Université Lyon I **Année 2003-2004**
Licence B.O.

Examen de BMS

Durée : 1 heure 30

Documents Autorisés

Exercice 1

A la roulette, en misant *s* euros sur un numéro allant de 0 à 36, un joueur gagne 36*s* euros (sa mise initiale + 35 fois la mise payée par la banque) si bien sûr le numéro a la chance de sortir. Dans tous les autres cas, le joueur perd sa mise.

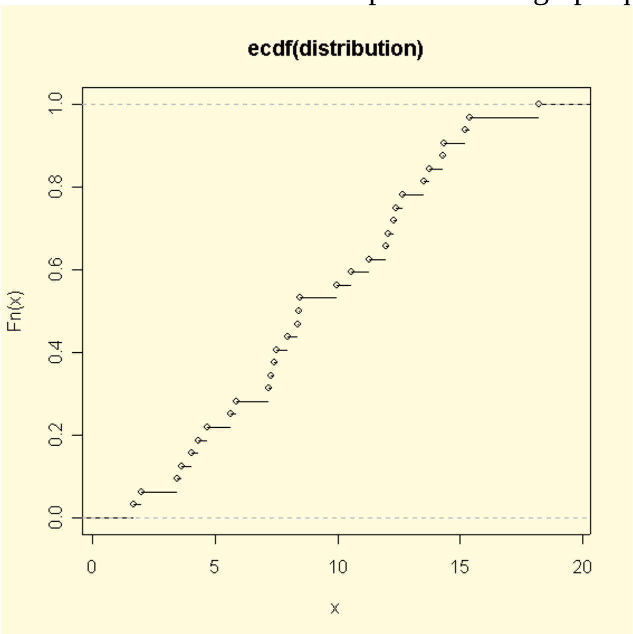
1) Quelle est la probabilité de sortie d'un numéro ?
2) Ecrire l'espérance de gain pour une mise de *s* euros.
3) On dit qu'un jeu est équitable si l'espérance est nulle. Quelle somme doit-on jouer pour que le jeu soit équitable ?

Exercice 2

Soient 15 lois normales centrées réduites. Que pouvez-vous dire de la somme des carrés de ces 15 lois ?

Exercice 3

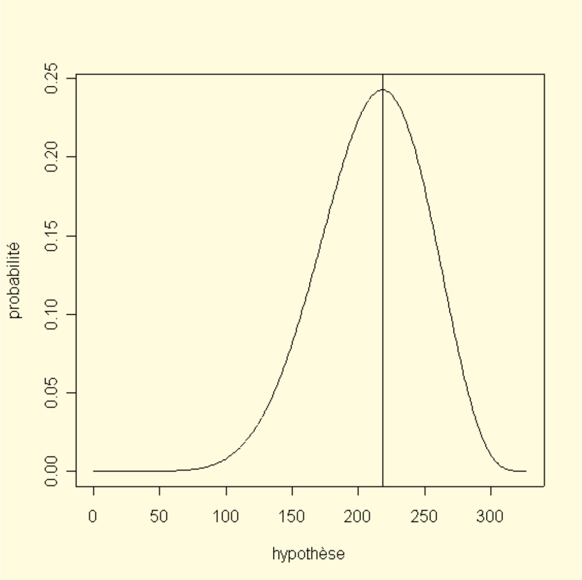
On étudie les notes de 32 étudiants. On réalise la représentation graphique ci-dessous.



1) Donner un titre à cette représentation graphique.
2) Donner la valeur de x associée à Fn(x)=0.8
3) Que signifie-t-elle ?

Exercice 4

Le directeur d'une école primaire décide de faire plaisir aux élèves en organisant une séance de cinéma. Il se demande si un film à conotation historique les intéresserait. L'école comprend 327 élèves et le directeur choisit d'interroger au hasard ceux qu'il croise dans la cours de récréation. Il pose la question à 12 élèves : 8 sont favorables à cette proposition.



Le directeur construit la représentation graphique de la valeur de la vraisemblance en fonction de l'hypothèse mais étourdi, il oublie de l'expliciter.

1) Que signifie ce mot « hypothèse » ?
2) Quelle probabilité a-t-il calculé sur l'axe vertical ?
3) Quelle est l'hypothèse la plus vraisemblable ? Est-elle juste ou fausse ?

Exercice 5

On considère n lois de Poisson X_i indépendantes et de même paramètre λ. Que pouvez-vous dire de $Z_n = \frac{\sum_{i=1}^{n} X_i}{n}$? de sa convergence ?

