

Comment utiliser BiostaTGV ?

Tutoriel rapide

<https://marne.u707.jussieu.fr/biostatgv/>

Sommaire

1. Ecran d'accueil.....	2
2. Croiser 2 variables qualitatives	3
2.1. Chi-2	3
2.2. Fisher.....	5
2.3. Cas particulier : variables appariées.....	6
3. Comparer 1 variable quantitative entre 2 groupes	7
3.1. Student, Mann-Whitney	7
3.2. Cas particulier : variables appariées.....	9
4. Comparer 1 variable quantitative entre >2 groupes	10
5. Comparer 2 variables quantitatives entre elles (corrélation)	12
6. Réaliser une courbe de survie (Kaplan-Meier).....	13
7. Faire un test du log-rank (comparaison de survie entre 2 groupes).....	15

1. Ecran d'accueil



1 : Cliquer sur « Tests statistiques ». Le tableau suivant illustre tous les tests disponibles sur le site. Dans chaque case :

- En haut : test « classique »
- En bas : test « alternatif », lorsque les effectifs sont trop faibles pour effectuer le test classique

En cas d'hésitation, cliquer sur « Besoin d'aide pour choisir votre test ? » (2).

Type de test à mettre en évidence		Variable de réponse			
Type de test		Qualitative nominale (2 groupes)	Qualitative nominale (plus de 2 groupes)	Qualitative ordinale	Quantitative
Facteur d'étude	Qualitatif (deux groupes)	Indépendants	Z de comparaison de proportions. Chi² (y,2) Test exact de Fisher.	Chi² (y,2)	Test de Cochran-Armitage Test de Mann-Whitney. t de Student. Test de Welch.
		Appariés	Test de McNemar. Test exact de Fisher.	Q de Cochran.	Tests des rangs signés de Wilcoxon. t de Student pour données appariées. Tests des rangs signés de Wilcoxon.
	Qualitatif (plus de deux groupes)	Indépendants	Chi² (y,2)	Chi² (y,2)	Test de Kruskal-Wallis. (ordinal) Analyse de la variance. Test de Kruskal-Wallis. (échelle quant)
		Appariés	Q de Cochran.	Q de Cochran.	Test de Friedman. Test de Friedman.
Quantitatif		Régression logistique.	Régression logistique multinomiale.	Corrélation de Spearman. Tau de Kendall. Corrélation de Pearson. Régression linéaire.	

* : La réalisation de ces tests n'est actuellement pas disponible sur biostaTGV

Attention, seuls les tests les plus fréquemment utilisés sont présentés dans ce tutoriel.

NB : en fin de chaque test, apparaîtra au bas de l'écran la « commande R ». Les analyses sont réalisées au moyen d'un logiciel libre et gratuit, R, téléchargeable ici : <https://www.r-project.org/>

2. Croiser 2 variables qualitatives

Ex : y a-t-il autant de fumeurs chez les hommes que chez les femmes ?

- **Chi-2** : lorsque tous effectifs théoriques ≥ 5
- **Fisher** : lorsqu'au moins un effectif théorique < 5 . Variante : Yates (Chi-2 corrigé) lorsque tous les effectifs théoriques sont ≥ 3 .

Pour connaître vos effectifs théoriques, commencer par un test du Chi-2 (1).

2.1. Chi-2

BiostaTGV
Tests statistiques en ligne

Tableau des tests statistiques d'hypothèse

Un test d'hypothèse est une démarche permettant d'évaluer la validité d'une hypothèse statistique en fonction d'un échantillon de données dont l'interprétation des résultats nécessite une bonne compréhension. Ce site n'a pas pour objectif de proposer un cours de statistique ni d'expliquer en détail cette démarche mais de permettre la réalisation de certains tests sans logiciel de statistique. Nous vous invitons donc fortement à consulter d'autres ressources pour bien comprendre les tests statistiques:

- Wikipedia Test d'hypothèses et Tests statistiques
- Des ouvrages comme ceux indiqués en page d'accueil

ⓘ Besoin d'aide pour choisir votre test ?

Calculez les statistiques de base d'une série de données
Analyse de survie **nouveau !**

Type de test à mettre en évidence		Variable de réponse			
Type de test		Qualitative nominale (2 groupes)	Qualitative nominale (plus de 2 groupes)	Qualitative ordinale	Quantitative
Facteur d'étude	Qualitatif (deux groupes)	Indépendants	Z de comparaison de proportions.* Chi² (x2.) Test exact de Fisher.	Chi² (x2.) 1	Cochran-Armitage* Test de Mann-Whitney. t de Student. Test de Welch.*
		Appariés	Test de McNemar. Test exact de Fisher.	Q de Cochran.*	Tests des signes.* Tests des rangs signés de Wilcoxon.
	Qualitatif (plus de deux groupes)	Indépendants	Chi² (x2.)	Chi² (x2.)	Test de Kruskal-Wallis. (ordinal) Tests des rangs signés de Wilcoxon.
		Appariés	Q de Cochran.*	Q de Cochran.*	t de Student pour données appariées. Tests des rangs signés de Wilcoxon.
Quantitatif		Régression logistique*	Régression logistique multinomiale*	Test de Friedman. Corrélation de Spearman. Tau de Kendall.	Analyse de la variance. Test de Kruskal-Wallis. (échelle quanti) Test de Friedman. Corrélation de Pearson. Régression linéaire.*

* : La réalisation de ces tests n'est actuellement pas disponible sur biostaTGV

⚠ Jusqu'en décembre 2011 le test réalisé incluait systématiquement la correction de Yates. Depuis cette date cette correction est une option du test

ETAPE 2 : Statistique de test Q, loi sous H0 et calcul de sa valeur observée Qobs à partir des données.

Statistique

Statistique du chi deux, somme des carrés des écarts à l'indépendance normalisés. Le principe est de calculer à partir des données un effectif attendu tel que l'hypothèse H0 soit vérifiée (dite "sous H0"). La statistique du Chi-Deux mesure dans quelle mesure les effectifs fournis dans les données sont proche de cette distribution théorique.

