

RECH 801

M1 MEEF NSI

Variables de recherche

Benjamin Wack

à partir de matériel de <https://explorable.com>



2023 – 2024

Avertissement

Les notions présentées ici constituent une **grille de construction** d'un protocole expérimental :

- ▶ Le cadre présenté est celui d'une étude « idéale ».
- ▶ Dans la pratique, il n'est pas toujours possible de s'y conformer complètement.
- ▶ Ce cadre doit aussi permettre d'identifier les limites de validité de l'expérience réalisée.

Plan

Construire un protocole expérimental

Variation de recherche

Causalité et corrélation

Randomisation

Construire un protocole expérimental

Par le biais d'une expérience, le chercheur tente d'apprendre **quelque chose de nouveau** sur le monde, d'expliquer « *pourquoi* » quelque chose se produit.

- ▶ La réalisation d'une expérience comporte plusieurs **étapes**, au cours desquelles les paramètres et la structure de l'expérience sont clairement définis.
- ▶ Suivre ces étapes de base garantit en général d'obtenir des résultats **valides**.
- ▶ Bien que tous les scientifiques utilisent le raisonnement, l'opérationnalisation et les étapes du processus scientifique, ce n'est **pas toujours un processus explicite**.

Limites et transparence

En pratique, il n'est pas toujours possible de suivre strictement chaque étape. **Toute aberration doit alors être justifiée**, qu'elle soit due au budget, à l'impraticabilité ou à l'éthique.

Hypothèses

Hypothèse nulle

Explication communément admise, que le chercheur essaye de **réfuter**.

- ▶ consiste à postuler une *absence d'effet* : deux groupes ayant des valeurs différentes pour un paramètre auront des résultats similaires
- ▶ est toujours testée dans une expérience, même indirectement

Hypothèse de recherche

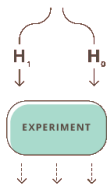
- ▶ postule une explication (différente) pour le phénomène observé
- ▶ ne sera **jamais « prouvée »** (uniquement vérifiée/confirmée)

Hypothèse de recherche : « Les tomates poussent plus vite dans le compost que dans la terre ».

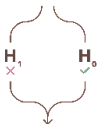
Hypothèse nulle : « Les tomates poussent à la même vitesse dans le compost ou dans la terre ».

Hypothèses : exemples

Research question:
DOES EATING ONLY FISH
INCREASE IQ?



Result:
NO SIGNIFICANT DIFFERENCE
BETWEEN GROUPS



Conclusion:
ONLY EATING FISH
DOES NOT INCREASE IQ

Les hypothèses suivantes sont-elles pertinentes ?
Sinon, comment pourraient-elles être améliorées ?

1. L'ajout de mica à un composé plastique diminuera sa viscosité.
2. Ceux qui boivent une tasse de thé vert tous les jours ressentent une amélioration de leur bien-être.
3. Le fait de regarder longuement les éclipses solaires confère des pouvoirs extrasensoriels.
4. Le déclin des valeurs familiales fait baisser le taux de mariages.
5. Les enfants manquant de sécurité affective sont plus susceptibles de s'engager dans la dissidence politique à l'âge adulte.
6. L'Afrique subsaharienne connaît plus de décès dus à la tuberculose, car le taux de SIDA y est plus élevé.

Construire un protocole expérimental (I)

Groupes d'échantillons

Après avoir choisi une hypothèse, on spécifie les groupes d'échantillons :

- ▶ suffisamment grands pour donner une étude statistiquement viable ;
- ▶ suffisamment petits pour être pratiques ;
- ▶ idéalement, sélectionnés au hasard parmi la population.

(Assez facile en sciences physiques, moins en biologie ou en sciences comportementales.)

Exemples

Les essais médicaux ne permettent pas souvent de trouver des groupes aléatoires car ils reposent sur

Construire un protocole expérimental (I)

Groupes d'échantillons

Après avoir choisi une hypothèse, on spécifie les groupes d'échantillons :

- ▶ suffisamment grands pour donner une étude statistiquement viable ;
- ▶ suffisamment petits pour être pratiques ;
- ▶ idéalement, sélectionnés au hasard parmi la population.

(Assez facile en sciences physiques, moins en biologie ou en sciences comportementales.)

Exemples

Les essais médicaux ne permettent pas souvent de trouver des groupes aléatoires car ils reposent sur des volontaires.

Dans un tel cas les résultats

Construire un protocole expérimental (I)

Groupes d'échantillons

Après avoir choisi une hypothèse, on spécifie les groupes d'échantillons :

- ▶ suffisamment grands pour donner une étude statistiquement viable ;
- ▶ suffisamment petits pour être pratiques ;
- ▶ idéalement, sélectionnés au hasard parmi la population.

