

Protocole de TP 2 :

Test de Contingence -Test de Khy-Deux

(Comparaison entre l'effectif observé et l'effectif théorique)

Notre objectif dans ce TP est de savoir comment faire une comparaison entre deux variables qualitatives, avec k modalités pour la variable explicative X, et l modalités pour la variable à expliquer Y. En utilisant pour cette raison les techniques du test de Khi-deux sous logiciel SPSS.

Pour cela on doit prendre un exemple d'explication :

Exemple 1

On considère de deux types de traitement contre la maladie (Coronavirus 19) qui sévit en Wilaya de Biskra pendant le mois de Mars en année 2021, on choisit deux type de traitement (SINOVAC, et SINOPHARM). On souhaite évaluer si la nature du traitement influe sur la guérison des personnes ayant contracté cette maladie.

Personnes	Traitement	Guérison
1	SINOVAC	Oui
2	SINOPHARM	Oui
3	SINOVAC	Oui
4	SINOPHARM	Non
5	SINOPHARM	Oui
6	SINOVAC	Oui
7	SINOPHARM	Oui
8	SINOPHARM	Non
9	SINOVAC	Oui
10	SINOVAC	Non
11	SINOPHARM	Oui
12	SINOVAC	Oui
13	SINOPHARM	Non
14	SINOVAC	Non
15	SINOVAC	Non
16	SINOVAC	Non
17	SINOPHARM	Oui
18	SINOPHARM	Oui
19	SINOVAC	Oui
20	SINOVAC	Non
21	SINOVAC	Oui

22	SINOPHARM	Oui
23	SINOPHARM	Non
24	SINOVAC	Oui
25	SINOVAC	Non
26	SINOPHARM	Non
27	SINOPHARM	Oui
28	SINOVAC	Oui
29	SINOPHARM	Oui
30	SINOVAC	Oui
31	SINOVAC	Oui
32	SINOVAC	Non
33	SINOVAC	Oui
34	SINOVAC	Oui

Le laboratoire a étudié les statistiques concernant les 34 personnes affectées par Coronavirus 19.

- 1) Déterminer l'objectif pour cette expérience.
- 2) Déterminer les variables qualitatives.
- 3) Déterminer l'hypothèse nulle et alternative.
- 4) Tracer le tableau de contingence pour l'effectif ainsi que le pourcentage pour chaque modalité.
- 5) Que peut-on conclure pour l'effet de type de traitement sur guérison des personnes.

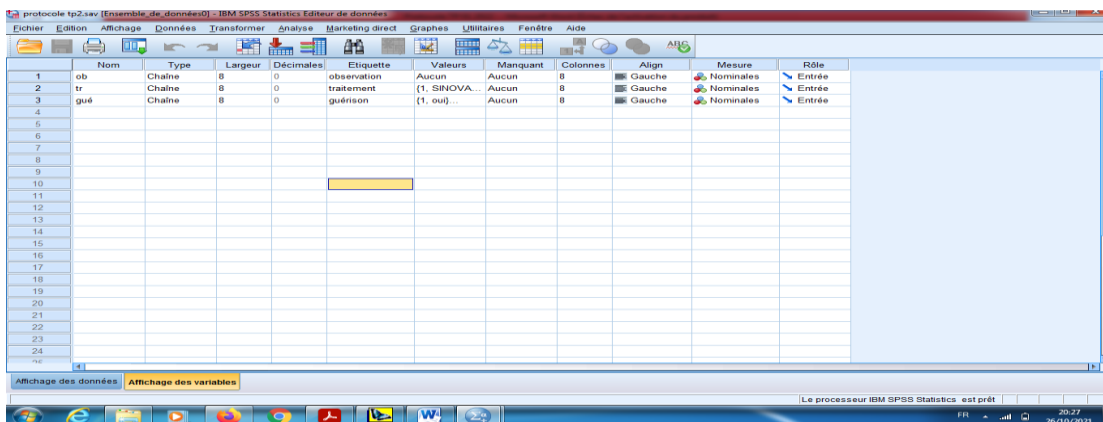
Solution

1- Introduction des Données

Tout d'abord il faut entrer ces données dans SPSS, on a déjà vu les techniques dans TP 1.

