

|                                      | Null Hypothesis (H <sub>0</sub> ) is true<br>He truly is not guilty | Alternative Hypothesis (H <sub>1</sub> ) is true<br>He truly is guilty |
|--------------------------------------|---|--|
| Accept Null Hypothesis<br>Acquittal  | Right decision  | Wrong decision<br>Type II Error  |
| Reject Null Hypothesis<br>Conviction | Wrong decision<br>Type I Error                                      | Right decision   |

$$T = \frac{\sum D}{\sqrt{((n \sum D^2) - (\sum D)^2) / (n-1)}}$$

$$t = 11 / \sqrt{((8 \times 123) - 11^2) / 7}$$

$$= 11 / \sqrt{(984 - 121) / 7}$$

$$= 11 / \sqrt{123.286}$$

$$= 11 / 11.103$$

$$= 0.991$$

Combien de df ?  
Df = n-1 = 7

Nous allons à la p. 412 et cherchons le point critique pour distribution t et obtenons t = 2.37. Alors... ?

Si la valeur dans le tableau est > notre valeur calculée, alors nous ne pouvons rejeter H<sub>0</sub>

**Type I** : Condamnation d'un innocent **Incorrectement rejeter l'hypothèse nulle**

**Type II** : Acquiescement d'un coupable **Ne pas avoir rejeté l'hypothèse nulle, alors qu'elle aurait dû être rejetée**

**Coefficient de Pearson = non ordonné** Toujours entre -1 et +1. Plus proche de +/- de 1 = relation linéaire parfaite. Seuils de considération de relation. Lien mais pas nécessairement un lien de causalité. Coefficient de corrélation : stats qui résume ampleur et direction

| Correlation Coefficient | Interpretation                        |
|-------------------------|---------------------------------------|
| .00 - .19               | Slight, almost negligible correlation |
| .20 - .39               | Low, quite small correlation          |
| .40 - .59               | Moderate correlation                  |
| .70 - .89               | High correlation                      |
| .90 - 1.00              | Very high correlation                 |

|            | Homme                             | Femme                             | Total                                |
|------------|-----------------------------------|-----------------------------------|--------------------------------------|
| Fumeur     | 10<br>50.0 %<br>20.0 %<br>10.0 %  | 10<br>50.0 %<br>20.0 %<br>10.0 %  | 20<br>100.0 %<br>20.0 %<br>20.0 %    |
| Non-fumeur | 40<br>50.0 %<br>80.0 %<br>40.0 %  | 40<br>50.0 %<br>80.0 %<br>40.0 %  | 80<br>100.0 %<br>80.0 %<br>80.0 %    |
| Total      | 50<br>50.0 %<br>100.0 %<br>50.0 % | 50<br>50.0 %<br>100.0 %<br>50.0 % | 100<br>100.0 %<br>100.0 %<br>100.0 % |

Tableaux croisés. **just a faire aXd x bXc (tableau de contingence pour RC).**

Risque relatif • **Le risque relatif** est le rapport des risques absolus dans les deux groupes • RR : RA<sub>E</sub> + RA<sub>NE</sub>  
Rapport cote = • RC = RC<sub>E</sub> + RC<sub>NE</sub>  
Étude de cohorte = débute avec l'exposition et ont le suit jusqu'au résultat (étudier si l'autre groupe peut développer le résultat des gens en exposition)

**prospectif**  
Étude cas témoins = commence avec le résultat et fonctionne à l'envers pour trouver le résultat  
cas = gens qui sont malade  
témoins = n'ont pas le résultat

**rétrospectif**  
Indice de risque = décrire résultat de risque et prendre meilleur décision clinique  
Indice absolu = quantifier le montant réel de risque lié à expo diff

Indice relatif = comparer le risque dans les 2 groupes d'expositions  
Relation entre 2 variable = fonction  
Désigne expérimental = groupe contrôle (référence)  
groupe expérimentale (résultat)

**One sample t test** = H<sub>0</sub> aucune diff. sig. compare le score moyen d'un échantillon à une valeur connue (usuellement moyenne de population). **df = n-1**. Plus petit que 0.05 = rejette hypothèse nulle).

$$T = \frac{(x - u)}{SEM} \quad SEM = \frac{SD}{\sqrt{n}}$$

x = moyenne d'échantillon.  
u = moyenne population. SEM = standard error of mean.

