MT4531/MT5731: (Advanced) Bayesian Inference

Bayes' Theorem (continuous case)

Nicolò Margaritella

School of Mathematics and Statistics, University of St Andrews

–

University of St Andrews

Outline

Outline

Bayes Theorem (continuous case)

- Bayes' Theorem for continuous quantities (usually model parameters) is the version of the theorem most often used.
- Suppose that we have a parameter $\theta \in \Theta$, on which we wish to make inference.
- We then observe data $\boldsymbol{x} = \{x_1, \ldots, x_n\}$ from some probability distribution $f(\boldsymbol{x}|\theta)$.
- Then Bayes' Theorem states that,

$$\pi(\theta|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\theta)p(\theta)}{f(\boldsymbol{x})}$$

- Functions $p(.)$, $\pi(.)$ and $f(.)$ are all just distributions, be it univariate or multivariate joint distributions, conditional distributions, pmfs or pdfs.

Bayes Theorem (continuous case)

- Bayes' Theorem for continuous quantities (usually model parameters) is the version of the theorem most often used.
- Suppose that we have a parameter $\theta \in \Theta$, on which we wish to make inference.
- We then observe data $\boldsymbol{x} = \{x_1, \ldots, x_n\}$ from some probability distribution $f(\boldsymbol{x}|\theta)$.
- Then Bayes' Theorem states that,

$$\pi(\theta|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\theta)p(\theta)}{f(\boldsymbol{x})}$$

- Functions $p(.)$, $\pi(.)$ and $f(.)$ are all just distributions, be it univariate or multivariate joint distributions, conditional distributions, pmfs or pdfs.

Bayes Theorem (continuous case)

- Bayes' Theorem for continuous quantities (usually model parameters) is the version of the theorem most often used.
- Suppose that we have a parameter $\theta \in \Theta$, on which we wish to make inference.
- We then observe data $\boldsymbol{x} = \{x_1, \ldots, x_n\}$ from some probability distribution $f(\boldsymbol{x}|\theta)$.
- Then Bayes' Theorem states that,

$$\pi(\theta|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\theta)p(\theta)}{f(\boldsymbol{x})}$$

- Functions $p(.)$, $\pi(.)$ and $f(.)$ are all just distributions, be it univariate or multivariate joint distributions, conditional distributions, pmfs or pdfs.

Bayes Theorem (continuous case)

- Bayes' Theorem for continuous quantities (usually model parameters) is the version of the theorem most often used.
- Suppose that we have a parameter $\theta \in \Theta$, on which we wish to make inference.
- We then observe data $\boldsymbol{x} = \{x_1, \ldots, x_n\}$ from some probability distribution $f(\boldsymbol{x}|\theta)$.
- Then Bayes' Theorem states that,

$$\pi(\theta|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\theta)p(\theta)}{f(\boldsymbol{x})}$$

- Functions $p(.)$, $\pi(.)$ and $f(.)$ are all just distributions, be it univariate or multivariate joint distributions, conditional distributions, pmfs or pdfs.

Bayes Theorem (continuous case)

- Bayes' Theorem for continuous quantities (usually model parameters) is the version of the theorem most often used.
- Suppose that we have a parameter $\theta \in \Theta$, on which we wish to make inference.
- We then observe data $\boldsymbol{x} = \{x_1, \ldots, x_n\}$ from some probability distribution $f(\boldsymbol{x}|\theta)$.
- Then Bayes' Theorem states that,

$$\pi(\theta|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\theta)p(\theta)}{f(\boldsymbol{x})}$$

- Functions $p(.)$, $\pi(.)$ and $f(.)$ are all just distributions, be it univariate or multivariate joint distributions, conditional distributions, pmfs or pdfs.

Distributions within the Theorem's formula

- Here the term $p(\theta)$ is referred to as the **prior** distribution.
- $\pi(\theta|\boldsymbol{x})$ is the **posterior** distribution.
- Essentially the prior represents the initial beliefs concerning the parameters prior to any data being observed, whereas the posterior distribution represents an update of these beliefs, following the data $\boldsymbol{x}$ being observed.
- The term $f(\boldsymbol{x}|\theta)$ is typically called the **likelihood**.
- In classical statistics, θ in $f(\boldsymbol{x}|\theta)$ is a single unknown **fixed value** (or vector of values) to estimate.
- In Bayesian statistics, $\pi(\theta|\boldsymbol{x})$ represents a **distribution of θ** after observing x.

