Université de Picardie Jules Verne **2017-2018**
IAE Amiens

Licence mention Gestion parcours Management et Marketing Vente - Semestre 5
Statistiques appliquées - Partiel du 15/11/2017
Durée : 2h - Tout document interdit - Calculatrice autorisée

Exercice 1.

Le département Marketing de la société Rola-Cola souhaite étudier le comportement des consommateurs vis-à-vis de Rola-Cola et de la marque concurrente Cora-Cola. Une enquête a été réalisé sur des clients choisis au hasard dans un centre commercial et auxquels on a posé les questions suivantes :

Question 1 : Quelle boisson préférez-vous ? Rola-Cola (RC) ou Cora-Cola (CC).
Question 2 : Avez-vous déjà acheté Rola-Cola ? Oui (O) ou Non (N).
Question 3 : Quelle est votre réaction à la phrase "J'aime les boissons au Cola sucrées" ?
D'accord (DA), Pas sûr (PS) ou Pas d'accord (PA).
Question 4 : Combien de litres de boisson au Cola votre famille a-t-elle bu le mois dernier ?

Les réponses des clients ont été les suivantes :

n° client	Q1	Q2	Q3	Q4	n° client	Q1	Q2	Q3	Q4	n° client	Q1	Q2	Q3	Q4
1	RC	O	DA	1	11	RC	O	DA	6	21	RC	O	PS	6
2	CC	N	PA	6	12	CC	O	PA	2	22	RC	O	DA	8
3	RC	O	DA	4	13	RC	O	PS	4	23	CC	N	PA	2
4	CC	O	PS	7	14	RC	O	DA	7	24	CC	N	PS	17
5	RC	O	PS	2	15	RC	O	PS	11	25	RC	O	DA	5
6	CC	O	DA	12	16	RC	O	PA	4	26	CC	N	PS	14
7	CC	N	PA	6	17	CC	O	DA	3	27	RC	O	PA	15
8	RC	O	PA	2	18	CC	N	PS	10	28	RC	O	PA	5
9	RC	O	PS	9	19	RC	O	DA	5	29	CC	O	PS	9
10	RC	O	PS	5	20	CC	O	PA	7	30	RC	O	PS	4

1) Préciser la population étudiée, les variables étudiées et leur nature, la taille de l'échantillon.
2) Présenter les données obtenues avec la question Q3 dans un tableau modalités/effectifs, puis graphiquement de deux façons différentes.

Exercice 2.

1) Des sachets de café sont conditionnés à l'entreprise MDD par une ensacheuse. On teste l'efficacité de l'ensacheuse sur un échantillon de 300 sachets en mesurant leur masse. On obtient les résultats suivants :

Masse en grammes	$]246;250]$	$]250;252]$	$]252;254]$	$]254;256]$	$]256;258]$
Nombre de sachets	6	60	150	78	6

a) Préciser la population étudiée, la variable étudiée et sa nature, la taille de l'échantillon.
b) Construire le diagramme des effectifs correspondant aux résultats présentés ci-dessus.
c) Calculer les fréquences cumulées croissantes de cette distribution et tracer le polygone correspondant.
d) En déduire par lecture graphique, puis par une formule d'interpolation linéaire, la valeur de la médiane. Interpréter le résultat obtenu.
e) Donner la moyenne et l'écart-type de la distribution ; préciser les données utilisées pour le calcul.
f) La machine a besoin d'un réglage dès que l'une des trois conditions suivantes n'est pas vérifiée :
- le pourcentage de sachets de l'échantillon ayant une masse inférieure à 250 g est inférieur à 4% ;
- la masse moyenne des sachets de l'échantillon est comprise entre 252 g et 254 g ;
- l'écart type de la série de l'échantillon est inférieur à 1,5 g.

Cette machine doit-elle être réglée ? Justifier la réponse.

g) Le coefficient d'asymétrie de Fisher est $S \simeq -0,356$. Cela est-il cohérent avec les résultats précédents ? Justifier la réponse.

2) L'entreprise Café Grand Père commercialise les sachets de café. On admet que la variable aléatoire X qui représente la masse d'un sachet suit la loi normale de moyenne $\mu = 253$ et d'écart type $\sigma = 1,5$.

a) Calculer $P(X \leqslant 250)$.

b) Un sachet est vendu pour un poids de 250 g. Calculer la probabilité que sa masse soit d'au moins 250 g.

c) La société voudrait que le taux de sachet dont la masse est inférieure à 250 g soit inférieur à 1%, sans changer la valeur de l'écart type σ. Quelle devrait être la valeur (à 0,1 g près) de la moyenne μ ?

