Question 1 **(6 points)**

On a recueilli 11 observations de sols pour lesquels on a mesuré la masse volumique (X) et la porosité (Y). On a obtenu les résultats suivants : $\sum_{i=1}^{11} x_i = 19.2$; $\sum_{i=1}^{11} y_i = 3,3$; $\sum_{i=1}^{11} x_i^2 = 39,8$; $\sum_{i=1}^{11} y_i^2 = 1,2$; $\sum_{i=1}^{11} x_i y_i = 5,1$

a) Calculez le coefficient de corrélation entre porosité et masse volumique.
b) Faites le test (bilatéral, niveau α =5%) pour vérifier si la corrélation est significativement différente de 0. Supposez que les variables X et Y suivent une distribution binormale. (Si vous n'avez pu calculer la valeur de r_{xy} en a) prenez r_{xy}= -0,6 comme valeur substitut).

Question 2 **(6 points)**

Sur une certaine route à sens unique, le temps moyen entre le passage de deux véhicules est de 15 s. Le temps entre deux véhicules suit une loi exponentielle.

a) Quel est la valeur du paramètre λ de la loi exponentielle?
b) Quelle est la probabilité qu'il s'écoule une minute ou plus entre le passage de deux véhicules?
c) Quelle est la distribution exacte du nombre de véhicules croisant un point de contrôle en une heure?
d) Utilisant une approximation normale, calculez la probabilité que le nombre de véhicules croisant le point de contrôle en une heure soit ≥ 250.

Question 3 **(6 points)**

Soit la relation C=$\frac{1}{2}$MV^2 où C est l'énergie cinétique (en kg(km/h)^2). On étudie l'énergie absorbée lors de collisions de véhicules avec des objets fixes. On considère que la masse M (en kg) et la vitesse V (en km/h) sont indépendantes et distribuées chacune suivant une loi lognormale. Les paramètres logarithmiques sont :

	μ	σ^2
M	7,25	0,10
V	3,98	0,22

a) Quelle est la distribution de C (type et paramètres)?
b) Quelle est l'énergie moyenne (i.e. E[C]) absorbée lors d'une collision?

Question 4 (8 points)

Soit les v.a. $X_1, X_2, \ldots X_{21}$, des v.a. aléatoires indépendantes, chacune distribuée selon une N(0,2).

a) Soit $\overline{X} = \frac{1}{21}\sum_{i=1}^{21} X_i$. Quelle est $P(\overline{X} < 0{,}3)$?

b) Calculez $P(\sum_{i=2}^{21} X_i^2 < 56{,}82)$

c) Trouvez « c » tel que $P(X_1/\sqrt{\sum_{i=2}^{21} X_i^2} > c) = 0{,}05$

d) Trouvez « d » tel que $P(\sum_{i=1}^{5} X_i^2 / \sum_{i=6}^{21} X_i^2 > d) = 0{,}05$

e) Pourquoi ne peut-on pas calculer facilement $P(\sum_{i=1}^{5} X_i^2 / \sum_{i=1}^{16} X_i^2 > 0{,}8)$?

Question 5 (6 points)

Lors de la construction d'un barrage en terre, on effectue un contrôle de qualité sur le taux de compaction du sol (mesuré en % par rapport au maximum Proctor). Le taux de compaction est mesuré in-situ, lors de la construction, à tous les 5 m. Le barrage est considéré parfaitement fiable si la proportion des matériaux du barrage montrant une compaction inférieure à 96% Proctor ne dépasse pas 0,02. On a effectué 1000 mesures de compaction et la proportion montrant une compaction inférieure à 96% Proctor est 0,018. Les mesures de compaction sont considérées indépendantes.

a) Construisez l'intervalle de confiance unilatéral à droite de niveau 5% (i.e. trouvez « a » tel que $P(p<a)=0{,}95$).
b) Doit-on considérer le barrage parfaitement fiable selon le résultat obtenu en a) ? Justifiez.

Question 6 (6 points)

Le tableau suivant montre les résultats d'une série d'échantillonnages de remblais sous les dalles de béton des résidences de plus de 20 ans en vu de déterminer l'IPPG (indice pétrographique du potentiel de gonflement) du remblai. On a aussi noté la présence de dommages à la dalle de béton selon deux niveaux d'intensité. Faites le test (niveau 5%) permettant de vérifier si la valeur d'IPPG d'un remblai est un bon prédicteur de dommages potentiels à la dalle de béton des résidences (i.e. les variables IPPG et « dommages observés» sont-elles liées significativement?).

		IPPG			
		IPPG ≤ 10	10<IPPG ≤ 40	40<IPPG	Total
Dommages observés à la dalle	aucun ou légers	30	20	10	60
	importants	10	30	30	70
	Total	40	50	40	130

Question 7 (6 points)

Trois routes d'accès (A,B et C) mènent à une autoroute. Le nombre de véhicules empruntant ces voies d'accès au cours d'une journée type peut être considéré comme étant distribué suivant une loi multinormale de paramètres (avec dans l'ordre A, B et C):

$$\mu = \begin{bmatrix} 5 \\ 10 \\ 15 \end{bmatrix} \text{(en milliers)} ; \ \Sigma = \begin{bmatrix} 1 & 1{,}5 & 1 \\ 1{,}5 & 4 & 2 \\ 1 & 2 & 4 \end{bmatrix} \text{(en milliers}^2\text{)}$$

a) Quelle est la probabilité que le nombre total de véhicules empruntant ces trois voies d'accès soit supérieur à 40 000?
b) Quelle est la corrélation entre le nombre de véhicules utilisant les voies d'accès A et B?

