Université Mohammed V
Faculté des Sciences. Rabat
Département de Mathématiques.

Année 2019-2020
Filière SMI, Module 18
Probabilité & Statistique

Contrôle Final Décembre 2019

(Durée 1H 30)

Il sera tenu compte de la présentation et de la rédaction. Aucun document n'est autorisé. Il est strictement interdit d'emprunter ou de prêter une calculatrice. Justifier les réponses.

Statistique (10 points)

Exercice : 1 On étudie la distribution des salaires horaires (en Dirhams) des employés d'une grande entreprise.

La distribution est donnée dans le tableau suivant où $a < 26$, f_3 et f_4 sont des inconnues.

Classes	$]a, 26]$	$]26, 30]$	$]30, 34]$	$]34, 40]$	$]40, 46]$	$]46, 50]$	$]50, 56]$
Fréquences	0.14	0.16	f_3	f_4	0.16	0.12	0.04

On prendra 3 chiffres après la virgule.

1) Calculer les fréquences cumulées en fonction de f_3 et f_4.

2) La moyenne arithmétique est égale à 35.62 DH. Calculer $\overline{x}$ en fonction de a, f_3 et f_4. En déduire une équation liant a, f_3 et f_4.

3) La médiane est égale à 34.6 DH.

a) Exprimer la médiane de cette série en fonction de f_3 et f_4. En déduire une équation liant f_3 et f_4.

b) Donner une deuxième équation liant seulement f_3 et f_4.

c) En déduire les valeurs de f_3 et f_4.

4) En déduire la valeur de l'inconnue a.

Exercice : 2 Dans une ville, on a sélectionné un échantillon de familles et l'on a compté le nombre d'adultes.

Nombre d'adultes	1	2	3	4	5	6	Total
Nombre de familles	22	36	34	22	18	8	140

On prendra 3 chiffres après la virgule.

2) Calculer la moyenne arithmétique, la variance et l'écart-type.

3) En utilisant le théorème de Tchebychev, déterminer l'intervalle qui contient plus de 75% de l'effectif.

4) Calculer le mode et étudier la symétrie de cette distribution.

Probabilité (10 points)

Soit l'expérience $\mathcal{E}$ = " tiré successivement deux boules d'une urne contenant 3 boules rouges et 4 bleues. **Après chaque tirage**, on remet dans l'urne la boule tirée ainsi qu'une autre boule de la même couleur"
On note C_1 la couleur de la boule du premier tirage et C_2 celle du second tirages.
Les boules dans l'urne sont indiscernables au toucher.
Les résultats seront donnés sous forme de fraction irréductible.

Partie I

1) Représenter cette situation par un arbre de probabilités.

2) Calculer $P(C_2 = \text{rouge})$.

3) Calculer $P(C_1 = \text{bleu}/C_2 = \text{rouge})$.

Parie II

On note X la variable aléatoire donnant le nombre de boules rouges dans l'urne après le deuxième tirage (donc après la remise des boules dans l'urne).

1) Donner les valeurs possibles pour X, c'est à dire $X(\Omega)$.

2) Donner la loi de X.

3) Calculer l'espérance mathématique de X.

Partie III

On répète l'expérience $\mathcal{E}$ n fois (lors de chaque premier tirage le nombre de boules rouges est 3 et de bleues est 4).

On définit la variable aléatoire Y= “nombre de boules bleus tiré lors du second tirage de chaque expérience.”

1) Quelle est la loi de probabilité de la variable aléatoire Y ?

2) Dans cette question, on suppose que $n = 150$.

a) Calculer l'espérance mathématique et la variance de la variable aléatoire Y.

b) Vérifier que la loi de la variable aléatoire Y peut être approchée par la loi gaussienne $\mathcal{N}(E(Y), V(Y))$.

c) On note T la variable aléatoire gaussienne $\mathcal{N}(E(Y), V(Y))$ et Z la variable aléatoire définie par : $Z = \dfrac{7T - 600}{30\sqrt{2}}$. Vérifier que la variable aléatoire Z suit la loi normale centrée réduite.

d) En utilisant une approximation, déterminer la probabilité qu'il y ait 90 boules bleues tirées.