Hypothèse nulle (H<sub>0</sub>) = les moyennes des deux groupes sont pareilles. Pas liées et pas significatif. (P > 0,05) (t tableau > t calculé)

Hypothèse alternative (H<sub>1</sub>) = moyennes sont différent. **Sont** liées et significatif. (p < or = 0,05) (t tableau < t calculé)

**T test à deux échantillons** = compare moyennes de deux échantillons différents pour voir si ce sont les mêmes.. Utilisé lorsque variable dépendante = intervalle ou de ratio (ex. poids). **Dépendant = tester 2x même groupe ou groupe lier. ex: before and after surgery.** **df = (nombres de paires) - 1**;  $t = \sum D \div SED$  ou = (oublie pas racine) D = diff. De score par paire

**Variable indépendante** est sur niveau nominal avec **seulement** deux niveaux (hommes vs femmes). **df = n1 + n2 - 2**. **Formule de variance mise en commun = conserve H<sub>0</sub>. Lorsque test F est non significatif.** p (Sig) > 0,05 la valeur nulle est conservé et les variances sont égales. **Formule de variance séparée** = Test Levene diff. significatif, then u need this formule to guess SED and calcul test t. **si p (Sig) < 0,05 la valeur nulle est rejeté et les variances sont inégales.**

Moyenne de distribution d'échantillon est toujours égale à la différence moyenne dans la population.

**L'erreur type de la différence (SED)** = erreur standard. Estimation

$$SED = \left[ \sqrt{(sd^2 \div na) + (sd^2 \div nb)} \right] \quad sd^2 = \text{variance de population. na} = \text{taille de l'échantillon A. nb} = \text{taille d'échantillon B.}$$

**Postulats**; échantillon au hasard indépendant dans les 2 populations. Variable de résultat est normalement distribué dans 2 pop. Variances des 2 pop sont égales (l'hypothèse de l'homogénéité de la variance).

**Robustesse** = n'a pas d'incidence sur la prise de décision statistique si; Taille de l'échantillon est grande (>40 per group) ou taille des échantillons dans les deux groupes sont similaires (<1,5 fois le nombre de personnes dans un groupe comme dans l'autre groupe)  
**Test Kolmogorov-Smirnov** = pour tester la normalité. (positif = 2 échantillons sont distribué normale). (Sig (valeur p) dans SPSS).

**Levene's F test** = test pour variance égal (nous indique si les variances des deux échantillons sont statistiquement différentes). Positif = variance des 2 échantillons est plus ou moins égale. (Sig (valeur p) dans SPSS). Détermine quel formule de variance prendre.

**Intervalles de confiance (IC)** = utilise différence entre les deux moyennes de groupe, pour indiquer la précision de l'estimation de l'échantillon. IC = (M1 - M2) +/- SED x t tableau. Et df pour corriger IC. (M1 - M2 = différence moyennes).

| t-Test for Equality of Means |        |      |                   |                 |                       |
|------------------------------|--------|------|-------------------|-----------------|-----------------------|
|                              | t      | df   | Sig. (Two-Tailed) | Mean Difference | Std. Error Difference |
| Equal variances assumed      | -1.847 | 18   | .081              | -10.00          | 5.41                  |
| Equal variances not assumed  | -1.847 | 17.9 | .081              | -10.00          | 5.41                  |

**SPSS** = Bilatéral, pas sig. Diff. Analyze - compare means - indépendant sample t test - test variable = variable dépendante - grouping variable = variable indépendant.

