

**L3 Economie - Gestion**  
**Cours de Mr Karcher 2012/2013**  
**T.D. n°1 : La régression simple**

I On dispose de données agrégées (en milliards de dollars) concernant les dépenses de consommation et le revenu disponible aux Etats – Unis sur 12 années :

Années	y Consommation	x revenu
1	102	114
2	106	118
3	108	126
4	110	130
5	122	136
6	124	140
7	128	148
8	130	156
9	142	160
10	148	164
11	150	170
12	154	178

- 1°) Construire un diagramme de dispersion pour les données agrégées du tableau, et indiquer s'il existe à l'observation une relation linéaire entre y et x ;
- 2°) Formuler la relation générale entre la consommation y et le revenu disponible x :
  - a) Sous une forme linéaire exacte,
  - b) Sous une forme aléatoire,
  - c) Pourquoi peut – on s'attendre à ce que la plupart des valeurs observées de y ne donnent pas des points situés exactement en ligne droite ?
- 3°) Etablir l'équation de régression et tracer la droite de régression correspondante, en indiquant l'écart spécifiant chaque couple  $(y_i, \hat{y}_i)$  ; Calculer les  $e_i$  et en faire la somme des carrés.
- 4°) Quelle est la différence entre les deux couples de termes  $(\alpha, \beta)$  et  $(\hat{\alpha}, \hat{\beta})$  ? Quelle est la différence entre  $u_i$  et  $e_i$  ? Ecrire les équations des droites correspondant aux deux régressions, vraie et estimée, de y par rapport à x.
- 5°) Quelle est la signification de  $\hat{\alpha}$  ? de  $\hat{\beta}$  ? Déterminer l'élasticité par rapport au revenu de la consommation.
- 6°) Calculer  $\hat{\sigma}_u^2$ ,  $\sigma_{\hat{\alpha}}^2$  et  $\sigma_{\hat{\beta}}^2$ .
- 7°) Tester, au seuil de signification de 5%,  $\alpha$  et  $\beta$ .
- 8°) Etablir les intervalles de confiance à 95% pour  $\alpha$  et  $\beta$ .
- 9°) Déterminer le  $R^2$ , par la formule de calcul usuelle et en utilisant la somme du carré des résidus.
- 10°) Présenter enfin le résultat de la régression de y par rapport à x.

II Un artiste peintre un peu pessimiste se dit qu'en fait le prix auquel il vend ses toiles est lié à la taille de l'œuvre. Son neveu lui ayant parlé d'économétrie, il décide d'appliquer les moindres carrés sur les surfaces et les prix de ventes de ses sept dernières œuvres. Soit :

y surface en cm <sup>2</sup>	x prix en euros
240	150
280	190
320	230
360	250
400	300
420	310
490	350

**A :** Notre artiste n'y connaissant pas grand chose, il vous demande d'effectuer la régression de  $y$  expliqué par  $x$ , soit :

$$y = \alpha x + \beta + u,$$

puis la régression de  $x$  expliquée par  $y$ , soit :

$$x = \alpha' y + \beta' + u'.$$

**B :** A partir des résultats obtenus dans la partie A :

- 1°) Déterminer le coefficient de régression  $R^2$  ;
- 2°) Le coefficient de chaque régression est – il significativement différent de zéro ?
- 3°) les régression sont –elles bonnes dans leurs ensembles ?
- 4°) Notre artiste peint une toile de 510 cm<sup>2</sup>. Quel prix peut – il en espérer ?
- 5°) Il doit rembourser une dette de 400 euros. S'il peint une toile pour cela, quelle surface doit-il peindre ?

**III** On dispose, pour un secteur donné et sur une période de 11 années, de la série du nombre de salariés  $Y$  et du chiffre d'affaires  $X$  du secteur.

