Propensity scores for multiple treatments: A tutorial for the `mnps` function in the `twang` package

Lane Burgette, Beth Ann Griffin and Dan McCaffrey *
RAND Corporation

October 10, 2013

1 Introduction

The Toolkit for Weighting and Analysis of Nonequivalent Groups, `twang`, was designed to make causal estimates in the binary treatment setting. In `twang` versions 1.3 and later, we have extended this software package to handle more than two treatment conditions through the new `mnps` function, which stands for multinomial propensity scores. McCaffrey et al. (2013) describe the methodology behind the `mnps` function; the purpose of this document is to describe the syntax and features related to the implementation in `twang`.

At a high level, the `mnps` function decomposes the propensity score estimation into several applications of the `ps` function, which was designed for the standard dichotomous treatment setting. For this reason, users who are new to `twang` are encouraged to learn about the `ps` function before using the `mnps` function. The other vignette that accompanies the package (Ridgeway et al., 2012) provides an extensive overview of the `ps` function, and much of that information will not be repeated here.

2 An ATE example

To demonstrate the package we utilize a random subset of the data described in McCaffrey et al. (2013). This truncated dataset is called `AOD`, and is included in the package. There are three treatment groups in the study, and the data include records for 200 youths in each treatment group of an alcohol and other drug treatment evaluation. We begin by loading the package and the data. Because there is a stochastic component to the subsequent model fits, we also set the random seed to ensure full replicability.

```
> library(twang)
> data(AOD)
> set.seed(1)
```

For the `AOD` dataset, the variable `treat` contains the treatment indicators, which have possible values `community`, `metcbt5`, and `scy`. The other variables included in the dataset are:

- `suf12`: outcome variable, substance use frequency at 12 month follow-up

*The development of this software and tutorial was funded by National Institute of Drug Abuse grant number 1R01DA015697 (PI: McCaffrey).

- `illact`: pretreatment covariate, illicit activities scale
- `crimjust`: pretreatment covariate, criminal justice involvement
- `subprob`: pretreatment covariate, substance use problem scale
- `subdep`: pretreatment covariate, substance use dependence scale
- `white`: pretreatment covariate, indicator for non-Hispanic white youth

In such an observational study, there are several quantities that one may be interested in estimating. The estimands that are most commonly of interest are the average treatment effect on the population (ATE) and the average treatment effect on the treated (ATT). The differences between these quantities are explained at length in McCaffrey et al. (2013), but in brief the ATE answers the question of how, on average, the outcome of interest would change if everyone in the population of interest had been assigned to a particular treatment relative to if they had all received another single treatment. The ATT answers the question of how the average outcome would change if everyone who received one particular treatment had instead received another particular treatment.

The main argument for the `mnps` function is a formula with the treatment variable on the left-hand side of a tilde, and pre-treatment variables on the right-hand side, separated by plus signs. Other key arguments are `data`, which simply tells the function the name of the dataframe that contains the variables for the propensity score estimation; the `estimand`, which can either be "ATT" or "ATE"; and `verbose`, which if set as `TRUE` instructs the function to print updates on the model fitting process, which can take a few minutes.

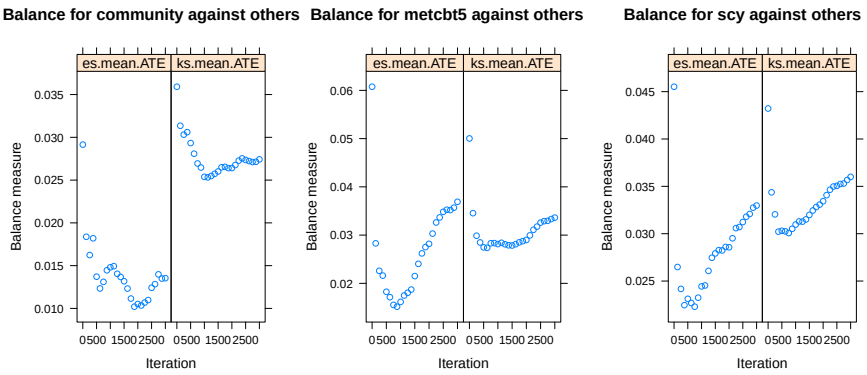
```
> mnps.AOD <- mnps(treat ~ illact + crimjust + subprob + subdep + white,
+                  data = AOD, estimand = "ATE", verbose = FALSE,
+                  stop.method = c("es.mean", "ks.mean"),
+                  n.trees = 3000)
```

The `twang` methods rely on tree-based regression models that are built in an iterative fashion. As the iterations or number of regression trees added to the model increases, the model becomes more complex. However, at some point, more complex models typically result in worse balance on the pre-treatment variables and therefore are less useful in a propensity score weighting context. The `n.trees` argument controls the maximum number of iterations.

