

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.