On prendra deux chiffres après la virgule.

Table de la loi normale centrée réduite

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
$0,0$	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
$0,1$	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
$0,2$	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
$0,3$	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
$0,4$	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
$0,5$	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
$0,6$	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
$0,7$	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
$0,8$	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
$0,9$	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
$1,0$	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
$1,1$	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
$1,2$	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
$1,3$	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
$1,4$	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
$1,5$	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
$1,6$	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
$1,7$	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
$1,8$	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
$1,9$	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670

Correction Contrôle Final Décembre 2019

Statistique (10 points)

Exercice 1 : 1) le tableau des fréquences et des fréquences cumulées.

Classes	$]a,\, 26]$	$]26,\, 30]$	$]30,\, 34]$	$]34,\, 40]$	$]40,\, 46]$	$]46,\, 50]$	$]50,\, 56]$
Fréquences	0.14	0.16	f_3	f_4	0.16	0.12	0.04
F_k	0.14	0.3	$0.3+f_3$	$0.3+f_3+f_4$	$0.46+f_3+f_4$	$0.58+f_3+f_4$	1
c_k	$\frac{a+26}{2}$	28	32	37	43	48	53

2) La moyenne arithmétique est :

$$\overline{x} = \sum_{k=1}^{n} c_k\, f_k = \frac{a+26}{2} \times 0.14 + 28 \times 0.16 + 32 \times f_3 + 37 \times f_4 + 43 \times 0.16 + 48 \times 0.12 + 53 \times 0.04.$$

$35.62 = 21.06 + a \times 0.07 + 32 \times f_3 + 37 \times f_4 \Longrightarrow a \times 0.07 + 32 \times f_3 + 37 \times f_4 = 14.56$

3) a) La médiane appartient à la classe médiane, $m_e = 34.6 \in]34,\, 40]$. $m_e = x_i + (x_{i+1} - x_i)\dfrac{0.5 - F(x_i)}{F(x_{i+1}) - F(x_i)}$

$x_i = 34$; $F_i = 0.3 + f_3$ et $x_{i+1} = 40$; $F_{i+1} = 0.3 + f_3 + +f_4 \Longrightarrow$

$m_e = 34 + (40-34) \times \dfrac{0.5 - 0.3 - f_3}{f_4 + 0.3 + f_3 - f_3 - 0.3} \Longrightarrow (m_e - 34) \times f_4 = 6 \times (0.2 - f_3) \Longrightarrow 6 \times f_3 + 0.6 \times f_4 = 1.2$

b) $\displaystyle\sum_{k=1}^{7} f_k = 1 \Longrightarrow 0.14 + 0.16 + f_3 + f_4 + 0.16 + 0.12 + 0.04 = 1 \Longrightarrow f_3 + f_4 = 0.38$

$$\begin{cases} f_3 + f_4 = 0.38 \\ 6 \times f_3 + 0.6 \times f_4 = 1.2 \end{cases} \Longrightarrow \begin{cases} 0.6\,(f_3 + f_4) = 0.6 \times 0.38 \\ 6 \times f_3 + 0.6 \times f_4 = 1.2 \end{cases} \Longrightarrow 5.4\, f_3 = 1.2 - 0.228 = 0.972$$

D'où $f_3 = \dfrac{0.972}{5.4} = 0.18$ et $f_4 = 0.38 - f_3 = 0.38 - 0.18 = 0.2$.

4) D'après la question 2), on a $a \times 0.07 + 32 \times f_3 + 37 \times f_4 = 14.56$ et d'après la question 3) on a $f_3 = 0.18$ et $f_4 = 0.2$. D'où

$$a = \frac{14.56 - 32 \times 0.18 - 37 \times 0.2}{0.07} = 20$$

Exercice : 2 **1)** Le tableau de distribution des effectifs, .

