Estatística

Sérgio Manuel Salazar dos Santos, N^o: 1020881

13 de Dezembro de 2019

Conteúdo

.1 Introdução 1
.2 O conjunto de dados 1
.3 Metodologia Estatística 4
  .3.1 TEE $IC_{95\%}$ 4
  .3.2 NEE 4
.4 Resultados e interpretação 4
.5 Conclusões 4

Resumo

Este trabalho consiste no estudo de Estatística das Entregas Expresso em duas regiões **A** e **B**, as variaveis em estudo é o tempo de demora das entregas e a variavel de numero de encomendas entregues num determinado unidade de tempo [u.t.]. Nestas situações foram retiradas 120 e 90 amostras nas duas regiões respectivamente.
A primeira é uma distribuição continua, o tempo, e a segunda uma distribuição discreta.
As materias abordadas vai ser **Amostragem**, **Estimação de parâmetros** e **Testes de Hipóteses**

.1 Introdução

As variáveis consideradas são:

- Regiao (REG): variável nominal com dois niveis
  Regiao A
  Região B
- Tempo de entrega (TEE), por encomenda: Variável expressa em u.t.
- Número de encomendas entregues (NEE) por u.t.

Admitindo que a amostra disponível é uma amostra aleatória representativa das populações.

Neste relatorio esta-se a trabalhar com duas grandezas precisamente o tempo (TEE) e quantidade por u.t (NEE), temos recolhidos 120 registos **TEE** na qual pela regra de sturges $c = int(1+3.3log(n))$, determina-se que é necesario sete [**7**] classes.
Podemos obter a amplitude de cada classe $h = b - a$ e sua marca $x_i = \frac{a+b}{2}$.

.2 O conjunto de dados

X_{Ai}- "Variavel aleatoria que representa o tempo de demora na Região **A** da entrega de uma encomenda Expresso em u.t." i=1,2,3, ...,117
X_{Bi}- "Variavel aleatoria que representa o tempo de demora na Região **B** da entrega de uma encomenda Expresso em u.t." i=1,2,3, ...,120
Abaixo o resultado da tabela TEE:

h_i	CLASSE	MARCA	nA_i	nB_i	$\frac{nA_i}{h_i}$	$\frac{nB_i}{h_i}$	$f_i(A)$	$f_i(B)$	$F_i(A)$	$F_i(B)$
4	[5,10[	7,5	5	0	1,25	0	0,0427	0	0,0427	0
4	[10,15[	12,5	16	19	4	4,75	0,1368	0,1583	0,1795	0,1583
4	[15,20[	17,5	40	28	10	7	0,3419	0,2333	0,5214	0,3917
4	[20,25[	22,5	25	41	6,25	10,25	0,2137	0,3417	0,7350	0,7333
4	[25,30[	27,5	26	22	6,5	5,5	0,2222	0,1833	0,9573	0,9167
4	[30,35[	32,5	5	8	1,25	2	0,0427	0,0667	1	0,9833
5	[35,40]	37,5	0	2	0	0,4	0	0,0167	1	1
			n=117	n=120						

Recorrendo ao excell obeteve-se os seguintes resultados:

Média aritmetica dados classificados	Variância de uma amostra dados classificados
$\bar{x} = \frac{1}{n}\sum_{i=1}^{c} x_i n_i = \sum_{i=1}^{c} x_i f_i$	$s^2 = \frac{1}{n-1}\sum_{i=1}^{c}(x_i - \bar{x})^2 n_i$

Estatística	X_A	X_B
Mínimo	7,5	12,5
Q_1:1^o Quartil	17,5	17,5
m_d: mediana	17,5	22,5
Q_3:3^o Quartil	27,5	27,5
Máximo	32,5	37,5
$\bar{X}$: Média	20,3205	21,5833
s : desvio-padrão	6,1020	6,0106
m_o: moda	17,5	22,5
Tamanho amostral [n]	117	120

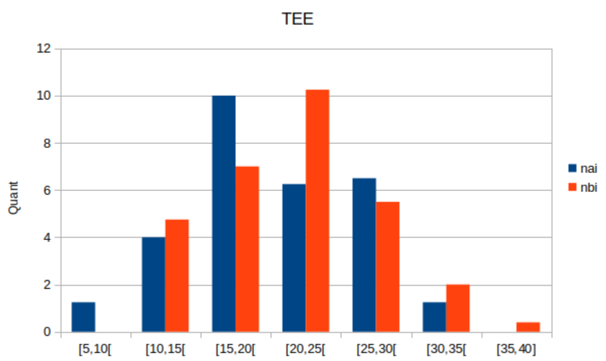

Figura 1: TEE

Na Região **A** a Média > Moda = Mediana
com skew = -0,02876 e kurt = -0,50909
Na Região **B** a Média < Moda = Mediana
com skew = 0,14205 e kurt = -212187

Na prática, considera-se que a qualidade da aproximação é suficientemente boa quando $n \geqslant 30$.
Pode-se tomar que $\delta \cong s$.

$$\left\{ \begin{array}{l} \mu \\ \delta \end{array} \right. \qquad \Longrightarrow \qquad \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \sim N\left(\mu; \frac{\delta^2}{n}\right)$$

