O dono da empresa, após verificar o resultado da implementação desta medida decide analisar o que se passa numa outra empresa do qual é proprietário. Com o objetivo de averiguar a possibilidade de aplicação da mesma medida, foram registadas as vendas médias de 12 trabalhadores:

13.2; 11.6; 14.1; 14.3; 12.7; 12.8; 14.9; 13.7; 12.6; 15.5; 14.7; 14.6

Admita a normalidade das vendas.

Teste de hipóteses para a igualdade das vendas médias dos vendedores das duas empresas:

- *Analyze → Compare Means → Independent-Samples T Test…*
  - ☐ Test Variable(s): Vendas
  - ☐ Grouping Variables: Empresa
    - Define Groups…: ⊙ use specified values…
    - Group 1: 1; Group 2: 2
    - Continue
  - ☐ Options… Confidence Interval: 95 → Continue
  - ☐ OK

X_1 - vendas da empresa 1
X_2 - " " " 2

Group Statistics

	Empresa (1 ou 2)	N	Mean	Std. Deviation	Std. Error Mean
vendas médias dos vendedores	1	$n_1 =$ 10	$\bar{x}_1 =$ 14,510	$s_1 =$ 1,3110	$s_1/\sqrt{n_1}$,4146
	2	$n_2 =$ 12	$\bar{x}_2 =$ 13,725	$s_2 =$ 1,1553	$s_2/\sqrt{n_2}$,3335

$H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					$IC_{95\%}(\mu_1-\mu_2)$ 95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference " erro padrão da diferença	Lower	Upper
vendas médias dos vendedores	Equal variances assumed	,185	,672	1,493	20	,151	,7850	,5257	] -,3116	; 1,8816 [
	Equal variances not assumed			1,475	18,185	,157	,7850	,5321	-,3320	1,9020

$\bar{x}_1 - \bar{x}_2$ $\quad s^*$ $\quad \bar{x}_1 - \bar{x}_2 \pm \underbrace{t_{n_1+n_2-2;\,0.975}\, s^*}_{\varepsilon}$

$\to \dfrac{\bar{x}_1 - \bar{x}_2}{\text{erro padrão da diferença}}$

graus de liberdade $= n_1 + n_2 - 2$

$H_0: \sigma_1^2 = \sigma_2^2$
$H_1: \sigma_1^2 \neq \sigma_2^2$

p-value = 67,2%

α: 1%
α: 5% $\ngtr$ 67,2%
α: 10%

$\Rightarrow$ não rejeitamos H_0 para qq um dos α's usuais $\Rightarrow \sigma_1^2 = \sigma_2^2$

p-value = 0,151 $\Rightarrow$ n̄ rej. H_0 para qq um dos α's usuais $\Rightarrow$ podemos admitir $\mu_1 = \mu_2$

A procura diária de um certo produto foi, em 40 dias escolhidos ao acaso, a seguinte:

Número de unidades	Número de dias
0	6
1	14
2	10
3	7
4	2
5	1

Será que se pode admitir que tais observações foram extraídas de uma população com distribuição Poisson, isto é, será de admitir que a procura diária segue uma distribuição de Poisson?

X – procura diária do produto.

$H_0: X \sim P_\lambda$

$H_1: X \not\sim P_\lambda$

$\hat{\lambda} = \bar{x}$ $\quad P(X=x_i) = e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} \quad x_i: 0,1,2,\ldots$

- *Analyze → Nonparametric Test…*
  - Legacy Dialogs → Chi-Square…
    - Test variable list: Procura
    - Values:
      - 0.1827 Add — $P(X=0)$
      - 0.3106 Add — $P(X=1)$
      - 0.2640 Add — $P(X=2)$
      - 0.1496 Add — $P(X=3)$
      - 0.0636 Add — $P(X=4)$
      - 0.0295 Add — $P(X\geq 5)$

    p_i
- OK

x_i $\quad O_i \quad E_i = np_i \quad O_i - E_i$

procura diária

	Observed N	Expected N	Residual
0	6	7,3	-1,3
1	14	12,4	1,6
2	10	10,6	-,6
3	7	6,0	1,0
4	2	2,5	-,5
5	1	1,2	-,2
Total	40		

K=6 $\quad$ n

Test Statistics

	procura diária
Chi-Square	,780[a]
df	5 *
Asymp. Sig.	,978

a. 2 cells (33,3%) have expected frequencies less than 5. The minimum expected cell frequency is 1,2.

graus de liberdade

p-value

$$\chi^2_{obs} = \sum_{i=1}^{K} \frac{(O_i - E_i)^2}{E_i}$$

$$p\text{-value} = P[\chi^2 \geq \chi^2_{obs}]$$

Condições de aplicabilidade
- No máximo 20% de $E_i < 5$ X
- Todas $E_i > 1$ ✓

falham!

$$\chi^2 = \sum_{i=1}^{K} \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{K-p-1}$$

K = nº de categorias de X

p = nº de parâmetros estimados

* Apenas no caso em que temos que estimar parâmetros (estimámos λ ⇒ p=1) os graus de liberdade são os do SPSS menos 1 ⇒ df = 4 neste caso

