DUT TC2 - Module OS 01 - PROBABILITÉS ET STATISTIQUE INFÉRENTIELLE

CORRECTION Exercices Chapitre 2 - Lois de probabilité continues usuelles.

Exercice 11 *Correction* : On a $X \rightsquigarrow \mathcal{N}(120, 14)$. On souhaite calculer $p(\{110 \leq X \leq 130\})$. On va travailler avec la loi centrée réduite associée à X. La variable aléatoire $T = \dfrac{X - 120}{14}$ suit la loi normale centrée réduite $\mathcal{N}(0, 1)$. On a alors

$$p(\{110 \leq X \leq 130\}) = p\left(\left\{-\frac{5}{7} \leq T \leq \frac{5}{7}\right\}\right) = 2\Pi(0,7143) - 1$$

où Π est la fonction de répartition de la loi centrée réduite. La table de la fonction de répartition de cette loi donne $\Pi(0,71) = 0,7611$ et $\Pi(0,72) = 0,7642$. Par interpolation linéaire, on obtient $p(\{50 \leq X \leq 60\}) \simeq 2 \times 0,7624 - 1 = 0,5428$ à 10^{-4} près. En effet, on a les encadrements

$0,71$	$0,7143$	$0,72$
$0,7611$	v	$0,7642$

$\Leftrightarrow \dfrac{0,7143 - 0,71}{0,72 - 0,71} = \dfrac{v - 0,7611}{0,7642 - 0,7611} \Leftrightarrow v = 0,7611 + 0,0031 \times \dfrac{0,0043}{0,01}$. On en déduit que $v \simeq 0,7624$.

Exercice 12 *Correction* :

1. *Statistique* - On a le tableau :

i	x_i	n_i	$n_i x_i$	$n_i x_i^2$
1	95	1	95	9025
2	96	3	288	27648
3	97	6	582	56454
4	98	8	784	76832
5	99	10	990	98010
6	100	13	1300	130000
7	101	18	1818	183618
8	102	14	1428	145656
9	103	9	927	95481
10	104	8	832	86528
11	105	6	630	66150
12	106	2	212	22472
13	107	2	214	22898
Total	–	100	10100	1020772

On en déduit que $\overline{x} = \dfrac{10100}{100} = 101$ et $V(X) = \dfrac{1020772}{100} - (101)^2 = 6,72$ ce qui donne l'écart-type $\sigma = \sqrt{6,72} = 2,59$ à 10^{-2} près.

2. *Probabilités* - On admet que $Y \rightsquigarrow \mathcal{N}(101; 2,59)$. On veut calculer la probabilité de l'événement "le contrat est rempli" c'est-à-dire $p(\{Y \geq 99,5\})$. On utilise la variable centrée réduite $T = \dfrac{Y - 101}{2,59}$.

$p(\{Y \geq 99,5\}) = p(\{T \geq -0,5792\}) = p(\{T \leq 0,5792\}) = \Pi(0,5792)$. On peut lire dans la table de la loi normale centrée réduite les valeurs $\Pi(0,57) = 0,7157$ et $\Pi(0,58) = 0,7190$. On a les encadrements

$0,57$	$0,5792$	$0,58$
$0,7157$	v	$0,7190$

$$\Leftrightarrow \frac{0,5792 - 0,57}{0,58 - 0,57} = \frac{v - 0,7157}{0,7190 - 0,7157} \Leftrightarrow v = 0,7157 + 0,0033 \times \frac{0,0092}{0,01} \simeq 0,7187.$$

À l'aide de cette interpolation affine, on obtient

$$p(\{Y \geq 99,5\}) \simeq 0,7187 \text{ à } 10^{-4} \text{ près.}$$

Exercice 13 *Correction* :

1. (a) L'univers Ω est défini par $\{p, f\}$ avec $Card(\Omega) = 2$.
   - Soit l'événement A : "obtenir 'face"'. En utilisant l'équiprobabilité,
   $$p(A) = \frac{1}{2}.$$
   - L'événement B est le complémentaire dans Ω de A. On a alors
   $$p(B) = 1 - p(A) = \frac{1}{2}.$$