$\bar{x}_{A_0}$ = 20,3205 $\quad$ $\bar{x}_{B_0}$ = 21,5833
δ_A = 6,1020 $\quad$ δ_B = 6,0106

Tratamento dos dados da Segunda Variavel Aleatótia

Y_{Ai}- "Variavel aleatoria que representa o numero de encomendas entregues pela Expresso na Regiao **A** por u.t." i=1,2,3, ...,90

Y_{Bi}- "Variavel aleatoria que representa a numero de encomendas entregues pela Expresso na Regiao **B** por u.t." i=1,2,3, ...,90

Abaixo o resultado da tabela NEE:

Y_i	nA_i	nB_i	$f_i(A)$	$f_i(B)$	$F_i(A)$	$F_i(B)$
3	6	3	0,0667	0,0333	0,0667	0,0333
4	8	6	0,0889	0,0667	0,1556	0,1
5	19	13	0,2111	0,1444	0,3677	0,2444
6	15	7	0,1667	0,0778	0,5333	0,3222
7	13	19	0,1444	0,2111	0,6778	0,5333
8	11	15	0,1222	0,1667	0,8	0,7
9	6	8	0,0667	0,0889	0,8667	0,7889
10	5	11	0,0556	0,1222	0,9222	0,9111
11	4	3	0,0444	0,0333	0,9667	0,9444
12	0	2	0	0,0222	0,9667	0,9667
13	2	1	0,0222	0,0111	0,9889	0.9778
14	1	0	0,0111	0	1	0,9778
15	0	1	0	0,0111	1	0,9889
16	0	1	0	0,0111	1	1

Estatística	Y_A	Y_B
Mínimo	3	3
Q_1:1^o Quartil	5	6
m_d: mediana	6	7
Q_3:3^o Quartil	8	9
Máximo	14	16
$\bar{Y}$: Média	6,6889	7,5111
s : desvio-padrão	2,4062	2,5139
m_o: moda	5	7
Tamanho amostral [n]	90	90

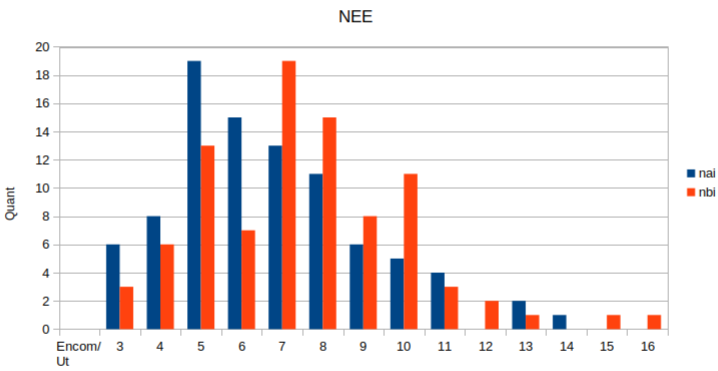

Figura 2: NEE

$$\begin{cases} \mu \\ \delta \end{cases} \Longrightarrow \qquad \bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} \sim N(\mu; \frac{\delta^2}{n})$$

$\bar{y}_{A_0} = 6,6889$ $\quad \bar{y}_{B_0} = 7,5111$

$\delta_A = 2,4062$ $\quad \delta_B = 2,5139$

.3 Metodologia Estatística

.3.1 $TEE \quad IC_{95\%}$

Estimação do tempo médio para as regiões **A** e **B** com um indice de confiança de 95%.

$IC_{1-\alpha} = [A, B]$; para $1-\alpha = 0.95$, $\alpha = 0.05$, $\frac{\alpha}{2} = 0.025$
Zona critica $Z_c = \Phi^{-1}(0.025) \approx 1.96$
$P(A \leqslant \mu \leqslant B) = 1 - \alpha$
$\triangle = Z_c \times \frac{\delta}{\sqrt{n}}$
$A = \bar{x} - \triangle \qquad and \qquad B = \bar{x} + \triangle$
$\therefore$
$IC_{A_{0.95}} = [\quad 19.2148\,,\, 21.4262 \quad]$ and $IC_{B_{0.95}} = [\quad 20.5078\,,\, 22.6587 \quad]$

Pode-se estimar que o tempo médio $[\,\mu\,]$ de entrega na população esta dentro dos intervalos acima mencionados com 95% de confiança.

.3.2 NEE

Verificar se os dados permitem afirmar que existe diferença significativa entre a % de períodos com menos de **6** entregas por u.t. na região **A** e na região **B**. Responda com base num intervalo de confiança de 97%.
P(X_A < 6)= ? e P(X_B < 6)= ? then $P_A - P_B$ then $IC_{97\%}$

.4 Resultados e interpretação

fazer tabela só com resultados

.5 Conclusões

Lista de Figuras

1 TEE 1
2 NEE 3
[]

Bibliografia

[1] *Probabilidades e estatística Volume 1*. McGraw-Hill, 1990.

[2] *Probabilidades e Processos Estocásticos*. Universidade de Aveiro, 2002.

[1]

[1] Apontamentos Estatistica