Likelihood ratio test for samples from gamma-distributed populations

Brenton Horne

10th October 2021

Contents

Definitions

Let Y_{ij} denote the jth observation from the ith treatment group, where $i = 1, 2, 3, ..., m$ and $j = 1, 2, 3, ..., n_i$.

Let:

$$n = \sum_{i=1}^{m} n_i$$
$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{m}\sum_{j=1}^{n_i} Y_{ij}$$
$$\overline{Y}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij}.$$

Hypotheses

H_0: $Y_{ij} \sim \text{Gamma}(\alpha, \beta)$

H_A: $Y_{ij} \sim \text{Gamma}(\alpha_i, \beta_i)$, where $\alpha_i \neq \alpha_k$ or $\beta_i \neq \beta_k$ for at least one combination of i and k values.

Let us denote the parameter space under the null hypothesis as $\Omega_0 = \{(\alpha, \beta) :\ 0 < \alpha, \beta < \infty\}$ and the parameter space under the alternative hypothesis as

$\Omega_a = \{(\alpha_i, \beta_i) :\ 0 < \alpha_i, \beta_i < \infty,\ \alpha_i \neq \alpha_k \text{ or } \beta_i \neq \beta_k \text{ for at least one pair of } i \text{ and } k \text{ values}\}$. The unrestricted parameter space is $\Omega = \Omega_0 \cup \Omega_a = \{(\alpha_i, \beta_i) :\ 0 < \alpha_i, \beta_i < \infty\}$.

Derivation of the maximum likelihood under the null

$$\begin{aligned}
L(\Omega_0) &= \prod_{i=1}^{m}\prod_{j=1}^{n_i} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} Y_{ij}^{\alpha-1} \exp\left(-\frac{Y_{ij}}{\beta}\right) \\
&= (\Gamma(\alpha)\beta^{\alpha})^{-n} \left(\prod_{i=1}^{m}\prod_{j=1}^{n_i} Y_{ij}\right)^{\alpha-1} \exp\left(-\frac{n\overline{Y}}{\beta}\right) \\
&= \left(\Gamma(\alpha)\beta^{\alpha} \exp\left(\frac{\overline{Y}}{\beta}\right)\right)^{-n} \left(\prod_{i=1}^{m}\prod_{j=1}^{n_i} Y_{ij}\right)^{\alpha-1}.
\end{aligned} \tag{1}$$

Taking the natural logarithm yields:

$$\ln L(\Omega_0) = -n \ln\left(\Gamma(\alpha)\beta^{\alpha}\exp\left(\frac{\overline{Y}}{\beta}\right)\right) + (\alpha-1)\sum_{i=1}^{m}\sum_{j=1}^{n_i}\ln Y_{ij}.$$

Next we will take the partial derivative with respect to α and set it to zero:

$$\frac{\partial \ln L(\Omega_0)}{\partial\alpha}\Big|_{\alpha=\widehat{\alpha},\beta=\widehat{\beta}} = -n\left(\frac{\Gamma'(\widehat{\alpha})\widehat{\beta}^{\widehat{\alpha}}\exp\left(\frac{\overline{Y}}{\widehat{\beta}}\right) + \Gamma(\widehat{\alpha})(\ln\widehat{\beta})\widehat{\beta}^{\widehat{\alpha}}\exp\left(\frac{\overline{Y}}{\widehat{\beta}}\right)}{\Gamma(\widehat{\alpha})\widehat{\beta}^{\widehat{\alpha}}\exp\left(\frac{\overline{Y}}{\widehat{\beta}}\right)}\right) + \sum_{i=1}^{m}\sum_{j=1}^{n_i}\ln Y_{ij} = 0$$

$$\therefore\;\; 0 = -n(\psi^{(0)}(\widehat{\alpha}) + \ln\widehat{\beta}) + \sum_{i=1}^{m}\sum_{j=1}^{n_i}\ln Y_{ij}. \tag{2}$$

Where $\psi^{(0)}(\widehat{\alpha})$ is the digamma function. Next we will take the partial derivative with respect to β and set it to zero:

$$\frac{\partial \ln L(\Omega_0)}{\partial\beta}\Big|_{\alpha=\widehat{\alpha},\beta=\widehat{\beta}} = -n\left(\frac{\Gamma(\widehat{\alpha})\widehat{\beta}^{\widehat{\alpha}-1}\widehat{\alpha}\exp\left(\frac{\overline{Y}}{\widehat{\beta}}\right) - \Gamma(\widehat{\alpha})\widehat{\beta}^{\widehat{\alpha}}\exp\left(\frac{\overline{Y}}{\widehat{\beta}}\right)\frac{\overline{Y}}{\widehat{\beta}^2}}{\Gamma(\widehat{\alpha})\widehat{\beta}^{\widehat{\alpha}}\exp\left(\frac{\overline{Y}}{\widehat{\beta}}\right)}\right) = 0$$

$$\therefore\;\; 0 = -n\left(\frac{\widehat{\alpha}}{\widehat{\beta}} - \frac{\overline{Y}}{\widehat{\beta}^2}\right). \tag{3}$$