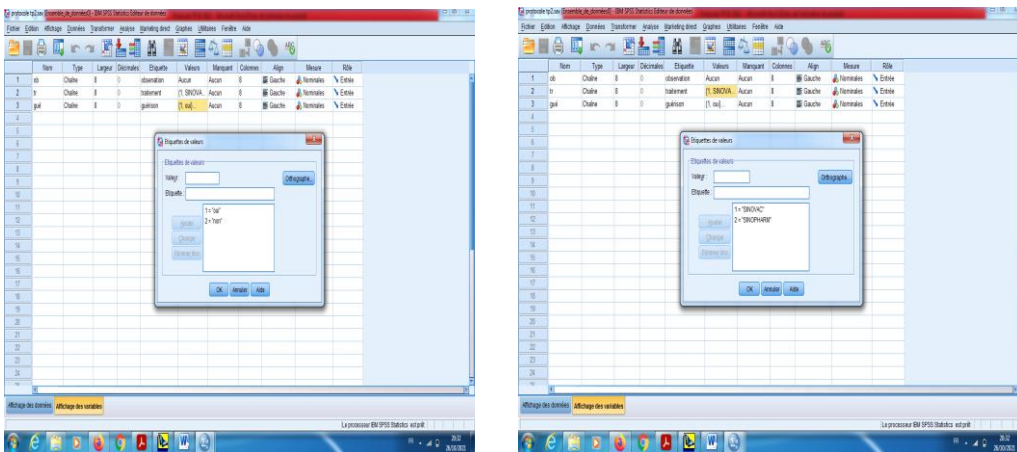
Nous suivant les étapes suivantes :

- a) Il faut définir les variables qualitatives suivantes : observations, type de traitement, et la guérison des personnes (de type chaîne, et de mesures nominales).

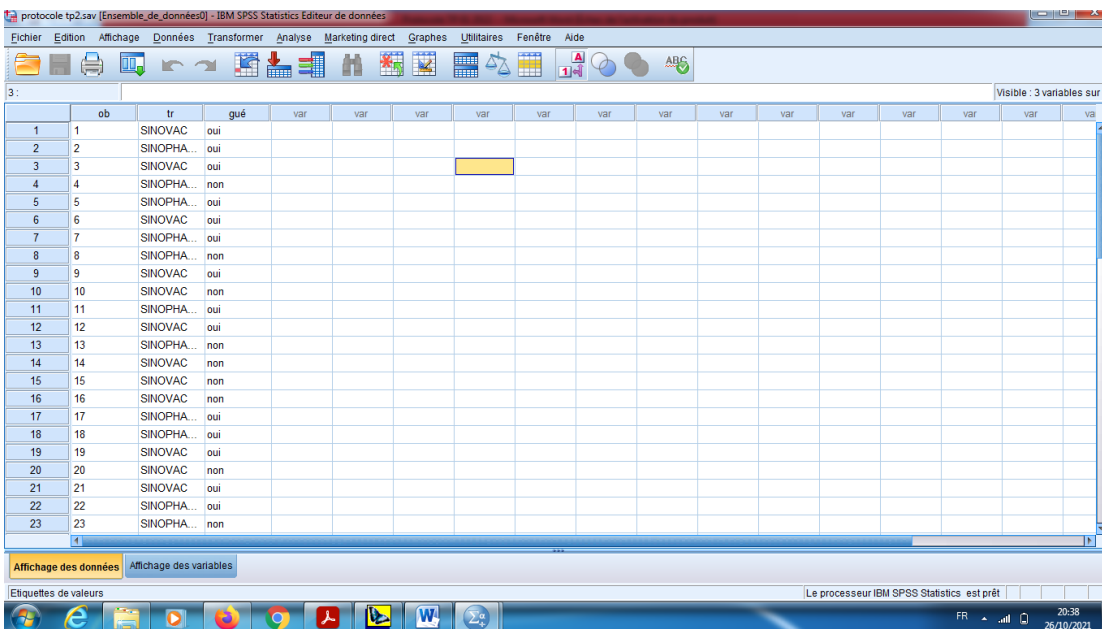


- b) On choisit les noms des modalités pour les deux variables dans « valeur », on peut prendre comme un exemple pour la variable type de traitement (1= pour le traitement « SINOVAC » et 2= pour le traitement « SINOPHARM »).

