3 - Statistique Déductive

Etude de la liaison entre 2 caractères qualitatifs.

1 - Comparaison de 2 pourcentages observés.

$$\varepsilon = \frac{p_A - p_B}{\sqrt{\frac{p_A q_A}{n_A} + \frac{p_B q_B}{n_B}}}$$

ε < 1,96 avec α <5 %

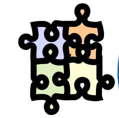
q = probabilité complémentaire de p = 1 – p

2 - Test du χ^2

$$\chi^2 = \sum \frac{(O_i - C_i)^2}{C_i}$$

χ^2 tabulé

Nb ddl = (nb lignes-1)(nb colonnes-1)



Exercice déjà traité : Test du $\chi 2$

On veut étudier l'efficacité d'un nouveau traitement T contre la leucémie.

On administre T à 50 souris et le traitement de référence R à 50 autres souris de la même espèce. On note au bout d'un mois, 33 morts dans le groupe T et 44 morts dans le groupe R.

Peut on conclure à la supériorité de ce nouveau traitement?

1. On dispose de 2 groupes indépendants traités T ou R : variables qualitatives
2. Dénombrements des morts dans chaque groupe (mort ou non) : variables qualitatives
3. Comparaison de % (vu précédemment), ou test du χ^2
4. H0 : Il n'y a pas de différence significative entre les 2 traitements.

Exercice déjà traité :Test du $\chi 2$

Ces calculs sont justes !!

	T	R	TOTAL
Morts	33 (38,5)	44 (38,5)	77
Vivants	17 (11,5)	6 (11,5)	23
TOTAL	50	50	100

77% DC

23% VIVANTS

77% DC

Gr T : 77% de 50 = 38,5

puis 23% de 50=11,5,

etc ..Gr R …

$$\chi^2 = \sum \frac{(O_i - C_i)^2}{C_i}$$

$\chi 2$ = 6,83 ddl = 1

$\chi 2$ théorique = 3,84 ; α = 5%

$\chi 2$ calculé > $\chi 2$ théorique : **on rejette H0**

Il existe une différence significative entre les 2 traitements (α < 5%).

On a le droit d'interpréter le contenu du tableau :

on peut conclure que T est meilleur que R sur cet échantillon.

PLAN GÉNÉRAL DU COURS

1 - La méthode Statistique en Médecine

2 - Statistique Descriptive

3 - Statistique Déductive

- ***Liaisons entre caractères qualitatifs***
- ***Liaisons entre caractères qualitatifs et quantitatifs***
- ***Liaisons entre caractères quantitatifs***
- ***Tests non paramétriques***

3 - Statistique Déductive

Liaison entre caractères qualitatifs et quantitatifs

Question : En moyenne la taille des individus d'une population *A* coïncide-t-elle avec la taille des individus d'une population *B* ?

1 - Comparaison de moyennes

n_1 et n_2 > 30 "Grands échantillons"

$$\varepsilon = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Table de l'écart réduit

$\varepsilon < 1{,}96$ avec $\alpha < 5\ \%$

2 - Test t de Student

n_1 ou n_2 < 30 "Petits échantillons"

$$t = \frac{m_1 - m_2}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}}$$

Table t de Student

Nb ddl = $(n_1-1)+(n_2-1)$

$$s = \sqrt{\frac{\sum (x_i - m_1)^2 + \sum (x_j - m_2)^2}{(n_1 - 1) + (n_2 - 1)}}$$

= écart type sur les 2 échantillons

Exercice Comparaison de moyennes

On cherche à comparer les taux de T3 libre (hormone thyroïdienne) chez des femmes prenant un contraceptif oral et chez des femmes n'en prenant pas.

On dispose, après TAS de 2 groupes de femmes :

Femmes sans c.o $n_1=50$ $m_1=2$ nmol $s_1=0{,}35$ nmol

Femmes avec c.o $n_2=33$ $m_2=2{,}5$ nmol $s_2=0{,}30$ nmol

Les taux de T3 libre peuvent ils être considérés comme « identiques »

dans les 2 groupes ou bien sont ils significativement différents ?

Exercice Comparaison de moyennes

1. H0 : m_1 et m_2 ne sont guères différentes. Ce sont 2 estimateurs de la valeur moyenne de T3 libre chez la femme, en général.