Exercice 6

On a enregistré le nombre de jours qui séparent deux tremblements de terre graves sur la surface de la terre. Un tremblement de terre est grave si sa magnitude est au moins égale à 7.5 sur l'échelle de Richter ou s'il a causé la mort d'au moins 1000 personnes. L'enregistrement couvre une période de 27300 jours entre 1902 et 1977. Il y a eu 63 tremblements de terre et donc 62 valeurs du temps d'attente du suivant.

```
> delai Temps d'attente du suivant
[1] 840 157 145 44 33 121 150 280 434 736 584 887 263 1901 695
[16] 294 562 721 76 710 46 402 194 759 319 460 40 1336 335 1354
[31] 454 36 667 40 556 99 304 375 567 139 780 203 436 30 384
[46] 129 9 209 599 83 832 328 246 1617 638 937 735 38 365 92
[61] 82 220
```

[Source : Hand, D.J., Daly, F., Lunn, A.D., McConway, K.J. & Ostrowski, E. (1994) A handbook of small data sets. Chapman & Hall, London. 1-458. p. 203.]

6a

La moyenne de la statistique `delai` vaut 437.21. On fait l'hypothèse que le temps d'attente d'une catastrophe aléatoire suit une loi exponentielle de paramètre α. Quelle est l'estimation de α au maximum de vraisemblance ?

6b

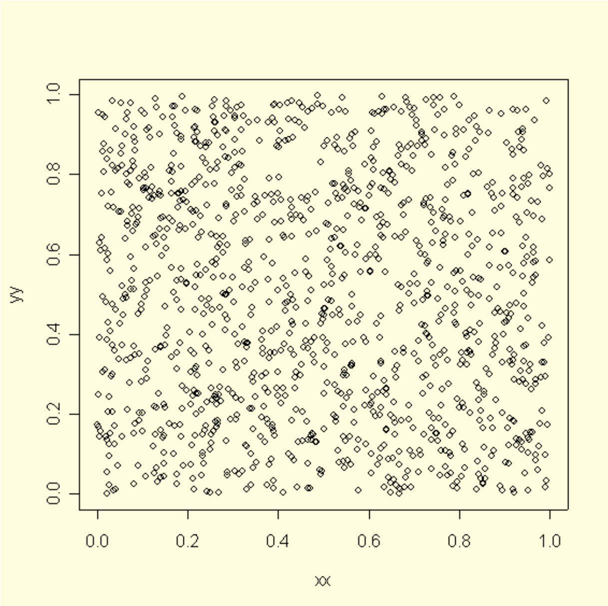
Les valeurs observées de `delai` sont rangées en 5 classes. Les fréquences, les probabilités et les effectifs théoriques associés à la loi exponentielle sont calculés.

Classes	]0,100]	]100,200]	]200,400]	]400,800]	>800	Total
Effectifs observés	14	7	14	19	8	62
Fréquences	0.226	0.113	0.226	0.306	0.129	1
Probabilités	0.204	0.163	0.232	0.240	0.160	1
Effectifs théoriques	12.7	10.1	14.4	14.9	9.9	62
(Obs-The)^2/The	0.138	0.943	0.011	1.136	0.381	2.611

La variable étudiée suit-elle une loi exponentielle ?

Exercice 7

On donne le nuage de points ci-dessous ainsi que le résultat du test du coefficient de corrélation.




```
> cor.test(xx,yy)

        Pearson's product-moment correlation

data:  xx and yy
t = -2.5013, df = 1198, p-value = 0.01250
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.12814683 -0.01555374
sample estimates:
        cor
-0.07207992
```

Qu'en pensez-vous ?

Exercice 8

Dans l'aide de R, on trouve un exemple de test de Wilcoxon provenant du livre de Hollander et Wolfe (1983). Il s'agit de mesurer l'état d'anxiété et de dépression de 9 patients selon l'échelle de dépression de Hamilton. Cet état a été mesuré lors d'une première visite. Un tranquillisant a été prescrit au patient. Puis l'état a été à nouveau testé lors de la deuxième visite.

```
x <- c(1.83,  0.50,  1.62,  2.48, 1.68, 1.88, 1.55, 3.06, 1.30)
y <- c(0.878, 0.647, 0.598, 2.05, 1.06, 1.29, 1.06, 3.14, 1.29)
> y-x
[1] -0.952  0.147 -1.022 -0.430 -0.620 -0.590 -0.490  0.080 -0.010

>  wilcox.test(y - x, alternative = "less")
        Wilcoxon signed rank test
data:  y - x
V = 5, p-value = 0.01953
```

1) Le test proposé est-il apparié ou non apparié ?
2) Expliquer comment `V=5` a été obtenu.
3) Que signifie l'alternative « less » pour le calcul de la p-value ?
4) Si par défaut, il y avait eu l'alternative « two-sided », quelle aurait été la valeur de p ?
5) Ecrire une autre manière de réaliser ce test.

