

The story so far

■ What is Bayesian Statistics?

■ Review:

- Calculating probabilities: Classical, frequentist and subjective (self-directed reading of Section 1.2)
- Conditional probability and Bayes Theorem – proofs and three examples
- Likelihood

■ Suppose we have a random sample $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ from a *Rayleigh* distribution, with PDF

$$f(x|\sigma) = \frac{x}{\sigma} e^{-x^2/2\sigma^2}, \quad x \geq 0, \quad \sigma > 0.$$

Find the likelihood function for σ . Do not maximise this function, and do not find the maximum likelihood estimate for σ !

The story so far

- What is Bayesian Statistics?
- Review:
 - Calculating probabilities: Classical, frequentist and subjective (self-directed reading of Section 1.2)
 - Conditional probability and Bayes Theorem – proofs and three examples
 - Likelihood
- Suppose we have a random sample $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ from a *Rayleigh* distribution, with PDF

$$f(x|\sigma) = \frac{x}{\sigma} e^{-x^2/2\sigma^2}, \quad x \geq 0, \quad \sigma > 0.$$

Find the likelihood function for σ . Do not maximise this function, and do not find the maximum likelihood estimate for σ !

Consider again the Poisson model in Example 1.5. The likelihood function is

$$\begin{aligned} f(\mathbf{x}|\theta) &= \frac{\theta^{n\bar{x}} e^{-n\theta}}{\prod_{i=1}^n x_i!} \\ &= \left(\prod_{i=1}^n x_i! \right)^{-1} \times \theta^{n\bar{x}} e^{-n\theta}. \end{aligned}$$

Notice that this likelihood function depends on the data only through $\prod_{i=1}^n (x_i!)^{-1}$ and $\bar{x}$.

Further, in $f(\mathbf{x}|\theta)$, θ only “interacts” with $\bar{x}$ — the other term simply scales $f(\mathbf{x}|\theta)$ — so that, for example, the point at which $f(\mathbf{x}|\theta)$ is maximized is determined only by $\bar{x}$.

Consider again the Poisson model in Example 1.5. The likelihood function is

$$\begin{aligned} f(\mathbf{x}|\theta) &= \frac{\theta^{n\bar{x}} e^{-n\theta}}{\prod_{i=1}^n x_i!} \\ &= \left(\prod_{i=1}^n x_i! \right)^{-1} \times \theta^{n\bar{x}} e^{-n\theta}. \end{aligned}$$

Notice that this likelihood function depends on the data only through $\prod_{i=1}^n (x_i!)^{-1}$ and $\bar{x}$.

Further, in $f(\mathbf{x}|\theta)$, θ only “interacts” with $\bar{x}$ — the other term simply scales $f(\mathbf{x}|\theta)$ — so that, for example, the point at which $f(\mathbf{x}|\theta)$ is maximized is determined only by $\bar{x}$.

Consider again the Poisson model in Example 1.5. The likelihood function is

$$\begin{aligned} f(\mathbf{x}|\theta) &= \frac{\theta^{n\bar{x}} e^{-n\theta}}{\prod_{i=1}^n x_i!} \\ &= \left(\prod_{i=1}^n x_i! \right)^{-1} \times \theta^{n\bar{x}} e^{-n\theta}. \end{aligned}$$

Notice that this likelihood function depends on the data only through $\prod_{i=1}^n (x_i!)^{-1}$ and $\bar{x}$.

Further, in $f(\mathbf{x}|\theta)$, θ only “interacts” with $\bar{x}$ — the other term simply scales $f(\mathbf{x}|\theta)$ — so that, for example, the point at which $f(\mathbf{x}|\theta)$ is maximized is determined only by $\bar{x}$.

Consider again the Poisson model in Example 1.5. The likelihood function is

$$\begin{aligned} f(\mathbf{x}|\theta) &= \frac{\theta^{n\bar{x}} e^{-n\theta}}{\prod_{i=1}^n x_i!} \\ &= \left(\prod_{i=1}^n x_i! \right)^{-1} \times \theta^{n\bar{x}} e^{-n\theta}. \end{aligned}$$

Notice that this likelihood function depends on the data only through $\prod_{i=1}^n (x_i!)^{-1}$ and $\bar{x}$.

Further, in $f(\mathbf{x}|\theta)$, θ only “interacts” with $\bar{x}$ — the other term simply scales $f(\mathbf{x}|\theta)$ — so that, for example, the point at which $f(\mathbf{x}|\theta)$ is maximized is determined only by $\bar{x}$.

Consider again the Poisson model in Example 1.5. The likelihood function is

$$\begin{aligned} f(\mathbf{x}|\theta) &= \frac{\theta^{n\bar{x}} e^{-n\theta}}{\prod_{i=1}^n x_i!} \\ &= \left(\prod_{i=1}^n x_i! \right)^{-1} \times \theta^{n\bar{x}} e^{-n\theta}. \end{aligned}$$

Notice that this likelihood function depends on the data only through $\prod_{i=1}^n (x_i!)^{-1}$ and $\bar{x}$.

