MT4531/MT5731: (Advanced) Bayesian Inference

Summaries of a posterior distribution

Nicolò Margaritella

School of Mathematics and Statistics, University of St Andrews

–

University of St Andrews

Outline

Outline

All information is within $\pi(\theta|\boldsymbol{x})$

- Assume a quantity of interest, for example some parameter θ.
- Once you have calculated the posterior distribution $\pi(\theta|\boldsymbol{x})$, you have a complete description of your beliefs and uncertainty about θ.
- Any statement on θ can be made, as long as you can integrate the posterior distribution, analytically, numerically, or with the help of some statistical software such as R.
- For example, your expectation for the true value of θ is,

$$\int_{\Theta} \theta \pi(\theta|\boldsymbol{x}) d\theta$$

- Or, the probability that θ is greater than 10, is,

$$\int_{10}^{\infty} \pi(\theta|\boldsymbol{x}) d\theta$$

- It is standard practice to report certain summaries of the posterior distribution that describe one's beliefs and uncertainty about θ.

All information is within $\pi(\theta|\boldsymbol{x})$

- Assume a quantity of interest, for example some parameter θ.
- Once you have calculated the posterior distribution $\pi(\theta|\boldsymbol{x})$, you have a complete description of your beliefs and uncertainty about θ.
- Any statement on θ can be made, as long as you can integrate the posterior distribution, analytically, numerically, or with the help of some statistical software such as R.
- For example, your expectation for the true value of θ is,

$$\int_{\Theta} \theta \pi(\theta|\boldsymbol{x}) d\theta$$

- Or, the probability that θ is greater than 10, is,

$$\int_{10}^{\infty} \pi(\theta|\boldsymbol{x}) d\theta$$

- It is standard practice to report certain summaries of the posterior distribution that describe one's beliefs and uncertainty about θ.

All information is within $\pi(\theta|\boldsymbol{x})$

- Assume a quantity of interest, for example some parameter θ.
- Once you have calculated the posterior distribution $\pi(\theta|\boldsymbol{x})$, you have a complete description of your beliefs and uncertainty about θ.
- Any statement on θ can be made, as long as you can integrate the posterior distribution, analytically, numerically, or with the help of some statistical software such as R.
- For example, your expectation for the true value of θ is,

$$\int_{\Theta} \theta \pi(\theta|\boldsymbol{x}) d\theta$$

- Or, the probability that θ is greater than 10, is,

$$\int_{10}^{\infty} \pi(\theta|\boldsymbol{x}) d\theta$$

- It is standard practice to report certain summaries of the posterior distribution that describe one's beliefs and uncertainty about θ.

All information is within $\pi(\theta|\boldsymbol{x})$

- Assume a quantity of interest, for example some parameter θ.
- Once you have calculated the posterior distribution $\pi(\theta|\boldsymbol{x})$, you have a complete description of your beliefs and uncertainty about θ.
- Any statement on θ can be made, as long as you can integrate the posterior distribution, analytically, numerically, or with the help of some statistical software such as R.
- For example, your expectation for the true value of θ is,

$$\int_{\Theta} \theta \pi(\theta|\boldsymbol{x}) d\theta$$

- Or, the probability that θ is greater than 10, is,

$$\int_{10}^{\infty} \pi(\theta|\boldsymbol{x}) d\theta$$

- It is standard practice to report certain summaries of the posterior distribution that describe one's beliefs and uncertainty about θ.

All information is within $\pi(\theta|\boldsymbol{x})$

- Assume a quantity of interest, for example some parameter θ.
- Once you have calculated the posterior distribution $\pi(\theta|\boldsymbol{x})$, you have a complete description of your beliefs and uncertainty about θ.
- Any statement on θ can be made, as long as you can integrate the posterior distribution, analytically, numerically, or with the help of some statistical software such as R.
- For example, your expectation for the true value of θ is,

$$\int_{\Theta} \theta \pi(\theta|\boldsymbol{x}) d\theta$$

- Or, the probability that θ is greater than 10, is,

$$\int_{10}^{\infty} \pi(\theta|\boldsymbol{x}) d\theta$$

- It is standard practice to report certain summaries of the posterior distribution that describe one's beliefs and uncertainty about θ.