3) On suppose que dans la production très importante, 2% des sachets ont une masse inférieure à 250 g. Les sachets sont conditionnés par lots de 100 sachets choisis au hasard dans la production. Soit Y la variable aléatoire qui à chaque lot associe le nombre de sachets de masse inférieure à 250 g.

a) Justifier que Y suit une loi binomiale dont on donnera les paramètres.

b) En moyenne, combien y a-t-il de sachets dont la masse est inférieure à 250 g dans un lot ? Justifier.

c) Calculer la probabilité que tous les sachets aient une masse supérieure à 250 g.

d) Par quelle loi peut-on approcher la loi de Y ? Justifier la réponse. Utiliser cette loi pour calculer une valeur approchée de la probabilité demandée au 3) c).

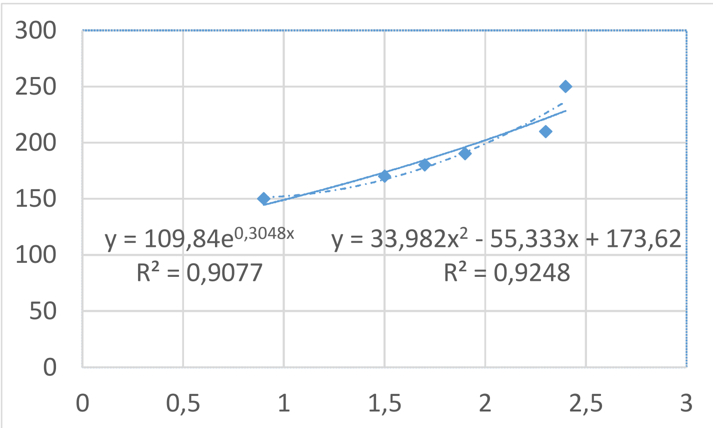
Exercice 3.

Le service marketing d'un centre commercial veut évaluer l'impact des frais engagés en publicité, par mois, sur le nombre de clients.

Pour cela, ce service s'appuie sur les données ci-dessous, relevées sur une période de 6 mois :

Frais publicitaires (en milliers d'euros) : x_i	1,9	2,4	1,5	0,9	2,3	1,7
Fréquentation (en milliers de clients) : y_i	190	250	170	150	210	180

A l'aide d'un tableur, on a obtenu le graphique et les résultats suivants :



Partie A - Ajustement affine

1) a) Considérant le nuage de points représentant la série statistique $(x_i;y_i)$, un ajustement affine vous semble-t-il approprié ? Justifier.

b) Donner le coefficient de corrélation linéaire entre x et y, arrondi au millième, de cette série. Le coefficient calculé confirme-t-il la réponse à la question précédente ? Justifier

2) a) Donner une équation de la droite de régression de y en x par la méthode des moindres carrés (les coefficients seront arrondis au millième).

b) En déduire une estimation de la fréquentation pour 4 000 euros de frais publicitaires engagés. Cette estimation est-elle fiable ? Justifier les réponses.

c) Le centre commercial décide d'engager 5 000 euros pour la campagne publicitaire pour un mois donné. A la fin de ce mois, on a dénombré 330 000 clients ayant fréquenté le site. Analyser ce résultat ?

3) a) En utilisant ce modèle, estimer le montant des frais publicitaires devant être engagés pour espérer 400 000 clients au cours d'un mois.

b) Etait-il cohérent d'utiliser ce modèle pour répondre à la question précédente ? Justifier la réponse en proposant une méthode alternative si nécessaire.

Partie B - Autres ajustements

La fonction « courbe de tendance » du tableur a donné les deux courbes représentées ci-dessus.

1) En considérant l'ajustement affine précédent et les deux ajustement présentés sur le graphique, lequel des trois ajustement est le meilleur ? Justifier la réponse.

2) Donner une équation de la courbe représentée en pointillés. Justifier la réponse.

Exercice 4.

La DGMIC (Direction générale des médias et des industries culturelles) a réalisé une étude auprès de 12 quotidiens d'information générale qui possèdent des applications numériques sur les trois supports que sont les tablettes, les smartphones et les ordinateurs.

Le taux de rebond désigne le pourcentage d'internautes qui sont entrés sur un site par une page web puis l'ont quitté sans consulter d'autres pages.

Cette étude donne les informations suivantes :

- 2 visites sur 5 se font depuis un smartphone, ces visites ayant un taux de rebond de 65 % ;
- 10% des visites se font depuis une tablette, ces visites ayant un taux de rebond de 53 % ;
- la moitié des visites ont lieu à partir d'un ordinateur, ces visites ayant un taux de rebond de 59 %.

