Reminders: Propagation of Error, and Standard Errors for Derived Quantities

36-402, Advanced Data Analysis

Suppose we are trying to estimate some quantity θ. We compute an estimate $\widehat{\theta}$, based on our data. Since our data is more or less random, so is $\widehat{\theta}$. One convenient way of measuring the purely statistical noise or uncertainty in $\widehat{\theta}$ is its standard deviation. This is the **standard error** of our estimate of θ.[1] Standard errors are not the only way of summarizing this noise, nor a completely sufficient way, but they are often useful.

Suppose that our estimate $\widehat{\theta}$ is a function of some intermediate quantities $\widehat{\psi_1}, \widehat{\psi_2}, \ldots, \widehat{\psi_p}$, which are also estimated:

$$\widehat{\theta} = f(\widehat{\psi_1}, \widehat{\psi_2}, \ldots \widehat{\psi_p}) \tag{1}$$

For instance, θ might be the difference in expected values between two groups, with ψ_1 and ψ_2 the expected values in the two groups, and $f(\psi_1, \psi_2) = \psi_1 - \psi_2$. If we have a standard error for each of the original quantities $\widehat{\psi_1}$, it would seem like we should be able to get a standard error for the **derived quantity** $\widehat{\theta}$. There is in fact a simple if approximate way of doing so, which is called **propagation of error**[2].

We start with (what else?) a Taylor expansion. We'll write ψ_i^* for the true (ensemble or population) value which is estimated by $\widehat{\psi_i}$.

$$\begin{aligned} f(\psi_1^*, \psi_2^*, \ldots \psi_p^*) &\approx f(\widehat{\psi_1}, \widehat{\psi_2}, \ldots \widehat{\psi_p}) + \sum_{i=1}^{p} (\psi_i^* - \widehat{\psi_i}) \left. \frac{\partial f}{\partial \psi_i} \right|_{\psi = \widehat{\psi}} & (2) \\ f(\widehat{\psi_1}, \widehat{\psi_2}, \ldots \widehat{\psi_p}) &\approx f(\psi_1^*, \psi_2^*, \ldots \psi_p^*) + \sum_{i=1}^{p} (\widehat{\psi_i} - \psi_i^*) \left. \frac{\partial f}{\partial \psi_i} \right|_{psi = \widehat{\psi}} & (3) \\ \hat{\theta} &\approx \theta^* + \sum_{i=1}^{p} (\widehat{\psi_i} - \psi_i^*) f_i'(\widehat{\psi}) & (4) \end{aligned}$$

introducing f_i' as an abbreviation for $\frac{\partial f}{\partial \psi_i}$. The left-hand side is now the quantity whose standard error we want. I have done this manipulation because now it is a lin-

[1] It is not, of course, to be confused with the standard deviation of the data. It is not even to be confused with the standard error of the mean, unless θ is the expected value of the data and $\widehat{\theta}$ is the sample mean.

[2] Or, sometimes, the **delta method**.

ear function (approximately!) of some random quantities whose variances we know, and some derivatives which we can calculate.

Remember the rules for arithmetic with variances: if X and Y are random variables, and a, b and c are constants,

$$\begin{aligned} \text{Var}[a] &= 0 && (5)\\ \text{Var}[a+bX] &= b^2\text{Var}[X] && (6)\\ \text{Var}[a+bX+cY] &= b^2\text{Var}[X]+c^2\text{Var}[Y]+2bc\text{Cov}[X,Y] && (7)\end{aligned}$$

While we don't know $f(\psi_1^*, \psi_2^*, \ldots \psi_p^*)$, it's constant, so it has variance 0. Similarly, $\text{Var}\left[\widehat{\psi_i} - \psi_i^*\right] = \text{Var}\left[\widehat{\psi_i}\right]$. Repeatedly applying these rules to Eq. 4,

$$\text{Var}\left[\widehat{\theta}\right] \approx \sum_{i=1}^{p}(f_i'(\widehat{\psi}))^2\text{Var}\left[\widehat{\psi_i}\right] + 2\sum_{i=1}^{p-1}\sum_{j=i+1}^{p} f_i'(\widehat{\psi})f_j'(\widehat{\psi})\text{Cov}\left[\widehat{\psi_i},\widehat{\psi_j}\right] \quad (8)$$

The standard error for $\widehat{\theta}$ would then be the square root of this.

If we follow this rule for the simple case of group differences, $f(\psi_1, \psi_2) = \psi_1 - \psi_2$, we find that

$$\text{Var}\left[\widehat{\theta}\right] = \text{Var}\left[\widehat{\psi_1}\right] + \text{Var}\left[\widehat{\psi_2}\right] - \text{Cov}\left[\widehat{\psi_1},\widehat{\psi_2}\right] \quad (9)$$

just as we would find from the basic rules for arithmetic with variances. The approximation in Eq. 8 comes from the nonlinearities in f.

If the estimates of the initial quantities are uncorrelated, Eq. 8 simplifies to

$$\text{Var}\left[\widehat{\theta}\right] \approx \sum_{i=1}^{p}(f_i'(\widehat{\psi}))^2\text{Var}\left[\widehat{\psi_i}\right] \quad (10)$$

and, again, the standard error of $\widehat{\theta}$ would be the square root of this. This special case is sometimes called *the* propagation of error formula, but that seems like a bad usage.