Further, in $f(\mathbf{x}|\theta)$, θ only “interacts” with $\bar{x}$ — the other term simply scales $f(\mathbf{x}|\theta)$ — so that, for example, the point at which $f(\mathbf{x}|\theta)$ is maximized is determined only by $\bar{x}$.

Consider again the Poisson model in Example 1.5. The likelihood function is

$$\begin{aligned} f(\mathbf{x}|\theta) &= \frac{\theta^{n\bar{x}} e^{-n\theta}}{\prod_{i=1}^n x_i!} \\ &= \left(\prod_{i=1}^n x_i! \right)^{-1} \times \theta^{n\bar{x}} e^{-n\theta}. \end{aligned}$$

Notice that this likelihood function depends on the data only through $\prod_{i=1}^n (x_i!)^{-1}$ and $\bar{x}$.

Further, in $f(\mathbf{x}|\theta)$, θ only “interacts” with $\bar{x}$ — the other term simply scales $f(\mathbf{x}|\theta)$ — so that, for example, the point at which $f(\mathbf{x}|\theta)$ is maximized is determined only by $\bar{x}$.

Informally, we think of all the information about θ in the data being contained in $\bar{x}$.

More formally, we can show that the distribution of the data given the value $\bar{x}$ does not depend on θ .

Informally, we think of all the information about θ in the data being contained in $\bar{x}$.

More formally, we can show that the distribution of the data given the value $\bar{x}$ does not depend on θ .

Definition (Statistic)

A **statistic** is any function of the data (and not of unknown parameters).

Definition (Sufficiency)

The statistic $T(\mathbf{X})$ is **sufficient** for θ if $f(\mathbf{x} | T(\mathbf{X}) = t)$ does not depend on θ .

Definition (Statistic)

A **statistic** is any function of the data (and not of unknown parameters).

Definition (Sufficiency)

The statistic $T(\mathbf{X})$ is **sufficient** for θ if $f(\mathbf{x} | T(\mathbf{X}) = t)$ does not depend on θ .

Example 1.8

Consider again the Poisson model in Example 1.5.

Suppose we had just two observations.

Then $n = 2$ and $X_i|\theta \sim Po(\theta)$, $i = 1, 2$ (independent).

Show that $T = X_1 + X_2$ is sufficient for θ . Note that $T|\theta \sim Po(2\theta)$.

Example 1.8

Consider again the Poisson model in Example 1.5.

Suppose we had just two observations.

Then $n = 2$ and $X_i|\theta \sim \text{Po}(\theta)$, $i = 1, 2$ (independent).

Show that $T = X_1 + X_2$ is sufficient for θ . Note that $T|\theta \sim \text{Po}(2\theta)$.

Example 1.8

Consider again the Poisson model in Example 1.5.

Suppose we had just two observations.

Then $n = 2$ and $X_i|\theta \sim Po(\theta)$, $i = 1, 2$ (independent).

Show that $T = X_1 + X_2$ is sufficient for θ . Note that $T|\theta \sim Po(2\theta)$.

Example 1.8

Consider again the Poisson model in Example 1.5.

Suppose we had just two observations.

Then $n = 2$ and $X_i|\theta \sim Po(\theta)$, $i = 1, 2$ (independent).

Show that $T = X_1 + X_2$ is sufficient for θ . Note that $T|\theta \sim Po(2\theta)$.

Solution to Example 1.8 (1/2)

Consider the conditional distribution of the data $\mathbf{X}$ on $T = t$.

As X_1 and X_2 are discrete random variables, we have

$$\begin{aligned}f(\mathbf{x}|T(\mathbf{X}) = t) &= \frac{\Pr(X_1 = x_1, X_2 = x_2, T = t)}{\Pr(T = t)} \\&= \frac{\Pr(X_1 = x_1, X_2 = t - x_1)}{\Pr(T = t)}.\end{aligned}$$

Now, as X_1 and X_2 are independent, for the numerator we have:

$$\begin{aligned}\Pr(X_1 = x_1, X_2 = t - x_1|\theta) &= \Pr(X_1 = x_1|\theta)\Pr(X_2 = t - x_1|\theta) \\&= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{t-x_1}}{(t-x_1)!} \\&= \frac{e^{-2\theta}\theta^{x_1+t-x_1}}{x_1!(t-x_1)!} = \frac{e^{-2\theta}\theta^t}{x_1!(t-x_1)!}.\end{aligned}$$

Solution to Example 1.8 (1/2)

Consider the conditional distribution of the data $\mathbf{X}$ on $T = t$.

As X_1 and X_2 are discrete random variables, we have

$$\begin{aligned}f(\mathbf{x}|T(\mathbf{X}) = t) &= \frac{\Pr(X_1 = x_1, X_2 = x_2, T = t)}{\Pr(T = t)} \\&= \frac{\Pr(X_1 = x_1, X_2 = t - x_1)}{\Pr(T = t)}.\end{aligned}$$

Now, as X_1 and X_2 are independent, for the numerator we have:

$$\begin{aligned}\Pr(X_1 = x_1, X_2 = t - x_1|\theta) &= \Pr(X_1 = x_1|\theta)\Pr(X_2 = t - x_1|\theta) \\&= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{t-x_1}}{(t-x_1)!} \\&= \frac{e^{-2\theta}\theta^{x_1+t-x_1}}{x_1!(t-x_1)!} = \frac{e^{-2\theta}\theta^t}{x_1!(t-x_1)!}.\end{aligned}$$

Solution to Example 1.8 (1/2)

Consider the conditional distribution of the data $\mathbf{X}$ on $T = t$.

