Causal Effects in Between-Group Experiments and Quasi-Experiments: Theory

Rolf Steyer Ivailo Partchev, Ulf Kröhne, Benjamin Nagengast & Christiane Fiege
University of Jena

– Basic Concepts and Definitions –

1 Variables and Probabilities

1. Variables of interest:

- Y is an outcome variable
- X is the treatment variable with $X = 0$ or $X = k$ the control group and $X = 1$ or $X = j$ the treatment group 1 or j
- Z is a manifest or latent one- or multidimensional covariate
- U is the unit variable (e. g. a person variable)

- **True-outcome variables**
  - True outcomes of unit u under control ($X = 0$):

$$\tau_0(u) \equiv E(Y \mid X = 0, U = u)$$

  - True outcomes of unit u under treatment ($X = 1$):

$$\tau_1(u) \equiv E(Y \mid X = 1, U = u)$$

2. Relevant probabilities:

- $P(U = u)$ is the sampling probability of unit u
- $P(U = u \mid X = j)$ is the probability that a randomly chosen observational unit given (in) the treatment group j is unit u
- $P(X = j \mid U = u)$ is the probability that unit u gets treatment j
- $P(X = j \mid Z = z)$ is the **propensity score**, the probability of being assigned to treatment j given a value of the covariate

2 Regressions

Note: The term regression refers here to a random variable which can take on different values, the conditional expected values. This definition of a regression does not imply a certain parametrization of the regression model! The following regressions are essential for the theory of causality depicted here.

- Treatment regression $E(Y \mid X)$ with the values $E(Y \mid X = x)$
- Unit treatment regression $E(Y \mid X, U)$ with the values $E(Y \mid X = x, U = u)$
- Covariate-treatment regression $E(Y \mid X, Z)$ with the values $E(Y \mid X = x, Z = z)$

3 Regressions describing the prima facie effect

Unconditional and conditional prima facie effects can be written by means of parameters of certain regression models. For that reason the following linear regressions will be used. For simplicity the variable X has only two levels, $X = 0$ the control group and $X = 1$ the treatment group:

1. The **treatment regression** as simple linear regression of Y on X:

$$E(Y \mid X) = \alpha_0 + \alpha_1 \cdot X$$

2. The **covariate-treatment regression** as multiple linear regression of Y on X and Z:

$$\begin{aligned} E(Y \mid X, Z) &= g_0(Z) + g_1(Z) \cdot X \\ &= (\gamma_{00} + \gamma_{01} \cdot Z) + (\gamma_{10} + \gamma_{11} \cdot Z) \cdot X \end{aligned}$$

Note: In the multiple regression $E(Y \mid X, Z)$ of Y on X and Z the intercept $g_0(Z)$ and the slope $g_1(Z)$ are conceived as functions of the covariate Z! Therefor, they are called the *intercept function* and the *effect function*. They can vary depending on the values of Z. If the slope $g_1(z)$ takes on different values for different levels z of Z, there is a *moderation, i.e. interaction*!

4 Three kinds of prima facie effects

In general, the prima facie effects are defined as differences in the expected outcomes of Y between treatment and control group. They can either be defined for a population or for subpopulations which are given by the values z of a covariate Z.

Three kinds of prima facie effect can be defined and expressed as parameters of the regressions described above:

1. The **unconditional prima facie effect** PFE_{10}:

$$\begin{aligned} PFE_{10} &= E(Y \mid X = 1) - E(Y \mid X = 0) \\ &= \alpha_1 \end{aligned}$$

2. The **conditional prima facie effect** PFE_{10} given a value z of the covariate Z:

$$\begin{aligned} PFE_{10}(z) &= E(Y \mid X = 1, Z = z) - E(Y \mid X = 0, Z = z) \\ &= g_1(z) \end{aligned}$$

3. The **average prima facie effect** $E[g_1(Z)]$ with respect to the covariate Z.
   - if Z is categorical:

$$\begin{aligned} E[g_1(Z)] &= \sum_z g_1(z) \cdot P(Z = z) \\ &= \sum_z \Big[E(Y \mid X = 1, Z = z) - E(Y \mid X = 0, Z = z) \Big] \cdot P(Z = z) \end{aligned}$$

   - if Z is continuous:

$$E[g_1(Z)] = \gamma_{10} + \gamma_{11} \cdot E(Z)$$

Note that usually $E[g_1(Z)] \neq PFE_{10}$!

5 Definitions of causal effects

Individual causal effect (ICE)

Individual causal effect $ICE_{10}(u)$ of the treatment as compared to the control of unit u:

$$\delta_{10} \equiv \tau_1(u) - \tau_0(u)$$

Average causal effect (ACE)

Average causal effect ACE_{10} of the treatment as compared to the control:

$$\begin{aligned} ACE_{10} &\equiv E(\delta_{10}) \\ &\equiv E(\tau_1 - \tau_0) \equiv E(\tau_1) - E(\tau_0) \end{aligned}$$

---

whereby:

- $E(\tau_1)$ is the **expected true outcome under treatment** $X = 1$:

$$E(\tau_1) = \sum_u \tau_1(u) \cdot P(U = u)$$

- $E(\tau_0)$ is the **expected true outcome under control** $X = 0$:

$$E(\tau_0) = \sum_u \tau_0(u) \cdot P(U = u)$$

Conditional average causal effects

- **Conditional average causal effect given a covariate Z**
  Average causal effect ACE_{10} of the treatment as compared to the control given a value of the covariate:

$$\begin{aligned} ACE_{10;Z=z} &\equiv E(\delta_{10} \,|\, Z = z) = \sum_u \delta_{10} \cdot P(U = u) \\ &\equiv E(\tau_1 - \tau_0 \,|\, Z = z) = E(\tau_1 \,|\, Z = z) - E(\tau_0 \,|\, Z = z) \end{aligned}$$

- **Conditional average causal effect given the treatment variable X**
  Average causal effect $ACE_{10;X=0}$ given control:

$$\begin{aligned} ACE_{10;X=0} &\equiv E(\delta_{10} \,|\, X = 0) = \sum_u \delta_{10} \cdot P(U = u \,|\, X = 0) \\ &\equiv E(\tau_1 - \tau_0 \,|\, X = 0) = E(\tau_1 \,|\, X = 0) - E(\tau_0 \,|\, X = 0) \end{aligned}$$

  Average causal effect $ACE_{10;X=1}$ given treatment:

$$\begin{aligned} ACE_{10;X=1} &\equiv E(\delta_{10} \,|\, X = 1) = \sum_u \delta_{10} \cdot P(U = u \,|\, X = 1) \\ &\equiv E(\tau_1 - \tau_0 \,|\, X = 1) = E(\tau_1 \,|\, X = 1) - E(\tau_0 \,|\, X = 1) \end{aligned}$$

- **Conditional average causal effect given the treatment and the covariate**
  Average causal effect $ACE_{10;X=i,Z=z}$ of the treatment $X = 1$ as compared to the control $X = 0$ given a value z of the covariate Z and given the treatment $X = i$:

$$\begin{aligned} ACE_{10;X=i,Z=z} &\equiv E(\delta_{10} \,|\, X = i, Z = z) = \sum_u \delta_{10} \cdot P(U = u \,|\, X = i, Z = z) \\ &\equiv E(\tau_1 - \tau_0 \,|\, X = i, Z = z) = E(\tau_1 \,|\, X = i, Z = z) - E(\tau_0 \,|\, X = i, Z = z) \end{aligned}$$

Note: All definitions are based on probability theory. They refer to an event which has not happened yet. Only from this point of view, the so called *pre factual perspective*, the definitions become meaningful and understandable. That means the subjects are not treated yet or will never be treated and therefore we have no empirical data! We only consider the values on the outcome variable they might get, if they were treated or not treated. Just regarding the population of the untreated we may theoretically consider their expected outcome under treatment and their expected outcome under control. Thus, the average causal effect for this population, the $ACE_{jk;X=0}$ may be defined. Note, that even if we only consider the population of the untreated we may consider the expected outcome under treatment as well as the average causal effect for this population.

