

Let  $X_1', X_2'$  be the number of heads and tails after  $T = X_1 + X_2 = t$  tosses of coin, of a fair coin ( $P(\text{head}) = \frac{1}{2}$ ). Then

$$(X_1', X_2') \mid T=t$$

is the same as

$$(X_1, X_2) \mid T=t$$

#

Ex on sufficient statistic for sample  $X_1, \dots, X_n$  of  $Un(0, \theta)$ , read on your own, note also that

$$\sum_{i=1}^n X_i$$

(which is proved to be the sufficient statistic in Lehman) is actually the ML estimator

#

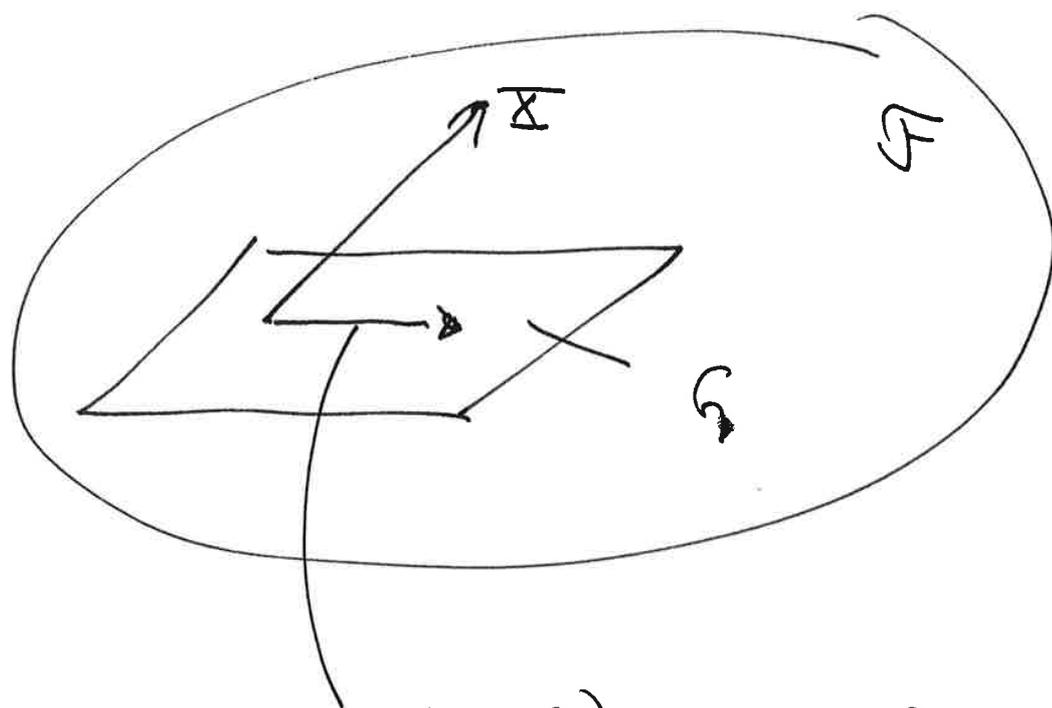
In measure theory and probability theory one defines conditioned expectation on the basic conditional entity.

$$E(X|Y) = ?$$

$$\mathcal{F} = \sigma(X), \quad \mathcal{G} = \sigma(Y), \quad \mathcal{G} \subseteq \mathcal{F}$$

Look at all r.v.  $Z$  are in  $\mathcal{G}$   
 s.t.  $E(Z^2) < \infty$  and all  $Z$  in  $\mathcal{F}$

s.t.  $E(Z^2) < \infty$



$$E(X|\mathcal{G}) = E(X|Y)$$

One defines (Kolmogorov 1933)

$E(X|\mathcal{G})$  as the projection of  $X$   
 on the sub-sigma-algebra  $\mathcal{G}$

(2)

$$E(X|G) = \underset{c \in G}{\operatorname{argmin}} E((X - c)^2)$$

Having defined a conditional expectation one defines a conditional probability by

$$P(A|G) = E(1_A | G)$$

$$P(A|B) = E(1_A | B)$$

How do we find a sufficient statistic?

