

Examen : Statistique appliquée

Exercice01 :

- On considère deux groupes d'étudiants. Nous relevons leurs notes d'examens dans les deux tableaux suivants :

Note (groupe A)	8	9	10	11
Effectif	2	2	1	1

Note (groupe B)	6	8	9	13	14
Effectif	2	2	2	1	1

-Calculer la moyenne \bar{x}_A et \bar{x}_B et l'écart type σ_X^A , σ_X^B de chaque groupe. Comparer les deux groupes.

Solution :

Solution Dans un premier temps, nous remarquons que l'effectif total du groupe A est égal à 6 et celui du groupe B est égal à 8.

En utilisant la formule de la moyenne, nous obtenons

$$\bar{x}_A = 9.2 \quad \text{et} \quad \bar{x}_B = 9.1.$$

On remarque que les moyennes sont très proches. Peut-on pour autant conclure que ces deux groupes ont des niveaux identiques ?

Nous répondons à cette question après le calcul des écarts type. Ils sont donnés par

$$\sigma_X^A = 1.11 \quad \text{et} \quad \sigma_X^B = 2.8.$$

Nous remarquons que même si les deux groupes ont des moyennes quasiment identiques, le groupe B est beaucoup plus dispersé que le groupe A car $\sigma_X^B > \sigma_X^A$. Les étudiants de ce groupe ont des notes plus irréguliers. On peut dire donc que le groupe B est moins homogènes que le groupe A. En observant les valeurs du tableau, on voit que c'est cohérent.

Exercice02 :

		X	Y
	Time	P (MCE)	Qfuite (m3/h)
1Jours	01:00-01 :30	0,011	1,364
	01:30-02 :00	0,013	1,889
	02:00-02 :30	0,017	3,333
	02:30-03 :00	0,013	1,81
	03:00-03 :30	0,01	1,151
	03:30-04 :00	0,011	1,6
	04:00-04 :30	0,009	1,24
	04:30-05 :00	0,008	1,083
2Jours	01:00-01 :30	0,01	1,5
	01:30-02 :00	0,013	3,4
	02:00-02 :30	0,014	2,222
	02:30-03 :00	0,011	1,462
	03:00-03 :30	0,009	1,094
	03:30-04 :00	0,008	0,941
	04:00-04 :30	0,007	0,816
	04:30-05 :00	0,006	0,65

Un doctorant chercheur en hydraulique veut utiliser la régression linéaire pour estimer les fuites du fond Q_{fuite} dans un réseau d'alimentation maillé (AEP) en fonction de la pression P des noueux de conduites :

$$Q_{fuite} = a (P)^b / a \text{ et } b \text{ constant}$$

A l'aide d'un réglage PRV1 au début du canal de distribution est varié toutes les 30 minutes entre 01 :00 H – 05 :00H pendant 2jours présenter dans le tableau ci-dessous :

- 1- choisir la bonne droite de régression $Y=f(X)$
- 2-la variance totale
- 3-La variance de régression
- 4- Le Coefficient de corrélation R^2

Solution

$$1/Y = 1875,1 X^{1,5659}$$

$$a = \frac{\sum_{i=1}^n x_i y_i - n \cdot \bar{x} \cdot \bar{y}}{\sum_{i=1}^n x_i^2 - n(\bar{x})^2} \quad b = \bar{y} - a\bar{x}$$

$$a=1875,1 / b= 1,5659$$

$$\text{Variance totale} = SY^2 = \sum Y^2/n - Y_{moy}^2 = 1.24$$

$$\text{Variance de régression} = S\hat{Y}^2 = \sum \hat{Y}^2/n - Y_{moy}^2 = 1.13$$

$$R^2 = S\hat{Y}^2 / SY^2 = 0.9 \quad r=0.94$$

Ce coefficient est compris entre 0 et 1. Plus il est proche de 1 et plus la qualité globale de la régression est bonne.

Bon Courage