Worked example of visualization tools when estimating effects in target populations

Karolinska Institutet
Tuesday, November 16, 2021

General breakdown

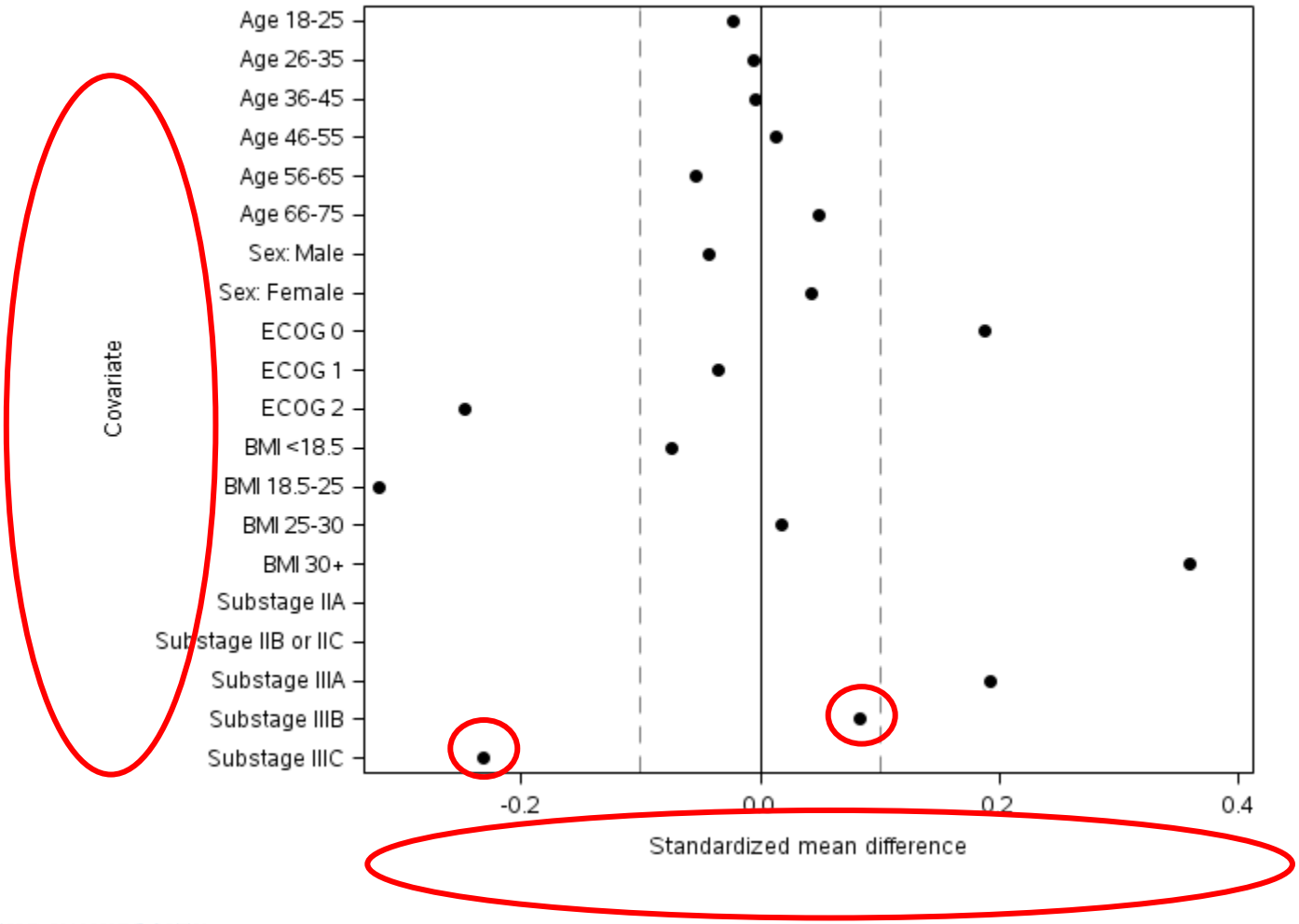
- Variable selection and importance figures
  - Breakout group session #1-10 minutes
- Diagnosing potential problems with weighted analyses
  - Breakout group session #2-10 minutes
- Performing and interpreting results
- General discussion

