ANNALES CORRIGÉES

est une collection dirigée par Alexandre MÉLISSOPOULOS.
Elle a pour objectif de mettre au service des élèves
l'expérience de MATHS MÉLISSO,
organisme spécialisé dans le soutien scolaire depuis 20 ans.

Épreuves corrigées des sessions de juin et de remplacement de septembre

Fiches de cours de mathématiques

ANNALES BTSA

MATHÉMATIQUES

[Traitement de données]

Volodia LECUYER
Enseignant en Mathématiques et Physique-Chimie
Enseigne depuis plus de 20 ans dans le secondaire
et le supérieur, notament en BTSA GPN.

16

Diverses propriétés avec $t \geq 0$:

- $P(T \leq -t) = P(T \geq t)$ par symétrie de la loi normale centrée réduite

Donc $P(T \leq -t) = 1 - P(T \leq t)$

- $P(-t \leq T \leq -t) = 2P(T \leq t) - 1$

t au lieu de -t

On introduit la variable aléatoire fréquence $F = \frac{X}{n}$ pour laquelle X est la variable aléatoire associée à la valeur observée x.

Il est clair que $X \hookrightarrow B\,(n\,;p)$ avec $E(X) = np$ et $\sigma(X) = \sqrt{np(1-p)}$

Or par définition, $E(aX) = a\,E(X)$ pour tout réel a et $\sigma(aX) = a\,\sigma(X)$ pour tout réel a positif

Ici $a = \frac{1}{n}$ donc $E(F) = \frac{1}{n}\,E(X) = \frac{1}{n}\,np$ d'où $\mathbf{E(F) = p}$

et $\sigma\,(F) = \frac{1}{n}\,\sigma\,(X) = \frac{1}{n}\,\sqrt{np(1-p)}$ d'où $\cancel{(F)} = \boxed{\sigma(F) = \sqrt{\frac{p(1-p)}{n}}}$

Alors, pour $n \geq 30$, $np \geq 5$ et $nq \geq 5$, on a $F \hookrightarrow N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$

Quand on vous demande par quelle loi on peut approcher la loi de F, vérifiez que $n \geq 30$ et rajoutez bien ces 2 conditions surlignées qui n'apparaissent pas dans mes corrections et qui sont indispensables .
$np \geq 5$ et $n(1 - p) \geq 5$

Si n < 30 (cas des petits échantillons)

- Si σ est connu, la théorie est la même que pour les grands échantillons.

- Si σ est inconnu. Soit T la variable centrée réduite correspondante à $\overline{x}$,
  Maintenant $T \hookrightarrow T_{n-1}$ (Loi de Student à n-1 degrés de liberté)

Concrètement, la théorie est la même qu'avant, mais il faut remplacer 1,96 et 2,58 par des valeurs lues dans la table de Student.

Ligne : $k = n - 1$ et Colonne : $p = 0{,}975$ pour 95% de confiance
$p = 0{,}995$ pour 99% de confiance

Estimation de la proportion p par intervalle de confiance à 95 % ou 99 %

Par définition, l'intervalle de confiance de p est : $\left[f - t\sqrt{\frac{f(1-f)}{n}} \; ; \; f + t\sqrt{\frac{f(1-f)}{n}}\right]$

t est défini tel que $P(-t \leq T \leq t) = 0{,}95$ (ou 0,99) avec $T \hookrightarrow N(0\,;1)$.

Ce qui équivaut à $P(T \leq t) = 0{,}975$ (ou 0,995).

Par lecture inverse de la table de la loi normale centrée réduite, on obtient t = **1,96 ou 2,58**.

Pensez à rajouter ces conditions pour justifier l'utilisation de cet intervalle de confiance :
" En effet, on a bien : $n \geq 30$ $nf \geq 5$ et $n(1 - f) \geq 5$ "

T2	92,45	107,50	230,05

Par exemple, on a : $t_1 = \dfrac{570 \times 215}{1000} = 122{,}55$

122,55 est l'effectif théorique de 125.

3. Calcul du Khi² observé : χ^2_{obs}

$\boxed{n = 1000}$

Par définition, on a : $\chi^2_{obs} = \sum_{i=1}^{j} \dfrac{(n_i - t_i)^2}{t_i}$, avec ici $j = 6$

Avec ni les effectifs observés et ti les effectifs théoriques.

Notons que l'on a bien $n \geq 50$ puisque $n =$ ~~200~~ et que les $t_i \geq 5$, conditions pour pouvoir appliquer le test.