6 Alternative definitions of causal effects

If there is more than one treatment condition and no control group, further definitions of causal effects can be established. The basic idea of the following alternative definitions of individual and average causal effects is the comparison of a treatment $X = j$ to the average of all $J + 1$ treatments. This might be a reasonable definition when there is no meaningful control group!

Individual causal effect ($ICE_{j,total}$)
Individual causal effect $ICE_{j,total}$ of treatment j as compared to all treatment conditions:

$$ICE_{j,total} \equiv \tau_j(U) - \sum_{k=0}^{J} \tau_k(U) \qquad j = 0, 1, ..., J$$

Average causal effect ($ACE_{j,total}$)
Average causal effect $ACE_{j,total}$ of treatment j as compared to all treatment conditions:

$$\begin{aligned} ACE_{j,total} &\equiv E(ICE_{j,total}) \\ &\equiv E(\tau_j) - \sum_{k=0}^{J} E(\tau_k) \qquad j = 0, 1, ..., J \end{aligned}$$

Corresponding alternative definitions of the conditional average causal effects

- **Conditional average causal effect of treatment j as compared to all treatment conditions given a value z of the covariate Z**

$$\begin{aligned} ACE_{j,total;Z=z} &\equiv E(ICE_{j,total} \,|\, Z = z) \\ &\equiv E(\tau_j \,|\, Z = z) - \sum_{k=0}^{J} E(\tau_k \,|\, Z = z) \qquad j = 0, 1, ..., J \end{aligned}$$

- **Conditional average causal effect of treatment j as compared to all treatment conditions given treatment i**

$$\begin{aligned} ACE_{j,total;X=i} &\equiv E(ICE_{j,total} \,|\, X = i) \\ &\equiv E(\tau_j \,|\, X = i) - \sum_{k=0}^{J} E(\tau_k \,|\, X = 0i) \qquad j = 0, 1, ..., J \end{aligned}$$

- **Conditional average causal effect of treatment j as compared to all treatment conditions given treatment i and a value z of the covariate z of Z**

$$\begin{aligned} ACE_{j,total;X=i,Z=z} &\equiv E(ICE_{j,total} \,|\, X = i, Z = z) \\ &\equiv E(\tau_j \,|\, X = i, Z = z) - \sum_{k=0}^{J} E(\tau_k \,|\, X = i, Z = z) \qquad j = 0, 1, ..., J \end{aligned}$$

7 Relationship between prima facie effects and average causal effects

Unconditional and conditional prima facie effects are defined as differences between the expectations of the outcome variable Y! By contrast, the unconditional and conditional average causal effects are defined as differences between the expectations of the true outcomes τ_j!

7.1 Prima facie effects and average causal effects: A comparison

Comparison of PFE an ACE	
Prima facie effect (PFE_{jk}) $$PFE_{jk} \equiv E(Y \mid X = j) - E(Y \mid X = k)$$	**Average causal effect** (ACE_{jk}) $$ACE_{jk} \equiv E(\tau_j) - E(\tau_k)$$
Note the difference!	
Expectation of the outcome Y in treatment j $$E(Y \mid X = j) = \sum_u \tau_j(u) \cdot P(U = u \mid X = j)$$	*Expectation of the true outcome τ_j in treatment j* $$E(\tau_j) = \sum_u \tau_j(u) \cdot P(U = u)$$
Comparison of conditional $PFE_{jk;Z=z}$ and $ACE_{jk;Z=z}$	
$PFE_{jk;Z=z}$ $$PFE_{jk;Z=z} \equiv E(Y \mid X = j, Z = z) - E(Y \mid X = k, Z = z)$$	$ACE_{jk;Z=z}$ $$ACE_{jk;Z=z} \equiv E(\tau_j \mid Z = z) - E(\tau_k \mid Z = z)$$
Note the difference!	
Expectation of the outcome Y in treatment j given the value z of the Covariate Z $$E(Y \mid X = j, Z = z) = \sum_u \tau_j(u) \cdot P(U = u \mid X = j, Z = z)$$	*Expectation of the true outcome τ_j in treatment j given the value z of the Covariate Z* $$E(\tau_j \mid Z = z) = \sum_u \tau_j(u) \cdot P(U = u \mid Z = z)$$

7.2 Unbiasedness of the regressions $E(Y \mid X)$ and $E(Y \mid X, Z)$

Unbiasedness of the treatment regression $E(Y \mid X)$

It can be deduced from the definitions, that the equality of prima facie effect and average causal effect follows from $E(Y \mid X = j) = E(\tau_j)$, i. e. the equality of the expectations of the outcomes and the true outcomes. This is true, e.g., if there is no stochastic dependency between the person variable U and the treatment X, because:

$$U \perp X \Rightarrow P(U = u \mid X = j) = P(U = u)$$

If the stochastic independency holds, the computation of the expectation $E(Y \mid X = j)$ reduces to the formula of the expectation of the true outcome $E(\tau_j)$,

$$\begin{aligned} E(Y \mid X = j) &= \sum_u \tau_j(u) \cdot P(U = u \mid X = j) \\ &= \sum_u \tau_j(u) \cdot P(U = u) \quad | \quad \text{if } U \perp X \\ &= E(\tau_j). \end{aligned}$$

If the equality $E(Y \mid X = j) = E(\tau_j)$ holds for all treatments $j \in (0, 1, ..., J)$, the treatment regression $E(Y \mid X)$ is called unbiased!

Unbiasedness of the covariate-treatment regression $E(Y \mid X, Z)$

The unbiasedness of the covariate-treatment regression $E(Y \mid X, Z)$ is defined analogous: The conditional prima facie effect and the conditional average causal effect given the covariate Z are equal if $E(Y \mid X = j, Z = z) = E(\tau_j \mid Z = z)$, i. e. if the expectations of the outcomes equal those of the true outcomes given a value z of the covariate Z. This is true, e.g., if there is no stochastic dependency between the person variable U and the treatment X given Z, because:

$$U \perp X \mid Z \Rightarrow P(U = u \mid X = j, Z = z) = P(U = u \mid Z = z).$$

If the stochastic independency holds, the computation of the expectation $E(Y \mid X = j, Z = z)$ reduces to the formula of the expectation of the true outcome $E(\tau_j \mid Z = z)$,

$$\begin{aligned} E(Y \mid X = j, Z = z) &= \sum_u \tau_j(u) \cdot P(U = u \mid X = j, Z = z) \\ &= \sum_u \tau_j(u) \cdot P(U = u \mid Z = z) \quad | \quad \text{if } U \perp X \mid Z \\ &= E(\tau_j \mid Z = z). \end{aligned}$$

If the equality $E(Y \mid X = j, Z = z) = E(\tau_j \mid Z = z)$ holds for all treatments $j \in (0, 1, ..., J)$ and all values z of Z, the covariate-treatment regression $E(Y \mid X, Z)$ is called unbiased!

