

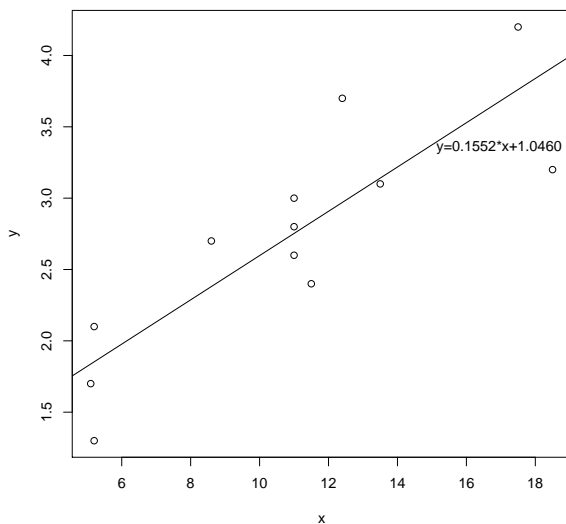
1. Régression linéaire

Une société fait une analyse du coût de publicité et la quantité de vente d'un produit par mois, pendant une année. Elle obtient les données suivantes (publicité,vente) en millions.

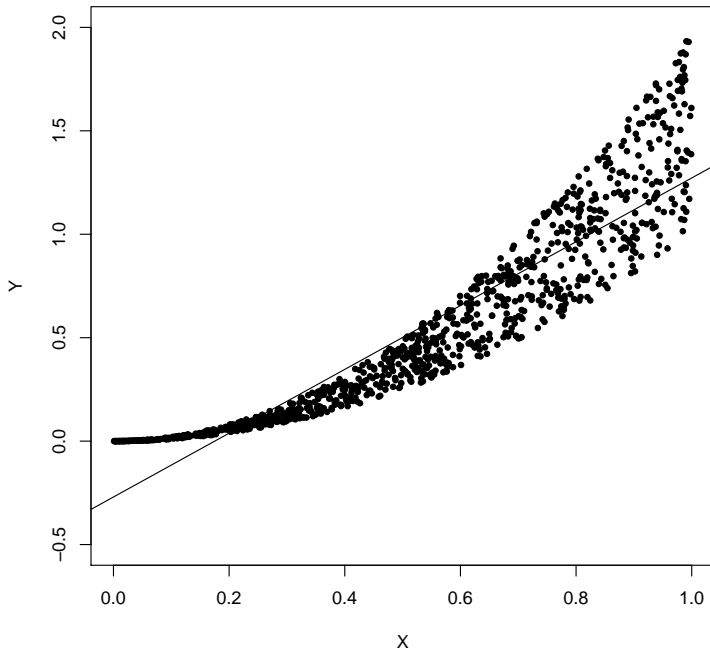
(5.2,1.3), (5.2,2.1), (11.5,2.4), (5.1,1.7), (13.5,3.1), (11,2.8),
(8.6,2.7), (17.5,4.2), (18.5,3.2), (12.4,3.7), (11,2.6), (11,3)

1. Calculer les moyennes du coût de publicité et de la vente mensuels.
2. Calculer les variances de publicité et de la vente, puis la covariance.
3. Dessiner le nuage des points, et marquer le centre de gravité du nuage.
4. Calculer l'équation de la régression linéaire.
5. Dessiner la droite sur le dessin.
6. Est-ce que le rapport entre le coût de la publicité et la vente est suffisamment fort pour que la droite soit représentative?
7. Supposons que le bénéfice marginal du produit vendu est de 5 euros par unité. D'après l'analyse, est-ce que la société aura plutôt intérêt de faire plus de publicité ou moins de publicité sur le produit ?

Solution :

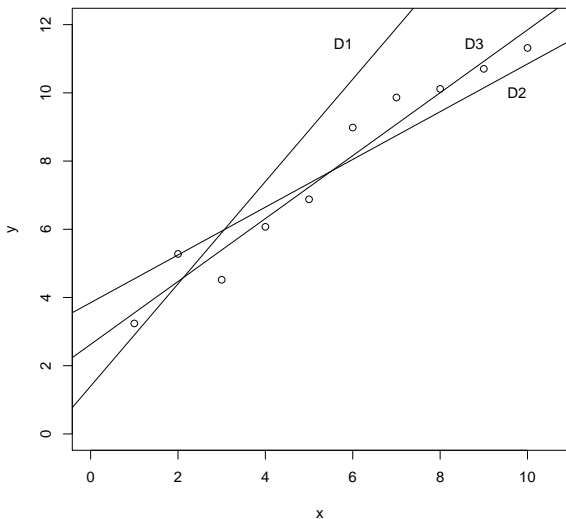


2. Une population de 1000 individus est étudiée à travers deux caractères quantitatifs X et Y . On représente ci-dessous le nuage des points $(X(i), Y(i))$ et la droite de régression $y = ax + b$. Pour chaque individu i on note $R(i)$ la différence entre la valeur de Y et la valeur prédite par X et la régression linéaire : $R(i) = Y(i) - aX(i) - b$.
- a. D'après le graphique, que valent approximativement les coefficients a et b de la droite de régression ?



- b. Sachant que la moyenne de X vaut $\frac{1}{2}$, que vaut approximativement la moyenne de Y ?
- c. Que vaut la moyenne de Y conditionnée à l'évènement " X est proche de 0" ? Que vaut la moyenne de R conditionnée à " X est proche de 0" ?
- d. Que vaut la moyenne de Y puis de R conditionnée à " X est proche de 1" ?
- e. La régression linéaire est elle satisfaisante ? Quel changement de variable sur X peut on essayer pour améliorer la régression ?

3. Le nuage de points ci-dessous représente les caractères x et y d'une population de 10 individus. Pour chacune des droites D1, D2, D3 sur le dessin, dites si la droite peut raisonnablement être la droite de régression associée au nuage de points. Argumentez.



4. Voir aussi le sujet d'examen 2ème session de 2009-10 sur la page web et les exercices 6.1 à 6.5 de [Hamon & Jégou].