All information is within $\pi(\theta|\boldsymbol{x})$

- Assume a quantity of interest, for example some parameter θ.
- Once you have calculated the posterior distribution $\pi(\theta|\boldsymbol{x})$, you have a complete description of your beliefs and uncertainty about θ.
- Any statement on θ can be made, as long as you can integrate the posterior distribution, analytically, numerically, or with the help of some statistical software such as R.
- For example, your expectation for the true value of θ is,

$$\int_{\Theta} \theta \pi(\theta|\boldsymbol{x}) d\theta$$

- Or, the probability that θ is greater than 10, is,

$$\int_{10}^{\infty} \pi(\theta|\boldsymbol{x}) d\theta$$

- It is standard practice to report certain summaries of the posterior distribution that describe one's beliefs and uncertainty about θ.

Outline

Point estimates

- Some point estimates that summarize posterior distributions can be the
  - mean of the posterior distribution (posterior mean)
  - mode of the posterior (posterior mode)
  - median of the posterior (posterior median)
- These are intuitively 'sensible' estimates.
- These estimates can also be derived theoretically, after one defines their personal 'loss' when estimating some quantity or parameter (see lecture notes)

Point estimates

- Some point estimates that summarize posterior distributions can be the
  - mean of the posterior distribution (posterior mean)
  - mode of the posterior (posterior mode)
  - median of the posterior (posterior median)
- These are intuitively 'sensible' estimates.
- These estimates can also be derived theoretically, after one defines their personal 'loss' when estimating some quantity or parameter (see lecture notes)

Point estimates

- Some point estimates that summarize posterior distributions can be the
  - mean of the posterior distribution (posterior mean)
  - mode of the posterior (posterior mode)
  - median of the posterior (posterior median)
- These are intuitively 'sensible' estimates.
- These estimates can also be derived theoretically, after one defines their personal 'loss' when estimating some quantity or parameter (see lecture notes)

Point estimates

- Some point estimates that summarize posterior distributions can be the
  - mean of the posterior distribution (posterior mean)
  - mode of the posterior (posterior mode)
  - median of the posterior (posterior median)
- These are intuitively 'sensible' estimates.
- These estimates can also be derived theoretically, after one defines their personal 'loss' when estimating some quantity or parameter (see lecture notes)

Point estimates

- Some point estimates that summarize posterior distributions can be the
  - mean of the posterior distribution (posterior mean)
  - mode of the posterior (posterior mode)
  - median of the posterior (posterior median)
- These are intuitively 'sensible' estimates.
- These estimates can also be derived theoretically, after one defines their personal 'loss' when estimating some quantity or parameter (see lecture notes)

Point estimates

- Some point estimates that summarize posterior distributions can be the
  - mean of the posterior distribution (posterior mean)
  - mode of the posterior (posterior mode)
  - median of the posterior (posterior median)
- These are intuitively 'sensible' estimates.
- These estimates can also be derived theoretically, after one defines their personal 'loss' when estimating some quantity or parameter (see lecture notes)

Outline

Classical/frequentist Confidence Intervals

- To obtain some perspective, consider first the classical $100(1-\alpha)\%$ confidence interval.
- Without loss of generality, consider $\alpha = 0.05$, i.e. a 95% conf. int.
- If the experiment is repeated 100 times, 95 out of 100 confidence intervals formed are expected to contain the (fixed) unknown parameter value.
- The 100 derived conf. intervals will typically all be slightly different.
- It is not straightforward to explain to a non-statistician what a 95% conf. int. describes...
- ... as a 95% conf. int. **does not** contain the true value of the parameter with probability 95%.
- Statisticians explaining inferences often talk about 'confidence' rather than probability.
- This muddles the message, often resulting in an audience that believes that the derived 95% conf. int. contains the true value of the parameter with probability 95%.

