# Propensity Score (PS) Method

## Propensity Score:

$$e(x) = \Pr(A = 1 | X = x; \gamma)$$

- Unit's probability of being treated, conditional on $X$=$x$

Rosenbaum and Rubin (1983) showed that

$$Y^a \perp A | X = x \rightarrow Y^a \perp A | e(x)$$

Conditional on the true PS, treatment assignment is independent of potential outcomes (above) or covariates (below)

$$X \perp A | e(X)$$

Treatment status is balanced (**Balance Property**), i.e.,

$$f\big(X \big| A = 1, e(X)\big) = f\big(X \big| A = 0, e(X)\big)$$

Among those units with same PS, X is identically distributed between the treated and untreated.

We have to know the **true PS** to have all these results work!

**Estimate PS:** $\hat{e}_i = pr(A = 1|X_i; \hat{\gamma})$

In R, for example, we could calculate the propensity scores:

```
glm(treatment~var, data=mydata, family=binomial())$fitted.values
```

Q: what variables do we include in the propensity score model?

Variables that block the backdoor paths from A to Y

**Check balance** within strata of $\hat{e}_i$, covariates should be balanced

$$f(X|A = 1, \hat{e}_i) = f(X|A = 0, \hat{e}_i)$$

**Question: How to obtain ACE with estimated PS?**

- Matching with PS
- Stratification with PS (Rubin 2007 SIM)
- Inverse PS Weighting

## **Matching with PS –** covered before

Basic idea: estimated PS is used as distance metric, which maps covariates vectors into a single number – Dimension reduction from $p$ to 1

Option 1: $d(i,j) = \left|\hat{e}_i - \hat{e}_j\right|$

Option 2: $d(i,j) = \left|logit(\hat{e}_i) - logit\left(\hat{e}_j\right)\right|$

# Stratification with PS

# Stratification with estimated PS

**Step 1:** Identify boundary points:$0 = b_0 < b_1 < \cdots < b_K = 1$.

**Step 2:** Create block indicators:

$$B_i(k) = \begin{cases} 1 & \text{if } b_{k-1} < \hat{e}_i < b_k \\ 0 & \text{otherwise} \end{cases}$$

**Step 3:** Calculate effect estimates within strata

$$\tau_k = E[Y_i | A_i = 1, B_i(k) = 1] - E[Y_i | A_i = 0, B_i(k) = 1]$$

**Step 4:** Take the average of stratum-specific ACE over strata

$$\tau = \sum_{k=1}^{K} \tau_k \, P[B_i(k) = 1],$$

where $P[B_i(k) = 1]$ proportion of units in block $k \left( = \frac{\sum_{i=1}^{N} B_i(k)}{N} \right)$

# How to create strata if X (or PS) is continuous?

# Classic example: Cigars/pipes versus cigarettes

- $A_i = 1$ for pipe/cigar smokers, $A_i = 0$ for cigarette smokers,
- $Y_i$ = Death in the first year of follow up
- Naïve positive effect: cigar/pipe smokers more likely to die.
  - Pipe/cigar smokers much older than cigarette smokers
  - AGE is the confounder here

- Cochran's approach: stratify based on coarsened continuous age:
  - Divide age into $k$ strata: $S_i \in s_1, s_2, \ldots, s_k$
  - $s_1$ might be 18-25, $s_2$ might be 26-35, and so on
  - Calculate effect within strata and then aggregate.

- Key assumption: no unmeasured confounders using stratified version of age, that is, $A_i \perp (Y_i^0, Y_i^1) | S_i$

## Example – Rubin 2007 SIM

In all the U.S. Tobacco litigation,

- Compare health-care costs and disease rates of smokers and 'like' never-smokers
- Data: 1987 National Medical Expenditure Survey (NMES)
  ~ Find a sample of male never-smokers in NMES who, as a group, look just like the male current smokers in NMES.

