estudante pretende obter informação sobre o decorrer daquelas naquele dia. Para tal decide telefonar a seus colegas até obter a informação procurada. Supõe-se que a probabilidade de obter informação de um determinado colega seja 0.70. Seja X a variável aleatória que indica o número de chamadas telefónicas necessárias até obter a informação. Encontre a probabilidade de que sejam necessárias mais de três chamadas para se obter a informação.

Resolução

$$Seja\ v.a: x = nº\ de\ chamadas\ telefónicas\ necessárias\ até\ obter\ a\ informação$$

$$x \sim Geom(p = 0.70)$$

$$P(x=k) = p * q^{k-1}$$

Pede-se $P(x > 3)$

$$P(x > 3) = 1 - P(x > 3) = 1 - \{P(x = 1) + P(x = 2) + P(x = 3)\}$$

Nota que na distribuíção geométrica o valor de k (n° de experimentos) é de pelo menos 1.

$$P(x > 3) = 1 - (p + pq + pq^2)$$

$$P(x > 3) = 1 - (0.7 + 0.3 * 0.7 + 0.7 * 0.3^2) = 0.027$$

Resposta: A probabilidade de que sejam necessárias mais de três chamadas para se obter a informação é de 2.7%.

71. Determinado tipo de parafuso é vendido em caixas com 1000 peças. É uma característica da fabricação produzir 10% de defeituosas. Normalmente, cada caixa é vendida por 50000.00 Mt. Um comprador faz a seguinte proposta para o produtor: de cada caixa, ele escolhe uma amostra de 20 peças; se ele não encontrar nenhuma defeituosa, ele paga 60000.00 Mt pela caixa; 1 ou 2 defeituosas, ele paga 40000.00 Mt; 3 ou mais defeituosas, ele paga 30000.00 Mt. Qual é a alternativa mais vantajosa para o fabricante?

Resolução

$$Seja\ v.a\ x = nº\ de\ peças\ defeitosas, onde\ P = 0.10$$

$$x \sim Bin\ (n = 20; P = 0.20)$$

$$Pela\ fórmula: P(x = k) = \binom{n}{k} * p^k * q^{n-k}\ \text{temos:}$$

$$P(x = 0) = \binom{20}{0} * 0.10^0 * 0.90^{20} = 0.12158$$

$$P(x = 1) = \binom{20}{1} * 0.10^1 * 0.90^{19} = 0.27017$$

$$P(x = 2) = \binom{20}{2} * 0.10^2 * 0.90^{18} = 0.28518$$

$$P(x \geq 3) = 1 - P(x < 3) = 1 - (0.12158 + 0.27017 + 0.28518) = 0.32307$$

Determinando fdp da $v.a\ x = nº\ de\ peças\ defeitosas$ temos:

x	0	1 ou 2	Mais que 3	Somatório
$P(x)$	0.12158	0.55535	0.32307	1

Seja agora a proposta do vendedor, onde $v.a\ y = valor\ a\ pagar\ por\ caixa$, determinando o f.d.p teremos:

y	60000	40000	30000	Somatório
$P(y)$	0.12158	0.55535	0.32307	1

$$E(y) = \sum y_i * P(y) = 60000 * 0.12158 + 40000 * 0.55535 + 0.32307 * 30000 = 39200.9$$

Logo

Resposta: A proposta, a alternativa mais vantajosa para o fabricante é manter o preço de 50000, rejeitando a proposta do vendedor.

72. Um distribuidor de gasolina tem capacidade de receber, nas condições atuais, no máximo, três camiões por dia. Se chegarem mais que três camiões, o excesso deve ser enviado a outro distribuidor, e, nesse caso, há uma perda média de 800 dólares, por dia, em que não se pode aceitar todos camiões. Sabendo-se que o número de camiões que chegam diariamente obedece à distribuição de Poisson de média 2.

    a) Qual é a probabilidade de, em certo dia, ter-se que mandar camiões para outro distribuidor.

Resolução

$$v.a: x \sim Poisson\ (\mu = 2)$$

$$P(X = k) = \frac{(\lambda t)^K * e^{-\lambda t}}{k!}$$

Para que sejam mandados camiões para outro distribuidor, é necessário que cheguem 4 ou mais camiões

$$P(X = k) = \frac{(\lambda t)^K * e^{-\lambda t}}{k!}$$

$$P(X > 3) = 1 - P(x \leq 3) = 1 - \{P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3)\}$$

$$P(X \geq 4) = 1 - e^{-2} * \left(\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!}\right) = 0.1428765$$

Resposta: A probabilidade de, em certo dia, ter-se que mandar camiões para outro distribuidor é de 14.3%.

    b) Qual é a perda média mensal (30 dias) por causa de camiões que não puderam ser aceitos.

Resolução

$$Seja\ v.a: y = a\ perda\ mensal\ em\ USD$$

$$v.a: y \sim Bin(n = 800 * 30; p = 0.143)$$

$$E(Y) = n * p = 800 * 30 * 0.143 = 3432$$

A perda média mensal (30 dias) por causa de camiões que não puderam ser aceitos é de 3432.00 USD.

73. Na manufactura de certo artigo, é sabido que 1 entre 10 artigos é defeituoso. Uma amostra de tamanho 4 é retirada com reposição, de um lote da produção. Qual a probabilidade de que a amostra contenha:

    a) Nenhum defeituoso?

<u>**Resolução**</u>

$$v.a: x \sim Bin\left(n = 4; p = \frac{1}{10} = 0.10\right)$$

$$P(x = k) = \binom{n}{k} * p^k * q^{n-k}$$

$$P(x = 0) = \binom{4}{0} * 0.10^0 * 0.90^4 = 0.0001$$

Respostra: A probabilidade de que a amostra contenha nenhum defeituoso é de 0.01%.

b) Exactamente 1 defeituoso?

<u>**Resolução**</u>

$$v.a: x \sim Bin\left(n = 4; p = \frac{1}{10} = 0.10\right)$$

$$P(x = k) = \binom{n}{k} * p^k * q^{n-k}$$

$$P(x = 1) = \binom{4}{1} * 0.10^1 * 0.90^3 = 0.2916$$

Respostra: A probabilidade de que a amostra contenha exactamente 1 defeituoso é de 29.2%.

74. ABC Association é uma empresa júnior de consultoria em contabilidade. A probabilidade de encontrar um cliente satisfeito com os seus serviços é de 80%. Em um determinado dia, contou com 6 clientes que procuravam seus serviços, determina a probabilidade de:

a) Mais de metade dos clientes estarem satisfeito com os serviços.

<u>**Resolução**</u>

Pelos dados temos:

$$p = 0.80; 1 - p = 0.20;\ n = 6; Pede - se\ P(x > 3)$$

$$onde\ v.a: x = nº\ de\ clientes\ satisfeitos\ com\ o\ serviço$$

Logo:

$$x \sim Bin(n, p)\ \ onde\ \ P(x = k) = \binom{n}{k} * p^k * q^{n-k}$$

Então:

$$P(x > 3) = P(x = 4) + P(x = 5) + P(x = 6)$$

$$P(x > 3) = \binom{6}{4} * 0.80^4 * 0.20^2 + \binom{6}{5} * 0.80^5 * 0.20^1 + \binom{6}{6} * 0.80^6 * 0.20^0 = 0.90112$$

Resposta: A probabilidade de mais de metade dos clientes estarem satisfeito com os serviços é de 90.1%.

b) No máximo 2 clientes não estarem satisfeitos com os serviços.

<u>**Resolução**</u>

Pelos dados temos:

$$p = 0.20; 1 - p = 0.80;\ n = 6; Pede - se\ P(x \leq 2)$$

$$onde\ v.a: x = nº\ de\ clientes\ insatisfeitos\ com\ o\ serviço$$

Logo:

$$x \sim Bin(n,p) \;\; onde \;\; P(x=k) = \binom{n}{k} * p^k * q^{n-k}$$

Então:

$$P(x \leq 2) = P(x=0) + P(x=1) + P(x=2)$$

$$P(x \leq 2) = \binom{6}{0} * 20^0 * 0.80^6 + \binom{6}{1} * 0.20^1 * 0.80^5 + \binom{6}{2} * 0.20^2 * 0.80^4 = 0.98304$$

Resposta: A probabilidade de no máximo 2 clientes não estarem satisfeitos com os serviços é de 98.3%.

75. O número médio das participações de corte de energia que chegam através da linha de cliente da EDM, vindo do bairro de Mafala na Cidade de Maputo é de 4, em um período de 1 minuto. Achar a probabilidade de que em 2 minutos sejam recebidas:

a) No máximo duas participações;

Resolução

$$X \sim Poisson\ (\lambda \mathrm{t});\ \mu = \lambda \mathrm{t};\ \lambda = 4; \mathrm{t} = 1 \text{ minuto}$$

Pelos dados temos:

$$\text{Aqui pede} - \text{se } \mathrm{P}(x \leq 2); \text{onde } \ \mathrm{t} = 2;\ \mu = 8$$

Pela fórmula de poisson teremos:

$$P(X=k) = \frac{\mu^K * e^{-\mu}}{k!}$$

$$\mathrm{P}(x \leq 2) = P(x=0) + P(x=1) + P(x=2)$$

$$\mathrm{P}(x \leq 2) = \frac{8^0 * e^{-8}}{0!} + \frac{8^1 * e^{-8}}{1!} + \frac{8^2 * e^{-8}}{2!}$$

$$\mathrm{P}(x \leq 2) = 0.0138$$

Resposta: A probabilidade de em 2 minutos sejam recebidas no máximo duas participações é de 1.38%.

b) Não menos de três participações.

Resolução

$$X \sim Poisson\ (\lambda \mathrm{t});\ \mu = \lambda \mathrm{t};\ \lambda = 4; \mathrm{t} = 1 \text{ minuto}$$

Pelos dados temos:

$$\text{Aqui pede} - \text{se } \mathrm{P}(x \geq 3); \text{onde } \ \mathrm{t} = 2;\ \mu = 8$$

Pela fórmula de poisson teremos:

$$P(X=k) = \frac{\mu^K * e^{-\mu}}{k!}$$

$$\mathrm{P}(x \geq 3) = 1 - P(x < 3) = 1 - [P(x=0) + P(x=1) + P(x=2)]$$

$$\mathrm{P}(x \geq 3) = 1 - \left[\frac{e^{-8} * 8^0}{0} + \frac{e^{-8} * 8^1}{1!} + \frac{e^{-8} * 8^2}{2!}\right]$$

$$\mathrm{P}(x \geq 3) = 1 - 0.0138 = 0.987$$

Resposta: A probabilidade de em 2 minutos sejam recebidas não menos de três participações é de 98.7%.

76. Uma loja vende, em média, 2,5 fogões por dia. Certo dia, ao encerrar o expediente, verifica-se existirem três fogões em estoque, e sabe-se que a nova remessa só chegará depois de dois dias. Qual a probabilidade de, no fim desses dois dias, a loja não ter deixado de atender, por falta de estoque, às pessoas que vierem comprar?

Resolução

$$X \sim Poisson\ (\lambda\mathrm{t});\ \mu = \lambda\mathrm{t};\ \lambda = 2.5; \mathrm{t} = 1\ \mathrm{dia}$$

Pelos dados temos:

$$\mathrm{Aqui\ pede - se\ P}(x \leq 3); \mathrm{onde}\ \ \mathrm{t} = 2;\ \mu = 5$$

Pela fórmula de poisson teremos:

$$P(X = k) = \frac{\mu^K * e^{-\mu}}{k!}$$

$$\mathrm{P}(x \leq 3) = P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3)$$

$$\mathrm{P}(x \leq 3) = \frac{5^0 * e^{-5}}{0!} + \frac{5^1 * e^{-5}}{1!} + \frac{5^2 * e^{-5}}{2!} + \frac{5^3 * e^{-5}}{3!}$$

$$\mathrm{P}(x \leq 3) = 0.4054$$

Resposta: A probabilidade de, no fim desses dois dias, a loja não ter deixado de atender, por falta de estoque, às pessoas que vierem comprar é de 40.54%.

77. Um fabricante de peças de automóveis garante que uma caixa de suas peças conterá, no máximo, 2 defeituosas. Se a caixa contém 8 peças e a experiência mostra que esse processo de fabricação produz 95% de peças perfeitas, qual a probabilidade de que uma caixa satisfaça a garantia?

Resolução

Pelos dados temos:

$$p = 0.05; 1 - p = 0.95;\ n = 18; Pede - se\ P(x \leq 2)$$

$$onde\ v.a{:}\ x = n^{\underline{o}}\ de\ peças\ defeituosas$$

Logo:

$$x \sim Bin(n, p)\ \ onde\ \ P(x = k) = \binom{n}{k} * p^k * q^{n-k}$$

Então:

$$P(x \leq 2) = P(x = 0) + P(x = 1) + P(x = 6)$$

$$P(x \leq 2) = \binom{8}{0} * 0.05^0 * 0.95^8 + \binom{8}{1} * 0.05^1 * 0.95^7 + \binom{8}{2} * 0.05^2 * 0.95^6 = 0.994$$

Resposta: A probabilidade de que uma caixa satisfaça a garantia é de 99.4%.

78. O número de pedidos de assistência em serviços agrários recebidos pela Direcção dos Serviços Agrários segue uma distribuição Poisson com uma média de 4 pedidos por hora.

a) Ache a probabilidade de exactamente 6 pedidos serem efectuados entre as 12 horas e as 13 horas da tarde.

Resolução

$$X \sim Poisson\ (\lambda t);\ \mu = \lambda t;\ \lambda = 4; t = 1\ \text{hora}$$

Pelos dados temos:

$$\text{Aqui pede} - \text{se } P(x = 6); \text{onde } t = 1;\ \mu = 4$$

Pela fórmula de poisson teremos:

$$P(X = k) = \frac{\mu^K * e^{-\mu}}{k!}$$

$$P(x = 6) = \frac{4^6 * e^{-4}}{6!} = 0.1042$$

Resposta: A probabilidade de exactamente 6 pedidos serem efectuados entre as 12 e as 13 horas da tarde é de 10.42%.

b) Qual é a probabilidade de serem recebidos nove ou menos pedidos de assistência entre as 12 horas e as 15 horas e 30 minutos da tarde?

Resolução

$$X \sim Poisson\ (\lambda t);\ \mu = \lambda t;\ \lambda = 4; t = 1\ hora$$

Pelos dados temos:

$$Aqui\ pede - se\ P(x \leq 9); onde\ \ t = 3.5\ horas;\ \mu = 14$$

Pela fórmula de poisson teremos:

$$P(X = k) = \frac{\mu^K * e^{-\mu}}{k!}$$

$$P(x \leq 9) = P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3) + P(x = 4) + P(x = 5) + P(x = 6) + P(x = 7) + P(x = 8) + P(x = 9)$$

$$P(x \leq 9) = e^{-14} * \left(\frac{14^0}{0!} + \frac{14^1}{1!} + \frac{14^2}{2!} + \frac{14^3}{3!} + \frac{14^4}{4!} + \frac{14^5}{5!} + \frac{14^6}{6!} + \frac{14^7}{7!} + \frac{14^8}{8!} + \frac{14^9}{9!}\right)$$

$$P(x \leq 9) = 0.1094$$

Resposta: A probabilidade de serem recebidos nove ou menos pedidos de assistência entre as 12 as 15:30 horas da tarde é de 10.94%.

79. Um graduado do ISCAM, possui 60% de chance de conseguir uma vaga de emprego quando concorre com os graduados das restantes instituições de ensino superior em moçambique. Se 8 candidatos a uma vaga de emprego, qual a probabilidade de:

a) Mais de 5 estudantes do ISCAM conseguirem vagas.

Resolução

Pelos dados temos:

$$p = 0.60; 1 - p = 0.40;\ n = 8; Pede - se\ P(x > 5)$$

$$onde\ v.a{:}\ x = nº\ de\ estudantes\ com\ vaga\ de\ emprego$$

Logo:

$$x \sim Bin(n, p)\ \ onde\ \ P(x = k) = \binom{n}{k} * p^k * q^{n-k}$$

Então:

$$P(x > 5) = P(x = 6) + P(x = 7) + P(x = 8)$$

$$P(x > 5) = \binom{8}{6} * 0.60^6 * 0.40^2 + \binom{8}{7} * 0.60^7 * 0.40^1 + \binom{8}{8} * 0.60^8 * 0.40^0 = 0.315395$$

Resposta: A probabilidade de 5 estudantes do ISCAM conseguirem vagas é de 31.5%.

b) No máximo dois estudantes do ISCAM não consigam vaga.

Resolução

Pelos dados temos:

$$p = 0.40; 1 - p = 0.60;\ n = 8; Pede - se\ P(x \leq 2)$$

$$onde\ v.a{:}\ x = nº\ de\ estudantes\ sem\ vaga\ de\ emprego$$

Logo:

$$x \sim Bin(n, p)\ \ onde\ \ P(x = k) = \binom{n}{k} * p^k * q^{n-k}$$

Então:

$$P(x \leq 2) = P(x = 0) + P(x = 1) + P(x = 2)$$

$$P(x \leq 2) = \binom{8}{0} * 0.40^0 * 0.60^8 + \binom{8}{1} * 0.40^1 * 0.60^7 + \binom{8}{2} * 0.40^2 * 0.60^6 = 0.315395$$

Resposta: A probabilidade de no máximo dois estudantes do ISCAM não consigam vaga é de 31.5%.

80. O número dos caros que passam pela portagem de Maputo segue uma distribuição de Poison com média de 12 carros em período de um minuto na hora de pico, e nas horas normais a media reduz para 1/3 em cada pista de leitura óptica. No meio do mês e nos períodos normais, chega a passar um máximo de 3 carros por minuto e a gerência da portagem reduz o numero de pistas para metade como forma de reduzir custo do seu funcionamento. Calcule a probabilidade de:

a) Mais de 2 carros passar na hora de pico, no final do mês em 1 minuto.

Resolução

$$X \sim Poisson\ (\lambda t);\ \mu = \lambda t;\ \lambda = 12; t = 1\ minuto$$

$$seja\ v.a{:} x = nº\ de\ carros\ que\ passam\ na\ hora\ do\ pico\ no\ final\ do\ mês\ em\ 1\ minuto$$

Pelos dados temos:

$$Aqui\ pede - se\ P(x \geq 2)$$

Pela fórmula de poisson teremos:

$$P(X = k) = \frac{\mu^K * e^{-\mu}}{k!}$$

$$\mathrm{P}(x > 2) = 1 - P(x \leq 2) = 1 - \{P(x = 0) + P(x = 1) + P(x = 2)\}$$

$$\mathrm{P}(x > 2) = 1 - e^{-12} * \left(\frac{12^0}{0!} + \frac{12^1}{1!} + \frac{12^2}{2!}\right)$$

$$\mathrm{P}(x > 2) = 0.9994$$

Resposta: A probabilidade de mais de 2 carros passar na hora de pico, no final do mês em 1 minuto é de 99.94%.

b) A gerência reduzir o numero de pista para metade.

Resolução

Aqui pede-se $P(x \leq 3)$

$$X \sim Poisson\ (\lambda t);\ \mu = \lambda t;\ \lambda = \frac{1}{3} * 12 = 46; t = 1\, minuto$$

Pela fórmula de poisson teremos:

$$P(X = k) = \frac{\mu^K * e^{-\mu}}{k!}$$

$$P(x \leq 3) = P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3)$$

$$P(x \leq 3) = e^{-4} * \left(\frac{4^0}{0!} + \frac{4^1}{1!} + \frac{4^2}{2!} + \frac{4^3}{3!}\right)$$

$$P(x \leq 3) = 0.4335$$

Resposta: A probabilidade da gerência reduzir o numero de pista para metade é de 43.35%.

81. Um laboratório farmacêutico produz seringas, das quais 0,5% são defeituosas. As seringas são vendidas em caixas com 20 unidades. Se a caixa tiver duas ou mais defeituosas o preço de venda é 1,00 USD; tendo uma, o preço é 2,50USD e não tendo defeituosa, o preço é 6,00USD. Qual o preço médio de uma caixa?

Resolução

$$Seja\ v.a\ x = nº\ de\ seringas\ defeitosas, onde\ P = 0.005$$

$$x \sim Bin\ (n = 20; P = 0.005)$$

$Pela\ fórmula: P(x = k) = \binom{n}{k} * p^k * q^{n-k}$ *temos:*

$$P(x = 0) = \binom{20}{0} * 0.005^0 * 0.995^{20} = 0.9046$$

$$P(x = 1) = \binom{20}{1} * 0.005^1 * 0.995^{19} = 0.0909$$

$$P(x \geq 2) = 1 - P(x < 2) = 1 - (0.9046 + 0.0909) = 0.0045$$

Determinando fdp da $v.a\ x = nº\ de\ seringas\ defeitosas$ temos:

x	0	1	≥ 2	Somatório
$P(x)$	0.9046	0.0909	0.0045	1

Seja agora a proposta do vendedor, onde $v.a\ y = preço\ da\ venda\ por\ caixa$, determinando o f.d.p teremos:

y	6	2.5	1	Somatório
$P(y)$	0.9046	0.0909	0.0045	1

$$E(y) = \sum y_i * P(y) = 6 * 0.9046 + 2.5 * 0.0909 + 1 * 0.0045 = 5.7$$

Logo

Resposta: O preço médio duma caixa será de 5.70 USD.

82. Um novo remédio tem efeito colateral indesejável em 5% das pessoas que o tomam. Se 16 pacientes tomam o remédio qual a probabilidade de:

a) Nenhuma reacção negativa?

Resolução

$$Seja\ v.a\ x = nº\ de\ seringas\ defeitosas, onde\ P = 0.05$$

$$x \sim Bin\ (n = 16; P = 0.05)$$

$Pela\ fórmula: P(x = k) = \binom{n}{k} * p^k * q^{n-k}$ temos:

$$P(x = 0) = \binom{16}{0} * 0.05^0 * 0.95^{16-0} = 0.44$$

Resposta: Se 16 pacientes tomam o remédio a probabilidade de nenhum deles ter reacção negativa é de 44%.

b) Uma reacção negativa?

Resolução

$$P(x = 1) = \binom{16}{1} * 0.05^1 * 0.95^{16-1} = 0.371$$

Resposta: Se 16 pacientes tomam o remédio a probabilidade de um deles ter reacção negativa é de 37.1%.

c) No máximo uma reacção negativa?

Resolução

$$P(x \leq 1) = P(x = 0) + P(x = 1) = 0.44 + 0.371 = 0.811$$

Resposta: Se 16 pacientes tomam o remédio a probabilidade de no máximo uma reacção negativa é de 81.1%.

d) No máximo uma reacção negativa?

Resolução

$$P(x \geq 1) = 1 - \{P(x = 0) + P(x = 1)\} = 1 - (0.44 + 0.371) = 18.9\%$$

83. Sabendo que o preço unitário de um certo produto é de 300,00 USD. Um supermercado faz a seguinte promoção: o cliente, ao passar pelo caixa, lança um dado. Se sair face 6 tem um desconto de 30% por produto. Se sair face 5 o desconto é de 20%. Se sair face 4 o desconto é de 10% e se ocorrerem faces 1, 2 ou 3, o desconto é de 5%. Seja X o preço final do produto com desconto.

a) Encontre a função de distribuição de probabilidade de X.

Resolução

Seja X = preço final do produto com desconto.

Como o dado é honesto, cada face tem probabilidade 1/6

Então a f.d.p sera dada por:

Nº da face	6	5	4	1,2 ou 3	
Desconto	30%	20%	10%	5%	Total
x	210	240	270	299.5	
$P\ (x)$	1/6	1/6	1/6	1/2	1

b) Calcule a probabilidade de que, num grupo de 5 clientes, pelo menos um consiga um preço maior que 270,00 USD.

Resolução

$$Seja\ v.a\ x = nº\ de\ clientes\ com\ preço\ superior\ a\ 270{,}00\ USD, onde\ P = \frac{1}{2}$$

$$x \sim Bin\ (n = 5; P = 0.05)$$

$$Pela\ fórmula: P(x = k) = \binom{n}{k} * p^k * q^{n-k} \text{ temos:}$$

$$P(x \geq 1) = 1 - P(x < 1) = 1 - \binom{5}{0} * 0.05^0 * 0.95^{5-0} = 0.96875$$

Resposta: A probabilidade de que, num grupo de 5 clientes, pelo menos um consiga um preço maior que 270,00 USD é de 96,9%.

c) Calcule a probabilidade de que o quarto cliente seja o primeiro a receber 30% de desconto.

Resolução

$$Seja\ v.a\ x = nº\ de\ clientes\ o\ que\ passam\ pelo\ caixa\ até\ primeiro\ desconto\ de\ 30\%$$

$$onde\ P = \frac{1}{6}$$

$$x \sim Geom\ \left(P = \frac{1}{6}\right)$$

$$Pela\ fórmula: P(x = k) = p^k * q^{n-1} \text{ temos:}$$

$$P(x = 4) = \frac{1}{6} * \left(\frac{5}{6}\right)^{4-1} = 0.096$$

Resposta: A probabilidade de que o quarto cliente seja o primeiro a receber 30% de desconto é de 9,6%.

84. As chegadas de petroleiros a uma refinaria em cada dia ocorrem segundo uma distribuição de Poisson, com parâmetro $\lambda = 2$. As atuais instalações podem atender, no máximo, a 3 petroleiros por dia. Se mais de 3 petroleiros chegarem num dia, o excesso é enviado a outro porto.

a) Em um dia, qual a probabilidade de se enviar petroleiros para outro porto?

Resolução

$$X \sim Poisson\ (\lambda t);\ \mu = \lambda t;\ \lambda = 2; t = 1\ dia$$

$$seja\ v.a: x = nº\ de\ petroleiros\ por\ dia$$

Para que sejam enviados petroleiros para outro porto mais de três petroleiros devem chegarem num dia. Então, pela fórmula de distribuição de posson temos:

$$Aqui\ pede - se\ P(x > 3)$$

Pela fórmula de poisson teremos:

$$P(X = k) = \frac{\mu^K * e^{-\mu}}{k!}$$

$$\mathrm{P}(x > 3) = 1 - P(x \leq 3) = 1 - \{P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3)\}$$

$$\mathrm{P}(x > 3) = 1 - e^{-2} * \left(\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!}\right)$$

$$\mathrm{P}(x > 3) = 0.143$$

Resposta: A probabilidade de se enviar petroleiros para outro porto em um dia é de 14.3%.

b) De quanto deverão ser aumentadas as instalações para permitir atender a todos os navios que chegarem em pelo menos 95% dos dias?

Resolução

Seja k a nova capacidade. Atender a essa capacidade significa que $x \leq k$. Como essa probabilidade tem que ser pelo menos 95%, temos que ter:

$$P(x \leq k) \geq 0,95$$

$x = k$	$P(x = k)$	$P(acumulada) = P(x \leq k$
0	0.135335283	0.135335283
1	0.270670566	0.406005850
2	0.270670566	0.676676416
3	0.180447044	0.857123460
4	0.090223522	0.947346983
5	0.036089409	0.983436392

Resposta: Pelas probabilidades acumuladas, o valor de k de modo que $P(x \leq k) \geq 0,95$ é de 5.

Logo: As instalações para permitir atenderem a todos os navios que chegarem em pelo menos 95% dos dias devem ser aumentadas para 5 petroleiros.

85. Entre os 16 programadores de uma empresa, 12 são do sexo masculino. A empresa decide sortear 5 programadores para fazer um curso avançado de programação. Qual é a probabilidade dos 5 sorteados serem do sexo masculino?

Resolução

Seja $v.a: x = nº\ de\ sorteados\ do\ sexo\ masculino$

$$Então, X \sim hiper(N = 16;\ r = 12; n = 5)$$

$$Pela\ fórmula:$$

$$P(X = k) = \frac{\binom{r}{k} * \binom{N-r}{n-k}}{\binom{N}{n}}$$

$$P(X = 5) = \frac{\binom{12}{5} * \binom{16-12}{5-5}}{\binom{16}{5}} = 0.181319$$

86. Deseja-se produzir 5 peças boas, em uma máquina que dá 20% de peças defeituosas. Qual é a probabilidade de ser necessário fabricar 8 peças para se conseguir as 5 peças boas?

Resolução

Seja $v.a: x = nº\ de\ peças\ fabricadas\ necessários\ para\ consiguir\ 5\ boas$

$$Então, X \sim BinNeg(r, p)$$

$$Pela\ fórmula: P(x = k) = \binom{k-1}{r-1} * p^r * q^{k-r}$$

$$onde\ p = 0.80\ e\ r = 5$$

$$P(x = 8) = \binom{8-1}{5-1} * 0.80^5 * 0.20^{8-5} = 0.0917504$$

87. Determinado tipo de parafuso é vendido em caixas com 50 peças. É uma característica da fabricação produzir 90% de boas. Qual é a probabilidade de serem necessárias 8 peças para encontrar a primeira defeituosa?

Resolução

Seja $v.a: x =$

$$nº\ de\ parafusos\ necessários\ até\ se\ encontrar\ o\ primeiro\ defeituoso$$

$$Então, x \sim Geom(p)$$

$$Pela\ fórmula: P(x = k) = p * q^{k-1}$$

$$onde\ p = 0.10\ e\ n = 8$$

$$P(x = 8) = 0.10 * 0.90^{8-1} = 0.048$$

DISTRIBUÍÇÃO DE PROBABILIDADES DE VARIÁVEL ALEATÓRIA CONTÍNUA

88. O consumo mensal em minutos por conta de celular em uma região é uma variável aleatória normal com média 36 e desvio padrão 12.

a) Qual é a probabilidade de uma pessoa desta região usar o telefone celular por menos de 48 minutos?

Resolução

$$X \sim N(\mu = 36; \sigma^2 = 144\,)$$

Padronizando a variável X e Z teremos:

$$Z \sim N(\mu = 0; \sigma^2 = 1\,)$$

Onde $Z = \frac{x-\mu}{\sigma}$

$$P(x < 48) = P\left(Z < \frac{48-36}{12}\right) = P(Z < 1) = 0.5 + P(0 < z < 1) = 0.5 + 0.3413$$
$$= 0.8413$$

Resposta: A probabilidade de uma pessoa desta região usar o telefone celular por menos de 48 minutos é de 84.1%.

b) Qual é a probabilidade de uma pessoa desta região usar o telefone celular por mais de 30 minutos?

Resolução

$$X \sim N(\mu = 36; \sigma^2 = 144\,)$$

Padronizando a variável X e Z teremos:

$$Z \sim N(\mu = 0; \sigma^2 = 1\,)$$

Onde $Z = \frac{x-\mu}{\sigma}$

$$P(x > 30) = P\left(Z > \frac{30-36}{12}\right) = P(Z > -0.333) = 0.5 + P(-0.333 < z < 0)$$
$$= 0.5 + 0.1293 = 0.6293$$

Resposta: A probabilidade de uma pessoa desta região usar o telefone celular por mais de 30 minutos é de 62.9%.

c) Qual o tempo mínimo que alguém deve gastar ao telefone no mês para estar entre os 5% que MAIS usam o celular?