As X_1 and X_2 are discrete random variables, we have

$$\begin{aligned}f(\mathbf{x}|T(\mathbf{X}) = t) &= \frac{\Pr(X_1 = x_1, X_2 = x_2, T = t)}{\Pr(T = t)} \\&= \frac{\Pr(X_1 = x_1, X_2 = t - x_1)}{\Pr(T = t)}.\end{aligned}$$

Now, as X_1 and X_2 are independent, for the numerator we have:

$$\begin{aligned}\Pr(X_1 = x_1, X_2 = t - x_1|\theta) &= \Pr(X_1 = x_1|\theta)\Pr(X_2 = t - x_1|\theta) \\&= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{t-x_1}}{(t-x_1)!} \\&= \frac{e^{-2\theta}\theta^{x_1+t-x_1}}{x_1!(t-x_1)!} = \frac{e^{-2\theta}\theta^t}{x_1!(t-x_1)!}.\end{aligned}$$

Solution to Example 1.8 (1/2)

Consider the conditional distribution of the data $\mathbf{X}$ on $T = t$.

As X_1 and X_2 are discrete random variables, we have

$$\begin{aligned}f(\mathbf{x}|T(\mathbf{X}) = t) &= \frac{\Pr(X_1 = x_1, X_2 = x_2, T = t)}{\Pr(T = t)} \\&= \frac{\Pr(X_1 = x_1, X_2 = t - x_1)}{\Pr(T = t)}.\end{aligned}$$

Now, as X_1 and X_2 are independent, for the numerator we have:

$$\begin{aligned}\Pr(X_1 = x_1, X_2 = t - x_1|\theta) &= \Pr(X_1 = x_1|\theta)\Pr(X_2 = t - x_1|\theta) \\&= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{t-x_1}}{(t-x_1)!} \\&= \frac{e^{-2\theta}\theta^{x_1+t-x_1}}{x_1!(t-x_1)!} = \frac{e^{-2\theta}\theta^t}{x_1!(t-x_1)!}.\end{aligned}$$

Solution to Example 1.8 (1/2)

Consider the conditional distribution of the data $\mathbf{X}$ on $T = t$.

As X_1 and X_2 are discrete random variables, we have

$$\begin{aligned}f(\mathbf{x}|T(\mathbf{X}) = t) &= \frac{\Pr(X_1 = x_1, X_2 = x_2, T = t)}{\Pr(T = t)} \\&= \frac{\Pr(X_1 = x_1, X_2 = t - x_1)}{\Pr(T = t)}.\end{aligned}$$

Now, as X_1 and X_2 are independent, for the numerator we have:

$$\begin{aligned}\Pr(X_1 = x_1, X_2 = t - x_1|\theta) &= \Pr(X_1 = x_1|\theta)\Pr(X_2 = t - x_1|\theta) \\&= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{t-x_1}}{(t-x_1)!} \\&= \frac{e^{-2\theta}\theta^{x_1+t-x_1}}{x_1!(t-x_1)!} = \frac{e^{-2\theta}\theta^t}{x_1!(t-x_1)!}.\end{aligned}$$

Solution to Example 1.8 (1/2)

Consider the conditional distribution of the data $\mathbf{X}$ on $T = t$.

As X_1 and X_2 are discrete random variables, we have

$$\begin{aligned}f(\mathbf{x}|T(\mathbf{X}) = t) &= \frac{\Pr(X_1 = x_1, X_2 = x_2, T = t)}{\Pr(T = t)} \\&= \frac{\Pr(X_1 = x_1, X_2 = t - x_1)}{\Pr(T = t)}.\end{aligned}$$

Now, as X_1 and X_2 are independent, for the numerator we have:

$$\begin{aligned}\Pr(X_1 = x_1, X_2 = t - x_1|\theta) &= \Pr(X_1 = x_1|\theta)\Pr(X_2 = t - x_1|\theta) \\&= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{t-x_1}}{(t-x_1)!} \\&= \frac{e^{-2\theta}\theta^{x_1+t-x_1}}{x_1!(t-x_1)!} = \frac{e^{-2\theta}\theta^t}{x_1!(t-x_1)!}.\end{aligned}$$

Solution to Example 1.8 (1/2)

Consider the conditional distribution of the data $\mathbf{X}$ on $T = t$.