Condition de validité

Les effectifs attendus doivent être au moins égal à 5. Si inférieur le Chi-Deux n'est pas valable. Il faut faire un autre test (par exemple le test de Fisher)

Loi de la statistique sous H0

Loi du chi-deux à $(n_x - 1) \times (n_y - 1)$ degrés de liberté

Question préliminaire

Saisissez le nombre de modalités pour :
la variable X ? (2 <= n <= 50) **2**
la variable Y ? (2 <= n <= 50) **2**

Saisie des données

Tableau de contingence 1

	Y modalité 1	Y modalité 2
X modalité 1	149	182
X modalité 2	106	118

Copiez vos données depuis Excel et collez-les ci-dessus **Générer** **Tableau**



Remplissez le tableau ci-dessus en indiquant dans chaque case le nombre d'observations (ou de sujets) qui présente les modalités correspondantes de X et Y

Options du test

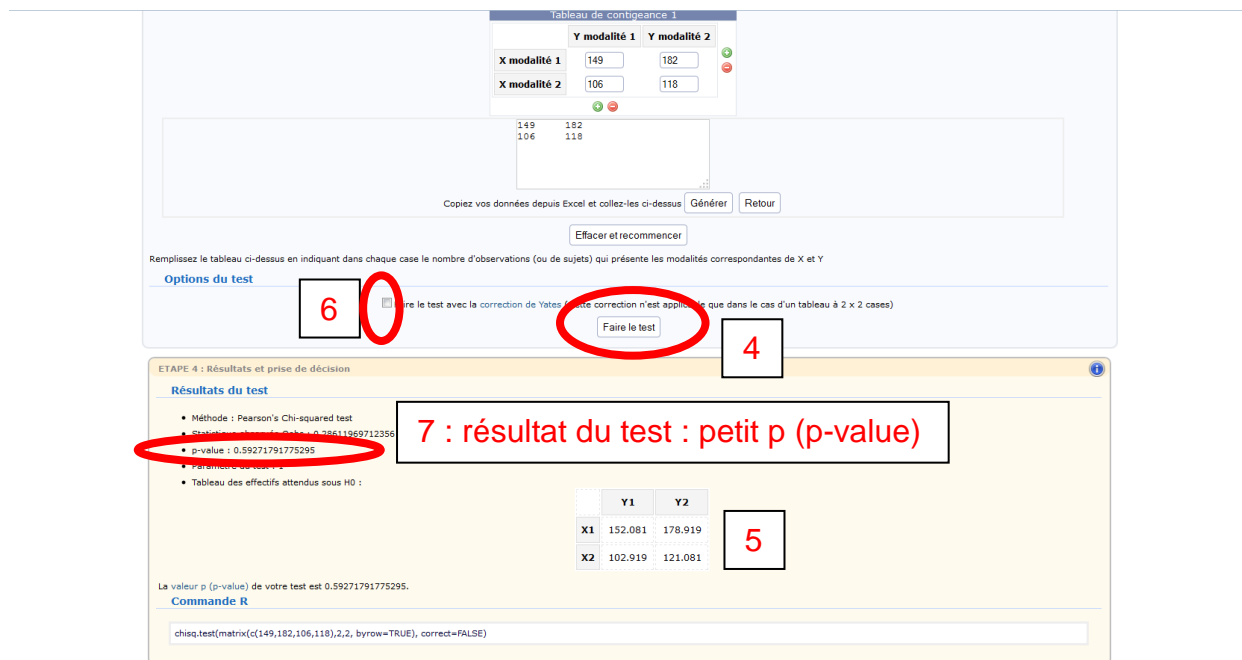
Faire le test avec la correction de Yates (Cette correction n'est applicable que dans le cas d'un tableau à 2 x 2 cases)

2 : renseigner le nombre de modalités de chaque variable. Ex : réponse oui/non = 2 modalités

3 : saisir les données :

- case par case. Vous pouvez encore ajouter des lignes ou colonnes en cliquant sur  
- ou cliquer sur « copier-coller via Excel » → copier les données excel dans le cadre qui apparaît → cliquer sur « générer »

Vous devrez donc au préalable avoir créé votre tableau « 4 cases » (ci-besoin via « tableau croisé dynamique » dans Excel)



The screenshot shows the BiostaTGV interface. At the top, there is a 'Tableau de contingence 1' with a 2x2 grid of input fields for 'X modalité 1', 'X modalité 2', 'Y modalité 1', and 'Y modalité 2'. Below this is a text area for pasting data from Excel, with 'Générer' and 'Retour' buttons. A 'Faire le test' button is circled in red and labeled '4'. Below the input section, there is a checkbox for 'Faire le test avec la correction de Yates' which is checked and circled in red, labeled '6'. The 'Résultats du test' section shows the p-value: 0.59271791775295, circled in red and labeled '7 : résultat du test : petit p (p-value)'. Below the p-value is a table of expected counts (effectifs théoriques) labeled '5':

	Y1	Y2
X1	152.081	178.919
X2	102.919	121.081

The command line at the bottom shows: `chisq.test(matrix(c(149,182,106,118),2,2), byrow=TRUE), correct=FALSE)`

4 : cliquer sur « Faire le test »

5 : vérifier les effectifs théoriques. Si effectifs théoriques <5, BiostaTGV vous le signalera :

Certaines cases des effectifs attendus sont inférieures à 5. Les conditions de validité du Chi2 ne sont pas remplies

Dans ce cas :

- Si tous vos effectifs théoriques sont ≥ 3 , cocher « faire le test avec la correction de Yates » (**6**).
- Sinon, revenir sur l'écran principal et sélectionner « **test exact de Fisher** »

7 : lire le résultat du test (valide si effectifs théoriques >5)

2.2. Fisher

BiostaTGV
Tests statistiques en ligne

Tableau des tests statistiques d'hypothèse

Un test d'hypothèse est une démarche permettant d'évaluer la validité d'une hypothèse statistique en fonction d'un échantillon de données dont l'interprétation des résultats nécessite une bonne compréhension. Ce site n'a pas pour objectif de proposer un cours de statistique ni d'expliquer en détail cette démarche mais de permettre la réalisation de certains tests sans logiciel de statistique. Nous vous invitons donc fortement à consulter d'autres ressources pour bien comprendre les tests statistiques:

- Wikipedia Test d'hypothèses et Tests statistiques
- Des ouvrages comme ceux indiqués en page d'accueil

🔗 Besoin d'aide pour choisir votre test ?

