

## Exercício TS.2

Nos postos de trabalho PT-23 e PT-37 de um *call center*, os tempos de atendimento por ordem cronológica de chamadas dirigidas à empresa “Simuluso” foram escrutinadas durante alguns dias para determinar qual a distribuição de probabilidade de melhor aderência, com o objectivo de caracterizar esta variável para integrar um modelo de simulação. Este modelo permitirá determinar quantos atendedores deverão ser contratados de modo a assegurar que, pelo menos, 90% das chamadas serão atendidas dentro de uma janela temporal de 2 minutos.

No PT-23 foram registados 70 tempos de atendimento por ordem cronológica escolhidos de forma aleatória durante alguns dias. O próximo Quadro mostra estes tempos em segundos.

Ordem	Tempo	Ordem	Tempo	Ordem	Tempo	Ordem	Tempo	Ordem	Tempo
1	169	16	114	31	95	46	107	61	123
2	155	17	91	32	97	47	117	62	118
3	125	18	123	33	170	48	107	63	129
4	108	19	107	34	113	49	93	64	94
5	84	20	164	35	121	50	89	65	103
6	81	21	82	36	109	51	78	66	88
7	97	22	137	37	87	52	97	67	78
8	141	23	47	38	139	53	116	68	118
9	120	24	133	39	108	54	97	69	69
10	104	25	150	40	85	55	83	70	108
11	78	26	113	41	110	56	174	71	-
12	57	27	62	42	125	57	107	72	-
13	101	28	136	43	85	58	93	73	-
14	101	29	105	44	96	59	47	74	-
15	54	30	116	45	129	60	89	75	-

No PT-37 foram registados 60 tempos de atendimento por ordem cronológica escolhidos de forma aleatória durante alguns dias. O próximo Quadro mostra estes tempos em segundos.

Ordem	Tempo	Ordem	Tempo	Ordem	Tempo	Ordem	Tempo	Ordem	Tempo
1	17	16	119	31	101	46	89	61	-
2	137	17	78	32	30	47	180	62	-
3	171	18	16	33	26	48	233	63	-
4	173	19	67	34	4	49	57	64	-
5	55	20	55	35	119	50	556	65	-
6	87	21	18	36	73	51	265	66	-
7	95	22	182	37	45	52	13	67	-
8	37	23	189	38	47	53	55	68	-
9	3	24	16	39	22	54	171	69	-
10	218	25	49	40	84	55	98	70	-
11	272	26	61	41	93	56	270	71	-
12	541	27	44	42	36	57	536	72	-
13	3	28	24	43	39	58	115	73	-
14	176	29	9	44	45	59	7	74	-
15	83	30	8	45	48	60	175	75	-

### Questões:

1. Trate em frequência as observações obtidas de cada PT, represente o histograma e identifique os *outliers* (se algum). Usar a aplicação EXCEL “Análise\_frequência”;
2. Teste estes dados quanto à sua aleatoriedade a um nível de significância de  $\alpha = 5\%$ , de modo a despistar a eventual existência de correlação com algum acontecimento específico. Use os métodos das “alternâncias” e da “recta de ajustamento”, mostre os gráficos da cronologia e da estatística de teste e descreva as conclusões de cada um dos dois métodos para cada um dos PT’s. Usar a aplicação EXCEL “Teste\_alternancias” e “Teste\_recta\_ajustamento”;

3. No PT23, se aceitássemos uma margem máxima de erro igual a 7 segundos, teriam sido necessárias as 70 observações realizadas? Se não, quantas teriam bastado? E qual teria sido o valor médio a considerar em futuros cálculos? Usar a aplicação EXCEL “**Observações\_continuas**”
4. No PT37, se aceitássemos uma margem máxima de erro igual a 7 segundos, teriam sido necessárias as 60 observações realizadas? Se sim, quantas mais teriam sido necessárias? Usar a aplicação EXCEL “**Observações\_continuas**”
5. Para realizar um novo estudo de tempos, gere um horário possível de observações instantâneas com as seguintes regras:

