# Supplementary File 1: Numerical example

To provide some additional context to the theoretical results in the main text, Table S1 includes a small numerical example with one covariate (education) to show how Alt. 1 and Alt. 2 lead to the same generalization weights and that using them results in a sample distribution that matches the distribution in the population. In the example, individuals with low education are less likely to participate than the average individual, whereas individuals with high education are more likely to participate. Individuals with low education therefore receive a weight that is larger than one, while those with high education receive a weight below one.

**Table S1.** A small numerical example showing how generalization weights can be used to re-weight a biased sample to match population with respect to a single covariate (education).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Education (Z) | $$Pr(Z)$$ | $$Pr(Z|S)$$ | $$Pr(P|S)$$ | $$\tilde{Pr}(S^{\*}|Z)$$ | $$\tilde{Odds}(S^{\*}|Z)$$ | Weight | Weighted $Pr(Z|S)$ |
| Low | 0.33 | 0.20 | 0.30 | 0.23 | 0.30 | 1.67 | 0.33 |
| Mid | 0.33 | 0.33 | 0.50 | 0.33 | 0.50 | 1.00 | 0.33 |
| High | 0.33 | 0.47 | 0.70 | 0.41 | 0.70 | 0.71 | 0.33 |

**Notes:** The generalization weight is calculated using Pr(S)/Pr(S|Z) under Alt. 1 and $\tilde{Odds}(S^{\*})/\tilde{Odds}(S^{\*}|Z)$ under Alt. 2 (see text for details). The marginal probability that occurs in the numerator is 0.5 in this example. The weighted estimates are obtained by multiplying Pr(Z|S) with the corresponding weights.

**Definitions:** S = Participants; S\* = Participation as observed when the population data contains duplicates of the participants members (see text for details).