1. Motivation.

	X covariates		A (Treatment)	R.
patient 1	x_{11} ……	x_{m1}	1	
patient 2.	x_{12} ……	x_{m2}	-1	
$\vdots$	$\vdots$		$\vdots$	
patient n.	x_{1n}, ……	x_{mn}	-1	

Training ~~set~~ set

A is randomly assigned to each patient, A could be not appropriate if we based on Training set to design machine learning, it will give us wrong conclusion for testing set.

Therefore, R (reward of patient) is introduced to measure whether A is correct or not,

previous literature, $\min P(A \neq \text{sign}(f(x)))$

In this paper, $\max_{f} R$

2. Model setup:

Covariate: X – a vector, $f_0(X)$ is unknown.

Treatment: A, $p(a|x)$ is known.

special case: $A = \begin{cases} 1 & \pi \\ -1 & 1-\pi. \end{cases}$ independent of X

Reward : R. $f_1(r \mid X, A)$ is unkmon.

ITR : $d : X \longrightarrow A$.

Joint density of (X, A, R) :

$$f(x, a, r) = f_0(x)\, p(a \mid x)\, f_1(r \mid x, a)$$

a measure on $((X, A, R), \mathcal{F}, \mathbb{P})$

Joint density of $(X, A = d(x), R)$:

$$f^d(x, a, r) = f_0(x)\, 1_{a=d(x)}\, f_1(r \mid x, a)$$

$$= f_0(x)\, f_1(r \mid x, d(x))$$

a measure on $((X, A, R), \mathcal{F}^d, \mathbb{P}^d)$

2. Goal : To maximize $E^d(R)$ w.r.t d

$$E^d(R) = \int r \, d\mathbb{P}^d = \int r \frac{d\mathbb{P}^d}{d\mathbb{P}} \cdot d\mathbb{P}$$

$$= \int r \frac{1_{a=d(x)}}{p(a \mid x)} dP$$

$$= E\left(R \frac{1_{A=d(X)}}{p(A \mid X)}\right)$$

$$= E\left[E\left(\frac{1_{A=d(X)}}{p(A \mid X)} R \,\Big|\, X\right)\right]$$

$$= E_X\left[E_A\left(E_R\left(\frac{1_{A=d(X)}}{p(A \mid X)} R \,\Big|\, A, X\right) \Big|\, X\right)\right]$$

$$= E_X \left[\sum_{a \in A} E_R \left(\frac{1_{a=d(x)}}{p(a|X)} R \,\middle|\, a, X \right) p(a|X) \,\middle|\, X \right]$$

$$= E_X \left[\sum_{a \in A} E_R (1_{a=d(x)}, R) \,\middle|\, X \right]$$

$$= E\left[E(R | X, d(x)) \right]$$

$\Rightarrow$ To maximize $E^d(R)$ is same as maximize $E(R|X.d(X))$

3. Method 1:

Renote $Q_0(X.A) = E(R|X.A)$

goal is to $\max_{d_0} Q_0(X, d_0(X))$

since find u to $\min E(X-u)^2 \Rightarrow u = EX$

$\Rightarrow$ Find $Q(X,A)$ to $\min E(R-Q(X,A))^2 \Rightarrow$
$Q(X.A) = E(R|X,A)$

Step 1: Find Q_0

To find Q_0 is to min $L(\theta) = E(R-Q(X,A))^2$

Guess $Q(X,A) = \Phi(X,A)^T \theta$

where $\Phi(X.A) = (\varphi_1(X.A), \varphi_2(X.A), \dots \varphi_m(X,A))$

φ_i is unknown

$\Rightarrow$ To estimate $\hat{\theta} \Rightarrow Q(X.A) = \Phi(X,A)^T \hat{\theta}$

Step 2: $\max Q_0 = \max_d Q(X, d(X)) = \max_d \left(\Phi(X, d(X))^T \hat{\theta} \right)$

Theorem: $E^{d_0}(R) - E^{d}(R) \le L$.

d is $\underbrace{\cancel{use}}_{\text{ITR found by}}$ above method

d_0 is true ITR

d is consistent to d_0

Method 2: Goal: $\max_d E^d(R) = E\left(\frac{1_{A=d(X)}}{P(A|X)} R \right)$

$$\Updownarrow$$

$$\min E\left(\frac{1_{A\neq d(X)}}{P(A|X)} R \right)$$

$$E\left(\frac{1_{A\neq d(X)}}{P(A|X)} R \right) = \frac{1}{n}\sum_{i=1}^{n} \frac{r_i}{P(a_i|X_i)} 1(a_i \neq d(X_i))$$

special case: $A = \begin{cases} 1 & \pi \\ -1 & 1-\pi \end{cases}$

$$\Rightarrow P(A|X) = A\pi + \frac{1-A}{2}$$

$$\Rightarrow \min_d \frac{1}{n}\sum_{i=1}^{n} \frac{r_i}{a_i\pi + \frac{1-a_i}{2}} 1(a_i \neq d(X_i)) \qquad *$$

Traditional problem is $\min P(A \neq \text{sign}(f(X))$

$$= E 1(A \neq \text{sign} f(X))$$

$$= \frac{1}{n}\sum_{i=1}^{n} 1(a_i \neq \text{sign} f(X_i))$$

if $\begin{cases} \pi = \frac{1}{2} \\ d(X_i) = \text{sign} f(X_i) \end{cases} \Rightarrow (*) = \min \frac{1}{n}\sum_{i=1}^{n} r_i 1(a_i \neq \text{sign} f(X_i))$