Classical/frequentist Confidence Intervals

- To obtain some perspective, consider first the classical $100(1-\alpha)\%$ confidence interval.
- Without loss of generality, consider $\alpha = 0.05$, i.e. a 95% conf. int.
- If the experiment is repeated 100 times, 95 out of 100 confidence intervals formed are expected to contain the (fixed) unknown parameter value.
- The 100 derived conf. intervals will typically all be slightly different.
- It is not straightforward to explain to a non-statistician what a 95% conf. int. describes...
- ... as a 95% conf. int. **does not** contain the true value of the parameter with probability 95%.
- Statisticians explaining inferences often talk about 'confidence' rather than probability.
- This muddles the message, often resulting in an audience that believes that the derived 95% conf. int. contains the true value of the parameter with probability 95%.

Classical/frequentist Confidence Intervals

- To obtain some perspective, consider first the classical $100(1-\alpha)\%$ confidence interval.
- Without loss of generality, consider $\alpha = 0.05$, i.e. a 95% conf. int.
- If the experiment is repeated 100 times, 95 out of 100 confidence intervals formed are expected to contain the (fixed) unknown parameter value.
- The 100 derived conf. intervals will typically all be slightly different.
- It is not straightforward to explain to a non-statistician what a 95% conf. int. describes...
- ... as a 95% conf. int. **does not** contain the true value of the parameter with probability 95%.
- Statisticians explaining inferences often talk about 'confidence' rather than probability.
- This muddles the message, often resulting in an audience that believes that the derived 95% conf. int. contains the true value of the parameter with probability 95%.

Classical/frequentist Confidence Intervals

- To obtain some perspective, consider first the classical $100(1-\alpha)\%$ confidence interval.
- Without loss of generality, consider $\alpha = 0.05$, i.e. a 95% conf. int.
- If the experiment is repeated 100 times, 95 out of 100 confidence intervals formed are expected to contain the (fixed) unknown parameter value.
- The 100 derived conf. intervals will typically all be slightly different.
- It is not straightforward to explain to a non-statistician what a 95% conf. int. describes...
- ... as a 95% conf. int. **does not** contain the true value of the parameter with probability 95%.
- Statisticians explaining inferences often talk about 'confidence' rather than probability.
- This muddles the message, often resulting in an audience that believes that the derived 95% conf. int. contains the true value of the parameter with probability 95%.

Classical/frequentist Confidence Intervals

- To obtain some perspective, consider first the classical $100(1-\alpha)\%$ confidence interval.
- Without loss of generality, consider $\alpha = 0.05$, i.e. a 95% conf. int.
- If the experiment is repeated 100 times, 95 out of 100 confidence intervals formed are expected to contain the (fixed) unknown parameter value.
- The 100 derived conf. intervals will typically all be slightly different.
- It is not straightforward to explain to a non-statistician what a 95% conf. int. describes...
- ... as a 95% conf. int. **does not** contain the true value of the parameter with probability 95%.
- Statisticians explaining inferences often talk about 'confidence' rather than probability.
- This muddles the message, often resulting in an audience that believes that the derived 95% conf. int. contains the true value of the parameter with probability 95%.

Classical/frequentist Confidence Intervals

- To obtain some perspective, consider first the classical $100(1-\alpha)\%$ confidence interval.
- Without loss of generality, consider $\alpha = 0.05$, i.e. a 95% conf. int.
- If the experiment is repeated 100 times, 95 out of 100 confidence intervals formed are expected to contain the (fixed) unknown parameter value.
- The 100 derived conf. intervals will typically all be slightly different.
- It is not straightforward to explain to a non-statistician what a 95% conf. int. describes...
- ... as a 95% conf. int. **does not** contain the true value of the parameter with probability 95%.
- Statisticians explaining inferences often talk about 'confidence' rather than probability.
- This muddles the message, often resulting in an audience that believes that the derived 95% conf. int. contains the true value of the parameter with probability 95%.

Classical/frequentist Confidence Intervals

- To obtain some perspective, consider first the classical $100(1-\alpha)\%$ confidence interval.
- Without loss of generality, consider $\alpha = 0.05$, i.e. a 95% conf. int.
- If the experiment is repeated 100 times, 95 out of 100 confidence intervals formed are expected to contain the (fixed) unknown parameter value.
- The 100 derived conf. intervals will typically all be slightly different.
- It is not straightforward to explain to a non-statistician what a 95% conf. int. describes...
- ... as a 95% conf. int. **does not** contain the true value of the parameter with probability 95%.
- Statisticians explaining inferences often talk about 'confidence' rather than probability.
- This muddles the message, often resulting in an audience that believes that the derived 95% conf. int. contains the true value of the parameter with probability 95%.