Multiplying both sides of Equation 3 by $-\frac{\widehat{\beta}^2}{n}$ yields:

$$0 = \widehat{\alpha}\widehat{\beta} - \overline{Y}$$

$$\therefore\;\; \widehat{\beta} = \frac{\overline{Y}}{\widehat{\alpha}}. \tag{4}$$

Re-writing Equation 2 using Equation 4 yields:

$$0 = -n\left(\psi^{(0)}(\widehat{\alpha}) + \ln\frac{\overline{Y}}{\widehat{\alpha}}\right) + \sum_{i=1}^{m}\sum_{j=1}^{n_i}\ln Y_{ij}. \tag{5}$$

This equation cannot be analytically solved, so $\widehat{\alpha}$ must be numerically approximated using Newton's method and $\widehat{\beta}$ must be estimated from our approximation of $\widehat{\alpha}$ using Equation 4.

Before we apply Newton's method to get an algorithm for estimating $\widehat{\alpha}$, let us use our MLEs to find our expression for maximum likelihood under the null. Substituting Equation 4 into Equation 1 yields:

$$
\begin{aligned}
L(\widehat{\Omega_0}) &= \left(\Gamma(\widehat{\alpha})\left(\frac{\overline{Y}}{\widehat{\alpha}}\right)^{\widehat{\alpha}}\exp(\widehat{\alpha})\right)^{-n}\left(\prod_{i=1}^{m}\prod_{j=1}^{n_i}Y_{ij}\right)^{\widehat{\alpha}-1} \\
&= \left(\Gamma(\widehat{\alpha})\left(\frac{\overline{Y}e}{\widehat{\alpha}}\right)^{\widehat{\alpha}}\right)^{-n}\prod_{i=1}^{m}\prod_{j=1}^{n_i}Y_{ij}^{\widehat{\alpha}-1}.
\end{aligned} \tag{6}
$$

Using Newton's method to obtain $\widehat{\alpha}$

To use Newton's method, we must find the Jacobian for our problem, which should in this case be a scalar as we have only one equation to solve (Equation 5). Labelling the right-hand side of Equation 5 as $f(\widehat{\alpha})$, and differentiating it with respect to $\widehat{\alpha}$ yields:

$$
\begin{aligned}
\frac{\partial f}{\partial \widehat{\alpha}} &= -n\left(\psi^{(1)}(\widehat{\alpha}) - \frac{1}{\widehat{\alpha}}\right) \\
&= n\left(\frac{1}{\widehat{\alpha}} - \psi^{(1)}(\widehat{\alpha})\right).
\end{aligned}
$$

Therefore we can refine our estimate of $\widehat{\alpha}$ from an initial guess $\widehat{\alpha}^{(0)}$ using:

$$
\begin{aligned}
\widehat{\alpha}^{(k+1)} &= \widehat{\alpha}^{(k)} - \frac{f(\widehat{\alpha}^{(k)})}{\left.\frac{\partial f}{\partial \widehat{\alpha}}\right|_{\widehat{\alpha}=\widehat{\alpha}^{(k)}}} \\
&= \widehat{\alpha}^{(k)} - \frac{-n\left(\psi^{(0)}(\widehat{\alpha}^{(k)}) + \ln\frac{\overline{Y}}{\widehat{\alpha}^{(k)}}\right) + \sum_{i=1}^{m}\sum_{j=1}^{n_i}\ln Y_{ij}}{n\left(\frac{1}{\widehat{\alpha}^{(k)}} - \psi^{(1)}(\widehat{\alpha}^{(k)})\right)}.
\end{aligned}
$$

Where $\psi^{(1)}(\widehat{\alpha}^{(k)})$ is the derivative of the digamma function, also known as the trigamma function.

