I) Un radar enregistre la vitesse des véhicules sur une route. Déterminer le nombre de données nécessaires afin d'obtenir une précision à ±1 km/h à un niveau de confiance de 97.5 % si l'écart-type est de 17 km/h.

$$n = \frac{k^2\sigma^2}{e^2} = \frac{2.24^2 17^2}{1^2} = 1450\ véhicules$$

J) La répartition modale des déplacements annuels au Canada est donnée dans le tableau suivant :

Mode de transport	**Répartition**
Automobile	81.0%
Transport en commun	10.4%
Bicyclette	1.3%
Motocyclette	0.1%
Taxi	0.2%
À pied	7.0%

On analyse 11 394 déplacements et on obtient la répartition suivante :

Mode de transport	**Nombre de déplacements**
Automobile (conducteur ou passager)	9021
Transport en commun	1367
Bicyclette	141
Motocyclette	7
Taxi	29
À pied	829

L'échantillon est-il représentatif au niveau de confiance de 95 % ? Qu'est-ce qui pourrait expliquer la différence s'il y en a ?

Le tableau suivant montre les déplacements attendus (E_i) et les déplacements observés (O_i) en fonction des différents modes.

ATTENDUS	OBSERVÉS
9229.14	9021
1184.976	1367
148.122	141
11.394	7
22.788	29
797.58	829

On calcule la variable de décision du test d'ajustement du khi-2 :

$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} = 37.62$$

qu'on compare avec la valeur de

$$\chi^2{}_{6-1}(0.05) = 11.07.$$

La valeur de la variable de décision est trop élevée pour que la distribution suive la répartition théorique à un niveau de confiance de 95 %.

La taille de l'échantillon peut être insuffisante : on pourrait prendre un plus grand échantillon de déplacements et refaire le test d'adéquation pour cette 2e série de données.

K) Le nombre d'accidents sur une route est comptabilisé dans le tableau suivant :

Jour	Nombre d'accidents
Jour 1	1
Jour 2	0
Jour 3	2
Jour 4	1
Jour 5	0
Jour 6	0
Jour 7	2
Jour 8	1
Jour 9	0
Jour 10	0
Jour 11	1
Jour 12	1
Jour 13	0
Jour 14	1
Jour 15	0

a) Calculer le nombre moyen d'accidents par jour

$$\bar{X} = 0.667\ accidents$$

b) Calculer un intervalle de confiance à 99 % sur cette moyenne si l'écart-type est connu et vaut 0.7

L'expression $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ *suit la loi normale centrée réduite, d'où on déduit l'intervalle de confiance suivant*

$$\mu = 0.667 \pm 2.576 \cdot \frac{0.7}{\sqrt{15}}$$

$$\mu \in [0.201\ ;\ 1.132]$$

c) Calculer le nombre de jours d'observation nécessaires afin d'obtenir le nombre d'accidents moyen à ± 0.1 pour un niveau de confiance de 99%.

$$n = \frac{2.576^2 * 0.7^2}{0.1^2} = 325.15$$

On aurait donc besoin de 326 jours d'observation.

L) L'enquête Origine-Destination de 2008[1] a révélé que les 2 213 000 déplacements AM en fonction du motif étaient distribués comme suit :

[1] Source : http://www.enquete-od.qc.ca/docs/EnqOD08_FaitsSaillants.pdf

Motif	**% déplacements**
TRAVAIL	51%
ÉTUDES	29%
AUTRES	16%
RETOUR	4%

Lors d'une seule journée de 2009, on enquête sur les motifs des déplacements du matin et on recueille les données suivantes :

Motif	**Nb déplacements**
TRAVAIL	4012
ÉTUDES	2319
AUTRES	1306
RETOUR	376

Cette distribution concorde-t-elle avec les résultats de l'enquête Origine-Destination ?