Resolução

$$P(Z \geq z_0) \geq 0.95$$

$$P(Z \geq z_0) = 0.95 - 0.50 = 0.45$$

$$Z_{0.45} = 1.64$$

$$Logo$$

$$1.64 \leq \frac{x - 36}{12}$$

$$x \geq 1.64 * 12 + 36 = 48.2$$

Resposta: O tempo mínimo que alguém deve gastar ao telefone no mês para estar entre os 5% que mais usam o celular é de 48.2 minutos.

d) Qual o tempo máximo que alguém deve gastar ao telefone no mês para estar entre os 12% que menos usam o celular?

Resolução

$$P(Z \leq z_0) \geq 0.12$$

$$P(Z \leq z_0) = 0.50 - 0.12 = 0.38$$

$$-Z_{0.38} = 1.175$$

$$Logo$$

$$-1.175 \geq \frac{x - 36}{12}$$

$$x \leq -1.75 * 12 + 36 = 15$$

Resposta: O tempo máximo que alguém deve gastar ao telefone no mês para estar entre os 12% que menos usam o celular é de 15 minutos.

89. Uma fábrica de carros sabe que os motores de sua fabricação têm duração normal com média de 150000 km e desvio-padrão de 5000 km. Qual a probabilidade de que um carro, escolhido ao acaso, dos fabricados por essa firma, tenha um motor que dure:

a) Menos de 170000 km?

Resolução

$$Do\ problemas\ temos:$$

$$\mu = 50\ e\ \sigma = 1.6$$

$$x \sim N(\mu = 150000;\ \sigma^2 = 5000^2) \rightarrow \quad Z \sim N(\mu = 0;\ \sigma = 1), onde\ Z = \frac{x - \mu}{\sigma}$$

$$P(X < 170000) = P\left(Z < \frac{170000 - 150000}{5000}\right) = P(Z < 4)$$

$$= 0.5 + P(0 \leq Z \leq 4) = 0.5 + 0.49998 = 0.99998$$

Resposta: *A probabilidade de que um carro, escolhido ao acaso, dos fabricados por essa firma, tenha um motor que dure menos de 170000 km é de aproximadamente 100%.*

b) Entre 140000 km e 165000 km?

Resolução

$$P(140000 < X < 165000) = P\left(\frac{140000 - 150000}{5000} < Z < \frac{165000 - 150000}{5000}\right)$$
$$= P(-2 < Z < 3) = P(-2 < Z < 0) + P(0 < Z < 3)$$
$$= 0.4772 + 0.4987$$
$$= 0.9759$$

Resposta: *A probabilidade de que um carro, escolhido ao acaso, dos fabricados por essa firma, tenha um motor que dure entre 140000 e 1650000 km é de 97.6%.*

c) Se a fábrica substitui o motor que apresenta duração inferior à garantia, qual deve ser esta garantia para que a percentagem de motores substituídos seja inferior a 0,2%?

<u>Resolução</u>

$$P(X < x) = 0.002$$
$$Z_{0.002} = -2.28$$
$$\frac{x - 150000}{5000} = -2.88$$
$$x = 150000 - 2.88 * 5000 = 135600$$

Resposta: A garantia para que a percentagem de motores substituídos seja inferior a 0,2% é de 135600 Km.

90. Seja X uma variável aleatória exponencial com média 4. Calcule

a) $P(x > 1)$

Solução

$$P(x > 1) = \int_1^{\infty} \frac{1}{4} * e^{-\frac{x}{4}} dx = 0 - (e^{-0.25}) = 0.7788$$

ou pelo uso directo da função de distribuição

$$P(x > 1) = 1 - P(x \le 1) = 1 - F(1) = 1 - \left(1 - e^{-\frac{1}{4}}\right) = 0.7788$$

b) $P(1 \le x \le 2)$

Solução

$$P(1 \le x \le 2) = P(x \le 2) - P(x < 1) = P(x \le 2) - P(x \le 1)$$
$$= F(2) - F(1) = \left(1 - e^{-\frac{2}{4}}\right) - \left(1 - e^{-\frac{1}{4}}\right) = 0.17227$$

91. Latas de coca-cola são enchidas num processo automático segundo uma distribuição uniforme no intervalo (em ml) [345;355].

a) Qual é a probabilidade de uma lata conter mais de 353 ml?

<u>Resolução</u>

$$Seja\ v.a\ X = \text{“conteúdo da lata de coca} - \text{cola”. Então,} X \sim U[345;\ 355]$$

$$Pela\ fórmula\ de\ distribuíção\ uniforme\ temos:$$

$$P(a \leq x \leq b) = \int_a^b f_x(x)dx$$

Onde

$$f_x(x) = \frac{1}{b-a}$$

$$P(x > 353) = \int_{353}^{355} f_x(x)dx = \frac{355-353}{355-345} = 0.2$$

b) Qual é a probabilidade de uma lata conter menos de 346 ml?

Resolução

$$P(x < 346) = \int_{345}^{346} f_x(x)dx = \frac{346-345}{355-345} = 0.1$$

c) Qualquer lata com volume 4 ml abaixo da média pode gerar reclamação do consumidor e com volume 4 ml acima da média pode transbordar no momento de abertura, devido à pressão interna. Qual é a proporção de latas problemáticas?

Resolução

$$P(x < 350-4) + P(x > 350+4) = \int_{345}^{346} f_x(x)dx + \int_{354}^{355} f_x(x)dx$$

$$= \frac{346-345}{355-345} + \frac{355-354}{355-345} = 0.2$$

92. O tempo necessário para o atendimento de uma pessoa em um escritório de um Banco tem distribuição aproximadamente normal com média de 130 segundos com um desvio padrão igual a 45 segundos. Qual é a probabilidade de que um indivíduo aleatoriamente escolhido:

a) Leve menos de 100 segundos para terminar o atendimento?

Resolução

$$Do\ problemas\ temos:$$

$$\mu = 130\ e\ \sigma = 45$$

$$x \sim N(\mu = 130;\ \sigma^2 = 45^2) \rightarrow \quad Z \sim N(\mu = 0;\ \sigma = 1), onde\ Z = \frac{x-\mu}{\sigma}$$

$$P(X < 100) = P\left(Z < \frac{100-130}{45}\right) = P(Z < -0.67)$$

$$= 0.5 - P(-0.67 \leq Z \leq 0) = 0.2514$$

Resposta: A probabilidade de que um indivíduo leve menos de 100 segundos para terminar o atendimento .

b) Leve entre 2 e 3 minutos no escritório?

Resolução

$$P(2 < X < 100) = P\left(\frac{120-130}{45} Z < \frac{180-130}{45}\right) = P(-0.22 < Z < 1.11)$$
$$= 0.8665 - 0.4129 = 0.4536$$

***Resposta:** A probabilidade de que um indivíduo leve entre 2 e 3 minutos segundos para terminar o atendimento é de 45.4%.*

c) Leve mais de 200 segundos.

Resolução

$$P(X < 100) = P\left(Z > \frac{200-130}{45}\right) = P(Z > 1.56)$$
$$= 0.5 - P(0 \leq Z \leq 1.56) = 0.0594$$

***Resposta:** A probabilidade de que um indivíduo leve mais de 200 segundos para terminar o atendimento .*

93. A ABC association, é um escritório de consultoria em contabilidade e auditoria que tem como receita média mensal de USD 1000 com desvio padrão de USD 150. Para realização dos das suas actividades o escritório tem um custo médio de USD 600 mensais. Em tempos de crise económica, o escritório não vai mais do que USD 700 de receitas. Sabendo que as receitas seguem uma distribuição normal determine.

a) A probabilidade das receitas em um mês de crise económica?

Resolução

$$v.a: X \sim N(\mu = 1000;\ \sigma^2 = 150^2)$$

Aqui pede-se $P(x \leq 700)$

$$Z = \frac{x-\mu}{\sigma}$$

$$P(x \leq 700) = P\left(Z \leq \frac{700-1000}{150}\right) = P(Z \leq -2) = 0.5 - 0.47725 = 0.02275$$

Resposta: A probabilidade das receitas em um mês de crise económica é de 2.3%

b) A probabilidade de que em certo mês haja prejuízo?

Resolução

$$P(x < 600) = P\left(Z \leq \frac{600-1000}{150}\right) = P(Z \leq -2.67) = 0.5 - 0.49621 = 0.00379$$

Resposta: A probabilidade de que em certo mês haja prejuízo é de 0.4%.

c) Qual é a probabilidade das vendas excederem USD 700 em determinado mês.

Resolução

$$P(x > 700) = P\left(Z > \frac{700-1000}{150}\right) = P(Z > -2) = 0.5 + 0.49621 = 0.99621$$

Resposta: A probabilidade das vendas excederem USD 700 em determinado mês é de 99.6%

94. Uma empresa produz televisores e garante a restituição da quantia paga se qualquer televisor apresentar algum defeito grave no prazo de 6 meses. Ela produz televisores do tipo A comum e do tipo B de luxo, com um lucro respectivo de 1100 u.m. e 2100 u.m. caso não haja restituição, e com prejuízo de 3100 u.m. e 8100 u.m., se houver restituição. Suponha que o tempo para ocorrência de algum defeito grave seja, em ambos os casos, uma variável aleatória com distribuição normal com médias 9 meses e 12 meses e desvios padrões 2 meses e 3 meses. Se tivesse que planear uma estratégia de marketing para a empresa, você incentivaria as vendas dos aparelhos tipo A ou tipo B?

Resolução

Televisor tipo A

$$\mu = 9;\ \sigma = 2$$

Se $X \sim N(\mu; \sigma^2)$, então:

$$Z = \frac{X - \mu}{\sigma}$$

Determinando a probabilidade de haver restituíção temos:

$$P(x \leq 6) = P\left(Z \leq \frac{6-9}{2}\right) = P(Z \leq -1.5) = 0.5 - P(-1.5 \leq Z \leq 0)$$

$$P(x \leq 6) = 0.5 - 0.43319 = 0.06681$$

Logo O Lucro médio será:

$$E(L) = 0.06681 * 3100 + (1 - 0.06681) * 1100 = 1233.62$$

Televisor tipo B

$$\mu = 12;\ \sigma = 3$$

Se $X \sim N(\mu; \sigma^2)$, então:

$$Z = \frac{X - \mu}{\sigma}$$

Determinando a probabilidade de haver restituíção temos:

$$P(x \leq 6) = P\left(Z \leq \frac{6-12}{3}\right) = P(Z \leq -2.0) = 0.5 - P(-2 \leq Z \leq 0)$$

$$P(x \leq 6) = 0.5 - 0.47725 = 0.02275$$

Logo O Lucro médio será:

$$E(L) = 0.02275 * 8100 + (1 - 0.02275) * 2100 = 2236.5$$

Resposta: A empresa deveria incentivar as vendas dos aparelhos do tipo B, pois o lucro médio de B é maior que o lucro médio de A.

95. O saldo médio dos clientes de um banco é uma v.a. normal com média 3400, 00 u.m e variança 67600,00 u.m. Os clientes com os 5% maiores saldos médios recebem tratamento VIP, enquanto aqueles com os 10% menores saldos médios receberão propaganda extra para estimular maior movimentação da conta.

   a) Quanto um cliente precisa de saldo médio para se tornar um VIP?

Resolução

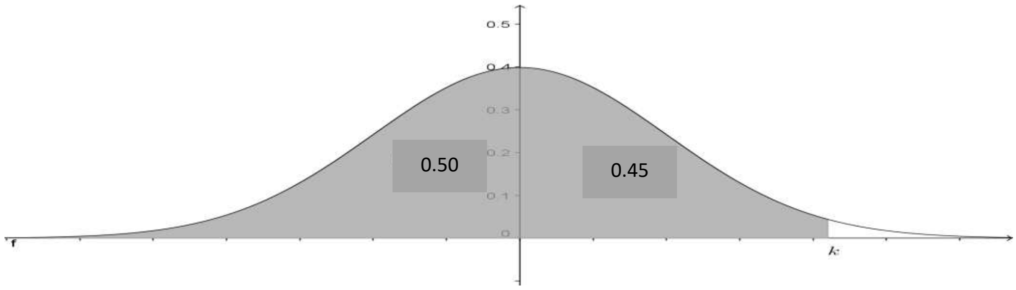
$$Pelo\ dados\ temos: \mu = 3400;\ \sigma^2 = 67600$$

Para encontrar o valor mínimo deve-se determinar o valor de k tal que:

$$P(x \leq k) = 0.95$$

Nota que queremos encontrar a abcissa k da normal padrão com 0,95 de área (probabilidade) à esquerda dela.

Graficamente temos:



Com a probabilidade à esquerda de k é maior que 0,5, resulta que k tem de ser maior que a média. O primeiro passo na solução é escrever a probabilidade dada em termos da normal padrão.

$$P(X \leq k) = 0.95$$

Se $X \sim N(\mu; \sigma^2)$, *então:*

$$Z = \frac{X - \mu}{\sigma}$$

Logo:

$$P\left(\frac{X - 3400}{260} \leq \frac{k - 3400}{260}\right) = 0.95$$

$$P\left(Z \leq \frac{k - 3400}{260}\right) = 0.95$$

$$P(Z \leq 0) + P\left(0 \leq Z \leq \frac{k - 3400}{260}\right) = 0.95$$

$$0.5 + P\left(0 \leq Z \leq \frac{k - 3400}{260}\right) = 0.95$$

$$tab\left(\frac{k - 3400}{260}\right) = 0.45$$

$$\frac{k - 3400}{260} = 1.645$$

$$k = 3827.7$$

Resposta: Para um cliente se tornar um VIP, precisa de um saldo mínimo de 3827.7 u.m.

b) Abaixo de qual saldo médio o cliente receberá a propaganda extra?

Resolução

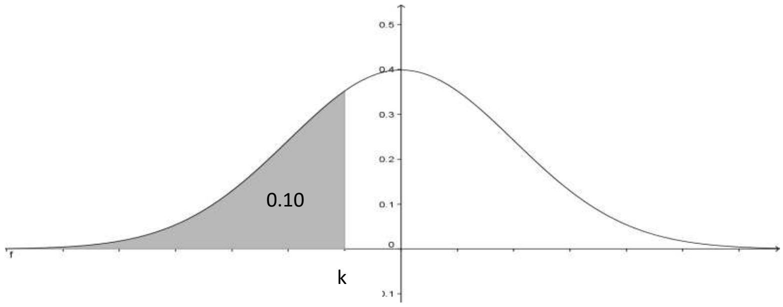
$$Pelo\ dados\ temos: \mu = 3400;\ \sigma^2 = 67600$$

Para encontrar o valor máximo deve-se determinar o valor de k tal que:

$$P(x \leq k) = 0.10$$

Nota que queremos encontrar a abcissa k da normal padrão com 0.10 de área (probabilidade) à esquerda dela.

Graficamente temos:



Com a probabilidade à esquerda de k é menrr que 0,5, resulta que k tem de ser menor que a média. O primeiro passo na solução é escrever a probabilidade dada em termos da normal padrão.

$$P(X \leq k) = 0.10$$

Se $X \sim N(\mu; \sigma^2)$, *então:*

$$Z = \frac{X - \mu}{\sigma}$$

Logo:

$$P\left(\frac{X - 3400}{260} \leq \frac{k - 3400}{260}\right) = 0.10$$

$$P\left(Z \leq \frac{k - 3400}{260}\right) = 0.10$$

$$P(Z \leq 0) - P\left(0 \leq Z \leq \frac{k - 3400}{260}\right) = 0.10$$

$$0.5 - P\left(0 \leq Z \leq \frac{k - 3400}{260}\right) = 0.10$$

$$tab\left(\frac{k - 3400}{260}\right) = 0.40$$

$$\frac{k - 3400}{260} = -1.28$$

$$k = 3067.2$$

Resposta: O cliente só receberá a propaganda extra se o saldo for inferior a 3067.2 u.m.

96. Suponha que o tempo necessário para atendimento de clientes em uma central de atendimento telefônico siga uma distribuição normal de média de 8 minutos e variança de 4 minutos.

a) Qual é a probabilidade de que um atendimento dure menos de 5 minutos?

Resolução

$$\mu = 8;\ \sigma = 2$$

Se $X \sim N(\mu; \sigma^2)$, então:

$$Z = \frac{X - \mu}{\sigma}$$

Aqui pede-se $P(x \leq 5)$

$$P(x \leq 5) = P\left(Z \leq \frac{5-8}{2}\right) = P(Z \leq -1.5) = 0.5 - P(-1.5 \leq Z \leq 0)$$

$$P(x \leq 5) = 0.5 - 0.43319 = 0.06681$$

Resposta: A probabilidade de que um atendimento dure menos de 5 minutos é de 6.7%.

b) E entre 7 e 10 minutos?

Resolução

$$\mu = 8;\ \sigma = 2$$

Se $X \sim N(\mu; \sigma^2)$, então:

$$Z = \frac{X - \mu}{\sigma}$$

Aqui pede-se $P(7 \leq x \leq 10)$

$$P(7 \leq x \leq 10) = P\left(\frac{7-8}{2} \leq Z \leq \frac{10-8}{2}\right) = P(-0.5 \leq Z \leq -1.0) = 0.19146 + 0.34134$$

$$= 0.5328$$

Resposta: A probabilidade de que um atendimento dure entre 7 e10 minutos é de 53.28%.

c) 75% das chamadas telefônicas requerem pelo menos quanto tempo de atendimento?

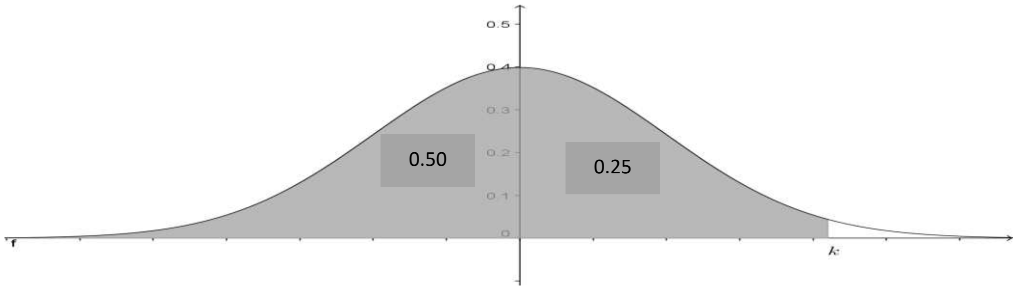
Resolução

$$Pelo\ dados\ temos: \mu = 8;\ \sigma^2 = 4$$

Para encontrar o valor mínimo deve-se determinar o valor de k tal que:

$$P(x \geq k) = 0.25\ ou\ \ P(x \leq k) = 0.75$$

Graficamente temos:



Com a probabilidade à esquerda de k é maior que 0,5, resulta que k tem de ser maior que a média. O primeiro passo na solução é escrever a probabilidade dada em termos da normal padrão.

$$P(X \leq k) = 0.75$$

Se $X \sim N(\mu; \sigma^2)$, *então:*

$$Z = \frac{X-\mu}{\sigma}$$

Logo:

$$P\left(\frac{X-8}{2} \leq \frac{k-8}{2}\right) = 0.75$$

$$P\left(Z \leq \frac{k-8}{2}\right) = 0.75$$

$$P(Z \leq 0) + P\left(0 \leq Z \leq \frac{k-8}{2}\right) = 0.75$$

$$0.5 + P\left(0 \leq Z \leq \frac{k-8}{2}\right) = 0.75$$

$$tab\left(\frac{k-8}{2}\right) = 0.25$$

$$\frac{k-8}{2} = 0.67$$

$$k = 9.34$$

Resposta: 75% das chamadas telefônicas requerem pelo menos 9 minutos e 34 segundos atendimento

97. Uma clínica de emagrecimento recebe pacientes adultos com peso $N(130\ Kg;\ 400\ Kg^2)$. Para efeito de determinar o tratamento mais adequado, os 25% pacientes de menor peso são classificados de "magros", enquanto os 25% de maior peso de "obesos". Determinar os pesos que delimitam cada classe.

Resolução

$$X \sim N(\mu = 130; \sigma^2 = 400\ )$$

Padronizando a variável X e Z teremos:

$$Z \sim N(\mu = 0; \sigma^2 = 1\ )$$

Onde $Z = \frac{x-\mu}{\sigma}$

Para os classificados como magros

$$P(Z \leq z_0) \geq 0.25$$

$$P(Z \leq z_0) = 0.50 - 0.25 = 0.0.25$$

$$-Z_{0.25} = 0.67$$

$$Logo$$

$$-0.67 \geq \frac{x - 130}{20}$$

$$x \leq -0.67 * 20 + 130 = 116.6$$

Para os classificados como obesos

$$P(Z \geq z_0) \geq 0.75$$

$$P(Z \geq z_0) = 0.75 - 0.50 = 0.25$$

$$Z_{0.25} = 0.67$$

$$Logo$$

$$0.67 \leq \frac{x - 130}{20}$$

$$x \geq 0.67 * 20 + 130 = 143.4$$

Resposta: Resposta: Os pesos que delimitam cada classe são 116.6Kg (classificados como magros) e 143.4Kg (classificados como obesos)

98. Em indivíduos sadios, o consumo geral de oxigénio tem distribuição normal com média 12 cm3/min e desvio padrão 2 cm3/min. Determine a proporção de indivíduos sadios com consumo:

a) Inferior a 10 cm^3/min.

Resolução

$$v.a: X \sim N(\mu = 12;\ \sigma^2 = 2^2)$$

Aqui pede-se $P(x < 10)$

$$Z = \frac{x - \mu}{\sigma}$$

$$P(x < 10) = P\left(Z \leq \frac{10 - 12}{2}\right) = P(Z \leq -1) = 0.5 - 0.3413 = 0.1587$$

b) Superior a 15 cm^3/min.

Resolução

$$P(x > 15) = P\left(Z > \frac{15 - 12}{2}\right) = P(Z > 1.5) = 0.5 - 0.4332 = 0.9104$$

c) Entre 8 cm^3/min e 15 cm^3/min

Resolução

$$P(8 \leq x \leq 15) = P\left(\frac{8-8}{2} \leq Z \leq \frac{15-8}{2}\right) = P(-2 \leq Z \leq -1.5) = 0.4772 + 0.4332$$
$$= 0.9104$$

99. Você está interessado em dar uma proposta em um leilão de um lote de terra. Você sabe que existe um outro licitante. Pelas regras estabelecidas para este leilão, a proposta mais alta acima de 100.000,00 u.m será aceite. Suponha que a proposta do seu competidor seja uma variável aleatória uniformemente distribuída entre 100.000,00 u.m e 150.000,00 u.m. Supondo que tens dois conselheiros identificados por A e B. O conselheiro A propõe lhe uma proposta de 120.000,00 u.m, enquanto que o B uma proposta de 140.000,00 u.m. Você optaria por proposta do conselheiro A ou B?

Resolução

$$Seja\ v.a\ X = \text{“}conteúdo\ da\ lata\ de\ coca - cola\text{”}. Então, X \sim U[100000;\ 150000]$$

Pela forma de distribuíção uniforme temos:

$$P(a \leq x \leq b) = \int_a^b f_x(x)dx$$

Onde

$$f_x(x) = \frac{1}{b-a}$$

Conselheiro A

$$P(x = 120000) = P(100000 \leq x \leq 120000) = \int_{100000}^{120000} f_x(x)dx = \frac{120000 - 100000}{150000 - 100000} = 0.4$$

Conselheiro B

$$P(x = 140000) = P(100000 \leq x \leq 140000) = \int_{100000}^{140000} f_x(x)dx = \frac{140000 - 100000}{150000 - 100000} = 0.8$$

Resposta: Optaria por B, pois a probabilidade de ficar com o lote é duas vezes maior que A.

ESTIMAÇÃO PONTUAL

100. Uma companhia fabrica cilindros que tem uma média de 2 polegadas de diâmetro. O desvio padrão dos diâmetros dos cilindros é de 10 polegadas. Os diâmetros de uma amostra de 4 cilindros são medidos todas as horas. A média amostral é usada para decidir se o processo de fabricação está operando satisfatoriamente ou não. A seguinte regra de decisão é aplicada: se diâmetro médio da amostra de 4 cilindros é maior ou igual a 2,15 polegadas, ou menor ou igual a 1,85 polegadas, interrompe-se o processo. Qual é a probabilidade de parar o processo se a média do processo permanece constante no valor de 2,00 polegadas?

Resolução

Pelos dados temos:

$$\mu = 2; \sigma = 10; n = 4;\ \bar{x}_1 = 2.15;\ \bar{x}_2 = 1.85$$

Nota que só para se:

$$\bar{x}_1 \geq 2.15\ ou\ \bar{x}_2 \leq 1.85$$

$$\bar{x} \sim N\left(\mu; \frac{\sigma^2}{n}\right) \to Z \sim N(0;1)\ Onde\ Z = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$$

Logo

$$P(\bar{x}_2 \leq 1.85) + P(\bar{x}_1 \geq 2.15) = P\left(Z \leq \frac{1.85-2}{10/\sqrt{4}}\right) + P\left(Z \geq \frac{2.15-2}{10/\sqrt{4}}\right)$$

$$= P(Z \leq -0.03) + P(Z \geq 0.03) = 0.5 - 2 * P(0 \leq Z \leq 0.3)$$

$$= 0.5 - 2 * 0.01197 = 0.47606$$

Resposta: A probabilidade de parar o processo se a média do processo permanece constante no valor de 2,00 polegadas é dse 47.6%

101. Numa linha de engarrafamento de azeite a quantidade deitada em cada garrafa é uma variável aleatória que se admite ter distribuição normal. O processo de enchimento considera se regulado se a garrafa conter exactamente 1 litro, não sendo de admitir grandes desvios. Para controlar processo de enchimento escolheram se ao acaso 60 garrafas da produção diária. Suponha que se obteve um desvio padrão de 0.08 litros, qual é a probabilidade de se obter uma média igual ou inferior a 0.965 litros?

Resolução

Do problema temos:

$$\mu = 1;\ n = 60;\ s = 0.08;\ \bar{x} = 0.965$$

Pede-se $P(\bar{x} \leq 0.965)$

$$\bar{x} \sim N\left(\mu; \frac{\sigma^2}{n}\right) \rightarrow Z \sim N(0;1)\ Onde\ Z = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

Então:

$$P(\bar{x} \leq 0.965) = P\left(Z \leq \frac{0.965 - 1}{0.08/\sqrt{60}}\right) = P(Z \leq -3.39) = 0.5 - P(-3.39 \leq Z \leq 0)$$

$$= 0.5 - 0.49965 = 0.00035$$

Resposta: A probabilidade de se obter uma média igual ou inferior a 0.965 litros 0.035%.

102. Suponha que 40% de todos os clientes dum banco possuem várias contas. Caso seja seleccionada uma amostra de 200 clientes, qual é a probabilidade de que a proporção amostral de clientes com várias contas fique entre 0.40 e 0.43?

Resolução

$$Seja\ v.a: nº\ de\ clientes\ com\ várias\ contas$$

Do problema temos:

$$\pi = 0.40; n = 200; Pede - se\ P(0.40 \leq p \leq 0.43)$$

Sabe se que:

$$p \sim N\left(\pi; \frac{\pi * (1 - \pi)}{n}\right)$$

Padronizando teremos:

$$Z \sim N(0;1)$$

Onde

$$Z = \frac{p - \pi}{\sqrt{\frac{\pi * (1 - \pi)}{n}}}$$

Então:

$$P(0.40 \leq p \leq 0.43) = P\left(\frac{0.40 - 0.40}{\sqrt{\frac{0.40 * 0.60}{200}}} \leq Z \leq \frac{0.43 - 0.40}{\sqrt{\frac{0.40 * 0.60}{200}}}\right)$$

$$P(0 \leq Z \leq 0.87) = 0.30785$$

Resposta: A probabilidade de que em uma amostra com 200 clientes com várias contas fique entre 0.40 e 0.43 é de 30.785%.

103. O gerente de marketing de uma fábrica de automóveis está interessado em determinar a proporção de novos proprietários de carros compactos que teriam adquirido um air-bag inflável para o lado do passageiro se o mesmo estivesse disponível a um custo adicional de 300,00 u.m. Por informações anteriores, o gerente acredita que a proporção é 30%. Suponha que é feito um levantamento com 200 novos proprietários de carros compactos, qual é a probabilidade de pelo menos 79 proprietários indicarem que teriam comprado os air-bags infláveis?

Resolução

$$Seja\ v.a: nº\ de\ proprietários\ favoráveis\ a\ compra\ de\ air - bags\ infláveis$$

Do problema temos:

$$\pi = 0.30; n = 200; p = \frac{x}{n} = \frac{79}{200} = 0.395;\ Pede - se\ P(p \geq 0.395)$$

Sabe se que:

$$p \sim N\left(\pi; \frac{\pi * (1-\pi)}{n}\right)$$

Padronizando teremos:

$$Z \sim N(0; 1)$$

Onde

$$Z = \frac{p - \pi}{\sqrt{\frac{\pi * (1-\pi)}{n}}}$$

Então:

$$P(p \geq 0.395) \quad = P\left(Z \geq \frac{0.395 - 0.30}{\sqrt{\frac{0.30 * 0.70}{200}}}\right)$$

$$P(Z \geq 2.93) = 0.5 - P(0 \leq Z \leq 2.93) = 0.5 - 0.49831$$

Resposta: A probabilidade de que em uma amostra com 200 proprietários, pelo menos 79 que teriam comprado os air-bags infláveis é de 0.169%.

104. Um fornecedor apresenta uma caixa, e afirma que o peso médio desta caixa é de 368 gramas. De experiências anteriores sabe-se que o desvio padrão vale 15 gramas e que os valores se comportam segundo a distribuição Normal. Para verificar se a afirmação é verdadeira, verifica-se uma amostra de 25 caixas. Qual é a probabilidade de o peso médio da amostra ser inferior a 372,5 gramas?

Resolução

Do problema temos:

$$\mu = 368;\ \sigma = 15;\ n = 25;\ ; Pede - se\ P(\bar{x} \leq 372.5)$$

Sabe se que:

$$\bar{x} \sim N\left(\mu; \frac{\sigma^2}{n}\right) \rightarrow Z \sim N(0; 1)\ Onde\ Z = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

$$P(\bar{x} \leq 372.5) = P\left(Z \leq \frac{372.5 - 368}{\frac{15}{\sqrt{25}}}\right) = P(Z \leq 1.5) = 0.5 + P(0 \leq Z \leq 1.5)$$

$$= 0.5 + 0.43319 = 0.93319$$

Resposta: A probabilidade de o peso médio da amostra ser inferior a 372,5 gramas é de 93.3%.

