

Example 2.6: Summary

If we have a random sample from a $N(\mu, 1/\tau)$ distribution (with τ known) and our prior beliefs about μ follow a $N(b, 1/d)$ distribution then, after incorporating the data, our (posterior) beliefs about μ follow a $N(B, 1/D)$ distribution.

Example 2.6: Summary

If we have a random sample from a $N(\mu, 1/\tau)$ distribution (with τ known) and our prior beliefs about μ follow a $N(b, 1/d)$ distribution then, after incorporating the data, our (posterior) beliefs about μ follow a $N(B, 1/D)$ distribution.

Example 2.6: Summary

The changes in our beliefs about μ are summarised in Table 2.6.

	Prior (2.19)	Likelihood (2.20)	Posterior (2.21)
$Mode(\mu)$	b	$\bar{x}$	$(db + n\tau\bar{x})/(d + n\tau)$
$E(\mu)$	b	—	$(db + n\tau\bar{x})/(d + n\tau)$
$Precision(\mu)$	d	—	$d + n\tau$

Example 2.6: Summary

The changes in our beliefs about μ are summarised in Table 2.6.

	Prior (2.19)	Likelihood (2.20)	Posterior (2.21)
$Mode(\mu)$	b	$\bar{x}$	$(db + n\tau\bar{x})/(d + n\tau)$
$E(\mu)$	b	—	$(db + n\tau\bar{x})/(d + n\tau)$
$Precision(\mu)$	d	—	$d + n\tau$

Example 2.6: Summary

Notice that the way prior information and observed data combine is through the parameters of the normal distribution:

$$b \rightarrow \frac{db + n\tau\bar{x}}{d + n\tau} \quad \text{and} \quad d \rightarrow d + n\tau.$$

Example 2.6: Summary

Notice that the way prior information and observed data combine is through the parameters of the normal distribution:

$$b \rightarrow \frac{db + n_{\tau}\bar{x}}{d + n_{\tau}} \quad \text{and} \quad d \rightarrow d + n_{\tau}.$$

Example 2.6: Summary

Notice also that the posterior variance (and precision) does not depend on the data, and the posterior mean is a convex combination of the prior and sample means, that is,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{db}{d + n\tau} + \frac{n\tau\bar{x}}{d + n\tau} \\ &= \frac{d}{d + n\tau}b + \frac{n\tau}{d + n\tau}\bar{x} \\ &= \left(\frac{d}{d + n\tau}\right)b + \left(\frac{d + n\tau - d}{d + n\tau}\right)\bar{x} \\ &= \alpha b + (1 - \alpha)\bar{x}, \quad \alpha = \frac{d}{d + n\tau}. \end{aligned}$$

Example 2.6: Summary

Notice also that the posterior variance (and precision) does not depend on the data, and the posterior mean is a convex combination of the prior and sample means, that is,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{db}{d + n\tau} + \frac{n\tau\bar{x}}{d + n\tau} \\ &= \frac{d}{d + n\tau}b + \frac{n\tau}{d + n\tau}\bar{x} \\ &= \left(\frac{d}{d + n\tau}\right)b + \left(\frac{d + n\tau - d}{d + n\tau}\right)\bar{x} \\ &= \alpha b + (1 - \alpha)\bar{x}, \quad \alpha = \frac{d}{d + n\tau}. \end{aligned}$$

Example 2.6: Summary

Notice also that the posterior variance (and precision) does not depend on the data, and the posterior mean is a convex combination of the prior and sample means, that is,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{db}{d + n\tau} + \frac{n\tau\bar{x}}{d + n\tau} \\ &= \frac{d}{d + n\tau}b + \frac{n\tau}{d + n\tau}\bar{x} \\ &= \left(\frac{d}{d + n\tau}\right)b + \left(\frac{d + n\tau - d}{d + n\tau}\right)\bar{x} \\ &= \alpha b + (1 - \alpha)\bar{x}, \quad \alpha = \frac{d}{d + n\tau}. \end{aligned}$$

