# الانحدار الخطي البسيط

## La régression linéaire simple



المتغيرات الاحصائية

المتغير المستقل variable Independent: هو المتغير الذي يتم بحث اثره في متغير اخر ويمكن للباحث التحكم فيه للكشف عن تباين هذا الاثر باختلاف قيم ذلك المتغير.

Itariable dépendent هو ذلك المتغير الذي يرغب الباحث في الكشف عن تأثير المتغير المستقل عليه.

## تعريف الانحدار الخطي

الانحدار أسلوب يمكن بواسطته تقدير قيمة أحد المتغيرين بمعلومية قيمة المتغير الآخر عن طريق معادلة(دالة) الانحدار.

Dépendent ويعرف المتغير الأول بالمتغير التابع Dépendent ويرمز له بالحرف Y، في حين يعرف المتغير الآخر بالمتغير المستقل Indépendant ويرمز له بالحرف X.

تهدف دراسة الانحدار التنبؤ بقيمة متغير Y بمعرفة متغير X. إذن الغاية من استخدام أسلوب تحليل الانحدار الخطي دراسة وتحليل أثر متغير كمي على متغير كمي آخر.

سندرس: الانحدار الخطي البسيط : فكلمة " بسيط " تعني أن المتغير التابع Y يعتمد على متغير مستقل واحد وهو X وكلمة " خطي" تعني أن العلاقة بين المتغيرين (X , Y) هي علاقة خطية.

$$\hat{y} = a + bx$$

$$ext{act} \quad a \quad \text{: ily: Product the formula of the equation of the equation$$



## حساب الانحدار الخطي البسيط باستخدام برنامج الـ SPSS :

#### Menu Analyse--> Régression--> Linéaire

stics	Editeur de	e données							
r	<u>A</u> nalyse	Marketing direct	<u>G</u> raphes	Utilit	taires	Fenêtre	Aide		
	Ra <u>p</u> Stati	ports stiques descriptive	s	•					ABC
	Ta <u>b</u> l	eaux		•					
	Com	parer les moyenne	es.	•		var	var	var	var
65	Mod Mod	ele lineaire <u>g</u> enera èles linéaires géné	ı éralisés						
64 18	Mod	èles Mi <u>x</u> tes							
7	<u>C</u> orr	élation		•					
1	<u>R</u> ég	ression		•	💻 м	odélis <u>a</u> tio	on linéaire au	tomatique	
6	Log	Linéaire				inéaire			
7	Rés	eaux neuronaux		•	- <b>A</b>	justemen	t de fon <u>c</u> tions	;	
	Clas	sification		•	S. M	oindres o	arrés partiels	s	
	Ré <u>d</u>	uction des dimensi	ions	•	R		binaire		
	Eche	elle		•		ogistique	binane		
	Test	s <u>n</u> on paramétriqu	es	•	HULT L	ogistique	multinomiale	<b>.</b>	
	Prev	ieione				r <u>d</u> inale			

بحيث يتم نقل المتغير التابع الى المجال Dependent والمتغير المستقل الى المجال Independent، ويتوفر الزر Method على عدة خيارات خاصة بطريقة التقدير، والطريقة التلقائية للبرنامج هي طريقة المربعات الصغرى العادية المشار اليها بـ Enter. أما المجال الخاص بالمتغيرات المستقلة فيتوفر فيه امكانية تقدير مجموعة من النماذج باستخدام نفس المتغير التابع ومتغيرات مختلفة مستقلة مثلا استخدام المتغير X1 في النموذج (1) و X2 في النموذج (2) بنفس الإجراء، وذلك بالاستعانة بالزر Next في المجال Block 1 Of 1، بحيث يتم نقل المتغير X1 في Block1 والمتغير X2 في Block2، ونتائج التقدير تكون في شكل نموذجين مختلفين. بينما ادخال المتغيرين معا في المجال Independent فإنه يؤدي الى تقدير نموذج انحدار متعدد (بمتغيرين مستقلين).





