

MTH2302D - TD 11

Vincent Perreault

Ex. 1 - Énoncé

Deux types différents de machines de moulage par injection sont utilisés pour former des pièces en plastique. Une pièce est considérée comme défectueuse si elle est trop rétrécie ou si elle est décolorée. Deux échantillons aléatoires, chacun de taille 500, sont choisis ; 32 pièces défectueuses sont trouvées dans l'échantillon de la machine 1 et 21 pièces défectueuses dans l'échantillon de la machine 2. Est-il raisonnable de conclure que les deux machines fabriquent le même pourcentage de pièces défectueuses au seuil de 5 % ?

- Si $H_0 : p_1 = p_2$, $H_1 : p_1 \neq p_2$,
alors
$$Z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

où $\hat{p}_i = \frac{X_i}{n_i}$ et $\hat{p} = \frac{X_1 + X_2}{n_1 + n_2}$,
on rejette H_0 si $|Z_0| > z_{\alpha/2}$

Ex. 1 - Réponse

Les 2 machines ont-elles la même proportion de pièces défectueuses?

$$H_0 : p_1 = p_2, \quad H_1 : p_1 \neq p_2$$

$$Z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}, \text{ rejeter } H_0 \text{ si}$$
$$|Z_0| > z_{\alpha/2} = z_{0.05/2} = 1.96$$

$$Z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{\frac{32}{500} - \frac{21}{500}}{\sqrt{\frac{32+21}{2 \cdot 500} \left(1 - \frac{32+21}{2 \cdot 500} \right) \left(\frac{1}{500} + \frac{1}{500} \right)}}$$
$$\approx 1.5526 \leq 1.96$$

On ne rejette pas H_0 .

$$n_1 = n_2 = 500$$

$$X_1 = 32, \quad X_2 = 21$$

$$H_0 : p_1 = p_2, \quad H_1 : p_1 \neq p_2,$$

$$\Rightarrow \hat{p}_i = \frac{X_i}{n_i}, \quad \hat{p} = \frac{X_1 + X_2}{n_1 + n_2}$$

$$\Rightarrow Z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

$$\Rightarrow \text{rejeter } H_0 \text{ si } |Z_0| > z_{\alpha/2}$$

Ex. 2 - Énoncé

Soit le processus du moulage par injection décrit à l'exercice 1.

- a) Présentez ces données sous la forme d'un tableau de contingence (deux lignes et deux colonnes).
- b) On décide de réaliser un test du khi-carré à partir de ce tableau. Quelle sera l'hypothèse testée : l'homogénéité ou l'indépendance ? Énoncez la règle de rejet de H_0 au seuil de 5 %, calculez la valeur de la statistique observée et tirez vos conclusions.
- c) Cette procédure équivaut-elle à la procédure de test utilisée à l'exercice 1 ?

- Pour tester $H_0 : X_1 \sim X_2$ contre $H_1 : X_1 \not\sim X_2$, étant donné un tableau d'effectifs observés O_{ij} pour toutes variables X_i et valeurs V_j avec

$$\sum_{i=1}^r \sum_{j=1}^c O_{ij} = n,$$

alors $\chi_0^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$ où

$$E_{ij} = \frac{1}{n} \left(\sum_{k=1}^c O_{ik} \right) \left(\sum_{l=1}^r O_{lj} \right),$$

on rejette H_0 si $\chi_0^2 > \chi_{\alpha; (r-1)(c-1)}^2$

Ex. 2 - Réponse - a)

Présenter les données sous forme de table de contingence.

O_{ij}	non défect.	défect.	Total
X_1	468	32	500
X_2	479	21	500
Total	947	53	1000

X_1 : 32 défect. / 500

X_2 : 21 défect. / 500

$$\begin{aligned} H_0 : X_1 &\sim X_2, \quad H_1 : X_1 \not\sim X_2 \\ \Rightarrow E_{ij} &= \frac{1}{n} \left(\sum_{k=1}^c O_{ik} \right) \left(\sum_{l=1}^r O_{lj} \right) \\ \Rightarrow \chi_0^2 &= \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \\ \Rightarrow \text{rejeter } H_0 &\text{ si } \chi_0^2 > \chi_{\alpha; (r-1)(c-1)}^2 \end{aligned}$$

Ex. 2 - Réponse - b)

Quelle hypothèse de test χ^2 (homogénéité ou indépendance)? Déterminer la règle, la statistique et le résultat.