- **Analyze → Nonparametric Test…**
  - Legacy Dialogs → Chi-Square…
    - Test variable list: Procura
    - Values: 0.1822 Add — $P(X=0)$
      0.3106 Add — $P(X=1)$
      0.2640 Add — $P(X=2)$
      0.1496 Add — $P(X=3)$
      0.0931 Add — $P(X\geq 4)$
- OK

Temos de agregar classes

Procura2

	Observed N	Expected N	Residual
0	6	7,3	-1,3
1	14	12,4	1,6
2	10	10,6	-,6
3	7	6,0	1,0
>=4	3	3,7	-,7
Total	40		

Test Statistics

	Procura2
Chi-Square	,777[a]
df	4
Asymp. Sig.	,942

a. 1 cells (20,0%) have expected frequencies less than 5. The minimum expected cell frequency is 3,7.

χ^2_{obs}

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{k-p-1} = \chi^2_3$$

$$p\text{-value} = P[\chi^2_3 \geq 0.777] = 1 - P[\chi^2_3 < 0.777]$$

tabela da distribuição χ^2 $\simeq 1 - 0,15 = 0.85$

As condições de aplicabilidade são verificadas.

$\alpha = 5\% \not> p\text{-value} = 0.85\ (85\%)$

$\Rightarrow$ Não rejeitamos Ho para $\alpha = 5\%$ $\Rightarrow$ Podemos admitir que a procura diária segue uma distribuição de Poisson

Foi conduzida uma experiência no âmbito da qual se procurou testar se existe alguma relação entre a qualidade da secagem de máquinas de lavar roupa de um certo tipo e a velocidade de rotação a que se eleva o tambor da roupa na fase de centrifugação. Os resultados desta experiência, efectuada com base no comportamento de 90 máquinas, estão apresentados na seguinte tabela:

		Qualidade da secagem			
		Medíocre	**Suficiente**	**Boa**	**Muito Boa**
Velocidade de Rotação (rpm)	**600**	12	8	7	3
	900	9	10	7	4
	1200	2	9	8	11

Teste, ao nível de significância de 5%, se as duas variáveis são ou não independentes.

- *Analyze → Descriptive Statistics → Crosstabs...*
  - Row(s): Velocidade de Rotação
  - Column(s): Qualidade da secagem
  - Statistics...
    - Chi-squared

      → Continue
  - Cells...
    - Observed
    - Expected

      → Continue
  - OK

X - velocidade de rotação

Y - qualidade da secagem

H_0: X e Y são independentes

H_1: X e Y não são independentes

Velocidade da rotação (RPM) * Qualidade da Secagem Crosstabulation

			Qualidade da Secagem				Total	
			Mediocre	Suficiente	Boa	Muito Boa		
Velocidade da rotação (RPM)	600	Count	$O_{11}=$ 12	$O_{12}=$ 8	$O_{13}=$ 7	$O_{14}=$ 3	30	$= O_{1\cdot}$
		Expected Count	E_{11} 7,7	E_{12} 9,0	E_{13} 7,3	$E_{14}=$ 6,0	30,0	
	900	Count	$O_{21}=$ 9	$O_{22}=$ 10	O_{23} 7	$O_{24}=$ 4	30	$= O_{2\cdot}$
		Expected Count	$E_{21}=$ 7,7	$E_{22}=$ 9,0	E_{23} 7,3	$E_{24}=$ 6,0	30,0	
	1200	Count	$O_{31}=$ 2	$O_{32}=$ 9	$O_{33}=$ 8	$O_{34}=$ 11	30	$= O_{3\cdot}$
		Expected Count	$E_{31}=$ 7,7	$E_{32}=$ 9,0	E_{33} 7,3	E_{34} 6,0	30,0	
Total		Count	$O_{\cdot 1}=$ 23	$O_{\cdot 2}=$ 27	$O_{\cdot 3}$ 22	$O_{\cdot 4}$ 18	90	$O_{\cdot\cdot} = n$
		Expected Count	23,0	27,0	22,0	18,0	90,0	

Chi-Square Tests

$$E_{ij} = \frac{O_{i\cdot} \times O_{\cdot j}}{n}$$

χ^2_{obs}

graus de liberdade $(L-1)(C-1)$

$$p\text{-value} = P\left[\chi^2_{(L-1)(C-1)} \geq \chi^2_{obs}\right]$$

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	13,516a	6	,036
Likelihood Ratio	14,509	6	,024
Linear-by-Linear Association	10,459	1	,001
N of Valid Cases	90		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 6,00.

L = nº de linhas =
= nº de categorias de X

C = nº de colunas =
= nº de categorias de Y

Condições de aplicabilidade:

- No máximo 20% de $E_{ij} < 5$ ✓
- Todos os $E_{ij} \geq 1$ ✓

$$\chi^2 = \sum_{i=1}^{L}\sum_{j=1}^{C} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi^2_{(L-1)(C-1)}$$

$\alpha = 5\% \geq p\text{-value} = 0{,}036$ (3,6%) $\Rightarrow$ Rejeitamos H_0 para $\alpha = 5\%$

$\Rightarrow$ Para $\alpha = 5\%$, existe evidência de que as variáveis não são independentes.