Topic 11

Central Limit Theorem

Propagation of Error and the Delta Method

Outline

Propagation of Error

Propagation of error or propagation of uncertainty is a strategy to estimate the impact on the standard deviation of the consequences of a *nonlinear* transformation of a measured quantity whose measurement is subject to some uncertainty.

Exercise. Show that

$$E[a + b(Y - \mu_Y)] = a \qquad \mathrm{Var}(a + b(Y - \mu_Y)) = b^2\mathrm{Var}(Y).$$

We will apply this to the linear approximation of $g(y)$ about the point μ_Y.

$$g(y) \approx g(\mu_Y) + g'(\mu_Y)(y - \mu_Y).$$

If we take expected values, then

$$Eg(Y) \approx E[g(\mu_Y) + g'(\mu_Y)(Y - \mu_Y)] = g(\mu_Y).$$

Propagation of Error

The variance

$$\text{Var}(g(Y)) \approx \text{Var}(g(\mu_Y) + g'(\mu_Y)(Y - \mu_Y)) = g'(\mu_Y)^2\sigma_Y^2.$$

Thus, the standard deviation $\sigma_{g(Y)} \approx |g'(\mu_Y)|\sigma_Y$ gives the formula for propagation of error.

Example. For Y, the measurement of a side of a cube with length q, Y^3 is an estimate of the volume. If the measurement error has standard deviation σ_Y, then, taking $g(y) = y^3$, we see that the standard deviation of the error in the measurement of the volume

$$\sigma_{Y^3} \approx 3q^2\sigma_Y.$$

If we estimate q with Y, then $\sigma_{Y^3} \approx 3Y^2\sigma_Y$.

Propagation of Error

In a effort to estimate the angle θ of the sun, the length s of a shadow from a 10 meter flag pole is measured. If $\sigma_{\hat{s}}$ is the standard deviation for the length measurement, we use propagation of error to estimate $\sigma_{\hat{\theta}}$, the standard deviation in the estimate of the angle. Using right triangle trigonometry, we have that

$$\theta = g(s) = \tan^{-1}\left(\frac{s}{10}\right). \quad \text{Thus,} \quad g'(s) = \frac{1}{1+(s/10)^2}\cdot\frac{1}{10} = \frac{10}{100+s^2}.$$

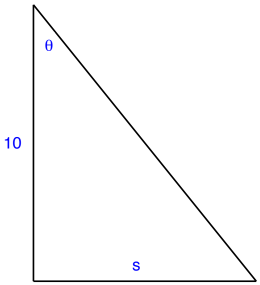
So,

$$\sigma_{\hat{\theta}} \approx \frac{10}{100+s^2}\cdot\sigma_s.$$

For example, set $\sigma_s = 0.1$ meter and $s = 1.26$. Then,

$$\hat{\theta} = \tan^{-1}(0.126) = 0.1253 \text{ radians} = 7.18^\circ$$

$$\sigma_{\hat{\theta}} \approx 0.0984 \cdot 0.1 = 0.00984 \text{ radians} = 0.564^\circ.$$



Propagation of Error

What happens if g is a function of several variables?

We will show the multivariate propagation of error in the two dimensional case noting that extension to the higher dimensional case is straightforward. For random variables Y_1 and Y_2 with

- means μ_1 and μ_2, and
- variances σ_1^2 and σ_2^2,

the linear approximation about the point (μ_1, μ_2) is

$$g(Y_1, Y_2) \approx g(\mu_1, \mu_2) + \frac{\partial g}{\partial y_1}(\mu_1, \mu_2)(Y_1 - \mu_1) + \frac{\partial g}{\partial y_2}(\mu_1, \mu_2)(Y_2 - \mu_2).$$

Take the expected value of this expression to see that

$$Eg(Y_1, Y_2) \approx g(\mu_1, \mu_2).$$

Propagation of Error

For Y_1 and Y_2 independent, we also have that the random variables

$$\frac{\partial g}{\partial y_1}(\mu_1, \mu_2)(Y_1 - \mu_1) \quad \text{and} \quad \frac{\partial g}{\partial y_2}(\mu_1, \mu_2)(Y_2 - \mu_2)$$

are independent. Using the Pythagorean identity and the quadratic identity for the sum of their variances, we have

