Predictive Probability for NB

Elad Zippory

September, 14, 2015

1 Negative Binomial

$$
\begin{aligned}
P(Y) &= \int P(Y|\lambda)P(\lambda)d\lambda \\
&= \int \frac{\lambda^y e^{-\lambda}}{y!}\frac{\alpha^\beta}{\Gamma(\beta)}\lambda^{\beta-1}e^{-\lambda\alpha}d\lambda \\
&= \frac{\alpha^\beta}{y!\Gamma(\beta)}\int \lambda^{y+\beta-1}e^{-\lambda(\alpha+1)}d\lambda
\end{aligned}
$$

Now, stare at the integral until you see the kernel of the Gamma and its parameters. The integral of the Gamma is the CDF which equals 1... So add the constant to make the kernel a proper CDF and get rid of the integral.

$$
= \frac{\alpha^\beta}{y!\Gamma(\beta)}\frac{\Gamma(y+\beta)}{(1+\alpha)^{y+\beta}} \quad = \frac{(y+\beta-1)!}{y!(\beta-1)!}\left(\frac{\alpha}{1+\alpha}\right)^\beta\left(\frac{1}{1+\alpha}\right)^y
$$

2 predictive probability

We need to derive the posterior distribution of a new predicted variable and show it is $\backsim$ NB. Remember that:

$$
\begin{aligned}
P(Y^*|y_1,..,y_n) &= \int P(Y^*,\theta|y_1,..y_n)d\theta \\
&= \int P(Y^*|\theta)P(\theta|y)d\theta
\end{aligned}
$$

What is $P(\lambda|y)$?

$$
\begin{aligned}
P(\lambda|y) &\propto P(y|\lambda)P(\lambda) \\
&= \frac{\lambda^{\sum y}e^{-n\lambda}}{\prod y_i!}\lambda^{\alpha-1}e^{-\lambda\beta}\frac{B^\alpha}{\Gamma(\alpha)} \\
&\propto \lambda^{\sum y+\alpha-1}e^{-\lambda(n+\beta)} \backsim \Gamma(\Sigma y+\alpha, n+\beta)
\end{aligned}
$$

Then the posterior of a new observation is:

$$
\begin{aligned}
P(Y^*|y) &= \int P(Y^*|\theta)P(\theta|y)d\theta \\
&= \int \frac{\lambda^{y^*}e^{-\lambda}}{y^*!} e^{-\lambda(n+\beta)} \lambda^{\sum y+\alpha-1} \frac{(n+\beta)^{\sum y+\alpha}}{\Gamma(\sum y+\alpha)} d\lambda \\
&= \int \lambda^{y^*+\sum y+\alpha-1} e^{-\lambda(n+\beta+1)} \frac{(n+\beta)^{\sum y+\alpha}}{y^*!\Gamma(\sum y+\alpha)} d\lambda
\end{aligned}
$$

For the sake of clarity, let's let $A = y^* + \sum y + \alpha$ and $B = n + \beta + 1$ for a couple of lines. And again we are going to multiply by 1... which is the twin brother of adding zero - The Two Pillars of Statistics.

$$
\begin{aligned}
&= \int \lambda^{A-1} e^{-\lambda B} \frac{B^A}{\Gamma(A)} \frac{\Gamma(A)}{B^A} \frac{(n+\beta)^{\sum y+\alpha}}{y^*!\Gamma(\sum y+\alpha)} d\lambda \\
&= \frac{\Gamma(A)}{B^A} \frac{(n+\beta)^{\sum y+\alpha}}{y^*!\Gamma(\sum y+\alpha)} \int \lambda^{A-1} e^{-\lambda B} \frac{B^A}{\Gamma(A)} d\lambda
\end{aligned}
$$

Again, give that integral a hard look and convince yourself it equals 1. Then let's get rid of A and B and re-arrange.

$$
= \frac{\Gamma(y^* + \sum y + \alpha)}{\Gamma(\sum y + \alpha)} \frac{(n+\beta)^{\sum y+\alpha}}{(n+\beta+1)^{y^*+\sum y+\alpha}}
$$

$$
\begin{aligned}
&= \binom{y^* + \sum y + \alpha - 1}{y^*} \left(\frac{n+\beta}{n+\beta+1}\right)^{\sum y+\alpha} \frac{1}{(n+\beta+1)^{y*}} \\
&\backsim NB\left(\Sigma y + \alpha, \frac{1}{n+\beta+1}\right) \square
\end{aligned}
$$

Verify that the new variance is larger than the NB of the observed data. Explain to yourself.