Remarque : En 2020-2021, les examens ont eu lieu en ligne. Ce document retranscrit un exemple d'énoncé du contrôle de première session du 05 mai.

Exercice 1 : Ensoleillement hivernal

On considère la durée d'ensoleillement par jour dans la ville de NEVEPOLIS, au cours des 979 journées d'hiver des 11 dernières années.

On obtient les données ci-dessous :

Ensoleillement X en minutes	[5 ; 37[	[37 ; 69[	[69 ; 101[	[101 ; 133[	[133 ; 165[	[165 ; 197[	[197 ; 229[
Nombre de journées	100	214	317	232	91	22	3

1. Sur ces données, combien valent la moyenne et l'écart type de la durée d'ensoleillement journalier ?
   moyenne : $m(X) = \frac{\sum c_i\, n_i}{n} = \frac{21\times100+53\times214+85\times317+\cdots+213\times3}{979} = \frac{85\,711}{979} \simeq 87{,}55$
   $m(X^2) = \frac{\sum c_i^2\, n_i}{n} = \frac{21^2\times100+53^2\times214+85^2\times317+\cdots+213^2\times3}{979} = \frac{8\,988\,539}{979}$
   $Var(X) = m(X^2) - m(X)^2 = \frac{8\,988\,539}{979} - \left(\frac{85\,711}{979}\right)^2 \simeq 1\,516{,}43$
   Écart-type : $s(X) = \sqrt{Var(X)} \simeq 38{,}94$
2. À titre de comparaison, quels auraient été les effectifs théoriques si l'ensoleillement suivait la loi $\mathcal{N}(90\,;\,40)$?
   On obtient les probabilités suivantes et les effectifs théoriques suivants :

Ensoleillement X en minutes	[5 ; 37[	[37 ; 69[	[69 ; 101[	[101 ; 133[	[133 ; 165[	[165 ; 197[	[197 ; 229[
Probabilité	0,0926	0,2072	0,3086	0,2505	0,1108	0,0267	0,0037
Effectif théorique	90,6	202,9	302,1	245,2	108,5	26,1	3,7

   - On a d'abord calculé la probabilité associée à chaque intervalle, en remplaçant le premier intervalle par $[-\infty\,;\,37[$, et le dernier intervalle par $[197\,;\,+\infty]$.
   - On a ensuite obtenu les effectifs théoriques en multipliant ces probabilités par l'effectif total qui est 979.
3. Si l'on voulait estimer l'ensoleillement moyen des journées d'hiver de NEVEPOLIS, à 15 minutes près avec une confiance de 99%, combien de journées d'observations (c'est à dire quelle taille d'échantillon) seraient nécessaires ?
   Pour avoir une précision 15 avec la confiance 0,99, il faut $n > z_\alpha^2 \frac{(s_e)^2}{h^2} \simeq 2{,}576^2 \frac{38{,}94^2}{15^2} \simeq 45$
4. Mr Lacombe est psychiatre et consulte les archives de ses 11 dernières années de consultations. Il compare le nombre de consultations pour dépressions qu'il a fait chaque année avec l'ensoleillement moyen enregistré dans sa ville de Névépolis : il obtient les données ci-dessous :

Année	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Ensoleillement X	81	73	57	100	103	74	86	123	40	125	98
Nombre de Consultations Y	44	35	53	34	34	44	46	35	55	34	42

(a) Calculer le coefficient de corrélation de Spearman entre cet ensoleillement et le nombre de consultations pour dépression.

Année	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
X	81	73	57	100	103	74	86	123	40	125	98
Y	44	35	53	34	34	44	46	35	55	34	42
rang X'	5	3	2	8	9	4	6	10	1,0	11	7
rang Y'	7,5	4,5	10	2	2	7,5	9	4,5	11	2	6
$(X'-Y')^2$	6,25	2,25	64	36	49	12,25	9	30,25	100	81	1,0

le coefficient de corrélation des rangs de Spearman est donc
$1-\left(6\times\frac{6{,}25+2{,}25+64+36+49+\cdots+1{,}0}{11\left(11^2-1\right)}\right)\simeq -0{,}777$

(b) Que peut-on en conclure de cette valeur du coefficient de Spearman ?
Il y a donc une très forte corrélation entre cet ensoleillement et le nombre de consultations pour dépression où le nombre de consultations tend à diminuer quand l'ensoleillement augmente.