(Assez facile en sciences physiques, moins en biologie ou en sciences comportementales.)

Exemples

Les essais médicaux ne permettent pas souvent de trouver des groupes aléatoires car ils reposent sur des volontaires.

Dans un tel cas les résultats ne doivent pas être appliqués à l'ensemble de la population. Ainsi, si les volontaires sont des étudiants masculins de 18 à 24 ans, les résultats ne peuvent être généralisés qu'à ce groupe démographique.

Construire un protocole expérimental (II)

Groupe contrôle et groupe test

Afin de réduire la possibilité de variables confondantes, on constitue :

- ▶ un groupe contrôle
- ▶ et un groupe expérimental.

La répartition des sujets dans les groupes doit :

- ▶ être aléatoire
- ▶ se faire en aveugle ou en double aveugle

ce qui réduira les risques de biais.

Obstacles

- ▶ Souvent éthiques (refus délibéré d'un traitement, cf étude de Tuskegee)
- ▶ En cas de compromis sur le caractère aléatoire, expliquer comment et pourquoi les groupes ont été sélectionnés.

Construire un protocole expérimental (III)

Organisation dans le temps

Déterminer :

- ▶ une échelle de temps (durée)
- ▶ une fréquence d'échantillonnage

adaptées à l'expérience.

Exemples

Étude de l'efficacité d'un remède :

- ▶ contre le rhume :

Construire un protocole expérimental (III)

Organisation dans le temps

Déterminer :

- ▶ une échelle de temps (durée)
- ▶ une fréquence d'échantillonnage

adaptées à l'expérience.

Exemples

Étude de l'efficacité d'un remède :

- ▶ contre le rhume : échantillons fréquents, sur une période de plusieurs jours
- ▶ contre la maladie de Parkinson :

Construire un protocole expérimental (III)

Organisation dans le temps

Déterminer :

- ▶ une échelle de temps (durée)
- ▶ une fréquence d'échantillonnage

adaptées à l'expérience.

Exemples

Étude de l'efficacité d'un remède :

- ▶ contre le rhume : échantillons fréquents, sur une période de plusieurs jours
- ▶ contre la maladie de Parkinson : tests plus espacés, sur une période de plusieurs mois ou années.

Construire un protocole expérimental (IV)

Réalisation

Expliciter :

- ▶ comment la variable indépendante est manipulée.
- ▶ quelles données (variables dépendantes) sont recueillies ;
- ▶ les méthodes de recueil des données ;

Construire un protocole expérimental (V)

Analyse

Analyse des données brutes issues des résultats par des moyens **statistiques** :

- ▶ établir s'il existe une relation entre les variables ;
- ▶ accepter ou rejeter l'hypothèse nulle (celle que l'on cherche à réfuter) ;
- ▶ éventuellement proposer des hypothèses alternatives.

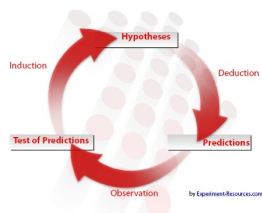


Schéma de coordination de l'ensemble du plan de recherche

- ▶ aide à planifier la recherche ;
- ▶ renforce la lisibilité du protocole ;
- ▶ peut permettre d'identifier la nécessité de décomposer les variables complexes en variables simples.

À vous

- ▶ Quelle est votre population ?
- ▶ Quels sont vos groupes d'échantillons ?
- ▶ Avez-vous un groupe contrôle ?

Plan

Construire un protocole expérimental

Variables de recherche

Causalité et corrélation

Randomisation

Variables de recherche

- ▶ peuvent prendre différentes valeurs
- ▶ devraient pouvoir être mesurées
- ▶ peuvent être
 - ▶ qualitatives ou quantitatives
 - ▶ ordinales ou non
 - ▶ discrètes ou continues
 - ▶ intensives ou extensives
 - ▶ ...

La température

- ▶ dépend de différents facteurs (pression, sources de chaleur...)
- ▶ est mesurable en différents points et instants
- ▶ est numérique, continue et intensive

Variables dépendantes, indépendantes, confondantes

Dans la recherche, on définit généralement les variables en fonction de ce que l'on mesure.

C'est au **chercheur** de **décider** quels éléments seront soit des variables dépendantes, soit des variables indépendantes.

- ▶ La **variable indépendante** est la variable que le chercheur souhaite mesurer (la cause).
- ▶ La **variable dépendante** est l'effet (ou l'effet supposé), dépendant de la variable indépendante.

Ces variables sont souvent énoncées dans l'hypothèse, par exemple « quel est l'effet de la personnalité sur le comportement d'aide ».