7.3 Unbiasedness of the prima facie effect PFE_{jk}

Unbiasedness of the prima facie effects PFE_{jk} and $PFE_{jk;Z=z}$
The prima facie effect is called unbiased, if

$$PFE_{jk} = ACE_{jk}.$$

The unbiasedness of the regression $E(Y \mid X)$ is sufficient for the unbiasedness of the prima facie effect PFE_{jk}, since it implies the equality $E(Y \mid X = j) = E(\tau_j)$. Therefore, the following equation holds:

$$\begin{aligned} PFE_{jk} &= E(Y \mid X = j) - E(Y \mid X = k) \\ &= E(\tau_j) - E(\tau_k) \\ &= ACE_{jk}. \end{aligned}$$

Unbiasedness of the conditional prima facie effect $PFE_{jk;Z=z}$

The conditional prima facie effect is called unbiased, if

$$PFE_{jk;Z=z} = ACE_{jk;Z=z}$$

The unbiasedness of the covariate-treatment regression $E(Y \mid X, Z)$ is sufficient for the unbiasedness of the conditional prima facie effect $PFE_{jk;Z=z}$, since it implies the equality $E(Y \mid X = j, Z = z) = E(\tau_j, Z = z)$ for all values z of Z. Therefore, the following equation holds:

$$\begin{aligned} PFE_{jk;Z=z} &= E(Y \mid X = j; Z = z) - E(Y \mid X = k; Z = z) \\ &= E(\tau_j \mid Z = z) - E(\tau_k \mid Z = z) \\ &= ACE_{jk;Z=z}. \end{aligned}$$

8 Bias of the prima facie effect

Two kinds of biases are to be distinguished: the *baseline bias* and the *effect bias*! They are defined for each, the unconditional and the conditional prima facie effect.

The two biases of the unconditional PFE_{jk}:

1. ***Baseline bias$_{jk}$***

$$\textit{baseline bias}_{jk} = E(\tau_k \mid X = j) - E(\tau_k \mid X = k),$$

with $E(\tau_k \mid X = j)$ being the expected true outcome of control given the treatment j. Note, that this is again just a theoretical value. We consider the expected outcome the population of the treated *would* get, if they *were* in the control group.
Substantive meaning: The *baseline bias* is the difference in the expectation of the true outcomes under control between the population of the treated and the untreated.

2. ***Effect bias$_{jk}$***:

$$\begin{aligned}\textit{effect bias}_{jk} &= E(\delta_{jk} \mid X = j) - ACE_{jk} \\ &= ACE_{jk;X=j} - ACE_{jk}\end{aligned}$$

Substantive meaning: The *effect bias* is the difference between the conditional average causal effect given the treatment j and the unconditional average causal effect! Note that $ACE_{jk;X=j}$ is the difference $E(\tau_j \mid X = j) - E(\tau_k \mid X = j)$.

The two biases of the conditional $PFE_{jk;Z}$ given a covariate Z:

1. ***Conditional baseline bias$_{jk;Z}$***

$$\textit{baseline bias}_{jk;Z} = E_{X=j}(\tau_k \mid Z) - E_{X=k}(\tau_k \mid Z),$$

Substantive meaning: If the covariate Z is a function $f(u)$ of the person and can be measured without measurement error, the conditional *baseline bias* given a value z of a covariate Z can be interpreted as the difference in the expectation of the true outcomes under control between the population of the treated and the untreated *within the subpopulation* $Z = z$.

2. ***Conditional effect bias$_{jk;Z}$***

$$\begin{aligned}\textit{effect bias}_{jk;Z} &= E_{X=j}(\delta_{jk} \mid Z) - ACE_{jk;Z} \\ &= ACE_{jk;X=j,Z} - ACE_{jk;Z}\end{aligned}$$

Substantive meaning: Under the condition $Z = f(u)$, the conditional *effect bias* is the difference between the conditional average causal effect given the treatment j and the unconditional average causal effect *within the subpopulation* $Z = z$!

9 The relation between ACE, PFE and the biases

The prima facie effect PFE_{jk} is the sum of the average causal effect ACE_{jk}, the *baseline bias*$_{jk}$ and the *effect bias*$_{jk}$.

$$PFE_{jk} = ACE_{jk} + baseline\ bias_{jk} + effect\ bias_{jk}$$

Proof:

$$\begin{aligned}
PFE_{jk} &= ACE_{jk} + baseline\ bias_{jk} + effect\ bias_{jk}\\
ACE_{jk} &= PFE_{jk} - baseline\ bias_{jk} - effect\ bias_{jk}\\
&= \Big[E(\tau_j\,|\,X=j) - E(\tau_k\,|\,X=k)\Big] - \Big[E(\tau_k\,|\,X=j) - E(\tau_k\,|\,X=k)\Big]\\
&\quad -\Big[E(\delta_{jk}\,|\,X=j) - E(\delta_{jk})\Big]\\
&= \Big[E(\tau_j\,|\,X=j) - E(\tau_k\,|\,X=k)\Big] - \Big[E(\tau_k\,|\,X=j) - E(\tau_k\,|\,X=k)\Big]\\
&\quad -\Big\{E(\tau_j\,|\,X=j) - E(\tau_k\,|\,X=j) - \Big[E(\tau_j) - E(\tau_k)\Big]\Big\}\\
&= E(\tau_j\,|\,X=j) - E(\tau_k\,|\,X=k) - E(\tau_k\,|\,X=j) + E(\tau_k\,|\,X=k)\\
&\quad -E(\tau_j\,|\,X=j) + E(\tau_k\,|\,X=j) + E(\tau_j) - E(\tau_k)\\
&= E(\tau_j) - E(\tau_k)
\end{aligned}$$

Accordingly, the conditional prima facie effect $PFE_{jk;Z=z}$ given the covariate Z is the sum of the average causal effect $ACE_{jk;Z=z}$, the conditional *baseline bias*$_{jk;Z=z}$ and the conditional *effect bias*$_{jk;Z=z}$.