Another key choice is the measure of balance that one uses when fitting these models. This is specified in the `stop.method` argument. As with the `ps` function, four `stop.method` objects are included in the package. They are `es.mean`, `es.max`, `ks.mean`, and `ks.max`. The four stopping rules are defined by two components: a balance metric for covariates and rule for summarizing across covariates. The balance metric summarizes the difference between two univariate distributions of a single pre-treatment variable (e.g., illicit activities scale). The default stopping rules in `twang` use two balance metrics: absolute standardized bias (also referred to as the absolute standardized mean difference or the effect size (ES)) and the Kolmogorov-Smirnov (KS) statistic. The stopping rule use two different rules for summarizing across covariates: the mean of the covariate balance metrics ("mean") or the maximum of the balance metrics ("max"). The first piece of the stopping rule name identifies the balance metric (ES or KS) and the second piece specifies the method for summarizing across balance metrics. For instance, `es.mean` uses the effect size or the absolute standardized bias and summarizes across variables with the mean and the `ks.max` uses the KS statistics to assess balances and summarizes using the maximum across variables and the other two stopping rules use the remaining two combinations of balance metrics and summary statistics. In this example, we chose to examine both `es.mean` and `ks.mean`, which is the default.

A first step is to make sure that we let the models run for a sufficiently large number of iterations in order to optimize the balance statistics of interest. We do this by seeing whether any of the balance measures of interest still appear to be decreasing after the number of iterations specified by the argument `n.trees` (10,000 iterations is the default).

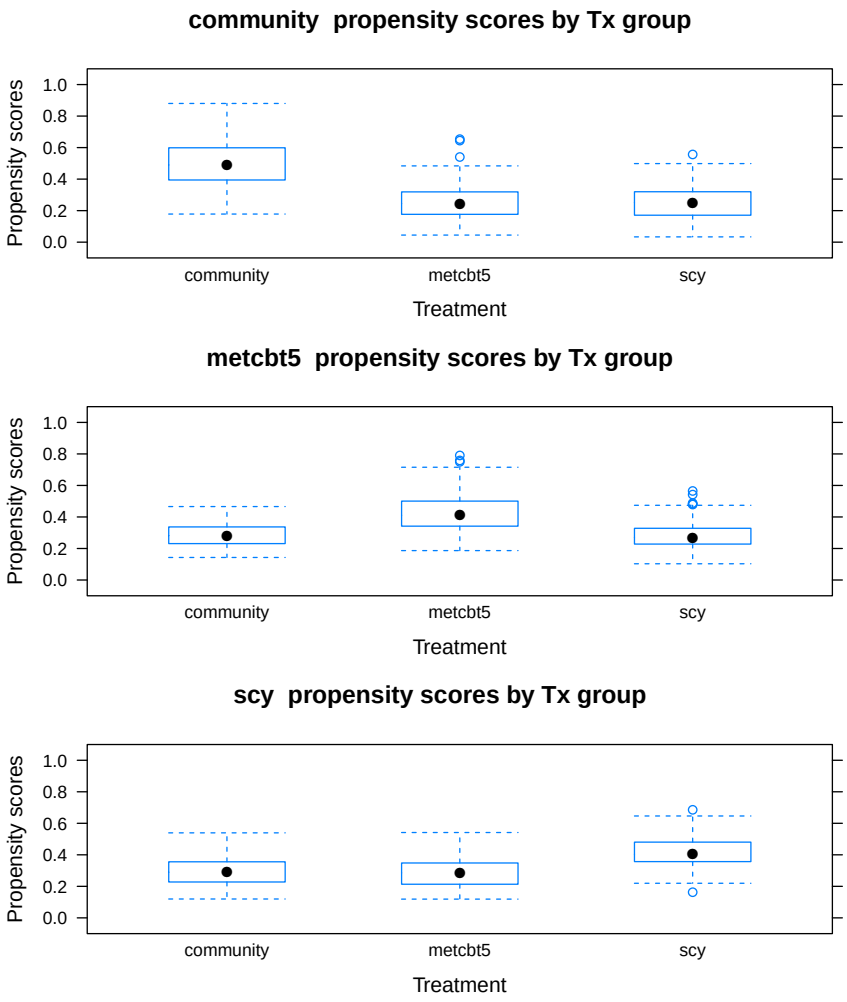
```
> plot(mnps.AOD, plots = 1)
```



In this figure, it appears that each of the balance measures are optimized with substantially fewer than 3,000 iterations, so we do not have evidence that we should re-run the `mnps()` call with a higher number of iterations or trees.

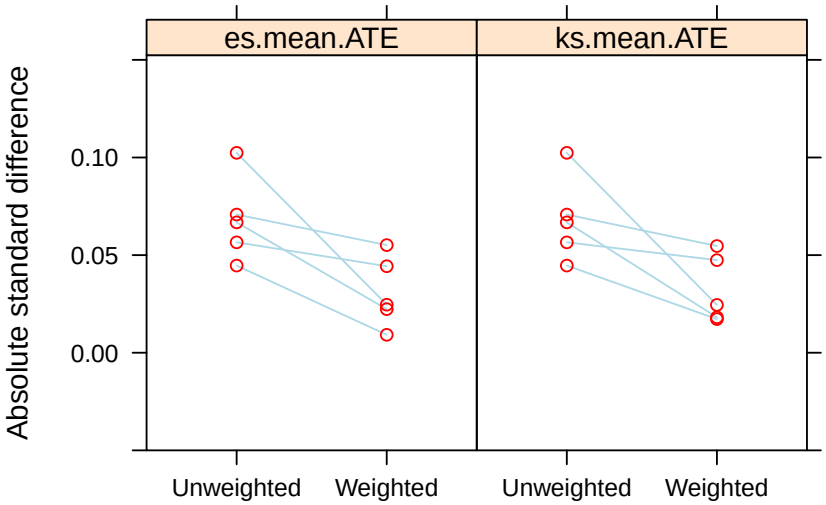
A key assumption in propensity score analyses is that each experimental unit has a non-zero probability of receiving each treatment. The plausibility of this assumption may be assessed by the overlap of the empirical propensity score distributions. This diagnostic is available using the `plots = 2` argument in the `plot` function. Here, the overlap assumption generally seems to be met, although there should be some concern that adolescents in the metcbt5 and scy conditions do not overlap well with the community group given the top most graphic. See McCaffrey et al. (2013) for more details on this issue.

```
>     plot(mnps.AOD, plots = 2)
```



