1 Gamma distribution as the conjugate prior for a Poisson likelihood

Here we show that the gamma distribution is the conjugate prior for a Poisson likelihood used in Bayes' theorem; that is, choosing a gamma-distributed prior also yields a gamma-distributed posterior if the likelihood is a Poisson distribution.

Assume a series of observations $\{n_1, n_2, \ldots, n_M\}$ are independent and identically distributed from a Poisson distribution $f_{Poisson(\lambda)}(n)$. This gives a Bayesian likelihood term

$$P(\{n_m\} \,|\, \lambda) = \prod_{m=1}^{M} P(n_m \,|\, \lambda) \tag{1}$$

$$= \prod_{m=1}^{M} \frac{e^{-\lambda}\lambda^{n_m}}{n_m!} \tag{2}$$

$$= \frac{e^{-M\lambda}\lambda^{\sum_{m=1}^{M} n_m}}{\prod_{m=1}^{M} n_m!}. \tag{3}$$

We choose the prior to be a gamma distribution over λ with shape and rate parameters α, β:

$$P(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}. \tag{4}$$

Bayes' theorem gives the posterior

$$P(\lambda \,|\, \{n_m\}) = \frac{P(\{n_m\} \,|\, \lambda)P(\lambda)}{\int_{\lambda'=0}^{\infty} P(\{n_m\} \,|\, \lambda')P(\lambda')\, d\lambda'} \tag{5}$$

$$= \frac{\frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\sum_{m=1}^{M} n_m+\alpha-1}e^{-(M+\beta)\lambda}}{\int_{\lambda'=0}^{\infty} \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda'^{\sum_{m=1}^{M} n_m+\alpha-1}e^{-(M+\beta)\lambda'}\, d\lambda'}. \tag{6}$$

The denominator of the above expression is an integral of the form

$$c\int x^a e^{bx}\, dx \tag{7}$$

where $c = \frac{\beta^\alpha}{\Gamma(\alpha)}$, $a = \sum_{m=1}^{M} n_m + \alpha - 1$, $b = -(M+\beta)$, and $x = \lambda'$. This integral can be treated by iteratively applying integration by parts:

$$\int x^a e^{bx}\,dx = \frac{1}{b}x^a e^{bx} - \frac{a}{b}x^{a-1}e^{bx}\,dx \tag{8}$$

$$= \frac{1}{b}x^a e^{bx} - \frac{a}{b}\left(\frac{1}{b}x^{a-1}e^{bx} - \frac{a-1}{b}\int x^{a-2}e^{bx}\,dx\right) \tag{9}$$

$$= \frac{1}{b}x^a e^{bx} - \frac{a}{b}\Big[\frac{1}{b}x^{a-1}e^{bx} - \frac{a-1}{b}\Big(\frac{1}{b}x^{a-2}e^{bx}\int x^{a-3}e^{bx}\,dx\Big)\Big] \tag{10}$$

$$\ldots.$$

Rewriting and combining terms, we see that the process of integration by parts proceeds indefinitely like so:

$$\int x^a e^{bx}\,dx = \frac{1}{b}x^a e^{bx} - \frac{a}{b^2}x^{a-1}e^{bx} + \frac{a(a-1)}{b^3}x^{a-2}e^{bx} - \frac{a(a-1)(a-2)}{b^4}x^{a-3}e^{bx} + \ldots$$
$$+ \frac{a(a-1)\ldots(a-a)}{b^{a+2}}e^{bx} + \ldots. \tag{11}$$

Note that the last term written above is in fact zero, and all terms after it vanish as well. So we can rewrite our integral as the sum

$$\int x^a e^{bx}\,dx = \sum_{n=0}^{a}(-1)^n \frac{a!}{b^{n+1}(a-n)!}x^{a-n}e^{bx}. \tag{12}$$

Substituting our original symbols for the placeholders a, b, c, and x into expression (7) and evaluating the integral on the limits 0 to ∞ gives

$$\frac{\beta^\alpha}{\Gamma(\alpha)}\int_{\lambda'=0}^{\infty} \lambda'^{\sum_{m=1}^{M} n_m+\alpha-1} e^{-(M+\beta)\lambda'}\,d\lambda'$$
$$= \frac{\beta^\alpha}{\Gamma(\alpha)}\sum_{n=0}^{N+\alpha-1}(-1)^n \frac{(N+\alpha-1)!}{\Big[-(M+\beta)\Big]^{n+1}(N+\alpha-1-n)!}\lambda'^{N+\alpha-1-n}e^{-(M+\beta)\lambda'}\Big|_{\lambda'=0}^{\infty}, \tag{13}$$

where we have let $N = \sum_{m=1}^{M} n_m$. To evaluate the right hand side of the above equation, we can first note that the lower bound of the limit vanishes at $\lambda' = 0$. We can compute the upper bound by noting that the final two terms after the fraction bar are of the form

