

# Caso Gamma revisto

## Relatório

Silvaneio Viera dos Santos Junior

2022-11-23

### Priori alternativa para o modelo Gamma

Neste relatório apresentaremos um *follow-up* das análises feitas sobre o GDLM k-paramétrico para o caso Gamma com parâmetro de forma e média desconhecidos. Durante a última reunião, observamos que a compatibilização de prioris estava gerando uma aproximação ruim e foi levantada a suspeita de que a isso estaria ocorrendo por causa de restrições feitas ao espaço paramétrico afim de obter soluções para o sistema de compatibilização de prioris.

Após a reunião, refletindo mais sobre o assunto, encontrei percebi que o problema estava na escolha da priori, especificamente, a escolha do modelo observacional  $X|\alpha, \beta \sim \mathcal{G}(\alpha, \beta)$  me levou a escolher uma parametrização ruim para a priori. Ao trabalhar com o modelo observacional  $X|\phi, \mu \sim \mathcal{G}\left(\phi, \frac{\phi}{\mu}\right)$  a priori “natural”, é bem mais tratável. Especificamente, com o modelo observacional  $X|\alpha, \beta \sim \mathcal{G}(\alpha, \beta)$ , a priori “natural” seria:

$$\pi(\alpha, \beta) \propto \exp \{n_0 \alpha \ln(\beta) - k_0 \ln(\Gamma(\alpha)) + \theta_0 \alpha - \tau_0 \beta\},$$

Já com o modelo observacional  $X|\phi, \mu \sim \mathcal{G}\left(\phi, \frac{\phi}{\mu}\right)$ , a priori “natural” seria:

$$\pi^*(\phi, \mu) \propto \exp \left\{ -n_0^* \phi \ln(\mu) + k_0^* (\phi \ln(\phi) - \ln(\Gamma(\phi))) + \theta_0 \phi - \tau_0 \frac{\phi}{\mu} \right\},$$

Nessa nova priori, temos que o vetor de estatística suficientes associado é  $H_{p^*} = \left( \phi, \frac{\phi}{\mu}, \phi \ln(\mu), \phi \ln(\phi) - \ln(\Gamma(\phi)) \right)'$ .

Na priori para  $\alpha$  e  $\beta$  temos  $n_0$  que controla a precisão associada a  $\beta$  e  $k_0$  que controla a precisão associada a  $\alpha$ , porém, como há uma forte associação entre  $\alpha$  e  $\beta$  (os dois parâmetros controlam a precisão), há também fortes restrições na escolha de  $n_0$  e  $k_0$ , sendo que, a única forma que encontramos de lidar com essa associação foi fixar  $n_0 = k_0$ .

Ao usar uma priori para  $\phi$  e  $\mu$ , contornamos esse problema, pois temos que a  $\phi$  e  $\mu$  não tem uma associação tão restritiva entre si, de modo que conseguimos escolher  $n_0^*$  e  $k_0^*$  de forma quase independente e ainda ter uma priori própria.

De forma análoga ao que fizemos antes, se  $X, Y$  tem densidade  $\pi^*$ , então diremos que  $X, Y \sim \Pi^*(n_0^*, k_0^*, \tau_0, \theta_0)$ . Se  $\phi, \mu \sim \Pi^*(n_0^*, k_0^*, \theta_0, \tau_0)$ , então:

$$\begin{aligned} f(\mu|\phi) &\propto \pi^*(\phi, \mu) \propto \exp \left\{ -n_0^* \phi \ln(\mu) + k_0^* (\phi \ln(\phi) - \ln(\Gamma(\phi))) + \theta_0 \phi - \tau_0 \frac{\phi}{\mu} \right\} \\ &\propto \exp \left\{ -n_0^* \phi \ln(\mu) - \tau_0 \frac{\phi}{\mu} \right\} \\ &= \mu^{-n_0^* \phi} e^{-\tau_0 \frac{\phi}{\mu}}, \end{aligned}$$

ou seja  $\mu|\phi \sim \mathcal{IG}(n_0^*\phi + 1, \phi\tau_0)$ , onde  $\mathcal{IG}$  representa a distribuição Inversa Gamma.