2. Relation entre caractères qualitatifs (prise ou non de c.o) et quantitatifs (dosages de T3 libre, ici valeur moyenne)

3. n_1 et $n_2 > 30$ ⇒ test de comparaison de moyennes

4. $$\varepsilon = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{2{,}5 - 2}{\sqrt{\frac{0{,}35^2}{50} + \frac{0{,}30^2}{33}}} = 6{,}94$$

α= 0,0001 ε= 3,89 Très significatif. **On rejette H0 avec p< 0,0001**

TAS : donc médicalement, résultat généralisable :

La prise de contraceptifs oraux augmente le taux de T3libre

Université Nice Sophia Antipolis

FACULTE DE MEDECINE – UNIVERSITE DE NICE SOPHIA ANTIPOLIS

Table de l'écart réduit

α

		0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	∞	2,576	2,326	2,17	2,054	1,96	1,881	1,812	1,751	1,695
0,1	1,645	1,598	1,555	1,514	1,476	1,44	1,405	1,372	1,341	1,311
0,2	1,282	1,254	1,227	1,2	1,175	1,15	1,126	1,103	1,08	1,058
0,3 α	1,036	1,015	0,994	0,974	0,954	0,935	0,915	0,896	0,878	0,86
0,4	0,842	0,824	0,806	0,789	0,772	0,755	0,739	0,722	0,706	0,69
0,5 ε	0,674	0,659	0,643	0,628	0,613	0,598	0,583	0,568	0,553	0,539
0,6	0,524	0,51	0,496	0,482	0,468	0,454	0,44	0,426	0,412	0,399
0,7	0,385	0,372	0,358	0,345	0,332	0,319	0,305	0,292	0,279	0,266
0,8	0,253	0,24	0,228	0,215	0,202	0,189	0,176	0,164	0,151	0,138
0,9	0,126	0,113	0,1	0,088	0,075	0,063	0,05	0,038	0,025	0,013

Table pour les petites valeurs de la probabilité

0,001	0,000 1	0,000 01	0,000 001	0,000 000 1	0,000 000 01	0,000 000 001
3,2905	3,89059	4,41717	4,89164	5,32672	5,73073	6,10941

Exercice Comparaison de moyennes

On teste un antiviral diminuant le nb de jours de symptômes cliniques chez des patients infectés par le virus de la grippe.

Soit 100 sujets non traités, atteints de grippe.

Nb moyen de jours avec symptômes m_1= 4,74 jours et s_1= 1.

Soit 100 autres sujets traités avec l'antiviral et atteints de grippe

Nb moyen de jours avec symptômes m_2 = 4,2 jours et s_2=1,7.

Parmi les propositions suivantes choisir celles qui sont exactes :

Exercice Comparaison de moyennes

A) Le test de comparaison de moyennes permet de rejeter/garder H0

B) Le test de comparaison de moyennes ne permet pas de rejeter/garder H0

C) On ne pourra pas conclure à l'efficacité du tt à cause du risque de 1ère espèce 5%

D) On ne pourra pas conclure à l'efficacité du tt à cause du risque de 2ème espèce inconnu

E) On ne pourra pas conclure à l'efficacité du tt car l'étude n'a pas été bien menée

Réponse A. H0 : m1 et m2 ne sont pas significativement différentes
En effet le test de comparaison de moyennes répond à la question : peut on accepter H0? Et éventuellement rejeter H0…

Réponse E. L'étude aurait due être randomisée (TAS) : Traitement contre Placebo

Les items **B, C, D** sont faux (B), ou sans rapport (C, D)

Table de l'écart réduit

α

		0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	∞	2,576	2,326	2,17	2,054	**1,96**	1,881	1,812	1,751	1,695
0,1	1,645	1,598	1,555	1,514	1,476	1,44	1,405	1,372	1,341	1,311
0,2	1,282	1,254	1,227	1,2	1,175	1,15	1,126	1,103	1,08	1,058
0,3	1,036	1,015	0,994	0,974	0,954	0,935	0,915	0,896	0,878	0,86
0,4	0,842	0,824	0,806	0,789	0,772	0,755	0,739	0,722	0,706	0,69
0,5	0,674	0,659	0,643	0,628	0,613	0,598	0,583	0,568	0,553	0,539
0,6	0,524	0,51	0,496	0,482	0,468	0,454	0,44	0,426	0,412	0,399
0,7	0,385	0,372	0,358	0,345	0,332	0,319	0,305	0,292	0,279	0,266
0,8	0,253	0,24	0,228	0,215	0,202	0,189	0,176	0,164	0,151	0,138
0,9	0,126	0,113	0,1	0,088	0,075	0,063	0,05	0,038	0,025	0,013

Table pour les petites valeurs de la probabilité

0,001	0,000 1	0,000 01	0,000 001	0,000 000 1	0,000 000 01	0,000 000 001
3,2905	3,89059	4,41717	4,89164	5,32672	5,73073	6,10941

3 - Statistique Déductive

Liaison entre caractères qualitatifs et quantitatifs

Question : En moyenne la taille des individus d'une population *A* coïncide-t-elle avec la taille des individus d'une population *B* ?