Example 2.6: Summary

Notice also that the posterior variance (and precision) does not depend on the data, and the posterior mean is a convex combination of the prior and sample means, that is,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{db}{d + n\tau} + \frac{n\tau\bar{x}}{d + n\tau} \\ &= \frac{d}{d + n\tau}b + \frac{n\tau}{d + n\tau}\bar{x} \\ &= \left(\frac{d}{d + n\tau}\right)b + \left(\frac{d + n\tau - d}{d + n\tau}\right)\bar{x} \\ &= \alpha b + (1 - \alpha)\bar{x}, \quad \alpha = \frac{d}{d + n\tau}. \end{aligned}$$

Example 2.6: Summary

Notice also that the posterior variance (and precision) does not depend on the data, and the posterior mean is a convex combination of the prior and sample means, that is,

$$\begin{aligned} B &= \frac{db + n_{\tau} \bar{x}}{d + n_{\tau}} \\ &= \frac{db}{d + n_{\tau}} + \frac{n_{\tau} \bar{x}}{d + n_{\tau}} \\ &= \frac{d}{d + n_{\tau}} b + \frac{n_{\tau}}{d + n_{\tau}} \bar{x} \\ &= \left(\frac{d}{d + n_{\tau}} \right) b + \left(\frac{d + n_{\tau} - d}{d + n_{\tau}} \right) \bar{x} \\ &= \alpha b + (1 - \alpha) \bar{x}, \quad \alpha = \frac{d}{d + n_{\tau}}. \end{aligned}$$

Example 2.6: Summary

This equation for the posterior mean, which can be rewritten as

$$E(\mu|\mathbf{x}) = \alpha E(\mu) + (1 - \alpha)\bar{x},$$

arises in other models and is known as the **Bayes linear rule**.

Example 2.6: Summary

Notice that the posterior mean is greater than the prior mean if and only if the likelihood mode (sample mean) is greater than the prior mean, that is

$$E(\mu|\mathbf{x}) > E(\mu) \quad \Longleftrightarrow \quad \text{Mode}[f(\mathbf{x}|\mu)] > E(\mu).$$

Also, the standard deviation of the posterior distribution is smaller than that of the prior distribution.

Example 2.6: Summary

Notice that the posterior mean is greater than the prior mean if and only if the likelihood mode (sample mean) is greater than the prior mean, that is

$$E(\mu|\mathbf{x}) > E(\mu) \quad \Longleftrightarrow \quad \text{Mode}[f(\mathbf{x}|\mu)] > E(\mu).$$

Also, the standard deviation of the posterior distribution is smaller than that of the prior distribution.

Example 2.7

The ages of *Ennerdale granophyre* rocks can be determined using the relative proportions of rubidium–87 and strontium–87 in the rock.

An expert in the field suggests that the ages of such rocks (in millions of years) $X|\mu \sim N(\mu, 8^2)$ and that a prior distribution $\mu \sim N(370, 20^2)$ is appropriate.

A rock is found whose chemical analysis yields $x = 421$. What is the posterior distribution for μ and what is the probability that the rock will be older than 400 million years?

Solution to Example 2.7 (1/2)

We have $n = 1$, $\bar{x} = x = 421$, $\tau = 1/64$, $b = 370$ and $d = 1/400$. Therefore, using the results in Example 2.6,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{370/400 + 421/64}{1/400 + 1/64} = 414.0 \quad \text{and} \end{aligned}$$

$$\begin{aligned} D &= d + n\tau \\ &= 1/400 + 1/64 = 1/7.43^2. \end{aligned}$$

Solution to Example 2.7 (1/2)

We have $n = 1$, $\bar{x} = x = 421$, $\tau = 1/64$, $b = 370$ and $d = 1/400$. Therefore, using the results in Example 2.6,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{370/400 + 421/64}{1/400 + 1/64} = 414.0 \quad \text{and} \end{aligned}$$

$$\begin{aligned} D &= d + n\tau \\ &= 1/400 + 1/64 = 1/7.43^2. \end{aligned}$$

Solution to Example 2.7 (1/2)

