

Exercise 1

A statistic $T = T(X)$ is a sufficient statistic for θ
if it can be computed without knowing the value of θ T F

Exercise 1

A statistic $T = T(X)$ is a sufficient statistic for θ
if it can be computed without knowing the value of θ T F

Exercise 2

Let $(X_1, \dots, X_n)$ be independent identically distributed random variables from the Cauchy distribution, with p.d.f.

$$f(x) = \frac{1}{\pi(1 + (x - \theta)^2)}$$

- 1 Find a sufficient statistics for θ

Exercise 2

$$f(x) = \frac{1}{\pi(1 + (x - \theta)^2)}$$
$$f(x_1, \dots, x_n) = \prod_i \frac{1}{\pi(1 + (x_i - \theta)^2)}$$

- 1 Find a sufficient statistics for θ

Exercise 2

$$f(x_1, \dots, x_n) = \prod_i \frac{1}{\pi(1 + (x_i - \theta)^2)}$$

Find a sufficient statistics for θ

Exercise 2

$$f(x_1, \dots, x_n) = \prod_i \frac{1}{\pi(1 + (x_i - \theta)^2)}$$

Find a sufficient statistics for θ

$(X_1, \dots, X_n)$ is a sufficient statistics for θ

Exercise 2

$$f(x_1, \dots, x_n) = \prod_i \frac{1}{\pi(1 + (x_i - \theta)^2)}$$

Find a sufficient statistics for θ

$(X_1, \dots, X_n)$ is a sufficient statistics for θ

Exercise 3

Let $(X_1, \dots, X_n)$ be a random sample of i.i.d. random variables distributed as follows:

$$f(x; \theta) = \frac{\theta 2^\theta}{x^{\theta+1}} \quad x > 2$$

Exercise 3

$$f(x_1, \dots, x_n; \theta) = \prod_i \frac{\theta 2^\theta}{x_i^{\theta+1}} \quad x_i > 2$$

$$f(x_1, \dots, x_n; \theta) = \frac{\theta^n 2^{n\theta}}{\prod x_i^{\theta+1}} \quad (x_i > 2?)$$

$$f(x_1, \dots, x_n; \theta) = \frac{\theta^n 2^{n\theta}}{(\prod x_i)^{\theta+1}} \quad (x_i > 2?)$$

Exercise 3

Let $(X_1, \dots, X_n)$ be a random sample of i.i.d. random variables distributed as follows:

$$f(x; \theta) = \frac{\theta 2^\theta}{x^{\theta+1}} \quad x > 2$$

$\prod X_i$ is a sufficient statistics for θ

$\sum \log(X_i)$ is a sufficient statistics for θ

Exercise 3

Let $(X_1, \dots, X_n)$ be a random sample of i.i.d. random variables distributed as follows:

$$f(x; \theta) = \frac{\theta 2^\theta}{x^{\theta+1}} \quad x > 2$$

$\prod X_i$ is a sufficient statistics for θ

$\sum \log(X_i)$ is a sufficient statistics for θ

Exercise 4

Suppose that X_1 and X_2 are two independent Bernoulli random variables with parameter p ; $0 < p < 1$. Show that the statistics $T = X_1 - X_2$ is not a sufficient statistics for p .

Exercise 4

Definition of Sufficient Statistics

A statistic $T(X)$ is a sufficient statistic for θ if the conditional distribution of the sample X given the value of $T(X)$ does not depend on θ .

Distribution of (X_1, X_2) , given $T = X_1 - X_2$

Exercise 4

Definition of Sufficient Statistics

A statistic $T(X)$ is a sufficient statistic for θ if the conditional distribution of the sample X given the value of $T(X)$ does not depend on θ .

Distribution of (X_1, X_2) , given $T = X_1 - X_2$

Exercise 4

Distribution of (X_1, X_2) , given $X_1 - X_2$

(x_1, x_2)	$x_1 - x_2$	$p(X_1, X_2 X_1 - X_2 = t)$
(0,0)	-1	0
(0,1)	-1	1
(1,0)	-1	0
(1,1)	-1	0
(0,0)	0	$(1-p)^2/(p^2 + (1-p)^2)$
(0,1)	0	0
(1,0)	0	0
(1,1)	0	$p^2/(p^2 + (1-p)^2)$
(0,0)	1	0
(0,1)	1	0
(1,0)	1	1
(1,1)	1	0

Exercise 4

Definition of Sufficient Statistics

A statistic $T(X)$ is a sufficient statistic for θ if the conditional distribution of the sample X given the value of $T(X)$ does not depend on θ .

