Maximizing Likelihood: Examining the Negative Binomial Model

Natraj Vairavan
University of California, Berkeley

April 2022

1 Introduction

The concept of Maximum Likelihood Estimate, or maximizing the probability of success to maximize the total probability of getting a certain number of successes, already exists. In this article, however, I try to show how, given a certain number of successes and the probability of success, there is a way to choose the number of trials so that we maximize total probability.

We use the following negative binomial model:

$$P(X = k) \sim \binom{n-1}{k-1} p^k (1-p)^{n-k}$$

k is the number of successes, n is the number of trials, and p is the probability of success.

2 Maximum Likelihood Estimate

In order to optimize p, or the probability of success, we can simply use a partial derivative with respect to p and set it equal to zero.

$$\frac{\partial}{\partial p} \frac{(n-1)!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} = 0$$

$$\frac{(n-1)!}{(k-1)!(n-k)!} [kp^{k-1}(1-p)^{n-k} + (n-k)p^k(1-p)^{n-k-1}] = 0$$

$$\frac{(n-1)!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} \left[\frac{k(1-p) - p(n-k)}{p(1-p)}\right] = 0$$

We can cancel out $\frac{(n-1)!}{(k-1)!(n-k)!}p^k(1-p)^{n-k}$ as long as $p \neq 1$. Thus, we are left with:

$$\frac{k-np}{p(1-p)} = 0$$

Here, it is clear we must also set the condition that $p \neq 0$. As a result, we can cancel out $p(1-p)$.

$$k - np = 0$$

$$p = \frac{k}{n}$$

That is, when our probability of success equals the number of successes divided by the number of trials, we maximize the total probability. However, this is given that p is not 0, 1, and $k < n$.

3 Maximizing Number of Trials

Now, we will go ahead and optimize n like we did for p in the section above. Here, p and k are held constant, and n is what we are optimizing.

3.1 Model

In order to maximize n, we must show that $f(n) \geq f(n-1)$ and $f(n) \geq f(n+1)$. This is because we assume that the negative binomial has a maximum probability, and so there must be an n in which the probability is greater than the points next to it. To start off, we look at:

$$f(n) \geq f(n-1)$$

This can then be expanded into:

$$\frac{(n-1)!}{(k-1)!(n-k)!}p^k(1-p)^{n-k} \geq \frac{(n-2)!}{(k-1)!(n-k-1)!}p^k(1-p)^{n-k-1}$$

$$\frac{n-1}{n-k} \geq \frac{1}{1-p}$$

$$n \leq \left\lfloor 1 + \frac{k-1}{p} \right\rfloor$$

Next, let us look at:

$$f(n) \geq f(n+1)$$

This can be expanded into:

$$\frac{(n-1)!}{(k-1)!(n-k)!}p^k(1-p)^{n-k} \geq \frac{n!}{(k-1)!(n-k+1)!}p^k(1-p)^{n-k+1}$$

$$1 \geq \frac{n(1-p)}{n-k+1}$$

$$n \geq \left\lfloor \frac{k-1}{p} \right\rfloor$$

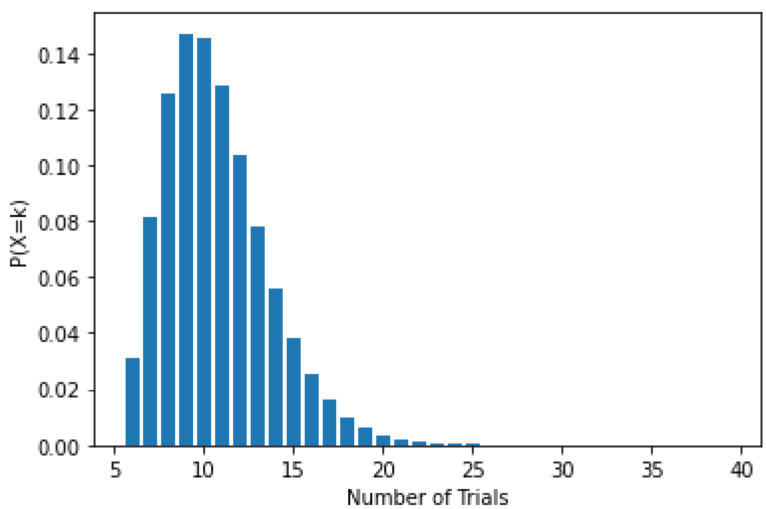
From this, we can compute two different n values and experiment to see which one will produce the greatest probability.

3.2 Empirical Data

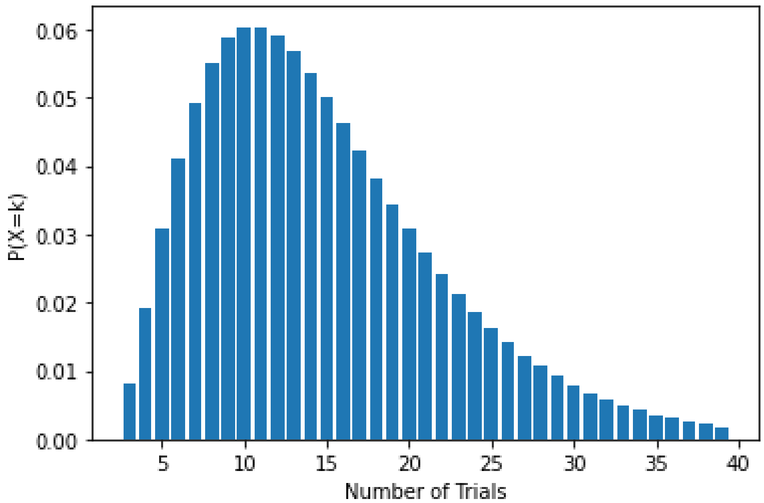
Below is the code used to create my graphs. The function tests multiple different values for n and then plots them.

```
def neg_bin(n, k, p):
    list = []
    for i in range(k, n):
        comb = math.factorial(i-1) / (math.factorial(k-1)*math.factorial(i-k))
        prob = comb*(p**k)*(1-p)**(i-k)
        list.append(prob)
    dic = {"n": [i for i in range(k, n)], "Probability": list}
    data = pd.DataFrame(data=dic)
    max_prob1 = data[data["Probability"] == max(data["Probability"])]
    max_prob = (max_prob1["n"]).tolist()[0]
    print("Number of trials that maximizes P(X=k): " + str(max_prob))
    plt.bar(data['n'], data["Probability"])
    plt.xlabel("Number of Trials")
    plt.ylabel("P(X=k)")
    plt.show()
```

Below is an example of a graph where $k = 6$ and $p = 0.56$. We try different n values ranging from 6 to 40, and find that the probability is maximized when the total number of trials is 9.



In another example, we try $k = 3$ and $p = 0.2$. Our n values range from 3 to 40. As can be seen below, the total probability is maximized when $n = 9$.



In both simulations, our answer does not deviate away from the formula we derived above.

4 Works Cited

Piegorsch, Walter W., " Maximum Likelihood Estimation for the Negative Binomial Dispersion Parameter". *Biometrics*, Sep. 1990, Vol. 46, No. 3, pp. 863-867. *JSTOR*, https://www.jstor.org/stable/2532104.