Il s'agit de déterminer si l'échantillon des motifs de déplacements suit la distribution de l'enquête O-D (hypothèse H_0) à l'aide du test du khi-2. On compare les effectifs attendus (E_i) et les effectifs observés (O_i) :

Motif	Oi	Ei	$\frac{(O_i - E_i)^2}{E_i}$
TRAVAIL	4012	4087	1.36
ÉTUDES	2319	2324	0.01
AUTRES	1306	1282	0.45
RETOUR	376	320	9.60
			11.42

La variable de décision

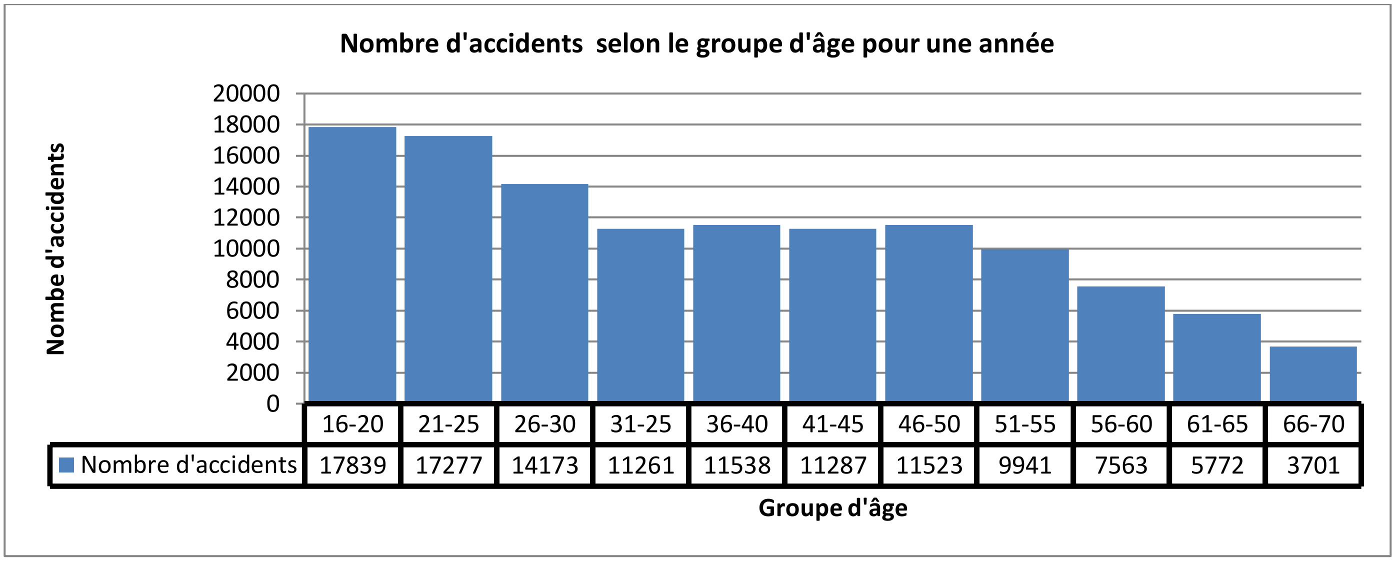
$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} = 11.42$$

est supérieure à la valeur de la loi du khi-2 pour un risque de première espèce de 0.05

$$\chi^2{}_{4-1}(0.05) = 7.81$$

On peut donc rejeter l'hypothèse nulle que l'échantillon observé sur une seule journée suit la distribution des déplacements par motif déterminée lors de l'enquête OD : il s'agit d'une journée différente.

M) Le nombre d'accidents selon le groupe d'âge[2] est donné dans le graphique suivant :



	16-20	21-25	26-30	31-25	36-40	41-45	46-50	51-55	56-60	61-65	66-70
Nombre d'accidents	17839	17277	14173	11261	11538	11287	11523	9941	7563	5772	3701

Soit X l'âge d'un conducteur impliqué dans un accident. Déterminer si cette variable aléatoire suit une loi exponentielle en utilisant un niveau de confiance de 95 %.

