Complete Proofs

1 Highest Posterior Density (HPD) Intervals

We report the complete proofs of Theorems 1-3.

Theorem 1. Given a prior distribution $\text{Beta}(a, b)$ for μ and a Knowledge Graph (KG) correctness annotation process, where $0 < \tau_{\mathcal{S}} < n_{\mathcal{S}}$, the $1 - \alpha$ Highest Posterior Density (HPD) interval is the smallest (l, u) interval satisfying the condition $F(u) - F(l) = 1 - \alpha$.

Proof. When $0 < \tau_{\mathcal{S}} < n_{\mathcal{S}}$, the beta posterior $\text{Beta}(a + \tau_{\mathcal{S}}, b + n_{\mathcal{S}} - \tau_{\mathcal{S}})$ is unimodal and continuous over the $[0, 1]$ interval, guaranteeing the existence of a Credible Interval (CrI). The method of Lagrange multipliers allows us to find the smallest interval satisfying $F(u) - F(l) = 1 - \alpha$. That is, we minimize the following Lagrangian function:

$$\mathcal{L} = (u - l) + \lambda\Big(\int_l^u f(\mu \mid G_{\mathcal{S}})\, d\mu - (1 - \alpha)\Big) \tag{1}$$

where λ is the Lagrange multiplier.

The first-order partial derivatives of $\mathcal{L}$ in ∂l and ∂u lead to:

$$\frac{\partial \mathcal{L}}{\partial l} = -1 - \lambda f(l \mid G_{\mathcal{S}}) = 0 \Rightarrow f(l \mid G_{\mathcal{S}}) = -\frac{1}{\lambda};$$
$$\frac{\partial \mathcal{L}}{\partial u} = 1 + \lambda f(u \mid G_{\mathcal{S}}) = 0 \Rightarrow f(u \mid G_{\mathcal{S}}) = -\frac{1}{\lambda}.$$

Since $f(\cdot \mid G_{\mathcal{S}}) > 0$, it follows that $\lambda < 0$. To confirm that this yields a minimum, we compute the second-order partial derivatives:

$$\frac{\partial^2 \mathcal{L}}{(\partial l)^2} = -\lambda \frac{\partial f(l \mid G_{\mathcal{S}})}{\partial l};$$
$$\frac{\partial^2 \mathcal{L}}{(\partial u)^2} = \lambda \frac{\partial f(u \mid G_{\mathcal{S}})}{\partial u};$$
$$\frac{\partial^2 \mathcal{L}}{(\partial l)(\partial u)} = \frac{\partial^2 \mathcal{L}}{(\partial u)(\partial l)} = 0.$$

The posterior distribution is unimodal, hence $\partial f(l \mid G_{\mathcal{S}})/\partial l > 0$, leading to $\partial^2 \mathcal{L}/(\partial l)^2 > 0$. Similarly, $\partial f(u \mid G_{\mathcal{S}})/\partial u < 0$ implies $\partial^2 \mathcal{L}/(\partial u)^2 > 0$. Therefore, the Hessian matrix is positive definite, confirming that the interval minimizing the Lagrangian in Eq. (1) is the smallest satisfying $F(u) - F(l) = 1 - \alpha$. □

Theorem 2. Under the assumptions of Theorem 1, the $1 - \alpha$ HPD interval is unique.

Proof. Ab absurdo, let (l^*, u^*) represent the HPD interval, and assume there exists another interval (l', u') such that:

(1) $|u' - l'| = |u^* - l^*|$;
(2) $\int_{l'}^{u'} f(\mu \mid G_{\mathcal{S}})\, d\mu = 1 - \alpha$.

Denote the mode of the posterior distribution as ω. Since the posterior is unimodal, Th. 1 ensures that the HPD interval contains the mode, meaning $l^* < \omega < u^*$. Following the same reasoning, the alternative interval (l', u') must also satisfy $l' < \omega < u'$.

