Université de Picardie Jules Verne **2010-2011**
UFR des Sciences

Licence mention Mathématiques et mention Informatique parcours MIAGE - Semestre 3
Statistique et Probabilités
Partiel du lundi 8 novembre 2010

Durée 2h00
Tout document interdit - Calculatrices autorisées

Exercice 1.

Pour la commune d'Amiens, en 1999, on a observé les données suivantes :

Nombre de résidences principales	57661
Nombre de logements occasionnels	722
Nombre de résidences secondaires	198
Nombre total de logements vacants	4597
Nombre total de logements	63178

Tableau 1

Age	Hommes	Femmes	Total
de 0 à 19 ans	17600	17430	35030
de 20 à 24 ans	7437	8706	16143
de 25 à 39 ans	15956	16477	32433
de 40 à 59 ans	14141	15766	29907
de 60 à 74 ans	5608	7769	13377
de plus de 75 ans	2862	5697	8559
Total	63604	71845	135449

Tableau 2

Source INSEE : Recensement de la population de 1999

1) On s'intéresse d'abord au tableau 1.
   a) Préciser la population étudiée, la variable étudiée et sa nature.
   b) Représenter graphiquement les données, de deux façons différentes.

2) On s'intéresse maintenant au tableau 2. Pour la présentation des différents calculs effectués dans les questions suivantes, on construira dès le début un unique tableau présentant l'ensemble des résultats demandés ou utiles.

a) Préciser la population étudiée, la variable étudiée et sa nature. Le type de présentation des données correspond-il à la nature de la variable étudiée ?

b) Présenter les données concernant l'ensemble de la population (sans distinguer hommes et femmes) dans un tableau classes/effectifs. On pourra considérer que les différentes classes d'âge du tableau 2 correspondent aux intervalles $[0;20[$, $[20;25[$, $[25;40[$, $[40;60[$, $[60;75[$ et $[75;100]$. Par exemple, les individus agés de 0 à 19 ans sont ceux qui ont moins de 20 ans, ce qui donne l'intervalle $[0;20[$; pour les plus de 75 ans, on considère que l'âge maximum est 100 ans.

c) Représenter graphiquement les résultats présentés dans le tableau construit au b). Préciser, en justifiant, la classe modale.

d) Calculer les fréquences cumulées (croissantes) de la distribution et tracer le polygone des fréquences cumulées.

e) En déduire par lecture graphique, puis par une formule d'interpolation linéaire, la valeur de la médiane et des quartiles de la distribution. Interpréter les résultats obtenus. Construire la boîte à moustaches correspondante.

f) A partir du tableau contruit au b), donner une valeur approchée à 10^{-2} près de la moyenne et l'écart-type de la distribution. Préciser les données à partir desquelles ces valeurs ont été obtenues.

Exercice 2.

Le tableau ci-dessous donne l'évolution, par période de 5 ans, de la population (en millions d'habitants) de l'Allemagne ; il s'agit de la population globale des deux Allemagnes (RDA et RFA) de 1958 à 1973, puis de la population de l'Allemagne réunifiée de 1993 à 2008.

Année	1958	1963	1968	1973	1993	1998	2003	2008
Rang de l'année : x_i	1	2	3	4	8	9	10	11
Population : y_i	71.5	74.4	77.0	78.8	81.0	82.1	82.5	82.2

1) Représenter graphiquement la série statistique (x_i, y_i).

2) On commence par chercher un ajustement affine.

a) Donner une équation de la droite de régression de y en x. Donner le coefficient de corrélation linéaire entre x et y. Interpréter le résultat obtenu.

b) En déduire une estimation de la population de l'Allemagne en 2018. L'estimation est-elle fiable ? Justifier votre réponse.

3) On cherche maintenant un ajustement de type logarithmique ; autrement dit à modéliser le phénomène étudié par une relation du type $y = a\ln x + b$. Pour cela, on considère la nouvelle variable $z = \ln x$.

a) Effectuer une régression permettant d'obtenir les coefficients a et b. Préciser les variables considérées et donner le coefficient de corrélation linéaire correspondant. Interpréter le résultat obtenu.

b) En déduire une estimation de la population de l'Allemagne en 2018. L'estimation est-elle fiable ?