الزر Statistiques نقوم بتحديد الخيار ات التالية: ✓ Estimations: لتقدير معلمتي نموذج الانحدار. 95% : Intervalles de confiance : لتقدير مجال الثقة €0 لمعلمي الانحدار ✓ Qualité de l'ajustement : لتقدير جودة النموذج (لعرض جدول تحليل التباين ANOVA و معامل التحديدR2).  $\checkmark$  Caractéristiques: لعرض الإحصاءات الوصفية. Mesure et corrélations partielles√ الحساب معاملات الارتباط البسبط و الجزئي.

## Poursuivre

Régression linéaire : Enregistrer	×		الزر Enregistrer
Prévisions	Résidus		
✓ Non standardisés	Non standardisés		
Standa <u>r</u> disés	Standardisés		
Ajustées	Studentisés		
Erreur standard prévision moyenne	Supprimées		
	Supprimés stud <u>e</u> ntisés		
Distances	Statistiques d'influence	Non standardisós	nrávicion
Ma <u>h</u> alanobis	Df <u>B</u> êta(s)	Non standaruises	
Cook	DfBêta(s) standardisés		
Valeurs influentes	Différence de prévision		
Intervalles de la prévision	Dfprévision standardisée		
Moyenne 🔲 Individuelle	Rapport de covariance		
Intervalle de <u>c</u> onfiance : 95 %		Standardisés	résidus 💶 🔤
Statistiques à coefficients			
Créer des statistiques à coefficients			
Oréer un ensemble de données			
Nom <u>d</u> e l'ensemble de données :			
Ecriture d'un nouveau fichier de données			
Fichier			
Exporter les informations du modèle dans un fic	hier XML		
	Pa <u>r</u> courir		Poursuivre
✓ Inclure la matrice de covariance			
Poursuivre	Aide		
Prévisions	لاختيارات الخاصة ب (القيم التنبؤية) ،	ی Non standardisésمن آ	في هذه النافذة نؤشر على
pours للعودة إلى النافذة الأساسية.	ة ب (البواقي)، تم نضغط على 🛛 uivre	Stanc من الاختيارات الخاص	ا ا ا ا ا ا ا ا ا ا ا ا ا ا ا ا ا ا ا
		J = U	



## الزر Diagrammes



## فيما يلي بيانات عن عدد ساعات المذاكرة في الأسبوع لعينه من 10 طلاب و مقدار الزيادة في التحصيل لمادة الرياضيات.

ساعات المذاكرة	10	11	14	15	20	25	46	50	59	70
الزيادة في التحصيل	10	10	12	12	13	13	19	15	16	20

#### المطلوب:

- ارسم مخطط الانتشار.
- قدر معادلة انحدار ساعات المذاكرة على التحصيل.
  - فسر معادلة الانحدار.