Classical/frequentist Confidence Intervals

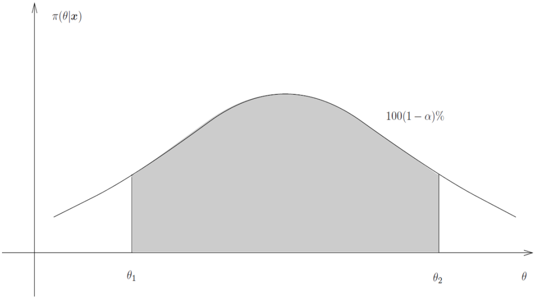
- To obtain some perspective, consider first the classical $100(1-\alpha)\%$ confidence interval.
- Without loss of generality, consider $\alpha = 0.05$, i.e. a 95% conf. int.
- If the experiment is repeated 100 times, 95 out of 100 confidence intervals formed are expected to contain the (fixed) unknown parameter value.
- The 100 derived conf. intervals will typically all be slightly different.
- It is not straightforward to explain to a non-statistician what a 95% conf. int. describes...
- ... as a 95% conf. int. **does not** contain the true value of the parameter with probability 95%.
- Statisticians explaining inferences often talk about 'confidence' rather than probability.
- This muddles the message, often resulting in an audience that believes that the derived 95% conf. int. contains the true value of the parameter with probability 95%.

Bayesian Credible Interval

- **Definition:** The interval (θ_1, θ_2) is defined as an $100(1-\alpha)\%$ credible interval if,

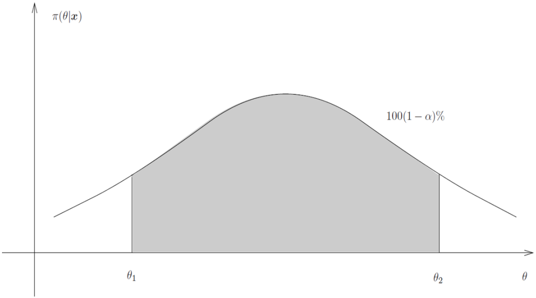
$$\int_{\theta_1}^{\theta_2} \pi(\theta|\boldsymbol{x})d\theta = 1-\alpha, \qquad 0 \leq \alpha \leq 1.$$

- For $\alpha = 0.05$, a 95% credible interval (CI) (θ_1, θ_2) contains the true value of θ with 95% probability.
- Easy to understand and explain!



Bayesian Credible Interval

- **Definition:** The interval (θ_1, θ_2) is defined as an $100(1-\alpha)\%$ credible interval if,
$$\int_{\theta_1}^{\theta_2} \pi(\theta|\boldsymbol{x})d\theta = 1-\alpha, \qquad 0 \leq \alpha \leq 1.$$
- For $\alpha = 0.05$, a 95% credible interval (CI) (θ_1, θ_2) contains the true value of θ with 95% probability.
- Easy to understand and explain!

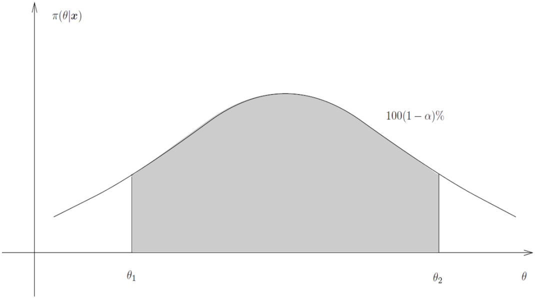


Bayesian Credible Interval

- **Definition:** The interval (θ_1, θ_2) is defined as an $100(1-\alpha)\%$ credible interval if,

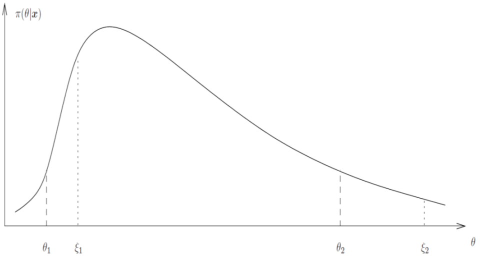
$$\int_{\theta_1}^{\theta_2} \pi(\theta|\boldsymbol{x})d\theta = 1-\alpha, \qquad 0 \le \alpha \le 1.$$

- For $\alpha = 0.05$, a 95% credible interval (CI) (θ_1, θ_2) contains the true value of θ with 95% probability.
- Easy to understand and explain!