La moyenne de la distribution est calculée avec les valeurs centrales des intervalles (18, 23, 28, 33, 38, 43, 48, 53, 58, 63 et 68) :

$$\bar{X} = \frac{18 * 17839 + 23 * 17277 + 28 * 14173 + 33 * 11261 + \cdots}{17839 + 17277 + 14173 + 11261 + \cdots} = 37.29\ ans$$

On calcule par la suite les fréquences théoriques de chacun des intervalles avec la loi exponentielle :

$$P(16 \le X \le 20) = F(X = 20) - F(X = 16)$$
$$P(16 \le X \le 20) = \left(1 - e^{\frac{-20}{37.29}}\right) - \left(1 - e^{\frac{-16}{37.29}}\right) = 0.0662$$

On effectue un test d'ajustement du khi-2 avec les hypothèses suivantes :

$H_0 : X \sim Exp\left(\lambda = \frac{1}{37.29}\right)$

$H_1 : X \nsim Exp\left(\lambda = \frac{1}{37.29}\right)$

On compare par la suite les valeurs observées O_i aux valeurs attendues E_i à l'aide de la variable de décision du test d'ajustement du khi-2:

$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} = 111732.7$$

[2] Source : Missouri Department of Transportation; Transportation Planning; Crash Statistics by Driver Age for Year 2008 http://www.modot.mo.gov/safety/documents/2008AccidentStatisticsManual.pdf

La valeur du khi-2 à 11-p-1 degrés de liberté au seuil α=5 % est de ${\chi^2}_{11-1-1}(0.05) = 16.92$*. Puisque* $X^2 \gg {\chi^2}_9(0.05)$*, on peut rejeter l'hypothèse* H_0 *et affirmer que cette distribution ne suit pas à la loi exponentielle.*

N) Des observations ont montré qu'en 10 h, une moyenne de 67 bicyclettes se retrouvait dans la circulation d'une certaine rue. Si le nombre de vélos X passant sur cette rue suit une loi de Poisson, calculer la probabilité qu'on dénombre :

a) 54 vélos en 10 heures sur cette même rue

La valeur du paramètre λ d'une loi de Poisson est égale à la valeur de la moyenne. On calcule la probabilité avec

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

$$P(X = 54) = \frac{67^{54} e^{-67}}{54!} = 0.014$$

b) Si la moyenne du nombre de bicyclettes était de 50 par 10 heures, que deviendrait la probabilité calculée en a) ?

$$P(X = 54) = \frac{50^{54} e^{-50}}{54!} = 0.0464$$

O) Une enquête de *Statistiques CANADA*[3] révèle l'utilisation du transport en commun pour le motif travail en 1996 et en 2001

Groupes d'âges	1996 %	2001 %
15 à 19	12,7	14,5
20 à 24	14,9	16,6
25 à 34	11,0	11,9
35 à 44	8,8	8,8
45 à 54	8,4	8,5
55 à 64	9,1	8,3
65 et plus	8,9	7,7

Y a-t-il eu un changement significatif entre les deux années au niveau de confiance de 95 % ?

Posons les hypothèses suivantes :

H_0 *: les distributions sont semblables*
H_1 *: les distributions ne sont pas semblables*

[3] (Source : *Où travaillent les Canadiens et comment s'y rendent-ils?* http://www12.statcan.gc.ca/francais/census01/products/analytic/companion/pow/publictrans_f.cfm

Pour comparer les deux distributions, il s'agit de calculer la variable de décision du test d'ajustement du khi-2, en prenant les observations de l'année 1996 comme référence

$$X^2 = \sum_{i=1}^{n} \frac{(O_{2001} - 0_{1996})^2}{O_{1996}} = 0.756$$

et de la comparer avec

$$\chi^2{}_{7-1}(0.05) = 12.59$$

La valeur de la variable de décision est inférieure à la valeur de $\chi^2{}_6(0.05)$*. On ne peut pas rejeter l'hypothèse* H_0 *: les deux échantillons suivent la même loi au niveau de confiance de 95 %.*

P) Les débits journaliers entre 6h et 18h dans une direction d'une autoroute sont classés dans le tableau suivant :

Heure	Débit (véh)
6h	1519
7h	2267
8h	2036
9h	1902
10h	2025
11h	2197
12h	2371
13h	2669
14h	3617
15h	5486
16h	5843
17h	6024
18h	5012

a) Calculer le débit moyen durant cette période de la journée.

$$\bar{X} = 3305.23\ véhicules$$

b) Calculer un intervalle de confiance à 95 % sur la moyenne si l'écart type vaut 1673 véhicules.

$$\mu = 3305.23 \pm 1.96 \cdot \frac{1673}{\sqrt{13}}$$

$$\mu \in [2395.8\,;4214.7\,]$$