Binomial Inference

Pedro Victor Brasil Ribeiro

2021-12-19 - Last changed in 2021-12-19

Supose you have a sample that comes from a binomial distribution, and the sample size is igual to m. How can you estimate the parameter p and n With only that sample?

To solve this problem we use a estimator deriving from the likelihood function, which is define as:

$$L(x;\theta) = \prod_{i=1}^{n} f(x_i;\theta) \tag{1}$$

And then we find the maximum point of the function, or in other words, the point that make the derivative of $f(x;\theta)$ igual to 0.

The Problem

Let $X_1, X_2, \cdots, X_m$ be a random variable i.i.d. (Independent and identically distributed) a binomial sample of m observations with parameters n and p, in other words $X \sim Bin(n,p)$. Then we know that the density function of a binomial is:

$$f(x;\{n,p\}) = \binom{n}{x} p^x (1-p)^{n-x}; \qquad n \in \mathbb{Z}^* \quad p \in (0,1) \tag{2}$$

So first of all we need to find the likehood function for a binomial distribution so using the equation 1, where $f(x;\theta)$ for $\theta = \{n,p\}$, expressed in the equation 2.

Estimation

Usually is more convenient to work with the derivative of $log[(L(x;\theta)]$, note that the maximum point of $L(x;\theta)$ and $log[L(x;\theta)]$ is the same point. Since we are looking at the maximum point then we know that $L'(x;\theta) = 0$, so:

$$\begin{aligned}
log'[L(x;\theta)] =& \frac{\partial log[L(x;\theta)]}{\partial \theta} \\
=& \frac{1}{L(x;\theta)} \frac{\partial L(x;\theta)}{\partial \theta}; \quad \text{but we know that in the maximum point} \frac{\partial L(x;\theta)}{\partial \theta} = 0 \\
\Rightarrow \frac{\partial log[L(x;\theta)]}{\partial \theta} =& 0
\end{aligned}$$

$\forall x_0 \quad$ That $f(x_0;\theta) \neq 0$.

$$
\begin{aligned}
L(x;\theta) &= \prod_{i=1}^{n} f(x_i;\theta) \\
&= \prod_{i=1}^{n} \binom{n}{x_i} p^{x_i}(1-p)^{n-x_i} \\
\Rightarrow l(x;\theta) &= \sum_{i=1}^{n} log\Big[\binom{n}{x_i}\Big] + \sum_{i=1} x_i log(p) + \sum_{i=1}^{n}(n-x_i)log(1-p)
\end{aligned}
$$

Estimate p, with n known

Estimate n, with p know

Both n and p unknow