Variables dépendantes, indépendantes, confondantes

Dans la recherche, on définit généralement les variables en fonction de ce que l'on mesure.

C'est au **chercheur** de **décider** quels éléments seront soit des variables dépendantes, soit des variables indépendantes.

- ▶ La **variable indépendante** est la variable que le chercheur souhaite mesurer (la cause).
- ▶ La **variable dépendante** est l'effet (ou l'effet supposé), dépendant de la variable indépendante.

Ces variables sont souvent énoncées dans l'hypothèse, par exemple « quel est l'effet de la personnalité sur le comportement d'aide ».

- ▶ Les **variables confondantes** sont des variables :
 - ▶ ayant un effet significatif sur la variable dépendante
 - ▶ mais que le chercheur n'a pas réussi à contrôler ou à éliminer parfois parce que le chercheur n'a pas conscience de cet effet.

Variable indépendante

- ▶ « indépendante » parce que sa variation ne dépend pas d'une autre variable de l'expérience
- ▶ **manipulée** (ou contrôlée) par le chercheur

C'est à la conception de la recherche que l'on décide quelles variables sont manipulées et lesquelles sont mesurées.

Effets d'un engrais sur la croissance des plantes

La variable indépendante est la quantité d'engrais utilisé.

Les éléments mesurables en tant que variable dépendantes pourraient être :

Variable indépendante

- ▶ « indépendante » parce que sa variation ne dépend pas d'une autre variable de l'expérience
- ▶ **manipulée** (ou contrôlée) par le chercheur

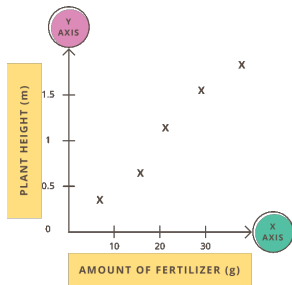
C'est à la conception de la recherche que l'on décide quelles variables sont manipulées et lesquelles sont mesurées.

Effets d'un engrais sur la croissance des plantes

La variable indépendante est la quantité d'engrais utilisé.

Les éléments mesurables en tant que variable dépendantes pourraient être :

- ▶ la taille de la plante ;
- ▶ le nombre de fruits ;
- ▶ le poids moyen des fruits produits...

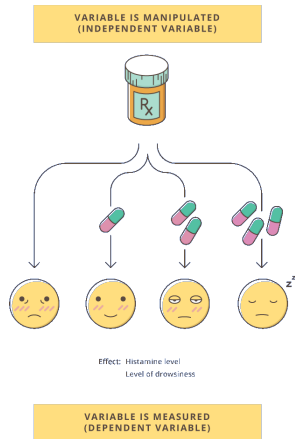


Autre exemple

Essai pharmaceutique

On cherche à déterminer la dose optimale pour un nouveau médicament (traitement antihistaminique).

- ▶ Variable indépendante = quantité de médicaments reçue
- ▶ Variable dépendante = gravité des symptômes d'allergie (yeux rouges, éternuements)



On pourrait également mesurer d'autres variables dépendantes : effets secondaires comme la somnolence.

Exemple - quizz

Dans une expérience sur l'efficacité d'un nouveau programme de formation :

Quelles peuvent être les variables de recherche ?

Exemple - quizz

Dans une expérience sur l'efficacité d'un nouveau programme de formation :

Quelles peuvent être les variables de recherche ?

- ▶ Résultat au test final
- ▶ Âge des participants
- ▶ Temps consacré au programme de formation
- ▶ Temps nécessaire pour terminer le test final
- ▶ Sexe des participants
- ▶ Évaluation du programme de formation par les participants

Pouvez-vous identifier la variable indépendante dans cette expérience ?

- ▶ Résultat au test final
- ▶ Temps consacré au programme de formation
- ▶ Participation au programme de formation

Expérience antiraciste de Jane Elliott

Jane Elliott divise sa classe en deux groupes :

- ▶ les enfants aux yeux bleus, à qui elle accorde des privilèges supplémentaires et souligne leur supériorité par rapport aux autres ;
- ▶ les enfants aux yeux bruns, qui constituent un « groupe minoritaire ».

Conséquences

- ▶ les enfants aux yeux bruns voient leur confiance en soi et leurs résultats scolaires baisser ;
- ▶ leur taux de harcèlement augmente.
- ▶ Cependant, lorsqu'elle qualifie plus tard le groupe des yeux bleus d'inférieur, ces effets s'inversent.

La variable indépendante est

Expérience antiraciste de Jane Elliott

Jane Elliott divise sa classe en deux groupes :

- ▶ les enfants aux yeux bleus, à qui elle accorde des privilèges supplémentaires et souligne leur supériorité par rapport aux autres ;
- ▶ les enfants aux yeux bruns, qui constituent un « groupe minoritaire ».