Homogénéité: $H_0 : X_1 \sim X_2, \quad H_1 : X_1 \not\sim X_2$

$$\chi_0^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \text{ où } E_{ij} = \frac{1}{n} \left(\sum_{k=1}^c O_{ik} \right) \left(\sum_{l=1}^r O_{lj} \right),$$

on rejette H_0 si $\chi_0^2 > \chi_{\alpha; (r-1)(c-1)}^2 = \chi_{0.05; 1 \cdot 1}^2 = 3.84$

E_{ij}	non défaut.	défect.	Total
X_1	473.5	26.5	500
X_2	473.5	26.5	500
Total	947	53	1000

$$\chi_0^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \approx 2.4108 \leq 3.84$$

On ne rejette pas l'hypothèse qu'ils sont homogènes.

X_1 : 32 défaut. / 500

X_2 : 21 défaut. / 500

$$\begin{aligned} H_0 : X_1 \sim X_2, \quad H_1 : X_1 \not\sim X_2 \\ \Rightarrow E_{ij} = \frac{1}{n} \left(\sum_{k=1}^c O_{ik} \right) \left(\sum_{l=1}^r O_{lj} \right) \\ \Rightarrow \chi_0^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \\ \Rightarrow \text{rejeter } H_0 \text{ si } \chi_0^2 > \chi_{\alpha; (r-1)(c-1)}^2 \end{aligned}$$

Ex. 2 - Réponse - c)

Cette procédure est-elle équivalente à l'exercice 1?.

Oui, U_0 est le carré du Z_0 de l'exercice 1.

X_1 : 32 défaut. / 500

X_2 : 21 défaut. / 500

$$\begin{aligned} H_0 : X_1 &\sim X_2, \quad H_1 : X_1 \not\sim X_2 \\ \Rightarrow E_{ij} &= \frac{1}{n} \left(\sum_{k=1}^c O_{ik} \right) \left(\sum_{l=1}^r O_{lj} \right) \\ \Rightarrow \chi_0^2 &= \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \\ \Rightarrow \text{rejeter } H_0 &\text{ si } \chi_0^2 > \chi_{\alpha; (r-1)(c-1)}^2 \end{aligned}$$

Ex. 3 - Énoncé

Nombre d'unités défectueuses par jour	Nombre d'observations O_i	Probabilité théorique \hat{p}_i	Fréquence espérée $\hat{E}_i = n\hat{p}_i$	$\frac{(O_i - \hat{E}_i)^2}{\hat{E}_i}$
0-10	6	0,054	6,27	0,012
11-15	11	0,092	10,76	0,005
16-20	16	0,163	19,06	0,493
21-25	28	0,214	24,99	0,362
26-30	22	0,207	24,25	0,209
31-35	19	?	?	?
36-40	11	?	?	?
41-45	4	0,043	4,97	0,191

- Pour tester $H_0 : X \sim \text{Loi}(\theta_1, \dots, \theta_p)$ contre $H_1 : X \not\sim \text{Loi}(\theta_1, \dots, \theta_p)$,

alors $\chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$ où $E_i = n \cdot P(X \in V_i | H_0)$

et on rejette H_0 si $\chi_0^2 > \chi_{\alpha; k-p-1}^2$

- Si $Y \sim N(\mu, \sigma^2)$, alors $F_Y(y) = \Phi\left(\frac{y-\mu}{\sigma}\right)$

On a compilé le nombre d'unités défectueuses trouvées chaque jour dans un procédé d'assemblage d'une carte de circuit imprimé. On veut savoir s'il est raisonnable de conclure que ces données viennent d'une distribution normale.