As with the `ps` function for the binary treatment setting, the default plotting function for `mnps`-class objects also displays information on commonly-used balance statistics. In particular, it provides comparisons of the absolute standard differences (setting the `plots` argument equal to 3) and t statistics (with the `plots` argument equal to 4), before and after weighting. However, whereas there is a single plot for these balance diagnostics in the binary treatment setting, in the multiple treatment case, one can either examine a plot for each of the treatment conditions, or summarize the balance statistics in some way, across the treatment conditions. As a default, the `plot` function for an `mnps` object returns the maximum of the balance statistics across treatment groups for each of the covariates:

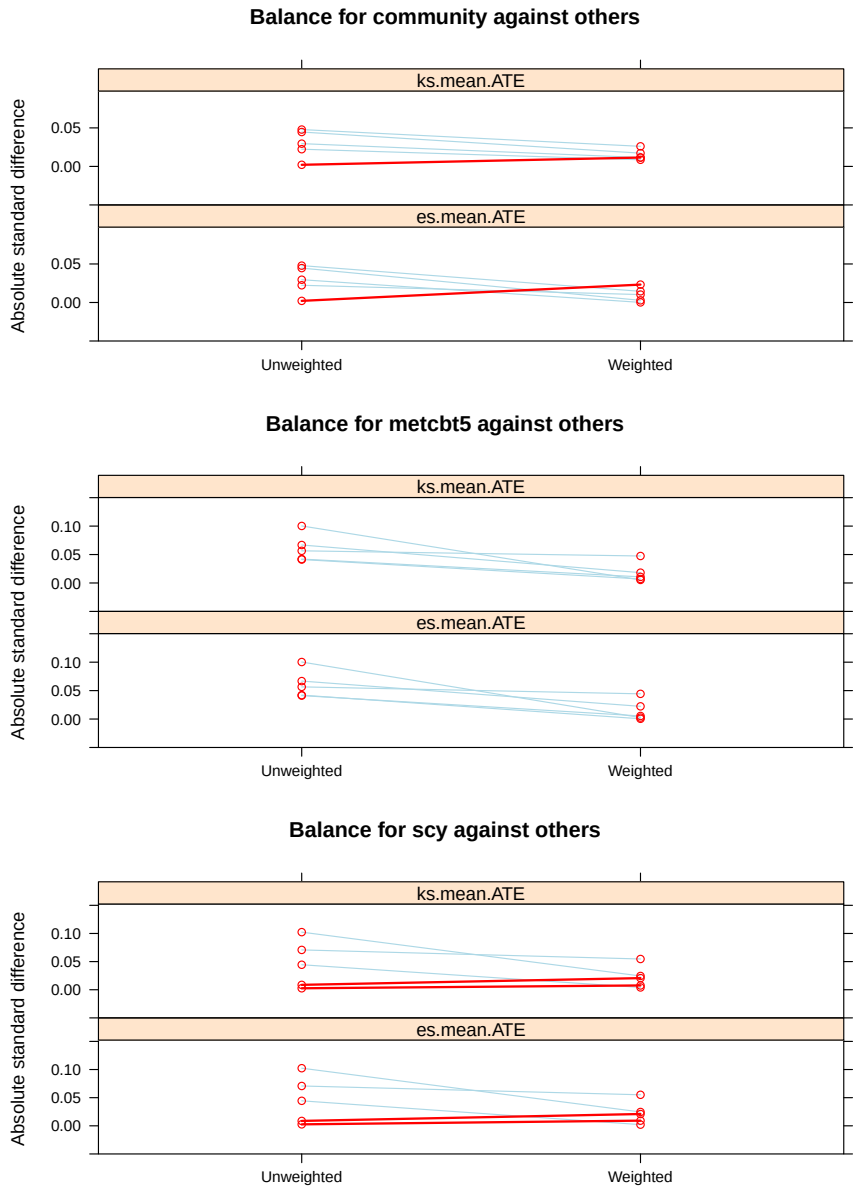
```
> plot(mnps.AOD, plots = 3)
```



If any of the differences had been statistically significant (before taking the maximum across treatment groups), the corresponding hollow circles in this plot would be solid.

