Cluster sampling with PPS

Jobs example continued

After creating a cumulative range to go with the data, we select four random numbers between 0 and 99. They are 22, 35, 58, and 65, so we select companies 1, 2, 5, and 7. Here are the calculations:

Company	employees	new jobs	$\overline{y}_i$	$M\overline{y}_i$
1	30	2.0	$2/30 = .0667$	6.67
2	10	.7	$.7/10 = .07$	7.0
5	5	.4	$.4/5 = .08$	8.0
7	5	.4	$.4/5 = .08$	8.0

Then

$$\widehat{\tau}_{pps} = \frac{M}{n}\sum_{i=1}^{n}\overline{y}_i = \frac{100}{4}[.2967] = 7.42 = \frac{1}{n}\sum_{i=1}^{n}M\overline{y}_i$$

To calculate $\widehat{V}(\widehat{\tau}_{pps})$, use $\widehat{\mu}_{pps} = \frac{1}{M}\widehat{\tau}_{pps}$ and $\sum(\overline{y}_i - \widehat{\mu}_{pps})^2 = .000142$:

$$\widehat{V}(\widehat{\tau}_{pps}) = \frac{M^2}{n(n-1)}\sum(\overline{y}_i - \widehat{\mu}_{pps})^2 = \frac{100^2}{4(3)}(.000142) = .118$$

So the bound is $B = .688$. Note that

$$\widehat{V}(\widehat{\tau}_{pps}) = \frac{M^2}{n(n-1)}\sum(\overline{y}_i - \widehat{\mu}_{pps})^2 = \frac{1}{n}\left[\frac{\sum(M\overline{y}_i - M\widehat{\mu}_{pps})^2}{n-1}\right] = \frac{1}{n}\left[\frac{\sum(\widehat{\tau}_i - \widehat{\tau}_{pps})^2}{n-1}\right]$$
$$= \frac{\widehat{\sigma}^2_{\hat{\tau}}}{n}$$