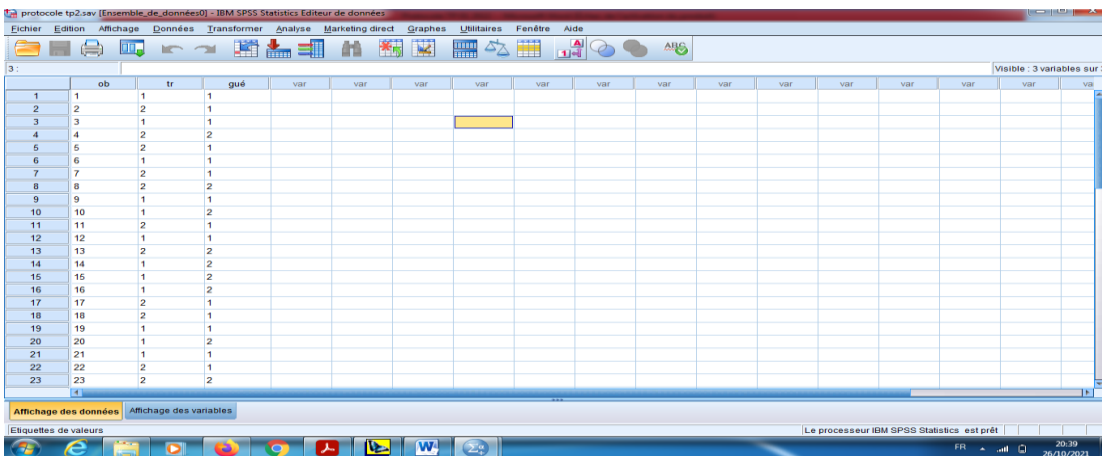
De plus pour la variable Guérison (1= pour la réponse « Oui » et 2= pour la réponse « Non »).



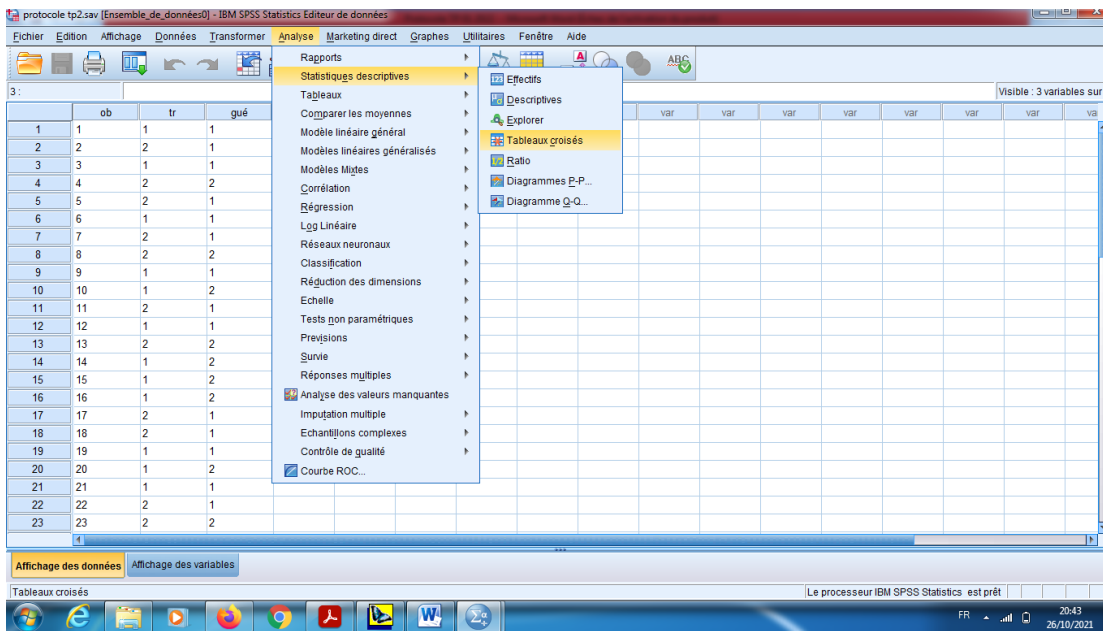
c) Introduire des données dans la barre « Affichage des données » par SPSS.