Usando a distribuição condicional de  $\mu$  podemos reescrever  $\mathbb{E}_{p^*}[H_{p^*}] = \mathbb{E}_{p^*}[\mathbb{E}_{p^*}[H_{p^*}|\phi]]$ , de onde obtemos:

$$\begin{aligned}\mathbb{E}_{p^*}\left[\frac{\phi}{\mu}\right] &= \mathbb{E}_{p^*}[\phi \mathbb{E}_{p^*}\left[\frac{1}{\mu}|\phi\right]] = \mathbb{E}_{p^*}\left[\phi \frac{n_0^*\phi + 1}{\phi\tau_0}\right] = \frac{n_0^*\mathbb{E}_{p^*}[\phi] + 1}{\tau_0} \\ \mathbb{E}_{p^*}[\phi \ln(\mu)] &= \mathbb{E}_{p^*}[\phi \mathbb{E}_{p^*}[\ln(\mu)|\phi]] = \mathbb{E}_{p^*}[\phi(-\psi(n_0\phi + 1) + \ln(\phi\tau_0))].\end{aligned}$$

Usando que  $\mathbb{E}_{p^*}[\phi] = \mathbb{E}_q[\phi]$  ( $\mathbb{E}_q[H_p]$  é suposto conhecido), temos que:

$$n_0^* = \frac{\mathbb{E}_q\left[\frac{\phi}{\mu}\right] \tau_0 - 1}{\mathbb{E}_q[\phi]}$$

Com as equações acima, conseguimos escrever  $\mathbb{E}_{p^*}[H_{p^*}]$  como valores esperados que dependem apenas da distribuição marginal de  $\phi$ , sendo que a distribuição marginal de  $\phi$  é tal que:

$$\begin{aligned}f(\phi) &\propto \int_0^{+\infty} \pi^*(\phi, \mu) d\mu \\ &\propto \int_0^{+\infty} \exp\left\{-n_0^*\phi \ln(\mu) + k_0^*(\phi \ln(\phi) - \ln(\Gamma(\phi))) + \theta_0\phi - \tau_0\frac{\phi}{\mu}\right\} d\mu \\ &= \exp\{k_0^*(\phi \ln(\phi) - \ln(\Gamma(\phi))) + \theta_0\phi\} \int_0^{+\infty} \exp\left\{-n_0^*\phi \ln(\mu) - \tau_0\frac{\phi}{\mu}\right\} d\mu \\ &= \exp\{k_0^*(\phi \ln(\phi) - \ln(\Gamma(\phi))) + \theta_0\phi\} \frac{\Gamma(n_0^*\phi + 1)}{\phi^{n_0^*\phi+1} \tau_0^{n_0^*\phi+1}} \\ &\propto \frac{\phi^{k_0^*\phi}}{\Gamma(\phi)^{k_0^*}} \frac{\Gamma(n_0^*\phi + 1)}{\phi^{n_0^*\phi+1} \tau_0^{n_0^*\phi+1}} \exp\{\theta_0\phi\} \\ &= \frac{\phi^{k_0^*\phi}}{\phi^{-k_0^*} \Gamma(\phi + 1)^{k_0^*}} \frac{\Gamma(n_0^*\phi + 1)}{\phi^{n_0^*\phi+1} \tau_0^{n_0^*\phi+1}} \exp\{\theta_0\phi\}.\end{aligned}$$

Novamente, a fórmula de Stirling nos diz que:

$$\Gamma(x + 1) \approx \sqrt{2\pi x} \left(\frac{x}{e}\right)^x.$$

Usando essa aproximação na densidade de  $\phi$  obtemos que:

$$\begin{aligned}f(\alpha) &\propto \frac{\phi^{k_0^*\phi}}{\phi^{-k_0^*} \Gamma(\phi + 1)^{k_0^*}} \frac{\Gamma(n_0^*\phi + 1)}{\phi^{n_0^*\phi+1} \tau_0^{n_0^*\phi+1}} \exp\{\theta_0\phi\} \\ &\approx \frac{\phi^{k_0^*\phi}}{\phi^{-k_0^*} (\sqrt{2\pi\phi} (\frac{\phi}{e})^\phi)^{k_0^*}} \frac{\sqrt{2\pi n_0^*\phi} (\frac{n_0^*\phi}{e})^{n_0^*\phi}}{\phi^{n_0^*\phi+1} \tau_0^{n_0^*\phi+1}} \exp\{\theta_0\phi\} \\ &\propto \frac{\phi^{k_0^*\phi}}{\phi^{-k_0^*} \phi^{\frac{k_0^*}{2}} \phi^{k_0^*\phi} e^{-k_0^*\phi}} \frac{n_0^{*n_0^*\phi} \phi^{\frac{1}{2}n_0^*\phi} e^{-n_0^*\phi}}{\phi^{n_0^*\phi+1} \tau_0^{n_0^*\phi}} \exp\{\theta_0\phi\} \\ &\propto \phi^{\frac{k_0^*+1}{2}-1} \exp\{-(n_0^* - k_0^* + n_0^* \ln\left(\frac{\tau_0}{n_0^*}\right) - \theta_0)\phi\}.\end{aligned}$$