On choisit au hasard un visiteur et on considère les évènements suivants :

- S : « Le visiteur utilise un smartphone »
- T : « Le visiteur utilise une tablette »
- O : « Le visiteur utilise un ordinateur »
- R : « Le visiteur quitte le site après avoir visité la première page »

1) a) Traduire les données de l'énoncé en termes de probabilité d'événements.

b) Calculer la probabilité que le visiteur utilise un smartphone et quitte le site après avoir visité la première page.

c) Calculer la probabilité que le visiteur quitte le site après avoir visité la première page. Justifier le calcul.

d) Calculer la probabilité que le visiteur utilise un ordinateur sachant qu'il a quitté le site après avoir consulté la première page.

2) a) Les évènements O et S sont-ils incompatibles ? indépendants ? Justifier la réponse.

b) Les évènements O et R sont-ils incompatibles ? indépendants ? Justifier la réponse.

Exercice 5.

1) Un supermarché dispose de plusieurs caisses. Un client qui se présente à une caisse doit attendre un certain temps avant d'être pris en charge par le caissier. On considère que ce temps d'attente X, exprimé en minute, est une variable aléatoire qui suit la loi uniforme sur l'intervalle [0; 12].

Rappel sur la loi Uniforme sur $[a,b]$.

$$F_X(x) = P(X \le x) = \begin{cases} 0 \text{ si } x < a \\ \dfrac{x-a}{b-a} \text{ si } x \in [a,b] \\ 1 \text{ si } x > b \end{cases}, \quad E(X) = \frac{a+b}{2} \text{ et } V(X) = \frac{(b-a)^2}{12}.$$

a) Calculer la probabilité qu'un client attende au moins 5 minutes avant d'être pris en charge.

b) Calculer la probabilité qu'un client attende entre 2 et 5 minutes avant d'être pris en charge.

c) Quel est le temps moyen d'attente à une caisse ?

d) Déterminer, en justifiant, le nombre réel x tel que $P(X \le x) = P(X \ge x)$. A quel indicateur statistique correspond le résultat obtenu ?

2) Le gérant du magasin décide de mettre à disposition des clients des caisses automatiques, de façon à réduire le temps d'attente pour les clients ayant un panier contenant peu d'articles.

Le temps d'attente Y, exprimé en minute, à chacune de ces caisses automatiques est modélisé par une variable aléatoire qui suit la loi exponentielle de paramètre $\theta = 0,5$.

Rappel sur la loi Exponentielle de paramètre θ, avec $\theta > 0$.

$$F_X(x) = P(X \le x) = \begin{cases} 1 - e^{-\theta x} \text{ si } x \ge 0 \\ 0 \text{ si } x < 0 \end{cases}, \quad E(X) = \frac{1}{\theta}, \quad Var(X) = \frac{1}{\theta^2}.$$

a) Donner le temps d'attente moyen à une caisse automatique.

b) Calculer la probabilité que le temps d'attente à une caisse automatique soit supérieur à 2 minutes.

c) Calculer la probabilité que le temps d'attente à une caisse automatique soit compris entre 30 secondes et 3 minutes.

3) Ces caisses automatiques tombent souvent en panne. On dispose des informations suivantes :

- le nombre de caisses automatiques est égal à 10 ;
- la probabilité qu'une caisse automatique tombe en panne pendant une journée donnée est égale à 0,1 ;
- une panne constatée sur une caisse automatique n'influence pas les autres caisses automatiques.

Soit Z la variable aléatoire correspondant au nombre de caisses automatiques qui tombent en panne pendant une journée donnée.

a) Quelle est la loi de probabilité suivie par Z ? Préciser ses paramètres.

b) Calculer la probabilité qu'aucune caisse automatique ne tombe en panne pendant une journée donnée.

TABLE 1

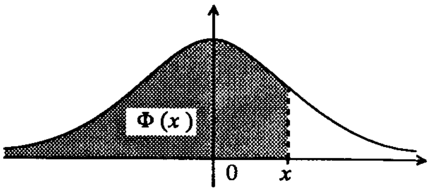
Fonction de répartition de la loi normale réduite

Si U suit la loi normale réduite, pour $x \geqslant 0$, la table donne la valeur :

$\phi(x) = P(U \leqslant x)$.