1 - Comparaison de moyennes

n_1 et $n_2 > 30$ "Grands échantillons"

$$\varepsilon = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Table de l'écart réduit

$\varepsilon < 1{,}96$ avec $\alpha < 5$ %

2 - Test t de Student

n_1 ou $n_2 < 30$ "Petits échantillons"

$$t = \frac{m_1 - m_2}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}}$$

Table t de Student

Nb ddl = $(n_1-1)+(n_2-1)$

$$s = \sqrt{\frac{\sum (x_i - m_1)^2 + \sum (x_j - m_2)^2}{(n_1 - 1) + (n_2 - 1)}}$$

= écart type sur les 2 échantillons

Exercice t de Student

Soient un groupe de 15 femmes obèses, et un autre groupe de 12 femmes de poids normal. On a mesuré le taux de corticoïdes sanguins moyens à l'intérieur de ces 2 groupes.

Gr 1 :	$n_1 = 15$	$m_1 = 6{,}3$	$s_1 = 1{,}8$
Gr 2 :	$n_2 = 12$	$m_2 = 4{,}5$	$s_2 = 1{,}6$

L'obésité a-t-elle une influence sur le taux de corticoïdes ?

Exercice t de Student

1. H0 : m_1 et m_2 ne sont pas différentes dans ces 2 groupes.

2. Relation entre caractères qualitatifs (obèses et non obèses), et quantitatifs (valeurs de dosages sanguins, valeurs moyennes)

3. n_1 et n_2 < 30 petits échantillons ⟹ t de student

4. Calcul de l'écart type commun aux 2 groupes. La formule s'écrit :

$$s^2 = \frac{(n_1-1) \times s_1^2 + (n_2-1) \times s_2^2}{(n_1+n_2-2)} = 2{,}53$$

nb ddl = (15+12) – 2 = 25

$$t = \frac{6{,}3 - 4{,}5}{\sqrt{\frac{2{,}53}{15} + \frac{2{,}53}{12}}} = 2{,}92$$

> 2,06 à 5% lu dans la table t Student

On rejette H0 avec α= 5%. Relation entre obésité et augmentation du taux de corticoïdes au niveau de ces échantillons.

TABLE DU t DE STUDENT

ddl \ α	0,9	0,5	0,3	0,2	0,1	**0,05**	0,02	0,01	
1	0,158	1	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,142	0,816	1,386	1,886	2,92	4,303	6,965	9,925	31,598
3	0,137	0,765	1,25	1,638	2,353	3,182	4,541	5,841	12,924
4	0,134	0,741	1,19	1,533	2,132	2,776	3,747	4,604	8,61
5	0,132	0,727	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,718	1,134	1,44	1,943	2,447	3,143	3,707	5,959
7	0,13	0,711	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,13	0,706	1,108	1,397	1,86	2,306	2,896	3,355	5,041
9	0,129	0,703	1,1	1,383	1,833	2,262	2,821	3,25	4,781
10	0,129	0,7	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,697	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,695	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,694	1,079	1,35	1,771	2,16	2,65	3,012	4,221
14	0,128	0,692	1,076	1,345	1,761	2,145	2,624	2,977	4,14
15	0,128	0,691	1,074	1,341	1,753	2,131	2,602	2,947	4,073
17	0,128	0,689	1,069	1,333	1,74	2,11	2,567	2,898	3,965
18	0,127	0,688	1,067	1,33	1,734	2,101	2,552	2,878	3,922
19	0,127	0,688	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,687	1,064	1,325	1,725	2,086	2,528	2,845	3,85
….									
25	0,127	0,684	1,058	1,316	1,708	**2,06**	2,485	2,787	3,725

Liaison entre caractères qualitatifs et quantitatifs.

Cas des séries appariées. Méthode des couples

Exemple

Comparer deux méthodes de dosage de la glycémie. On dispose de n patients, auxquels on prélève 2 tubes de sang. On dose la glycémie dans chacun de ces tubes par une méthode différente.

On souhaite comparer les valeurs moyennes de ces 2 séries de n résultats. La question posée est :

Ces 2 méthodes de dosage fournissent elles des résultats identiques ?

On calcule **si n>30** $\varepsilon = m_d \,/\, \sqrt{\frac{s^2}{n}}$ **si n<30** $t = m_d \,/\, \sqrt{\frac{s^2}{n}}$

Avec d=différence des résultats pour un même sujet, m_d=moyenne des d, n=nb de couples, s=variance des différences

Puis la méthodologie est identique aux tests déjà vus : on compare cette valeur calculée aux valeurs dans la table adaptée, et la conclusion se fait de la même manière **en fixant un risque α.**

Cas des séries appariées. Méthode des couples

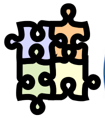
On souhaite évaluer l'intérêt d'une substance S capable de désintoxiquer les fumeurs. On constitue par T.A.S. 2 groupes de 40 fumeurs, un reçoit S, l'autre reçoit un placebo P. Le traitement dure 2 mois pour les 2 groupes. La consommation de cig/jours C est notée avant et après traitement.

	S (n=40)		P (n=40)	
	m_1	s_1^2	m_2	s_2^2
C avant tt	19,5	54,2	16,5	35,6
C après tt	5,4	30,4	3,8	20,1
Variation de C	14,1	9,1	12,7	8,9

1)Quelle est la première précaution à prendre ?