Ou seja,  $\phi$  tem distribuição aproximada  $\mathcal{G}\left(\frac{k_0^*+1}{2}, n_0^* - k_0^* + n_0^* \ln\left(\frac{\tau_0}{n_0^*}\right) - \theta_0\right)$ . Vale enfatizar que, dessa vez **não** foi necessário fazer nenhum tipo de restrição sob o espaço paramétrico.

A aproximação acima é muito útil para destacar a condição na qual  $\Pi^*$  é própria: Para que  $\Pi^*$ , basta que  $n_0^* - k_0^* + n_0^* \ln\left(\frac{\tau_0}{n_0^*}\right) - \theta_0 > 0$ .

Usando a aproximação acima, podemos calcular obter que:

$$\mathbb{E}_{p^*}[\phi] = \frac{k_0^* + 1}{2(n_0^* - k_0^* + n_0^* \ln\left(\frac{\tau_0}{n_0^*}\right) - \theta_0)}$$

Usando que  $\psi(x) \approx \ln(x) - \frac{1}{2x} - \frac{1}{12x^2}$  ( $\psi$  é a função digamma) e que  $\psi(x+1) = \psi(x) + \frac{1}{x}$ , temos que:

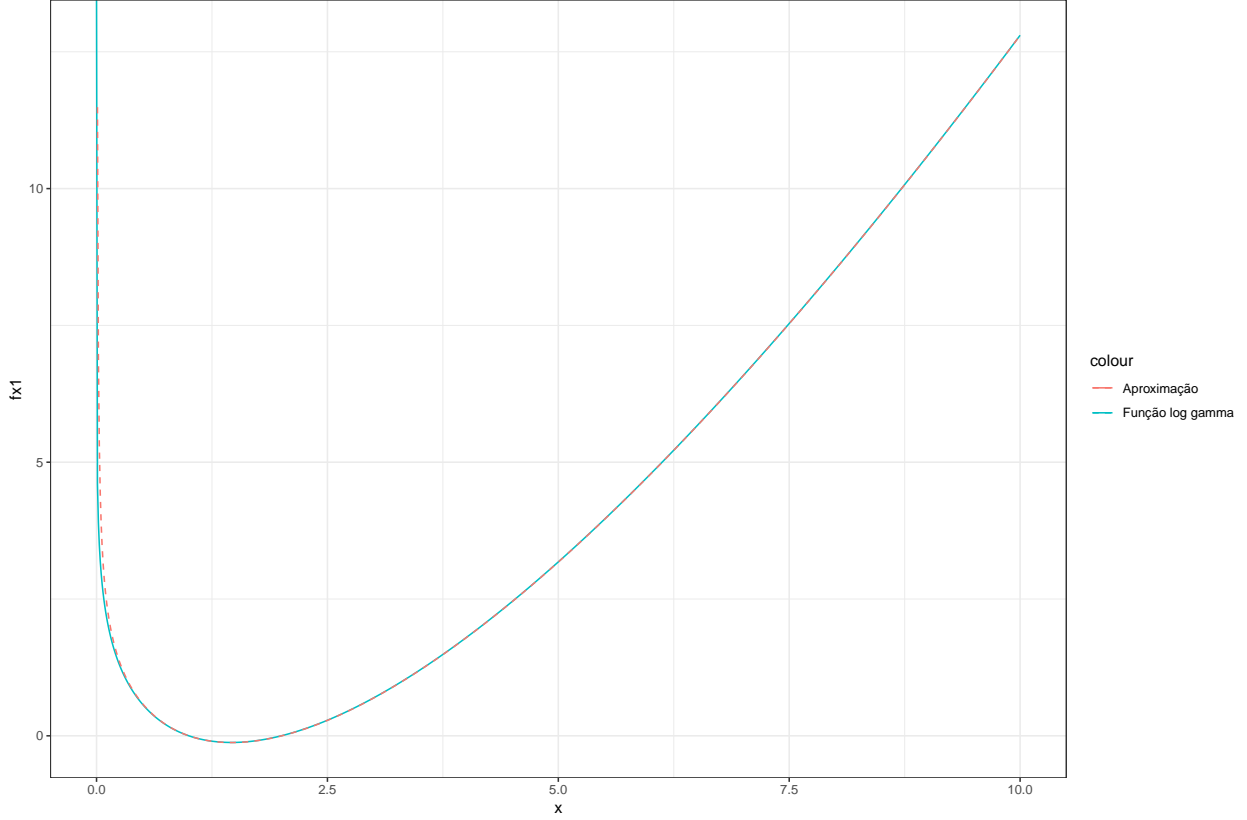
$$\begin{aligned} \mathbb{E}_{p^*}[-\phi(\psi(n_0^*\phi + 1) - \ln(\phi\tau_0))] &= \mathbb{E}_{p^*} \left[ -\phi \left( \psi(n_0^*\phi) + \frac{1}{n_0^*\phi} - \ln(\phi\tau_0) \right) \right] \\ &\approx \mathbb{E}_{p^*} \left[ -\phi \left( \ln(n_0^*\phi) - \frac{1}{2n_0^*\phi} - \frac{1}{12n_0^{*2}\phi^2} + \frac{1}{n_0^*\phi} - \ln(\phi\tau_0) \right) \right] \\ &= \mathbb{E}_{p^*} \left[ -\phi \left( -\ln\left(\frac{\tau_0}{n_0^*}\right) + \frac{1}{2n_0^*\phi} - \frac{1}{12n_0^{*2}\phi^2} \right) \right] \\ &= \frac{k_0^* + 1}{2(n_0^* - k_0^* + n_0^* \ln\left(\frac{\tau_0}{n_0^*}\right) - \theta_0)} \ln\left(\frac{\tau_0}{n_0^*}\right) - \frac{1}{2n_0^*} + \frac{1}{12n_0^{*2}} \frac{2(n_0^* - k_0^* + n_0^* \ln\left(\frac{\tau_0}{n_0^*}\right) - \theta_0)}{k_0^* - 1}. \end{aligned}$$