Derivation of the unrestricted maximum likelihood

$$\begin{aligned} L(\Omega) &= \prod_{i=1}^{m}\prod_{j=1}^{n_i} \frac{1}{\Gamma(\alpha_i)\beta_i^{\alpha_i}} Y_{ij}^{\alpha_i-1}\exp\left(-\frac{Y_{ij}}{\beta_i}\right) \\ &= \left(\prod_{i=1}^{m} (\Gamma(\alpha_i)\beta_i^{\alpha_i})^{-n_i}\right)\left(\prod_{i=1}^{m}\prod_{j=1}^{n_i} Y_{ij}^{\alpha_i-1}\right)\exp\left(-\sum_{i=1}^{m}\frac{n_i\overline{Y}_i}{\beta_i}\right). \end{aligned} \tag{7}$$

Taking the natural logarithm yields:

$$\ln L(\Omega) = -\sum_{i=1}^{m} n_i \ln\left(\Gamma(\alpha_i)\beta_i^{\alpha_i}\right) + \sum_{i=1}^{m}(\alpha_i - 1)\sum_{j=1}^{n_i}\ln Y_{ij} - \sum_{i=1}^{m}\frac{n_i\overline{Y}_i}{\beta_i}.$$

Differentiating our log-likelihood with respect to α_k and setting the derivative to zero:

$$\begin{aligned} \frac{\partial \ln L(\Omega)}{\partial \alpha_k}\Big|_{\alpha_i=\widehat{\alpha}_i,\beta_i=\widehat{\beta}_i} &= -\sum_{i=1}^{m} n_i\left(\frac{\Gamma'(\widehat{\alpha}_i)\widehat{\beta}_i^{\widehat{\alpha}_i} + \Gamma(\widehat{\alpha}_i)(\ln\widehat{\beta}_i)\widehat{\beta}_i^{\widehat{\alpha}_i}}{\Gamma(\widehat{\alpha}_i)\widehat{\beta}_i^{\widehat{\alpha}_i}}\right)\delta_{ik} + \sum_{i=1}^{m}\delta_{ik}\sum_{j=1}^{n_i}\ln Y_{ij} = 0 \\ \therefore\ 0 &= -n_k(\psi^{(0)}(\widehat{\alpha}_k) + \ln\widehat{\beta}_k) + \sum_{j=1}^{n_k}\ln Y_{kj}. \end{aligned} \tag{8}$$

Where δ_{ik} is the Kronecker delta. Differentiating our log-likelihood with respect to β_k and setting the derivative to zero:

$$\begin{aligned} \frac{\partial \ln L(\Omega)}{\partial \beta_k}\Big|_{\alpha_i=\widehat{\alpha}_i,\beta_i=\widehat{\beta}_i} &= -\sum_{i=1}^{m} n_i\left(\frac{\Gamma(\widehat{\alpha}_i)\widehat{\beta}_i^{\widehat{\alpha}_i-1}\widehat{\alpha}_i}{\Gamma(\widehat{\alpha}_i)\widehat{\beta}_i^{\widehat{\alpha}_i}}\right)\delta_{ik} + \sum_{i=1}^{m}\frac{n_i\overline{Y}_i}{\widehat{\beta}_i^2}\delta_{ik} = 0 \\ \therefore\ 0 &= -\frac{n_k\widehat{\alpha}_k}{\widehat{\beta}_k} + \frac{n_k\overline{Y}_k}{\widehat{\beta}_k^2}. \end{aligned} \tag{9}$$

Multiplying Equation 9 by $\dfrac{\widehat{\beta}_k^2}{n_k}$ yields:

$$\begin{aligned} -\widehat{\alpha}_k\widehat{\beta}_k + \overline{Y}_k &= 0 \\ \widehat{\beta}_k &= \frac{\overline{Y}_k}{\widehat{\alpha}_k}. \end{aligned} \tag{10}$$

Substituting Equation 10 into Equation 8 yields:

$$0 = -n_k \left(\psi^{(0)}(\widehat{\alpha}_k) + \ln \left(\frac{\overline{Y}_k}{\widehat{\alpha}_k} \right) \right) + \sum_{j=1}^{n_k} \ln Y_{kj}. \tag{11}$$