C identique dans les 2 groupes? les 2 groupes doivent être comparables vis-à-vis paramètres susceptibles d'influencer la réponse au traitement (âge, sexe, CSP, conso/jour etc..).

Si ce n'est pas le cas, il faut en tenir compte lors des conclusions.



Cas des séries appariées. Méthode des couples

Comparaison des conso moyennes avant tt dans les 2 groupes:

1. **H0 = les moyennes des conso sont équivalentes dans les 2 groupes.**
2. **Etude liaison entre variables qualitatives (S ou P) et quantitatives(nb cig/j) dans 2 échantillons indépendants**
3. **n> 30** **Test de comparaison de moyennes**
4. $$\varepsilon = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{19{,}5 - 16{,}5}{\sqrt{\frac{54{,}2}{40} + \frac{35{,}6}{40}}} = 2{,}00 > 1{,}96 \quad (\varepsilon \text{ pour } \alpha = 5\%)$$

On rejette H0, avec un risque α = 5%. Il existe donc une différence significative entre les conso moyennes des 2 groupes : Il faudra en tenir compte lors de l'étude de la variation de cette conso avant/après traitement.

Université Nice Sophia Antipolis

Cas des séries appariées. Méthode des couples

2) Dans le groupe Placebo, la conso moyenne après tt diffère t elle de la valeur avant tt ? Interpréter le résultat.

P (n=40)	
m_2	s_2^2
16,5	35,6
3,8	20,1
12,7	8,9

1. Liaison entre variable qualitative (avant / après tt) et quantitative (nb cig/j)

2. Echantillons non indépendants (méthode des couples)

3. n >30 ➡ Test de comparaison de moyennes

$$\varepsilon = m_d \,/\, \sqrt{\frac{s^2}{n}} = 12{,}7 \,/\, \sqrt{\frac{8{,}9}{40}} = 26{,}9 \qquad > 1{,}96 \text{ au risque } \alpha = 5\%$$

<u>On rejette H0.</u> Il existe une différence très significative (p < 0,001) entre les consommations avant / après tt, dans le groupe P.

Effet psychologique, envie de profiter de l'étude pour arrêter de fumer ?

Cas des séries appariées. Méthode des couples

3) Les 2 groupes diffèrent ils pour leur conso moyenne après tt?

S (n=40)		P (n=40)	
m_1	s_1^2	m_2	s_2^2
19,5	54,2	16,5	35,6
5,4	30,4	3,8	20,1
14,1	9,1	12,7	8,9

1. **H0 = les moyennes des conso sont équivalentes dans les 2 groupes.**

2. **Liaison entre variables qualitatives (S ou P) et quantitatives(nb cig/j) dans 2 échantillons indépendants**

3. **n> 30 → Test de comparaison de moyennes**

4. $$\varepsilon = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{5{,}4 - 3{,}8}{\sqrt{\frac{30{,}4}{40} + \frac{20{,}1}{40}}} = 1{,}42 \quad < 1{,}96 \text{ avec } \alpha = 5\%$$

On ne rejette pas H0 : il n'existe pas de différence significative entre les 2 groupes pour la consommation après tt.

Cas des séries appariées. Méthode des couples

4) Les 2 groupes diffèrent ils pour la variation de conso avant/après tt ?

Il faut comparer les variations avant / après tt dans les 2 groupes afin de prouver l'intérêt de la substance S

1. **H0 : Il n'existe pas de différence entre les variations de consommation dans les 2 groupes**
2. **Etude liaison entre variables qualitatives (S ou P) et quantitatives(nb cig/j) dans 2 échantillons indépendants**
3. **n> 30 >> Test de comparaison de moyennes**
4. $$\varepsilon = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{14{,}1 - 12{,}7}{\sqrt{\frac{9{,}1}{40} + \frac{8{,}9}{40}}} = 2{,}09 \quad > \quad 1{,}96 \text{ au risque } 5\%$$

<u>On rejette H0 :</u> il existe une différence significative entre les variations de conso dans les 2 groupes (p<5%) . Conclusion : efficacité de S. Il y avait eu TAS, donc résultat généralisable.

