STA 6351, Report.1.13

Carson Slater *Baylor University*

1.13

1. Proof that the Quadratic Approximation (1.11.1) is Exact for the Normal Case

Let $X_1, \ldots, X_n \overset{iid}{\sim} N(\theta, \sigma^2)$ (with σ^2 known). The log-likelihood is $\ell(\theta) = \text{const} - \frac{1}{2\sigma^2}\sum_{i=1}^n (X_i - \theta)^2$. The MLE is $\hat{\theta} = \bar{X}$. The sum of squares can be decomposed as:

$$\sum_{i=1}^n (X_i - \theta)^2 = \sum_{i=1}^n [(X_i - \hat{\theta}) + (\hat{\theta} - \theta)]^2 = \sum_{i=1}^n (X_i - \hat{\theta})^2 + n(\hat{\theta} - \theta)^2$$

The log-likelihood becomes:

$$\ell(\theta) = \left[\text{const} - \frac{1}{2\sigma^2}\sum_{i=1}^n (X_i - \hat{\theta})^2\right] - \frac{n}{2\sigma^2}(\theta - \hat{\theta})^2 = \ell(\hat{\theta}) - \frac{n}{2\sigma^2}(\theta - \hat{\theta})^2$$

The log-likelihood ratio is exactly:

$$\log \frac{L(\theta)}{L(\hat{\theta})} = \ell(\theta) - \ell(\hat{\theta}) = -\frac{n}{2\sigma^2}(\theta - \hat{\theta})^2 \quad (*)$$

The observed information is $I(\theta) = -l''(\theta)$. The derivatives are:

$$l'(\theta) = \frac{n}{\sigma^2}(\bar{X} - \theta) \implies l''(\theta) = -\frac{n}{\sigma^2}$$

Thus, $I(\hat{\theta}) = \frac{n}{\sigma^2}$. Substituting $I(\hat{\theta})$ into the quadratic approximation (1.11.1):

$$-\frac{1}{2}I(\hat{\theta})(\theta - \hat{\theta})^2 = -\frac{1}{2}\left(\frac{n}{\sigma^2}\right)(\theta - \hat{\theta})^2 = -\frac{n}{2\sigma^2}(\theta - \hat{\theta})^2$$

Comparing with $(*)$, the approximation is **exact**.

2. Establishment of (1.11.2)

In the Normal case, the MLE $\hat{\theta} = \bar{X}$ has the exact sampling distribution:

$$\hat{\theta} \sim N\left(\theta, \frac{\sigma^2}{n}\right)$$

Standardizing $\hat{\theta}$ gives:

$$\frac{\hat{\theta}-\theta}{\sqrt{\mathbb{V}(\hat{\theta})}} = \frac{\hat{\theta}-\theta}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

From Section 1, the observed information at the MLE is $I(\hat{\theta}) = \frac{n}{\sigma^2}$. Thus, $\sqrt{\mathbb{V}(\hat{\theta})} = \sqrt{\frac{\sigma^2}{n}} = \frac{1}{\sqrt{I(\hat{\theta})}} = I^{-1/2}(\hat{\theta})$. Substituting this, we get:

$$\frac{\hat{\theta}-\theta}{I^{-1/2}(\hat{\theta})} = I^{1/2}(\hat{\theta})(\hat{\theta}-\theta) \sim N(0,1)$$

Since the normal distribution is symmetric about zero:

$$I^{1/2}(\hat{\theta})(\theta-\hat{\theta}) \sim N(0,1)$$

3. Check that (1.11.4) is Exactly True in the Normal Case

Equation (1.11.4) is $-I^{-1/2}(\hat{\theta})S(\theta) \approx I^{1/2}(\hat{\theta})(\theta-\hat{\theta})$. In the Normal case, $S(\theta) = \frac{n}{\sigma^2}(\hat{\theta}-\theta) = -I(\hat{\theta})(\theta-\hat{\theta})$. Substituting the exact score into the Left-Hand Side (LHS):

$$\begin{aligned}\text{LHS} &= -I^{-1/2}(\hat{\theta})S(\theta) = -I^{-1/2}(\hat{\theta})\left[-I(\hat{\theta})(\theta-\hat{\theta})\right] \\ &= I^{-1/2}(\hat{\theta})I(\hat{\theta})(\theta-\hat{\theta})\end{aligned}$$

Using $a^{-1/2}a^1 = a^{1/2}$:

$$\text{LHS} = I^{1/2}(\hat{\theta})(\theta-\hat{\theta}) = \text{RHS}$$

The relation is an **exact equality** for the Normal case.

