

This note explains the calculation of the probability information inside the original TADA method for the alternative model (Model 1 or M_1).

Let X , μ and γ be the gene-level count, the mutation rate and the relative risk of one variant category at the tested gene, and let N be the number of trios.

- Gene-level de novo count (X) for model 1 (a risk gene)

$$X \sim \text{Poisson}(2N\mu\gamma)$$

$$P(X|\gamma) = \frac{e^{-2N\mu\gamma}(2N\mu\gamma)^x}{x!} \quad (1)$$

- Prior information for γ :

$$\gamma \sim \text{Gamma}(\bar{\gamma}\beta, \beta)$$

$$\text{Let } \alpha = \bar{\gamma}\beta, \text{ then } P(\gamma) = \frac{\beta^\alpha}{\Gamma(\alpha)} \gamma^{\alpha-1} e^{-\beta\gamma} \quad (2)$$

- We have $P(X|M_1) = \int P(X, \gamma) d\gamma = \int P(X|\gamma) P(\gamma) d\gamma$

Use (1) and (2)

$$P(X|M_1) = \int \frac{e^{-2N\mu\gamma}(2N\mu\gamma)^x}{x!} \frac{\beta^\alpha}{\Gamma(\alpha)} \gamma^{\alpha-1} e^{-\beta\gamma} d\gamma$$

$$P(X|M_1) = \frac{\beta^\alpha (2N\mu)^x}{\Gamma(\alpha) x!} \int \gamma^{x+\alpha-1} e^{-\gamma(2N\mu+\beta)} d\gamma$$

We add some parts to create a density function for a Gamma distribution inside the integral.

$$P(X|M_1) = \frac{\beta^\alpha (2N\mu)^x}{\Gamma(\alpha) x!} \frac{\Gamma(x+\alpha)}{(\beta+2N\mu)^{x+\alpha}} \left(\int \frac{(\beta+2N\mu)^{x+\alpha}}{\Gamma(x+\alpha)} \gamma^{x+\alpha-1} e^{-\gamma(\beta+2N\mu)} d\gamma \right)$$

The expression inside the bracket is equal to 1 (the integral of a probability density function); therefore, we have:

$$P(X|M_1) = \frac{\beta^\alpha (2N\mu)^x}{\Gamma(\alpha) x!} \frac{\Gamma(x+\alpha)}{(\beta+2N\mu)^{x+\alpha}} = \left(\frac{\beta}{\beta+2N\mu} \right)^\alpha \left(\frac{2N\mu}{\beta+2N\mu} \right)^x \frac{\Gamma(x+\alpha)}{\Gamma(\alpha) x!}$$

- In the R language (<https://stat.ethz.ch/R-manual/R-devel/library/stats/html/NegBinomial.html>), the negative binomial distribution with size = n and prob = p has density $p^n (1-p)^x \frac{\Gamma(x+n)}{\Gamma(n) x!}$

and the R function is `dnbinom(x, size, prob, mu, log = FALSE)`

We can use the `dnbinom` function for X in which $\frac{\beta}{\beta+2N\mu}$ is a main parameter:

$$\text{dnbinom}(x, \text{size} = \alpha, \text{prob} = \frac{\beta}{\beta+2N\mu}) \text{ OR } \text{dnbinom}(x, \text{size} = \bar{\gamma}\beta, \text{prob} = \frac{\beta}{\beta+2N\mu}).$$