The Delta Method and Applications

This handout introduces the "Delta Method" of finding approximations based on Taylor series expansions to the variance of functions of random variables (whose variances are known). The basic idea will be illustrated with some simple examples, and the use of the Delta Method in sampling applications will be indicated.

For estimators, the delta method is important because while the theoretical variance of a mean of an SRS is known, there is no general theoretical expression for the variance of most functions of a mean, such as the inverse of a mean, or the ratio of two means. In the previous handout on ratio estimation, the variance of the ratio estimator r of the population ratio R was approximated as:

$$\text{Var}(r) \;=\; \text{Var}\left(\frac{\overline{y}}{\overline{x}}\right) \;\approx\; \left(\frac{N-n}{N}\right)\frac{1}{\mu_x^2}\cdot\frac{\sigma_r^2}{n}.$$

There is no exact expression for the variance of a ratio of two random variables in terms of the variances and covariance of the random variables.

Taylor Series Expansion: The Taylor series expansion of a function $f(\cdot)$ about a value a is given as:

$$f(x) = f(a) + f'(a)(x-a) + f''(a)\frac{(x-a)^2}{2!} + \cdots,$$

where we can often drop the higher order terms to give the approximation:

$$f(x) \approx f(a) + f'(a)(x-a).$$

Letting $a = \mu_x$, the mean of X, a Taylor series expansion of $y = f(x)$ about μ_x gives the approximation:

$$y = f(x) \approx f(\mu_x) + f'(\mu_x)(x-\mu_x).$$

Taking the variance of both sides yields:

$$\text{Var}(Y) = \text{Var}(f(X)) \approx [f'(\mu_x)]^2\text{Var}(X).$$

- So, if Y is any function of a random variable X, we need only calculate the variance of X and the first derivative of the function to approximate the variance of Y.

Example: Suppose $Y = X^2$. Then $f(x) = x^2$ and $f'(x) = 2x$, so that:

$$\text{Var}(Y) \approx (2\mu_x)^2\text{Var}(X) = 4\mu_x^2\sigma_x^2.$$

Example: Suppose $Y = 1/X$. Then $f(x) = 1/x$ and $f'(x) = -1/x^2$, so that:

$$\text{Var}(Y) \approx \left[-\frac{1}{\mu_x^2}\right]^2\text{Var}(X) = \frac{\sigma_x^2}{\mu_x^4}.$$

Two-Variable Taylor Series Expansion: Suppose now we have random variables X, Y. A Taylor series expansion of $f(x,y)$ about the values (x_0, y_0) is given by:

$$f(x,y) = f(x_0,y_0) + \left.\frac{\partial f(x,y)}{\partial x}\right|_{(x_0,y_0)} (x-x_0) + \left.\frac{\partial f(x,y)}{\partial y}\right|_{(x_0,y_0)} (y-y_0) + \left(\begin{array}{l}\text{2nd and higher}\\ \text{order terms}\end{array}\right)$$

Example: Suppose $f(x,y) = \frac{y}{x}$. Then: $\frac{\partial f(x,y)}{\partial x} = \frac{-y}{x^2}$, $\frac{\partial f(x,y)}{\partial y} = \frac{1}{x}$

$$\Longrightarrow \quad f(x,y) = \frac{y}{x} \approx \frac{\mu_y}{\mu_x} + \frac{-\mu_y}{\mu_x^2}(x-\mu_x) + \frac{1}{\mu_x}(y-\mu_y)$$

$$\Longrightarrow \quad \text{Var}\left(\frac{Y}{X}\right) \approx \frac{\mu_y^2}{\mu_x^4}\text{Var}(X) + \frac{1}{\mu_x^2}\text{Var}(Y) - \frac{2\mu_y}{\mu_x^3}\text{Cov}(X,Y).$$

(using the fact that the variance of the sum of two random variables is the sum of the variances plus two times the covariance). By analogy with the above result then, the approximate variance of the ratio estimator is:

$$\text{Var}\left(\frac{\overline{y}}{\overline{x}}\right) \approx \left[\frac{\mu_y^2}{\mu_x^4}\cdot\frac{\sigma_x^2}{n} + \frac{1}{\mu_x^2}\cdot\frac{\sigma_y^2}{n} - \frac{2\mu_y}{\mu_x^3}\cdot\frac{\rho\sigma_x\sigma_y}{n}\right]\cdot \underbrace{\left(\frac{N-n}{N}\right)}_{\substack{\text{if the fpc}\\ \text{is required}}},$$

where: $\text{Cov}(\overline{X},\overline{Y}) = \frac{\text{Cov}(X,Y)}{n} = \frac{\rho\sigma_x\sigma_y}{n}$.