It is possible to adjust the summarizing function using the `summaryFcn` argument. For example, one might consider the mean absolute standard differences rather than the maximum by setting `summaryFcn = mean`. Note that the function name should be provided without quotes. Regardless of the summary function, the circles at the end of the line segments will be hollow if none of the differences is statistically significant, and will be solid if at least one is significant. Another useful option is setting that argument equal to `NULL` which avoids the summary step altogether, and displays the balance statistics for each of the treatment conditions separately:

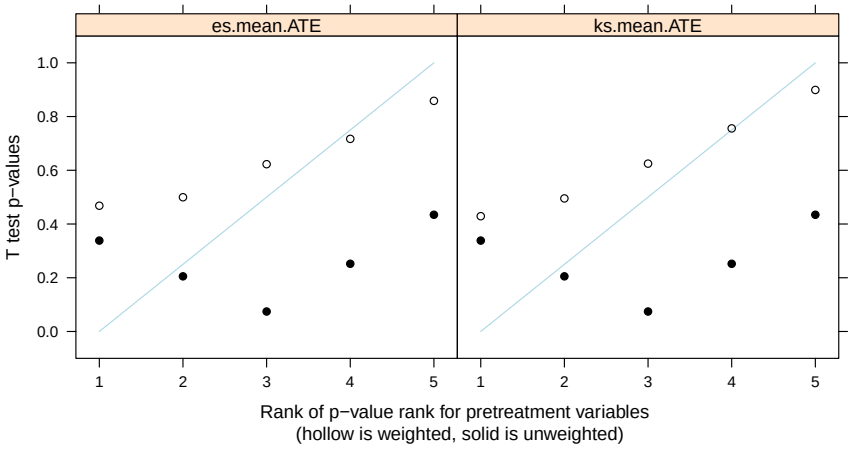
```
> plot(mnps.AOD, plots = 3, summaryFcn = NULL, figureRows = 3)
```



The additional `figureRows` argument instructs the function to spread the plots over three rows; by default the plots would be arranged in a single row rather than a column.

Setting the `plots` argument equal to 4 compares weighted and unweighted t-test or χ^2 statistic p-values for differences between each of the individual treatment groups and observations in all other treatment groups. Note that KS p-values are not available for ATE in the multiple treatment setting, and the `plots` argument therefore may not be set to 5.

```
>     plot(mnps.AOD, plots = 4)
```



Beyond graphics, there are several other functions that may be of interest to `mnps` users. The first is `means.table` which provides a nice, simple summary of balance across the groups. When `estimand` is set as `'ATE'`, the table shows the population means for each pretreatment covariate in the first column as well as each treatment group's unweighted and ATE weighted means and corresponding unweighted and weighted population standardized mean differences. As shown in the table below, incorporation of the ATE propensity score weights improves each treatment groups overall balance with the population means for each pretreatment covariate. The function also includes an argument called `includeSD` whose default is `FALSE`; changing it to `TRUE` returns standard deviations for each of the treatment conditions (not shown).

```
> means.table(mnps.AOD, stop.method = "es.mean", digits = 3)
```

```
         pop.mean unwt.community.mean wt.community.mean unwt.community.smd
illact      0.075                0.097             0.085               0.022
crimjust   -0.068               -0.065            -0.092               0.002
subprob    -0.016               -0.060            -0.013              -0.045
subdep      0.015                0.046             0.015               0.030
white       0.178                0.160             0.173              -0.048
         wt.community.smd unwt.metcbt5.mean wt.metcbt5.mean unwt.metcbt5.smd
illact              0.010              0.007           0.052           -0.067
crimjust           -0.023              0.037          -0.065            0.100
subprob             0.003              0.026          -0.016            0.042
subdep              0.000              0.058           0.021            0.041
white              -0.015              0.200           0.195            0.057
         wt.metcbt5.smd unwt.scy.mean wt.scy.mean unwt.scy.smd wt.scy.smd
illact           -0.022         0.120       0.077        0.044      0.002
crimjust          0.003        -0.174      -0.093       -0.102     -0.025
subprob           0.000        -0.013      -0.007        0.003      0.009
subdep            0.005        -0.058      -0.042       -0.071     -0.055
white             0.044         0.175       0.170       -0.009     -0.021
```

More extensive balance information is given by the `bal.table` function. For propensity score analyses with multiple treatments, this function returns a lot of information. For each outcome

