Interactions et modifications d'effet en Epidémiologie

CERPOP, INSERM, EQUITY Team

Last compiled on 20 septembre, 2023

Contents

Chapter 1

Présentation

Ce document a été rédigé en tant que document de synthèse du travail du groupe "Interaction" de l'équipe EQUITY, CERPOP. Ce travail a consisté en une revue de la littérature et en une application détaillée des méthodes sur des analyses illustratives, dans un but d'auto-formation et pédagogique.

Les participant.e.s du groupe de travail sont :

- Hélène COLINEAUX

- Léna BONIN
- Camille JOANNES
- Benoit LEPAGE
- Lola NEUFCOURT
- Ainhoa UGARTECHE

The online version of this book is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Chapter 2

Introduction

Comment telle prédisposition génétique et telle exposition environnementale *inter-agissent*-elles ? L'effet de tel traitement varie-t-il selon les circonstances ? Selon les caractéristiques du patient ? Telle intervention peut-elle être bénéfique pour un groupe social et délétère pour un autre ?

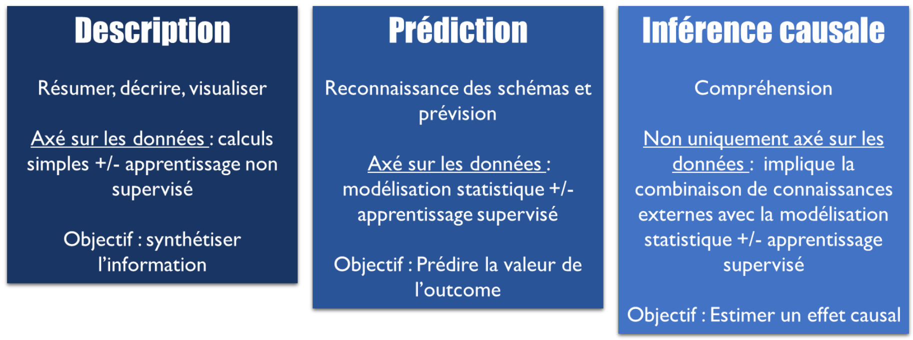
De nombreuses questions épidémiologiques impliquent des mécanismes d'interactions ou de modifications d'effet. Pourtant, étudier ces mécanismes restent encore complexe aujourd'hui sur le plan méthodologique : quelle démarche adopter ? sur quelle échelle mesurer cette interaction ? comment interpréter les coefficients ? et cetera.

Dans ce document, nous proposons une synthèse de la littérature et une démarche progressive et appliquée pour explorer ces questions.

2.1 Quand étudier les interactions ?

2.1.1 *Prediction* versus *causalité*

La science des données cherche à répondre à 3 types d'objectifs Hernán et al. [2019] :



Selon le type d'objectif, la démarche d'analyse et les enjeux méthodologiques ne vont pas être les mêmes. Si l'objectif est prédictif, la démarche va être centrée sur la *prédiction de l'outcome*, à partir de covariables sélectionnées afin d'optimiser les performances de la prédiction, tout en prenant en compte leur disponibilité en pratique et la parcimonie du modèle.

Dans une démarche explicative, ou *étiologique*, au contraire, la démarche va être centrée sur l'*estimation d'un effet causal*, en prenant en compte les covariables en fonction de leur rôle vis-à-vis de l'effet d'intérêt (facteurs de confusion, colliders, médiateurs…).

En épidémiologie, à l'exception des cas où l'on souhaite développer un test ou score diagnostique ou pronostique, les objectifs sont le plus souvent explicatifs. On cherche en effet, la plupart du temps, à identifier des liens de cause à effet, afin de pouvoir agir sur les causes pour modifier les effets.

Finalement, pour répondre à la question "quand doit-on prendre en compte les interactions ?", il est d'abord nécessaire d'identifier dans quel type de démarche l'on s'inscrit :

- **Démarche prédictive** : on ajoutera alors les interactions dans le modèle de prédiction, pour le rendre plus *flexible*, si cela améliore les performances de la prédiction VanderWeele and Knol [2014].
- **Démarche explicative/étiologique** : on étudiera les interactions ou modifications d'effet, si cela répond directement à l'objectif. Par exemple :
  - Si l'objectif est du type "l'effet de X sur Y varie-t-il en fonction de V ?", on prendra en compte l'interaction entre X et V.
  - Les objectifs qui nécessitent la prise en compte de l'interaction peuvent aussi être du type : "Quel est l'effet conjoint de X et V sur Y ?" ou "Quel part de l'effet de X sur Y disparaît quand V est modifié ?", etc.
  - Par contre, si l'objectif est simplement d'estimer l'effet de X sur Y, ou l'effet médié par un médiateur M, la prise en compte des interactions entre X et des covariables (facteurs de confusion ou médiateurs)

n'est pas indispensable pour répondre à la question scientifique. Un effet "moyen" pourra être estimé. Des termes d'interactions peuvent cependant être ajoutés (mais non interprétés), si cela améliore la précision de l'estimation (enjeu d'optimisation du modèle).

2.1.2 Types d'objectifs

Dans ce document, nous nous intéresserons principalement aux interactions et modifications d'effet dans une démarche étiologique/ explicative.

Les objectifs pouvant nécessiter l'étude de l'interaction/modification d'effet sont VanderWeele and Knol [2014] :

- **Cibler des sous-groupes**. Par exemple, identifier des sous-groupes pour lesquels l'intervention aura le plus d'effet afin de pouvoir cibler l'intervention en cas de ressources limitées, ou s'assurer que l'intervention est bénéfique pour tous les groupes et pas délétères pour certains groupes.
- **Explorer les mécanismes d'un effet**. Par exemple, en cas d'intervention qui n'a d'effet qu'en présence ou absence d'une caractéristiques particulière (définition mécanistique de l'interaction) ou seulement conjointement à une autre intervention.
- **Etudier l'effet d'une intervention pour éliminer une partie de l'effet d'une exposition non modifiable**. Par exemple, quelle part de l'effet du niveau d'éducation des parents sur la mortalité disparaîtrait si on intervenait sur le tabagisme à l'adolescence ? Ce type d'objectif est proche d'un objectif ciblant la *médiation* d'un effet, par exemple la médiation de l'effet du niveau d'éducation des parents *par* le tabagisme, mais les mécanismes envisagés et explorés ne sont pas exactement les mêmes. Explorer ces deux types de mécanismes peut nécessiter des approches spécifiques (voir chapitre 17)

2.2 Les points les plus importants

La première étape importante consiste donc à **définir précisément l'objectif** :

- L'objectif est-il de type descriptif, prédictif ou explicatif ?
- Si l'on est dans une démarche explicative, d'inférence causale, est-ce que la mesure d'un effet d'interaction est nécessaire pour y répondre ? (identifier précisément l'effet que l'on cherche à estimer, ou *estimand*).

Ensuite, de **nombreuses questions** se posent pour réaliser une analyse d'interaction, auxquelles nous tentons de répondre dans ce document :

- S'agit-il d'une interaction ou une modification d'effet ? (Chapitre 4)
- Sur quelle échelle la mesure-t-on ? Un effet d'interaction peut en effet être défini sur une échelle multiplicative ou additive, et les résultats entre ces échelles peuvent être contradictoires. (Chapitre 5)
- Quels paramètres présenter et comment les interpréter ? (Chapitre 6)
- Comment estimer ces paramètres ? (Chapitre 9 et Chapitre 10)
- Comment représenter cette interaction graphiquement ? (Chapitre 11)

2.3 Avertissement

Les analyses d'effets d'interaction (ou de modifications d'effets) sont peu puissantes. Pour observer un effet d'interaction "statistiquement signitificatif", le nombre de sujets nécessaire est habituellement beaucoup plus élevé que le nombre de sujets nécessaire permettant d'observer une différence global entre 2 moyennes ou pourcentages.

A titre d'exemple, Brookes S et al (2004) Brookes et al. [2004] décrivent que dans un contexte d'essai contrôlé randomisé à deux bras parallèles, équilibrés, incluant un nombre de sujets N optimisé pour observer une différence Δ significative entre un groupe exposé et un groupe non-exposé avec une puissance de 80%, la puissance pour observer un effet d'interaction de taille similaire ($\approx \Delta$) ne sera que de 29%.

Pour observer une effet de taille similaire ($\approx \Delta$) de manière significative, le nombre de sujets à recruter sera 4 fois plus élevé ($4 \times N$). Si l'on cherche à mesurer de manière significative des effets d'interaction plus petits que l'effet global entre groupe exposé et non-exposé, le nombre de sujets nécessaires augmente de manière spéctaculaire :

- 6 fois plus élevé ($6 \times N$) pour rechercher une interaction un peu plus petite correspondant à 80% de l'effet global ($0,80 \times \Delta$),
- 15 fois plus élevé ($15 \times N$) pour rechercher une interaction égale à la moitié de l'effet global $\left(\frac{\Delta}{2}\right)$,
- 100 fois plus élevé ($100 \times N$) pour rechercher une petite interaction égale à 20% de l'effet global $(\left(\frac{\Delta}{5}\right))$.

Part I

Synthèse de la littérature

Chapter 3

Définitions préalables

3.1 Variables et probabilités

On note :

- un outcome : Y,
- deux expositions : X et V

La probabilité de l'outcome Y dans chaque strate définie par les 2 expositions est notée :

- $p_{xv} = P(Y = 1|X = x, V = v)$

  Exemple

  On a deux expositions X, le tabagisme actif à 20 ans, et V, le fait d'avoir vécu un évènement traumatique pendant l'enfance. L'outcome Y est binaire et représente le fait d'avoir au moins une pathologie chronique à 60 ans ($Y = 1$) ou aucune ($Y = 0$).

  On décrit (données complètement fictives) :

X \ V	**V = 0**	**V = 1**
X = 0	P_{00} = 0,1	P_{10} = 0,2
X = 1	P_{01} = 0,4	P_{11} = 0,9

  Interprétation : La probabilité d'avoir au moins une pathologie chronique à 60 ans quand on n'a pas vécu d'événement traumatique pendant l'enfance et pas fumé à 20 ans est de 10%, tandis qu'elle est de 90% quand on a vécu un événement traumatique et fumé.

3.2 Mesures d'effets

L'effet d'une variable X sur Y peut être mesuré sur deux échelles : additive (différence de risques ou de probabilités) ou multiplicative (rapport de risques ou de probabilités).

Concernant les différences de risques (DR, effets additifs)

On va noter $P(Y = 1|do(X = 1))$ la probabilité d'observer $Y = 1$ sous une intervention contrefactuelle où la totalité de la population étudiée est exposée à $X = 1$ (notée $do(X = 1)$).

De même, on va noter $P(Y = 1|do(X = 1, V = 1))$ la probabilité d'observer $Y = 1$ sous une intervention contrefactuelle conjointe à la fois sur X et sur V où la totalité de la population étudiée est exposée à $X = 1$ et $V = 1$ (notée $do(X = 1, V = 1)$).

- L'effet d'une exposition X binaire sur Y est : $DR(X) = P(Y = 1|do(X = 1)) - P(Y = 1|do(X = 0))$
  - qu'on peut estimer, si les conditions d'identifiabilité sont réunies,
  - par $P(Y = 1|X = 1) - P(Y = 1|X = 0) = p_1 - p_0$
- L'effet conjoint de X et V est : $DR(X, V) = p_{11} - p_{00}$
- L'effet de X sur Y pour chaque valeur fixée de V est : $DR(X, V = 0) = p_{10} - p_{00}$ et $DR(X, V = 1) = p_{11} - p_{01}$

Exemple

Différences de risques pour l'exemple 1

X \ V	**V = 0**	**V = 1**
X = 0	$P_{00} = 0{,}1$	$P_{10} = 0{,}2$
X = 1	$P_{01} = 0{,}4$	$P_{11} = 0{,}9$

- $DR(X, V) = p_{11} - p_{00} = 0,90 - 0,10 = +0,80$
- $DR(X, V = 0) = p_{10} - p_{00} = 0,40 - 0,10 = +0,30$
- $DR(X, V = 1) = p_{11} - p_{01} = 0,90 - 0,20 = +0,70$

Le fait d'être doublement exposé (tabagisme + événement traumatique) par rapport à pas du tout augmente le risque d'avoir au moins une pathologie chronique à 60 ans de +80%. Dans une population n'ayant pas vécu d'événement traumatique, le fait de fumer à 20 ans augmente le risque d'avoir au moins une pathologie chronique à 60 ans de +30%, alors que dans une population ayant vécu un événement traumatique, il est augmenté de +70%.

Concernant les rapports de risques (RR, effets multiplicatifs)

On peut notamment utiliser les **risques relatifs** (RR). On donc :

- L'effet d'une exposition X binaire sur Y est :
  - $RR(X) = \frac{P(Y=1|do(X=1))}{P(Y=1|do(X=0))}$
  - qu'on peut estimer, si les conditions d'identifiabilité sont réunies, par :
  - $\frac{P(Y=1|do(X=1))}{P(Y=1|do(X=0))} = \frac{p_1}{p_0}$
- L'effet conjoint de X et V est : $RR(X,V) = \frac{p_{11}}{p_{00}}$
- L'effet de X sur Y pour chaque valeur fixée de V est :
  - $RR(X, V=0) = \frac{p_{10}}{p_{00}}$
  - et $RR(X, V=1) = \frac{p_{11}}{p_{01}}$

  Exemple

  Risques relatifs pour l'exemple 1

X \ V	V = 0	V = 1
X = 0	P_{00} = 0,1	P_{10} = 0,2
X = 1	P_{01} = 0,4	P_{11} = 0,9

  - $RR(X,V) = \frac{0{,}9}{0{,}1} = \times 9$
  - $RR(X, V=0) = \frac{0{,}4}{0{,}1} = \times 4$
  - $RR(X, V=1) = \frac{0{,}9}{0{,}2} = \times 4,5$

  Le risque d'avoir au moins une pathologie chronique à 60 ans quand on est doublement exposé (tabagisme + événement traumatique) par rapport à pas du tout est multiplié par 9. Dans une population n'ayant pas vécu d'événement traumatique, le fait de fumer à 20 ans multiplie le risque par 4, alors que dans une population ayant vécu un événement traumatique, il est multiplié par 4,5.

Une autre échelle multiplicative fréquemment utilisée est l'échelle des odds-ratios (OR)

L'échelle des **Odds-Ratios** (OR) est fréquemment utilisée car on peut l'obtenir facilement à partir d'un modèle de régression logistique (en utilisant l'exponentielle des coefficients de la régression logistique). L'**odds** correspond à la cote d'une probabilité p et est définie par $\text{odds}(p) = \frac{p}{1-p}$. L'odds-ratio est le rapport de la cote dans le groupe exposé divisée par la cote dans le groupe non-exposé.

Si on reprend l'exemple précédent :

- L'effet d'une exposition X binaire sur Y est :
  - $OR(X) = \frac{P(Y=1|do(X=1))/[1-P(Y=1|do(X=1))]}{P(Y=1|do(X=0))/[1-P(Y=1|do(X=0))]}$
  - qu'on peut estimer, si les conditions d'identifiabilité sont réunies, par :
  - $\frac{P(Y=1|do(X=1))/[1-P(Y=1|do(X=1))]}{P(Y=1|do(X=0))/[1-P(Y=1|do(X=0))]} = \frac{p_1/(1-p_1)}{p_0/(1-p_0)}$
- L'effet conjoint de X et V est : $OR(X,V) = \frac{p_{11}/(1-p_{11})}{p_{00}/(1-p_{00})}$
- L'effet de X sur Y pour chaque valeur fixée de V est :
  - $OR(X, V=0) = \frac{p_{10}/(1-p_{10})}{p_{00}/(1-p_{00})}$
  - et $OR(X, V=1) = \frac{p_{11}/(1-p_{11})}{p_{01}/(1-p_{01})}$

Exemple

Odds-ratios pour l'exemple 1

X \ V	V = 0	V = 1
X = 0	P_{00} = 0,1	P_{10} = 0,2
X = 1	P_{01} = 0,4	P_{11} = 0,9

- $OR(X,V) = \frac{0{,}9/(1-0{,}9)}{0{,}1/(1-0{,}1)} = \times 81$
- $OR(X,V=0) = \frac{0{,}4/(1-0{,}4)}{0{,}1/(1-0{,}1)} = \times 6$
- $OR(X,V=1) = \frac{0{,}9/(1-0{,}9)}{0{,}2/(1-0{,}2)} = \times 36$

La cote d'avoir au moins une pathologie chronique à 60 ans quand on est doublement exposé (tabagisme + événement traumatique) par rapport à pas du tout est multiplié par 81. Dans une population n'ayant pas vécu d'événement traumatique, le fait de fumer à 20 ans multiplie par 6 la cote d'avoir au moins une pathologie chronique à 60 ans, alors que dans une population ayant vécu un événement traumatique, elle est multipliée par 36.

3.3 Effets conditionnels et marginaux

Dans une analyse de l'effet d'interaction entre deux expositions binaires X et V sur un outcome Y, il sera parfois nécessaire de prendre en compte un ensemble de facteurs de confusion pour estimer les effets causaux. On note L cet ensemble qui peut être constitué par exemple de 3 facteurs de confusion $L = \{L_1 = age, L_2 = sexe, L_3 = comorbidits\}$.

Au-delà de l'échelle des mesures d'association (additive pour les DR, multiplicative pour les RR et OR), il faudra choisir si on présente des mesures d'associations :

- **conditionnelles** c'est-à-dire estimées dans des strates définies par l'ensemble (ou par un sous-ensemble) des facteurs de confusion
- ou **marginales**, c'est à dire un effet moyen estimé pour l'ensemble de la population (une moyenne pondérée des associations observées dans les différentes strates de la population).

Par exemple, sur l'échelle des odds-ratios, une méthode classiquement utilisée pour estimer l'effet d'interaction entre X et V est d'appliquer une régression logistique de Y en fonction de X et V, de leur interaction, ajustée sur les 3 facteurs de confusion (et on suppose que le modèle est correctement spécifié) :

$$\text{logit}P(Y = 1 \mid X, V, L_1, L_2, L_3) = \beta_0+\beta_X X+\beta_V V+\beta_{X*V} X*V+\beta_{L1} L_1+\beta_{L2} L_2+\beta_{L3} L_3$$

A partir de ce modèle, il est possible d'estimer directement :

- l'interaction conjointe de X et V : $OR(X, V)\big|L_1, L_2, L_3 = \exp(\beta_X + \beta_V + \beta_{X*V})$
- L'effet de X sur Y pour chaque valeur fixée de V :
  - $OR(X, V = 0)\big|L_1, L_2, L_3 = \exp(\beta_X)$
  - $OR(X, V = 1)\big|L_1, L_2, L_3 = \exp(\beta_X + \beta_{X*V})$

Il s'agit d'**OR conditionnels**, c'est-à-dire "toutes choses égales par ailleurs au niveau individuel", conditionnellement au sexe, à l'âge et aux comorbidités de chaque individu : d'après ce modèle, l'odds ratio obtenu est indépendant du sexe, de l'âge et des comorbidités. Sa valeur sera identique chez un homme de 35 ans sans comorbidités et chez une femmes de 60 ans avec comorbidités.

A partir du même modèle, on peut également estimer des associations **marginales**, en calculant l'effet moyen observé dans la population. Par exemple pour l'interaction conjointe de X et V,

- on calcule d'abord l'effet populationnel associé à une double exposition $X = 1$ et $V = 1$, comme une moyenne pondérée des probabilités attendues dans chaque strate $\{l_1, l_2, l_3\}$ définie par les facteurs de confusion :

  $$p_{11} = \sum_{l_1,l_2,l_3} P(Y = 1 \mid X = 1, V = 1, L_1 = l_1, L_2 = l_2, L_3 = l_3)\times P(L_1 = l_1, L_2 = l_2, L_3 = l_3)$$

- puis on calcul l'effet populationnel associé à une double absence d'exposition $X = 0$ et $V = 0$:

  $$p_{00} = \sum_{l_1,l_2,l_3} P(Y = 1 \mid X = 0, V = 0, L_1 = l_1, L_2 = l_2, L_3 = l_3)\times P(L_1 = l_1, L_2 = l_2, L_3 = l_3)$$

- l'**odds-ratio marginal** peut-être obtenu à partir de ces deux probabilités populationnelles $OR(X, V) = \frac{p_{11}/(1-p_{11})}{p_{00}/(1-p_{00})}$. C'est la méthode qui est appliquée en *G-computation* (cf. paragraphe 10.1). Si l'on a bien pris en compte les facteurs de confusion, l'interprétation se fait comme une mesure d'association causale moyennée au niveau de l'ensemble de la population ("toutes choses égales par ailleurs au niveau populationnel", la population étant caractérisée par sa distribution de sexe, d'âge et de comorbidités).

Selon le même principe, on peut calculer des risques relatifs (RR) conditionnels ou marginaux, et des différences de risques (DR) conditionnelles ou marginales.

Une propriété intéressantes des RR et des DR est que se sont des mesures d'associations **collapsibles** (anglicisme venant du terme anglais *collapsibility*) : la mesure conditionnelle est la même que la mesure marginale. Whitcomb and Naimi [2021]

En revanche, les odds-ratios (OR) sont des mesures d'associations **non-collapsibles**, c'est-à-dire qu'un OR conditionnel sera différent d'un OR marginal (en dehors de cas particuliers où l'exposition n'a aucun effet causal sur l'outcome ou bien lorsqu'aucun des facteurs de confusion potentiel n'a d'effet sur l'outcome Y). Cela est parfois source de confusion car :

- il s'agit de deux *estimands* différents (par définition $OR_{\text{marginal}} \neq OR_{\text{conditionnel}}$, sauf cas particulier),
- mais l'OR marginal comme l'OR conditionnel sont tous les deux des mesures d'association causales valides (à partir du moment où les facteurs de confusion ont bien été pris en compte dans le calcul de l'OR conditionnel ou de l'OR marginal). Daniel et al. [2021]

Le choix de présenter une association marginale ou une association conditionnelle va donc influencer la valeur du résultat présenté, en particulier si l'on présente des mesures d'association "non-collapsibles" comme les OR.