As X_1 and X_2 are discrete random variables, we have

$$\begin{aligned}f(\mathbf{x}|T(\mathbf{X}) = t) &= \frac{\Pr(X_1 = x_1, X_2 = x_2, T = t)}{\Pr(T = t)} \\&= \frac{\Pr(X_1 = x_1, X_2 = t - x_1)}{\Pr(T = t)}.\end{aligned}$$

Now, as X_1 and X_2 are independent, for the numerator we have:

$$\begin{aligned}\Pr(X_1 = x_1, X_2 = t - x_1|\theta) &= \Pr(X_1 = x_1|\theta)\Pr(X_2 = t - x_1|\theta) \\&= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{t-x_1}}{(t-x_1)!} \\&= \frac{e^{-2\theta}\theta^{x_1+t-x_1}}{x_1!(t-x_1)!} = \frac{e^{-2\theta}\theta^t}{x_1!(t-x_1)!}.\end{aligned}$$

Solution to Example 1.8 (1/2)

Consider the conditional distribution of the data $\mathbf{X}$ on $T = t$.

As X_1 and X_2 are discrete random variables, we have

$$\begin{aligned}f(\mathbf{x}|T(\mathbf{X}) = t) &= \frac{\Pr(X_1 = x_1, X_2 = x_2, T = t)}{\Pr(T = t)} \\&= \frac{\Pr(X_1 = x_1, X_2 = t - x_1)}{\Pr(T = t)}.\end{aligned}$$

Now, as X_1 and X_2 are independent, for the numerator we have:

$$\begin{aligned}\Pr(X_1 = x_1, X_2 = t - x_1|\theta) &= \Pr(X_1 = x_1|\theta)\Pr(X_2 = t - x_1|\theta) \\&= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{t-x_1}}{(t-x_1)!} \\&= \frac{e^{-2\theta}\theta^{x_1+t-x_1}}{x_1!(t-x_1)!} = \frac{e^{-2\theta}\theta^t}{x_1!(t-x_1)!}.\end{aligned}$$

Solution to Example 1.8 (1/2)

Consider the conditional distribution of the data $\mathbf{X}$ on $T = t$.

As X_1 and X_2 are discrete random variables, we have

$$\begin{aligned}f(\mathbf{x}|T(\mathbf{X}) = t) &= \frac{\Pr(X_1 = x_1, X_2 = x_2, T = t)}{\Pr(T = t)} \\&= \frac{\Pr(X_1 = x_1, X_2 = t - x_1)}{\Pr(T = t)}.\end{aligned}$$

Now, as X_1 and X_2 are independent, for the numerator we have:

$$\begin{aligned}\Pr(X_1 = x_1, X_2 = t - x_1|\theta) &= \Pr(X_1 = x_1|\theta)\Pr(X_2 = t - x_1|\theta) \\&= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{t-x_1}}{(t-x_1)!} \\&= \frac{e^{-2\theta}\theta^{x_1+t-x_1}}{x_1!(t-x_1)!} = \frac{e^{-2\theta}\theta^t}{x_1!(t-x_1)!}.\end{aligned}$$

Solution to Example 1.8 (2/2)

For the denominator we know that

$$\Pr(T = t|\theta) = \frac{e^{-2\theta}(2\theta)^t}{t!} \quad (\text{sum of two Poissons}).$$

Therefore

$$\begin{aligned} f(\mathbf{x}|T(\mathbf{X}) = t) &= \frac{e^{-2\theta}\theta^t}{x_1!(t-x_1)!} \bigg/ \frac{e^{-2\theta}(2\theta)^t}{t!} \\ &= \frac{t!e^{-2\theta}\theta^t}{x_1!(t-x_1)!e^{-2\theta}(2\theta)^t} \\ &= \frac{t!}{2^t x_1!(t-x_1)!} \end{aligned}$$

which does not depend on θ . Hence, by Definition 1.4, $T = X_1 + X_2$ is sufficient for θ .

Solution to Example 1.8 (2/2)

For the denominator we know that

$$\Pr(T = t|\theta) = \frac{e^{-2\theta}(2\theta)^t}{t!} \quad (\text{sum of two Poissons}).$$

Therefore

$$\begin{aligned} f(\mathbf{x}|T(\mathbf{X}) = t) &= \frac{e^{-2\theta}\theta^t}{x_1!(t-x_1)!} \bigg/ \frac{e^{-2\theta}(2\theta)^t}{t!} \\ &= \frac{t!e^{-2\theta}\theta^t}{x_1!(t-x_1)!e^{-2\theta}(2\theta)^t} \\ &= \frac{t!}{2^t x_1!(t-x_1)!} \end{aligned}$$

which does not depend on θ . Hence, by Definition 1.4, $T = X_1 + X_2$ is sufficient for θ .

Solution to Example 1.8 (2/2)

For the denominator we know that

$$\Pr(T = t|\theta) = \frac{e^{-2\theta}(2\theta)^t}{t!} \quad (\text{sum of two Poissons}).$$

Therefore

$$\begin{aligned} f(\mathbf{x}|T(\mathbf{X}) = t) &= \frac{e^{-2\theta}\theta^t}{x_1!(t-x_1)!} \bigg/ \frac{e^{-2\theta}(2\theta)^t}{t!} \\ &= \frac{t!e^{-2\theta}\theta^t}{x_1!(t-x_1)!e^{-2\theta}(2\theta)^t} \\ &= \frac{t!}{2^t x_1!(t-x_1)!} \end{aligned}$$

which does not depend on θ . Hence, by Definition 1.4, $T = X_1 + X_2$ is sufficient for θ .

