Criteria for drug/placebo performance measurement

2020-07-12

We consider the criteria to measure the performance of the drug or placebo through the score v.s. time trajectory. We can calculate the:

- Change score: $Y(t=8) - Y(t=0)$ (the start time is $t=0$, the end time is $t=8$)
- Integral: $\int_0^8 f(t)dt$
  - We set the trajectroies as quadratic functions, therefore, $f(t)$ can be formularized as

$$f(t) = a_1 t^2 + a_2 t + a_3 \tag{1}$$

$$\int_0^8 f(t)dt = \int_0^8 a_1 t^2 + a_2 t + a_3 dt = a_1 \frac{8^3}{3} + 32a_2 + 8a_3 \tag{2}$$

- Weighted intergal: $\int_0^8 w(t)f(t)dt$
  - We set the weight function as $w(t) = c\exp(\mu'_{pbo}(t))$, where $\mu_{pbo}(t)$ is the population level average trajectory of placebo-treated patients. c is the constant to make the integral of $w(t) = 1$.
  - Suppose $\mu_{pbo}(t) = a_{pbo,1}t^2 + a_{pbo,2}t + a_{pbo,3}$, then

$$\mu'_{pbo}(t) = 2a_{pbo,1}t + a_{pbo,2} \tag{3}$$

$$\begin{aligned}\int_0^8 w(t)f(t)dt =& c\int_0^8 \exp(2a_{pbo,1}t + a_{pbo,2})(a_1t^2 + a_2t + a_3)dt \\ =& \frac{c}{4a_{pbo,1}^3}\big\{\exp(a_{pbo,2})(((2a_3 + 16a_2 + 128a_1)a_{pbo,1}^2 + \\ & (-a_2 - 16a_1)a_{pbo,1} + a_1)\exp(16a_{pbo,1}) - 2a + 3a_{pbo,1}^2 + a_2a_{pbo,1} - a_1)\big\}\end{aligned} \tag{4}$$

$$c = \frac{2a_{pbo,1}}{(\exp(16a_{pbo,1}) - 1)\exp(a_{pbo,2})} \tag{5}$$

Simulation Procedure

1. Data generation

The outcome is generated following

$$Y_k = S(\beta_k + b_k + \Gamma_k(\alpha_k' x)) + \epsilon_k, k = \{1, 2\} \text{ presents drug group and placebo group.} \tag{6}$$

The parameter settings are:

- dimension of the predictors $p = 3, 10$
- $\beta_{drg} = \beta_{pbo} = (1, -0.05, -0.02)'$
- $\Gamma_{drg} = (0, -\sin(\frac{\pi}{3}, -\cos(\frac{\pi}{3}))$
- $\Gamma_{pbo} = (0, \cos(\frac{\pi}{3}, -\sin(\frac{\pi}{3}))$

- $S = [1, t, t^2]$, $t = [0, 1, 2, 3, 4, 6, 8]$ is the design matrix for fixed effect and random effect
- $x \sim MVN(\mu_x, \Sigma_x)$, $\mu_x = \mathbf{0}_p$, Σ_x has diagonal equals to 1 and 0.5 everywhere else.
- $D_{drg} = \begin{pmatrix} 1.45 & -0.11 & 0.2 \\ -0.11 & 0.17 & -0.08 \\ 0.2 & -0.08 & 0.23 \end{pmatrix}$, $D_{pbo} = \begin{pmatrix} 1.03 & -0.23 & -0.15 \\ -0.23 & 0.68 & 0.25 \\ -0.15 & 0.25 & 1.36 \end{pmatrix}$
- $\epsilon_{drg}, \epsilon_{pbo} \sim N(0, 10^2)$
- $\alpha = \alpha_1 + \delta\alpha_{2k}, k = 1, 2$, $\delta = 0, 1 \ldots$
  - $\alpha_1 = (1, 1, 1)$
  - drg: $\alpha_{21} = (1, ..., p)$
  - pbo: $\alpha_{22} = (-p, ..., -1)$

If $\delta = 0$, $\alpha_{drg} = \alpha_{pbo} = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$. The cosine similarity is 1.

If $\delta = 1$,

- drg: $\alpha = (0.37, 0.56, 0.74)$
- pbo: $\alpha = (-0.89, -0.45, 0)$
- cosine similarity is -0.58

A data set with $n = 100$ subjects in each treatment group (200 subjects in total) is then generated. Each subject has 7 time points ($t = 0, 1, 2, 43, 4, 6, 8$).