105. O proprietário de uma empresa de venda de produtos metálicos, está preocupado com o tempo gasto pelos operários estagiários no processamento de ferro. O proprietário decide submeter os estagiários a um teste de aptidão para a execução da actividade. Para passar no teste, o operário deve completá-lo em, no máximo, 80 minutos. Admita que o tempo, em minutos, para completar a prova seja uma variável aleatória normal com média 90 minutos e desvio padrão 22 minutos. Selecciona-se aleatoriamente 28 operários, qual é a chance de serem reprovados?

Resolução

Pelos dados temos:

$$\sigma = 22;\ \mu = 90; n = 28;\ Pede - se\ o\ P(\bar{x} > 80)$$

$$Sabe - se\ que:$$

$$\bar{x} \sim N\left(\mu;\ \frac{\sigma^2}{n}\right) \rightarrow\ Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

$$P(\bar{x} > 80) = P\left(Z > \frac{80 - 90}{\frac{22}{\sqrt{28}}}\right) = P(Z > -2.41\ ) = 0.5 + P(-2.41 \leq Z \leq 0\ )$$
$$= 0.5 + 0.49202 = 0.99202$$

Resposta: A chance de 28 operários serem reprovados é de 99.2%.

106. Numa certa cidade, a duração de conversas telefónicas em minutos, segue um modelo de Poisson com parâmetro 3. Observando se uma amostra aleatória de 50 dessas chamadas, qual será a probabilidade de em média, a duração de conversas telefónicas não ultrapassarem 4 minutos.

Resolução

$$x\ {\sim}Poisson\ (\lambda t)\ \rightarrow \begin{cases} E(x) = \mu = \lambda t \\ Var(x) = \mu = \lambda t \end{cases}$$

Pelos dados temos:

$$\lambda = 3; \sigma = 3;\ \mu = 3; n = 50;\ Pede - se\ o\ P(\bar{x} \leq 4)$$

$$Sabe - se\ que:$$

$$\bar{x} \sim N\left(\mu;\ \frac{\sigma^2}{n}\right) \rightarrow\ Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

$$P(\bar{x} \leq 4) = P\left(Z \leq \frac{4 - 3}{\frac{3}{\sqrt{50}}}\right) = P(Z \leq 2.36\ ) = 0.5 + P(0 \leq Z \leq 2.36\ ) = 0.5 + 0.49086$$
$$= 0.99086$$

Resposta: A probabilidade de em média, a duração de conversas telefónicas não ultrapassarem 4 minutos 99.1%.

INTERVALO DE CONFIANÇA

107. Para uma amostra de 50 firmas tomada de uma determinada indústria, o número médio de empregados por firma é 420.4 com um desvio padrão de 55.7. Nesta indústria, há um total de 380 firmas.

a) Determinar o intervalo de confiança de 90%, para estimar o número médio de trabalhadores por firma da indústria.

Resolução:

Do problema se tem:

$$N = 380;\ n = 50;\ S = 55.7;\ \bar{\boldsymbol{x}} = 420.4;\ (1 - \alpha).100\% \ = \ 90\%\ onde\ \alpha = 0.10$$

Pelo teorema de limite central, como:

$$n \geq 30\ e\ \frac{n}{N} \geq 0.05\ então\ \bar{X} \sim N\left(\mu;\ \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)\right)$$

$$onde\ a\ variável\ padronizada\ de\ \bar{x}\ será\ Z\ e\ Z_{\frac{\alpha}{2}} = Z_{0.05} = 1.645$$

Lembre-se que para descobrir a abcissa 1.645, entrou-se na tabela de distribuição normal com $0.45 = \ 45\%$*, já que a tabela è da faixa central.*

A fórmula geral do intervalo de confiança para média será:

$$\bar{x} - Z_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}} \leq \mu \leq \bar{x} + Z_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}}$$

Substituindo os dados pela fórmula temos:

$$420.4 - 1.645 * \frac{55.7}{\sqrt{50}} / \sqrt{\frac{380-50}{380-1}} \leq \mu \leq 420.4 + 1.645 * \frac{55.7}{\sqrt{50}} \sqrt{\frac{380-50}{380-1}}$$

$$408.3 \leq \mu \leq 432.5$$

Interpretação: A um nível de confiança de 90% *pode se afirmar que o intervalo* $[408.3; 432.5]$ *contêm o verdadeiro número médio de trabalhadores por firma da indústria.*

b) Qual a amplitude do intervalo de confiança (a 99% de confiança) ao estimar μ por $\bar{x} = 420.4$?

Resolução:

$$(1 - \alpha).100\% \ = \ 99\%;\ \alpha = 0.01\ onde\ Z_{\frac{\alpha}{2}} = Z_{0.005} = 2.576$$

$$420.4 - 2.576 * \frac{55.7}{\sqrt{50}} * \sqrt{\frac{380-50}{380-1}} \leq \mu \leq 420.4 + 2.576 * \frac{55.7}{\sqrt{50}} * \sqrt{\frac{380-50}{380-1}}$$

$$401.5 \leq \mu \leq 439.3$$

Logo

$$(At) = lim.sup - lim.inf = 439.3 - 401.5 = 37.8$$

Logo a amplitude do intervalo de confiança será de 37.8.

c) Qual seria o erro absoluto de estimação obtido para estimar o número médio de trabalhadores por firma da indústria a um nível de confiança de 95%?

<u>Resolução:</u>

$$(1 - \alpha).100\% = 95\%;\ \alpha = 0.05\ onde\ Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$$

Recorrendo a fórmula para cálculo do erro (ε) *através do intervalo de confiança terremos:*

$$\varepsilon = Z_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}}$$

$$\varepsilon = 1.96 * \frac{55.7}{\sqrt{50}} * \sqrt{\frac{380-50}{380-1}}$$

$$\varepsilon = 14.4$$

Logo o erro absoluto de estimação obtido para estimar o número médio de trabalhadores por firma da indústria a um nível de confiança de 95% é de 14.4

OU

Sabe-se que:

$$\varepsilon = \frac{At}{2};\ e\ At = lim.sup - lim.inf$$

Construindo o *intervalo de confiança* com 95% *de confiança teremos:*

$$420.4 - 1.96\frac{55.7}{\sqrt{50}}\sqrt{\frac{380-50}{380-1}} \leq \mu \leq 420.4 + 1.96\frac{55.7}{\sqrt{50}}\sqrt{\frac{380-50}{380-1}}$$

$$401.5 \leq \mu \leq 439.3$$

Então

$$At = 434.8 - 406 = 28.8$$

$$\varepsilon = \frac{At}{2} = \frac{28.8}{2} = 14.4$$

Logo o erro absoluto de estimação obtido para estimar o número médio de trabalhadores por firma da indústria a um nível de confiança de 95% *é de 14.4*

d) Qual o efeito de variar o grau de confiança mantendo a amostra?

<u>Resposta:</u>

Quanto maior for o grau de confiança $\begin{cases} maior\ será\ a\ amplitude\ do\ intervalo \\ e \\ maior\ o\ erro\ de\ estimação \end{cases}$

108. Num estudo de mercado foi encontrado o intervalo de confiança de $[0.52;\ 0.61]$ a 95% para a proporção de pessoas receptivas a um novo tipo de espuma de banho a lançar em breve no mercado:. Comente as seguintes afirmações, indicando se estas lhe parecem correctas ou incorrectas:

a) 95% das pessoas vão passar a usar a nova espuma de banho.

Afirmação incorrecta:

Os 95% de confiança significam o seguinte:

O processo que se utiliza para calcular os intervalos, é um processo tal que se o utilizasse com todas as amostras possíveis (da mesma dimensão) selecionadas da mesma população, cerca de 95% das vezes produziria intervalos que contêm a proporção de pessoas receptivas a um novo tipo de espuma de banho e cerca de 5% das vezes intervalos que não o contêm.

b) A probabilidade da nova espuma de banho alcançar uma quota de mercado de 50%, é de 0.95.

Afirmação incorrecta:

Do estudo conclui-se que a probabilidade de a proporção de pessoas receptivas a um novo tipo de espuma de banho estar entre 52% e 61% é de 0.95.

c) A quota de mercado poderá ser, com 95% de confiança, de 56.5% (valor intermédio do intervalo).

Afirmação incorrecta:

Recorrendo aos cometários das alíneas a) e b)

d) O resultado obtido indica apenas que é oportuno proceder ao lançamento da nova espuma de banho.

Afirmação subjectiva:

Pois, para se lançar a nova espuma de banho ao mercado depende de vários factores que o pesquisador achar lhes relevantes como intervalo de confiança esperado, tamanho da amostra, o grau de confiança, a margem do erro, etc.

109. Pretende-se investigar o nível de remuneração salarial dos homens e mulheres duma certa categoria profissional. De duas amostras obtidas entre dois grupos, destacam-se os seguintes resultados (em unidades monetárias):

$$Amostra\ de\ 18\ homens: \bar{x} = 34.8\ e\ s^2 = 5.7$$

$$Amostra\ de\ 11\ mulheres: \bar{x} = 31\ e\ s^2 = 10.3$$

a) Construa um intervalo de confiança a 99% para as diferenças salariais médias entre homens e mulheres, sabendo que as varianças saõ iguais.

<u>Resolução:</u>

Seja $n_1 = homens\ e\ n_2 = mulheres$

Como $\sigma_1{}^2\ e\ \sigma_2{}^2$ *são desconhecidos e o tamanho da amostra é pequeno, então* $t \sim t_{(n_1+n_2-2)}$

Sendo (1 – α*).100%* $= 99\%;\ \alpha = 1\%$ e *graus de liberdade (G.L)=v=*$n_1 + n_2 - 2 = 27$

da tabela distribuíção de student temos: $t_{\frac{\alpha}{2};(n_1+n_2-2)} = t_{0.005;27} = 2.771$

A fórmula geral do intervalo de confiança para a diferença das médias será:

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2}(n_1+n_2-2)}\sqrt{\frac{S_p{}^2}{n_1} + \frac{S_p{}^2}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + t_{\frac{\alpha}{2}(n_1+n_2-2)}\sqrt{\frac{S_p{}^2}{n_1} + \frac{S_p{}^2}{n_2}}$$

$$S_p{}^2 = \frac{(n_1 - 1)S_1{}^2 + (n_2 - 1)S_2{}^2}{n_1 + n_2 - 2}$$

$$S_p{}^2 = \frac{(18 - 1) * 5.7 + (11 - 1) * 10.3}{18 + 11 - 2} = 7.4$$

Substituíndo os dados na fórmula do intervalo de confiança teremos:

$$(34.8 - 31) - 2.771\sqrt{\frac{7.4}{18} + \frac{7.4}{11}} \le \mu_1 - \mu_2 \le (34.8 - 31) + 2.771\sqrt{\frac{7.4}{18} + \frac{7.4}{11}}$$

$$0.915 \le \mu_1 - \mu_2 \le 6.685$$

Interpretação: A um nível de confiança de 99% pode se afirmar que o intervalo $[0.915; 6.685]$ *contêm o verdadeiro valor para as diferenças salariais médias entre os dois sexos.*

b) Determine o intervalo de confiança para o quociente entre as variâncias dos salários de mulheres e homens, e conclua sobre a possível existência de discriminação sexual na atribuição de remunerações.

<u>Resolução:</u>

Seja agora $n_1 = mulheres\ e\ n_2 = homens$

Sendo os valores médios $\mu_1 e\ \mu_2$ e *as variâncias* $\sigma_1{}^2\ e\ \sigma_2{}^2$ *desconhecidos, então,* $F \sim F_{(n_1-1;n_2-1)}$

Como (1 – α*).100%* $= 99\%;\ \alpha = 1\%$ e,

graus de liberdade (G.L): $v_1 = n_{1-1} = 10\ e\ v_2 = n_{2-1} = 17$

A fórmula geral do intervalo de confiança para quociente entre as variâncias será:

$$\frac{S_1{}^2}{S_2{}^2} * F_{inf} \le \frac{\sigma_1{}^2}{\sigma_2{}^2} \le \frac{S_1{}^2}{S_2{}^2} * F_{sup}$$

Neste caso é necessário inicialmente determinar os valores da distribuição, F de modo, que F_{inf} tenha uma área (probabilidade) à esquerda igual a 99.5% e F_{sup} tenha uma área (probabilidade) à direita igual a 0.5%.

Nota que as tabelas F fornecem apenas os valores de $F_{\alpha;v1;v2}$, mas o valor de $F_{1-\alpha;v1;v2}$ pode ser obtido fazendo $\frac{1}{F_{\alpha;v1;v2}}$; assim teremos:

$$F_{sup} = F_{\frac{\alpha}{2};v1;v2} = F_{0.005;10;17} = 4.14$$

$$F_{inf} = F_{1-\frac{\alpha}{2};v1;v2} = F_{0.995;10;17} = \frac{1}{F_{0.005;17;10}} = \frac{1}{4.20+4.12} = 0.1202$$

Substituindo os dados na fórmula do intervalo de confiança teremos:

$$\frac{10.3}{5.7} * 0.1202 \leq \frac{{\sigma_1}^2}{{\sigma_2}^2} \leq \frac{10.3}{5.7} * 4.142$$

$$0.2172 \leq \frac{{\sigma_1}^2}{{\sigma_2}^2} \leq 7.485$$

Interpretação: A um nível de confiança de 99% pode se afirmar que o intervalo [0.2172;7.485] contêm o verdadeiro valor para o quociente entre as variâncias dos salários de mulheres e homens.

Conclusão: Como o intervalo inclui o valor 1, não pode ser descartada a hipótese de que as variâncias de mulheres e homens seja a mesma. Logo conclui-se que os salários dos dois grupos são homoéneos.

110. Um teste de auditoria, para estabelecer com que frequência ocorrem falhas no processamento de determinado procedimento de controlo interno, está para ser feito. O auditor decide que a taxa máxima de erro tolerável permitida é de 5%.

a) Que tamanho de amostra é necessário para atingir uma precisão de 2%, com 99% de confiança?

Resolução:

Do problema se tem:

$$p = 0.05;\ q = 1 - 0.05 = 0.95;\ \varepsilon = 0.02;\ (1 - \alpha).100\% = 99\%\ onde\ \alpha = 0.1\%$$

$$Z_{\frac{\alpha}{2}} = Z_{0.005} = 2.576$$

Partindo da fórmula do erro de estimação do I.C para as proporções teremos:

$$\varepsilon = Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$$

isolando o n teremos:

$$n = \frac{{Z_{\frac{\alpha}{2}}}^2 * p * q}{\varepsilon^2}$$

$$n = \frac{2.576^2 * 0.05 * 0.95}{0.02^2}$$

$$n = 788$$

Para atingir uma precisão de amostra de 2%, com 99% de confiança e taxa máxima tolerável de 5% é necessário um tamanho de amostra de 788.

b) Qual seria sua resposta em (a) se a taxa máxima tolerável de erro fosse 10%?

Resolução:

Do problema se tem:

$$p = 0.1;\ q = 1 - 0.1 = 0.9;\ \varepsilon = 0.02$$

Seguindo os passos da alinea a) teremos:

$$n = \frac{Z_{\frac{\alpha}{2}}{}^{2} * p * q}{\varepsilon^2}$$

$$n = \frac{2.576^2 * 0.01 * 0.90}{0.02^2} = 1493$$

Para atingir uma precisão de amostra de 2%, com 99% de confiança e taxa máxima tolerável de 10% é necessário um tamanho de amostra de 1493.

111. Sabe-se por experiência que o nº de cheques gravados por hora pelos operadores de terminal do Serviço de Tesouraria dum certo Banco tem distribuição aproximadamente normal. Em 9 dias ao acaso foi feito, durante uma hora, o controlo do nº de cheques gravados, tendo-se obtido os seguintes resultados:

$$\sum_{i=1}^{9} x_i = 10.206 \qquad \sum_{i=1}^{9} (x_i - \bar{x})^2 = 25.688$$

a) Construa um intervalo de confiança a 95% para o número médio de cheques gravados por hora nesse terminal.

Resolução:

Seja n= 9

Como σ^2 é desconhecido e o tamanho da amostra é pequeno, entã

$$t \sim t_{(n-1)}$$

Sendo $(1 - \alpha) * 100\% = 90\%;\ \alpha = 10\%$ *e grau de liberdade* $Gl = v = n - 1 = 8$

da tabela distribuíção de student temos: $t_{\frac{\alpha}{2}(n-1)} = t_{0.05;8} = 1.860$

Calculando os estimadores de μ e σ^2 teremos:

$$\bar{x} = \frac{\sum_{i=1}^{9} xi}{n} = \frac{10.206}{9} = 1.134$$

$$S^2 = \frac{\sum_{i=1}^{9} (xi - \bar{x})^2}{n - 1} = \frac{25.688}{8} = 3.211$$

$$S = \sqrt{3.211} = 1.792$$

A fórmula geral do intervalo de confiança para a média será:

$$\bar{x} - t_{\frac{\alpha}{2}(n-1)} * \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\frac{\alpha}{2}(n-1)} * \frac{s}{\sqrt{n}}$$

Substituindo os dados pela fórmula temos:

$$1.134 - 1.860\frac{1.792}{\sqrt{9}} \leq \mu \leq 1.134 + 1.860\frac{1.792}{\sqrt{9}}$$

$$0.023 \leq \mu \leq 2.245$$

Interpretação: A um nível de confiança de 90% pode se afirmar que o intervalo $[0.023; 2.245]$ *contêm o verdadeiro número médio de cheques gravados por hora nesse terminal.*

b) Construa um intervalo de confiança a 90% para o mesmo parâmetro, supondo agora que o desvio-padrão da população é 2.

Como σ^2 é conhecido, então $\bar{x} \sim N\left(\mu; \frac{\sigma^2}{n}\right)$

e $\sigma = 2;\ Z_{0.05} = 1.645$

A fórmula geral do intervalo de confiança para a média será:

$$\bar{x} - Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}}$$

Substituindo os dados pela fórmula temos:

$$1.134 - 1.645 * \frac{2}{\sqrt{9}} \leq \mu \leq 1.134 + 1.645 * \frac{2}{\sqrt{9}}$$

$$0.037 \leq \mu \leq 2.231$$

Interpretação: A um nível de confiança de 90% pode se afirmar que o intervalo $[0.037; 2.231]$ *contêm o verdadeiro número médio* de cheques gravados por hora nesse terminal.

c) Determine o intervalo de confiança a 90% para o desvio padrão.

Resolução

$$\chi^2 \sim {\chi^2}_{n-1}$$

Sendo $(1-\alpha).100\% = 90\%;\ \alpha = 10\%;\ \frac{\alpha}{2} = 0.05\ e\ 1 - \frac{\alpha}{2} = 0.95$ *e,*

graus de liberdade (G.L): $v = n - 1 = 8$

A fórmula geral do intervalo de confiança para o desvio padrão será:

$$\sqrt{\frac{(n-1) * S^2}{{\chi^2}_{sup}}} \leq \sqrt{\sigma^2} \leq \sqrt{\frac{(n-1) * S^2}{{\chi^2}_{inf}}}$$

Neste caso é necessário inicialmente determinar os valores da distribuição, χ^2 *de modo, que* ${\chi^2}_{inf}$ *tenha uma área (probabilidade) à esquerda igual a 95% e* ${\chi^2}_{sup}$ *tenha uma área (probabilidade) à direita igual a 5%.*

Recorrendo a tabela Chi-Quadrado teremos:

$$\chi^2{}_{sup} = \chi^2{}_{0.05;8} = 15.5 \text{ e } \chi^2{}_{inf} = \chi^2{}_{0.95;8} = 2.73$$

Substituindo os dados pela fórmula temos:

$$\sqrt{\frac{(8-1)*3.211}{15.5}} \leq \sigma \leq \sqrt{\frac{(8-1)*3.211}{2.73}}$$

$$1.204 \leq \sigma \leq 2.869$$

Interpretação: A um nível de confiança de 90% pode se afirmar que o intervalo [1.204;2.869] contêm o verdadeiro desvio padrão de cheques gravados por hora nesse terminal.

112. O gerente de uma rede de hipermercados está a analisar os desvios observados no volume de vendas mensais. Este gerente sabe que o volume de vendas mensais segue uma lei Normal com média μ. Após ter recolhido 20 meses de observações, resumiu-os no seguinte resultado:

$$\sum_{i=1}^{20}(x_i - 20)^2 = 800$$

a) Determine a margem do erro a um nível de significância de 99%.

<u>*Resolução:*</u>

Seja μ e σ^2 *desconhecidos, então,* $t \sim t_v$

Sendo $(1-\alpha).100\% = 99\%;\ \alpha = 1\%$ $(\alpha/2 = 0.005\ e\ 1-\alpha/2 = 0.995)$ e,

graus de liberdade (G.L): $v = n - 1 = 20 - 1 = 19$

e $t_{0.005} = 2.871$

Calculando os estimador de σ^2 *teremos:*

$$S^2 = \frac{\sum_{i=1}^{20}(xi - \bar{x})^2}{n-1} = \frac{800}{19} = 42.11$$

Recorrendo a fórmula para cálculo do erro (ε) *através do intervalo de confiança terremos:*

$$\varepsilon = t_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}}$$

$$\varepsilon = 2.871 * \frac{6.49}{\sqrt{20}}$$

$$\varepsilon = 4.17$$

Logo o erro absoluto de estimação obtido para estimar o número médio de de desvios de vendas por mês a um nível de confiança de 99% é de 4.17.

b) Determine o intervalo de confiança de 90%, para estimar o volume médio de vendas por mês.

<u>Resolução:</u>

Seja $n = 20$

Como σ^2 *é desconhecido e o tamanho da amostra é pequeno, então* $t \sim t_{(n-1)}$

Sendo $(1 - \alpha).100\% = 90\%$; $\alpha = 10\%$ *e grau de liberdade* $(G.L) = v = n - 1 = 19$

da tabela distribuição de student temos: $t_{\frac{\alpha}{2}(n-1)} = t_{0.05;19} = 1.729$

Calculando os estimadores de μ *e* σ *teremos:*

$$\bar{x} = 20$$

$$S = \sqrt{42.11} = 6.49$$

A fórmula geral do intervalo de confiança para a média será:

$$\bar{x} - t_{\frac{\alpha}{2}(n-1)} \frac{S}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\frac{\alpha}{2}(n-1)} \frac{S}{\sqrt{n}}$$

Substituindo os dados pela fórmula temos:

$$20 - 1.729 * \frac{6.49}{\sqrt{20}} \leq \mu \leq 20 + 1.729 * \frac{6.49}{\sqrt{20}}$$

$$17.49 \leq \mu \leq 22.51$$

Interpretação: A um nível de confiança de 90% pode se afirmar que o intervalo $[17.49; 22.51]$ *contêm o verdadeiro número médio de cheques gravados por hora nesse terminal.*

113. Uma agência de propaganda, que atende a uma das principais estações de rádio, gostaria de calcular a quantidade média de tempo que a audiência gasta diariamente ouvindo radio. A partir de estudos do passado, o desvio padrão é calculado em 45 minutos. Que tamanho de amostra é necessário se a agência quiser ter 90% de confiança de estar correta num intervalo de ± 5 minutos?

Resolução:

Seja $\sigma = 45$ *e* $\varepsilon = 5, então$ então $\bar{x} \sim N\left(\mu; \frac{\sigma^2}{n}\right)$

$(1 - \alpha).100\% = 90\%;\ \alpha = 0.10$ onde $Z_{\frac{\alpha}{2}} = Z_{0.05} = 1.645$

Recorrendo a fórmula para cálculo do erro (ε) *através do intervalo de confiança terremos:*

$$\varepsilon = Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$n = \left(\frac{Z_{\frac{\alpha}{2}} * \sigma}{\varepsilon}\right)^2$$

$$n = \left(\frac{1.645 * 45}{5}\right)^2 \approx 219$$

Interpretação: Se a agência quiser ter 90% de confiança de estar correta num intervalo de ± 5 *minutos é necessário um tamanho de amostra de aproximadamente 219.*

114. Pesaram-se 16 sacos de café e com os pesos observados, em gramas, construiu-se o seguinte intervalo de confiança a 95%, para o valor médio do peso de um saco: $[1000.74; 1009.26]$.

a) Deduza o valor médio e o desvio padrão (corrigido) do peso dos sacos que constituem a amostra, admitindo a normalidade da população.

<u>*Resolução:*</u>

Do problema temos:

$$\mathrm{n} = 25 \text{ e } \alpha = 0.05$$

$$Seja\ \sigma\ desconhecido\ e\ n < 30, então\ t \sim t_v, \text{ onde } t_{\frac{\alpha}{2}} = t_{0.025;15} = 2.131$$

Partindo da fórmula de erro para intervalos de confiança pode-se determinar o valor do desvio padrão amostral, então:

$$\varepsilon = t_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}}, onde\ \varepsilon = \frac{At}{2} = \frac{1009.26 - 1000.74}{2} = 4.26$$

Insolando o S teremos:

$$S = \frac{\varepsilon * \sqrt{n}}{t_{\frac{\alpha}{2}}} = \frac{4.26 * 4}{2.131} = 8$$

e partindo da fórmula de intervalo de confiança para média quando

σ desconhecido e $n < 30$, pode-se determinar o valor da média amostral, então:

$$\bar{x} - t_{\frac{\alpha}{2}(n-1)} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\frac{\alpha}{2}(n-1)} \frac{s}{\sqrt{n}}$$

Recorrendo a um dos limites (inferior ou superior) teremos:

$$\bar{x} - t_{\frac{\alpha}{2}(n-1)} \frac{s}{\sqrt{n}} = 1000.74$$

$$\bar{x} = 1000.74 + 2.131 * \frac{7.996}{4} = 1005$$

Logo, os valores da média e desvio padrão amostral são 1005g e 8 aproximadamente.

b) Para construir um intervalo de confiança com uma amplitude de 3 gramas, qual deverá ser a dimensão da amostra, mantendo-se o grau de confiança do intervalo?

<u>**Resolução:**</u>

Do problema temos:

$$\mathrm{At} = 3 \text{ e } \alpha = 0.05$$

$$Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$$

Recorrendo a resolução em a) temos:

$$\varepsilon = Z_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}}$$

$$n = \left(\frac{Z_{\frac{\alpha}{2}} * S}{\varepsilon}\right)^2 = \left(\frac{1.96 * 8}{1.5}\right)^2 \approx 109$$

Interpretação: Para construir um intervalo de confiança com uma amplitude de 3 gramas, a dimensão da amostra deve ser de aproximadamente 109 sacos.

115. Em uma indústria química, os engenheiros desejam saber a nível de confiança de 95% se o alongamento de um composto de borracha permanece inalterado ao passar por uma máquina estrussora. Como o alongamento do composto depende do lote de matéria-prima usado na sua confecção, os dados foram colectados aos pares:

Lote	1	2	3	4	5	6	7	8	9	10
Antes	360	370	380	345	365	380	390	395	385	410
Depois	360	365	355	340	350	370	390	375	375	395

a) Construa um intervalo de confiança para a diferença entre os pares observados e conclua a respeito;

Resolução:

1° Passo: Determinar o valor da diferença d em módulo:

Lote	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Antes	*360*	*370*	*380*	*345*	*365*	*380*	*390*	*395*	*385*	*410*
Depois	*360*	*365*	*355*	*340*	*350*	*370*	*390*	*375*	*375*	*395*
d	*0*	*5*	*25*	*5*	*15*	*10*	*0*	*20*	*10*	*15*

2° Passo: Determinar o valor da média e desvio padrão da média d

$$\bar{d} = \frac{\sum_{i=1}^{10} d_i}{n} = \frac{105}{10} = 10.5$$

$$S_d = \sqrt{\frac{\sum_{i=1}^{10}\left(d_i - \bar{d}\right)^2}{n-1}} = 8.32$$

Como $n < 30$, então $t \sim t_v$, onde $t_v = t_{\frac{\alpha}{2};n-1} = t_{0.025;9} = 2.262$

Recorrendo a fórmula de intervalo de confiança para a diferença entre os pares observados, temos:

$$\bar{d} - t_{\frac{\alpha}{2}(n-1)} \frac{S_d}{\sqrt{n}} \le \mu \le \bar{d} + t_{\frac{\alpha}{2}(n-1)} \frac{S_d}{\sqrt{n}}$$

$$10.5 - 2.262 * \frac{8.32}{\sqrt{10}} \le \mu \le 10.5 + 2.262 * \frac{8.32}{\sqrt{10}}$$

$$4.57 \le \mu \le 16.45$$

Conlusão: Como o valor zero não está incluído no intervalo, rejeita-se a hipótese de que o alongamento de um composto de borracha permanece inalterado ao passar por uma máquina estrussora.

b) Calcula o quociente entre as variâncias dos alongamentos medidos antes e depois do composto passar pela estrussora. Depois construa um intervalo de confiança para esse quociente.

<u>Resolução:</u>

A fórmula para cálculo do quociente entre as varianças é:

$$F = \frac{S^2{}_1}{S^2{}_2}\ , onde\ 1\ (antes)\ e\ 2\ (depois)$$

$$então, F = \frac{S^2{}_1}{S^2{}_2} = \frac{351.1}{295.8} = 1.19$$

Intervalo de confiança:

A fórmula geral do intervalo de confiança para quociente entre as variâncias será:

$$\frac{{S_1}^2}{{S_2}^2} * F_{inf} \leq \frac{{\sigma_1}^2}{{\sigma_2}^2} \leq \frac{{S_1}^2}{{S_2}^2} * F_{sup}$$

Neste caso é necessário inicialmente determinar os valores da distribuição, F de modo, que F_{inf} tenha uma área (probabilidade) à esquerda igual a 97.5% e F_{sup} tenha uma área (probabilidade) à direita igual a $2,5\%$.

Nota que as tabelas F fornecem apenas os valores de $F_{\alpha;v1;v2}$, mas o valor de $F_{1-\alpha;v1;v2}$ pode ser obtido fazendo $\frac{1}{F_{\alpha;v1;v2}}$; assim teremos:

$$F_{sup} = F_{\frac{\alpha}{2};v1;v2} = F_{0.025;9;9} = 4.026 \text{ e } F_{inf} = F_{1-\frac{\alpha}{2};v1;v2} = F_{0.975;9;9} = \frac{1}{4.026} = 0.248$$

Substituindo os dados na fórmula do intervalo de confiança teremos:

$$1.19 * 0.248 \leq \frac{{\sigma_1}^2}{{\sigma_2}^2} \leq 1.19 * 4.026$$

$$0.295 \leq \frac{{\sigma_1}^2}{{\sigma_2}^2} \leq 4.791$$

Interpretação: A um nível de confiança de 99% pode se afirmar que o intervalo $[0.387; 7.485]$ contêm o verdadeiro valor para o quociente entre as variâncias dos alongamentos medidos antes e depois do composto passar pela estrussora.

Conclusão: Como o intervalo inclui o valor 1, não pode ser descartada a hipótese de que as variâncias dos alongamentos medidos antes e depois do composto passar pela estrussora seja a mesma.