Manhã: Entre as 09.00 e as 12.00 horas

Tarde: Entre as 14.00 e as 18.00 horas

Intervalo de tempo mínimo entre cada duas observações: 10 minutos

Intervalo de tempo máximo entre cada duas observações: 30 minutos

Usar a aplicação EXCEL “**Horario\_observações**”

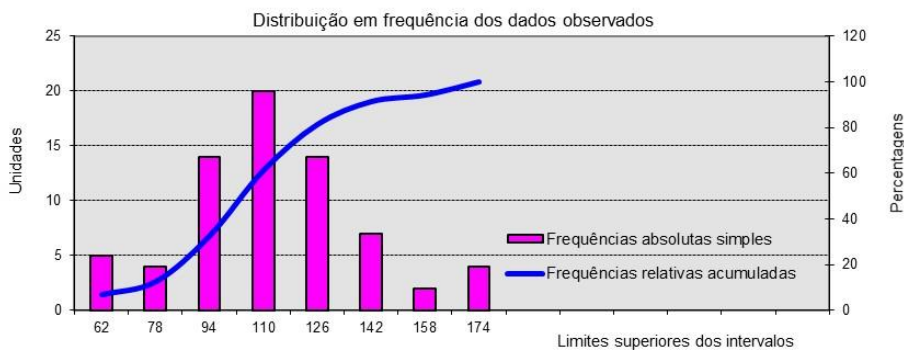
**Nota:** Em teoria das probabilidades e estatística, uma sequência ou outra coleção de variáveis aleatórias é **independente e identicamente distribuída** (*i.i.d.* ou *iid* ou **IID**) se cada variável aleatória tiver a mesma distribuição de probabilidade das outras e todas forem mutuamente independentes. Em consequência, uma sequência cronológica de variáveis aleatórias não apresenta tendência – nem crescente nem decrescente.

Uma coleção de variáveis aleatórias apenas identicamente distribuídas é frequentemente abreviada como **IID**.

**1.**  
**PT-23**

Depois de introduzirmos os dados (70 observações de tempos de operação em segundos) na coluna B:B da aplicação “Análise frequencia.XLSX” e aceitarmos os valores dos parâmetros sugeridos para a construção do Histograma, obtemos o seguinte Quadro e correspondente gráfico. Este mostra a distribuição em frequências absolutas  $f(x)$  sob a forma de polígonos, cujas amplitudes se encontram representadas nas ordenadas do lado esquerdo e a distribuição em frequências relativas acumuladas  $F(x)\%$  sob a forma de uma curva, cujas amplitudes se encontram representadas nas ordenadas do lado direito.

Definição dos intervalos de classe					
Qual o limite superior do 1º intervalo ?	62			62	
Qual o nº de intervalos pretendido ?	8			8	
Qual a amplitude de cada intervalo ?	16			16	
Análise em frequência					
Lim.inf. intervalo	Lim.sup. intervalo	$f(x)$	$f(x)\%$	$F(x)$	$F(x)\%$
47	62	5	7,142857143	5	7,142857143
63	78	4	5,714285714	9	12,85714286
79	94	14	20	23	32,85714286
95	110	20	28,57142857	43	61,42857143
111	126	14	20	57	81,42857143
127	142	7	10	64	91,42857143
143	158	2	2,857142857	66	94,28571429
159	174	4	5,714285714	70	100



Outliers	
LIA =	41,375
LSA =	168,375
Nº existente =	3

Pela regra de Sturges  $K = 1 + \log_2(n)$  o nº de intervalos de classe deve ser igual a 8.

A figura supra mostra a distribuição em frequência dos dados (absolutas e relativas, simples e acumuladas), a qual apresenta uma forma genericamente simétrica.