Chapter 4

Interaction ou modification d'effets

Dans le champ des analyses d'interaction, deux termes peuvent être rencontrés : "interaction" et "modification d'effet". Quel est la différence entre ces deux termes ?

4.1 Modification d'effets

La question de la modification d'effet consiste à identifier si un scénario contrefactuel modifiant le traitement ou l'exposition X donne un résultat différent dans différents groupes V de patients (estimer l'effet d'une exposition séparément en fonction d'une autre variable) Corraini et al. [2017].

Si l'on compare avec un essai d'intervention, c'est comme s'il y avait une seule intervention X et que l'analyse était stratifiée sur V. On analyse donc l'effet du scénario $do(X)$ dans chaque groupe de V.

En observationnel, l'effet causal qui nous intéresse est donc celui de X mais pas celui de V. On ajustera sur les facteurs de confusion de la relation $X \rightarrow Y$.

On ne fait pas d'hypothèse sur les mécanismes de la modification d'effet, qui peut être causale (de façon directe ou indirecte), ou non-causale (présence d'une modification d'effet par proxy ou cause commune, sans qu'il existe d'effet direct ou indirect du modificateur d'effet vers le critère de jugement, comme dans la figure en bas de page) VanderWeele and Robins [2007].

> Exemples d'objectifs : identifier des groupes pour lesquels le traitement ne serait pas utile, ou explorer si l'effet du traitement est homogène/hétérogène en fonction de l'âge, du sexe, etc.

On a une modification de l'effet de X par V si l'effet de X est différent dans deux strates définies par V:

- en additif : $DR(X|V=0) \neq DR(X|V=1)$
  - soit $p_{10} - p_{00} \neq p_{11} - p_{01}$

- en multiplicatif : $RR(X|V=0) \neq RR(X|V=1)$
  - soit $\frac{p_{10}}{p_{00}} \neq \frac{p_{11}}{p_{01}}$

  Exemple

  Modification d'effet dans l'exemple 1 L'objectif serait formulé ainsi : l'effet du tabagisme X sur le risque de maladie chronique Y est-il différent lorsqu'on a ou non vécu un événement traumatique V antérieurement ?

  Les données (fictives) :

X \ V	**V = 0**	**V = 1**
X = 0	$P_{00} = 0{,}1$	$P_{10} = 0{,}2$
X = 1	$P_{01} = 0{,}4$	$P_{11} = 0{,}9$

  En additif :

  - effet dans le groupe $V = 0$: $DR(X|V=0) = 0,40 - 0,10 = +0,30$
  - effet dans le groupe $V = 1$: $DR(X|V=1) = 0,90 - 0,20 = +0,70$
  - donc $DR(X|V=0) \neq DR(X|V=1)$

  En multiplicatif :

  - effet dans le groupe $V = 0$: $RR(X|V=0) = \frac{0{,}40}{0{,}10} = \times 4,0$
  - effet dans le groupe $V = 1$: $RR(X|V=1) = \frac{0{,}90}{0{,}20} = \times 4,5$
  - donc $RR(X|V=0) \neq RR(X|V=1)$

  Ici l'effet du tabagisme est différent selon que les personnes ont vécu un événement traumatique ou non, sur l'échelle additive et multiplicative (données fictives). On peut donc dire que le fait d'avoir vécu un événement traumatique modifie l'effet du tabac. Attention, dans cet exemple, on fait l'hypothèse de l'absence de facteurs de confusion entre le tabagisme et l'outcome Y, ce qui est en réalité peu probable.

Lorsqu'on utilise les approches causales pour estimer l'effet de X sur Y, on va intervenir seulement sur X. En G-computation, le code serait :

```r
#modèle
Q.model <- glm(data=bootData, formula = Y ~ X + V + X*V + L,family = binomial)

# Scénarios #
data.X0 <-  data.X1 <- bootData
data.X0$X <- 0
data.X1$X <- 1

# Y contrefactuel
bootData$Y.X0.pred <- predict(Q.model, newdata = data.X0, type = "response")
bootData$Y.X1.pred <- predict(Q.model, newdata = data.X1, type = "response")

# Modification d'effet, échelle additive
simu.base$est.AI[simu.base$i.simu==i] = round(
 # effet de X quand V==1
 mean(bootData$Y.X1.pred[which(bootData$V == 1),]) -
   bootData$Y.X0.pred[which(bootData$V == 1),]) -
 # effet de X quand V==0
 mean(bootData$Y.X1.pred[which(bootData$V == 0),]) -
  bootData$Y.X0.pred[which(bootData$V == 0),]),4)
```

Remarque : on ne peut pas considérer V comme un modificateur de l'effet de X si X est une cause de V. Par exemple, si X était le tabagisme à 20 ans, V le fait de souffrir de bronchite chronique obstructive à 50 ans et Y la mortalité. Ca n'aurait pas de sens de demander si l'effet du tabac sur la mortalité varie en fonction de la présence ou non de BPCO, car V est un descendant de X (le tabagisme augmente le risque de BPCO). Lorsqu'on intervient sur X, $do(X)$, on modifie donc aussi V car $X \rightarrow V$, on est donc obligé d'intervenir aussi sur V (en faisant une analyse de médiation ou d'interaction) pour étudier l'effet de X en fonction de V, nous ne sommes donc plus dans le cadre d'une modification d'effet.

4.2 Interaction

Quand on s'intéresse à l'interaction, on s'intéresse plutôt à l'effet conjoint de 2 expositions (ou plus) sur un outcome. Il y a une interaction synergique si l'effet conjoint est supérieur à la somme de l'effet individuels. Il y a une interaction antagoniste lorsque l'effet conjoint est inférieur à la somme des effets individuels Corraini et al. [2017].

Si l'on compare avec un essai d'intervention, c'est comme s'il y avait plusieurs interventions, selon le nombre de combinaisons. On analyse donc l'effet du scénario $do(X, V)$. Ici l'effet causal d'intérêt est vraiment l'effet conjoint des deux variables.

Dans un schéma observationnel, l'effet causal qui nous intéresse est donc celui de l'interaction $X * V$. On ajustera sur les facteurs de confusion des deux relations $X \to Y$ et $V \to Y$. On fait l'hypothèse que les mécanismes de l'effet conjoint de X et V sont causaux.

Par définition, on a une interaction si l'effet conjoint de X et V sur Y ($DR(X,V)$) est différent de la somme (ou du produit sur l'échelle multiplicative) :

- de l'effet isolé de X sur Y (où V est constant, fixé à $V = 0$), noté $DR(X, V = 0)$ (ou $RR(X, V = 0)$)
- et de l'effet isolé de V sur Y (où X est constant, fixé à $X = 0$), noté $DR(V, X = 0)$ (ou $RR(V, X = 0)$)

On a ainsi,

- en additif : $DR(X,V) \neq DR(X, V = 0) + DR(V, X = 0)$
  - $p_{11} - p_{00} \neq (p_{10} - p_{00}) + (p_{01} - p_{00})$
  - $p_{11} \neq p_{10} + p_{01} - p_{00}$
  - $p_{11} - p_{10} - p_{01} + p_{00} \neq 0$
- en multiplicatif $RR(X,V) \neq RR(X, V = 0) \times RR(V, X = 0)$
  - $\frac{p_{11}}{p_{00}} \neq \frac{p_{10}}{p_{00}} \times \frac{p_{01}}{p_{00}}$
  - $p_{11} \neq \frac{\frac{p_{10}}{p_{01}}}{p_{00}}$
  - $\frac{p_{00} \times p_{11}}{p_{10} \times p_{01}} \neq 1$

  Exemple

  Interaction dans l'exemple 1 L'objectif serait formulé ainsi : le tabagisme X et le vécu d'un événement traumatique V se potentialisent-il l'un autre pour augmenter le risque de maladie chronique Y ?

  En additif :

  - effet joint : $DR(X,V) = 0,90 - 0,10 = +0.80$
  - somme des effets individuels : $DR(X, V = 0) + DR(V, X = 0) = +0,30 + 0,10 = +0,40$
  - donc $DR(X,V) \neq DR(X, V = 0) + DR(V, X = 0)$

  En multiplicatif :

  - effet joint : $RR(X,V) = \frac{0,9}{0,1} = \times 9$
  - produit des effets individuels : $RR(X, V = 0) \times RR(V, X = 0) = 4 \times 2 = \times 8$
  - donc $DR(X,V) \neq DR(X, V = 0) \times DR(V, X = 0)$

> Ici l'effet joint des 2 expositions est supérieur à la somme ou au produit des effets individuels, il y a donc une interaction synergique entre les deux expositions. On peut conclure que l'expérience d'un événement traumatique et le tabagisme *se potentialisent* pour aboutir à une augmentation du risque de maladies chroniques : ces expositions ont un effet plus fort lorsqu'elles sont présentes toutes les deux que la somme/le produit des deux.

Lorsqu'on utilise les approches causales pour estimer l'effet de X sur Y, on va intervenir sur X et sur Y, contrairement à l'approche précédente ou l'on intervenait seulement sur X. En G-computation, le code serait :

```
#modèle
Q.model <- glm(data=bootData, formula = Y ~ X + V + X*V + L,family = binomial)

# Scénarios #
data.X0V0 <-  data.X0V1 <-  data.X1V0 <-  data.X1V1 <- bootData
data.X0V0$X <-  data.X0V1$X <- 0
data.X1V0$X <-  data.X1V1$X <- 1
data.X0V0$V <-  data.X1V0$V <- 0
data.X0V1$V <-  data.X1V1$V <- 1

# Y contrefactuel
Y.X0V0.pred <- predict(Q.model, newdata = data.X0V0, type = "response")
Y.X1V0.pred <- predict(Q.model, newdata = data.X1V0, type = "response")
Y.X0V1.pred <- predict(Q.model, newdata = data.X0V1, type = "response")
Y.X1V1.pred <- predict(Q.model, newdata = data.X1V1, type = "response")

# Interaction additive
simu.base$est.AI[simu.base$i.simu==i] = round(
 # effet joint
 mean(bootData$Y.X1V1.pred - bootData$Y.X0V0.pred) -
 # somme des effets individuels
 mean(bootData$Y.X0V1.pred - bootData$Y.X0V0.pred) +
 mean(bootData$Y.X1V0.pred - bootData$Y.X0V0.pred),4)
```

4.3 Synthèse

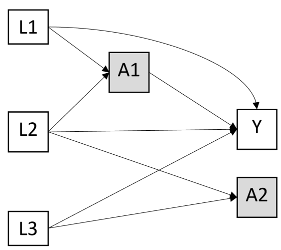
Mathématiquement, les formulations sont équivalentes :

- échelle additive: $p_{10} - p_{00} \neq p_{11} - p_{01} \iff p_{11} \neq (p_{10} + p_{01}) - p_{00}$
- échelle multiplicative : $p_{10}/p_{00} \neq p_{11}/p_{01} \iff p_{11} \neq (p_{10} \times p_{01})/p_{00}$

La différence se joue plutôt sur :

- la façon dont la question est posée (effet de X selon V, *versus* effet conjoint de X et V),
- les hypothèses causales formulées (scénario $do(X)|V$ *versus* $do(X,V)$)
- et donc sur les sets de facteurs de confusion à considérer (seulement sur la relation $X \rightarrow Y$ *versus* les deux relations $X \rightarrow Y$ et $V \rightarrow Y$)
- et sur l'intervention contrefactuelle que l'on va réalisée si l'on utilise des approches causales ($do(X)|V$ *versus* $do(X,V)$).

Il existe des cas où l'identification d'une interaction ou d'une modification d'effet ne conduira pas à la même démarche et donc au même résultat VanderWeele [2009]. Prenons le DAG suivant :



Dans ce cas, il n'y a pas d'interaction entre $A1$ et $A2$, car il n'y a pas d'effet direct ni indirect de $A2 \rightarrow Y$. L'effet de $A1 \rightarrow Y$ restera le même quelle que soit la valeur que l'on pourrait attribuer à $A2$:

$$P\left[Y=1|do(A1=1,A2=0)\right]-P\left[Y=1|do(A1=0,A2=0)\right]=P\left[Y=1|do(A1=1,A2=1)\right] - P\left[Y=1|do(A1=0,A2=1)\right]$$

Par contre, il peut y avoir une modification de l'effet de $A1$ par $A2$, en particulier s'il existe une interaction $A1 * L2 \rightarrow Y$ ou $A1 * L3 \rightarrow Y$, on s'attend à ce que les contrastes suivants soient différents :

$$P[Y=1|do(A1=1),A2=1]-P[Y=1|do(A1=2),A2=1] \neq P[Y=1|do(A1=1),A2=0] - P[Y=1|do(A1=2),A2=0]$$

Chapter 5

La question des échelles

5.1 Mesures des interactions

Echelle additive

Une façon simple de mesurer l'interaction est de mesurer à quel point l'effet conjoint de deux facteurs est différents de la somme de leurs effets individuels VanderWeele and Knol [2014] :

- $AI = DR(X,V) - [DR(X|V=0) + DR(V|X=0)]$
- $AI = (p_{11} - p_{00}) - [(p_{10} - p_{00}) + (p_{01} - p_{00})]$
- soit $AI = p_{11} - p_{10} - p_{01} + p_{00}$

Exemple

Mesure de l'interaction dans l'exemple 1

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	+0,1
X = 1	P10 = +0,4	P11 = +0,9	+0,5
Effet X	+0,3	+0,7	

+0,4

+0,4

+0,4

On retrouve l'effet d'interaction, calculé/exprimé de différentes façon,

Soit la différence entre l'effet joint et la somme des effets individuels (flèche rouge) :

- $DR(X,V) - [DR(X|V=0) + DR(V|X=0)] = 0.8 - (0,3 + 0,1) = +0,4$

- $p_{11} - p_{10} - p_{01} + p_{00} = 0,9 - 0,4 - 0,2 + 0,1 = +0,4$

Soit la différence entre l'effet de X quand V = 1 et quand V = 0 (flèche verte) :

- $(p_{11} - p_{01}) - (p_{10} - p_{00}) = (0,9 - 0,2) - (0,4 - 0,1) = 0,7 - 0,3 = +0,4$

Soit la différence entre l'effet de V quand X = 1 et quand X = 0 (flèche bleue) :

- $(p_{11} - p_{10}) - (p_{01} - p_{00}) = (0,9 - 0,4) - (0,2 - 0,1) = 0,5 - 0,1 = +0,4$

On peut l'interpréter ainsi : la probabilité d'avoir une maladie chronique quand on fume augmente de +30% quand on n'a pas vécu d'évémement traumatique (40% contre 10%), et de +70% quand on a vécu un événement traumatique (de 20 à 90%). Donc l'effet du tabac est augmenté de +40% (0,70 - 0,30) quand on a vécu un événement traumatique par rapport à l'effet du tabac quand on n'a pas vécu d'événement traumatique.

Echelle multiplicative

En cas d'outcome binaire, c'est souvent le RR ou l'OR qui est utilisé pour mesurer les effets. La mesure de l'interaction sur une échelle multiplicative serait donc VanderWeele and Knol [2014] :

- $MI = \frac{RR_{11}}{RR_{10} \times RR_{01}}$
- soit $MI = \frac{p_{11}/p_{00}}{(p_{10}/p_{00}) \times (p_{01}/p_{00})}$
- soit $MI = \frac{p_{11} \times p_{00}}{p_{10} \times p_{01}}$

Exemple

Mesure de l'nteraction dans l'exemple 1

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	x2
X = 1	P10 = +0,4	P11 = +0,9	X2,25
Effet X	x4	X4,5	

X1,1

X1,1

X1,1

On retrouve l'effet d'interaction, calculé/exprimé de différentes façon,

Soit le rapport entre l'effet joint et le produit des effets individuels (flèche rouge) :

- $\frac{RR(X,V)}{RR(X|V=0)*RR(V|X=0)} = \frac{9}{4\times 2} = \times 1,1$
- $\frac{p_{11}/p_{00}}{(p_{10}+p_{01})/p_{00}} = \frac{0,9/0,1}{(0,4\times 0,2)/0,1} = \times 1,1$

Soit le produit de l'effet de X quand V = 1 et quand V = 0 (flèche verte) :

- $\frac{p_{11}/p_{01}}{p_{10}/p_{00}} = \frac{0,9/0,2}{0,4/0,1} = \frac{\times 4,5}{\times 4} = \times 1,1$

Soit le produit de l'effet de V quand X = 1 et quand X = 0 (flèche bleue):

- ou $\frac{p_{11}/p_{10}}{p_{01}/p_{00}} = \frac{0,9/0,4}{0,2/0,1} = \frac{\times 2,25}{\times 2} = \times 1,1$

On peut l'interpréter ainsi : la probabilité d'avoir une maladie chronique quand on fume est multiplier par 4 quand on n'a pas vécu d'évémement traumatique (40% contre 10%), et par 4,5 quand on a vécu un événement traumatique (90% contre 20%). Donc l'effet du tabac est multiplié par 1,1 ($\frac{4,5}{4}$) quand on a vécu un événement traumatique par rapport à l'effet du tabac quand on n'a pas vécu d'événement traumatique.

5.2 Lien entre les deux échelles

Un apparent paradoxe

Mesurer l'interaction sur une seule échelle peut être trompeur Mathur and VanderWeele [2018]. On peut régulièrement observer une interaction positive dans une échelle (par exemple $p11 - p10 - p01 + p00 > 0$) et négative dans l'autre (par exemple $(p11 \times p00)/(p10 \times p01) < 1$).

Exemple

Dans cet exemple (on modifie seulement la probabilité p_{11}, en jaune dans le tableau), on observe une interaction additive positive (l'effet de X augmente de +20% quand $V = 1$ par rapport à $V = 0$) mais une interaction multiplicative négative (l'effet de X est multiplié par 0,9 - donc diminue - quand $V = 1$ par rapport à $V = 0$).

Additif

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	+0,1
X = 1	P10 = +0,4	P11 = +0,7	+0,3
Effet X	+0,3	+0,5	

+0,2

Multiplicatif

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	x2
X = 1	P10 = +0,4	P11 = +0,7	x1,75
Effet X	x4	x3,5	

x0,9

Remarque : on retrouverait les mêmes résultats en comparant les effets de V dans les strates de X ou les effets conjoints et somme/produit des effets individuels.

Il a même été démontré que si on n'observe pas d'interaction sur une échelle, alors on en observera obligatoirement sur l'autre échelle… VanderWeele and Knol [2014].

Exemple

Dans cet exemple, il n'y a pas d'interaction multiplicative (effet de X identique quelque soit V), mais sur l'echelle additive, on observe une interaction positive.

Additif

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	+0,1
X = 1	P10 = +0,4	P11 = +0,8	+0,4
Effet X	+0,3	+0,6	

+0,3

Multiplicatif

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	x2
X = 1	P10 = +0,4	P11 = +0,8	x2
Effet X	x4	x4	

x1,0

Dans cet autre exemple, il n'y a pas d'interaction additive (effet de X identique quelque soit V), mais sur l'echelle multiplicative, on observe une interaction négative.

Additif

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	+0,1
X = 1	P10 = +0,4	P11 = +0,5	+0,1
Effet X	+0,3	+0,3	

+0,0

Multiplicatif

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	x2
X = 1	P10 = +0,4	P11 = +0,5	X1,25
Effet X	x4	X2,5	

x0,6

Le continuum

Dans un article de 2019 VanderWeele [2019], Vanderweele décrit le continuum existant entre les 2 échelles.

Par exemple, dans l'exemple 1, l'interaction additive et multiplicative sont positives. Mais si l'on fait varier la probabilité p_{11} en la diminuant, l'interaction multiplicative devient négative alors que l'interaction additive reste positive. Puis, lorsque la probabilité diminue encore, l'interaction devient négative sur les deux échelles :

Additif

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	+0,1
X = 1	P10 = +0,4	P11 = +Δ	DR(V|X=1)
Effet X	+0,3	DR(X|V=1)	

Multiplicatif

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	X2,0
X = 1	P10 = +0,4	P11 = +Δ	RR(V|X=1)
Effet X	X4,0	RR(V|X=1)	

	P11	RR(X|V=1) DR(X|V=1)	RR(V|X=1) DR(V|X=1)	MI	AI	
1	0.9	x4.5 +0.7	x2.3 +0.5	1.1	+0.4	M+ positive-multiplicative A+ positive-additive
2	0.8	x4.0 +0.6	x2.0 +0.4	1.0	+0.3	M_0 no-multiplicative A+ positive-additive
3	0.7	x3.5 +0.5	x1.75 +0.3	0.9	+0.2	M- negative-multiplicative A+ positive-additive
4	0.5	x2.5 +0.3	x1.25 +0.1	0.6	+0.0	M- negative-multiplicative A_0 zero-additive
5	0.45	x2.3 +0.25	x1.1 +0.05	0.56	-0.05	M- negative-multiplicative A- negative-additive

Interactions pures et qualitatives, interactions inversées

Dans ce continuum, si l'on continue à faire varier p_{11}, des cas particuliers d'interaction peuvent être retrouvés :

	P11	RR(X|V=1) DR(X|V=1)	RR(V|X=1) DR(V|X=1)	MI	AI	
6	0.4	x2.0 +0.2	x1.0 0.0	0.5	-0.1	M- **single** A- **pure interaction**
7	0.3	x1.5 +0.1	x0.75 -0.1	0.4	-0.2	M- **single** A- **qualitative interaction**
8	0.2	x1.0 0.0	x0.5 -0.2	0.3	-0.3	M- **single-qualitative** A- **single-pure interaction**
9	0.15	x0.8 -0.05	x0.4 -0.25	0.2	-0.35	M- **double** A- **qualitative interaction**
10	0.1	x0.5 -0.1	x0.3 -0.3	0.1	-0.4	M- A- **perfect antagonism**
11	0.05	x0.3 -0.15	x0.1 -0.35	0.06	-0.45	M- A- **inverted interaction**

- **Interaction pure** de X en fonction de V, si X n'a un effet que dans une seule strate de V. Par exemple, $p_{10} = p_{00}$ et $p_{11} \neq p_{01}$.

