

# Lab 4: The anchored multiplier

## A Bayesian approach for synthesizing estimates of the size of hidden populations

PB HLTH 250C (Spring 2024)

February 29, 2024

In their paper, Wesson, Mirzazadeh, and McFarland (2018) introduce a novel Bayesian approach for synthesizing *multiplier method* estimates of the size of hidden populations. Hidden populations are populations from which sampling is impossible, unethical, or simply too difficult. Often times, vulnerable populations are also hidden populations (e.g. queer youth, people living with HIV, survivors of abuse).

A common method for population size estimation is the *multiplier method*, which synthesizes programmatic data with a survey of the target population. Programmatic data provide the benchmark  $n$ , the number of individuals served, and survey data provide the multiplier  $p$ , the proportion of people in the sample who report receipt of service. The *multiplier method* estimate of the target population size is

$$e = \frac{n}{p}.$$

Unfortunately, this method may be particularly sensitive to biased estimates. The *anchored multiplier* of Wesson, Mirzazadeh, and McFarland (2018) may be more robust to the effect of biased data by “anchoring” a pooled estimate to an *a priori* belief.

### General steps

0. Gather *multiplier method* point estimates of the target population size to be combined as a synthesized estimate as well as their uncertainty bounds. Express as a fraction of the total population.
1. Specify a beta distribution to match each point estimate  $\mu_i$  for  $i = 1, \dots, k$  and their corresponding upper uncertainty bounds.
  - a) Compute the variance:

$$\text{variance}_i = \sigma_i^2 = \left( \frac{(\text{upper bound})_i - \mu_i}{1.96} \right)^2.$$

- b) If desired, compute the mixed effects adjusted variance:

$$(\text{adjusted variance})_i = \sigma_i^2 + \tau^2$$

where

$$\begin{aligned}\tau^2 &= \frac{Q - df}{C} = \frac{Q - (k - 1)}{C} \\ Q &= \sum_i \frac{\mu_i^2}{\sigma_i^2} - \frac{(\sum_i \mu_i / \sigma_i^2)^2}{\sum_i 1 / \sigma_i^2} \\ C &= \sum_i \frac{1}{\sigma_i^2} - \frac{\sum_i [1 / (\sigma_i^2)^2]}{\sum_i 1 / \sigma_i^2}.\end{aligned}$$

- c) Compute the shape parameters of the beta distribution:

$$\begin{aligned}\alpha &= -\frac{\mu(\sigma^2 + \mu^2 - \mu)}{\sigma^2} \\ \beta &= \frac{(\sigma^2 + \mu^2 - \mu)(\mu - 1)}{\sigma^2}.\end{aligned}$$

Use the adjusted variance for computing variance adjusted shape parameters.

2. Specify the likelihood

$$\text{round}(\alpha) \sim \text{binomial}(\theta, n = \text{round}(\alpha + \beta)).$$

3. Specify an appropriate prior for  $\theta$ .  
4. Compute or approximate the posterior distribution

## Application: Population size estimation for people who inject drugs

0. Several multiplier method estimates of the PWID population sizes for San Francisco in the years 2005, 2009, and 2012, were collected by Wesson, Mirzazadeh, and McFarland (2018). They may be found in the file `multiplier_method_estimates.csv`. Express the estimates and uncertainty bounds as a fraction of the total population size of San Francisco ( $N = 777\,660$  from the 2005 US Census).

```
estimates <- fread("multiplier_method_estimates.csv")
N <- 777660
```

1. Specify the beta distribution parameters for each point estimate and its associated upper bound (columns `p` and `ub`, respectively). Then, specify the beta distribution parameters using the adjusted variance, stratified by year.
2. Suppose we have prior mean  $\mu_0$  and prior 95% interval  $lb_0$  to  $ub_0$ . We can specify a beta distribution to serve as a prior using the formulas above. We will consider two beta priors, specified using the following prior means and prior intervals:  
0.026 (0.018 to 0.033) and  
0.003 (0.0019 to 0.0041).

We will also consider a uniform prior over 0 to 0.1288.

Compute the appropriate shape parameters for the two beta priors.

3. Run the JAGS script using `R2jags`.

```
library(R2jags)
```

Warning: package 'R2jags' was built under R version 4.2.3

```
binom_lik_beta_prior <- function() {
  # SAMPLING DISTRIBUTION
  for (i in 1:N) {alpha[i] ~ dbinom(theta, n[i])}
  # PRIORS
  theta ~ dbeta(alpha0, beta0);
}
binom_lik_uniform_prior <- function() {
  # SAMPLING DISTRIBUTION
  for (i in 1:N) {alpha ~ dbinom(theta, n[i])};
  # PRIORS
  theta ~ dunif(a0, b0);
```

}

---

Wesson, Paul D, Ali Mirzazadeh, and Willi McFarland. 2018. “A Bayesian Approach to Synthesize Estimates of the Size of Hidden Populations: The Anchored Multiplier.” <https://doi.org/10.1093/ije/dyy132>.