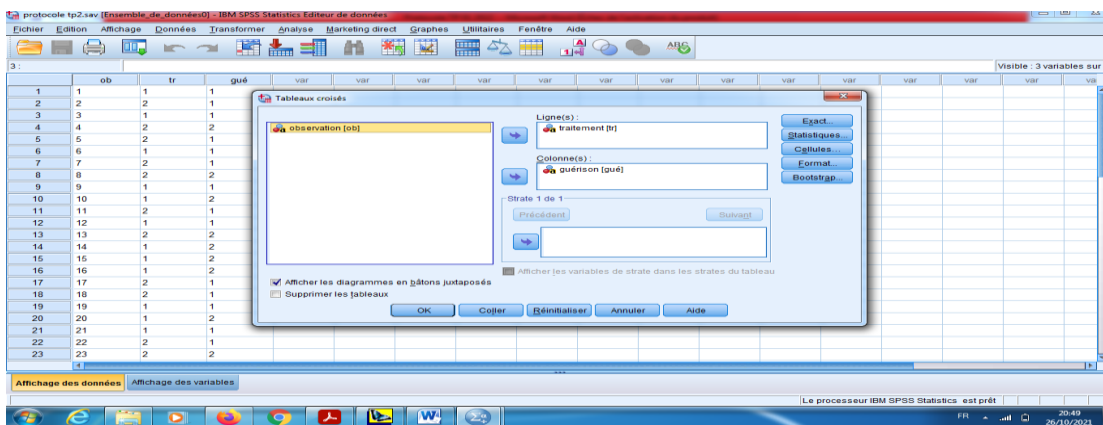
d) Pour le codage en cliquant sur le bouton « Etiquettes de valeurs » qui se trouve dans la barre de menu.



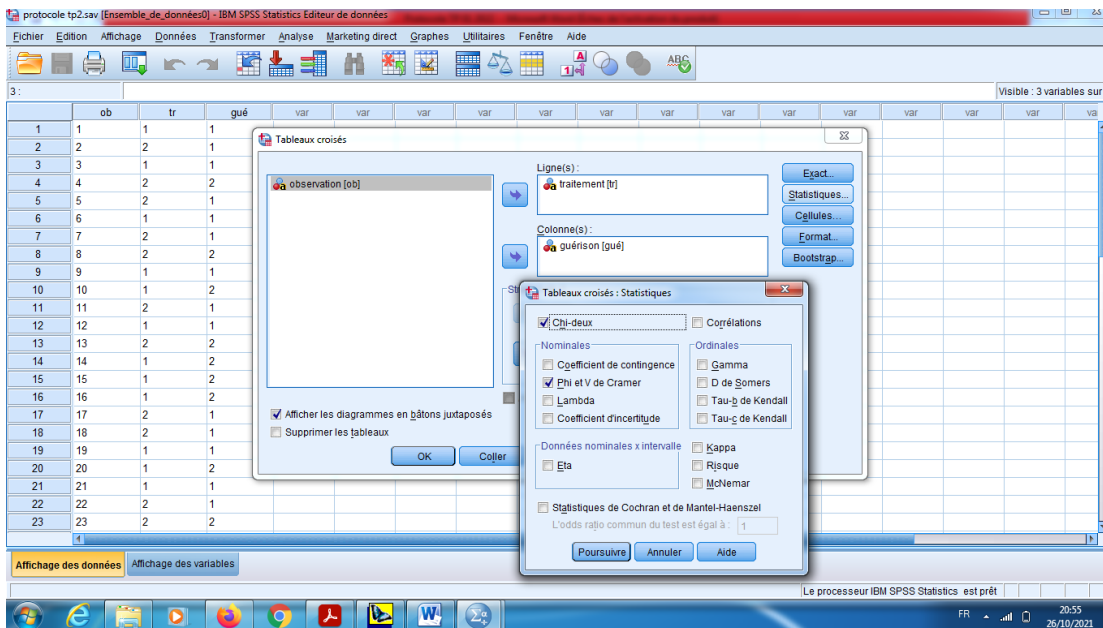
e) En suite en cliquant sur le bouton « Analyse » qui se trouve dans la barre des outils, et choisir « statistique descriptive », et puis « Tableau croisés ».



