SUFFICIENCY. CHAP. 6 SEC. 6.1-6.2

GIVEN A RANDOM SAMPLE $X_1, X_2, X_3 \ldots X_n$ WHERE EACH X FOLLOWS A DIST. $f(x|\theta)$, WE WANT TO MAKE INFERENCES ON THE UNKNOWN θ.

$T(X_1, X_2, X_3 \ldots X_n)$ WILL ASSIST US IN SUMMARIZING THE INFO. IN THE SAMPLE.

SOME NOTATION:

$\underline{X} = (X_1, X_2 \ldots X_n)$ RANDOM SAMPLE

$\underline{x} = (x_1, x_2 \ldots x_n)$ OBS. VALUES (DATA)

$T(\underline{X})$ IS A STATISTIC.

$T(\underline{x})$ OBS. VALUE OF A STATISTIC. ALSO REPRESENTED BY t, i.e. $T(\underline{x}) = t$. SO THE INFORMATION IN THE SAMPLE IS SUMMARIZED BY t.

$\mathcal{X}$ IS THE SAMPLE SPACE. T GENERATES A PARTITION OVER $\mathcal{X}$ SINCE $A_t = \{\underline{x} : T(\underline{x}) = t\}$ ARE THE DEFINING ELEMENTS OF THIS PARTITION

EX: CONSIDER X_1, X_2, X_3 iid WHERE X_i FOLLOWS A BERNOULLI (θ) SO $f(x|\theta) = \theta^x (1-\theta)^{1-x}$, $x = 0, 1$, $0 < \theta < 1$

ALSO CONSIDER 3 STATISTICS. $T_1 = \sum_{i=1}^{3} X_i$; $T_2 = X_1$.

$T_3 = X_{(3)} = \max\{X_1, X_2, X_3\}$.

WHAT ARE THE VALUES $\underline{x}$ IN THE SAMPLE SPACE $\mathcal{X}$?

SINCE X JUST TAKES THE VALUES 0 OR 1, THEN

$\mathcal{X} = \{(0,0,0), (0,0,1), (0,1,0), \text{ETC.} \ldots\}$

TABLE

(X_1, X_2, X_3)	$T_1 = \sum X_i$	$T_2 = X_1$	$T_3 = X_{(3)}$	PROB. OF SAMPLE VAL
$(0,0,0)$	0	0	0	$(1-\theta)^3$
$(0,0,1)$	1	0	1	$\theta(1-\theta)^2$
$(0,1,0)$	1	0	1	$\theta(1-\theta)^2$
$(1,0,0)$	1	1	1	$\theta(1-\theta)^2$
$(1,1,0)$	2	1	1	$\theta^2(1-\theta)$
$(0,1,1)$	2	0	1	$\theta^2(1-\theta)$
$(1,0,1)$	2	1	1	$\theta^2(1-\theta)$
$(1,1,1)$	3	1	1	θ^3

(the 8 sample points together: $\mathcal{X}$)

UNDER T1, THE SET OF VALUES FOR T ARE $\mathcal{T} = \{0,1,2,3\}$

PARTITION SET:

$A_0 = \{(0,0,0)\}$; $A_1 = \{(0,0,1),(0,1,0),(0,0,1)\}$;

$A_2 = \{(1,1,0),(0,1,1),(1,0,1)\}$; $A_3 = \{(1,1,1)\}$

NOTICE $A_i \cap A_j = \emptyset$ $\quad \bigcup A_i = \mathcal{X}$

FOR T_2: $\mathcal{T} = \{0,1\}$

$A_0 = \{(0,0,0),(0,0,1),(0,1,0),(0,1,1)\}$

$A_1 = \{(1,0,0),(1,1,0),(1,0,1),(1,1,1)\}$

FOR T_3: $\mathcal{T} = \{0,1\}$

$A_0 = \{(0,0,0)\}$, $A_1 = \{$ OTH. ELEMENTS $\} = A_0^c$

IF $\underline{x}, \underline{y} \in A_t \Rightarrow T(\underline{x}) = T(\underline{y}) = t$ AND THEN THE INFERENCE ABOUT θ IS THE SAME WITH BOTH $\underline{x} + \underline{y}$ (SUFFICIENCY PRINCIPLE).

LETS CONSIDER THE FOLLOWING FOR T_1

$Pr[\underline{X} = \underline{x} \mid T_1 = t] = Pr[X_1 = x_1, X_2 = x_2, X_3 = x_3 \mid T_1(x_1, x_2, x_3) = t$

AND $t = 0$

$Pr[\underline{X} = \underline{x} \mid T_1(\underline{x}) = 0]$. NOTICE THAT IF $\underline{x}$ IS SUCH THAT $T_1(\underline{x}) \neq 0 \Rightarrow Pr[\underline{X} = \underline{x} \mid T_1(\underline{x}) = 0] = 0$

WE'RE ONLY INTERESTED IN $\underline{x}$ SUCH THAT $T_1(\underline{x}) = 0 \Rightarrow$ $\underline{x} = (0,0,0)$.

$$Pr[\underline{X} = (0,0,0) \mid T_1(\underline{x}) = 0] = \frac{Pr[X = (0,0,0) \text{ AND } T_1(\underline{x}) = 0]}{Pr[T_1(\underline{x}) = 0]} =$$