arrow de la connéeso - IBM SPSS Statistics Editeur de données

Ragnorts         X       Y         X       Y         Modèles linéaire général         X       Y         X       Y         Modèles linéaire général         Modèles linéaire         Qurrélation         Modèles linéaire général         Modèles linéaire         Generonaux         Cligstique multineautomatique         Logistique planaire         Modèles de de dine de dimensions         Echelle         Test non paramétriques         Previsions        Qurvie       Modèles de cloux binaire         Modèles de valeurs manquantes       Modèles de cloux binaire         Modèles de valeurs manquantes       Modèles de cloux bi	<u>Fichier</u> Ec	dition Affichage	<u>D</u> onnées <u>T</u> ransformer	<u>Analyse</u> <u>Marketing direct</u> <u>Graphes</u>	Util	itaires	s Fenêtre	Aide				
X         Y         Comparer les moyennes           1         10,00         10,00           2         11,00         10,00           3         14,00         12,00           4         15,00         12,00           6         25,00         13,00           7         46,00         19,00           8         50,00         15,00           9         59,00         16,00           7         46,00         19,00           Régression         Image: Comparer les moyennes           Régression         Image: Comparer les moyennes           8         50,00         13,00           Comparer les moyennes         Image: Comparer les moyennes           Modèles linéaires général         Image: Comparer les moyennes           Régression         Ligg Linéaire           Régression         Ligg Linéaire           Régression         Ligg Linéaire           Régression         Image: Linéaire           Régression         Image: Linéaire           Régression         Image: Linéaire           Image: Linéaire         Image: Linéaire           9         59,000         16,00           Tests inon paramétriques         Ima				Rapports Statistiqu <u>e</u> s descriptives Tableaux	* *	4	۵ 🎹 🍐			6		
1       10,00       10,00       Modèle linéaire général       Modèles linéaires généralisés         2       11,00       10,00       Modèles linéaires généralisés       Modèles linéaires généralisés         3       14,00       12,00       Corrélation       Modèles linéaires         4       15,00       12,00       Régression       Implication         5       20,00       13,00       Log Linéaire       Implication         6       25,00       13,00       Régression       Implication         7       46,00       19,00       Réguidion des dimensions       Implication         8       50,00       15,00       Echelle       Implication       Implication         9       59,00       16,00       Tests non paramétriques       Implication       Implication       Implication         11       Réponses multiples       Implication multiple       Implication multiple       Implication multiple       Implication multiple       Implication multiple       Implication	,	X	Y	Co <u>m</u> parer les moyennes	*	r	var	var	var	var	var	
2       11,00       10,00       Modèles linéaires généralisés       Modèles Migtes         3       14,00       12,00       Corrélation       Modèles Migtes         4       15,00       12,00       Régression       Modèlis giton linéaire automatique         5       20,00       13,00       Log Linéaire       Modèlis giton linéaire automatique         6       25,00       13,00       Classification       Modèles dimensions         7       46,00       19,00       Régression       Modèles dimensions         8       50,00       15,00       Echelle       Modèles de choix binaire         9       59,00       16,00       Tests non paramétriques       Modèles de choix binaire         10       70,00       20,00       Previsions       Modèles de choix binaire         11       Survie       Modèles valeurs manquantes       Modèles de choix binaire         13       Imputation multiple       Echantilions complexes       Modèles moindres carrés         14       Contrôle de gualité       Contrôle de gualité       Codeage optimal (CATREG)         16       Courbe ROC       Courbe ROC       Modèles de choix de carrés	1	10,00	10,00	Modèle linéaire <u>a</u> énéral	►							
3       14,00       12,00       Orrélation       Imodeles Mages       Imodeles Mages         4       15,00       12,00       Corrélation       Imodeles Mages       Imodeles Mages         5       20,00       13,00       Log Linéaire       Imodeles Mages       Imodeles Mages         6       25,00       13,00       Log Linéaire       Imodeles Mages       Imodeles Mages         7       46,00       19,00       Réseaux neuronaux       Imodeles Mages       Imodeles de fongtions         8       50,00       15,00       Echelle       Imodeles fongtique binaire       Imodeles de choix binaire         9       59,00       16,00       Tests non paramétriques       Imodeles de choix binaire       Imodeles de choix binaire         10       70,00       20,00       Previsions       Imodeles de valeurs manquantes       Imodeles de valeurs manquantes         11       Imodeles de valeurs manquantes       Imodeles de valeurs carrés       Imodeles de gualité       Imodeles de gualité         14       Contrôle de gualité       Imodeles de choc       Codeage optimal (CATREG)       Codeage optimal (CATREG)	2	11,00	10,00	Modèles linéaires généralisés	•							