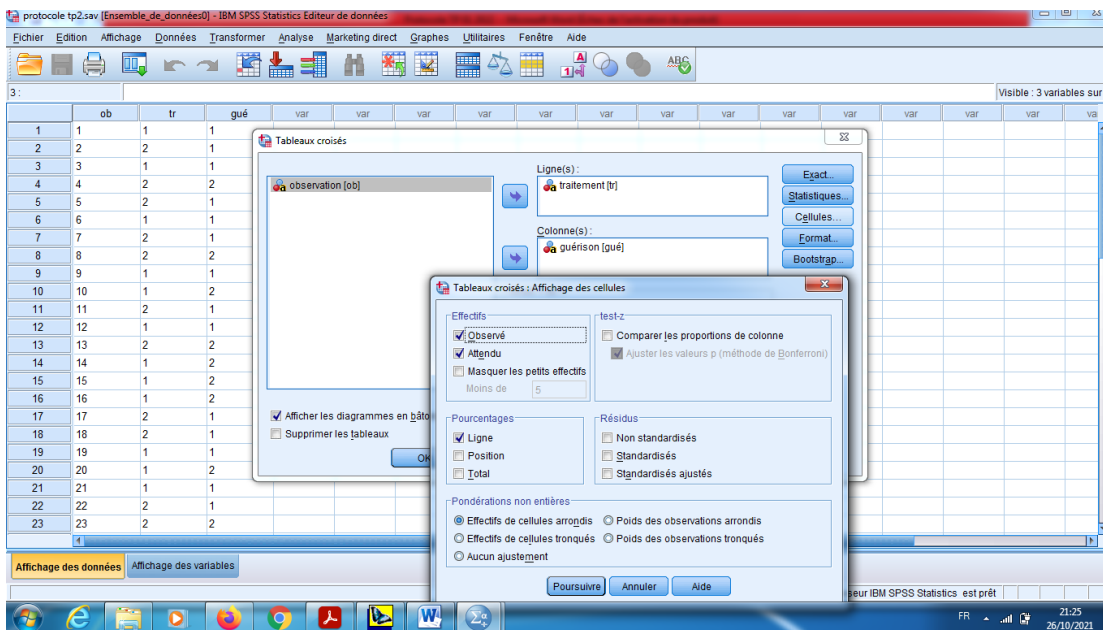
- e-1) On pose la variable qualitative explicative (le type de traitement) dans la barre « Ligne ».
- e-2) On pose la variable qualitative expliquée (la guérison) dans la barre « Colonne ».
- e-3) On choisit aussi « Afficher les diagrammes en bâtons juxtaposés ».



e-4) Dans le Boutons Statistique vous pouvez choisir le type de test « Test de Khi-deux » car les variables sont de types qualitatifs, et on choisit aussi le test de « Phi et V de Cramer » comme le meilleur test. . En cliquant en fin sur « poursuivre ».



e-5) Dans le bouton « Cellules » on peut prendre le choix « observé » pour l'effectif, et le choix « Ligne » pour le pourcentage, et « Attendu » pour l'effectif théorique. En cliquant en fin sur « poursuivre ».



e-6) On garde dans le « Format » l'ordre de ligne croissant.

f) En fin en cliquant sur « Ok ».

2- Affichages des Résultats

Dans la table 1 «_Récapitulatif du traitement des observations » on remarque qu'il y a des 34 valeurs observés et 0 valeurs manquants.

Récapitulatif du traitement des observations

	Observations					
	Valide		Manquante		Total	
	N	Pourcent		Pourcent	N	Pourcent
traitement * guérison	34	100,0%		0,0%	34	100,0%

Dans la table 2 « Tableau croisé traitement * guérison », on remarque la variable qualitative « type de traitement » se trouve dans la ligne, mais dans la colonne « Guérison ».

Par exemple pour l'effectif du traitement « SINOVAC » il y a 13 qui sont « oui guéries » et 7 non guéries.

Pour des pourcentages on peut calculer le pourcentage total pour la guérison « oui » par la formule suivante :

$$64,7\% = \frac{22}{34} \times 100,$$

Avec 22 c'est l'effectifs des personnes qui sont guéries, et 34 c'est l'effectif total pour l'échantillon.