Conséquences

- ▶ les enfants aux yeux bruns voient leur confiance en soi et leurs résultats scolaires baisser ;
- ▶ leur taux de harcèlement augmente.
- ▶ Cependant, lorsqu'elle qualifie plus tard le groupe des yeux bleus d'inférieur, ces effets s'inversent.

La variable indépendante est le statut du groupe (privilegié ou non), **pas la couleur des yeux**.

Identifier une variable dépendante possible :

- ▶ taux de harcèlement
- ▶ résultats scolaires
- ▶ niveau de confiance

Combien de variables tester ?

- ▶ Normalement, pas plus d'une ou deux variables indépendantes (sinon difficile de déterminer l'influence de chacune)
- ▶ Il peut y avoir plusieurs variables dépendantes, car la manipulation de la variable indépendante peut influencer beaucoup de choses.

Dans le cas de plusieurs variables manipulées :

- ▶ Chacune constitue une expérience à part entière.
- ▶ Pour manipuler simultanément deux variables expérimentales, il faudrait en réalité **quatre** expériences distinctes.



À vous

- ▶ Quelle est votre variable indépendante ?
- ▶ Quelles sont vos variables dépendantes ?

Plan

Construire un protocole expérimental

Variation de recherche

Causalité et corrélation

Randomisation

Variable confondante

- ▶ influence à la fois la variable indépendante et la variable dépendante
- ▶ peut amener à une **analyse erronée**
(fausse corrélation entre les variables indépendantes et dépendantes)

Étude pour déterminer si les gros buveurs meurent à un âge plus précoce

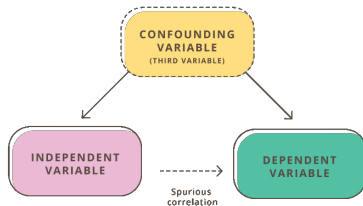
- ▶ Résultats : en effet, les personnes qui boivent excessivement sont susceptibles de mourir plus jeunes.
- ▶ Mais on constate que : leurs proches ou leurs collègues non buveurs meurent eux aussi plus tôt que la moyenne.

Si on ne considère que la consommation excessive d'alcool et la longévité, on en déduit une association fallacieuse.

Comment expliquer ces observations ?

L'erreur commise est dans la **conception de l'expérience** :

- ▶ les variables confondantes ne sont pas prises en compte
- ▶ ou on n'a pas essayé de les éliminer ou de les contrôler

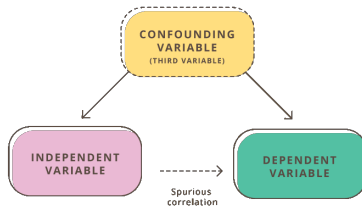


Variables confondantes possibles :

Comment expliquer ces observations ?

L'erreur commise est dans la **conception de l'expérience** :

- ▶ les variables confondantes ne sont pas prises en compte
- ▶ ou on n'a pas essayé de les éliminer ou de les contrôler



Variables confondantes possibles :

- ▶ Les gros buveurs peuvent être issus d'un groupe social particulier.
- ▶ Ce groupe pourrait, pour d'autres raisons (non mesurées), avoir une durée de vie plus courte que les autres.
(par exemple, une tendance au tabagisme ou à la malbouffe)
- ▶ Cet autre facteur peut influencer à la fois la consommation d'alcool et la longévité.

Dans tous les cas, c'est l'appartenance à ce groupe qui diminue leur longévité, et non la consommation excessive d'alcool.

Variable externe

= Variable que les chercheurs n'étudient pas **délibérément** et qui pourrait menacer la validité des résultats.

Une variable externe devient confondante lorsque...

elle varie en même temps que les variables qui vous intéressent réellement.
Dans un tel cas :

- ▶ elle pourrait aussi expliquer le résultat observé ;
- ▶ il devient donc difficile de distinguer quel effet provient de quelle variable.

On parle alors de **biais de confusion**, en particulier lorsque la variable confondante agit différemment sur les différents groupes comparés.

Consommation d'alcool et longévité

- ▶ Variables externes : âge, sexe, habitudes alimentaires, état civil...
- ▶ On pourrait contrôler l'âge en s'assurant que tous les participants à l'expérience ont le même âge.

Variables cachées

Un exemple

On constate que les fumeurs se rétablissent mieux suite à leur première attaque cardiaque que le reste de la population.

Peut-on en conclure qu'il vaut mieux fumer pour se prémunir des attaques cardiaques ?

Variables cachées

Un exemple

On constate que les fumeurs se rétablissent mieux suite à leur première attaque cardiaque que le reste de la population.

Peut-on en conclure qu'il vaut mieux fumer pour se prémunir des attaques cardiaques ?

