

TP4 STATISTIQUES BIVARIEES

Réalisé par :

- PELTIER Alexandre
- THIAM Boubacar Roger

SOMMAIRE

PAGE 1_3 : Exercice 1 : TABAC

PAGE 4_6 : Exercice 2 : HERBICIDE

Exercice 1 : TABAC

Question 1

- 1) Voir Excel
- 2) **Tableau de Contingence :**

Tableau d'effectifs

	Nombre de Fumeur		
probleme_pulmonaire	fumeur	non fumeur	Total général
FALSE	19	46	65
TRUE	30	5	35
Total général	49	51	100

Tableau de Fréquences

	Fréquences de fumeur		
probleme_pulmonaire	fumeur	non fumeur	Total général
FALSE	0,19	0,46	0,65
TRUE	0,3	0,05	0,35
Total général	0,49	0,51	1

- 3) Le pourcentage de l'échantillon qui est fumeur est de 49%.
Le pourcentage de l'échantillon qui présente des problèmes pulmonaires est de 35%.
Le pourcentage de l'échantillon qui est non fumeur et avec des troubles pulmonaires est de 5%.

Question 2

1)

Profil Ligne	Fumeurs	Pas Fumeurs	Total
Problèmes Pulmonaires	0,857	0,143	1,000
Pas Problèmes Pulmonaires	0,292	0,708	1,000
Fréquence Marginale	0,490	0,510	1,000

2) Parmi les personnes présentant un trouble pulmonaire, il y a 85,7% de fumeurs

3) Dans la population on a 49% de fumeurs. Dans ces fumeurs on a 85,7% qui ont des problèmes pulmonaires. Donc il ya un lien important entre le fait de fumer et d'avoir des problèmes pulmonaires.

Dans la population on a 51% de non fumeurs. Sur ces non fumeurs seuls 14% ont des problèmes pulmonaires. Donc on peut en conclure que ne pas fumer réduit la probabilité d'avoir des problèmes pulmonaires.

Test du CHI DEUX

Le CHI DEUX vaut 29,04

Le Coefficient de Contingence vaut 0,474

Le Coefficient de Cramer vaut 0,539

De ces résultats on ne peut rien en conclure quant à la dépendance des deux caractères, il nous faudrait effectuer un test statistique.

Exercice 2 : Herbicide

Question 1 :

1) Tableau de contingence

Moyenne de taux de survivants	plante			
herbicide	ble	chiendent	liseron	Total général
aucun	0,928	0,946	0,938	0,938
herbicide1	0,849	0,225	0,776	0,617
herbicide2	0,740	0,689	0,250	0,560
herbicide3	0,114	0,156	0,066	0,112
Total général	0,658	0,504	0,507	0,556

- 2) Le plus efficace est de ne rien mettre dans la mesure où 93,8% des plants survivent. Et si on doit en mettre un il est plus judicieux d'utiliser l'herbicide 1 car 61,7% des plants survivent.

Le plus néfaste est l'herbicide 3 car seulement 11,2% des plants survivent.

3) Tableau de Contingence TAUX-PLANTE

	Données		
plante	Nombre de taux de survivants	Moyenne de taux de survivants	Var de taux de survivants
ble	80,000	65,779%	10,954%
chiendent	80,000	50,416%	13,059%
liseron	80,000	50,735%	14,194%
Total général	240,000	55,644%	13,145%

La variance inter vaut $Se^2=0.005$

Le coefficient de corrélation vaut $Cr=0.198$

D'après cela, on peut dire que Cr est plus proche de 0 que de 1. Donc on ne peut pas conclure que le taux de survivants dépend de la plante.