Solution to Example 1.8 (2/2)

For the denominator we know that

$$\Pr(T = t|\theta) = \frac{e^{-2\theta}(2\theta)^t}{t!} \quad (\text{sum of two Poissons}).$$

Therefore

$$\begin{aligned} f(\mathbf{x}|T(\mathbf{X}) = t) &= \frac{e^{-2\theta}\theta^t}{x_1!(t-x_1)!} \bigg/ \frac{e^{-2\theta}(2\theta)^t}{t!} \\ &= \frac{t!e^{-2\theta}\theta^t}{x_1!(t-x_1)!e^{-2\theta}(2\theta)^t} \\ &= \frac{t!}{2^t x_1!(t-x_1)!} \end{aligned}$$

which does not depend on θ . Hence, by Definition 1.4, $T = X_1 + X_2$ is sufficient for θ .

Solution to Example 1.8 (2/2)

For the denominator we know that

$$\Pr(T = t|\theta) = \frac{e^{-2\theta}(2\theta)^t}{t!} \quad (\text{sum of two Poissons}).$$

Therefore

$$\begin{aligned} f(\mathbf{x}|T(\mathbf{X}) = t) &= \frac{e^{-2\theta}\theta^t}{x_1!(t-x_1)!} \bigg/ \frac{e^{-2\theta}(2\theta)^t}{t!} \\ &= \frac{t!e^{-2\theta}\theta^t}{x_1!(t-x_1)!e^{-2\theta}(2\theta)^t} \\ &= \frac{t!}{2^t x_1!(t-x_1)!} \end{aligned}$$

which does not depend on θ . Hence, by Definition 1.4, $T = X_1 + X_2$ is sufficient for θ .

Definition (Joint sufficiency)

The statistics $\underline{T}(\mathbf{X}) = (T_1(\mathbf{X}), T_2(\mathbf{X}), \dots, T_k(\mathbf{X}))^T$ are **jointly sufficient** for θ if $f(\mathbf{x} | \underline{T}(\mathbf{X}) = \mathbf{t})$ does not depend on θ .

Theorem (Factorisation Theorem)

Under certain regularity conditions

$$\underline{T}(\mathbf{X}) \text{ is sufficient for } \theta \iff f(\mathbf{x} | \theta) = h(\mathbf{x}) g(\underline{t}(\mathbf{x}), \theta)$$

for some functions h and g .

Definition (Joint sufficiency)

The statistics $\underline{T}(\mathbf{X}) = (T_1(\mathbf{X}), T_2(\mathbf{X}), \dots, T_k(\mathbf{X}))^T$ are **jointly sufficient** for θ if $f(\mathbf{x} | \underline{T}(\mathbf{X}) = \mathbf{t})$ does not depend on θ .

Theorem (Factorisation Theorem)

Under certain regularity conditions

$$\underline{T}(\mathbf{X}) \text{ is sufficient for } \theta \iff f(\mathbf{x} | \theta) = h(\mathbf{x}) g(\underline{t}(\mathbf{x}), \theta)$$

for some functions h and g .

Example 1.9

Consider again the Poisson model in Example 1.5 with $n = 2$: $X_i|\theta \sim Po(\theta)$, $i = 1, 2$ (independent). Determine a sufficient statistic for θ .

Solution to Example 1.9 (1/1)

The (joint) probability function is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{x_2}}{x_2!} \\&= \frac{e^{-2\theta}\theta^{x_1+x_2}}{x_1!x_2!} \\&= \frac{1}{x_1!x_2!} \times e^{-2\theta}\theta^{x_1+x_2} \\&= h(\mathbf{x}) g(x_1 + x_2, \theta),\end{aligned}$$

where $h(\mathbf{x}) = 1/(x_1!x_2!)$ and $g(t, \theta) = e^{-2\theta}\theta^t$.

Therefore, by the Factorisation Theorem, $T = X_1 + X_2$ is sufficient for θ .

Solution to Example 1.9 (1/1)

The (joint) probability function is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{x_2}}{x_2!} \\&= \frac{e^{-2\theta}\theta^{x_1+x_2}}{x_1!x_2!} \\&= \frac{1}{x_1!x_2!} \times e^{-2\theta}\theta^{x_1+x_2} \\&= h(\mathbf{x}) g(x_1 + x_2, \theta),\end{aligned}$$

where $h(\mathbf{x}) = 1/(x_1!x_2!)$ and $g(t, \theta) = e^{-2\theta}\theta^t$.

Therefore, by the Factorisation Theorem, $T = X_1 + X_2$ is sufficient for θ .

Solution to Example 1.9 (1/1)

The (joint) probability function is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{x_2}}{x_2!} \\&= \frac{e^{-2\theta}\theta^{x_1+x_2}}{x_1!x_2!} \\&= \frac{1}{x_1!x_2!} \times e^{-2\theta}\theta^{x_1+x_2} \\&= h(\mathbf{x}) g(x_1 + x_2, \theta),\end{aligned}$$

where $h(\mathbf{x}) = 1/(x_1!x_2!)$ and $g(t, \theta) = e^{-2\theta}\theta^t$.