$\frac{(1-\theta)^3}{(1-\theta)^3} = 1$ DOES NOT DEPEND ON θ

NOW TAKE $t = 1$, THE ELEMENTS $\underline{x}$ WE ARE INTERESTED CORRESPONDING TO A_1, (FOR OTHERS COND. PROB $= 0$)

IF $\underline{x} \in A_1 \Rightarrow T_1(\underline{x}) = 1$ AND $Pr[T_1(\underline{x}) = 1] = 3\theta(1-\theta)^2$

$$\Rightarrow Pr[\underline{X} = \underline{x} \mid T_1(\underline{x}) = 1] = \frac{Pr[\underline{X} = \underline{x}]}{Pr[T_1(\underline{x}) = 1]} = \frac{\theta(1-\theta)^3}{3(1-\theta)^3} = \frac{1}{3}$$

FOR $t = 2$, IF $\underline{x} \in A_2$

$$Pr[\underline{X} = \underline{x} \mid T_1(\underline{x}) = 2] = \frac{Pr[\underline{X} = \underline{x}]}{Pr[T_1(\underline{x}) = 2]} = \frac{\theta^2(1-\theta)}{3\theta^2(1-\theta)} = \frac{1}{3}$$

FOR $t = 3$, IF $\underline{x} \in A_3$

$$Pr[\underline{X} = \underline{x} \mid T_1(\underline{x}) = 3] = \theta^3/\theta^3 = 1$$

THIS LEADS TO THE FOLLOWING DEFINITION:

<u>DEF</u> $T(\underline{X})$ IS A SUFFICIENT STATISTIC FOR θ IF AND ONLY IF $Pr[\underline{X} = \underline{x} \mid T(\underline{x}) = t]$ DOES NOT DEPEND ON θ.

NOTE: DEF. APPLIES FOR CONT. CASE $f_{\underline{X}|T=t}(\underline{x}|T=t)$

IS T2 SUFFICIENT?

IF WE TAKE $t=0$ AND $\underline{x}=(0,0,0) \Rightarrow$

$$P_r[\underline{X}=\underline{x} \mid T_2(\underline{x})=0] = \frac{P_r[\underline{X}=(0,0,0)]}{P_r[T_2(\underline{x})=0]} = \frac{(1-\theta)^3}{(1-\theta)^3+2\theta(1-\theta)^2}$$

DEPENDS ON θ! $\Rightarrow$ T2 IS NOT SUFFICIENT

HOW ABOUT T3?

TAKE $t=1$ AND $\underline{x}=(0,0,1) \Rightarrow$

$$P_r[\underline{X}=\underline{x} \mid T_3(\underline{x})=1] = \frac{P_r[\underline{X}=(0,0,1)]}{P_r[T_3(\underline{x})=1]} = \frac{\theta(1-\theta)^2}{3\theta(1-\theta)^2+3\theta^2(1-\theta)+\theta^3}$$

$\Rightarrow$ T3 IS NOT SUFFICIENT.

IN GENERAL, TO CHECK SUFFICIENCY, WE TAKE AN $\underline{x}$ SUCH THAT $T(\underline{x})=t$. THEN, WE OBTAIN

$$P_\theta[\underline{X}=\underline{x} \mid T(\underline{x})=t] = \frac{P_\theta[\underline{X}=\underline{x} \text{ AND } T(\underline{x})=t]}{P_\theta[T(\underline{x})=t]} \quad (1)$$

SINCE $T(\underline{x})=t$ THE EVENT $\{\underline{X}=\underline{x}\} \subset \{T(\underline{X})=t\}$

$$\Rightarrow P_\theta[\underline{X}=\underline{x} \text{ AND } T(\underline{X})=t] = P_\theta[\underline{X}=\underline{x}] = \prod_{i=1}^{n} f(x_i|\theta)$$

CHECK IF (1) DOES OR DOES NOT DEPEND ON θ.

T IS SUFFICIENT FOR $\theta \iff \dfrac{f(\underline{x}|\theta)}{f(T(x)=t|\theta)}$ DOES NOT DEPEND ON θ

(21)

MORE EXAMPLES:

TAKE $X_1, X_2 . \; X_n$ BE iid POISSON(λ) RVS.

LET $T(X_1, X_2 . \; X_n) = \sum_{i=1}^{n} X_i$. SHOW THAT T IS SUFFICIENT

$$f(\underline{x}|\lambda) = \prod_{i=1}^{n} f(x_i|\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = \frac{\lambda^{\sum_{i}^{n} x_i} e^{-n\lambda}}{\prod_{i=1}^{n} x_i!}$$

SINCE X'S ARE INDEPENDENT $\Rightarrow T = \sum_{i=1}^{n} X_i \sim$ POISSON$(n\lambda)$

LETS MAKE $\sum_{i=1}^{n} X_i = t$

$$\Rightarrow f(T(\underline{x}) = t \,|\, \lambda) = \frac{(n\lambda)^t e^{-n\lambda}}{t!}$$

$$\text{RATIO} = \frac{f(\underline{x}|\lambda)}{f(T(\underline{x}) = t|\lambda)} = \frac{\lambda^{\sum x_i} e^{-n\lambda}}{\prod x_i!} \Big/ \frac{(n\lambda)^t e^{-n\lambda}}{t!} = \frac{\lambda^t e^{-n\lambda}}{\prod_{i=1}^{n} x_i!} \Big/ \frac{(n\lambda)^t e^{-n\lambda}}{t!}$$

$$= \frac{1}{\prod x_i!} \Big/ \frac{n^t}{t!} = \frac{t!}{\prod x_i! \, n^t}$$ DOES NOT DEPEND ON λ

$\Rightarrow \sum_{i=1}^{n} X_i$ IS SUFFICIENT.