- Table 1 in the paper provides a long list of variables that define 'similar' current smokers and never smokers
  - Include (about 50) covariates that define 'similar' current smokers and never smokers
  - Exclude variables that have no possible connections to the outcomes

Table II. Original propensity score analyses in NMES: balance between male current smokers ($N = 3510$) and all male never smokers ($N = 4297$), and between male current smokers and the roughly comparable subset of male never smokers ($N = 3510$).

| Analysis of current smokers *versus* | Distributional differences in propensity scores | | Percent of covariates with specified variance ratio after adjustment for propensity score | | |
|---|---|---|---|---|---|
| | *B* | *R* | Good | Of concern | Bad |
| All never | 1.09 | 1.00 | 57 | 34 | 9 |
| Subset never | 0.08 | 1.16 | 90 | 9 | 1 |

*Note*: $B$ = number of standard deviations between means of propensity score in current and never smokers; $R$ = ratio of current-smoker to never-smoker variance on the propensity score; also displayed is the percentage of the covariates orthogonal to the propensity score with the specified variance ratios: good = between 4/5 and 5/4; of concern = not good but between 1/2 and 2.

**B (=*smd*)**: Mean of $X_i\hat{\beta}$ in the treated group minus the mean of $X_i\hat{\beta}$ in the control group, divided by the within group standard deviation of $X_i\hat{\beta}$, defined as $\sqrt{\frac{S_t^2+S_c^2}{2}}$, where $S_t^2$ is the variance of $X_i\hat{\beta}$ in the treated group and analogously for $S_c^2$.

**First row – All never:**

- smd=1.09 implies never-smoker group includes individuals who look nothing like any current smokers with respect to the propensity score. Those individuals should simply be discarded because they carry essentially no information about what a current smoker's health outcomes might be like.
- Large percentage (43%=34%+9%) of covariates with variance ratios "of concern" or "bad" – Linear modeling may not be able to adjust for those covariates "of concern" or "bad".

**Second row – subset never**

- 3510 individuals who were chosen by a 'propensity score caliper with Mahalanobis metric matching' procedure
- Discard 787 never smokers who were least like the smokers

<u>After discarding 787 never smokers:</u>

- The treated and matched control samples are only *smd*=0.08 standard deviations apart
- Have a ratio of variances on this score equal to R=1.16, and 90% of covariates have good variance.

→Balance of the covariates are greatly improved

**Check Balance** – re-estimated the propensity score in the matched samples

Table III. Re-estimated propensity score analyses in NMES: assessing achieved balance after subclassification on re-estimated propensity scores for male current smokers ($N = 3510$) *versus* subset male never smokers ($N = 3510$).

| Number of subclasses | Distributional differences in propensity scores | | Percent of covariates with specified variance ratio after adjustment for propensity score | | |
|---|---|---|---|---|---|
| | *B* | *R* | Good | Of concern | Bad |
| 1 | 0.39 | 1.33 | 88 | 12 | 0 |
| 2 | 0.18 | 1.36 | 98 | 2 | 0 |
| 4 | 0.10 | 1.25 | 99 | 1 | 0 |
| 6 | 0.09 | 1.30 | 100 | 0 | 0 |
| 8 | 0.08 | 1.16 | 100 | 0 | 0 |
| 10 | 0.07 | 1.12 | 100 | 0 | 0 |

*Note*: $B$ = number of standard deviations between means of propensity score in current and never smokers; $R$ = ratio of current-smoker to never-smoker variance on the propensity score; also displayed is the percentage of the covariates orthogonal to the propensity score with the specified variance ratios: good = between 4/5 and 5/4; of concern = not good but between 1/2 and 2; bad = less than 1/2 or greater than 2.

For example, for the second row, the matched samples are split at the median propensity score,

- Compared treated and controls who have low propensity scores,
- Compared treated and controls who have high propensity scores,
- Weighted each comparison by the number of treated in each subclass

As the result, diagnostics for two subclasses are substantially better than the diagnostics for the first row (smd=.18<.39).

If the propensity scores are well estimated, within a narrow bin of propensity score values (i.e., within any subclass)

→ Joint distributions of all covariates $X$ that entered the propensity score estimation should be the same in the treatment and control groups in expectation (as shown in "10 subclasses").

The stratified analyses after matching:

- Subclass-specific analyses:
  - Simple: a comparison of mean total medical expenses for current smokers and never smokers within each of the ten subclasses; or
  - Regression: a comparison, adjusting for all *X* variables.

  ❖ Above 2 analyses would yield essentially the same point estimates because within each subclass, the distribution of *X* is nearly the same.

