Tasdawit n'Bgayet
Université de Béjaïa

Série de TD N° 1

Exercice N°1

le tableau suivant représente la série des chiffres d'affaires trimestriels (en millions de dinars) sur les 6 dernières années d'une grande entreprise.
Désignons par y_t le chiffre d'affaires relatif au trimestre t et par x_t le rang du trimestre t.
On fait l'hypothèse que ces données sont liées par un modèle de régression linéaire simple de type $y_t = a + bx_t + \varepsilon_t$.

	Trim1	Trim2	Trim3	Trim4
2011	526.76	535.08	459.68	557.96
2012	559.00	570.96	489.32	603.72
2013	601.12	619.84	508.04	625.56
2014	598.52	621.40	525.72	662.48
2015	648.96	670.80	568.36	694.72
2016	672.88	685.36	573.04	709.28

1. Représenter graphiquement les données.Que peut-on conclure ?
2. Calculer le coefficient de corrélation r. Que peut-on conclure ?
3. Estimer les paramètres a et b du modèle et tracer la droite de régression correspondante.
4. Tester l'hypothèse H_0 : "$b = 0$" contre H_1 : "$b \neq 0$" et donner un intervalle de confiance à 95% pour b. ($t_{(\frac{0.05}{2},22)} = 2.074$).
5. Tester l'hypothèse H_0 : "$a = 0$" contre H_1 : "$a \neq 0$" et donner un intervalle de confiance à 95% pour a.
6. Au vu de ces résultats, quelles prévisions pouvait-on faire du chiffre d'affaires de l'année 2017.

Exercice N°2

Le tableau ci dessous contient la liste de 14 pays d'Amérique du nord et centrale, dont la population dépassait le million d'habitants en 1985. Pour chaque pays, on a le taux de natalité y_i (le nombre de naissances par année pour 1000 personnes) ainsi que le taux d'urbanisation x_i (le pourcentage de la population vivant dans des villes de plus de 100000 habitants) en 1980.
On fait l'hypothèse que ces données sont liées par un modèle de régression linéaire simple de type $y_i = a + bx_i + \varepsilon_i$. C'est à dire que le taux de natalité d'un pays dépend de son taux d'urbanisation.

Obs	Pays	Taux d'urbanisation x_i	taux de natalité y_i
1	Canada	55	16.2
2	Costa rica	27.3	30.5
3	Cuba	33.3	16.9
4	USA	56.5	16
5	El Salvador	11.5	40.2
6	Guatemala	14.2	38.4
7	Haiti	13.9	41.3
8	Honduras	19	43.9
9	Jamaique	33.1	28.3
10	Mexique	43.2	33.9
11	Nicaragwa	28.5	44.2
12	Trinidade-Tobago	6.8	24.6
13	Panama	37.7	28
14	Rép Dominicaine	37.1	33.1

1. Représenter graphiquement les données.
2. Estimer les paramètres a et b du modèle et tracer la droite de régression correspondante.
3. Calculer SCT, SCE et SCR puis R^2.
4. Tester l'hypothèse H_0 : "$b = 0$" contre H_1 : "$b \neq 0$" et donner un intervalle de confiance à 95% pour b. ($t_{(\frac{0.05}{2},12)} = 2.179$).
5. Tester l'hypothèse H_0 : "$a = 0$" contre H_1 : "$a \neq 0$" et donner un intervalle de confiance à 95% pour a.

Exercice N°3

On dispose des données suivantes au sujet de deux variables d'intérêt x et y :

x_i	7	9	9	10	13	17	19	20	21	25
y_i	5	4	6	4	1	2	0	1	1	0

$\sum x_i = 150$; $\sum y_i = 24$; $\sum x_i.y_i = 253$; $\sum x_i^2 = 2596$; $\sum y_i^2 = 100$. On se refère au modèle linéaire $y_i = a + bx_i + \varepsilon_i$.