$f(x_i|\theta) = \frac{1}{\theta} \exp\left(-\frac{x_i}{\theta}\right)$

WITH CONT RVS. LET $X_1, X_2 . \; X_n$ BE iid AND Exp(θ), $T(X_1, X_2 \; X_n) = \sum_{i=1}^{n} X_i / n = \bar{X}$. IN A PREVIOUS CLASS, WE SHOWED THAT T HAS A GAMMA DIST $\alpha = n$, $\beta = \theta/n$

$$\Rightarrow f(T(\underline{x}) = t|\theta) = \frac{1}{\Gamma(n)(\theta/n)^n} t^{n-1} e^{-\frac{tn}{\theta}}$$

$$f(\underline{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-x_i/\theta} = \left(\frac{1}{\theta}\right)^n e^{-\sum_{i=1}^{n} x_i/\theta} =$$

IF $t = \bar{X}$

$$\left(\frac{1}{\theta}\right)^n e^{-\frac{n\bar{x}}{\theta}} = \left(\frac{1}{\theta}\right)^n e^{-\frac{nt}{\theta}}$$

IF WE TAKE THE RATIO $\dfrac{f(\underline{x}|\theta)}{f(T(\underline{x})=t|\theta)} = \dfrac{\left(\frac{1}{\theta}\right)^n e^{-nt/\theta}}{\frac{1}{\Gamma(n)(\theta/n)^n} t^{n-1} e^{-nt/\theta}}$

$= \left(\frac{1}{n}\right)^n \Gamma(n) / t^{n-1} \Rightarrow \bar{X}$ IS SUFFICIENT.

IN GENERAL, IT TAKES ~~QUITE A BIT OF~~ SOME WORK TO FIND. SUFFICIENT STATISTICS.

1ST. WE NEED TO PROPOSE A STATISTIC $T \sim f_T(t|\theta)$

2ND. DERIVE $f(T(\underline{x})=t|\theta)$ (DIST. OF T).

3RD. CHECK THAT RATIO $\dfrac{f(\underline{x}|\theta)}{f(T(\underline{x})|\theta)}$ DOES NOT DEPEND ON θ.

HOWEVER, ~~WE MAY FIND~~ IT IS POSSIBLE TO FIND A SUFFICIENT STATISTIC FOR θ BY SIMPLE INSPECTION OF $f(\underline{x}|\theta)$

FACTORIZATION THEOREM (PAG 276).

LET $f(\underline{x}|\theta)$ DENOTE THE JOINT pmf OF $\underline{X}$. $T(\underline{x})$ IS A SUFFICIENT FOR θ IF AND ONLY IF

$$f(\underline{x}|\theta) = g(T(\underline{x})|\theta)\, h(\underline{x})$$ FOR ALL VALUE OF $\underline{x}$ AND θ. WHERE $g(\cdot|\cdot)$ DEPENDS ON THE SAMPLE ONLY THROUGH $T(\cdot)$ AND $h(\cdot)$. EXCLUSIVELY DEPENDS ON $\underline{x}$.

HOW DOES THIS THEOREM WORKS ?

IN THE POISSON EX.

$f(\underline{x}|\lambda) = (\lambda)^{\sum_{i=1}^{n} x_i} e^{-n\lambda} / \prod_{i=1}^{n} x_i!$. MAKE $g(T(\underline{x})|\lambda) = \lambda^{\sum x_i} e^{-n\lambda}$ AND $h(\underline{x}) = 1/\prod_{i=1}^{n} x_i$ FACTORIZATION APPLIES

FOR THE EXPONENTIAL EX:

$$f(\underline{x}|\theta) = \left(\tfrac{1}{\theta}\right)^n \exp\left(-\sum_{i=1}^{n} x_i/\theta\right) = \underbrace{\left(\tfrac{1}{\theta}\right)^n \exp\left(-\tfrac{n\bar{x}}{\theta}\right)}_{g(T(\underline{x})|\theta)}$$

AND $h(\underline{x}) = 1$. $\Rightarrow$ $\bar{X}$ SUFFICIENT. THIS ARGUMENT ALSO APPLIES TO $\sum_{i=1}^{n} X_i$

ACTUALLY IF

$$S(X_1, X_2 \ldots X_n) = Z(T(X_1 \ldots X_n)) \text{ WHERE } Z$$

HAS INVERSE AND T SUFFICIENT $\Rightarrow$ S IS SUFFICIENT

$$g(T(\underline{x})|\theta)\, h(\underline{x}) = g(Z^{-1}(S(X_1, X_2 \ldots X_n))|\theta)\, h(\underline{x}).$$

ANOTHER EXAMPLE. (EX 6.2.4 IN BOOK)