PLAN GÉNÉRAL DU COURS

1 - La méthode Statistique en Médecine

2 - Statistique Descriptive

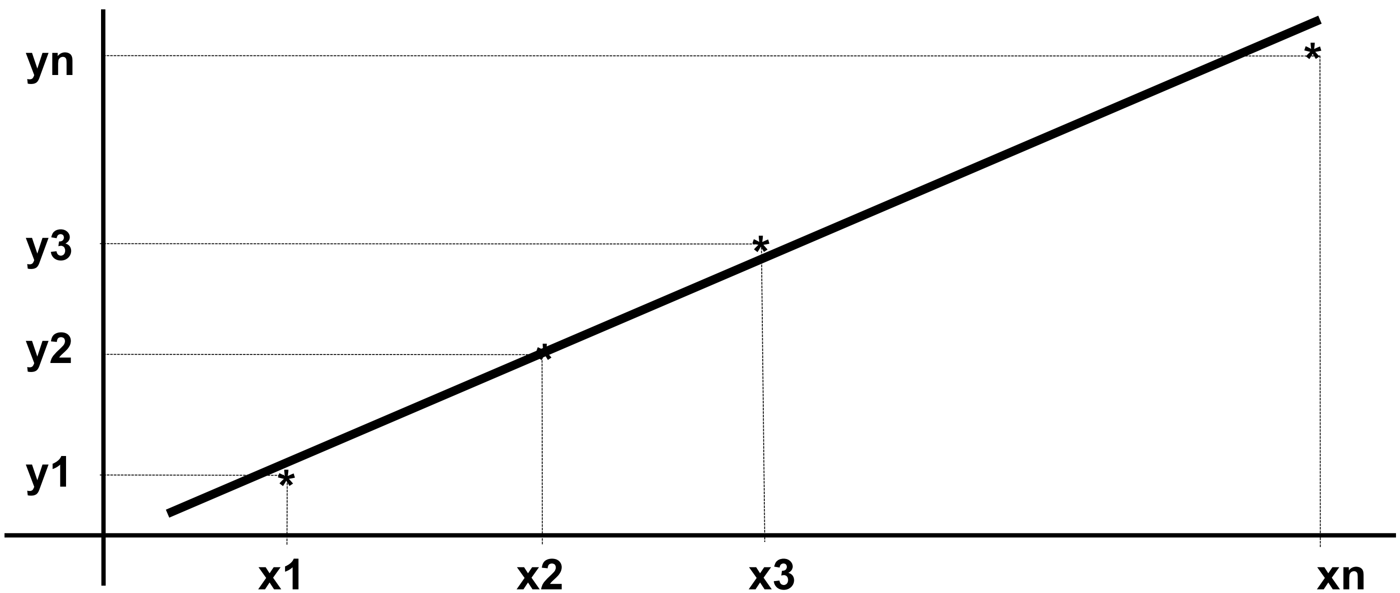
3 - Statistique Déductive

- ***Liaisons entre caractères qualitatifs***
- ***Liaisons entre caractères qualitatifs et quantitatifs***
- ***Liaisons entre caractères quantitatifs***
- ***Tests non paramétriques***

Etude de la liaison entre caractères quantitatifs

- consommation de cigarettes et capacité respiratoire sont-elles liées ?
- poids des bébés à la naissance et âge de la mère sont ils liés?

si x et y liées alors ⇒ y = f(x) *droite de régression de y en x*



Droite de régression:

Droite des moindres carrés, passe au «plus près» de chaque point du graphe

Etude de la liaison entre caractères quantitatifs

coefficient de corrélation = **pente** de cette droite. Soient (x_i,y_i) les couples dont on cherche à étudier la corrélation, m_x et m_y sont les moyennes des x_i et y_i respectivement

$$r = \frac{\sum (x_i - m_x)(y_i - m_y)}{\sqrt{\sum (x_i - m_x)^2 \sum (y_i - m_y)}}$$

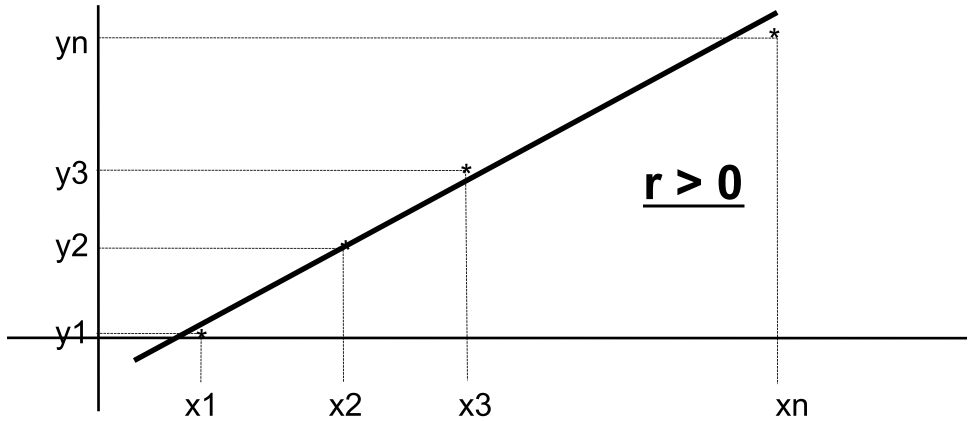
Ou bien

$$r = \frac{\sum xy - \frac{\sum x . \sum y}{n}}{\sqrt{(\sum x^2 - \frac{(\sum x)^2}{n})(\sum y^2 - \frac{(\sum y)^2}{n})}}$$