**ANOVA** = compare moyennes de groupe lorsqu'il y a plus de deux groupes.

**One-way** = km d'auto varie entre usine. 1 caractéristique mais 3 groupe indépendant.

**Two-way** = km d'auto varie ET modèle d'auto entre usine. 2 caractéristiques.

**N-way** = Encore plus de facteurs. 3 ou plus caractéristique.. N = nombre de différents facteurs.

**RM-ANOVA à mesures répétée** = compare moyennes du même groupe testé plus d'une fois dans plusieurs scénario ou période de temps.

**Variables dépendante** = Intervalle ou ratio ( km )

**Indépendantes** = Niveau nominale (3 niveau). Niveau ordinaire (poids normal, surpoids, obésité)

**Hypothèse nulle** = moyennes entre les groupes sont égales

**Hypothèse alternative** = au moins une des moyennes de groupe n'est pas égale

**Postulats**; Échantillonnage aléatoire vient des populations à l'étude. Variable dépendante est normalement distribué dans les populations Variances

dans les populations sont égales (homogénéité des variances); tested with Levene's F test.

**Robustesse** = Oublie postulats si n par group > 20. Need each échantillon to have similar amounts of people (n). **Partitioning variance** = isoler les raisons que les scores peuvent différer l'un à l'autre.

**Variance entre les groupes** = différences entres groupes comparés (indépendant).

**Variance intra-groupe** = différences individuelles entre personnes dans les groupes.

**One-way ANOVA**

F = t<sup>2</sup>. F = variance entre les groupes ÷ variance intra-groupe. ou F = Effet de VI + erreur d'échantillonnage ÷ l'erreur d'échantillonnage. Si VI n'as aucun effet; la test f = 1.0. Distributions d'échantillonnage F sont asymétriques autour de la valeur de 1. **Df<sup>B</sup>** = pour comparaison entre groupes (n de catégories nominales - 1). **Df<sup>B</sup> = k-1**. **Df<sup>W</sup>** = intragroupes. (n total de sujets - n de groupes ou catégories). **Df<sup>W</sup> = n-k**. K = # de groupes. N = # de sujets.

**Sommes des carrés** = somme des écarts des carrés.

**F = (SS<sup>B</sup> ÷ df<sup>B</sup>) ÷ (SS<sup>W</sup> ÷ df<sup>W</sup>) or F = MS<sup>B</sup> ÷ MS<sup>W</sup>** (moyennes carrés). Si F calculé > F du tableau, les valeurs sont statistiquement significatif donc hypothèse nulle est rejetée. **SS<sup>B</sup>** = somme des carrés entre les groupes. **SS<sup>W</sup>** = sommes des carrés intra groupes. **MS<sup>W</sup>** = moyenne carré intra-groupe; la somme des carrés intra, divisé par df intra. **df<sup>W</sup> = N - nombres de groupes.** **MS<sup>B</sup>** = moyenne carré inter groupes; la somme des carrés inter, divisé par df inter. **df<sup>B</sup> = nombres de groupes - 1.**

|                      | Groupe 1 (musique) | Groupe 2 (relaxation) | Le groupe 3 (contrôle) |
|----------------------|--------------------|-----------------------|------------------------|
| Les scores de stress | 0                  | 1                     | 5                      |
|                      | 6                  | 4                     | 6                      |
|                      | 2                  | 3                     | 10                     |
|                      | 4                  | 2                     | 8                      |
|                      | 3                  | 0                     | 6                      |
| Somme                | 15                 | 10                    | 35                     |
| Moyenne              | 3.0                | 2.0                   | 7.0                    |

Nous calculons maintenant quelque chose qu'on appelle la "moyenne globale" qui est la moyenne de toutes les moyennes

$$(3.0 + 2.0 + 7.0) / 3 = 12 / 3 = 4,0$$

| Moyenne de chaque groupe (j) | Moyenne - moyenne globale (=4,0) | (Moyenne - moyenne globale) <sup>2</sup> | N (nombre de personnes dans chaque groupe) | Multiplier deux dernières colonnes |
|------------------------------|----------------------------------|--|--|------------------------------------|
| 3.0                          | -1                               | 1  | 5  | 5,0                                |
| 2.0                          | -2                               | 4  | 5  | 20,0                               |
| 7.0                          | 3                                | 9  | 5  | 45,0                               |
| Somme                        |                                  |  |  | 70,0                               |