Uma vez os dados filtrados pelos limites de aceitação inferior e superior:

$$LIA = Q_1 - 1,5 \cdot (Q_3 - Q_1) = 89 - 1,5 \times (120,75 - 89) = 41,375 \text{ segundos}$$

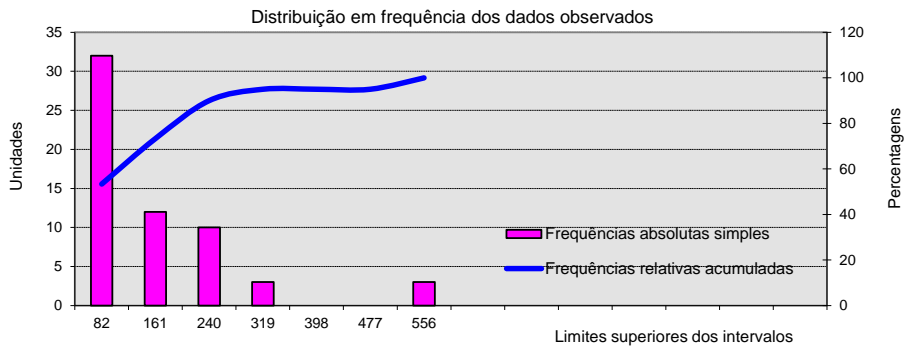
$$LSA = Q_3 + 1,5 \cdot (Q_3 - Q_1) = 120,75 + 1,5 \times (120,75 - 89) = 168,375 \text{ segundos}$$

Obtemos os *outliers* (valores atípicos ou particularmente elevados ou baixos): 169, 170 e 174

**PT-37**

Repetindo o procedimento que seguimos no caso do PT-23, obtemos:

Definição dos intervalos de classe					
Qual o limite superior do 1º intervalo ?	82	82			
Qual o nº de intervalos pretendido ?	7	7			
Qual a amplitude de cada intervalo ?	79	79			
Análise em frequência					
Lim.inf. intervalo	Lim.sup. intervalo	f(x)	f(x)%	F(x)	F(x)%
3	82	32	53,33333333	32	53,33333333
83	161	12	20	44	73,33333333
162	240	10	16,66666667	54	90
241	319	3	5	57	95
320	398	0	0	57	95
399	477	0	0	57	95
478	556	3	5	60	100



Outliers	
LIA =	-170,25
LSA =	375,75
Nº existente =	3

Pela regra de Sturges  $K = 1 + \log_2(n)$  o nº de intervalos de classe deve ser igual a 7.

A figura supra mostra a distribuição em frequência dos dados (absolutas e relativas, simples e acumuladas), a qual apresenta uma forma marcadamente assimétrica.

Uma vez os dados filtrados pelos limites de aceitação inferior e superior:

$$LIA = Q_1 - 1,5 \cdot (Q_3 - Q_1) = 34,5 - 1,5 \cdot (171 - 34,5) = -170,25 \text{ segundos}$$

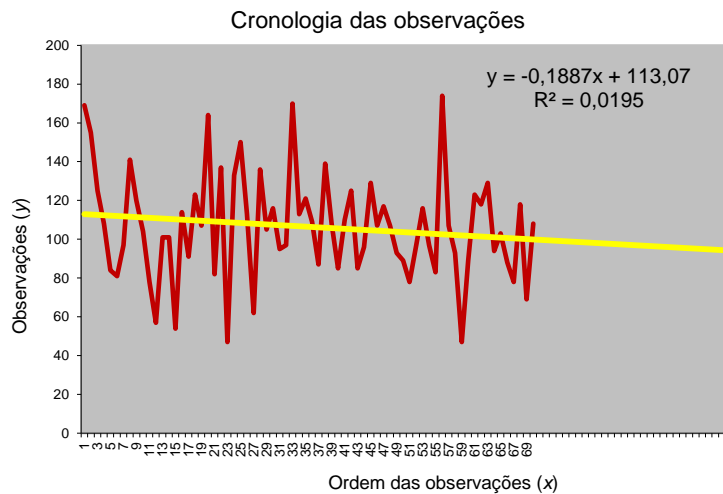
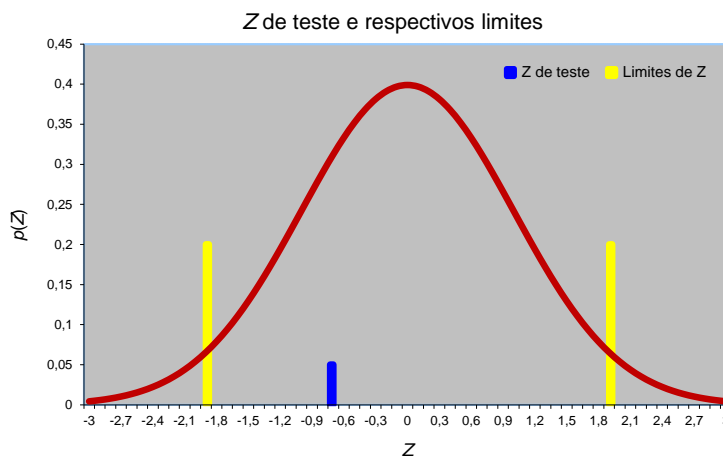
$$LSA = Q_3 + 1,5 \cdot (Q_3 - Q_1) = 171 + (171 - 34,5) = 375,75 \text{ segundos}$$