To conduct a 10-fold cross validation, we separate the data set into 10 parts. We use 9 parts as the training set and 1 part as the test set.

2. Optimization

To estimate the $\hat{\alpha}$, we consider optimization of purity function and loglikelihood function for the training set.

- For purity function
  - 0. Set an initial α^* value, calculate $w^* = \alpha'^* x$
  - 1. Fit equation (6) and estimate $\hat{\beta}_1, \hat{\Gamma}_1, \hat{D}_1, \hat{\beta}_2, \hat{\Gamma}_2, \hat{D}_2$.
  - 2. Input the estimated value and $\hat{\alpha}$ in the purity function to calculate the purity.
  - 3. Repeat step 1, 2 and apply Nelder-Mead algorithm to find the α that maximizes the purity function.
- For loglikelihood function
  - 0. Set an initial α^* value, calculate $w^* = \alpha'^* x$
  - 1. Fit equation (6) and estimate $\hat{\beta}_1, \hat{\Gamma}_1, \hat{D}_1, \hat{\epsilon}_1, \hat{\beta}_2, \hat{\Gamma}_2, \hat{D}_2, \hat{\epsilon}_2$.
  - 2. Input the estimated value and $\hat{\alpha}$ in the loglikelihood function to calculate the loglikelihood.
  - 3. Repeat step 1, 2 and apply Nelder-Mead algorithm to find the α that maximizes the loglikelihood function.

Besides, I also consider a scenario, where the α is generated randomly from a standard normal distribution

3. Estimate group assingment

After the optimization, we get the estimated $\hat{\alpha}$, and then we may fit the linear mixed effect model and estimated the corresponding $\hat{\beta}_1, \hat{\Gamma}_1, \hat{D}_1, \hat{\beta}_2, \hat{\Gamma}_2, \hat{D}_2$.

- For the change score criteria:
  - Calculate the trajectroy coefficients for each subject (only consider the fixed effects)

$$\text{estimated drug group coefficient: } coef_1 = \hat{\beta}_1 + \hat{\Gamma}_1(\hat{\alpha}' x_i)$$

$$\text{estimated placebo group coefficient: } coef_2 = \hat{\beta}_2 + \hat{\Gamma}_2(\hat{\alpha}' x_i)$$

  - Calculate the estimated outcome $\hat{Y}_k = Xcoef_k$
  - If the $\hat{Y}_1(t=8) - \hat{Y}_1(t=0) < \hat{Y}_2(t=8) - \hat{Y}_2(t=0)$, the subject will be assigned to drug group, otherwise will be assigned to placebo group.
- For the intergal criteria:
  - Calculate the trajectroy coefficients for each subject (only consider the fixed effects)

$$\text{estimated drug group coefficient: } coef_1 = \hat{\beta}_1 + \hat{\Gamma}_1(\hat{\alpha}' x_i)$$

$$\text{estimated placebo group coefficient: } coef_2 = \hat{\beta}_2 + \hat{\Gamma}_2(\hat{\alpha}' x_i)$$

  - Input the estimated coefficient into the equation (2) to calculate the integral.
  - If integral of drug group is smaller than the integral of placebo group, the subject is assigned to drug group. Otherwise, the subject is assigned to placebo group.
- For the weighted intergal critera:
  - Calculate the trajectroy coefficients for each subject (only consider the fixed effects)

$$\text{estimated drug group coefficient: } coef_1 = \hat{\beta}_1 + \hat{\Gamma}_1(\hat{\alpha}' x_i)$$

$$\text{estimated placebo group coefficient: } coef_2 = \hat{\beta}_2 + \hat{\Gamma}_2(\hat{\alpha}' x_i)$$

  - Input the estimated coefficient into the equation (4) to calculate the weighted integral.
  - For the population level placebo-treated trajectory, μ_{pbo},
    - I tried to use the true average value, which is β_{pbo}
    - I tried to estimate the $\hat{\mu}_{pbo}$ from the simulated data, by fitting linear mixed effect model without predictors,

$$Y = X\beta_{pbo} + Xb_{pbo} + \epsilon_{pbo}$$

  - If integral of drug group is smaller than the integral of placebo group, the subject is assigned to drug group. Otherwise, the subject is assigned to placebo group.