TAKE $X_1, X_2, \ldots X_n$ A RANDOM SAMPLE SUCH THAT $X_i \sim N(\mu, \sigma^2)$ WITH σ^2 KNOWN $\Rightarrow f(x_i|\mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2}$

$$\Rightarrow f(\underline{x}|\mu, \sigma^2) = \prod_{i=1}^{n} \left(\tfrac{1}{2\pi\sigma^2}\right)^{1/2} \exp\left(-\tfrac{1}{2\sigma^2}(x_i-\mu)^2\right).$$

$$= \left(\tfrac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(-\tfrac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

IN THE EXP $\sum_{i=1}^{n}(x_i-\mu)^2 = \sum_{i=1}^{n}(x_i - \bar{x} - \mu + \bar{x})^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2 - 2 \sum_{i=1}^{n}(x_i-\bar{x})(\mu-\bar{x}) + n(\bar{x}-\mu)^2$ CROSS-PRODUCT TERM $= 0$

$\Rightarrow$

$$f(\underline{x}|\mu, \sigma^2) = \left(\tfrac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(\tfrac{-1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\mu-\bar{x})^2\right)\right) \quad (1)$$

MAKE $g(T(\underline{x}); \mu) = \exp\left(-\frac{1}{2\sigma^2} n(\mu-\bar{x})^2\right)$ AND

$h(\underline{x}) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)$ THIS SHOWS THAT $\bar{X}$ IS SUFFICIENT

WITH σ^2 UNKNOWN

$$g(T(\underline{x}),\mu,\sigma^2) = \left(\frac{1}{\sigma^2}\right)^{n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2\right) \exp\left(-\frac{n}{2\sigma^2}(\mu-\bar{x})^2\right)$$

$$h(\underline{x}) = \left(\frac{1}{2\pi}\right)^{n/2}$$

THE SUFFICIENT STATISTIC IS NOW $T = \left(\bar{x}, \sum_{i=1}^{n}(x_i-\bar{x})^2\right)$.

EX:

LET $X_1, X_2 .\ X_n$ BE A RANDOM SAMPLE SUCH THAT $X_i \sim U(0,\theta) \Rightarrow f(x_i|\theta) = \frac{1}{\theta}$ IF $0 < x_i < \theta$.

$$f(\underline{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \frac{1}{\theta} = \frac{1}{\theta^n} \quad \text{WHEN } 0 < x_i < \theta$$

FOR ALL i.

$0 < x_i < \theta \ \forall i \iff$ ALL X'S ARE IN $(0,\theta)$

$\iff 0 < \min\{x_1 .\ x_n\}$ AND $\max\{x_1, x_2 .\ x_n\} < \theta$

THEN,

$f(\underline{x}|\theta) = \frac{1}{\theta^n} I_{(0,\infty)}(x_{(1)}) \, I_{(0,\theta)}(x_{(n)})$. THE FACTORIZATION

FOLLOWS BY MAKING $g(T(\underline{x});\theta) = \frac{1}{\theta^n} I_{[0,\theta]}(x_{(n)})$; $h(\underline{x}) = I_{(0,\infty)}(x_{(1)})$

$\Rightarrow X_{(n)} = \max\{x_1, x_2 .\ x_n\}$ IS A SUFFICIENT STATISTIC.

FOR θ.

PROOF OF FACTORIZATION THEOREM. (ONLY FOR THE DISCRETE CASE)

SUPPOSE $T(\underline{X})$ IS SUFFICIENT AND $\underline{x}$ IS SUCH $T(\underline{x}) = t$

$$f(\underline{x}|\theta) = P_\theta[\underline{X}=\underline{x} \text{ AND } T(\underline{X})=t] = \underbrace{P_\theta[\underline{X}=\underline{x} \mid T(\underline{X})=t]}_{\text{DOES NOT DEPEND ON } \theta \text{ BECAUSE } T \text{ IS SUFFICIENT}}$$

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

CHOOSE $g(T(\underline{x})=t;\theta) = P_\theta[T(\underline{x})=t]$ AND
$h(\underline{x}) = P[\underline{X}=\underline{x} \mid T(\underline{x})=t]$
$\Rightarrow f(\underline{x}\mid\theta) = g(T(\underline{x})=t;\theta)\, h(\underline{x})$ FACTORIZATION APPLIES

NOW ASSUME THE FACTORIZATION EXISTS. NEED TO SHOW THAT T IS SUFFICIENT

$$\frac{f(\underline{x}\mid\theta)}{f(T(\underline{x})=t\mid\theta)} = \frac{g(T(\underline{x})=t\mid\theta)\, h(\underline{x})}{f(T(\underline{x})=t\mid\theta)}$$

$$= \frac{g(T(\underline{x})=t\mid\theta)\, h(\underline{x})}{\sum_{\underline{y}\in A_t} f(\underline{y}\mid\theta)} \qquad A_t = \{\underline{y} \mid T(\underline{y})=t\}$$

$$= g(T(\underline{x})=t\mid\theta)\, h(\underline{x}) \Big/ \sum_{\underline{y}\in A_t} g(T(\underline{y})=t\mid\theta)\, h(\underline{y})$$

t IS CONSTANT IN A_t $\Rightarrow$

$$\text{RATIO} = g(T(\underline{x})=t\mid\theta)\, h(\underline{x}) \Big/ g(T(\underline{x})=t\mid\theta) \sum_{\underline{y}\in A_t} h(\underline{y})$$

$$= \frac{h(\underline{x})}{\sum_{\underline{y}\in A_t} h(\underline{y})}$$ DOES NOT DEPEND ON θ.