Explication :

- ▶ les fumeurs ont tendance à faire leur première attaque cardiaque plus jeunes que le reste de la population
- ▶ ils sont donc en « meilleure forme » que les autres (malgré le tabac) et se remettent mieux

Variables cachées

Un exemple

On constate que les fumeurs se rétablissent mieux suite à leur première attaque cardiaque que le reste de la population.

Peut-on en conclure qu'il vaut mieux fumer pour se prémunir des attaques cardiaques ?

Explication :

- ▶ les fumeurs ont tendance à faire leur première attaque cardiaque plus jeunes que le reste de la population
- ▶ ils sont donc en « meilleure forme » que les autres (malgré le tabac) et se remettent mieux

Variable cachée : l'âge de la première crise cardiaque

- ▶ conditionnée par la variable *fumeur*
- ▶ qui conditionne le taux de rétablissement

Minimiser les effets des variables confondantes

Idéalement : en laboratoire, contrôle précis de l'environnement.

Hors laboratoire :

- ▶ bonne conception de l'expérience (par exemple randomisation)
- ▶ contrôles réguliers (groupes contrôles bien choisis)

Effet d'un antidépresseur à base de plantes

- ▶ On prélève un échantillon de participants à un groupe de médecine alternative en ligne.
- ▶ On leur demande de prendre le traitement pendant un mois.
- ▶ Les participants remplissent un questionnaire sur la dépression avant et après pour mesurer l'amélioration de leur humeur.
- ▶ On constate en effet que celle-ci est meilleure après traitement.

Pouvez-vous identifier des variables qui auraient pu fausser ce résultat ?

Minimiser les effets des variables confondantes

Idéalement : en laboratoire, contrôle précis de l'environnement.

Hors laboratoire :

- ▶ bonne conception de l'expérience (par exemple randomisation)
- ▶ contrôles réguliers (groupes contrôles bien choisis)

Effet d'un antidépresseur à base de plantes

- ▶ On prélève un échantillon de participants à un groupe de médecine alternative en ligne.
- ▶ On leur demande de prendre le traitement pendant un mois.
- ▶ Les participants remplissent un questionnaire sur la dépression avant et après pour mesurer l'amélioration de leur humeur.
- ▶ On constate en effet que celle-ci est meilleure après traitement.

Pouvez-vous identifier des variables qui auraient pu fausser ce résultat ?

- ▶ Croyance des participants en l'efficacité du traitement
- ▶ Amélioration de l'humeur pour d'autres raisons : météo, vacances, situation personnelle...

Variables contrôlées

Objectif :

- ▶ Réduire la possibilité que tout autre facteur influence les changements de la variable dépendante.
- ▶ Autrement dit : rendre **constantes** les variables confondantes.

Retour sur les exemples précédents

- ▶ Engrais et croissance des plantes :
elles doivent toutes recevoir la même quantité d'eau.
(*normalisation de l'effet*)

Variables contrôlées

Objectif :

- ▶ Réduire la possibilité que tout autre facteur influence les changements de la variable dépendante.
- ▶ Autrement dit : rendre **constantes** les variables confondantes.

Retour sur les exemples précédents

- ▶ Engrais et croissance des plantes :
elles doivent toutes recevoir la même quantité d'eau.
(*normalisation de l'effet*)
- ▶ Essai d'antihistaminiques : comment contrôler les effets de la météo ?

Variables contrôlées

Objectif :

- ▶ Réduire la possibilité que tout autre facteur influence les changements de la variable dépendante.
- ▶ Autrement dit : rendre **constantes** les variables confondantes.

Retour sur les exemples précédents

- ▶ Engrais et croissance des plantes :
elles doivent toutes recevoir la même quantité d'eau.
(normalisation de l'effet)
- ▶ Essai d'antihistaminiques : comment contrôler les effets de la météo ?
Constituer un groupe contrôle qui ne reçoit pas le traitement.
Autre variable à contrôler : par exemple, arrêter tout autre traitement médicamenteux pendant l'expérience.
(élimination de l'effet)

Ces variables sont en fait beaucoup plus importantes que les variables dépendantes ou indépendantes, puisqu'elles peuvent compromettre toute l'expérience.

Exercice

On cherche à comparer l'efficacité de différentes lessives.

Variable indépendante

Exercice

On cherche à comparer l'efficacité de différentes lessives.

Variable indépendante

la marque du produit utilisé

Variables à contrôler

Exercice

On cherche à comparer l'efficacité de différentes lessives.