116. Uma votação será feita entre os residentes de uma cidade e a região rural ao redor desta cidade para determinar se um projecto químico deverá ser construído. A construção é dentro dos limites da cidade e por esta razão muitos eleitores do campo sentem que o projecto passará por causa da grande proporção dos eleitores da cidade, os quais são favoráveis. Para determinar se existe diferença na proporção de eleitores da cidade e do campo a favor do projecto, uma amostragem foi feita. Supondo que 80 de 200 eleitores da cidade não são a favor do projecto e 240 de 500 eleitores do campo são a favor, construa um intervalo a 97% de confiança para a diferença das proporções de eleitores favoráveis ao projecto.

Resolução

Do problema temos:

$$n_A = 200;\ x_A = 120;\ p_A = \frac{120}{200} = 0.60;\ n_B = 500;\ x_B = 240;\ p_B = \frac{240}{500} = 0.48; 1 - \alpha = 0.97$$

$$Z_{\frac{\alpha}{2}} = Z_{0.015} = 2.17$$

Pela fórmula temos:

$$(p_1 - p_2) - Z_{\frac{\alpha}{2}} * \sqrt{\frac{p_1 * q_1}{n_1} + \frac{p_2 * q_2}{n_2}} \leq \pi_1 - \pi_2 \leq (p_1 - p_2) + Z_{\frac{\alpha}{2}} * \sqrt{\frac{p_1 * q_1}{n_1} + \frac{p_2 * q_2}{n_2}}$$

$$(0.60 - 0.48) \pm 2.17 * \sqrt{\frac{0.60 * 0.40}{200} + \frac{0.48 * 0.52}{500}} \in \pi_1 - \pi_2$$

$$0.031 \leq \pi_1 - \pi_2 \leq 0.209$$

Interpretação: A um nível de confiança de 98% pode se afirmar que o intervalo $[-0.077; 0.137]$ contém a verdadeira proporção de clientes favoráveis ao produto.

117. Uma das maneiras de medir o grau de satisfação dos empregados de uma mesma categoria quanto a política salarial é por meio da variabilidade dos seus salários. A Fábrica A diz ser mais coerente na política salarial do que a Fabrica B. Para verificar essa afirmação, sorteou-se uma amostra de 17 funcionários não especializados da fábrica A e 15 da fábrica B, obtendo-se variânças de 1000 u.m e 1200 u.m respectivamente. Usando um nível de confiança de 95%, qual seria a sua conclusão?

Resolução

Do problema temos:

$$n_A = 17;\ s^2{}_A = 1000;\ \ n_B = 15;\ s^2{}_B = 1200; 1 - \alpha = 0.95$$

A partir do intervalo de confiança para o quociente das variançcas temos:

$$\frac{S^2{}_1}{S^2{}_2} * F_{1-\frac{\alpha}{2};\, n_1-1; n_2-1} \leq \frac{\sigma^2{}_1}{\sigma^2{}_2} \leq \frac{S^2{}_1}{S^2{}_2} * F_{\frac{\alpha}{2};\, n_1-1; n_2-1}$$

Usando a tabela F, obteve se os seguintes valores críticos:

$$F_{\frac{\alpha}{2};\, n_1-1;n_2-1} = F_{0.025;\, 16;14} = 2.92$$

$$F_{1-\frac{\alpha}{2};\, n_1-1;n_2-1} = \frac{1}{F_{\frac{\alpha}{2};\, n_2-1;n_1-1;}} = \frac{1}{F_{0.025;\, 14;16}} = \frac{1}{2.82} = 0.3546$$

Logo:

$$\frac{1000}{1200} * 0.3546 \leq \frac{\sigma^2{}_1}{\sigma^2{}_2} \leq \frac{1000}{1200} * 2.92$$

$$0.2955 \leq \frac{\sigma^2{}_1}{\sigma^2{}_2} \leq 2.4333$$

Conclusão: Como o valor $1 \in [0.2955; 2.4333]$, pode se concluir com 95% de confiança que não há diferença entre as empresas A e B quanto ao grau de satisfação dos empregados de uma mesma categoria quanto a política salarial.

118. Um teste de auditoria, para estabelecer com que frequências ocorrem falhas no processamento de determinado procedimento de controlo interno, está para ser feito. O auditor decide que a taxa máxima de erro tolerável permitida é de 5%.

a) Que tamanho de amostra é necessário para atingir uma precisão de amostra de 2%, com 99% de confiança?

<u>Resolução</u>

$$p = 0.05;\ q = 1 - 0.05 = 0.95;\ \varepsilon = 0.02;\ (1 - \alpha).100\% \ = \ 99\%\ onde\ \alpha = 0.1\%$$

$$Z_{\frac{\alpha}{2}} = Z_{0.005} = 2.58$$

Partindo da fórmula do erro de estimação do I.C para as proporções teremos:

$$\varepsilon = Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$$

isolando o n teremos:

$$n = \frac{Z_{\frac{\alpha}{2}}{}^2 * p * q}{\varepsilon^2} = \frac{2.58^2 * 0.05 * 0.95}{0.02^2} = 788$$

Para atingir uma precisão de amostra de 2%, com 99% de confiança e taxa máxima tolerável de 5% é necessário um tamanho de amostra de 788.

b) Qual seria sua resposta em (a) se a taxa máxima tolerável de erro fosse 10%?

<u>Resolução:</u>

Do problema se tem:

$$p = 0.1;\ q = 1 - 0.1 = 0.9;\ \varepsilon = 0.02$$

$$Z_{\frac{\alpha}{2}}{=}Z_{0.005} = 2.58$$

Seguindo os passos da alinea a) teremos:

$$n = \frac{Z_{\frac{\alpha}{2}}^{\ 2} * p * q}{\varepsilon^2} = \frac{2.58^2 * 0.10 * 0.90}{0.02^2} = 14936$$

Para atingir uma precisão de amostra de 2%, com 99% de confiança e taxa máxima tolerável de 10% é necessário um tamanho de amostra de 1493.

119. Um fabricante de insecticida descobriu, em uma pesquisa, que muitos consumidores da categoria achavam que o cheiro do seu produto era suave demais, e enfraquecia o seu apelo publicitário de produto fortíssimo contra os insectos. Após alguns sniff tests, foi seleccionado um novo aroma para o produto. Faltava testar o novo aroma no produto em uso pela consumidora. O fabricante solicitou à empresa de pesquisa que fizesse um teste de produto in home em que uma amostra de 15 consumidores usaria durante uma semana o produto com o cheiro actual, enquanto uma outra amostra, também de 12 consumidores com o mesmo perfil, usaria o produto com o novo cheiro. A avaliação em escala de 7 pontos resultou nos seguintes parâmetros.

	Aroma actual	Aroma novo
Média	5.94	6.27
Desvio Padrão	5.22	2.39

Assumindo que as populações são normais de variânças iguais, qual é a probabilidade de o intervalo $[-3.868; 3.208]$ conter a diferença entre os valores esperados das duas populações?

<u>*Resolução*</u>

Assumindo que $\sigma_1^{\ 2}$ *e* $\sigma_2^{\ 2}$ *são desconhecidas* e *iguais*, e n_1 *e* n_2 *pequeno*, o *intervalo acima foi construído usando a seguinte fórmula:*

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\frac{\alpha}{2};n_1+n_2-2} * \sqrt{\left(\frac{(n_1-1) * S_1^{\ 2} + (n_2-1) * S_2^{\ 2}}{n_1+n_2-2}\right) * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \in \mu_1 - \mu_2$$

Sabe se que:

$$\varepsilon = \frac{At}{2} = \frac{3.208 + 3.868}{2} = 3.538$$

$$\varepsilon = t_{\frac{\alpha}{2};n_1+n_2-2} * \sqrt{\left(\frac{(n_1-1) * S_1^{\ 2} + (n_2-1) * S_2^{\ 2}}{n_1+n_2-2}\right) * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$3.538 = t_{\frac{\alpha}{2};25} * \sqrt{\left(\frac{14 * 5.22^2 + 11 * 2.39^2}{15+12-2}\right) * \left(\frac{1}{15} + \frac{1}{12}\right)}$$

$$t_{\frac{\alpha}{2};25} = \frac{3.538}{1.633} = 2.167$$

Recorrendo a tabela t temos:

$$\frac{\alpha}{2} = 0.02$$

$$\alpha = 0.04$$

$$1-\alpha=1-0.04=0.96$$

Resposta: A probabilidade de o intervalo $[-3.868; 3.208]$ conter a diferença entre os valores esperados das duas populações é de 96%.

120. Um teste de auditoria, para estabelecer com que frequências ocorrem falhas no processamento de determinado procedimento de controlo interno, está para ser feito. O auditor decide que a precisão deve ser igual a 2% do desvio padrão máximo do erro tolerável permitido. Qual deve ser o tamanho de amostra a um nível de confiança de 97%, sabendo que as falhas no processamento seguem uma distribuição normal?

Resolução

Pelo problema temos:

$$\varepsilon=0.02*\sigma;\ (1-\alpha)=0.97; Z_{\frac{\alpha}{2}}=Z_{0.015}=2.17;\ n=?$$

A fórmula para o cálculo de amostra, será:

$$n=\left(\frac{Z_{\frac{\alpha}{2}}*\sigma}{\varepsilon}\right)^2=\left(\frac{2.17*\sigma}{0.02*\sigma}\right)^2=11772$$

Resposta: O tamanho de amostra a um nível de confiança de 97%, sabendo que as falhas no processamento seguem uma distribuição normal deve ser de 11772.

121. Um fabricante de insecticida descobriu, em uma pesquisa, que muitos consumidores da categoria achavam que o cheiro do seu produto era suave demais, e enfraquecia o seu apelo publicitário de produto fortíssimo contra os insectos. Após alguns sniff tests, foi seleccionado um novo aroma para o produto. Faltava testar o novo aroma no produto em uso pela consumidora. O fabricante solicitou à empresa de pesquisa que fizesse um teste de produto in home em que uma amostra de 15 consumidores usaria durante uma semana o produto com o cheiro actual, enquanto uma outra amostra, também de 12 consumidores com o mesmo perfil, usaria o produto com o novo cheiro. A avaliação em escala de 7 pontos resultou nos seguintes parâmetros.

	Aroma actual	Aroma novo
Média	5.94	6.27
Desvio Padrão	5.22	2.39

Assumindo que as populações são normais e as variânças são diferentes, qual é a probabilidade de o intervalo $[0.1126; 5.5474]$ conter a diferença entre os valores esperados das duas populações?

Resolução

Assumindo que ${\sigma_1}^2$ e ${\sigma_2}^2$ são desconhecidas e diferentes, e n_1 e n_2 pequeno, o intervalo acima foi construído usando a seguinte fórmula:

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2};Gl} * \sqrt{\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}} \leq \mu_1 - \mu_2 \leq (\bar{x}_1 - \bar{x}_2) + t_{\frac{\alpha}{2};Gl} * \sqrt{\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}}$$

$$Onde$$

$$Gl = \frac{\left(\frac{S^2{}_1}{n_1} + \frac{S^2{}_2}{n_2}\right)^2}{\frac{\left(\frac{S^2{}_1}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S^2{}_2}{n_2}\right)^2}{n_2 - 1}} = \frac{\left(\frac{5.22^2}{15} + \frac{2.39^2}{12}\right)^2}{\frac{\left(\frac{5.22^2}{15}\right)^2}{14} + \frac{\left(\frac{2.39^2}{12}\right)^2}{11}} = 20$$

Sabe se que:

$$\varepsilon = \frac{At}{2} = \frac{5.5474 - 0.1126}{2} = 2.7174$$

$$\varepsilon = t_{\frac{\alpha}{2};Gl} * \sqrt{\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}}$$

$$2.7174 = t_{\frac{\alpha}{2};21} * \sqrt{\frac{5.22^2}{15} + \frac{2.39^2}{12}}$$

$$t_{\frac{\alpha}{2};21} = 1.914$$

Recorrendo a tabela t temos:

$$\frac{\alpha}{2} = 0.035$$

$$\alpha = 0.07$$

$$1 - \alpha = 1 - 0.07 = 0.93$$

Resposta: A probabilidade de o intervalo $[0.1126; 5.5474]$ conter a diferença entre os valores esperados das duas populações é de 96%.

122. Uma empresa que presta serviços de acessória a outras empresas está interessada em comparar a taxa de reclamações sobre os seus serviços em dois dos seus escritórios em duas cidades diferentes. Suponha que a empresa tenha seleccionado aleatoriamente serviços realizados pelo escritório da cidade A e da cidade B, tendo obtido a 95% de confiança as seguintes taxas de reclamações:

Cidade	Cidade A	Cidade B
$Taxas\ de\ reclamações$	0.54 ± 0.0977	0.50 ± 0.0033

Construa um intervalo de confiança a 98% para a diferença entre as taxas de reclamações das duas cidades.

Resolução

A fórmula para o cálculo do intervalo de confiança para a diferença das proporções é:

$$(p_1 - p_2) - Z_{\frac{\alpha}{2}} * \sqrt{\frac{p_1 * q_1}{n_1} + \frac{p_2 * q_2}{n_2}} \leq \pi_1 - \pi_2 \leq (p_1 - p_2) + Z_{\frac{\alpha}{2}} * \sqrt{\frac{p_1 * q_1}{n_1} + \frac{p_2 * q_2}{n_2}}$$

Conforme os dados, as taxas foram obtidas através do intervalo de confiança para as proporções, onde:

$$p \pm Z_{\frac{\alpha}{2}} * \sqrt{\frac{p * (1 - p)}{n}} \in \pi$$

A partir da fórmula acima pode-se determinar os tamanhos de amostra para cada cidade, sendo assim teremos:

$$n = \frac{Z^2_{\frac{\alpha}{2}} * p * (1 - p)}{\varepsilon^2}$$

$$n_1 = \frac{1.96^2 * 0.54 * 0.46}{0.0977^2} = 100$$

$$n_2 = \frac{1.96^2 * 0.50 * 0.50}{0.0033^2} = 88191$$

Logo:

$$(0.54 - 0.50) \pm 2.33 * \sqrt{\frac{0.54 * 0.46}{100} + \frac{0.50 * 0.50}{88191}} \in \pi_1 - \pi_2$$

$$0.076 \leq \pi_1 - \pi_2 \leq 0.156$$

Interpretação: A um nível de confiança de 98% pode se afirmar que o intervalo $[0.076; 0.156]$ *contém a verdadeira diferença entre as taxas de reclamações das duas cidades.*

123. Um grupo de pesquisadores pretende estimar o real salário médio de trabalhadores bancários. Para tal seleccionou-se aleatoriamente 100 trabalhadores de diferentes bancos, tendo-se obtido um salário médio de 1500 u.m e desvio padrão de 1200 u.m. Os responsáveis pela pesquisa consideram que uma boa estimativa para o salário médio mensal deve ter no máximo um erro de 200 u.m. Determine a probabilidade do erro máximo ser de 200 u.m.

<u>Resolução</u>

Do problema temos:

$$n = 100;\ \bar{x} = 1500; s = 1200;\ \varepsilon \leq 200$$

Nota que a variança populacional é desconhecida e tamanho de amostra grande. Então pelo teorema de limite central temos que:

$$Z \sim N(0; 1)$$

$$\varepsilon = Z_{\frac{\alpha}{2}} * \frac{s}{\sqrt{n}}$$

$$Z_{\frac{\alpha}{2}} = \frac{\varepsilon * \sqrt{n}}{s} = \frac{200 * \sqrt{100}}{1200} = 1.67$$

Usando a tabela Z de distribuição normal temos:

$$\frac{\alpha}{2} = 0.5 - 0.45254 = 0.04746$$

$$\alpha = 2 * 0.04746 = 0.09492$$

$$1 - \alpha = 1 - 0.09492 = 0.90508$$

Resposta: A probabilidade do erro máximo ser de 200 u.m é de 90.5%.

OU

$$P(-200 \le \varepsilon \le 200) = P\left(\frac{-200 * \sqrt{n}}{s} \le Z \le \frac{200 * \sqrt{n}}{s}\right)$$

$$= P\left(\frac{-200 * \sqrt{100}}{1200} \le Z \le \frac{200 * \sqrt{100}}{1200}\right)$$

$$P(-1.67 \le Z \le 1.67) = 2 * 0.45254 = 0.90508$$

Resposta: A probabilidade do erro máximo ser de 200 u.m é de 90.5%.

124. Para estudar a viabilidade de lançamento de um novo produto no mercado, o gerente de uma grande empresa contrata uma firma de consultoria estatística para estudar a aceitação do produto entre os clientes potenciais. O gerente deseja obter uma estimativa com um erro máximo de 1% com probabilidade 80% e pede ao consultor estatístico que forneça o tamanho de amostra necessário.

a) De posse das informações dadas, o consultor calcula o tamanho da amostra necessário no pior cenário. O que significa "pior cenário" nesse caso? Qual o tamanho de amostra obtido pelo consultor?

Resolução

Do problema temos:

$$p = 0.5; \varepsilon \le 0.01; 1 - \alpha = 0.80$$

$$Z_{\frac{\alpha}{2}} = Z_{0.10} = 1.28$$

Pior cenário significa atribuir a probabilidade p de aceitação do produto de 50% a população dos clientes favoráveis ao produto.

Pela fórmula de tamanho de amostra temos:

$$n = \frac{Z^2_{\frac{\alpha}{2}} * p * (1-p)}{\varepsilon^2} = \frac{1.28^2 * 0.5 * 0.5}{0.01^2} = 4096$$

O tamanho de amostra obtido pelo consultor é de pelo menos 4096.

b) O gerente acha que o custo de tal amostra seria muito alto e autoriza o consultor a realizar um estudo piloto com uma amostra de 150 pessoas para obter uma estimativa da verdadeira proporção. O resultado desse estudo piloto é uma estimativa $P = 76\%$ de aceitação do novo produto. Com base nessa estimativa, o consultor recalcula o tamanho da amostra necessário. Qual é esse tamanho?

Resolução

Do problema temos:

$$p = 0.76; \varepsilon \leq 0.01; 1 - \alpha = 0.80$$

$$Z_{\frac{\alpha}{2}} = Z_{0.10} = 1.28$$

$$n = \frac{Z^2{}_{\frac{\alpha}{2}} * p * (1-p)}{\varepsilon^2} = \frac{1.28^2 * 0.76 * 0.24}{0.01^2} = 2988$$

O tamanho de amostra obtido pelo consultor é de pelo menos 2988.

c) Seleccionada a amostra com o tamanho obtido no item anterior (b), obteve-se uma proporção de 72% de clientes favoráveis ao produto. Construa um intervalo de confiança para a verdadeira proporção com nível de confiança de 90%.

Resolução

Do problema temos:

$$p = 0.72; n = 2988; 1 - \alpha = 0.90$$

$$Z_{\frac{\alpha}{2}} = Z_{0.05} = 1.645$$

$$p \pm Z_{\frac{\alpha}{2}} * \sqrt{\frac{p * (1-p)}{n}} \in \pi$$

$$0.72 \pm 1.645 * \sqrt{\frac{0.72 * 0.28}{2988}} \in \pi$$

$$0.71 \leq \pi \leq 0.73$$

Interpretação: A um nível de confiança de 90% pode se afirmar que o intervalo $[0.71 - 0.73]$ contém a verdadeira proporção de clientes favoráveis ao produto.

125. Uma empresa que presta serviços de acessória a outras empresas está interessada em comparar a taxa de reclamações sobre os seus serviços em dois dos seus escritórios em duas cidades diferentes. Suponha que a empresa tenha seleccionado aleatoriamente 100 serviços realizados pelo escritório da cidade A e foi constatado que em 12 deles houve algum tipo de reclamação. Já do escritório da cidade B foram seleccionados 120 serviços e 18 receberam algum tipo de reclamação. Construa um intervalo de confiança a 98% para a diferença entre as taxas de reclamações das duas cidades.

Resolução

Do problema temos:

$$n_A = 100;\ x_A = 12;\ p_A = \frac{12}{100} = 0.12;\ n_B = 120;\ x_B = 18;\ p_B = \frac{18}{120} = 0.15; 1 - \alpha = 0.98$$

$$Z_{\frac{\alpha}{2}} = Z_{0.01} = 2.33$$

Pela fórmula temos:

$$(p_1 - p_2) - Z_{\frac{\alpha}{2}} * \sqrt{\frac{p_1*q_1}{n_1} + \frac{p_2*q_2}{n_2}} \leq \pi_1 - \pi_2 \leq (p_1 - p_2) + Z_{\frac{\alpha}{2}} * \sqrt{\frac{p_1*q_1}{n_1} + \frac{p_2*q_2}{n_2}}$$

$$(0.12 - 0.15) \pm 2.33 * \sqrt{\frac{0.12*0.88}{100} + \frac{0.15*0.85}{120}} \in \pi_1 - \pi_2$$

$$-0.077 \leq \pi_1 - \pi_2 \leq 0.137$$

Interpretação: A um nível de confiança de 98% pode se afirmar que o intervalo $[-0.077; 0.137]$ contém a verdadeira diferença entre as taxas de reclamações das duas cidades.

126. Suponha que as produções (em gramas) de castanhas, em intervalos de tempo fixos, aleatoriamente seleccionados de duas máquinas M1 e M2 de uma fábrica se podem considerar normais. Os pesos obtidos em duas amostras permitiram determinar as quantidades seguintes:

M1	M2
$\sum_{i=1}^{8} x_i = 80.8$	$\sum_{i=1}^{9} y_i = 96.3$
$\sum_{i=1}^{8} x^2{}_i = 816.664$	$\sum_{i=1}^{9} (y_i - \bar{y})^2 = 0.549$

a) Indique uma estimativa pontual para a produção média de cada máquina e respectiva variância.

Resolução

$$\bar{x}_1 = \frac{\sum x_i}{n} = \frac{80.8}{8} = 10.1$$

$$S^2{}_1 = \frac{\sum x^2{}_i - n * \bar{x}^2}{n-1} = \frac{816.664 - 8 * 10.1^2}{7} = 0.0834$$

$$\bar{x}_2 = \frac{\sum y_i}{n} = \frac{96.3}{9} = 10.7$$

$$S^2{}_2 = \frac{\sum (y_i - \bar{y})^2}{n-1} = \frac{0.549}{8} = 0.0686$$

b) Verifique se é plausível considerar que a variabilidade em gramas da produção das duas máquinas é idêntica (use 95% de confiança).

Resolução

A partir do intervalo de confiança para o quociente das variancas temos:

$$\frac{S^2{}_1}{S^2{}_2} * F_{1-\frac{\alpha}{2};\, n_1-1; n_2-1} \leq \frac{\sigma^2{}_1}{\sigma^2{}_2} \leq \frac{S^2{}_1}{S^2{}_2} * F_{\frac{\alpha}{2};\, n_1-1; n_2-1}$$

Usando a tabela F, obteve se os seguintes valores críticos:

$$F_{\frac{\alpha}{2};\, n_1-1; n_2-1} = F_{0.025;\, 7;8} = 4.53$$

$$F_{1-\frac{\alpha}{2};\, n_1-1; n_2-1} = \frac{1}{F_{\frac{\alpha}{2};\, n_2-1; n_1-1;}} = \frac{1}{F_{0.025;\, 8;7}} = \frac{1}{4.90} = 0.204$$

Logo:

$$\frac{0.0834}{0.0686} * 0.204 \leq \frac{\sigma^2{}_1}{\sigma^2{}_2} \leq \frac{0.0834}{0.0686} * 4.53$$

$$0.248 \leq \frac{\sigma^2{}_1}{\sigma^2{}_2} \leq 5.507$$

Conclusão: Como o valor $1 \in [0.248; 5.507]$, *a* 95% *de confiança pode-se concluir que a variabilidade em gramas da produção das duas máquinas é idêntica.*

c) Construa um intervalo de confiança a 95% para a diferença das produções médias entre as duas máquinas.

<u>Resolução</u>

Pelo intem b) temos que $\sigma_1{}^2$ *e* $\sigma_2{}^2$ *são desconhecidas e iguais, e* n_1 *e* n_2 *pequeno então:*

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\frac{\alpha}{2}; n_1+n_2-2} * \sqrt{\left(\frac{(n_1-1)*S_1{}^2+(n_2-1)*S_2{}^2}{n_1+n_2-2}\right) * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \in \mu_1 - \mu_2$$

$$t_{0.025;15}=2.131$$

$$(10.1 - 10.7) \pm 2.131 * \sqrt{\left(\frac{7*0.0834+8*0.0686}{18}\right) * \left(\frac{1}{8} + \frac{1}{9}\right)} \in \mu_1 - \mu_2$$

$$-0.9597 \leq \mu_1 - \mu_2 \leq -0.4403$$

Resposta: A probabilidade de a diferença de produções médias entre as duas máquinas estar no intervalo de $[-0.9597; -0.4403]$ *é de* 95%.

127. Suponha-se em presença de uma população normal, com parâmetros desconhecidos. Com base numa amostra casual, com 16 observações, foi construído o seguinte intervalo de confiança para a média da população:

$$[7.398, 12.602]$$

a) Sabendo que, com a informação da amostra, obteve-se $s = 4$, qual o grau de confiança que pode atribuir ao intervalo atrás referido?

<u>Resolução</u>

$$\varepsilon = t_{\frac{\alpha}{2}; n-1} * \frac{s}{\sqrt{n}}$$

$$\frac{12.602 - 7.398}{2} = t_{\frac{\alpha}{2};15} * \frac{4}{\sqrt{16}}$$

$$t_{\frac{\alpha}{2};15} = 2.602$$

$$\alpha = 0.02$$

$$1 - \alpha = 0.98$$

Resposta: O grau de confiança que pode atribuir ao intervalo atrás referido é de 98%

b) Com base na mesma amostra construa um intervalo de confiança a 95% para a variância da população.

<u>Resolução</u>

$$\frac{(n-1) * S^2}{\chi^2{}_{\frac{\alpha}{2};n-1}} \le \sigma^2 \le \frac{(n-1) * S^2}{\chi^2{}_{1-\frac{\alpha}{2};n-1}}$$

$$\frac{(16-1) * 4^2}{\chi^2{}_{0.025;15}} \le \sigma^2 \le \frac{(16-1) * 4^2}{\chi^2{}_{1-0.025;15}}$$

$$\frac{(16-1) * 4^2}{24.996} \le \sigma^2 \le \frac{(16-1) * 4^2}{6.262}$$

$$9.6 \le \sigma^2 \le 38.3$$

Interpretação: A um nível de confiança de 95% pode se afirmar que o intervalo $[9.6; 38.3]$ *contém a verdadeira variança populacional.*

128. Suponha-se um arquivo contendo 5000 ordens bancárias, normalmente distribuídas, cujo o montante médio é de 1076.39 u.m e desvio padrão de 273.62 u.m. O gerente dum banco pretende estimar o montante médio das contas à ordem. Para tal seleccionou aleatoriamente 100 contas. Qual é a probabilidade de o erro máximo do montante das contas à ordem ser de 52.283 u.m?

<u>Resolução</u>

Do problema temos:

$$N = 5000;\ \mu = 1076.39;\ \sigma = 273.62; n = 100; \varepsilon = 52.283;\ Pede - se\ o\ (1-\alpha)$$

Como a população é finita, partindo do intervalo de confiança para a média quando a variança é conhecida temos:

$$\varepsilon = Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}}$$

$$52.283 = Z_{\frac{\alpha}{2}} * \frac{273.62}{\sqrt{100}} * \sqrt{\frac{5000-100}{5000-1}}$$

$$Z_{\frac{\alpha}{2}} = \frac{52.283 * 10 * \sqrt{4999}}{273.62 * \sqrt{4900}} = 1.93$$

Pela tabela normal padrão teremos:

$$\frac{\alpha}{2} = 0.5 - 0.47320 = 0.0268$$

$$\alpha = 2 * 0.0268 = 0.0536$$

Logo

$$(1-\alpha) = 1 - 0.0536 = 0.9464$$

Resposta: A probabilidade de o erro máximo ser de 52.283 u.m é de 94.64%.

129. Uma fábrica de relógios de alta precisão pretende estudar a fiabilidade da sua produção. É escolhida uma amostra aleatória de 10 relógios. Ao fim de um mês estes relógios são confrontados com um relógio padrão e os seus dados são registados; resulta que a média da amostra é de 0,7 segundos. Admitindo que a distribuição dos erros dos relógios (relativamente ao relógio padrão) é normal com desvio padrão de 0.4 segundos, construa o intervalo de confiança a 90% para a fiabilidade média dos relógios da fábrica?

Resolução

Do problema temos:

$$n = 10;\ \sigma = 0.4; \bar{x} = 0.7; (1-\alpha) = 0.90;\ Z_{\frac{\alpha}{2}} = Z_{0.05} = 1.64;\ Pede\ IC_{\mu}$$

Uma vez que a variança populacional é conhecida e a população é infinita, a fórmula para o cálculo de intervalo de confiança da média é:

$$\bar{x} - Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}}$$

$$0.7 - 1.64 * \frac{0.4}{\sqrt{10}} \leq \mu \leq 0.7 + 1.64 * \frac{0.4}{\sqrt{10}}$$

$$0.4926 \leq \mu \leq 0.9074$$

Resposta: A um nível de confiança de 90% pode-se afirmar que o intervalo [0.4926;0.9074] contém a verdadeira fiabilidade média dos relógios da fábrica.

130. Realizou-se uma pesquisa de mercado em duas cidades distintas (A e B) com o objectivo de estudar o tempo necessário para um consumidor tomar uma decisão sobre o que comprar entre embalagens alternativas do mesmo produto. Eliminaram-se as marcas das embalagens, a fim de reduzir o efeito da preferência por uma ou outra marca. Os consumidores fizeram suas escolhas somente com base na descrição do produto, anotada nas embalagens pelos fabricantes. Suponha que o tempo necessário para os consumidores decidir sobre o que comprar segue uma distribuição normal.

a) Quantos consumidores devem ser seleccionados em cada cidade para que o erro máximo não exceda 3.3 segundos com um nível de confiança de 95%? Considere que dados de um estudo piloto revelaram varianças de 25 segundos para cidade A e 36 segundos na Cidade B.

Resolução

Do problema temos:

$$\sigma_A{}^2 = 25;\ \sigma_B{}^2 = 36;\ \varepsilon = 3.3; (1-\alpha) = 0.95;\ Z_{0.025} = 1.96;\ Pede\ se\ n_A\ e\ n_B$$

A fórmula para o cálculo de tamanho de amostra quando a variança é conhecida e população infinita é:

$$n = \left(\frac{Z_{\frac{\alpha}{2}} * \sigma}{\varepsilon}\right)^2$$

$$Então:$$

$$n_A = \left(\frac{1.96 * 5}{3.3}\right)^2 = 8.8$$

$$n_B = \left(\frac{1.96 * 6}{3.3}\right)^2 = 12.7$$

Resposta: Para que o erro máximo não exceda 3.3 segundos com um nível de confiança de 95%, devem ser seleccionados 9 consumidores da cidade A e 13 da cidade B.

b) Dos dados da amostragem dos consumidores da cidade A obteve-se um tempo médio de 16 segundos e variança de 23 segundos, enquanto que na cidade B o tempo médio foi de 14.5 segundos com variança de 28 segundos. Construa um intervalo de confiança a 95% para a diferença dos tempos médios necessários que os consumidores das duas cidades levam para a tomar a decisão.