Obtemos os outliers (valores atípicos ou particularmente elevados ou baixos): 541, 536 e 556

2.  
PT-23

Pelo método das alternâncias:

Alternâncias			
Nº de observações (N) =	70	O.K.	
Nº de alternâncias (V) =	44		
Valor esperado ( $\mu_v$ ) =	46,33333333		
Desvio Padrão ( $\sigma_v$ ) =	3,481698181		
Estatística de teste ( $Z_t$ ) =	-0,670171052		
Nível de significância ( $\alpha$ ) =	5,00%	$Z_{(\alpha/2)}$ = -1,959963985	$Z_{(1-\alpha/2)}$ = 1,959963985
Conclusão:	As alternâncias são aleatórias		
Valor de prova =	50,27%		

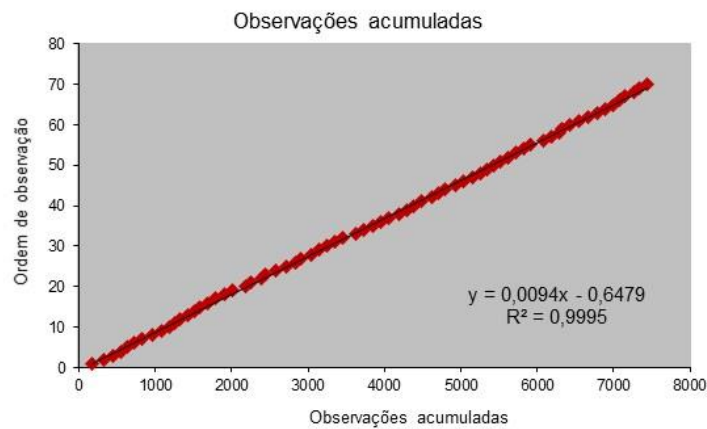
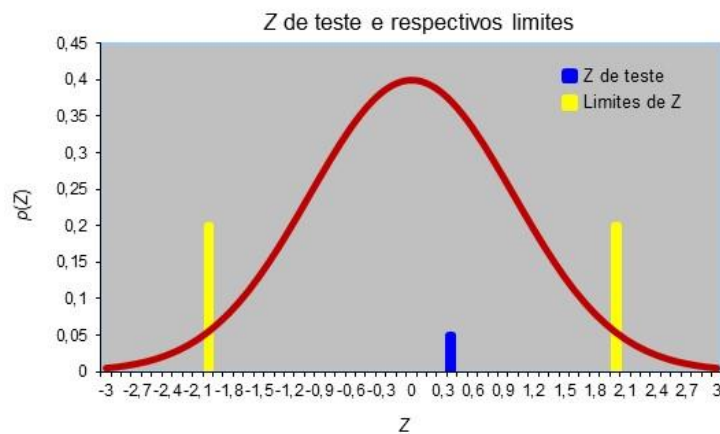


Sendo a hipótese nula  $H_0$ : “Os dados não apresentam ciclicidade” e tendo resultado do teste de aderência (ou de ajustamento) o facto da estatística de teste  $Z_t = -0,67$  encontrar-se dentro do intervalo de aceitação  $Z_{\alpha/2} = -1,96$  e  $Z_{1-\alpha/2} = 1,96$ , podemos concluir o seguinte:

“Não existe evidência estatística suficiente para rejeitar a hipótese nula  $H_0$ , pelo que podemos concluir que, na perspectiva das alternâncias, os dados são aleatórios, isto é, são i.i.d. (independentes e identicamente distribuídos)”.