Calculez les statistiques de base d'une série de données

Analyse de survie ■ nouveau !

Type de test à mettre en évidence		Variable de réponse				
Type de test		Qualitative nominale (2 groupes)	Qualitative nominale (plus de 2 groupes)	Qualitative ordinale	Quantitative	
Facteur d'étude	Qualitatif (deux groupes)	Indépendants	Z de comparaison de proportions.* Chi² (x2,1) Test exact de Fisher*	Chi² (x2,)	Test de Cochran-Armitage*	Test de Mann-Whitney. t de Student. Test de Welch.*
		Appariés	Test de McNemar. Test exact de Fisher.	Q de Cochran.*	Tests des signes.* Tests des rangs signés de Wilcoxon.	t de Student pour données appariées. Tests des rangs signés de Wilcoxon.
	Qualitatif (plus de deux groupes)	Indépendants	Chi² (x2,)	Chi² (x2,)	Test de Kruskal-Wallis. (ordinal)	Analyse de la variance. Test de Kruskal-Wallis. (échelle quant)
		Appariés	Q de Cochran.*	Q de Cochran.*	Test de Friedman.	Test de Friedman.
Quantitatif		Régression logistique.*	Régression logistique multinomiale.*	Corrélation de Spearman. Tau de Kendall.	Corrélation de Pearson. Régression linéaire.*	

* : La réalisation de ces tests n'est actuellement pas disponible sur biostaTGV

Et reproduire la même procédure, à partir de **3**.

2.3. Cas particulier : variables appariées

Ex : les patients fumeurs fument-ils autant après que avant un infarctus du myocarde ?

Cas le plus fréquent : mesure **avant/après** (traitement,...) → Test de Mc Nemar

Tableau des tests statistiques d'hypothèse

Un test d'hypothèse est une démarche permettant d'évaluer la validité d'une hypothèse statistique en fonction d'un échantillon de données dont l'interprétation des résultats nécessite une bonne compréhension. Ce site n'a pas pour objectif de proposer un cours de statistique ni d'expliquer en détail cette démarche mais de permettre la réalisation de certains tests sans logiciel de statistique. Nous vous invitons donc fortement à consulter d'autres ressources pour bien comprendre les tests statistiques:

- Wikipedia Test d'hypothèses et Tests statistiques
- Des ouvrages comme ceux indiqués en page d'accueil

ⓘ Besoin d'aide pour choisir votre test ?

Calculez les statistiques de base d'une série de données

Analyse de survie **nouveau !**

Type de test à mettre en évidence		Variable de réponse			
Type de test		Qualitative nominale (2 groupes)	Qualitative nominale (plus de 2 groupes)	Qualitative ordinale	Quantitative
Facteur d'étude	Qualitatif (deux groupes)	<input type="checkbox"/> Z de comparaison de proportions.* <input type="checkbox"/> Chi ² (x2.) <input type="checkbox"/> Test exact de Fisher. <input checked="" type="checkbox"/> Test de McNemar. <input type="checkbox"/> Test exact de Fisher.	<input type="checkbox"/> Chi ² (x2.)	<input type="checkbox"/> Test de Cochran-Armitage*	<input type="checkbox"/> Test de Mann-Whitney. <input type="checkbox"/> t de Student. <input type="checkbox"/> Test de Welch.*
		<input type="checkbox"/> Chi ² (x2.)	<input type="checkbox"/> Q de Cochran.*	<input type="checkbox"/> Tests des signes.* <input type="checkbox"/> Tests des rangs signés de Wilcoxon.	<input type="checkbox"/> t de Student pour données appariées. <input type="checkbox"/> Tests des rangs signés de Wilcoxon.
	Qualitatif (plus de deux groupes)	<input type="checkbox"/> Chi ² (x2.)	<input type="checkbox"/> Chi ² (x2.)	<input type="checkbox"/> Test de Kruskal-Wallis. (ordinal)	<input type="checkbox"/> Analyse de la variance. <input type="checkbox"/> Test de Kruskal-Wallis. (échelle quant)
		<input type="checkbox"/> Q de Cochran.*	<input type="checkbox"/> Q de Cochran.*	<input type="checkbox"/> Test de Friedman.	<input type="checkbox"/> Test de Friedman.
Quantitatif		<input type="checkbox"/> Régression logistique*	<input type="checkbox"/> Régression logistique multinomiale*	<input type="checkbox"/> Corrélation de Spearman. <input type="checkbox"/> Tau de Kendall.	<input type="checkbox"/> Corrélation de Pearson. <input type="checkbox"/> Régression linéaire.*

* : La réalisation de ces tests n'est actuellement pas disponible sur biostaTGV

2 : cliquer sur « Test de McNemar »

Présentation

Il s'agit de tester l'indépendance entre une variable explicative X type avant/après traitement et une variable réponse Y type malade/non-malade. Les données sont donc des mesures d'une variable réponse binaire Y (malade/non-malade) sur les mêmes sujets avant et après le traitement, c'est-à-dire que nous sommes dans le cadre de données appariées. Le façon la plus explicite de présenter ces données est un tableau de contingence croisant en ligne le statut malade/non malade avant le traitement (en ligne) avec le statut malade/non malade après le traitement (en colonne).

	Malade après	Non malade après	Total
Malade avant	a	b	a + b
Non malade avant	c	d	c + d
Total	a + c	b + d	n

Définition de l'hypothèse nulle

HO : la variable X (avant / après traitement) est indépendante de la variable Y (malade/non malade) autrement dit les proportions de sujets dont le statut Y avant et après traitement change dans un sens ou dans l'autre sont égales ($p_0 = p_1$).