C'est SS<sub>B</sub>

Groupe 1 (moyenne = 3,0)

| x     | x - moyenne | Carré |
|-------|-------------|-------|
| 0     | -3          | 9     |
| 6     | 3           | 9     |
| 2     | -1          | 1     |
| 4     | 1           | 1     |
| 3     | 0           | 0     |
| Somme |             | 20,0  |

Groupe 3 (moyenne = 7,0)

| x     | x - moyenne | Carré |
|-------|-------------|-------|
| 5     | -2          | 4     |
| 6     | -1          | 1     |
| 10    | 3           | 9     |
| 8     | 1           | 1     |
| 6     | -1          | 1     |
| Somme |             | 16,0  |

Groupe 2 (moyenne = 2,0)

| x     | x - moyenne | Carré |
|-------|-------------|-------|
| 1     | -1          | 1     |
| 4     | 2           | 4     |
| 3     | 1           | 1     |
| 2     | 0           | 0     |
| 0     | -2          | 4     |
| Somme |             | 10,0  |

$$SS_W = 20,0 + 10,0 + 16,0 = 46,0$$

$$F = (SS_B / df_B) / (SS_W / df_W)$$

$$F = (70,0 / 2) / (46,0 / 12)$$

$$= 9,13$$

**SPSS** =

Valeur F est sig.

|                | Sum of Squares | df | Mean Square | F     | Sig. |
|----------------|----------------|----|-------------|-------|------|
| Between Groups | 70,000         | 2  | 35,000      | 9,130 | ,004 |
| Within Groups  | 46,000         | 12 | 3,833       |       |      |
| Total          | 116,000        | 14 |             |       |      |

**Chi carrée**

Pour tester une relation entre variable catégorielle (hiérarchique). **Hypothèse nulle = variables catégorielles sont indépendantes.** Non lié  
**Hypothèse alternative = Variables ne sont pas indépendant. sont liés.**

**Statistiques** : Échantillon aléatoire. Observation indépendante. (pas de avant après). fréquence prévue = plus que zero. fréquence observé = plus que 5; (si plus bas que 5 = the test exacte de Fisher). Si valeur attendue plus bas que 5 = the correction de continuité de Yates. Réduit la valeur du  $\chi^2$  calculé et permet d'avoir un estimé plus conservateur).

Calculer le Chi Carrée

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}$$

where  
 $\chi^2$  = Pearson's cumulative test statistic,  
 $O_i$  = an observed frequency,  
 $E_i$  = an expected (theoretical) frequency, asserted by the null hypothesis,  
 $n$  = the number of cells in the table.

Les données attendues.

|               | Bas       | Moyen     | Haut      | Les marginaux |
|---------------|-----------|-----------|-----------|---------------|
| Bas           | 32x41/171 | 32x61/171 | 32x69/171 | 32            |
| Moyen         | 47x41/171 | 47x61/171 | 47x69/171 | 47            |
| Haut          | 41x92/171 | 61x92/161 | 69x92/171 | 92            |
| Les marginaux | 41        | 61        | 69        | 171           |

| Cellule | O  | A    | (O-A) | (O-A) <sup>2</sup> | (O-A) <sup>2</sup> /A |
|---------|----|------|-------|--------------------|-----------------------|
| 1       | 13 | 7.7  | 5.3   | 28.1               | 3.6                   |
| 2       | 11 | 11.4 | -0.4  | 0.2                | 0                     |
| 3       | 8  | 12.9 | -4.9  | 24.0               | 1.9                   |
| 4       | 10 | 11.3 | -1.3  | 1.7                | 0.2                   |
| 5       | 18 | 16.8 | 1.2   | 1.4                | 0.1                   |
| 6       | 19 | 19   | 0     | 0                  | 0                     |
| 7       | 18 | 22.1 | -4.1  | 16.8               | 0.8                   |
| 8       | 32 | 32.8 | -0.8  | 0.6                | 0                     |
| 9       | 42 | 37.1 | 4.9   | 24.0               | 0.6                   |

$\chi^2 = \text{total} = 7.2$

**Df** = (r-1) x (c-1). r = rows. c = colonnes.

Hypothèse nulle = valeur calculé < valeur tableau

Hypothèse alternative = valeur calculé > valeur tableau.