   En notant X la variable aléatoire qui, à chaque partie de 10 épreuves, associe le nombre de fois où A peut être réalisé, on a
   $$X = X_1 + X_2 + \ldots + X_{10}$$
   où X_i est la variable aléatoire définie par $X_i(A) = 1$ et $X_i(B) = 0$ pour $i \in \{1, 2, \ldots, 10\}$. La variable X_i est donc une variable de Bernoulli de loi de probabilité

x_j	0	1	**Total**
$p_j = p(\{X_i = x_j\})$	$\frac{1}{2}$	$\frac{1}{2}$	1

   ainsi $X_i \rightsquigarrow \mathcal{B}\left(\frac{1}{2}\right)$. X étant la somme de 10 variables de Bernoulli de paramètre $\frac{1}{2}$ indépendantes deux à deux, on a finalement $X \rightsquigarrow \mathcal{B}\left(10, \frac{1}{2}\right)$.

   (b) On veut calculer $p(\{3 \leq X \leq 6\})$:
   $$p(\{3 \leq X \leq 6\}) = p(\{X \leq 6\}) - p(\{X \leq 2\}) = \sum_{k=3}^{6} C_{10}^{k} \left(\frac{1}{2}\right)^{k} \left(\frac{1}{2}\right)^{10-k} = \left(\frac{1}{2}\right)^{10} \sum_{k=3}^{6} C_{10}^{k}$$
   $$= \left(\frac{1}{2}\right)^{10} (C_{10}^{3} + C_{10}^{4} + C_{10}^{5} + C_{10}^{6}) = \left(\frac{1}{2}\right)^{10} (120 + 210 + 252 + 210) = \frac{792}{2^{10}} \simeq 0,7734.$$

2. On suppose que $X \rightsquigarrow \mathcal{N}(5, \sqrt{2,5})$. On a

   (a) . $m = np = 10 \times \frac{1}{2} = 5$,

   . $\sigma = \sqrt{npq} = \sqrt{10 \times \frac{1}{2} \times \frac{1}{2}} = \sqrt{2,5}$.

   (b) $Y \rightsquigarrow \mathcal{N}(0,5; \sqrt{2,5})$. On veut calculer la probabilité de l'événement "le nombre de 'face' est compris entre 3 et 6 (bornes incluses)", c'est-à-dire $p(\{2,5 \leq Y \leq 6,5\})$.

   On utilise la variable aléatoire normale centrée réduite $T = \dfrac{Y - 5}{\sqrt{2.5}}$ ainsi que la table du formulaire. On obtient $p(\{2,5 \leq Y \leq 6,5\}) = \Pi(0,95) - \Pi(-1,58) = \Pi(0,95) - (1 - \Pi(1,58))$
   $= 0,8289 - (1 - 0,9429) = 0,7718$.

Exercice 14 *Correction* : On sait que $Y \rightsquigarrow \mathcal{N}(80, 60)$.

1. Z est la somme de 25 variables aléatoires indépendantes de même loi $\mathcal{N}(80, 60)$ donc Z suit la loi normale de moyenne $25E(Y) = 2000$, de variance $25V(Y)$ et d'écart-type $5\sigma(Y) = 300$; Z suit la loi $\mathcal{N}(2000, 300)$. On a alors $p(\{Z > 2300\}) = 1 - p(\{Z \leq 2300\}) = 1 - 0,8413 = 0,1587$ en considérant la loi centrée réduite $T = \dfrac{Z - 2000}{300}$.
2. $p(\{Z > a\}) = 1 - p(\{Z \leq a\}) = 0,05 \Leftrightarrow p(\{Z \leq a\}) = 0,95$. Or $p(\{T \leq b\}) = 0,95$ pour $b = 1,65$ donc $a = 1,65 \times 300 + 2000 = 2495$.

Exercice 15 *Correction* : X suit la loi normale $\mathcal{N}(20, 5)$. La variable aléatoire $T = \dfrac{X - 20}{5}$ suit la loi normale centrée réduite $\mathcal{N}(0, 1)$ dont la table est dans l'annexe A.