Année	Nombre de salariés (en milliers)	Chiffres d'affaires (en millions)
1992	294	624
1993	271	661
1994	314	728
1995	356	782
1996	383	819
1997	369	869
1998	402	938
1999	422	1023
2000	475	1136
2001	475	1227
2002	486	1317

On veut tester la réalité d'une relation linéaire entre les variables  $Y$  et  $X$ , soit :

$$Y_n = \alpha X_n + \beta + \varepsilon_n \quad (n = 1, \dots, 11).$$

Les hypothèses classiques du modèle linéaire simple sont supposées réalisées, i.e. que  $\varepsilon_n$  est une variable aléatoire suivant une loi normale centrée d'écart - type  $\sigma$ ,  $N(0, \sigma)$ .

- 1°) Représenter le nuage de points.
- 2°) Donner les estimations  $a$  et  $b$  de  $\alpha$  et  $\beta$  obtenues par la méthode des moindres carrés ordinaires.
- 3°) Calculer l'estimation de la variance  $\sigma^2_{u}$ , l'estimation de la variance de l'estimateur  $a$ , et le coefficient de détermination de la régression  $R^2$ .

4°) On se pose le problème de test  $\begin{cases} H_0 : \alpha = 0 \\ H_1 : \alpha \neq 0 \end{cases}$ . A quelle question ce test répond t-il ? Peut-on dire que  $a$  est

significativement différent de zéro, au risque  $\alpha = 0,05$  ?

5°) On suppose que, connu de façon exogène, le chiffre d'affaires en 2003 est de 1772. Sur la seule base de la relation linéaire estimée jusque là à partir des données fournies, à combien de salariés peut-on s'attendre en 2003 ?

Trouver un intervalle de confiance de niveau 0,95 pour l'effectif salarié du secteur. A posteriori, on dénombre, en 2003, 514 000 salariés. Que pouvez vous déduire de cette observation supplémentaire ?

#### Exercice IV

Un chef d'entreprise étudie les ventes d'un de ses plus récents produits. Il a à sa disposition les montants des ventes au cours des sept derniers mois, le montant des investissements lourds qu'il a du réaliser, ainsi que les frais de publicité. Les résultats des sept mois de vente, ainsi que les investissements et les frais de publicité, sont résumés dans le tableau suivant :

Mois	Ventes	Investissement
1	40	100
2	50	200
3	50	300
4	70	400
5	65	500
6	65	600
7	80	700

*En milliers d'euros.*

Le relevé d'observations ci-dessus permet d'envisager l'hypothèse d'une relation linéaire entre le montant des ventes, et celui de l'investissement. L'hypothèse s'écrit :

$$V = \alpha + \beta.I + u$$

Avec : V : montant des ventes ; I : montant de l'investissement ; u variable aléatoire représentant les erreurs.

L'ajustement est réalisé par la méthode des moindres carrés ordinaires, au moyen d'un logiciel adapté.

On obtient les résultats suivants :

## RAPPORT DÉTAILLÉ

<i>Statistiques de la régression</i>	
Coefficient de détermination multiple	0,91950905
Coefficient de détermination R <sup>2</sup>	?
Coefficient de détermination R <sup>2</sup>	?
Erreur-type	5,96118397
Observations	7

## ANALYSE DE VARIANCE

	<i>Degré de liberté</i>	<i>Somme des carrés</i>	<i>Moyenne des carrés</i>	<i>F</i>
Régression	1	972,321429	972,321429	?
Résidus	5	177,678571	35,5357143	
Total	6	1150		

	<i>Coefficients</i>	<i>Erreur-type</i>	<i>Statistique t</i>	<i>Probabilité</i>
Constante	36,4285714	5,03811999	?	0,0007894 3
Investissement	0,05892857	0,01126558	?	0,0033793 5

1°) Déterminez le coefficient de corrélation R<sup>2</sup> ;

2°) Le coefficient  $\alpha$  est-il significativement différent de 0 ? Commentez le résultat.

3°) La régression est-elle bonne dans son ensemble ? Commentez le résultat.