  Par exemple (ligne 6) ici, V a un effet (sur les deux échelles) si $X = 0$ mais pas si $X = 1$:

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	+0,1
X = 1	P10 = +0,4	P11 = +0,4	+0,0
Effet X	+0,3	+0,2	

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	x2
X = 1	P10 = +0,4	P11 = +0,4	x1
Effet X	x4	x2	

- **Interaction qualitative** de X en fonction de V, si l'effet de X dans une strate de V va dans la direction opposée de l'autre strate de V.

  Par exemple (ligne 7), V a un effet positif si $X = 0$ mais négatif si $X = 1$:

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	+0,1
X = 1	P10 = +0,4	P11 = +0,3	-0,1
Effet X	+0,3	+0,1	

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	x2
X = 1	P10 = +0,4	P11 = +0,3	x0,75
Effet X	x4	x1,5	

- **Antagonisme parfait** : l'effet joint est nul $p_{11} - p_{00} = 0$, alors que les effets individuels sont positifs.

  Par exemple (ligne 10), $p_{11} - p_{00} = 0$ alors que $p_{01} - p_{00} > 0$ et $p_{10} - p_{00} > 0$

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	+0,1
X = 1	P10 = +0,4 +0,0	P11 = +0,1	-0,3
Effet X	+0,3	-0,1	

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	x2
X = 1	P10 = +0,4 x1,0	P11 = +0,1	x0,25
Effet X	x4	x0,5	

- **Interaction inversée** (ligne 11): l'effet joint est négatif, alors que les effets individuels sont positifs.

  Par exemple (ligne 10), $p_{11} - p_{00} < 0$ alors que $p_{01} - p_{00} > 0$ et $p_{10} - p_{00} > 0$

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	+0,1
X = 1	P10 = +0,4 -0,05	P11 = +0,05	-0,35
Effet X	+0,3	-0,15	

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	x2
X = 1	P10 = +0,4 x0,5	P11 = +0,05	x0,125
Effet X	x4	x0,25	

5.3 Synthèse

Quelle échelle choisir pour mesurer un effet d'interaction ?

Même si en pratique l'échelle multiplicative est plus utilisée, car les outcomes sont souvent binaires en épidémiologie et donc les modèles logistiques sont souvent utilisés Knol and VanderWeele [2012], il semble y avoir un consensus pour privilégier plutôt l'échelle additive, plus appropriée pour évaluer l'utilité en santé publique VanderWeele and Knol [2014] Knol and VanderWeele [2012].

Si on reprend l'exemple ci dessous :

Additif

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	+0,1
X = 1	P10 = +0,4	P11 = +0,7	+0,3
Effet X	+0,3	+0,5	

+0,2

Multiplicatif

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	x2
X = 1	P10 = +0,4	P11 = +0,7	x1,75
Effet X	x4	x3,5	

x0,9

X représente un traitement dont on ne dispose que de 100 doses et Y un outcome de santé favorable (guérison). Il faut choisir si on donne 100 doses au groupe $V = 0$ ou au groupe $V = 1$.

Si on donne 100 doses :

- au groupe $V = 0$, 40 personnes seront guéries, soit 30 personnes de plus que l'évolution naturelle (40 - 10)
- au groupe $V = 1$, 70 personnes seront guéries, soit 50 personnes de plus que l'évolution naturelle (70 - 20).

Il semble donc préférable d'allouer les doses au groupe $V = 1$, car on guéri 20 personnes de plus (50 - 30).

N guéries, pour 100 personnes

Groupes	V = 0	V = 1
Evolution naturelle	10	20
Traitement	40	70
Différence	+30	+50

Pourtant si on avait réfléchi à partir de l'échelle multiplicative, on aurait choisi le groupe $V = 0$ car :

- l'effet du traitement est de RR=4 dans le groupe $V = 0$ ($\frac{40}{10} =$ 4x plus de personnes guéries par rapport à l'évolution naturelle)
- et de RR=3,5 dans le groupe $V = 1$ ($\frac{70}{20}$) = 3.5x plus de personnes guéries par rapport à l'évolution naturelle.

On peut donc conclure à un effet multiplicatif plus fort d'un traitement dans un groupe alors qu'en terme d'utilité (nombre de personnes favorablement impactées), l'échelle additive nous conduirait à choisir l'autre groupe…

Idéalement, les interactions devraient cependant être reportées sur les 2 échelles Knol and VanderWeele [2012] VanderWeele and Knol [2014].

Chapter 6

Types de paramètres

Plusieurs paramètres peuvent être utilisés pour décrire une interaction, sur l'échelle additive ou multiplicative.

6.1 Sur l'échelle multiplicative

Avec les risques relatifs (MI)

On a déjà défini précédemment un paramètre d'interaction sur l'échelle multiplicative (MI),défini à partir des risques relatifs VanderWeele and Knol [2014] :

- $MI = \frac{RR_{11}}{RR_{10} \times RR_{01}}$
- soit $MI = \frac{p_{11}/p_{00}}{(p_{10}/p_{00}) \times (p_{01}/p_{00})}$
- soit $MI = \frac{p_{11} \times p_{00}}{p_{10} \times p_{01}}$

Avec les Odds Ratio (MI)

Souvent en épidémiologie, lorsque l'outcome Y est binaire, les effets sont mesurés par des odds ratios estimés à partir de modèles de régression logistique.

Un paramètre d'interaction sur l'échelle multiplicative (MI_{OR}) peut être estimé à partir de ces OR VanderWeele and Knol [2014] :

- $MI_{OR} = \frac{OR_{11}}{OR_{10} \times OR_{01}}$

En général, la mesure MI_{OR} et MI_{RR} seront proches si l'outcome est rare VanderWeele and Knol [2014].

6.2 Sur l'échelle additive

Avec les différences de risques (AI)

On a déjà défini un paramètre d'interaction sur l'échelle additive (AI) à partir des différences d'effets VanderWeele and Knol [2014] :

- $AI = DR(X, V) - [DR(X|V = 0) + DR(V|X = 0)]$
- $AI = (p_{11} - p_{00}) - [(p_{10} - p_{00}) + (p_{01} - p_{00})]$
- soit $AI = p_{11} - p_{10} - p_{01} + p_{00}$

Excès de risque, à partir des RR (RERI)

Lorsque seulement les risques relatifs sont donnés mais que l'on souhaite évaluer l'interaction sur l'échelle additive, "l'excès de risque du à l'interaction" (RERI) ou "interaction contrast ratio" (ICR), peut être estimé à partir des risques relatifs VanderWeele and Knol [2014] :

- $RERI = \frac{AI}{p_{00}} = \frac{p_{11} - p_{10} - p_{01} + p_{00}}{p_{00}}$
- $RERI = RR_{11} - RR_{10} - RR_{01} + 1$

On voit que le RERI correspond à l'interaction mesurée sur l'échelle additive, rapportée au risque de base p_{00}.

Il faut noter que, bien que le RERI donne la direction (positive, négative ou nulle) de l'interaction additive, nous ne pouvons pas utiliser le RERI pour évaluer l'ampleur de l'interaction additive, à moins de connaître au moins p_{00}, dans ce cas on retrouve $AI = p_{00} \times RERI$.

Si l'on a seulement l'OR et que l'outcome est rare, les OR peuvent approximer les RR, on a donc :

- $RERI_{OR} = OR_{11} - OR_{10} - OR_{01} + 1 \approx RERI_{RR}$

Le "Synergie index" (SI)

Il s'agit d'un paramètre explorant aussi l'interaction additive VanderWeele and Knol [2014]. Il est définit à partir des Augmentation Relatif du Risque (ARR).

Pour rappel, **l'Augmentation Relative du Risque** liée à l'exposition jointe correspond à l'augmentation absolue du risque (différence de risques), exprimée en pourcentage par rapport au risque de base p_{00}.

- $ARR(X, V) = \frac{DR(X,V)}{p_{00}} = \frac{p_{11} - p_{00}}{p_{00}} = RR_{11} - 1$

L'augmentation relative du risque liée à l'exposition X ou V, exprimées en pourcentage par rapport au risque de base p_{00} sont respectivement :

- $ARR(X|V=0) = \frac{p_{10}-p_{00}}{p_{00}} = RR_{10} - 1$
- et $ARR(V|X=0) = \frac{p_{01}-p_{00}}{p_{00}} = RR_{01} - 1$

L'index synergique correspond à l'augmentation relative du risque liée à l'exposition jointe, rapportée à la somme des augmentations relatives du risque liées à la 1ère et la 2ème exposition.

- $SI = \frac{RR_{11}-1}{(RR_{10}-1)+(RR_{01}-1)}$.

On peut aussi l'interpréter ainsi : la différence liée à l'effet joint $DR(X,V)$ est égale à SI fois la somme des différences liées aux effets individuels $DR(X|V = 0) + DR(V|X = 0)$, car :

- $SI = \frac{p_{11}-p_{00}}{(p_{10}-p_{00})+(p_{01}-p_{00})}$

Si le dénominateur est positif:

- si $SI > 1$, alors $AI > 0$ et $RERI_{RR} > 0$
- si $SI < 1$, alors $AI < 0$ et $RERI_{RR} < 0$

L'interprétation de l'indice de synergie devient difficile dans les cas où l'effet de l'une des expositions a un effet négatif et que le dénominateur de S est inférieur à 1.

Proportion attribuable (AP)

Il s'agit aussi d'un paramètre explorant l'interaction additive :

- $AP = \frac{RR_{11}-RR_{10}-RR_{01}+1}{RR_{11}}$.

Ce paramètre mesure la proportion du risque dans le groupe doublement exposé qui est due à l'interaction.

L'AP est en lien avec le $RERI_{RR}$:

- AP > 0 si et seulement si $RERI_{RR} > 0$
- AP < 0 si et seulement si $RERI_{RR} < 0$.

En fait $AP = \frac{RERI_{RR}}{RR_{11}-1}$.

Part II

Estimations, Interprétations, Présentations

Chapter 7

Présentation des résultats

7.1 Recommandations

Knol et VanderWeele ont émis des recommandations concernant la présentation des résultats d'une analyse d'interaction Knol and VanderWeele [2012]. Ces recommandations sont :

Pour une analyse d'une modification d'effet de X sur Y par V

- Présenter les effectifs dans chaque catégorie
  - avec et sans l'outcome ($N_{x,v}(Y = 1)$ et $N_{x,v}(Y = 0)$)
- Présenter les risques relatifs (RR), les OR ou les différences de risques (RD)
  - avec les intervalles de confiance (IC)
  - pour chaque strate de X et de V avec une seule catégorie de référence
  - (éventuellement prise comme la strate $X \cap V$ présentant le plus faible risque de Y).
- Présenter les RR, OR ou RD avec les IC
  - de l'effet de X sur Y dans les strates de V
- Présenter les mesures de la modification de l'effet avec les IC, sur des échelles
  - additives (par exemple, RERI)
  - et multiplicatives.

- Énumérer les facteurs de confusion pour lesquels la relation entre X et Y a été ajustée.

Exemple de présentation avec les données fictives de l'exemple 1, modification de l'effet de X par V :

	X = 0		X = 1		OR dans les strates de V
	n+/n-	**OR [IC95%]**	**n+/n-**	**OR [IC95%]**	
V = 0	10 / 90	*ref*	40 / 60	5.8 [… à …]	5.8 [… à …]
V = 1	20 / 80	2.1 [… à …]	90 /10	80.6 [… à …]	34.5 [… à …]
Modification d'effet additive (RERI) = 68.5 [… à …]					
Modification d'effet multiplicative (MI_{OR}) = 6.0 [… à …]					
L'OR est ajusté sur L1 et L2					

Interaction $X * V$ sur Y

- Présenter les effectifs dans chaque catégorie
  - avec et sans l'outcome ($N_{x,v}(Y = 1)$ et $N_{x,v}(Y = 0)$)
- Présenter les risques relatifs (RR), les OR ou les différences de risques (RD)
  - avec les intervalles de confiance (IC)
  - pour chaque strate de X et de V avec une seule catégorie de référence
  - (éventuellement prise comme la strate $X \cap V$ présentant le plus faible risque de Y).
- Présenter les RR, OR ou RD avec les IC
  - de l'effet de X sur Y dans les strates de V
  - **et de V sur Y dans les strates de X.**
- Présenter les mesures de la modification de l'effet d'interaction avec les IC sur des échelles
  - additives (par exemple, RERI)
  - et multiplicatives.
- Énumérer les facteurs de confusion pour lesquels la relation entre X et Y **et la relation entre V et Y** ont été ajustées.

Exemple de présentation avec les données fictives de l'exemple 1, interaction entre X et V :

	X = 0		X = 1		OR dans les strates de V
	n+/n-	OR [IC95%]	n+/n-	OR [IC95%]	
V = 0	10 / 90	*ref*	40 / 60	5.8 [… à …]	5.8 [… à …]
V = 1	20 / 80	2.1 [… à …]	90 /10	80.6 [… à …]	34.5 [… à …]
OR dans les strates de X		2.1 [… à …]		12.7 [… à …]	

Interaction additive (RERI) = 68.5 [… à …]
Interaction multiplicative (MI_{OR}) = 6.0 [… à …]
L'OR est ajusté sur L1, L2 et L3

7.2 Proposition

Ces recommandations sont très utiles lorsque les interactions ont été évaluées à partir de modèles de régression (logistiques, log-linéaires ou linéaires) permettant d'estimer directement des OR, des RR ou des DR, condionnellement aux facteurs de confusion.

En inférence causale, des assocations marginales plutôt que conditionnelles sont souvent estimées (que ce soit en termes de difference de risques, de risques relatifs ou d'odds ratio). Dans la suite de ce document, nous proposons une variante des recommandations de Knol et VanderWeele, adaptée à des estimations marginales. Nous proposons en effet :

- De présenter les effets marginaux ou proportions prédites de Y dans chaque strate $X \cap V$,
  - plutôt les effectifs avec et sans l'outcome
- Ne pas forcément présenter une différence de risques ou un rapport de risques
  - pour chaque strate de X et de V avec une seule catégorie de référence
- Mais présenter les effets
  - de X dans chaque strate de V
  - et de V dans chaque strate de X (si analyse d'interaction)
  - sur une échelle multiplicative **et** additive.

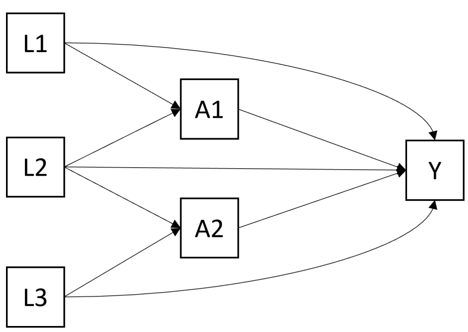
Exemple de présentation avec les données fictives de l'exemple 1, interaction entre X et V :

	X = 0	X = 1	OR, strates de V	DR, strates de V
V = 0	p00 = 0.10	p10 = 0.40	5.8 [… à …]	+0.30 [… à …]
V = 1	p01 = 0.19	p11 = 0.89	34.5 [… à …]	+0.69 [… à …]
OR, strates de X	2.1 [… à …]	12.7 [… à …]		
DR, strates de X	+0.09 [… à …]	+0.49 [… à …]		
Interaction multiplicative (MI_{OR}) = 6.0 [… à …]				
Interaction additive (RERI) = 68.5 [… à …]				
Interaction additive (AI) = +0.40 [… à …]				
Les modèles sont ajustés sur L1, L2 et L3				

Chapter 8

Simulations

Pour la description des différents types d'estimation, on a simulé des données selon le DAG suivant (toutes les variables sont binaires). Les deux expositions d'intérêt sont A_1 et A_2, l'outcome est Y, et L_1, L_2 et L_3 sont 3 facteurs de confusion :



Les équations structurelles associées au DAG sont décrites ci-dessous, les paramètres correspondent aux paramètres renseignés dans le code de simulation.

$$
\begin{aligned}
P(L1 = 1) =& p_{L_1} \\
P(L2 = 1) =& p_{L_2} \\
P(L3 = 1) =& p_{L_3} \\
P(A1 = 1 \mid L1, L2) =& \beta_{A_1} + \beta_{L_1,A_1} L1 + \beta_{L_2,A_1} L2 \\
P(A2 = 1 \mid L1, L3) =& \beta_{A_2} + \beta_{L_1,A_2} L1 + \beta_{L_3,A_2} L3 \\
P(Y = 1 \mid L1, L2, L3, A1, A2) =& \beta_Y + \beta_{L_1,Y} L1 + \beta_{L_2,Y} L2 + \beta_{L_3,Y} L3 \\
& + \beta_{A_1,Y} A1 + \beta_{A_2,Y} A2 + \beta_{A_1 * A_2,Y} (A1 * A2)
\end{aligned}
$$

Le code ayant permis de simuler les données est le suivant :

```
  rm(list=ls())

  param.causal.model <- function(p_L1 = 0.50, # baseline confounders
                                 p_L2 = 0.20, # baseline confounders
                                 p_L3 = 0.70, # baseline confounders
                                 b_A1 = 0.10,    # modèle de A1
                                 b_L1_A1 = 0.15, # modèle de A1
                                 b_L2_A1 = 0.25, # modèle de A1
                                 b_A2 = 0.15,    # modèle de A2
                                 b_L1_A2 = 0.20, # modèle de A2
                                 b_L3_A2 = 0.20, # modèle de A2
                                 b_Y = 0.10,      # modèle de Y
                                 b_L1_Y = 0.02,   # modèle de Y
                                 b_L2_Y = 0.02,   # modèle de Y
                                 b_L3_Y = -0.02,  # modèle de Y
                                 b_A1_Y = 0.3,    # modèle de Y
                                 b_A2_Y = 0.1,    # modèle de Y
                                 b_A1A2_Y = 0.4 ) { # <- effet d'interaction Delta)

  # coefficients pour simuler l'exposition
  # exposition A1  # vérif
  try(if(b_A1 + b_L1_A1 + b_L1_A1 > 1)
    stop("la somme des coefficient du modèle A1 dépasse 100%"))

  # exposition A2  # vérif
  try(if(b_A2 + b_L1_A2 + b_L3_A2 > 1)
    stop("la somme des coefficients du modèle A2 dépasse 100%"))

  # coefficients pour simuler l'outcome, vérif
  try(if(b_Y + b_L1_Y + b_L2_Y + b_L3_Y + b_A1_Y + b_A2_Y + b_A1A2_Y > 1)
    stop("la somme des coefficients du modèle Y dépasse 100%"))
  try(if(b_Y + b_L1_Y + b_L2_Y + b_L3_Y + b_A1_Y + b_A2_Y + b_A1A2_Y < 0)
    stop("la somme des coefficients du modèle Y est inférieure à 0%"))

  coef <- list(c(p_L1 = p_L1, p_L2 = p_L2, p_L3 = p_L3),
               c(b_A1 = b_A1, b_L1_A1 = b_L1_A1, b_L2_A1 = b_L2_A1),
               c(b_A2 = b_A2, b_L1_A2 = b_L1_A2, b_L3_A2 = b_L3_A2),
               c(b_Y = b_Y, b_L1_Y = b_L1_Y, b_L2_Y = b_L2_Y, b_L3_Y = b_L3_Y,
                 b_A1_Y = b_A1_Y, b_A2_Y = b_A2_Y, b_A1A2_Y = b_A1A2_Y))
  return(coef)
}

generate.data <- function(N, b =  param.causal.model()) {
```

A2	label	levels	value
0	A1	0	0.10 (0.30)
0		1	0.41 (0.49)
1	A1	0	0.20 (0.40)
1		1	0.90 (0.30)

```r
  L1 <- rbinom(N, size = 1, prob = b[[1]]["p_L1"])
  L2 <- rbinom(N, size = 1, prob = b[[1]]["p_L2"])
  L3 <- rbinom(N, size = 1, prob = b[[1]]["p_L3"])
  A1 <- rbinom(N, size = 1, prob = b[[2]]["b_A1"] +
               (b[[2]]["b_L1_A1"] * L1) + (b[[2]]["b_L2_A1"] * L2))
  A2 <- rbinom(N, size = 1, prob = b[[3]]["b_A2"] +
               (b[[3]]["b_L1_A2"] * L1) + (b[[3]]["b_L3_A2"] * L3))
  Y <- rbinom(N, size = 1, prob = (b[[4]]["b_Y"] +
                                   (b[[4]]["b_L1_Y"] * L1) +
                                   (b[[4]]["b_L2_Y"] * L2) +
                                   (b[[4]]["b_L3_Y"] * L3) +
                                   (b[[4]]["b_A1_Y"] * A1) +
                                   (b[[4]]["b_A2_Y"] * A2) +
                                   (b[[4]]["b_A1A2_Y"] * A1 * A2)) )
  data.sim <- data.frame(L1, L2, L3, A1, A2, Y)
  return(data.sim)
}

#### On simule une base de données
  set.seed(12345)
  # b =  param.causal.model(b_A1A2_Y = -0.45)
  b = param.causal.model()
  df <- generate.data(N = 10000, b = b)
  summary(df)
  prop.table(table(df$Y, df$A1, df$A2, deparse.level = 2))
```

Au final, les probabilités de l'outcome P(Y=1), dans chaque catégorie sont :

Les paramètres utilisés pour simuler les données ont été choisis de sorte que les "vraies" valeurs des paramètres de la distribution correspondent au tableau présenté au paragraphe 5 "Mesure des interactions".

Chapter 9

A partir de modèles de régression

Dans une première étape exploratoire, on peut simplement utiliser les modèles de régression habituels : les modèles de régression logistique et linéaire.