Pelo método da **recta de ajustamento** (ou de tendência):

$N-1 =$	69	
$\sum_{1}^{N-1} X_j =$	262 280	
$X_N =$	7 446	
<b>Estadística de teste</b>		
Estadística de teste ( $Z_T$ ) =	0,302046262	$Z_{(\alpha/2)} = -1,959963985$
Nível de significância ( $\alpha$ ) =	5%	$Z_{(1-\alpha/2)} = 1,959963985$
Conclusão:	A taxa de acumulação é constante e igual a 0,009381	
Valor de prova =	76,26%	



Sendo a hipótese nula  $H_0$ : “Os dados não apresentam tendência” e tendo resultado do teste de aderência (ou de ajustamento) o facto da estatística de teste  $Z_t = 0,302$  encontrar-se dentro do intervalo de aceitação  $Z_{\alpha/2} = -1,96$  e  $Z_{1-\alpha/2} = 1,96$ , podemos concluir o seguinte:

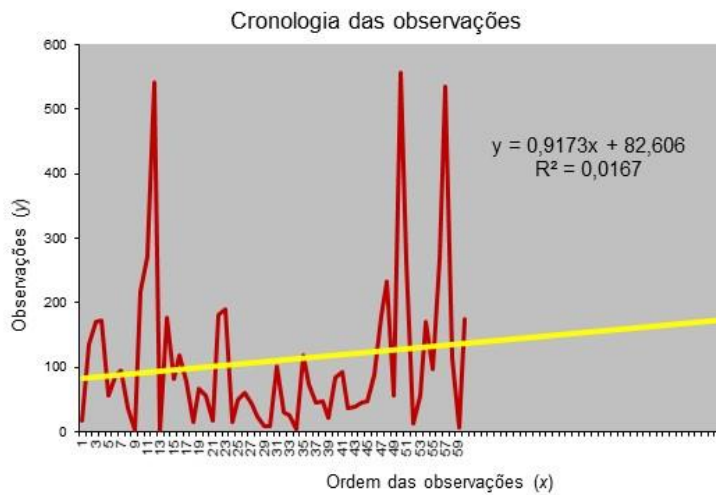
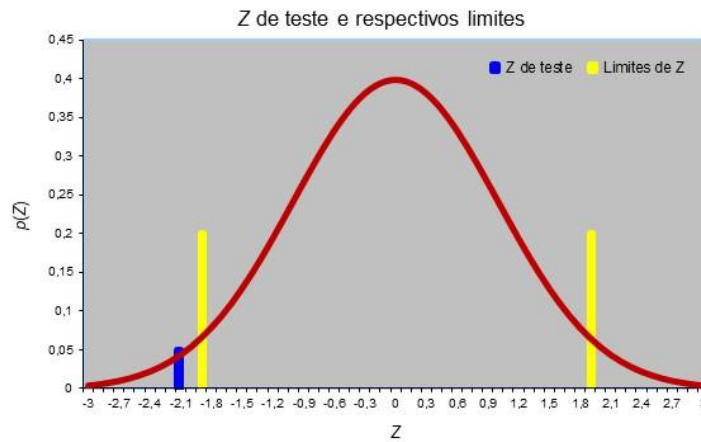
“Não existe evidência estatística suficiente para rejeitar a hipótese nula  $H_0$ , pelo que podemos concluir que, na perspectiva da tendência, os dados são aleatórios, isto é, são i.i.d. (independentes e identicamente distribuídos)”

Em conclusão: Como ambos os testes resultaram favoráveis à hipótese nula, os dados são de facto aleatórios, isto é, não se encontram correlacionados ou são i.i.d. (independentes e identicamente distribuídos)”.