$$\begin{aligned}\sigma^2_{g(Y_1,Y_2)} = \mathrm{Var}(g(Y_1, Y_2)) &\approx \mathrm{Var}\left(\frac{\partial g}{\partial y_1}(\mu_1, \mu_2)(Y_1 - \mu_1)\right) + \mathrm{Var}\left(\frac{\partial g}{\partial y_2}(\mu_1, \mu_2)(Y_2 - \mu_2)\right) \\ &= \left(\frac{\partial g}{\partial y_1}(\mu_1, \mu_2)\right)^2 \sigma_1^2 + \left(\frac{\partial g}{\partial y_2}(\mu_1, \mu_2)\right)^2 \sigma_2^2.\end{aligned}$$

and consequently, the standard deviation,

$$\sigma_{g(Y_1,Y_2)} \approx \sqrt{\left(\frac{\partial g}{\partial y_1}(\mu_1, \mu_2)\right)^2 \sigma_1^2 + \left(\frac{\partial g}{\partial y_2}(\mu_1, \mu_2)\right)^2 \sigma_2^2}.$$

Propagation of Error

In a effort to estimate the angle θ of the sun, now assume that the height h of the flagpole is also unknown. Then

$$\theta = g(h,s) = \tan^{-1}\left(\frac{s}{h}\right).$$

Exercise. Show that $$\frac{\partial}{\partial h} g(h,s) = -\frac{s}{s^2+h^2} \quad \text{and} \quad \frac{\partial}{\partial s} g(h,s) = \frac{h}{s^2+h^2}.$$

Let H and S be, respectively, the random variables for the measurement of flagpole height and shadow length. Then,

$$\sigma^2_{\hat{\theta}} = \mathrm{Var}(g(H,S)) \approx \left(\frac{\partial}{\partial h} g(h,s)\right)^2 \sigma^2_H + \left(\frac{\partial}{\partial s} g(h,s)\right)^2 \sigma^2_S = \frac{s^2\sigma^2_H + h^2\sigma^2_S}{(s^2+h^2)^2}$$

Note that if $\sigma_H = \sigma_S = \sigma$, then $\sigma_{\hat{\theta}} \approx \frac{\sigma}{\sqrt{s^2+h^2}}$.

Delta Method

Let's use repeated independent measurements, $Y_1, Y_2, \ldots Y_n$ to estimate a quantity q by its sample mean $\bar{Y}$. If each measurement has mean μ_Y and variance σ_Y^2, then $\bar{Y}$ has mean $q = \mu_Y$ and variance σ_Y^2/n. By the central limit theorem

$$Z_n = \frac{\bar{Y} - \mu_Y}{\sigma_Y/\sqrt{n}}$$

has a distribution that can be a approximated by a standard normal. We can apply the propagation of error analysis based on a linear approximation of $g(\bar{Y})$ to obtain

$$g(\bar{Y}) \approx g(\mu_Y), \quad \text{and} \quad \mathrm{Var}(g(\bar{Y})) \approx g'(\mu_Y)^2 \frac{\sigma_Y^2}{n}.$$

The reduction in the variance in the estimate of q propagates to a reduction in variance in the estimate of $g(q)$.

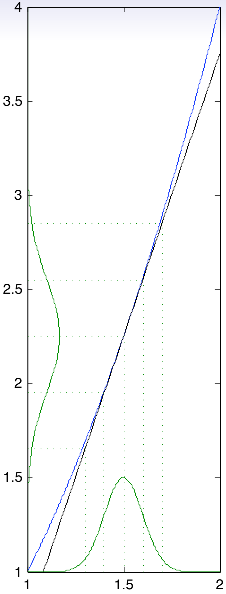
Delta Method

The delta method combines the central limit theorem and the propagation of error. To see use to write,

$$\frac{g(\bar{Y}) - g(\mu_y)}{\sigma_{g(\bar{Y})}} \approx \frac{g'(\mu_Y)(\bar{Y} - \mu_Y)}{|g'(\mu_Y)|\sigma_Y/\sqrt{n}} = \pm Z_n.$$

The $\pm$ sign depends on the sign of the derivative $g'(\mu_Y)$. Because the negative of a standard normal is itself also a standard normal, we have the desired approximation.

In this way, the delta method greatly extends the applicability of the central limit theorem.