Therefore, by the Factorisation Theorem, $T = X_1 + X_2$ is sufficient for θ .

Solution to Example 1.9 (1/1)

The (joint) probability function is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{x_2}}{x_2!} \\&= \frac{e^{-2\theta}\theta^{x_1+x_2}}{x_1!x_2!} \\&= \frac{1}{x_1!x_2!} \times e^{-2\theta}\theta^{x_1+x_2} \\&= h(\mathbf{x})g(x_1 + x_2, \theta),\end{aligned}$$

where $h(\mathbf{x}) = 1/(x_1!x_2!)$ and $g(t, \theta) = e^{-2\theta}\theta^t$.

Therefore, by the Factorisation Theorem, $T = X_1 + X_2$ is sufficient for θ .

Solution to Example 1.9 (1/1)

The (joint) probability function is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{x_2}}{x_2!} \\&= \frac{e^{-2\theta}\theta^{x_1+x_2}}{x_1!x_2!} \\&= \frac{1}{x_1!x_2!} \times e^{-2\theta}\theta^{x_1+x_2} \\&= h(\mathbf{x})g(x_1 + x_2, \theta),\end{aligned}$$

where $h(\mathbf{x}) = 1/(x_1!x_2!)$ and $g(t, \theta) = e^{-2\theta}\theta^t$.

Therefore, by the Factorisation Theorem, $T = X_1 + X_2$ is sufficient for θ .

Solution to Example 1.9 (1/1)

The (joint) probability function is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \frac{e^{-\theta}\theta^{x_1}}{x_1!} \times \frac{e^{-\theta}\theta^{x_2}}{x_2!} \\&= \frac{e^{-2\theta}\theta^{x_1+x_2}}{x_1!x_2!} \\&= \frac{1}{x_1!x_2!} \times e^{-2\theta}\theta^{x_1+x_2} \\&= h(\mathbf{x}) g(x_1 + x_2, \theta),\end{aligned}$$

where $h(\mathbf{x}) = 1/(x_1!x_2!)$ and $g(t, \theta) = e^{-2\theta}\theta^t$.

Therefore, by the Factorisation Theorem, $T = X_1 + X_2$ is sufficient for θ .

Example 1.10

Suppose we have a random sample from an exponential distribution: $X_i|\theta \sim \text{Exp}(\theta)$, $i = 1, 2, \dots, n$ (independent).

Determine a sufficient statistic for θ .

Solution to Example 1.10 (1/1)

The (joint) probability density function is

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

where $h(\mathbf{x}) = 1$ and $g(t, \theta) = \theta^n e^{-\theta t}$. Therefore, by the Factorisation Theorem, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Solution to Example 1.10 (1/1)

The (joint) probability density function is

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

where $h(\mathbf{x}) = 1$ and $g(t, \theta) = \theta^n e^{-\theta t}$. Therefore, by the Factorisation Theorem, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Solution to Example 1.10 (1/1)

The (joint) probability density function is

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

where $h(\mathbf{x}) = 1$ and $g(t, \theta) = \theta^n e^{-\theta t}$. Therefore, by the Factorisation Theorem, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Solution to Example 1.10 (1/1)

The (joint) probability density function is

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

where $h(\mathbf{x}) = 1$ and $g(t, \theta) = \theta^n e^{-\theta t}$. Therefore, by the Factorisation Theorem, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Solution to Example 1.10 (1/1)

The (joint) probability density function is

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

where $h(\mathbf{x}) = 1$ and $g(t, \theta) = \theta^n e^{-\theta t}$. Therefore, by the Factorisation Theorem, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Definition (Joint sufficiency)

The statistics $\underline{T}(\mathbf{X}) = (T_1(\mathbf{X}), T_2(\mathbf{X}), \dots, T_k(\mathbf{X}))^T$ are **jointly sufficient** for θ if $f(\mathbf{x} | \underline{T}(\mathbf{X}) = \mathbf{t})$ does not depend on θ .

Theorem (Factorisation Theorem)

Under certain regularity conditions

$$\underline{T}(\mathbf{X}) \text{ is sufficient for } \theta \iff f(\mathbf{x} | \theta) = h(\mathbf{x}) g(\underline{t}(\mathbf{x}), \theta)$$

for some functions h and g .

Definition (Joint sufficiency)

The statistics $\underline{T}(\mathbf{X}) = (T_1(\mathbf{X}), T_2(\mathbf{X}), \dots, T_k(\mathbf{X}))^T$ are **jointly sufficient** for θ if $f(\mathbf{x} | \underline{T}(\mathbf{X}) = \mathbf{t})$ does not depend on θ .

Theorem (Factorisation Theorem)

Under certain regularity conditions

$$\underline{T}(\mathbf{X}) \text{ is sufficient for } \theta \iff f(\mathbf{x} | \theta) = h(\mathbf{x}) g(\underline{t}(\mathbf{x}), \theta)$$

for some functions h and g .