**PT-37**

Pelo método das **alternâncias**:

Alternâncias			
Nº de observações (N) =	60	O.K.	
Nº de alternâncias (V) =	33		
Valor esperado ( $\mu_V$ ) =	39,66666667		
Desvio Padrão ( $\sigma_V$ ) =	3,216278042		
Estatística de teste ( $Z_t$ ) =	-2,072789286		
Nível de significância ( $\alpha$ ) =	5,00%	$Z_{(\alpha/2)}$	$Z_{(1-\alpha/2)}$
		-1,959963985	1,959963985
Conclusão:	As alternâncias não são aleatórias		
Valor de prova =	-		

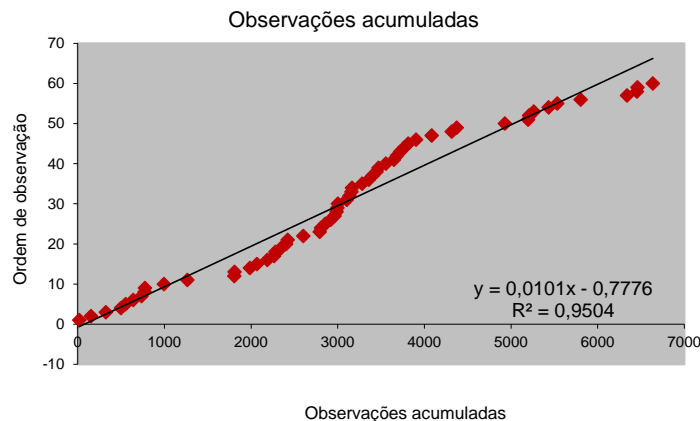
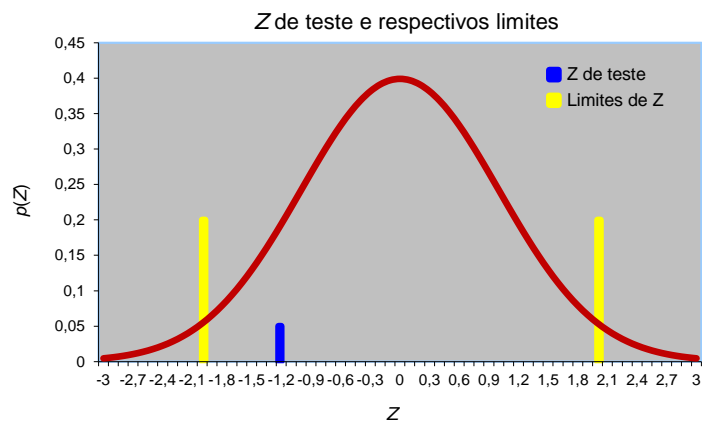


Sendo a hipótese nula  $H_0$ : “Os dados não apresentam ciclicidade” e tendo resultado do teste de aderência (ou de ajustamento) o facto da estatística de teste  $Z_t = -2,073$  encontrar-se fora do intervalo de aceitação  $Z_{\alpha/2} = -1,96$  e  $Z_{1-\alpha/2} = 1,96$ , podemos concluir o seguinte:

“Existe evidência estatística suficiente para rejeitar a hipótese nula  $H_0$ , pelo que podemos concluir que, na perspectiva das alternâncias, os dados não são aleatórios, isto é, não são i.i.d. (independentes e identicamente distribuídos)”.

Pelo método da **recta de ajustamento** (ou de tendência):

$N-1 =$	59		
$\sum_{i=1}^{N-1} X_i =$	179 226		
$X_N =$	6 635		
<b>Estatística de teste</b>			
Estatística de teste ( $Z_T$ ) =	-1,121963391	$Z_{(\alpha/2)} =$	-1,959963985
Nível de significância ( $\alpha$ ) =	5%	$Z_{(1-\alpha/2)} =$	1,959963985
Conclusão: A taxa de acumulação é constante e igual a 0,010097			
Valor de prova =	26,19%		



Sendo a hipótese nula  $H_0$ : “Os dados não apresentam tendência” e tendo resultado do teste de aderência (ou de ajustamento) o facto da estatística de teste  $Z_t = -1,122$  encontrar-se dentro do intervalo de aceitação  $Z_{\alpha/2} = -1,96$  e  $Z_{1-\alpha/2} = 1,96$ , podemos concluir o seguinte:

“Não existe evidência estatística suficiente para rejeitar a hipótese nula  $H_0$ , pelo que podemos concluir que, na perspectiva da tendência, os dados são aleatórios, isto é, são i.i.d. (independentes e identicamente distribuídos)”.