Variable selection figures

Variable selection/importance and external validity

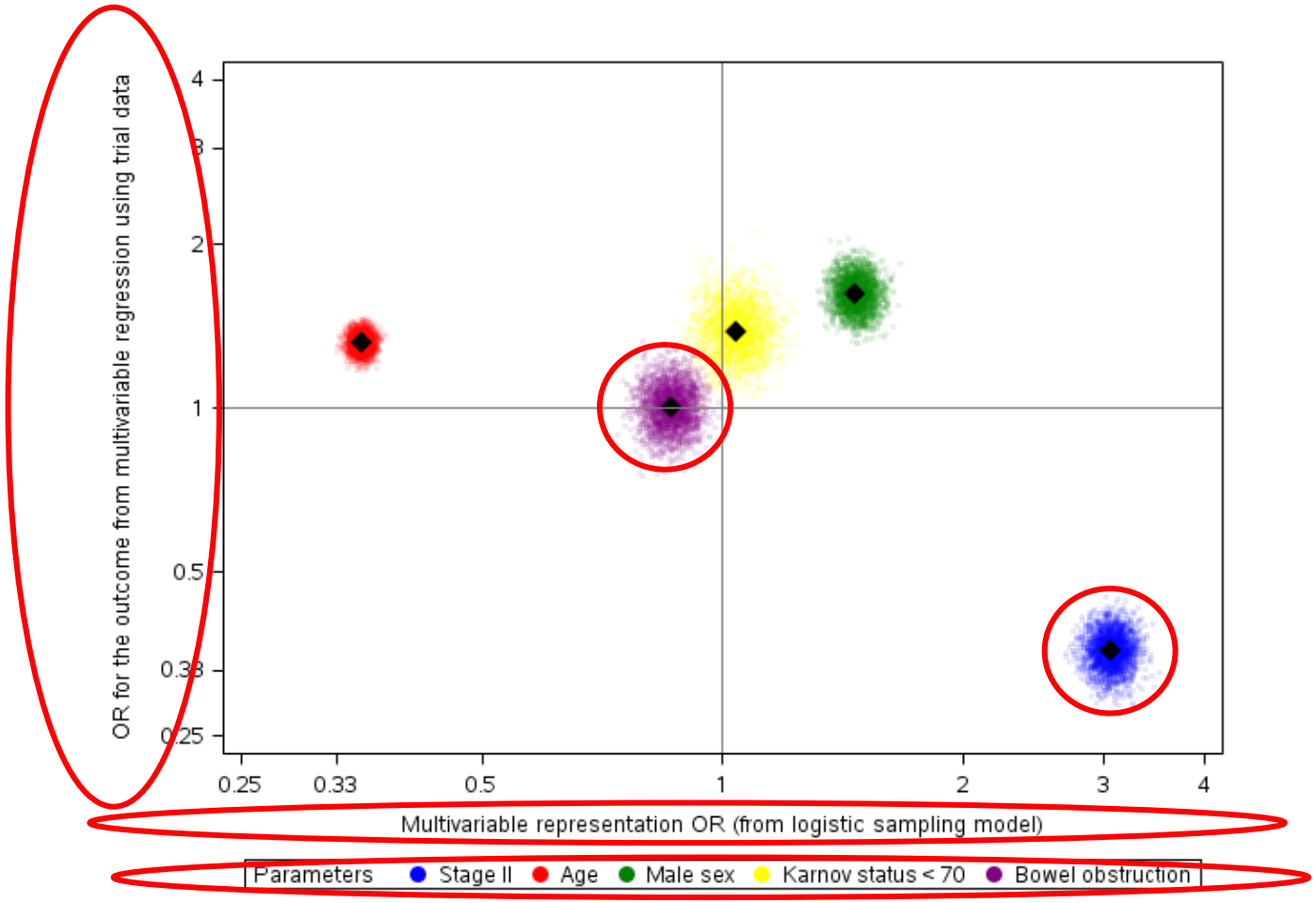
- Remember that whichever method we use, we need to select variables for an adjustment set
- There is ongoing work on using directed acyclic graphs and selection diagrams for this purpose
- Due to their non-parametric nature, these tools say very little about the relative importance of each variable
- Two plots can give some insight into variable importance
  - Love plots of standardized mean differences
  - A new plot (VITT plot) of multivariable relationships
  - These plots are **method-agnostic**

