

QUARTO BIG NOTES

Table of contents

1	Résumé statistique	1
1.1	Position	1
1.2	Dispersion	2
1.3	Dépendance, liaison et association	3
2	Tests statistiques	10
2.1	Comment les utiliser ?	10
2.2	Types de tests	10
2.3	Calcul de β et du nombre de sujets à inclure	16
2.4	Tests uni- ou bilatéraux	17
2.5	Intervalles de confiance	18
3	Modèles	19

1 Résumé statistique

1.1 Position

1.1.1 Moyenne

- = centre de gravité des valeurs
- suppose **équivalence de la quantité** : 1 euro voit 1 euro quelque soit sa position sur la droite des réels.

1.1.2 Médiane

- Partage la distribution en 2 parts égales
- Si distribution symétrique : moyenne = médiane

1.1.3 Mode

- = Valeur la plus fréquente

1.2 Dispersion

1.2.1 Empan

- = différence entre valeur minimale et valeur maximale

1.2.2 Écart interquartile

- = Q3 - Q1

1.2.3 Variance

- = Variance = moyenne des carrés des écarts à la moyenne

$$v = \frac{(\text{ecart}_1 - \mu)^2 + (\text{ecart}_2 - \mu)^2 + \dots + (\text{ecart}_n - \mu)^2}{\text{nombre d'observation}}$$

1.2.4 Écart type

- = racine carrée de la variance
- = racine carrée des carrés des écarts à la moyenne
- Avantages variances et écarts types :
 - Carrés : positive écarts négatifs et accentue écarts importants
 - Variance de deux variables indépendantes est égale à la somme de leurs variances

- Dans le cas d'une distribution normale :
 - * 2/3 des observations se situent dans un intervalle de ± 1 écart type autour de la moyenne
 - * 95% des observations se situent dans un intervalle de ± 2 écart type autour de la moyenne

1.2.5 Fonction R pour tout ça :

- `summary()` donne la moyenne, médiane, min, max, 1er et 3e quartile
- `describe()` pas mal de trucs aussi

1.3 Dépendance, liaison et association

- Définition de la dépendance : deux variables sont dépendantes si la connaissance de la valeur de l'une permet de mieux prédire la valeur de l'autre
- Par exemple : taille et poids sont dépendantes (plus on est grand, plus on pèse lourd)

1.3.1 Variables quantitatives

- 3 types de liaisons entre variables quantitatives :
 - Dépendance : connaître X permet de mieux prédire Y
 - Dépendance monotone : X et Y varient dans le même sens (croissant ou décroissant)
 - * Dépendance linéaire : relation de proportionnalité entre X et Y
 - * Corrélation linéaire (coefficient de corrélation linéaire de Pearson) = r : mesure de la force de la dépendance linéaire entre X et Y
 - * Coefficient de détermination = variance partagée r^2 : proportion de la variance de Y expliquée par la variance de X dans une relation linéaire
 - Concordance : X et Y varient dans le même sens mais pas forcément de façon linéaire
 - * Coefficient intraclass : mesure de la concordance entre plusieurs mesures de la même variable

Mesure	Paramètre	Symbole	Interprétation
Dépendance	Coefficient de corrélation linéaire	r	Mesure la direction et la force de la dépendance linéaire entre deux variables quantitatives. Varie entre -1 et 1.
Dépendance	Coefficient de détermination = variance partagée	r^2	Mesure la proportion de la variance de Y expliquée par la variance de X dans une relation linéaire. Varie entre 0 et 1.
Concordance	Coefficient intraclasse	ICC	Mesure la concordance entre plusieurs mesures de la même variable. Varie entre 0 et 1.
Autre	Covariance	$Cov(X, Y)$	Mesure comment deux variables varient ensemble (positive: ensemble ; négative : sens inverse)

1.3.1.1 Dépendance entre variables quantitatives

- Pas de paramètre estimant parfaitement la dépendance ou l'indépendance entre deux variables quantitatives