We have $n = 1$, $\bar{x} = x = 421$, $\tau = 1/64$, $b = 370$ and $d = 1/400$. Therefore, using the results in Example 2.6,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{370/400 + 421/64}{1/400 + 1/64} = 414.0 \quad \text{and} \end{aligned}$$

$$\begin{aligned} D &= d + n\tau \\ &= 1/400 + 1/64 = 1/7.43^2. \end{aligned}$$

Solution to Example 2.7 (1/2)

We have $n = 1$, $\bar{x} = x = 421$, $\tau = 1/64$, $b = 370$ and $d = 1/400$. Therefore, using the results in Example 2.6,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{370/400 + 421/64}{1/400 + 1/64} = 414.0 \quad \text{and} \end{aligned}$$

$$\begin{aligned} D &= d + n\tau \\ &= 1/400 + 1/64 = 1/7.43^2. \end{aligned}$$

Solution to Example 2.7 (1/2)

We have $n = 1$, $\bar{x} = x = 421$, $\tau = 1/64$, $b = 370$ and $d = 1/400$. Therefore, using the results in Example 2.6,

$$\begin{aligned} B &= \frac{db + n\tau\bar{x}}{d + n\tau} \\ &= \frac{370/400 + 421/64}{1/400 + 1/64} = 414.0 \quad \text{and} \end{aligned}$$

$$\begin{aligned} D &= d + n\tau \\ &= 1/400 + 1/64 = 1/7.43^2. \end{aligned}$$

Solution to Example 2.7 (2/2)

Thus, the posterior distribution is $\mu|x = 421 \sim N(414.0, 7.43^2)$.

The (posterior) probability that the rock will be older than 400 million years is

$$\Pr(\mu > 400|x = 421) = 0.9702$$

calculated using the R commands `1-pnorm(400, 414, 7.43)` or `1-pnorm(-1.884)` (or indeed by hand using tables).

Without the chemical analysis, the only basis for determining the age of the rock is via the prior distribution.

The (prior) probability that the rock will be older than 400 million years is $\Pr(\mu > 400) = 0.0668$ calculated using the R command `1-pnorm(400, 370, 20)`.

Solution to Example 2.7 (2/2)

Thus, the posterior distribution is $\mu|x = 421 \sim N(414.0, 7.43^2)$.

The (posterior) probability that the rock will be older than 400 million years is

$$\Pr(\mu > 400|x = 421) = 0.9702$$

calculated using the R commands `1-pnorm(400, 414, 7.43)` or `1-pnorm(-1.884)` (or indeed by hand using tables).

Without the chemical analysis, the only basis for determining the age of the rock is via the prior distribution.

The (prior) probability that the rock will be older than 400 million years is $\Pr(\mu > 400) = 0.0668$ calculated using the R command `1-pnorm(400, 370, 20)`.

Solution to Example 2.7 (2/2)

Thus, the posterior distribution is $\mu|x = 421 \sim N(414.0, 7.43^2)$.

The (posterior) probability that the rock will be older than 400 million years is

$$\Pr(\mu > 400|x = 421) = 0.9702$$

calculated using the R commands `1-pnorm(400, 414, 7.43)` or `1-pnorm(-1.884)` (or indeed by hand using tables).

Without the chemical analysis, the only basis for determining the age of the rock is via the prior distribution.

The (prior) probability that the rock will be older than 400 million years is $\Pr(\mu > 400) = 0.0668$ calculated using the R command `1-pnorm(400, 370, 20)`.

Solution to Example 2.7 (2/2)

Thus, the posterior distribution is $\mu|x = 421 \sim N(414.0, 7.43^2)$.

The (posterior) probability that the rock will be older than 400 million years is

$$\Pr(\mu > 400|x = 421) = 0.9702$$

calculated using the R commands `1-pnorm(400, 414, 7.43)` or `1-pnorm(-1.884)` (or indeed by hand using tables).

Without the chemical analysis, the only basis for determining the age of the rock is via the prior distribution.

The (prior) probability that the rock will be older than 400 million years is $\Pr(\mu > 400) = 0.0668$ calculated using the R command `1-pnorm(400, 370, 20)`.

