First-Order Taylor Approximations

Taylor's theorem for a smooth function f of one variable says, for $x \approx a$,

$$f(x) = T_n(x) + R_n(x),$$

where

$$T_n(x) \;=\; f(a) + f'(a)(x-a) + \frac{1}{2!}f''(a)(x-a)^2 + \cdots + \frac{1}{n!}f^{(n)}(a)(x-a)^n,$$

and $R_n(x)$ is a remainder term which is not always small. Nevertheless, when f behaves like a line near the center $x = a$ of the Taylor expansion, it is reasonable to neglect the remainder term $R_1(x)$ and approximate f by the 1st-degree Taylor polynomial $T_1(x)$:

$$f(x) \;\approx\; f(a) + f'(a)(x-a), \qquad \text{for } x \text{ near } a. \tag{1}$$

In cases where f is a function of two variables x and y, then choosing a center (a, b) near which f behaves like a plane, it is reasonable to approximate $f(x, y)$

$$f(x,y) \;\approx\; f(a,b) + (x-a)\frac{\partial f}{\partial x}(a,b) + (y-b)\frac{\partial f}{\partial y}(a,b), \qquad \text{for } (x,y) \text{ near } (a,b). \tag{2}$$

Equations (1) and (2) are called first-order Taylor approximations of f near $x = a$ or (x, y) near (a, b) respectively, depending on whether f is a function of one variable or two. Naturally, an extension of these equations can be written for smooth functions f of k variables.

The delta method, and the propagation of uncertainty

One underlying variable. Suppose we know mean μ and variance σ^2 for a random variable X, but are interested in another random variable Y which is a transformation of X: $Y = g(X)$. In Chapter 3, we encountered the cdf method which, when the pdf or cdf of X is known, allows us (at least in the case of elementary transformations g) to find the cdf of Y. From that, we could potentially calculate expected value and variance for Y.

The delta method uses Taylor approximations to bypass all that. Using the first-order approximation (1) above, and choosing to center on μ (i.e., our choice of a in (1) is μ), we have

$$g(X) \approx g(\mu) + g'(\mu)(X-\mu)$$

$$\begin{aligned} \mathrm{E}(Y) &= \mathrm{E}(g(X)) \\ &\approx \mathrm{E}(g(\mu) + g'(\mu)(X-\mu)) \\ &= \mathrm{E}(g(\mu)) + \mathrm{E}(g'(\mu)(X-\mu)) \\ &= g(\mu) + g'(\mu)\,\underbrace{\mathrm{E}(X-\mu)}_{= E(X) - E(\mu) = \mu - \mu = 0.} \\ &= g(\mu), \end{aligned}$$

$$E(b + aX) = b + aE(X)$$

$Var(b + aX)$
$= a^2 Var(X)$

$$\begin{aligned}
\text{Var}(Y) &= \text{Var}(g(X)) \\
&\approx \text{Var}(g(\mu) + g'(\mu)(X-\mu)) \\
&= \text{Var}(g'(\mu)(X-\mu) \\
&= [g'(\mu)]^2 \,\text{Var}(X-\mu) \\
&= [g'(\mu)]^2 \,\text{Var}(X).
\end{aligned}$$

Now, the standard deviation is a common measure of uncertainty. Our approximation above can be seen as relating the uncertainty $\sqrt{\text{Var}(Y)}$ in Y to the uncertainty σ in X:

$$\sqrt{\text{Var}(Y)} \approx |g'(\mu)|\,\sigma.$$

Two underlying variables. Now suppose W is a random variable of interest, and it relies on two other random variables X, Y through a transformation $W = g(X,Y)$. Taking (μ_X, μ_Y) as our center in (2), the delta method leads to approximations for mean and variance of W

$$\begin{aligned}
\text{E}(W) &\approx g(\mu_X, \mu_Y), \\
\text{Var}(W) &\approx \text{Var}(X)\left[\frac{\partial g}{\partial X}(\mu_X,\mu_Y)\right]^2 + \text{Var}(Y)\left[\frac{\partial g}{\partial Y}(\mu_X,\mu_Y)\right]^2 + 2\,\text{Cov}(X,Y)\left[\frac{\partial g}{\partial X}(\mu_X,\mu_Y)\right]\left[\frac{\partial g}{\partial Y}(\mu_X,\mu_Y)\right] \\
&= \text{Var}(X)\left[\frac{\partial g}{\partial X}(\mu_X,\mu_Y)\right]^2 + \text{Var}(Y)\left[\frac{\partial g}{\partial Y}(\mu_X,\mu_Y)\right]^2
\end{aligned}$$

with the latter expression for $\text{Var}(W)$ in effect if X, Y are independent. Once again, we can turn this into an approximate formula relating the uncertainty $\sqrt{\text{Var}(W)}$ in W to the uncertainties σ_X, σ_Y in independent random variables X and Y:

$$\sqrt{\text{Var}(W)} \approx \sqrt{\left[\frac{\partial g}{\partial X}(\mu_X,\mu_Y)\right]^2 \sigma_X^2 + \left[\frac{\partial g}{\partial Y}(\mu_X,\mu_Y)\right]^2 \sigma_Y^2}.$$

Some caveats:

- These approximations rely on first-order Taylor approximations, and are only as good those approximations allow them to be.
- The specific transformation g may be relatively simple: $Y = aX + b$, $Y = X^2$, $W = X + Y$, in which case we may have already found *exact* formulas for expected value and variance using another approach (perhaps the cdf method). See, for instance, Lemma 2.5.4 and Theorem 3.8.9.