1. $p(\{X \leq 28\}) = p\left(\left\{T \leq \dfrac{28 - 20}{5}\right\}\right) = p(\{T \leq 1,6\}) = \Pi(1,6) \simeq 0,9452$
2. $p(\{X \geq 28\}) = 1 - \Pi(1.6) \simeq 0,0548$
3. $p(\{X \geq 12\}) = p\left(\left\{T \geq \dfrac{12 - 20}{5}\right\}\right) = p(\{T \leq -1,6\}) = \Pi(1,6) \simeq 0,9452$
4. $p(\{X \leq 12\}) = 1 - \Pi(1,6) \simeq 0,0548$

5. $p(\{12 \leq X \leq 28\}) = p\left(\left\{\frac{12-20}{5} \leq T \leq \frac{28-20}{5}\right\}\right) = 2\Pi(1,6) - 1 \simeq 0,8904.$

Exercice 16 *Correction* : Pour utiliser le test de Student-Fischer, il faut faire l'hypothèse que les populations sont normales.

1. Les deux échantillons sont indépendants
   - Pour le premier échantillon : $\sum_i x_i = 84+92+72+91+84 = 423$, $\sum_i x_i^2 = 84^2+92^2+72^2+91^2+84^2 = 36041$.
     La moyenne de cet échantillon est $\frac{423}{5} = 84,6$.
     On admettra que la variance estimée est calculée par $\sigma_1^2 = \frac{1}{4}(36041 - \frac{1}{5}423^2) = 63,8$.
   - Pour le second échantillon : $\sum_i x_i = 81+88+74+81+90 = 414$, $\sum_i x_i^2 = 81^2+88^2+74^2+81^2+90^2 = 34442$.
     La moyenne de cet échantillon est $\frac{414}{5} = 82,8$.
     On admettra que la variance estimée est calculée par $\sigma_2^2 = \frac{1}{4}(34442 - \frac{1}{5}414^2) = 40,7$.

   L'intervalle de confiance à 95% se calcule avec la loi de Student à 4 degrés de liberté : on récupère à l'aide de l'annexe C2, $t_{0,95} = 2,776$. Par conséquent l'intervalle de confiance à 95% est donné par

   $$\left[\overline{x} - 2,776\frac{\sigma}{\sqrt{n}}; \overline{x} + 2,776\frac{\sigma}{\sqrt{n}}\right]$$

   soit :
   (a) pour le premier échantillon (avant) : $[84,6 - 9,92; 84,6 + 9,92] = [74,68; 94,52]$,
   (b) pour le second échantillon (après) : $[82,8 - 7,92; 82,8 + 7,92] = [74,88; 90,72]$.
   (c) La perte de masse corporelle y est la valeur avant régime à laquelle on soustrait la valeur après régime.
   Pour étudier la compatibilité des écart-types donc des variances, on calcule $\frac{\sigma_1^2}{\sigma_2^2} = \frac{63,8}{40,7} = 1,57$ or, la table de Fischer-Snedecor (voir annexe D) à $(4,4)$ degrés de liberté donne la valeur seuil de $9,60$ donc $\frac{\sigma_1^2}{\sigma_2^2} < 9,60$ ce qui permet d'affirmer que les variances sont compatibles et on peut parler dans ce ce cas de variance estimée commune. Cette variance vaut $\sigma^2 = \frac{4\times 63,8 + 4\times 40,7}{8} = \frac{1}{2}(63,8+40,7) = 52,25$ et l'écart-type estimé commun est égal à $\sigma = 7,23$. L'intervalle de confiance à 95% se calcule avec la loi de Student à 8 degrés de liberté, on obtient à l'aide de la table de l'annexe C2 $t_{0,95} = 2,306$ donc l'intervalle de confiance à 95% est donné par

   $$\left[\overline{y} - 2,306\sigma\sqrt{\frac{1}{\nu_1}+\frac{1}{\nu_2}}; \overline{y} + 2,306\sigma\sqrt{\frac{1}{\nu_1}+\frac{1}{\nu_2}}\right]$$
   $$= \left[(84,6-82,8) - 2,306\times 7,23\times\sqrt{\tfrac{2}{5}}; (84,6-82,8) - 2,306\times 7,23\times\sqrt{\tfrac{2}{5}}\right].$$

   Finalement, on trouve l'intervalle $[1,8 - 10,54; 1,8 + 10,54] = [-8,74; 12,34]$.

