\[ \text{ATE} = E_x \left( \frac{(d_i - p(x_i)) y_i}{p(x_i) E[1-p(x_i)]} \right) \]

\text{Numerator}

\[ = (d_i - p(x_i)) y_i \]

\[ = (d_i - p(x_i)) (d_i y_{i1} + (1-d_i) y_{i0}) \]

\[ = d_i y_{i1} - p(x_i) d_i y_{i1} - (1-d_i) p(x_i) y_{i0} \]

\[ = d_i y_{i1} (1-p(x_i)) - p(x_i) (1-d_i) y_{i0} \]

For

\[ m_t(x_i) = E[y_{i1} \mid x_i], \]

\[ m_0(x_i) = E[y_{i0} \mid x_i] \]

\[ E \left( \text{Numerator} \mid x_i, d_i \right) \]

\[ = E \left[ y_{i1} \mid d_i, x_i \right] d_i (1-p(x_i)) - E[y_{i0} \mid x_i, y_{i1}] p(x_i) (1-d_i) \]

\[ = m_t(x_i) d_i (1-p(x_i)) - m_0(x_i) (1-d_i) p(x_i) \]
\[ E \left[ \text{Numerator} \mid x_i \right] \]

\[ = m(x_i) \cdot p(x_i) / p(x_i) - m_0(x_i) (1 - d_i) \cdot p(x_i) \]

\[ = \left( m(x_i) - m_0(x_i) \right) \cdot p(x_i) (1 - p(x_i)) \]

Therefore:

\[ ATE \]

\[ = E_x \left( m(x_i) - m_0(x_i) \right) \]

\[ = E \left[ E \left[ y_{i1} \mid x_i \right] - E \left[ y_{i0} \mid x_i \right] \right] \]

\[ = \left( E y_{i1} - E y_{i0} \right) \]
from Robins & Rosenbaum

Lemma. If \( Y_i, Y_0 \perp D \mid X \)

( unconfoundedness )

and \( b(x) \) is a balancing score,

meaning that \( D \perp X \mid b(x) \)

then \( Y_i, Y_0 \perp D \mid b(x) \)

Proof: Want to show

\[
\begin{align*}
f(Y_i \mid D=1, b(x)) &= f(Y_i \mid b(x)) & (1) \\
f(Y_0 \mid D=0, b(x)) &= f(Y_0 \mid b(x)) & (2) \\
f(Y_i \mid D=1, b(x)) &= f(Y_0 \mid b(x)) & (3) \\
f(Y_0 \mid D=0, b(x)) &= f(Y_0 \mid b(x)) & (4)
\end{align*}
\]
Consider (1) first

\[ f(y, \{D=1, b(x)\}) \]

= \int f(y \mid D=1, x) f(x \mid D=1, b(x)) \, dy

\[ \uparrow \text{unconfoundedness} \quad \uparrow \text{definition of balancing score} \]

= \int f(y \mid x) f(x \mid b(x)) \, dy

\[ \downarrow \text{x is more informative than } b(x) \]

= \int f(y \mid x, b(x)) f(x \mid b(x)) \, dx \, dy

= f(y \mid b(x)) \quad \text{Law of Iterated Expectation.}

These other three relations are similar.
\[(2)\]
\[
f(y_i \mid D = 0, b(x))
\]
\[
= \int f(y_i \mid D = 0, x) f(x \mid D = 0, b(x)) \, dx
\]
\[
= \int f(y_i \mid D = 0, x) f(x \mid D = 0, b(x)) \, dx
\]
\[
= \int f(y_i \mid x, b(x)) f(x \mid b(x)) \, dx
\]
\[
= f(y_i \mid b(x))
\[ (2) \]
\[
f(Y_0 | D=1, b(x))
\]
\[
= \int f(Y_0 | D=1, X, b(x)) f(x | D=1, b(x)) \, dx
\]
\[
= \int f(Y_0 | D=1, X) f(x | D=1, b(x)) \, dx
\]
\[
= \int f(Y_0 | X) f(x | b(x)) \, dx
\]
\[
= \int f(Y_0 | X, b(x)) f(x | b(x)) \, dx
\]
\[
= \int f(Y_0 | b(x))
\]
(4) \[ f ( \gamma_0 \mid D=0, b(x)) = \int f(y_0 \mid D=0, x, b(x)) f(x \mid D=0, b(x)) \, dx \]

= \int f(y_0 \mid D=0, x) f(x \mid D=0, b(x)) \, dx

= \int f(y_0 \mid x) f(x \mid b(x)) \, dx

= \int f(y_0 \mid x, b(x)) f(x \mid b(x)) \, dx

= f (\gamma_0 \mid b(x))
\[ E[y_1 - y_0 \mid p(x) = 0.3] \]

\[ = E[y_1 \mid D=1, p(x) = 0.3] - E[y_1 \mid D=0, p(x) = 0.3] \]

\[ = E[y \mid D=1, p(x) = 0.5] - E[y \mid D=0, p(x) = 0.5] \]