<u>QUESTION 1</u> (20 points)

À l'échelle mondiale, l'intervalle de temps (T en jours) entre des séismes de magnitude supérieure à 7 sur l'échelle Richter est réputé suivre une loi <u>lognormale</u>. La moyenne et la variance du <u>logarithme naturel</u> de T sont : $\mu = 2.71$, $\sigma^2 = 0.69$.

a) Quelle est E[T] ?

b) Quel est l'écart-type de T ?

c) Quelle est la probabilité que T < 60 jours ?

d) Déterminer T_0 tel que $P(T<T_0) = 0.95$?

<u>QUESTION 2</u> (18 points)

Une ville compte 20 000 logements. La ville est alimentée en électricité par une centrale thermique de puissance 350 000 kW. On définit X_i (i=1,...20 000) les v.a. donnant la consommation d'électricité instantanée (en kW) au mois de janvier pour chaque logement. La loi de distribution des X_i est inconnue, la moyenne de la distribution est 5 kW et l'écart-type est 10 kW.

a) Pourquoi la loi normale est-elle inadéquate pour décrire la distribution des X_i ?

b) Si l'on suppose l'indépendance des X_i, quelle est la probabilité que la demande totale d'électricité excède la capacité de la centrale thermique ?

c) Comme la consommation de chaque résidence augmente les jours plus froids, l'hypothèse d'indépendance des X_i n'est sûrement pas réaliste. On a pu déterminer une corrélation de 0.7 entre les consommations pour tout couple de résidences. Que devient la probabilité que la demande totale d'électricité excède la capacité de la centrale thermique ?

<u>QUESTION 3</u> (14 points)

On définit X : précipitation (de pluie) quotidienne maximale (en cm) obtenue au cours d'une année à la station pluviométrique de Dorval. On considère que X est distribuée suivant une loi normale (de paramètres inconnus). Un relevé des 31 dernières années a fourni les valeurs suivantes :

$$\sum_{i=1}^{31} x_i = 96 \text{ cm}$$

$$\sum_{i=1}^{31} x_i^2 = 300 \text{ cm}^2$$

a) Construisez l'intervalle de confiance bilatéral de niveau $1-\alpha = 0.95$ pour la moyenne de X.

b) Construisez l'intervalle de confiance bilatéral de niveau $1-\alpha = 0.95$ pour l'écart-type de X.

QUESTION 4 (16 points)

La résistance d'une éprouvette de béton dépend, entre autres facteurs, de la composition chimique du ciment. Soit X : la résistance en compression d'un béton (en MPa) et Y la concentration en C_3S du ciment (en %). On sait que ces variables suivent conjointement une loi binormale avec :

espérance : $E\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 25 \\ 60 \end{bmatrix}$ et matrice de variance-covariance : $K = \begin{bmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{bmatrix} = \begin{bmatrix} 4 & 5 \\ 5 & 9 \end{bmatrix}$

a) Calculez la corrélation entre X et Y.

b) Le ciment d'une éprouvette de béton montre à l'analyse Y=62 %. Quelle est la distribution conditionnelle de la variable X sachant cette information ?

QUESTION 5 (14 points)

Soit une variable aléatoire Z distribuée suivant une loi N(0,1). La fonction de densité de Z est :

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-0.5z^2} \qquad -\infty < z < \infty$$

a) Déterminez la fonction de densité de Y = | Z |.

b) Évaluez E[Y] .

QUESTION 6 (18 points)

Le sol d'un site considéré pour une transaction immobilière est contaminé au plomb. Vendeur et acheteur échantillonnent séparément le site :

	taille de l'échantillon	Moyenne échant. $\overline{X}$	Variance échant. (ppm^2)
Vendeur	$n_1 = 21$	$\overline{X}_1$	S_1^2
Acheteur	$n_2 = 11$	$\overline{X}_2$	S_2^2

On fait l'hypothèse que les deux échantillons sont aléatoires et indépendants entre eux et que les variables aléatoires X_{1i} (i=1...21) et X_{2i} (i=1 :11) suivent une même loi N(μ, σ^2).

a) Quelle est $P\left(\frac{S_1^2}{S_2^2} > 1.52\right)$?

b) Quelle est la loi de distribution (type et paramètres) de $\overline{X}_2 - \overline{X}_1$?

c) On suppose σ^2 = 40 000 ppm^2. Quelle est la probabilité que $|\overline{X}_2 - \overline{X}_1| > 100$ ppm ?

MTH2302C – Probabilités et statistique Contrôle périodique # 2 – Hiver 2010

Corrigé

Question 1 (20 pts; 5 pts pour a, b, c, d):

a) E[T] = exp(2.71+0.69/2) = 21.22 j.

b) σ_T = $\{21.22^2$ (exp(0.69)-1)$\}^{0.5}$ = 21.15 j.

c) P(T<60) = P(Z < (ln(60)-2.71)/$0.69^{0.5}$) = P(Z<1.67) = 0.953

d) P(Z < (ln(T_0)-2.71)/$0.69^{0.5}$) =0.95 => (ln(T_0)-2.71)/$0.69^{0.5}$ = 1.645 =>
T_0 = exp(1.645*$0.69^{0.5}$+2.71) = 58.94 j.