THEN T IS SUFFICIENT

EX

WE SAY THAT $f(x\mid\theta)$ BELONGS TO THE EXPONENTIAL FAMILY IF

$$f(x\mid\theta) = h(x)\, c(\theta) \exp\left(\sum_{i=1}^{K} w_i(\theta)\, t_i(x)\right)$$

θ MAY BE A VECTOR OR A SCALAR

KNOWN MODELS BERNOULLI, BIN. POISSON, NORMAL, BELONG TO THE EXPONENTIAL FAMILY.

(26)

IF $X \sim \text{POISSON}(\lambda)$

$$f(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-\lambda}e^{x\log\lambda}}{x!} = \underbrace{\frac{1}{x!}}_{h(x)}\underbrace{e^{-\lambda}}_{c(\lambda)} e^{x\log\lambda}$$

$w(\lambda) = \log\lambda$

$t_1(x) = x$

IF $X \sim N(\mu, \sigma^2)$

$$f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) =$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}x^2\right) \exp\left(+\frac{2x\mu}{2\sigma^2}\right) \exp\left(-\frac{\mu^2}{2\sigma^2}\right)$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-\mu^2}{2\sigma^2}\right) \exp\left(-\frac{1}{2\sigma^2}x^2 + \frac{x\mu}{\sigma^2}\right)$$

HOW TO FIND A SUFFICIENT T?

$X_1, X_2, \ldots X_n$ IID OBSERVATIONS SUCH THAT X_i FOLLOWS AN EXPONENTIAL FAMILY.

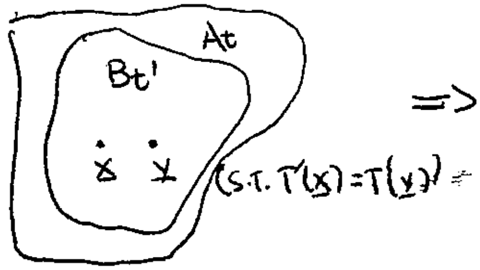
$$f(\underline{x}|\theta) = \prod_{j=1}^{n} f(x_j|\theta) = \prod_{j=1}^{n} h(x_j)\, c(\theta) \exp\left(\sum_{i=1}^{k} w_i(\theta)\, t_i(x_j)\right)$$

$$= \underbrace{\prod_{j=1}^{n} h(x_j)}_{h(\underline{x})} \underbrace{c^n(\theta) \exp\left(\sum_{i=1}^{k} w_i(\theta) \sum_{j=1}^{n} t_i(x_j)\right)}_{g(T(\underline{x})|\theta)}$$

WHERE $T(\underline{x}) = \left(\sum_{j=1}^{n} t_1(x_j), \sum_{j=1}^{n} t_2(x_j), \ldots \sum_{j=1}^{n} t_k(x_j)\right)$

IS A SUFFICIENT STATISTIC.

GRAPHICAL:



$\Rightarrow Bt' \subseteq At$

VERY HARD DEFINITION TO CHECK IN PRACTICE!

FORTUNATELY, WE HAVE THE FOLLOWING THEOREM BY LEHMANN AND SCHEFFÉ

THEOREM: SUPPOSE WE HAVE $T(\underline{x})$ SUCH THAT FOR ANY TWO POINTS $\underline{x}, \underline{y} \in \mathcal{X}$, THE RATIO $\frac{f(\underline{x}|\theta)}{f(\underline{y}|\theta)}$ IS CONSTANT AS A FUNCTION OF $\theta \Leftrightarrow T(\underline{x}) = T(\underline{y})$. THEN $T(\underline{x})$ IS A MINIMAL SUFFICIENT STATISTIC FOR θ (SEE PROOF PAG 281)

EXAMPLES:

IF $X_1, X_2, \dots X_n$ ARE iid POISSON(λ) RVs THEN

$f(\underline{x}|\lambda) = \lambda^{\sum_1^n x_i} e^{-n\lambda} / \prod_{i=1}^n x_i!$. FOR ANOTHER POINT $\underline{y} \in \mathcal{X}$

$f(\underline{y}|\lambda) = \lambda^{\sum_{i=1}^n y_i} e^{-n\lambda} / \prod_{i=1}^n y_i!$

CONSIDER THE RATIO:

$$\frac{f(\underline{x}|\lambda)}{f(\underline{y}|\lambda)} = \frac{\lambda^{\sum x_i} e^{-n\lambda} / \prod x_i!}{\lambda^{\sum y_i} e^{-n\lambda} / \prod y_i!} = \left(\lambda^{\sum_{i=1}^n x_i - \sum_{i=1}^n y_i}\right)\left(\frac{\prod y_i!}{\prod x_i!}\right) = c$$

THIS RATIO IS CONSTANT AS A FUNCTION OF $\lambda \Leftrightarrow \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$

$\Rightarrow T(\underline{x}) = \sum_{i=1}^n x_i$ IS SUFFICIENT AND MINIMAL

COMMENT:

IN A GIVEN PROBLEM, THERE ARE MANY SUFFICIENT STATISTICS. (IF $\sum_{i=1}^{n} x_i \to \bar{x}$ SUFFICIENT)

IN FACT, THE FULL SAMPLE $\underline{X} = (X_1, X_2, \ldots X_n)$ IS A SUFFICIENT STATISTIC.