category, and each stopping rule (in addition to the unweighted analysis) the `bal.table` function gives balance statistics such as weighted and unweighted means by treatment group. Note in this case that the "control" columns (labeled `ct`) refer to every treatment group except the one that is considered the treatment for a particular output. For example, in the first table that follows, `tx` refers to `community` treatment, and `ct` refers to all treatments except `community`.

```
> bal.table(mnps.AOD)

$community
$community$unw
          tx.mn tx.sd  ct.mn ct.sd std.eff.sz   stat     p    ks ks.pval
illact    0.097 1.045  0.075 1.014      0.022  0.382 0.703 0.037      NA
crimjust -0.065 1.050 -0.068 1.041      0.002  0.036 0.971 0.038      NA
subprob  -0.060 0.965 -0.016 0.985     -0.045 -0.782 0.434 0.058      NA
subdep    0.046 1.079  0.015 1.031      0.030  0.501 0.617 0.028      NA
white     0.160 0.368  0.178 0.383     -0.048 -0.847 0.397 0.018      NA

$community$es.mean.ATE
          tx.mn tx.sd  ct.mn ct.sd std.eff.sz   stat     p    ks ks.pval
illact    0.085 1.001  0.075 1.014      0.010  0.139 0.889 0.024      NA
crimjust -0.092 1.018 -0.068 1.041     -0.023 -0.254 0.799 0.036      NA
subprob  -0.013 0.938 -0.016 0.985      0.003 -0.178 0.858 0.044      NA
subdep    0.015 1.046  0.015 1.031      0.000  0.083 0.934 0.023      NA
white     0.173 0.379  0.178 0.383     -0.015 -0.462 0.644 0.006      NA

$community$ks.mean.ATE
          tx.mn tx.sd  ct.mn ct.sd std.eff.sz   stat     p    ks ks.pval
illact    0.083 1.001  0.075 1.014      0.009  0.136 0.892 0.025      NA
crimjust -0.080 1.018 -0.068 1.041     -0.011 -0.193 0.847 0.031      NA
subprob   0.001 0.948 -0.016 0.985      0.017 -0.023 0.982 0.032      NA
subdep    0.003 1.052  0.015 1.031     -0.012 -0.031 0.975 0.027      NA
white     0.168 0.375  0.178 0.383     -0.026 -0.577 0.564 0.010      NA


$metcbt5
$metcbt5$unw
         tx.mn tx.sd  ct.mn ct.sd std.eff.sz   stat     p    ks ks.pval
illact   0.007 1.035  0.075 1.014     -0.067 -1.147 0.252 0.065      NA
crimjust 0.037 1.038 -0.068 1.041      0.100  1.743 0.082 0.077      NA
subprob  0.026 1.019 -0.016 0.985      0.042  0.716 0.474 0.047      NA
subdep   0.058 1.047  0.015 1.031      0.041  0.709 0.478 0.042      NA
white    0.200 0.401  0.178 0.383      0.057  0.958 0.338 0.022      NA

$metcbt5$es.mean.ATE
          tx.mn tx.sd  ct.mn ct.sd std.eff.sz   stat     p    ks ks.pval
illact    0.052 0.997  0.075 1.014     -0.022 -0.363 0.717 0.038      NA
crimjust -0.065 1.003 -0.068 1.041      0.003  0.283 0.777 0.029      NA
subprob  -0.016 0.997 -0.016 0.985      0.000  0.062 0.950 0.028      NA
subdep    0.021 1.037  0.015 1.031      0.005  0.085 0.932 0.030      NA
white     0.195 0.397  0.178 0.383      0.044  0.726 0.468 0.017      NA
```

```
$metcbt5$ks.mean.ATE
          tx.mn tx.sd  ct.mn ct.sd std.eff.sz   stat     p    ks ks.pval
illact    0.056 0.999  0.075 1.014     -0.018 -0.311 0.756 0.036      NA
crimjust -0.073 1.008 -0.068 1.041     -0.005  0.178 0.859 0.026      NA
subprob  -0.026 1.000 -0.016 0.985     -0.011 -0.067 0.947 0.025      NA
subdep    0.022 1.034  0.015 1.031      0.007  0.097 0.923 0.032      NA
white     0.197 0.398  0.178 0.383      0.047  0.791 0.429 0.018      NA


$scy
$scy$unw
          tx.mn tx.sd  ct.mn ct.sd std.eff.sz   stat     p    ks ks.pval
illact    0.120 0.963  0.075 1.014      0.044  0.790 0.430 0.057      NA
crimjust -0.174 1.028 -0.068 1.041     -0.102 -1.788 0.074 0.062      NA
subprob  -0.013 0.972 -0.016 0.985      0.003  0.045 0.964 0.037      NA
subdep   -0.058 0.964  0.015 1.031     -0.071 -1.268 0.205 0.058      NA
white     0.175 0.381  0.178 0.383     -0.009 -0.151 0.880 0.003      NA

$scy$es.mean.ATE
          tx.mn tx.sd  ct.mn ct.sd std.eff.sz   stat     p    ks ks.pval
illact    0.077 0.992  0.075 1.014      0.002  0.052 0.958 0.033      NA
crimjust -0.093 1.010 -0.068 1.041     -0.025 -0.492 0.623 0.036      NA
subprob  -0.007 0.973 -0.016 0.985      0.009  0.149 0.881 0.029      NA
subdep   -0.042 0.973  0.015 1.031     -0.055 -0.675 0.500 0.044      NA
white     0.170 0.377  0.178 0.383     -0.021 -0.427 0.669 0.008      NA

$scy$ks.mean.ATE
          tx.mn tx.sd  ct.mn ct.sd std.eff.sz   stat     p    ks ks.pval
illact    0.079 0.992  0.075 1.014      0.004  0.071 0.943 0.033      NA
crimjust -0.093 1.011 -0.068 1.041     -0.024 -0.489 0.625 0.036      NA
subprob  -0.008 0.972 -0.016 0.985      0.008  0.127 0.899 0.029      NA
subdep   -0.041 0.974  0.015 1.031     -0.055 -0.682 0.495 0.044      NA
white     0.170 0.377  0.178 0.383     -0.021 -0.419 0.675 0.008      NA
```