2. On a le tableau suivant :

Individu	1	2	3	4	5	**Total**
Avant	84	92	72	91	84	423
Après	81	88	74	81	90	414
Perte δ	3	4	-2	10	-6	9
δ^2	9	16	4	100	36	165

   On a donc $\overline{\delta} = \frac{9}{5} = 1,8$ et $\sigma^2 = \frac{1}{4}(165 - \frac{1}{5}9^2) = 37,2$ soit $\sigma = 6,1$. La loi de Student à 4 degrés de liberté donne comme précédemment $t_{0,95} = 2,776$. L'intervalle de confiance à 95% est donc

   $$\left[\overline{\delta} - 2,776\frac{\sigma}{\sqrt{n}}; \overline{\delta} + 2,776\frac{\sigma}{\sqrt{n}}\right] = [1,8 - 7,57; 1,8 + 7,57] = [-5,77; 9,37].$$

   En comparant cet intervalle de confiance à 95% avec le précédent, on voit que celui qu'on obtient avec les séries appariées est plus précis ; l'appariement est une caractéristique qu'il est souhaitable d'introduire chque fois que c'est possible.

   La conclusion de ce test est que 0 appartient à l'intervalle de confiance de l'amaigrissement : on ne peut pas rejeter l'hypothèse que l'amaigrissement est nul.

$\boxed{\textbf{Exercice 17}}$ *Correction* :

1. . $p_1 = p(\{8,45 \leq X \leq 8,70\})$: pour pouvoir se référer aux tables de la loi normale, on pose
$$T = \frac{X - m_X}{\sigma_X} = \frac{X - 8,55}{0,05}.$$
Donc $p_1 = p(\{-2 \leq T \leq 3\}) = \Pi(3) - \Pi(-2) = \Pi(3) - (1 - \Pi(2))$ de par la symétrie de la fonction de densité. On utilise ensuite la table de la loi normale centrée réduite et on récupère les valeurs $\Pi(3) = 0,99865$ et $\Pi(2) = 0,9772$. Par conséquent, $p_1 = 0,99865 - (1 - 0,9772) = 0,97585$.

   . $p_2 = p(\{5,07 \leq Y \leq 5,33\})$: on pose
$$W = \frac{Y - m_Y}{\sigma_Y} = \frac{Y - 5,2}{0,05}.$$
Donc $p_2 = p(\{-2,6 \leq W \leq 2,6\}) = \Pi(2,6) - \Pi(-2,6) = \Pi(2,6) - (1 - \Pi(2,6)) = 2\Pi(2,6) - 1$. On utilise ensuite la table de la loi normale centrée réduite et on récupère la valeur $\Pi(2,6) = 0,9953$. Par conséquent, $p_2 = 2 \times 0,9953 - 1 = 0,9906$.

2. (a) Calculons le pourcentage de pilules conformes à la sortie de la chaîne de fabrication. On cherche donc la probabilité de l'intersection des deux événements $8,45 \leq X \leq 8,70$ et $5,07 \leq Y \leq 5,33$ soit
$$p = p(\{(8,45 \leq X \leq 8,70) \cap (5,07 \leq Y \leq 5,33)\}).$$
Comme les événements X et Y sont indépendants,
$$p = p(\{8,45 \leq X \leq 8,70\}) \times p(\{5,07 \leq Y \leq 5,33\}) = 0,97585 \times 0,9906 = 0,966677$$
soit un pourcentage de $96,6677\%$. Le pourcentage de pilules qui seront hors normes à la sortie de la chaîne de fabrication vaut par conséquent $100 - 96,6677 \simeq 3,33\%$.

   (b) Si on veut que le pourcentage de pilules défectueuses ne dépasse pas 3%, le procédé ne peut être retenu puisque $3,33\% > 3\%$. On cherche σ_Y pour que le pourcentage de pilules défectueuses soit inférieur à 3%. On veut déterminer la valeur de $\tilde{p}$ telle que
$$. \ 1 - 0,97585 \times \tilde{p} < 0,03 \Leftrightarrow \tilde{p} > 0,9940052$$
ce qui revient à déterminer a tel que
$$p(\{5,07 \leq Y \leq 5,33\}) = 2\Pi(a) - 1 > 0,9940052 \Leftrightarrow \Pi(a) > 0,9970026.$$
À l'aide de la table de la loi normale centrée réduite, on trouve
$$2,75 < a < 2,76$$
On utilise ensuite l'interpolation linéaire, on a les encadrements

$2,75$	a	$2,76$
$0,9970$	$0,9970026$	$0,9971$

ce qui nous permet d'écrire que
$$\frac{a - 2,75}{0,9970026 - 0,9970} = \frac{2,76 - 2,75}{0,9971 - 0,9970} \Leftrightarrow a = 2,75026.$$
Finalement comme $a = \dfrac{5,33 - 5,2}{\sigma_Y}$, on trouve $\sigma_Y \simeq 0,047$.