54

3. $y = 0{,}802\,t - 0{,}701$

$\ln x = 0{,}802\,t - 0{,}701$

$\cancel{y} =$ **$e^{0{,}802\,t - 0{,}701}$**

$x =$

$x =$

4. 5h30 correspond à $t = 5{,}5$ h, alors $\cancel{y} = e^{0{,}802 \times 5{,}5 - 0{,}701}$

D'où x $\cancel{y}$ **$= 40{,}85$** à 10^{-1} près

Notons que l'on a bien $n \geq 50$ puisque $n =$ ~~200~~ et que les $t_i \geq 5$, conditions pour pouvoir appliquer le test.

On obtient alors $\chi^2_{obs} = \dfrac{(22-23{,}00)^2}{23{,}00} + \dfrac{(152-133{,}33)^2}{133{,}33} + \ldots$

D'où $\chi^2_{obs} = 37{,}45$

$\boxed{n = 600}$

On peut également répertorier les termes de la somme donnant le χ^2_{obs} dans un tableau pour avoir à taper une somme plus agréable à la calculatrice :

Intensité du défaut / Gamme de bouchon	importante	faible	nulle
A	0,043	2,613	7,148
B	6,261	0,213	1,018
C	5,261	1,333	13,560

$$\Leftrightarrow a\frac{\sqrt{n}}{\sigma} = 1{,}96 \text{ par lecture inverse de la table de la loi N }(0\,;1)$$

$$\Leftrightarrow a = 1{,}96\frac{\sqrt{n}}{\sigma}$$

$$\Leftrightarrow \mathbf{a = 2{,}51}$$

EXERCICE 3

Pour répondre à cette problématique, nous allons utiliser le test du Khi^2.

1. Hypothèses

• Soit H_0 l'hypothèse nulle : la qualité gustative des œufs ne dépend pas du mode d'élevage des poules.

• Soit H_1 l'hypothèse alternative : la qualité gustative des œufs dépend du mode d'élevage des poules.

2. Calcul des effectifs théoriques : t_i

Par définition, l'effectif théorique $= \dfrac{\text{total ligne} \times \text{total colonne}}{\text{total général}}$

Avec ni les effectifs observés et t_i les effectifs théoriques.

Notons que l'on a bien $n \geq 50$ puisque $n =$ ~~200~~ et que les $t_i \geq 5$, conditions pour pouvoir appliquer le test.

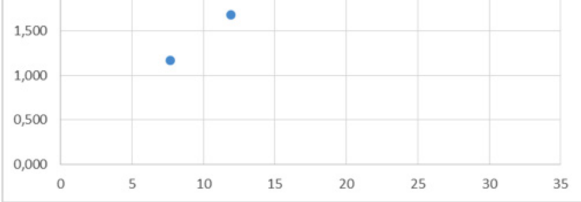
On obtient alors $\chi^2\text{obs} = \dfrac{(14-13)^2}{13} + \dfrac{(8-8)^2}{8} + \ldots$

D'où **χ^2obs = 0,38**

$\boxed{n = 60}$

On peut également répertorier les termes de la somme donnant le χ^2obs dans un tableau pour avoir à taper une somme plus agréable à la calculatrice :

Envahissement des parcelles	Peu envahies	Moyennement envahies	Très envahies



c. **$r(v, z) = 0{,}997$** à 10^{-3} près

d. **$z = 0{,}099\, v + 0{,}452$**

e. Détaillons le calcul de e_1 :

$$\begin{aligned} e_1 &= z_1 - (0{,}099\, v_1 + 0{,}452) \\ &= 1{,}169 - (0{,}099 \times 7{,}7 + 0{,}452) \\ &= -\,0{,}045 \end{aligned}$$

En procédant de même pour les autres résidus, on obtient les valeurs consignées dans le tableau ci-dessus.

0,052

f. La moyenne des résidus valant – 0,009 et l'écart type ~~0,0027~~, ces paramètres sont suffisamment faibles pour légitimement retenir cet ajustement. De plus, r est très proche

Avec ni les effectifs observés et ti les effectifs théoriques.

Notons que l'on a bien n ≥ 50 puisque n = ~~200~~ et que les ti ≥ 5, conditions pour pouvoir appliquer le test.