4. Calculate the IPWE

The mean value of change score, integral or weighted integral of the test dataset is calculated for subjects whose estimated group assignment is consistent with their observed group assignment.

- If we estimate the assignment by using the change score criteria, the mean value is chosen to be the mean value of the observed change scores.
- If we estimate the assignment by using the integral criteria, the mean value is chosen to be the mean value of the observed integrals.
- If we estimate the assignment by using the weighted integral criteria, the mean value is chosen to be the mean value of the observed weighted integrals.

Note: I used different integral and weighted integral here, i.e. they are different from the ones for group assignment calculation. Since the change score for each subject does not changed for different estimation of α, the integral and weighted integral here are also calculated without affection of α.

We fit the linear mixed effect model without predictors as

$$Y_k = X\beta_k + Xb_k + \epsilon$$

For each subject, we have his or her estimated coefficients from the above LME. Then by pluging in the equation (2) and euqaiont (4), we can calculate the subject's integral and weighted integral.

Results

The above procedures are repeated for 100 times.

For each p and δ scenario, six plots are drawn. In the first row, the plots from left to right are plots for

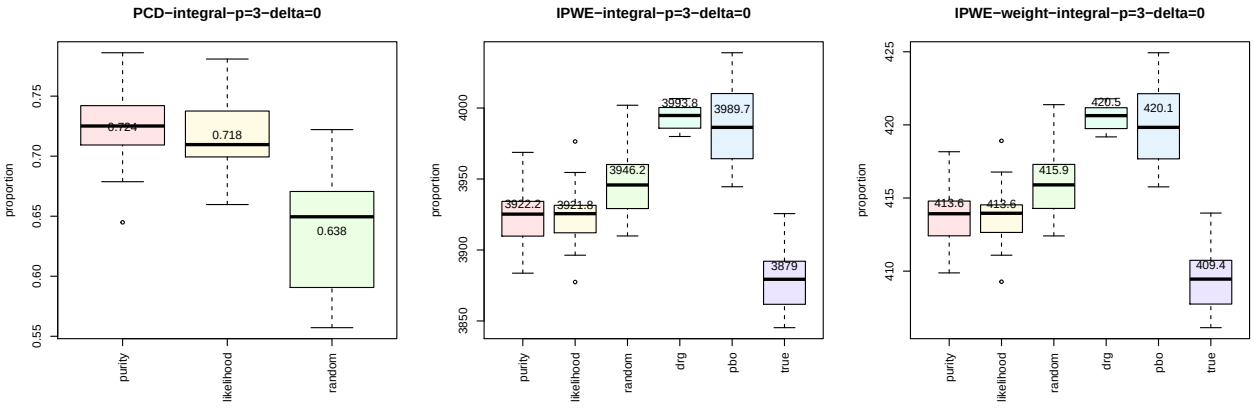
- boxplot of proportion of correct decision, i.e. the proportion of how many subjcets have correct estimated assignment.
  - the boxs from left to right present the results calcualted by purity method, loglikelihood method, and a randomly selected α.
- boxplot of IPWE, with intergal as the calculation critera.
  - the boxs from left to right present the results calcualted by purity method, loglikelihood method, a randomly selected α, only consider drug group, only consider placebo group, and with true assignment.
- boxplot of IPWE, with weighted intergal as the calculation critera.
  - the average placebo trajectory is calculated with $\beta_p bo$
  - the boxs from left to right present the results calcualted by purity method, loglikelihood method, a randomly selected α, only consider drug group, only consider placebo group, and with true assignment.

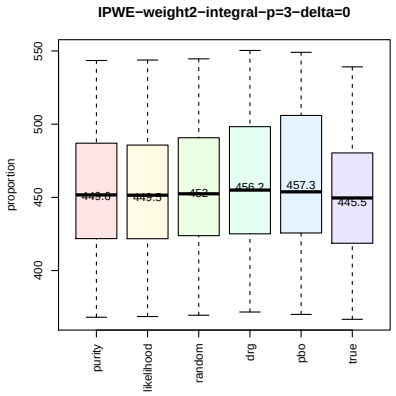
The plots at the second row, from left to right are