- Take the average of subclass-specific ACE weighted by the number of current smokers in the subclasses.

# Inverse PS Weighting

# Inverse PS Weighting

BASIC IDEA: turn observational study into a pseudo-randomized trial by re-weighting samples.

RECALL:

- With a completely randomized experiment (exchangeability), we can just use the difference in means to estimate ACE:

$$E(Y_i^1) - E(Y_i^0) = E(Y_i|A_i = 1) - E(Y_i|A_i = 0)$$

- In observational studies, with all confounders $X$ observed, mimicking the conditional randomized experiment

(conditional exchangeability)

$$E(Y_i^1) - E(Y_i^0) = \sum_x E(Y_i|A_i = 1, X_i = x)P(X_i = x) - \sum_x E(Y_i|A_i = 0, X_i = x)P(X_i = x)$$

Because $E(Y_i^a) = E_X[E(Y_i^a|X_i)]$ by Law of Iterated Expectation (LIE)

$=\sum_x E(Y_i^a|X_i = x)P(X_i = x)$ by definition of expectation

$=\sum_x E(Y_i^a|A_i = a, X_i = x)P(X_i = x)$ by cond. exchangeability

$=\sum_x E(Y_i|A_i = a, X_i = x)P(X_i = x)$ by consistency

In observational studies, $E(Y_i|A_i = a)$ is identifiable. What is it?

$$E(Y_i|A_i = a)$$
$$= \sum_x E(Y_i|A_i = a, X_i = x)P(X_i = x|A_i = a) \text{ by L. I. E.}$$

$$E(Y_i|A_i = a)$$
$$= \sum_x E(Y_i|A_i = a, X_i = x)\frac{\boldsymbol{P(A_i = a|X_i = x)}P(X_i = x)}{\boldsymbol{P(A_i = a)}} \text{ by Bayes' Theorem}$$

What is the difference between $E(Y_i|A_i = a)$ and $E(Y_i^a)$?

Recall: $E(Y_i^a) = \sum_x E(Y_i|A_i = a, X_i = x)P(X_i = x)$

**Question:** What if we were to reweight the data by $w_i = \frac{1}{p(A_i = a|X_i = x)}$ or $w_i^s = \frac{p(A_i=a)}{p(A_i = a|X_i = x)}$? Are the weighted treatment group and control group balanced in X distribution?

Table 1. Frequency

| | X=0 | X=1 |
|---|---|---|
| A=0 | 4 | 3 |
| A=1 | 4 | 9 |

Table 2.1 Weight

| $w_i$ | **X=0** | **X=1** |
|---|---|---|
| **A=0** | $\frac{1}{p(A=0\|X=0)} = \frac{1}{0.5} = 2$ | $\frac{1}{p(A=0\|X=1)} = \frac{1}{0.25} = 4$ |
| **A=1** | $\frac{1}{p(A=1\|X=0)} = \frac{1}{0.5} = 2$ | $\frac{1}{p(A=1\|X=1)} = \frac{1}{0.75} = 1.33$ |

Table 2.2 Stabilized weight

| $w_i^s$ | **X=0** | **X=1** |
|---|---|---|
| **A=0** | $\frac{p(A=0)}{p(A=0\|X=0)} = \frac{0.35}{0.5} = 0.7$ | $\frac{p(A=0)}{p(A=0\|X=1)} = \frac{0.35}{0.25} = 1.4$ |
| **A=1** | $\frac{p(A=1)}{p(A=1\|X=0)} = \frac{0.65}{0.5} = 1.3$ | $\frac{p(A=1)}{p(A=1\|X=1)} = \frac{0.65}{0.75} = 0.86$ |

Correspondingly,

Table 3.1 Weighted Frequency

| | X=0 | X=1 |
|---|---|---|
| A=0 | 8 | 12 |
| A=1 | 8 | 12 |

Table 3.2 Stabilized-weighted Frequency

| | X=0 | X=1 |
|---|---|---|
| A=0 | 2.8 | 4.2 |
| A=1 | 5.2 | 7.8 |

# Conclusions:

- *Weighted* frequencies sum up to the population size within each treatment group (N = 8+12 = 8+12 = 20), i.e.