On obtient alors $\chi^2 obs = \dfrac{(15-22,5)^2}{22,5} + \dfrac{(75-54)^2}{54} + ...$

D'où χ^2**obs = 22,60** $\boxed{n = 400}$

On peut également répertorier les termes de la somme donnant le χ^2obs dans un tableau pour avoir à taper une somme plus agréable à la calculatrice :

Situation familiale / Consommation de produits biologiques	Célibataire	En couple sans enfant	En couple avec enfant
Oui	2,500	8,167	1,761
Non	2,045	6,682	1,441

Ainsi on retrouve $\chi^2 obs = 2,500 + 8,167 + ... + 1,441 = 22,60$

Avec ni les effectifs observés et t_i les effectifs théoriques.

Notons que l'on a bien n ≥ 50 puisque n = ~~200~~ et que les ti ≥ 5, conditions pour pouvoir appliquer le test.

$\boxed{n = 90}$

On obtient alors $\chi^2\text{obs} = \frac{(24-20{,}70)^2}{20{,}70} + \frac{(3-6{,}30)^2}{6{,}30} + \ldots.$

D'où $\chi^2\text{obs} = 6{,}46$

On peut également répertorier les termes de la somme donnant le χ^2obs dans un tableau pour avoir à taper une somme plus agréable à la calculatrice :

	Vêlage	Avortement
Race 1	0,526	1,729
Race 2	0,871	2,861
Race 3	0,110	0,360

Ainsi on retrouve $\chi^2\text{obs} = 0{,}526 + 1{,}729 + \ldots + 0{,}360 = 6{,}46$

4. Détermination du Khi² théorique : χ^2th

• Le degré de liberté : ddl = k = (nombre de lignes – 1) x (nombre de colonnes - 1)
Ici, k = 2 x 1 = 2

3. Calcul du Khi² observé : χ^2obs

Par définition, on a : $\chi^2\text{obs} = \sum_{i=1}^{j} \frac{(n_i - t_i)^2}{t_i}$, avec ici j = 4

$\boxed{n = 100}$

Avec ni les effectifs observés et t_i les effectifs théoriques.
Notons que l'on a bien n ≥ 50 puisque n = ~~200~~ et que les ti ≥ 5, conditions pour pouvoir appliquer le test.

On obtient alors $\chi^2\text{obs} = \frac{(35-37{,}5)^2}{\ldots} + \frac{(15-12{,}5)^2}{\ldots} + \ldots$

tement linéaire n’est pas envisageable.
stement ne pourraient être cohérentes.

1970	1980	1990	2000	2010
70	80	90	100	110
200	175	160	153	130
,298	5,165	5,075	5,030	4,868

a une corrélation linéaire quasi parfaite
tion est plus forte entre les variables T et
rmée de Tuckey. Tukey

Par définition, on a : $\chi^2 obs = \sum_{i=1} \frac{\ldots}{t_i}$, avec ici j = 12

Avec ni les effectifs observés et t_i les effectifs théoriques.

Notons que l'on a bien n ≥ 50 puisque n = ~~200~~ et que les ti ≥ 5, conditions pour pouvoir appliquer le test.

$$\text{On obtient alors } \chi^2 obs = \frac{(12-11{,}50)^2}{11{,}50} + \frac{(9-11{,}50)^2}{11{,}50} + \ldots$$

D'où **χ^2obs = 4,84**

$\boxed{n = 120}$

On peut également répertorier les termes de la somme donnant le χ^2obs dans un tableau pour avoir à taper une somme plus agréable à la calculatrice :

Qualité \ Origine géographique	Alaska	Ecosse	Irlande	Norvège
Excellente	0,022	0,543	0,196	1,065
Bonne	0,139	0,006	0,272	0,006
Moyenne	0,078	1,043	0,009	1,457

$y = 1,831e^{0,547x12}$

7. 2017 correspond à x = 12

d'où $\cancel{y = 1,831\,x}$? $\Rightarrow$ **$y = 1298$**

On peut donc estimer le nombre de pieds atteints en 2017 à 1298 soit environ 1300, si la tendance ne change pas d'ici 2017.

EXERCICE 4

1.a. Soit X la variable aléatoire qui, à un échantillon aléatoire simple de n = 120 escargots, associe le nombre d'escargots de couleur jaune.