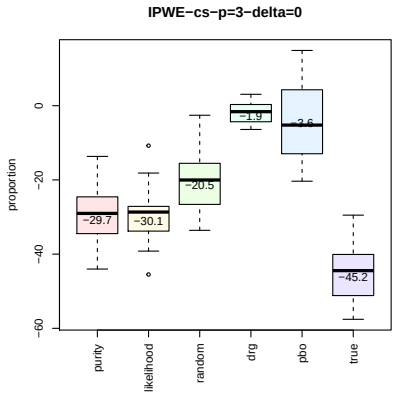
- boxplot of IPWE, with weighted intergal as the calculation critera.
  - the average placebo trajectory is calculated with estimated coefficients from LME
  - the boxs from left to right present the results calcualted by purity method, loglikelihood method, a randomly selected α, only consider drug group, only consider placebo group, and with true assignment.
- boxplot of IPWE, with change score as the calculation critera.
  - the boxs from left to right present the results calcualted by purity method, loglikelihood method, a randomly selected α, only consider drug group, only consider placebo group, and with true assignment.
- histogram of the proportions that purity method and likelihood method give same group assignment.

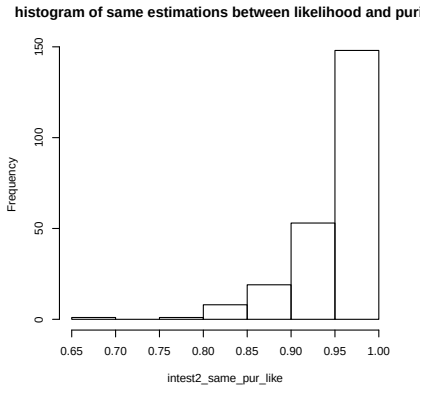
Besides, I also consider a scenario, where the α is generated randomly from a standard normal distribution