Finally, there is also `summary` method for `mnps` objects which gives some information on balance measures as well as the number of iterations (trees) selected for each model under each stopping rule.

```
> summary(mnps.AOD)

Summary of mnps object:
Summary of community against others.
            n.treat n.ctrl ess.treat ess.ctrl     max.es    mean.es     max.ks
unw             200    400  200.0000 400.0000 0.04785366 0.02930366 0.05833333
es.mean.ATE     200    400  184.5124 389.3660 0.02322317 0.01018732 0.04353362
ks.mean.ATE     200    400  188.3758 392.3995 0.02607402 0.01499874 0.03192988
            max.ks.p    mean.ks iter
unw               NA 0.03600000   NA
es.mean.ATE       NA 0.02639685 1874
ks.mean.ATE       NA 0.02507839 1082
```

```
Summary of metcbt5 against others.
            n.treat n.ctrl ess.treat ess.ctrl     max.es    mean.es     max.ks
unw             200    400  200.0000 400.0000 0.10027943 0.06137739 0.07666667
es.mean.ATE     200    400  186.1874 394.2979 0.04431685 0.01501965 0.03800035
ks.mean.ATE     200    400  188.3648 395.4704 0.04744407 0.01774854 0.03562304
            max.ks.p    mean.ks iter
unw               NA 0.05033333   NA
es.mean.ATE       NA 0.02830985  880
ks.mean.ATE       NA 0.02724992  581

Summary of scy against others.
            n.treat n.ctrl ess.treat ess.ctrl     max.es    mean.es     max.ks
unw             200    400  200.0000 400.0000 0.10238503 0.04578066 0.06166667
es.mean.ATE     200    400  189.5017 393.4114 0.05514367 0.02240795 0.04425948
ks.mean.ATE     200    400  190.0047 393.7790 0.05466257 0.02231721 0.04368363
            max.ks.p    mean.ks iter
unw               NA 0.04333333   NA
es.mean.ATE       NA 0.03013078  754
ks.mean.ATE       NA 0.02992878  712
```

After examining the graphical and tabular diagnostics provided by `twang`, we can analyze the outcome variable using the propensity scores generated by the `mnps` function. Although two stop methods were specified initially (`es.mean` and `ks.mean`), at this point we have to commit to a single set of weights. From the `bal.table` call above, we see that the balance properties are very similar for the two stopping rules, and from the `summary` statement, we see that the effective sample sizes (`ess.treat`) are similar as well. Hence, we expect the two stop methods to give similar results; we choose to analyze the data with the `es.mean` weights.

In order to analyze the data using the weights, it is recommended that one use the `survey` package, which performs weighted analyses. We can add the weights to the dataset using the `get.weights` function and specify the survey design as follows:

```
> require(survey)
> AOD$w <- get.weights(mnps.AOD, stop.method = "es.mean")
> design.mnps <- svydesign(ids=~1, weights=~w, data=AOD)
```

As shown in the `ps` vignette, we can then perform the propensity score-adjusted regression using the `svyglm` function:

```
> glm1 <- svyglm(suf12 ~ as.factor(treat), design = design.mnps)
> summary(glm1)

Call:
svyglm(formula = suf12 ~ as.factor(treat), design = design.mnps)

Survey design:
svydesign(ids = ~1, weights = ~w, data = AOD)

Coefficients:
                      Estimate Std. Error t value
(Intercept)            -0.09913    0.06736  -1.472
as.factor(treat)metcbt5  0.14858    0.10502   1.415
as.factor(treat)scy      0.06464    0.09998   0.647
```

```
                          Pr(>|t|)
(Intercept)                  0.142
as.factor(treat)metcbt5      0.158
as.factor(treat)scy          0.518

(Dispersion parameter for gaussian family taken to be 1.002082)

Number of Fisher Scoring iterations: 2
```

Using this small subset of the data, we are unable to detect differences in the treatment group means. However, the coefficient for the metcbt5 term represents the causal effect of metcbt5 vs. community and the coefficient for the scy term represents the causal effect of scy vs. community assuming there are no unobserved confounders. In the context of this application, the signs of the estimates correspond to higher substance use frequency for youths exposed to metcbt5 or scy relative to community. More details on how to obtain all relevant pairwise differences can be found in McCaffrey et al. (2013).

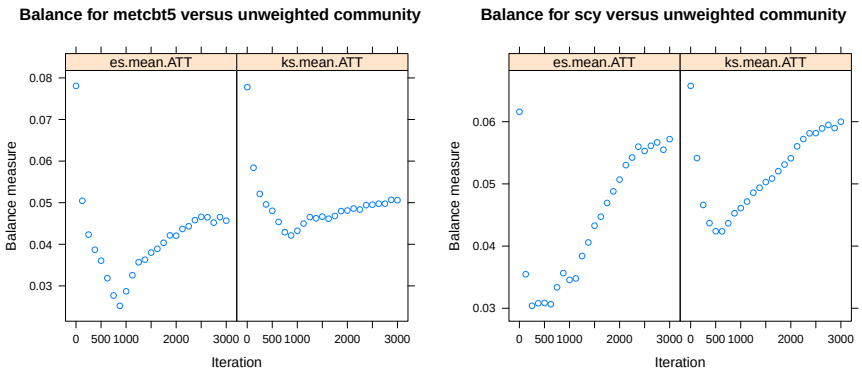
3 An ATT example

It is also possible to explore treatment effects on the treated (ATTs) using the `mnps` function. A key difference in the multiple treated setting is that we must be clear as to which treatment condition "the treated" refers to. This is done through the `treatATT` argument. Here, we define the treatment group of interest to be the community group; thus, we are trying to draw inferences about the relative effectiveness of the three treatment groups for individuals like those who were enrolled in the community program.

