Therefore, we have proved that:

$$\hat{\text{ATE}} = \hat{\mu}_{HT} = \frac{1}{N}\sum_{i=1}^{N}\frac{y_i}{\pi_i}$$

The *Inverse Probability Reweighting* estimation of ATE is thus equivalent to the Horvitz–Thompson estimator, due to Daniel G. Horvitz and Donovan J. Thompson in 1952. In sampling theory it is a method for estimating the *total* and *mean* of a super-population in a stratified sample. Inverse probability weighting is applied to account for "different proportions of observations within strata in a target population". The Horvitz-Thompson estimator is frequently applied in survey analyses and can be used also to account for missing data.

Similarly, we can also calculate the ATET by considering that:

$$[w - p(\mathbf{x})]\, y = [w - p(\mathbf{x})] \cdot [y_0 + w \cdot (y_1 - y_0)] = [w - p(\mathbf{x})] \cdot y_0 + w \cdot [w - p(\mathbf{x})] \cdot (y_1 - y_0) =$$
$$= [w - p(\mathbf{x})] \cdot y_0 + w \cdot [1 - p(\mathbf{x})] \cdot (y_1 - y_0)\text{ , since } w^2 = w.$$

Thus, by dividing the previous expression by $[1 - p(\mathbf{x})]$:

$$\frac{[w - p(\mathbf{x})]y}{[1 - p(\mathbf{x})]} = \frac{[w - p(\mathbf{x})]y_0}{[1 - p(\mathbf{x})]} + w(y_1 - y_0) \tag{4}$$

Consider now the quantity $[w - p(\mathbf{x})]y_0$ in the RHS of (4). We have that:

$$[w - p(\mathbf{x})]\, y_0 = \text{E}\{[w - p(\mathbf{x})]\, y_0 \mid \mathbf{x}\} = \text{E}(\text{E}\{[w - p(\mathbf{x})]\, y_0 \mid \mathbf{x}, w\} \mid \mathbf{x}) = \text{E}([w - p(\mathbf{x})] \cdot \text{E}\{y_0 \mid \mathbf{x}, w\} \mid \mathbf{x}) =$$
$$\text{E}([w - p(\mathbf{x})] \cdot \text{E}\{y_0 \mid \mathbf{x}\} \mid \mathbf{x}) = \text{E}([w - p(\mathbf{x})] \cdot g_0(\mathbf{x}) \mid \mathbf{x}) = g_0(\mathbf{x}) \cdot \text{E}([w - p(\mathbf{x})] \mid \mathbf{x}) = g_0(\mathbf{x}) \cdot [\,\text{E}(w \mid \mathbf{x}) -$$
$$\text{E}(p(\mathbf{x}) \mid \mathbf{x})] = g_0(\mathbf{x}) \cdot [p(\mathbf{x}) - p(\mathbf{x})] = 0.$$

Taking relation (3), and applying the expectation conditional on **x**:

$$\text{E}\left\{\frac{[w - p(\mathbf{x})]y}{[1 - p(\mathbf{x})]} \mid \mathbf{x}\right\} = \text{E}\left\{\frac{[w - p(\mathbf{x})]y_0}{[1 - p(\mathbf{x})]} \mid \mathbf{x}\right\} + \text{E}\{w(y_1 - y_0) \mid \mathbf{x}\} = \text{E}\{w(y_1 - y_0) \mid \mathbf{x}\}$$

since we proved that $[w - p(\mathbf{x})]\, y_0$ is zero. By LIE 1, we can get that:

$$\begin{cases} \text{E}_{\mathbf{x}}\text{E}\left\{\frac{[w - p(\mathbf{x})]y}{[1 - p(\mathbf{x})]} \mid \mathbf{x}\right\} = \text{E}\left\{\frac{[w - p(\mathbf{x})]y}{[1 - p(\mathbf{x})]}\right\} \\ \text{E}_{\mathbf{x}}\text{E}\{w(y_1 - y_0) \mid \mathbf{x}\} = \text{E}\{w(y_1 - y_0)\} \end{cases}$$