$$PFE_{jk;Z} = ACE_{jk;Z} + baseline\ bias_{jk;Z} + effect\ bias_{jk;Z}$$

Proof:

$$\begin{aligned}
PFE_{jk;Z} &= ACE_{jk;Z} + baseline\ bias_{jk;Z} + effect\ bias_{jk;Z}\\
ACE_{jk;Z} &= PFE_{jk;Z} - baseline\ bias_{jk;Z} - effect\ bias_{jk;Z}\\
&= \Big[E_{X=j}(\tau_j\,|\,Z) - E_{X=k}(\tau_k\,|\,Z)\Big] - \Big[E_{X=j}(\tau_k\,|\,Z) - E_{X=k}(\tau_k\,|\,Z)\Big]\\
&\quad -\Big[E(\delta_{jk}\,|\,Z) - E(\delta_{jk})\Big]\\
&= \Big[E_{X=j}(\tau_j\,|\,Z) - E_{X=k}(\tau_k\,|\,Z)\Big] - \Big[E_{X=j}(\tau_k\,|\,Z) - E_{X=k}(\tau_k\,|\,Z)\Big]\\
&\quad -\Big\{E_{X=j}(\tau_j\,|\,Z) - E_{X=j}(\tau_k\,|\,Z) - \Big[E(\tau_j) - E(\tau_k)\Big]\Big\}\\
&= E_{X=j}(\tau_j\,|\,Z) - E_{X=k}(\tau_k\,|\,Z) - E_{X=j}(\tau_k\,|\,Z) + E_{X=k}(\tau_k\,|\,Z)\\
&\quad -E_{X=j}(\tau_j\,|\,Z) + E_{X=j}(\tau_k\,|\,Z) + E(\tau_j\,|\,Z) - E(\tau_k\,|\,Z)\\
&= E(\tau_j\,|\,Z) - E(\tau_k\,|\,Z)
\end{aligned}$$

Unbiasedness of prima facie effects and unbiasedness of the regressions $E(Y\,|\,X)$ and $E(Y\,|\,X,Z)$. Unbiasedness of the different kinds of prima facie effects and unbiasedness of the treatment regression and the covariate-treatment regression are closely related! Since the PFE_{jk} is the difference between two values of the treatment regression, the conditional expectations $E(Y\,|\,X=j)$ and $E(Y\,|\,X=k)$, the unbiasedness of the treatment regression implies the unbiasedness of PFE_{jk}! The same relation holds for the conditional prima facie effect $PFE_{jk;Z=z}$. If the covariate-treatment regression is unbiased, the $PFE_{jk;Z=z}$ is unbiased as well!

10 Prima facie effects - conditions for unbiasedness

1. There are four sufficient conditions for the **unbiasedness of the treatment regression and the prima facie effect** PFE_{jk}!
   a) Stochastic independence of units and treatments $U \perp X$
   b) Unit treatment-homogeneity $Y \vdash U \mid X$
   c) Stochastic independence of true outcomes and treatments $\tau \perp X$
   d) Regressive independence of true outcomes and treatments $\tau \vdash X$

2. Accordingly, there are four sufficient conditions for the **unbiasedness of the covariate-treatment regression and the conditional prima facie effects** $PFE_{jk;Z=z}$ given the covariate Z!
   a) Stochastic independence of units and treatments given the covariate Z: $U \perp X \mid Z$
   b) Unit treatment-homogeneity given the covariate Z: $Y \vdash U \mid X, Z$
   c) Stochastic independence of true outcomes and treatments given the covariate Z: $\tau \perp X \mid Z$
   d) Regressive independence of true outcomes and treatments given the covariate Z: $\tau \vdash X \mid Z$

In detail:

1. **Stochastic independence of units and treatments**
   a) Stochastic independence of units and treatments $U \perp X$ is given if the following equation holds:

   $$P(X = j \mid U) \quad = \quad P(X = j).$$

   The assignment to a treatment j does not dependent on the person (unit u) and his characteristics! This assumption is *essential for designs of experiments*! It can be created by the experimenter via *random assignment*. Stochastically independence $U \perp X$ implies the unbiasedness of the treatment regression,

   $$E(Y \mid X = j) \quad = \quad E(\tau_j) \qquad \text{for each } j = 0, 1, ..., J,$$

   and therefore the unbiasedness of the prima facie effects, too.
   b) Conditional stochastic independence of units and treatments $U \perp X \mid Z$ given the covariate Z can be expressed as follows:

   $$P(X = j \mid U, Z) \quad = \quad P(X = j \mid Z).$$

   In cases where Z is measured without measurement error, the conditional independence means that the assignment to a treatment j does not dependent on the person (unit u) within subpopulations indicated by the value z of Z! Under conditional independence the following equation applies:

   $$E_{X=j}(Y \mid Z) \quad = \quad E(\tau_j \mid Z) \qquad \text{for each } j = 0, 1, ..., J.$$

   Conditional stochastic independence can be assured by the design of experiments. It can be established by *randomization within subpopulations*, if the subpopulations can be perfectly determined (e.g. gender). If Z can only be measured with measurement error the subpopulations can not be indicated by Z. In these cases *conditional randomization* may be used to ensure conditional stochastic independence.
2. **Unit treatment-homogeneity**
   a) *Unit treatment-homogeneity* $Y \vdash U \mid X$:
   The regression $E(Y \mid X)$ is called homogenous, if

   $$E(Y \mid X, U) \quad = \quad E(Y \mid X).$$

   This implies that each unit u has the same expectation $E(Y \mid X = j)$ of the outcome variable Y within each treatment condition! In this case, there are *no interindividual differences* between the units in their expectations of the outcome variable Y.

b) *Conditional unit treatment-homogeneity* $Y \vdash U \mid X, Z$:
The regression $E(Y \mid X, Z)$ is called conditionally homogenous, if

$$E(Y \mid X, Z, U) \quad = \quad E(Y \mid X, Z).$$

Is Z a function $f(U)$ of the units, all units u within a subpopulation z of Z have the same expectation $E_{X=j}(Y \mid Z = z)$ of the outcome variable Y within each treatment j. There are *no interindividual differences* between the units with respect to their conditional expectation of Y in each treatment condition j given a value z of Z.

3. **Stochastic independence of true outcomes and treatments**

a) *Stochastic independence of true outcomes and treatments*, $\tau \perp X$
Let $\boldsymbol{\tau} \equiv (\tau_0, \tau_1, ..., \tau_J)$ be the vector of the true outcome variables. The treatment variable X and $\boldsymbol{\tau}$ are called stochastically independent, if

$$P(X = j \mid \boldsymbol{\tau}) \quad = \quad P(X = j).$$

That means that the expectations of the true outcome variables τ_j (for all $j = 0, 1, ..., J$) do not differ between the treatment conditions:

$$E(\tau_j \mid X = i) \quad = \quad E(\tau_j \mid X = k) \qquad \text{for all } j, k \text{ and } i = 0, 1, ..., J$$

b) *Conditional stochastic independence of true outcomes and treatments*, $\tau \perp X \mid Z$
The treatment variable X and $\boldsymbol{\tau}$ are called conditionally stochastically independent, if

$$P(X = j \mid \boldsymbol{\tau}, Z) \quad = \quad P(X = j \mid Z).$$

This is equivalent to the following equation for all values z of Z and for all treatment conditions $j = 0, 1, ..., J$):

$$E_{X=i}(\tau_j \mid Z) \quad = \quad E_{X=k}(\tau_j \mid Z) \qquad \text{for all } j, k \text{ and } i = 0, 1, ..., J.$$

4. **Regressive independence of true outcomes and treatments**

a) *Regressive independence of true outcomes and treatments* $\tau \vdash X$
Let $\boldsymbol{\tau} \equiv (\tau_0, \tau_1, ..., \tau_J)$ denote the vector of true-outcome variables, which is called regressively independent from the treatment, if

$$E(\boldsymbol{\tau} \mid X) \quad = \quad E(\boldsymbol{\tau}).$$

The regressive independence $\tau \vdash X$ of the true-outcomes is the weakest condition for unbiasedness, because the three other conditions depicted above imply this last one! Regressive independence as well as stochastic independence of true outcomes and treatments $\tau \perp X$ implies that the expectations $E(\tau_j \mid X = i)$ of the true outcome variables τ_j (for all $j = 0, 1, ..., J$) do not differ between the treatment conditions.

b) *Conditional regressive independence of true outcomes and treatments* $\tau \vdash X \mid Z$
The treatment variable X and $\boldsymbol{\tau}$ are called conditionally regressively independent, if

$$E(\boldsymbol{\tau} \mid X, Z) \quad = \quad E(\boldsymbol{\tau} \mid Z).$$

The regressive independence of treatment variable X and the true-outcome $\boldsymbol{\tau}$ holds for the given levels z of the covariate Z.