Love plots of standardized mean differences

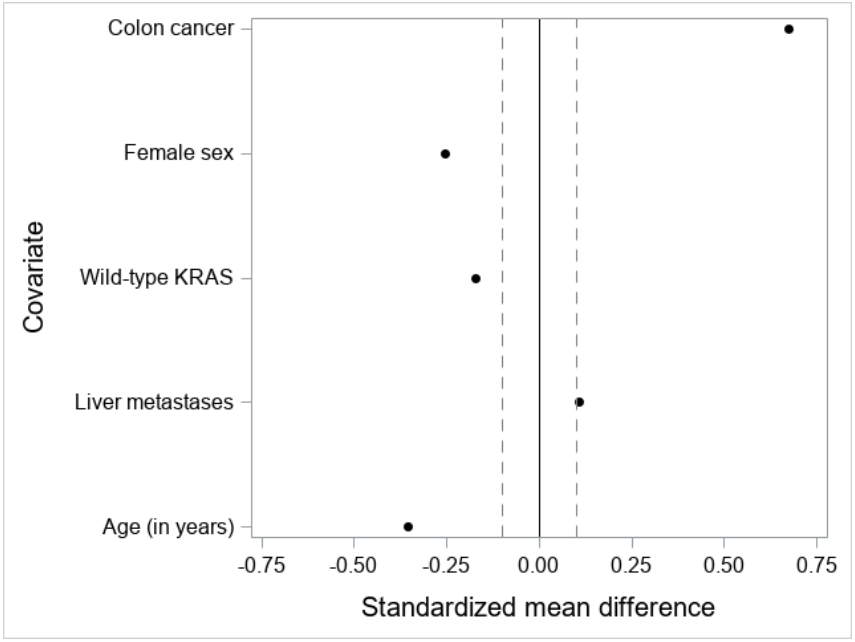


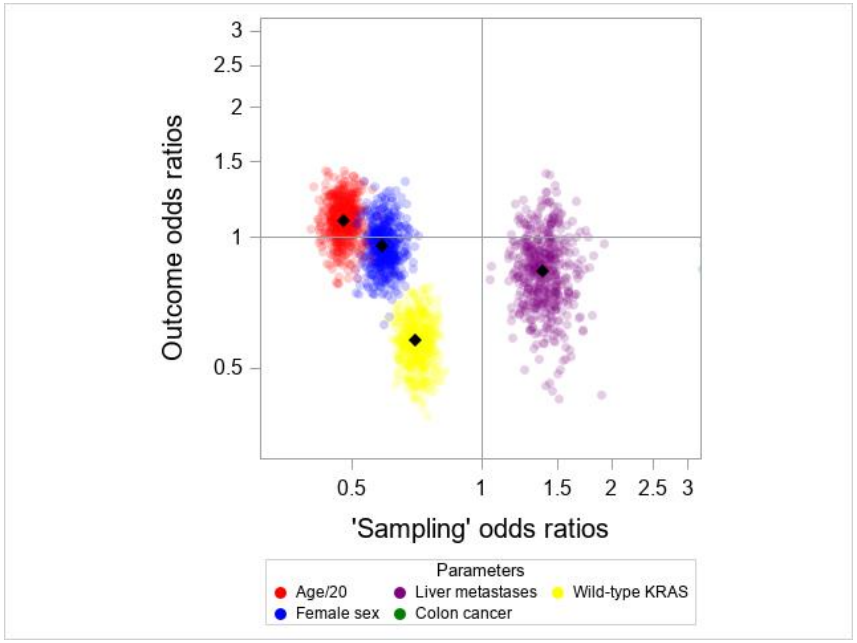
Code to reference: Love_plot_code

VITT plots of "sampling" and outcome ORs



Discussion time: LOVE and VITT plots





Code to create this plot: Love_plot_code

Code to create this plot: VITT_plot_code

Questions:

1) What variables seem most important for inclusion in the adjustment set?
2) What variables seem the least important?
3) Are there any variables you would consider dropping completely?