Equation 11 has no closed-form solution and hence $\widehat{\alpha}_k$ must be numerically approximated using a technique like Newton's method. Before we obtain an iterative formula based on Newton's method to approximate $\widehat{\alpha}_k$, we will use Equation 10 to obtain the unrestricted maximum likelihood by substituting it into Equation 7:

$$\begin{aligned} L(\widehat{\Omega}) &= \left(\prod_{i=1}^{m} \left(\Gamma(\widehat{\alpha}_i) \left(\frac{\overline{Y}_i}{\widehat{\alpha}_i} \right)^{\widehat{\alpha}_i} \right)^{-n_i} \right) \left(\prod_{i=1}^{m} \prod_{j=1}^{n_i} Y_{ij}^{\widehat{\alpha}_i - 1} \right) \exp \left(-\sum_{i=1}^{m} \frac{n_i \overline{Y}_i}{\frac{\overline{Y}_i}{\widehat{\alpha}_i}} \right) \\ &= \left(\prod_{i=1}^{m} \left(\Gamma(\widehat{\alpha}_i) \left(\frac{\overline{Y}_i}{\widehat{\alpha}_i} \right)^{\widehat{\alpha}_i} \right)^{-n_i} \right) \left(\prod_{i=1}^{m} \prod_{j=1}^{n_i} Y_{ij}^{\widehat{\alpha}_i - 1} \right) \exp \left(-\sum_{i=1}^{m} n_i \widehat{\alpha}_i \right) \\ &= \left(\prod_{i=1}^{m} \left(\Gamma(\widehat{\alpha}_i) \left(\frac{\overline{Y}_i e}{\widehat{\alpha}_i} \right)^{\widehat{\alpha}_i} \right)^{-n_i} \right) \prod_{i=1}^{m} \prod_{j=1}^{n_i} Y_{ij}^{\widehat{\alpha}_i - 1}. \end{aligned} \tag{12}$$

Using Newton's method to obtain $\widehat{\alpha}_k$

Calling the right-hand side of Equation 11, $g_k(\{\widehat{\alpha}_j\})$, and taking its partial derivative with respect to $\widehat{\alpha}_i$ so as to create the Jacobian matrix:

$$\begin{aligned} \frac{\partial g_k(\{\widehat{\alpha}_j\})}{\partial \widehat{\alpha}_i} &= -n_k \left(\psi^{(1)}(\widehat{\alpha}_k) - \frac{1}{\widehat{\alpha}_k} \right) \delta_{ik} \\ &= n_k \delta_{ik} \left(\frac{1}{\widehat{\alpha}_k} - \psi^{(1)}(\widehat{\alpha}_k) \right). \end{aligned}$$

Which means our Jacobian $\mathrm{J} = \left(\frac{\partial g_k(\{\widehat{\alpha}_j\})}{\partial \widehat{\alpha}_i} \right)$ will be a diagonal matrix. Hence $\mathrm{J}^{-1} = \left(\left(\frac{\partial g_k(\{\widehat{\alpha}_j\})}{\partial \widehat{\alpha}_i} \right)^{-1} \right)$. If we let $\widehat{\boldsymbol{\alpha}} = (\widehat{\alpha}_i)$ and $\mathbf{F} = (g_k(\{\widehat{\alpha}_j\}))$, then:

$$\widehat{\boldsymbol{\alpha}}^{(l+1)} = \widehat{\boldsymbol{\alpha}}^{(l)} - \mathrm{J}^{-1}\mathbf{F}.$$

Likelihood ratio

Therefore the likelihood ratio is:

$$
\begin{aligned}
\lambda &= \frac{L(\widehat{\Omega_0})}{L(\widehat{\Omega})} \\
&= \frac{\left(\Gamma(\widehat{\alpha})\left(\frac{\overline{Y}e}{\widehat{\alpha}}\right)^{\widehat{\alpha}}\right)^{-n} \prod_{i=1}^{m}\prod_{j=1}^{n_i} Y_{ij}^{\widehat{\alpha}-1}}{\left(\prod_{i=1}^{m}\left(\Gamma(\widehat{\alpha}_i)\left(\frac{\overline{Y}_i e}{\widehat{\alpha}_i}\right)^{\widehat{\alpha}_i}\right)^{-n_i}\right)\prod_{i=1}^{m}\prod_{j=1}^{n_i} Y_{ij}^{\widehat{\alpha}_i-1}} \\
&= \left(\Gamma(\widehat{\alpha})\left(\frac{\overline{Y}}{\widehat{\alpha}}\right)^{\widehat{\alpha}}\right)^{-n}\left(\prod_{i=1}^{m}\left(\Gamma(\widehat{\alpha}_i)\left(\frac{\overline{Y}_i}{\widehat{\alpha}_i}\right)^{\widehat{\alpha}_i}\right)^{n_i}\right)\prod_{i=1}^{m}\prod_{j=1}^{n_i} Y_{ij}^{\widehat{\alpha}-\widehat{\alpha}_i}
\end{aligned}
$$

And we know that asymptotically under the null hypothesis $-2\ln\lambda \sim \chi^2_{2m-2}$, so we will test $-2\ln\lambda$ against the χ^2_{2m-2} distribution and determine how likely our result is under the null.