Example 1.11

Suppose we have a random sample from a Normal distribution:
 $X_i | \mu, \sigma \sim N(\mu, \sigma^2), i = 1, 2, \dots, n$ (independent).

Determine sufficient statistics for (μ, σ) .

Solution to Example 1.11 (1/1)

The (joint) probability density function is

$$\begin{aligned}f(\mathbf{x}|\mu, \sigma) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x_i - \mu)^2}{2\sigma^2} \right\} \\&= (2\pi)^{-n/2} \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right\} \\&= (2\pi)^{-n/2} \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \left(\sum x_i^2 - 2\mu \sum x_i + n\mu^2 \right) \right\} \\&= h(\mathbf{x}) g(\sum x_i, \sum x_i^2, \mu, \sigma),\end{aligned}$$

where $h(\mathbf{x}) = (2\pi)^{-n/2}$ and

$g(t_1, t_2, \mu, \sigma) = \sigma^{-n} \exp\{-(t_2 - 2\mu t_1 + n\mu^2)/(2\sigma^2)\}$.

Therefore, by the Factorisation Theorem, $T_1 = \sum_{i=1}^n X_i$ and $T_2 = \sum_{i=1}^n X_i^2$ are (jointly) sufficient for μ and σ .

Solution to Example 1.11 (1/1)

The (joint) probability density function is

$$\begin{aligned}f(\mathbf{x}|\mu, \sigma) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x_i - \mu)^2}{2\sigma^2} \right\} \\&= (2\pi)^{-n/2} \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right\} \\&= (2\pi)^{-n/2} \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \left(\sum x_i^2 - 2\mu \sum x_i + n\mu^2 \right) \right\} \\&= h(\mathbf{x}) g(\sum x_i, \sum x_i^2, \mu, \sigma),\end{aligned}$$

where $h(\mathbf{x}) = (2\pi)^{-n/2}$ and

$g(t_1, t_2, \mu, \sigma) = \sigma^{-n} \exp\{-(t_2 - 2\mu t_1 + n\mu^2)/(2\sigma^2)\}$.

Therefore, by the Factorisation Theorem, $T_1 = \sum_{i=1}^n X_i$ and $T_2 = \sum_{i=1}^n X_i^2$ are (jointly) sufficient for μ and σ .

Solution to Example 1.11 (1/1)

The (joint) probability density function is

$$\begin{aligned}f(\mathbf{x}|\mu, \sigma) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x_i - \mu)^2}{2\sigma^2} \right\} \\&= (2\pi)^{-n/2} \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right\} \\&= (2\pi)^{-n/2} \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \left(\sum x_i^2 - 2\mu \sum x_i + n\mu^2 \right) \right\} \\&= h(\mathbf{x}) g(\sum x_i, \sum x_i^2, \mu, \sigma),\end{aligned}$$

where $h(\mathbf{x}) = (2\pi)^{-n/2}$ and

$g(t_1, t_2, \mu, \sigma) = \sigma^{-n} \exp\{-(t_2 - 2\mu t_1 + n\mu^2)/(2\sigma^2)\}$.

Therefore, by the Factorisation Theorem, $T_1 = \sum_{i=1}^n X_i$ and $T_2 = \sum_{i=1}^n X_i^2$ are (jointly) sufficient for μ and σ .

Solution to Example 1.11 (1/1)

The (joint) probability density function is

$$\begin{aligned}f(\mathbf{x}|\mu, \sigma) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x_i - \mu)^2}{2\sigma^2} \right\} \\&= (2\pi)^{-n/2} \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right\} \\&= (2\pi)^{-n/2} \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \left(\sum x_i^2 - 2\mu \sum x_i + n\mu^2 \right) \right\} \\&= h(\mathbf{x}) g(\sum x_i, \sum x_i^2, \mu, \sigma),\end{aligned}$$

where $h(\mathbf{x}) = (2\pi)^{-n/2}$ and

$g(t_1, t_2, \mu, \sigma) = \sigma^{-n} \exp\{-(t_2 - 2\mu t_1 + n\mu^2)/(2\sigma^2)\}$.

Therefore, by the Factorisation Theorem, $T_1 = \sum_{i=1}^n X_i$ and $T_2 = \sum_{i=1}^n X_i^2$ are (jointly) sufficient for μ and σ .

Solution to Example 1.11 (1/1)

The (joint) probability density function is

$$\begin{aligned}f(\mathbf{x}|\mu, \sigma) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x_i - \mu)^2}{2\sigma^2} \right\} \\&= (2\pi)^{-n/2} \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right\} \\&= (2\pi)^{-n/2} \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \left(\sum x_i^2 - 2\mu \sum x_i + n\mu^2 \right) \right\} \\&= h(\mathbf{x}) g(\sum x_i, \sum x_i^2, \mu, \sigma),\end{aligned}$$

where $h(\mathbf{x}) = (2\pi)^{-n/2}$ and

$$g(t_1, t_2, \mu, \sigma) = \sigma^{-n} \exp \{ -(t_2 - 2\mu t_1 + n\mu^2) / (2\sigma^2) \}.$$

Therefore, by the Factorisation Theorem, $T_1 = \sum_{i=1}^n X_i$ and $T_2 = \sum_{i=1}^n X_i^2$ are (jointly) sufficient for μ and σ .

