Mitigating the Effect of Selection Bias

What is selection bias?
Assessing student success programs often includes challenges of mitigating the effects of selection bias. Selection bias occurs when randomization of a sample or population being assessed is not achieved. Programs and services are not assigned randomly to students; rather, students self-select into participation or gain access given institutional policies or other systematic conditions (e.g. course enrollment). Therefore, selection bias is inherent to programs and services at the University of Arizona. These biases need to be addressed within the assessment. Failure to not account for this will produce biased results. Fortunately, there are strategies available to mitigate these effects.

Strategy: Propensity Score Matching

Propensity score matching (PSM) creates a statistical comparison group with similar characteristics to program participants to better assess the effectiveness of intended outcomes. Propensity scores are calculated with logistic regression, which accounts for student characteristics influencing self-selection in a program. Propensity score values therefore reflect the likelihood of students participating in a program and are used to match participants with non-participants to estimate a program’s effect while reducing the role of selection bias.

A six step walkthrough to using PSM is available to UA community members here. 


Strategy: Heckman Sample Selection Model

The Heckman model accounts for selection bias in evaluations of programs with a two-step estimation procedure. First, the selection analysis uses regression to predict the actual group membership from a set of presumed factors. With a predicted group membership score, the second analysis also uses regression to evaluate the impact of a program while including the newly created membership score as a covariate. This in turn controls for selection bias that may be within the program assessment.

Nonresponse Bias: Selection Bias when Using Surveys

Overview
Surveys are often utilized to evaluate the impact of participation in programs and services on student success outcomes. Response rates vary across student demographics which introduces a form of selection bias called nonresponse bias. To diagnose if nonresponse bias exists, it is recommended to run a comparison of survey respondents to the population to determine how representative they are. A statistical difference indicates the presence of nonresponse bias and this bias should be mitigated. An example of this procedure can be found in the UA Student Services Fee and Green Fund Survey.

Strategies to reduce nonresponse bias:
- A proactive way to limit nonresponse bias is to follow-up with initial non-respondents during data collection. In some cases, an incentive may be warranted to support their participation in the survey.
- Post survey adjustments are used after data collection ends to reduce nonresponse bias. These include imputation and survey weighting procedures.

Contact Assessment, Research, & Grant Development for support when considering strategies to mitigate selection bias.