Exercice 2 : Coloriage

Mr Simon, instituteur en maternelle, a développé une méthode novatrice pour développer les qualités artistiques de ses élèves.

Il enseigne à 13 enfants selon cette méthode, puis leur fait passer un test de coloriage, et obtient les scores ci-dessous :

12	13	12	10	12	12	8	16	13	12	12	13	10

On supposera dans la suite que les notes obtenues à ce test de coloriage suivent une loi normale.

1. (a) Calculer le score moyen et l'écart type de ce groupe d'enfants.
moyenne : $m(X)=\frac{\sum x_i}{n}=\frac{12+13+12+\cdots+10}{13}=\frac{155}{13}\simeq 11{,}92$
$m(X^2)=\frac{\sum x_i^2}{n}=\frac{12^2+13^2+12^2+\cdots+10^2}{13}=\frac{1\,891}{13}$
$Var(X)=m(X^2)-m(X)^2=\frac{1\,891}{13}-\left(\frac{155}{13}\right)^2\simeq 3{,}3$
Écart-type : $s(X)=\sqrt{Var(X)}\simeq 1{,}82$

(b) Estimer le score moyen de l'ensemble des élèves de maternelle, s'ils étaient soumis à la méthode d'apprentissage développée par Mr Simon. *On demande de déterminer un intervalle de confiance, avec la confiance* 90%.
Comme $n=13\leqslant 30$, on cherche t_α à partir de la table inverse de Student avec $p=\frac{\alpha}{2}=0{,}05$ et $n-1=12$ degrés de liberté (ddl)
On lit $t_\alpha\simeq 1{,}7823$ d'où $a_\alpha=t_\alpha\frac{s_e}{\sqrt{n-1}}\simeq 0{,}936\,4$.
On estime donc que le score moyen des élèves de maternelle soumis à cette

méthode d'apprentissage est dans l'intervalle $[11{,}92 - 0{,}936\,4\,;\, 11{,}92 + 0{,}936\,4] \simeq [10{,}98\,;\, 12{,}86]$ avec la confiance $c = 0{,}9$

2. À titre de comparaison, on considère 40 autres élèves de maternelle avec lesquels on n'a pas utilisé cette méthode. On leur fait passer le même test de coloriage, et on obtient les scores suivants :

Score	[5 ; 7[	[7 ; 9[	[9 ; 11[	[11 ; 13[	[13 ; 14[	[14 ; 15[	[15 ; 17[
Effectif	1	1	8	11	13	3	3

   (a) Calculer le score moyen et l'écart type au sein de cet échantillon.
   moyenne : $m(X) = \frac{\sum c_i\, n_i}{n} = \frac{6\times1+8\times1+10\times8+\cdots+16\times3}{40} = \frac{493}{40} \simeq 12{,}32$
   $m(X^2) = \frac{\sum c_i^2\, n_i}{n} = \frac{6^2\times1+8^2\times1+10^2\times8+\cdots+16^2\times3}{40} = \frac{6\,252}{40}$
   $Var(X) = m(X^2) - m(X)^2 = \frac{6\,252}{40} - \left(\frac{493}{40}\right)^2 \simeq 4{,}39$
   Écart-type : $s(X) = \sqrt{Var(X)} \simeq 2{,}1$

   (b) Estimer le score moyen des élèves de maternelle, s'ils ne sont pas soumis à la méthode d'apprentissage développée par Mr Simon. *On demande de déterminer un intervalle de confiance, avec la confiance 90%.*
   Comme n= 40 >30, on cherche z_α tel que $F(z_\alpha) = \frac{0{,}9+1}{2} = 0{,}95$
   On a $F(1{,}645) \simeq 0{,}95$ d'où $z_\alpha \simeq 1{,}645$ d'où $a_\alpha = z_\alpha \frac{s_e}{\sqrt{n-1}} \simeq 0{,}5532$
   On estime donc que le score moyen des élèves de maternelle sans cette méthode d'apprentissage est dans l'intervalle $[12{,}32 - 0{,}553\,2\,;\, 12{,}32 + 0{,}553\,2] \simeq [11{,}77\,;\, 12{,}87]$, avec la confiance $c = 0{,}9$.