Todavia, como o método das alternâncias rejeitou a hipótese nula  $H_0$ , concluímos que os dados não são de facto aleatórios, isto é, não são i.i.d. (independentes e identicamente distribuídos)”. Encontram-se de facto correlacionados. Para o serem, teriam de ter passado ambos os testes.



**3.**  
**PT-23**

OBJECTIVOS								Seleccionar uma opção	
Margem tolerada de erro ( $\epsilon$ ) ≤	7	<input type="text" value="x"/>	População =	100000000					
Margem tolerada de erro ( $\epsilon\%$ ) ≤	3,00%	<input type="text" value="x"/>					Valor da média a considerar	107,746	
Nível de confiança ( $\delta$ ) =	95,00%								
Nível de significância ( $\alpha$ ) =	5,00%								

n	Observações	Média acumulada $\bar{x}$	Desvio padrão s	t de Student ou Z = N(0,1) t ou Z	Margem erro $\epsilon$	Margem erro $\epsilon\%$	Observações suficientes?	Observações adicionais necessárias
58	93	108,310	27,930	1,960	7,188	6,64%	Não	3
59	47	107,271	28,816	1,960	7,353	6,85%	Não	6
60	89	106,967	28,668	1,960	7,254	6,78%	Não	4
61	123	107,230	28,502	1,960	7,152	6,67%	Não	3
62	118	107,403	28,300	1,960	7,044	6,56%	Não	1
63	129	107,746	28,203	1,960	6,964	6,46%	Sim	

Depois de introduzirmos os dados na coluna B, constatamos que teriam bastado 63 observações.

O valor médio a considerar seria  $107,746 \cong 108$  segundos

**4.**  
**PT-37**

OBJECTIVOS								Seleccionar uma opção	
Margem tolerada de erro ( $\epsilon$ ) ≤	7	<input type="text" value="x"/>	População =	1000000000					
Margem tolerada de erro ( $\epsilon\%$ ) ≤	3,00%	<input type="text" value="x"/>					Valor da média a considerar	-	
Nível de confiança ( $\delta$ ) =	95,00%								
Nível de significância ( $\alpha$ ) =	5,00%								

n	Observações	Média acumulada $\bar{x}$	Desvio padrão s	t de Student ou Z = N(0,1) t ou Z	Margem erro $\epsilon$	Margem erro $\epsilon\%$	Observações suficientes?	Observações adicionais necessárias
55	98	100,582	112,151	1,960	29,639	29,47%	Não	931
56	270	103,607	113,409	1,960	29,703	28,67%	Não	952
57	536	111,193	126,143	1,960	32,747	29,45%	Não	1 190
58	115	111,259	125,032	1,960	32,178	28,92%	Não	1 168
59	7	109,492	124,691	1,960	31,817	29,06%	Não	1 160
60	175	110,583	123,919	1,960	31,355	28,35%	Não	1 144
61								
62								
63								

Depois de introduzirmos os dados na coluna B, constatamos na coluna I que seriam ainda necessárias aproximadamente 1.144 observações.

**5.**

Horário de observações			
1º período:	entre as:	9	e as: 12 horas
2º período:	entre as:	14	e as: 18 horas

Intervalo de tempo entre cada duas observações	
Mínimo =	10 minutos
Máximo =	30 minutos

	Observações	
	horas	minutos
1	9	18
2	9	41
3	9	52
4	10	13
5	10	30
6	10	46
7	11	15
8	11	39

14	14	12
15	14	28
16	14	41
17	14	54
18	15	6
19	15	30
20	15	49
21	16	18
22	16	29
23	16	48
24	17	6
25	17	33
26	17	46

O número de observações a realizar é aproximadamente igual a 20. Obviamente que obteremos novos valores cada vez que recalcularmos a aplicação.