- Is this the same as the approximate variance for the ratio estimator given earlier?

The corresponding estimated variance of the ratio estimator is given by:

$$\widehat{\text{Var}}\left(\frac{\overline{y}}{\overline{x}}\right) \approx \frac{1}{n}\left[\frac{\overline{y}^2}{\overline{x}^4}s_x^2 + \frac{1}{\overline{x}^2}s_y^2 - \frac{2\overline{y}}{\overline{x}^3}\widehat{\rho}s_x s_y\right].$$

Some Useful Approximations: The linear approximation via a Taylor series expansion gives the approximate variance for the following three useful functions of random variables X and Y where ρ is the correlation between X and Y.

1. $\text{Var}\left(\frac{1}{X}\right) = \left(\frac{1}{\mu_X^4}\right)\sigma_X^2$.
2. $\text{Var}\left(\frac{Y}{X}\right) = \left(\frac{\mu_Y^2}{\mu_X^4}\right)\sigma_X^2 + \left(\frac{1}{\mu_X^2}\right)\sigma_Y^2 - 2\left(\frac{\mu_Y}{\mu_X^3}\right)\rho\sigma_X\sigma_Y$.
3. $\text{Var}(XY) = \mu_Y^2\sigma_X^2 + \mu_X^2\sigma_Y^2 + 2\mu_X\mu_Y\rho\sigma_X\sigma_Y$.
   Note: If X and Y are independent, then an exact expression for $\text{Var}(XY)$ can be derived: $\text{Var}(XY) = \mu_Y^2\sigma_X^2 + \mu_X^2\sigma_Y^2 + \sigma_X^2\sigma_Y^2$.

To obtain estimates of these variances, simply substitute sample values of the means, variances and correlation.

Some applications:

- Estimating the Population Ratio: To estimate the population ratio, $R = \mu_Y/\mu_X$, from a simple random sample (x_1, y_1), ..., (x_n, y_n), the estimator is $r = \overline{y}/\overline{x}$. Using formula 2 above, and letting $\overline{X}$ and $\overline{Y}$ be the random variables, this formula yields the variance given earlier in this handout, and in equation (4) of page 60 of the text, although some algebra is involved to attain this simplified form.

- PPS sampling with replacement (Chapter 6, Section 1): The following summarizes the results in the notes on pp. 24-25, "Summary of results for PPS sampling", and includes estimated variances from the Delta Method where needed.

  As before, consider the following notation. Let:

$$\begin{aligned} N &= \text{the population size,} \\ n &= \text{the sample size,} \\ x_i &= \text{the size of the } i^{th} \text{ unit in the population,} \\ p_i &= x_i/\tau_x = \text{the probability of selecting the } i^{th} \text{ unit on each draw,} \\ y_i &= \text{the response variable (variable of interest).} \end{aligned}$$

  The estimates and approximate standard errors, using the Hansen-Hurwitz estimator where possible, of relevant population quantities are given below.

  1. Estimating N:

$$\widehat{N} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{p_i} \text{ (unbiased)}, \qquad \widehat{\text{Var}}(\widehat{N}) = \frac{1}{n(n-1)}\sum_{i=1}^{n}\left(\frac{1}{p_i} - \widehat{N}\right)^2 \text{ (Hansen-Hurwitz)}$$

  2. Estimating μ_x and τ_x:

$$\widehat{\mu}_x = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}} \text{ (biased)}, \qquad \widehat{\text{Var}}(\widehat{\mu}_x) = \left(\frac{1}{\overline{v}^4}\right)\frac{s_v^2}{n} \text{ (Delta Method)}$$

  where $v_i = \dfrac{1}{x_i}$, and $\overline{v}$ and s_v^2 are the sample mean and variance of the v_i's. Note that we can calculate the variance of the denominator because it is the mean of the $1/x_i$.