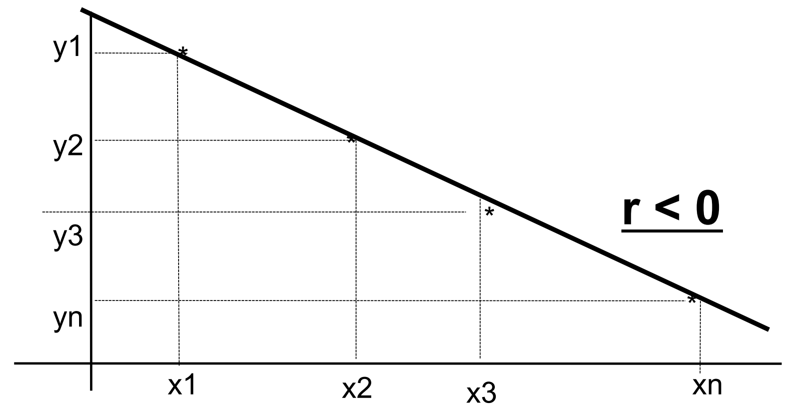
r est toujours < 1

Si r > 0 liaison positive : x et y varient dans le même sens
Si r < 0 liaison négative : x et y varient en sens inverse

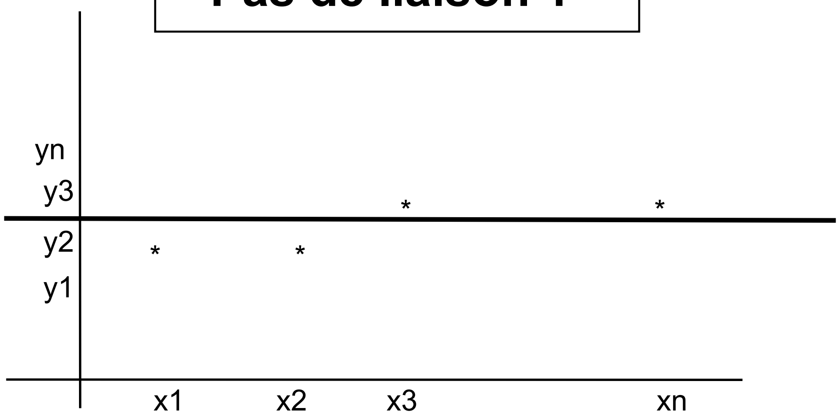
Liaison positive



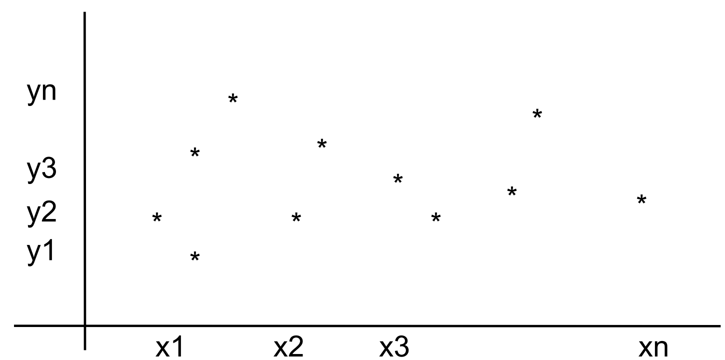
Liaison négative



Pas de liaison 1



Pas de liaison 2

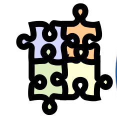


Sur un échantillon de 10 sujets d'âges différents, on recueille les données suivantes : âge (années) et concentration de cholestérol dans le sang (g/L)

X âges	30	60	40	20	50	30	40	20	70	60
Y chol	1,6	2,5	2,2	1,4	2,7	1,8	2,1	1,5	2,8	2,6

Le taux de cholestérol est il lié à l'âge ?

Exercice Coefficient de corrélation



Existe-t-il un lien entre ces 2 séries de données? Ou bien s'agit t il de 2 séries totalement indépendantes?

H0 = Le taux de cholestérol est indépendant de l'âge

H1 = Le taux de cholestérol est lié à l'âge

2 variables quantitatives >>> **Test du coefficient de corrélation.**

$$r = \frac{\sum xy - \frac{\sum x . \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)\left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}}$$

X âges	30	60	40	20	50	30	40	20	70	60
Y chol	1,6	2,5	2,2	1,4	2,7	1,8	2,1	1,5	2,8	2,6