4       15,00       12,00       Régression       Image: Constraint of the second s	3	14,00	12,00	Corrélation	- P-							
5       20,00       13,00       Lgg Linéaire       Image: Classification       Imag	4	15,00	12,00				Modélis <u>a</u> tion	linéaire autor	matique			
6       25,00       13,00       Réseaux neuronaux       Ajustement de fongtions         7       46,00       19,00       Régaution des dimensions       Moindreg carrés partiels       Moindreg carrés partiels         8       50,00       15,00       Echelle       Logistique binaire       Logistique multinomiale         9       59,00       16,00       Tests non paramétriques       Moidreg carrés partiels       Moidreg carrés partiels         10       70,00       20,00       Previsions       Moidreg carrés de choix binaire       Moidreg carrés de choix binaire         11       Képonses multiples       Moidreg carrés       Moidreg carrés de choix binaire       Moidreg carrés de choix binaire         12       Aluge des valeurs manquantes       Moidreg carrés       Moidreg carrés       Moidreg carrés         13       Contrôle de gualité       Moidreg carrés       Moidreg carrés de choix binaire       Moidreg carrés de choix binaire         14       Contrôle de gualité       Moidreg carrés       Codage optimal (CATREG)       Moidreg carrés         16       Courbe ROC       Courbe ROC       Moidreg carrés de carrés       Moidreg carrés de carrés	5	20,00	13,00	L <u>og</u> Linéaire	•	R	<u>L</u> inéaire					
7       46,00       19,00       Réduction des dimensions       Moindreg carrés partiels         8       50,00       15,00       Echelle       Logistique binaire         9       59,00       16,00       Tests non paramétriques       Logistique multinomiale         10       70,00       20,00       Previsions       Modèles de choix binaire         11       Survie       Moindreg carrés partiels       Moindreg carrés partiels         12       Survie       Modèles de choix binaire         13       Imputation multiple       Moindreg carrés         14       Contrôle de gualité       Codage optimal (CATREG)         16       Courbe ROC       Imputation multiple	6	25,00	13,00	Réseaux neuronaux	- F-	محسر	Ajustement	de fon <u>c</u> tions				
8       50,00       15,00       Echelle       Logistique binaire         9       59,00       16,00       Tests non paramétriques       Logistique multinomiale       Logistique binaire         10       70,00       20,00       Previsions       Modèles de choix binaire       Modèles de choix binaire         11       Képonses multiples       Modèles de choix binaire       Modèles de choix binaire       Modèles de choix binaire         12       Analyse des valeurs manquantes       Modèles multiple       Modèles de choix binaire       Modèles de choix binaire         13       Contrôle de gualité       Contrôle de gualité       Codage optimal (CATREG)       Codage optimal (CATREG)         16       Courbe ROC       Courbe ROC       Doubles       Doubles       Doubles       Doubles	7	46,00	19,00	Réduction des dimensions	*	Moindres carrés partiels						
9       59,00       16,00       Tests non paramétriques       Image: Construction multinomiale         10       70,00       20,00       Previsions       Image: Construction multiples       Image: Construction multiple	8	50,00	15,00	Echelle	۲		Logistique b	inaire				
10       70,00       20,00       Previsions       Image: Commander of the c	9	59,00	16,00	Tests <u>n</u> on paramétriques	►	R	Ordinale	ultinomiale				
11       Surve       Réponses multiples         12       Malyse des valeurs manquantes       Mon linéaire         13       Imputation multiple       Mon diréaire         14       Echantillons complexes       Codage optimal (CATREG)         15       Courbe ROC       Imputation	10	70,00	20,00	Previsions	- F	R	Modèles de	choix binaire				
12       Image: Analyse des valeurs manquantes imputation multiple       Imputation multiple <td< th=""><th>11</th><th></th><th></th><th>Réponses multiples</th><th>- P</th><th>R</th><th>Non linéaire</th><th></th><th></th><th></th><th></th><th></th></td<>	11			Réponses multiples	- P	R	Non linéaire					
13     Imputation multiple       14     Echantillons complexes       15     Contrôle de gualité       16     Courbe ROC	12			Analyse des valeurs manquantes		R	Pondération	estimée				
14     Echantillons complexes     Codage optimal (CATREG)       15     Contrôle de gualité     Image: Courbe ROC       16     Image: Courbe ROC     Image: Courbe ROC	13			Imputation multiple	*	R	Doubles mo	indres carrés.				
15 16 1	14			Echantillons complexes	•		C <u>o</u> dage opti	mal (CATREG	)			
	15			Controle de <u>q</u> ualite								_
4	16					]						
		1										