1. Guessing.
2. Factorization criterion.

Theorem.

$T$  is a sufficient statistic for  $\mathcal{F} = \{F_\theta : \theta \in \Omega\}$ , dominated by a ~~sigma~~  $\sigma$ -finite measure

$\mu$ , it and only if there are non-negative functions,  $g_\theta$  and  $h$ , such that

$$(2) \quad p_{\theta}^{opt} = \frac{dF_{\theta}}{d\mu} = g_{\theta}(T(x)) h(x)$$

#

$\sigma$ -finite measure means...

$\mu$  finite on  $A$  (whole set, universal set) means  $\mu(A) < \infty$

$\mu$   $\sigma$ -finite on  $A$  means  $\exists$  partition  $A = \bigcup_{i=1}^{\infty} A_i$  s.t.  $\mu(A_i) < \infty, \forall i$ .

All "nice" measures are  $\sigma$ -finite.

Recall, we only have length measure, and counting measure.

Note that length measure  $\lambda$  on  $\mathbb{R}$  is not finite, but it is  $\sigma$ -finite.

$$\mathbb{R} = \bigcup_{n=-\infty}^{\infty} (n, n+1]$$

and  $\mu((n, n+1]) = 1$ .

Proof.

Assume that  $X$  is a discrete r.v., so  
 $p_{\theta}(x) = P_{\theta}(X=x)$

(II') Assume that (2) holds. Let  $T(x) = t$  arbitrary but fixed

$$\begin{aligned} P_{\theta}(X=x | T=t) &= \frac{P_{\theta}(X=x, T(x)=t)}{P_{\theta}(T=t)} \\ &= \frac{h(x) g_{\theta}(t)}{\sum_{x': T(x')=t} p_{\theta}(x')} \\ &= \frac{h(x) g_{\theta}(t)}{\sum_{x': T(x')=t} h(x') g_{\theta}(t)} = \frac{h(x)}{\sum_{x': T(x')=t} h(x')} \end{aligned}$$

(5)

and this does not depend on  $\theta$ ,  
therefore  $T$  is sufficient

( $\Downarrow$ )  $T$  sufficient so that

$$P_{\theta}(X=x | T(x)=t)$$

does not depend on  $\theta$ , say it is  
equal to  $k(x, t)$ . Then

$$P_{\theta}(X=x) = P_{\theta}(X=x | T(X)=t)$$

$$\cdot P_{\theta}(T(X)=t)$$

$$= \underbrace{k(x, t)}_{\text{does not}} \cdot \underbrace{P_{\theta}(T(X)=t)}_{\text{depend on } \theta}$$

$$= h(x) = f_{\theta}(T(X))$$

#

EX: (Poisson)

Assume  $X_1, \dots, X_n$  i.i.d.  $Po(\lambda)$ . The

$$\begin{aligned} P_\lambda(X_1 = x_1, \dots, X_n = x_n) \\ &= \prod_{i=1}^n \lambda \frac{e^{-\lambda}}{x_i!} \\ &= \prod_{i=1}^n e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} = e^{-n\lambda} \frac{\lambda^{\sum x_i}}{\prod x_i!} \\ &= \underbrace{\lambda^{\sum x_i} e^{-n\lambda}}_{g_\lambda(\sum x_i)} \underbrace{\frac{1}{\prod x_i!}}_{h(x)} \end{aligned}$$

$\therefore T = \sum x_i$  is sufficient

$\neq 1$

## EX. (order statistic)

Suppose i.i.d.  $X_1, \dots, X_n \sim F$  unknown  
and the  $X_i$ 's are cont. r.v.'s.

$$\text{Let } T(X_{(1)}, \dots, X_{(n)}) = (X_{(1)}, \dots, X_{(n)})$$

The ~~statistic~~ <sup>values are</sup> ~~is~~ unique almost surely,  
so there are no ties in the definition  
of the statistic.

$$\text{Given } T(X_{(1)}, \dots, X_{(n)}) = (X_{(1)}, \dots, X_{(n)})$$

there are  $n!$  equally probable,  
values for  $X_{(1)}, \dots, X_{(n)}$ , so the  
joint probability of  $X_{(1)}, \dots, X_{(n)}$   
given  $T$  is  $\frac{1}{n!}$ , does not  
depend on  $F$ .

