# 1    Notation

Let $C$ be a set of observed covariates, $A$ be an indicator of treatment, and $Y_a$ be the outcome observed under $A = a$. The observed outcome is given by $Y = AY_1 + (1 - A)Y_0$.

We are interested in comparing features of the marginal distributions of the potential outcomes. Let $f_{Y_a}$ denote the marginal distribution of potential outcome $Y_a$ (with CDF given by $F_{Y_a}$). We are interested in features of these distributions, which we will either describe through the quantiles or the mean.

$$E(Y_a) = \int y f_{Y_a}(y) \partial y$$

$$Q_{Y_a}(\tau) = F_{Y_a}^{-1}(\tau)$$

Examples of comparisons are the average causal effect (ATE), $E(Y_1 - Y_0)$, the median causal effect, $Q_{Y_1}(0.5) - Q_{Y_0}(0.5)$, and more generally, the $p^{th}$ quantile causal effect, $Q_{Y_1}(p) - Q_{Y_0}(p)$.

A few notes. First, the expected value is the integral of the quantile function:

$$E(Y_a) = \int_0^1 Q_{Y_a}(p) \partial p$$

Second, if we view the quantile causal effect as a function, then we can consider estimating the quantile effect for all $p$:

$$QE(p) = Q_{Y_1}(p) - Q_{Y_0}(p)$$

Finally, combining the first two notes, we get:

$$ATE = E(Y_1 - Y_0) = \int_0^1 QE(p) \partial p$$

# 2    Main assumptions

The primary assumption will be strong ignorability,

$$(Y_1, Y_0) \perp A | C$$

Based on Rubin, we know that this implies

$$(Y_1, Y_0) \perp A | e(C)$$

where $e(C) = P(A = 1|C)$ is known as the propensity score. The implication of strong ignorability is that the potential outcomes are independent of the treatment assignment conditional on the propensity score. That is, conditional on the propensity score, we can ignore the treatment assignment mechanism when describing features of the potential outcomes. In most applications of causal inference, we focus on the ATE. However, strong ignorability makes a statement about the entire distribution of the potential outcomes, so that we can make comparisons between any feature of the potential outcomes of interest (e.g. quantiles).

We will also need SUTVA. No big deal.

# 3   Local quantile regression

We will use local quantile regression to estimate $QE(p)$ around a given value of the propensity score. In other words, we will estimate the conditional quantile effect $QE(p|e(C) = \psi)$ using local quantile regression, where locality is based on the propensity score.

For a given value of the propensity score $\psi$, we seek to minimize

$$\sum_i w_i \rho_\tau(Y_i - f(X_i))$$

where $w_i$ is a weight and $\rho_\tau(u) = u(\tau - I(u < 0))$. We choose the weight based on a kernel function, $K(.)$, and the bandwidth, $h$. Specifically, $w_i = K(\frac{\psi - e(C_i)}{h})$. Since the kernel function weights more heavily the observations that have propensity score close to $\psi$, we are estimating a local (in the PS) quantile effect. I'm sure we could somehow modify strong ignorability to explicity take into account the localness of the procedure.

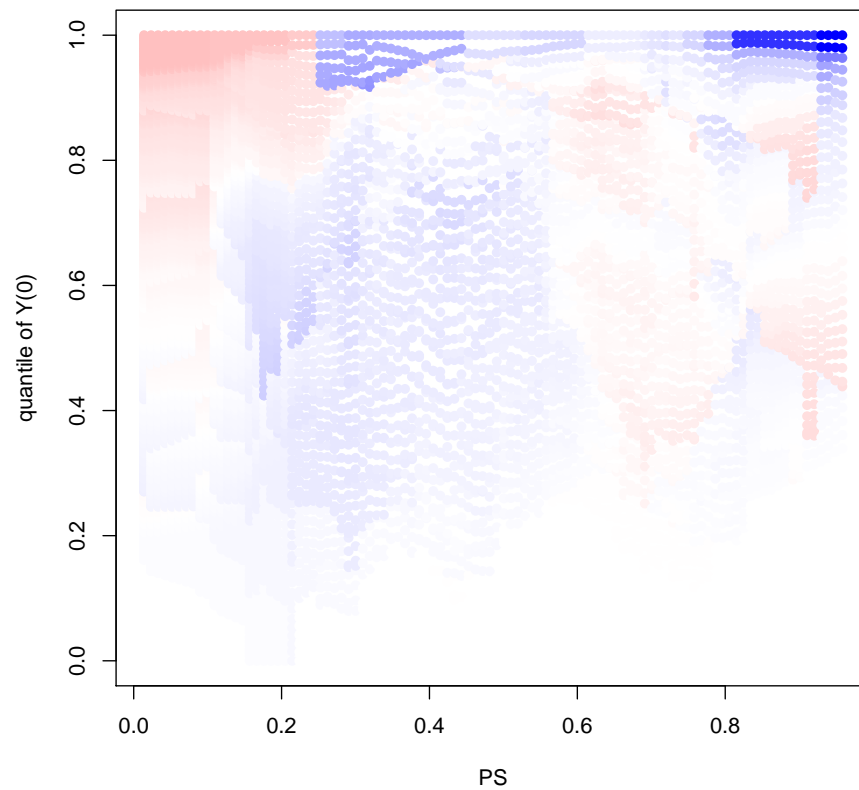So we have the $\tau^{th}$ quantile effect around $e(C) = \psi$