Given that $(l', u') \neq (l^*, u^*)$ but $|u' - l'| = |u^* - l^*|$, there are two possible cases:

(i) $l' < l^*$ and $u' < u^*$;

(ii) $l' > l^*$ and $u' > u^*$.

We prove case (i); the proof for case (ii) is analogous.

Since both intervals satisfy the condition $F(u) - F(l) = 1 - \alpha$

$$\int_{l'}^{u'} f(\mu \mid G_{\mathcal{S}})\, d\mu = \int_{l^*}^{u^*} f(\mu \mid G_{\mathcal{S}})\, d\mu$$

Under case (i), we can rewrite the left-hand side as:

$$\int_{l'}^{u'} f(\mu \mid G_{\mathcal{S}})\, d\mu = \int_{l'}^{l^*} f(\mu \mid G_{\mathcal{S}})\, d\mu + \int_{l^*}^{u'} f(\mu \mid G_{\mathcal{S}})\, d\mu$$

and the right-hand side as:

$$\int_{l^*}^{u^*} f(\mu \mid G_{\mathcal{S}})\, d\mu = \int_{l^*}^{u'} f(\mu \mid G_{\mathcal{S}})\, d\mu + \int_{u'}^{u^*} f(\mu \mid G_{\mathcal{S}})\, d\mu$$

Eliding the common term $\int_{l^*}^{u'} f(\mu \mid G_{\mathcal{S}})\, d\mu$ leads to:

$$\int_{l'}^{l^*} f(\mu \mid G_{\mathcal{S}})\, d\mu = \int_{u'}^{u^*} f(\mu \mid G_{\mathcal{S}})\, d\mu$$

that is a contradiction. Indeed, the posterior distribution is unimodal; the Probability Density Function (PDF) $f(\mu \mid G_{\mathcal{S}})$ is increasing before the mode and decreasing after it. Therefore, $\int_{l'}^{l^*} f(\mu \mid G_{\mathcal{S}})\, d\mu < \int_{u'}^{u^*} f(\mu \mid G_{\mathcal{S}})\, d\mu$, as l' lies outside the HPD region while u' lies inside. Despite l' and u' are equidistant from l^* and u^*, respectively, the densities of (l', l^*) and $(u', u*)$ differ due to their positions relative to the HPD region. Therefore, no other interval can have the same properties as the HPD interval, proving its uniqueness. □

Theorem 3. When the beta posterior for μ is unimodal and symmetric, the $1 - \alpha$ HPD interval is equivalent to the $1 - \alpha$ Equal-Tailed (ET) interval.

Proof. Consider a unimodal and symmetric beta posterior. Let ω represent the mode of this posterior. By the symmetry property, for any $\delta > 0$, we have $f(\omega - \delta \mid G_{\mathcal{S}}) = f(\omega + \delta \mid G_{\mathcal{S}})$. Therefore, the $1 - \alpha$ HPD interval is symmetric around ω – i.e., the lower and upper bounds, l and u, are equidistant from ω.

Let us rewrite $\int_{l}^{u} f(\mu \mid G_{\mathcal{S}})\, d\mu = 1 - \alpha$ as $\int_{l}^{\omega} f(\mu \mid G_{\mathcal{S}})\, d\mu + \int_{\omega}^{u} f(\mu \mid G_{\mathcal{S}})\, d\mu = 1 - \alpha$. Due to the symmetry of the posterior:

$$\int_{l}^{\omega} f(\mu \mid G_{\mathcal{S}})\, d\mu = \int_{\omega}^{u} f(\mu \mid G_{\mathcal{S}})\, d\mu = \frac{1 - \alpha}{2}$$

By construction, it follows that:

$$\int_{0}^{l} f(\mu \mid G_{\mathcal{S}})\, d\mu = \int_{u}^{1} f(\mu \mid G_{\mathcal{S}})\, d\mu = \frac{\alpha}{2}$$

corresponding to the Cumulative Distribution Function (CDF) of the $\alpha/2$ and $1 - \alpha/2$ quantiles, respectively. Thus, the $1 - \alpha$ HPD interval coincides with the $1 - \alpha$ ET interval. □