1.3.1.2 Dépendance monotone ou linéaire entre variables quantitatives

- Mesure de la direction dépendance linéaire : coefficient de corrélation linéaire de Pearson (ρ ou r)
 - Varie entre -1 (X et Y parfaitement négativement corrélées) et 1 (X et Y parfaitement positivement corrélées)
 - 0 : pas de corrélation linéaire entre X et Y = indépendance linéaire
 - **NB : le coefficient de corrélation ne mesure pas vraiment la force de la dépendance entre X et Y mais son type de dépendance !!** *en pratique, mesure un peu la force quand même mais il n'y a pas d'unité pour l'exprimer*
- Mesure de la force de la dépendance linéaire : **coefficient de détermination** (r^2)
 - Varie entre 0 (pas de dépendance) et 1 (dépendance parfaite)
 - Indique la proportion de la variance de Y expliquée par la variance de X dans une relation linéaire
 - Par exemple : $r = 0.8 \Rightarrow r^2 = 0.64$: 64% de la variance de Y est expliquée par la variance de X dans une relation linéaire (et ce n'est pas r^2 des valeurs de Y sont expliquées par X !!)
- Petit tableau récapitulatif coefficient de corrélation / variance partagée / covariante

Paramètre	Symbole	Interprétation
Coefficient de corrélation linéaire	r	Mesure la direction et la force de la dépendance linéaire entre deux variables quantitatives. Varie entre -1 et 1.
Coefficient de détermination	r^2	Mesure la proportion de la variance de Y expliquée par la variance de X dans une relation linéaire. Varie entre 0 et 1.
Covariance	$Cov(X, Y)$	Mesure la tendance conjointe de deux variables à varier ensemble. Peut être positive, négative ou nulle.

1.3.1.3 Concordance entre variables quantitatives

- Mesure de la concordance entre variables quantitatives : **coefficient intraclasse** (ICC)
- Varie entre 0 (pas de concordance) et 1 (concordance parfaite)
- Indique la proportion de la variance totale attribuable à la variance entre les sujets
- Calcul : $ICC = \frac{\text{variance entre sujets}}{\text{variance entre sujets} + \text{variance entre examinateurs} + \text{variance résiduelle}}$
 - En gros : $ICC = \frac{\text{vrai signal}}{\text{vrai signal} + \text{bruit}}$
 - Vaut 1 quand aucun bruit (variance examinateurs et résiduelle = 0); vaut 0 quand tout est bruit (variance entre sujets = 0)
 - Par exemple : ICC = 0.75 : 75% de la variance totale est attribuable à la variance entre les sujets
- ICC le plus classiquement utilisé : 2-way random effects, absolute agreement, single rater/measurement

1.3.2 Variables catégorielles

1.3.2.1 Dépendance entre variables catégorielles

- Mesure de la dépendance entre variables catégorielles : **chi-carré de Pearson** (χ^2)
 - Existe-t-il une association entre deux variables catégorielles ?
 - Le *test du Chi-carré d'indépendance* permet de tester l'hypothèse nulle d'indépendance entre deux variables catégorielles, c'est à dire l'absence de relation entre elles.

– Méthode :

- Pour évaluer la force de la relation : **transformation normalisée du chi-carré**

– **Coefficient phi (ϕ) :**

- * Utile surtout pour un tableau 2×2 (deux variables binaires)
- * $|\phi| = \sqrt{\chi^2/n}$ où n = taille de l'échantillon
- * Correspond à la corrélation de Pearson entre deux variables binaires codées 0/1 (version signée entre -1 et 1)

– **V de Cramér (V) :**

- * Utile pour un tableau $r \times c$ (variables nominales, possiblement > 2 modalités)
- * $V = \sqrt{\chi^2/(n(k-1))}$ avec $k = \min(r, c)$
- * Varie entre 0 (pas d'association) et 1 (association parfaite)