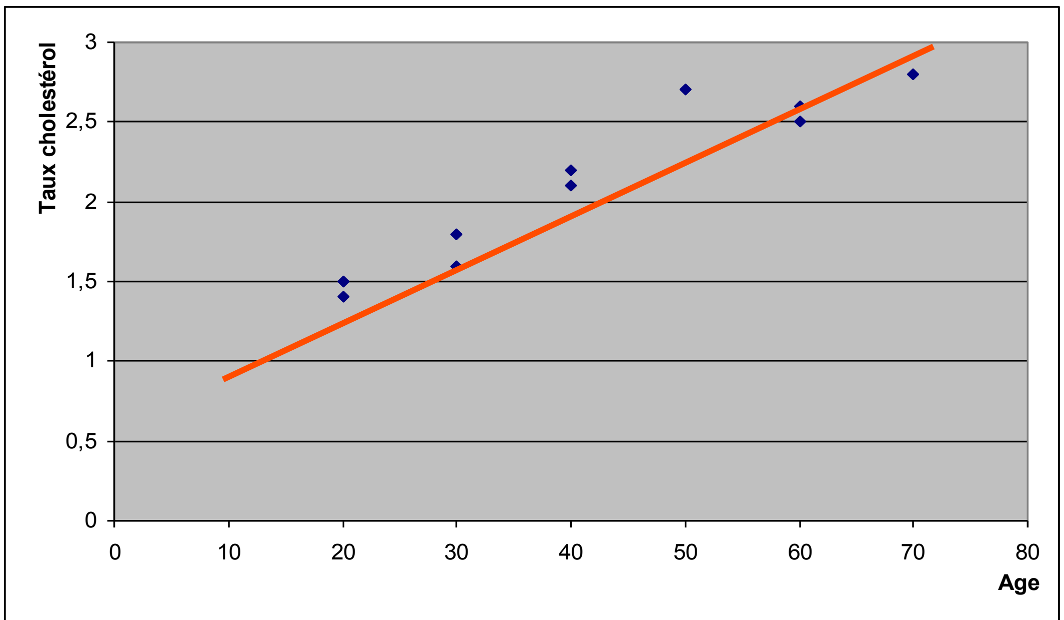
Exercice Coefficient de corrélation (suite)

r calculé = 0,955

r théorique (α = 5%) lu dans la table avec 10-2 = 8 ddl = 0,632

r calculé > r théorique **Rejet de H0**

Il existe une relation significative (α= 5%) entre l'âge et le taux de cholestérol



Plus l'âge augmente, plus le taux de cholestérol augmente

Exercice Coefficient de corrélation

On teste un traitement favorisant la baisse de tension artérielle.

On cherche à savoir si l'effet de ce traitement est lié à l'âge des patients.
Dans ce but, on effectue 2 prises de tension :

- une avant traitement,
- et une autre 1h après le traitement.

On note la différence entre les 2 valeurs.

Soit X la série des âges des patients
Soit Y la série des différences de T.A avant – après traitement.

La question posée peut se traduire par **Y = f(X)** ?

Posons tout d'abord H0..

H0 : les 2 séries x et y sont indépendantes, et il n'existe pas de relation entre elles.

Exercice Coefficient de corrélation

x = âge	y = diff	xy
64	1	64
52	9	468
56	9	504
60	6	360
62	2	124
58	4	232
57	3	171
61	1	61
58	6	348
55	9	495

$$\underbrace{\sum x = 583 \quad \sum y = 50 \quad \sum xy = 2827}$$

$$r = \frac{\sum xy - \frac{\sum x . \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)\left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}}$$

Soit $r = -0{,}83$

avec un ddl =10-2 = 8

r théorique au risque 1% : **r= 0,76**

r calculé **-0,83** est très > (en valeur absolue).

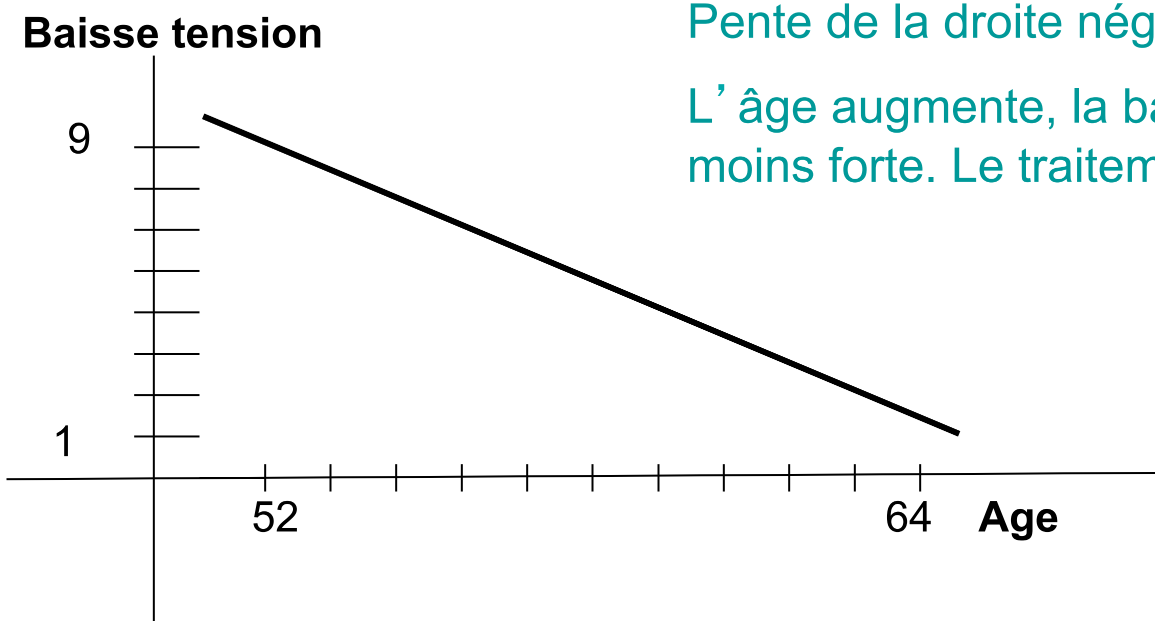
Exercice Coefficient de corrélation (suite)