**A aproximação acima é usada apenas quando  $\frac{k_0^*+1}{2}$  é grande (maior que 3 já é o suficiente). Quando  $\frac{k_0^*+1}{2}$  é pequeno, esse valor esperado é calculado usando quadratura Gaussiana.**

Usando que  $\psi(x) \approx \ln(x) - \frac{1}{2x} - \frac{1}{12x^2}$  junto ao Teorema Fundamental do Cálculo, obtemos:

$$\ln(\Gamma(x)) = \int_1^x \psi(t)dt \approx \int_1^x \ln(t) - \frac{1}{2t} - \frac{1}{12t^2} dt = \left( t \ln(t) - t - \frac{1}{2} \ln(t) + \frac{1}{12t} \right)_1^x = x \ln(x) - x - \frac{1}{2} \ln(x) + \frac{1}{12x} + \frac{11}{12}$$

Podemos verificar que a qualidade da aproximação acima no gráfico a seguir:



Em geral a aproximação parece boa, porém há certas ressalvas sobre o seu uso, conforme discutirei mais à frente.

Calculemos então o valor esperado da última estatística suficiente restante:

$$\begin{aligned}
 \mathbb{E}_{p^*}[\phi \ln(\phi) - \ln(\Gamma(\phi))] &\approx \mathbb{E}_{p^*} \left[ \cancel{\phi \ln(\phi)} - \cancel{\phi \ln(\phi)} + \phi + \frac{1}{2} \ln(\phi) - \frac{1}{12\phi} - \frac{11}{12} \right] \\
 &= \frac{k_0^* + 1}{2(n_0^* - k_0^* + n_0^* \ln\left(\frac{\tau_0}{n_0^*}\right) - \theta_0)} + \frac{1}{2} \left( \psi\left(\frac{k_0^* + 1}{2}\right) - \ln\left(n_0^* - k_0^* + n_0^* \ln\left(\frac{\tau_0}{n_0^*}\right) - \theta_0\right) \right) \\
 &\quad - \frac{1}{12} \frac{2(n_0^* - k_0^* + n_0^* \ln\left(\frac{\tau_0}{n_0^*}\right) - \theta_0)}{k_0^* - 1} - \frac{11}{12}.
 \end{aligned}$$