ETAPE 2 : Statistique de test Q, loi sous HO et calcul de sa valeur observée Qobs à partir des données.

Statistique

Statistique de Mac Nemar avec correction de Yates.

Loi de la statistique sous HO

Loi du chi-deux à 1 degré de liberté

Saisie des données

Tableau de contingence 1

	Y 1	Y 2
X 1		
X 2		

Importer un CSV | Importer un tableau | Importer un fichier Excel

Effacer et recommencer

Faire le test

3 : Saisir les données comme indiqué dans le tableau en haut de l'écran, de la même manière que pour le test du chi-2 : soit écrire directement, soit copier-coller via excel. Le résultat du test apparaîtra de la même manière que pour le Chi-2.

3. Comparer 1 variable quantitative entre 2 groupes

Ex : le BMI des hommes est-il différent de celui des femmes ?

- **Test t de Student** : lorsque $n > 30$ dans chaque groupe
- **Test non-paramétrique de Wilcoxon (=Mann-Whitney)** : lorsque $n < 30$ dans au moins un des groupes

Attention cette distinction est volontairement un peu grossière, pour vous permettre de sélectionner rapidement un test adapté.

Avant d'effectuer le test de Student, il est toutefois fortement conseillé de vérifier les conditions d'application : distribution normale de la variable dans chaque groupe, égalité des variances.

- **Distribution normale de la variable dans chaque groupe** : faire un histogramme (diagramme en barres), et vérifier visuellement si la distribution suit une loi normale.
- **Egalité des variances** : calculer la variance dans chaque groupe : excel dispose d'un test d'égalité des variances : fonction **test.f()**. Approximation : si la variance d'un groupe est plus de 2 fois plus importante que celle de l'autre groupe, on considère que les variances sont inégales. Si cette condition n'est pas remplie mais que les distributions sont normales dans chaque groupe, choisir le test t de Student, en prenant bien en compte l'étape 6

Si la distribution n'est pas normale, réaliser un test non-paramétrique de Mann-Whitney. Si seule la condition d'égalité des variances n'est pas respectée, il est possible de réaliser un test t de Student.

3.1. Student, Mann-Whitney

Tableau des tests statistiques d'hypothèse

Un test d'hypothèse est une démarche permettant d'évaluer la validité d'une hypothèse statistique en fonction d'un échantillon de données dont l'interprétation des résultats nécessite une bonne compréhension. Ce site n'a pas pour objectif de proposer un cours de statistique ni d'expliquer en détail cette démarche mais de permettre la réalisation de certains tests sans logiciel de statistique. Nous vous invitons donc fortement à consulter d'autres ressources pour bien comprendre les tests statistiques:

- Wikipedia Test d'hypothèses et Tests statistiques
- Des ouvrages comme ceux indiqués en page d'accueil

ⓘ Besoin d'aide pour choisir votre test ?

Calculez les statistiques de base d'une série de données

Analyse de survie • nouveau !

Type de test à mettre en évidence	Variable de réponse			
	Qualitative nominale (2 groupes)	Qualitative nominale (plus de 2 groupes)	Qualitative ordinale	Quantitative
Type de test -Tous-	Indépendants	Z de comparaison de proportions. ^{**} Chi ² (χ ²). Test exact de Fisher.	Chi ² (χ ²).	Test de Mann-Whitney. t de Student. Test de Welch.
		Appariés	Test de McNemar. Test exact de Fisher.	Q de Cochran. ^{**}
Facteur				Tests des rangs signés de Wilcoxon. ^{**}
			Tests des rangs signés de Wilcoxon.	

1 : choisir le test approprié : Mann-Whitney ou t de Student

Test de Student

ETAPE 1 : Présentation du test et définition de l'hypothèse nulle

Présentation

Ce test permet de comparer les mesures d'une variable quantitative effectuées sur deux groupes de sujets indépendants définis par les modalités de la variable qualitative.

Définition de l'hypothèse nulle

H0 : les moyennes sont égales dans les deux groupes

ETAPE 2 : Statistique de test Q, loi sous H0 et calcul de sa valeur observée Qobs à partir des données.

Statistique

t, déviation de la moyenne calculée avec une variance commune aux deux groupes

Loi de la statistique sous H0

Loi du t à (n-1) degrés de liberté

Question préliminaire

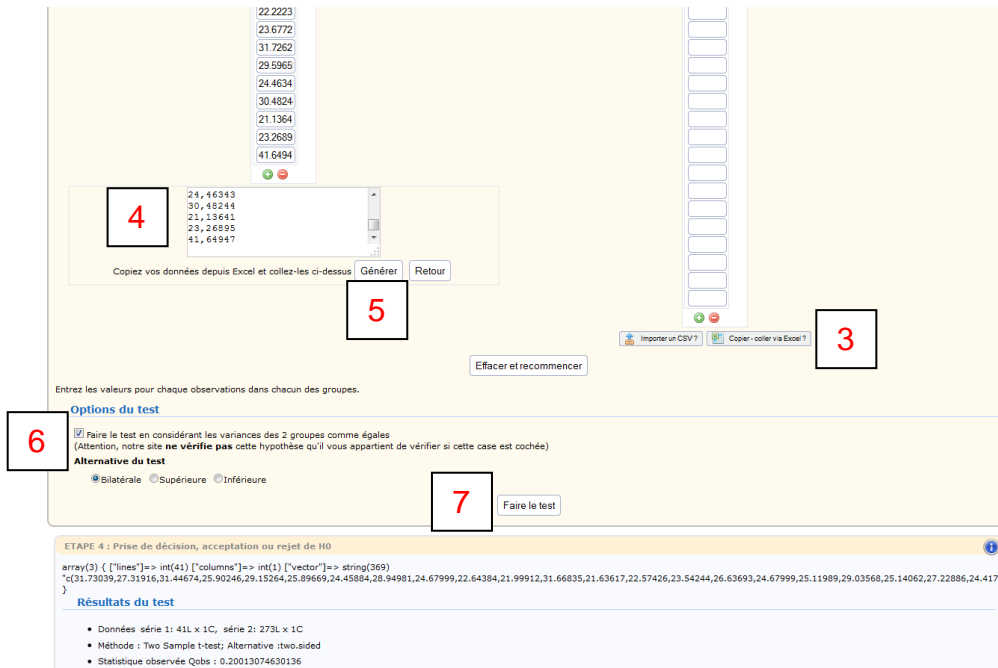
Quel est le nombre d'observations dans :

le groupe 1 ? 41

le groupe 2 ? 273

Envoyer

2 : entrer le nombre d'observations dans chaque groupe, puis « envoyer »