– **Coefficient de contingence de Pearson (C) :**

- * $C = \sqrt{\chi^2/(\chi^2 + n)}$
- * Utilisable pour des tableaux $r \times c$, mais l'échelle dépend de la taille du tableau : C ne peut pas atteindre 1 et n'est pas directement comparable entre tableaux de dimensions différentes
- * Varie entre 0 et C_{max} avec $C_{max} = \sqrt{(k-1)/k}$ (donc $C < 1$), où $k = \min(r, c)$
- * Lien avec ϕ (cas 2×2) : C est juste un re-calage de $|\phi|$: $C = \frac{|\phi|}{\sqrt{1+\phi^2}}$ (donc même ordre/monotonie, mais valeurs plus "compressées")

– **Coefficient de corrélation de Spearman ρ_s :**

- * Pour des variables **ordinales** (ou une relation **monotone** non linéaire entre deux quantitatives)
- * = corrélation de Pearson calculée sur les rangs, varie entre -1 et 1

Mesure	Paramètre	Interprétation
Dépendance	Chi-carré de Pearson	Mesure l'association entre deux variables catégorielles. Valeur plus élevée indique une association plus forte.

Mesure	Paramètre	Interprétation
Dépendance (2 variables binaires)	Coefficient phi (ϕ)	Varie entre -1 et 1.
Dépendance (variables nominales)	V de Cramér (V)	Mesure la force de l'association entre deux variables nominales. Varie entre 0 et 1.
Dépendance (variables nominales)	Coefficient de contingence (C)	Varie entre 0 et C_max (dépendant de la taille du tableau).
Dépendance (variables ordinales)	Coefficient de corrélation de Spearman (ρ_s)	Mesure la force et la direction de l'association monotone entre deux variables ordinales. Varie entre -1 et 1.

1.3.2.2 Dépendance monotone

- Uniquement pour des variables ordinales (ou une relation monotone non linéaire entre deux quantitatives)
- Mesure de la dépendance monotone entre variables ordinales : **coefficient de corrélation de Spearman** (ρ_s)
 - = corrélation de Pearson calculée sur les rangs
 - Varie entre -1 (dépendance monotone négative parfaite) et 1 (dépendance monotone positive parfaite)
 - 0 : pas de dépendance monotone entre X et Y = indépendance monotone

1.3.2.3 Concordance entre variables catégorielles

- Mesure de la concordance entre variables catégorielles : **kappa de Cohen** (κ)
 - Pour deux évaluateurs classant des sujets dans des catégories (variables nominales ou ordinales)
 - Varie entre -1 (désaccord parfait) et 1 (accord parfait)
 - 0 : accord équivalent au hasard
 - $\kappa = \frac{\text{concordance observée} - \text{concordance due au hasard}}{1 - \text{concordance due au hasard}}$

1.3.2.4 OR et RR pour variables binaires

Pour un tableau de contingence 2×2 :

	Événement présent	Événement absent
Groupe exposé	a	b
Groupe non exposé	c	d

- **Odds Ratio (OR) :**

- Mesure la force de l'association entre deux variables binaires
- = (odds d'un événement dans le groupe exposé) / (odds d'un événement dans le groupe non exposé)
- $= \frac{a/b}{c/d} = \frac{a \times d}{b \times c}$
- Interprétation :
 - * OR = 1 : pas d'association entre exposition et événement
 - * OR > 1 : exposition associée à une augmentation des odds de l'événement
 - * OR < 1 : exposition associée à une diminution des odds de l'événement
- Attention 1 : les odds ne sont pas des probabilités !
- Attention 2 : pour des événements fréquents, l'OR peut surestimer la force de l'association par rapport au RR

