

Variances of Fitted Category Probabilities and Logits by the Delta Method for the Nested Logit Model

John Fox

2023-05-16

1 Notation

Let $\phi^{(j)}$ represent the probability that the dichotomous response in the j th nested dichotomous logit model is $Y_j = 1$ (i.e., a “success”), $j = 1, \dots, m-1$, where m is the number of response categories for the polytomy. Then $1 - \phi^{(j)}$ is the probability that $Y_j = 0$ (i.e., a “failure”). Let ϕ_k , $k = 1, \dots, m$ represent the probability that the polytomous response is $Y = k$. Let $\hat{\phi}^{(j)}$ and $\hat{\phi}_k$ represent the estimates of these probabilities. In the sequel, which involves only the estimates, I’ll omit the hats so as to simplify the notation.

In the nested logit model, the polytomous probabilities ϕ_k are each products of probabilities $\phi^{(j)}$ or $1 - \phi^{(j)}$ for $j = 1, \dots, m-1$. Let $\phi^{(j,k_j)}$ represent either $\phi^{(j)}$ or $1 - \phi^{(j)}$, as appropriate for category k of the polytomous response. Then

$$\begin{aligned}\phi_k &= \phi^{(1,k_1)} \times \dots \times \phi^{(m-1,k_{m-1})} \\ &= \prod_{j=1}^{m-1} \phi^{(j,k_j)}\end{aligned}$$

for $k = 1, \dots, m$.

Let $\lambda^{(j)}$ represent the (estimated) logit (log-odds) for the j th dichotomous logit model, with variance $V(\lambda)$.

Finally, the individual-category probabilities ϕ_k can be converted into logits, $\lambda_k = \log[\phi_k/(1 - \phi_k)]$. The estimates of these logits should approach asymptotic normality more rapidly than the estimates of the corresponding probabilities.

2 Variances of the Estimated Probabilities

The estimated probability of success $\phi^{(j)}$ for the j th dichotomous logit model is

$$\phi^{(j)} = \frac{1}{1 + e^{-\lambda^{(j)}}}$$

where $\lambda^{(j)} = \alpha^{(j)} + \beta_1^{(j)} x_1 + \dots + \beta_p^{(j)} x_p$ is a function of the regression coefficients, and the probability of failure is

$$1 - \phi^{(j)} = \frac{1}{1 + e^{\lambda^{(j)}}}$$

Then, the derivatives of $\phi^{(j)}$ and $1 - \phi^{(j)}$ with respect to $\lambda^{(j)}$ are

$$\begin{aligned} \frac{d\phi^{(j)}}{d\lambda^{(j)}} &= \frac{e^{-\lambda^{(j)}}}{(1 + e^{-\lambda^{(j)}})^2} \\ \frac{d(1 - \phi^{(j)})}{d\lambda^{(j)}} &= -\frac{e^{\lambda^{(j)}}}{(1 + e^{\lambda^{(j)}})^2} \end{aligned}$$

And by the univariate delta method

$$\begin{aligned} V(\phi^{(j)}) &\approx \left(\frac{d\phi^{(j)}}{d\lambda^{(j)}} \right)^2 V(\lambda^{(j)}) \\ &= \left[\frac{e^{-\lambda^{(j)}}}{(1 + e^{-\lambda^{(j)}})^2} \right]^2 V(\lambda^{(j)}) \\ V(1 - \phi^{(j)}) &\approx \left[\frac{d(1 - \phi^{(j)})}{d\lambda^{(j)}} \right]^2 V(\lambda^{(j)}) \\ &= \left[\frac{e^{\lambda^{(j)}}}{(1 + e^{\lambda^{(j)}})^2} \right]^2 V(\lambda^{(j)}) \end{aligned}$$

The variances of the estimated response-category probabilities for the polytomous response can be obtained similarly by the multivariate delta method, recognizing that these probabilities are products of the dichotomous probabilities. The result is greatly simplified because the dichotomies are independent, and so the covariance matrix of the estimated dichotomous probabilities is diagonal.

The required derivatives are

$$\frac{\partial \phi_k}{\partial \phi^{(j, k_j)}} = \prod_{j' \neq j} \phi^{(j', k_{j'})}$$

for $j = 1, \dots, m-1$ and $k = 1, \dots, m$. Then

$$\begin{aligned} V(\phi_k) &\approx \sum_{j=1}^{m-1} \left(\frac{\partial \phi_k}{\partial \phi^{(j, k_j)}} \right)^2 V(\phi^{(j, k_j)}) \\ &= \sum_{j=1}^{m-1} \left(\prod_{j' \neq j} \phi^{(j', k_{j'})} \right)^2 V(\phi^{(j, k_j)}) \end{aligned}$$

for $k = 1, \dots, m$.

Yet another application of the delta method produces approximate variances for the individual-category logits. The relevant derivative is

$$\frac{d\lambda_k}{d\phi_k} = \frac{1}{\phi_k(1 - \phi_k)}$$

for $k = 1, \dots, m$, and so

$$\begin{aligned} V(\lambda_k) &\approx \left(\frac{d\lambda_k}{d\phi_k} \right)^2 V(\phi_k) \\ &= \left[\frac{1}{\phi_k(1 - \phi_k)} \right]^2 V(\phi_k) \end{aligned}$$