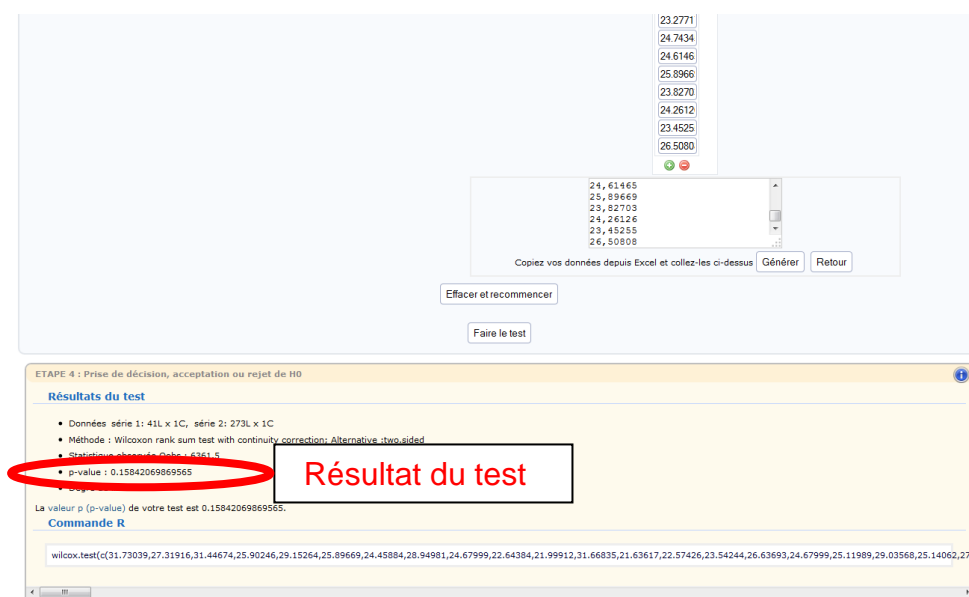
- 3** : cliquer sur « copier-coller via Excel ? »
- 4** : coller vos données dans l'encadré qui s'ouvre
- 5** : cliquer sur « générer »

Reproduire la manœuvre pour le groupe 2

Si test de Student : 6

- Si nécessaire, cocher la case « faire le test en considérant les variances des 2 groupes comme égales », (voir en-tête du paragraphe). *Si vous ne cochez pas la case, le test t de Welch sera réalisé.*
- Conserver l'alternative du test à « bilatérale » (valeur par défaut)

7 : cliquer sur « Faire le test »



3.2. Cas particulier : variables appariées

Cas le plus fréquent : mesure avant/après (traitement,...) :

- **Test t de Student pour variables appariées** : lorsque $n > 30$
- **Test non-paramétrique des rangs signés de Wilcoxon** : lorsque $n < 30$

Attention cette distinction est volontairement un peu grossière, pour vous permettre de sélectionner rapidement un test adapté.

Attention, vous devez avoir le même nombre d'observations (individus) avant et après +++.

Tableau des tests statistiques d'hypothèse

Un test d'hypothèse est une démarche permettant d'évaluer la validité d'une hypothèse statistique en fonction d'un échantillon de données dont l'interprétation des résultats nécessite une bonne compréhension. Ce site n'a pas pour objectif de proposer un cours de statistique ni d'expliquer en détail cette démarche mais de permettre la réalisation de certains tests sans logiciel de statistique. Nous vous invitons donc fortement à consulter d'autres ressources pour bien comprendre les tests statistiques:

- Wikipedia Test d'hypothèses et Tests statistiques
- Des ouvrages comme ceux indiqués en page d'accueil

ⓘ Besoin d'aide pour choisir votre test ?

Calculez les statistiques de base d'une série de données

Analyse de survie **nouveau !**

Type de test à mettre en évidence		Variable de réponse				
Type de test		Qualitative nominale (2 groupes)	Qualitative nominale (plus de 2 groupes)	Qualitative ordinale	Quantitative	
Facteur d'étude	Qualitatif (deux groupes)	Indépendants	Z de comparaison de proportions.* Chi² (x2.) Test exact de Fisher.	Chi² (x2.)	Test de Cochran-Armitage*	Test de Mann-Whitney. t de Student. Test de Welch.*
			Appariés	Test de McNemar. Test exact de Fisher.	Q de Cochran.*	Tests des signes.* Tests des rangs signés de Wilcoxon.
	Qualitatif (plus de deux groupes)	Indépendants		Chi² (x2.)	Chi² (x2.)	Test de Kruskal-Wallis. (ordinal)
			Appariés	Q de Cochran.*	Q de Cochran.*	Test de Friedman.
Quantitatif		Régression logistique.*		Régression logistique multinomiale.*	Corrélation de Spearman. Tau de Kendall.	Corrélation de Pearson. Régression linéaire.*

* : La réalisation de ces tests n'est actuellement pas disponible sur biostaTGV

2 : choisir le test approprié

Test du Student pour échantillons appariés

ETAPE 1 : Présentation du test et définition de l'hypothèse nulle

Présentation
Ce test permet de comparer deux mesures d'une variable quantitative effectuées sur les mêmes sujets (mesures définies par les modalités de la variable qualitative). En fait ce test traite les deux échantillons appariés comme un seul sur lequel on aurait mesuré la différence d entre les deux mesures.
Définition de l'hypothèse nulle
HO : la différence moyenne \bar{d} entre les deux mesures est nulle

ETAPE 2 : Statistique de test Q_0 , loi sous H_0 et calcul de sa valeur observée Q_{obs} à partir des données.

