

Sum and Product Operators, Partial Derivatives, Likelihood Functions

Leif E. Peterson, Ph.D.

Dept. of Medicine; Dept. of Molec. & Human Genetics

Baylor College of Medicine

Faculty Center, Office 10.57, ext. 8-5386

Summation and Product Operators

We begin this section by looking at summation and product operators that are commonly used in likelihood equations. Recall that most likelihood functions equate to the product of many probabilities in the form

$$f(x) = \prod_{i=1}^n p_i = p_1 \times p_2 \times \cdots \times p_n. \quad (1)$$

We already know that $\log(AB) = \log(A) + \log(B)$, so for example, we can rewrite this as $\log(p_1 p_2) = \log(p_1) + \log(p_2)$. The product operator

then becomes the summation symbol over the same range of x values

$$\log f(x) = \sum_{i=1}^n p_i = p_1 + p_2 + \cdots + p_n. \quad (2)$$

Partial Derivatives

In order to review partial derivatives, it is essential to have a strong foothold in solving derivatives of functions. Let's assume the function

$$f(x) = 3x^2 + 4y^5 \quad (3)$$

The first derivative is

$$f(x)' = 6x + 20y^4 \quad (4)$$

and the second derivative is

$$f(x)'' = 6 + 80y^3 \quad (5)$$

Now, the partial derivative of the function with respect to (w.r.t) y is

$$\frac{\partial(f(x))}{\partial y} = 20y^4 \quad (6)$$

As one can see partial derivatives focus on taking the derivative of whatever term is in the denominator. As another example, assume the function

$$f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \quad (7)$$

the partial derivative of the function with respect to β_1 is

$$\frac{\partial(f(x))}{\partial \beta_1} = x_1 \quad (8)$$

Likelihood Functions

Consider the exponential function that commonly appears in likelihood functions for many of the multiplicative regression models. The derivative of any exponentiated function is simply the derivative of the function times the original exponentiated function. In notation, this is written in

functional form as

$$\frac{\partial e^u}{\partial u} = e^u du \quad (9)$$

Let us now work with a likelihood “kernel” function replacing the function u with

$$p(x_i) = e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}} \quad i = 1, 2, \dots, n; j = 1, 2, \dots, p \quad (10)$$

where x_{ij} are the covariate values for p covariates ($j = 1, 2, \dots, p$) for the i th subject ($i = 1, 2, \dots, n$). We already know that complete the likelihood function for all subjects in the data set is

$$f(x) = \prod_{i=1}^n p_i \quad (11)$$

A general property of likelihood functions is that they are not linear. A commonly used method for “linearizing” likelihood functions is to take the natural logarithm of the kernel, which makes it more amenable for

solving. Taking the log of our function now results in the form

$$\begin{aligned}\log f(x) &= \sum_{i=1}^n p_i \\ &= \sum_{i=1}^n e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}}\end{aligned}\tag{12}$$

Next, recall that our goal is to maximize the likelihood, so we will need to find values of the β parameters at the maxima of the log likelihood function. First, to find the maximum of the likelihood due to the contribution of β_1 , we must find first partial derivative of the function w.r.t β_1 is known as the *score*, which is

$$\frac{\partial(f(x))}{\partial\beta_1} = x_{i1} e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}}\tag{13}$$

and the second derivative of $f(x)$ w.r.t β_2 , called the *information*, is

$$\frac{\partial(f(x))}{\partial\beta_1\partial\beta_2} = x_{i1}x_{i2} e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}}\tag{14}$$

Scores, or first partial derivatives of the log likelihood function w.r.t coefficients are typically arranged in a p -dimension vector $U(\beta_0, \beta_1, \dots, \beta_p)$, whereas the second derivatives form a $p \times p$ information matrix $\mathbf{I}(\beta_j, \beta_k)$. Last, another common element seen in likelihood functions involves the partial derivatives of the log function.

$$f(x) = \log(y) \tag{15}$$

The partial derivative of log function is

$$\frac{\partial \log(y)}{\partial y} = \frac{1}{y} \tag{16}$$