- **Risque relatif (RR) :**

- Mesure la force de l'association entre deux variables binaires
- = (risque d'un événement dans le groupe exposé) / (risque d'un événement dans le groupe non exposé)
- $= \frac{a/(a+b)}{c/(c+d)}$
- Interprétation :
 - * RR = 1 : pas d'association entre exposition et événement
 - * RR > 1 : exposition associée à une augmentation du risque de l'événement
 - * RR < 1 : exposition associée à une diminution du risque de l'événement
- Le RR est souvent plus intuitif que l'OR, surtout pour des événements fréquents

- Cependant, le RR ne peut être estimé directement que dans des études de cohorte ou essais cliniques, pas dans des études cas-témoins
- Rapport entre OR et RR :
 - Pour des événements rares (incidence $\leq 10\%$), $OR \approx RR$
 - Pour des événements fréquents, OR peut diverger significativement de RR

1.3.2.5 Sensibilité et spécificité, VPP et VPN

- Le kappa de Cohen mesure une **concordance globale et symétrique** entre deux évaluateurs, mais ne distingue pas les types d'erreurs (faux positifs vs faux négatifs)
- Pour évaluer la performance d'un test diagnostique binaire, on utilise des mesures spécifiques :
 - Sensibilité = proportion de vrais positifs parmi les cas positifs réels = $P(Y = 1|X = 1)$
 - Spécificité = proportion de vrais négatifs parmi les cas négatifs réels = $P(Y = 0|X = 0)$
- Le problème de la sensibilité et spécificité est qu'elles ne tiennent pas compte de la prévalence de la condition dans la population (par exemple, dans le dépistage d'une maladie rare, un test peut avoir une haute sensibilité et spécificité mais produire beaucoup de faux positifs)
 - VPP = proportion de vrais positifs parmi les résultats positifs du test = $P(X = 1|Y = 1)$
 - VPN = proportion de vrais négatifs parmi les résultats négatifs du test = $P(X = 0|Y = 0)$

Dans un tableau de contingence 2×2 :

	Condition présente	Condition absente
Test positif	a	b
Test négatif	c	d

- Sensibilité = $a / (a + c)$
- Spécificité = $d / (b + d)$
- VPP = $a / (a + b)$
- VPN = $d / (c + d)$

2 Tests statistiques

2.1 Comment les utiliser ?

2.1.1 Intérêt

- Outils d'inférence : passer de ce qui est observé sur un échantillon à une affirmation sur une population plus large
- Langage commun de la communauté scientifique

2.1.2 Légitimité

- Usage légitime si répond à une question scientifique précise, formulée a priori
- Usage illégitime si utilisé pour "fouiller" les données à la recherche d'effets significatifs :
 - Test de normalité (Shapiro-Wilk, Kolmogorov-Smirnov, etc.) sur chaque variable
 - Tester les caractéristiques des perdus de vue...
- Donc n'est pas un substitut au raisonnement scientifique

2.2 Types de tests

2.2.1 Comparaison de pourcentages

2.2.1.1 Test du Chi-carré d'indépendance

- Compare la **répartition des fréquences** observées dans un tableau de contingence à celles attendues sous l'hypothèse d'indépendance entre les variables
- En fait, il teste l'hypothèse nulle d'indépendance entre deux variables catégorielles
- Conditions de validité :
 - Effectifs théoriques ≥ 5 dans chaque case des tableaux
 - Observations indépendantes

- Distribution des effectifs suit une loi du Chi-carré (c'est à dire que les effectifs sont suffisamment grands pour que la distribution asymptotique du Chi-carré soit une bonne approximation de la distribution réelle des effectifs sous l'hypothèse nulle)
- Si conditions non remplies : test exact de Fisher (pour tableaux 2×2)
- Rapport Chi-2 et OR :
 - Le test du Chi-carré teste l'hypothèse nulle d'indépendance entre deux variables catégorielles, ce qui est équivalent à tester si l'odds ratio (OR) est égal à 1 dans un tableau de contingence 2×2 .
 - Donc la p-value du test du chi-carré et celle du test "OR=1" sont identiques.
 - Si chi-carré est significatif, ça vaut le coup de calculer l'OR pour quantifier la force de l'association entre les variables.