$$\widehat{\tau}_x = \frac{N}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}} \text{ (biased)}, \qquad \widehat{\text{Var}}(\widehat{\tau}_x) = N^2\widehat{\text{Var}}(\widehat{\mu}_x) = \left(\frac{N^2}{\overline{v}^4}\right)\frac{s_v^2}{n} \text{ (Delta Method)}$$

3. Estimating τ_y and μ_y:

$$\widehat{\tau}_y = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{p_i} \text{ (unbiased)}, \quad \widehat{\text{Var}}(\widehat{\tau}_y) = \frac{1}{n(n-1)}\sum_{i=1}^{n}\left(\frac{y_i}{p_i} - \widehat{\tau}_y\right)^2 = \frac{s_w^2}{n} \text{ (Hansen-Hurwitz)}$$

where $w_i = \dfrac{y_i}{p_i}$ and s_w^2 is the sample variance of the w_i's (note that $\widehat{\tau}_y = \overline{w}$).

Estimating μ_y, N known:

$$\widehat{\mu}_y = \frac{\widehat{\tau}_y}{N} \text{ (unbiased)}, \quad \widehat{\text{Var}}(\widehat{\mu}_y) = \frac{1}{N^2}\widehat{\text{Var}}(\widehat{\tau}_y) \text{ (Hansen-Hurwitz)}$$

Estimating μ_y, N unknown:

$$\widehat{\mu}_y = \frac{\widehat{\tau}_y}{\widehat{N}} = \frac{\frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{p_i}}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{p_i}} = \frac{\sum_{i=1}^{n}\frac{y_i}{x_i}}{\sum_{i=1}^{n}\frac{1}{x_i}} \text{ (biased)}$$

$$\widehat{\text{Var}}(\widehat{\mu}_y) = \left(\frac{\overline{t}^2}{\overline{v}^4}\right)\frac{s_v^2}{n} + \left(\frac{1}{\overline{v}^2}\right)\frac{s_t^2}{n} - 2\left(\frac{\overline{t}}{\overline{v}^3}\right)\frac{\widehat{\rho}_{t,v}s_t s_v}{n} \text{ (Delta Method)}$$

where $t_i = \dfrac{y_i}{x_i}$, $v_i = \dfrac{1}{x_i}$, and $\overline{t}, \overline{v}, s_t^2, \& s_v^2$ are the sample means and variances, and $\widehat{\rho}_{t,v}$ is the sample correlation between the t_i's and the v_i's.

- Other Applications: Suppose we want to estimate $R = \mu_y/\mu_x$ and we already have independent estimates $\widehat{\mu}_x$ and $\widehat{\mu}_y$ of the means available (from two different studies, for example) along with estimated variances $\widehat{\text{Var}}(\widehat{\mu}_x)$ and $\widehat{\text{Var}}(\widehat{\mu}_y)$ (i.e., the standard errors squared). It doesn't matter what sampling plans or sample sizes were used to generate these independent estimates as long as valid standard errors can be calculated. Then, we can estimate R by $\widehat{R} = \widehat{\mu}_y/\widehat{\mu}_x$ and, by the Delta Method, estimate $\text{Var}(\widehat{R})$ by:

$$\widehat{\text{Var}}(\widehat{R}) = \left(\frac{\widehat{\mu}_y^2}{\widehat{\mu}_x^4}\right)\widehat{\text{Var}}(\widehat{\mu}_x) + \left(\frac{1}{\widehat{\mu}_x^2}\right)\widehat{\text{Var}}(\widehat{\mu}_y).$$

  - Note that the covariance term drops out because these are assumed to be independent estimates.

Example: Suppose it is desired to estimate the average expenditure per day for visitors to Yellowstone National Park. One study estimated the average expenditures per trip, but did not obtain trip length information. The estimate was \$240 with a standard error of \$60. Another study estimated the average length of a trip but did not gather expenditure data. The estimate was 2.3 days with a standard error of 0.5 days. With these two studies then, the estimated expenditure per day is 240/2.3 = \$104.3 per person. The estimated variance of the estimate is

$$\left(\frac{240^2}{(2.3)^4}\right)(.5)^2 + \left(\frac{1}{(2.3)^2}\right)(60)^2 = 1195.1,$$

so the estimated standard error is $\sqrt{1195.1}$ = \$34.60.

If we were interested in estimating the product of two means for which we already had independent estimates of the individual means, we could use the exact formula for $\text{Var}(XY)$ in formula 3 on page 40 of this handout.