Distribution of (X_1, X_2) , given $T = X_1 - X_2$

$T = X_1 - X_2$ is not a sufficient statistics for p

Exercise 4

Definition of Sufficient Statistics

A statistic $T(X)$ is a sufficient statistic for θ if the conditional distribution of the sample X given the value of $T(X)$ does not depend on θ .

Distribution of (X_1, X_2) , given $T = X_1 - X_2$

$T = X_1 - X_2$ is not a sufficient statistics for p

Exercise 5

Let consider the pdf

$$f(x; \theta) = \frac{1}{\theta} \exp\left(1 - \frac{x}{\theta}\right) \quad 0 < \theta < x$$

Find a sufficient statistic for parameter θ .

Factorization Theorem

$$\begin{aligned}f(x_1, x_2, \dots, x_n; \theta) &= \prod_{i=1}^n \left(\frac{1}{\theta} \exp \left(1 - \frac{x_i}{\theta} \right) l_{(\theta, \infty)}(x_i) \right) \\&= \frac{1}{\theta^n} \exp \left(\sum_{i=1}^n \left(1 - \frac{x_i}{\theta} \right) \right) \prod_{i=1}^n l_{(\theta, \infty)}(x_i) \\&= \frac{1}{\theta^n} \exp \left(n - \frac{1}{\theta} \sum_{i=1}^n x_i \right) l_{(\theta, \infty)}(x_{(1)})\end{aligned}$$

Exercise 5

Therefore we can set: $h(x) = 1$ and

$$g(T_1(x), T_2(x); \theta) = \frac{1}{\theta^n} \exp \left(n - \frac{1}{\theta} \sum_{i=1}^n x_i \right) l_{(\theta, \infty)}(x_{(1)})$$

Sufficient Statistics: Sample sum and Sample minimum

$$T_1(x) = \sum_{i=1}^n x_i \text{ and } T_2(x) = x_{(1)}$$

Two jointly sufficient statistics for the parameter θ

The dimension of a minimal sufficient statistic (two) is greater than the dimension of the parameter space (one).

Exercise 5

Therefore we can set: $h(x) = 1$ and

$$g(T_1(x), T_2(x); \theta) = \frac{1}{\theta^n} \exp \left(n - \frac{1}{\theta} \sum_{i=1}^n x_i \right) I_{(\theta, \infty)}(x_{(1)})$$

Sufficient Statistics: Sample sum and Sample minimum

$$T_1(x) = \sum_{i=1}^n x_i \text{ and } T_2(x) = x_{(1)}$$

Two jointly sufficient statistics for the parameter θ

The dimension of a minimal sufficient statistic (two) is greater than the dimension of the parameter space (one).

Solution 6 (Truncated Poisson, $X_1 + X_2$ not sufficient).

$$p(x; \theta) = \begin{cases} e^{-\theta} & x = 0 \\ \theta e^{-\theta} & x = 1 \\ 1 - e^{-\theta} - \theta e^{-\theta} & x = 2 \\ 0 & x \neq 0, 1, 2 \end{cases}$$

Let $T = X_1 + X_2 \in \{0, 1, 2, 3, 4\}$.

To show T is *not* sufficient for θ , it suffices to find one value of T at which the conditional distribution of (X_1, X_2) given T depends on θ .

Take $T = 2$, reached by $(0, 2)$, $(2, 0)$, or $(1, 1)$:

Solution 6 (Truncated Poisson, $X_1 + X_2$ not sufficient).

$$P(X_1 = 1, X_2 = 1) = \theta^2 e^{-2\theta}$$

$$\begin{aligned} P(X_1 = 0, X_2 = 2) &= P(X_1 = 2, X_2 = 0) \\ &= e^{-\theta}(1 - e^{-\theta} - \theta e^{-\theta}) \end{aligned}$$

$$P(T = 2) = \theta^2 e^{-2\theta} + 2e^{-\theta}(1 - e^{-\theta} - \theta e^{-\theta}).$$

Hence

$$P(X_1 = 1, X_2 = 1 \mid T = 2) = \frac{\theta^2 e^{-2\theta}}{\theta^2 e^{-2\theta} + 2e^{-\theta}(1 - e^{-\theta} - \theta e^{-\theta})}.$$

it depends on θ !