Mais pour l'effectif théorique sous l'hypothèse nulle H0 en utilisant la formule suivante

$$12,9 = \frac{64,7 \times 20}{100},$$

et

$$7,1 = \frac{35,3 \times 20}{100}.$$

De même pour le type de traitement SINOPHARM, on calcul tout d'abord le pourcentage total pour le choix « Oui », et en suit on calcul effectif théorique sous l'hypothèse nulle H0, et on utilise les formules suivantes

$$64,7 = \frac{22}{34} \times 100,$$

$$9,1 = \frac{64,7 \times 14}{100},$$

Ainsi de suite on peut lire les autres valeurs dans le tableau croisé traitement et guérison. On veut comparer les valeurs d'effectifs observé « la répartition observé avec l'effectifs théoriques « la répartition calculés ou bien théoriques », c'est pour ça il faut choisir le test de Khi-deux.

La formule du Khi-deux utilisé c'est

$$\chi_{obs}^2 = \sum_{i=1}^2 \sum_{j=1}^2 \frac{(o_{ij} - c_{ij})^2}{c_{ij}},$$

Avec

$O_{ij} = c'$ est l'effectif observés, $C_{ij} = c'$ est l'effectifs calculés

Alors

$$\chi_{obs}^2 = \frac{(13 - 12,9)^2}{12,9} + \frac{(7 - 7,1)^2}{7,1} + \frac{(9 - 9,1)^2}{9,1} + \frac{(5 - 4,9)^2}{4,9} = 0,002$$

Tableau croisé traitement * guérison					
			guérison		Total
			Oui	non	
Traitement	SINOVAC		13	7	20
		Effectif théorique	12,9	7,1	20,0
		% compris dans traitement	65,0%	35,0%	100,0%
	SINOPHARM		9	5	14
		Effectif théorique	9,1	4,9	14,0
		% compris dans traitement	64,3%	35,7%	100,0%
			22	12	34
Effectif théorique			22,0	12,0	34,0
% compris dans traitement			64,7%	35,3%	100,0%

Pour cela on peut voir dans les tableaux obtenues dans les résultats de SPSS.

Tests du Khi-deux					
	Valeur	Ddl	Signification Asymptotique (bilatérale)	Signification exacte (bilatérale)	Signification Exacte (unilatérale)
Khi-deux de Pearson	,002 ^a	1	,966		
Correction pour la continuité ^b	,000	1	1,000		
Rapport de vraisemblance	,002	1	,966		
Test exact de Fisher				1,000	,623
Nombre d'observations valides	34				

3- Interprétation des résultats

Hypothèse nulle- H0 : « La guérison au bout de 15 jours de la maladie ne dépend pas du traitement suivi.

c-à-d que les deux critères (type du traitement et la guérison) sont indépendants ».

Hypothèse alternative- H1 : « La guérison au bout de 15 jours de la maladie est liée avec le traitement suivi. c-à-d que les deux critères (type du traitement et la guérison) sont dépendants ».

Dans le test de Khi-deux « test indépendance » de deux variables qualitatives :
 L'acceptation d'hypothèse nulle permettrait de mettre en évidence une indépendance entre les deux variables, c'est-à-dire pas d'influence entre type du traitement et la guérison.
 Pour cela on compare la valeur « Signification asymptotique (bilatérale) » avec le taux de signification 5% =0,05.