#### Statistiques descriptives

	Moyenne	Ecart-type	И
الزيادة في المحصيل الدراسي	14,0000	3,46410	10
ساعات المذاكرة	32,0000	22,17105	10

<u>الجدول الأول</u>: يوضح الإحصاءات الوصفية للمتغيرين حيث يظهر ان المتوسط الحسابي لمتغير التحصيل قد بلغ 14بانحراف معياري3.46410 في حين أن المتوسط الحسابي لمتغير ساعات الدراسة قد بلغ 32 بانحراف معياري 22.17105 و بما ان كل البيانات كانت كاملة لكل وحدات العينة فان حجم العينة لم يتغير و هو يساوي 10.

COLLEIGUOUS
-------------

		الزيادة في المحصيل الاراسي	ساعات المذاكرة
Corrélation de Pearson	الزيادة في التحصيران الدراسي	1,000	,913
	ساعات المذاكرة	,913	1,000
Sig. (unilatérale)	الزبادة في المحصبك الاراسي		,000
	ساعات المذاكرة	,000	
N	الزيادة في التحصيران الدراسي	10	10
	ساعات المذاكرة	10	10

#### Variables introduites/supprimées<sup>b</sup>

Modèle	Variables introduites	Variables supprimées	Méthode
1	<sup>a</sup> ساعات المداكرة		Entrée

Toutes variables requises saisies.

الزيادة في المحصيل الدراسي : b. Variable dépendante

<u>الجدول الثالث ( جدول نوع الطريقة )</u>: يبين أن طريقة المربِعات الصغرى هي المتبعة في تحليل الانحدار الخطي وكذلك تم ادخال متغير مفسر واحد (مستقل) و هو متغير ساعات المذاكرة باعتبار انه المتغير الوحيد المدرج في النموذج المفسر للانحدار و بأن المتغير التابع هو متغير التحصيل.

جدول معامل التحديد :

#### Récapitulatif des modèles<sup>5</sup>

Modèle	R	R-deux	R-deux ajusté	Erreur standard de l'estimation
1	,913ª	,833	,813	1,49999

a. Valeurs prédites : (constantes), ساعات المذاكرة b. Variable dépendante : الزيادة في المحصيل الاراسي

<u>الجدول الرابع:</u> يتضمن الجدول أهم مؤشرات نموذج الانحدار و المتمثل في معامل التحديد و الذي يرمز له بالرمز R-deux)R<sup>2</sup>) و الذي يعتبر مقياسا لدقة النموذج و هو يساوي 0.833 و تفسير ذلك ان 83.3% من التغيرات التي تحدث في المتغير التابع (التحصيل الدراسي سببه المتغير المستقل ( ساعات الدراسة) و النسبة الباقية 16.7% ترجع إلى عوامل أخرى، وكلما اقتربت قيمة R<sup>2</sup> من 100% دل ذلك على دقة النموذج و ذلك لان :

جدول تحليل التباين ANOVA:

#### ANOVA<sup>b</sup>

Modèle		Somme des carrés	ddl	Moyenne des carrés	D	Sig.
1	Régression	90,000	1	90,000	40,001	,000ª
	Résidu	18,000	8	2,250		
	Total	108,000	9			

a. Valeurs prédites : (constantes), ساعات المذاكرة b. Variable dépendante : الزيادة في المحصيل الاراسي

جدول معاملات الانحدار:

**Coefficients**<sup>a</sup>

		Coefficients no	n standardisés	Coefficients standardisés			95,0% % ini confiance	tervalles de e pour B	Co	rrélations	
	Modèle	A	Erreur standard	Bêta	t	Sig.	Borne inférieure	Limite supérieure	Corrélation simple	Partielle	Partie
	1 (Constante)	9,436	,864		10,926	,000	7,444	11,427			
	ساعات المذاكرة	,143	,023	,913	6,325	,000	,091	,195	,913	,913	,913
	a. Variable dépenda	فتصبرك الاراسي : ante	الزيادة في ا	ل من	ن هي أقل مدنيا	ة بالثابت	0.00 الخاص	بة Sig=00=00	لجدول أن قيم حمد مدا	ح من ا۔	كها يتض
				ر معنوي.	الانحدار	ې نموذج	ار الثابت فج	دل أن المقد	، 0.05 مما ي	، المعنوية	مستوى
				ل أن	.0 مما يد	من 05	0 وهي أقل	ر هي 000.	لخاصة بالمتغير	L Sig ä	ونجد قيم
							ار معنوي.	فوذج الانحا	مدار b في ن	مل الانح	قيمة معا
				<u>راسي</u>	صيل اله	في التح	في الزيادة	يمكنه التنبأ	ات المذاكرة	فان ساء	بالتالي و
التالية:	نموذج الانحدار	تنتج معادلة	الجدول نسن	حسب							
y=9.43	6+0.143 (x )	منه	y=a و	+bx							

التحصيل الدراسي= 0.143+9.436 (ساعات المذاكرة)