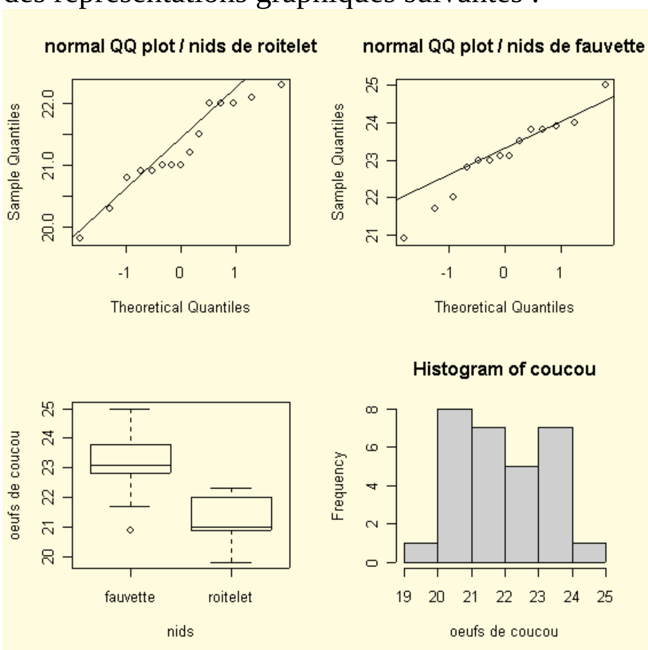
Exercice 9

Le biologiste Latter[1] donne la longueur (en mm) des œufs de coucou trouvés dans les nids de deux espèces d'oiseaux : le roitelet (nids de petite taille), la fauvette (nids de grande taille).

```
roitelet=c(19.8,22.1,21.5,20.9,22,21,22.3,21,20.3,20.9,22,22,20.8,21.2,21)
fauvette=c(22,23.9,20.9,23.8,25,24,23.8,21.7,22.8,23.1,23.5,23,23,23.1)
```

9a

On donne l'ensemble des représentations graphiques suivantes :



[1] In Biometrika, d'après Couty, Debord et Fredon « Probabilités et Statistiques pour biologistes », coll U-Flash, Armand Colin, 1990

Donner le type d'information générale que chaque graphique apporte à une étude statistique et à cette étude en particulier.

9b

La variable longueur des œufs de coucou est supposée gaussienne et les variances des populations sont égales.

```
> summary(roitelet)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  19.80   20.90   21.00   21.25   22.00   22.30 
> summary(fauvette)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  20.90   22.85   23.10   23.11   23.80   25.00 
> var(roitelet);var(fauvette)
[1] 0.5155238
[1] 1.101319

> var.test(roitelet,fauvette)
        F test to compare two variances
data:  roitelet and fauvette 
F = 0.4681, num df = 14, denom df = 13, p-value = 0.1722
alternative hypothesis: true ratio of variances is not equal to 1 
95 percent confidence interval:
 0.1518881 1.4098580 
sample estimates:
ratio of variances 
         0.4680969
```

On donne les deux résultats ci-dessous :

```
> t.test(roitelet,fauvette,var.equal=T)
        Two Sample t-test
data:  roitelet and fauvette 
t = -5.6074, df = 27, p-value = 5.988e-06
alternative hypothesis: true difference in means is not equal to 0 
95 percent confidence interval:
 -2.541904 -1.180001 
sample estimates:
mean of x mean of y 
 21.25333  23.11429 

> coucou=c(roitelet,fauvette)
> nids=as.factor(rep(c("roitelet","fauvette"),c(15,14))
> anova(lm(coucou~nids))
Analysis of Variance Table
Response: coucou
          Df  Sum Sq Mean Sq F value    Pr(>F)    
nids       1 25.0779 25.0779  31.443 5.988e-06 ***
Residuals 27 21.5345  0.7976
```

Quels sont les objectifs de ces deux tests ?
Y-a-t-il une relation entre les deux ? Si oui, laquelle.

9c

A l'aide des informations ci-dessus, répondre à la question : le coucou adapte-t-il la taille de ces œufs à la taille des nids dans lequel il pond ?