Nombre d'adultes	1	2	3	4	5	6	Total
Nombre de familles	22	36	34	22	18	8	140
$n_k\, x_k$	22	72	102	88	90	48	422
$n_k\, x_k^2$	22	144	306	352	450	288	1562

2) $\overline{x} = \dfrac{1}{N}\displaystyle\sum_{k=1}^{n} n_k\, x_k = \dfrac{422}{140} = 3.014.$

$V(x) = \dfrac{1}{N}\displaystyle\sum_{k=1}^{n} n_k\, x_k^2 - \overline{x}^2 = \dfrac{1562}{140} - 3.014^2 = 11.157 - 9.084 = 2.073.$

$\sigma_x = \sqrt{V(x)} = \sqrt{2.073} = 1.440.$

3) D'après le théorème de Tchebychev,

$100 \times (1 - \dfrac{1}{k^2}) = 75 \Longrightarrow 25 = \dfrac{100}{k^2} \Longrightarrow k^2 = \dfrac{100}{25} \Longrightarrow k = \sqrt{4} = 2.$

D'où l'intervalle qui contient plus de 75% de l'effectif est

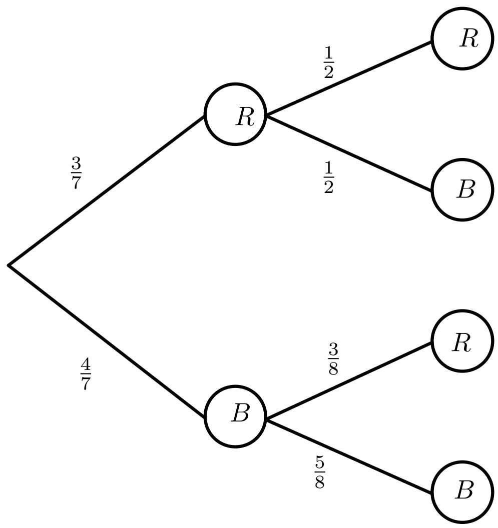
$$[\overline{x} - k\,\sigma_x,\, \overline{x} + k\,\sigma_x,] = [3.014 - 2 \times 1.44,\, 3.014 + 2 \times 1.44] = [0.134,\, 5,894]$$

4) On a $m_d = 2$. Le deuxième cœfficient d'asymétrie de Pearson $A_{P2} = \dfrac{\overline{x} - m_d}{\sigma_x} = \dfrac{3.014 - 2}{1.44} = 0.704 > 0$

D'où la distribution est asymétrique à droite.

Probabilité (10 points)

Partie 1 **1)**



2) On a $\{C_1 = R,\, C_1 = B\}$ forment un S.C.E. ou une partition, d'où d'après la formule des probabilités totales on a :

$P(C_2 = R) = P(C_2 = R/C_1 = R) \times P(C_1 = R) + P(C_2 = R/C_1 = B) \times P(C_1 = B) = \dfrac{3}{7} \times \dfrac{1}{2} + \dfrac{3}{8} \times \dfrac{4}{7} = \dfrac{3}{7}$

3) D'après la formule de Bayes, on a : $P(C_1 = B/C_2 = R) = \dfrac{P(C_2 = R/C_1 = B)P(C_1 = B)}{P(C_2 = R)} = \dfrac{(\frac{3}{8} \times \frac{4}{7})}{\frac{3}{7}} = \dfrac{1}{2}$.

Partie II

1) $X(\Omega) = \{3,\, 4,\, 5\}$.

2)

k	3	4	5
$P(X = k)$	$\frac{5}{14}$	$\frac{3}{7}$	$\frac{3}{14}$

$P(X = 3) = P(C_1 = B, C_2 = B) = P(C_2 = B/C_1 = B) \times P(C_1 = B) = \dfrac{5}{8} \times \dfrac{4}{7} = \dfrac{5}{14}$

$P(X = 4) = P(C_1 = B, C_2 = R) + P(C_1 = R, C_2 = B) = P(C_2 = R/C_1 = B) \times P(C_1 = B) + P(C_2 = B/C_1 = R) \times P(C_1 = R) = \dfrac{3}{8} \times \dfrac{4}{7} + \dfrac{1}{2} \times \dfrac{3}{7} = \dfrac{3}{7}$

$P(X = 5) = P(C_1 = R, C_2 = R) = P(C_2 = R/C_1 = R) \times P(C_1 = R) = \dfrac{1}{2} \times \dfrac{3}{7} = \dfrac{3}{14}$

3) $E(X) = \displaystyle\sum_{k \in X(\Omega)} k\, P(X = k) = 3 \times \frac{5}{14} + 4 \times \frac{3}{7} + 5 \times \frac{3}{14} = \frac{54}{14} = \frac{27}{7}$.