$p = 3, \delta = 0$

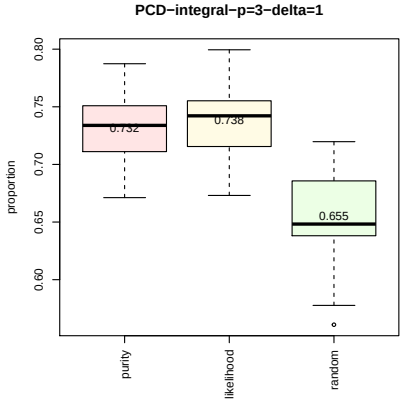


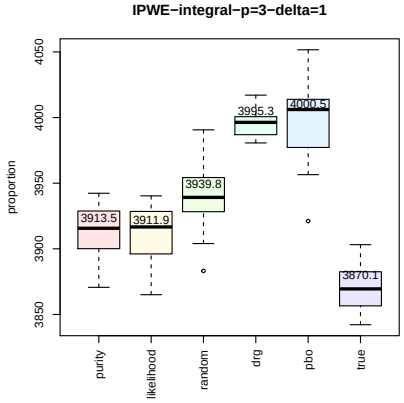


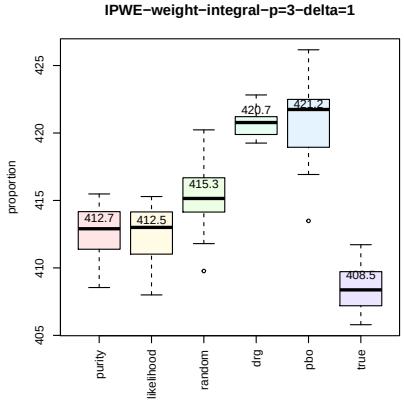


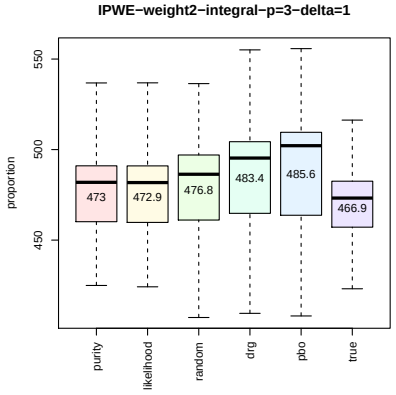


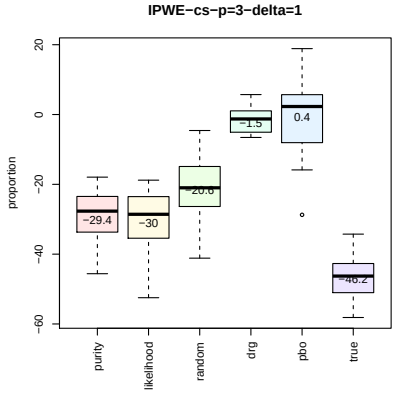
$p = 3, \delta = 1$

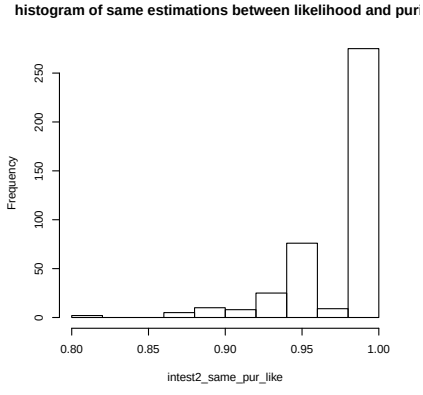




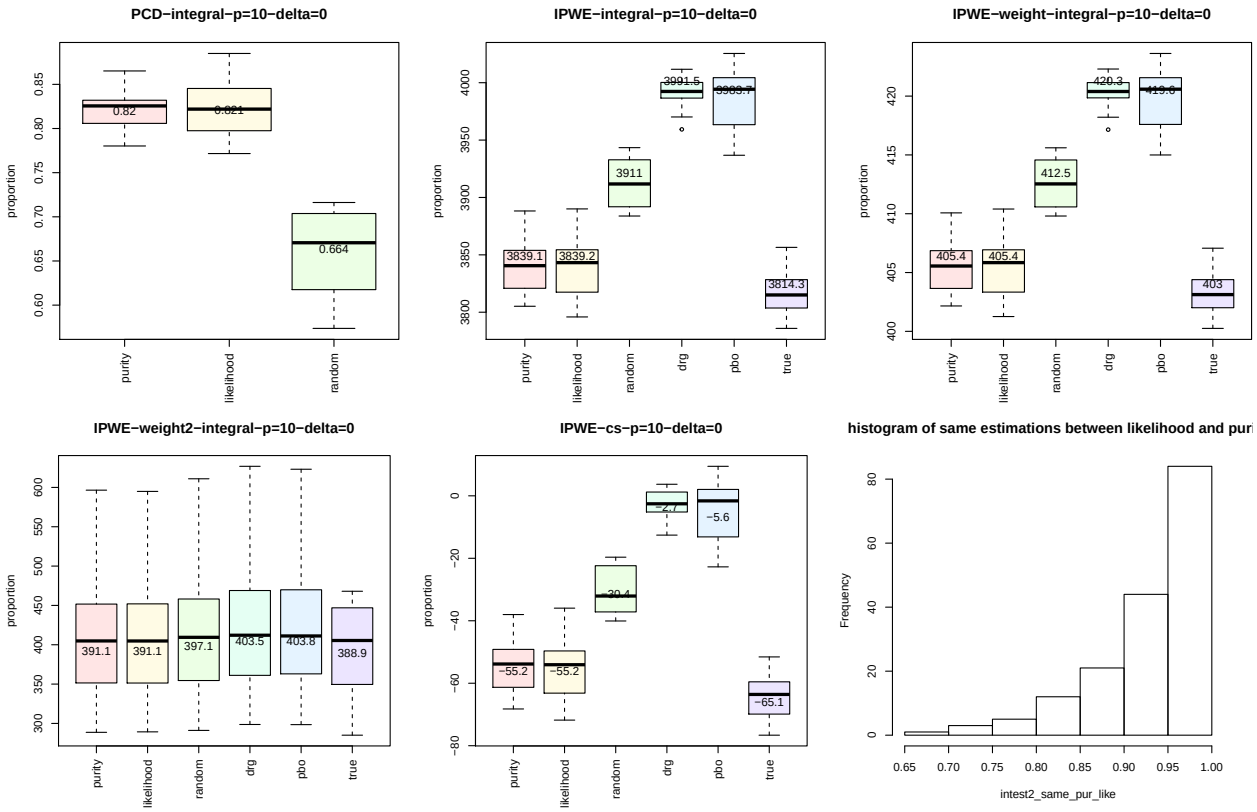




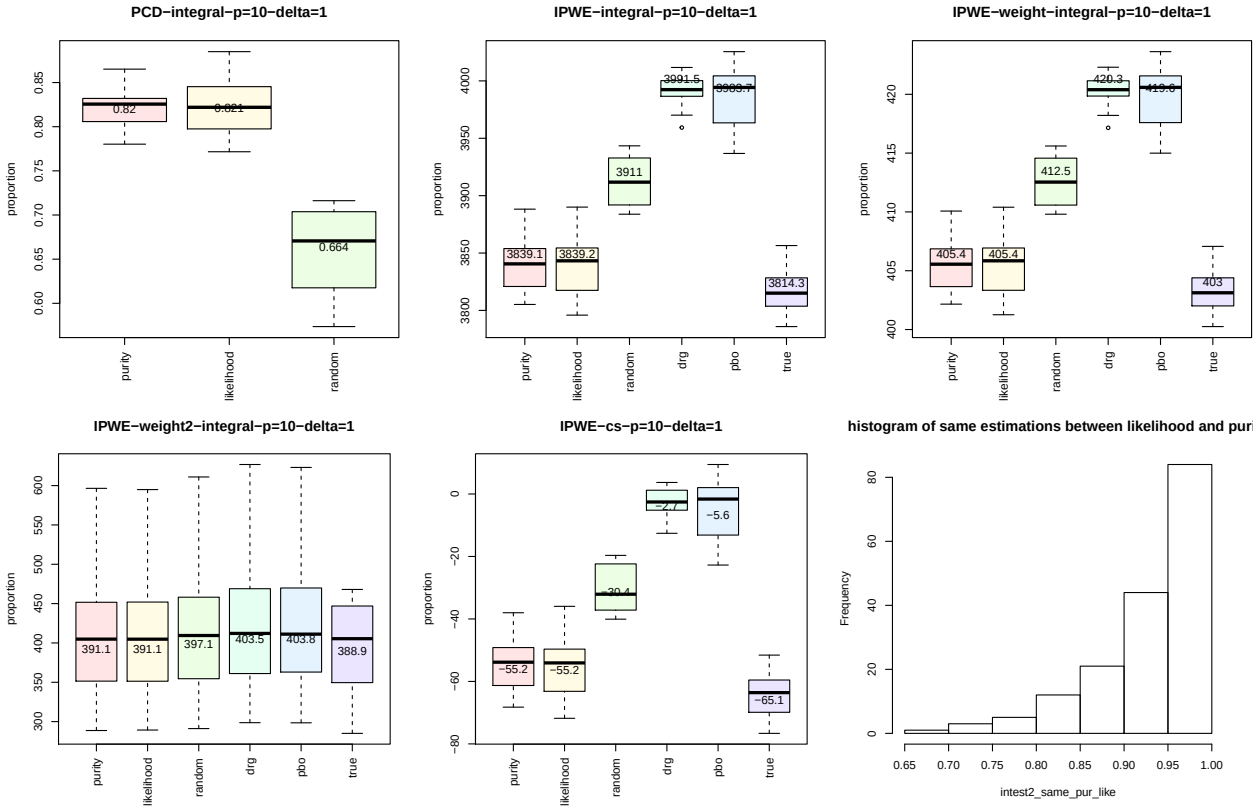




$p = 10, \delta = 0$


PCD−integral−p=10−delta=0
IPWE−integral−p=10−delta=0
IPWE−weight−integral−p=10−delta=0
IPWE−weight2−integral−p=10−delta=0
IPWE−cs−p=10−delta=0
histogram of same estimations between likelihood and pur
proportion
Frequency
purity
likelihood
random
drg
pbo
true
intest2_same_pur_like
0.82
0.821
0.664
3839.1
3839.2
3911
3991.5
3969.7
3814.3
405.4
405.4
412.5
420.3
419.6
403
391.1
391.1
397.1
403.5
403.8
388.9
−55.2
−55.2
−39.4
−2.7
−5.6
−65.1

$p = 10, \delta = 1$


PCD−integral−p=10−delta=1
IPWE−integral−p=10−delta=1
IPWE−weight−integral−p=10−delta=1
IPWE−weight2−integral−p=10−delta=1
IPWE−cs−p=10−delta=1
histogram of same estimations between likelihood and pur
proportion
Frequency
purity
likelihood
random
drg
pbo
true
intest2_same_pur_like
0.82
0.821
0.664
3839.1
3839.2
3911
3991.5
3969.7
3814.3
405.4
405.4
412.5
420.3
419.6
403
391.1
391.1
397.1
403.5
403.8
388.9
−55.2
−55.2
−39.4
−2.7
−5.6
−65.1