Significant.

Pearson chi-square is only important **SPSS** line, look at sig 2 tailed.

**Test U de Mann-Whitney** = tests hypothèse que les deux échantillons sont identiques en population. (plus de 20 par groupe). Utilise quand variable catégorielle est ordinale. Ont regarde juste valeur p (sig). Valeur p indique association, si plus petit que 0,05 = significant. (équivalent non-paramétrique de test t indépendants).

**Test de Kruskal-Wallis (H)** = tests hypothèse nulle que trois distributions ou plus sont identiques dans le cas de variables ordinales. (équivalent non-paramétrique de l'ANOVA à une voie).

| Nombre de groupes     | Niveau de mesure (variable dépendante) |                        |
|-----------------------|--|------------------------|
|                       | Mesures nominales                      | Des mesures ordinales  |
| Deux groupes          | Test du khi-carré                      | Mann-Whitney U Test    |
| Trois groupes ou plus | Test du Khi-carré                      | Test de Kruskal-Wallis |

**Test McNemar** = basically Chi carrée mais peut être lier (ex. Jumeaux ou avant après). Tout en valeur absolue.

|       | Oui | Non | Total   |
|-------|-----|-----|---------|
| Oui   | A   | B   | A+B     |
| non   | C   | D   | C+D     |
| TOTAL | A+C | B+D | A+B+C+D |

$\chi^2 = \frac{(|B-C| - 1)^2}{B+C}$  ← Test de McNemar utilise le  $\chi^2$

|       | Oui | Non | Total |
|-------|-----|-----|-------|
| Oui   | 15  | 0   | 15    |
| Non   | 5   | 30  | 35    |
| TOTAL | 20  | 30  | 50    |

$\chi^2 = \frac{(|B-C| - 1)^2}{B+C} = \frac{(|5-1|)^2}{0+5} = 16/5 = 3.2$

**Régression linéaire**

| Test            | Symbole | Type de test     | Type de données                           |
|-----------------|---------|------------------|---|
| Rho de Pearson  | $R_p$   | Paramétrique     | Ratio/intervalle                          |
| Rho de Spearman | $R_s$   | non-paramétrique | Ordinale                                  |
| Tau de Kendall  | T       | non-paramétrique | ordinaire ou ordinaire + Ratio/intervalle |

**Corrélation ordonné = Rho de Spearman** ; variables ordinales ou étendues entre variables inégale ou valeurs extrêmes. Test de corrélation de rang pour voir si 2 variable sont corrélé. Classer sujets en ordre de chaque variable. Hypothèse nulle; Rho = 0 (notre Rho < Rho tableau). Hypothèse alternative; Rho n'est pas égale à 0. (notre Rho > Rho tableau) en valeur absolue.

**RHO SPEARMAN**

$$1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

| Détenu | Échelle d'attitud e (0-100) | Rang Attitude | De temps en prison (mois) | Rang temps Prison | La différence en rangs | La différence au carré |
|--------|-----------------------------|---------------|---------------------------|-------------------|------------------------|------------------------|
| Bob    | 47                          | 3             | 14                        | 2                 | 1                      | 1                      |
| Frank  | 65                          | 2             | 35                        | 3                 | -1                     | 1                      |
| Raoul  | 85                          | 1             | 6                         | 1                 | 0                      | 0                      |
| Dan    | 6                           | 4             | 65                        | 4                 | 0                      | 0                      |

