

HPD set for the posterior $\pi(\theta | y) = 2\theta$ on $(0, 1)$

Posterior model

Consider a one-dimensional parameter Θ with posterior density

$$\pi(\theta | y) = \begin{cases} 2\theta, & 0 < \theta < 1, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

We first check that this is a proper density:

$$\int_0^1 2\theta d\theta = [\theta^2]_0^1 = 1. \quad (2)$$

We want the $(1 - \alpha)$ highest posterior density (HPD) set $H_{1-\alpha}$ using the *level set* definition:

$$H_{1-\alpha} = \{\theta : \pi(\theta | y) \geq k_\alpha\}, \quad (3)$$

where the threshold k_α is chosen such that

$$\int_{H_{1-\alpha}} \pi(\theta | y) d\theta = 1 - \alpha. \quad (4)$$

Level sets of the posterior

For a fixed level $k > 0$, the level set of $\pi(\theta | y)$ is

$$L(k) := \{\theta \in (0, 1) : \pi(\theta | y) \geq k\} = \{\theta \in (0, 1) : 2\theta \geq k\}. \quad (5)$$

Solving the inequality $2\theta \geq k$ gives

$$\theta \geq \frac{k}{2}. \quad (6)$$

Intersecting with the support $0 < \theta < 1$ yields, for $0 < k \leq 2$,

$$L(k) = \left[\frac{k}{2}, 1 \right]. \quad (7)$$

(For $k > 2$ the set is empty; for $k \leq 0$ the set is $(0, 1)$, but the interesting case is $0 < k \leq 2$.)

Thus, for $0 < k \leq 2$,

$$L(k) = \left[\frac{k}{2}, 1 \right]. \quad (8)$$

Posterior probability of a level set

Define the posterior probability contained in the level set $L(k)$ by

$$\varphi(k) := P(\Theta \in L(k) \mid y) = \int_{L(k)} \pi(\theta \mid y) d\theta. \quad (9)$$

For $0 < k \leq 2$, using $L(k) = [k/2, 1]$, we obtain

$$\varphi(k) = \int_{k/2}^1 2\theta d\theta \quad (10)$$

$$= [\theta^2]_{\theta=k/2}^{\theta=1} = 1 - \left(\frac{k}{2}\right)^2. \quad (11)$$

Thus

$$\varphi(k) = 1 - \left(\frac{k}{2}\right)^2, \quad 0 < k \leq 2. \quad (12)$$

Determining the HPD threshold k_α

By definition, the $(1 - \alpha)$ HPD set corresponds to the level k_α such that

$$\varphi(k_\alpha) = 1 - \alpha. \quad (13)$$

Using the explicit expression for $\varphi(k)$,

$$1 - \left(\frac{k_\alpha}{2}\right)^2 = 1 - \alpha, \quad (14)$$

which implies

$$\left(\frac{k_\alpha}{2}\right)^2 = \alpha \implies k_\alpha = 2\sqrt{\alpha}, \quad (15)$$

where we take the positive root since $k_\alpha \geq 0$.

HPD region

The HPD region $H_{1-\alpha}$ is the level set at height k_α :

$$H_{1-\alpha} = L(k_\alpha) = \left[\frac{k_\alpha}{2}, 1\right] = \left[\frac{2\sqrt{\alpha}}{2}, 1\right] = [\sqrt{\alpha}, 1]. \quad (16)$$

Therefore, the $(1 - \alpha)$ HPD credible set for this posterior is

$$H_{1-\alpha} = [\sqrt{\alpha}, 1]. \quad (17)$$

Verification of the posterior mass

We can check directly that this set has posterior probability $1 - \alpha$:

$$P(\Theta \in H_{1-\alpha} \mid y) = \int_{\sqrt{\alpha}}^1 2\theta d\theta \quad (18)$$

$$= [\theta^2]_{\theta=\sqrt{\alpha}}^{\theta=1} \quad (19)$$

$$= 1 - (\sqrt{\alpha})^2 = 1 - \alpha. \quad (20)$$

Comparison with the equal-tailed credible interval

The CDF associated with $\pi(\theta \mid y) = 2\theta$ on $(0, 1)$ is

$$F(\theta) = P(\Theta \leq \theta \mid y) = \int_0^\theta 2t dt = \theta^2. \quad (21)$$

A $(1 - \alpha)$ equal-tailed credible interval (a, b) satisfies

$$F(a) = \frac{\alpha}{2}, \quad F(b) = 1 - \frac{\alpha}{2}. \quad (22)$$

Thus

$$a^2 = \frac{\alpha}{2} \implies a = \sqrt{\frac{\alpha}{2}}, \quad (23)$$

and

$$b^2 = 1 - \frac{\alpha}{2} \implies b = \sqrt{1 - \frac{\alpha}{2}}. \quad (24)$$

Hence the equal-tailed $(1 - \alpha)$ credible interval is

$$(a, b) = \left(\sqrt{\frac{\alpha}{2}}, \sqrt{1 - \frac{\alpha}{2}} \right), \quad (25)$$

which is clearly different from the HPD region

$$H_{1-\alpha} = [\sqrt{\alpha}, 1]. \quad (26)$$

$$H_{1-\alpha} = [\sqrt{\alpha}, 1] \quad \text{for} \quad \pi(\theta \mid y) = 2\theta, 0 < \theta < 1.$$

Highest Posterior Density (HPD) Sets: General Definition

Setup

Let (Θ, \mathcal{B}) be a parameter space (for instance $\Theta \subseteq \mathbb{R}^d$) with posterior density

$$\pi(\theta | y), \quad \theta \in \Theta, \quad (27)$$

with respect to some reference measure μ (typically Lebesgue measure), so that

$$\int_{\Theta} \pi(\theta | y) d\mu(\theta) = 1. \quad (28)$$

We fix a credibility level $1 - \alpha$ with $0 < \alpha < 1$ (e.g. $\alpha = 0.05$ for a 95% credible level). A *credible set* $C \subseteq \Theta$ of level $1 - \alpha$ satisfies

$$P(\Theta \in C | y) = \int_C \pi(\theta | y) d\mu(\theta) = 1 - \alpha. \quad (29)$$

Among all such sets, the *highest posterior density (HPD) set* is the one that contains the $(1 - \alpha)$ most probable parameter values, in the sense of posterior density.