Variable indépendante

la marque du produit utilisé

Variables à contrôler

- ▶ degré de saleté
- ▶ type de saleté ou de tache
- ▶ température de l'eau
- ▶ durée du cycle de nettoyage
- ▶ (entre autres)

Groupes contrôle

Efficacité d'un programme de lecture à l'école

- ▶ On mesure le niveau de lecture des enfants avant et après avoir suivi le programme ; celui-ci s'améliore de 50%.
- ▶ Comment savoir si leur niveau de lecture ne se serait pas amélioré de toute façon ?

On peut constituer un groupe contrôle :

- ▶ ne reçoit aucun traitement ou intervention, ou bien reçoit un traitement « standard »
- ▶ identique au groupe expérimental en tout autre point
- ▶ donnera une **mesure de référence** pour les variables inconnues.

Groupes contrôle

Efficacité d'un programme de lecture à l'école

- ▶ On mesure le niveau de lecture des enfants avant et après avoir suivi le programme ; celui-ci s'améliore de 50%.
- ▶ Comment savoir si leur niveau de lecture ne se serait pas amélioré de toute façon ?

On peut constituer un groupe contrôle :

- ▶ ne reçoit aucun traitement ou intervention, ou bien reçoit un traitement « standard »
- ▶ identique au groupe expérimental en tout autre point
- ▶ donnera une **mesure de référence** pour les variables inconnues.

Attention à l'*effet Hawthorne* : une personne consciente d'être le sujet d'une expérience modifie son comportement.

- ▶ mentir sur l'objectif du protocole (mais problème éthique)
- ▶ proposer un protocole alternatif au groupe contrôle

Groupes contrôle (cont.)

Groupe contrôle positif

- ▶ reçoit un traitement connu pour produire un effet ;
- ▶ permet de vérifier si le nouveau traitement produit les effets escomptés (ou mieux).
- ▶ Si le contrôle échoue également, on détecte un problème dans la conception de l'expérience. (Réduction du risque de faux négatifs)

Groupe contrôle négatif

- ▶ n'est censé produire aucun effet (dans ce cas, aucune variable confondante n'a affecté les résultats) ;
- ▶ permet également une mesure différentielle (exemple : les enfants n'ayant pas suivi le programme de lecture ont progressé de 10%).

Là encore, les groupes contrôle sont presque plus importants que les échantillons réels.

À vous

- ▶ Quelles sont les variables confondantes pour votre étude ?
(générales, spécifiques)
- ▶ Quelles mesures de contrôle pouvez-vous mettre en place ?
- ▶ Quels biais subsisteront ?

Causalité et corrélation

L'identification des variables confondantes est lié à la distinction entre corrélation et causalité.

- ▶ Un test statistique permet de constater une **corrélation** (variation simultanée de deux variables).
- ▶ Mais il ne permet **pas** d'en déduire un lien de causalité entre elles.

Effet cigogne

- ▶ Plus il y a de cigognes dans une région, plus le taux de natalité est haut : les cigognes apportent-elles les bébés ?
- ▶ Plus il y a de pompiers combattant un incendie, plus les dégâts sont importants : faut-il diminuer les effectifs ?
- ▶ Moins il y a de pirates, plus la température terrestre augmente.

Causalité et corrélation

L'identification des variables confondantes est lié à la distinction entre corrélation et causalité.

- ▶ Un test statistique permet de constater une **corrélation** (variation simultanée de deux variables).
- ▶ Mais il ne permet **pas** d'en déduire un lien de causalité entre elles.

Effet cigogne

- ▶ Plus il y a de cigognes dans une région, plus le taux de natalité est haut : les cigognes apportent-elles les bébés ?
- ▶ Plus il y a de pompiers combattant un incendie, plus les dégâts sont importants : faut-il diminuer les effectifs ?
- ▶ Moins il y a de pirates, plus la température terrestre augmente.

Même en cas de corrélation :

- ▶ il se peut qu'on confonde l'effet et la cause
- ▶ les deux phénomènes observés peuvent avoir une cause commune
- ▶ il se peut même qu'ils ne soient liés que par pure coïncidence

Les schémas de causalité dans l'esprit humain

Inconsciemment

L'esprit humain a tendance naturellement à essayer d'établir des liens entre les informations qu'il reçoit, de façon à construire des modèles.

Consciemment

Un sondage, une statistique, des données numériques nombreuses revêtent facilement un aspect « scientifique donc correct ».

Volontairement

Un média, un groupe politique ou un lobby peuvent utiliser une corrélation perçue pour « prouver » leurs propres convictions.

D'autres biais à propos de causalité

Biais émotionnel

- ▶ Un chercheur montre une corrélation entre le mauvais comportement des enfants et le tabagisme des parents.
- ▶ Il propose d'inscrire des avertissements sur les questions sociales sur les paquets de cigarettes.

Le chercheur ayant mené l'enquête est opposé au tabagisme, il espérait trouver un lien qui soutiendrait son propre programme.