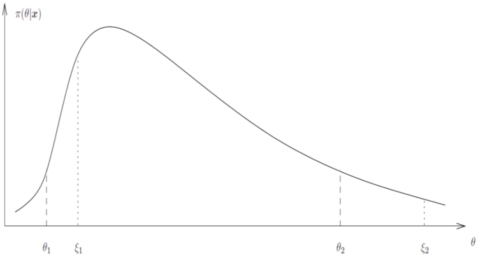
Credible Intervals are not unique

- A $100(1-\alpha)\%$ credible interval is not unique.
- In general, there will be many choices of θ_1 and θ_2, such that, $\int_{\theta_1}^{\theta_2} \pi(\theta|\boldsymbol{x})d\theta = 1-\alpha$.
- Below, we show two $100(1-\alpha)\%$ credible intervals, $[\theta_1, \theta_2]$ and $[\xi_1, \xi_2]$, for some parameter θ, and its distribution $p(.)$ (not necessarily a posterior distribution).



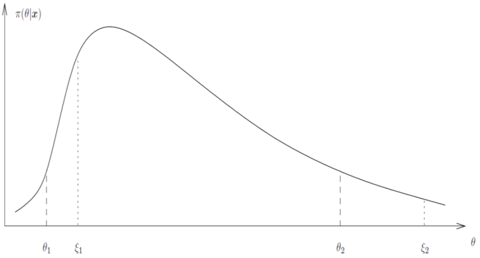
Credible Intervals are not unique

- A $100(1-\alpha)\%$ credible interval is not unique.
- In general, there will be many choices of θ_1 and θ_2, such that, $\int_{\theta_1}^{\theta_2} \pi(\theta|\boldsymbol{x})d\theta = 1-\alpha$.
- Below, we show two $100(1-\alpha)\%$ credible intervals, $[\theta_1, \theta_2]$ and $[\xi_1, \xi_2]$, for some parameter θ, and its distribution $p(.)$ (not necessarily a posterior distribution).



Credible Intervals are not unique

- A $100(1-\alpha)\%$ credible interval is not unique.
- In general, there will be many choices of θ_1 and θ_2, such that, $\int_{\theta_1}^{\theta_2} \pi(\theta|\boldsymbol{x})d\theta = 1-\alpha$.
- Below, we show two $100(1-\alpha)\%$ credible intervals, $[\theta_1, \theta_2]$ and $[\xi_1, \xi_2]$, for some parameter θ, and its distribution $p(.)$ (not necessarily a posterior distribution).



Symmetric Credible Intervals

- Often, a symmetric $100(1-\alpha)\%$ credible interval (θ_1, θ_2) is used. This credible interval is unique, and is defined such that,

$$\int_{-\infty}^{\theta_1} \pi(\theta|\boldsymbol{x}) d\theta = \frac{\alpha}{2} = \int_{\theta_2}^{\infty} \pi(\theta|\boldsymbol{x}) d\theta,$$

- For example, a 95% credible interval is such that θ_1 corresponds to the lower 2.5% quantile, and θ_2 to the upper 97.5% quantile of the posterior distribution $\pi(\theta|\boldsymbol{x})$.

Symmetric Credible Intervals

- Often, a symmetric $100(1-\alpha)\%$ credible interval (θ_1, θ_2) is used. This credible interval is unique, and is defined such that,

$$\int_{-\infty}^{\theta_1} \pi(\theta|\boldsymbol{x})d\theta = \frac{\alpha}{2} = \int_{\theta_2}^{\infty} \pi(\theta|\boldsymbol{x})d\theta,$$

- For example, a 95% credible interval is such that θ_1 corresponds to the lower 2.5% quantile, and θ_2 to the upper 97.5% quantile of the posterior distribution $\pi(\theta|\boldsymbol{x})$.