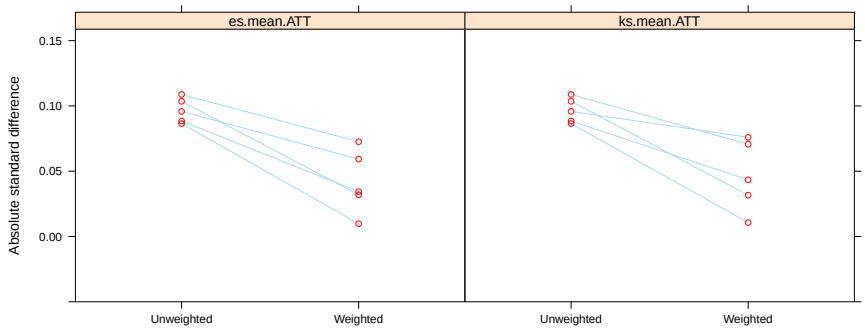
```
> mnps.AOD.ATT <- mnps(treat ~ illact + crimjust + subprob + subdep + white,
+                  data = AOD, estimand = "ATT", treatATT = "community",
+                  verbose = FALSE, n.trees = 3000)
```

The same array of visual and numerical summaries are available as they were in the ATE analysis.

```
>     plot(mnps.AOD.ATT, plots = 1)
```



```
> plot(mnps.AOD.ATT, plots = 3)
```



Although the same basic graphical descriptions are available as in the ATE case, note that the population means above are replaced with the means of the `treatATT` category in the `means.table` call.

```
> means.table(mnps.AOD.ATT, digits = 3)
```

```
         community.mean unwt.metcbt5.mean wt.metcbt5.mean unwt.metcbt5.smd
illact           0.097             0.007           0.086            0.087
crimjust        -0.065             0.037          -0.032           -0.097
subprob         -0.060             0.026          -0.062           -0.088
subdep           0.046             0.058           0.057           -0.011
white            0.160             0.200           0.186           -0.109
         wt.metcbt5.smd unwt.scy.mean wt.scy.mean unwt.scy.smd wt.scy.smd
illact           0.011         0.120       0.098       -0.021     -0.001
crimjust        -0.032        -0.174      -0.041        0.104     -0.023
subprob          0.002        -0.013      -0.018       -0.048     -0.043
subdep          -0.011        -0.058      -0.036        0.096      0.076
white           -0.071         0.175       0.163       -0.041     -0.008
```

The `bal.table` output is similar to the ATE case. However, for ATT, we do not need tables that give balance for the `treatATT` category against itself.

```
> bal.table(mnps.AOD.ATT)
```

```
Note that `tx` refers to the category specified as the treatATT, community.

$metcbt5
$metcbt5$unw
          tx.mn tx.sd ct.mn ct.sd std.eff.sz   stat     p    ks ks.pval
illact    0.097 1.045 0.007 1.035      0.087  0.870 0.385 0.100   0.270
crimjust -0.065 1.050 0.037 1.038     -0.097 -0.980 0.328 0.105   0.221
subprob  -0.060 0.965 0.026 1.019     -0.088 -0.861 0.390 0.090   0.394
subdep    0.046 1.079 0.058 1.047     -0.011 -0.113 0.910 0.055   0.924
white     0.160 0.368 0.200 0.401     -0.109 -1.041 0.298 0.040   0.997

$metcbt5$ks.mean.ATT
          tx.mn tx.sd  ct.mn ct.sd std.eff.sz   stat     p    ks ks.pval
```

```
illact    0.097 1.045  0.086 1.023      0.011  0.102 0.919 0.042   0.995
crimjust -0.065 1.050 -0.032 0.997     -0.032 -0.313 0.754 0.051   0.959
subprob  -0.060 0.965 -0.062 0.988      0.002  0.018 0.986 0.039   0.997
subdep    0.046 1.079  0.057 1.048     -0.011 -0.104 0.917 0.050   0.963
white     0.160 0.368  0.186 0.390     -0.071 -0.662 0.509 0.026   1.000

$metcbt5$es.mean.ATT
          tx.mn tx.sd  ct.mn ct.sd std.eff.sz   stat     p    ks ks.pval
illact    0.097 1.045  0.087 1.024      0.010  0.094 0.925 0.041   0.995
crimjust -0.065 1.050 -0.032 0.998     -0.032 -0.317 0.752 0.051   0.957
subprob  -0.060 0.965 -0.062 0.989      0.003  0.025 0.980 0.039   0.998
subdep    0.046 1.079  0.058 1.049     -0.012 -0.112 0.911 0.051   0.959
white     0.160 0.368  0.187 0.391     -0.073 -0.680 0.497 0.027   1.000


$scy
$scy$unw
          tx.mn tx.sd  ct.mn ct.sd std.eff.sz   stat     p    ks ks.pval
illact    0.097 1.045  0.120 0.963     -0.021 -0.223 0.823 0.060   0.866
crimjust -0.065 1.050 -0.174 1.028      0.104  1.048 0.295 0.080   0.545
subprob  -0.060 0.965 -0.013 0.972     -0.048 -0.481 0.631 0.090   0.394
subdep    0.046 1.079 -0.058 0.964      0.096  1.012 0.312 0.085   0.466
white     0.160 0.368  0.175 0.381     -0.041 -0.401 0.688 0.015   1.000

$scy$ks.mean.ATT
          tx.mn tx.sd  ct.mn ct.sd std.eff.sz   stat     p    ks ks.pval
illact    0.097 1.045  0.098 1.036     -0.001 -0.006 0.995 0.050   0.960
crimjust -0.065 1.050 -0.041 0.973     -0.023 -0.235 0.814 0.039   0.998
subprob  -0.060 0.965 -0.018 0.979     -0.043 -0.402 0.688 0.045   0.987
subdep    0.046 1.079 -0.036 0.994      0.076  0.744 0.457 0.074   0.664
white     0.160 0.368  0.163 0.370     -0.008 -0.077 0.939 0.003   1.000

$scy$es.mean.ATT
          tx.mn tx.sd  ct.mn ct.sd std.eff.sz   stat     p    ks ks.pval
illact    0.097 1.045  0.100 1.005     -0.002 -0.023 0.982 0.056   0.902
crimjust -0.065 1.050 -0.064 0.995     -0.002 -0.016 0.988 0.052   0.941
subprob  -0.060 0.965 -0.027 0.967     -0.034 -0.336 0.737 0.055   0.904
subdep    0.046 1.079 -0.018 0.993      0.059  0.596 0.551 0.069   0.707
white     0.160 0.368  0.176 0.382     -0.045 -0.433 0.665 0.016   1.000
```