Level sets of the posterior density

For each *density level* $k \geq 0$ we define the *level set* of $\pi(\theta | y)$ at height k by

$$L(k) := \{\theta \in \Theta : \pi(\theta | y) \geq k\}. \quad (30)$$

Intuitively, $L(k)$ is the subset of parameter values for which the posterior density is at least k . As k decreases, the set $L(k)$ becomes larger.

We then define the posterior *probability content* of the level set $L(k)$ as

$$\varphi(k) := P(\Theta \in L(k) | y) = \int_{L(k)} \pi(\theta | y) d\mu(\theta). \quad (31)$$

Since lowering k can only enlarge $L(k)$, the function $\varphi(k)$ is nonincreasing in k :

$$k_1 > k_2 \implies L(k_1) \subseteq L(k_2) \implies \varphi(k_1) \leq \varphi(k_2). \quad (32)$$

Moreover, as $k \rightarrow 0$ we recover the whole parameter space, so $\varphi(k) \rightarrow 1$, while as k increases towards the maximum of $\pi(\theta | y)$ we have $\varphi(k) \rightarrow 0$.

Definition of the $(1 - \alpha)$ HPD set

The $(1 - \alpha)$ HPD set $H_{1-\alpha}$ is defined via a *density threshold* k_α such that the level set at height k_α contains exactly $(1 - \alpha)$ posterior probability:

$$\varphi(k_\alpha) = \int_{L(k_\alpha)} \pi(\theta | y) d\mu(\theta) = 1 - \alpha. \quad (33)$$

Whenever φ is continuous and strictly decreasing in k over the relevant range, the value k_α is uniquely determined by this equation.

The $(1 - \alpha)$ HPD set is then

$$H_{1-\alpha} = L(k_\alpha) = \{\theta \in \Theta : \pi(\theta | y) \geq k_\alpha\}. \quad (34)$$

By construction:

1. $H_{1-\alpha}$ has posterior probability

$$P(\Theta \in H_{1-\alpha} | y) = \int_{H_{1-\alpha}} \pi(\theta | y) d\mu(\theta) = 1 - \alpha. \quad (35)$$

2. For any $\theta_1 \in H_{1-\alpha}$ and $\theta_2 \notin H_{1-\alpha}$, we have

$$\pi(\theta_1 | y) \geq k_\alpha > \pi(\theta_2 | y), \quad (36)$$

i.e. every point inside $H_{1-\alpha}$ has posterior density at least as large as any point outside.

This expresses formally the idea that $H_{1-\alpha}$ is the set of “most probable” parameter values with total probability $1 - \alpha$.

Univariate case and HPD intervals

In the one-dimensional case ($\Theta \subseteq \mathbb{R}$), $H_{1-\alpha}$ is called an *HPD interval* when it is connected. If the posterior density $\pi(\theta | y)$ is strictly unimodal and sufficiently regular, then for each k the level set $L(k)$ is typically an interval

$$L(k) = [a(k), b(k)], \quad (37)$$

and the HPD set $H_{1-\alpha}$ is simply the shortest interval centered around the mode that has posterior probability $1 - \alpha$.

If the posterior density is *multimodal*, the level set $L(k)$, and hence the HPD set $H_{1-\alpha}$, can be a union of disjoint intervals:

$$H_{1-\alpha} = \bigcup_{j=1}^J I_j, \quad I_j \subset \Theta, \quad (38)$$

each I_j containing one or more modes, with the property that the union has total probability $1 - \alpha$ and consists only of points whose density is at least k_α .

Relation to equal-tailed credible intervals

Let $F(\theta | y)$ denote the posterior cumulative distribution function (CDF),

$$F(\theta | y) := P(\Theta \leq \theta | y) = \int_{(-\infty, \theta]} \pi(t | y) dt \quad (\text{univariate case}). \quad (39)$$

A $(1 - \alpha)$ *equal-tailed* credible interval (a, b) is defined by the quantile conditions

$$F(a | y) = \frac{\alpha}{2}, \quad F(b | y) = 1 - \frac{\alpha}{2}. \quad (40)$$

Both (a, b) and $H_{1-\alpha}$ satisfy

$$P(a < \Theta < b | y) = 1 - \alpha, \quad P(\Theta \in H_{1-\alpha} | y) = 1 - \alpha, \quad (41)$$

but they need not coincide. They do coincide in many common symmetric and unimodal cases (e.g. Normal posteriors), but in general:

- Equal-tailed intervals are determined by *tail probabilities*;
- HPD sets are determined by *contours of constant density*.

Summary

Given a posterior density $\pi(\theta | y)$ and level $1 - \alpha$, the HPD set is defined as

$$H_{1-\alpha} = \{\theta \in \Theta : \pi(\theta | y) \geq k_\alpha\}, \quad (42)$$

where k_α is chosen such that

$$\int_{\{\theta : \pi(\theta | y) \geq k_\alpha\}} \pi(\theta | y) d\mu(\theta) = 1 - \alpha. \quad (43)$$

This set has posterior probability $1 - \alpha$ and contains only points of highest posterior density relative to the rest of the parameter space.