STAT 231 — LECTURE 10

Bartosz Antczak *Instructor: Michael Wallace* *September 29, 2017*

10.1 Likelihood Functions for Independent Experiments

Today we'll look at determining the likelihood functions of multiple datasets.

Example 10.1.1. *Two surveys*

Assume we have data $n_1 = 2000$, $y_1 = 520$ (consider this the successes), and $n_2 = 2000$, $y_2 = 540$. We define our probability functions using a binomial distribution:

$$P(Y_1 = y_1; \theta) = \binom{2000}{520} \theta^{520} (1 - \theta)^{1480}$$

$$P(Y_2 = y_2; \theta) = \binom{2000}{540} \theta^{520} (1 - \theta)^{1460}$$

We assume that these surveys are indepdendent, then we define the maximum likelihood function as

$$L(\theta) = P(Y_1 = y_1; \theta) P(Y_2 = y_2; \theta)$$

Doing the math, we arrive at $\widehat{\theta} = 0.265$.
In general, if we observe data $Y = (Y_1, Y_2, \ldots, Y_n)$ that are all independent and identically distributed each with probability function $P(Y_i = y_i; \theta)$, then

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

Example 10.1.2. *Hockey goals*

We're given data about the number of goals the Toronto Maple Leafs scored over 10 games:

$$[4, 4, 4, 2, 4, 4, 3, 3, 1, 1]$$

For this dataset, we choose a Poisson distribution:

$$P(Y_i = y_i; \theta) = \frac{\theta^{y_i} e^{-\theta}}{y_i!}$$

Assuming that our data is independent and distributed the same, we have:

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

Now let's solve for $\widehat{\theta}$:

$$L(\theta) = \left[\prod_{i=1}^{n} \frac{1}{y_i!}\right] \left[\prod_{i=1}^{n} \theta^{y_i}\right] \left[\prod_{i=1}^{n} e^{-\theta}\right] \tag{10.1}$$

$$= \left[\prod_{i=1}^{n} \frac{1}{y_i!}\right] \theta^{\sum_{i=1}^{n} y_i} e^{-n\theta} \tag{10.2}$$

Observe that the y_i value is constant, so we can ignore it since it has no effect on the value of θ. Observe that we now must differentiate $L(\theta)$, which will result in a pretty complex equation, so let's use our handy-dandy log function!

$$\ell(\theta) = \log\left(\theta^{n\bar{y}} e^{-n\theta}\right) \tag{10.3}$$

$$= \log\left(\theta^{n\bar{y}}\right) - n\theta \tag{10.4}$$

$$= n\bar{y}\log(\theta) - n\theta \tag{10.5}$$

Now deriving, we have:

$$\frac{d}{d\theta}\ell(\theta) = \frac{n\bar{y}}{\theta} - n \tag{10.6}$$

So $\ell'(\theta) = 0 \implies \widehat{\theta} = \bar{y}$. $\square$

10.1.1 Likelihood Function for a Random Sample

The past two examples focused on data from different processes under the same distribution. Let's now consider the data from the *same* process repeated n times. In this case, nothing changes. Our likelihood function remains the same:

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