Highest Posterior Density Intervals

- A narrower $100(1-\alpha)\%$ credible interval is clearly more informative compared to a wider $100(1-\alpha)\%$ one of the same probability coverage.
- This motivates the following definition.
- **Definition:** The interval $[\theta_1, \theta_2]$ is a $100(1-\alpha)\%$ **highest posterior density interval** (HPDI) if:
  1. $[\theta_1, \theta_2]$ is a $100(1-\alpha)\%$ credible interval; and
  2. for all $\theta' \in [\theta_1, \theta_2]$ and $\theta'' \notin [\theta_1, \theta_2]$, $\pi(\theta'|\boldsymbol{x}) \geq \pi(\theta''|\boldsymbol{x})$.

Highest Posterior Density Intervals

- A narrower $100(1-\alpha)\%$ credible interval is clearly more informative compared to a wider $100(1-\alpha)\%$ one of the same probability coverage.
- This motivates the following definition.
- **Definition:** The interval $[\theta_1, \theta_2]$ is a $100(1-\alpha)\%$ **highest posterior density interval** (HPDI) if:
  1. $[\theta_1, \theta_2]$ is a $100(1-\alpha)\%$ credible interval; and
  2. for all $\theta' \in [\theta_1, \theta_2]$ and $\theta'' \notin [\theta_1, \theta_2]$, $\pi(\theta'|\boldsymbol{x}) \geq \pi(\theta''|\boldsymbol{x})$.

Highest Posterior Density Intervals

- A narrower $100(1-\alpha)\%$ credible interval is clearly more informative compared to a wider $100(1-\alpha)\%$ one of the same probability coverage.
- This motivates the following definition.
- **Definition:** The interval $[\theta_1, \theta_2]$ is a $100(1-\alpha)\%$ **highest posterior density interval** (HPDI) if:
  1. $[\theta_1, \theta_2]$ is a $100(1-\alpha)\%$ credible interval; and
  2. for all $\theta' \in [\theta_1, \theta_2]$ and $\theta'' \notin [\theta_1, \theta_2]$, $\pi(\theta'|\boldsymbol{x}) \geq \pi(\theta''|\boldsymbol{x})$.

Highest Posterior Density Intervals

- A narrower $100(1-\alpha)\%$ credible interval is clearly more informative compared to a wider $100(1-\alpha)\%$ one of the same probability coverage.
- This motivates the following definition.
- **Definition:** The interval $[\theta_1, \theta_2]$ is a $100(1-\alpha)\%$ **highest posterior density interval** (HPDI) if:
  1. $[\theta_1, \theta_2]$ is a $100(1-\alpha)\%$ credible interval; and
  2. for all $\theta' \in [\theta_1, \theta_2]$ and $\theta'' \notin [\theta_1, \theta_2]$, $\pi(\theta'|\boldsymbol{x}) \geq \pi(\theta''|\boldsymbol{x})$.

Highest Posterior Density Intervals

- A narrower $100(1-\alpha)\%$ credible interval is clearly more informative compared to a wider $100(1-\alpha)\%$ one of the same probability coverage.
- This motivates the following definition.
- **Definition:** The interval $[\theta_1, \theta_2]$ is a $100(1-\alpha)\%$ **highest posterior density interval** (HPDI) if:
  1. $[\theta_1, \theta_2]$ is a $100(1-\alpha)\%$ credible interval; and
  2. for all $\theta' \in [\theta_1, \theta_2]$ and $\theta'' \notin [\theta_1, \theta_2]$, $\pi(\theta'|\boldsymbol{x}) \geq \pi(\theta''|\boldsymbol{x})$.

Highest Posterior Density Intervals (2)

- The HPDI is the narrowest possible interval having a given credible level $1 - \alpha$.
- It centres the interval around the mode, in the uni-modal case.
- If the distribution is symmetric about the mean, such as the Normal distribution, the $100(1-\alpha)\%$ symmetric credible interval is identical to the $100(1-\alpha)\%$ HPDI.
- In the case where the posterior distribution is multi-modal, the corresponding HPDI may consist of several disjoint intervals.

