A. Fisher information matrix for the Normal Distribution

Under regularity conditions (Wasserman, 2013), the Fisher information matrix can also be obtained from the second-order partial derivatives of the log-likelihood function

$$I(\theta) = -\mathbb{E}[\frac{\partial^2 l(\theta)}{\partial \theta^2}], \tag{D1}$$

where $l(\theta) = \log \pi_\theta(a|s)$. This gives us the Fisher information for the Normal distribution

$$\begin{aligned} \mathcal{I}(\mu, \sigma) &= -\mathbb{E}_{a\sim\pi_\theta} \begin{bmatrix} \frac{\partial^2 l}{\partial \mu^2} & \frac{\partial^2 l}{\partial \mu \partial \sigma} \\ \frac{\partial^2 l}{\partial \sigma \partial \mu} & \frac{\partial^2 l}{\partial \sigma^2} \end{bmatrix} \\ &= -\mathbb{E}_{a\sim\pi_\theta} \begin{bmatrix} -\frac{1}{\sigma^2} & -2\frac{(a-\mu)}{\sigma^3} \\ -2\frac{(a-\mu)}{\sigma^3} & \frac{-3(a-\mu)^2}{\sigma^4} + \frac{1}{\sigma^2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{bmatrix}. \end{aligned} \tag{D2}$$

B. Fisher information matrix for the Beta Distribution

To see how variance changes as the policy converges and becomes more deterministic, let us first compute the partial derivative of $\log \pi_\theta(a|s)$ with respect to shape parameter α

$$\begin{aligned} \frac{\partial \log \pi_\theta(a|s)}{\partial \alpha} &= \frac{\partial}{\partial \alpha} \log \left(\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} a^{\alpha-1}(1-a)^{\beta-1} \right) \\ &= \frac{\partial}{\partial \alpha}(\log \Gamma(\alpha+\beta) - \log \Gamma(\alpha) + (\alpha - 1)\log a) \\ &= \log a + \psi(\alpha+\beta) - \psi(\alpha), \end{aligned}$$

where $\psi(\cdot) = \frac{d}{dz}\log\Gamma(\cdot)$ is the *digamma* function. Similar results can also be derived for β. From (D1), we have

$$\begin{aligned} \mathcal{I}(\alpha, \beta) &= -\mathbb{E}_{a\sim\pi_\theta} \begin{bmatrix} \frac{\partial^2 l}{\partial \alpha^2} & \frac{\partial^2 l}{\partial \alpha \partial \beta} \\ \frac{\partial^2 l}{\partial \beta \partial \alpha} & \frac{\partial^2 l}{\partial \beta^2} \end{bmatrix} \\ &= -\mathbb{E}_{a\sim\pi_\theta} \begin{bmatrix} \frac{\partial}{\partial \alpha}(\psi(\alpha+\beta) - \psi(\alpha)) & \frac{\partial}{\partial \beta}(\psi(\alpha+\beta)) \\ \frac{\partial}{\partial \alpha}(\psi(\alpha+\beta)) & \frac{\partial}{\partial \beta}(\psi(\alpha+\beta) - \psi(\beta)) \end{bmatrix} \\ &= \begin{bmatrix} \psi'(\alpha) - \psi'(\alpha+\beta) & -\psi'(\alpha+\beta) \\ -\psi'(\alpha+\beta) & \psi'(\beta) - \psi'(\alpha+\beta) \end{bmatrix}, \end{aligned}$$

where $\psi'(z) = \psi^{(1)}(z)$ and $\psi^{(m)}(z) = \frac{d^{m+1}}{dz^{m+1}}\log\Gamma(z)$ is the *polygamma* function of order m. Figure 7 shows how $\psi'(z)$ goes to 0 as $z \to \infty$.

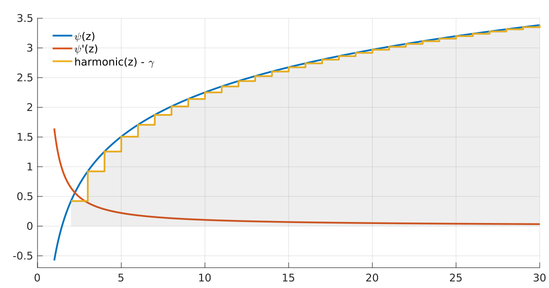

Figure 7. Graphs of the *digamma* function, the *polygamma* function, and the harmonic series. The harmonic series and the digamma function are related by $\sum_{k=1}^{z} \frac{1}{k} = \psi(z+1) + \gamma$, where $\gamma = 0.57721\cdots$ is the Euler-Mascheroni constant.