9.1 Régression logistique

Lorsque l'on étudie un outcome binaire, on utilise souvent les modèles de régression logistique.

```
##
## Call:
## glm(formula = Y ~ as.factor(A1) + as.factor(A2) + as.factor(A1) *
##     as.factor(A2) + as.factor(L1) + as.factor(L2) + as.factor(L3),
##     family = binomial, data = df_f)
##
## Coefficients:
##                             Estimate Std. Error z value Pr(>|z|)
## (Intercept)                 -2.16540    0.06708 -32.281  < 2e-16 ***
## as.factor(A1)1               1.75607    0.07604  23.093  < 2e-16 ***
## as.factor(A2)1               0.75332    0.06831  11.028  < 2e-16 ***
## as.factor(L1)1               0.15753    0.05702   2.763  0.00573 **
## as.factor(L2)1               0.14128    0.06878   2.054  0.03996 *
## as.factor(L3)1              -0.14926    0.06141  -2.431  0.01507 *
## as.factor(A1)1:as.factor(A2)1  1.78587    0.14131  12.638  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 11037.7  on 9999  degrees of freedom
## Residual deviance:  8460.4  on 9993  degrees of freedom
## AIC: 8474.4
##
## Number of Fisher Scoring iterations: 4
```

A partir de cette sortie, on peut extraire :

- **A1|A2=0**
  - à partir du coefficient `as.factor(A1)1`
  - qui correspond à l'effet de A1 dans la catégorie de référence de A2,
  - soit $OR_{A1|A2=0} = exp(1.756) = 5.789$.
- **A1|A2=1**
  - à partir du coefficient `as.factor(A1)1:as.factor(A2)1`,
  - qui correspond à la différence d'effet de A1 quand on passe dans l'autre catégorie de A2.
  - L'effet de A1 dans la catégorie A2=1 est donc
  - $OR_{A1|A2=1} = exp(1.756 + 1.786) = 34.536$.
- **L'interaction multiplicative (IM)**
  - peut être estimée à partir du coefficient `as.factor(A1)1:as.factor(A2)1`
  - par $IM = exp(1.786) = 5.966$,
  - qu'on peut retrouver en faisant $OR_{A1|A2=1}/OR_{A1|A2=0}$.
  - Ici l'interaction est significative (p-value >0.05).
- **A2|A1=0** et **A2|A1=1**
  - On aurait aussi pu décrire l'interaction à partir de l'effet d'A2 dans chaque strate de A1
  - à partir de `as.factor(A2)1` et `as.factor(A1)1:as.factor(A2)1`,
  - avec : $OR_{A2|A1=0} = exp(0.753) = 2.123$
  - et $OR_{A2|A1=1} = exp(0.753 + 1.786) = 12.667$
- **L'interaction additive**
  - On peut explorer l'interaction sur l'échelle additive en estimant le RERI par
  - $RERI \approx OR_{11} - OR_{10} - OR_{01} + 1 =$
  - $OR_{A1,A2} - OR_{A1|A2=0} - OR_{A2|A1=0} + 1 =$
  - $exp(1.786 + 0.753 + 1.786) - exp(1.786) - exp(0.753) + 1 = 68.477$.

En résumé, (le package `finalfit` permet de sortir quelques résultats proprement) :

names	OR
A1|A2=0	5.79 (4.99-6.72, p<0.001)
A2|A1=0	2.12 (1.86-2.43, p<0.001)
Interaction	5.96 (4.54-7.90, p<0.001)

```
explanatory = c("as.factor(A1)",
                "as.factor(A2)",
                "as.factor(A1)*as.factor(A2)",
                "as.factor(L1)",
                "as.factor(L2)",
                "as.factor(L3)")
dependent = "Y"

df_f %>%
  finalfit(dependent, explanatory)-> t

# le tableau t entier peut être imprimé, mais ici je sélectionne seulement les effets d'intéret
# pour éviter la table 2 fallacy (les coefficient des facteurs de confusion L ne sont pas interpr

cbind(names = c("A1|A2=0", "A2|A1=0", "Interaction"),
      OR = t[c(12,14,13),6]) %>%
  as.data.frame %>%
      kbl() %>%
      kable_classic()
```

Attention, les modèles de régressions logistiques sont ici biaisés car les données sont générées à partir de modèles additifs.

9.2 Régression lineaire

Même si l'outcome binaire, on peut en théorie utiliser un modèle de régression linéaire et explorer les effets sur une échelle additive. Si l'outcome est quantitatif, on utilise aussi, en général, les modèles de régression linéaire.

```
##
## Call:
## lm(formula = Y ~ as.factor(A1) + as.factor(A2) + as.factor(A1) *
##     as.factor(A2) + as.factor(L1) + as.factor(L2) + as.factor(L3),
##     data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -0.93110 -0.19602 -0.10494 -0.08426  0.91574
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    0.103835   0.008146  12.746  < 2e-16 ***
## as.factor(A1)1                 0.300796   0.011592  25.948  < 2e-16 ***
## as.factor(A2)1                 0.092280   0.008671  10.642  < 2e-16 ***
## as.factor(L1)1                 0.020677   0.007495   2.759  0.00581 **
## as.factor(L2)1                 0.019476   0.009410   2.070  0.03851 *
## as.factor(L3)1                -0.019574   0.008085  -2.421  0.01549 *
## as.factor(A1)1:as.factor(A2)1  0.394034   0.017854  22.070  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3615 on 9993 degrees of freedom
## Multiple R-squared:  0.2856, Adjusted R-squared:  0.2852
## F-statistic: 665.8 on 6 and 9993 DF,  p-value: < 2.2e-16
```

A partir de cette sortie, on peut extraire :

- **A1|A2=0**
  - à partir du coefficient `as.factor(A1)1`
  - qui correspond à l'effet de A1 dans la catégorie de référence de A2,
  - soit $DR = +30,08\%$.
- **A1|A2=1**
  - à partir du coefficient `as.factor(A1)1:as.factor(A2)1`,
  - qui correspond à la différence d'effet de A1 quand on passe dans l'autre catégorie de A2.
  - L'effet de A1 dans la catégorie A2=1 est donc
  - $DR = 30.08 + 39.40 = 69.48$ %.
- **L'interaction additive**
  - à partir du coefficient `as.factor(A1)1:as.factor(A2)1`
  - avec $AI = +39.40\%$,
  - qu'on peut retrouver en faisant $DR(A1|A2 = 1) - DR(A1|A2 = 0)$.
  - Ici l'interaction est significative (p-value >0.05).
- **A2|A1=0** et **A2|A1=1**
  - On aurait aussi pu décrire cette interaction à partir de l'effet d'A2 dans chaque strate de A1
  - à partir de `as.factor(A2)1` et `as.factor(A1)1:as.factor(A2)1`,
  - avec : $DR_{A1|A2=0} = +9.23\%$
  - et $DR_{A1|A2=1} = 9.23 + 39.40 = 48.63\%$.

names	DR
A1|A2=0	0.30 (0.28 to 0.32, p<0.001)
A2|A1=0	0.09 (0.08 to 0.11, p<0.001)
Interaction	0.39 (0.36 to 0.43, p<0.001)

En résumé, (le package `finalfit` permet de sortir quelques résultats proprement) :

```
explanatory = c("as.factor(A1)",
                "as.factor(A2)",
                "as.factor(A1)*as.factor(A2)",
                "as.factor(L1)",
                "as.factor(L2)",
                "as.factor(L3)")
dependent = "Y"
df %>%
  finalfit(dependent, explanatory)-> t

cbind(names = c("A1|A2=0", "A2|A1=0", "Interaction"), DR = t[c(12,14,13),6]) %>%
  as.data.frame %>%
      kbl() %>%
      kable_classic()
```

Chapter 10

Approches causales

10.1 Estimation par G-computation

Il s'agit d'une "G-méthode" qui peut être décrite comme une "standardisation" par régression (Hernàn Hernán and Robins [2020]). Le principe est le suivant :

i) Modéliser le critère de jugement en fonction des deux expositions d'intérêt, de leur interaction et des facteurs de confusion

$$\overline{Q}(A_1, A_2, L) = \mathbb{E}(Y|A_1, A_2, L_1, L_2, L_3)$$

ii) A partir de ce modèle $\overline{Q}$, estimer pour chaque individu i les valeurs moyennes attendues (contrefactuelles) $\overline{Q}_i(0,0,L)$, $\overline{Q}_i(1,0,L)$, $\overline{Q}_i(0,1,L)$ et $\overline{Q}_i(1,1,L)$, sous les 4 scénarios possibles $\{A_1 = 0, A_2 = 0\}$, $\{A_1 = 1, A_2 = 0\}$, $\{A_1 = 0, A_2 = 1\}$ et $\{A_1 = 1, A_2 = 1\}$.
iii) Utiliser les moyennes contrefactuelles estimées pour calculer les différents indicateurs d'intérêt pour l'analyse d'interaction :

- Les moyennes marginales pour chaque case du tableau d'interaction

$$p_{00} = \frac{1}{n}\sum \overline{Q}_i(0,0,L)$$
$$p_{10} = \frac{1}{n}\sum \overline{Q}_i(1,0,L)$$
$$p_{01} = \frac{1}{n}\sum \overline{Q}_i(0,1,L)$$
$$p_{11} = \frac{1}{n}\sum \overline{Q}_i(1,1,L)$$

- Les différences de risques marginales

$$
\begin{aligned}
DR(A_1, A_2 = 0) &= p_{10} - p_{00} \\
DR(A_1, A_2 = 1) &= p_{11} - p_{01} \\
DR(A_2, A_1 = 0) &= p_{01} - p_{00} \\
DR(A_2, A_1 = 1) &= p_{11} - p_{10}
\end{aligned}
$$

- Les risques relatifs marginaux

$$
\begin{aligned}
RR(A_1, A_2 = 0) &= p_{10}/p_{00} \\
RR(A_1, A_2 = 1) &= p_{11}/p_{01} \\
RR(A_2, A_1 = 0) &= p_{01}/p_{00} \\
RR(A_2, A_1 = 1) &= p_{11}/p_{10}
\end{aligned}
$$

- Les mesures d'interaction

$$
\begin{aligned}
MI &= \frac{p_{11} \times p_{00}}{p_{10} \times p_{01}} \\
AI &= p_{11} - p_{10} - p_{01} + p_{00} \\
RERI &= \frac{p_{11} - p_{10} - p_{01} + p_{00}}{p_{00}}
\end{aligned}
$$

```
## i) Modéliser le critère de jugement en fonction des deux expositions d'intérêt,
##    de leur interaction et des facteurs de confusion
    model.Y <- glm(Y ~ L1 + L2 + L3 + A1 + A2 + A1:A2, data = df, family = "binomial")

    # dans cette exemple, les données étaient simulées à partir d'un modèle additif,
    # le modèle plus adapté serait donc plutôt le suivant :
    # model.Y <- glm(Y ~ L1 + L2 + L3 + A1 + A2 + A1:A2, data = df,
    #                family = "gaussian")
    # (en pratique, on ne connaît pas la nature du modèle générateur des données)

## ii) Estimer les valeurs attendues sous les 4 scenarios contrefactuels
## ii.a) on crée 4 bases de données correspondant aux 4 scenarios contrefactuels
    df.A1_0.A2_0 <- df.A1_1.A2_0 <- df.A1_0.A2_1 <- df.A1_1.A2_1 <- df

    # scénario do(A1 = 0, A2 = 0) pour toute la population
    df.A1_0.A2_0$A1 <- df.A1_0.A2_0$A2 <- rep(0, nrow(df))

    # scénario do(A1 = 1, A2 = 0) pour toute la population
    df.A1_1.A2_0$A1 <- rep(1, nrow(df))
    df.A1_1.A2_0$A2 <- rep(0, nrow(df))

    # scénario do(A1 = 0, A2 = 1) pour toute la population
    df.A1_0.A2_1$A1 <- rep(0, nrow(df))
```

```r
    df.A1_0.A2_1$A2 <- rep(1, nrow(df))

    # scénario do(A1 = 1, A2 = 1) pour toute la population
    df.A1_1.A2_1$A1 <- df.A1_1.A2_1$A2 <- rep(1, nrow(df))

## ii.b) on prédit le critère de jugement sous les scenarios contrefactuels
    Qbar_00 <- predict(model.Y, newdata = df.A1_0.A2_0, type = "response")
    Qbar_10 <- predict(model.Y, newdata = df.A1_1.A2_0, type = "response")
    Qbar_01 <- predict(model.Y, newdata = df.A1_0.A2_1, type = "response")
    Qbar_11 <- predict(model.Y, newdata = df.A1_1.A2_1, type = "response")


## iii) Utiliser les moyennes contrefactuelles estimées pour calculer les différents
##      indicateurs d'intérêt pour l'analyse d'interaction
## iii.a) on va enregistrer l'ensemble des résultats pertinents dans une table 'int.r'
##        de longueur 2 x 2
    int.r <- matrix(NA,
                    ncol = 26,
                    nrow = nlevels(as.factor(df$A1)) * nlevels(as.factor(df$A2)))
    int.r <- as.data.frame(int.r)
    names(int.r) <- c("A1","A2","p","p.lo","p.up",
                      "RD.A1","RD.A1.lo","RD.A1.up","RD.A2","RD.A2.lo","RD.A2.up",
                      "RR.A1","RR.A1.lo","RR.A1.up","RR.A2","RR.A2.lo","RR.A2.up",
                      "a.INT", "a.INT.lo", "a.INT.up","RERI","RERI.lo","RERI.up",
                      "m.INT", "m.INT.lo", "m.INT.up" )
    int.r[,c("A1","A2")] <- expand.grid(c(0,1), c(0,1))

## iii.b) Les moyennes marginales pour chaque case du tableau d’interaction
# dans chaque case de la table 2 x 2
    # A1 = 0 et A2 = 0
    int.r$p[int.r$A1 == 0 & int.r$A2 == 0] <- mean(Qbar_00)
    # A1 = 1 et A2 = 0
    int.r$p[int.r$A1 == 1 & int.r$A2 == 0] <- mean(Qbar_10)
    # A1 = 0 et A2 = 1
    int.r$p[int.r$A1 == 0 & int.r$A2 == 1] <- mean(Qbar_01)
    # A1 = 1 et A2 = 1
    int.r$p[int.r$A1 == 1 & int.r$A2 == 1] <- mean(Qbar_11)

## iii.c) Les différences de risques marginales
    # RD.A1.A2is0
    int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 0] <- mean(Qbar_10) - mean(Qbar_00)
    # RD.A1.A2is1
    int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 1] <- mean(Qbar_11) - mean(Qbar_01)
    # RD.A2.A1is0
    int.r$RD.A2[int.r$A1 == 0 & int.r$A2 == 1] <- mean(Qbar_01) - mean(Qbar_00)
```

```r
    # RD.A2.A1is1
    int.r$RD.A2[int.r$A1 == 1 & int.r$A2 == 1] <- mean(Qbar_11) - mean(Qbar_10)

## iii.d) Les risques relatifs marginaux
    # RR.A1.A2is0
    int.r$RR.A1[int.r$A1 == 1 & int.r$A2 == 0] <- mean(Qbar_10) / mean(Qbar_00)
    # RR.A1.A2is1
    int.r$RR.A1[int.r$A1 == 1 & int.r$A2 == 1] <- mean(Qbar_11) / mean(Qbar_01)
    # RR.A2.A1is0
    int.r$RR.A2[int.r$A1 == 0 & int.r$A2 == 1] <- mean(Qbar_01) / mean(Qbar_00)
    # RR.A2.A1is1
    int.r$RR.A2[int.r$A1 == 1 & int.r$A2 == 1] <- mean(Qbar_11) / mean(Qbar_10)

## iii.e) Les mesures d'interaction
    # additive interaction
    int.r$a.INT[int.r$A1 == 1 & int.r$A2 == 1] <- mean(Qbar_11) -
                                                   mean(Qbar_10) -
                                                   mean(Qbar_01) +
                                                   mean(Qbar_00)
    # RERI
    int.r$RERI[int.r$A1 == 1 & int.r$A2 == 1] <- (mean(Qbar_11) -
                                                    mean(Qbar_10) -
                                                    mean(Qbar_01) +
                                                    mean(Qbar_00)) /
                                                    mean(Qbar_00)
    # multiplicative interaction
    int.r$m.INT[int.r$A1 == 1 & int.r$A2 == 1] <- (mean(Qbar_11) *
                                                    mean(Qbar_00)) /
                                                   (mean(Qbar_10) *
                                                    mean(Qbar_01))


## iii.f) Calcul des intervalles de confiance
    # Le script ci-dessous présente une méthode de bootstrap
    # pour calculer les intervalles de confiance à 95% des différents indicateurs
    set.seed(5678)
    B <- 1000 # on utilise 1000 échantillons boostrap dans cet exemple
    bootstrap.est <- data.frame(matrix(NA, nrow = B, ncol = 15))
    colnames(bootstrap.est) <- c("p.A1is0.A2is0", "p.A1is1.A2is0",
                                 "p.A1is0.A2is1", "p.A1is1.A2is1",
                                 "RD.A1.A2is0", "RD.A1.A2is1",
                                 "RD.A2.A1is0", "RD.A2.A1is1",
                                 "lnRR.A1.A2is0", "lnRR.A1.A2is1",
                                 "lnRR.A2.A1is0", "lnRR.A2.A1is1",
                                 "INT.a", "lnRERI", "lnINT.m")
```

```r
for (b in 1:B){ # on refait toute l'analyse dans chaque échantillon bootstrap
  # constituer un échantillon bootstrap (échantillonnage avec remise)
  bootIndices <- sample(1:nrow(df), replace=T)
  bootData <- df[bootIndices,]

  if ( round(b/100, 0) == b/100 ) print(paste0("bootstrap number ",b))

  # étape (i) de modélisation du critère de jugement
  model.Y <- glm(Y ~ L1 + L2 + L3 + A1 + A2 + A1:A2,
                 data = bootData,                    # use BootData here +++
                 family = "binomial")

  # étape (ii) d'estimation des valeurs attendus sous les 4 scénarios contrefactuels
  boot.A1_0.A2_0 <- boot.A1_1.A2_0 <- boot.A1_0.A2_1 <- boot.A1_1.A2_1 <- bootData
  boot.A1_0.A2_0$A1 <- boot.A1_0.A2_0$A2 <- rep(0, nrow(df))
  boot.A1_1.A2_0$A1 <- rep(1, nrow(df))
  boot.A1_1.A2_0$A2 <- rep(0, nrow(df))
  boot.A1_0.A2_1$A1 <- rep(0, nrow(df))
  boot.A1_0.A2_1$A2 <- rep(1, nrow(df))
  boot.A1_1.A2_1$A1 <- boot.A1_1.A2_1$A2 <- rep(1, nrow(df))

  # prédire les résultats contrefactuels sous les différents scénarios
  Qbar_00 <- predict(model.Y, newdata = boot.A1_0.A2_0, type = "response")
  Qbar_10 <- predict(model.Y, newdata = boot.A1_1.A2_0, type = "response")
  Qbar_01 <- predict(model.Y, newdata = boot.A1_0.A2_1, type = "response")
  Qbar_11 <- predict(model.Y, newdata = boot.A1_1.A2_1, type = "response")

  # étape (iii) de calcul des différents indicateurs d'intérêt
  # on sauve les résultats dans la table bootstrap.est
  bootstrap.est[b,"p.A1is0.A2is0"] <- mean(Qbar_00)
  bootstrap.est[b,"p.A1is1.A2is0"] <- mean(Qbar_10)
  bootstrap.est[b,"p.A1is0.A2is1"] <- mean(Qbar_01)
  bootstrap.est[b,"p.A1is1.A2is1"] <- mean(Qbar_11)

  bootstrap.est[b,"RD.A1.A2is0"] <- mean(Qbar_10) - mean(Qbar_00)
  bootstrap.est[b,"RD.A1.A2is1"] <- mean(Qbar_11) - mean(Qbar_01)
  bootstrap.est[b,"RD.A2.A1is0"] <- mean(Qbar_01) - mean(Qbar_00)
  bootstrap.est[b,"RD.A2.A1is1"] <- mean(Qbar_11) - mean(Qbar_10)

  # les Ic95% des RR sont d'abord calculés sur une échelle logarithmique
  bootstrap.est[b,"lnRR.A1.A2is0"] <- log(mean(Qbar_10) / mean(Qbar_00))
  bootstrap.est[b,"lnRR.A1.A2is1"] <- log(mean(Qbar_11) / mean(Qbar_01))
  bootstrap.est[b,"lnRR.A2.A1is0"] <- log(mean(Qbar_01) / mean(Qbar_00))
  bootstrap.est[b,"lnRR.A2.A1is1"] <- log(mean(Qbar_11) / mean(Qbar_10))
```

```
      bootstrap.est[b,"INT.a"] <- (mean(Qbar_11) -
                                  mean(Qbar_10) - mean(Qbar_01) + mean(Qbar_00))
      # les IC95% du RERI et de l'interaction multiplicative sont d'abord calculés
      # sur une échelle logarithmique
      bootstrap.est[b,"lnRERI"] <- log((mean(Qbar_11) -
                                       mean(Qbar_10) - mean(Qbar_01) + mean(Qbar_00)
                                     / mean(Qbar_00))
      bootstrap.est[b,"lnINT.m"] <- log((mean(Qbar_11) * mean(Qbar_00)) /
                                        (mean(Qbar_10) * mean(Qbar_01)))
    }

    ## On peut vérifier la normalité des distributions bootstrap avec le script
    ## proposé en commentaire ci-dessous :

    # par(mfrow = c(4,4))
    # for(c in 1:ncol(bootstrap.est)) {
    #   hist(bootstrap.est[,c], freq = FALSE, main = names(bootstrap.est)[c])
    #   lines(density(bootstrap.est[,c]), col = 2, lwd = 3)
    #   curve(1/sqrt(var(bootstrap.est[,c]) * 2 * pi) *
    #           exp(-1/2 * ((x-mean(bootstrap.est[,c])) / sd(bootstrap.est[,c]))^2),
    #         col = 1, lwd = 2, lty = 2, add = TRUE)
    # par(mfrow = c(1,1))
    # }

    # Les distributions ont bien une allure normale, on peut donc utiliser la
    # déviation standard des distributions pour calculer les IC95%.
    # En cas de distributions asymétriques, il serait préférable d'utiliser les
    # percentiles 2.5% et 97.5% (ou un méthode bootstrap BCa)

## on calcule les IC95% et on les intègre dans la table de synthèse 'int.r'
    # A1 = 0 et A2 = 0
    int.r$p.lo[int.r$A1 == 0 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 0]
      qnorm(0.975) * sd(bootstrap.est$p.A1is0.A2is0)
    int.r$p.up[int.r$A1 == 0 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 0]
      qnorm(0.975) * sd(bootstrap.est$p.A1is0.A2is0)
    # A1 = 1 et A2 = 0
    int.r$p.lo[int.r$A1 == 1 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 0]
      qnorm(0.975) * sd(bootstrap.est$p.A1is1.A2is0)
    int.r$p.up[int.r$A1 == 1 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 0]
      qnorm(0.975) * sd(bootstrap.est$p.A1is1.A2is0)
    # A1 = 0 et A2 = 1
    int.r$p.lo[int.r$A1 == 0 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 1]
      qnorm(0.975) * sd(bootstrap.est$p.A1is0.A2is1)
    int.r$p.up[int.r$A1 == 0 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 1]
      qnorm(0.975) * sd(bootstrap.est$p.A1is0.A2is1)
```

```
    # A1 = 1 et A2 = 1
    int.r$p.lo[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 1] -
      qnorm(0.975) * sd(bootstrap.est$p.A1is1.A2is1)
    int.r$p.up[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 1] +
      qnorm(0.975) * sd(bootstrap.est$p.A1is1.A2is1)

# risk difference
    # RD.A1.A2is0
    int.r$RD.A1.lo[int.r$A1 == 1 &
                     int.r$A2 == 0] <- (int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 0] -
                                          qnorm(0.975) * sd(bootstrap.est$RD.A1.A2is0))
    int.r$RD.A1.up[int.r$A1 == 1 &
                     int.r$A2 == 0] <- (int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 0] +
                                          qnorm(0.975) * sd(bootstrap.est$RD.A1.A2is0))
    # RD.A1.A2is1
    int.r$RD.A1.lo[int.r$A1 == 1 &
                     int.r$A2 == 1] <- (int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 1] -
                                          qnorm(0.975) * sd(bootstrap.est$RD.A1.A2is1))
    int.r$RD.A1.up[int.r$A1 == 1 &
                     int.r$A2 == 1] <- (int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 1] +
                                          qnorm(0.975) * sd(bootstrap.est$RD.A1.A2is1))
    # RD.A2.A1is0
    int.r$RD.A2.lo[int.r$A1 == 0 &
                     int.r$A2 == 1] <- (int.r$RD.A2[int.r$A1 == 0 & int.r$A2 == 1] -
                                          qnorm(0.975) * sd(bootstrap.est$RD.A2.A1is0))
    int.r$RD.A2.up[int.r$A1 == 0 &
                     int.r$A2 == 1] <- (int.r$RD.A2[int.r$A1 == 0 & int.r$A2 == 1] +
                                          qnorm(0.975) * sd(bootstrap.est$RD.A2.A1is0))
    # RD.A2.A1is1
    int.r$RD.A2.lo[int.r$A1 == 1 &
                     int.r$A2 == 1] <- (int.r$RD.A2[int.r$A1 == 1 & int.r$A2 == 1] -
                                          qnorm(0.975) * sd(bootstrap.est$RD.A2.A1is1))
    int.r$RD.A2.up[int.r$A1 == 1 &
                     int.r$A2 == 1] <- (int.r$RD.A2[int.r$A1 == 1 & int.r$A2 == 1] +
                                          qnorm(0.975) * sd(bootstrap.est$RD.A2.A1is1))

# relative risk
    # RR.A1.A2is0
    int.r$RR.A1.lo[int.r$A1 == 1 &
                     int.r$A2 == 0] <- exp(log(int.r$RR.A1[int.r$A1 == 1 & int.r$A2 == 0]) -
                                                          qnorm(0.975) *
                                             sd(bootstrap.est$lnRR.A1.A2is0))
    int.r$RR.A1.up[int.r$A1 == 1 &
                     int.r$A2 == 0] <- exp(log(int.r$RR.A1[int.r$A1 == 1 & int.r$A2 == 0]) +
                                                          qnorm(0.975) *
```

```
                                                sd(bootstrap.est$lnRR.A1.A2is0))
    # RR.A1.A2is1
    int.r$RR.A1.lo[int.r$A1 == 1 &
                     int.r$A2 == 1] <- exp(log(int.r$RR.A1[int.r$A1 == 1 & int.r$A2 ==
                                                           qnorm(0.975) *
                                                sd(bootstrap.est$lnRR.A1.A2is1))
    int.r$RR.A1.up[int.r$A1 == 1 &
                     int.r$A2 == 1] <- exp(log(int.r$RR.A1[int.r$A1 == 1 & int.r$A2 ==
                                                           qnorm(0.975) *
                                                sd(bootstrap.est$lnRR.A1.A2is1))
    # RR.A2.A1is0
    int.r$RR.A2.lo[int.r$A1 == 0 &
                     int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 0 & int.r$A2 ==
                                                           qnorm(0.975) *
                                                sd(bootstrap.est$lnRR.A2.A1is0))
    int.r$RR.A2.up[int.r$A1 == 0 &
                     int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 0 & int.r$A2 ==
                                                           qnorm(0.975) *
                                                sd(bootstrap.est$lnRR.A2.A1is0))
    # RR.A2.A1is1
    int.r$RR.A2.lo[int.r$A1 == 1 &
                     int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 1 & int.r$A2 ==
                                                           qnorm(0.975) *
                                                sd(bootstrap.est$lnRR.A2.A1is1))
    int.r$RR.A2.up[int.r$A1 == 1 &
                     int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 1 & int.r$A2 ==
                                                           qnorm(0.975) *
                                                sd(bootstrap.est$lnRR.A2.A1is1))

# additive interaction
    int.r$a.INT.lo[int.r$A1 == 1 &
                     int.r$A2 == 1] <- (int.r$a.INT[int.r$A1 == 1 & int.r$A2 == 1] -
                                             qnorm(0.975) * sd(bootstrap.est$INT.a))
    int.r$a.INT.up[int.r$A1 == 1 &
                     int.r$A2 == 1] <- (int.r$a.INT[int.r$A1 == 1 & int.r$A2 == 1] +
                                             qnorm(0.975) * sd(bootstrap.est$INT.a))
# RERI
    int.r$RERI.lo[int.r$A1 == 1 &
                    int.r$A2 == 1] <- exp(log(int.r$RERI[int.r$A1 == 1 & int.r$A2 == 1]
                                               qnorm(0.975) * sd(bootstrap.est$lnRERI))
    int.r$RERI.up[int.r$A1 == 1 &
                    int.r$A2 == 1] <- exp(log(int.r$RERI[int.r$A1 == 1 & int.r$A2 == 1]
                                               qnorm(0.975) * sd(bootstrap.est$lnRERI))
# multiplicative interaction
    int.r$m.INT.lo[int.r$A1 == 1 &
```

```
                      int.r$A2 == 1] <- exp(log(int.r$m.INT[int.r$A1 == 1 & int.r$A2 == 1]) -
                                                qnorm(0.975) * sd(bootstrap.est$lnINT.m))
    int.r$m.INT.up[int.r$A1 == 1 &
                      int.r$A2 == 1] <- exp(log(int.r$m.INT[int.r$A1 == 1 & int.r$A2 == 1]) +
                                                qnorm(0.975) * sd(bootstrap.est$lnINT.m))
```