Resolução:

Do problema temos:

$$n_A = 9;\ \bar{x}_A = 16;\ S_A{}^2 = 23;\ n_B = 13;\ \bar{x}_B = 14.5;\ S_B{}^2 = 28;\ (1-\alpha) = 0.95$$

Nota que nada sabe-se pelas varianças populacionais. Assim para determinar o intervalo de confiança para diferença das médias temos que verificar se as populações são homogéneas ou não, através de intervalo de confiança para o quociente das varianças.

Assim:

$$\frac{S^2{}_A}{S^2{}_B} * F_{1-\frac{\alpha}{2};\, n_A-1;\, n_B-1} \le \frac{\sigma^2{}_A}{\sigma^2{}_B} \le \frac{S^2{}_A}{S^2{}_B} * F_{\frac{\alpha}{2};\, n_A-1;\, n_B-1}$$

Usando a tabela F, obteve-se os seguintes valores críticos:

$$F_{\frac{\alpha}{2};\, n_A-1; n_B-1} = F_{0.025;\, 8;12} = 3.51$$

$$F_{1-\frac{\alpha}{2};\, n_A-1;n_B-1} = \frac{1}{F_{\frac{\alpha}{2};\, n_B-1;n_A-1}} = \frac{1}{F_{0.025;\,12;8}} = \frac{1}{4.20} = 0.2381$$

Logo:

$$\frac{23}{28} * 0.2381 \leq \frac{{\sigma^2}_1}{{\sigma^2}_2} \leq \frac{23}{28} * 3.51$$

$$0.1956 \leq \frac{{\sigma^2}_1}{{\sigma^2}_2} \leq 2.8832$$

Conclusão: Como o valor $1 \in [0.1956; 2.8832]$*, a* 95% *de confiança pode-se concluir que as populações das duas cidades apresentam tempos homogéneos, isto é,* ${\sigma^2}_A = {\sigma^2}_B$ *.*

Como ${\sigma_1}^2$ *e* ${\sigma_2}^2$ *são desconhecidas e iguais, e* n_1 *e* n_2 *pequeno então:*

$$(\bar{x}_A - \bar{x}_B) \pm t_{\frac{\alpha}{2};n_A+n_B-2} * \sqrt{\left(\frac{(n_A-1)*{S_A}^2+(n_B-1)*{S_B}^2}{n_A+n_B-2}\right) * \left(\frac{1}{n_A}+\frac{1}{n_B}\right)} \in \mu_A - \mu_B$$

$$t_{0.025;20} = 2.086$$

$$(16 - 14.5) \pm 2.086 * \sqrt{\left(\frac{8*23+12*28}{20}\right) * \left(\frac{1}{9}+\frac{1}{13}\right)} \in \mu_A - \mu_B$$

$$-3.1123 \leq \mu_A - \mu_B \leq 6.1123$$

Resposta: A probabilidade de a diferença do desempenho médio entre as duas marcas de veículos estar no intervalo de $[-3.1123; 6.1123]$ *é de* 95%.

131. Duas espécies de um certo tipo de cereal estão sendo testadas quanto ao seu crescimento. O experimento foi feito escolhendo 10 blocos de terreno e plantando em cada bloco mudas de ambas as espécies. Os resultados a seguir são as alturas medidas ao final do primeiro mês.

Terreno	1	2	3	4	5	6	7	8	9	10
Espécie 1	22	27	18	33	25	21	15	33	21	24
Espécie 2	21	31	24	32	29	23	19	37	22	27

Os dados deste experimento foram colectados aos pares para impedir que as diferenças de fertilidade entre os blocos de terreno (que podem ser grandes) mascarem os resultados. Conclua a 94% de confiança quanto a diferença das alturas medidas entre as duas espécies.

Resolução:

Nota que aqui trata-se de amostras emparelhadas de amostra pequena. A fórmula para o intervalo de confiança para a diferença é:

$$\bar{d} - t_{\frac{\alpha}{2};n-1} * \frac{S_d}{\sqrt{n}} \leq \mu_d \leq \bar{d} + t_{\frac{\alpha}{2};n-1} * \frac{S_d}{\sqrt{n}}$$

A tabela a seguir fornece os valores da diferença d:

Terreno	1	2	3	4	5	6	7	8	9	10
Espécie 1	22	27	18	33	25	21	15	33	21	24
Espécie 2	21	31	24	32	29	23	19	37	22	27
Diferença (d)	1	4	6	1	4	2	4	4	1	3

$$Dos\ dados\ temos:$$

$$\bar{d} = 3;\ S_d = 1.7; n = 10;\ t_{\frac{\alpha}{2};n-1} = t_{0.03;9} = 2.150$$

$$3 - 2.150 * \frac{1.7}{\sqrt{10}} \leq \mu_d \leq 3 + 2.150 * \frac{1.7}{\sqrt{10}}$$

$$1.8442 \leq \mu_d \leq 4.1558$$

Conclusão: Como o valor zero não pertence ao intervalo $[1.8442; 4.1558]$, a um nível de confiança de 94% pode se concluir que as alturas medidas entre as duas espécies não é a mesma.

132. Um fornecedor apresenta uma caixa, e afirma que o peso médio desta caixa é de 368 gramas. De experiências anteriores sabe-se que o desvio padrão vale 15g e que os valores se comportam segundo a distribuição Normal. Para verificar se a afirmação é verdadeira, verifica-se uma amostra de 25 caixas, pesa-se e calcula-se o peso médio da amostra, achando 372,5g. Construa um intervalo de confiança a 95% para o verdadeiro peso médio das caixas?

<u>Resolução</u>

Do problema temos:

$$\mu = 368;\ \ \sigma = 15; n = 25;\ \bar{x} = 372.5; (1-\alpha) = 0.95;\ Z_{\frac{\alpha}{2}} = Z_{0.05} = 1.96;\ Pede\ IC_{\mu}$$

Uma vez que a variança populacional é conhecida e a população é infinita, a fórmula para o cálculo de intervalo de confiança da média é:

$$\bar{x} - Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}}$$

$$372.5 - 1.96 * \frac{15}{\sqrt{25}} \leq \mu \leq 372.5 + 1.96 * \frac{15}{\sqrt{25}}$$

$$366.6 \leq \mu \leq 378.4$$

Resposta: A um nível de confiança de 95% pode-se afirmar que o intervalo [366.6; 378.4] contém o verdadeiro peso médio das caixas.

133. O gestor de uma consultoria multinacional de auditoria pretende comparar as taxas de satisfação entre funcionários a tempo integral e parcial quanto a sua política de remuneração salarial. Como é impossível entrevistar todos os funcionários a tempo integral e tempo parcial em tempo razoável, o gestor decide fazer uma amostragem aleatória simples dos funcionários.

a) Determinar o tamanho de amostra para cada grupo de funcionários necessário para estimar a taxa de satisfação com uma precisão 4% e nível de confiança de 95%. Assumir que um estudo piloto revelou 68% e 55% de satisfação de funcionários a tempo integral e parcial respectivamente.

Resolução

Uma vez que a população é infinita, a fórmula para o cálculo de amostra é:

$$n = \frac{Z^2{}_{\frac{\alpha}{2}} * p * (1-p)}{\varepsilon^2}$$

Seja 1 funcionários a tempo integral e 2 funcionários a tempo inteiro;

$$v.a: x = \text{nº } de\ funcionários\ satisfeitos\ com\ politica\ de\ remuneração\ sarial$$

Do problema temos:

$$p_1 = 0.68;\ \ p_2 = 0.55;\ \varepsilon = 0.04; 1-\alpha = 0.95;\ Z_{0.025} = 1.96$$

Então:

$$n_1 = \frac{1.96^2 * 0.68 * (1-0.68)}{0.04^2} = 522.4576$$

$$n_2 = \frac{1.96^2 * 0.55 * (1-0.55)}{0.04^2} = 594.2475$$

Resposta: O tamanho de amostra necessário para estimar a taxa de satisfação com uma precisão 4% e nível de confiança de 95% é de 522 para funcionários a tempo inteiro e 594 a tempo parcial.

b) Supondo que a partir da amostragem obteve-se 193 funcionários do tempo integral insatisfeitos e 271 do tempo parcial satisfeitos, determine o intervalo de confiança a 97% para a diferença das proporções de funcionários a tempo integral e parcial, satisfeitos com a política de remuneração salarial.

Resolução

Do problema temos:

$$n_1 = 522;\ x_1 = 290;\ p_1 = \frac{x_1}{n_1} = \frac{329}{522} = 0.630;\ n_2 = 594;\ x_2 = 271;\ p_2 = \frac{x_2}{n_2} = \frac{271}{594} = 0.456;$$

$$1-\alpha = 0.95;\ Z_{0.015} = 2.17$$

Pela fórmula temos:

$$(p_1 - p_2) - Z_{\frac{\alpha}{2}} * \sqrt{\frac{p_1 * q_1}{n_1} + \frac{p_2 * q_2}{n_2}} \leq \pi_1 - \pi_2 \leq (p_1 - p_2) + Z_{\frac{\alpha}{2}} * \sqrt{\frac{p_1 * q_1}{n_1} + \frac{p_2 * q_2}{n_2}}$$

$$(0.630 - 0.456) \pm 2.17 * \sqrt{\frac{0.630 * 0.370}{522} + \frac{0.456 * 0.544}{594}} \in \pi_1 - \pi_2$$

$$0.1102 \leq \pi_1 - \pi_2 \leq 0.2378$$

Interpretação: A um nível de confiança de 97% pode se afirmar que o intervalo $[0.1102; 0.2378]$ contém a verdadeira diferença das proporções de funcionários a tempo integral e parcial, satisfeitos com a política de remuneração salarial.

134. Um órgão fiscalizador de impostos pretende comparar o número de renda pessoal tributável de funcionários de duas empresas distintas de construção civil. Para tal seleccionou aleatoriamente 16 funcionários da empresa A e 13 da empresa B tendo obtido os seguintes números de rendas tributáveis (em milhares):

Empreza A	192	80	162	100	164	122	99	142
	111	132	124	137	98	104	120	131
Empreza B	98	112	104	168	121	86	108	115
	109	130	127	110	152			

Construa um intervalo de confiança a 95% para a diferença do número de renda pessoal tributável de funcionários das duas empresas.

Resolução:

Do problema temos:

$n_A = 16;\ \bar{x}_A = 126.125;\ S_A{}^2 = 840.25;\ n_B = 13;\ \bar{x}_B = 118.4615;\ S_B{}^2 = 483.1026;\ (1-\alpha) = 0.95$

Nota que nada sabe-se sobre as varianças populacionais. Assim para determinar o intervalo de confiança para diferença das médias temos que verificar se as populações são homogéneas ou não, através de intervalo de confiança para o quociente das varianças.

Assim:

$$\frac{S^2{}_A}{S^2{}_B} * F_{1-\frac{\alpha}{2};\, n_A-1;\, n_B-1} \leq \frac{\sigma^2{}_A}{\sigma^2{}_B} \leq \frac{S^2{}_A}{S^2{}_B} * F_{\frac{\alpha}{2};\, n_A-1;\, n_B-1}$$

Usando a tabela F, obteve-se os seguintes valores críticos:

$$F_{\frac{\alpha}{2};\, n_A-1; n_B-1} = F_{0.025;\, 15;12} = 3.18$$

$$\mathrm{F}_{1-\frac{\alpha}{2};\, n_A-1; n_B-1} = \frac{1}{\mathrm{F}_{\frac{\alpha}{2};\, n_B-1; n_A-1}} = \frac{1}{F_{0.025;\, 12;15}} = \frac{1}{2.960} = 0.3378$$

Logo:

$$\frac{840.25}{483.1026} * 0.3378 \leq \frac{\sigma^2{}_1}{\sigma^2{}_2} \leq \frac{840.25}{483.1026} * 3.18$$

$$0.3358 \leq \frac{\sigma^2{}_1}{\sigma^2{}_2} \leq 5.5309$$

Conclusão: Como o valor $1 \in [0.3358; 5.5309]$, a 95% de confiança pode-se concluir que as populações das números de rendas tributáveis (em milhares) dos funcionários das duas empresas são homogéneas, isto é, $\sigma^2{}_1 = \sigma^2{}_2$.

Como $\sigma_1{}^2$ e $\sigma_2{}^2$ são desconhecidas e iguais, e n_1 e n_2 *pequeno então:*

$$(\bar{x}_A - \bar{x}_B) \pm t_{\frac{\alpha}{2};n_A+n_B-2} * \sqrt{\left(\frac{(n_A-1)*S_A{}^2+(n_B-1)*S_B{}^2}{n_A+n_B-2}\right) * \left(\frac{1}{n_A} + \frac{1}{n_B}\right)} \in \mu_A - \mu_B$$

$$t_{0.025;27} = 2.052$$

$$(126.125 - 118.4615) \pm 2.052 * \sqrt{\left(\frac{15*840.25+12*0.835}{27}\right) * \left(\frac{1}{16} + \frac{1}{13}\right)} \in \mu_A - \mu_B$$

$$-8.8975 \leq \mu_A - \mu_B \leq 24.2245$$

Resposta: A probabilidade de a diferença do desempenho médio entre o número de renda pessoal tributável de funcionários das duas empresas *estar no intervalo de* $[-8.8975; 24.2245]$ *é de 95%.*

135. Um clube deseja estimar a proporção de crianças que tem um animal de estimação. Se o clube deseja que a estimativa esteja no máximo afastada 3 % da proporção populacional, quantas crianças devem conter a amostra? Assuma um intervalo de confiança de 95 % e que o clube estimou, com base em experiência anterior, que aproximadamente 30 % das crianças têm um animal de estimação.

<u>Resolução</u>

Uma vez que a população é infinita, a fórmula para o cálculo de amostra é:

$$n = \frac{Z^2{}_{\frac{\alpha}{2}} * p * (1-p)}{\varepsilon^2}$$

Seja 1 funcionários a tempo integral e *2 funcionários a tempo inteiro;*

$$v.a: x = \text{nº } de\ crianças\ têm\ um\ animal\ de\ estimação$$

Do problema temos:

$$p = 0.30;\ \ q = 0.70;\ \varepsilon = 0.03; 1 - \alpha = 0.95;\ Z_{0.025} = 1.96$$

Então:

$$n = \frac{1.96^2 * 0.30 * (1 - 0.30)}{0.03^2} = 894$$

136. O proprietário de uma empresa de venda de produtos metálicos, está preocupado com o tempo gasto pelos operários estagiários no processamento de ferro. O proprietário decide submeter os estagiários a um teste de aptidão para a execução da actividade. Para tal foi seleccionada uma amostra de 25 operários, tendo se obtido um tempo médio de execução de 88.2 minutos e desvio

padrão de 20 minutos. Construa um intervalo de confiança para a variança. Use nível de confiança de 95%.

Resolução

$$n = 25;\ \bar{x} = 88.2;\ s = 20;\ \mu_0 = 80;\ 1 - \alpha = 0.95; Pede\ se\ IC_{\sigma^2}$$

Sabe se que:

$$\chi^2 \sim {\chi^2}_{n-1}$$

Recorrendo a fórmula do intervalo de confiança para variança temos:

$$\frac{(n-1)*S^2}{{\chi^2}_{\frac{\alpha}{2};n-1}} \leq \sigma^2 \leq \frac{(n-1)*S^2}{{\chi^2}_{1-\frac{\alpha}{2};n-1}}$$

Nota que trata-se de um intervalo bilateral, pela tabela χ^2 temos:

$${\chi^2}_{\frac{\alpha}{2};n-1} = {\chi^2}_{0.025;24} = 39.364\ e\ {\chi^2}_{1-\frac{\alpha}{2};n-1} = {\chi^2}_{0.975;24} = 12.401$$

Então:

$$\frac{24*400}{39.364} \leq \sigma^2 \leq \frac{24*400}{12.401}$$

$$243.9 \leq \sigma^2 \leq 774.1$$

Resposta: A um nível de confiança de 95% pode-se afirmar que o intervalo [243.9-774.1] contém a verdadeira dispersão do tempo da execução da actividade.

137. Um analista de um departamento de pessoal de uma firma com 2000 trabalhadores, pretende estimar a taxa média dos salários por hora desses trabalhadores. Os registos mostram uma variabilidade dos salários de 100 u.m. Supõe-se que os salários da firma sejam distribuídos normalmente.

a) Quantos trabalhadores devem ser seleccionados, a fim de que o erro de estimativa dos salários médio seja no máximo de 2 u.m, com probabilidade igual a 96%?

Resolução

Pelos dados temos:

$$N = 2000;\ \sigma^2 = 100; \varepsilon \leq 2; (1-\alpha) = 0.96;\ Z_{0.02} = 2.05;\ Pede - se\ o\ n:$$

Como a população é finita, partindo do intervalo de confiança para a média quando a variança é conhecida temos:

$$n = \frac{{z^2}_{\frac{\alpha}{2}} * \sigma^2 * N}{\varepsilon^2 * (N-1) + {z^2}_{\frac{\alpha}{2}} * \sigma^2}$$

$$n \geq \frac{2.05^2 * 100 * 2000}{2^2 * (2000-1) + 2.05^2 * 100} \geq 99.9$$

Resposta: A fim de que o erro de estimativa dos salários médio seja no máximo de 2 u.m, com probabilidade igual a 96%, devem ser seleccionados pelo menos 100 trabalhadores.

b) Seleccionada a amostra com o tamanho obtido no item anterior, obteve-se a taxa de salários por hora de 75 u.m. Construa um intervalo de confiança para a verdadeira taxa média dos salários com nível de confiança de 90%.

Resolução

Pelos dados temos:

$$\bar{x} = 75;\ N = 2000;\ \sigma = 10; (1-\alpha) = 0.90;\ Z_{0.10} = 1.645;\ Pede\ IC_{\mu}$$

Como a população é finita, pela fórmula do intervalo de confiança para a média quando a variança é conhecida temos:

$$\bar{x} - Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}} \leq \mu \leq \bar{x} + Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}}$$

$$75 - 1.645 * \frac{10}{\sqrt{100}} * \sqrt{\frac{2000-100}{2000-1}} \leq \mu \leq 75 - 1.645 * \frac{10}{\sqrt{100}} * \sqrt{\frac{2000-100}{2000-1}}$$

$$73.4 \leq \mu \leq 76.6$$

Interpretação: A um nível de confiança de 90% pode-se afirmar que o intervalo [73.4-76.6] contém a verdadeira taxa média dos salários.

c) Qual é a probabilidade de a verdadeira taxa média dos salários estar entre 73.6 e 76.4 u.m?

Resolução

Do problema temos:

$$N = 2000;\ \sigma = 10; n = 100; \varepsilon = \frac{At}{2} = \frac{76.4 - 73.6}{2} = 1.4;\ Pede - se\ o\ (1-\alpha)$$

Como a população é finita, partindo do intervalo de confiança para a média quando a variança é conhecida temos:

$$\varepsilon = Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} * \sqrt{\frac{N-n}{N-1}}$$

$$1.4 = Z_{\frac{\alpha}{2}} * \frac{10}{\sqrt{100}} * \sqrt{\frac{2000-100}{2000-1}}$$

$$Z_{\frac{\alpha}{2}} = \frac{1.4 * \sqrt{100} * \sqrt{1999}}{10 * \sqrt{1900}} = 1.436$$

Pela tabela normal padrão teremos:

$$\frac{\alpha}{2} = 0.5 - 0.42507 = 0.07493$$

$$\alpha = 2 * 0.07493 = 0.14986$$

Logo

$$(1-\alpha) = 1 - 0.14986 = 0.85014$$

Resposta: A probabilidade de a verdadeira taxa média dos salários estar entre 73.6 e 76.4 u.m é de 85%.

138. A Leishmaniose Visceral é uma doença perigosa e que, se não for tratada correctamente, pode levar a óbito. Todo caso diagnosticado de Leishmaniose Visceral deve ser notificado às autoridades de saúde. Em estudo sobre o número de dias entre o início dos sintomas da Leishmaniose Visceral e a notificação do caso às autoridades, uma pesquisadora deseja estimar o número médio de dias entre os sintomas e a notificação usando um intervalo de 90% de confiança. Sabendo que ela gostaria que o erro de estimação fosse a metade do desvio-padrão do número de dias e supondo que o número de dias entre o início dos sintomas e a notificação tenha distribuição Gaussiana. Estima-se que a população com casos de Leishmaniose Visceral é de 2600. Quantos casos de Leishmaniose Visceral, no mínimo, ela deve estudar?

Resolução

Pelos dados temos:

$$N = 2600; \varepsilon = \frac{\sigma}{2}; (1-\alpha) = 0.90;\ Z_{0.02} = 1.645;\ Pede - se\ o\ n:$$

Como a população é finita, partindo do intervalo de confiança para a média quando a variança é conhecida temos:

$$n = \frac{z^2{}_{\frac{\alpha}{2}} * \sigma^2 * N}{\varepsilon^2 * (N-1) + z^2{}_{\frac{\alpha}{2}} * \sigma^2}$$

$$n \geq \frac{1.645^2 * \sigma^2 * 2600}{\left(\frac{\sigma}{2}\right)^2 * (2600-1) + 1.96^2 * \sigma^2} = 10.8$$

Resposta: No mínimo, ela deve estudar 11 casos de Leishmaniose Visceral.

139. As distribuições da pressão arterial sistólica e diastólica para mulheres entre 30 e 34 anos têm distribuições normais de médias desconhecidas. Uma amostra aleatória de 10 mulheres é seleccionada dessa população obtendo-se uma pressão arterial sistólica média de 130 mmHg e desvio padrão de 9.1 mmHg; enquanto que para mesma amostra, a pressão arterial diastólica média foi de 84 mmHg com desvio padrão de 11.8 mmHg. Usando um nível de confiança de 95% determine:

   a) O intervalo de confiança para o quociente das varianças entre a pressão arterial sistólica e diastólica desse grupo de mulheres e conclua quanto a homogeneidade.

Resolução

Seja

$$1 \ = \ pressão\ arterial\ sistólica\ e\ 2 \ = \ pressão\ arterial\ diastólica$$

Pelos dados temos:

$$n_1 = n_2 = 10; s_1 = 9.1;\ s_2 = 11.8;\ (1-\alpha) = 0.95$$

A partir do intervalo de confiança para o quociente das varianças temos:

$$\frac{S^2{}_1}{S^2{}_2} * F_{1-\frac{\alpha}{2};\, n_1-1;n_2-1} \leq \frac{\sigma^2{}_1}{\sigma^2{}_2} \leq \frac{S^2{}_1}{S^2{}_2} * F_{\frac{\alpha}{2};\, n_1-1;n_2-1}$$

Usando a tabela F, obteve se os seguintes valores críticos:

$$F_{\frac{\alpha}{2};\, n_1-1;n_2-1} = F_{0.025;\,9;9} = 4.026$$

$$F_{1-\frac{\alpha}{2};\, n_1-1;n_2-1} = \frac{1}{F_{\frac{\alpha}{2};\, n_2-1;n_1-1;}} = \frac{1}{F_{0.025;\,9;9}} = \frac{1}{4.026} = 0.248$$

Logo:

$$\left(\frac{9.1}{11.8}\right)^2 * 0.248 \leq \frac{\sigma^2{}_1}{\sigma^2{}_2} \leq \left(\frac{9.1}{11.8}\right)^2 * 4.026$$

$$0.1475 \leq \frac{\sigma^2{}_1}{\sigma^2{}_2} \leq 2.3944$$

Interpretação: A um nível de confiança de 95%, pode se concluir que o intervalo $[0.1475 - 2.3944]$ contém o quociente das variranças entre a pressão arterial sistólica e diastólica desse grupo de mulheres. Como o valor $1\ \in\ [0.1475\ -\ 2.3944]$, não se pode descartar a hipótese de as varianças serem as mesmas, isto é, as varianças são homogéneas.

b) Construa um intervalo de confiança a 95% para a diferença das produções médias entre as duas máquinas.

<u>Resolução</u>

Pelos dados temos:

$$\bar{x}_1 = 130;\ \bar{x}_2 = 84;\ n_1 = n_2 = 10; s_1 = 9.1;\ s_2 = 11.8;\ (1-\alpha) = 0.95$$

Pelo intem a) temos que ${\sigma_1}^2$ e ${\sigma_2}^2$ são desconhecidas e iguais, e n_1 e n_2 pequeno então:

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\frac{\alpha}{2};n_1+n_2-2} * \sqrt{\left(\frac{(n_1-1)*{S_1}^2+(n_2-1)*{S_2}^2}{n_1+n_2-2}\right) * \left(\frac{1}{n_1}+\frac{1}{n_2}\right)} \in \mu_1 - \ \mu_2$$

$$t_{0.025;18}=2.101$$

$$(130-84) \pm 2.101 * \sqrt{\left(\frac{9*9.1^2+9*11.8^2}{18}\right) * \left(\frac{1}{9}+\frac{1}{9}\right)} \in \mu_1 - \ \mu_2$$

$$36.1 \leq \mu_1 - \ \mu_2 \leq 55.9$$

Resposta: Temos 95% de confiança de que a verdadeira diferença da pressão arterial sistólica e diastólica desse grupo de mulheres esta entre 36.1 a 55.9 mmHg.

140. O intervalo $[35.21; 35.99]$ é o intervalo de confiança de 95%, construído a partir de uma amostra de tamanho 25, para a média µ de uma população Normal com variança desconhecida.

a) Qual o valor do desvio padrão?

Resolução

Recorrendo a fórmula de ε para o intervalo de confiança para média quando a variança é desconhecida e tamanho de amostra pequena, temos:

$$\varepsilon = t_{\frac{\alpha}{2};n-1} * \frac{s}{\sqrt{n}}$$

Pelos dados temos:

$$t_{\frac{0.05}{2};25-1} = 2.064;\ \varepsilon = \frac{35.99 - 35.21}{2} = 0.39$$

$$s = \frac{\varepsilon * \sqrt{n}}{t_{\frac{\alpha}{2};n-1}} = \frac{0.39 * \sqrt{25}}{2.064} = 0.945$$

b) Determine o intervalo de confiança para variança.

Resolução

Recorrendo a fórmula do intervalo de confiança para variança temos:

$$\frac{(n-1) * S^2}{{\chi^2}_{\frac{\alpha}{2};n-1}} \leq \sigma^2 \leq \frac{(n-1) * S^2}{{\chi^2}_{1-\frac{\alpha}{2};n-1}}$$

Nota que trata-se de um intervalo bilateral, pela tabela χ^2 temos:

$${\chi^2}_{\frac{\alpha}{2};n-1} = {\chi^2}_{0.025;24} = 39.364 \ e\ {\chi^2}_{1-\frac{\alpha}{2};n-1} = {\chi^2}_{0.975;24} = 12.401$$

Então:

$$\frac{24 * 0.945^2}{39.364} \leq \sigma^2 \leq \frac{24 * 0.945^2}{12.401}$$

$$0.5444 \leq \sigma^2 \leq 1.7282$$

Resposta: A um nível de confiança de 95% pode-se afirmar que o intervalo [243.9-774.1] contém a verdadeira dispersão.

141. Um empresário deseja conhecer a satisfação de seus clientes em relação aos serviços prestados por sua empresa. Em uma amostra aleatória de 100 clientes entrevistados, 4 pessoas demonstraram insatisfação com os serviços prestados. Construa um intervalo de 95% de confiança para a proporção de clientes insatisfeitos.

Resolução

Pelos dados temos:

$$Z = \frac{p-\pi}{\sqrt{\frac{\pi*(1-\pi)}{n}}} \sim N(0;1)$$

Já que $1-\alpha = 0.95$, *temos da tabela normal padrão* $Z_{0.975} = 1.96$

$$p = \frac{4}{100} = 0.04\,; n = 100$$

Substituindo na fórmula temos:

$$p - Z_{\frac{\alpha}{2}} * \sqrt{\frac{p*(1-p)}{n}} \leq \pi \leq p + Z_{\frac{\alpha}{2}} * \sqrt{\frac{p*(1-p)}{n}}$$

$$0.04 - 1.96 * \sqrt{\frac{0.04*(1-0.04)}{100}} \leq \pi \leq 0.04 + 1.96 * \sqrt{\frac{0.04*(1-0.04)}{100}}$$

$$0.03 \leq \pi \leq 0.05$$

142. Um fornecedor alega que entrega 10% de produtos defeituosos. Qual o tamanho de amostra suficiente para estimar a proporção de produtos defeituosos entregues por este fornecedor com precisão de 0.03 e 95% de confiança?

Resolução

Já que $1-\alpha = 0.95$, *temos da tabela normal padrão* $Z_{0.975} = 1.96$

$$p = 0.10\,; \varepsilon = 0.03$$

$$n = \frac{Z^2{}_{\frac{\alpha}{2}} * p*(1-p)}{\varepsilon^2} = \frac{1.96^2 * 0.10 ** (1-0.10)}{0.03^2} = 384.16$$

Logo, é necessário uma amostra de 385 produtos.

143. Uma empresa quer verificar se o conhecimento de seus alunos a respeito de um determinado assunto melhorou após 30 horas de treinamento. Para isso foi realizado com os quinze alunos do treinamento um teste antes e após o treinamento. Os dados a seguir representam as notas obtidas pelos alunos. Conclua a respeito da eficiência do treinamento com 95% de confiança.

Alunos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Antes	6.5	6.7	7.0	7.0	6.5	7.3	7.8	6.9	6.7	7.2	7.5	7.5	7.2	7.0	6.8
Depois	7.5	7.7	7.9	8.0	7.4	8.3	8.8	8.9	7.7	8.2	8.5	8.5	8.2	8.0	8.8
Diferença	1.0	1.0	0.9	1.0	0.9	1.0	1.0	2.0	1.0	1.0	1.0	1.0	1.0	1.0	2.0

Resolução

Pelos dados temos:

$$\bar{d} = 1.12\,;\ S_d = 0.36;\ t_{0.025;14} = 2.145$$

$$\bar{d} - t_{\frac{\alpha}{2};n-1} * \frac{S_d}{\sqrt{n}} \leq \mu_d \leq \bar{d} - t_{\frac{\alpha}{2};n-1} * \frac{S_d}{\sqrt{n}}$$

$$1.12 - 2.145 * \frac{0.36}{\sqrt{15}} \leq \mu_d \leq 1.12 + 2.145 * \frac{0.36}{\sqrt{15}}$$

$$0.92 \leq \mu_d \leq 1.32$$

Conclusão: Como o valor zero não está incluído no intervalo, rejeita-se a hipótese de que as notas antes e depois sejam as mesmas, logo conclui-se que o treinamento foi eficiente.