Somme = 2

$R_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} = 1 - (6x2)/(4x15) = 1 - 0,20 = 0,80$

| Étudiant | Note de mi-session (X) | X = (X - moyenne de X) | X <sup>2</sup> | Examen final (Y) | Y = (Y - moyenne de Y) | xy |
|----------|------------------------|------------------------|----------------|------------------|------------------------|----|
| 1        | 2                      | -3                     | 9              | 3                | -3                     | 9  |
| 2        | 6                      | -1                     | 1              | 7                | 1                      | 1  |
| 3        | 5                      | 0                      | 0              | 6                | 0                      | 0  |
| 4        | 9                      | 4                      | 16             | 8                | 2                      | 8  |
| 5        | 7                      | 2                      | 4              | 9                | 3                      | 6  |
| 6        | 9                      | 4                      | 16             | 10               | 4                      | 16 |
| 7        | 3                      | -2                     | 4              | 4                | -2                     | 4  |
| 8        | 4                      | -1                     | 1              | 6                | 0                      | 0  |
| 9        | 1                      | -4                     | 16             | 2                | -4                     | 16 |
| 10       | 4                      | -1                     | 1              | 5                | -1                     | 1  |
| Somme    | 50                     | 0                      | 68             | 60               | 0                      | 61 |

$b = \frac{\sum xy}{\sum x^2}$

**RÉGRESSION**

$a = (\text{moyenne Y}) - b (\text{moyenne X})$

**Tau de Kendall (T)** = calcul corrélations ordonnées.

**Régression** = processus d'observations individuelles d'être contraint de se conformer à une modèle. Souvent linéaire. Pour analyser relation entre variables et permet de faire prédictions sur valeurs et variables. (grande relation entre regression et correlation) ; Y = mx + b. m = pente, (m = changement de y/ changement de x) b is ordonnée à l'origine or Y intercept.

Équation pour ligne droite; Y = a + bX. Y is valeurs sur variable dépendante. X is valeurs sur variable indépendante. a is intercepte (constante ou ligne coupe axe vertical Y). b is pente de la ligne.

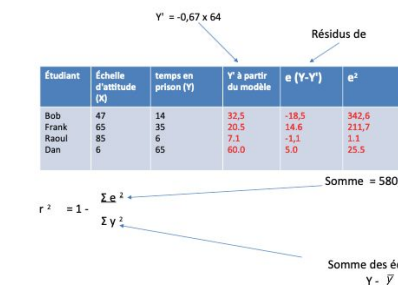
**Équation de régression** (ligne droite qui passe à travers le nuage de données le meilleure);  $Y' = a + bX$

Y' is valeur prévue de Y. X is valeur réelle de X. a is ordonnée. b is pente, mais dans context de coefficient de régression.

**Méthode des moindre carrés** = calcul distances entre chaque point et estime une ligne pour avoir la minime distance totale.

Pour calculer b et a; last colonne is the red columns multiplied. Y' = 64 - 0.67X

**CALCUL r2 (diagnostic de ligne, interprétation)**



$r^2 = 1 - \text{somme } e^2 / \text{somme } y^2$  r = racine carrée de  $r^2$ . Interprétation avec tableau Pearson.

| Détenu | Échelle d'attitude (X) | X - X̄ | X <sup>2</sup> | temps en prison (Y) | Y - Ȳ | Yy       |
|--------|------------------------|--------|----------------|---------------------|-------|----------|
| Bob    | 47                     | -3.75  | 14.06          | 14                  | -16   | 60       |
| Frank  | 65                     | 14.25  | 203.06         | 35                  | 5     | 71.25    |
| Raoul  | 85                     | 34.25  | 1173.06        | 6                   | -24   | -822     |
| Dan    | 6                      | -44.75 | 2002.56        | 65                  | 35    | -1566.25 |
| Somme  | 203                    |        | 3392.75        | 120                 |       | -2257    |

$\bar{X} = 203/4 = 50.75$

$\bar{Y} = 120/4 = 30$

$b = \frac{\sum Yy}{\sum Xx^2} = -2257/3392.75 = -0.67$

$a = \text{moyenne (Y)} - b (\text{moyenne X}) = 30 - (-0.67) x (50.75) = 64$

(temps en prison) = 64 - 0,67 (attitude score)