2.2.1.2 Test exact de Fisher

- Compare la **répartition des fréquences** dans des petits échantillons
- Utilisé quand les effectifs théoriques sont trop petits pour le test du Chi-carré

2.2.1.3 Comparaison d'un pourcentage à une valeur théorique

- Test exact de binomial
- Situation peu fréquente en pratique ! Par exemple équiprobabilité des naissances garçons/filles

2.2.1.4 Comparaison de 3 pourcentages ou plus

- Situation peu fréquente non plus !
- Idem avec un test du Chi-carré d'indépendance
- Le problème pour l'interprétation de la p-value : si significatif, on ne sait pas où est la différence
- Il faut dans ce cas faire un *Chi-carré de tendance*

2.2.2 Comparaison de moyennes

- Test t de Student
- test de Welch
- test t de Student apparié
- test non paramétrique de Wilcoxon
 - test de Mann-Whitney pour échantillons indépendants
 - test de Wilcoxon pour échantillons appariés

Note

Résumé des tests de comparaison de moyennes

- Comparaison de moyennes avec test t de Student :
 - Effectif : > 30 par groupe ou à peu près égaux (11 et 12 par exemple)
 - Distribution : symétrique (approximativement normale) et variances à peu près égales
- Si échantillons appariés : test t apparié, ou Wilcoxon apparié (= signed rank)
- Si distribution très asymétrique : test non paramétrique de Mann-Whitney (= Wilcoxon non apparié = rank sum)
- Si effectifs très inégaux ou variances très différentes : test de Welch

2.2.2.1 Test t de Student

- Compare les moyennes de deux échantillons indépendants
- Conditions de validité :
 - Variable continue
 - Distribution normale dans chaque groupe (en pratique : $n \geq 30$ par groupe)
 - Homogénéité des variances entre les groupes = homoscedasticité, **appréciée par QQplot !!**
- Si conditions non remplies : test de Welch (si variances inégales) ou test non paramétrique de Mann-Whitney (si distribution non normale)

2.2.2.2 Test de Welch

- Compare les moyennes de deux échantillons indépendants
- Variante du test t de Student qui ne suppose pas l'homogénéité des variances entre les groupes

2.2.2.3 Test t de Student apparié

- Compare les moyennes de deux échantillons appariés (mêmes sujets mesurés deux fois, ou sujets appariés par paires)

2.2.2.4 Test non paramétrique de Wilcoxon

- Compare les distributions de deux échantillons indépendants ou appariés
- Ne suppose pas la normalité des distributions : utilise les rangs des données plutôt que les valeurs brutes
- Échantillons indépendants : test de Mann-Whitney
- Échantillons appariés : test de Wilcoxon

2.2.2.5 Comparaison de 3 moyennes ou plus

- Comparaison pas si fréquente !
- Comparaison de moyennes avec ANOVA :
 - Si une variable à une distribution égale dans plusieurs groupes, alors la variance intra-groupe est proche de la variance inter-groupe
 - On passe par la comparaison des variances car pas de test t pour plus de 2 groupes
 - ANOVA : calcule le ratio de la variance inter-groupe sur la variance intra-groupe
- Conditions de validité :
 - Indépendance des observations
 - Normalité des distributions dans chaque groupe et homogénéité des variances entre les groupes (homoscédasticité appréciée par QQplot)
- Si conditions non remplies : test non paramétrique de Kruskal-Wallis
- Sur R : il faut passer par une régression linéaire (`lm` ou `lmer`) pour faire une ANOVA car :
 - ANOVA est un cas particulier de régression linéaire quand les variables explicatives sont catégorielles

- La régression linéaire et l'ANOVA utilisent toutes les deux le même principe : analyser la variance totale en variance expliquée par le modèle et variance résiduelle.