Tableau de Contingence TAUX-HERBICIDE

	Données		
herbicide	Nombre de taux de survivants	Moyenne de taux de survivants	Var de taux de survivants
aucun	60,000	0,938	0,002
herbicide1	60,000	0,617	0,104
herbicide2	60,000	0,560	0,067
herbicide3	60,000	0,112	0,006
Total général	240,000	0,556	0,131

La variance inter vaut $Se^2=0.087$

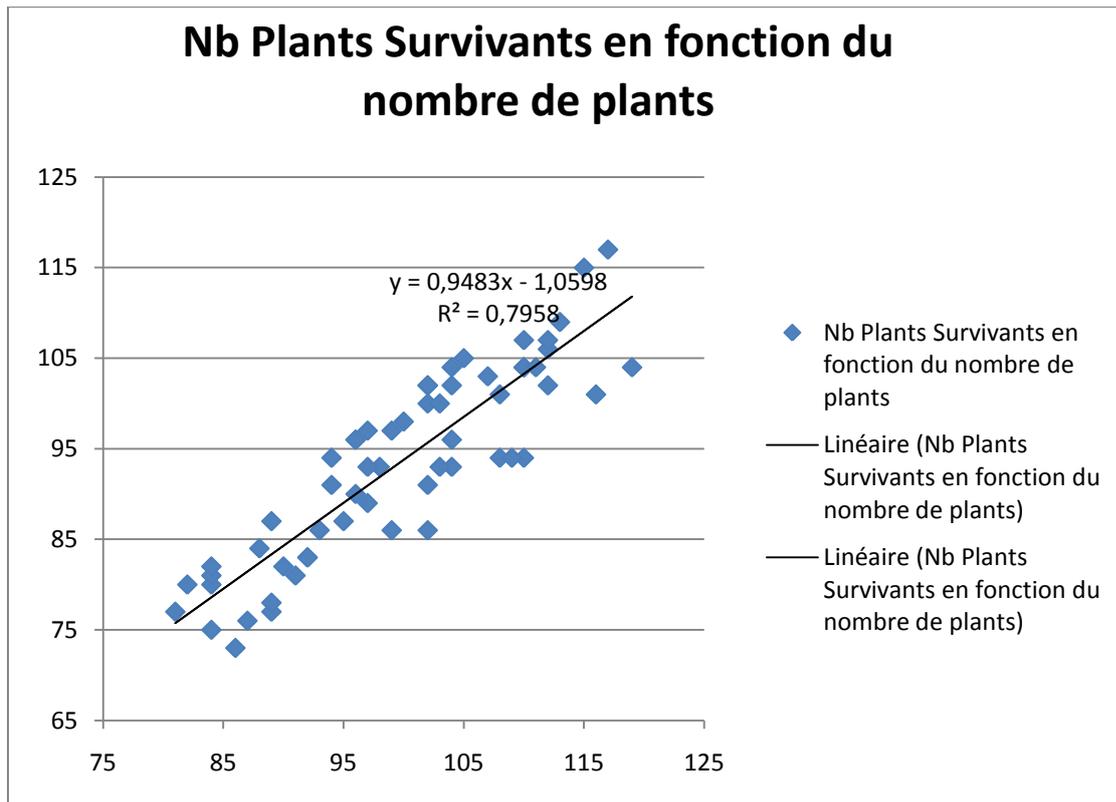
Le coefficient de corrélation vaut $Cr=0.812$

Ce coefficient de corrélation est proche de 1. Ainsi on peut en déduire que la survie des plants dépend de l'herbicide utilisé.

Pour conclure, on peut voir que ces résultats confirment nos précédentes observations.

Question 2

1)



L'équation de la droite de tendance approchant le mieux de façon linéaire notre relation est :

$$Y=0.9483X-1.0598$$

Donc pour 100 plants il y aura $Y=93$ plants qui survivront

2) Le pourcentage de la variance de y est expliqué par le coefficient de détermination R^2 . Ici il vaut 0,7958. Il est proche de 1. C'est donc une très bonne approche de notre relation. ON peut donc faire une approximation linéaire de notre courbe $Y=F(X)$.