Example analysis

Linear model, smoking is the exposure, birthweight in grams the outcome

```r
infunc_confint <- function(ofit, wfit, est, exposurename, outcomename) {

  ofit <- glm(ofit$formula, family = ofit$family,
              data = ofit$data)

  data <- ofit$data
  phat <- predict(wfit, type = "response")

  data1 <- data0 <- data
  data1[[exposurename]] <- 1
  data0[[exposurename]] <- 0

  est1 <- predict(ofit, newdata = data1, type = "response")
  est0 <- predict(ofit, newdata = data0, type = "response")


  eifterms1 <- data[[exposurename]] / phat *
    (data[[outcomename]] - est1) + (est1 - mean(est1))

  eifterms0 <- (1 - data[[exposurename]]) / (1 - phat) *
    (data[[outcomename]] - est0) + (est0 - mean(est0))

  ifweight <- t(vcov(wfit) %*% t(sandwich::estfun(wfit)))
  ifout <- t(vcov(ofit) %*% t(sandwich::estfun(ofit)))

  XXw <- model.matrix(wfit)
  XXo <- model.matrix(ofit)
  hdot <- XXw * family(wfit)$mu.eta(predict(wfit, type = "link"))
  gdot0 <- model.matrix(terms(ofit), data = data0) *
```

```r
        family(ofit)$mu.eta(predict(ofit, newdata = data0, type = "link"))
      gdot1 <- model.matrix(terms(ofit), data = data1) *
        family(ofit)$mu.eta(predict(ofit, newdata = data1, type = "link"))

    fullif1 <- cbind(eifterms1,
                -diag((data[[exposurename]] * hdot / phat^2 *
                            (data[[outcomename]] - est1)) %*% t(ifweight)),
                diag((gdot1 / phat * (phat - data[[exposurename]])) %*% t(ifout)))

    fullif0 <- cbind(eifterms0,
                diag(((1 - data[[exposurename]]) * hdot / (1 - phat)^2 *
                        (data[[outcomename]] - est0)) %*% t(ifweight)),
                diag((gdot0 / (1 - phat) * (data[[exposurename]] - phat)) %*% t(ifout)))


    estse <- sd(rowSums(fullif1) - rowSums(fullif0)) / sqrt(nrow(data))

    c(lower = est - 1.96 * estse, upper = est + 1.96 * estse)


}

data <- AF::clslowbwt
head(data)
```

```
  id birth  smoke      race    age lwt  bwt    low lbw smoker
1 1      1 1. Yes 3. Other  28 120 2865  0. No   0      1
2 1      2 1. Yes 3. Other  33 141 2609  0. No   0      1
3 2      1  0. No 1. White  29 130 2613  0. No   0      0
4 2      2  0. No 1. White  34 151 3125  0. No   0      0
5 2      3  0. No 1. White  37 144 2481 1. Yes   1      0
6 3      1 1. Yes 2. Black  31 187 1841 1. Yes   1      1
```

```r
## propensity score fit

pwfit <- glm(smoker ~ race * age * lwt + I(age^2) + I(lwt^2), data = data,
             family = "binomial")
phat <- predict(pwfit, type = "response")

data$weight <- data$smoker / phat + (1 - data$smoker) / (1 - phat)
## outcome model
```

```r
outfit <- glm(bwt ~ smoker * (race + age + lwt) + I(age^2) + I(lwt^2),
              data = data, family = "gaussian", weights = weight)

## dummy data, where we set X to 0 and 1

data0 <- data1 <- data
data0$smoker <- 0
data1$smoker <- 1

## predicted couterfactuals
Yhat0 <- predict(outfit, newdata = data0, type = "response")
Yhat1 <- predict(outfit, newdata = data1, type = "response")

ATEmean <- mean(Yhat1) - mean(Yhat0)
ATEmean
```

```
[1] -223.6736
```

```r
infunc_confint(outfit, pwfit, ATEmean, "smoker", "bwt")
```

```
      lower        upper
-365.07104   -82.27607
```

```r
## CAN Richard Funk?

## separate outcome models, unweighted

ofit1 <- glm(bwt ~ (race + age + lwt) + I(age^2) + I(lwt^2),
             data = subset(data, smoker == 1), family = "gaussian")

ofit0 <- glm(bwt ~ (race + age + lwt) + I(age^2) + I(lwt^2),
             data = subset(data, smoker == 0), family = "gaussian")


dr1 <- data$bwt * (data$smoker == 1) / phat -
  predict(ofit1, newdata = data1, type = "response") *
  ((data$smoker == 1) - phat) / phat
dr0 <- data$bwt * (data$smoker == 0) / (1 - phat) +
  predict(ofit0, newdata = data0, type = "response") *
```

```
    ((data$smoker == 1) - phat) / (1 - phat)

  ATEfunk <- mean(dr1) - mean(dr0)
  ATEfunk
```

```
[1] -225.5489
```

Bootstrap for confidence interval, not echoed

```
      2.5%        97.5%
-387.12347   -46.48662
```

Logistic model, smoking is the exposure, low birthweight (birth weight < 2500g) the outcome

```
## propensity score fit

pwfit <- glm(smoker ~ race * age * lwt + I(age^2) + I(lwt^2), data = data,
             family = "binomial")
phat <- predict(pwfit, type = "response")

data$weight <- data$smoker / phat + (1 - data$smoker) / (1 - phat)
## outcome model

outfit <- glm(lbw ~ smoker * (race + age + lwt) + I(age^2) + I(lwt^2),
              data = data, family = "binomial", weights = weight)
```

```
Warning in eval(family$initialize): non-integer #successes in a binomial glm!
```

```
## dummy data, where we set X to 0 and 1

data0 <- data1 <- data
data0$smoker <- 0
data1$smoker <- 1

## predicted couterfactuals
Yhat0 <- predict(outfit, newdata = data0, type = "response")
Yhat1 <- predict(outfit, newdata = data1, type = "response")
```

```
ATEmean <- mean(Yhat1) - mean(Yhat0)
ATEmean
```

[1] 0.134932

```
infunc_confint(outfit, pwfit, ATEmean, "smoker", "lbw")
```

```
     lower      upper
0.04285596 0.22700794
```

```
RRmean <- mean(Yhat1) / mean(Yhat0)
RRmean
```

[1] 1.478392

```
## separate outcome models, unweighted

ofit1 <- glm(lbw ~ (race + age + lwt) + I(age^2) + I(lwt^2),
             data = subset(data, smoker == 1), family = "binomial")

ofit0 <- glm(lbw ~ (race + age + lwt) + I(age^2) + I(lwt^2),
             data = subset(data, smoker == 0), family = "binomial")


dr1 <- data$lbw * (data$smoker == 1) / phat -
  predict(ofit1, newdata = data1, type = "response") *
  ((data$smoker == 1) - phat) / phat
dr0 <- data$lbw * (data$smoker == 0) / (1 - phat) +
  predict(ofit0, newdata = data0, type = "response") *
  ((data$smoker == 1) - phat) / (1 - phat)

ATEfunk <- mean(dr1) - mean(dr0)
RRfunk <- mean(dr1) / mean(dr0)
ATEfunk
```

[1] 0.1372512

```
RRfunk
```

[1] 1.488187

Bootstrap for confidence interval, not echoed

```
     2.5%      97.5%
0.0279252 0.2411590
```

```
    2.5%     97.5%
1.082286 2.004029
```