STA 6384, Report 3.9

Carson Slater *Baylor University*

Problem:

On p. 74, Agresti informally justifies the needed convergence to normality of the right-hand side of (3.1.26) when the MLE is substituted for σ. Provide a formal version of his "proof," justifying each of his steps. (Hint: Use Slutsky's theorem.)

Let $\hat{\pi}$ be the vector of sample proportions from a multinomial(n, π) distribution. Let $g(\pi)$ be a function with continuous first partial derivatives in a neighborhood of the true parameter vector π.

1. **Asymptotic Normality of $\hat{\pi}$:** By the Multivariate Central Limit Theorem, the sample proportion vector $\hat{\pi}$ is asymptotically normal:

$$\sqrt{n}(\hat{\pi} - \pi) \xrightarrow{d} N(0, \Sigma)$$

   where Σ is the covariance matrix of the multinomial proportions.

2. **Delta Method:** Applying the Delta Method to the function $g(\pi)$, we get:

$$\sqrt{n}(g(\hat{\pi}) - g(\pi)) \xrightarrow{d} N(0, \sigma^2)$$

   where the asymptotic variance is $\sigma^2 = \nabla g(\pi)^T \Sigma \nabla g(\pi)$. We assume $\sigma^2 > 0$.

3. **Standardized Statistic:** Dividing by the constant (but unknown) standard deviation σ, we obtain a pivotal quantity that converges to a standard normal distribution:

$$A_n = \frac{\sqrt{n}(g(\hat{\pi}) - g(\pi))}{\sigma} \xrightarrow{d} Z \sim N(0, 1)$$

Formal Justification using Slutsky's Theorem

Our goal is to show the effect of replacing the unknown σ with its consistent estimator $\hat{\sigma}$. The statistic of interest is:

$$T_n = \frac{g(\hat{\pi}) - g(\pi)}{\hat{\sigma}/\sqrt{n}}$$

We can rewrite this statistic as a product, as suggested by Agresti:

$$T_n = \left(\frac{\sqrt{n}(g(\hat{\pi}) - g(\pi))}{\sigma}\right) \cdot \left(\frac{\sigma}{\hat{\sigma}}\right)$$

Let's analyze the two terms in this product.

The First Term As established by the Delta Method in step 3 above, the first term converges in distribution to a standard normal random variable:

$$A_n = \frac{\sqrt{n}(g(\hat{\pi}) - g(\pi))}{\sigma} \xrightarrow{d} N(0,1)$$

The Second Term To analyze the second term, $B_n = \sigma/\hat{\sigma}$, we first need to show that $\hat{\sigma}$ is a consistent estimator of σ.

1. By the Weak Law of Large Numbers (WLLN), the sample proportion vector $\hat{\pi}$ converges in probability to the true proportion vector π:

$$\hat{\pi} \xrightarrow{p} \pi$$

2. The asymptotic standard deviation σ is a function of π, i.e., $\sigma = h(\pi) = \sqrt{\nabla g(\pi)^T \Sigma(\pi) \nabla g(\pi)}$. The estimator $\hat{\sigma}$ is obtained by substituting $\hat{\pi}$ for π: $\hat{\sigma} = h(\hat{\pi})$. Since the components of $\nabla g(\cdot)$ are assumed continuous and the elements of $\Sigma(\cdot)$ are polynomial functions of its arguments, the function $h(\cdot)$ is continuous.
3. By the **Continuous Mapping Theorem**, since $\hat{\pi} \xrightarrow{p} \pi$ and h is a continuous function, we have:

$$\hat{\sigma} = h(\hat{\pi}) \xrightarrow{p} h(\pi) = \sigma$$

This shows that $\hat{\sigma}$ is a consistent estimator for σ.
4. Now consider the ratio $B_n = \sigma/\hat{\sigma}$. Since $\hat{\sigma} \xrightarrow{p} \sigma$ and σ is a non-zero constant, we can again apply the Continuous Mapping Theorem (with the function $f(x) = \sigma/x$, which is continuous at $x = \sigma$) to conclude:

$$B_n = \frac{\sigma}{\hat{\sigma}} \xrightarrow{p} \frac{\sigma}{\sigma} = 1$$

Conclusion via Slutsky's Theorem

We have shown that:

- $A_n = \frac{\sqrt{n}(g(\hat{\pi}) - g(\pi))}{\sigma} \xrightarrow{d} N(0,1)$
- $B_n = \frac{\sigma}{\hat{\sigma}} \xrightarrow{p} 1$

Slutsky's Theorem states that if $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{p} c$ (where c is a constant), then $X_n Y_n \xrightarrow{d} cX$.

Applying Slutsky's Theorem to our statistic $T_n = A_n \cdot B_n$, with $X_n = A_n$, $X \sim N(0,1)$, $Y_n = B_n$, and $c = 1$, we get:

$$T_n = A_n B_n \xrightarrow{d} Z \cdot 1, \quad \text{where } Z \sim N(0,1)$$

Therefore, we have formally shown that:

$$\frac{g(\hat{\pi}) - g(\pi)}{\hat{\sigma}/\sqrt{n}} \xrightarrow{d} N(0,1)$$

This justifies the use of $\hat{\sigma}$ in place of σ for constructing large-sample Wald-type confidence intervals and hypothesis tests.