Question 8 (6 points)

On détermine annuellement le débit maximal d'une certaine rivière. Les débits maximaux obtenus pour 16 années consécutives sont considérés indépendants. On obtient $\overline{x}$=200 m^3/s. On suppose que le débit maximal est approximativement distribué comme une loi normale ayant $\sigma = 50 m^3/s$.

a) Doit-on rejeter l'hypothèse H_0 $\mu = 210 m^3/s$ en faveur de l'alternative H_1 $\mu < 210 m^3/s$ (test unilatéral niveau $\alpha = 5\%$)?
b) Quelle est la puissance du test face à l'alternative H_1 $\mu = 180 m^3/s$?
c) Quelle taille d'échantillon aurait-il fallu pour assurer une puissance de 90% en b)?

Corrigé

Question 1 :

a) On calcule $\overline{x} = 1{,}745$, $\overline{y} = 0{,}3$; s_x^2=0,629; s_y^2=0,021; s_{xy}= -0,066 et r_{xy}= -0,57

b) On obtient $t_{calcule}$ = -0,57*3/(1-0.57^2)^{0,5}= -2,08>-2,26 donc on accepte H_0 , i.e. le coefficient de corrélation n'est pas significatif.

Question 2 :

a) $\lambda = 1/15$

b) P(T>60) = 1-(1-e$^{-60/15}$)=1,8%

c) Loi de Poisson de paramètre $\lambda = 1/15 * 3600 = 240$

d) X : nombre de véhicules en une heure : N(240,240)-> P(X>250)=P(Z>10/240$^{0.5}$)=P(Z>0,645)=0,2595

Question 3 :

a) C est aussi distribuée comme une loi lognormale. ln(C)=ln(0,5)+ln(M)+2ln(V)

E[ln(C)]=ln(0,5)+7,25+2*3,98=14,517

Var(ln(C))=0,10+4*0,22=0,98

b) E[C]=e$^{14{,}517+0{,}98/2}$=3,29 x 10^{6}kg(km/h)^2

Question 4 :

a) $\overline{X}$ ~N(0,2/21) ->P($\overline{X}$ <0,3)=P(Z<0,3(21)^{0,5}/2^{0,5}=P(Z<0.9721)=0,8345

b) $\sum_{i=2}^{21} X_i^2 /2 \sim \chi^2_{20}$. P($\sum_{i=2}^{21} X_i^2$ <56,82)=P($\sum_{i=2}^{21} X_i^2$ /2<28,41)=1-0,10=0,9

c) $X_1 / \sqrt{\left(\sum_{i=2}^{21} X_i^2\right)/20} \sim t_{20}$. P($X_1 / \sqrt{\left(\sum_{i=2}^{21} X_i^2\right)}$ >c)= P($X_1(20)^{0,5} / \sqrt{\left(\sum_{i=2}^{21} X_i^2\right)}$ >c(20)^{0,5})=P(t_{20}>c(20)^{0,5})

On trouve de la table t : c(20)^{0,5}= 1,725 ->c=0,3857

d) En divisant les sommes de carrés par leurs degrés de liberté respectifs (5 et 16) on obtient une distribution Fisher.

$\dfrac{\sum_{i=1}^{5} X_i^2 /5}{\sum_{i=6}^{21} X_i^2 /16} \sim F_{5,16}$. P($P(\sum_{i=1}^{5} X_i^2 / \sum_{i=6}^{21} X_i^2 > d)$ = P($F_{5,16}$>d 16/5) On lit $F_{5,16,5\%}$=2,85=d 16/5. d=0,89

e) les χ^2_5 et χ^2_{16} ne sont pas indépendantes puisque X_1 à X_5 apparaissent dans les deux termes. La distribution n'est donc pas Fisher.

Question 5 :

a) L'intervalle de confiance est p< $\hat{p} + z_{0,05}(\hat{p}(1-\hat{p})/n)^{0,5}$ ->p<0,018+1,645(0,018(1-0,018)/1000)$^{0.5}$ donc p<0,0249

b) Comme la valeur de l'intervalle inclut la valeur 2, on ne peut considérer le barrage totalement fiable.

Question 6 :
On calcule le tableau des fréquences théoriques

	IPPG		
	IPPG$\leq$10	10<IPPG$\leq$40	40<IPPG
aucun ou légers	18,46	23,08	18,46
importants	21,54	26,92	21,54

On calcule Q=21,4 comparé à une $\chi^2_{2;\,0,05} = 5,99$. Comme 21,4>5,99, on rejette H_0, donc il y a une dépendance entre ces variables. (il y a plus de dommages associés aux IPPG élevés).

Question 7 :
a) Soit X le nombre total de véhicules. Comme X_1, X_2 et X_3 sont multinormal, X est normal, de moyenne 30 et de variance 1+1,5+1+1,5+4+2+1+2+4=18.
$P(X>40)=P(Z>(40-30)/18^{0,5})=P(Z>2,36)=0,009$
b) $\rho_{XY} = 1,5/(1*4)^{0,5} = 0,75$

Question 8 :
a) On calcule $z_{\text{calculé}} = (200-210)*16^{0,5}/50=-0,8$. $z_{0,95}=-1,645$. Comme –0,8>-1,645, on ne rejette pas H_0.
b) On rejette H_0 si $\overline{X} < 210-1,645(50/4) = 189,4$.
Sous H_1, $P(\overline{X} < 189,4)=P(Z<(189,4-180)*4/50)=P(Z<0,752)=0,774$
c) Pour avoir une puissance de 90% (β=10% et Z_β=1,28); $n=(1,645+1,28)^2 50^2/(210-180)^2=23,76$ ->24.