1. Estimer les paramètres a et b par les MCO.
2. Pour chacun de ces deux paramètres, trouver un intervalle de confiance à $\alpha = 99\%$. on donne ($t_{(\frac{0.01}{2},8)} = 3.355$)

Exercice N°4

On dispose des données concernant les salaires des Hommes et des Femmes à l'intérieur de 30 ménages. Le but est de faire une comparaison de ses salaires à l'aide d'un modèle de régression linéaire simple **sans constante** $y_i = a.x_i + \varepsilon_i$ avec Y : représente le salaire des hommes et X : le salaire des femmes.
On veut tester l'hypothèse "$H_0 : a = 1$", l'homme a le même salaire que sa conjointe contre "$H_1 : a > 1$", l'homme a un salaire plus élevé que sa conjointe.

Numéro	Salaire H	Salaire F	Numéro	Salaire H	Salaire F
1	7.83	7.20	16	7.80	7.38
2	6.83	7.06	17	7.57	7.53
3	6.97	7.10	18	6.02	6.03
4	7.85	7.39	19	7.28	7.05
5	7.48	6.97	20	8.42	8.01
6	7.86	7.50	21	7.42	7.25
7	7.44	7.16	22	7.47	7.59
8	7.83	7.77	23	7.14	7.20
9	7.38	7.78	24	7.29	6.93
10	7.28	7.47	25	8.28	7.85
11	7.53	7.51	26	6.98	7.29
12	8.40	8.07	27	8.03	7.94
13	7.48	7.25	28	7.69	7.11
14	7.46	6.79	29	6.67	6.76
15	7.33	7.14	30	7.92	7.72

1. Estimer le coefficient de la droite d'équation :$Y = a.X$.
2. Tester au seuil de 5% la validation de l'hypothèse H_0.

Exercice N°5

Un Econométre a mené une étude sur le salaire horaire Y d'ouvriers, à partir d'un échantillon comportant 250 hommes et 280 femmes. On considère la variable indicatrice définie par :

$$M = \begin{cases} 1, & \text{si l'ouvrier est un homme}\,; \\ 0, & \text{s'il s'agit d'une femme.} \end{cases}$$

Les résultats de l'estimation du modèle $Y = a_0 + a_1 M$ sont les suivants :

$$\begin{cases} \hat{a}_0 = 12.52 \ et \ S(\hat{a}_0) = 0.23, \\ \hat{a}_1 = \ 2.12 \ et \ S(\hat{a}_1) = 0.36, \end{cases} \quad \text{et on note } \hat{Y} = 12.52 + 2.12M \text{ avec } \hat{S}_\epsilon = 4.2$$

1. Donner la signification des paramètres $\hat{a}_0$ et $\hat{a}_1$.
2. La différence de salaire dans les deux sous échantillons (hommes et femmes) est-elle significativement différente de zéro au niveau 5%. On donne $t_{(\frac{0.05}{2},528)} = 1.96$
3. Un autre économètre a mené une étude similaire en regressant le salaire, sur la variable indicatrice Femme définie par :

$$M = \begin{cases} 1, & \text{si l'ouvrier est une femme}\,; \\ 0, & \text{s'il s'agit d'un homme.} \end{cases} \quad \text{Pour le modèle : } \hat{Y} = \hat{b}_0 + \hat{b}_1 M.$$

– Déterminer $\hat{b}_0$ et $\hat{b}_1$.

Exercice N°6 :(à traiter dans le cours)

Les experts d'un ministère ont constaté une élévation importante du nombre de saisies opérées par les services de répression des fraudes concernant un produit. Ces experts pensent que cette évolution peut être imputée à l'augmentation du prix du produit considéré. Les données présentées dans le tableau ci-après indiquent le nombre d'unités du produit saisies dans les ports et son prix pour les mêmes années.

Année	Nombres d'unités saisies (milliers)	Prix unitaire TTC
1	50	4
2	120	16
3	350	42
4	800	82
5	1600	200
6	2800	420
7	5000	900

1. Représenter graphiquement les variations du nombre de saisies (Y) en fonction du prix unitaire (X). Commenter le graphique.
2. On se propose d'ajuster sur ces données un modèle du type :

$$Y_t = \alpha . X_t^{\beta}$$

- Utiliser la transformation log pour linéariser le modèle.
- Estimer les coefficients α et β du modèle .
3. Estimer le nombre d'unités saisies pour l'année 8 sachant que le prix unitaire sera de 1000.