Statistique
 t , déviation de la moyenne (calculée avec la variance empirique de la différence d)
Loi de la statistique sous H_0
Loi du t à $(n-1)$ degrés de liberté

Question préliminaire
quel est le nombre d'observations dans chaque échantillon : Envoyer

3 : entrer le nombre d'individus / observations. Attention, ne compter qu'une fois chaque individu. Ex : si 35 individus et 70 observations (une avant et une après), entrer « 35 ».

Puis procéder comme pour les tests de Student et Mann-Whitney.

4. Comparer 1 variable quantitative entre >2 groupes

Ex : la glycémie est-elle la même chez les personnes obèses ($BMI > 30$), en surpoids ($25 < BMI < 30$), normo-pondérées ($18,5 < BMI < 25$) ou en insuffisance pondérale ($BMI < 18,5$) ?

- **Analyse de la variance (ANOVA)** : lorsque $n > 30$ dans chaque groupe
- **Test non-paramétrique de Kruskal-Wallis** : lorsque $n < 30$ dans au moins un des groupes

Attention cette distinction est volontairement un peu grossière, pour vous permettre de sélectionner rapidement un test adapté. BiostaTGV précise bien que le test ANOVA est très restrictif. Dans le doute, effectuer le test non-paramétrique de Kruskal-Wallis.

Attention, 75 observations maximum par groupe.

Choisir le test adapté (1)

Tableau des tests statistiques d'hypothèse

Un test d'hypothèse est une démarche permettant d'évaluer la validité d'une hypothèse statistique en fonction d'un échantillon de données dont l'interprétation des résultats nécessite une bonne compréhension. Ce site n'a pas pour objectif de proposer un cours de statistique ni d'expliquer en détail cette démarche mais de permettre la réalisation de certains tests sans logiciel de statistique. Nous vous invitons donc fortement à consulter d'autres ressources pour bien comprendre les tests statistiques:

- Wikipedia Test d'hypothèses et Tests statistiques
- Des ouvrages comme ceux indiqués en page d'accueil

Ⓞ Besoin d'aide pour choisir votre test ?

Calculez les statistiques de base d'une série de données
Analyse de survie * nouveau !

Type de test à mettre en évidence		Variable de réponse				
		Qualitative nominale (2 groupes)	Qualitative nominale (plus de 2 groupes)	Qualitative ordinale	Quantitative	
Facteur d'étude	Qualitatif (deux groupes)	Indépendants	Z de comparaison de proportions. ^{**} Chi ² (x2.) Test exact de Fisher.	Chi ² (x2.)	Test de Cochran-Armitage. ^{**}	Test de Mann-Whitney. t de Student. Test de Welch. ^{**}
		Appariés	Test de McNemar. Test exact de Fisher.	Q de Cochran. ^{**}	Tests des signes. ^{**} Tests des rangs signés de Wilcoxon.	t de Student pour données appariées. Tests des rangs signés de Wilcoxon.
	Qualitatif (plus de deux groupes)	Indépendants	Chi ² (x2.)	Chi ² (x2.)	Test de Kruskal-Wallis. (ordinal)	Analyse de la variance. Test de Kruskal-Wallis. (échelle quant)
		Appariés	Q de Cochran. ^{**}	Q de Cochran. ^{**}	Test de Friedman.	Test de Friedman.
Quantitatif		Régression logistique. ^{**}	Régression logistique multinomiale. ^{**}	Corrélation de Spearman. Tau de Kendall.	Corrélation de Pearson. Régression linéaire. ^{**}	

^{**} : La réalisation de ces tests n'est actuellement pas disponible sur biostaTGV

[Réseau Sentinelles](#) • [Institut Pierre Louis UMR S 1136](#) • [Inserm UPMC](#) • [Credits & Mentions légales](#)

Entrer le nombre de groupes : de 2 à 10 groupes (2). Le principe est ensuite le même que celui du test t de Student.

BiostaTGV
Tests statistiques en ligne

Analyse de variance

ETAPE 1 : Présentation du test

Présentation

L'analyse de la variance (ANOVA) à un facteur permet d'étudier le comportement d'une variable réponse Y quantitative en fonction d'une variable explicative catégorielle X à p modalités définissant p groupes indépendants. Elle repose sur une décomposition de la variance totale en une variance intra et inter classes.

Définition de l'hypothèse de test

H0 : les moyennes de la variable quantitative sont toutes égales entre les p groupes

Statistique

Statistique de Fisher F, rapport des variances inter et intra classes

Loi de la statistique sous H0

loi de Fisher à (p-1, n-p) degrés de liberté, où n est le nombre total de sujets.

Note Bene

Ce test repose sur des hypothèses très restrictives (normalité et égalité des variances des distributions entre les groupes), parfois considérées comme irréalistes. Avant de procéder à une ANOVA, il convient donc de vérifier, ou mieux, de tester ces hypothèses par :

- Des tests de normalité (p. ex. test de Kolmogorov-Smirnov ou test de Shapiro-Wilks), ou un simple diagramme quantile-quantile (ou "Q-Q plot).
- Des tests d'homogénéité des variances, comme le test de Bartlett ou le test de Levene.

S'il apparaît que les données sont incompatibles avec les hypothèses d'ANOVA, il restera cependant possible de tester l'hypothèse d'égalité des moyennes en recourant à un test non paramétrique, le test de Kruskal-Wallis.

ETAPE 2 : Statistique de test Q, loi sous H0 et calcul de sa valeur observée Qobs à partir des données.

Statistique

Moyennes des rangs des observations dans les différents groupes

Loi de la statistique sous H0

loi du chi deux à k-1 groupes

Questions préliminaires

Combien de groupes sont définis par les modalités de la variable nominale X (1 à 10) ? Envoyer

Le résultat du test se lit en 3.