La valeur x s'obtient par addition des nombres inscrits en marge.
Pour $x < 0$, on a :

$\phi(x) = 1 - \phi(-x)$.



x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,500 0	0,504 0	0,508 0	0,512 0	0,516 0	0,519 9	0,523 9	0,527 9	0,531 9	0,535 9
0,1	0,539 8	0,543 8	0,547 8	0,551 7	0,555 7	0,559 6	0,563 6	0,567 5	0,571 4	0,575 3
0,2	0,579 3	0,583 2	0,587 1	0,591 0	0,594 8	0,598 7	0,602 6	0,606 4	0,610 3	0,614 1
0,3	0,617 9	0,621 7	0,625 5	0,629 3	0,633 1	0,636 8	0,640 6	0,644 3	0,648 0	0,651 7
0,4	0,655 4	0,659 1	0,662 8	0,666 4	0,670 0	0,673 6	0,677 2	0,680 8	0,684 4	0,687 9
0,5	0,691 5	0,695 0	0,698 5	0,701 9	0,705 4	0,708 8	0,712 3	0,715 7	0,719 0	0,722 4
0,6	0,725 7	0,729 1	0,732 4	0,735 7	0,738 9	0,742 2	0,745 4	0,748 6	0,751 7	0,754 9
0,7	0,758 0	0,761 1	0,764 2	0,767 3	0,770 4	0,773 4	0,776 4	0,779 4	0,782 3	0,785 2
0,8	0,788 1	0,791 0	0,793 9	0,796 7	0,799 5	0,802 3	0,805 1	0,807 8	0,810 6	0,813 3
0,9	0,815 9	0,818 6	0,821 2	0,823 8	0,826 4	0,828 9	0,831 5	0,834 0	0,836 5	0,838 9
1,0	0,841 3	0,843 8	0,846 1	0,848 5	0,850 8	0,853 1	0,855 4	0,857 7	0,859 9	0,862 1
1,1	0,864 3	0,866 5	0,868 6	0,870 8	0,872 9	0,874 9	0,877 0	0,879 0	0,881 0	0,883 0
1,2	0,884 9	0,886 9	0,888 8	0,890 7	0,892 5	0,894 4	0,896 2	0,898 0	0,899 7	0,901 5
1,3	0,903 2	0,904 9	0,906 6	0,908 2	0,909 9	0,911 5	0,913 1	0,914 7	0,916 2	0,917 7
1,4	0,919 2	0,920 7	0,922 2	0,923 6	0,925 1	0,926 5	0,927 9	0,929 2	0,930 6	0,931 9
1,5	0,933 2	0,934 5	0,935 7	0,937 0	0,938 2	0,939 4	0,940 6	0,941 8	0,942 9	0,944 1
1,6	0,945 2	0,946 3	0,947 4	0,948 4	0,949 5	0,950 5	0,951 5	0,952 5	0,953 5	0,954 5
1,7	0,955 4	0,956 4	0,957 3	0,958 2	0,959 1	0,959 9	0,960 8	0,961 6	0,962 5	0,963 3
1,8	0,964 1	0,964 9	0,965 6	0,966 4	0,967 1	0,967 8	0,968 6	0,969 3	0,969 9	0,970 6
1,9	0,971 3	0,971 9	0,972 6	0,973 2	0,973 8	0,974 4	0,975 0	0,975 6	0,976 1	0,976 7
2,0	0,977 2	0,977 8	0,978 3	0,978 8	0,979 3	0,979 8	0,980 3	0,980 8	0,981 2	0,981 7
2,1	0,982 1	0,982 6	0,983 0	0,983 4	0,983 8	0,984 2	0,984 6	0,985 0	0,985 4	0,985 7
2,2	0,986 1	0,986 4	0,986 8	0,987 1	0,987 5	0,987 8	0,988 1	0,988 4	0,988 7	0,989 0
2,3	0,989 3	0,989 6	0,989 8	0,990 1	0,990 4	0,990 6	0,990 9	0,991 1	0,991 3	0,991 6
2,4	0,991 8	0,992 0	0,992 2	0,992 5	0,992 7	0,992 9	0,993 1	0,993 2	0,993 4	0,993 6
2,5	0,993 8	0,994 0	0,994 1	0,994 3	0,994 5	0,994 6	0,994 8	0,994 9	0,995 1	0,995 2
2,6	0,995 3	0,995 5	0,995 6	0,995 7	0,995 9	0,996 0	0,996 1	0,996 2	0,996 3	0,996 4
2,7	0,996 5	0,996 6	0,996 7	0,996 8	0,996 9	0,997 0	0,997 1	0,997 2	0,997 3	0,997 4
2,8	0,997 4	0,997 5	0,997 6	0,997 7	0,997 7	0,997 8	0,997 9	0,997 9	0,998 0	0,998 1
2,9	0,998 1	0,998 2	0,998 2	0,998 3	0,998 4	0,998 4	0,998 5	0,998 5	0,998 6	0,998 6

Quelques valeurs supplémentaires

x	$\phi(x)$
2,99	0,998605
3,00	0,998650
3,50	0,999767
4,00	0,999968
4,50	0,999997
5,00	1,000000