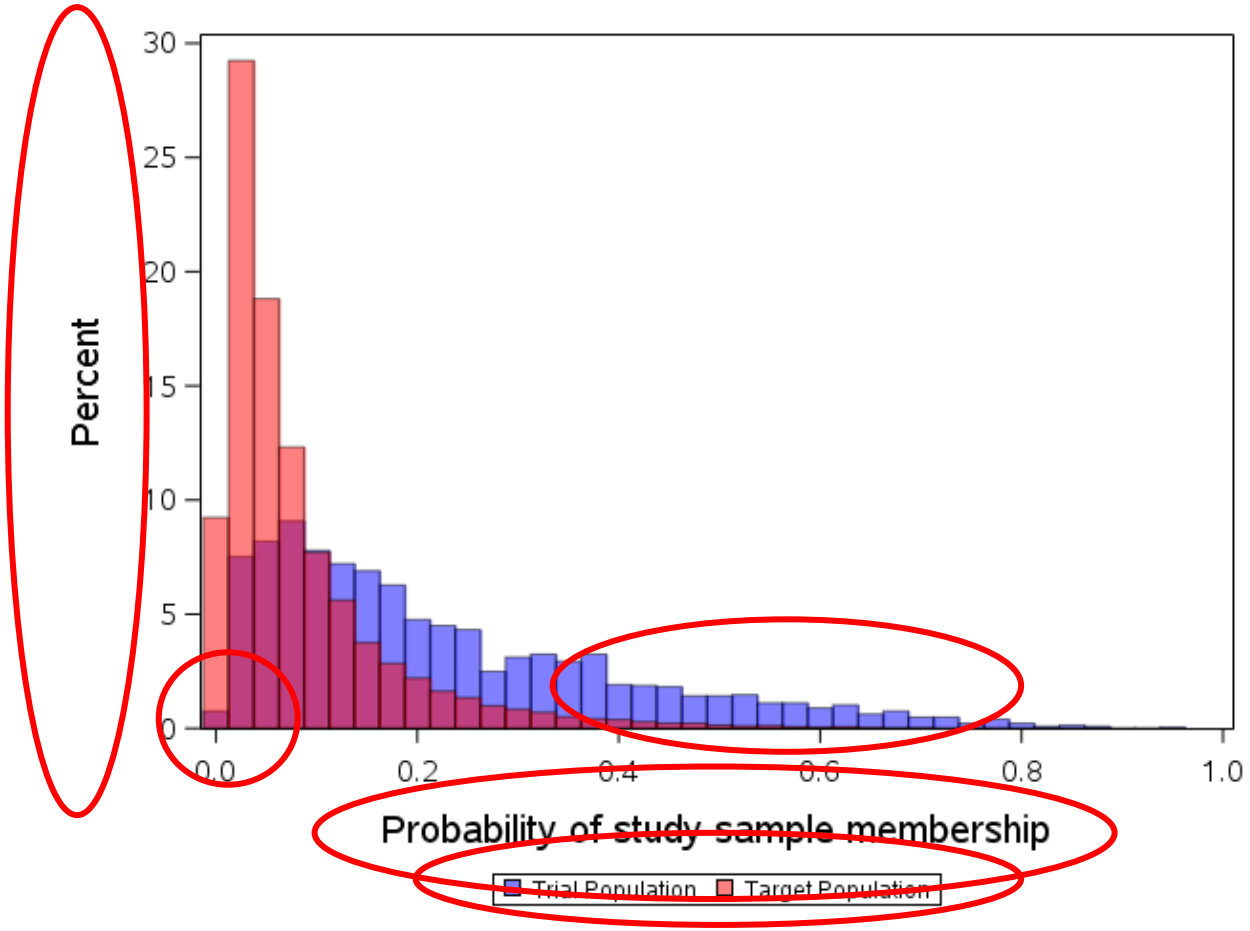
THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Diagnosing potential problems