D'autres biais à propos de causalité

Biais émotionnel

- ▶ Un chercheur montre une corrélation entre le mauvais comportement des enfants et le tabagisme des parents.
- ▶ Il propose d'inscrire des avertissements sur les questions sociales sur les paquets de cigarettes.

Le chercheur ayant mené l'enquête est opposé au tabagisme, il espérait trouver un lien qui soutiendrait son propre programme.

Problématisation a posteriori

Les hypothèses sont choisies **après** le recueil des données.

D'autres biais à propos de causalité

Biais émotionnel

- ▶ Un chercheur montre une corrélation entre le mauvais comportement des enfants et le tabagisme des parents.
- ▶ Il propose d'inscrire des avertissements sur les questions sociales sur les paquets de cigarettes.

Le chercheur ayant mené l'enquête est opposé au tabagisme, il espérait trouver un lien qui soutiendrait son propre programme.

Problématisation a posteriori

Les hypothèses sont choisies **après** le recueil des données.

Raisonnement *post hoc*

Puisqu'un événement suit un autre (dans le temps), le premier doit être la cause du second.

Plan

Construire un protocole expérimental

Variation de recherche

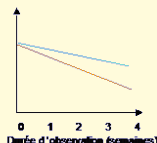
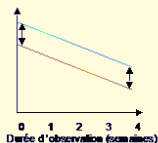
Causalité et corrélation

Randomisation

Le biais de sélection

Utilisation du groupe contrôle

- ▶ Mesurer la différence entre les groupes expérimental et contrôle
- ▶ Cette différence doit être uniquement due à l'action sur laquelle porte la recherche
- ▶ Il faut donc que les deux groupes soient comparables auparavant !



Quand les deux groupes sont dissemblables initialement, on parle de **biais de sélection**.

Remarque

Même si on corrigeait la différence finale en fonction de la différence initiale :

- ▶ Il n'est pas sûr que les deux groupes évoluent parallèlement
- ▶ Le niveau initial peut influencer l'évolution

Les mauvais groupes contrôles

Contrôles historiques

Par exemple, élèves ayant suivi le même cours l'année précédente :

- ▶ pas forcément comparables aux nouveaux
- ▶ l'enseignant, ou les autres enseignements reçus ont pu changer eux aussi de méthodes

Un groupe contrôle devrait être contemporain du groupe expérimental.

Contrôle géographique

Par exemple, élèves du lycée A comparés aux élèves du lycée B :

- ▶ rien ne garantit que les élèves soient similaires
- ▶ ni qu'ils soient pris en charge de la même manière

Essai contrôlé randomisé

Comment assurer la comparabilité initiale des deux groupes ?

Randomisation

Il s'agit de choisir quels participants à la recherche sont affectés au groupe expérimental ou au groupe de contrôle, sans tenir compte du niveau de l'élève, de ses caractéristiques, du contexte...

- ▶ toute méthode de *choix* risque d'induire des biais
- ▶ on affecte donc chaque participant à un groupe **au hasard**

(On parle de **groupes parallèles**.)

Sur une population suffisamment nombreuse, la randomisation a tendance à produire des groupes comparables pour toutes les variables externes, et donc à lisser les biais.

(Les groupes créés sont comparables *en moyenne*.)

On peut également vérifier (mesurer) la comparabilité des groupes en début d'expérience.

La randomisation en pratique

Dans le détail, plusieurs procédures sont possibles :

- ▶ randomisation simple : chacun son tour, chaque participant est affecté à un groupe
- ▶ randomisation par paires : les participants sont réunis par paires similaires (pour certaines variables identifiées) et on ventile aléatoirement chaque paire dans les deux groupes

Attention encore aux biais

Pour être la plus efficace possible, la randomisation doit être imprévisible :

- ▶ Si le chercheur a connaissance à l'avance de la répartition des participants dans les groupes,

La randomisation en pratique

Dans le détail, plusieurs procédures sont possibles :

- ▶ randomisation simple : chacun son tour, chaque participant est affecté à un groupe
- ▶ randomisation par paires : les participants sont réunis par paires similaires (pour certaines variables identifiées) et on ventile aléatoirement chaque paire dans les deux groupes

Attention encore aux biais

Pour être la plus efficace possible, la randomisation doit être imprévisible :

- ▶ Si le chercheur a connaissance à l'avance de la répartition des participants dans les groupes, il peut inconsciemment essayer de privilégier « l'intérêt » d'un participant au détriment de l'essai.
- ▶ Une répartition basée sur l'âge, le sexe, l'enseignant en charge du groupe... ou tout autre critère « pseudo-aléatoire » introduit