$$\widehat{\beta}_\tau(\psi) = \widehat{QE}(p|e(C) = \psi) = \arg\min_\beta \sum_i K(\frac{\psi - e(C_i)}{h})\rho_\tau(Y_i - \beta_0 - \beta_1 A_i)$$

# 4   R code example

First, we fit the model and suppress all output. So we don't see anything here!
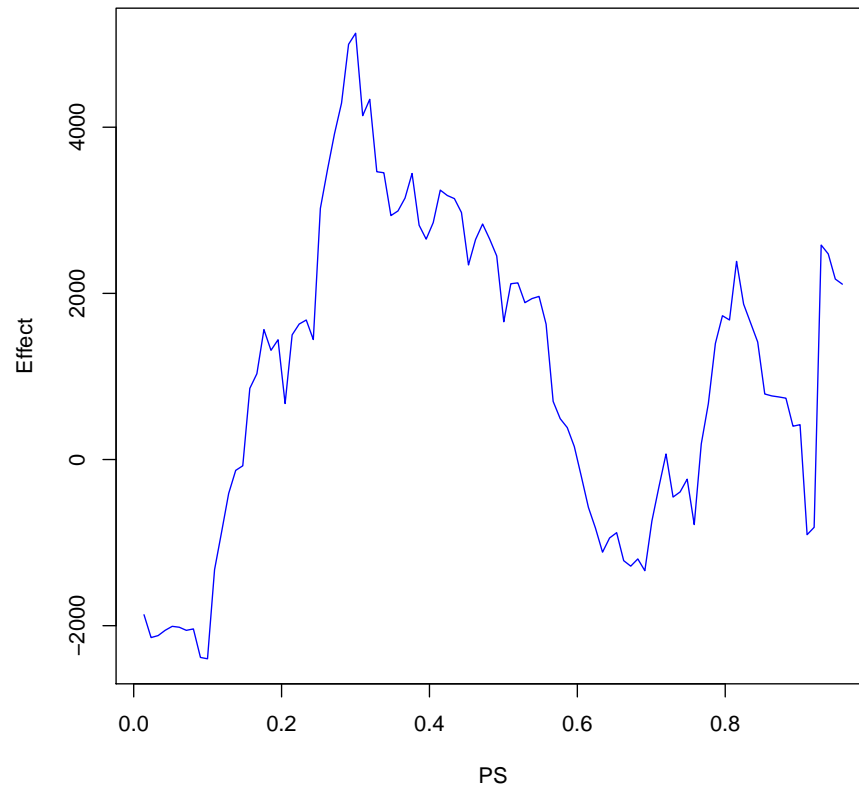
Next, we run the code to produce the heatmaps!

```
heatmap.localQ(fit=out , Ytype="quantile" , pch=16 , xlab="PS")
```

Next, we run the code to produce the mean plot!

```
plot.Qmean(fit=out , add=F , col='blue' , xlab="PS" , ylab="Effect")
```

Next, we run the code to produce the quantile plot!

```
plot.Qavg(fit=out , Ytype="quantile" , type='l' , col='blue' , xlab="quantile of Y(0)" , yla
```

4