Suppression Criterion for Substate Region by Age Group Tables from the 2004-2006 National Surveys on Drug Use and Health[1]

The suppression criterion used for the 1999-2001 and 2002-2004 substate by age group small area estimates also has been used for the 2004-2006 substate by age group estimates. Suppressed estimates are noted by asterisks (*). The suppression criterion is based on a combination of the relative standard error (RSE) of $[-\ln(p)]$ or $[-\ln(1-p)]$ and the effective sample size (EFN) where p denotes the substate by age group level small area estimate and $\ln(p)$ denotes the natural logarithm of p. For $p \leq 50$ percent, an RSE of $[-\ln(p)]$ is used, and for $p > 50$ percent, an RSE of $[-\ln(1-p)]$ is used. The separate formulae for $p \leq 50$ percent and $p > 50$ percent produce a symmetric suppression rule; that is, if p is suppressed, then so will $(1-p)$. By using the first order Taylor series approximation method, an estimate of an RSE of $[-\ln(p)]$ and an RSE of $[-\ln(1-p)]$ is given by

$$\text{RSE of } [-\ln(p)] = \frac{\sqrt{\text{var}(p)}}{p[-\ln(p)]}$$

$$\text{RSE of } [-\ln(1-p)] = \frac{\sqrt{\text{var}(1-p)}}{(1-p)[-\ln(1-p)]}$$

where $\text{var}(p)$ denotes the posterior variance of p. The EFN is defined as $\text{EFN} = \dfrac{n}{design\ effect}$, where n denotes the raw sample size and design effect is defined as

$$design\ effect = \frac{\text{var}(p)}{[p(1-p)/n]}; \text{ hence, } \text{EFN} = \frac{p(1-p)}{\text{var}(p)}.$$ A lower bound of 0.2 also was imposed on the design effects (i.e., all design effects that were less than 0.2 were changed to 0.2) to avoid publishing substate by age group estimates with very small sample sizes. The following criterion was used in Tables 1 to 23:

When $p < 5.23$ percent, then suppress if RSE of $[-\ln(p)] > 10$ percent;

when 5.23 percent $\leq p \leq 94.77$ percent, then suppress if $EFN \leq 208$; and

when $p > 94.77$ percent, then suppress if RSE of $[-\ln(1-p)] > 10$ percent.

---

[1] This is a companion piece to the age group tables found in the report on *Substate Estimates from the 2004-2006 National Surveys on Drug Use and Health*. Estimates are based on a survey-weighted hierarchical Bayes estimation approach.

A graph describing the relationship between p and the EFN for an RSE of $[-\ln(p)]=10$ percent when $p \leq 50$ percent and for an RSE of $[-\ln(1-p)]=10$ percent when $p > 50$ percent is shown below (see Figure 1). The suppression criterion switches to EFN between 5.23 and 94.77 percent so that the EFN is not allowed to fall below the EFN of 208 required at $p = 50$ percent.

Figure 1. Small Area Estimate versus Effective Sample Size When the Relative Standard Error Equals 10 Percent.