Au final, on a :

```
## On peut présenter les différents résultats dans un tableau des effets marginaux
    out.table <- data.frame(c1 = rep("",4), c2 = rep("",4), c3 = rep("",4), c4 = rep("",4))
    names(out.table) <- c("A2=0", "A2=1", "RD.A2|A1", "RR.A2|A1")
    rownames(out.table) <- c("A1=0", "A1=1", "RD.A1|A2", "RR.A1|A2")
# p
    out.table["A1=0","A2=0"] <- paste0("$p_{00}$=",
                                       round(int.r$p[which(int.r$A1==0 & int.r$A2==0)],
                                             digits = 3), " [",
                                       round(int.r$p.lo[which(int.r$A1==0 & int.r$A2==0)],
                                             digits = 3), ",",
                                       round(int.r$p.up[which(int.r$A1==0 & int.r$A2==0)],
                                             digits = 3), "]")
    out.table["A1=0","A2=1"] <- paste0("$p_{01}$=",round(int.r$p[which(int.r$A1==0 &
                                                                         int.r$A2==1)],
                                                         digits = 3), " [",
                                       round(int.r$p.lo[which(int.r$A1==0 & int.r$A2==1)],
                                             digits = 3), ",",
                                       round(int.r$p.up[which(int.r$A1==0 & int.r$A2==1)],
                                             digits = 3), "]")
    out.table["A1=1","A2=0"] <- paste0("$p_{10}$=",round(int.r$p[which(int.r$A1==1 &
                                                                         int.r$A2==0)],
                                                         digits = 3), " [",
                                       round(int.r$p.lo[which(int.r$A1==1 & int.r$A2==0)],
                                             digits = 3), ",",
                                       round(int.r$p.up[which(int.r$A1==1 & int.r$A2==0)],
                                             digits = 3), "]")
    out.table["A1=1","A2=1"] <- paste0("$p_{11}$=",round(int.r$p[which(int.r$A1==1 &
                                                                         int.r$A2==1)],
                                                         digits = 3), " [",
                                       round(int.r$p.lo[which(int.r$A1==1 & int.r$A2==1)],
                                             digits = 3), ",",
                                       round(int.r$p.up[which(int.r$A1==1 & int.r$A2==1)],
                                             digits = 3), "]")
# RD
    out.table["A1=0","RD.A2|A1"] <- paste0(round(int.r$RD.A2[which(int.r$A1==0 &
                                                                     int.r$A2==1)],
                                                 digits = 3), " [",
```

```
                                           round(int.r$RD.A2.lo[which(int.r$A1==0 &
                                                                         int.r$A2==1)],
                                                 digits = 3), ",",
                                           round(int.r$RD.A2.up[which(int.r$A1==0 &
                                                                         int.r$A2==1)],
                                                 digits = 3), "]")
    out.table["A1=1","RD.A2|A1"] <- paste0(round(int.r$RD.A2[which(int.r$A1==1 &
                                                                      int.r$A2==1)],
                                                 digits = 3), " [",
                                           round(int.r$RD.A2.lo[which(int.r$A1==1 &
                                                                         int.r$A2==1)],
                                                 digits = 3), ",",
                                           round(int.r$RD.A2.up[which(int.r$A1==1 &
                                                                         int.r$A2==1)],
                                                 digits = 3), "]")
    out.table["RD.A1|A2","A2=0"] <- paste0(round(int.r$RD.A1[which(int.r$A1==1 &
                                                                      int.r$A2==0)],
                                                 digits = 3), " [",
                                           round(int.r$RD.A1.lo[which(int.r$A1==1 &
                                                                         int.r$A2==0)],
                                                 digits = 3), ",",
                                           round(int.r$RD.A1.up[which(int.r$A1==1 &
                                                                         int.r$A2==0)],
                                                 digits = 3), "]")
    out.table["RD.A1|A2","A2=1"] <- paste0(round(int.r$RD.A1[which(int.r$A1==1 &
                                                                      int.r$A2==1)],
                                                 digits = 3), " [",
                                           round(int.r$RD.A1.lo[which(int.r$A1==1 &
                                                                         int.r$A2==1)],
                                                 digits = 3), ",",
                                           round(int.r$RD.A1.up[which(int.r$A1==1 &
                                                                         int.r$A2==1)],
                                                 digits = 3), "]")

# RR
    out.table["A1=0","RR.A2|A1"] <- paste0(round(int.r$RR.A2[which(int.r$A1==0 &
                                                                      int.r$A2==1)],
                                                 digits = 2), " [",
                                           round(int.r$RR.A2.lo[which(int.r$A1==0 &
                                                                         int.r$A2==1)],
                                                 digits = 2), ",",
                                           round(int.r$RR.A2.up[which(int.r$A1==0 &
                                                                         int.r$A2==1)],
                                                 digits = 2), "]")
    out.table["A1=1","RR.A2|A1"] <- paste0(round(int.r$RR.A2[which(int.r$A1==1 &
```

```
                                                          int.r$A2==1)],
                                            digits = 2), " [",
                                      round(int.r$RR.A2.lo[which(int.r$A1==1 &
                                                             int.r$A2==1)],
                                            digits = 2), ",",
                                      round(int.r$RR.A2.up[which(int.r$A1==1 &
                                                             int.r$A2==1)],
                                            digits = 2), "]")
    out.table["RR.A1|A2","A2=0"] <- paste0(round(int.r$RR.A1[which(int.r$A1==1 &
                                                          int.r$A2==0)],
                                            digits = 2), " [",
                                      round(int.r$RR.A1.lo[which(int.r$A1==1 &
                                                             int.r$A2==0)],
                                            digits = 2), ",",
                                      round(int.r$RR.A1.up[which(int.r$A1==1 &
                                                             int.r$A2==0)],
                                            digits = 2), "]")
    out.table["RR.A1|A2","A2=1"] <- paste0(round(int.r$RR.A1[which(int.r$A1==1 &
                                                          int.r$A2==1)],
                                            digits = 2), " [",
                                      round(int.r$RR.A1.lo[which(int.r$A1==1 &
                                                             int.r$A2==1)],
                                            digits = 2), ",",
                                      round(int.r$RR.A1.up[which(int.r$A1==1 &
                                                             int.r$A2==1)],
                                            digits = 2), "]")
    interaction.effects <- c(paste0("additive Interaction = ",
                               round(int.r$a.INT[which(int.r$A1==1 & int.r$A2==1)],
                                     digits = 3), " [",
                               round(int.r$a.INT.lo[which(int.r$A1==1 & int.r$A2==1)],
                                     digits = 3), ";",
                               round(int.r$a.INT.up[which(int.r$A1==1 & int.r$A2==1)],
                                     digits = 3), "]"),
                          paste0("RERI = ",
                                 round(int.r$RERI[which(int.r$A1==1 & int.r$A2==1)],
                                       digits = 2), " [",
                                 round(int.r$RERI.lo[which(int.r$A1==1 & int.r$A2==1)],
                                       digits = 2), ";",
                                 round(int.r$RERI.up[which(int.r$A1==1 & int.r$A2==1)],
                                       digits = 2), "]"),
                          paste0("multiplicative Interaction = ",
                                 round(int.r$m.INT[which(int.r$A1==1 & int.r$A2==1)],
                                       digits = 2), " [",
                                 round(int.r$m.INT.lo[which(int.r$A1==1 & int.r$A2==1)],
                                       digits = 2), ";",
```

	A2=0	A2=1	RD.A2|A1	RR
A1=0	p_{00}=0.104 [0.094,0.113]	p_{01}=0.197 [0.182,0.212]	0.093 [0.076,0.11]	1.9
A1=1	p_{10}=0.4 [0.373,0.427]	p_{11}=0.893 [0.872,0.915]	0.494 [0.46,0.527]	2.23
RD.A1|A2	0.296 [0.268,0.325]	0.697 [0.67,0.723]		
RR.A1|A2	3.86 [3.44,4.33]	4.54 [4.19,4.91]		

Note:
additive Interaction = 0.4 [0.362;0.438]
RERI = 3.86 [3.45;4.33]
multiplicative Interaction = 1.18 [1.03;1.34]

```
                                        round(int.r$m.INT.up[which(int.r$A1==1 & int.r$A2=
                                            digits = 2), "]"))

  library(kableExtra)
  kbl(out.table) %>%
    kable_classic() %>%
    footnote(general = interaction.effects)
```

10.2 Estimation par Modèle Structurel Marginal

Les différentes quantités causales d'intérêt dans une analyse d'interaction peuvent être définies à partir des paramètres d'un modèle structurel marginal (*marginal structural model*, noté MSM).

Un modèle structurel marginal est un modèle qui permet de synthétiser le lien entre les potential outcomes (critères de jugement contrefactuels) et les différentes variables d'exposition d'intérêt. Ils sont particulièrement utiles dès que l'exposition a une dimensionnalité qui est plus grande que la dimensionnalité d'une variable binaire (par exemple si l'exposition est une variable quantitative continue, ou lorsqu'il existe plusieurs variables d'exposition à prendre en compte, comme dans les analyses de la médiation ou les expositions longitudinales répétées dans le temps).

Dans l'analyse d'interaction entre deux variables binaires, on s'intéresse à 4 niveaux d'exposition $\{A_1 = 0, A_2 = 0\}$, $\{A_1 = 1, A_2 = 0\}$, $\{A_1 = 0, A_2 = 1\}$ et $\{A_1 = 1, A_2 = 1\}$. On peut synthétiser les liens entre le potential outcome Y_{A_1,A_2} et les deux variables d'exposition à l'aide du MSM suivant, défini à partir de 4 paramètres β :

$$\mathbb{E}(Y_{A_1,A_2}) = \beta_0 + \beta_1 A_1 + \beta_2 A_2 + \beta_3 (A_1 \times A_2) \tag{10.1}$$

On peut redéfinir nos quantités d'intérêt pour l'analyse d'interaction en fonction des paramètres de ce modèles structurel marginal :

- Les moyennes marginales pour chaque case du tableau d'interaction

$$\begin{aligned}
p_{00} &= \beta_0 \\
p_{10} &= \beta_0 + \beta_1 \\
p_{01} &= \beta_0 + \beta_2 \\
p_{11} &= \beta_0 + \beta_1 + \beta_2 + \beta_3
\end{aligned}$$

- Les différences de risque, risques relatifs et mesures d'interaction peuvent ensuite se calculer comme indiqué précédemment pour la g-computation.

Les *potential outcomes* Y_{A_1,A_2} ne peuvent pas être observés directement dans une étude observationnelles, on n'est pas possible d'appliquer directement l'équation (10.1) du MSM sur la base de donnée observée pour estimer les paramètres β. Différentes méthodes peuvent être appliquées pour estimer ces coefficients β, tout en prenant en compte les biais de confusion :

- estimation par IPTW (*Inverse probability of Treatment Weighting*), il s'agit probablement de la méthode la plus utilisée dans la littérature. Les biais de confusion sont pris en compte en appliquant une méthode de pondération aux différents participants. La pondération repose sur l'estimation des scores de propension $g(A_1)$ et $g(A_2)$ (les probabilités d'exposition à chaque variable d'intérêt A_1 et A_2).
- estimation par g-computation. Elle repose sur l'estimation de fonctions $\overline{Q}$ (modèles du critère de jugement)
- estimation par TMLE (*Targeted maximum likelihood estimation*), il s'agit d'une méthode à double robustesse qui repose à la fois l'estimation des fonctions $\overline{Q}$ et des scores de propension ($g(A_1)$ et $g(A_2)$). Cette méthode sera non-biaisée si au moins les fonctions $\overline{Q}$ ou les fonction g ont été correctement estimées. L'approche par TMLE utilise des méthodes de machine learning pour estimer ces fonctions.

10.2.1 Estimation du MSM par g-computation

Nous pouvons utiliser l'estimation de la fonction $\overline{Q}(A_1, A_2, L)$ réalisée précédemment au paragraphe 10.1 ainsi que les valeurs estimées sous les 4 scénarios contrefactuels.

1) Pour estimer les paramètres d'un MSM par g-computation, il est d'abord nécessaire de constituer une base de données constituée :

- d'une colonne contenant les estimations individuelles de $\mathbb{E}(Y_{A_1=a_1,A_2=a_2})(i) = \overline{Q}_i(a_1, a_2, L)$ pour chaque scénario $\{A_1 = a_1, A_2 = a_2\}$

- d'une colonne pour chaque variable d'exposition d'intérêt. Ici, on a besoin de deux colonnes : une colonne où $A_1 = a_1$ pour tous les individus, et une colonne où $A_2 = a_2$ pour tous les individus.

Initialement, la base de données contient 10 000 individus. Les *potential outcomes* $\overline{Q}_i(a_1, a_2, L)$ ont été estimés à partir de 4 scénarios contrefactuels. On va donc constituer une base de données de 40 000 lignes.