4) Comparer les deux estimations des 2)b) et 3)b), et commenter les résultats obtenus. Lequel des deux ajustements est le meilleur ?

Exercice 3

L'examen du code de la route est composé de 40 questions numérotées de 1 à 40. Pour chaque question, quatre réponses sont proposées, dont une seule est bonne. Un candidat est déclaré reçu à l'examen s'il donne au moins 35 bonnes réponses.

1) Un candidat, totalement ignorant du code de la route, décide de tenter sa chance en donnant, pour chaque question, une réponse choisie au hasard parmi les quatre réponses proposées. On désigne par X la variable aléatoire égale au nombre de bonnes réponses données par le candidat.

a) Quelle est la loi de probabilité de la variable aléatoire X ? Justifier votre réponse.

b) Quelle est la probabilité que le candidat donne au moins une bonne réponse ?

c) Quelle est la probabilité que le candidat soit reçu à l'examen ?

2) Le modèle précédent est un peu trop simpliste ! En effet, le candidat n'est pas complêtement ignorant, et il connaît la réponse à certaines questions. Ainsi, pour chaque question, il y a deux possibilités :

- s'il connait la réponse, alors il donne la bonne réponse à coup sûr ;
- s'il ignore la réponse, alors il choisit au hasard une des quatre réponses proposées.

On suppose que pour une question choisie au hasard, la probabilité que le candidat connaisse la réponse est égale à p, avec $p \in]0,1[$.

a) Traduire les données de l'énoncé de la question 2) en termes de probabilité d'événements.

b) Calculer, en fonction de p, la probabilité que le candidat donne la bonne réponse à la question ; on notera $f(p)$ cette probabilité, et on pourra vérifier que $f(p) = \dfrac{3}{4}p + \dfrac{1}{4}$.

c) Lorsque le candidat donne la bonne réponse, quelle est la probabilité qu'il ait répondu au hasard ?

3) On considère toujours le modèle de la question 2) et on désigne par Y la variable aléatoire égale au nombre de bonnes réponses données par le candidat.

a) Quelle est, en fonction de p, la loi de probabilité de la variable aléatoire Y ?

b) Donner, en fonction de p, l'espérance mathématique $E(Y)$ de Y.

c) Quelle condition doit vérifier p pour qu'en moyenne, le candidat donne au moins 35 bonnes réponses ?

Exercice 4

1) Pour rendre visite à son ami Alexis, Benjamin emprunte une navette maritime qui effectue une liaison entre son île de résidence et le continent. La variable aléatoire X mesurant la durée du trajet suit la loi normale de moyenne 315 minutes (soit 5 heures 15 minutes) et d'écart-type 20 minutes.

Quelle est la probabilité que le trajet dure plus de 5 heures ?

2) Alexis va chercher son ami au port et prend sa voiture depuis son domicile. La variable aléatoire Y mesurant la durée du trajet suit la loi normale de moyenne 30 minutes et d'écart-type 12 minutes.

Quelle est la probabilité que le trajet dure moins de 40 minutes ?

3) Benjamin a quitté son île à 13h30 et Alexis a quitté son domicile à 17h50. En supposant que les durées des trajets des deux amis sont indépendantes, quelle est la probabilité qu'Alexis constate à 18h30 que son ami n'est toujours pas arrivé au port ?

TABLE 1

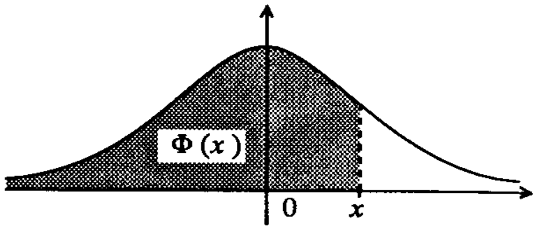
Fonction de répartition de la loi normale réduite

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