Pour la décision, on utilise souvent la règle suivante

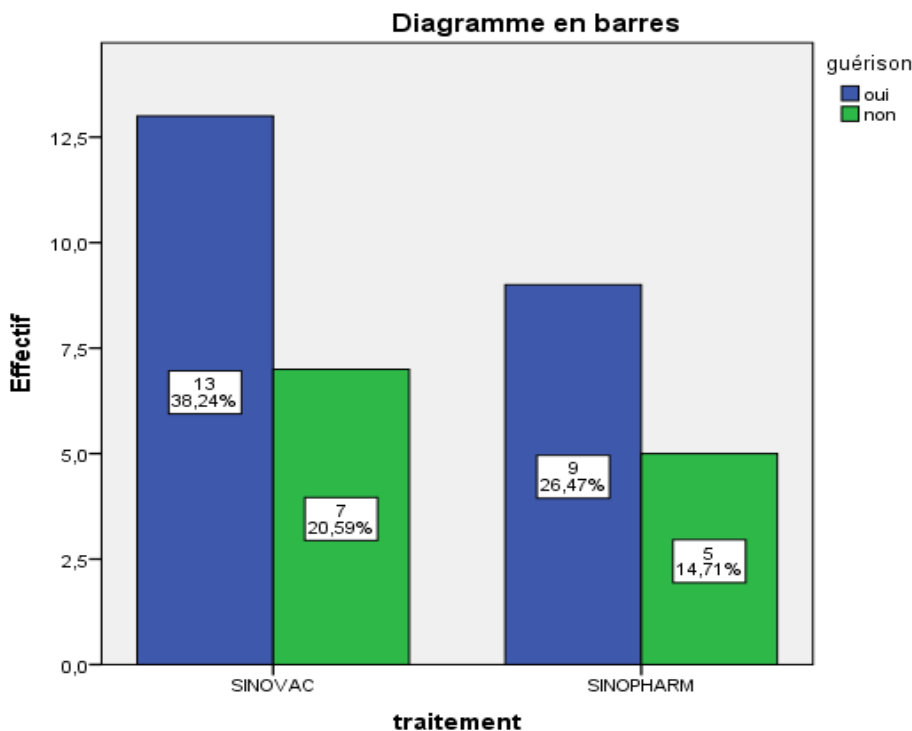
Si Signification asymptotique (bilatérale) inférieure à 0,05, alors on rejette H0 (Il y a une liaison).

Si Signification asymptotique (bilatérale) supérieure à 0,05, alors on accepte H0 (Pas de Liaison).

Tests du Khi-deux					
	Valeur	Ddl	Signification asymptotique (bilatérale)	Signification exacte (bilatérale)	Signification exacte (unilatérale)
Khi-deux de Pearson	,002 ^a	1	,966		

On remarque que la signification (bilatérale) égale à 0,966 c'est supérieure à 0,05. Donc on constate que on accepte de H0, alors les deux critères ne sont pas liés.

Finalement la représentation graphique nous donne



Pour le traitement SINOVAC on a 13 des personnes sont guéries avec pourcentage de 38,24%, et 7 personnes non guéries avec pourcentage de 20,59%.

D'autre part pour le traitement SINOPHARM on remarque qu'il y a 9 personnes qui sont guéries avec un pourcentage de 26,47% et 5 personnes non guéries avec 14,71%.

De plus $13+7+9+5=34$.

$38,24\%+20,59\%+26,47\%+14,71\%=100\%$.

Exercice 2

Un sondage est réalisé sur 44 étudiants universitaires pour savoir si les étudiants sont pour ou contre le durcissement de la vérification de sécurité à la cité universitaire en fonction son niveau. Il y a deux réponses possibles (Favorable ou Défavorable), et trois tranches de niveau possibles (L1 et L2, L3, M1 et M2), les résultats des sondages sont dans le tableau ci-dessous

Etudiants	Réponse	Niveau universitaire
1	Favorable	L1
2	Favorable	L1
3	Favorable	L1
4	Favorable	L3
5	Défavorable	L1
6	Défavorable	L3
7	Défavorable	L1
8	Favorable	L1
9	Favorable	L3
10	Défavorable	M1
11	Défavorable	L1
12	Favorable	L1
13	Favorable	L3
14	Favorable	L1
15	Défavorable	M1
16	Défavorable	L1
17	Défavorable	L1
18	Favorable	L1
19	Favorable	L3
20	Défavorable	M1
21	Favorable	L1
22	Défavorable	L1
23	Défavorable	M1
24	Favorable	L3
25	Défavorable	M1
26	Favorable	M1
27	Favorable	L3
28	Favorable	M1
29	Favorable	M1
30	Favorable	M1
31	Défavorable	L1
32	Défavorable	L1
33	Favorable	M1

34	Favorable	M1
35	Défavorable	L3
36	Défavorable	L1
37	Favorable	L1
38	Favorable	M1
39	Défavorable	L1
40	Favorable	L1
41	Favorable	M1
42	Favorable	M1
43	Défavorable	L3
44	Favorable	M1

- 1) Déterminer l'objectif pour cette expérience.
- 2) Déterminer les variables qualitatives.
- 3) Déterminer l'hypothèse nulle et alternative.
- 4) Tracer le tableau de contingence des effectifs (effectifs observés et effectifs théoriques, ainsi que le pourcentage pour chaque modalité.
- 5) Que peut-on conclure pour l'effet des niveaux universitaires pour l'étudiant sur ses réponses.