- En considérant que la moyenne échantillonnale est de 25 unités et que l'écart-type échantillonnal est de 9 unités, remplissez la colonne des fréquences espérées dans le tableau 10.8.
- On veut réaliser un test d'ajustement du khi-carré pour vérifier s'il est raisonnable de conclure que ces données viennent d'une distribution normale. Combien de degrés de liberté la statistique du test comptera-t-elle ?
- Quelle est la règle de rejet de H_0 , au seuil de 5 % ?
- Quelle est votre conclusion ?

Ex. 3 - Réponse - a)

Remplissez les fréquences espérées du tableau.

si $\bar{X} = 25$ et $S = 9$.

Soit $Y \sim N(25, 9^2)$

Nombre d'unités défectueuses par jour	Nombre d'observations O_i	Probabilité théorique \hat{p}_i	Fréquence espérée $\hat{E}_i = n\hat{p}_i$	$\frac{(O_i - \hat{E}_i)^2}{\hat{E}_i}$
0-10	6	0,054	6,27	0,012
11-15	11	0,092	10,76	0,005
16-20	16	0,163	19,06	0,493
21-25	28	0,214	24,99	0,362
26-30	22	0,207	24,25	0,209
31-35	19	?	?	?
36-40	11	?	?	?
41-45	4	0,043	4,97	0,191

$$\hat{p}_6 = P(Y \in [30.5, 35.5)) = F_Y(35.5) - F_Y(30.5)$$

$$= \Phi\left(\frac{35.5 - 25}{9}\right) - \Phi\left(\frac{30.5 - 25}{9}\right) = \Phi(1.1\bar{6}) - \Phi(0.6\bar{1})$$

$$= 0.87698 - 0.72907 = 0.14791$$

$$\hat{E}_6 = n \cdot \hat{p}_6 = 117 \cdot 0.14791 \approx 17.31$$

$$\hat{p}_7 = P(Y \in [35.5, 40.5)) = F_Y(40.5) - F_Y(35.5)$$

$$= \Phi\left(\frac{40.5 - 25}{9}\right) - \Phi\left(\frac{35.5 - 25}{9}\right) = \Phi(1.7\bar{2}) - \Phi(1.1\bar{6})$$

$$= 0.95728 - 0.87698 = 0.08030$$

$$\hat{E}_7 = n \cdot \hat{p}_7 = 117 \cdot 0.08030 \approx 9.400$$

$$H_0 : X \sim \text{Loi}(\theta_1, \dots, \theta_p),$$

$$H_1 : X \not\sim \text{Loi}(\theta_1, \dots, \theta_p)$$

$$\Rightarrow E_i = n \cdot P(X \in V_i | H_0 \text{ est vraie}),$$

$$\Rightarrow \chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

$$\Rightarrow \text{rejeter } H_0 \text{ si } \chi_0^2 > \chi_{\alpha; k-p-1}^2$$

$$Y \sim N(\mu, \sigma^2)$$

$$\Rightarrow F_Y(y) = \Phi\left(\frac{y-\mu}{\sigma}\right)$$

Ex. 3 - Réponse - b), c), d)

Combien de degrés de libertés la statistique χ_0^2 a-t-elle?

$$k - p - 1 = 8 - 2 - 1 = 5$$

Quelle est la règle de rejet de H_0 au seuil 5%?

$$\chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i},$$

on rejette H_0 si $\chi_0^2 > \chi_{\alpha; k-p-1}^2 = \chi_{0.05; 5}^2 = 11.07$

Quelle est la conclusion du test?

$$\chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \approx 1.709 \leq 11.07$$

On ne rejette pas H_0 . La loi normale pourrait être un bon modèle.