144. Um eixo deve ser montado no interior de um rolamento. Uma amostra de doze unidades indicou para o diâmetro inteiro do rolamento $\bar{x}_1 = 2.538\ cm\ e\ S_1 = 0.008\ cm$; e para o diâmetro do eixo $\bar{x}_2 = 2.520\ cm\ e\ S_2 = 0.006\ cm$. Calcule o intervalo de confiança de 99% para a folga de montagem.

<u>Resolução</u>

Supondo que as variânças são iguais têm-se:

$$S_p{}^2 = \frac{(n_1 - 1) * S_1{}^2 + (n_2 - 1) * S_2{}^2}{n_1 + n_2 - 2}$$

$$S_p{}^2 = \frac{(12 - 1) * 0.008^2 + (12 - 1) * 0.006^2}{12 + 12 - 2} = 0.000050$$

$$Graus\ de\ liberdade = v = n_1 + n_2 - 2 = 22$$

$$t_{0.005;22} = 2.82$$

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2};v} * \sqrt{S_p{}^2 * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \leq \mu_1 - \mu_2 \leq (\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2};v} * \sqrt{S_p{}^2 * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$(2.538 - 2.52) \pm 2.82 * \sqrt{0.000050 * \left(\frac{1}{12} + \frac{1}{12}\right)} \in \mu_1 - \mu_2$$

$$0.00986 \leq Folga \leq 0.026$$

145. Uma firma construtora deseja estimar a resistência média das barras de aço utilizadas na construção de casas. Qual o tamanho amostral necessário para garantir que haja um risco de 0,001 de ultrapassar um erro de 5 kg ou mais na estimação? O desvio padrão da resistência para este tipo de barra é de 25 kg.

<u>Resolução</u>

Pelos dados temos:

$$\sigma = 25;\ \alpha = 0.001;\ \varepsilon = 5;\ Z_{0.9995} = 3.29$$

$$\varepsilon = Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} \rightarrow n = \left(\frac{3.29 * 25}{5}\right)^2 \approx 271$$

TESTE DE HIPÓTESES

146. Um auditor deseja testar a hipótese de que o valor médio de todas as contas a receber de uma dada firma é, no mínimo 26,000.00 Mt. Tomando para tanto uma amostra de n = 36 e supondo que a média amostral é 24,000.00 Mt. Testar a hipótese ao nível de confiança de 98%, dado que se conhece o desvio padrão dos valores das contas a receber, isto é, $\sigma = 4{,}300.00\ Mt$.

Resolução:

Do problema temos:

$$\mu_0 = 26000; \sigma_0 = 4300;\ n = 36;\ \bar{x} = 24000\ e\ \alpha = 0.02$$

1° passo: Formular hipóteses

$$\begin{cases} H_0: \mu_0 \geq 26000\ (o\ valor\ médio\ de\ todas\ as\ contas\ que\ o\ auditor\ irá\ receber\ é\ no\ mínimo\ 26{,}000.00Mt) \\ H_1: \mu_0 < 26000\ (o\ valor\ médio\ de\ todas\ as\ contas\ que\ o\ auditor\ irá\ receber\ é\ inferior\ a\ 26{,}000.00Mt) \end{cases}$$

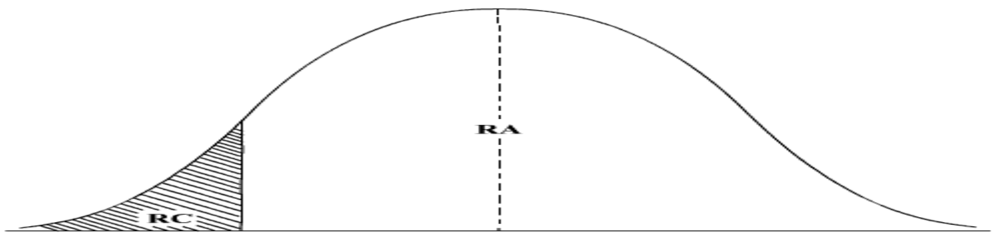
2° passo: Fixar o limite de erro α, e identificar a variável de teste

Para $\alpha = 0.02;\ \bar{x} \sim N\left(\mu =; \frac{\sigma^2}{n}\right)$

Z: variável normal padrão

3° passo: Com o auxílio da tabela de distribuíção normal padrão, determinar a RC (região crítica) e RA (região de aceitação) para H_0

Trata-se de um teste unicaudal esquerdo, logo $-Z_\alpha = -Z_{crítico} = -Z_{0.02} = -2.06$



Se $Z_{calculado} > -Z_{crítico}$, não-se rejeita o H_0

4° passo: *Com os elementos amostrais, calcular o valor da variável do teste*

$$Z_{calculado} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{24000 - 26000}{\frac{4300}{\sqrt{36}}} = -2.79$$

5° passo: Tomada de decisão

Como $Z_{calculado}$ =-2.79< $Z_{crítico} = -2.06$, rejeita-se o H_0

6° passo: Conclusão

A um nível de significância de 2%, há indícios muito fortes para afirmar que o valor médio de todas as contas que o auditor irá receber é inferior a 26,000.00Mt.

147. Uma rede distribuídora de combustíveis alega que seus preços são maiores que os da concorrência porque seu produto possui um melhor rendimento. Foi realizado um teste com 150 litros desse combustível e o valor médio de quilométros percorridos com um litro foi de 10,6 e desvio-padrão de 0,4. Para a comparação, utilizaram-se 100 litros de combustível de uma outra distribuídora e os seguintes valores foram obtidos: média de 10,3 km/l e desvio-padrão de 0,6 km/l. O teste foi realizado em igualdade de condições (mesmo veículo, mesmo trajecto, etc). Podemos dizer, ao nível de significância de 5%, que vale a pena pagar mais caro pelo combustível?

Resolução:

Do problema temos:

Seja 1 distribuídora e 2 concorrência, então:

$\bar{x}_1 = 10.6; \mathrm{S}_1 = 0.4;\ n_1 = 150;\ \bar{x}_2 = 10.3; \mathrm{S}_2 = 0.6;\ n_2 = 100$ e $\alpha = 0.05$

1° passo: Formular hipóteses

$$\begin{cases} H_{0}: \mu_1 \leq \mu_2 \ (A\ distância\ percorrida\ com\ 1\ litro\ é\ menor\ ou\ igual\ a\ da\ concorrência) \\ H_{1}: \mu_1 > \mu_2 \ (A\ distância\ percorrida\ com\ 1\ litro\ é\ superior\ a\ da\ concorrência) \end{cases}$$

2° passo: Fixar o limite de erro α, e identificar a variável de teste

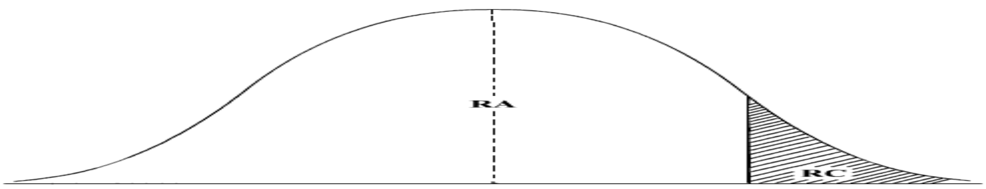
$\text{Como } n_1\ e\ n\ é\ maior\ que\ 30, então$

Para $\alpha = 0.05;\ (\bar{x}_1 - \bar{x}_2) \approx N\left(\mu_1 - \mu_2;\ \frac{\sigma^2{}_1}{n_1} + \frac{\sigma^2{}_2}{n_2}\right)$

Z: variável normal padrão

3° passo: Com o auxílio da tabela de distribuíção normal padrão, determinar a RC (região crítica) e RA (região de aceitação) para H_0

Trata-se de um teste unicaudal direito, logo $Z_\alpha = Z_{crítico} = Z_{0.05} = 1.645$



Se $Z_{calculado} < Z_{crítico}$, não-se rejeita o H_0

4° passo: Com os elementos amostrais, calcular o valor da variável do teste

$$Z_{calculado} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{S^2{}_1}{n_1} + \frac{S^2{}_2}{n_2}}} = \frac{10.6 - 10.3}{\sqrt{\frac{0.4^2}{150} + \frac{0.6^2}{100}}} = 4.39$$

5° passo: Tomada de decisão

Como $Z_{calculado} = 4.39 > Z_{crítico} = 1.645$, rejeita-se o H_0

6° passo: Conclusão

A um nível de significância de 5%, podemos dizer que vale a pena pagar mais caro pelo combustível.

148. Em períodos de pico, os clientes de um banco são obrigados a enfrentar longas filas para tirar dinheiro nas caixas eletrônicos. Dados históricos de vários anos de operação indicam que o tempo de transação nessas caixas tem distribuição normal com média igual a 270 segundos. Para aliviar essa situação o banco resolve instalar, em caráter experimental, algumas caixas eletrônicos de concepção mais avançada. Após o período de experiência, o banco pretende examinar o tempo médio obtido em uma amostra casual simples das transações realizadas nesses caixas. Em 64 transações obteve-se uma média de 262,3 segundos e desvio padrão de 21,4 segundos.

a) Que tipo de informação o banco pretende obter com esse conjunto de dados? (Formule as hipóteses);

Resolução:

O banco deseja obter informações que dê suporte á conjectura de que o tempo médio de transacção nas novas máquinas são inferiores a 270 segundos. Isso serviria como base objectiva para a decisão de substituir as máquinas antigas pelas novas.

b) Formule as hipóteses estatísticas inerentes ao problema;

$$\begin{cases} H_0\text{: } \mu = 270 \text{ (O tempo médio nas caixas é igual a 270 segundos} - \text{máquinas antigas)} \\ H_1\text{: } \mu < 270 \text{ (O tempo médio nas caixas é inferior a 270 segundos} - \text{máquinas novas)} \end{cases}$$

c) Baseando-se na resposta em a) e b) indique os possíveis erros tipo I e II que o banco possa cometer;

Resposta:

Erro tipo I: *Não havendo melhoria significativa do tempo das máquinas novas em relação as antigas, Optar por substituír as máquinas antigas pelas novas;*

Erro tipo I: *Optar por manter as máquinas antigas em dentrimento das novas, mesmo com uma redução significativa do tempo médio das transacções;*

d) A partir de qual nível de significância, essa amostra certamente proporcionará aceitação de Ho?

Resolução:

Do problema temos:

$$\mu_0 = 270;\ n = 64;\ \bar{x} = 262.3; S = 21.4$$

$$n > 30, então\ \bar{x} \approx N\text{:}\left(\mu; \frac{\sigma^2}{n}\right)$$

$$P - value = P(\bar{x} \leq 262.3) = P\left(Z \leq \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}}\right) P\left(Z \leq \frac{262.3-270}{\frac{21.4}{\sqrt{64}}}\right)$$

$$P(Z \leq -2.88) = 0.002$$

$$P - value = 0.2\%$$

Proporcionará aceitação de Ho a partir de um nível de significância superior a 0.2%.

e) Recorrendo a resposta em d), há evidências suficiente para que o banco substitua as máquinas atuais pelas mais modernas?

Solução

Sendo $P - value = 0.2\%$, podemos afirmar que há evidencias mais que suficiente para que o banco substitua as máquinas atuais pelas máquinas modernas para qualquer nível de confiança inferior a 99.8%.

149. O gerente de pessoal de uma grande seguradora deseja avaliar a eficácia de dois diferentes programas de treinamento de vendas, desenvolvidos para novos empregados. Um grupo de 32 alunos recém-formados na faculdade é aleatoriamente indicado para os dois programas, de modo que existam 16 sujeitos em cada programa. Ao final do período de treinamento, cuja duração foi um mês, é aplicado um exame padrão aos 32 sujeitos; os resultados são apresentados na tabela a seguir:

	Programa A	Programa B
$\bar{x}$	72.2	66.9
S^2	128.7	85.3

a) A varibilidade dentro dos grupos parece ser semelhante para ambos os grupos? (use nível de significância de 5%)

Seja 1 (Programa A) e 2 (Programa B)

1° passo: Formular hipóteses

$$\begin{cases} H_0: \sigma^2{}_A = \sigma^2{}_B \ (A\ variabilidade\ entre\ os\ Programas\ de\ treinamento\ de\ vendas\ A\ e\ B\ é\ o\ mesmo) \\ H_1: \ \sigma^2{}_A \neq \sigma^2{}_B \ \ (Há\ diferenças\ entre\ a\ variabilidade\ entre\ os\ Programas\ de\ treinamento\ de\ vendas\ A\ e\ B) \end{cases}$$

2° passo: Fixar o limite de erro α, e identificar a variável de teste

$$Seja\ \alpha = 0.05, então\ F \sim F_{n_1-1;n_2-1}$$

3° passo: Com o auxílio da tabela F, determinar a RC (região crítica) e RA (região de aceitação) para H_0

Trata-se de um teste bicaudal

Recorrendo a tabela F teremos:

$$F_{superior} = F_{\frac{\alpha}{2}} = F_{0.025;15;15} = 2.862 \quad e\ F_{inferior} = F_{1-\frac{\alpha}{2}} = F_{0.975;15;15} = \frac{1}{2.979} = 0.349$$

Se $F_{inferior} \leq F_{calculado} \leq F_{superior}$, não se rejeita o H_0

4° passo: Com os elementos amostrais, calcular o valor da variável do teste

$$F_{calculado} = \frac{S^2{}_1}{S^2{}_2} = \frac{128.7}{85.3} = 1.509$$

5° passo: Tomada de decisão

$$Como\ F_{inferior} = 0.349 < F_{calculado} = 1.509 < F_{superior} = 2.862$$

Não se rejeita o Ho

6° passo: Conclusão

Com base nesses resultados conclui-se que não há evidências suficientes de que as variâncias dos dois programas diferem-se significativamente, a um nível de significância de 5%.

b) Se as condições forem apropiadas, em um nível de significância de 5%, determine se há evidências de diferença nos dois programas de treinamento de vendas.

1° passo: Formular hipóteses

$$\begin{cases} H_{0}: \mu_A = \mu_B \ (O\ rendimento\ médio\ entre\ os\ programas\ A\ e\ B\ é\ o\ mesmo) \\ H_{1}: \mu_A \neq \mu_B \ (O\ rendimento\ médio\ entre\ os\ programas\ A\ e\ B\ é\ diferente) \end{cases}$$

2° passo: Fixar o limite de erro α, e identificar a variável de teste

$$Seja\ \alpha = 0.05, então\ t \sim t_v$$

Recorrendo ao resultado em a) temos ${\sigma^2}_A = {\sigma^2}_B$

Então $v = n_1 + n_2 - 2 = 30$

3° passo: Com o auxílio da tabela t de Student, determinar a RC (região crítica) e RA (região de aceitação) para H_0

$$t_{crítico} = \pm t_{\frac{\alpha}{2};n_1+n_2-2} = t_{0.025;30} = \pm 2.042$$

Se $-t_{crítico} \leq t_{calculado} \leq t_{crítico}$, *não-se rejeita o H_0*

4° passo: Com os elementos amostrais, calcular o valor da variável do teste

Recorrendo ao resultado em a) temos ${\sigma^2}_A \neq {\sigma^2}_B$

Então

$$t_{calculado} = \frac{\bar{\bar{x}}_A - \bar{\bar{x}}_B}{\sqrt{\frac{{S^2}_p}{n_1} + \frac{{S^2}_p}{n_2}}}$$

onde

$${S_p}^2 = \frac{(n_1 - 1){S_1}^2 + (n_2 - 1){S_2}^2}{n_1 + n_2 - 2} = \frac{(16 - 1) * 128.7 + (16 - 1) * 85.3}{30} = 107$$

$$t_{calculado} = \frac{72.2 - 69.9}{\sqrt{\frac{107}{16} + \frac{107}{16}}} = 1.45$$

5° passo: Tomada de decisão

Como $t_{crítico} = -2.042 < t_{calculado} = 1.45 < t_{crítico} = 2.042$; *não se rejeita o H_0*

6° passo: Conclusão

Com base nesses resultados conclui-se que não há evidências suficientes de que o rendimento médio entre os grupos diferem-se significativamente, a um nível de significância de 5%.

150. Uma máquina de enchimento automático é usada para encher garrafas com detergente liquido. Uma amostra aleatória de 20 garrafas resulta em uma variância de volume de enchimento de 0,0153. Se a variância do volume de enchimento exceder 0,01; existirá proporção inaceitável de garrafas cujo enchimento não foi completo e cujo enchimento foi em demasia. Há evidências nos dados da amostra que sugira que o fabricante tenha um problema com garrafas cheias com falta e excesso de detergente? Use $\alpha = 0{,}05$ e considere que o volume de enchimento tenha uma distribuição normal.

<u>Resolução:</u>

Do problema temos:

$$S^2 = 0.0153;\ n = 20;\ {\sigma^2}_0 = 0.01; e\ \alpha = 5\%$$

1° passo: Formular hipóteses

$$\begin{cases} H_{0:\ {\sigma^2}_0 \leq 0.01} \\ H_{1:\ {\sigma^2}_0 > 0.01} \end{cases}$$

2° passo: *Fixar o limite de erro α, e identificar a variável de teste*

$seja\ \alpha = 5\%, então\ \chi^2 \sim \chi^2$*(n-1), onde graus de liberdade é igual a 19*

3° passo: Com o auxílio da tabela Qui-Quadrado, determinar a RC (região crítica) e RA (região de aceitação) para H_0

Trata-se de um teste unicaudal esquero

Recorrendo a tabela Chi-Quadrado teremos:

$${\chi^2}_{crítico} = {\chi^2}_{sup} = {\chi^2}_{0.05;19} = 30.1$$

Não se rejeita H_0, se ${\chi^2}_{calculado} < {\chi^2}_{crítico}$

4° passo: *Com os elementos amostrais, calcular o valor da variável do teste*

$${\chi^2}_{calculado} = \frac{(n-1) * S^2}{\sigma^2} = \frac{19 * 0.0153}{0.01} = 29.07$$

5° passo: *Tomada de decisão*

Como ${\chi^2}_{calculado} = 29.07 < {\chi^2}_{crítico} = 30.1$ não se rejeita o H_0

6° passo:Conclusão

Com base nesses resultados conclui-se que não há evidências suficientes de que a variância no volume de enchimento excede 0,01, a um nível de significância de 5%.

151. Uma empresa que presta serviços de acessória a outras empresas esta interessada em comparar a taxa de reclamações sobre os seus serviços em dois dos seus escritórios em duas cidades diferentes. Suponha que a empresa tenha seleccionado aleatoriamente 100 serviços realizados

pelo escritório da cidade A e foi constatado que em 12 deles houve algum tipo de reclamação. Já do escritório da cidade B foram seleccionados 120 serviços e 18 receberam algum tipo de reclamação. A empresa deseja saber a um nível de signicância de 5%, se estes resultados são suficientes para se concluir que os dois escritórios apresentam diferenças significativas entre suas taxas de reclamações.

Resolução:

Do problema temos:

$$n_A = 100;\ P_A = \frac{12}{100}; n_B = 120;\ P_B = \frac{18}{120};\ \alpha = 0.05$$

1° passo: Formular hipóteses

$$\begin{cases} H_{0}: \pi_A = \pi_B \ (As\ taxas\ de\ reclamações\ entre\ as\ cidades\ A\ e\ B\ não\ se\ diferem) \\ H_{1}: \pi_A \neq \pi_B \ (Há\ diferenças\ entre\ as\ taxas\ de\ reclamações\ entre\ as\ cidades\ A\ e\ B) \end{cases}$$

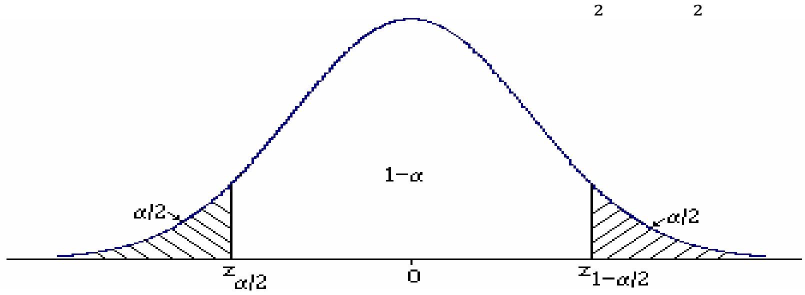
2° passo: Fixar o *limite de erro* α, *e identificar a variável de teste*

Para $\alpha = 0.05$; $(P_A - P_B) \approx N\left(\pi_A - \pi_B; \frac{\bar{p}*\bar{q}}{n_A} + \frac{\bar{p}*\bar{q}}{n_B}\right)$

Z: variável normal padrão

3° passo: Com o auxílio da tabela de distribuíção normal padrão, determinar a RC (região crítica) e RA (região de aceitação) para H_0

Trata-se de um teste bicaudal, onde $Z_{crítico} = Z_{\frac{\alpha}{2}} = -Z_{\frac{\alpha}{2}} = 1.96$



Se $-Z_{crítico} \leq Z_{calculado} \leq Z_{crítico}$, não-se rejeita o H_0.

4° passo: *Com os elementos amostrais, calcular o valor da variável do teste*

$$Z_{cal} = \frac{P_A - P_B}{\sqrt{\frac{\bar{P}*(1-\bar{P})}{n_A} + \frac{\bar{P}*(1-\bar{P})}{n_B}}}$$

Onde

$$\bar{P} = \frac{n_A * P_A + n_B * P_B}{n_A + n_B} = \frac{12+18}{220} = 0.136$$

$$Z_{cal} = \frac{P_A - P_B}{\sqrt{\frac{\bar{P}*(1-\bar{P})}{n_A} + \frac{\bar{P}*(1-\bar{P})}{n_B}}} = \frac{0.12 - 0.15}{\sqrt{0.136 * 0.864 * \left(\frac{1}{100} + \frac{1}{120}\right)}} = -0.646$$

5° passo: Tomada de decisão

$Como, Z_{crítico} = -1.96 < Z_{calculado} = -0.646 < Z_{crítico} = 1.96$, não-se rejeita o H_0

6° passo: Conclusão

A um nível de significância de 5%, há evidências suficientes para concluír que não há diferenças significativas entre as taxas de reclamação dos escritórios das cidades.

152. Uma rede distribuídora de combustíveis alega que seus preços são maiores que os da concorrência porque seu produto possui um melhor rendimento. Foi realizado um teste com 21 litros desse combustível e o valor médio de quilométros percorridos com um litro foi de 10,6 e desvio padrão de 0,4. Para a comparação, utilizaram-se 31 litros de combustível de uma outra distribuídora e os seguintes valores foram obtidos: média de 10,3 km/l e desvio-padrão de 0,6 km/l. O teste foi realizado em igualdade de condições (mesmo veículo, mesmo trajecto, etc). Podemos dizer, ao nível de significância de 5%, que vale a pena pagar mais caro pelo combustível?

Resolução:

Do problema temos:

Seja 1 distribuídora e 2 concorrência, então:

$\bar{x}_1 = 10.6; S_1 = 0.4;\ n_1 = 21;\ \bar{x}_2 = 10.3; S_2 = 0.6;\ n_2 = 31$ e $\alpha = 0.05$

1° passo: Formular hipóteses

$$\begin{cases} H_{0:}\ \mu_1 \leq \mu_2\ (A\ distância\ percorrida\ com\ 1\ litro\ é\ menor\ ou\ igual\ a\ da\ concorrência) \\ H_{1:}\ \mu_1 > \mu_2\ (A\ distância\ percorrida\ com\ 1\ litro\ é\ superior\ a\ da\ concorrência) \end{cases}$$

2° passo: Fixar o limite de erro α, e identificar a variável de teste

$$Como\ n_1\ é\ menor\ que\ 30, então$$

$$então\ t \sim t_v$$

Cálculo de graus de liberdade: 1° Verificar o pressuposto de homogeneidade

$$\begin{cases} H_{0:}\ \sigma^2{}_A = \sigma^2{}_B \\ H_{1:}\ \sigma^2{}_A \neq \sigma^2{}_B \end{cases}$$

$$Seja\ \alpha = 0.05, então\ F \sim F_{n_1-1;n_2-1}$$

Com o auxílio da tabela F, determinar a RC (região crítica) e RA (região de aceitação) para H_0

Trata-se de um teste bicaudal

Recorrendo a tabela F teremos:

$$F_{superior} = F_{\frac{\alpha}{2}} = F_{0.025;20;30} = 2.195\ \ e\ \ F_{inferior} = F_{1-\frac{\alpha}{2}} = F_{0.975;20;30} = \frac{1}{2.195} = 0.456$$

Se $F_{inferior} \leq F_{calculado} \leq F_{superior}$, não se rejeita o H_0

Com os elementos amostrais, calcular o valor da variável do teste

$$F_{calculado} = \frac{S^2{}_1}{S^2{}_2} = \frac{0.4}{0.6} = 0.667$$

Tomada de decisão

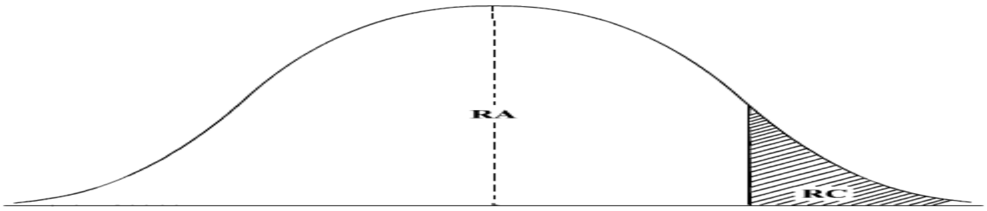
$$Como\ F_{inferior} = 0.456 < F_{calculado} = 0.667 < F_{superior}=2.195$$

Não se rejeita o Ho e conclui-se que as varianças são homogéneas

$$Logo\ Gl = v = n_1 + n_2 - 2 = 21 + 31 - 2 = 50$$

3° passo: Com o auxílio da tabela t de Student, determinar a RC (região crítica) e RA (região de aceitação) para H0

Trata-se de um teste unicaudal direito, logo $t_\alpha = t_{crítico} = t_{0.05;50} = 1.676$



Se $t_{calculado} < t_{crítico}$, não-se rejeita o H_0

$4^{\underline{o}}\ passo$: Com os elementos amostrais, calcular o valor da variável do teste

$$t_{calculado} = \frac{(\bar{\bar{x}}_1 - \bar{\bar{x}}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S^2{}_p}{n_1} + \frac{S^2{}_p}{n_2}}}$$

onde

$$S_p{}^2 = \frac{(n_1 - 1)S_1{}^2 + (n_2 - 1)S_2{}^2}{n_1 + n_2 - 2} = \frac{(21-1)*0.16 + (31-1)*0.36}{50} = 0.28$$

$$t_{calculado} = \frac{10.6 - 10.3}{\sqrt{\frac{0.28}{21} + \frac{0.28}{31}}} = 2$$

5° passo: Tomada de decisão

Como $t_{calculado} = 2 > t_{crítico} = 1.676$, rejeita-se o H_0

6° passo: Conclusão

A um nível de significância de 5%, podemos dizer que vale a pena pagar mais caro pelo combustível.

153. Um auditor deseja testar a hipótese de que o valor médio de todas as contas a receber de uma dada firma é, no mínimo 26,000.00 Mt. Tomando para tanto uma amostra de n = 29 e supondo que a média amostral é 24,000.00 Mt. Testar a hipótese ao nível de confiança de 98%, dado que se conhece o desvio padrão dos valores das contas a receber, isto é, 4,300.00 Mt.

Resolução:

Do problema temos:

$$\mu_0 = 26000; \sigma_0 = 4300;\ n = 29;\ \bar{x} = 24000\ e\ \alpha = 0.02$$

1° passo: Formular hipóteses

$$\begin{cases} H_0: \mu_0 \geq 26000 \ (o\ valor\ médio\ de\ todas\ as\ contas\ que\ o\ auditor\ irá\ receber\ é\ no\ mínimo\ 26{,}000.00Mt) \\ H_1: \mu_0 < 26000 \ (o\ valor\ médio\ de\ todas\ as\ contas\ que\ o\ auditor\ irá\ receber\ é\ inferior\ a\ 26{,}000.00Mt) \end{cases}$$

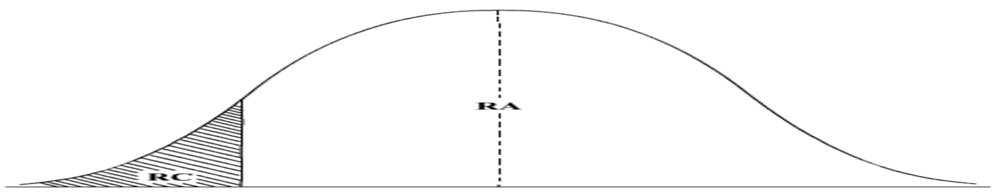
2° passo: Fixar o limite de erro α, e identificar a variável de teste

Para $\alpha = 0.02;\ \bar{x} \sim N\left(\mu = 26000; \frac{\sigma^2 = 4300^2}{n = 29}\right)$

Z: variável normal padrão

3° passo: Com o auxílio da tabela de distribuíção normal padrão, determinar a RC (região crítica) e RA (região de aceitação) para H0

Trata-se de um teste unicaudal esquerdo, logo $-Z_\alpha = -Z_{crítico} = -Z_{0.02} = -2.06$



Se $Z_{calculado} > -Z_{crítico}$, não-se rejeita o H_0

4° passo: Com os elementos amostrais, calcular o valor da variável do teste

$$Z_{calculado} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{24000 - 26000}{\frac{4300}{\sqrt{29}}} = -2.505$$

5° passo: Tomada de decisão

Como $Z_{calculado} = -2.505 < Z_{crítico} = -2.06$, rejeita-se o H_0

6° passo: Conclusão

A um nível de significância de 2%, há indícios muito fortes para afirmar que o valor médio de todas as contas que o auditor irá receber é inferior a 26,000.00Mt.

154. Uma rede de postos de gasolina afirma que, em seus estabelecimentos não se vende gasolina adulterada. Sabe-se que, de acordo com os padrões de qualidade, a gasolina não pode conter mais de 240 ml de álcool por litro. O órgão de fiscalização colheu 25 medições do produto nos postos dessa rede, obtendo a partir delas uma média de 240,75 ml de álcool/litro. Admitindo-se que a quantidade de álcool presente na gasolina tem uma distribuição normal com desvio-padrão de 2,5 ml/litro. Ao nível de significância 5%, pode-se afirmar que a gasolina é adulterada?