Solving the external validity problem with weights

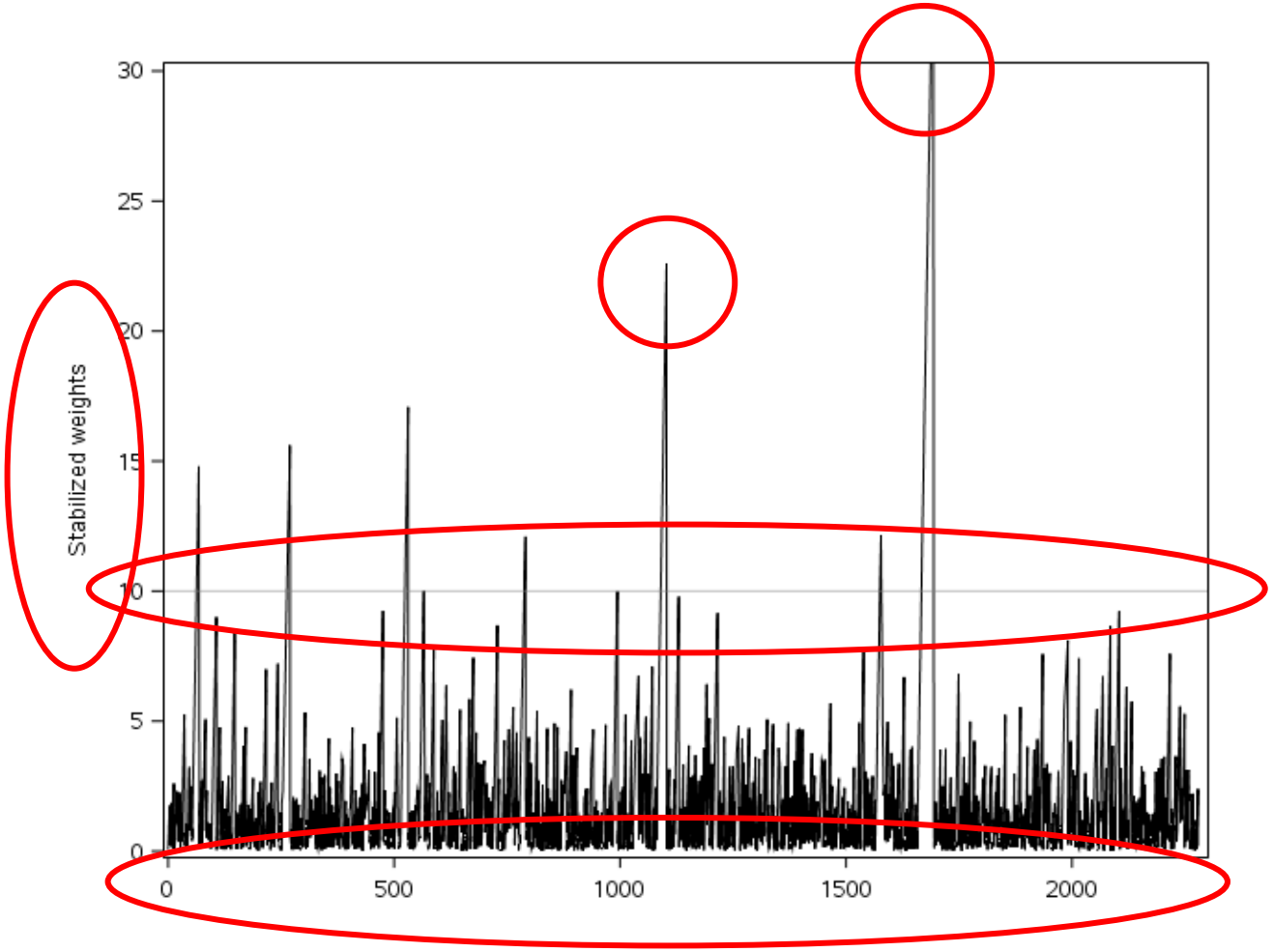
- Once you have identified the key variables, it is necessary to choose an analytic method
- Here we'll be focusing on weights and potential issues that can arise (especially the lack of transport positivity)
- When using these weights, some preliminary visualizations can help diagnose potential problems
  - Histograms and density plots of estimated probabilities
  - "Skyscraper" plots of the weights to be used