Nombre d'unités défectueuses par jour	Nombre d'observations O_i	Probabilité théorique p_i	Fréquence espérée $E_i = n\hat{p}_i$	$\frac{(O_i - E_i)^2}{E_i}$
0-10	6	0,054	6,27	0,012
11-15	11	0,092	10,76	0,005
16-20	16	0,163	19,06	0,493
21-25	28	0,214	24,99	0,362
26-30	22	0,207	24,25	0,209
31-35	19	?	?	?
36-40	11	?	?	?
41-45	4	0,043	4,97	0,191

$$H_0 : X \sim \text{Loi}(\theta_1, \dots, \theta_p),$$

$$H_1 : X \not\sim \text{Loi}(\theta_1, \dots, \theta_p)$$

$$\Rightarrow E_i = n \cdot P(X \in V_i | H_0 \text{ est vraie}),$$

$$\Rightarrow \chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

$$\Rightarrow \text{rejeter } H_0 \text{ si } \chi_0^2 > \chi_{\alpha; k-p-1}^2$$

$$Y \sim N(\mu, \sigma^2)$$

$$\Rightarrow F_Y(y) = \Phi\left(\frac{y-\mu}{\sigma}\right)$$

Ex. 4 - Énoncé

- Pour tester $H_0 : X$ et Y indép. contre $H_1 : X$ et Y non indép., étant donné un tableau d'effectifs observés O_{ij} pour tout x_i et y_j avec $\sum_{i=1}^r \sum_{j=1}^c O_{ij} = n$,

$$\text{alors } \chi_0^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \text{ où}$$

$$E_{ij} = \frac{1}{n} \left(\sum_{k=1}^c O_{ik} \right) \left(\sum_{l=1}^r O_{lj} \right)$$

on rejette H_0 si $\chi_0^2 > \chi_{\alpha; (r-1)(c-1)}^2$

- Si $Y \sim \chi_\nu^2$, alors $E(Y) = \nu$

10.33 Une entreprise exploite quatre machines avec trois équipes d'employés chaque jour. Les rapports de production présentent les données suivantes sur le nombre de pannes.

Équipe	Machines			
	A	B	C	D
1	41	20	12	16
2	31	11	9	14
3	15	17	16	10

On veut vérifier si les pannes sont indépendantes de l'équipe en place à l'aide d'un test du khi-carré.

- Quelles sont les hypothèses H_0 et H_1 ?
- Quelle est la règle de rejet de l'hypothèse nulle, si on utilise un seuil de 1 % ?
- Si l'hypothèse nulle est vraie, quelle est l'espérance (la moyenne théorique) de la statistique du test U_0 ?
- Sachant que la valeur observée de la statistique du test est de 11,649, que pouvez-vous conclure ?

Ex. 4 - Réponse - a), b)

Quelles sont les hypothèses?

H_0 : équipes et machines indép.,

H_1 : équipes et machines non indép.

Quelle est la règle de rejet de H_0 au seuil 1%?

$$\chi_0^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \text{ où}$$

$$E_{ij} = \frac{1}{n} \left(\sum_{k=1}^c O_{ik} \right) \left(\sum_{l=1}^r O_{lj} \right),$$

on rejette H_0 si $\chi_0^2 > \chi_{\alpha; (r-1)(c-1)}^2 = \chi_{0.01; 2.3}^2 = 16.81$

H_0 : X et Y indép.,

H_1 : X et Y non indép.

$$\Rightarrow E_{ij} = \frac{1}{n} \left(\sum_{k=1}^c O_{ik} \right) \left(\sum_{l=1}^r O_{lj} \right)$$

$$\Rightarrow \chi_0^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

\Rightarrow rejeter H_0 si $\chi_0^2 > \chi_{\alpha; (r-1)(c-1)}^2$

$$Y \sim \chi_\nu^2 \Rightarrow E(Y) = \nu$$

Ex. 4 - Réponse - c), d)

Si H_0 est vraie, quelle est $E(\chi_0^2)$?

$$E(\chi_0^2) = (r - 1)(c - 1) = 2 \cdot 3 = 6$$

Sachant que la valeur observée de χ_0^2 est 11.649, quelle est la conclusion du test?

$$11.649 \leq 16.81$$

On ne rejette pas H_0 .