Distributions within the Theorem's formula

- Here the term $p(\theta)$ is referred to as the **prior** distribution.
- $\pi(\theta|\boldsymbol{x})$ is the **posterior** distribution.
- Essentially the prior represents the initial beliefs concerning the parameters prior to any data being observed, whereas the posterior distribution represents an update of these beliefs, following the data $\boldsymbol{x}$ being observed.
- The term $f(\boldsymbol{x}|\theta)$ is typically called the **likelihood**.
- In classical statistics, θ in $f(\boldsymbol{x}|\theta)$ is a single unknown **fixed value** (or vector of values) to estimate.
- In Bayesian statistics, $\pi(\theta|\boldsymbol{x})$ represents a **distribution of θ** after observing x.

Distributions within the Theorem's formula

- Here the term $p(\theta)$ is referred to as the **prior** distribution.
- $\pi(\theta|\boldsymbol{x})$ is the **posterior** distribution.
- Essentially the prior represents the initial beliefs concerning the parameters prior to any data being observed, whereas the posterior distribution represents an update of these beliefs, following the data $\boldsymbol{x}$ being observed.
- The term $f(\boldsymbol{x}|\theta)$ is typically called the **likelihood**.
- In classical statistics, θ in $f(\boldsymbol{x}|\theta)$ is a single unknown **fixed value** (or vector of values) to estimate.
- In Bayesian statistics, $\pi(\theta|\boldsymbol{x})$ represents a **distribution of θ** after observing x.

Distributions within the Theorem's formula

- Here the term $p(\theta)$ is referred to as the **prior** distribution.
- $\pi(\theta|\boldsymbol{x})$ is the **posterior** distribution.
- Essentially the prior represents the initial beliefs concerning the parameters prior to any data being observed, whereas the posterior distribution represents an update of these beliefs, following the data $\boldsymbol{x}$ being observed.
- The term $f(\boldsymbol{x}|\boldsymbol{\theta})$ is typically called the **likelihood**.
- In classical statistics, θ in $f(\boldsymbol{x}|\theta)$ is a single unknown **fixed value** (or vector of values) to estimate.
- In Bayesian statistics, $\pi(\theta|\boldsymbol{x})$ represents a **distribution of θ** after observing x.

Distributions within the Theorem's formula

- Here the term $p(\theta)$ is referred to as the **prior** distribution.
- $\pi(\theta|\boldsymbol{x})$ is the **posterior** distribution.
- Essentially the prior represents the initial beliefs concerning the parameters prior to any data being observed, whereas the posterior distribution represents an update of these beliefs, following the data $\boldsymbol{x}$ being observed.
- The term $f(\boldsymbol{x}|\boldsymbol{\theta})$ is typically called the **likelihood**.
- In classical statistics, $\boldsymbol{\theta}$ in $f(\boldsymbol{x}|\boldsymbol{\theta})$ is a single unknown **fixed value** (or vector of values) to estimate.
- In Bayesian statistics, $\pi(\theta|\boldsymbol{x})$ represents a **distribution of θ** after observing x.

Distributions within the Theorem's formula

- Here the term $p(\theta)$ is referred to as the **prior** distribution.
- $\pi(\theta|\boldsymbol{x})$ is the **posterior** distribution.
- Essentially the prior represents the initial beliefs concerning the parameters prior to any data being observed, whereas the posterior distribution represents an update of these beliefs, following the data $\boldsymbol{x}$ being observed.
- The term $f(\boldsymbol{x}|\boldsymbol{\theta})$ is typically called the **likelihood**.
- In classical statistics, $\boldsymbol{\theta}$ in $f(\boldsymbol{x}|\boldsymbol{\theta})$ is a single unknown **fixed value** (or vector of values) to estimate.
- In Bayesian statistics, $\pi(\theta|\boldsymbol{x})$ represents a **distribution of** $\boldsymbol{\theta}$ after observing x.