On a $X \hookrightarrow B(120\,;p)$ avec $E(X) = np = 120\,p$

Or, $F = \dfrac{X}{n} = \dfrac{X}{120}$ et par définition, $E(aX) = a\,E(X)$ pour tout réel a

Ici $a = \dfrac{1}{120}$ donc $E(F) = \dfrac{1}{120}\ E(X) = \dfrac{1}{120} \times 120\,p$ d'où **$E(F) = p$**

b. n = 120 donc $n \geq 30$. n est donc suffisamment grand pour avoir :

$F \hookrightarrow N(E(F)\,;\sigma(F))$

Alors on a : **$F \hookrightarrow N(p\,;\ \sigma(F))$**

c. $F \hookrightarrow N(0,3\,;0,04)$ en effet, $\sigma(F) = \sqrt{\dfrac{0,3 \times 0,7}{120}} = 0,04$

Par définition, on a : $\chi^2 obs = \sum_{i=1} \frac{\dots}{t_i}$, avec ici j = 6

Avec ni les effectifs observés et t_i les effectifs théoriques.

Notons que l'on a bien n ≥ 50 puisque n = ~~200~~ et que les ti ≥ 5, conditions pour pouvoir appliquer le test.

$\boxed{n = 315}$

On obtient alors $\chi^2 obs = \frac{(47-45,27)^2}{45,27} + \frac{(15-16,73)^2}{16,73} + \dots$

D'où $\chi^2 obs = 6,93$

On peut également répertorier les termes de la somme donnant le $\chi^2 obs$ dans un tableau pour avoir à taper une somme plus agréable à la calculatrice :

Situation géographique / Habitudes de consommation	Nord	Sud

$n \geq 30$ donc n est suffisamment grand pour que l'on approche la loi de F par :

$$F \hookrightarrow N\left(p\,;\ \sqrt{\frac{p(1-p)}{n}}\right)$$

$\sigma(F) = 0{,}008$ soit $0{,}01$ à 10^{-2} près

b. $n = 300$ donc **$E(F) = 0{,}98$** à 10^{-2} près et **$\sigma(F) = \cancel{0{,}08}$** à 10^{-2} près

$P(F \leq 0{,}97) = P\left(T \leq \dfrac{0{,}97 - 0{,}98}{\cancel{0{,}08}\ 0{,}01}\right)$ avec $T = \dfrac{F - \mu}{\sigma}$ et $T \hookrightarrow N(0\,;1)$

$= P(T \leq \cancel{-0{,}13})$ $-1{,}00$

$= P(T \geq \cancel{0{,}13})$ $1{,}00$ par symétrie de la courbe de la loi normale centrée réduite

$= 1 - P(T < \cancel{0{,}13})$ $1{,}00$

$= 1 - P(T \leq \cancel{0{,}13})$ $1{,}00$ car $P(T = \cancel{0{,}13}\ 1{,}00) = 0$

$= 1 - \cancel{0{,}5517}$ $0{,}8413$ par lecture de la table de la loi normale centrée réduite

$= \cancel{0{,}4483}$ $0{,}1587$

$P(F \leq 0{,}97) = \cancel{0{,}45}$ $0{,}16$ à 10^{-2} près

$\overset{6\,-\,a}{}$

3. On cherche a tel que $P(\cancel{6 + a} \le X \le 6 + a) \ge 0{,}9$

$$P(\underset{6\,-\,a}{\cancel{6 + a}} \le X \le 6 + a) \ge 0{,}9 \Leftrightarrow P\left(\frac{6-a-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{6+a-\mu}{\sigma}\right) \ge 0{,}9$$

$$\Leftrightarrow P\left(-\frac{a}{0{,}4} \le U \le \frac{a}{0{,}4}\right) \ge 0{,}9 \text{ avec } U = \frac{X-\mu}{\sigma} \text{ et } U \hookrightarrow N(0\,;1)$$

$$\Leftrightarrow 2 \times P\left(U \le \frac{a}{0{,}4}\right) - 1 \ge 0{,}9$$

$$\Leftrightarrow P\left(U \le \frac{a}{0{,}4}\right) \ge 0{,}95$$

$$\Leftrightarrow \frac{a}{0{,}4} \ge \underset{1{,}645}{\cancel{1{,}96}} \text{ par lecture inverse de la table de la loi } N(0\,;1)$$

$$\Leftrightarrow a \ge \cancel{1{,}96} \times 0{,}4$$

$$\Leftrightarrow \mathbf{a \ge \underset{0{,}66}{\cancel{0{,}78}}} \quad \text{à } 10^{-2} \text{ près}$$