IF $T(\underline{x}) = \underline{x}$

BY THE FACT. THEOREM: $f(\underline{x}|\theta) = f(T(\underline{x})|\theta) \cdot 1$, $h(\underline{x}) = 1$

ALSO, IF $T(\underline{x})$ IS SUFFICIENT ANY ONE-TO-ONE FUNCTION IS ALSO SUFFICIENT

$T^*(\underline{x}) = r(T(\underline{x}))$ r ONE-TO-ONE $\Rightarrow T(\underline{x}) = r^{-1}(T^*(\underline{x}))$

BY THE FACT. THEOREM: IF $T(\underline{x})$ SUFFICIENT $\Leftrightarrow$

$$f(\underline{x}|\theta) = g(T(\underline{x})|\theta)\, h(\underline{x}) = g(r^{-1}(T^*(\underline{x}))|\theta)\, h(\underline{x})$$
$$= g^*(T^*(\underline{x})|\theta)\, h(\underline{x})$$ AND T^* IS ALSO SUFFICIENT

HOW DO WE FIND THE SUFFICIENT STATISTIC THAT ACHIEVES THE MOST DATA REDUCTION? ANS.: MINIMAL SUFFICIENT STATISTIC.

<u>DEFINITION</u>: A SUFFICIENT $T(\underline{X})$ IS CALLED MINIMAL SUFFICIENT IF FOR ANY OTHER SUFFICIENT STATISTIC $T'(\underline{x})$, $T(\underline{x})$ IS A FUNCTION OF $T'(\underline{x})$

TWO IMPLICATIONS:

- IF $T(\underline{x})$ IS A FUNCTION OF $T'(\underline{x})$, IF FOR SAMPLE POINTS $\underline{x}, \underline{y}$ $T'(\underline{x}) = T'(\underline{y}) \Rightarrow T(\underline{x}) = T(\underline{y})$
- •• IF $\{B_{t'}, t' \in T'\}$ ARE THE PARTITION SETS FOR T' AND $\{A_t, t \in T\}$ ARE THE PARTITION SETS FOR T NECESSARILY $B_{t'}$ IS CONTAINED IN ONE A_t

T LEADS TO THE "COARSEST" PARTITION SET.

NOW, SUPPOSE $X_1, X_2 \ldots X_n$ ARE iid $N(\mu, \sigma^2)$. FOR A SAMPLE POINT $\underline{x}$:

$$f(\underline{x}|\mu,\sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2\right) \exp\left(\frac{-n}{2\sigma^2}(\mu-\bar{x})^2\right)$$

THEN, FOR A POINT $\underline{y}$

$$f(\underline{y}|\mu,\sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\bar{y})^2\right) \exp\left(\frac{-n}{2\sigma^2}(\mu-\bar{y})^2\right)$$

RATIO:

$$\frac{f(\underline{x}|\mu,\sigma^2)}{f(\underline{y}|\mu,\sigma^2)} = \exp\left(-\frac{1}{2\sigma^2}\left[\sum_1^n(x_i-\bar{x})^2 - \sum_1^n(y_i-\bar{y})^2\right]\right) \exp\left(-\frac{n}{2\sigma^2}\left[(\mu-\bar{x})^2 - (\mu-\bar{y})^2\right]\right)$$

THIS RATIO IS CONSTANT AS A FUNCTION OF (μ, σ^2) IFF $\sum_{i=1}^{n}(x_i-\bar{x})^2 = \sum_{i=1}^{n}(y_i-\bar{y})^2$ AND $\bar{x} = \bar{y}$. $T = (\bar{X}, \sum(X_i-\bar{X})^2)$ IS MINIMAL SUFFICIENT.

LET $X_1, X_2 \ldots X_n$ BE iid $U(0,\theta)$ RVs.

$$\Rightarrow f(x|\theta) = \frac{1}{\theta}; \quad 0 < x < \theta \quad \text{OR} \quad f(x|\theta) = \frac{1}{\theta} I_{(0,\theta)}(x)$$

FOR A POINT $\underline{x}$, $f(\underline{x}|\theta) = \prod_{i=1}^{n} \frac{1}{\theta} I_{(0,\theta)}(x_i) = \frac{1}{\theta^n} \prod_{i=1}^{n} I_{(0,\theta)}(x_i)$

NOTICE THAT $\prod_{i=1}^{n} I_{(0,\theta)}(x_i) = 1 \iff$ ALL X's ARE IN $(0,\theta)$

$\iff X_{(1)} > 0$ AND $X_{(n)} < \theta \Rightarrow I_{(0,\infty)}(x_{(1)}) \, I_{(-\infty,\theta)}(x_{(n)}) = 1$

$\Rightarrow f(\underline{x}|\theta) = \frac{1}{\theta^n} I_{(0,\infty)}(x_{(1)}) \, I_{(-\infty,\theta)}(x_{(n)})$, FOR A POINT $\underline{y}$,

$f(\underline{y}|\theta) = \frac{1}{\theta^n} I_{(0,\infty)}(y_{(1)}) \, I_{(-\infty,\theta)}(y_{(n)})$. RATIO $= \dfrac{f(\underline{x}|\theta)}{f(\underline{y}|\theta)} = \dfrac{I_{(0,\infty)}(x_{(1)}) \, I_{(-\infty,\theta)}(x_{(n)})}{I_{(0,\infty)}(y_{(1)}) \, I_{(-\infty,\theta)}(y_{(n)})}$

WILL NOT DEPEND ON $\theta \iff X_{(n)} = Y_{(n)}$

$X_{(n)}$ IS MINIMAL SUFFICIENT.