Order statistic is sufficient for  
estimating  $F$  in  $\{F \text{ d.h. for}$   
cont. r.v.'s  $\}$ .

(8)

Recall that

$$F_n(t) = \frac{1}{n} \sum_{i=1}^n 1_{\{\Delta_i \leq t\}}$$

$\downarrow$

$$(\Delta_{(1)}, \dots, \Delta_{(n)})$$

#

Question What happens if we have dependent data?

EX: (Different statistics)

$X_1, \dots, X_n$  i.i.d.  $N(0, \sigma^2)$

$$T_1(X) = (X_1, \dots, X_n)$$

$$T_2(X) = (X_1^2, \dots, X_n^2)$$

$$T_3(X) = X_1^2 + \dots + X_n^2$$

are all sufficient statistics  
(Problem 6.5).

Two statistics  $T, T'$  are equivalent if  $\exists f, g$ .

$$T = f(T')$$

$$T' = g(T)$$

(almost surely).

If  $T$  sufficient and  $T = f(U)$  for some function  $f$  then  $U$  sufficient

know  $U \Rightarrow$  know  $T$

If  $f$  is not invertible ~~is~~ so it does not hold that

$$U = f^{-1}(T)$$

then  $T$  reduces data more than  $U$ ,  
or  $T$  is finer than  $U$ .

Def. If  $T$  is sufficient and for any other sufficient statistic  $U$ , there exists  $h, f$  such that

$$T = f(U)$$

then  $T$  is called minimal sufficient. #

Interpretation: A minimal sufficient statistic reduces the data the most while still being sufficient.

Lemma.

If  $\mathcal{F} = \{F_\theta : \theta \in \Omega\}$  set of d.f. dominated by a  $\sigma$ -finite measure  $\mu$ , then  $U$  is sufficient if and only if

$$\frac{p_\theta(x)}{p_{\theta_0}(x)}$$

is a function of only  $U(x)$ , for all fixed  $\theta_0, \theta$ .

Proof Problem 6.6.

#

(1)

## Theorem

Let  $\mathcal{F} = \{p_0, \dots, p_k\}$  be a finite family of densities with the same support.

Then

$$T(x) = \left( \frac{p_1(x)}{p_0(x)}, \dots, \frac{p_k(x)}{p_0(x)} \right)$$

is minimal sufficient.

## Proof

Lemma says that  $U$  sufficient if and only if

$$\frac{p_k(x)}{p_0(x)}$$

is a function of  $U(x)$  for all  $k$ .

That means that

$$T(x)$$

is a function of  $U(x)$ , sufficient

statistic. Therefore  $T$  is minimal sufficient

#

(12)

Proof and statement are also for countable families.

Lemma

$\mathcal{F}$  family of d.b. with common support and  $F_0 \in \mathcal{F}$ .

Assume

$T$  minimal sufficient for  $F_0$

$T$  sufficient for  $\mathcal{F}$

Then

$T$  is minimal sufficient for  $\mathcal{F}$ .

Proof

Start with arbitrary sufficient statistic for  $F_0$ ,  $U$ .  $U$  sufficient for  $\mathcal{F}$  therefore also for  $F_0$ .  $T$  is minimal sufficient on  $F_0$  therefore  $\exists \phi$

$$T = \phi(U)$$

$\Rightarrow T$  minimal sufficient on  $\mathcal{F}$ . (13)