$H_0 : X$ et Y indép.,

$H_1 : X$ et Y non indép.

$$\Rightarrow E_{ij} = \frac{1}{n} \left(\sum_{k=1}^c O_{ik} \right) \left(\sum_{l=1}^r O_{lj} \right)$$

$$\Rightarrow \chi_0^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

\Rightarrow rejeter H_0 si $\chi_0^2 > \chi_{\alpha; (r-1)(c-1)}^2$

$$Y \sim \chi_\nu^2 \Rightarrow E(Y) = \nu$$

Ex. 5 - Énoncé

Exercice n° 5 Un générateur de nombres est utilisé pour générer des nombres entiers compris entre 1 et 5. Les 1000 premiers nombres générés présentent les fréquences (effectifs) d'apparition suivantes :

Nombre i	1	2	3	4	5
Fréquence f_i	220	190	205	180	205

Peut-on affirmer au seuil 5% que les nombres sont générés de manière aléatoire ?
Répondre à l'aide d'un test en précisant les hypothèses H_0 et H_1 .

► Pour tester $H_0 : X \sim \text{Loi}(\theta_1, \dots, \theta_p)$ contre $H_1 : X \not\sim \text{Loi}(\theta_1, \dots, \theta_p)$, étant donné

Valeurs (x_i)	V_1	V_2	...	V_k	Total
Effectifs observés (O_i)	O_1	O_2	...	O_k	n
Effectifs attendus (E_i)	E_1	E_2	...	E_k	n

où $E_i = n \cdot P(X \in V_i | H_0 \text{ est vraie})$,

alors $\chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$ et on rejette H_0 si $\chi_0^2 > \chi_{\alpha; k-p-1}^2$

Ex. 5 - Réponse

Les nombres sont-ils générés de manière aléatoire?

$H_0 : X \sim \text{loi uniforme discrète sur } \{1, 2, 3, 4, 5\},$

$H_1 : X \not\sim \text{loi uniforme discrète sur } \{1, 2, 3, 4, 5\}$

$$\chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \text{ où } E_i = n \cdot P(X \in V_i | H_0),$$

on rejette H_0 si $\chi_0^2 > \chi_{\alpha; k-p-1}^2 = \chi_{0.05; 5-0-1}^2 = 9.49$

$$E_i = n \cdot P(X \in V_i | H_0) = 1000 \cdot \frac{1}{5} = 200$$

$$\chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

$$= \frac{(220-200)^2}{200} + \frac{(190-200)^2}{200} + \frac{(205-200)^2}{200} + \frac{(180-200)^2}{200} + \frac{(205-200)^2}{200}$$

$$= 4.75 \leq 9.49$$

Oui, on ne rejette pas cette hypothèse.

Nb i	1	2	3	4	5	Total
Freq f_i	220	190	205	180	205	1000

$H_0 : X \sim \text{Loi}(\theta_1, \dots, \theta_p),$

$H_1 : X \not\sim \text{Loi}(\theta_1, \dots, \theta_p)$

$\Rightarrow E_i = n \cdot P(X \in V_i | H_0 \text{ est vraie}),$

$$\Rightarrow \chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

$\Rightarrow \text{rejeter } H_0 \text{ si } \chi_0^2 > \chi_{\alpha; k-p-1}^2$

Ex. 6 - Énoncé

Exercice n°6 Le tableau suivant présente la distribution des fréquences de la durée de service (en minute) au comptoir d'une entreprise.

Durée de service	[0, 1)	[1, 2)	[2, 3)	[3, 4)	≥ 4
Fréquence	53	36	28	12	21

La moyenne observée dans cet échantillon est de 2,5 minutes.

Ces données permettent-elles de supporter l'hypothèse selon laquelle la durée de service est distribuée selon une loi exponentielle? Effectuer le test approprié au seuil critique 5%.