**A aproximação acima é usada apenas quando  $\frac{k_0^* + 1}{2}$  é grande (maior que 3 já é o suficiente). Quando  $\frac{k_0^* + 1}{2}$  é pequeno, esse valor esperado é calculado usando quadratura Gaussiana.**

As expressões aproximadas para  $\mathbb{E}_{p^*}[-\phi(\psi(k_0^* \phi + 1) - \ln(\phi \tau_0))]$  e  $\mathbb{E}_{p^*}[\phi \ln(\phi) - \ln(\Gamma(\phi))]$  parecem bastante úteis, porém elas tem um defeito muito grave: Elas não são válidas para  $k_0^* \leq 1$ . De fato, se  $k_0^* \leq 1$ ,  $\mathbb{E}_{p^*}[1/\phi]$  não está definido, o que invalida as equações acima. Vale destacar que  $\mathbb{E}_{p^*}[-\phi(\psi(k_0^* \phi + 1) - \ln(\phi \tau_0))]$  e  $\mathbb{E}_{p^*}[\phi \ln(\phi) - \ln(\Gamma(\phi))]$  **existem** e são finitos, porém a aproximação da função  $\psi$  é ruim para valores baixos de  $k_0^*$ , especificamente, a aproximação de  $\psi$  é ruim para valores pequenos de  $\phi$  e, quando  $k_0^*$  é pequeno, a região onde a aproximação de  $\psi$  é ruim tem bastante massa de probabilidade, de modo que a aproximação do valor esperado é ruim.

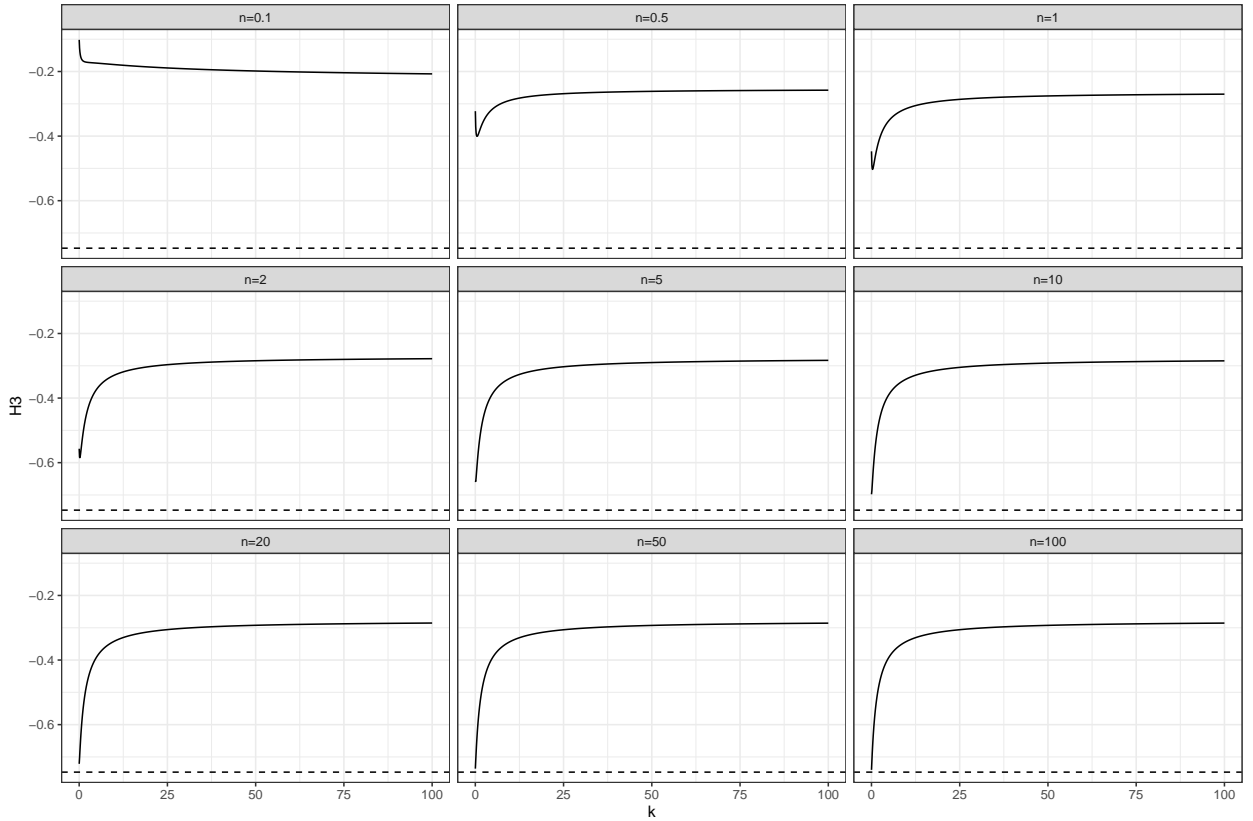
No caso em que trabalhamos no último relatório,  $\phi$  tinha distribuição aproximada  $\mathcal{G}\left(\frac{n_0 + 3}{2}, n_0 \ln\left(\frac{\tau_0}{n_0}\right) - \theta_0\right)$ , de modo que o parâmetro de forma era pelo menos 1.5, nesse caso, a aproximação de  $\psi$  é sempre boa, mas, no