Bayes formula denominator

- We can write the denominator of Bayes' Theorem as,

$$f(\boldsymbol{x}) = \int_{\Theta} f(\boldsymbol{x}, \theta) d\theta = \int_{\Theta} f(\boldsymbol{x}|\theta) p(\theta) d\theta.$$

- $f(\boldsymbol{x})$ is usually referred to as the marginal likelihood, and $f(\boldsymbol{x})^{-1}$ as the normalization constant.
- In many cases the integration for the marginal likelihood may be analytically intractable.
- $f(\boldsymbol{x})$ is often found by inspection (see later examples).
- Another option is to calculate it using stochastic simulation (see second part of this course).
- However, the calculation of $f(\boldsymbol{x})$ is often not necessary.

Bayes formula denominator

- We can write the denominator of Bayes' Theorem as,

$$f(\boldsymbol{x}) = \int_{\Theta} f(\boldsymbol{x}, \theta) d\theta = \int_{\Theta} f(\boldsymbol{x}|\theta) p(\theta) d\theta.$$

- $f(\boldsymbol{x})$ is usually referred to as the marginal likelihood, and $f(\boldsymbol{x})^{-1}$ as the normalization constant.
- In many cases the integration for the marginal likelihood may be analytically intractable.
- $f(\boldsymbol{x})$ is often found by inspection (see later examples).
- Another option is to calculate it using stochastic simulation (see second part of this course).
- However, the calculation of $f(\boldsymbol{x})$ is often not necessary.

Bayes formula denominator

- We can write the denominator of Bayes' Theorem as,

$$f(\boldsymbol{x}) = \int_{\Theta} f(\boldsymbol{x}, \theta) d\theta = \int_{\Theta} f(\boldsymbol{x}|\theta) p(\theta) d\theta.$$

- $f(\boldsymbol{x})$ is usually referred to as the marginal likelihood, and $f(\boldsymbol{x})^{-1}$ as the normalization constant.
- In many cases the integration for the marginal likelihood may be analytically intractable.
- $f(\boldsymbol{x})$ is often found by inspection (see later examples).
- Another option is to calculate it using stochastic simulation (see second part of this course).
- However, the calculation of $f(\boldsymbol{x})$ is often not necessary.

Bayes formula denominator

- We can write the denominator of Bayes' Theorem as,

$$f(\boldsymbol{x}) = \int_{\Theta} f(\boldsymbol{x}, \theta) d\theta = \int_{\Theta} f(\boldsymbol{x}|\theta) p(\theta) d\theta.$$

- $f(\boldsymbol{x})$ is usually referred to as the marginal likelihood, and $f(\boldsymbol{x})^{-1}$ as the normalization constant.
- In many cases the integration for the marginal likelihood may be analytically intractable.
- $f(\boldsymbol{x})$ is often found by inspection (see later examples).
- Another option is to calculate it using stochastic simulation (see second part of this course).
- However, the calculation of $f(\boldsymbol{x})$ is often not necessary.

Bayes formula denominator

- We can write the denominator of Bayes' Theorem as,

$$f(\boldsymbol{x}) = \int_{\Theta} f(\boldsymbol{x}, \theta) d\theta = \int_{\Theta} f(\boldsymbol{x}|\theta) p(\theta) d\theta.$$

- $f(\boldsymbol{x})$ is usually referred to as the marginal likelihood, and $f(\boldsymbol{x})^{-1}$ as the normalization constant.
- In many cases the integration for the marginal likelihood may be analytically intractable.
- $f(\boldsymbol{x})$ is often found by inspection (see later examples).
- Another option is to calculate it using stochastic simulation (see second part of this course).
- However, the calculation of $f(\boldsymbol{x})$ is often not necessary.