2) On peut ensuite estimer les paramètres β du MSM en appliquant directement l'équation (10.1) sur la nouvelle base de données de 40000 lignes.

```
## 1) Constituer la base de données pour estimer les paramètres du MSM
# On récupère les Qbar_i prédits précédents, que l'on fusionne dans un
# même vecteur
Y <- c(Qbar_00, Qbar_10, Qbar_01, Qbar_11)
length(Y)
# on a un vecteur de 40000 lignes

# On récupère les valeurs d'exposition qui ont servi dans les scénarios contrefactuels
# (garder le même ordre que pour les Qbar)
X <- rbind(subset(df.A1_0.A2_0, select = c("A1", "A2")), # A1 = 0, A2 = 0
           subset(df.A1_1.A2_0, select = c("A1", "A2")), # A1 = 1, A2 = 0
           subset(df.A1_0.A2_1, select = c("A1", "A2")), # A1 = 0, A2 = 1
           subset(df.A1_1.A2_1, select = c("A1", "A2"))) # A1 = 1, A2 = 1
#  dim(X) # il s'agit d'une matrice de 40000 lignes et 2 colonnes

## 2) Estimer les paramètres du MSM en appliquant l'équation (1.10) sur cette
##    nouvelle base de donnée 'data.frame(Y,X)'
# Modèle structurel marginal
msm <- glm(Y ~ A1 + A2 + A1:A2,
           data = data.frame(Y,X),
           family = "gaussian")
# note : il n'est pas utile d'ajuster sur les facteurs de confusion L
# car les Y ont été prédits à partir de modèles qui les déjà ont pris en compte
msm

## 3) Utilisation des paramètres du MSM afin d'estimer les quantités d'intérêt
##    pour l'analyse d'interaction
# présentation des résultats dans un tableau des effets marginaux
results.MSM <- matrix(NA, ncol = 4, nrow = 4)
colnames(results.MSM) <- c("A2 = 0", "A2 = 1",
                           "RD within strata of A1",
                           "RR within strata of A1")
rownames(results.MSM) <- c("A1 = 0", "A1 = 1",
                           "RD within strata of A2",
```

```
                           "RR within strata of A2")

# Estimation de p_00, p_10, p_01 et p_11 à partir des coefficients du MSM
results.MSM["A1 = 0","A2 = 0"] <- msm$coefficients["(Intercept)"]
results.MSM["A1 = 0","A2 = 1"] <- (msm$coefficients["(Intercept)"] +
                                     msm$coefficients["A2"])
results.MSM["A1 = 1","A2 = 0"] <- (msm$coefficients["(Intercept)"] +
                                     msm$coefficients["A1"])
results.MSM["A1 = 1","A2 = 1"] <- (msm$coefficients["(Intercept)"] +
                                     msm$coefficients["A2"] + msm$coefficients["A1"] +
                                     msm$coefficients["A1:A2"])

# RD et RR dans les strates de A2
results.MSM["RR within strata of A2", "A2 = 0"] <- (results.MSM["A1 = 1","A2 = 0"] /
                                                      results.MSM["A1 = 0","A2 = 0"])
results.MSM["RD within strata of A2", "A2 = 0"] <- (results.MSM["A1 = 1","A2 = 0"] -
                                                      results.MSM["A1 = 0","A2 = 0"])
results.MSM["RR within strata of A2", "A2 = 1"] <- (results.MSM["A1 = 1","A2 = 1"] /
                                                      results.MSM["A1 = 0","A2 = 1"])
results.MSM["RD within strata of A2", "A2 = 1"] <- (results.MSM["A1 = 1","A2 = 1"] -
                                                      results.MSM["A1 = 0","A2 = 1"])

# RD et RR dans les strates de A1
results.MSM["A1 = 0", "RR within strata of A1"] <- (results.MSM["A1 = 0","A2 = 1"] /
                                                      results.MSM["A1 = 0","A2 = 0"])
results.MSM["A1 = 0", "RD within strata of A1"] <- (results.MSM["A1 = 0","A2 = 1"] -
                                                      results.MSM["A1 = 0","A2 = 0"])
results.MSM["A1 = 1", "RR within strata of A1"] <- (results.MSM["A1 = 1","A2 = 1"] /
                                                      results.MSM["A1 = 1","A2 = 0"])
results.MSM["A1 = 1", "RD within strata of A1"] <- (results.MSM["A1 = 1","A2 = 1"] -
                                                      results.MSM["A1 = 1","A2 = 0"])

results.MSM <- round(results.MSM,3)
RD.interaction <- msm$coefficients["A1:A2"]
RR.interaction <- ((results.MSM["A1 = 1","A2 = 1"] * results.MSM["A1 = 0","A2 = 0"]) /
                     (results.MSM["A1 = 0","A2 = 1"] * results.MSM["A1 = 1","A2 = 0"]))
RERI <- ((results.MSM["A1 = 0","A2 = 0"] - results.MSM["A1 = 1","A2 = 0"] -
           results.MSM["A1 = 0","A2 = 1"] + results.MSM["A1 = 1","A2 = 1"]) /
           results.MSM["A1 = 0","A2 = 0"])
```

Au final, on a (sans les IC):

	A2 = 0	A2 = 1	RD within strata of A1	RR within strata of A1
A1 = 0	0.106	0.198	0.092	1.872
A1 = 1	0.404	0.892	0.488	2.206
RD within strata of A2	0.298	0.694	NA	NA
RR within strata of A2	3.816	4.498	NA	NA

Note:
additive Interaction = 0.395
RERI = 3.736
multiplicative Interaction = 1.18

10.2.2 Estimation du MSM par IPTW

Les paramètres du MSM de l'équation (10.1) peuvent également être estimée par une méthode de pondération inverse (IPTW).

Pour cela, il suffit d'appliquer une régression pondérée correspondant à l'équation (10.1), où chaque individu i est pondéré par w_i le produit l'inverse des scores de propension de l'exposition A_1 et de l'exposition A_2 (pour rappel, d'après le DAG simulant les données, les parents de A_1 sont L_1 et L_2, et les parents de A_2 sont L_2 et L_3) :

$$w_i = \frac{g^*(A_{1,i})}{g(A_{1,i} \mid L_1, L_2)} \times \frac{g^*(A_{2,i})}{g(A_{2,i} \mid L_2, L_3)}$$

où

- $g^*(A_i)$ sont des fonctions choisies par l'utilisateur, par exemple $g^*(A_i) = 1$ correspondant à une pondération IPTW simple, ou encore $g^*(A_i) = g(A_i)$ qui permet de stabiliser les estimations (*stabilized IPTW*)
- $g(A_i \mid L)$ sont les scores de propension : ils modélisent la probabilité d'observer sa propre exposition.

$$g(A_i \mid L) = \begin{cases} P(A = 1 \mid L_i) & \text{si } A_i = 1 \\ 1 - P(A = 1 \mid L_i) & \text{si } A_i = 0 \end{cases}$$

Cette pondération permet de créer une "pseudo-population" dans laquelle les facteurs de confusion des relations $A_1 \to Y$ et $A_2 \to Y$ sont équilibrés entre les strates d'exposition.

```
### Par exemple pour calculer des poids stabilisés :
## 1) estimer les fonctions g(A_1)*, g(A_1 | L1, L2), g(A_2*), g(A_2 | L2, L3)
model_gA1_star <- glm(A1 ~ 1, data = df, family = "binomial")
model_gA1 <- glm(A1 ~ L1 + L2, data = df, family = "binomial")
p_A1is1_star <- predict(model_gA1_star, type = "response")
```

```r
p_A1is1 <- predict(model_gA1, type = "response")

model_gA2_star <- glm(A2 ~ 1, data = df, family = "binomial")
model_gA2 <- glm(A2 ~ L2 + L3, data = df, family = "binomial")
p_A2is1_star <- predict(model_gA2_star, type = "response")
p_A2is1 <- predict(model_gA2, type = "response")

gA1_star <- gA1 <- rep(NA, nrow(df))
gA1_star[df$A1 == 1] <- p_A1is1_star[df$A1 == 1]
gA1_star[df$A1 == 0] <- 1 - p_A1is1_star[df$A1 == 0]
gA1[df$A1 == 1] <- p_A1is1[df$A1 == 1]
gA1[df$A1 == 0] <- 1 - p_A1is1[df$A1 == 0]

gA2_star <- gA2 <- rep(NA, nrow(df))
gA2_star[df$A2 == 1] <- p_A2is1_star[df$A2 == 1]
gA2_star[df$A2 == 0] <- 1 - p_A2is1_star[df$A2 == 0]
gA2[df$A2 == 1] <- p_A2is1[df$A2 == 1]
gA2[df$A2 == 0] <- 1 - p_A2is1[df$A2 == 0]

# On peut alors calculer les poids individuels stabilisés
sw <- (gA1_star / gA1) * (gA2_star / gA2)

## 2) Estimer les coefficients du MSM à l'aide d'une régression pondérée
# L'équation (1.10) correspond à un modèle additif,
# nous allons utiliser une régression linéaire (family = gaussian)
msm_iptw <- glm(Y ~ A1 + A2 + A1:A2, data = df, family = "gaussian",
                weights = sw)

msm_iptw

## 3) Utilisation des paramètres du MSM afin d'estimer les quantités d'intérêt
##    pour l'analyse d'interaction
# présentation des résultats dans un tableau des effets marginaux
results.MSM <- matrix(NA, ncol = 4, nrow = 4)
colnames(results.MSM) <- c("A2 = 0", "A2 = 1",
                           "RD within strata of A1",
                           "RR within strata of A1")
rownames(results.MSM) <- c("A1 = 0", "A1 = 1",
                           "RD within strata of A2",
                           "RR within strata of A2")

# Estimation de p_00, p_10, p_01 et p_11 à partir des coefficients du MSM
results.MSM["A1 = 0","A2 = 0"] <- msm_iptw$coefficients["(Intercept)"]
results.MSM["A1 = 0","A2 = 1"] <- (msm_iptw$coefficients["(Intercept)"] +
                                     msm_iptw$coefficients["A2"])
```

```
results.MSM["A1 = 1","A2 = 0"] <- (msm_iptw$coefficients["(Intercept)"] +
                                     msm_iptw$coefficients["A1"])
results.MSM["A1 = 1","A2 = 1"] <- (msm_iptw$coefficients["(Intercept)"] +
                                     msm_iptw$coefficients["A2"] +
                                     msm_iptw$coefficients["A1"] +
                                     msm_iptw$coefficients["A1:A2"])

# RD et RR dans les strates de A2
results.MSM["RR within strata of A2", "A2 = 0"] <- (results.MSM["A1 = 1","A2 = 0"] /
                                                      results.MSM["A1 = 0","A2 = 0"])
results.MSM["RD within strata of A2", "A2 = 0"] <- (results.MSM["A1 = 1","A2 = 0"] -
                                                      results.MSM["A1 = 0","A2 = 0"])
results.MSM["RR within strata of A2", "A2 = 1"] <- (results.MSM["A1 = 1","A2 = 1"] /
                                                      results.MSM["A1 = 0","A2 = 1"])
results.MSM["RD within strata of A2", "A2 = 1"] <- (results.MSM["A1 = 1","A2 = 1"] -
                                                      results.MSM["A1 = 0","A2 = 1"])

# RD et RR dans les strates de A1
results.MSM["A1 = 0", "RR within strata of A1"] <- (results.MSM["A1 = 0","A2 = 1"] /
                                                      results.MSM["A1 = 0","A2 = 0"])
results.MSM["A1 = 0", "RD within strata of A1"] <- (results.MSM["A1 = 0","A2 = 1"] -
                                                      results.MSM["A1 = 0","A2 = 0"])
results.MSM["A1 = 1", "RR within strata of A1"] <- (results.MSM["A1 = 1","A2 = 1"] /
                                                      results.MSM["A1 = 1","A2 = 0"])
results.MSM["A1 = 1", "RD within strata of A1"] <- (results.MSM["A1 = 1","A2 = 1"] -
                                                      results.MSM["A1 = 1","A2 = 0"])

results.MSM <- round(results.MSM,3)
RD.interaction <- msm_iptw$coefficients["A1:A2"]
RR.interaction <- ((results.MSM["A1 = 1","A2 = 1"] * results.MSM["A1 = 0","A2 = 0"]) /
                     (results.MSM["A1 = 0","A2 = 1"] * results.MSM["A1 = 1","A2 = 0"]))
RERI <- ((results.MSM["A1 = 0","A2 = 0"] - results.MSM["A1 = 1","A2 = 0"] -
           results.MSM["A1 = 0","A2 = 1"] + results.MSM["A1 = 1","A2 = 1"]) /
           results.MSM["A1 = 0","A2 = 0"])
```

Au final, on a (sans les IC):

10.3 Estimation avec TMLE

Le package R `ltmle` contient la fonction `ltmleMSM` qui permet d'estimer les paramètres d'un modèle structurel marginal par TMLE, par IPTW ou par g-computation. Par contre, le MSM utilisé est nécessairement une régression logistique. Les résultats que l'on peut en déduire sont donc des Odds Ratios ou une approximation du RERI, comme cela a été fait au paragraphe 9.1.

	A2 = 0	A2 = 1	RD within strata of A1	RR within strata of A1
A1 = 0	0.102	0.199	0.097	1.952
A1 = 1	0.409	0.900	0.491	2.202
RD within strata of A2	0.307	0.701	NA	NA
RR within strata of A2	4.000	4.514	NA	NA

Note:
additive Interaction = 0.394
RERI = 3.863
multiplicative Interaction = 1.13

Il est possible d'estimer les différences de risque, risques relatifs, et les différentes mesures d'interaction à partir de ce MSM défini sous forme de régression logistique, mais les calculs manuels sont plus fastidieux. Par ailleurs, les calculs d'intervalles de confiance reposent sur une delta-méthode.

Nous avons développé un package `MargIntTmle` disponible sur cette page Github, qui permet de faire ces calculs automatiquement. Vous trouverez la documentation complète du package sur la page.

Ci-dessous nous illustrons l'utilisation du package `MargIntTmle` pour obtenir des estimations par g-computation, IPTW, ou TMLE.

```
# L'installation du package peut se faire depuis la commande
# devtools::install_github("benoitlepage/MargIntTmle")
library(MargIntTmle)

# le package 'MargIntTmle' repose sur les package 'ltmle' et 'SuperLearner'
require(ltmle)
require(SuperLearner)

## Parmi les arguments nécessaires, il faut indiquer les fonctions Q et g
# Fonction Q
Q_formulas = c(Y="Q.kplus1 ~ A1 + A2 + A1 * A2 + L1 + L2 + L3")
# Les deux fonctions g
g_formulas = c("A1 ~ L1 + L2",
               "A2 ~ L2 + L3")

## Choisir des librairies de machine learning disponibles dans le package 'SuperLearner':
SL.library = list(Q=list("SL.glm"),g=list("SL.glm"))

## Utilisation de la fonction 'int.ltmleMSM()'
# Cette fonction va estimer les paramètres du MSM soit par g-computation,
# soit par TMLE et IPTW

# Pour obtenir une estimation des paramètres du MSM par TMLE et par IPTW, l'argument
```

```
# 'gcomp' doit égal à FALSE
interaction.ltmle <- int.ltmleMSM(data = df,
                                  Qform = Q_formulas,
                                  gform = g_formulas,
                                  Anodes = c("A1", "A2"),
                                  Lnodes = c("L1", "L2", "L3"),
                                  Ynodes = c("Y"),
                                  SL.library = SL.library,
                                  gcomp = FALSE, # pour l'estimation par TMLE et IPTW
                                  iptw.only = FALSE,
                                  survivalOutcome = FALSE,
                                  variance.method = "ic")

# Pour obtenir une estimation des paramètres du MSM par g-computation, l'argument
# 'gcomp' doit être égal à TRUE.
# Les IC95% seront alors estimés par bootstrap, vous pouvez préciser le nombre
# d'échantillons bootstrap et donner une valeur de seed
interaction.gcomp <- int.ltmleMSM(data = df,
                                  Qform = Q_formulas,
                                  gform = g_formulas,
                                  Anodes = c("A1", "A2"),
                                  Lnodes = c("L1", "L2", "L3"),
                                  Ynodes = c("Y"),
                                  SL.library = SL.library,
                                  gcomp = TRUE, # pour l'estimation par g-computation
                                  iptw.only = FALSE,
                                  survivalOutcome = FALSE,
                                  variance.method = "ic",
                                  B = 1000, # number of bootstrap samples
                                  boot.seed = 42) # seed for bootstrap
```

```
## [1] "bootstrap number 100"
## [1] "bootstrap number 200"
## [1] "bootstrap number 300"
## [1] "bootstrap number 400"
## [1] "bootstrap number 500"
## [1] "bootstrap number 600"
## [1] "bootstrap number 700"
## [1] "bootstrap number 800"
## [1] "bootstrap number 900"
## [1] "bootstrap number 1000"
```

```
## La fonction 'estim.int.effects()' permet d'estimer les différentes quantités
## d'intérêt pour l'analyse d'interaction à partir du MSM estimé précédemment.
# Les IC95% sont calculés par Delta méthode pour la TMLE et l'IPTW, et par
```

```
# bootstrap pour la g-computation.

# pour obtenir les estimations par TMLE :
est.tmle <- estim.int.effects(interaction.ltmle, estimator = "tmle")

# pour obtenir les estimations par IPTW :
est.iptw <- estim.int.effects(interaction.ltmle, estimator = "iptw")

# pour obtenir les estimations par g-computation :
est.gcomp <- estim.int.effects(interaction.gcomp, estimator = "gcomp")

## Les résultats peuvent ensuite être présentés sous forme de table, en suivant
## les recommandations de Knol et al. (2012) à l'aide de la fonction 'out.int.table()'

# par exemple pour les estimations par TMLE :
table_inter <- out.int.table(int.r = est.tmle)
table_inter$out.table
```

```
##                                  A2=0                        A2=1
## A1=0     $p_{00}$=0.103 [0.095,0.112] $p_{01}$=0.196 [0.181,0.212]
## A1=1      $p_{10}$=0.411 [0.38,0.442] $p_{11}$=0.898 [0.877,0.919]
## RD.A1|A2           0.308 [0.276,0.34]          0.702 [0.676,0.728]
## RR.A1|A2             3.98 [3.55,4.46]             4.58 [4.22,4.97]
##                     RD.A2|A1         RR.A2|A1
## A1=0      0.093 [0.075,0.11]  1.9 [1.69,2.13]
## A1=1     0.487 [0.449,0.524] 2.18 [2.02,2.36]
## RD.A1|A2
## RR.A1|A2
```

```
# les mesures d'interaction sont disponible dans l'objet
table_inter$interaction.effects
```

```
## [1] "additive Interaction = 0.394 [0.353;0.435]"
## [2] "RERI = 3.81 [3.38;4.3]"
## [3] "multiplicative Interaction = 1.15 [1;1.32]"
```

On peut également se servir du package `KableExtra` pour mettre en forme les tables (selon les recommandations Knol et al. Knol and VanderWeele [2012]) :

Effets d'interaction estimés par TMLE :

```
library(knitr)
knitr::kable(table_inter$out.table,) %>%
  kable_classic() %>%
```

	A2=0	A2=1	RD.A2|A1	R
A1=0	p_{00}=0.103 [0.095,0.112]	p_{01}=0.196 [0.181,0.212]	0.093 [0.075,0.11]	1.9
A1=1	p_{10}=0.411 [0.38,0.442]	p_{11}=0.898 [0.877,0.919]	0.487 [0.449,0.524]	2.1
RD.A1|A2	0.308 [0.276,0.34]	0.702 [0.676,0.728]		
RR.A1|A2	3.98 [3.55,4.46]	4.58 [4.22,4.97]		

[a] additive Interaction = 0.394 [0.353;0.435]
[b] RERI = 3.81 [3.38;4.3]
[c] multiplicative Interaction = 1.15 [1;1.32]

	A2=0	A2=1	RD.A2|A1	R
A1=0	p_{00}=0.102 [0.093,0.111]	p_{01}=0.199 [0.184,0.215]	0.097 [0.079,0.115]	1.9
A1=1	p_{10}=0.409 [0.378,0.44]	p_{11}=0.9 [0.879,0.921]	0.491 [0.454,0.529]	2.2
RD.A1|A2	0.307 [0.274,0.339]	0.701 [0.675,0.727]		
RR.A1|A2	4 [3.57,4.48]	4.51 [4.16,4.89]		

[a] additive Interaction = 0.394 [0.353;0.436]
[b] RERI = 3.86 [3.42;4.35]
[c] multiplicative Interaction = 1.13 [0.98;1.3]

```
  add_footnote(c(table_inter$interaction.effects[1],
                 table_inter$interaction.effects[2],
                 table_inter$interaction.effects[3]),
               escape = TRUE)
```

Résultats obtenus en appliquant l'estimation par IPTW :

```
table_inter <- MargIntTmle::out.int.table(int.r = est.iptw)
knitr::kable(table_inter$out.table,) %>%
  kable_classic() %>%
  add_footnote(c(table_inter$interaction.effects[1],
                 table_inter$interaction.effects[2],
                 table_inter$interaction.effects[3]),
               escape = TRUE)
```

Résultats obtenus en appliquant l'estimation par g-computation :

```
table_inter <- MargIntTmle::out.int.table(int.r = est.gcomp)
knitr::kable(table_inter$out.table,) %>%
  kable_classic() %>%
  add_footnote(c(table_inter$interaction.effects[1],
                 table_inter$interaction.effects[2],
                 table_inter$interaction.effects[3]),
               escape = TRUE)
```

	A2=0	A2=1	RD.A2|A1	RR.A2|A1
A1=0	p_{00}=0.104 [0.095,0.112]	p_{01}=0.197 [0.183,0.211]	0.093 [0.076,0.11]	1.9 [1.7,2.13]
A1=1	p_{10}=0.4 [0.372,0.427]	p_{11}=0.893 [0.872,0.915]	0.494 [0.458,0.529]	2.23 [2.08,2.4]
RD.A1|A2	0.296 [0.267,0.326]	0.697 [0.67,0.723]		
RR.A1|A2	3.86 [3.45,4.32]	4.54 [4.2,4.9]		

[a] additive Interaction = 0.4 [0.361;0.439]
[b] RERI = 3.86 [3.46;4.32]
[c] multiplicative Interaction = 1.18 [1.03;1.35]

Chapter 11

Représentations graphiques

Part III

En pratique

Chapter 12

Proposition d'étapes

1. **Formuler l'objectif**
    - Est-ce un objectif prédictif ou explicatif ?
    - Si démarche explicative, s'agit-il plutot d'une analyse d'interaction ou de modification d'effet?
2. **Stratégies et méthodes**
    - Poser les hypothèses sur un DAG ou schéma conceptuel
    - Identifier le ou les estimand(s), c'est-à-dire l'effet ou le paramètre que l'on va chercher à estimer pour répondre à l'objectif, par exemple :
        - effet conjoint de X et V sur Y, sur l'échelle multiplicative = $OR_{X,V}$
        - effet de X sur Y dans chaque strate de Y, sur l'échelle additive = $DR_{X|V=0}$ et $DR_{X|V=1}$
        - effet d'interaction sur l'échelle additive et multiplicative AI et MI
    - Elaborer l'estimateur, notamment :
        - quelles est(sont) l'exposition(s) d'intérêt ?
        - quels sont les facteurs de confusion +/- les médiateurs si besoin ?
        - quels types de modélisation va être utilisée (linéaire, logistique, autre) ?
3. **Analyses descriptives**
    - Description habituelle de la population
    - Décrire, dans un tableau croisé,
        - le Y moyen ou la proportion de Y = 1
        - pour chaque catégorie de X et V
4. **Analyses exploratoires**

- Analyses stratifiées
  - pour une analyse de modification d'effet,
  - il est possible en exploratoire, d'estimer l'effet de X sur Y
  - de façon stratifiée sur V (on découpe la population)
  - les effets ne seront directement pas comparables

5. **Analyses confirmatoires**

- Régressions avec terme d'interaction (voir Chapitre 9)
  - un modèle dans la population totale peut être utilisé
  - avec un terme d'interaction entre X et V
  - les différents paramètres peuvent être déduits des résultats du modèle
- Approches causales (voir Chapitre 10)
  - G-computation
  - MSM
  - TMLE

Chapter 13

Exemple 1 - Y binaire

13.1 Formuler les objectifs

Dans cet exemple, on s'intéresse à :

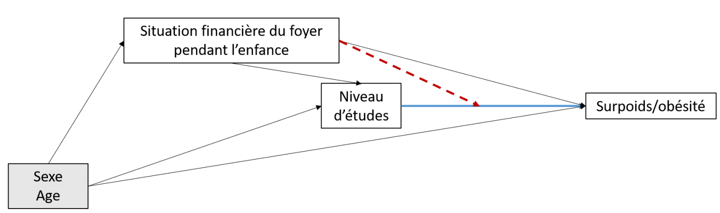
- Comment l'effet du niveau d'études (X) sur le surpoids/obésité à l'âge adulte (Y) varie en fonction de la défavorisation sociale précoce (D), mesurée par la situation financière du foyer pendant l'enfance.