Resolução:

Do problema temos:

$$\mu_0 = 240;\ n = 25;\ \bar{x} = 240.75;\ \sigma = 2.5\ e\ \alpha = 0.05$$

1° passo: Formular hipóteses

$$\begin{cases} H_0: \mu \leq 240 \ (A\ gasolina\ não\ está\ adulterada) \\ H_1: \mu > 240 \ (A\ gasolina\ \ está\ adulterada) \end{cases}$$

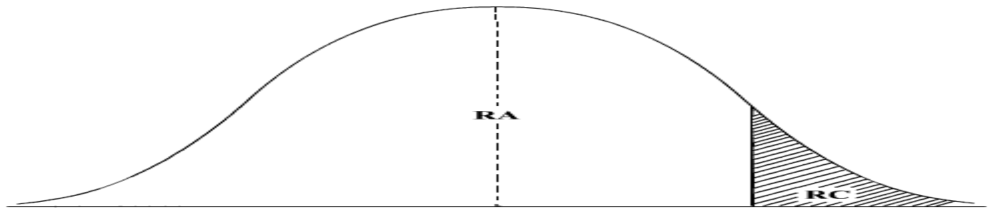
2° passo: Fixar o limite de erro α, e identificar a variável de teste

Para $\alpha = 0.05$*;* $\bar{x} \sim N\left(\mu; \frac{\sigma^2}{n}\right)$

Z: variável normal padrão

3° passo: Com o auxílio da tabela de distribuíção normal padrão, determinar a RC (região crítica) e RA (região de aceitação) para H_0

Trata-se de um teste unicaudal direito, logo $Z_\alpha = Z_{crítico} = Z_{0.05} = 1.645$



Se $Z_{calculado} < Z_{crítico}$*, não se rejeita o* H_0

4° passo: Com os elementos amostrais, calcular o valor da variável do teste

$$Z_{calculado} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{240.75 - 240}{\frac{2.5}{\sqrt{25}}} = 1.5$$

5° passo: Tomada de decisão

Como $Z_{calculado} = 1.5 < Z_{crítico} = 1.645$ *, Não se rejeita o* H_0

6° passo: Conclusão

A um nível de significância de 5%, há indícios muito fortes para afirmar que o o combústivel não está adulterada.

155. Um docente universitário está indeciso entre dois modelos de exame. A sua escolha recairá naquele que demorar menos tempo a concluir. Para efectuar a sua escolha, o docente aplicou o teste A, a uma amostra de 8 alunos e o teste B, a uma amostra de10 alunos, tendo obtido os seguintes tempos, em minutos, para a conclusão do exame:

- Teste A: 120; 90; 110; 100; 80; 85; 95; 80;
- Teste B: 110; 95; 100; 85; 90; 95; 110; 80; 90; 110.

Admitindo que os tempos de resolução se comportam segundo uma lei normal, teste a um nível de significância de 5%, se existe diferença na variabilidade do tempo de resolução dos dois tipos de exame.

<u>Resolução</u>

1. $\begin{cases} H_0: \ {\sigma^2}_1 = {\sigma^2}_2 \\ H_1: {\sigma^2}_1 \neq {\sigma^2}_2 \end{cases}$

2. $F \sim F_{n_1-1;n_2-1}$

$$F_{calculado} = \frac{S^2{}_1}{S^2{}_2} = \frac{207.14}{116.94} = 1.77$$

3. $\alpha = 0.05 \rightarrow \frac{\alpha}{2} = 0.025\ teste\ bicaudal.$

4. $F_{\frac{\alpha}{2};n_1-1;n_2-1} = F_{0.025;7;9} = 4.197$

$$F_{1-\alpha/2;n_1-1;n_2-1} = F_{0.9755;7;9} = \frac{1}{F_{0.025;9;7}} = \frac{1}{4.823} = 0.207$$

5. Como $F_{0.9755;7;9} = 0.207 < F_{calculado} = 1.77 < F_{0.025;7;9} = 4.197$ *, não se rejeita-se a hipótese nula* H_0.

6. *Conclusão: Não existe diferença na variabilidade do tempo de resolução dos dois tipos de exame, ao nível de significância de* 5%.

156. Uma empresa pretende lançar um novo produto numa cidade de um milhão de habitantes. No estudo de mercado realizado foram inquiridas 1000 pessoas, tendo 800 delas afirmado que muito dificilmente iriam utilizar aquele novo produto. Teste, ao nível de significância de 1%, se a verdadeira proporção de habitantes que utilizarão o novo produto pode ser considerada no máximo igual a 0, 18.

Resolução

1. $\begin{cases} H_0: \pi \leq 0.18 \\ H_1: \mu > 0.18 \end{cases}$

2. $\alpha = 0.05$

3. $Z \sim N(0{,}1)$

$$Z_0 = \frac{p - \pi_0}{\sqrt{\frac{\pi_0 * (1 - \pi_0)}{n}}} = \frac{0.2 - 0.18}{\sqrt{\frac{0.18 * 0.82}{1000}}} = 1.6462$$

4. $\alpha = 0.01$ *pois o teste é unilateral direito.*

Para uma área de 0.49, $Z_{crítico} = Z_\alpha = Z_{0.01} = 2.33$

5. Como $Z_{calculado} = 1.6462 < Z_{crítico} = 2.33$, *aceita–se a hipótese nula* H_0.

6. *Conclusão: conclui-se que a verdadeira proporção de habitantes que utilizarão o novo produto pode ser considerada no máximo igual a 0,18, ao nível de significância de 1%.*

157. Um candidato em campanha eleitoral deseja saber se a sua intenção de votos é a mesma em duas cidades importantes no cenário político (Atlântida e Flórida), e definir as estratégias de campanha. Para isso ele contratou um instituto de pesquisa que realizou pesquisas nas duas cidades. Em Atlântida foram entrevistados 500 eleitores, dos quais 116 afirmaram votar no candidato,

enquanto que, em Flórida, dos 600 entrevistados, 105 foram favoráveis a ele. Qual a conclusão que se pode tirar com essas informações? ($assumir\ \alpha = 0.05$).

Resolução

1. $\begin{cases} H_0:\ \pi_1 = \pi_2 \\ H_{1:}\ \pi_1 \neq \pi_2 \end{cases}$

2. $\alpha = 0.05$

3. $Z \sim N(0,1)$

$$Z = \frac{p_1 - p_2}{\sqrt{p * (1-p) * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.232 - 0.175}{\sqrt{0.201 * 0.799 * \left(\frac{1}{500} + \frac{1}{600}\right)}} = 2.35$$

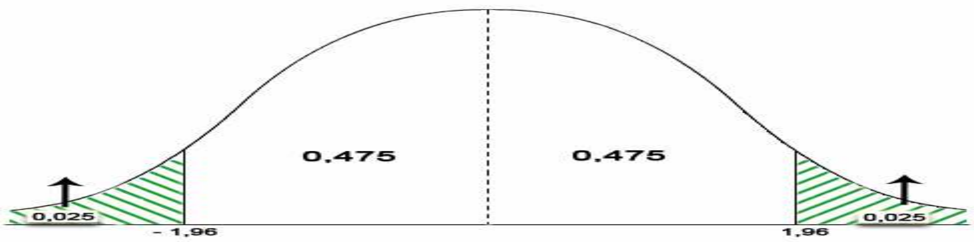
Sendo

$$p = \frac{n_1 * p_1 + n_2 * p_2}{n_1 + n_2} = \frac{500 * 0.232 + 600 * 0.175}{500 + 600} = 0.201$$

4. $\alpha = 0.05\ \rightarrow \frac{\alpha}{2} = 0.025\ pois\ o\ teste\ é\ bilateral.$ *Para uma área de 0.475,*

$$Z_{crítico} = Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$$

Desenhando a curva, temos:



5. Como $Z_{calculado} = 2.35 > Z_{crítico} = 1.96$, *Rejeita-se a hipótese nula* H_0.
6. *Conclusão: A um nível de significância de 5%, pode-se concluir que a intenção de votos nas duas cidades não é a mesma (a intenção de votos na Flórida é menor).*

158. Num estudo de mercado sobre a audiência de 2 jornais semanais e de 1 revista semanal foram inquiridos 1000 leitores de ambos os sexos sobre o semanário que compram preferencialmente, tendo-se encontrado os seguintes resultados:

Sexo Semanário	Savana	Domingo	Zambézia
Feminino	150	50	150
Masculino	350	200	100

Será de admitir que a preferência pelos vários semanários é influenciada pelo sexo dos leitores? (Admita um nível de significância de 5%).

Resolução

Pretende-se saber se a preferência pelos vários semanários é independente do sexo dos leitores. Para isso vai aplicar-se o teste do qui-quadrado da independência entre duas variáveis. Então:

- H_0*: A preferência pelos vários semanários não depende do sexo do eleitor*
- H_1*: A preferência pelos vários semanários depende do sexo do eleitor*

Cálculos auxiliares: *Os* $\boldsymbol{o_{ij}}$ *correspondem ao número de indivíduos observados das células* ij*. O cálculo do número esperado de indivíduos esperados* $\boldsymbol{e_{ij}}$*, pressupõe que a hipótese nula,* H_0 *é verdadeira, isto é, as variáveis são independentes:*

$$e_{ij} = n * p_{ij} = n * p_{i.} * p_{.j} = n * \frac{o_{i.}}{n} * \frac{o_{.j}}{n} = \frac{o_{i.}}{n} * o_{.j} * n$$

Neste caso ter-se-á:

- $o_{1.} = 150 + 50 + 150 = 350$
- $o_{2.} = 350 + 200 + 100 = 650$
- $o_{.1} = 150 + 350 = 500$
- $o_{.2} = 50 + 200 = 250$
- $o_{.3} = 150 + 100 = 250$
- $e_{11} = \frac{350*500}{1000} = 175$
- $e_{12} = \frac{350*250}{1000} = 87.5$
- $e_{13} = \frac{350*250}{1000} = 87.5$
- $e_{21} = \frac{650*500}{1000} = 325$
- $e_{22} = \frac{6500*250}{1000} = 162.5$
- $e_{23} = \frac{6500*250}{1000} = 162.5$

Na tabela seguinte estão incluídos para além dos o_{ij} *os* e_{ij} *obtidos:*

Sexo Semanário	Savana	Domingo	Zambézia	Total
Feminino	150 (175)	50 (87.5)	150 (87.5)	350
Masculino	350 (325)	200 (162.5)	100 (162.5)	650
Total	500	250	250	1000

$$\chi^2{}_0 = \sum_{i=1}^{l}\sum_{j=1}^{c} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

$$= \frac{(150-175)^2}{175} + \frac{(50-87.5)^2}{87.5} + \frac{(150-87.5)^2}{87.5} + \frac{(350-325)^2}{325} + \frac{(200-162.5)^2}{62.5} + \frac{(100-162.5)^2}{162.5} = 98.89$$

Valor crítico:

$$\chi^2{}_{(l-1)*(c-1);1-\alpha} = \chi^2{}_{(2-1)*(3-1);1-0.05} = \chi^2{}_{2;0.95} = 5.991$$

Decisão: *Como o valor da estatística de teste* ${\chi^2}_0 = 98.89 > {\chi^2}_{2;0.95} = 5.991$*, rejeita-se a hipótese nula, isto é, devemos concluir que a preferência pelos semanários depende sexo do eleitor.*

159. A produção do produto M em uma máquina no laboratório Y tem uma média de 72/h e um desvio-padrão de 2/h com distribuição normal. Recentemente a máquina foi ajustada. A fim de determinar o efeito do ajuste, 25 amostras foram testadas. Os testes apresentaram uma produção média de 75/h. Considere que o desvio-padrão não mudou. Com base nesses dados, com um nível de significância de 1% é possível afirmar que o valor médio não mudou?

Resolução

Pelos dados temos:

$$\mu_0 = 72;\ \sigma_0 = 2; n = 25;\ \bar{x} = 75$$

1° Passo: Formulação de hipóteses:

$$\begin{cases} H_0: \mu = 72 \\ H_1: \mu \neq 72 \end{cases}$$

2° Passo: Identificação da distribuição amostral

Como a variança populacional é conhecida, então:

$$Z \sim N(0; 1)$$

3° Passo: Determinação dos valores críticos e região de aceitação ou rejeição da Hipótese nula

Sabe se que o nível de significância é de 1% e o teste é bi-caudal, pela tabela de distribuição normal temos:

$$Z_{\frac{\alpha}{2}} = Z_{0.005} = \pm 2.58$$

$Se\ -Z_{\frac{\alpha}{2}} \leq Z_{calculado} \leq Z_{\frac{\alpha}{2}}$ *não se rejeita a hipótese nula*

4° Passo: Determinação da estatística do teste

$$Z_{calculado} = \frac{\bar{x} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} = \frac{75 - 72}{2/\sqrt{25}} = 7.5$$

5° Passo: Decisão

Como $Z_{calculado} = 7.5 > Z_{\frac{\alpha}{2}} = 2.58$, *rejeita a hipótese nula*

6ª passo: Conclusões

A um nível de significância de 1%, pode se afirmar que o valor médio mudou.

160. Pretende-se introduzir um novo processo na produção de um certo tipo de esferas para uso industrial. O novo processo mantém a pressão média mas espera-se que consiga reduzir a sua variabilidade, que até agora era de 14,5. Como a introdução completa do novo processo acarreta custos, resolveu-se proceder a um teste, para o qual foram obtidas 16 esferas produzidas de

acordo com o novo método. Ao nível de significância de 0,05 e sabendo que a variância da amostra é 6.8, qual a decisão a tomar?

Resolução

Pelos dados temos:

$$\sigma_0{}^2 = 14.5; n = 16;\ S^2 = 6.8$$

1° Passo: Formulação de hipóteses:

$$\begin{cases} \mathrm{H}_0: \sigma^2 = 14.5 \\ \mathrm{H}_1: \sigma^2 < 14.5 \end{cases}$$

2° Passo: Identificação da distribuição amostral

$$\chi^2 \sim \chi^2{}_{n-1}$$

3° Passo: Determinação dos valores críticos e região de aceitação ou rejeição da Hipótese nula

Sabe se que o nível de significância é de 5% e o teste é uni-caudal esquerdo, pela tabela de distribuição Qui-Quadrado temos:

$$\chi^2{}_{1-\alpha;n-1} = \chi^2{}_{1-0.05;16-1} = \chi^2{}_{0.95;15} = 7.261$$

$Se\ \chi^2{}_{calculado} > \chi^2{}_{1-\alpha;n-1}$ *não se rejeita a hipótese nula*

4° Passo: Determinação da estatística do teste

$$\chi^2{}_{calculado} = \frac{(n-1) * S^2}{\sigma^2{}_0} = \frac{15 * 6.8}{14.5} = 7.034$$

5° Passo: Decisão

Como $\chi^2{}_{calculado} = 7.034 < \chi^2{}_{0.95;15} = 7.261$, *rejeita a hipótese nula*

6ª Passo: Conclusões

A um nível de significância de 5%, pode se afirmar que a variabilidade diminuiu.

161. O consumidor de um certo produto acusou o fabricante, dizendo que mais de 20% das unidades fabricadas apresentam defeito. Para confirmar sua acusação, ele usou uma amostra de tamanho 50, onde 27% das peças eram defeituosas. Mostre como o fabricante poderia refutar a acusação. Utilize um nível de significância de 2%.

Resolução

Pelos dados temos:

$$\pi_0 = 0.20; n = 50;\ p = 0.27$$

1° Passo: Formulação de hipóteses:

$$\begin{cases} H_0: \pi \leq 0.20 \\ H_1: \pi > 0.20 \end{cases}$$

2° Passo: Identificação da distribuição amostral

$$Z \sim N(0; 1)$$

3° Passo: Determinação dos valores críticos e região de aceitação ou rejeição da Hipótese nula

Sabe-se que o nível de significância é de 2% e o teste é uni-caudal direito, pela tabela de distribuição Normal temos:

$$Z_\alpha = Z_{0.02} = 2.05$$

$Se\ Z_{calculado} < Z_{0.02}$ *não se rejeita a hipótese nula*

4° Passo: Determinação da estatística do teste

$$Z_{calculado} = \frac{p - \pi_0}{\sqrt{\frac{\pi_0 * (1 - \pi_0)}{n}}} = \frac{0.27 - 0.20}{\sqrt{\frac{0.20 * (1 - 0.20)}{50}}} = 1.237$$

5° Passo: Decisão

Como $Z_{calculado} = 1.237 < Z_{0.02} = 2.05$, *não se rejeita a hipótese nula.*

6ª Passo: Conclusões

A um nível de significância de 2%, pode se afirmar que acusação do consumidor não é verdadeira.

162. Um fabricante de insecticida descobriu, em uma pesquisa, que muitos consumidores da categoria achavam que o cheiro do seu produto era suave demais, o enfraquecia o seu apelo publicitário de produto fortíssimo contra os insectos. Após alguns sniff tests, foi seleccionado um novo aroma para o produto. Faltava testar o novo aroma no produto em uso pela consumidora. O fabricante solicitou à empresa de pesquisa que fizesse um teste de produto in home em que uma amostra de 15 consumidores usaria durante uma semana o produto com o cheiro actual, enquanto uma outra amostra, também de 12 consumidores com o mesmo perfil, usaria o produto com o novo cheiro. Definiu como padrão de acção: mudar o aroma de sua marca se, e somente se, o novo produto for superior ao actual. Suponha que o fabricante nada sabe sobre as variânças e tenha escolhido o nível de significância 0,05. A avaliação em escala de 7 pontos resultou nos seguintes parâmetros.

	Aroma actual	Aroma novo
Média	5.94	6.27
Desvio Padrão	5.22	2.39

Qual é a decisão do fabricante?

<u>Resolução</u>

Nota que o fabricante nada sabe sobre as variânças, então primeiros temos que testar a igualdade das varianças, então:

1° Passo: Formulação de hipóteses:

$$\begin{cases} H_0: \sigma^2{}_1 = \sigma^2{}_2 \\ H_1: \sigma^2{}_1 \neq \sigma^2{}_2 \end{cases}$$

2° Passo: Identificação da distribuição amostral

$$F \sim F_{n_1-1;n_1-1}$$

3° Passo: Determinação dos valores críticos e região de aceitação ou rejeição da Hipótese nula

Sabe-se que o nível de significância é de 5% e o teste é bi-caudal direito, pela tabela de distribuição F temos:

$$F_{\frac{\alpha}{2};\,n_1-1;n_2-1} = F_{0.025;\,14;11} = 3.36$$

$$\mathrm{F}_{1-\frac{\alpha}{2};\,n_1-1;n_2-1} = \frac{1}{\mathrm{F}_{\frac{\alpha}{2};\,n_2-1;n_1-1;}} = \frac{1}{F_{0.025;\,11;14}} = \frac{1}{3.05} = 0.02295$$

$Se\ \mathrm{F}_{1-\frac{\alpha}{2};\,n_1-1;n_2-1} \leq F_{calculado} \leq F_{\frac{\alpha}{2};\,n_1-1;n_2-1}$ *não se rejeita a hipótese nula*

4° Passo: Determinação da estatística do teste

$$F_{calculado} = \frac{S^2{}_1}{S^2{}_2} = \frac{5.22^2}{2.39^2} = 4.7703$$

5° Passo: Decisão

Como $F_{calculado} = 4.7703 > F_{0.025;\,14;11} = 3.36$*, rejeita-se a hipótese nula*

6ª Passo: Conclusões

A um nível de significância de 5%, pode se afirmar que as varianças são desconhecidas e diferentes.

Em seguida o teste de igualdade das média:

1° Passo: Formulação de hipóteses:

$$\begin{cases} \mathrm{H}_0: \mu_1 = \mu_2 \\ \mathrm{H}_1: \mu_1 < \mu_2 \end{cases}$$

2° Passo: Identificação da distribuição amostral

$$t \sim t_{Gl}$$

3° Passo: Determinação dos valores críticos e região de aceitação ou rejeição da Hipótese nula

Sabe-se que o nível de significância é de 5% e o teste é uni-caudal esquerdo com varianças desconhecidas e diferentes, então:

$$Gl = \frac{\left(\frac{S^2{}_1}{n_1} + \frac{S^2{}_2}{n_2}\right)^2}{\frac{\left(\frac{S^2{}_1}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{S^2{}_2}{n_2}\right)^2}{n_2-1}} = \frac{\left(\frac{5.22^2}{15} + \frac{2.39^2}{12}\right)^2}{\frac{\left(\frac{5.22^2}{15}\right)^2}{14} + \frac{\left(\frac{2.39^2}{12}\right)^2}{11}} = 21$$

$$-t_{\alpha;\,Gl} = -t_{0.05;21} = -1.721$$

$Se - t_{0.05;21} \leq t_{calculado}$ *não se rejeita a hipótese nula*

4° Passo: Determinação da estatística do teste

$$t_{calculado} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S^2{}_1}{n_1} + \frac{S^2{}_2}{n_2}}} = \frac{(5.94 - 6.27)}{\sqrt{\frac{5.22^2}{15} + \frac{2.39^2}{12}}} = -0.2179$$

5° Passo: Decisão

Como $t_{calculado} = -0.2179 > -t_{0.05;21} = -1.721$, *não se rejeita a hipótese nula*

6ª Passo: Conclusões

A um nível de significância de 5%, não há evidências suficientes para o fabricante mudar do aroma.

163. Para avaliar o nível de satisfação quanto a política de nomeação de gestores de uma empresa nos últimos 10 anos, entrevistou-se 500 funcionários, tendo-se encontrado os seguintes resultados:

- dos 300 funcionários do sexo feminino, 130 estão satisfeitos; e
- 200 funcionários do sexo masculino, 90 estão insatisfeitos.

Há evidências suficientes a um nível de significância de 5%, para afirmar que o nível de satisfação quanto a política de nomeação dos gestores é independente do sexo dos funcionários?

Resolução

$$\begin{cases} H_0: \text{ O nível de satisfação é independente do sexo dos funcionários} \\ H_1: \text{O nível de satisfação depende do sexo dos funcionários} \end{cases}$$

Assim, organizando os dados em tabelas temos tabela temos:

	Satisfeito	Insatisfeito	Total
Feminino	130	170	300
Masculino	110	90	200
Total	240	260	500

Nota que na tabela acima estão representados os valores observados, então:

Coluna	Linha	o_{ij}	e_{ij}	$\frac{(o_{ij} - e_{ij})^2}{e_{ij}}$
1	1	130	$\frac{300 * 240}{500} = 144$	1.361
1	2	110	$\frac{200 * 240}{500} = 96$	2.042
2	1	170	$\frac{300 * 260}{500} = 156$	1.256
2	2	90	$\frac{200 * 260}{500} = 104$	1.885
Total				6.544

Logo:

$$\chi^2{}_{calculado} = \sum_{i=1}^{l}\sum_{j=1}^{c}\frac{\left(O_{ij}-E_{ij}\right)^2}{E_{ij}} = 6.544$$

Valor crítico:

$$\chi^2{}_{(l-1)*(c-1);\alpha} = \chi^2{}_{(2-1)*(2-1);0.05} = \chi^2{}_{1;0.95} = 3.841$$

Decisão: Como o valor da estatística de teste $\chi^2{}_{calculado} = 6.544 > \chi^2{}_{1;0.05} = 3.841$, rejeita-se a hipótese nula, isto é, devemos concluir que o nível de satisfação quanto a política de nomeação dos gestores depende do sexo dos funcionários.

164. Pretende-se investigar o nível de remuneração salarial dos homens e mulheres de certa categoria profissional. De duas amostras obtidas entre dois grupos, destacam-se os seguintes resultados (em unidades monetárias):

Amostra de 15 homens	$\bar{x} = 33.8$	$S^2 = 5.7$
Amostra de 10 mulheres	$\bar{x} = 31$	$S^2 = 10.3$

Usando um nível de significância de 1%, conclua sobre a possível existência de discriminação sexual na atribuição de remunerações. Nota que nada sabe-se sobre as varianças.

Resolução

Nota que nada sabe-se sobre as variânças, então primeiros temos que testar a igualdade das varianças, então:

1° Passo: Formulação de hipóteses:

$$\begin{cases} H_0: \sigma^2{}_1 = \sigma^2{}_2 \\ H_1: \sigma^2{}_1 \neq \sigma^2{}_2 \end{cases}$$

2° Passo: Identificação da distribuição amostral

$$F \sim F_{n_1-1;n_1-1}$$

3° Passo: Determinação dos valores críticos e região de aceitação ou rejeição da Hipótese nula

Sabe-se que o nível de significância é de 5% e o teste é *bi-caudal direito, pela tabela de distribuição F temos:*

$$F_{\frac{\alpha}{2};\, n_1-1;n_2-1} = F_{0.005;\,14;9} = 6.09$$

$$F_{1-\frac{\alpha}{2};\, n_1-1;n_2-1} = \frac{1}{F_{\frac{\alpha}{2};\, n_2-1;n_1-1;}} = \frac{1}{F_{0.025;\,9;14}} = \frac{1}{4.72} = 0.21186$$

$Se\ F_{1-\frac{\alpha}{2};\, n_1-1;n_2-1} \leq F_{calculado} \leq F_{\frac{\alpha}{2};\, n_1-1;n_2-1}$ *não se rejeita a hipótese nula*

4° Passo: Determinação da estatística do teste

$$F_{calculado} = \frac{S^2{}_1}{S^2{}_2} = \frac{5.7}{10.3} = 0.5534$$

5° Passo: Decisão

Como $F_{0.975;\,14;9} = 0.21186 \, < F_{calculado} = 0.5534 < F_{0.025;\,14;9} = 6.09$ *, rejeita-se a hipótese nula.*

6ª Passo: Conclusões

A um nível de significância de 1%, pode se afirmar que as varianças são desconhecidas e iguais.

Em seguida o teste de igualdade das média:

1° Passo: Formulação de hipóteses:

$$\begin{cases} \mathrm{H_0}: \mu_1 = \mu_2 \\ \mathrm{H_1}: \mu_1 \neq \mu_2 \end{cases}$$

2° Passo: Identificação da distribuição amostral

$$t \sim t_{n_1+n_2-2}$$

3° Passo: Determinação dos valores críticos e região de aceitação ou rejeição da Hipótese nula

Sabe-se que o nível de significância é de 1% e o teste é bi-caudal com varianças desconhecidas e iguais, então:

$$t_{\frac{\alpha}{2};\, n_1+n_2-2} = t_{0.005;23} = \pm 2.807$$

$$Se - t_{0.005,22} \leq t_{calculado} \leq t_{0.005,22} \text{ não se rejeita a hipótese nula}$$

4° Passo: Determinação da estatística do teste

$$t_{calculado} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{(n_1-1)*{S_1}^2 + (n_2-1)*{S_2}^2}{n_1+n_2-2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{33.8 - 31.0}{\sqrt{7.5 * \left(\frac{1}{15}+\frac{1}{10}\right)}} = 2.504$$

5° Passo: Decisão

Como $t_{calculado} = 2.504 < t_{0.005,22} = 2.807$, *não se rejeita a hipótese nula*

6ª Passo: Conclusões

A um nível de significância de 1%, pode-se concluir que não há evidencias suficientes para afirmar que existe uma discriminação sexual na atribuição de remunerações.

165. Uma votação será feita entre os residentes de uma cidade e a região rural ao redor desta cidade para determinar se um projecto químico deverá ser construído. A construção é dentro dos limites da cidade e por esta razão muitos eleitores do campo sentem que o projecto passará por causa da grande proporção dos eleitores da cidade, os quais são favoráveis. Para determinar se existe diferença significante na proporção de eleitores da cidade e do campo a favor do projecto, uma amostragem foi feita. Se 110 de 200 eleitores da cidade são a favor do projecto e metade de 500 eleitores do campo são a favor, você concordaria que a proporção de eleitores da cidade favoráveis ao projecto é maior do que a proporção de eleitores do campo favoráveis ao projecto? Use $\alpha = 0{,}025$.

Resolução

Pelos dados temos:

$$n_1 = 200; x_1 = 110;\ p_1 = 0.55;\ n_2 = 500; x_2 = 250;\ p_2 = 0.50$$

1° Passo: Formulação de hipóteses:

$$\begin{cases} H_0: \pi_1 = \pi_2 \\ H_1: \pi_1 > \pi_2 \end{cases}$$

2° Passo: Identificação da distribuição amostral

$$Z \sim N(0;1)$$

3° Passo: Determinação dos valores críticos e região de aceitação ou rejeição da Hipótese nula

Sabe-se que o nível de significância é de 2.5% e o teste é uni-caudal direito, pela tabela de distribuição Normal Padrão temos:

$$Z_\alpha = Z_{0.025} = 1.96$$

Se $Z_{calculado} \leq Z_\alpha$ não se rejeita a hipótese nula

4° Passo: Determinação da estatística do teste

$$Z_{calculado} = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\frac{p_1 * (1 - p_1)}{n_1} + \frac{p_2 * (1 - p_2)}{n_2}}} = \frac{(0.55 - 0.50)}{\sqrt{\frac{0.55 * 0.45}{200} + \frac{0.50 * 0.50}{500}}} = 1.1995$$

5° Passo: Decisão

Como $Z_{calculado} = 1.1995 < Z_{0.025} = 1.96$, *não se rejeita a hipótese nula.*

6ª Passo: Conclusões

A um nível de significância de 2%, pode se concluir que a diferença não é significativa.

166. Num estudo de mercado sobre a audiência de 3 jornais semanais foram inquiridos 1500 leitores de ambos os sexos sobre o semanário que compram preferencialmente, tendo-se encontrado os seguintes resultados:

- Dos 600 leitores do semanário Savana, 200 são do sexo feminino;
- 400 leitores do semanário Domingo, 300 são do sexo masculino;
- 500 leitores do semanário Canal de Moçambique, 300 são feminino

Será de admitir a um nível de significância de 5%, que a preferência pelos vários semanários é influenciada pelo sexo dos leitores?

Resolução

$$\begin{cases} H_0: \text{ O nível de satisfação é independente do sexo dos funcionários} \\ H_1: \text{O nível de satisfação depende do sexo dos funcionários} \end{cases}$$

Assim, organizando os dados em tabelas temos tabela temos:

	Savana	*Domingo*	*Canal MZ*	*Total*
Feminino	*200*	*100*	*300*	*600*
Masculino	*400*	*300*	*200*	*900*
Total	*600*	*400*	*500*	*1500*

Nota que na tabela acima estão representados os valores observados, então:

Coluna	*Linha*	o_{ij}	e_{ij}	$\frac{(o_{ij}-e_{ij})^2}{e_{ij}}$
1	*1*	*200*	$\frac{600*600}{1500}=240$	*6.667*
1	*2*	*400*	$\frac{900*600}{1500}=360$	*4.444*
2	*1*	*100*	$\frac{600*400}{1500}=160$	*22.5*
2	*2*	*300*	$\frac{900*400}{1500}=240$	*15.0*
3	*1*	*300*	$\frac{600*500}{1500}=200$	*50.0*
3	*2*	*200*	$\frac{900*500}{1500}=300$	*33.333*
Total				*131.944*

Logo:

$$\chi^2{}_{calculado}=\sum_{i=1}^{l}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}=131.944$$

Valor crítico:

$$\chi^2{}_{(l-1)*(c-1);\alpha}=\chi^2{}_{(2-1)*(3-1);0.05}=\chi^2{}_{2;0.95}=5.991$$

Decisão: *Como o valor da estatística de teste* $\chi^2{}_{calculado}=131.944>\chi^2{}_{2;0.05}=5.991$*, rejeita-se a hipótese nula, isto é, devemos concluir que a preferência pelos vários semanários é influenciada pelo sexo dos leitores.*

167. Em uma indústria química, os engenheiros desejam saber se o alongamento de um composto de borracha permanece inalterado ao passar por uma máquina extrusora. Como o alongamento do composto depende do lote de matéria-prima usado na sua confecção, os dados foram colectados aos pares:

Lote	1	2	3	4	5	6	7	8	9	10
Antes	360	370	380	345	365	380	390	395	385	410
Depois	360	365	355	340	350	370	390	375	375	395

A um nível de significância de 5%, que conclusões foram tiradas?