$$\sum_{i=1}^{N} A_i w_i = \sum_{i=1}^{N} (1 - A_i) w_i = N$$

- *Stabilized-weighted* frequencies sum up to the population size across all treatment groups (N = 2.8+4.2+5.2+7.8 = 20), i.e.

$$\sum_{i=1}^{N} w_i^s = N$$

- Variance of Weights = var(2,2,2,2,4,4,4,2,2,2,2,1.33,...1.33)=0.844
  Variance of Stabilized Weights=
  var(.7,.7,.7,.7,1.4,1.4,1.4,1.3,1.3,1.3,1.3, .86,...,.86)=0.072
- Relationship between A and X is changed by reweighting the sample (Table 3.1 vs. Table 1; or Table 3.2 vs. Table 1)

- Distributions of $X$ are same (balanced) across treatment groups after weighting, that is, A is independent of $X$ after weighting

  Table 3.1: *weighted X* dist'n in treated vs. untreated

$$p(X = 1|A = 0) = p(X = 1|A = 1) = \frac{12}{20} = 60\%$$

  Table 3.2: *stabilized-weighted X* dist'n in treated vs. untreated

$$p(X = 1|A = 0) = \frac{4.2}{7} = p(X = 1|A = 1) = \frac{7.8}{13} = 60\%$$

Stabilized weights $\boldsymbol{w_i^s = w_i p(A_i = a)}$, where $p(A_i = a) = n_a/n$ constant for units within treatment groups and therefore won't change the $X$ dist'n within treatment groups (Table 3.1 vs. Table 3.2)

**Bridging Observed Outcomes and Potential Outcomes**

Prove: $E(w_i A_i Y_i) = E(Y_i^1)$. For a treated unit ($A_i$=1) weighted by $w_i$, the expected observed outcome is equivalent to expected potential outcome under the treatment, where $w_i = \frac{1}{p(A_i = 1|X_i = x)} = \frac{1}{e(X_i)}$.

$E(w_i A_i Y_i) = E\left(\frac{A_i Y_i}{e(X_i)}\right)$ by definition for $w_i$,

$= E_X\left[E\left(\frac{A_i Y_i}{e(X_i)}\middle| X_i = x\right)\right]$ by law of iterated expectations,

$= E_X\left[\frac{E(A_i Y_i|X_i = x)}{e(X_i)}\right]$ by moving $e(X_i)$ out of the 2nd expectation

$= E_X\left[\frac{E(A_i|X_i = x)E(Y_i|X_i = x, A_i = 1)}{e(X_i)}\right]$ by joint vs. cond. expectation

$= E_X\left[\frac{E(A_i|X_i = x)E\left(Y_i^1\middle|X_i = x, A_i = 1\right)}{e(X_i)}\right]$ by consistency

$= E_X\left[\frac{E(A_i|X_i = x)E\left(Y_i^1\middle|X_i = x\right)}{e(X_i)}\right]$ by cond. exchangeability

$= E\left(Y_i^1\right)$ by the law of iterated expectation

**Prove**: for an untreated unit weighted by $w_i = \frac{1}{p(A_i = 0|X_i = x)}$, the expected observed outcome is equivalent to expected potential outcome under the treatment $A_i$=0, i.e., $E\left((1 - A_i)w_i Y_i\right) = E\left(Y_i^0\right)$.

**Estimator $\hat{\tau}$ – unbiased estimator for ACE**

Since

$$\tau = E\left(Y_i^1\right) - E\left(Y_i^0\right) = E(A_i w_i Y_i) - E\left((1 - A_i) w_i Y_i\right).$$

Unbiased estimator:

$$\hat{\tau} = \frac{1}{n} \sum_{i=1}^{n} [w_i A_i Y_i - w_i (1 - A_i) Y_i],$$

where $w_i = \frac{1}{p(A_i = a | X_i = x)}$, that is,

$$w_i = \begin{cases} \dfrac{1}{p(A_i = 1 | X_i = x)} = \dfrac{1}{e(X_i)} & \text{for treated } i \\ \dfrac{1}{p(A_i = 0 | X_i = x)} = \dfrac{1}{1 - e(X_i)} & \text{for untreated } i \end{cases}$$

Alternative unbiased estimator:

$$\hat{\tau} = \frac{\sum_{i=1}^{n} w_i A_i Y_i}{\sum_{i=1}^{n} w_i A_i} - \frac{\sum_{i=1}^{n} w_i (1 - A_i) Y_i}{\sum_{i=1}^{n} w_i (1 - A_i)}$$

-- $\hat{\tau}$ is sometimes called the Horvitz-Thompson estimator

Prove: IPW estimates with stabilized weights $\boxed{w_i^S = w_i P(A_i = a)}$ are unbiased estimator for ACE. For simplicity, consider binary treatment.