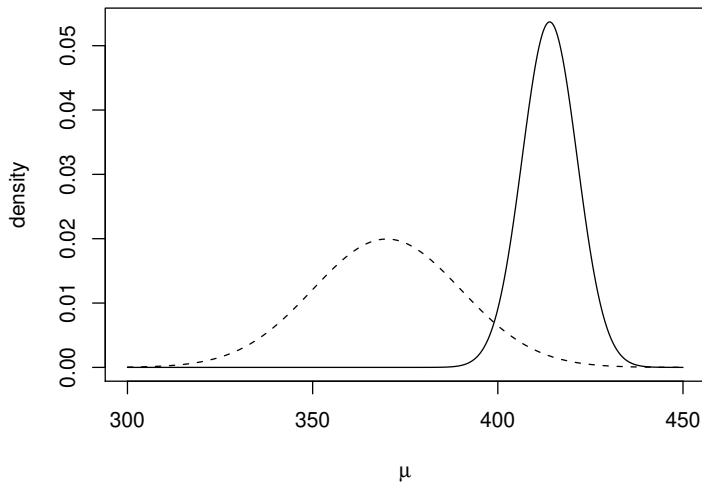
Example 2.7: Summary

This highlights the benefit of taking the chemical measurements.

Note that the large difference between these probabilities is not necessarily due to the expert's prior distribution being inaccurate, *per se*.

It is probably due to the large prior uncertainty about rock ages, as shown in Figure 2.10.

Example 2.7: Summary



We have already met the concept of sufficient statistics. Not surprisingly they also play a role in Bayesian Inference.

Suppose that we have data $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ available and we want to make inferences about the parameter θ in the statistical model $f(\mathbf{x}|\theta)$.

If T is a sufficient statistic then by the Factorisation Theorem

$$f(\mathbf{x}|\theta) = h(\mathbf{x}) g(t, \theta).$$

Posterior distributions and sufficient statistics

We have already met the concept of sufficient statistics. Not surprisingly they also play a role in Bayesian Inference.

Suppose that we have data $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ available and we want to make inferences about the parameter θ in the statistical model $f(\mathbf{x}|\theta)$.

If T is a sufficient statistic then by the Factorisation Theorem

$$f(\mathbf{x}|\theta) = h(\mathbf{x}) g(t, \theta).$$

We have already met the concept of sufficient statistics. Not surprisingly they also play a role in Bayesian Inference.

Suppose that we have data $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ available and we want to make inferences about the parameter θ in the statistical model $f(\mathbf{x}|\theta)$.

If T is a sufficient statistic then by the Factorisation Theorem

$$f(\mathbf{x}|\theta) = h(\mathbf{x}) g(t, \theta).$$

Therefore, using Bayes Theorem,

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\ &\propto \pi(\theta) h(\mathbf{x}) g(t, \theta) \\ &\propto \pi(\theta) g(t, \theta).\end{aligned}$$

It can be shown that, up to a constant not depending on θ , $g(t, \theta)$ is equal to the probability (density) function of T ; that is

$$g(t, \theta) \propto f_T(t|\theta).$$

Hence

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta) f_T(t|\theta).$$

Therefore, using Bayes Theorem,

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\ &\propto \pi(\theta) h(\mathbf{x}) g(t, \theta) \\ &\propto \pi(\theta) g(t, \theta).\end{aligned}$$

It can be shown that, up to a constant not depending on θ , $g(t, \theta)$ is equal to the probability (density) function of T ; that is

$$g(t, \theta) \propto f_T(t|\theta).$$

Hence

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta) f_T(t|\theta).$$

Therefore, using Bayes Theorem,

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\ &\propto \pi(\theta) h(\mathbf{x}) g(t, \theta) \\ &\propto \pi(\theta) g(t, \theta).\end{aligned}$$

It can be shown that, up to a constant not depending on θ , $g(t, \theta)$ is equal to the probability (density) function of T ; that is

$$g(t, \theta) \propto f_T(t|\theta).$$

Hence

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta) f_T(t|\theta).$$

Therefore, using Bayes Theorem,

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\ &\propto \pi(\theta) h(\mathbf{x}) g(t, \theta) \\ &\propto \pi(\theta) g(t, \theta).\end{aligned}$$