En utilisant ce modèle, si nous voulons prédire combien d'années en prison une personne passera si elle obtient un score de 43 sur l'échelle d'attitude, nous avons juste à faire....

temps en prison = 64 - 0,67 (43)  
 = 64 - 28.8  
 = 35,2 mois

**Erreur type d'estimation (SE<sup>estimate</sup>)** = indice de savoir à quel point une valeur prévue de Y est mal prédite. Plus le coefficient de corrélation (B) entre les deux variables dans la régression est grande, plus l'erreur type d'estimation sera petite.

$SE_{\text{Estimate}} = \sqrt{\frac{\sum (Y - Y')^2}{N-2}}$

| Étudiant | Note de mi-session (X) | Examen final (Y) | Y' à partir du modèle | e (Y-Y') | e <sup>2</sup> |
|----------|------------------------|------------------|-----------------------|----------|----------------|
| 1        | 2                      | 3                | 3.3                   | 0.3      | 0.09           |
| 2        | 6                      | 7                | 6.9                   | 0.1      | 0.01           |
| 3        | 5                      | 6                | 6.0                   | 0        | 0              |
| 4        | 9                      | 8                | 9.8                   | -1.6     | 2.56           |
| 5        | 7                      | 9                | 7.8                   | 1.2      | 1.44           |
| 6        | 9                      | 10               | 9.8                   | 0.4      | 0.16           |
| 7        | 3                      | 4                | 4.2                   | 0.2      | 0.04           |
| 8        | 4                      | 6                | 5.1                   | 0.9      | 0.81           |
| 9        | 1                      | 2                | 2.4                   | -0.4     | 0.16           |
| 10       | 4                      | 5                | 5.1                   | -0.1     | 0.01           |

$Y' = 1.5 + 0.90X$  Somme = 5,28

$SE_{\text{Estimate}} = \text{SQRT} [5.28/(10-2)] = \text{SQRT}[0.66] = 0.81$

SQRT = square root.

**Erreur moyen au carré (MSE)** = somme de  $e^2 \div n-2$ . (erreur standard [d'estimation] ou écart type des erreurs est racine carrée du MSE).

Hypothèse nulle = pente = 0.

Hypothèse alternative = pente pas = 0. Relation significative entre X et Y.

**Test de significations**; tester si pente = 0, ont calculé stats t selon  $T = b \div SE_{\text{estimate}}$ .  $df = N-2$ . Ex:  $T = 0.9 \div 0.81 = 1,11$ .

**Diagnostic de ligne** = Test de signification pour voir si pente = 0. ou statistique goodness of fit. ou  $r^2$ , le plus utiliser. Interprétation de  $r^2$ ; si  $r^2 > 0,60$  il est bon, sinon c'est faible ou mauvais. Trouver le  $r^2$ ;

[Modèle linéaire =  $Y' + bX + \text{erreur(s)}$ . Erreur ou résiduels est effet des variables omis de l'équation (pas constant).  $e' = Y' - Y''$ ]

**Régression multiple**

$Y = a + b^1 + b^2 + b^3 + b^4 \dots$

C'est plusieurs régression linéaire car facteurs sont combinés de façon linéaire sur plan multidimensionnel. analyse statistiques multivariées (+ 3 variables). Utilise critères des moindres carrés pour résoudre a et b (somme des carrés des termes d'erreur (résidus) est réduite au minimum). SPSS = valeur B pour faire équation. Diagramme de dispersion résiduelle pour voir violation de postulat.

R = coefficient de regression multiple. Indique ampleur de relation entre variable, mais pas la direction. (comment variables indépendantes prédisent ou expliquent une variable dépendante). Entre 0 and 1. Cannot be lower than r entre les prédicteurs et la VD. R diminue avec l'ajout de prédicteurs.

$R^2$  = coefficient de détermination multiple. C'est la proportion de variation du variable dépendant, expliquée par les prédicteurs. Évalue ajustement du modèle de régression multiple.