#### Statistiques des résidus<sup>a</sup>

	Minimum	Maximum	Moyenne	Ecart-type	N
Prévision	10,8621	19,4200	14,0000	3,16228	10
Résidu	-1,85104	3,00316	,00000	1,41420	10
Erreur Prévision	-,992	1,714	,000	1,000	10
Erreur Résidu	-1,234	2,002	,000	,943	10

a. Variable dépendante : الزيادة في المحصيرك الاراسي

الجدول بعنوان Statistiques des Residus يستخدم لمعرفه بعض المقاييس الخاصة بالبواقي.



ta Explorer  $\times$ Liste Variables dépendantes Statistiques.. ساعات المذاكرة [X] 🏈 🔗 Standardized Residual [. + الزيادة في التحصيل الدراسي [Y] 🏈 Diagrammes. 🔗 Unstandardized Pre... Options Liste des facteurs : Bootstrap. Etiqueter les observations par -Afficher Les deux O Statistiques O Diagrammes OK Coller Réinitialiser Annuler Aide



كما يمكن الحكم على مدى اعتدالية البواقي اعتمادا على الطريقة الحسابية وذلك من خلال اختبار (كلومجروف-سمنروف) واختبار (شابيرو-ويليك)، ونحصل عليهما من خلال اتباع الخطوات التالية Analyse---Statistiques descriptives---Explorer

فتظهر لنا النافذة التالية:

نقوم بإدراج المتغير Standardized Ressiduals Liste variables dépendantes

> ثم ننتقل إلى Diagrammes فتظهر نافذة جديدة في نفس النافذة كما يلي:

في هذه النافذة نؤشر على Boites à moustaches نؤشر على Boites à moustaches نؤشر على Avec tests نؤشر على Aucun Aucun ونلغي التأشير على Tige et feuille بهدف تقليل المخرجات التي لسنا بحاجة إليها. ثم نضغط على Continue ثم ok نحصل على الجدول التالي:



#### Tests de normalité

	Koln	nogorov-Smi	irnov <sup>a</sup>	Shapiro-Wilk			
	Statistique	ddl	Signification	Statistique	ddl	Signification	
Standardized Residual	,207	10	,200*	,919	10	,348	

a. Correction de signification de Lilliefors

\*. Il s'agit d'une borne inférieure de la signification réelle.

يتضح من الجدول السابق أن قيمة Sig في كلا الاختبارين أكبر من 0.05 وهو ما يدل على أن البواقي تتبع التوزيع الطبيعي. وبالعودة إلى نافذة البيانات Affichage de données وبالعودة إلى نافذة البيانات E\_Affichage de المعيارية (Standardized Ressiduals) بعنوان IE\_1 وعمود آخر للقيم المقدرة للمتغير التابع Yبعنوان PRE\_1 ويمكن الاعتماد على البيانات الخاصة بالقيم المعيارية للبواقي عند دراسة اعتدالية التوزيع الاحتمالي للبواقي حسابيا.

X	Y	PRE_1	ZRE_1
10,00	10,00	10,86212	-,57475
11,00	10,00	11,00475	-,66984
14,00	12,00	11,43264	,37824
15,00	12,00	11,57527	,28315
20,00	13,00	12,28843	,47439
25,00	13,00	13,00158	-,00105
46,00	19,00	15,99684	2,00212
50,00	15,00	16,56736	-1,04491
59,00	16,00	17,85104	-1,23403
70,00	20,00	19,41998	,38668

## تفسير معادلة الانحدار • يدل على أنه كلما زادت كمية ساعات المذاكرة ساعة واحدة حدث زيادة في التحصيل بمقدار 0.143 درجة



في الجدول التالي ،المتغير x هو تكلفة إنتاج فيلم سينمائي (بملايين الدولارات) والمتغير y هو مردود هذا الفيلم:

التكلفة (x)	62	90	50	35	200	100	95
المردود (y)	65	64	48	57	601	146	47

أرسم مخطط الانتشار

a) أوجد معادلة خط الانحدار.

- b قدّر مردود فيلم بلغت تكلفته 55 مليون دولار.
- 💽 أوجد مقدار الخطأ لفيلم بلغت تكلفته 90 مليون دولار.