La démarche ici est explicative : on cherche à comprendre des mécanismes causaux.

A partir de la formulation des objectifs, on pourrait dire qu'on s'intéresse ici plutôt à une modification d'effets: on analyse l'effet du scénario $do(X)$ dans chaque groupe de défavorisation sociale précoce (D). On ajustera sur les facteurs de confusion de la relation X $\rightarrow$ Y

13.2 Stratégies et méthodes

Le DAG (sans les médiateurs) était :



Niveau d'études	Défavorisation	% surpoids/obésité
Elevé	Non	38.4
Elevé	Oui	45.2
Faible	Non	50.2
Faible	Oui	54.6

Avec :

- X, le Niveau d'études : 0 = élevé / 1 = faible (réf)
- D, la Situation financière pendant l'enfance : 0 = bonne / 1 = difficile (réf)
- Y, le Surpoids/obésité : 0 = IMC < 25kg/m^2 / 1 = IMC   25kg/m^2

Les estimands étaient définis sur l'échelle multiplicative par :

- La modification de l'effet du niveau d'études sur le surpoids/obésité en fonction par la défavorisation sociale précoce :
  - $(Y_{x=1|d=1}/Y_{x=0|d=1})/(Y_{x=1|d=0}/Y_{x=0|d=0})$
  - Ce qui est équivalent à $(Y_{x=1|d=1} \times Y_{x=0|d=0})/(Y_{x=1|d=0} \times Y_{x=0|d=1})$

L'estimateur : Les effets ont été estimés par g-computation (*standardisation par régression*) Hernán and Robins [2020]. Des régressions linéaires ont été utilisées pour estimer les *potential outcomes* pour chaque scénario. A partir des fonctions estimées, nous avons prédit la valeur de l'outcome Y pour chaque individu i pour chaque scénario. Les valeurs moyennes de Y dans chaque scénario vont ensuite nous permettre d'estimer les *estimands* selon leurs définitions précisées ci-dessus.Ces modèles vont comprendre 4 variables : le niveau d'études et la défavorisation sociale précoce, ainsi que deux facteurs de confusion, le sexe et l'âge.

13.3 Analyse descriptive

Dans cette population (N=23 495), il y avait 61.1% d'individus avec un niveau d'études faible et 31.1% de personnes ayant été précocement défavorisées.

On peut commencer par décrire les proportions de personnes en surpoids/obésité dans chaque catégorie de niveau d'études et de défavorisation sociale :

13.4 Analyse exploratoire

La sortie d'un modèle logistique simple serait :

```
# Call:
# glm(formula = overw_obesity ~ EDUCATION_2CL.f * CHILDHOOD_ECONOMY_2CL.f + SEX.f +
#     AGE, family = binomial(link = "logit"))
#
# Coefficients:
#                                                  Estimate Std. Error t value   Pr(>|t|)
# (Intercept)                                     -1.3731389  0.0621930 -22.079  < 2e-16 ***
# EDUCATION_2CL.fHigh                             -0.1752537  0.0537373  -3.261  0.00111 **
# CHILDHOOD_ECONOMY_2CL.fGood                     -0.0190075  0.0361206  -0.526  0.59873
# SEX.fMale                                        0.5882549  0.0270502  21.747  < 2e-16 ***
# AGE                                              0.0234627  0.0009856  23.807  < 2e-16 ***
# EDUCATION_2CL.fHigh:CHILDHOOD_ECONOMY_2CL.fGood  -0.1312722  0.0623235  -2.106  0.03518 *
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#
```

On peut en déduire (échelle multiplicative) que :

- L'effet du niveau d'études (élevé plutôt que faible) sur le risque de surpoids/obésité est :
  - Quand on est défavorisé pendant l'enfance: $OR(X|D = 0) = exp(-0.175) = 0.84$
  - Quand on est favorisé pendant l'enfance: $OR(X|D = 1) = exp(-0.175 - 0.131) = 0.74$
- L'effet d'avoir un niveau d'étude élevé et d'être favorisé pendant l'enfance
  - plutot qu'avoir un niveau d'étude faible et être défavorisé pendant l'enfance est
  - $OR(X, D) = exp(-0.175 - 0.019 - 0.131) = 0.72$
- **La modification d'effet** est de:
  - Sur l'échelle multiplicative: $MI = exp(-0.131) = 0.88$ (interaction multiplicative <1 donc négative)
  - Sur l'échelle additive: $RERI = exp(-0.175 - 0.019 - 0.131) - exp(-0.175)$–$exp(-0.019) + 1 = -0.098$ (interaction additive négative)

13.5 Analyse confirmatoire

Si l'on utilise le package proposé par B Lepage pour réaliser cette analyse avec la g-computation, les résultats sont :

Les résulats peuvent être intérprétés ainsi :

	A2=0	A2=1	RD.A2|A1
A1=0	p_{00}=0.5 [0.487,0.513]	p_{01}=0.496 [0.486,0.506]	-0.005 [-0.021,0.012]
A1=1	p_{10}=0.459 [0.438,0.479]	p_{11}=0.423 [0.412,0.435]	-0.035 [-0.059,-0.012]
RD.A1|A2	-0.042 [-0.065,-0.018]	-0.072 [-0.088,-0.057]	
RR.A1|A2	0.92 [0.87,0.96]	0.85 [0.83,0.88]	

[a] additive Interaction = -0.031 [-0.059;-0.003]
[b] RERI = -0.061 [-0.123;4e-04]
[c] multiplicative Interaction = 0.932 [0.877;0.989]

- l'effet d'un niveau d'études élevé (par rapport à faible) sur le risque de surpoids/obésité est moins fort de 3% lorsqu'on est défavorisé précocement
- ou encore, un niveau d'études élevé joue un rôle protecteur contre le surpoids/obésité moins important chez les personnes ayant grandi dans un foyer défavorisé

Chapter 14

Exemple 2 - Y quantitatif

14.1 Formuler les objectifs

Dans cette étude Colineaux et al. [2023], on s'est intéressé à :

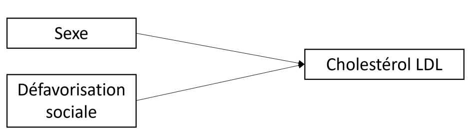
- comment l'effet du sexe sur le taux de cholestérol LDL vers 45 ans varie en fonction de la défavorisation sociale précoce,
- comment l'effet de la défavorisation sociale précoce sur le taux de cholestérol LDL varie en fonction du sexe.

La démarche ici est explicative : on cherche à comprendre des mécanismes causaux.

A partir de la formulation des objectifs, on pourrait dire qu'on s'intéresse ici plutot à deux modifications d'effet. On va donc devoir à la fois agir sur le sexe $do(S)$ et sur la défavorisation sociale $do(D)$. Donc la démarche, en fait, sera plutôt une analyse d'interaction $do(S, D)$

14.2 Stratégies et méthodes

Le DAG (sans les médiateurs) était :



Les estimands étaient définis sur l'échelle additive par :

Sexe	Défavorisation	Mean(Chol LDL)
Male	Non	3.57
Male	Oui	3.60
Female	Non	3.24
Female	Oui	3.37

- La modification de l'effet du sexe en fonction par la défavorisation sociale précoce :
  - $(Y_{s=1|d=0} - Y_{s=0|d=0}) - (Y_{s=1|d=1} - Y_{s=0|d=1})$
  - ou $(Y_{s=1,d=0} - Y_{s=0,d=0}) - (Y_{s=1,d=1} - Y_{s=0,d=1})$
- La modification de l'effet de la défavorisation sociale précoce par la sexe
  - $(Y_{d=1|s=0} - Y_{d=0|s=0}) - (Y_{d=1|s=1} - Y_{d=0|s=1})$
  - ou $(Y_{d=1,s=0} - Y_{d=0,s=0}) - (Y_{d=1,s=1} - Y_{d=0,s=1})$

Les deux formulations sont ici équivalentes car il n'y pas de facteurs de confusion, donc, par exemple, $Y_{d=1|s=0} = Y_{s=0|d=1} = Y_{d=1,s=0}$

L'estimateur : Les effets ont été estimés par g-computation (*standardisation par régression*) Hernán and Robins [2020]. Des régressions linéaires ont été utilisées pour estimer les *potential outcomes* pour chaque scénario, désignées par $\overline{Q}(S,D) = E(Y|S,D)$. A partir des fonctions $\overline{Q}(S,D)$ estimées, nous avons prédit la valeur de l'outcome Y pour chaque individu i pour chaque scénario. Les valeurs moyennes de Y dans chaque scénario vont ensuite nous permettre d'estimer les estimands selon leurs définitions précisées ci-dessus. Ces modèles$\overline{Q}(S,D)$ vont comprendre 2 variables : le sexe et la défavorisation sociale précoce (il n'y a pas ici de facteurs de confusion).

14.3 Analyse descriptive

Dans cette population (N=17 272), il y avait 51,4% d'hommes et 60,5% de personnes ayant été précocement défavorisées.

On peut commencer par décrire les moyennes de cholestérol dans chaque catégorie de sexe et de défavorisation sociale :

14.4 Analyse exploratoire

La sortie d'une modèle linéaire simple serait :

```
# Call:
# lm(formula = t8_ldl ~ as.factor(sex) + as.factor(soc_group) +
#     as.factor(sex) * as.factor(soc_group), data = ba_1)
#
# Coefficients:
#                                            Estimate Std. Error t value Pr(>|t|)
# (Intercept)                                 3.24270    0.01594 203.475  < 2e-16 ***
# as.factor(sex)1                             0.32553    0.02227  14.616  < 2e-16 ***
# as.factor(soc_group)2.Défav                 0.12614    0.02052   6.148 8.02e-10 ***
# as.factor(sex)1:as.factor(soc_group)2.Défav -0.09473    0.02863  -3.308 0.000941 ***
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#
```

On peut en déduire (échelle additive) que :

- L'effet du sexe (d'être homme plutot que femme) est :
  - Quand on est favorisé : $DR(S|D=0) = +0.326$ mmol/L
  - Quand on est défavorisé : $DR(S|D=1) = 0.326 - 0.095 = +0.231$ mmol/L
- L'effet de la défavorisation est :
  - Quand on est une femme : $DR(D|S=0) = +0.126$ mmol/L
  - Quand on est un homme : $DR(D|S=1) = 0.126 - 0.095 = +0.031$ mmol/L
- L'effet d'être un homme et défavorisé
  - plutot que femme et favorisé est
  - $DR(D,S) = 0.326 + 0.126 - 0.095 = +0.357$ mmol/L
- **L'effet d'interaction/modification d'effet** est : $AI = -0.095$ mmol/L

On pourrait aussi déduire :

- $Y_{00} = 3.24$ mmol/L
- $Y_{10} = 3.243 + 0.326 = 3.57$ mmol/L
- $Y_{01} = 3.243 + 0.126 = 3.37$ mmol/L
- $Y_{11} = 3.243 + 0.326 + 0.126 - 0.095 = 3.6$ mmol/L

14.5 Analyse confirmatoire

Si l'on utilise le package proposé par B Lepage pour réaliser cet analyse avec la TMLE (effets d'intéraction calculés à partir des paramètres d'une modèle structurel marginal estimé à l'aide du package R ltmle), les résultats sont :

	A2=0	A2=1	RD.A2|A1
A1=0	p_{00}=3.243 [3.213,3.273]	p_{01}=3.369 [3.344,3.394]	0.126 [0.087,0.165]
A1=1	p_{10}=3.568 [3.538,3.598]	p_{11}=3.6 [3.574,3.625]	0.031 [-0.008,0.071]
RD.A1|A2	0.326 [0.283,0.368]	0.231 [0.195,0.267]	

[a] additive Interaction = -0.095 [-0.15;-0.039]

On retrouve des résulats qui peuvent être intérprétés ainsi :

- l'effet d'être un homme (ou "la différence H-F) est moins fort de additive Interaction = -0.095 [-0.15;-0.039] mmol/L lorsqu'on est défavorisé précocement
- l'effet de la défavorisation est moins fort de additive Interaction = -0.095 [-0.15;-0.039] mmol/L chez les hommes

---

En réalité, on a réalisé cette analyse par g-computation (voir chapitre 10) sur des données imputées et boostrappées (l'exemple ci-dessus a été réalisé sur une seule des bases bootstrappées, ce qui explique les différences), et les résultats, présentées selon les recommandations modifiées de Knol et VanderWeele, étaient:

	Né·e·s-Avantagé·e·s		Né·e·s désavantagé·e·s		ET du désavantage précoce	
Nés-Hommes (moyenne)	3,48	[3,44 à 3,52]	3,49	[3,45 à 3,52]	+0,01	[-0,04 à 0,05]
Nées-Femmes (moyenne)	3,24	[3,20 à 3,28]	3,33	[3,29 à 3,36]	+0,09	[0,04 à 0,13]
ET d'être né homme	+0,24	[0,19 à 0,29]	+0,16	[0,12 à 0,20]	**(-0,07)**	**[-0,13 à -0,02]**

Chapter 15

Exemple 4 - X quantitatif

Les articles qui se consacrent aux interactions présentent souvent des méthodes applicables lorsque les deux expositions X et V sont binaires. Or, en épidémiologie, les expositions peuvent aussi être continues et, si dichotomiser ces variables peut simplifier l'approche de l'interaction, cela conduit à une perte d'information qui n'est pas souhaitable et pose la question complexe du choix des seuils Royston et al. [2006] Knol et al. [2007] Cadarso-Suárez et al. [2006].

Nous présentons ici un exemple dans lequel l'une des expositions, l'âge, est analysée en tant que variable quantitative continue.

15.1 Formuler les objectifs

Dans cette étude fictive, on s'est intéressé à la consommation de cannabis : comment le fait d'avoir déjà fumé du cannabis Y varie avec l'âge A et le sexe S.

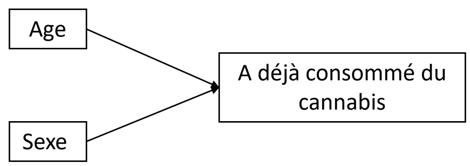
La démarche est explicative : on cherche à comprendre les mécanismes causaux de ce comportement de santé.

Ici, on adoptera une démarche d'analyse d'interaction $do(S, A)$

15.2 Stratégies et méthodes

Le DAG (sans les médiateurs) était :

Sexe	Age	P(Cannabis), %
Male	20-	51,1
Male	]20 à 40]	66,3
Male	]40 à 60]	40,4
Male	60+	12,1
Female	20-	44,2
Female	]20 à 40]	52,7
Female	]40 à 60]	26,7
Female	60+	12,1



Les estimands était définis par :

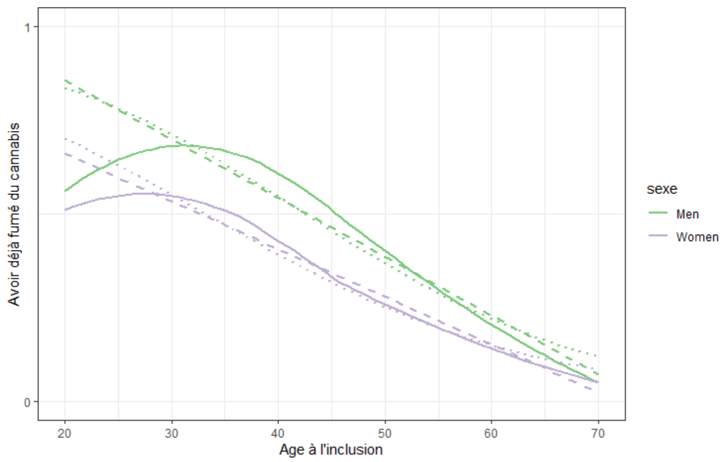
- L'effet de l'âge ("avoir 10 ans de plus") chez les hommes
  - $DR = Y_{S=0,A=a+10} - Y_{S=0,A=a}$
  - $RR = \frac{Y_{S=0,A=a+10}}{Y_{S=0,A=a}}$
- L'effet de l'âge ("avoir 10 ans de plus") chez les femmes :
  - $DR = Y_{S=1,A=a+10} - Y_{S=1,A=a}$
  - $RR = \frac{Y_{S=1,A=a+10}}{Y_{S=1,A=a}}$
- L'effet d'interaction entre l'âge et le sexe (l'effet du sexe est-il différent en fonction de l'âge et l'effet de l'âge est-il différent en fonction du sexe ?)
  - sur l'échelle additive : $AI = Y_{S=1,A=a+10} - Y_{S=0,A=a+10} - Y_{S=1,A=a} + Y_{S=0,A=a+10}$
  - sur l'échelle multiplicative : $MI = \frac{Y_{S=1,A=a+10} \times Y_{S=0,A=a}}{Y_{S=1,A=a} \times Y_{S=0,A=a+10}}$

15.3 Analyse descriptive

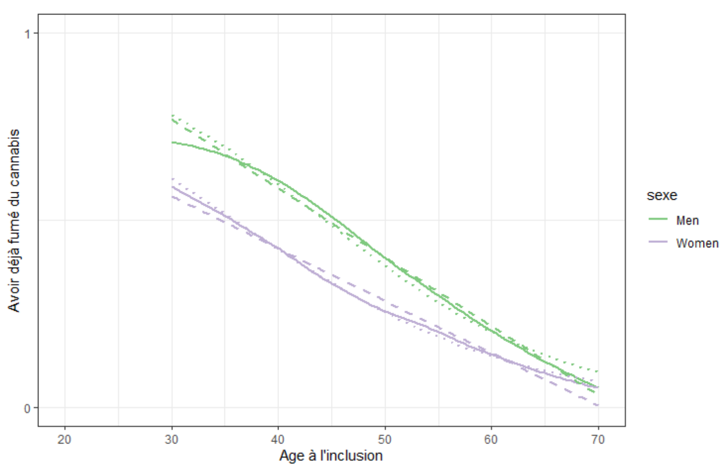
Dans cette population (N=202 768), il y avait 53,7% d'hommes et la moyenne d'âge était de 47,1 ans.

On peut commencer par décrire la proportion de personnes ayant déjà fumé du cannabis par sexe et classe d'âge :

Il semble y avoir une interaction entre l'âge et le sexe sur la probabilité d'avoir déjà fumé du cannabis. Cependant, la relation entre l'âge et l'outcome ne semble pas linéaire, ce qui est confirmé graphiquement :



Pour simplifier les analyses, nous n'allons inclure que les plus de 30 ans (N = 177 940), pour lesquels la relation est linéaire :



Le modèle de régression logistique (···) semble être plus proche de la modélisation non paramétrique sur données observées (loess, ——) que la modélisation linaire (— — —) . D'ailleurs, le R^2 du modèle logistique est de 0,168 contre 0,139 pour le modèle linéaire.

