Sampling distribution

Stat 422

μ and σ^2

The population consists of `{1,2,3,4}`, and we will take samples of $n = 2$ to look at the sampling distribution (all possible samples)

$$\mu = \frac{\sum Y_i}{N} = \frac{1+2+3+4}{4} = 2.5$$

$$\sigma^2 = \frac{\sum (Y_i - \mu)^2}{N} = \frac{(1-2.5)^2 + (2-2.5)^2 + (3-2.5)^2 + (4-2.5)^2}{4} = \frac{5}{4} = 1.25$$

Sampling without replacement (swor)

Sampling without replacement implies dependence and it also the way we treat an finite population.

$E(\overline{y})$ and $\hat{V}(\overline{y})$

$$E(\overline{y}) = \sum \overline{y} p(\overline{y})$$

$$= \frac{1}{6}(1.5 + 2 + 2.5 + 2.5 + 3 + 3.5) = 2.5$$

Showing that $E(\overline{y}) = \mu$. This is the mean of the sampling distribution, which is equal to the population mean.

$$\hat{V}(\overline{y}) = \sum (\overline{y} - \mu)^2 p(\overline{y}) = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) = \left(1 - \frac{n}{N}\right)\frac{s^2}{n}$$

Where $E(s^2) = \frac{s_i^2}{n}$ where s_i^2 is the i^{th} sample variance and n here is the number of samples

$$E(s^2) = \frac{0.5 + 2 + 4.5 + 0.5 + 2 + 0.5}{6} = \frac{5}{3}$$

Sampling without replacement

Sample	**Frequency**	$P(sample)$	$\overline{y}_i$	s_i^2	$\hat{\tau}$	$\hat{V}(\overline{y})$
{1,2}	2	$\frac{1}{6}$	1.5	0.5	6	0.125
{1,3}	2	$\frac{1}{6}$	2	2	8	0.5
{1,4}	2	$\frac{1}{6}$	2.5	4.5	10	1.125
{2,3}	2	$\frac{1}{6}$	2.5	0.5	10	.125
{2,4}	2	$\frac{1}{6}$	3	2	12	0.5
{3,4}	2	$\frac{1}{6}$	3.5	0.5	14	0.125

So now

$$\hat{V}(\overline{y}) = \left(1 - \frac{n}{N}\right)\frac{s^2}{n} = \left(1 - \frac{2}{4}\right)\left(\frac{5/3}{2}\right) = \frac{5}{12}$$

Which is the same as

$$\hat{V}(\overline{y}) = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) = \frac{5/4}{2}\left(\frac{4-2}{4-1}\right) = \frac{5}{12}$$

Handy but only if you know the population variance.

Sampling with replacement

Sample	**Frequency**	$P(sample)$	$\overline{y}_i$	s_i^2	$\hat{\tau}$	$\hat{V}(\overline{y})$
{1,1}	1	$\frac{1}{16}$	1	0	4	0
{1,2}	2	$\frac{2}{16}$	1.5	0.5	6	0.125
{1,3}	2	$\frac{2}{16}$	2	2	8	0.5
{1,4}	2	$\frac{2}{16}$	2.5	4.5	10	1.125
{2,2}	1	$\frac{1}{16}$	2	0	8	0
{2,3}	2	$\frac{2}{16}$	2.5	0.5	10	.125
{2,4}	2	$\frac{2}{16}$	3	2	12	0.5
{3,3}	1	$\frac{1}{16}$	3	0	12	0
{3,4}	2	$\frac{2}{16}$	3.5	0.5	14	0.125
{4,4}	1	$\frac{1}{16}$	4	0	16	0

Sampling with replacement (swr)

Sampling with replacement implies a few things. One is independence (not always but often), another is that it will produce more "extreme" samples, and it also the way we would treat an infinite population.

$E(\overline{y})$ and $\hat{V}(\overline{y})$

$$E(\overline{y}) = \sum \overline{y}p(\overline{y})$$

$$= \frac{1}{16}(1+2+3+4) + \frac{2}{16}(1.5+2+2.5+2.5+3+3.5) = 2.5$$

Showing that again, $E(\overline{y}) = \mu$. This is the mean of the sampling distribution, which is equal to the population mean.

$$\hat{V}(\overline{y}) = \sum (\overline{y} - \mu)^2 p(\overline{y}) = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) = \left(1 - \frac{n}{N}\right)\frac{s^2}{n}$$

Where $E(s^2) = \frac{s_i^2}{n}$ where s_i^2 is the i^{th} sample variance and n here is the number of samples

$$E(s^2) = \frac{0+0.5+2+4.5+0+0.5+2+0+0.5+0}{10} = 1$$

So now

$$\hat{V}(\overline{y}) = \left(1 - \frac{n}{N}\right)\frac{s^2}{n} = \left(1 - \frac{2}{4}\right)\left(\frac{1}{2}\right) = \frac{1}{4}$$

A larger variance for sampling with replacement.