$$x^a e^{-bx}, \tag{14}$$

where $a = N+\alpha-1-n$, $b = M+\beta$, and $x = \lambda'$. Its limit as $x \to \infty$ can be evaluated using the Taylor series expansion of the exponential function:

$$\lim_{x\to\infty} x^a e^{-bx} = \lim_{x\to\infty} \frac{x^a}{e^{bx}} \tag{15}$$

$$= \lim_{x\to\infty} \frac{x^a}{\sum_{n=0}^{\infty} \frac{b^n x^n}{n!}} \tag{16}$$

$$= \lim_{x\to\infty} \frac{x^a}{\frac{b^x x^x}{x!}} \tag{17}$$

$$= \lim_{x\to\infty} \frac{x^a x!}{(bx)^x}. \tag{18}$$

Since the x^x term in the denominator increases more rapidly than the x! term in the numerator,

$$\lim_{x\to\infty} x^a e^{-bx} = 0. \tag{19}$$

Thus it appears at first that the entire sum $\sum_{n=0}^{N+\alpha-1}$ in equation (13) collapses to zero! This is not the case, however: the single term in the sum where $N+\alpha-1-n=0$ leaves $\lambda'^{N+\alpha-1-n} e^{-(M+\beta)\lambda'}\Big|_{\lambda'=0}^{\infty}$ undefined. We see that at the lower bound, the λ' term is of the form

$$\lim_{\lambda'\to 0} \lambda'^0 = \lim_{\lambda'\to 0} 1 \tag{20}$$

$$= 1, \tag{21}$$

so

$$\lambda'^{N+\alpha-1-n} e^{-(M+\beta)\lambda'}\Big|_{\lambda'=0}^{\infty} = e^{-(M+\beta)\lambda'}\Big|_{\lambda'=0}^{\infty} \tag{22}$$

$$= -1. \tag{23}$$

We now see that the sum in the right hand side of equation (13) is nonzero for the single term where $n = N+\alpha-1$, and that this term is equal to

$$\frac{\beta^\alpha}{\Gamma(\alpha)} (-1)^{N+\alpha-1} \frac{(N+\alpha-1)!}{(-1)^{N+\alpha}(M+\beta)^{N+\alpha}} (-1). \tag{24}$$

Noting that all (-1) terms in the above expression cancel with one another, we find that both sides of equation (13), and thus also the denominator in equation (6), are equivalent to

$$\frac{\beta^\alpha}{(M+\beta)^{N+\alpha}} \frac{\Gamma(N+\alpha)}{\Gamma(\alpha)}. \tag{25}$$

Substituting this back into equation (6) gives

$$P(\lambda \,|\, \{n_m\}) = \frac{\frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{N+\alpha-1}e^{-(M+\beta)\lambda}}{\frac{\beta^\alpha}{\Gamma(\alpha)}\frac{\Gamma(N+\alpha)}{(M+\beta)^{N+\alpha}}} \tag{26}$$

$$= \frac{(M+\beta)^{N+\alpha}}{\Gamma(N+\alpha)}\lambda^{N+\alpha-1}e^{-(M+\beta)\lambda} \tag{27}$$

$$= f_{gamma(N+\alpha,M+\beta)}(\lambda). \tag{28}$$

Altogether, then, we see that for a series of sequential observations $\{n_1, n_2, \ldots, n_M\}$, Bayes' theorem

$$P(\lambda \,|\, \{n_m\}) = \frac{P(\{n_m\} \,|\, \lambda)P(\lambda)}{\int_{\lambda'=0}^{\infty} P(\{n_m\} \,|\, \lambda')P(\lambda')\, d\lambda'} \tag{29}$$

is equivalent to

$$f_{gamma(\sum_{m=1}^{M} n_m+\alpha,M+\beta)}(\lambda) = \frac{\Big[\prod_{m=1}^{M} f_{Poisson(\lambda)}(n_m)\Big] f_{gamma(\alpha,\beta)}(\lambda)}{\int_{\lambda'=0}^{\infty}\Big[\prod_{m=1}^{M} f_{Poisson(\lambda')}(n_m)\Big] f_{gamma(\alpha,\beta)}(\lambda')\, d\lambda'} \tag{30}$$

for a Poisson likelihood and a gamma-distributed prior. Note that this gives a posterior distribution of the same type as the prior distribution. Therefore, the conjugate prior to a Poisson likelihood is a gamma distribution.

2 Mode of a gamma distribution

Here we derive the expression for the mode of a gamma distribution as a function of its parameters.