Partie III

1) Calculons la probabilité d'avoir une boule bleu au second tirage, $p = P(C_2 = B)$.

On a $\{C_1 = R,\, C_1 = B\}$ forment un S.C.E. ou une partition, d'où d'après la formule des probabilités totales on a :

$$\begin{aligned} P(C_2 = B) &= P(C_2 = B/C_1 = B)P(C_1 = B) \\ &\quad + P(C_2 = B/C_1 = R)P(C_1 = R) \\ &= \frac{5}{8} \times \frac{4}{7} + \frac{1}{2} \times \frac{3}{7} = \frac{4}{7} \end{aligned}$$

ou $p = P(C_2 = B) = 1 - P(C_2 = R) = 1 - \dfrac{3}{7} = \dfrac{4}{7}$.

On définit la variable aléatoire $Y =$ “ nombre de boules bleus.”
Y suit la loi binomiale de paramètre n et $p = \dfrac{4}{7}$, $Y \sim \mathcal{B}(n,\ \dfrac{4}{7})$.

En effet, n expériences identiques et indépendantes $\mathcal{E}_i$ sont effectuées (tiré successivement deux boules n fois), chaque expérience a deux issues à savoir “ la boule du second tirage est bleu.” avec une probabilité p et “la boule du second tirage n'est pas bleu.” avec une probabilité $1 - p$, qui correspond à n variables $\{Y_k\}_{k=1}^{n}$ de Bernoulli $Y_k = \begin{cases} 0 & \text{si} \quad \text{la boule du second tirage est bleu.} \\ 1 & \text{si} \quad \text{la boule du second tirage n'est pas bleu.} \end{cases}$

$P(Y_k = 1) = p = \dfrac{4}{7}$, $k = 1, \cdots, n$ et $Y = Y_1 + \cdots + Y_n$

2) Dans cette question, on suppose que $n = 150$.

a) $E(X) = n \times p = 150 \times \dfrac{4}{7} = \dfrac{600}{7}$ et $V(X) = n \times p \times (1-p) = 150 \times \dfrac{4}{7} \times (1 - \dfrac{4}{7}) = \dfrac{1800}{49}$.

b) $\begin{cases} n = 150 > 30 \\ np = \dfrac{600}{7} > 15 \\ np(1-p) = \dfrac{1800}{49} > 5 \end{cases}$

D'où La loi $\mathcal{B}(150,\ \dfrac{4}{7})$ peut être approchée par la loi normale $\mathcal{N}(\dfrac{600}{7},\ \dfrac{1800}{49})$.

c) Si on note Z_1 la loi de la v.a $\mathcal{N}(\dfrac{600}{7},\ \dfrac{1800}{49})$, alors on a $\dfrac{Z_1 - \frac{600}{7}}{\sqrt{\frac{1800}{49}}} = \dfrac{7 \times Z_1 - 600}{30\sqrt{2}} = Z$.

D'où $Z \sim \mathcal{N}(0,\ 1)$

d)

$$\begin{aligned}
P(Y = 90) &\approx P(90 - 0.5 \leq Z_1 \leq 90 + 0.5) \\
&= P(\frac{7 \times 89.5 - 600}{30\sqrt{2}} \leq \frac{7\,Z_1 - 600}{30\sqrt{2}} \leq \frac{7 \times 90.5 - 600}{30\sqrt{2}}) \\
&= P(\frac{26,5}{30\sqrt{2}} \leq Z \leq \frac{33.5}{30\sqrt{2}}) \\
&= P(Z \leq \frac{33,5}{30\sqrt{2}}) - P(Z \leq \frac{26.5}{30\sqrt{2}}) \\
&= P(Z \leq 0.79) - P(Z \leq 0.62) \\
&= 0.78524 - 0.73237 = 0,05287 = 0,05
\end{aligned}$$