STAT 231 — Lecture 11

Bartosz Antczak *Instructor: Michael Wallace* *October 2, 2017*

Recall from Last Lecture

If we observe data $\{y_1, y_2, \ldots, y_n\}$ that are all distributed using the Poisson model. The likelihood function for that distribution is:

$$L(\theta) = \prod_{i=1}^{n} P(Y = y_i; \theta)$$

Calculating the function, we see $\widehat{\theta} = \bar{y}$.

11.1 More Likelihood Functions

11.1.1 Likelihood Function for Exponential Distribution

Recall that the exponential distribution has probability density function

$$f(y; \theta) = \frac{1}{\theta} e^{-\frac{y}{\theta}}$$

If we observe data $\{y_1, y_2, \ldots, y_n\}$. The likelihood function is then

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-\frac{y_i}{\theta}} \tag{11.1}$$

$$= \frac{1}{\theta^n} e^{-\frac{\sum_{i=1}^{n} y_i}{\theta}} \tag{11.2}$$

$$= \theta^{-n} e^{-n\bar{y}/\theta} \tag{11.3}$$

Maximizing this function yields

$$\ell(\theta) = -n \left(\log(\theta) + \frac{\bar{y}}{\theta} \right) \implies \frac{d}{d\theta} \ell(\theta) = -n \left(\frac{1}{\theta} - \frac{\bar{y}}{\theta^2} \right)$$

Solving $\frac{d}{d\theta} \ell(\theta) = 0$ shows that $\widehat{\theta} = \bar{y}$. $\square$

11.1.2 Likelihood Function for Gaussian Data

The probability distribution function for a dataset $Y \sim G(\mu, \sigma)$ is

$$f(y; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2} (y - \mu)^2 \right]$$

To derive the max likelihood function, we take the product as usual:

$$L(\theta) = \prod_{i=1}^{n} f(y; \mu, \sigma) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2}(y_i - \mu)^2\right] \tag{11.4}$$

$$= (2\pi)^{-n/2}\sigma^{-n} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2\right] \tag{11.5}$$

Now we compute $\ell(\theta) = \log L(\theta)$:

$$\ell(\theta) = -n\log(\sigma) - \left[\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2\right] \tag{11.6}$$

Observe now that we must maximize for both σ and μ. So we solve for each variable individually:

$$\frac{d}{d\mu}\ell(\theta) = \frac{n}{\sigma^2}(\bar{y} - \mu) \tag{11.7}$$

$$\frac{d}{d\sigma}\ell(\theta) = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(y_i - \mu)^2 \tag{11.8}$$

Solving $\frac{d\ell}{d\mu} = 0$ and $\frac{d\ell}{d\sigma} = 0$ yields $\widehat{\mu} = \bar{y}$ and $\widehat{\sigma} = \left[\frac{1}{n}\sum_{i=1}^{n}(y_i - \mu)^2\right]^{\frac{1}{2}}$.

11.1.3 Invariance Property

The **invariance property** is

If $\widehat{\theta}$ is the maximum likelihood estimate of θ, then $g(\widehat{\theta})$ is the maximum likelihood estimate of $g(\theta)$