2.2.3 Test de nullité d'un coefficient de corrélation

- Les coefficients de corrélation (Pearson, Spearman, etc.) mesurent la force et la direction de la relation entre deux variables
- Coefficients de corrélation :
 - Variables quantitatives :
 - * Coefficient de corrélation linéaire de Pearson (variables continues). Varie entre -1 et 1.
 - * Coefficient de corrélation de Spearman (variables ordinales). Varie entre -1 et 1. Basé sur les rangs.
 - Variables catégorielles :
 - * Coefficient phi (ϕ) pour tableaux 2×2 . Varie entre -1 et 1.
 - * V de Cramér (V) pour tableaux $r \times c$. Varie entre 0 et 1.
 - Accord entre évaluateurs :
 - * Coefficient de corrélation intraclasse (ICC(2,1)). Varie entre $-\infty$ et 1 (en pratique 0 à 1).
 - * Kappa de Cohen (κ). Varie entre -1 et 1.
- Test de nullité d'un coefficient de corrélation :
 - Permet de tester l'hypothèse nulle selon laquelle il n'y a pas de relation entre les variables
 - Coefficient : pas forcément égal à 0 dans la population !!

2.2.4 Tests appariés

- Utilisés lorsque les observations sont liées ou dépendantes
- Exemple : mesures répétées sur les mêmes sujets, ou sujets appariés par paires
- Tests appariés courants :
 - Sur des moyennes :
 - * Test t de Student apparié : compare les moyennes de deux échantillons appariés
 - * Test de Wilcoxon apparié : compare les distributions de deux échantillons appariés (non paramétrique)
 - Sur des pourcentages :
 - * Test du Chi2 de McNemar : compare les proportions dans des échantillons appariés (variables binaires)
 - * Test exact du Chi2 de McNemar : version exacte du test de McNemar pour petits échantillons

2.3 Calcul de β et du nombre de sujets à inclure

- Dans la théorie de Neyman et Pearson, on fixe un seuil de risque de première espèce (α) et un seuil de risque de deuxième espèce (β)
- α : probabilité de rejeter l'hypothèse nulle alors qu'elle est vraie (faux positif)
- β : probabilité de ne pas rejeter l'hypothèse nulle alors qu'elle est fausse (faux négatif)
- La puissance d'un test est égale à $1 - \beta$: probabilité de rejeter l'hypothèse nulle quand elle est fausse (vrai positif)
- β varie selon :
 - La différence vraie : plus la différence vraie est grande, plus la puissance est élevée (et β faible)
 - Le nombre de sujets inclus : plus le nombre de sujets est grand, plus la puissance est élevée (et β faible)
 - La variabilité des données : plus la variabilité est faible, plus la puissance est élevée (et β faible)
 - H_1 choisi : plus H_1 est éloigné de H_0 , plus la puissance est élevée (et β faible).
 - ✱ S'ils sont proches, il est plus difficile de les distinguer donc β est plus élevé

2.4 Tests uni- ou bilatéraux

- Selon Fisher :
 - L'inférence statistique est **bilatérale** par nature :
 - * on cherche à apprécier l'adéquation entre un échantillon observé et une population hypothétique infinie sous-jacente
 - * donc test si l'estimateur issu de l'échantillon est supérieur ou inférieur à la valeur hypothétique
- Selon Neyman et Pearson :
 - inférence statistique **unilatérale** ou **bilatérale** selon le choix des hypothèses 0 et 1

2.5 Intervalles de confiance

- Définitions :
 - IC = intervalle estimé à partir des données d'un échantillon, qui a une probabilité donnée (le niveau de confiance) de contenir la vraie valeur du paramètre dans la population.
 - IC à 95% = si on répétait l'expérience de prélèvement d'échantillons et de calcul d'IC plusieurs fois, environ 95% des IC ainsi obtenus contiendraient la vraie moyenne de la population.
- **Ça ne veut pas dire "95% des valeurs sont dans l'IC" !!**

3 Modèles