ANOTHER EXAMPLE

$X_1, X_2 . \; X_n$ BE iid $U(\theta, \theta+1)$ FIND A MINIMAL SUFFICIENT STATISTIC FOR θ.

$\Rightarrow f(x|\theta) = I_{(\theta,\theta+1)}(x)$

THEN $f(\underline{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} I_{(\theta,\theta+1)}(x_i)$. FOR ALL INDICATORS

TO BE ONE, ALL $\theta < x_i < \theta+1 \iff \theta < x_{(1)}$ & $x_{(n)} < \theta+1$

$\iff x_{(n)}-1 < \theta < x_{(1)}$

$\Rightarrow f(\underline{x}|\theta) = I_{(x_{(n)}-1,\, x_{(1)})}(\theta)$. FOR A SAMPLE POINT $\underline{y}$,

$f(\underline{y}|\theta) = I_{(y_{(n)}-1,\, y_{(1)})}(\theta) \Rightarrow \dfrac{f(\underline{x}|\theta)}{f(\underline{y}|\theta)} = \dfrac{I_{(x_{(n)}-1,\, x_{(1)})}(\theta)}{I_{(y_{(n)}-1,\, y_{(1)})}(\theta)}$ IS CONSTANT

$\iff x_{(1)} = y_{(1)}$ AND $x_{(n)} = y_{(n)} \Rightarrow T = (X_{(1)}, X_{(n)})$ IS MINIMAL SUFFICIENT

FINAL NOTE ON M.S:

NEEDS TO BE SUFFICIENT

IF T IS MINIMAL SUFFICIENT AND $T^* = r(T)$ WHERE r IS ONE TO ONE, T^* IS ALSO MINIMAL SUFFICIENT

PROOF:

WE SAW LAST TIME THAT T^* IS SUFFICIENT

SINCE T IS MINIMAL SUFFICIENT, FOR ANY OTHER T'

$T = g(T') \Rightarrow T^* = r \circ g(T') \Rightarrow T^*$ IS ALSO MINIMAL SUFFICIENT.

DEFINITION: $T(\underline{X})$ IS A COMPLETE STATISTIC IF $E_\theta(g(T)) = 0$ FOR ALL θ IMPLIES $P_\theta(g(T)=0) = 1$, FOR ALL θ (UNDER A FAMILY OF pdfs OR pmfs $f(t|\theta)$ FOR $T(\underline{X})$).

EX: LET $X_1, X_2 \ldots X_n$ iid BERNOULLI(θ) RVS.
IF $T = \sum_{i=1}^{n} X_i \Rightarrow T$ IS BINOMIAL (n,θ). $0 < \theta < 1$
LETS CONSIDER $g(T)$ A GENERIC FUNCTION OF T
$\Rightarrow$

$$E(g(T)) = \sum_{t=0}^{n} g(t) \binom{n}{t} \theta^t (1-\theta)^{n-t}.$$

$$E(g(T)) = 0 \Leftrightarrow \underbrace{\sum_{t=0}^{n} g(t) \binom{n}{t} \theta^t (1-\theta)^{n-t}}_{\text{POLINOMIAL OF DEGREE } n \text{ IN } \theta} = 0$$

THE POLYNOMIAL IS ZERO $\Leftrightarrow g(t)\binom{n}{t} = 0 \Leftrightarrow g(t) = 0$
$\Rightarrow T$ IS A COMPLETE STATISTIC.
AN EXAMPLE OF SOMETHING THAT IS NOT COMPLETE
X_1, X_2 TWO BERNOULLI (θ) RVS.
$T = X_2 - X_1$ NOTICE THAT $E(T) = E(X_1) - E(X_2) = \theta - \theta = 0$

$X_2 \backslash X_1$	0	1
0	$(1-\theta)^2$	$\theta(1-\theta)$
1	$\theta(1-\theta)$	θ^2

$E(T) = 0$

DIST OF T:

T	0	1	-1
	$\theta^2(1-\theta)^2$	$\theta(1-\theta)$	$\theta(1-\theta)$

$\Rightarrow P(T=0) = \theta^2(1-\theta) < 1 \Rightarrow T$ IS NOT COMPLETE.

(31)

WHY IS THIS IMPORTANT ?

$E(W) = \theta$

IF WE HAVE W AND UNBIASED ESTIMATOR OF θ AND T SUFFICIENT $\Rightarrow$ $W^* = E(W|T)$ IS UNBIASED AND $Var(W^*) \leq Var(W)$ (RAO-BLACKWELL).

IF T IS ALSO COMPLETE $\Rightarrow$ W^* IS UNIQUE (UMVUE) (EX 6.36 HW).

ANCILLARY AND COMPLETE STATISTICS.

DEF A STATISTIC $S(\underline{X})$ IS AN ANCILLARY STATISTIC IF AND ONLY IF THE DISTRIBUTION OF $S(\underline{X})$ DOES NOT DEPEND ON θ

* S ALONE DOES NOT PROVIDE INFO ON θ BUT JOINTLY WITH OTHER STATISTICS, IT DOES !