If $X_1, X_2, \dots, X_n$ are independent random variables with probability density function

$$f(x|\theta) = \theta(1 - \theta)^{x-1}, \quad x = 1, 2, \dots,$$

- (a) Form the likelihood function $f(\mathbf{x}|\theta)$;
- (b) Use the Factorisation Theorem to obtain a sufficient statistic for θ .

Quick quiz: solution

The joint probability function is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta(1 - \theta)^{x_i-1} \\&= \theta^n(1 - \theta)^{x_1-1}(1 - \theta)^{x_2-1} \times \dots \times (1 - \theta)^{x_n-1} \\&= \theta^n(1 - \theta)^{\sum x_i - n} \\&= 1 \times \theta^n(1 - \theta)^{\sum x_i - n} \\&= h(\mathbf{x}) g\left(\sum x_i, \theta\right)\end{aligned}$$

where $h(\mathbf{x}) = 1$ and $g(t, \theta) = \theta^n(1 - \theta)^{t-n}$. Therefore, by the Factorisation Theorem, $T = \sum X_i$ is sufficient for θ .

Quick quiz: solution

The joint probability function is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta(1 - \theta)^{x_i-1} \\&= \theta^n(1 - \theta)^{x_1-1}(1 - \theta)^{x_2-1} \times \dots \times (1 - \theta)^{x_n-1} \\&= \theta^n(1 - \theta)^{\sum x_i - n} \\&= 1 \times \theta^n(1 - \theta)^{\sum x_i - n} \\&= h(\mathbf{x}) g\left(\sum x_i, \theta\right)\end{aligned}$$

where $h(\mathbf{x}) = 1$ and $g(t, \theta) = \theta^n(1 - \theta)^{t-n}$. Therefore, by the Factorisation Theorem, $T = \sum X_i$ is sufficient for θ .

Quick quiz: solution

The joint probability function is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta(1 - \theta)^{x_i-1} \\&= \theta^n(1 - \theta)^{x_1-1}(1 - \theta)^{x_2-1} \times \dots \times (1 - \theta)^{x_n-1} \\&= \theta^n(1 - \theta)^{\sum x_i - n} \\&= 1 \times \theta^n(1 - \theta)^{\sum x_i - n} \\&= h(\mathbf{x}) g\left(\sum x_i, \theta\right)\end{aligned}$$

where $h(\mathbf{x}) = 1$ and $g(t, \theta) = \theta^n(1 - \theta)^{t-n}$. Therefore, by the Factorisation Theorem, $T = \sum X_i$ is sufficient for θ .

Quick quiz: solution

The joint probability function is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta(1 - \theta)^{x_i-1} \\&= \theta^n(1 - \theta)^{x_1-1}(1 - \theta)^{x_2-1} \times \dots \times (1 - \theta)^{x_n-1} \\&= \theta^n(1 - \theta)^{\sum x_i - n} \\&= 1 \times \theta^n(1 - \theta)^{\sum x_i - n} \\&= h(\mathbf{x}) g\left(\sum x_i, \theta\right)\end{aligned}$$

where $h(\mathbf{x}) = 1$ and $g(t, \theta) = \theta^n(1 - \theta)^{t-n}$. Therefore, by the Factorisation Theorem, $T = \sum X_i$ is sufficient for θ .

Quick quiz: solution

The joint probability function is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta(1 - \theta)^{x_i-1} \\&= \theta^n(1 - \theta)^{x_1-1}(1 - \theta)^{x_2-1} \times \dots \times (1 - \theta)^{x_n-1} \\&= \theta^n(1 - \theta)^{\sum x_i - n} \\&= 1 \times \theta^n(1 - \theta)^{\sum x_i - n} \\&= h(\mathbf{x}) g\left(\sum x_i, \theta\right)\end{aligned}$$

where $h(\mathbf{x}) = 1$ and $g(t, \theta) = \theta^n(1 - \theta)^{t-n}$. Therefore, by the Factorisation Theorem, $T = \sum X_i$ is sufficient for θ .

Quick quiz: solution

The joint probability function is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta(1 - \theta)^{x_i-1} \\&= \theta^n(1 - \theta)^{x_1-1}(1 - \theta)^{x_2-1} \times \dots \times (1 - \theta)^{x_n-1} \\&= \theta^n(1 - \theta)^{\sum x_i - n} \\&= 1 \times \theta^n(1 - \theta)^{\sum x_i - n} \\&= h(\mathbf{x}) g\left(\sum x_i, \theta\right)\end{aligned}$$

where $h(\mathbf{x}) = 1$ and $g(t, \theta) = \theta^n(1 - \theta)^{t-n}$. Therefore, by the Factorisation Theorem, $T = \sum X_i$ is sufficient for θ .