Note: The regressive independence $\tau \vdash X$ and $\tau \vdash X \mid Z$ is fundamental because it is the theoretical basis for *adjusting techniques* which correct for biases! If at least the conditional regressive independence given an one- or multidimensional covariate Z holds, the treatment covariate regression is unbiased. In this case, we can estimate all kinds of unconditional and conditional average causal effects from the expectations $E(Y \mid X = j, Z = z)$ and the conditional prima facie effects $PFE_{jk;Z}$, despite the fact that the treatment regression is possibly biased (see: Adjusting for bias).

Testability of the four conditions of unbiasedness . The stochastic and regressive independence of the true-outcome variables τ and the treatment variable X are not empirically testable! However, the unconditional and conditional stochastic independence as well as the unconditional and conditional unit-treatment homogeneity can be tested in empirical applications in the sense of *falsification*! This means that at least the consequences of a particular sufficient condition can be checked!

1. **Testability of independence of units and treatments**
   In order to test the independence covariates W which are functions $f(U)$ of the units, such as gender or educational status, are needed. If $W = f(U)$, independence of units and treatment implies: $W \perp X$! This can be checked by the statistical comparison of the treatment probabilities given all the G different values w of W which are estimated by the relative frequencies! The following null hypothesis can be tested:

   $$P(X = j \,|\, W = w_1) = P(X = j \,|\, W = w_2) = ... = P(X = j \,|\, W = w_G).$$

   The conditional independence of units and treatment given a covariate Z can be tested similarly! The following null hypothesis can be tested on the basis of empirical data:

   $$P_{W=w_1}(X = j \,|\, Z) = P_{W=w_2}(X = j \,|\, Z) = ... = P_{W=w_G}(X = j \,|\, Z).$$

2. **Testability of unit treatment-homogeneity**
   In applications, we can never be sure that conditional unit-treatment homogeneity holds! The only way is to try to falsify this assumption by showing that for a certain variable W, which is a function of U, i.e. $W = f(U)$, the following equation holds:

   $$E(Y \,|\, X, Z, W) \neq E(Y \,|\, X, Z).$$

   This equation can not be true if the unit treatment-homogeneity holds!

11 Unconfoundedness

3 Reasons for the importance of Unconfoundedness.

1. Unconfoundedness is **the weakest testable condition** which implies unbiasedness! However, it does not imply that the both biases (baseline bias and effect bias) are zero! It is testable by falsification!
2. Unlike unbiasedness, unconfoundedness of the treatment regression $E(Y \mid X)$ implies unconfoundedness in all subpopulations (**generalizability**). Thus, unbiasedness holds in all subpopulations and allows valid causal inference!
3. Furthermore, unbiasedness implies that the expectation of the average causal effects in the subpopulations (that is the expectation of the conditional average causal effects given Z, whereby the values z of the covariate indicate the subpopulations) is always the average causal effect in the total population (**average stability**)!

3 formulations of unconfoundedness:
To define the unconfoundedness we use a further variable W, which is a mapping (function) $f(U)$ of the observational units!

1. **Unconfoundedness of the treatment regression**:
   a) $P(X = j \mid U) = P(X = j)$ or $E_{X=j}(Y \mid U) = E(Y \mid X = j)$, for each value j of X
   b) For every mapping $W = f(U)$ of the observational unit variable U:
   $P(X = j \mid W) = P(X = j)$ or $E_{X=j}(Y \mid W) = E(Y \mid X = j)$, for each value j of X
   c) For every mapping $W = f(U)$ of the observational unit variable U:
   $E(Y \mid X = j) = E[E_{X=j}(Y \mid W)]$
2. **Unconfoundedness of the covariate-treatment-regression**:
   a) $P_{Z=z}(X = j \mid U) = P_{Z=z}(X = j)$ or $E_{X=j,Z=z}(Y \mid U) = E(Y \mid X = j)$, for each value j of X
   b) For every mapping $W = f(U)$ of the observational unit variable U:
   $P_{Z=z}(X = j \mid W) = P_{Z=z}(X = j)$ or $E_{X=j}(Y \mid W) = E(Y \mid X = j)$, for each value j of X
   c) For every mapping $W = f(U)$ of the observational unit variable U:
   $E_{Z=z}(Y \mid X = j) = E_{Z=z}[E_{X=j}(Y \mid W)]$

3 conditions implying unconfoundedness:

1. Each of the following conditions is sufficient for **unconfoundedness of the treatment regression** $E(Y \mid X)$:
   a) Independence of units and treatment: $X \perp U$
   b) Unit treatment-homogeneity: $E(Y \mid X, U) \neq E(Y \mid X)$
   c) Strong causality: For every mapping $W = f(U)$, there exists a function h such that
   $E(Y \mid X, W) = E(Y \mid X) + h(W)$
2. Each of the following conditions is sufficient for **unconfoundedness of the covariate-treatment-regression** $E(Y \mid X, Z)$:
   a) Conditional independence of units and treatment: $X \perp U \mid Z$
   b) Conditional unit treatment-homogeneity: $E(Y \mid X, Z, U) \neq E(Y \mid X, Z)$
   c) Conditional strong causality: For every mapping $W = f(U)$, there exists a function h such that
   $E(Y \mid X, Z, W) = E(Y \mid X, Z) + h(W, Z)$

Note that each of the conditions (a) to (c) is sufficient but not necessary for unconfoundedness of the treatment regression $E(Y \mid X)$, i.e., each of them implies unconfoundedness of $E(Y \mid X)$, but unconfoundedness of $E(Y \mid X)$ implies neither (a), nor (b), nor (c). The same relation between the conditions (a) to (c) holds for the covariate treatment regression $E(Y \mid X, Z)$!