For binary treatment: We just proved that $\hat{\tau}$ is unbiased estimate of ACE. Next, we need to prove

$$\hat{\tau} = \frac{\sum w_i A_i Y_i}{\sum w_i A_i} - \frac{\sum w_i (1 - A_i) Y_i}{\sum w_i (1 - A_i)} = \frac{\sum w_i^S A_i Y_i}{\sum w_i^S A_i} - \frac{\sum w_i^S (1 - A_i) Y_i}{\sum w_i^S (1 - A_i)}$$

Proof:

$$\hat{\tau} = \frac{\sum w_i A_i Y_i}{\sum w_i A_i} - \frac{\sum w_i (1 - A_i) Y_i}{\sum w_i (1 - A_i)}$$

$$= \frac{\sum w_i P(A_i = 1) A_i Y_i}{\sum w_i P(A_i = 1) A_i} - \frac{\sum w_i P(A_i = 0)(1 - A_i) Y_i}{\sum w_i P(A_i = 0)(1 - A_i)}$$

$$= \frac{\sum w_i^S A_i Y_i}{\sum w_i^S A_i} - \frac{\sum w_i^S (1 - A_i) Y_i}{\sum w_i^S (1 - A_i)}$$

Since $P(A_i = a) = n_a/n$ are constant for all units within the same treatment group.

## Natural questions: How to estimate $e(X_i)$?

Discrete covariates → estimate the within-strata PS:

$$\hat{e}(X_i) = \hat{p}(A_i = a|X_i = x) = N_{xa}/N_x$$

- Non-parametric estimate of the propensity score in each stratum (defined by $X$) of the data

Continuous covariates → logistic regression $A_i$ on $X_i$.

## Disadvantage of PS weighting method:

- $p_w(A_i = 1|X_i = x)$ can be very small – violation of positivity assumption!
- Small changes to PS model lead to big changes in the weights – big weight variation

# R Codes – Propensity Score Weighting

## R packages:

tableone, sandwich (for robust variance estimate), ipw, survey

## Functions:

svyCreateTableOne() in tableone

ipwpoint() and ipwplot() in ipw

svydesign() in survey package

## Obtain PS weights

```
weightmodel<-ipwpoint(exposure= treatment, family = "binomial", link ="logit",
                      denominator= ~ age + female + meanbp1+ARF+CHF+Cirr+colcan+
                       Coma+lungcan+MOSF+sepsis, data=mydata)
```

| | |
|---|---|
| exposure | a vector, representing the exposure variable of interest. Both numerical and categorical variables can be used. A binomial exposure variable should be coded using values 0/1. |
| family | is used to specify a family of link functions, used to model the relationship between the variables in numerator or denominator and exposure, respectively. Alternatives are "binomial", "multinomial", "ordinal" and "gaussian". A specific link function is then chosen using the argument link, as explained below. Regression models are fitted using glm, multinom, polr or glm, respectively. |

| | |
|---|---|
| link | specifies the link function between the variables in numerator or denominator and exposure, respectively. For family = "binomial" (fitted using glm) alternatives are "logit", "probit", "cauchit", "log" and "cloglog". For family = "multinomial" this argument is ignored, and multinomial logistic regression models are always used (fitted using multinom). For family = "ordinal" (fitted using polr) alternatives are "logit", "probit", "cauchit", and "cloglog". For family = "gaussian" this argument is ignored, and a linear regression model with identity link is always used (fitted using glm). |
| denominator | is a formula, specifying the right-hand side of the model used to estimate the elements in the denominator of the inverse probability weights. This typically includes the variables specified in the numerator model, as well as confounders for which to correct. |
| numerator | is a formula, specifying the right-hand side of the model used to estimate the elements in the numerator of the inverse probability weights. When left unspecified, unstabilized weights with a numerator of 1 are estimated. |
| trunc | Optional truncation percentile (0-0.5). E.g. when trunc = 0.01, the left tail is truncated to the 1st percentile, and the right tail is truncated to the 99th percentile. |

```
ipwplot(weights = weightmodel$ipw.weights, logscale = FALSE, main =
          "weights", xlim = c(0, 22))
```