It can be shown that, up to a constant not depending on θ , $g(t, \theta)$ is equal to the probability (density) function of T ; that is

$$g(t, \theta) \propto f_T(t|\theta).$$

Hence

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta) f_T(t|\theta).$$

However, applying Bayes Theorem to the data t gives

$$\pi(\theta|t) \propto \pi(\theta) f_T(t|\theta)$$

and so, since

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta|t)$$

and both are probability (density) functions, we have

$$\pi(\theta|\mathbf{x}) = \pi(\theta|t).$$

However, applying Bayes Theorem to the data t gives

$$\pi(\theta|t) \propto \pi(\theta) f_T(t|\theta)$$

and so, since

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta|t)$$

and both are probability (density) functions, we have

$$\pi(\theta|\mathbf{x}) = \pi(\theta|t).$$

However, applying Bayes Theorem to the data t gives

$$\pi(\theta|t) \propto \pi(\theta) f_T(t|\theta)$$

and so, since

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta|t)$$

and both are probability (density) functions, we have

$$\pi(\theta|\mathbf{x}) = \pi(\theta|t).$$

Therefore, our (posterior) beliefs about θ having observed the full data $\mathbf{x}$ are the same as if we had observed only the sufficient statistic T .

This is what we would expect if all the information about θ in the data were contained in the sufficient statistic.

Therefore, our (posterior) beliefs about θ having observed the full data $\mathbf{x}$ are the same as if we had observed only the sufficient statistic T .

This is what we would expect if all the information about θ in the data were contained in the sufficient statistic.

Example 2.8

Suppose we have a random sample from an exponential distribution with a gamma prior distribution, that is,
 $X_i|\theta \sim \text{Exp}(\theta)$, $i = 1, 2, \dots, n$ (independent) and $\theta \sim \text{Ga}(g, h)$.

Determine a sufficient statistic T for θ and verify that
 $\pi(\theta|\mathbf{x}) = \pi(\theta|t)$.

Solution to Example 2.8 (1/2)

The density of the data is

$$\begin{aligned}f_{\mathbf{X}}(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta e^{-\theta x_i} \\&= \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= 1 \times \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= h(\mathbf{x}) g(\sum x_i, \theta)\end{aligned}$$

and therefore, by the Factorisation Theorem, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Solution to Example 2.8 (1/2)

The density of the data is

$$\begin{aligned}f_{\mathbf{X}}(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta e^{-\theta x_i} \\&= \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= 1 \times \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= h(\mathbf{x}) g(\sum x_i, \theta)\end{aligned}$$

and therefore, by the Factorisation Theorem, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Solution to Example 2.8 (1/2)

The density of the data is

$$\begin{aligned}f_{\mathbf{X}}(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta e^{-\theta x_i} \\&= \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= 1 \times \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= h(\mathbf{x}) g(\sum x_i, \theta)\end{aligned}$$

and therefore, by the Factorisation Theorem, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Solution to Example 2.8 (1/2)

The density of the data is

$$\begin{aligned}f_{\mathbf{X}}(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta e^{-\theta x_i} \\&= \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= 1 \times \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= h(\mathbf{x}) g(\sum x_i, \theta)\end{aligned}$$

and therefore, by the Factorisation Theorem, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Solution to Example 2.8 (1/2)

The density of the data is

$$\begin{aligned}f_{\mathbf{X}}(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta e^{-\theta x_i} \\&= \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= 1 \times \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\&= h(\mathbf{x}) g(\sum x_i, \theta)\end{aligned}$$

and therefore, by the Factorisation Theorem, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Solution to Example 2.8 (2/2)

Now $T|\theta \sim \text{Ga}(n, \theta)$ and so

$$f_T(t|\theta) = \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0.$$

Also

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0.$$

Therefore, by Bayes Theorem,

$$\begin{aligned} \pi(\theta|t) &\propto \pi(\theta)f(t|\theta) \\ &\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0 \\ &\propto \theta^{g+n-1} e^{-(h+t)\theta}, \quad \theta > 0 \end{aligned}$$

and so the posterior distribution is $\theta|t \sim \text{Ga}(g+n, h+t)$. This is the same as the result we obtained previously for $\theta|x$.