15.4 Analyse exploratoire

15.4.1 Régression logistique

L'outcome étant binaire, il est plus classique d'utiliser un modèle logistique, dont les résultats seraient :

```
# Call:
# glm(formula = cannabis ~ sexe + age + sexe * age, family = binomial,
#     data = data)
#
# Coefficients:
#                  Estimate Std. Error z value Pr(>|z|)
# (Intercept)     3.9144609  0.0372560  105.07   <2e-16 ***
# sexeWomen      -1.1644706  0.0511834  -22.75   <2e-16 ***
# age            -0.0882928  0.0007566 -116.70   <2e-16 ***
# sexeWomen:age   0.0117238  0.0010623   11.04   <2e-16 ***
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ce qui, en terme d'OR, donnerait :

```
#                       OR      2.5 %     97.5 %
# (Intercept)   50.1220409 46.5985990 53.9259952
# sexeWomen      0.3120878  0.2822910  0.3450107
# age            0.9154927  0.9141333  0.9168485
# sexeWomen:age  1.0117927  1.0096884  1.0139018
```

Les modèles de régression logistique donnent des résulats sur l'échelle multiplicative :

- L'effet du sexe (d'être femme plutot que homme) est :
  - "A 0 ans" (à l'origine) : $OR(S|A=0) = \times 0.31$
  - "A 1 ans" : $OR(S|A=1) = exp(-1,164+0,012)\times$ 0.32
  - A 40 ans (par exemple) : $OR(S|A=40) = exp(-1,164+0,012\times 40) = \times$ 0.5
- L'effet de l'age est :
  - Quand on est un homme : $OR(A|S=0) = exp(-0,088\times 10) = \times$ 0.41 par 10 ans
  - Quand on est une femme : $OR(A|S=1) = exp(-0,088\times 10 + 0,012\times 10) = \times$ 0.47 par 10 ans
- L'effet d'être une femme et d'avoir 10 ans de plus

- plutot que homme "et 0 ans"
- $OR(A, S) = exp(-1,164 - 0,088 \times 10 + 0,012 \times 10) = \times$ 0.15

- **L'effet d'interaction/modification d'effet** est :
  - $MI = \times 1,01$ sur 1 an
  - $MI_{10} = exp(0,012 \times 10) = \times$ 1.13 sur 10 ans
- **Un effet d'interaction additif**
  - $RERI_{OR} = OR_{11} - OR_{01} - OR_{10} + 1 =$ 0.047 pour 1 ans
  - $RERI_{OR,10} =$ 0.362

On a donc une interaction multiplicative positive (MI>1) et significative et une interaction additive aussi positive (RERI >0).

15.4.2 Régression linéaire

La sortie d'une modèle linéaire simple serait :

```
#
# Call:
# lm(formula = cannabis ~ sexe + age + sexe * age, data = data)
#
# Coefficients:
#                  Estimate Std. Error t value Pr(>|t|)
# (Intercept)    1.3197103  0.0066820  197.50   <2e-16 ***
# sexeWomen     -0.3373482  0.0091573  -36.84   <2e-16 ***
# age           -0.0183730  0.0001294 -142.03   <2e-16 ***
# sexeWomen:age  0.0044248  0.0001781   24.85   <2e-16 ***
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

On peut en déduire, ici sur une échelle additive, que :

- L'effet du sexe (d'être femme plutot que homme) est :
  - "A 0 ans" (à l'origine) : $DR(S|A = 0) = -33,73\%$
  - A 20 ans (par exemple) : $DR(S|A = 20) = -33,73 + 0,44 \times 20 =$ -24.93%
  - A 40 ans (par exemple) : $DR(S|A = 40) = -33,73 + 0,44 \times 40 =$ -16.13%
  - A 60 ans (par exemple) : $DR(S|A = 60) = -33,73 + 0,44 \times 60 =$ -7.33%
- L'effet de l'age est :
  - Quand on est un homme : $DR(A|S = 0) = -1,84 \times 10 =$ -18.4% par 10 ans

  - Quand on est une femme : $DR(A|S = 1) = -1,84 \times 10 + 0.44 \times 10 =$ -14 % par année d'âge
- L'effet d'être une femme et d'avoir 10 ans de plus
  - plutot que homme et un certain âge
  - $DR(A, S) = -33,73 - 1,84 \times 10 + 0,44 \times 10 =$ -47.73 %
- **L'effet d'interaction/modification d'effet** est :
  - $AI = +0.44$'%
  - $AI_{10} = +0.44 \times 10 = 4.4\%$

On retrouve une interaction additive significative et positive. Les expositions ayant un effet négatif et l'effet d'interaction étant positif, cet effet est difficile à interpréter, mais on pourrait le formuler plus simplement en changeant la catégorie de référence du sexe de homme à femme.

Ainsi : globalement, la probabilité d'avoir déjà fumer du cannabis diminue avec l'âge chez les hommes (-1,8% par an) et chez les femmes (-1,4% par an). Cette probabilité est plus élevée chez les hommes (de 16% par exemple à 40 ans), mais cet écart diminue avec l'âge, de 4,4% tous les 10 ans.

15.4.3 Effets prédits

A partir des modèles, on peut déduire les effets prédits pour certaines catégories. Par exemple, avec le modèle logistique :

- $Y_{S=0,A=30} = \frac{exp(3,914-0,088\times 30)}{1+exp(3,914-0,088\times 30)} = 78.1\%$
- $Y_{S=0,A=50} = \frac{exp(3,914-0,088\times 50)}{1+exp(3,914-0,088\times 50)} = 38.1\%$
- $Y_{S=1,A=30} = \frac{exp(3,914-1,164-0,088\times 30+0,012\times 30)}{1+exp(3,914-1,164-0,088\times 30+0,012\times 30)} = 61.5\%$
- $Y_{S=1,A=50} = \frac{exp(3,914-1,164-0,088\times 50+0,012\times 50)}{1+exp(3,914-1,164-0,088\times 50+0,012\times 50)} = 25.9\%$

Avec le modèle linéaire, on aurait :

- $Y_{S=0,A=30} = 131,97 - 1,84 \times 30 = 76.8\%$
- $Y_{S=0,A=50} = 131,97 - 1,84 \times 50 = 40\%$
- $Y_{S=1,A=30} = 131,97 - 33,73 - 1,84 \times 30 + 0,44 \times 30 = 56.2\%$
- $Y_{S=1,A=50} = 131,97 - 33,73 - 1,84 \times 50 + 0,44 \times 50 = 28.2\%$

15.5 Analyse confirmatoire

Nous avons calculé les effets d'intérêt avec une méthode de modèle structurel marginal (Intervalles de confiance estimé par bootstrap, 200 répétitions), le modèle utilisé pour prédire les outcomes contrefactuels étaient un modèle de régression logistique.

Le code était :

```
B=200

simu.base <- data.frame(i.simu=c(1:B))

for (i in 1:B){
  # sample the indices 1 to n with replacement
  bootIndices <- sample(1:nrow(data), replace=T) ;    set.seed(01062023+i*12)
  bootData <- data[bootIndices,]

  #modèle
  Q.model <- glm(data=bootData, formula = cannabis ~ sexe+
                   age+  sexe*age,family = binomial)

  # Scénarios #
  data.S1 <-  data.S2 <- bootData
  data.S1$sexe <- "Women"
  data.S2$sexe <- "Men"
  data.S1A30 <-  data.S1A40 <- data.S1A50 <- data.S1A60 <- data.S1A70 <- data.S1
  data.S1A35 <-  data.S1A45 <- data.S1A55 <- data.S1A65 <- data.S1
  data.S2A30 <-  data.S2A40 <- data.S2A50 <- data.S2A60 <- data.S2A70 <- data.S2
  data.S2A35 <-  data.S2A45 <- data.S2A55 <- data.S2A65 <- data.S2
  data.S1A30$age <- data.S2A30$age <- 30
  data.S1A35$age <- data.S2A35$age <- 35
  data.S1A40$age <- data.S2A40$age <- 40
  data.S1A45$age <- data.S2A45$age <- 45
  data.S1A50$age <- data.S2A50$age <- 50
  data.S1A55$age <- data.S2A55$age <- 55
  data.S1A60$age <- data.S2A60$age <- 60
  data.S1A65$age <- data.S2A65$age <- 65
  data.S1A70$age <- data.S2A70$age <- 70


  # Y contrefactuel
  Y.S1A30.pred <- predict(Q.model, newdata = data.S1A30, type = "response")
  Y.S1A40.pred <- predict(Q.model, newdata = data.S1A40, type = "response")
  Y.S1A50.pred <- predict(Q.model, newdata = data.S1A50, type = "response")
  Y.S1A60.pred <- predict(Q.model, newdata = data.S1A60, type = "response")
  Y.S1A70.pred <- predict(Q.model, newdata = data.S1A70, type = "response")
  Y.S2A30.pred <- predict(Q.model, newdata = data.S2A30, type = "response")
  Y.S2A40.pred <- predict(Q.model, newdata = data.S2A40, type = "response")
  Y.S2A50.pred <- predict(Q.model, newdata = data.S2A50, type = "response")
```

```r
  Y.S2A60.pred <- predict(Q.model, newdata = data.S2A60, type = "response")
  Y.S2A70.pred <- predict(Q.model, newdata = data.S2A70, type = "response")

  Y.S1A35.pred <- predict(Q.model, newdata = data.S1A35, type = "response")
  Y.S1A45.pred <- predict(Q.model, newdata = data.S1A45, type = "response")
  Y.S1A55.pred <- predict(Q.model, newdata = data.S1A55, type = "response")
  Y.S1A65.pred <- predict(Q.model, newdata = data.S1A65, type = "response")
  Y.S2A35.pred <- predict(Q.model, newdata = data.S2A35, type = "response")
  Y.S2A45.pred <- predict(Q.model, newdata = data.S2A45, type = "response")
  Y.S2A55.pred <- predict(Q.model, newdata = data.S2A55, type = "response")
  Y.S2A65.pred <- predict(Q.model, newdata = data.S2A65, type = "response")

  Y <- c(Y.S1A30.pred, Y.S1A40.pred, Y.S1A50.pred, Y.S1A60.pred, Y.S1A70.pred,
         Y.S1A35.pred, Y.S1A45.pred, Y.S1A55.pred, Y.S1A65.pred,
         Y.S2A30.pred, Y.S2A40.pred, Y.S2A50.pred, Y.S2A60.pred, Y.S2A70.pred,
         Y.S2A35.pred, Y.S2A45.pred, Y.S2A55.pred, Y.S2A65.pred)

# On récupère les valeurs d'exposition qui ont servi dans les scénarios contrefac
# (garder le même ordre que pour les Y.A1.A2)

X <- rbind(subset(data.S1A30, select = c("sexe", "age")),
           subset(data.S1A40, select = c("sexe", "age")),
           subset(data.S1A50, select = c("sexe", "age")),
           subset(data.S1A60, select = c("sexe", "age")),
           subset(data.S1A70, select = c("sexe", "age")),
           subset(data.S1A35, select = c("sexe", "age")),
           subset(data.S1A45, select = c("sexe", "age")),
           subset(data.S1A55, select = c("sexe", "age")),
           subset(data.S1A65, select = c("sexe", "age")),
           subset(data.S2A30, select = c("sexe", "age")),
           subset(data.S2A40, select = c("sexe", "age")),
           subset(data.S2A50, select = c("sexe", "age")),
           subset(data.S2A60, select = c("sexe", "age")),
           subset(data.S2A70, select = c("sexe", "age")),
           subset(data.S2A35, select = c("sexe", "age")),
           subset(data.S2A45, select = c("sexe", "age")),
           subset(data.S2A55, select = c("sexe", "age")),
           subset(data.S2A65, select = c("sexe", "age")))

## Modèle structurel marginal
   # logistique
    msm.glm <- glm(Y ~ age + sexe + age:sexe,
                  data = data.frame(Y,X),
                  family = "binomial")
   #linéaire pour l'interaction additive
```

```
      msm.lm <- glm(Y ~ age + sexe + age:sexe,
                    data = data.frame(Y,X),
                    family = "gaussian")

   # Tous les effets
      simu.base$est.Y0_30[simu.base$i.simu==i] = round(mean(Y.S2A30.pred),4)
      simu.base$est.Y0_40[simu.base$i.simu==i] = round(mean(Y.S2A40.pred),4)
      simu.base$est.Y0_50[simu.base$i.simu==i] = round(mean(Y.S2A50.pred),4)
      simu.base$est.Y0_60[simu.base$i.simu==i] = round(mean(Y.S2A60.pred),4)
      simu.base$est.Y0_70[simu.base$i.simu==i] = round(mean(Y.S2A70.pred),4)
      simu.base$est.Y1_30[simu.base$i.simu==i] = round(mean(Y.S1A30.pred),4)
      simu.base$est.Y1_40[simu.base$i.simu==i] = round(mean(Y.S1A40.pred),4)
      simu.base$est.Y1_50[simu.base$i.simu==i] = round(mean(Y.S1A50.pred),4)
      simu.base$est.Y1_60[simu.base$i.simu==i] = round(mean(Y.S1A60.pred),4)
      simu.base$est.Y1_70[simu.base$i.simu==i] = round(mean(Y.S1A70.pred),4)

      simu.base$est.RD_30[simu.base$i.simu==i] = round(mean(Y.S1A30.pred - Y.S2A30.pred),4)
      simu.base$est.RD_40[simu.base$i.simu==i] = round(mean(Y.S1A40.pred - Y.S2A40.pred),4)
      simu.base$est.RD_50[simu.base$i.simu==i] = round(mean(Y.S1A50.pred - Y.S2A50.pred),4)
      simu.base$est.RD_60[simu.base$i.simu==i] = round(mean(Y.S1A60.pred - Y.S2A60.pred),4)
      simu.base$est.RD_70[simu.base$i.simu==i] = round(mean(Y.S1A70.pred - Y.S2A70.pred),4)

      simu.base$est.RR_30[simu.base$i.simu==i] = round(mean(Y.S1A30.pred / Y.S2A30.pred),4)
      simu.base$est.RR_40[simu.base$i.simu==i] = round(mean(Y.S1A40.pred / Y.S2A40.pred),4)
      simu.base$est.RR_50[simu.base$i.simu==i] = round(mean(Y.S1A50.pred / Y.S2A50.pred),4)
      simu.base$est.RR_60[simu.base$i.simu==i] = round(mean(Y.S1A60.pred / Y.S2A60.pred),4)
      simu.base$est.RR_70[simu.base$i.simu==i] = round(mean(Y.S1A70.pred / Y.S2A70.pred),4)

      simu.base$est.RD_Sm[simu.base$i.simu==i] = round(msm.lm$coefficients["age"]*10,4)
      simu.base$est.RR_Sm[simu.base$i.simu==i] = round(exp(msm.glm$coefficients["age"]*10),4)
      simu.base$est.RD_Sw[simu.base$i.simu==i] = round(msm.lm$coefficients["age"]*10 +
                                                       msm.lm$coefficients["age:sexeWomen"]*10
      simu.base$est.RR_Sw[simu.base$i.simu==i] = round(exp(msm.glm$coefficients["age"]*10 +
                                                           msm.glm$coefficients["age:sexeWome

      simu.base$est.AI[simu.base$i.simu==i] = round(msm.lm$coefficients["age:sexeWomen"]*10,4)
      simu.base$est.MI[simu.base$i.simu==i] = round(exp(msm.glm$coefficients["age:sexeWomen"]*
      simu.base$est.RERI[simu.base$i.simu==i] = round(exp(msm.glm$coefficients["age"]*10 +
                                                          msm.glm$coefficients["sexeWomen"]
                                                          msm.glm$coefficients["age:sexeWome
                                                    exp(msm.glm$coefficients["age"]*10 +
                                                          msm.glm$coefficients["age:sexeWome
                                                      exp(msm.glm$coefficients["sexeWomen"]
                                                            msm.glm$coefficients["age:sexeW
```

	Sex = Men	Sex = Women	RD within strata of Ag
Age = 30	0.62 [0.61 to 0.62]	0.78 [0.78 to 0.79]	-0.16 [-0.17 to -0.16]
Age = 40	0.42 [0.42 to 0.43]	0.6 [0.59 to 0.6]	-0.17 [-0.18 to -0.17]
Age = 50	0.25 [0.25 to 0.25]	0.38 [0.37 to 0.38]	-0.13 [-0.13 to -0.12]
Age = 60	0.13 [0.13 to 0.13]	0.2 [0.19 to 0.2]	-0.07 [-0.07 to -0.06]
Age = 70	0.06 [0.06 to 0.07]	0.09 [0.09 to 0.09]	-0.03 [-0.03 to -0.02]
RD (10 y) within strata of Sex	-0.18 [-0.18 to -0.18]	-0.14 [-0.14 to -0.14]	NA
OR (10 y) within strata of Sex	0.41 [0.4 to 0.41]	0.45 [0.45 to 0.46]	NA

[a] Additive interaction (10 years) =0.04 [0.04 to 0.04]
[b] Multiplicative Interaction (10 years) =1.11 [1.09 to 1.13]
[c] RERI (10 years) =0.33 [0.32 to 0.34]

```
        }


    effect <- round(colMeans(simu.base),2)
    confint <- apply(simu.base, 2, function(x) round(quantile(x,probs = c(0.025, 0.975
    tab_all <- as.data.frame(rbind(effect,confint))
```

Au final, les résultats étaient :

On retrouve :

- une interaction additive significative et positive : l'écart entre les hommes et les femmes diminue avec l'âge, de 4% tous les 10 ans, ou l'effet d'avoir 10 ans est plus faible de 4% chez les hommes par rapport au femmes. Le RERI est aussi positif et significatif (l'OR augmente de 33% tous les 10 ans).
- une interaction multiplicative significative et positive : l'effet d'être un homme plutôt qu'une femme sur le risque d'avoir consommer du cannabis est moins fort quand l'âge augmente, ou l'effet d'avoir 10 ans est multiplié par 1,11 chez les femmes par rapport aux hommes

Part IV

Conclusion

Chapter 16

Synthèse générale

La première étape importantes consiste à **définir précisément l'objectif**. Et, si l'on est dans une démarche explicative, d'inférence causale, il s'agit de définir si la mesure d'un effet d'interaction est nécessaire pour y répondre (identifier précisément l'effet que l'on cherche à estimer, ou *estimand*).

Le fait de choisir une **démarche d'analyse d'interaction ou de modification d'effet** repose sur :

- la façon dont la question est posée (effet de X selon V ou effet conjoint de X et V),
- sur les hypothèses causales formulées (scénarii $do(X)$ ou $do(X, V)$)
- et donc sur les sets de facteurs de confusion à considérer (seulement sur $X \rightarrow Y$ ou $X.V \rightarrow Y$).

Concernant le **choix de l'échelle**, idéalement, les interactions devraient être reportées sur les 2 échelles Knol and VanderWeele [2012] VanderWeele and Knol [2014]. Cependant, l'échelle additive est plus appropriée pour évaluer l'utilité en santé publique VanderWeele and Knol [2014] Knol and VanderWeele [2012].

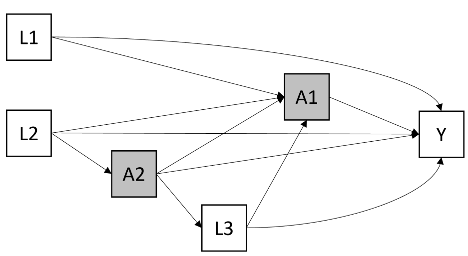
Concernant les paramètres,

Chapter 17

Pour aller plus loin...

17.1 Ajouter de la complexité

A1 et A2 sont rarement indépendants. Scénario plus probable :



17.2 Interaction avec confusion intermédiaire

17.3 Interaction et médiation

VanderWeele [2013]

VanderWeele [2014]

Chapter 18

Références

Bibliography

Sara T. Brookes, Elise Whitely, Matthias Egger, George Davey Smith, Paul A. Mulheran, and Tim J. Peters. Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *Journal of Clinical Epidemiology*, 47:229–236, 2004.

Carmen Cadarso-Suárez, Javier Roca-Pardiñas, and Adolfo Figueiras. Effect measures in non-parametric regression with interactions between continuous exposures. *Statistics in Medicine*, 25(4):603–621, 2006.

Hélène Colineaux, Lola Neufcourt, Cyrille Delpierre, Michelle Kelly-Irving, and Benoit Lepage. Explaining biological differences between men and women by gendered mechanisms. *Emerging Themes in Epidemiology*, 20(1):2, 2023.

Priscila Corraini, Morten Olsen, Lars Pedersen, Olaf M Dekkers, and Jan P Vandenbroucke. Effect modification, interaction and mediation: an overview of theoretical insights for clinical investigators. *Clinical Epidemiology*, 9:331–338, June 2017. ISSN 1179-1349. doi: 10.2147/CLEP.S129728.

Rhian Daniel, Jingjing Zhang, and Daniel Farewell. Making apples from oranges: Comparing noncollapsible effect estimators and their standard errors after adjustment for different covariate sets. *Biometrical Journal*, 63(3):528–557, 2021. doi: 10.1002/bimj.201900297.

Miguel A Hernán and James M Robins. *Causal Inference: What If - PREPRINT*. Chapman & Hall/CRC Boca Raton, FL, 2020.

Miguel A Hernán, John Hsu, and Brian Healy. A second chance to get causal inference right: a classification of data science tasks. *Chance*, 32(1):42–49, 2019.

Mirjam J. Knol and Tyler J. VanderWeele. Recommendations for presenting analyses of effect modification and interaction. *International Journal of Epidemiology*, 41(2):514–520, April 2012. ISSN 1464-3685. doi: 10.1093/ije/dyr218.

Mirjam J Knol, Ingeborg van der Tweel, Diederick E Grobbee, Mattijs E Numans, and Mirjam I Geerlings. Estimating interaction on an additive scale between continuous determinants in a logistic regression model. *International journal of epidemiology*, 36(5):1111–1118, 2007.

Maya B Mathur and Tyler J VanderWeele. R function for additive interaction measures. *Epidemiology (Cambridge, Mass.)*, 29(1):e5, 2018.

Patrick Royston, Douglas G Altman, and Willi Sauerbrei. Dichotomizing continuous predictors in multiple regression: a bad idea. *Statistics in medicine*, 25(1):127–141, 2006.

Tyler J. VanderWeele. On the distinction between interaction and effect modification. *Epidemiology (Cambridge, Mass.)*, 20(6):863–871, November 2009. ISSN 1531-5487. doi: 10.1097/EDE.0b013e3181ba333c.

Tyler J. VanderWeele. A Three-way Decomposition of a Total Effect into Direct, Indirect, and Interactive Effects. *Epidemiology*, 24(2):224–232, March 2013. ISSN 1044-3983. doi: 10.1097/EDE.0b013e318281a64e.

Tyler J. VanderWeele. A unification of mediation and interaction: a 4-way decomposition. *Epidemiology (Cambridge, Mass.)*, 25(5):749–761, September 2014. ISSN 1531-5487. doi: 10.1097/EDE.0000000000000121.

Tyler J. VanderWeele. The Interaction Continuum. *Epidemiology*, 30(5):648–658, September 2019. ISSN 1044-3983. doi: 10.1097/EDE.0000000000001054.

Tyler J. VanderWeele and Mirjam J. Knol. A Tutorial on Interaction. *Epidemiologic Methods*, 3(1):33–72, December 2014. ISSN 2161-962X. doi: 10.1515/em-2013-0005. Publisher: De Gruyter.

Tyler J. VanderWeele and James M. Robins. Four types of effect modification: a classification based on directed acyclic graphs. *Epidemiology (Cambridge, Mass.)*, 18(5):561–568, September 2007. ISSN 1044-3983. doi: 10.1097/EDE.0b013e318127181b.

Brian W Whitcomb and Ashley I Naimi. Defining, quantifying, and interpreting ”noncollapsibility” in epidemiologic studies of measures of ”effect”. *American Journal of Epidemiology*, 190(5):697–700, 2021. doi: 10.1093/aje/kwaa267.