La randomisation en pratique

Dans le détail, plusieurs procédures sont possibles :

- ▶ randomisation simple : chacun son tour, chaque participant est affecté à un groupe
- ▶ randomisation par paires : les participants sont réunis par paires similaires (pour certaines variables identifiées) et on ventile aléatoirement chaque paire dans les deux groupes

Attention encore aux biais

Pour être la plus efficace possible, la randomisation doit être imprévisible :

- ▶ Si le chercheur a connaissance à l'avance de la répartition des participants dans les groupes, il peut inconsciemment essayer de privilégier « l'intérêt » d'un participant au détriment de l'essai.
- ▶ Une répartition basée sur l'âge, le sexe, l'enseignant en charge du groupe... ou tout autre critère « pseudo-aléatoire » introduit un biais de sélection.

Essais en aveugle

L'aveuglement sert à éliminer davantage les biais.

Essai en simple aveugle

Les participants ne savent pas dans quel groupe ils se trouvent et quelle intervention ils reçoivent jusqu'à la fin de l'étude.

Exemple : On donne aux patients des gélules de même forme, taille, couleur, certaines contenant une molécule et les autres un placebo.

Essais en aveugle

L'aveuglement sert à éliminer davantage les biais.

Essai en simple aveugle

Les participants ne savent pas dans quel groupe ils se trouvent et quelle intervention ils reçoivent jusqu'à la fin de l'étude.

Exemple : On donne aux patients des gélules de même forme, taille, couleur, certaines contenant une molécule et les autres un placebo.

Essai en double aveugle

Ni les participants ni le chercheur ne savent à quel groupe le participant appartient.

Exemple : Les gélules sont préparées à l'avance dans des enveloppes numérotées, un système informatisé dit au chercheur quelles enveloppes administrer.

Essais en aveugle

L'aveuglement sert à éliminer davantage les biais.

Essai en simple aveugle

Les participants ne savent pas dans quel groupe ils se trouvent et quelle intervention ils reçoivent jusqu'à la fin de l'étude.

Exemple : On donne aux patients des gélules de même forme, taille, couleur, certaines contenant une molécule et les autres un placebo.

Essai en double aveugle (impossible pour nous)

Ni les participants ni le chercheur ne savent à quel groupe le participant appartient.

Exemple : Les gélules sont préparées à l'avance dans des enveloppes numérotées, un système informatisé dit au chercheur quelles enveloppes administrer.

Essais en aveugle (cont.)

Essai ouvert

Le chercheur et le participant savent à quel groupe le participant appartient.

(réservé au cas où il n'est pas possible de « simuler » une intervention dans le groupe contrôle)

Essais en aveugle (cont.)

Essai ouvert

Le chercheur et le participant savent à quel groupe le participant appartient.

(réservé au cas où il n'est pas possible de « simuler » une intervention dans le groupe contrôle)

Mitigation possible : faire évaluer les participants par un évaluateur indépendant, en aveugle de l'intervention reçue.

(utilisé dans le *triple aveugle*)

Limites des essais contrôlés randomisés

- ▶ Les participants aux groupes expérimental et contrôle sont tous recrutés dans un même échantillon de la population.
Si celui-ci n'est pas représentatif de la population générale, il devient difficile de généraliser les résultats.
- ▶ Pour que les mesures effectuées soient fiables en moyenne, il faut de grands groupes, ce qui demande des ressources importantes.
- ▶ Difficulté à établir des groupes recevant une action distincte au sein d'une même classe.

Des variantes intéressantes

Essai croisé

Chaque participant participe au groupe expérimental puis au groupe contrôle dans un ordre aléatoire.

- ▶ on remplace la variabilité inter-individus par une variabilité intra-individu
- ▶ on « double » l'effectif de chaque groupe

Limites :

- ▶ les mesures doivent pouvoir être répétées chez le même sujet
- ▶ pas d'effet d'entraînement, réversibilité de l'action prise sur le groupe expérimental
- ▶ respecter une période de « lavage » entre les deux phases

Des variantes intéressantes (cont.)

Essai par grappes

Au lieu de répartir les participants individuellement, on affecte des groupes entiers dans les groupes expérimental et contrôle :

- ▶ groupes pré-existants (classes, établissements, villes...)
- ▶ permet de tester une action collective
- ▶ contrôle la « contamination » entre participants
- ▶ mais augmente l'hétérogénéité inter-groupes et demande donc des méthodes statistiques plus poussées

À vous

- ▶ Quelle forme de randomisation serait possible dans votre expérience ?
- ▶ Que proposer au groupe « contrôle » ?
- ▶ Pouvez-vous facilement proposer des activités ON/OFF ?
- ▶ Y a-t-il une forme d'aveugle applicable à votre expérience ?