A gamma distribution over λ with shape and rate parameters α, β is given by the probability density function

$$f_{gamma(\alpha,\beta)}(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}. \tag{31}$$

To find the maximum we take the partial derivative with respect to λ and set it equal to zero:

$$\frac{\partial f_{gamma(\alpha,\beta)}(\lambda)}{\partial\lambda} = \frac{\partial}{\partial\lambda}\Big(\frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}\Big) \tag{32}$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)}\Big[(\alpha-1)\lambda^{\alpha-2}e^{-\beta\lambda} - \beta\lambda^{\alpha-1}e^{-\beta\lambda}\Big] \tag{33}$$

$$= 0. \tag{34}$$

Since $\frac{\beta^\alpha}{\Gamma(\alpha)}$ is assumed to be nonzero, this relation is satisfied when

$$(\alpha-1)\lambda^{\alpha-2}e^{-\beta\lambda} - \beta\lambda^{\alpha-1}e^{-\beta\lambda} = 0; \tag{35}$$

that is, when

$$(\alpha-1)\lambda^{\alpha-2}e^{-\beta\lambda} = \beta\lambda^{\alpha-1}e^{-\beta\lambda} \tag{36}$$

$$(\alpha-1) = \beta\lambda \tag{37}$$

or when

$$\frac{\alpha-1}{\beta} = \lambda. \tag{38}$$

So we take $\hat{\lambda} = \frac{\alpha-1}{\beta}$ to be the mode of a gamma distribution over λ.

3 Negative binomial distribution as a gamma-Poisson mixture

Here we show that a continuous mixture of Poisson distributions where the mixing distribution is a gamma distribution yields a negative binomial distribution.

Consider the continuous mixture for positive λ of a Poisson distribution over n with mean rate λ and a gamma distribution over λ with shape and rate parameters α, β:

$$\int_{\lambda=0}^{\infty} f_{Poisson(\lambda)}(n) f_{gamma(\alpha,\beta)}(\lambda)\, d\lambda = \int_{\lambda=0}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}\, d\lambda \tag{39}$$

$$= \frac{\beta^\alpha}{n!\,\Gamma(\alpha)} \int_{\lambda=0}^{\infty} \lambda^{\alpha+n-1} e^{-(1+\beta)\lambda}\, d\lambda. \tag{40}$$

The right hand side of the above equation contains an integral of the form dealt with in equations (7-25). Using the same approach as before, we find that

$$\int_{\lambda=0}^{\infty} \lambda^{\alpha+n-1} e^{-(1+\beta)\lambda}\, d\lambda = (n+\alpha-1)!\,(1+\beta)^{-(\alpha+n)}. \tag{41}$$

Substituting this into the right hand side of equation (40) and performing further manipulations gives

$$\frac{\beta^\alpha}{n!\,\Gamma(\alpha)}(n+\alpha-1)!\,(1+\beta)^{-(\alpha+n)} = \frac{\Gamma(\alpha+n)}{n!\,\Gamma(\alpha)} \frac{\beta^{\alpha+n}}{(1+\beta)^{\alpha+n}\beta^n} \tag{42}$$

$$= \frac{\Gamma(\alpha+n)}{n!\,\Gamma(\alpha)} \Big(\frac{\beta}{1+\beta}\Big)^{\alpha+n} \Big(\frac{1}{\beta}\Big)^n. \tag{43}$$

Isolating and performing manipulations on the two rightmost terms of the above equation gives

$$\left(\frac{\beta}{1+\beta}\right)^{\alpha+n}\left(\frac{1}{\beta}\right)^{n} = \left(\frac{\beta}{1+\beta}\right)^{\alpha}\left(\frac{\beta}{1+\beta}\right)^{n}\left(\frac{1}{\beta}\right)^{n} \tag{44}$$

$$= \left(\frac{\beta}{1+\beta}\right)^{\alpha}\left(\frac{1}{1+\beta}\right)^{n} \tag{45}$$

$$= \left(\frac{1}{1+\beta}\right)^{n}\left(1-\frac{1}{1+\beta}\right)^{\alpha}, \tag{46}$$

once it is seen that $\frac{\beta}{1+\beta} = \frac{1+\beta-1}{1+\beta} = 1 - \frac{1}{1-\beta}$.

Substituting this back into the right hand side of equation (43) and reducing further, we have

$$\frac{\Gamma(\alpha+n)}{n!\,\Gamma(\alpha)}\left(\frac{1}{1+\beta}\right)^{n}\left(1-\frac{1}{1+\beta}\right)^{\alpha} = \frac{(n+\alpha-1)!}{n!\,(\alpha-1)!}\left(\frac{1}{1+\beta}\right)^{n}\left(1-\frac{1}{1+\beta}\right)^{\alpha} \tag{47}$$

$$= \binom{n+\alpha-1}{n}\left(\frac{1}{1+\beta}\right)^{n}\left(1-\frac{1}{1+\beta}\right)^{\alpha} \tag{48}$$

$$= f_{NB(\alpha,\frac{1}{1+\beta})}(n). \tag{49}$$

The result is a negative binomial distribution with parameters $\alpha, \frac{1}{1+\beta}$. So, altogether we can write

$$\int_{\lambda=0}^{\infty} f_{Poisson(\lambda)}(n) f_{gamma(\alpha,\beta)}(\lambda)\, d\lambda = f_{NB(\alpha,\frac{1}{1+\beta})}(n). \tag{50}$$

Therefore, a continuous mixture of Poisson distributions where the mixing distribution is a gamma distribution yields a negative binomial distribution.