ETAPE 4 : Prise de décision

Résultats du test Anova

- Statistique observée Qobs : 0.55862356758462
- g :
- p-value : 0.57306092836782

Test d'homogénéité de variance de Bartlett

- Méthode : Bartlett test of homogeneity of variances
- p-value : 0.018277326436882

La valeur p (p-value) de votre test est 0.57306092836782.

Commande R

```
X =  
c(23.87631,20.98504,22.01154,28.75702,23.07662,36.43161,31.73039,21.78854,27.31916,38.18727,24.08119,49.12033,23.07662,25.18267,28.94981,21.03125,22.51215,27.49272,21.67837,21.32906,33.74417,32.51111,21.03125,22.51215,27.49272,21.67837,21.32906,33.74417,32.51111)  
Y = c(rep(1,60),rep(2,70),rep(3,40))  
facY = as.factor(Y)  
r = list()  
r$anova = anova(aov(X ~ facY))  
r$bartlett = bartlett.test(X,Y)  
r
```

5. Comparer 2 variables quantitatives entre elles (corrélation)

Ex : y a-t-il une corrélation entre le BMI et le nombre de calories ingérées ?

- Test du coefficient de corrélation de Pearson

Tableau des tests statistiques d'hypothèse

Un test d'hypothèse est une démarche permettant d'évaluer la validité d'une hypothèse statistique en fonction d'un échantillon de données dont l'interprétation des résultats nécessite une bonne compréhension. Ce site n'a pas pour objectif de proposer un cours de statistique ni d'expliquer en détail cette démarche mais de permettre la réalisation de certains tests sans logiciel de statistique. Nous vous invitons donc fortement à consulter d'autres ressources pour bien comprendre les tests statistiques:

- [Wikipedia Test d'hypothèses et Tests statistiques](#)
- Des ouvrages comme ceux indiqués en page d'accueil

📖 Besoin d'aide pour choisir votre test ?

Calculez les statistiques de base d'une série de données
Analyse de survie [nouveau !](#)

Type de test à mettre en évidence		Variable de réponse				
Type de test		Qualitative nominale (2 groupes)	Qualitative nominale (plus de 2 groupes)	Qualitative ordinale	Quantitative	
Facteur d'étude	Qualitatif (deux groupes)	Indépendants	Z de comparaison de proportions.* Chi² (x2.) Test exact de Fisher.	Chi² (x2.)	Test de Cochran-Armitage*	Test de Mann-Whitney. t de Student. Test de Welch.*
		Appariés	Test de McNemar. Test exact de Fisher.	Q de Cochran.*	Tests des signes.* Tests des rangs signés de Wilcoxon.	t de Student pour données appariées. Tests des rangs signés de Wilcoxon.
	Qualitatif (plus de deux groupes)	Indépendants	Chi² (x2.)	Chi² (x2.)	Test de Kruskal-Wallis. (ordinal)	Analyse de la variance. Test de Kruskal-Wallis. (échelle quant)
		Appariés	Q de Cochran.*	Q de Cochran.*	Test de Friedman.	Test de Friedman.
	Quantitatif		Régression logistique*	Régression logistique multinomiale*	Corrélation de Spearman. Tau de Kendall.	Corrélation de Pearson. Régression linéaire.

* : La réalisation de ces tests n'est actuellement pas disponible sur biostaTGV

1

Le principe est ensuite similaire à celui du test t apparié.

6. Réaliser une courbe de survie (Kaplan-Meier)

Tableau des tests statistiques d'hypothèse

Un test d'hypothèse est une démarche permettant d'évaluer la validité d'une hypothèse statistique en fonction d'un échantillon de données dont l'interprétation des résultats nécessite une bonne compréhension. Ce site n'a pas pour objectif de proposer un cours de statistique ni d'expliquer en détail cette démarche mais de permettre la réalisation de certains tests sans logiciel de statistique. Nous vous invitons donc fortement à consulter d'autres ressources pour bien comprendre les tests statistiques :

- Wikipedia Test d'hypothèses et Tests statistiques
- Des ouvrages comme ceux indiqués en page d'accueil

🔗 Besoin d'aide pour choisir votre test ?

Calculez les statistiques de base d'une série de données

Analyse de survie **nouveau !**

Type de test à mettre en évidence		Variable de l'étude				
Type de test		Qualitative nominale (2 groupes)	Qualitative nominale (plus de 2 groupes)	Qualitative ordinale	Quantitative	
Facteur d'étude	Qualitatif (deux groupes)	Indépendants	Z de comparaison de proportions.* Chi² (χ²). Test exact de Fisher.	Chi² (χ²).	Test de Cochran-Armitage*	Test de Mann-Whitney. t de Student. Test de Welch.*
			Appariés	Test de McNemar. Test exact de Fisher.	Q de Cochran.*	Tests des signes.* Tests des rangs signés de Wilcoxon.
	Qualitatif (plus de deux groupes)	Indépendants		Chi² (χ²).	Chi² (χ²).	Test de Kruskal-Wallis. (ordinal)
			Appariés	Q de Cochran.*	Q de Cochran.*	Test de Friedman.
Quantitatif		Régression logistique*		Régression logistique multinomiale*	Corrélation de Spearman. Tau de Kendall.	Corrélation de Pearson. Régression linéaire.*

* : La réalisation de ces tests n'est actuellement pas disponible sur biostaTGV

réseau.Santinelles® • Institut Pierre Louis UMR S 1136 • Inserm S MÉDECINE SORBONNE UNIVERSITÉ • Crédits & Mentions légales

ETAPE 2 : Entrée des données.

Question préliminaire

Quel type d'analyse voulez vous conduire ?

- Estimation d'une fonction de survie (1 seule série de données)
- Comparaison de 2 courbes de survie (2 séries de données)

Nombre d'individus :

Série : 8

Envoyer

Saisie des données

	Temps	Status (0/1)
Individu 1		
Individu 2		
Individu 3		
Individu 4		
Individu 5		
Individu 6		
Individu 7		
Individu 8		

Importer un CSV ? Copier - coller via Excel ?