The process to analyze the outcome variable is also similar:

```
> require(survey)
> AOD$w.ATT <- get.weights(mnps.AOD.ATT, stop.method = "es.mean")
> design.mnps.ATT <- svydesign(ids=~1, weights=~w.ATT, data=AOD)

> glm1 <- svyglm(suf12 ~ as.factor(treat), design = design.mnps.ATT)
> summary(glm1)

Call:
svyglm(formula = suf12 ~ as.factor(treat), design = design.mnps.ATT)
```

```
Survey design:
svydesign(ids = ~1, weights = ~w.ATT, data = AOD)

Coefficients:
                         Estimate Std. Error t value
(Intercept)              -0.10505    0.06383  -1.646
as.factor(treat)metcbt5   0.20071    0.10409   1.928
as.factor(treat)scy       0.08076    0.09901   0.816
                         Pr(>|t|)
(Intercept)                0.1003
as.factor(treat)metcbt5    0.0543 .
as.factor(treat)scy        0.4150
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.9746663)

Number of Fisher Scoring iterations: 2
```

Note in this case that the estimated treatment effect of community on those exposed to the community treatment is slightly stronger than in the ATE case (high numbers are bad for the outcome variable). Although not statistically significant, such differences are compatible with the notion that the youths who actually received the community treatment responded more favorably to it than the "average" youth would have (where the average is taken across the whole collection of youths enrolled in the study).

The discussion in McCaffrey et al. (2013) may be useful for determining whether the ATE or ATT is of greater interest in a particular application.

4 Conclusion

Often, more than two treatments are available to study participants. If the study is not randomized, analysts may be interested in using a propensity score approach. Previously, few tools existed to aide the analysis of such data, perhaps tempting analysts to ignore all but two of the treatment conditions. We hope that this extension to the `twang` package will encourage more appropriate analyses of observational data with more than two treatment conditions.

Acknowledgements

The random subset of data was supported by the Center for Substance Abuse Treatment (CSAT), Substance Abuse and Mental Health Services Administration (SAMHSA) contract # 270-07-0191 using data provided by the following grantees: Adolescent Treatment Model (Study: ATM: CSAT/SAMHSA contracts # 270-98-7047, # 270-97-7011, #2 77-00-6500, # 270-2003-00006 and grantees: TI-11894, TI-11874, TI-11892), the Effective Adolescent Treatment (Study: EAT; CSAT/SAMHSA contract # 270-2003-00006 and grantees: TI-15413, TI-15433, TI-15447, TI-15461, TI-15467, TI-15475, TI-15478, TI-15479, TI-15481, TI-15483, TI-15486, TI-15511, TI15514, TI-15545, TI-15562, TI-15670, TI-15671, TI-15672, TI-15674, TI-15678, TI-15682, TI-15686, TI-15415, TI-15421, TI-15438, TI-15446, TI-15458, TI-15466, TI-15469, TI-15485,

TI-15489, TI-15524, TI-15527, TI-15577, TI-15584, TI-15586, TI-15677), and the Strengthening Communities-Youth (Study: SCY; CSAT/SAMHSA contracts # 277-00-6500, # 270-2003-00006 and grantees: TI-13305, TI-13308, TI-13313, TI-13322, TI-13323, TI-13344, TI-13345, TI-13354). The authors thank these grantees and their participants for agreeing to share their data to support the development of the `mnps` functionality.

References

[1] McCaffrey, D.F., B.A. Griffin, D. Almirall, M.E. Slaughter, R. Ramchand, and L.F. Burgette (2013). "A tutorial on propensity score estimation for multiple treatments using generalized boosted models." Forthcoming at *Statistics in Medicine*.

[2] Ridgeway, G., D. McCaffrey, B.A. Griffin, and L. Burgette (2006). "twang: Toolkit for weighting and analysis of non-equivalent groups." Available at http://cran.fyxm.net/web/packages/twang/.