| | |
|---|---|
| weights | numerical vector of inverse probability weights to plot. |
| logscale | logical value. If TRUE, weights are plotted on a logarithmic scale. |

# R codes/outputs

```
#look at a table 1
table1<- CreateTableOne(vars=xvars, strata="treatment", data=mydata,
test=FALSE)
## include standardized mean difference (SMD)
print(table1,smd=TRUE)
```

| | Stratified by treatment | | | | |
|---|---|---|---|---|---|
| | 0 | | 1 | | SMD |
| n | 3551 | | 2184 | | |
| ARF (mean (sd)) | 0.45 | 0.5 | 0.42 | 0.49 | 0.059 |
| CHF (mean (sd)) | 0.07 | 0.25 | 0.1 | 0.29 | 0.095 |
| Cirr (mean (sd)) | 0.05 | 0.22 | 0.02 | 0.15 | **0.145** |
| colcan (mean (sd)) | 0 | 0.04 | 0 | 0.02 | 0.038 |
| Coma (mean (sd)) | 0.1 | 0.29 | 0.04 | 0.2 | **0.207** |
| lungcan (mean (sd)) | 0.01 | 0.1 | 0 | 0.05 | 0.095 |
| MOSF (mean (sd)) | 0.07 | 0.25 | 0.07 | 0.26 | 0.018 |
| sepsis (mean (sd)) | 0.15 | 0.36 | 0.32 | 0.47 | **0.415** |
| age (mean (sd)) | 61.76 | 17.29 | 60.75 | 15.63 | 0.061 |
| female (mean (sd)) | 0.46 | 0.5 | 0.41 | 0.49 | 0.093 |
| meanbp1 (mean (sd)) | 84.87 | 38.87 | 68.2 | 34.24 | **0.455** |

```
#propensity score model
psmodel<-
glm(treatment~age+female+meanbp1+ARF+CHF+Cirr+colcan+Coma+lun
gcan+MOSF+sepsis,data=mydata, family  = binomial(link ="logit"))

## value of propensity score for each subject
ps <-predict(psmodel, type = "response")

#create IP weights
weight<-ifelse(treatment==1,1/ps,1/(1-ps))

#apply weights to data
weighteddata<-svydesign(ids = ~ 1, data =mydata, weights = ~ weight)

#weighted table 1
weightedtable <-svyCreateTableOne(vars = xvars, strata="treatment",
data=weighteddata, test=FALSE)

## Show table with SMD
```

print(weightedtable, smd = TRUE)

| | Stratified by treatment 0 | | 1 | | SMD |
|---|---|---|---|---|---|
| n | 5732 | | 5745 | | |
| ARF (mean (sd)) | 0.44 | 0.5 | 0.44 | 0.5 | 0.01 |
| CHF (mean (sd)) | 0.08 | 0.27 | 0.08 | 0.27 | 0.005 |
| Cirr (mean (sd)) | 0.04 | 0.19 | 0.04 | 0.19 | 0.**001** |
| colcan (mean (sd)) | 0 | 0.04 | 0 | 0.06 | 0.042 |
| Coma (mean (sd)) | 0.08 | 0.26 | 0.07 | 0.25 | **0.023** |
| lungcan (mean (sd)) | 0.01 | 0.08 | 0.01 | 0.09 | 0.014 |
| MOSF (mean (sd)) | 0.07 | 0.26 | 0.07 | 0.26 | 0.004 |
| sepsis (mean (sd)) | 0.21 | 0.41 | 0.22 | 0.41 | **0.002** |
| age (mean (sd)) | 61.36 | 17.56 | 61.43 | 15.33 | 0.004 |
| female (mean (sd)) | 0.45 | 0.5 | 0.45 | 0.5 | 0.001 |
| meanbp1 (mean (sd)) | 78.6 | 37.58 | 79.26 | 40.31 | **0.017** |