Histograms of “sampling” probabilities



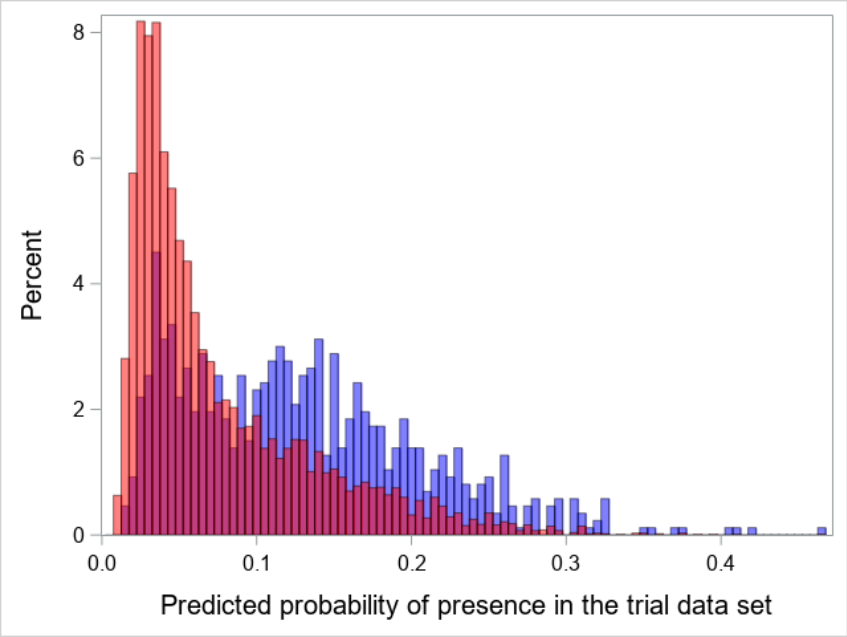
Code to reference: Density_plot_code

“Skyscraper” plots to identify high weights

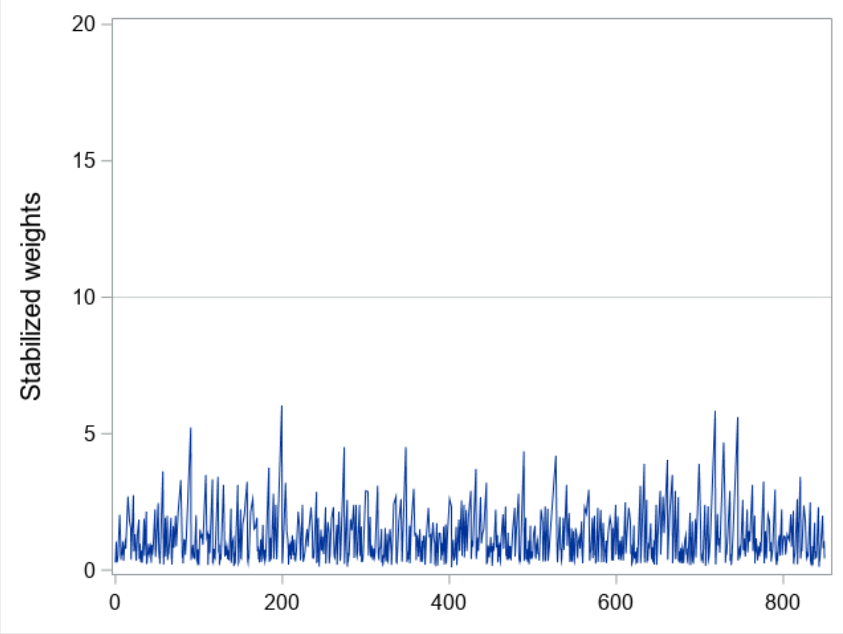


Code to reference: Skyscraper_plot_code

Discussion time: probability and skyscraper plots



Code to create this plot: Density_plot_code



Code to create this plot: Skyscraper_plot_code

Questions:

1) Do you see any potential problems with using this model based on the density plot?
2) What about the skyscraper plot? Do you see any potentially problematic weights?
3) Purely based on these plots, would you be comfortable proceeding with the analysis?

Analyzing data

Weighted analytics

- Having completed our diagnostics, how do we actually estimate a treatment effect?
- PROC PHREG conveniently allows use of a "WEIGHT" statement, as does many R packages for estimating hazard ratios
- Notably, traditional tools for estimating variance can ignore variability from sampling your target population
- Instead, we have bootstrapped, re-drawing from **both** the trial and target each iteration

Final results

Target population	Progression-free survival hazard ratio	95% confidence limits	Confidence limit ratio
PRIME trial	0.97	0.84, 1.11	1.33
SEER* population	**0.88**	**0.74, 1.06**	**1.43**

Questions?

mawc@live.unc.edu