Implications of unconfoundedness of the treatment regression $E(Y \mid X)$:

- *Implications for causal inference*
  1. The treatment regression $E(Y \mid X)$ and the prima facie effects PFE_{jk} in the total population are unbiased, i.e.: $PFE_{jk} = ACE_{jk}$
  2. The treatment regression $E_{W=w}(Y \mid X)$ and the prima facie effects $PFE_{jk;W=w}$ in each subpopulation represented by $W = w$ are unconfounded and unbiased, i.e.: $PFE_{jk;W} = ACE_{jk;W}$
  3. For each mapping $W = f(U)$ of the observational-unit variable U: $PFE_{jk} = E[PFE_{jk;W}] = E[E_{X=j}(Y \mid W) - E_{X=k}(Y \mid W)]$
- *Implications for testing unconfoundedness*
  1. For every mapping $W = f(U)$ of the observational unit variable U: (a) $P(X = j \mid U) = P(X = j)$ or (b) $E_{X=j}(Y \mid U) = E(Y \mid X = j)$
  2. For each mapping $W = f(U)$ of the observational-unit variable U: $E(Y \mid X = j) = E[E_{X=j}(Y \mid W)]$

Implications of unconfoundedness of the covariate -treatment regression $E(Y \mid X, Z)$:

- *Implications for causal inference*
  1. The treatment regression $E(Y \mid X, Z)$ and the prima facie effect function $PFE_{jk;Z}$ in the total population are unbiased, i.e.: $PFE_{jk;Z} = ACE_{jk;Z}$
  2. The covariate-treatment regression $E_{W=w}(Y \mid X, Z)$ and the prima facie effect functions $PFE_{jk;Z,W=w}$ are unconfounded and unbiased in each subpopulation represented by $W = w$, i.e.: $PFE_{jk;Z,W=w} = ACE_{jk;Z,W=w}$
  3. For each mapping $W = f(U)$ of the observational-unit variable U: $PFE_{jk;Z=z} = E_{Z=z}[PFE_{jk;Z=z,W}] = E_{Z=z}[E_{X=j,Z=z}(Y \mid W) - E_{X=k,Z=z}(Y \mid W)]$ for P_z-almost all values z of Z.
- *Implications for testing unconfoundedness*
  1. For every mapping $W = f(U)$ of the observational unit variable U: (a) $P_{Z=z}(X = j \mid U) = P_{Z=z}(X = j)$ or (b) $E_{X=j;Z=z}(Y \mid U) = E_{Z=z}(Y \mid X = j)$ for P_z-almost all values z of Z.
  2. For each mapping $W = f(U)$ of the observational-unit variable U: $E_{Z=z}(Y \mid X = j) = E_{Z=z}[E_{X=j;Z=z}(Y \mid W)]$ for P_z-almost all values z of Z.

12 Adjusting for bias

The theory of causal effects presented here is the theoretical foundation for a number of data analysis procedures, which aim at estimating the average causal effects. These may even be applied when the prima facie effects are biased! Examples are the analysis of covariance, matching, propensity score analysis, and weighting the outcome variable.

Procedures for adjusting bias.

The procedures for adjusting bias can be classified in two major categories:

1. **Procedures using a covariate Z:**
   a) Choose a (possibly multivariate) covariate Z such that the covariate-treatment regression $E(Y \mid X, Z)$ is unbiased. The prima facie effect functions $PFE_{jk}(Z)$ are also the conditional ACE-functions $ACE_{jk}(Z)$, and the expectations of the PFE-functions $PFE_{jk}(Z)$ are also the average causal effects ACE_{jk} of treatment j compared to treatment k.
   b) Choose a (possibly multivariate) covariate Z such that treatments and the vector of propensities $\pi \equiv (\pi_1, ..., \pi_J)$ are conditionally independent given Z, where $\pi_1 = P(X = j \mid Z)$. The prima facie effect functions $PFE_{jk}(\pi)$ are also the conditional ACE-functions $ACE_{jk}(\pi)$, and the expectations of the PFE-functions $PFE_{jk}(\pi)$ are also the average causal effects ACE_{jk} of treatment j compared to treatment k.
2. **Procedures using the individual treatment probabilities:**
   a) Replace the outcome variable Y by

$$Y_w \quad \equiv \quad Y \cdot \sum_{j=0}^{J} I_{X=j} \cdot \frac{P(X = j)}{P(X = j|U)}.$$

   The prima facie effect $E(Y_w \mid X = j) - E(Y_w \mid X = k)$ of Y_w is the average causal effect ACE_{jk} of treatment j compared to treatment k on the original outcome variable Y.
   b) Use the individual treatment probability functions $\varphi_j = P(X = j \mid U)$ as covariates in the regression $E(Y|X, \varphi)$, where $\varphi \equiv (\varphi_1, ..., \varphi_J)$. This regression is always unbiased. The prima facie effect functions $PFE_{jk}(\varphi)$ are also the conditional ACE-functions $ACE_{jk}(\varphi)$, and the expectations of the PFE-functions $PFE_{jk}(\varphi)$ are also the average causal effects ACE_{jk} of treatment j compared to treatment k.

The different procedures are explained in detail in the following sections.

12.1 Procedures using a covariate Z

This class of procedures is based on choosing a covariate Z such that the covariate-treatment regression $E(Y \mid X, Z)$ is unbiased.

This following adjustment theorem is the theoretical foundation of procedures using a covariate Z for adjusting bias.

Adjustment theorem

- **If the covariate-treatment regression $E(Y \mid X, Z)$ is unbiased, then the following general implications apply**:
  1. Regressions of the true outcomes on the covariate Z within treatments:
  $$E(\tau_j \mid Z) \quad = \quad E_{X=j}(Y \mid Z) \qquad\qquad j = 0, 1, \ldots, J$$
  2. Z-conditional ACE-functions:
  $$ACE_{jk}(Z) \quad = \quad PFE_{jk}(Z) \qquad \text{for each pair } (j,k), \quad , j \neq k, \quad j,k = 0, 1, \ldots, J$$
  3. $(Z = z)$-conditional ACEs:
  $$ACE_{jk;Z=z} \quad = \quad PFE_{jk;Z=z} \qquad \text{for each pair } (j,k), \quad , j \neq k, \quad j,k = 0, 1, \ldots, J$$
  4. Expectations of the true outcomes:
  $$E(\tau_j) \quad = \quad E[E_{X=j}(Y \mid Z)] \qquad\qquad j = 0, 1, \ldots, J$$
  5. Average causal effects in the total population:
  $$ACE_{jk} \quad = \quad E[PFE_{jk}(Z)] \qquad \text{for each pair } (j,k), \quad , j \neq k, \quad j,k = 0, 1, \ldots, J$$
- **Furthermore, if true outcomes are conditionally regressively independent from treatments given Z, i.e.**
  $$E(\tau_j \mid X, Z) \quad = \quad E(\tau_j \mid Z), \qquad \textbf{for each pair } j = 0, 1, \ldots, J,$$
  the following additional implications hold:
  1. Baseline bias of the conditional PFE-functions $PFE_{jk}(Z)$:
  $$\textit{baseline bias}_{jk}(Z) \quad = \quad 0 \qquad \text{for each pair } (j,k), \quad , j \neq k, \quad j,k = 0, 1, \ldots, J$$
  2. Effect bias of the conditional PFE-functions $PFE_{jk}(Z)$:
  $$\textit{effect bias}_{jk}(Z) \quad = \quad 0 \qquad \text{for each pair } (j,k), \quad , j \neq k, \quad j,k = 0, 1, \ldots, J$$
  3. Regressions of the true outcomes on X:
  $$E(\tau_j \mid X) \quad = \quad E[E(\tau_j \mid Z) \mid X], \qquad \text{for each pair } j = 0, 1, \ldots, J$$
  4. Conditional ACE-functions given the treatment variable X:
  $$ACE_{jk}(X) \quad = \quad E[PFE_{jk}(Z) \mid X] \qquad \text{for each pair } (j,k), \quad , j \neq k, \quad j,k = 0, 1, \ldots, J$$

Adjusted means.
If the covariate-treatment regression $E(Y \mid X, Z)$ is unbiased, the regression $E(\tau_j \mid Z)$ of the true outcomes τ_j under treatment j on the covariate Z is equal to the regression $E_{X=j}(Y \mid Z)$ of the outcome variable Y on the covariate within treatment j. This equality can be used for the computation of the expectations $E(\tau_j)$,

$$E(\tau_j) = E[E_{X=j}(Y \mid Z)] \qquad j = 0, 1, \ldots, J.$$

The expectations $E[E_{X=j}(Y \mid Z)]$ (for each $j = 0, 1, \ldots, J$) are also called adjusted means of the outcome variable Y or means of Y adjusted for the covariate Z. These are the estimates of the expected true outcomes, under the condition that $E(Y \mid X, Z)$ is unbiased! In empirical applications, it is typically these expectations $E(\tau_j)$ that are of interest, and not the biased expectations $E(Y \mid X = j)$ of the outcome variable in the treatment conditions.

