L1-MATH - STATISTIQUES DESCRIPTIVES

———-

EXAMEN FINAL

Enseignants : H. El-Otmany & V. Darrigrand **A.U.** : 2014-2015

*****

La qualité de la rédaction sera prise en compte dans la note. Les réponses devront être justifiées.
Documents interdits, calculatrice UPPA autorisée.
On donnera les résultats à 10^{-2} près.
Le barème est donné à titre indicatif.

****

Exercice n°1 *[3 points]* Soit $(x_1, y_1), \ldots, (x_n, y_n)$ une série bivariée correspondant à deux variables quantitatives.

1. Rappeler la formule de la covariance et du coefficient de corrélation linéaire ρ_{xy}. Quelles sont les valeurs possibles pour ρ_{xy} ?
2. Qu'est-ce qu'un nuage de points ? Représenter graphiquement la forme du nuage de points dans les trois cas suivants : $\rho_{xy} < 0$, $\rho_{xy} = 0$, $\rho_{xy} > 0$.
3. Énoncer l'équation d'analyse de la variance. Quel est l'intérêt de cette équation pour la régression linéaire ?

Exercice n°2 *[5 points]* Le centre de transfusion sanguine de Pau a observé la répartition suivante sur 10000 donneurs en 2014.

Facteur \ Groupe	O	A	B	AB
Rhésus +	2592	4387	587	194
Rhésus -	939	996	232	73

1. Quelle est la population étudiée ? Quelles sont les variables étudiées ? Préciser leur nature.
2. Calculer la distribution du rhésus
   (a) sachant que le donneur possède un groupe sanguin de type O ;
   (b) sachant que le donneur possède un groupe sanguin de type A ;
   (c) sachant que le donneur possède un groupe sanguin de type B ;
   (d) sachant que le donneur possède un groupe sanguin de type AB ;
3. Calculer la distribution conditionnelle du Groupe sachant que le Rhésus est négatif.
4. Calculer la fréquence conditionnelle du Groupe sachant que le Rhésus est positif.
5. Déterminer le mode de la distribution bivariée.
6. Calculer le coefficient Φ^2 de Pearson. Peut-on affirmer que les deux variables sont indépendantes ?

Exercice n°3 *[12 points]* On considère une automobile roulant sur une route mouillée à une certaine vitesse v. Cette automobile freine brusquement pour simuler un freinage d'urgence. On note d la distance de freinage qui est fonction de la vitesse v. Les données recueillies sont les suivantes :

Vitesse v (Km/h)	50	60	70	80	90	100	110
Distance d (m)	10	21	34	49	66	85	106

Le plan est muni d'un repère orthogonal. On prendra pour unité sur l'axe des abscisses 1 cm pour 10 Km/h et sur l'axe des ordonnées 1 cm pour 15m.

1. Quelles sont les natures des deux variables étudiées ?
2. Tracer le nuage de points. Le nuage de points vous semble-t-il aligné le long d'une droite ? Quel est le signe du coefficient de corrélation linéaire (sans effectuer de calculs) ?
3. On suppose qu'il existe une relation linéaire entre la vitesse v (Km/h) et la distance d (m) de la forme : $d = \alpha v + \beta$.
   (a) Trouver la droite de régression des distances en fonction des vitesses.
   (b) Tracer la droite obtenue dans le graphique précédent.
   (c) Quelle est la distance estimée de freinage si la vitesse est de 95 Km/h ? 130 Km/h ? 150 Km/h ?
   (d) Que deviennent α et β si on exprime la distance en kilomètres et non pas en mètres ?
   (e) Calculer le coefficient de corrélation linéaire entre d et v. Commenter.
   (f) Rappeler la **définition** du coefficient de détermination. Donner sa valeur sans effectuer de calcul.
   (g) Déterminer la variance des résidus.
4. On suppose qu'il existe une relation quadratique entre les deux variables, i.e. une relation de la forme $d = av^2 + b$.
   (a) Quelles sont les unités de a et b.
   (b) Déterminer a et b par la méthode des moindres carrés.
   (c) Représenter la courbe obtenue sur le graphique de la question**??**.
   (d) Quelle est la distance estimée de freinage si la vitesse est de 95 Km/h ? 130 Km/h ? 150 Km/h ? Comparer les résultats avec ceux de la question **??**.
   (e) Que deviennent a et b si on exprime la distance en kilomètres et non pas en mètres ?
   (f) Que vaut le le coefficient de corrélation linéaire ? Commenter.
   (g) Que vaut le coefficient de détermination ?
   (h) Déterminer la variance des résidus pour ce nouveau modèle.
5. Trois autres expériences supplémentaires ont été menées :

Vitesse v (Km/h)	110	120	130
Distance d (m)	129	154	181

   (a) Déterminer les nouvelles valeurs de a et de b par la méthode des moindres carrés, en limitant le nombre de calculs. On expliquera la démarche adoptée.
   (b) Que vaut le coefficient de détermination ? Comparer avec celui précédemment obtenu et commenter.
   (c) Quelle est la distance estimée de freinage si la vitesse est de 95 Km/h ? 130 Km/h ? 150 Km/h ? Comparer avec les résultats obtenus précédemment et commenter.
6. Dans cette question, on compare les deux ajustements à l'aide de tableau suivant :

Vitesse v_i (Km/h)	50	60	70	80	90	100	110	
Distance d_i (m)	10	21	34	49	66	85	106	
$\lvert d_i - (\alpha v_i + \beta)\rvert$								$T_1 =?$
$\lvert d_i - (a v_i^2 + b)\rvert$								$T_2 =?$

Les deux totaux calculés T_1 et T_2 évaluent, pour chaque ajustement, la somme des écarts entre les ordonnées des points du nuage et les ordonnées des points de même abscisse de l'ajustement.

(a) Donnez les arrondis à 10^{-2} près des deux totaux T_1 et T_2 calculés ci-dessus.

(b) Déduisez l'ajustement qui paraît le mieux adapté.