► Pour tester $H_0 : X \sim \text{Loi}(\theta_1, \dots, \theta_p)$ contre $H_1 : X \not\sim \text{Loi}(\theta_1, \dots, \theta_p)$, étant donné

Valeurs (x_i)	V_1	V_2	...	V_k	Total
Effectifs observés (O_i)	O_1	O_2	...	O_k	n
Effectifs attendus (E_i)	E_1	E_2	...	E_k	n

où $E_i = n \cdot P(X \in V_i | H_0 \text{ est vraie})$,

alors $\chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$ et on rejette H_0 si $\chi_0^2 > \chi_{\alpha; k-p-1}^2$

► Si $Y \sim \text{Exp}(\lambda)$, alors $E(Y) = \frac{1}{\lambda}$ et $F_Y(y) = 1 - e^{-\lambda y}$

Ex. 6 - Réponse

La durée de service suit-elle une loi exponentielle?

$$H_0 : X \sim \text{Exp}(\lambda), \quad H_1 : X \not\sim \text{Exp}(\lambda)$$

$$\chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \text{ où } E_i = n \cdot P(X \in V_i | H_0),$$

on rejette H_0 si $\chi_0^2 > \chi_{\alpha; k-p-1}^2 = \chi_{0.05; 5-1-1}^2 = 7.81$

$$\text{Si } H_0 \text{ vraie, } E(X) = \frac{1}{\lambda} \Leftrightarrow \hat{\lambda} = \frac{1}{\bar{X}} = \frac{1}{2.5} = 0.4$$

$$\begin{aligned} E_1 &= n \cdot P(X \in [0, 1] | H_0) = n \cdot (F_X(1) - F_X(0)) \\ &= n \cdot (1 - e^{-\lambda \cdot 1} - (1 - e^{-\lambda \cdot 0})) = n \cdot (e^{-\lambda \cdot 0} - e^{-\lambda \cdot 1}) \\ &= 150 \cdot (e^{-0.4 \cdot 0} - e^{-0.4 \cdot 1}) \approx 49.46 \end{aligned}$$

$$E_2 = 150 \cdot (e^{-0.4 \cdot 1} - e^{-0.4 \cdot 2}) \approx 33.15$$

$$E_3 = 150 \cdot (e^{-0.4 \cdot 2} - e^{-0.4 \cdot 3}) \approx 22.22$$

$$E_4 = 150 \cdot (e^{-0.4 \cdot 3} - e^{-0.4 \cdot 4}) \approx 14.89$$

Durée	[0,1)	[1,2)	[2,3)	[3,4)	≥ 4	Total
Freq	53	36	28	12	21	150

moyenne: $\bar{X} = 2.5$

$$H_0 : X \sim \text{Loi}(\theta_1, \dots, \theta_p),$$

$$H_1 : X \not\sim \text{Loi}(\theta_1, \dots, \theta_p)$$

$$\Rightarrow E_i = n \cdot P(X \in V_i | H_0 \text{ est vraie}),$$

$$\Rightarrow \chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

$$\Rightarrow \text{rejeter } H_0 \text{ si } \chi_0^2 > \chi_{\alpha; k-p-1}^2$$

$$Y \sim \text{Exp}(\lambda)$$

$$\Rightarrow E(Y) = \frac{1}{\lambda}$$

$$\Rightarrow F_Y(y) = 1 - e^{-\lambda y}$$

Ex. 6 - Réponse

$$\begin{aligned}
 E_5 &= n \cdot P(X \geq 4 | H_0) = n \cdot (1 - F_X(4)) \\
 &= n \cdot (1 - (1 - e^{-\lambda^4})) = n \cdot e^{-\lambda^4} \\
 &= 150 \cdot e^{-0.4 \cdot 4} \approx 30.28
 \end{aligned}$$

Durée	[0,1)	[1,2)	[2,3)	[3,4)	≥ 4	Total
Freq	53	36	28	12	21	150

moyenne: $\bar{X} = 2.5$

Durée	[0,1)	[1,2)	[2,3)	[3,4)	≥ 4	Total
O_i	53	36	28	12	21	150
E_i	49.46	33.15	22.22	14.89	30.28	150

$$\begin{aligned}
 \chi_0^2 &= \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \\
 &= \frac{(53-49.46)^2}{49.46} + \frac{(36-33.15)^2}{33.15} + \frac{(28-22.22)^2}{22.22} + \frac{(12-14.89)^2}{14.89} + \frac{(21-30.28)^2}{30.28} \\
 &\approx 5.41 \leq 7.81
 \end{aligned}$$