12.1.1 Analysis of covariances (ANCOVA)

In traditional analysis of covariance, the effect estimate of the treatment variable is an estimate of the average causal effect ACE_{jk} and the adjusted means are estimates of the expectations $E(\tau_j)$, provided that the regression $E(Y \mid X, Z)$ is unbiased and the model assumptions about the functional form of the regression $E(Y \mid X, Z)$ hold. The model of this type of ANCOVA can be expressed as a linear regression model with indicator variables $I_{X=j}$ for J treatment groups:

$$E(Y \mid X, Z) = \gamma_{00} + \gamma_{01} Z + \gamma_{10} \cdot I_{X=1} + \gamma_{J0} \cdot I_{X=J}.$$

Note: There is no product term between Z (or functions of Z) and the indicator variables for the values of X. This means that **no interactions (in the sense of analysis of variance) are allowed** between the covariate and the treatment variable!

Adjusted means in traditional ANCOVA.

$$E(\tau_0) = E[E_{X=0}(Y \mid Z)] = E(\gamma_{00} + \gamma_{01} \cdot Z) = \gamma_{00} + \gamma_{01} \cdot E(Z)$$
$$E(\tau_j) = E[E_{X=j}(Y \mid Z)] = E(\gamma_{00} + \gamma_{01} \cdot Z + \gamma_{j0}) = (\gamma_{00} + \gamma_{j0}) + \gamma_{01} \cdot E(Z) \qquad j = 1, \ldots, J$$

Average causal effects in traditional ANCOVA.

$$\begin{aligned} ACE_{j0} &= PFE_{j0} = E[E_{X=j}(Y \mid Z) - E_{X=0}(Y \mid Z)] = E(\tau_j) - E(\tau_0) \\ &= \gamma_{00} + \gamma_{j0} + \gamma_{01} \cdot E(Z) - \gamma_{00} + \gamma_{01} \cdot E(Z) \\ &= \gamma_{j0} \qquad j = 1, \ldots, J \end{aligned}$$

12.1.2 Generalized ANCOVA

Contrary to the traditional ANCOVA, the **generalized analysis of covariances does allow for interaction** between the covariates and the treatment variable!

The fundamental equation for the generalized analysis of covariance is:

$$E(Y \mid X, Z) = g_0(Z) + g_1(Z) \cdot I_{X=1} + \ldots + g_J(Z) \cdot I_{X=J}$$

The intercept function $g_0(Z)$ and the effect functions $g_1(Z), ..., g_J(Z)$ are unknown functions of the (possibly multivariate) covariate Z, which need to be specified and estimated in applications.

Adjusted means in generalized ANCOVA.

$$
\begin{aligned}
E(\tau_0) &= E[E_{X=0}(Y\,|\,Z)] = E[g_0(Z)] \\
E(\tau_j) &= E[E_{X=j}(Y\,|\,Z)] = E[g_0(Z) + g_1(Z)] \qquad j = 1, \ldots, J
\end{aligned}
$$

Average causal effects in generalized ANCOVA.

$$
\begin{aligned}
ACE_{j0} &= PFE_{j0} = E[E_{X=j}(Y\,|\,Z) - E_{X=0}(Y\,|\,Z)] = E(\tau_j) - E(\tau_0) \\
&= E[g_0(Z) + g_1(Z)] - E[g_0(Z)] \\
&= E[g_1(Z)] \qquad j = 1, \ldots, J
\end{aligned}
$$

Conditional average causal effects in generalized ANCOVA.
The functions $g_j(Z)$, $j = 1, ..., J$, are called effect functions, because a value of such a function is the prima facie effect of treatment j compared to the control ($X = 0$) given a value z of the covariate Z.

$$
g_j(Z) = E(Y\,|\,X = j, Z = z) - E(Y\,|\,X = j, Z = z) = PFE_{jk;Z=z} \qquad j = 1, \ldots, J
$$

Such a conditional prima facie effect is also the average causal effect of treatment j compared to the control given the value z of the covariate Z, provided that $E(Y\,|\,X, Z)$ is unbiased!

$$
\begin{aligned}
ACE_{j0;Z=z} &= PFE_{j0;Z=z} = E_{Z=z}(Y\,|\,X = j) - E_{Z=z}(Y\,|\,X = 0) = E_{Z=z}(\tau_j) - E_{Z=z}(\tau_0) \\
&= g_j(z)
\end{aligned}
$$

12.1.3 Generalized ANCOVA with latent variables

The generalized analysis of covariance described in the last subsection can also be extended to latent covariates and/or latent outcome variables. This means that we may also analyze conditional and average effects given a latent covariate or (a vector of) several latent covariates.

The covariate-treatment regression with the uni- or multivariate latent covariate ξ are:

$$
E(Y\,|\,X, \xi) = g_0(\xi) + g_1(\xi) \cdot I_{X=1} + \ldots + g_J(\xi) \cdot I_{X=J}.
$$

Measurement model. The use of latent variables always needs the specification of a *measurement model*. The measurement model reflects the relation of the latent variable, that is intended to be measured, and the manifest variables in terms of regressions, $E(Z_m\,|\,\xi)$. Z_m ($m = 1, \ldots, M$) denotes one of M manifest variables which is indicative for the latent covariate ξ! The specification of the measurement model depends on the kind of data and the number of variables. If the regressions $E(Z_m\,|\,\xi)$ are linear, the measurement models of classical test theory are suited! These can be written in the following general form:

$$
\mathbf{Z} = \boldsymbol{\alpha} + \boldsymbol{\Lambda}\boldsymbol{\xi} + \boldsymbol{\zeta},
$$

where $\boldsymbol{\xi}$ is a column vector with P covariates ξ_p. $\boldsymbol{\alpha}$ is a $M \times 1$ vector of the measurement intercepts and $\boldsymbol{\zeta}$ a $M \times 1$ vector of residuals. $\boldsymbol{\Lambda}$ is the $M \times P$ Matrix of regression coefficients (the loadings in terms of factor analysis).

Latent outcome variables. Latent outcome variables can be specified and used just in the same way as latent covariates. This means that all analyses considered before, i.e., the analysis of conditional and average causal effects given manifest covariates or latent covariates, can also be done with a latent outcome variable.