Example: $X \sim Unif(1, 2)$

So $\mu_X = 1.5$, $\sigma_X = \frac{1}{\sqrt{12}}$ (Obtained from back cover)

Consider $Y = \sqrt{X} = g(X)$

From our formulas (delta method)

$$E(Y) \approx g(\mu_X) = \sqrt{1.5} = 1.225$$

$$Var(Y) = [g'(\mu_X)]^2 \, Var(X)$$

$g(x) = \sqrt{x} = x^{1/2}$, $\quad g'(x) = \frac{1}{2} x^{-1/2}$

$$g'(\mu_X) = \frac{1}{2\sqrt{\mu_X}} = \frac{1}{2\sqrt{1.5}}$$

$$Var(Y) \approx \left(\frac{1}{2\sqrt{1.5}}\right)^2 \cdot \left(\frac{1}{\sqrt{12}}\right)^2 = \frac{1}{72} \doteq 0.0139$$

Our transformation $g(x) = \sqrt{x}$ isn't overly complicated.

Could use methods from earlier: cdf method.

Know: pdf for X $\quad f_X(x) = \begin{cases} 1, & \text{for } 1 \le x \le 2 \\ 0, & \text{otherwise} \end{cases}$

Can use to obtain cdf for Y

Denote cdf for Y as $F_Y(y)$

$$\underline{F_Y(y)} = P(Y \le y)$$

$$= P(\sqrt{X} \le y)$$

$$= P(X \le y^2)$$

$$= \begin{cases} 0, & y^2 < 1 \\ \int_1^{y^2} 1 \cdot dx, & 1 \le y^2 \le 2 \\ 1, & y^2 > 2 \end{cases}$$

$$\underline{F_Y(y)} = \begin{cases} 0, & \text{if } y < 1 \\ y^2 - 1, & \text{if } 1 \le y \le \sqrt{2} \\ 1, & \text{if } y > \sqrt{2} \end{cases}$$

True values for $\underline{E(Y)}$, $\underline{Var(Y)}$

$$E(Y) = \int_{-\infty}^{\infty} y \cdot \underbrace{f_Y(y)\,dy}_{\text{need pdf}}$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \begin{cases} 2y, & \text{if } 1 \le y \le \sqrt{2} \\ 0, & \text{otherwise} \end{cases}$$

$$E(Y) = \int_{-\infty}^{\infty} y\, f_X(y)\,dy = \int_{1}^{\sqrt{2}} y \cdot 2y\, dy$$

$$= \frac{2}{3} y^3 \Big|_1^{\sqrt{2}} = \boxed{\underset{\text{exact}}{\frac{2}{3}\left(\sqrt{8} - 1\right)}} \doteq 1.219$$

$$\text{Var}(Y) = E(Y^2) - \left[E(Y)\right]^2$$

and $\underline{E(Y^2)} = \int_{-\infty}^{\infty} y^2 f_Y(y)\,dy = \int_{1}^{\sqrt{2}} y^2 \cdot 2y\, dy$

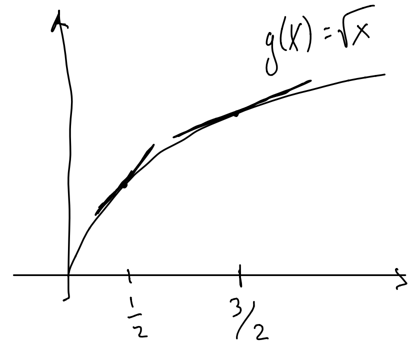
$$= \frac{1}{2} y^4 \Big|_1^{\sqrt{2}} = \frac{1}{2}(4 - 1) = \frac{3}{2}$$

$$= 1.5$$

$$\text{Var}(Y) = \underset{\text{exactly correct}}{\boxed{\frac{3}{2} - \left[\frac{2}{3}\left(\sqrt{8} - 1\right)\right]^2}} \doteq \underline{0.01416}$$

In 4.12, Rice explores

$$X \sim \underline{\text{Unif}(0,1)}$$
$$Y = \sqrt{X}$$



He shows that

$E(Y)$, $Var(Y)$ from delta method

are not very good approxs to exact ones obtained using cdf method.