#to get a weighted mean for a single covariate directly:
mean(weight[treatment==1]*age[treatment==1])/(mean(weight[treatment ==1]))

#get causal risk difference
glm.obj<-glm(died~treatment,weights=weight, family= quasibinomial ( link = "**identity**"))
#summary(glm.obj)
betaiptw<-coef(glm.obj)
SE<-sqrt(diag(vcovHC(glm.obj, type="HC0")))
causalrd<-(betaiptw[2])
lcl<-(betaiptw[2]-1.96*SE[2])
ucl<-(betaiptw[2]+1.96*SE[2])
c(lcl,causalrd,ucl)

**0.052 (0.023, 0.080)**

#get causal relative risk. Weighted GLM
glm.obj<-glm(died~treatment, weights=weight, family=quasibinomial( link=**log**))
#summary(glm.obj)
betaiptw<-coef(glm.obj)
#to properly account for weighting, use asymptotic (sandwich) variance
SE<-sqrt(diag(vcovHC(glm.obj, type="HC0")))

```
#get point estimate and CI for relative risk (need to exponentiate)
causalrr<-exp(betaiptw[2])
lcl<-exp(betaiptw[2]-1.96*SE[2])
ucl<-exp(betaiptw[2]+1.96*SE[2])
c(lcl,causalrr,ucl)
```

**1.08 (1.04, 1.13)**

```
#truncate weights at 10
truncweight<-replace(weight, weight>10,10)
#get causal risk difference
glm.obj<-glm(died~treatment,weights=truncweight, family=
            quasibinomial(link="identity"))
#summary(glm.obj)
betaiptw<-coef(glm.obj)
SE<-sqrt(diag(vcovHC(glm.obj, type="HC0")))

causalrd<-(betaiptw[2])
lcl<-(betaiptw[2]-1.96*SE[2])
ucl<-(betaiptw[2]+1.96*SE[2])
c(lcl,causalrd,ucl)
```

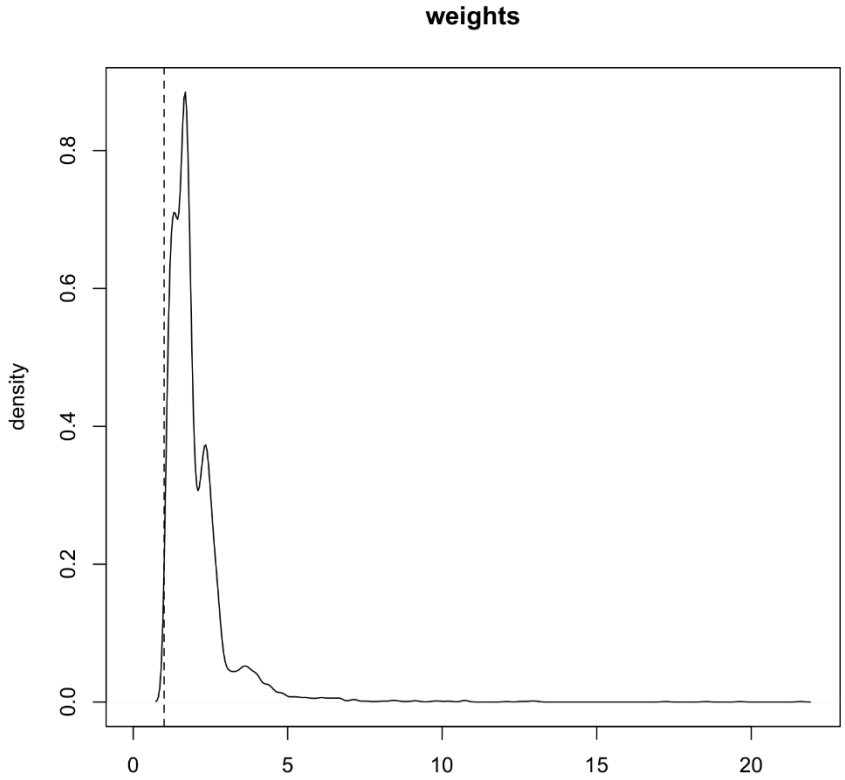
**0.055 (0.028, 0.082)**

```
##############################
#alternative: use ipw package
##############################

#first fit propensity score model to get weights
weightmodel<-ipwpoint(
      exposure= treatment, family = "binomial", link ="logit",
      denominator=~age+female+meanbp1+ARF+CHF+Cirr+colcan+Com
      a+lungcan+MOSF+sepsis, data=mydata)

#numeric summary of weights
summary(weightmodel$ipw.weights)
```

| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
|---|---|---|---|---|---|
| 1.046 | 1.405 | 1.721 | 2.001 | 2.280 | 21.606 |

```
#plot of weights
```

**ipwplot**(weights = weightmodel$ipw.weights, logscale = FALSE, main = "weights", xlim = c(0, 22))




mydata$wt<-weightmodel$ipw.weights

#Estimate risk difference
rd <- (svyglm(died ~ treatment, design = svydesign(~1, weights = ~wt, data =mydata)))
coef(rd); confint(rd)

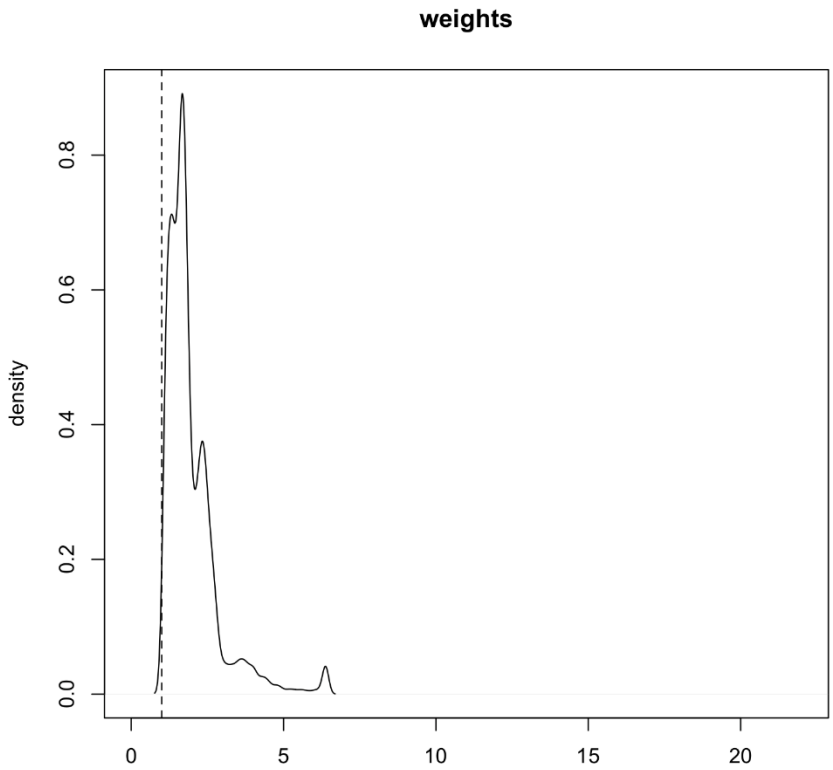
**0.052 (0.023, 0.080)**

# fit propensity score model to get weights, but truncated
weightmodel<-
**ipwpoint**(exposure= treatment, family = "binomial", link ="logit", denominator=~age+female+meanbp1+ARF+CHF+Cirr+colcan+Coma+lungcan+MOSF+sepsis, data=mydata, **trunc=.01**)

#numeric summary of weights
summary(weightmodel$weights.trun)

| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
|---|---|---|---|---|---|
| 1.081 | 1.405 | 1.721 | 1.972 | 2.280 | 6.379 |

#plot of weights
**ipwplot**(weights = weightmodel$weights.trun, logscale = FALSE, main = "weights", xlim = c(0, 22))




mydata$wt<-weightmodel$weights.trun

#estimate risk difference with truncation
rd.trk <- (svyglm(died ~ treatment, design = svydesign(~ 1, weights = ~wt, data =mydata)))
coef(rd.trk); confint(rd.trk)

**0.055 (0.028, 0.082)**