Bayes formula denominator

- We can write the denominator of Bayes' Theorem as,

$$f(\boldsymbol{x}) = \int_{\Theta} f(\boldsymbol{x}, \theta) d\theta = \int_{\Theta} f(\boldsymbol{x}|\theta) p(\theta) d\theta.$$

- $f(\boldsymbol{x})$ is usually referred to as the marginal likelihood, and $f(\boldsymbol{x})^{-1}$ as the normalization constant.
- In many cases the integration for the marginal likelihood may be analytically intractable.
- $f(\boldsymbol{x})$ is often found by inspection (see later examples).
- Another option is to calculate it using stochastic simulation (see second part of this course).
- However, the calculation of $f(\boldsymbol{x})$ is often not necessary.

Bayes Theorem and proportionality

- The denominator $f(\boldsymbol{x})$ in Bayes' Theorem is just a constant with respect to the distribution of interest, $\pi(\theta|\boldsymbol{x})$.
- For observed data $\boldsymbol{x}$, $f(\boldsymbol{x})$ is just a number.
- Dividing by $f(\boldsymbol{x})$ just 'shifts' the posterior distribution function up or down, so that it integrates to one.
- So, the denominator $f(\boldsymbol{x})$ does not change the shape of the posterior distribution $\pi(\theta|\boldsymbol{x})$. Crucially, it is the shape of $\pi(\theta|\boldsymbol{x})$ that informs on which parts of the parameter space of θ are more likely than other parts and by how much.
- This is why, often, Bayes' Theorem is quoted as,

$$\pi(\theta|\boldsymbol{x}) \propto f(\boldsymbol{x}|\theta)p(\theta).$$

- This formula forms the essential core of Bayesian inference.

Bayes Theorem and proportionality

- The denominator $f(\boldsymbol{x})$ in Bayes' Theorem is just a constant with respect to the distribution of interest, $\pi(\theta|\boldsymbol{x})$.
- For observed data $\boldsymbol{x}$, $f(\boldsymbol{x})$ is just a number.
- Dividing by $f(\boldsymbol{x})$ just 'shifts' the posterior distribution function up or down, so that it integrates to one.
- So, the denominator $f(\boldsymbol{x})$ does not change the shape of the posterior distribution $\pi(\theta|\boldsymbol{x})$. Crucially, it is the shape of $\pi(\theta|\boldsymbol{x})$ that informs on which parts of the parameter space of θ are more likely than other parts and by how much.
- This is why, often, Bayes' Theorem is quoted as,

$$\pi(\theta|\boldsymbol{x}) \propto f(\boldsymbol{x}|\theta)p(\theta).$$

- This formula forms the essential core of Bayesian inference.

Bayes Theorem and proportionality

- The denominator $f(\boldsymbol{x})$ in Bayes' Theorem is just a constant with respect to the distribution of interest, $\pi(\theta|\boldsymbol{x})$.
- For observed data $\boldsymbol{x}$, $f(\boldsymbol{x})$ is just a number.
- Dividing by $f(\boldsymbol{x})$ just 'shifts' the posterior distribution function up or down, so that it integrates to one.
- So, the denominator $f(\boldsymbol{x})$ does not change the shape of the posterior distribution $\pi(\theta|\boldsymbol{x})$. Crucially, it is the shape of $\pi(\theta|\boldsymbol{x})$ that informs on which parts of the parameter space of θ are more likely than other parts and by how much.
- This is why, often, Bayes' Theorem is quoted as,

$$\pi(\theta|\boldsymbol{x}) \propto f(\boldsymbol{x}|\theta)p(\theta).$$

- This formula forms the essential core of Bayesian inference.

Bayes Theorem and proportionality

- The denominator $f(\boldsymbol{x})$ in Bayes' Theorem is just a constant with respect to the distribution of interest, $\pi(\theta|\boldsymbol{x})$.
- For observed data $\boldsymbol{x}$, $f(\boldsymbol{x})$ is just a number.
- Dividing by $f(\boldsymbol{x})$ just 'shifts' the posterior distribution function up or down, so that it integrates to one.
- So, the denominator $f(\boldsymbol{x})$ does not change the shape of the posterior distribution $\pi(\theta|\boldsymbol{x})$. Crucially, it is the shape of $\pi(\theta|\boldsymbol{x})$ that informs on which parts of the parameter space of θ are more likely than other parts and by how much.
- This is why, often, Bayes' Theorem is quoted as,

$$\pi(\theta|\boldsymbol{x}) \propto f(\boldsymbol{x}|\theta)p(\theta).$$

- This formula forms the essential core of Bayesian inference.