3. Peut-on affirmer, avec la confiance 90% que les élèves obtiennent en moyenne des scores plus élevés à l'épreuve de coloriage s'ils sont confrontés à la méthodes de Mr Simon ? *Vous conclurez en vous référant aux réponses des questions précédentes.*
   Les intervalles que l'on trouvé aux questions 1.b. et 2.b. se chevauchent, donc on ne peut pas en déduire dans laquelle des deux situations la moyenne est la plus élevée.

Exercice 3 : Collection de dessins

Mr Marie est psychothérapeute, et utilise souvent des dessins auprès de ses patients. Il possède 8 cartes où sont imprimés des dessins : 2 cartes sont des dessins bleus, 3 des dessins rouges, et 3 des dessins verts.

L'objet de cet exercice est de savoir quelle est la probabilité, quand on pioche deux dessins au hasard, qu'ils aient la même couleur.

1. Combien vaut $\binom{11}{8}$?
   $\binom{11}{8} = 165$, comme l'indiquent les calculatrices.

2. Au total, combien y a-t-il de façons de choisir 2 cartes au sein du paquet de cartes de Mr Marie ?
   Il y a $\binom{8}{2} = 28$ façons des les choisir.
3. Parmi ces choix, combien correspondent à deux dessins qui sont tous les deux verts ? *Vous indiquerez quel calcul vous avez effectué (et si c'est la calculatrice qui a effectué le calcul, vous indiquerez ce que vous avez demandé à la calculatrice).*
   Cette fois-ci on considère les choix deux cartes parmi les 3 cartes verts. Il y en a donc $\binom{3}{2}$, et la calculatrice indique que $\binom{3}{2} = 3$.
   Donc il y a 3 choix où les deux dessins sont tous les deux vert.
4. De même, combien correspondent à deux dessins qui sont tous les deux bleus ? Et à deux dessins qui sont rouges ? *(Cette fois-ci, ont ne vous impose pas de détailler la façon dont vous obtenez le résultat).*
   De même il y a $\binom{2}{2} = 1$ choix de deux dessins bleus et $\binom{3}{2} = 3$ choix de deux dessins rouges.
5. Conclure : Quand on pioche deux dessins au hasard, quelle est la probabilité qu'ils soient de la même couleur ?
   Ainsi le nombre de choix où les deux cartes ont la même couleur est $3+1+3 = 7$, tandis que le nombre total de choix de deux cartes est 28. La probabilité demandée vaut donc $\frac{7}{28} = 0{,}25$.

Exercice 4 : Reconversion professionnelle

On cherche à déterminer dans cet exercice si les « reconversions professionnelles » concernent plus les personnes qui ont fait des « études courtes » (moins que Bac+2) ou celles qui ont fait des « études longues » (Bac+3 et plus). Pour cela on remarque tout d'abord que parmi les français·es de 32 ans, il y en a 27% qui ont fait des études longues. On va comparer cette proportion à celle que l'on observe chez un échantillon de personnes de 32 ans qui ont fait une reconversion professionnelle (c'est à dire des personnes qui, après avoir déjà exercé une activité professionnelle, ont repris des études pour changer de secteur d'activité).

1. On considère un échantillon de 9 personnes de 32 ans, et on désigne par X le nombre de personnes de l'échantillon qui avaient initialement fait des études longues. *Certaines d'entre elles peuvent depuis avoir fait une reconversion professionnelle, mais on ne le prend pas en compte dans cette question.*
   (a) Dans ces conditions, X suit une loi binomiale. Indiquez les paramètres le cette loi binomiale, ainsi que sa moyenne.
   $X \sim \mathcal{B}(9\,;\,0{,}27)$ donc $m(X) = 9 \times 0{,}27 = 2{,}43$.