Solution to Example 2.8 (2/2)

Now $T|\theta \sim \text{Ga}(n, \theta)$ and so

$$f_T(t|\theta) = \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0.$$

Also

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0.$$

Therefore, by Bayes Theorem,

$$\begin{aligned} \pi(\theta|t) &\propto \pi(\theta)f(t|\theta) \\ &\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0 \\ &\propto \theta^{g+n-1} e^{-(h+t)\theta}, \quad \theta > 0 \end{aligned}$$

and so the posterior distribution is $\theta|t \sim \text{Ga}(g+n, h+t)$. This is the same as the result we obtained previously for $\theta|x$.

Solution to Example 2.8 (2/2)

Now $T|\theta \sim Ga(n, \theta)$ and so

$$f_T(t|\theta) = \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0.$$

Also

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0.$$

Therefore, by Bayes Theorem,

$$\begin{aligned}\pi(\theta|t) &\propto \pi(\theta)f(t|\theta) \\ &\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0 \\ &\propto \theta^{g+n-1} e^{-(h+t)\theta}, \quad \theta > 0\end{aligned}$$

and so the posterior distribution is $\theta|t \sim Ga(g+n, h+t)$. This is the same as the result we obtained previously for $\theta|x$.

Solution to Example 2.8 (2/2)

Now $T|\theta \sim Ga(n, \theta)$ and so

$$f_T(t|\theta) = \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0.$$

Also

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0.$$

Therefore, by Bayes Theorem,

$$\pi(\theta|t) \propto \pi(\theta)f(t|\theta)$$

$$\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0$$

$$\propto \theta^{g+n-1} e^{-(h+t)\theta}, \quad \theta > 0$$

and so the posterior distribution is $\theta|t \sim Ga(g+n, h+t)$. This is the same as the result we obtained previously for $\theta|x$.

Solution to Example 2.8 (2/2)

Now $T|\theta \sim \text{Ga}(n, \theta)$ and so

$$f_T(t|\theta) = \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0.$$

Also

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0.$$

Therefore, by Bayes Theorem,

$$\begin{aligned} \pi(\theta|t) &\propto \pi(\theta)f(t|\theta) \\ &\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0 \\ &\propto \theta^{g+n-1} e^{-(h+t)\theta}, \quad \theta > 0 \end{aligned}$$

and so the posterior distribution is $\theta|t \sim \text{Ga}(g+n, h+t)$. This is the same as the result we obtained previously for $\theta|x$.

Solution to Example 2.8 (2/2)

Now $T|\theta \sim Ga(n, \theta)$ and so

$$f_T(t|\theta) = \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0.$$

Also

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0.$$

Therefore, by Bayes Theorem,

$$\pi(\theta|t) \propto \pi(\theta)f(t|\theta)$$

$$\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0$$

$$\propto \theta^{g+n-1} e^{-(h+t)\theta}, \quad \theta > 0$$

and so the posterior distribution is $\theta|t \sim Ga(g+n, h+t)$. This is the same as the result we obtained previously for $\theta|x$.

Solution to Example 2.8 (2/2)

Now $T|\theta \sim \text{Ga}(n, \theta)$ and so

$$f_T(t|\theta) = \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0.$$

Also

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0.$$

Therefore, by Bayes Theorem,

$$\begin{aligned} \pi(\theta|t) &\propto \pi(\theta)f(t|\theta) \\ &\propto \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)} \times \frac{\theta^n t^{n-1} e^{-\theta t}}{\Gamma(n)}, \quad \theta > 0 \\ &\propto \theta^{g+n-1} e^{-(h+t)\theta}, \quad \theta > 0 \end{aligned}$$

and so the posterior distribution is $\theta|t \sim \text{Ga}(g+n, h+t)$. This is the same as the result we obtained previously for $\theta|\mathbf{x}$.