Bayes Theorem and proportionality

- The denominator $f(\boldsymbol{x})$ in Bayes' Theorem is just a constant with respect to the distribution of interest, $\pi(\theta|\boldsymbol{x})$.
- For observed data $\boldsymbol{x}$, $f(\boldsymbol{x})$ is just a number.
- Dividing by $f(\boldsymbol{x})$ just 'shifts' the posterior distribution function up or down, so that it integrates to one.
- So, the denominator $f(\boldsymbol{x})$ does not change the shape of the posterior distribution $\pi(\theta|\boldsymbol{x})$. Crucially, it is the shape of $\pi(\theta|\boldsymbol{x})$ that informs on which parts of the parameter space of θ are more likely than other parts and by how much.
- This is why, often, Bayes' Theorem is quoted as,

$$\pi(\theta|\boldsymbol{x}) \propto f(\boldsymbol{x}|\theta)p(\theta).$$

- This formula forms the essential core of Bayesian inference.

Bayes Theorem and proportionality

- The denominator $f(\boldsymbol{x})$ in Bayes' Theorem is just a constant with respect to the distribution of interest, $\pi(\theta|\boldsymbol{x})$.
- For observed data $\boldsymbol{x}$, $f(\boldsymbol{x})$ is just a number.
- Dividing by $f(\boldsymbol{x})$ just 'shifts' the posterior distribution function up or down, so that it integrates to one.
- So, the denominator $f(\boldsymbol{x})$ does not change the shape of the posterior distribution $\pi(\theta|\boldsymbol{x})$. Crucially, it is the shape of $\pi(\theta|\boldsymbol{x})$ that informs on which parts of the parameter space of θ are more likely than other parts and by how much.
- This is why, often, Bayes' Theorem is quoted as,

$$\pi(\theta|\boldsymbol{x}) \propto f(\boldsymbol{x}|\theta)p(\theta).$$

- This formula forms the essential core of Bayesian inference.

Outline

Example: Exponential likelihood - Gamma prior

- Suppose that we observe data $\boldsymbol{x} = \{x_1, \ldots, x_n\}$, such that, given λ, each $X_i \overset{iid}{\sim} Exp(\lambda)$.
- Now, $E(X_i|\lambda) = 1/\lambda$.
- We place the following prior on λ, namely that,

$$\lambda \sim \Gamma(\alpha, \beta),$$

where α and β are known. Note: $E(\lambda) = \alpha/\beta$.

- What is the posterior distribution of λ, given the observed data $\boldsymbol{x} = \{x_1, \ldots, x_n\}$?

Example: Exponential likelihood - Gamma prior

- Suppose that we observe data $\boldsymbol{x} = \{x_1, \ldots, x_n\}$, such that, given λ, each $X_i \overset{iid}{\sim} Exp(\lambda)$.
- Now, $E(X_i|\lambda) = 1/\lambda$.
- We place the following prior on λ, namely that,

$$\lambda \sim \Gamma(\alpha, \beta),$$

where α and β are known. Note: $E(\lambda) = \alpha/\beta$.
- What is the posterior distribution of λ, given the observed data $\boldsymbol{x} = \{x_1, \ldots, x_n\}$?

Example: Exponential likelihood - Gamma prior

- Suppose that we observe data $\boldsymbol{x} = \{x_1, \ldots, x_n\}$, such that, given λ, each $X_i \stackrel{iid}{\sim} Exp(\lambda)$.
- Now, $E(X_i|\lambda) = 1/\lambda$.
- We place the following prior on λ, namely that,

$$\lambda \sim \Gamma(\alpha, \beta),$$

where α and β are known. Note: $E(\lambda) = \alpha/\beta$.

- What is the posterior distribution of λ, given the observed data $\boldsymbol{x} = \{x_1, \ldots, x_n\}$?