Highest Posterior Density Intervals (2)

- The HPDI is the narrowest possible interval having a given credible level $1 - \alpha$.
- It centres the interval around the mode, in the uni-modal case.
- If the distribution is symmetric about the mean, such as the Normal distribution, the $100(1 - \alpha)\%$ symmetric credible interval is identical to the $100(1 - \alpha)\%$ HPDI.
- In the case where the posterior distribution is multi-modal, the corresponding HPDI may consist of several disjoint intervals.

Highest Posterior Density Intervals (2)

- The HPDI is the narrowest possible interval having a given credible level $1 - \alpha$.
- It centres the interval around the mode, in the uni-modal case.
- If the distribution is symmetric about the mean, such as the Normal distribution, the $100(1 - \alpha)\%$ symmetric credible interval is identical to the $100(1 - \alpha)\%$ HPDI.
- In the case where the posterior distribution is multi-modal, the corresponding HPDI may consist of several disjoint intervals.

Highest Posterior Density Intervals (2)

- The HPDI is the narrowest possible interval having a given credible level $1 - \alpha$.
- It centres the interval around the mode, in the uni-modal case.
- If the distribution is symmetric about the mean, such as the Normal distribution, the $100(1-\alpha)\%$ symmetric credible interval is identical to the $100(1-\alpha)\%$ HPDI.
- In the case where the posterior distribution is multi-modal, the corresponding HPDI may consist of several disjoint intervals.

Credible Intervals and parameter transformations

- Suppose that we have a symmetric credible interval $[\theta_1, \theta_2]$ for a given parameter θ.
- If we consider a bijective and monotonic transformation of the parameters, such as $g(\theta)$, the corresponding symmetric credible interval is given by $[g(\theta_1), g(\theta_2)]$,
- However, this is not always true for HPDI's.
- The corresponding HPDI on $g(\theta)$ is $[g(\theta_1), g(\theta_2)]$, if and only if, g is a linear transformation; else, the HPDI needs to be recalculated for $g(\theta)$.

Credible Intervals and parameter transformations

- Suppose that we have a symmetric credible interval $[\theta_1, \theta_2]$ for a given parameter θ.
- If we consider a bijective and monotonic transformation of the parameters, such as $g(\theta)$, the corresponding symmetric credible interval is given by $[g(\theta_1), g(\theta_2)]$,
- However, this is not always true for HPDI's.
- The corresponding HPDI on $g(\theta)$ is $[g(\theta_1), g(\theta_2)]$, if and only if, g is a linear transformation; else, the HPDI needs to be recalculated for $g(\theta)$.

Credible Intervals and parameter transformations

- Suppose that we have a symmetric credible interval $[\theta_1, \theta_2]$ for a given parameter θ.
- If we consider a bijective and monotonic transformation of the parameters, such as $g(\theta)$, the corresponding symmetric credible interval is given by $[g(\theta_1), g(\theta_2)]$,
- However, this is not always true for HPDI's.
- The corresponding HPDI on $g(\theta)$ is $[g(\theta_1), g(\theta_2)]$, if and only if, g is a linear transformation; else, the HPDI needs to be recalculated for $g(\theta)$.

Credible Intervals and parameter transformations

- Suppose that we have a symmetric credible interval $[\theta_1, \theta_2]$ for a given parameter θ.
- If we consider a bijective and monotonic transformation of the parameters, such as $g(\theta)$, the corresponding symmetric credible interval is given by $[g(\theta_1), g(\theta_2)]$,
- However, this is not always true for HPDI's.
- The corresponding HPDI on $g(\theta)$ is $[g(\theta_1), g(\theta_2)]$, if and only if, g is a linear transformation; else, the HPDI needs to be recalculated for $g(\theta)$.

Estimates and stochastic simulation

- For non-trivial problems, stochastic simulation is used for obtaining summary estimates of posterior distributions.
- A random sample from the posterior distribution provides estimates of summary statistics.
- For example, the mean of the posterior distribution is estimated via the average of the random sample.
- Consider for example the posterior probability that $\theta > 10$.
- We can estimate $\mathbb{P}_\pi(\theta > 10)$ by calculating the proportion of the sampled θ values that are greater than 10.
- Credible interval estimates are calculated in a similar fashion.
- These (and other) ideas will be discussed in greater detail later.