On ne rejette pas H_0 . Donc, oui, les données supportent l'hypothèse d'une loi exponentielle.

$$H_0 : X \sim \text{Loi}(\theta_1, \dots, \theta_p),$$

$$H_1 : X \not\sim \text{Loi}(\theta_1, \dots, \theta_p)$$

$$\Rightarrow E_i = n \cdot P(X \in V_i | H_0 \text{ est vraie}),$$

$$\Rightarrow \chi_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

$$\Rightarrow \text{rejeter } H_0 \text{ si } \chi_0^2 > \chi_{\alpha; k-p-1}^2$$

$$Y \sim \text{Exp}(\lambda)$$

$$\Rightarrow E(Y) = \frac{1}{\lambda}$$

$$\Rightarrow F_Y(y) = 1 - e^{-\lambda y}$$

Ex. 7 - Énoncé

Exercice n° 7 On a effectué un relevé du nombre de contrats octroyés par 3 municipalités voisines (M_1, M_2, M_3) à un ensemble de 4 firmes de construction du secteur (C_1, C_2, C_3, C_4) au cours des 20 dernières années. Le relevé a donné :

	M_1	M_2	M_3	Total
C_1	30	20	10	60
C_2	20	30	10	60
C_3	20	10	30	60
C_4	10	20	30	60
Total	80	80	80	240

A-t-on évidence d'un lien significatif dans l'octroi des contrats entre municipalités et firmes de construction basé sur ces données ?

Effectuer le test requis, pour vous en assurer, au niveau critique 5%.

► Pour tester $H_0 : X$ et Y indép. contre $H_1 : X$ et Y non indép., étant donné un tableau d'effectifs observés O_{ij} pour tout x_i et y_j avec $\sum_{i=1}^r \sum_{j=1}^c O_{ij} = n$,

$$\text{alors } \chi_0^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \text{ où } E_{ij} = \frac{1}{n} \left(\sum_{k=1}^c O_{ik} \right) \left(\sum_{l=1}^r O_{lj} \right)$$

on rejette H_0 si $\chi_0^2 > \chi_{\alpha; (r-1)(c-1)}^2$

Ex. 7 - Réponse

Y a-t-il un lien significatif entre les C_i et les M_j ?

$H_0 : X$ et Y indép., $H_1 : X$ et Y non indép.

$$\chi_0^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \text{ où } E_{ij} = \frac{1}{n} \left(\sum_{k=1}^c O_{ik} \right) \left(\sum_{l=1}^r O_{lj} \right),$$

on rejette H_0 si $\chi_0^2 > \chi_{\alpha; (r-1)(c-1)}^2 = \chi_{0.05; 3 \cdot 2}^2 = 12.59$

	M_1	M_2	M_3	Total
C_1	30	20	10	60
C_2	20	30	10	60
C_3	20	10	30	60
C_4	10	20	30	60
Total	80	80	80	240

$H_0 : X$ et Y indép.,

$H_1 : X$ et Y non indép.

$$\Rightarrow E_{ij} = \frac{1}{n} \left(\sum_{k=1}^c O_{ik} \right) \left(\sum_{l=1}^r O_{lj} \right)$$

$$\Rightarrow \chi_0^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

\Rightarrow rejeter H_0 si $\chi_0^2 > \chi_{\alpha; (r-1)(c-1)}^2$

$$E_{ij} = \frac{1}{n} \left(\sum_{k=1}^c O_{ik} \right) \left(\sum_{l=1}^r O_{lj} \right) = \frac{1}{240} \cdot 60 \cdot 80 = 20$$

$$\begin{aligned} \chi_0^2 &= \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \\ &= \frac{(30-20)^2}{20} + \frac{(20-20)^2}{20} + \frac{(10-20)^2}{20} + \frac{(20-20)^2}{20} + \dots \\ &= 40 > 12.59 \end{aligned}$$

Oui, on rejette l'hypothèse qu'ils sont indépendants.