Example: Exponential likelihood - Gamma prior

- Suppose that we observe data $\boldsymbol{x} = \{x_1, \ldots, x_n\}$, such that, given λ, each $X_i \overset{iid}{\sim} Exp(\lambda)$.
- Now, $E(X_i|\lambda) = 1/\lambda$.
- We place the following prior on λ, namely that,

$$\lambda \sim \Gamma(\alpha, \beta),$$

  where α and β are known. Note: $E(\lambda) = \alpha/\beta$.
- What is the posterior distribution of λ, given the observed data $\boldsymbol{x} = \{x_1, \ldots, x_n\}$?

Example - Posterior for λ

- The posterior distribution for λ is given by,

$$\pi(\lambda|\boldsymbol{x}) \quad \propto \quad f(\boldsymbol{x}|\lambda)p(\lambda) = f(x_1|\lambda) \times \ldots f(x_n|\lambda)p(\lambda)$$

Example - Posterior for λ

- The posterior distribution for λ is given by,

$$\begin{aligned}
\pi(\lambda|\boldsymbol{x}) &\propto f(\boldsymbol{x}|\lambda)p(\lambda) = f(x_1|\lambda) \times \ldots f(x_n|\lambda)p(\lambda) \\
&= \prod_{i=1}^{n} \lambda \exp(-x_i\lambda) \times \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}\exp(-\lambda\beta) \\
&\propto \lambda^n \exp\left(-\lambda\sum_{i=1}^{n} x_i\right) \times \lambda^{\alpha-1}\exp(-\lambda\beta) \\
&= \lambda^{n+\alpha-1}\exp(-\lambda[n\bar{x}+\beta]) \\
&\propto \frac{(n\bar{x}+\beta)^{n+\alpha}}{\Gamma(n+\alpha)}\lambda^{n+\alpha-1}\exp(-\lambda[n\bar{x}+\beta]) \\
\Rightarrow \lambda|\boldsymbol{x} &\sim \Gamma(n+\alpha, n\bar{x}+\beta).
\end{aligned}$$

Example - Constant of proportionality

- Given this, we can state that the constant of proportionality (the constant we multiply with to obtain a density that integrates to one) is equal to,

$$\frac{(n\bar{x}+\beta)^{n+\alpha}}{\Gamma(n+\alpha)}.$$

- Or, by inspection, we can write that,

$$f(x) = \frac{\Gamma(n+\alpha)\beta^{\alpha}}{(n\bar{x}+\beta)^{n+\alpha}\Gamma(a)}.$$

Example - Constant of proportionality

- Given this, we can state that the constant of proportionality (the constant we multiply with to obtain a density that integrates to one) is equal to,

$$\frac{(n\bar{x}+\beta)^{n+\alpha}}{\Gamma(n+\alpha)}.$$

- Or, by inspection, we can write that,

$$f(\boldsymbol{x}) = \frac{\Gamma(n+\alpha)\beta^{\alpha}}{(n\bar{x}+\beta)^{n+\alpha}\Gamma(a)}.$$

Example - Tiger cubs

- We are interested in estimating λ, the number of cubs a tiger is expected to give birth at a time in a rescue facility for big cats in the US.
- Assume that prior beliefs about λ are described by $\lambda \sim \Gamma(2, 1)$, so that $E(\lambda) = 2$, and $Var(\lambda) = 2$.
- Eight tigers gave birth at the facility in 2020, and the average number of cubs observed was $\bar{x} = 1.5$.
- We model the number of cubs as independent draws from a Poisson distribution so that $f(x_1, \ldots, x_8|\lambda) = \prod_{i=1}^{8} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$
- Then $\lambda|\boldsymbol{x} \sim \Gamma(14, 9)$ (show that is true).

Example - Tiger cubs

- We are interested in estimating λ, the number of cubs a tiger is expected to give birth at a time in a rescue facility for big cats in the US.
- Assume that prior beliefs about λ are described by $\lambda \sim \Gamma(2,1)$, so that $E(\lambda) = 2$, and $Var(\lambda) = 2$.
- Eight tigers gave birth at the facility in 2020, and the average number of cubs observed was $\bar{x} = 1.5$.
- We model the number of cubs as independent draws from a Poisson distribution so that $f(x_1, \ldots, x_8|\lambda) = \prod_{i=1}^{8} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$
- Then $\lambda|\boldsymbol{x} \sim \Gamma(14, 9)$ (show that is true).