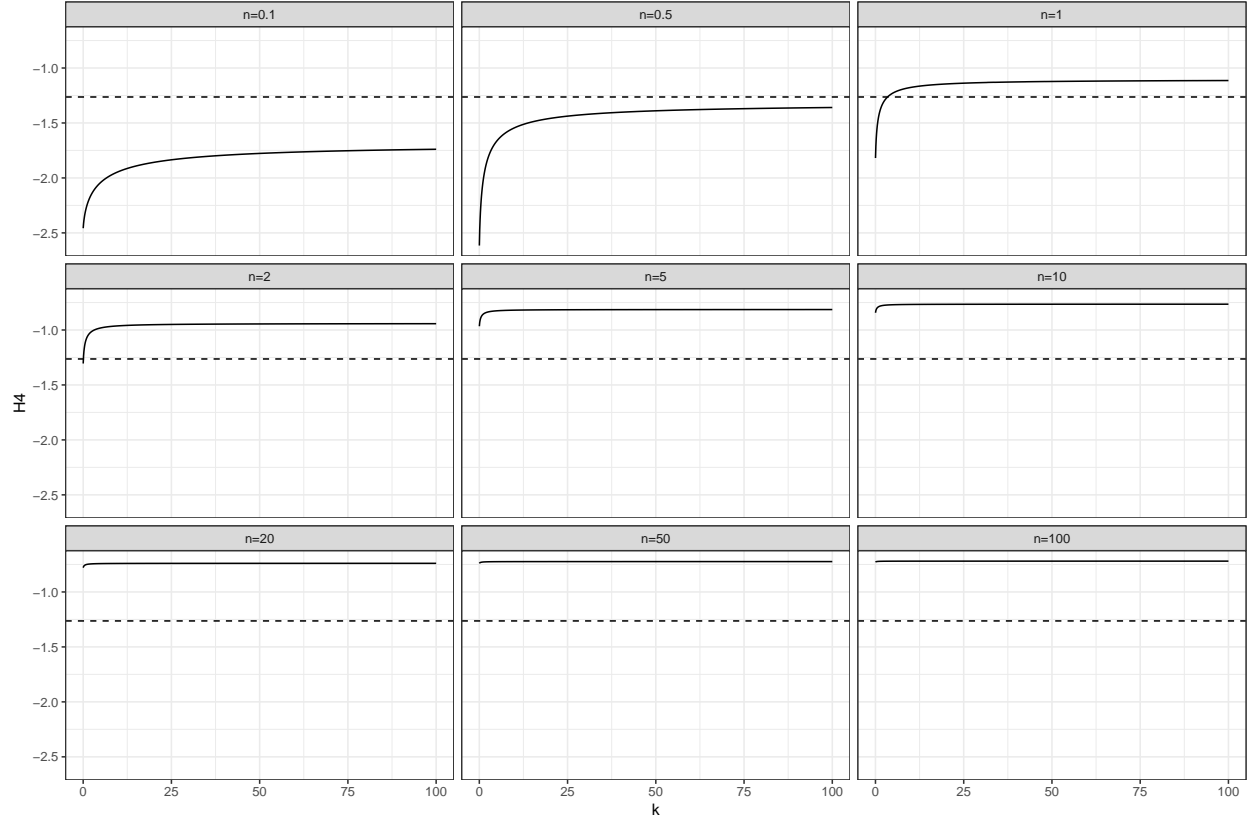
caso atual,  $\phi$  tem distribuição aproximada  $\mathcal{G}\left(\frac{k_0^*+1}{2}, n_0^* - k_0^* + n_0^* \ln\left(\frac{\tau_0}{n_0^*}\right) - \theta_0\right)$ , de modo que o parâmetro de forma é no mínimo 0.5, mas pode ser menor que 1.5, de modo que a aproximação de  $\psi$  pode ser ruim.