3. (a) On a $S = X + Y$. Par conséquent,
   . $E(S) = E(X+Y) = E(X)+E(Y) = 8,55+5,20 = 13,75$ car X et Y sont des variables indépendantes.
   . $V(S) = V(X + Y) = V(X) + V(Y)$ car X et Y sont des variables indépendantes. On en déduit que $V(S) = 2 \times (0,05)^2 = 0,005$. L'écart-type de S vaut alors $\sigma_S = 0,0707$.

   (b) $S \rightsquigarrow \mathcal{N}(13,75; 0,0707)$. On pose
$$V = \frac{S - 13,75}{0,0707}$$
ce qui implique que $p = p(\{13,6 \leq S \leq 13,8\}) = p(\{-2,122 \leq V \leq 0,707\}) = \Pi(0,707) - \Pi(-2,122)$. Toujours en utilisant la symétrie de la fonction de densité, on obtient
$$p = \Pi(0,707) + \Pi(2,122) - 1.$$
Déterminons $\Pi(0,707)$ et $\Pi(2,122)$ par interpolation linéaire.
   . Pour $\Pi(0,707)$ on a les encadrements

$0,70$	$0,707$	$0,71$
$0,7580$	$\Pi(0,707)$	$0,7611$

   ce qui nous permet d'écrire que
$$\frac{\Pi(0,707) - 0,7580}{0,7611 - 0,7580} = \frac{0,707 - 0,70}{0,71 - 0,70} \Leftrightarrow \Pi(0,707) = 0,7582.$$
   . Pour $\Pi(2,122)$ on a les encadrements

2,12	2,122	2,13
0,9830	$\Pi(2,122)$	0,9834

ce qui nous permet d'écrire que

$$\frac{\Pi(2,122) - 0,9830}{0,9834 - 0,9830} = \frac{2,122 - 2,12}{2,13 - 2,12} \Leftrightarrow \Pi(2,122) = 0,9831.$$

Finalement, $p = 0,7582 + 0,9831 - 1 = 0,7413$.

4. Soit $p_{HN} = 0,01$.

   (a) On tire sans remise 100 pilules dans un ensemble où chaque tirage est un tirage de Bernoulli (soit pilule hors norme de probabilité p, soit pilule conforme de probabilité $q - 1 = p$). Le nombre de pilules qui sont hors-norme est une variable binomiale de paramètres $n = 100$ et $p = 0,01$.

   (b) On a $n \geq 30$, $p \leq 0,1$ et $np = 1 \leq 10$ qui justifient l'approximation de Poisson, le paramètre de cette variable étant $np = 1$. Le complémentaire de l'événement $Z \geq 5$ est $Z \in \{0,1,2,3,4\}$. On approche la probabilité de cet événement par la somme

   $$e^{-1}\left(1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24}\right) = \frac{65}{24e} = 0,9963$$

   puisque pour la loi de Poisson, $p(\{X = k\}) = e^{-\lambda}\dfrac{\lambda^k}{k!}$. On récupère finalement la valeur approximative cherchée en calculant $p(\{Z \geq 5\}) = 1 - 0,9963 = 0,0037$.

5. (a) On a $U \rightsquigarrow \mathcal{B}(n,p)$. On cherche $p(\{X + Y < 13,8\}) = p(\{U < 13,8\})$. On pose

   $$R = \frac{U - 13,75}{0,0707} = 0,707$$

   En utilisant la table de la loi normale, on trouve $p(\{U < 13,8\}) = p(\{R < 0,707\}) = 0,7582$.

   (b) On a ici $n = 100 \geq 20$, $np = 75,82 \geq 10$ et $nq = 24,18 \geq 10$ et l'approximation par une variable normale est donc justifiée, ce ne serait pas le cas pour la loi de Poisson car, par exemple, $p = 0,75 \geq 0,1$. En utilisant une formule donnée dans le cours, la probabilité considérée a pour valeur approchée

   $$\Pi\left(\frac{8,55 - 75,82}{\sqrt{npq}}\right) - \Pi\left(\frac{69,5 - 75,82}{\sqrt{npq}}\right) = \Pi(2,2607) - \Pi(-1,476) = 0,9181$$