Effacer et recommencer

Pour chaque individu (sur chaque ligne) vous devez indiquer 2 informations :

- Dans la colonne "Temps", la durée d'observation de chaque individu, l'unité n'a pas d'importance (nombre de mois, nombre de jour, nombre de semaine) mais l'unité de temps doit être la même pour tous les individus et toutes les séries de données. Pour chaque individu, cette valeur indique la durée avant la survenue de l'évènement (le décès par exemple) de cet individu ou la durée du suivi si aucun évènement n'est survenu durant le suivi.
- Dans la colonne "Status", le status de l'observation qui prend la valeur soit 1 si l'évènement est vraiment observé, soit 0 si l'évènement n'est pas survenu pour cet individu (pour cet individu l'observation est "censurée" à droite).

Faire le test

6

Probabilité de survie à chaque temps

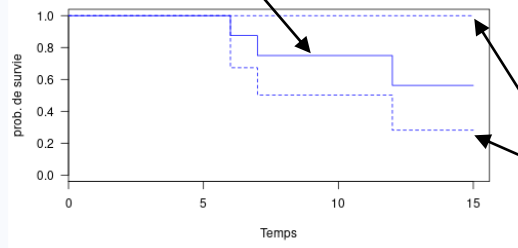
Résultats de l'analyse de survie

Données

time	n.risk	n.event	n.censor	surv	std.err	lower	upper
6	8	1	0	0.875	0.11692679333669	0.67338193650596	1
7	7	1	0	0.75	0.15309310892395	0.50270184129405	1
12	4	1	3	0.5625	0.19887378220872	0.28130492459566	1

Courbe de survie

Courbe de Survie



Intervalle de confiance à 95 %

Commande R

```
library(survival)
data = data.frame(matrix(c(12,1,10,0,15,0,8,0,6,1,7,1,12,0,13,0),8,2, byrow=TRUE))
names(data) <- c("time", "status")
data$status = data$status == 1
m = survfit(Surv(time,status) ~ 1, data=data)
summary(m)
```

7. Faire un test du log-rank (comparaison de survie entre 2 groupes)

Voir paragraphe 6., puis :

ETAPE 2 : Entrée des données.

Question préliminaire

Quel type d'analyse voulez vous conduire ?

Estimation d'une fonction de survie (1 seule série de données)

Comparaison de 2 courbes de survie (2 séries de données) 2

Nombre d'individus :

Série 1 :

Série 2 :

Saisie des données

Série 1		
	Temps	Status (0/1)
Individu 1	<input type="text"/>	<input type="text"/>
Individu 2	<input type="text"/>	<input type="text"/>
Individu 3	<input type="text"/>	<input type="text"/>
Individu 4	<input type="text"/>	<input type="text"/>
Individu 5	<input type="text"/>	<input type="text"/>
Individu 6	<input type="text"/>	<input type="text"/>
Individu 7	<input type="text"/>	<input type="text"/>
Individu 8	<input type="text"/>	<input type="text"/>

Série 2		
	Temps	Status (0/1)
Individu 1	<input type="text"/>	<input type="text"/>
Individu 2	<input type="text"/>	<input type="text"/>
Individu 3	<input type="text"/>	<input type="text"/>
Individu 4	<input type="text"/>	<input type="text"/>
Individu 5	<input type="text"/>	<input type="text"/>
Individu 6	<input type="text"/>	<input type="text"/>
Individu 7	<input type="text"/>	<input type="text"/>
Individu 8	<input type="text"/>	<input type="text"/>
Individu 9	<input type="text"/>	<input type="text"/>
Individu 10	<input type="text"/>	<input type="text"/>

Pour chaque individu (sur chaque ligne) vous devez indiquer 2 informations :

- Dans la colonne "Temps", la durée d'observation de chaque individu, l'unité n'a pas d'importance (nombre de mois, nombre de jour, nombre de semaine) mais l'unité de temps doit être la même pour tous les individus et toutes les séries de données.
Pour chaque individu, cette valeur indique la durée avant la survenue de l'évènement (le décès par exemple) de cet individu ou la durée du suivi si aucun évènement n'est survenu durant le suivi.
- Dans la colonne "Status", le status de l'observation qui prend la valeur soit 1 si l'évènement est vraiment observé, soit 0 si l'évènement n'est pas survenu pour cet individu (pour cet individu l'observation est "censurée" à droite).

Puis procéder comme au paragraphe 6.

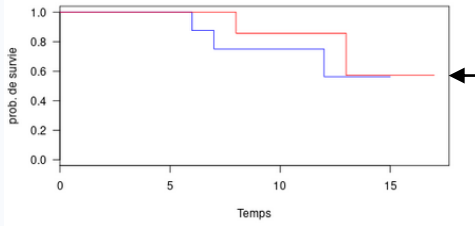
Vous obtenez le résultat page suivante :

Résultats de l'analyse de survie

Données

strata	time	n.risk	n.event	n.censor	surv	std.err	lower	upper
1	6	8	1	0	0.875	0.11692679333669	0.67338193650596	1
1	7	7	1	0	0.75	0.15309310892395	0.50270184129405	1
1	12	4	1	3	0.5625	0.19887378220872	0.28130492459566	1
2	8	7	1	4	0.85714285714286	0.13226001425322	0.63344652901215	1
2	13	3	1	2	0.57142857142857	0.24939187455542	0.24292216411775	1

Courbe de Survie



Courbes de survie

Pour retrouver à quel groupe chacune correspond, se reporter aux probabilités de survie (3) de chaque groupe (4), selon le temps (5).
Ici : bleu = groupe 1 ; rouge = groupe 2

Test du log-rank

• Statistique: 1.16230202249166917549
• p-value: 0.58862821680497

Résultat du test du log-rank

Effectif	Observés	Attendus
8	3	2.3996336996337
10	2	2.6003663003663

Nombre d'évènements dans chaque groupe