Estimates and stochastic simulation

- For non-trivial problems, stochastic simulation is used for obtaining summary estimates of posterior distributions.
- A random sample from the posterior distribution provides estimates of summary statistics.
- For example, the mean of the posterior distribution is estimated via the average of the random sample.
- Consider for example the posterior probability that $\theta > 10$.
- We can estimate $\mathbb{P}_{\pi}(\theta > 10)$ by calculating the proportion of the sampled θ values that are greater than 10.
- Credible interval estimates are calculated in a similar fashion.
- These (and other) ideas will be discussed in greater detail later.

Estimates and stochastic simulation

- For non-trivial problems, stochastic simulation is used for obtaining summary estimates of posterior distributions.
- A random sample from the posterior distribution provides estimates of summary statistics.
- For example, the mean of the posterior distribution is estimated via the average of the random sample.
- Consider for example the posterior probability that $\theta > 10$.
- We can estimate $\mathbb{P}_\pi(\theta > 10)$ by calculating the proportion of the sampled θ values that are greater than 10.
- Credible interval estimates are calculated in a similar fashion.
- These (and other) ideas will be discussed in greater detail later.

Estimates and stochastic simulation

- For non-trivial problems, stochastic simulation is used for obtaining summary estimates of posterior distributions.
- A random sample from the posterior distribution provides estimates of summary statistics.
- For example, the mean of the posterior distribution is estimated via the average of the random sample.
- Consider for example the posterior probability that $\theta > 10$.
- We can estimate $\mathbb{P}_\pi(\theta > 10)$ by calculating the proportion of the sampled θ values that are greater than 10.
- Credible interval estimates are calculated in a similar fashion.
- These (and other) ideas will be discussed in greater detail later.

Estimates and stochastic simulation

- For non-trivial problems, stochastic simulation is used for obtaining summary estimates of posterior distributions.
- A random sample from the posterior distribution provides estimates of summary statistics.
- For example, the mean of the posterior distribution is estimated via the average of the random sample.
- Consider for example the posterior probability that $\theta > 10$.
- We can estimate $\mathbb{P}_\pi(\theta > 10)$ by calculating the proportion of the sampled θ values that are greater than 10.
- Credible interval estimates are calculated in a similar fashion.
- These (and other) ideas will be discussed in greater detail later.

Estimates and stochastic simulation

- For non-trivial problems, stochastic simulation is used for obtaining summary estimates of posterior distributions.
- A random sample from the posterior distribution provides estimates of summary statistics.
- For example, the mean of the posterior distribution is estimated via the average of the random sample.
- Consider for example the posterior probability that $\theta > 10$.
- We can estimate $\mathbb{P}_\pi(\theta > 10)$ by calculating the proportion of the sampled θ values that are greater than 10.
- Credible interval estimates are calculated in a similar fashion.
- These (and other) ideas will be discussed in greater detail later.

Estimates and stochastic simulation

- For non-trivial problems, stochastic simulation is used for obtaining summary estimates of posterior distributions.
- A random sample from the posterior distribution provides estimates of summary statistics.
- For example, the mean of the posterior distribution is estimated via the average of the random sample.
- Consider for example the posterior probability that $\theta > 10$.
- We can estimate $\mathbb{P}_\pi(\theta > 10)$ by calculating the proportion of the sampled θ values that are greater than 10.
- Credible interval estimates are calculated in a similar fashion.
- These (and other) ideas will be discussed in greater detail later.

Finally...

- As discussed in this lecture, when describing a posterior distribution via summary statistics, a variety of point and interval estimates can be given to encapsulate the main properties of the distribution.
- In practice, the mean, median and 95% symmetric CI are the most common choices.

Finally...

- As discussed in this lecture, when describing a posterior distribution via summary statistics, a variety of point and interval estimates can be given to encapsulate the main properties of the distribution.
- In practice, the mean, median and 95% symmetric CI are the most common choices.