4. Proof that Changing the Scale of θ has No Effect on the Approximation

Consider the reparameterization $\varphi = a\theta$ with $a > 0$. This implies $\theta = \varphi/a$ and $\hat{\varphi} = a\hat{\theta}$. Let $\ell_\theta(\theta)$ be the log-likelihood function. The reparameterized log-likelihood is $\ell_\varphi(\varphi) = \ell_\theta(\theta) = \ell_\theta(\varphi/a)$. The chain rule is used to relate the derivatives, noting that $\frac{d\theta}{d\varphi} = \frac{1}{a}$:

1. **Score Function ($S(\theta)$):**

$$S_\varphi(\varphi) = \frac{\partial \ell}{\partial \varphi} = \frac{\partial \ell}{\partial \theta}\frac{\partial \theta}{\partial \varphi} = S_\theta(\theta)\cdot\frac{1}{a}$$

2. **Observed Information ($I(\theta)$):**

$$I_\varphi(\varphi) = -\frac{\partial^2 \ell}{\partial \varphi^2} = -\frac{\partial}{\partial \varphi}\left(\frac{1}{a}S_\theta(\theta)\right) = -\frac{1}{a}\frac{\partial S_\theta(\theta)}{\partial \theta}\frac{\partial \theta}{\partial \varphi} = -\frac{1}{a^2}\ell''_\theta(\theta) = \frac{1}{a^2}I_\theta(\theta)$$

3. **Maximum Likelihood Estimator (MLE):**

$$\hat{\varphi} = a\hat{\theta}$$

Invariance of the Scaled Score Function (Left Side)

We examine the term $-I^{-1/2}(\hat{\theta})S(\theta)$ under the reparameterization, substituting $I_\varphi(\hat{\varphi}) = \frac{1}{a^2}I_\theta(\hat{\theta})$ and $S_\varphi(\varphi) = \frac{1}{a}S_\theta(\theta)$:

$$\begin{aligned} -I_\varphi^{-1/2}(\hat{\varphi})S_\varphi(\varphi) &= -\left(\frac{1}{a^2}I_\theta(\hat{\theta})\right)^{-1/2}\left(\frac{1}{a}S_\theta(\theta)\right) \\ &= -\left(a^2 I_\theta^{-1}(\hat{\theta})\right)^{1/2}\frac{1}{a}S_\theta(\theta) \\ &= -\left(aI_\theta^{-1/2}(\hat{\theta})\right)\frac{1}{a}S_\theta(\theta) = -I_\theta^{-1/2}(\hat{\theta})S_\theta(\theta) \end{aligned}$$

The left side is unchanged.

Invariance of the Scaled Parameter Difference (Right Side)

We examine the term $I^{1/2}(\hat{\theta})(\theta - \hat{\theta})$ under the reparameterization, substituting $I_\varphi(\hat{\varphi}) = \frac{1}{a^2}I_\theta(\hat{\theta})$ and $\varphi - \hat{\varphi} = a(\theta - \hat{\theta})$:

$$\begin{aligned} I_\varphi^{1/2}(\hat{\varphi})(\varphi - \hat{\varphi}) &= \left(\frac{1}{a^2}I_\theta(\hat{\theta})\right)^{1/2}(a\theta - a\hat{\theta}) \\ &= \left(\frac{1}{a}I_\theta^{1/2}(\hat{\theta})\right)a(\theta - \hat{\theta}) \\ &= I_\theta^{1/2}(\hat{\theta})(\theta - \hat{\theta}) \end{aligned}$$

The right side is unchanged.

Since both sides of the approximation (1.11.4) remain identical after reparameterizing θ to $\varphi = a\theta$, the approximation is dimensionless and unaffected by the scale of the parameter.