(b) Calculer la probabilité $\mathbb{P}[2 \leqslant X \leqslant 4]$. *Vous indiquerez quels calculs vous avez effectués (et si c'est la calculatrice qui a effectué des calculs, vous indiquerez sommairement ce que vous avez demandé à la calculatrice).*
$\mathbb{P}[2 \leqslant X \leqslant 4] = \mathbb{P}[X \leqslant 4] - \mathbb{P}[X \leqslant 1] \simeq 0{,}933\,8 - 0{,}254\,8 \simeq 0{,}678\,96$, où l'on a utilisé la calculatrice pour calculer $\mathbb{P}[X \leqslant 4] \simeq 0{,}933\,8$ et $\mathbb{P}[X \leqslant 1] \simeq 0{,}254\,84$

2. On considère désormais un échantillon de 85 personnes de 32 ans, et on désigne à nouveau par X le nombre de personnes de l'échantillon qui avaient initialement fait des études longues. *À nouveau, peu importe si elles ont fait une reconversion professionnelle depuis.*

(a) X suit à nouveau une loi binomiale. Justifier que l'on peut l'approximer par une loi normale, et précisez sa moyenne et son écart type.
On a $n = 85 > 30$, et $n\,p \simeq 22{,}9 > 5$, et $n(1-p) \simeq 62 > 5$ donc on peut approximer X par $\mathcal{N}(22{,}9; 4{,}09)$ où $4{,}09 \simeq \sqrt{n\,p(1-p)}$

(b) En utilisant la loi normale de la question précédente, déterminez $\mathbb{P}[19 \leqslant X \leqslant 26]$. *On vous demande de faire une correction de continuité, et d'indiquer votre calcul en utilisant la table du formulaire.*

$$\begin{aligned}\mathbb{P}[19 \leqslant X \leqslant 26] &= \mathbb{P}[18{,}5 \leqslant X \leqslant 26{,}5] = \mathbb{P}[\tfrac{18{,}5-22{,}9}{4{,}09} < \tfrac{X-22{,}9}{4{,}09} < \tfrac{26{,}5-22{,}9}{4{,}09}] \\ &\simeq \mathbb{P}[-1{,}08 < Z < 0{,}88] \simeq F(0{,}88) - F(-1{,}08) \\ &\simeq 0{,}8106 - (1 - 0{,}859\,9) \simeq 0{,}810\,6 - 0{,}140\,1 \\ &\simeq 0{,}67\end{aligned}$$

3. On considère un échantillon de 85 personnes de 32 ans qui ont déjà fait (ou sont en train de faire) une reconversion professionnelle. On observe qu'au sein de cette échantillon il y en a 34 qui avaient initialement fait des études longues.

(a) À partir de cet échantillon, estimer avec la confiance 90% la proportion de personnes qui avaient fait des études longues parmi l'ensemble des personnes en reconversion. *Vous établirez un intervalle de confiance, avec la confiance* 90%.
On a $n = 85 > 30$, et $p_e = \frac{34}{85} = 0{,}4$, d'où $n\,p_e = 34 > 5$, et $n(1-p_e) = 51 > 5$ donc on peut utiliser la procédure du formulaire pour estimer cette proportion .
On a $F(1{,}645) \simeq 0{,}95$ d'où $z_\alpha \simeq 1{,}645$
d'où $a_\alpha = z_\alpha\sqrt{\frac{p_e(1-p_e)}{n}} \simeq 0{,}0874$.
On estime donc que est dans l'intervalle $[0{,}3126; 0{,}4874]$ avec la confiance $c = 0{,}9$.

(b) Conclure : Avec la confiance 90%, peut-on affirmer que parmi les personnes de 32 ans ayant fait une reconversion, il y en a plus que 27% qui avaient initialement fait des études longues ?

Oui, car toutes les valeurs de l'intervalle trouvé en question 3.a sont plus grandes que 27%.