Example - Tiger cubs

- We are interested in estimating λ, the number of cubs a tiger is expected to give birth at a time in a rescue facility for big cats in the US.
- Assume that prior beliefs about λ are described by $\lambda \sim \Gamma(2,1)$, so that $\mathrm{E}(\lambda) = 2$, and $\mathrm{Var}(\lambda) = 2$.
- Eight tigers gave birth at the facility in 2020, and the average number of cubs observed was $\bar{x} = 1.5$.
- We model the number of cubs as independent draws from a Poisson distribution so that $f(x_1, \ldots, x_8|\lambda) = \prod_{i=1}^{8} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$
- Then $\lambda|\boldsymbol{x} \sim \Gamma(14, 9)$ (show that is true).

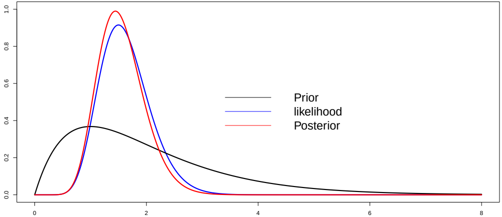
Example - Tiger cubs

- We are interested in estimating λ, the number of cubs a tiger is expected to give birth at a time in a rescue facility for big cats in the US.
- Assume that prior beliefs about λ are described by $\lambda \sim \Gamma(2, 1)$, so that $\mathrm{E}(\lambda) = 2$, and $\mathrm{Var}(\lambda) = 2$.
- Eight tigers gave birth at the facility in 2020, and the average number of cubs observed was $\bar{x} = 1.5$.
- We model the number of cubs as independent draws from a Poisson distribution so that $f(x_1, \ldots, x_8|\lambda) = \prod_{i=1}^{8} \dfrac{\lambda^{x_i} e^{-\lambda}}{x_i!}$
- Then $\lambda|\boldsymbol{x} \sim \Gamma(14, 9)$ (show that is true).

Example - Tiger cubs

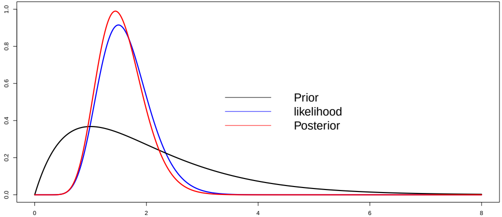
- We are interested in estimating λ, the number of cubs a tiger is expected to give birth at a time in a rescue facility for big cats in the US.
- Assume that prior beliefs about λ are described by $\lambda \sim \Gamma(2, 1)$, so that $\mathrm{E}(\lambda) = 2$, and $\mathrm{Var}(\lambda) = 2$.
- Eight tigers gave birth at the facility in 2020, and the average number of cubs observed was $\bar{x} = 1.5$.
- We model the number of cubs as independent draws from a Poisson distribution so that $f(x_1, \ldots, x_8|\lambda) = \prod_{i=1}^{8} \dfrac{\lambda^{x_i} e^{-\lambda}}{x_i!}$
- Then $\lambda|\boldsymbol{x} \sim \Gamma(14, 9)$ (show that is true).

Example - Tiger cubs



- Note: $E(\lambda) = 2$, $\bar{x} = 1.5$, and $E(\lambda|\boldsymbol{x}) = 14/9 = 1.556$
- $\text{Var}(\lambda) = 2$, $\text{Var}(\lambda|\boldsymbol{x}) = 14/9^2 = 0.173$

Example - Tiger cubs



- Note: $E(\lambda) = 2$, $\bar{x} = 1.5$, and $E(\lambda|\boldsymbol{x}) = 14/9 = 1.556$
- $\text{Var}(\lambda) = 2$, $\text{Var}(\lambda|\boldsymbol{x}) = 14/9^2 = 0.173$

Finally...

- **Task**: Complete the example of a prior to posterior calculation in the lecture notes (Section 1.2.2), for a Geometric likelihood, with a Beta prior distribution.