Resolução

Pelos dados temos:

$$n = 10;\ \bar{x}_d = 10.5; S_d = 8.317$$

1° Passo: Formulação de hipóteses:

$$\begin{cases} \mathrm{H}_0: \mu_d = 0 \\ \mathrm{H}_1: \mu_d \neq 0 \end{cases}$$

2° Passo: Identificação da distribuição amostral

Como o tamanho de amostra é pequena, então:

$$t \sim t_{n-1}$$

3° Passo: Determinação dos valores críticos e região de aceitação ou rejeição da Hipótese nula

Sabe se que o nível de significância é de 1% e o teste é bi-caudal, pela tabela de distribuição normal temos:

$$t_{\frac{\alpha}{2};n-1} = t_{0.025;9} = \pm 2.262$$

Se $-t_{0.025;9} \leq -t_{calculado} \leq t_{0.025;9}$ *não se rejeita a hipótese nula*

4° Passo: Determinação da estatística do teste

$$t_{calculado} = \frac{\bar{x}_d - \mu_d}{\frac{S_d}{\sqrt{n}}} = \frac{10.5}{8.317/\sqrt{10}} = 3.9923$$

5° Passo: Decisão

Como $t_{calculado} = 3.9923 > t_{0.025;9} = 2.262$, *rejeita a hipótese nula*

6ª passo: Conclusões

A um nível de significância de 1%, pode se concluir que o alongamento de um composto de borracha não permanece inalterado ao passar por uma máquina extrusora.

168. Seleccionou-se aleatoriamente uma amostra da cotação diária, em euros, de uma empresa, em relação aos dois últimos meses. Os dados obtidos, após tratamento resultaram na seguinte informação:

$$10,1;\ 10,3;\ 9,9;\ 9,8;\ 10,0;\ 10,2;\ 10,4;\ 10,6;\ 10,1.$$

A cotação diária da empresa é normalmente distribuída. Afirma-se que a variância das cotações da empresa é inferior a 0,04. Considerando um nível de significância de 5%, verifique a validade desta afirmação.

<u>Resolução</u>

Pelos dados temos:

$${\sigma_0}^2 = 0.04; n = 9;\ \bar{x} = 10.2; S^2 = 0.063$$

1° Passo: Formulação de hipóteses:

$$\begin{cases} H_0: \sigma^2 \geq 0.04 \\ H_1: \sigma^2 < 0.04 \end{cases}$$

2° Passo: Identificação da distribuição amostral

$$\chi^2 \sim {\chi^2}_{n-1}$$

3° Passo: Determinação dos valores críticos e região de aceitação ou rejeição da Hipótese nula

Sabe se que o nível de significância é de 5% e o teste é uni-caudal esquerdo, pela tabela de distribuição Qui-Quadrado temos:

$${\chi^2}_{1-\alpha;n-1} = {\chi^2}_{1-0.05;9-1} = {\chi^2}_{0.95;8} = 2.733$$

Se ${\chi^2}_{calculado} > {\chi^2}_{1-\alpha;n-1}$ não se rejeita a hipótese nula

4° Passo: Determinação da estatística do teste

$${\chi^2}_{calculado} = \frac{(n-1) * S^2}{{\sigma^2}_0} = \frac{8 * 0.063}{0.04} = 12.6$$

5° Passo: Decisão

Como ${\chi^2}_{calculado} = 12.6 > {\chi^2}_{0.95;15} = 2.733$, *não se rejeita a hipótese nula*

6ª Passo: Conclusões

A um nível de significância de 5%, pode se concluir que a afirmação não é verdadeira.

169. Em períodos de pico, os clientes de um banco são obrigados a enfrentar longas filas para tirar dinheiro nos caixas electrónicos. Dados históricos de vários anos de operação indicam que o tempo de transacção nessas caixas tem distribuição normal com média igual a 270 segundos. Para aliviar essa situação o banco resolve instalar, em carácter experimental, algumas caixas electrónicos de concepção mais avançada. Após o período de experiência, o banco pretende examinar o tempo médio obtido em uma amostra casual simples das transacções realizadas nesses caixas. Em 24 transacções obteve-se uma média de 262,3 segundos e desvio padrão de 21,4 segundos. Usando um nível de significância de 1%, há evidências suficientes para que o banco substitua as máquinas actuais pelas mais modernas?

Resolução

Pelos dados temos:

$$\mu_0 = 270;\ n = 24;\ \bar{x} = 262.3; S = 21.4$$

1° Passo: Formulação de hipóteses:

$$\begin{cases} H_0: \mu = 270 \\ H_1: \mu < 270 \end{cases}$$

2° Passo: Identificação da distribuição amostral

Como a variança populacional é desconhecida e n pequena, então:

$$t \sim t_{n-1}$$

3° Passo: Determinação dos valores críticos e região de aceitação ou rejeição da Hipótese nula

Sabe se que o nível de significância é de 1% e o teste é bi-caudal, pela tabela de distribuição normal temos:

$$-t_{\alpha;n-1} = -t_{0.01;23} = -2.500$$

$Se\ -t_{\alpha;n-1} \leq -t_{calculado}$ *não se rejeita a hipótese nula*

4° Passo: Determinação da estatística do teste

$$t_{calculado} = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{262.3 - 270}{21.4/\sqrt{24}} = -1.763$$

5° Passo: Decisão

Como $t_{calculado} = -1.763 > t_{\alpha;n-1} = -2.500$, *não se rejeita a hipótese nula*

6ª passo: Conclusões

A um nível de significância de 1%, pode se concluir que não há evidências suficientes para que o banco substitua as máquinas actuais pelas mais modernas.

170. Duas marcas de veículos pretendem comparar o desempenho de seus modelos populares. Para isso, a marca A seleccionou $\mathrm{n} = 12$ veículos de sua produção e fez um teste de consumo. A marca B também retirou uma amostra com $\mathrm{n} = 12$ veículos e realizou o mesmo teste. Os dados de desempenho obtidos pelas marcas em km/litro encontram-se na tabela a seguir:

A	13.5	12.8	11.4	10.9	11.9	12.3	10.7	11.9	10.9	11.5	11.8	12.1
B	12.8	12.8	13.6	13.8	10.1	11.1	11.9	11.4	10.8	12.2	12.4	12.5

a) As empresas afirmam que ambas têm a mesma variabilidade, $\sigma^2 = 0.81$ (os dados estão em km/litro). Conclua a um nível de confiança de 5%.

<u>Resolução</u>

Pelos dados temos:

$$n_1 = n_2 = 12; \bar{x}_1 = 12.12; \bar{x}_2 = 11.81; {\sigma^2}_1 = {\sigma^2}_1 = 0.81$$

1° Passo: Formulação de hipóteses:

$$\begin{cases} \mathrm{H_0}: \mu_1 = \mu_2 \\ \mathrm{H_1}: \mu_1 \neq \mu_2 \end{cases}$$

2° Passo: Identificação da distribuição amostral

$$Z \sim N(0; 1)$$

3° Passo: Determinação dos valores críticos e região de aceitação ou rejeição da Hipótese nula

Sabe-se que o nível de significância é de 2% e o teste é bi-caudal, pela tabela de distribuição Normal Padrão temos:

$$\pm Z_{\frac{\alpha}{2}} = \pm Z_{0.025} = \pm 1.96$$

$$Se\ -Z_{\frac{\alpha}{2}} \leq Z_{calculado} \leq +Z_{\frac{\alpha}{2}}$$ *não se rejeita a hipótese nula*

4° Passo: Determinação da estatística do teste

$$Z_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma^2{}_1}{n_1} + \frac{\sigma^2{}_2}{n_2}}} = \frac{12.12 - 11.81}{\sqrt{\frac{0.81}{12} + \frac{0.81}{12}}} = 0.844$$

5° Passo: Decisão

Como $Z_{calculado} = 0.844 < Z_{0.025} = 1.96$, *não se rejeita a hipótese nula*

6ª Passo: Conclusões

A um nível de significância de 5%, pode-se concluir que as duas marcas de veículos têm o mesmo desempenho.

b) O Estatístico contratado para fazer os cálculos não confiando no valor da variância $\sigma^2 = 0.81$, decidiu refazer as contas considerando as variâncias desconhecidas, porém iguais. Supondo que usou o mesmo nível de significância que conclusão obteve?

Resolução

Pelos dados temos:

$$n_1 = n_2 = 12; \bar{x}_1 = 12.12; \bar{x}_2 = 11.81; S^2{}_1 = 1.1118;\ S^2{}_2 = 0.8163$$

1° Passo: Formulação de hipóteses:

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

2° Passo: Identificação da distribuição amostral

$$t \sim t_{n_1+n_2-2}$$

3° Passo: Determinação dos valores críticos e região de aceitação ou rejeição da Hipótese nula

Sabe-se que o nível de significância é de 5% e o *teste é bi-caudal com varianças desconhecidas e iguais, então:*

$$t_{\frac{\alpha}{2};\, n_1+n_2-2} = t_{0.025;22} = \pm 2.0739$$

$$Se - t_{0.025,22} \leq t_{calculado} \leq t_{0.025,22}$$ *não se rejeita a hipótese nula*

4° Passo: Determinação da estatística do teste

$$t_{calculado} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{(n_1 - 1) * S_1{}^2 + (n_2 - 1) * S_2{}^2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{12.12 - 11.81}{\sqrt{\frac{0.9512}{12} + \frac{0.9512}{12}}} = 0.7786$$

5° Passo: Decisão

Como $t_{calculado} = 0.7786 < t_{0.025,22} = 2.0739$, *não se rejeita a hipótese nula*

6ª Passo: Conclusões

A um nível de significância de 5%, pode-se concluir que as duas marcas de veículos têm o mesmo desempenho.

c) Alguns inspectores, ainda mais desconfiados, foram investigar as informações e descobriram que a empresa A havia omitido 4 valores de sua amostra. São eles: 10.0; 10.5; 10.6 e 11.5. Considerando, agora, variâncias diferentes e desconhecidas, testar para a igualdade de médias com um nível significância de 0.05.

Resolução

Pelos dados temos:

$$n_1 = 16;\ n_2 = 12; \bar{x}_1 = 11.75; \bar{x}_2 = 11.81; {S^2}_1 = 1.1894;\ {S^2}_2 = 0.8163$$

1° Passo: Formulação de hipóteses:

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

2° Passo: Identificação da distribuição amostral

$$t \sim t_{GL}$$

3° Passo: Determinação dos valores críticos e região de aceitação ou rejeição da Hipótese nula

Sabe-se que o nível de significância é de 5% e o teste é uni-caudal esquerdo com varianças desconhecidas e diferentes, então:

$$Gl = \frac{\left(\frac{{S^2}_1}{n_1} + \frac{{S^2}_2}{n_2}\right)^2}{\frac{\left(\frac{{S^2}_1}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{{S^2}_2}{n_2}\right)^2}{n_2 - 1}} = \frac{\left(\frac{1.1894^2}{16} + \frac{0.8163^2}{12}\right)^2}{\frac{\left(\frac{(1.1894)^2}{16}\right)^2}{15} + \frac{\left(\frac{(0.8163)^2}{12}\right)^2}{11}} \approx 26$$

$$t_{\frac{\alpha}{2};\,Gl} = t_{0.025,26} = \pm 2.056$$

$Se - t_{0.025,26} \leq t_{calculado} \leq t_{0.025,26}$ *não se rejeita a hipótese nula*

4° Passo: Determinação da estatística do teste

$$t_{calculado} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{{S^2}_1}{n_1} + \frac{{S^2}_2}{n_2}}} = \frac{11.75 - 11.81}{\sqrt{\frac{1.1894}{16} + \frac{0.8163}{12}}} = -0.1590$$

5° Passo: Decisão

Como $t_{calculado} = -0.1590 > -t_{0.025;26} = -2.056$, *não se rejeita a hipótese nula*

6ª Passo: Conclusões

REGRESSÃO LINEAR SIMPLES

171. Assinal com F (Falsa) ou V (Verdade) as afirmações seguintes:
   a) Um coeficiente de correlação de +1,0 entre duas variáveis X e Y indica que X causa Y, mas um coeficiente de correlação de -1,0 significa que X não causa Y. Falsa
   b) Se o coeficiente de determinação é zero, o coeficiente de correlação é também zero. Verdade
   c) Se o coeficiente angular é 1 (um), isto significa que existe perfeita correlação entre X e Y. Falsa
   d) É possível que o coeficiente de correlação amostral seja positivo, quando não existe, de facto, nenhuma correlação entre as variáveis X e Y. Verdade
   e) Se o coeficiente linear é zero, isto significa que não existe regressão entre X e Y. Falsa
   f) Não se pode utilizar a técnica da regressão pelo método dos mínimos quadrados quando a relação básica entre X e Y não for linear. Falsa

172. A análise de 13 pares de valores indicou que existe uma relação linear entre as vendas (em u.m) e o Lucro (em u.m). A equação obtida foi $\widehat{Y} = 13.4812 + 0.0196X$. Os dados desses pares forneceram os seguintes resultados:

$$S_{VV} = 485063.2;\ S_{VL} = 9522.3;\ S_{LL} = 215.1;\ \sum V_i = 5896:\ \sum L_i = 291$$

$$\text{onde V} = \text{vendas e L} = \text{lucro}$$

   a) Determine a dispersão dos termos erro em torno da linha de regressão.

Resolução

Seja $Y = Lucro\ e\ X = vendas$

Pela fórmula temos:

$$S = \sqrt{\frac{s_{yy} - \hat{\beta}_1 * s_{xy}}{n-2}}$$

$$S = \sqrt{\frac{215.1 - 0.0196 * 9522.3}{13-2}} = 1.5995$$

Resposta: O desvio padrão do modelo é igual a 1.5995, isto é, em média o lucro previsto desvia-se das vendas em aproximadamente 1.5995 u.m.

   b) Faça a análise de Variança e conclua a respeito da significância do modelo;

Resolução

Hipóteses do Modelo:

$$\begin{cases} H_0: Y = \beta_0 + \varepsilon \\ H_1: Y = \beta_0 + \beta_1 X + \varepsilon \end{cases}$$

Tabela ANOVA (Análise de Variança)

Fonte de variação	GL	$Soma\ dos\ Quadrados$	$Quadrado\ Médio$	F
Regressão	$p-1$	$\widehat{\beta_1} * S_{xy}$	$\dfrac{\widehat{\beta_1} * S_{xy}}{p-1}$	$\dfrac{QM_{Reg}}{QM_{Res}}$
Resíduos	$n-p$	$S_{yy} - \widehat{\beta_1} * S_{xy}$	$\dfrac{S_{yy} - \widehat{\beta_1} * S_{xy}}{n-p}$	
Total	$n-1$	S_{yy}		

Substituíndo teremos:

Fonte de variação	GL	$Soma\ dos\ Quadrados$	$Quadrado\ Médio$	F
Regressão	1	186.933	186.933	73.0625
Resíduos	11	28.144	2.5585	
Total	12	215.1		

$$F_{\alpha;p-1;n-2} = F_{0.05;1;11} = F_{0.05;1;11} = 4.84$$

Como $F_{0.05;1;10} = 4.75 < F_{calculado} = 73.0625$, confirma-se a significância do modelo, isto é, o modelo completo (o que contém a variável X) é melhor do que o modelo reduzido (o que não contém a variável X).

c) Determine a percentagem da variação total do Lucro não explicada pelo modelo.

Resolução

Aqui pede-se $1 - R^2$:

Então

$$1 - R^2 = 1 - \frac{SQ_{reg}}{SQ_{total}}$$

$$1 - R^2 = 1 - \frac{186.933}{215.1} = 1 - 0.869 = 0.131$$

Logo:

A percentagem da variação total do Lucro não explicada pelo modelo é de 13.1%.

d) Faça uma previsão, através de um intervalo, para o lucro total médio, para uma venda de 400 u.m, utilizando uma confiança de 95%.

Resolução

Pela fórmula temos:

$$\hat{Y}_0 \pm t_{\frac{\alpha}{2};n-2} * S * \sqrt{\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{S_{xx}}} \ \epsilon \ Y$$

$$(13.4812 + 0.0196 * 400) \pm 2.201 * 1.5995 * \sqrt{\frac{1}{13} + \frac{(400 - 453.2)^2}{485063.2}} \ \epsilon \ Y_{x=400}$$

$$20.3084 \leq Y_{x=400} \leq 22.3339$$

Resposta: A um nível de confiança de 95%, pode-se concluir que a previsão para o lucro total médio, para uma venda de 400 u.m encontra-se entre 20.3084 e 22.3339 u.m.

173. O sucesso de um programa de investimento em papéis de outros países depende, em grande parte, do controlo do risco soberano dos países constantes no portfólio. Um meio de se avaliar o risco é através da taxa de juros praticada pelo país. Uma amostra de 10 países forneceu os valores das taxas de risco e de juros praticadas em 1997.

Taxa de risco	87	74.1	64.8	53.7	47.5	80.7	90.3	41.4	39.6	38.4
Taxa de juros	72.7	61.2	53.7	40.8	49.1	43.3	72.9	30.6	42.9	29.6

Considere os seguintes resultados:

$$S_{\varepsilon i} = 11.68; \ S_{RR} = 3691.425; \ S_{JJ} = 2151.076$$

Onde : RR=Taxa de risco e JJ=Taxa de Juros

a) Indique a variável independente e dependente.

Resolução

Variável independenete: Taxa de juros

Variável dependenete: Taxa de risco

b) É possível afirmar, com uma significância de 1%, que o coeficiente de correlação é significativamente diferente de zero?

Resolução:

$$\begin{cases} H_0: \rho = 0 \\ H_1: \rho \neq 0 \end{cases}$$

$$\alpha = 0.01;$$

Como a amostra contém n pares de dados, então:

$$t_{\frac{\alpha}{2};n-2} = t_{0.005;8} = \pm 3.355$$

Estatística do teste:

$$t_{calculado} = r * \sqrt{\frac{n-2}{1-r^2}}$$

$$t_{calculado} = 0.839 * \sqrt{\frac{10-2}{1-0.704}} = 4.362$$

Decisão:

$$t_{calculado} = 4.362 > t_{0.025;8} = 3.355$$

Rejeita-se H_0

Conclusão:

*Conclui-se que o valor de $r = 0.839$ obtido para a amostra é significante e que existe correlação **r** entre as variáveis taxa de juros e taxa de risco com nível de significância igual a 1%.*

c) Determine a equação de regressão de Y em função de X, caso o coeficiente de correlação seja significativamente diferente de zero e diga qual o significado económico das estimativas "β_0" e "β_1"?

<u>Resolução</u>

Pelos dados temos que:

$$\widehat{\beta_0} = 7.131 \text{ e } \widehat{\beta_1} = 1.099$$

Enttão:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$$

$$\hat{Y} = 7.131 + 1.099X$$

Interpretação do modelo:

$$\widehat{\beta_0} = 7.131$$

Se a Taxa de Juros (X) for igual a zero, a Taxa de Risco será igual a 7.131 %.

$$\widehat{\beta_1} = 1.099$$

A cada unidade adicional na variável Taxa de Juros (X), a variável Taxa de Risco (Y) aumenta em 1.099%.

d) Determine um intervalo de confiança de 90% para a taxa de risco marginal desses países.

<u>Resolução</u>

$$\hat{\beta}_1 \sim N\left(\beta_1; \sigma * \sqrt{\frac{1}{S_{xx}}}\right)$$

$$\widehat{\beta_1} \pm t_{\frac{\alpha}{2};n-2} * S * \sqrt{\frac{1}{S_{xx}}} \in \beta_1$$

$$1.099 \pm 1.860 * 11.68 * \sqrt{\frac{1}{2151.076}} \in \beta_1$$

$$0.6306 \leq \beta_1 \leq 1.5674$$

Resposta: Existe 95% de confiança de a verdadeira taxa de risco marginal pertencer ao intervalo (0.6306 - 1.5674).

e) Testar se a taxa de risco fixa desses países pode ser considerada diferente do zero (Use significância de 5%).

Resolução:

1° Passo: Formulação de hipóteses

$$\begin{cases} H_0: \beta_0 = 0 \\ H_1: \beta_0 \neq 0) \end{cases}$$

2° Passo: Fixar o nível de significância

$$\alpha = 0.05$$

3° Passo: Identificar a distribuição amostral e determinar a estatística do teste

$$\widehat{\beta}_0 N\left(\beta_0; \sigma * \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}}\right)$$

$$t_0 = \frac{\hat{\beta}_0 - \beta_0}{S * \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}}}$$

$$t_{calculado} = \frac{7.131}{11.68 * \sqrt{\frac{1}{10} + \frac{49.68^2}{2151.076}}} = 0.547$$

4° Passo: Com base ao tipo do teste, determinar os valores críticos e as regiões de rejeição

Como trata-se dum teste bilateral, então:

$$t_{\frac{\alpha}{2}; n-2} = t_{0.025; 8} = 2.306$$

Se $-t_{\frac{\alpha}{2}; n-2} \leq t_{calculado} \leq t_{\frac{\alpha}{2}; n-2}$ *não se rejeita a* H_0

5° Passo: Decisão

$-t_{\frac{\alpha}{2}; n-2} = -2.306 < t_{calculado} = 0.547 < t_{\frac{\alpha}{2}; n-2} = 2.036$ *não se rejeita a* H_0

6° Passo: Conclusão

A um nível de significância de 5%, a taxa de risco fixa desses países não pode ser considerada diferente do zero.

174. Com a finalidade de se tentar explicar o consumo de combustível de automóveis de passeio através de seus pesos, seleccionou-se aleatoriamente 14 automóveis do mesmo ano de fabricação, pertencentes à frota de uma grande empresa. Foram registrados o peso (kg) e o consumo (l) dos automóveis durante um certo trecho de uma determinada estrada. Os dados obtidos foram os seguintes:

Consumo	10.4	14.4	10.4	13.2	14.0	12.4	9.2	6.8	9.2	12.0	10.0	8.0	7.6	11.2
Peso	1232	880	1320	880	880	880	1320	1760	1320	1232	1760	1760	1232	1232

Cálculos Parciais:

$$\sum P_i = 17688; \sum C_i = 149; \sum {P_i}^2 = 23688896; \sum {C_i}^2 = 1654; \sum P_i * C_i = 179942.4$$

$$\text{onde P} = \text{Peso e C} = \text{Consumo}$$

a) Ajuste um Modelo de Regressão Linear Simples de Y em X, e interprete as estimativas dos parâmetros.

<u>_Resolução_</u>

$$Seja\ X = Peso\ e\ Y = Consumo$$

$$s_{xy} = \sum_{i=1}^{n} y_i x_i - n\bar{y}\bar{x} = 179942.4 - 14 * 1263.4 * 10.6 = -7546.16$$

$$s_{xx} = \sum_{i=1}^{n} {x_i}^2 - n\,\bar{x}^2 = 23688896 - 14 * 1263.4^2 = 1342382.16$$

$$s_{yy} = \sum_{i=1}^{n} {y_i}^2 - n\bar{y}^2 = 1654 - 14 * 10.6^2 = 80.96$$

Logo

$$\hat{\beta}_1 = \frac{s_{xy}}{s_{xx}} = \frac{-7546.16}{1342382.16} = -0.006$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = 10.6 - (-0.006) * 1263.4 = 18.2$$

Então:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X = 18.2 - 0.006X$$

Interpretação do modelo:

$$\widehat{\beta_0} = 18.2$$

$\widehat{\beta_0} = 18.2$: *Se o peso for igual a zero, o consumo será de 18.2 litros. Nota que esta interpretação não tem nenhum significado económico.*

$\widehat{\beta_1}$ = *-0.006: A cada 1Kg adicional no peso do automóvel , o consumo diminue em 0.006 litros.*

b) Faça a análise de Variança e conclua a respeito da significância do modelo.

Resolução

Hipóteses do Modelo:

$$\begin{cases} H_0: Y = \beta_0 + \varepsilon \\ H_1: Y = \beta_0 + \beta_1 X + \varepsilon \end{cases}$$

Tabela ANOVA (Análise de Variança)

Fonte de variação	GL	$Soma\ dos\ Quadrados$	$Quadrado\ Médio$	F
Regressão	$p-1$	$\widehat{\beta_1} * S_{xy}$	$\frac{\widehat{\beta_1} * S_{xy}}{p-1}$	$\frac{QM_{Reg}}{QM_{Res}}$
Resíduos	$n-p$	$S_{yy} - \widehat{\beta_1} * S_{xy}$	$\frac{S_{yy} - \widehat{\beta_1} * S_{xy}}{n-p}$	
Total	$n-1$	S_{yy}		

Substituíndo teremos:

Fonte de variação	GL	$Soma\ dos\ Quadrados$	$Quadrado\ Médio$	F
Regressão	1	48.4	48.4	23.7
Resíduos	12	24.5	2.04	
Total	13	72		

$$F_{\alpha;p-1;n-2} = F_{0.05;1;11} = F_{0.05;1;12} = 4.75$$

Como $F_{0.05;1;10} = 4.75 < F_{calculado} = 23.7$, confirma-se a significância do modelo, isto é, o modelo completo (o que contém a variável X) é melhor do que o modelo reduzido (o que não contém a variável X).

c) Determine o coeficiente de determinação e interprete-o.

Resolução

$$R^2 = \frac{SQ_{reg}}{SQ_{total}} = \frac{48.4}{72} = 67.2$$

Interpretação: 67.2% da variação total do Consumo de combustível de automóveis de passeio é explicado pelo seu peso, e os restantes 32.8% por outros factores.

175. Realizou-se uma pesquisa de mercado com o objectivo de estudar a relação entre o tempo necessário para um consumidor tomar uma decisão (sobre o que comprar) e o número de embalagens alternativas do mesmo produto apresentadas a esse consumidor. Eliminaram-se as marcas das embalagens, a fim de reduzir o efeito da preferência por uma ou outra marca. Os consumidores fizeram suas escolhas somente com base na descrição do produto, anotada nas embalagens pelos fabricantes. A análise de regressão do tempo necessário em minutos, para que cada um tomasse sua decisão foi anotado para 15 participantes. As tabelas abaixo, mostram os resultados da análise de regressão feita.

Regression Statistics	
Multiple R	0.728464397
R Square	0.530660377
Adjusted R Square	0.494557329
Standard Error	1.237242592
Observations	15

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	4.3	1.216447	3.534884	0.003661
Número de alternativas	1.5	0.39125	3.833861	0.00207

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	22.5	22.5	14.7	0.0021
Residual	13	19.9	1.530769		
Total	14	42.4			

a) Desenvolva uma equação de regressão e interprete os respectivos parâmetros.

<u>Resolução</u>

Pelos dados temos que:

$$\widehat{\beta_0} = 4.3 \text{ e } \widehat{\beta_1} = 1.5$$

Então:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$$

$$\hat{Y} = 4.3 + 1.5X$$

Interpretação do modelo:

$$\widehat{\beta_0} = 4.3$$

Se o nº de alternativas (X) for igual a zero, o tempo necessário para um consumidor tomar uma decisão será igual a 4.3 minutos. Nota que este valor não tem nenhum significado económico.

$$\widehat{\beta_1} = 1.5$$

A cada unidade adicional na variável nº de alternativas (X), tempo necessário para um consumidor tomar uma decisão aumenta em 1.5 minutos.

b) Teste a significância geral do modelo.

<u>Resolução</u>

Hipóteses do Modelo:

$$\begin{cases} H_0: Y = \beta_0 + \varepsilon \quad ou \;\; \beta_0 = \beta_1 = 0 \\ H_1: Y = \beta_0 + \beta_1 X + \varepsilon \quad ou \; \nexists \beta_i \neq 0 \end{cases}$$

Recorrendo a tabela ANOVA temos:

$$P - valor = 0.0021 < 0.05$$

Rejeita-se a H_0

Conclusão: Confirma-se a significância do modelo, isto é, o modelo completo (o que contém a variável X) é melhor do que o modelo reduzido (o que não contém a variável X).

c) Teste a significância de β_0 e β_1 , separadamente.

Resolução

$$\begin{cases} H_0: \beta_0 = 0 \\ H_1: \beta_0 \neq 0) \end{cases}$$

Recorrendo a tabela Coeficientes temos:

$$P - valor = 0.0037 < 0.05$$

Rejeita-se a H_0

Conclusão: A um nível de significância de 5%, o tempo fixo necessário para um consumidor tomar uma decisão não pode ser considerado igual do zero.

$$\begin{cases} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0) \end{cases}$$

Recorrendo a tabela Coeficientes temos:

$$P - valor = 0.00207 < 0.05$$

Rejeita-se a H_0

Conclusão: A um nível de significância de 5%, o tempo marginal necessário para um consumidor tomar uma decisão não pode ser considerado igual do zero.

d) O que pode-se concluir sobre o poder explicativo do modelo?

Resolução

Recorrendo a tabela Estatísticas de Regressão temos:

$R^2 = 53.1\%$ *o que significa que: 53.1% da variação total do tempo fixo necessário para um consumidor tomar uma decisão é explicado pelo modelo.*

e) Tente prever o tempo necessário para um consumidor tomar uma decisão (sobre o que comprar) para um número de alterntivas igual a 15.

Resolução

Recorrendo a equação:

$$\hat{Y} = 4.3 + 1.5X$$

$$temos:$$

$$\hat{Y} = 4.3 + 1.5 * 15 = 6.55$$

Resposta: O tempo necessário para um consumidor tomar uma decisão (sobre o que comprar) para um número de alterntivas igual a 15 é de 6.55 minutos.

BIBLIOGRAFIA

1. Mahaluça, F.(2016). Estatística Aplicada. Disponível em: WWW.ResearchGate.net
2. Mahaluça, F.(2016). Exercícios de Estatística Aplicada. Disponível em: WWW.ResearchGate.net
3. Mahaluça, F.(2011). Livro de Estatística