De todo modo, a observação acima não é necessariamente um empecilho, uma vez que podemos usar outros métodos para calcular  $\mathbb{E}_{p^*}[-\phi(\psi(k_0^*\phi + 1) - \ln(\phi\tau_0))]$  e  $\mathbb{E}_{p^*}[\phi \ln(\phi) - \ln(\Gamma(\phi))]$  sem aproximações.

Usando o novo sistema obtido, tentamos fazer o ajuste do modelo, porém não conseguimos um ajuste funcional. Atualmente o problema com que lidamos é que há casos onde o sistema  $\mathbb{E}_{p^*}[H_{p^*}] = \mathbb{E}_q[H_{p^*}]$  tem solução. Especificamente, há casos onde não existem parâmetros  $n_0^*$ ,  $k_0^*$ ,  $\tau_0$  e  $\theta_0$  tais que a equação  $\mathbb{E}_{p^*}[\phi \ln(\mu)] = \mathbb{E}_q[\phi \ln(\mu)]$  seja verdadeira.

A partir de tudo que foi discutido, conseguimos escrever  $\theta_0$  e  $\tau_0$  como funções de  $n_0^*$  e  $k_0^*$ , de modo que as duas primeiras equações do sistema sejam resolvidas (isto é, dado  $n_0^*$  e  $k_0^*$ , sabemos quem são os  $\theta_0$  e  $\tau_0$  que resolvem as duas primeiras equações), assim precisamos apenas buscar  $n_0^*$  e  $k_0^*$  que resolvam as duas últimas equações dado que  $\theta_0$  e  $\tau_0$  são tais que as duas primeiras equações estão resolvidas. Vamos analisar então  $\mathbb{E}_{p^*}[\phi \ln(\mu)]$  como funções de  $n_0^*$  e  $k_0^*$ :





Temos, no final das contas, que  $\mathbb{E}_{p^*}[\phi \ln(\mu)] < 0$ , para qualquer escolha de  $n_0^*$  ou  $k_0^*$ , por outro lado, temos que:

$$\mathbb{E}_q[\phi \ln(\mu)] = -(f_2 + q_{12}) \exp \left\{ f_1 + \frac{q_1}{2} \right\},$$

onde  $f_2$  é a média de  $\ln(\mu)$  na distribuição  $q$  (Normal bivariada),  $q_{12}$  é a covariância de  $\ln(\phi)$  e  $\ln(\mu)$ ,  $f_1$  é a média de  $\ln(\phi)$  e  $q_1$  é a variância de  $\ln(\phi)$ .

Veja que  $\mathbb{E}_q[\phi \ln(\mu)]$  pode ser positivo, basta que  $(f_2 + q_{12}) > 0$ , o que não pode ser evitado facilmente, de fato, mesmo inicializando o processo em um ponto onde esta condição é válida, pode acontecer de que esta condição seja violada eventualmente. Nesse casos, o sistema  $\mathbb{E}_{p^*}[H_{p^*}] = \mathbb{E}_q[H_{p^*}]$  **não** tem solução, pois sempre podemos diminuir a divergência entre  $p$  e  $q$  tomando  $k_0^*$  cada vez maior. Na prática, o que acontece ao longo do processo de resolução numérica do sistema é que o algoritmo de Newton-Raphson diverge, com  $k_0^*$  indo para infinito.

No momento estou tentando encontrar alguma solução para o problema acima.