EXERCICE 3

1. L'expérience qui consiste à voir si un sac de granules est de qualité optimale ou non

$\quad\quad 36 \quad\quad\quad\quad 35{,}5$

3. $P(X \geq \cancel{35}) = P(Y \geq \cancel{34{,}5})$ par correction de continuité

$$= P\left(T \geq \frac{34{,}5 - 40}{2{,}83}\right) \quad \text{avec } T = \frac{Y-\mu}{\sigma} \quad \text{et} \quad T \hookrightarrow N(0\,;1)$$

$$= P(T \geq -\cancel{1{,}94}) \quad -1{,}59$$

$$= P(T \leq \cancel{1{,}94}) \quad 1{,}59$$ par symétrie de la courbe de la loi normale centrée réduite

$$= \cancel{0{,}9738} \quad 0{,}9441$$ par lecture de la table de la loi normale centrée réduite

$P(X \geq \cancel{35}) = \cancel{0{,}97}$ à 10^{-2} près

$36 \quad\quad 0{,}94$

Il y a environ ~~97~~ % de chances d'avoir plus de 35 sacs de granules de qualité optimale sur 50 sacs prélevés.

94 %

Type B	28,8	11,2

Par exemple, on a : $t_1 = \dfrac{60 \times 72}{100} = 43{,}2$

43,2 est l'effectif théorique de 48.

$\boxed{n = 100}$

3. Calcul du Khi² observé : $\chi^2 obs$

Par définition, on a : $\chi^2 obs = \displaystyle\sum_{i=1}^{j} \frac{(n_i - t_i)^2}{t_i}$, avec ici $j = 4$

Avec n_i les effectifs observés et t_i les effectifs théoriques.

Notons que l'on a bien $n \geq 50$ puisque $n = \cancel{200}$ et que les $t_i \geq 5$, conditions pour pouvoir appliquer le test.

Partie A

1. ~~$z = 0,547x + 0,605$~~ $y = 0,017x + 0,158$

2. Si $x = 30$, $y = 0,668$

Le taux de dégâts d'une forêt de pins maritimes de haut
à des vents violents serait donc de 0,67.

Partie B

1. Nous étudierons, dans un premier temps, graphiqu
quement les coefficients de corrélation des 2 séries.
Concernant les résidus, les points des résidus de la séri
des abscisses donc les résidus sont plus faibles. Malgré
très élevé, présageant d'une croissance exponentielle
raison, on peut déjà retenir la série (x, z). De plus, les p
sont plus alignés que les points du nuage de la premiè

donc $\overline{x} = 6{,}98$ kg

Par conséquent, $s = \sqrt{\frac{1}{n}\sum(x_i - \overline{x})^2} = \sqrt{\cancel{45{,}62}}$ 0,027 donc $\hat{\sigma}^2 = \frac{16}{15} \times \cancel{45{,}62}$ 0,027

d'où $\hat{\sigma}^2 = \cancel{48{,}66}$ 0,029

En prenant $\sigma = \sqrt{\cancel{48{,}66}}$, on obtient :
0,029

$$\mu \in \left[\, 6{,}98 - 2{,}13\frac{\sqrt{\cancel{48{,}66}}}{\sqrt{16}} \;;\; 6{,}98 + 2{,}13\,\frac{\sqrt{\cancel{48{,}66}}}{\sqrt{16}} \,\right] \quad 0{,}029$$

$$\Rightarrow \qquad \mu \in \cancel{[3{,}27\,;\,10{,}69]}$$

[6,89 ; 7,07]

Partie B

1. On cherche σ tel que $P(X \leq 7{,}4) \geq 0{,}996$

$$P(X \leq 7{,}4) \geq 0{,}996 \;\Leftrightarrow\; P\left(U \leq \frac{7{,}4-7}{\sigma}\right) \geq 0{,}996 \qquad \text{avec } U = \frac{X-\mu}{\sigma}$$

$$\Leftrightarrow\; P\left(U \leq \frac{0{,}4}{\sigma}\right) \geq 0{,}996$$

Par lecture inverse de la table de la loi normale centrée réduite, on ob

$$\text{d'où}\quad \sigma \leq \frac{0{,}4}{2{,}65} = 0{,}151 \;\Rightarrow\; \sigma_{Max} = 0{,}15 \quad \text{à } 10^{-2} \text{ près}$$

$$\frac{6{,}77-7}{\ } \qquad \frac{X-\mu}{\ }$$