Adjusted means in generalized ANCOVA with the latent covariate ξ.

$$E(\tau_0) = E[E_{X=0}(Y \mid \xi)] = E[g_0(\xi)]$$
$$E(\tau_j) = E[E_{X=j}(Y \mid \xi)] = E[g_0(\xi) + g_1(\xi)] \qquad j = 1, \dots, J$$

Average causal effects in generalized ANCOVA with the latent covariate ξ.

$$\begin{aligned} ACE_{j0} &= PFE_{j0} = E[E_{X=j}(Y \mid \xi) - E_{X=0}(Y \mid \xi)] = E(\tau_j) - E(\tau_0) \\ &= E[g_0(\xi) + g_1(\xi)] - E[g_0(\xi)] \\ &= E[g_1(\xi)] \qquad j = 1, \dots, J \end{aligned}$$

Conditional average causal effects in generalized ANCOVA with the latent covariate ξ.
The functions $g_j(\xi)$, $j = 1, ..., J$, are called effect functions.

$$g_j(\xi) = E(Y \mid X = j, \xi) - E(Y \mid X = j, \xi) = PFE_{jk;}(\xi) \qquad j = 1, \dots, J$$

Such a conditional prima facie effect function is also the average causal effect function of treatment j compared to the control given the covariate ξ, provided that $E(Y \mid X, \xi)$ is unbiased!

$$\begin{aligned} ACE_{j0}(\xi) &= PFE_{j0}(\xi) = E(Y \mid X = j, \xi) - E(Y \mid X = 0, \xi) = E(\tau_j \mid \xi) - E(\tau_0 \mid \xi) \\ &= g_j(\xi) \end{aligned}$$

12.2 Procedures using the individual treatment probabilities

In this class of procedures the individual treatment probabilities are used for adjustment. These are determined by the experimenter in the true experiment or have to be estimated using relevant covariates Z.

12.2.1 Individual treatment probabilities

One adjustment method uses the probabilities $P(X = j \mid U = u)$, $j = 1, \dots, J$ for the adjustment, which are the values of the treatment probability function $\varphi_j \equiv P(X = j \mid U)$. These individual treatment probability functions can replace the covariates in the Adjustment Theorem!

The treatment probability functions φ_j have two favorable properties:

1. The Regression $E(Y \mid X, \varphi)$, where $\varphi = (\varphi_1, \dots, \varphi_J)$, is always unbiased! Therefore the following equations hold:

$$E(\tau_j) = E[E_{X=j}(Y \mid \varphi)] \qquad \text{for each } j = 1, \dots, J$$

$$ACE_{jk} = E[PFE_{jk}(\varphi)] \qquad \text{for each } j = 1, \dots, J.$$

   Hence, the Adjustment Theorem will also apply to φ taking the role of the covariate.

2. Using the individual treatment probabilities in adjustment procedures can be useful, because, compared to conditioning on covariates, they *may reduce the number of variables* with respect to which we have to condition in order to have causal unbiasedness.

The crucial prerequisite is that φ is known. But we will extend these ideas to the case in which φ can be computed (in empirical applications: estimated) from a (possibly multivariate) covariate Z.

12.2.2 Propensity scores

In *quasi-experiments*, the individual treatment probabilities $P(X = j \,|\, U = u)$ are usually unknown. This is what distinguishes a quasi-experiment from the true experiment in between-group designs. Nevertheless, in empirical applications, we may try to estimate the individual treatment probabilities $P(X = j \,|\, U = u)$ via covariates.

If all covariates which are determining the individual treatment probability are taken into account, the following equation holds:

$$P(X = j \,|\, U = u) = P(X = j \,|\, Z = z) \qquad \text{for each } j = 1, \dots, J,$$

whereby the values $P(X = j \,|\, Z = z)$ of the treatment probability function $P(X = j \,|\, Z)$ are called **propensity scores** of treatment j given $Z = z$. Furthermore, we call

$$P(X = j \,|\, Z) = \pi_j \qquad \text{for each } j = 1, \dots, J$$

the **propensity of treatment** j. Hence, we can gather the propensities in the vector $\pi \equiv (\pi_1, \dots \pi_J)$, keeping in mind: $\pi_0 = 1 - \sum_{j=1}^{J} \pi_j$.

If propensities have to be estimated, there are two sources of errors:

1. The first source of error concerns the **selection of the covariates**. The critical question is: "Did we select all relevant covariates such that treatments and true outcomes are independent given these covariates?"
2. The second source of error is in the **specification of the function** $P(X = j \,|\, Z)$. The critical question is: "Did we correctly specify the function P(X=j | Z)?"

An example of the estimation of the propensity of treatment j: If we assume that the propensity is a linear logistic regression, the Propensity of treatment j can be estimated by

$$P(X = j \,|\, Z) = \frac{exp(\alpha_0 + \alpha_1 \cdot Z)}{1 + exp(\alpha_0 + \alpha_1 \cdot Z)}.$$

The linearity of the logistic regression is an assumption, which may be wrong.

If all covariates which are determining the individual treatment probability are taken into account, the same favorable properties that hold for φ_j, also hold for π_j: Particularly, the regression $E(Y \,|\, X, \pi_j)$ is always unbiased! Therefore the following equations hold:

$$E(\tau_j) = E[E_{X=j}(Y \,|\, \pi)] \qquad \text{for each } j = 1, \dots, J$$

$$ACE_{jk} = E[PFE_{jk}(\pi)] \qquad \text{for each } j = 1, \dots, J.$$

Hence, the Adjustment Theorem also applies to π being the covariate.

12.3 Weighting the outcome variable

The average causal effects may also be computed using weighted outcome variables. In order to do this, the original outcome variable Y needs to be replaced by the weighted outcome variable Y_w,

$$Y_w = Y \cdot \left(\sum_{j=0}^{J} I_{X=j} \cdot \frac{P(X = j)}{P(X = j \,|\, U)} \right).$$

The weighting variable, which is the term in parentheses, is multiplied with the outcome variable Y. It is a function of X and U. The crucial ingredients of this weighting variable are the individual treatment probabilities $\varphi_j(u) = P(X = j \,|\, U = u)$ of the the conditional probability functions $\varphi_j = P(X = j \,|\, U)$.

The Weighting Theorem for the general case of $J + 1$ treatment conditions.

Let X and Y be the random variables defined in the single unit trial. The outcome variable Y has a finite expectation $E(Y)$ and a finite variance $Var(Y)$. The treatment variable X is finite with values $0, 1, ..., J$, and it is assumed that $0 < P(X = j \,|\, U = u) < 1$, for each pair (j, u) of values of X and U. If this conditions hold, the following equations are always true:

$$E(Y_w \,|\, X = j) \quad = \quad E(\tau_j) \qquad \text{for each } j = 1, \dots, J$$

and

$$\begin{aligned} E(Y_w \,|\, X = j) - E(Y_w \,|\, X = k) &= E(\tau_j) - E(\tau_j) \\ &= ACE_{jk} \qquad \text{for each pair } (j, k), \quad j \neq k, \quad j = 1, \dots, J. \end{aligned}$$

For using this procedure the individual treatment probabilities $P(X = j \,|\, U = u)$ either need to be known or have to be estimated using covariates that determine the individual treatment probabilities.