Question 2 (18 pts, 6 pts pour a, b, c)

a) Parce que si la loi était normale, il y aurait une forte probabilité d'observer des consommations négatives d'électricité. En effet P(Z<-5/10)=0.31

b) Soit T la consommation totale. Par le théorème central limite, T ~ N(5*20000, 100*20000)

⇨ P(T> 350000)=P(Z>(350000-100000)/$(2000000)^{0.5}$) = P(Z>176.78) = 0

c) Deux réponses sont possibles selon que l'on suppose ou non que la consommation totale suit une loi normale. En effet, la corrélation entre les consommations empêche d'appliquer le théorème central limite.

i. Si l'on suppose la normalité de la consommation totale T
On calcule $\sigma_{ij} = 0.7*10*10 = 70 \quad \forall i \neq j$, i=1...20000, j=1....20000;

Var(T)= 100*20000 + 70 * 20000*19999 = 2.8 x 10^{10} et

P(T>350000)=P(Z>(350000-100000)/$(2.8*10^{10})^{0.5}$) =P(Z>1.494) = 0.068

ii. Sinon, on doit répondre « on ne peut calculer cette probabilité car la corrélation entre les consommations empêche d'appliquer le théorème central limite ». (La distribution de T va dépendre en bonne partie de la distribution de la température en janvier et du lien entre consommation et température).

Question 3 (14 pts, 5 pts pour les deux intervalles, 2 pts pour les calculs de $\bar{x}$ et s^2)

$\bar{x}$ =96/31 = 3.1
s^2 = (1/30)*(300-96^2/31)=0.09

a) l'intervalle est $\bar{x} \pm t_{30,0.025}$ s/$n^{0.5}$ => 3.1 ± 2.042 $(0.09/31)^{0.5}$ = 3.1 ± 0.11 = [2.99 , 3.21]

b) l'intervalle pour σ^2 est $\left[\frac{(n-1)s^2}{\chi^2_{30,0.025}}, \frac{(n-1)s^2}{\chi^2_{30,0.975}}\right]$ => (30*0.09/46.98 , 30*0.09/16.79)
=> (0.057, 0.16).
Pour σ, ce sera : ($0.057^{0.5}$, $0.16^{0.5}$) = (0.24, 0.4)

Question 4 (16 pts, 4 pts pour a, et 12 pour b :
en b) 5 pts pour espérance cond et 5 pts pour var. cond, 2 pts pour indiquer loi normale)

a) $r_{XY} = 5/(4*9)^{0.5} = 5/6$

b) $E[X|\ y=62] = 25 + 5/9\ (62 - 60) = 26.11$

$var(X \mid y=62) = 4 - 5^2/9 = 1.22$

Donc $X_{|Y} \sim N(26.11, 1.22)$

Question 5 (14 pts, 7 pts pour a et b)

a) Par la symétrie de la loi normale, on a directement :

$$f_Y(y) = \begin{cases} \dfrac{2}{\sqrt{2\pi}} e^{-0.5y^2} & y \geq 0 \\ 0 & y<0 \end{cases}$$

(puisque l'intégrale de la fonction de densité de y doit donner 1)

On arrive au même résultat par la fonction de densité :
$f_Y(y)=f_z(z)|dz/dy|$. La fonction $y=|z|$ n'étant pas monotone, il faut additionner les contributions de la partie décroissante ($z<0$) et croissante ($z>0$) comme $|dz/dy|=1$, le résultat suit, $f_Y(y) = 2f_Z(y)$.
Une autre approche est par la fonction de répartition :
$F_Y(y)=P(Y<y)=P(-y<z<y) = F_Z(y)-F_Z(-y) = F_Z((y)-(1-F_Z(y))=2F_Z(y)-1$. Dérivant, $f_Y(y) = 2dF_Z(y)/dy =2f_Z(y)$

b) $E[Y]= \dfrac{2}{\sqrt{2\pi}} \displaystyle\int_0^\infty ye^{-0.5y^2}\,dy = \dfrac{2}{\sqrt{2\pi}}(-\exp(-0.5y^2)\,|_0^\infty = \dfrac{2}{\sqrt{2\pi}} = 0.798$

Question 6 (18 pts, 6 pts pour a, b et c)

a) $\dfrac{S_1^2}{S_2^2} \sim F_{n1-1,\ n2-1}$

On cherche $P(F_{20,10}>1.52)$ => Table 5 p. 551, correspond à α =0.25, donc 0.25

b) $N(0,\ \sigma^2\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right))$ => $N(0,\ (1/21+1/11)\,\sigma^2)$

c) Soit $\overline{D} = \overline{X}_2 - \overline{X}_1$. On cherche $P(\overline{D} >100) + P(\overline{D} < -100) = 2\ P(\overline{D} >100)$

$2\ P(Z> (100-0)/(1/21+1/11)*40000)^{0.5} = 2P(Z>1.34) =2*0.089 = 0.18$