Nous pouvons rejeter H0.

Il existe une relation entre x et y (elles sont « corrélées »), avec p < 1%

r calculé **-0,83** :

le signe **–** indique que plus les valeurs d'une série augmentent, plus les valeurs de l'autre série diminuent. Elles varient en sens inverse.



Pente de la droite négative.

L'âge augmente, la baisse de tension est moins forte. Le traitement est moins efficace

Table du coefficient de corrélation

ddl \ α	0,1	0,05	0,02	0,01
1	0,9877	0,9969	0,9995	0,9999
2	0,9	0,95	0,98	0,99
3	0,8054	0,8783	0,9343	0,9587
4	0,7293	0,8114	0,8822	0,9172
5	0,6694	0,7545	0,8329	0,8745
6	0,6215	0,7067	0,7887	0,8343
7	0,5822	0,6664	0,7498	0,7977
8	0,5494	0,6319	0,7155	0,7646
9	0,5214	0,6021	0,6851	0,7348
10	0,4973	0,576	0,6581	0,7079
11	0,4762	0,5529	0,6339	0,6835
12	0,4575	0,5324	0,612	0,6614
13	0,4409	0,5139	0,5923	0,6411
14	0,4259	0,4973	0,5742	0,6226
15	0,4124	0,4821	0,5577	0,6055
16	0,4	0,4683	0,5425	0,5897
17	0,3887	0,4555	0,5285	0,5751
18	0,3783	0,4438	0,5155	0,5614
….				
100	0,1638	0,1946	0,2301	0,254

PLAN GÉNÉRAL DU COURS

1 - La méthode Statistique en Médecine

2 - Statistique Descriptive

3 - Statistique Déductive

- ***Liaisons entre caractères qualitatifs***
- ***Liaisons entre caractères qualitatifs et quantitatifs***
- ***Liaisons entre caractères quantitatifs***
- ***Tests non paramétriques***

TESTS NON PARAMÉTRIQUES

- Utilisation obligatoire si les effectifs sont trop faibles
- (< 5), avec des caractères quantitatifs. Dans ce cas, les populations ne se distribuent pas normalement.
- Présentent une excellente robustesse

1) LIAISON QUANTITATIFS / QUALITATIFS

- U de Mann & Whitney

2) LIAISON ENTRE QUANTITATIFS

- Spearman (= Corrélation)

3 - Statistique Déductive

	TESTS	
	Paramétriques	**Non paramétriques**
Comparaison de 2 échantillons indépendants	**t Student** **Comp moyennes.**	**Mann-Whitney**
Comparaison de 2 échantillons appariés	**t Student** **Comp moyennes**	**Wilcoxon**
Test de corrélation	**Coefficient r**	**Coefficient r' Spearman**

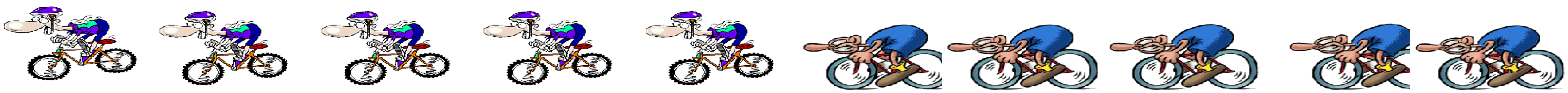
PRINCIPE

Il est nécessaire de transformer les données quantitatives en données de mesures ordinales (les rangs).

Exemple : Patients du plus jeune au plus âgé :	**14**	**17**	**28**	**30**
Rang des âges	**1**	**2**	**3**	**4**

Exemple : Cyclisme et dopage

Première possibilité



1 2 3 4 5 6 7 8 9 10

équipe A **équipe B**

L'équipe A se serait elle dopée ?

Seconde possibilité

1 2 3 4 5 6 7 8 9 10

Coureurs intercalés. Equipes de même niveau?

Le test de Mann-Whitney permet de tester si 2 groupes indépendants sont extraits d'une population unique (mêmes caractéristiques) ou non.