Delta Method

Returning to the estimate the angle θ of the sun using measurements of the length s of a shadow from a 10 meter flag pole. Using right triangle trigonometry,

$$\theta = g(s) = \tan^{-1}\left(\frac{s}{10}\right) \quad \text{and} \quad g'(s) = \frac{10}{100 + s^2}.$$

Exercise. For $\sigma_s = 0.1$ meter and $s = 1.26$. Estimate $\sigma_{\hat{\theta}}$ for $n = 16$ measurements. Use the delta method to estimate

$$P\{\hat{\theta} \leq 7^\circ\}.$$

Delta Method

Example. The wavelength ℓ of sound at (unknown) frequency f and (known) speed of sound s is given by the formula

$$\ell = g(f) = \frac{s}{f}.$$

If we make n measurements $f_1, \ldots f_n$ of a frequency f_0, we can estimate ℓ by

$$\hat{\ell} = g(\bar{f}) = \frac{s}{\bar{f}} = \frac{340\text{m}/\text{sec}}{442/\text{sec}} = 0.7692\text{m},$$

assuming that the estimated speed of sound is 340 meters/second, and $\bar{f} = 442$ Hertz.

If the standard deviation of a single measure of the frequency is σ_f, then the standard deviation of $\hat{\ell}$

$$\sigma_{\hat{\ell}} \approx |g'(f_0)| \frac{\sigma_f}{\sqrt{n}} = \frac{s}{f_0^2} \frac{\sigma_f}{\sqrt{n}} \approx \frac{340\text{m}/\text{sec}}{442^2/\text{sec}^2} \cdot \frac{10/\text{sec}}{\sqrt{12}} = 0.005\text{m}$$

for $\sigma_f = 10$ Hertz and $n = 12$ measurements.

Delta Method

For propagation of error, the delta method,

Y_1 $Y_1 = (Y_{1,1}, \ldots, Y_{1,n_1})$ independent each with mean μ_1 and standard deviation σ_1
and
Y_2 $Y_2 = (Y_{2,1}, \ldots, Y_{2,n_2})$ independent each with mean μ_2 and standard deviation σ_2
independent observations vectors,

$$\sigma^2_{g(Y_1,Y_2)} \sigma^2_{g(\bar{Y}_1,\bar{Y}_2)} \approx \left(\frac{\partial g}{\partial y_1}(\mu_1,\mu_2)\right)^2 \sigma_1^2 \frac{\sigma_1^2}{n_1} + \left(\frac{\partial g}{\partial y_2}(\mu_1,\mu_2)\right)^2 \sigma_2^2 \frac{\sigma_2^2}{n_2}$$

and

$$Z_n = \frac{g(\bar{Y}_1, \bar{Y}_2) - g(\mu_1, \mu_2)}{\sigma_{g(\bar{Y}_1,\bar{Y}_2)}}$$

has a distribution that can be a approximated by a standard normal.

Delta Method

In avian biology, the fecundity B is defined as the number of female fledglings per female per year. $B > 1$ indicates a growing population and $B < 1$, a declining population. B is a product of three quantities,

$$B = F \cdot p \cdot N,$$

where

- F equals the mean number of female fledglings per successful nest,
- p equals nest survival probability, and
- N equals the mean number of nests built per female per year.

Let's

- collect measurement $F_1, \ldots, F_{n_F}$ on n_F nests to count female fledglings in a successful nest, and determine the sample mean $\bar{F}$,
- check n_p nests for survival probability, and determine the sample proportion $\hat{p}$, and
- follow n_N females to count the number $N_1, \ldots, N_{n_N}$ of successful nests per year and determine the sample mean $\bar{N}$.

Delta Method

Our experimental design is structured so that measurements are independent. Then, taking the appropriate partial derivatives for $B = g(F, p, N) = F \cdot p \cdot N$, we obtain an estimate for the variance of $\hat{B} = g(\bar{F}, \hat{p}, \bar{N})$,

$$\sigma_{\hat{B}}^2 \approx (pN)^2 \frac{\sigma_F^2}{n_F} + (FN)^2 \frac{\sigma_p^2}{n_p} + (pF)^2 \frac{\sigma_N^2}{n_N}.$$

The checks of nest survival form a sequence of Bernoulli trials. Thus, $\sigma_p^2 = p(1-p)$ for a Bernoulli random variable, we can write the expression above upon dividing by B^2 as

$$\begin{aligned} \left(\frac{\sigma_{\hat{B}}}{B}\right)^2 &\approx \frac{1}{n_F}\left(\frac{\sigma_F}{F}\right)^2 + \frac{1}{n_p}\left(\frac{\sigma_p}{p}\right)^2 + \frac{1}{n_N}\left(\frac{\sigma_N}{N}\right)^2 \\ &= \frac{1}{n_F}\left(\frac{\sigma_F}{F}\right)^2 + \frac{1}{n_p}\left(\frac{1-p}{p}\right) + \frac{1}{n_N}\left(\frac{\sigma_N}{N}\right)^2. \end{aligned}$$