EX. SUPPOSE N IS A RANDOM VARIABLE TAKING VALUES 1, 2, 3 . . WITH PROBABILITIES $p_1, p_2, p_3, p_4 \ldots$ $\sum p_i = 1$

HAVING OBSERVED $N = n$, WE PERFORM n BERNOULLI TRIALS WITH PROB. θ, GETTING X SUCCESSES. θ PARAMETER OF INTEREST.

NOTICE THAT N IS ANCILLARY $Pr[N=n] = p_n$ - DO NOT DEPEND ON θ.

IF WE CONSIDER $T = (X, N)$, T IS MINIMAL SUFFICIENT

$X|N=n$ IS A BINOMIAL (n, θ) $\binom{n}{x} \theta^x (1-\theta)^{n-x}$

$\Rightarrow$

$f(X=x, N=n) = \binom{n}{x} \theta^x (1-\theta)^{n-x} p_n$. FOR ANOTHER SAMPLE POINT $X = x_1, N = n_1$ $f(X=x_1, N=n_1) = \binom{n_1}{x_1} \theta^{x_1} (1-\theta)^{n_1 - x_1} p_{n_1}$

THE RATIO

$$\frac{f(X=x, N=n)}{f(X=x_1, N=n_1)} = \frac{\binom{n}{x}}{\binom{n_1}{x_1}} \theta^{x-x_1} (1-\theta)^{(n-n_1)-(x-x_1)} \frac{p_n}{p_{n_1}} = C \quad \text{IFF} \quad \begin{matrix} x = x_1 \\ n = n_1 \end{matrix}$$

$\Rightarrow$ T IS MINIMAL SUFF.

IN GENERAL, AN ANCILLARY STATISTIC $S(\underline{X})$ IS NOT INDEPENDENT OF A SUFFICIENT STATISTIC.

EX: SKIP.

LET $X_1, X_2 \ldots X_n$ BE IID OBSERVATIONS SUCH THAT $X_i \sim f(x_i - \theta)$ LOCATION FAMILY OF DIST. (THINK OF $N(\theta, 1)$)

SHOW $R = X_{(n)} - X_{(1)}$ IS ANCILLARY.

NOTICE THAT $Z_1 = X_1 - \theta$; $Z_2 = X_2 - \theta, \ldots Z_n = X_n - \theta$ ARE SUCH THAT $Z \sim f(z)$ (THINK OF $N(0,1)$) WHERE $f(z)$ DOES NOT DEPEND ON θ.

$$\Rightarrow P(R \le r) = P(X_{(n)} - X_{(1)} \le r)$$
$$= P((X_{(n)} - \theta) - (X_{(1)} - \theta) \le r)$$
$$= P(Z_{(n)} - Z_{(1)} \le r)$$

THIS LAST PROB. ONLY DEPENDS ON $f(z) \Rightarrow R$ IS ANCILLARY

A STATISTIC THAT IS NOT ANCILLARY.

IF $X_1, X_2 \ldots X_n$ ARE IID $N(\mu, \sigma^2)$ $T = \bar{X}$

$\Rightarrow \bar{X} \sim N(\mu, \sigma^2/n) \Rightarrow \bar{X}$ IS NOT ANCILLARY.

EX. SUPPOSE X_1, X_2 ARE IID OBSERVATIONS FROM THE pdf $f(x|\alpha) = \alpha x^{\alpha-1} e^{-x^\alpha}$ (EX 6.13)

SHOW THAT $T = \log X_1 / \log X_2$ IS AN ANCILLARY STATISTIC

LETS CONSIDER $Y_1 = \alpha \log X_1$, $Y_2 = \alpha \log X_2$. FIND $\hat{f}(y_1, y_2)$

$$f(x_1, x_2 | \alpha) = \alpha^2 \left(\prod_{i=1}^{2} x_i\right)^{\alpha-1} e^{-\sum_1^2 x_i^\alpha}$$

$\Rightarrow x_1 = e^{y_1/\alpha}, x_2 = e^{y_2/\alpha}$ $\quad J = 1/\alpha^2 \, e^{(y_1+y_2)/\alpha}$

$$\Rightarrow f(y_1, y_2) = \cancel{\alpha^2} \left(e^{y_1+y_2/\alpha}\right)^{\alpha - \cancel{1}} e^{-\sum (e^{y_i/\alpha})^\alpha} \frac{1}{\cancel{\alpha^2}} \cancel{e^{(y_1+y_2)/\alpha}}$$

$$= e^{y_1+y_2} e^{-(e^{y_1}+e^{y_2})} \leftarrow \text{DOES NOT DEPEND ON } \alpha$$

$\Rightarrow$ THE DIST OF $\dfrac{\alpha \log x_1}{\alpha \log x_2} = \dfrac{\log x_1}{\log x_2}$ WE'LL NOT DEPEND ON α.

IN GENERAL, IT IS HARD TO CHECK IF A STATISTIC T IS COMPLETE. A USEFUL RESULT:

IF WE HAVE OBSERVATIONS FROM THE EXPONENTIAL FAMILY

$$f(x|\theta) = h(x)\, c(\theta) \exp\left(\sum_{j=1}^{k} w_j(\theta) t_j(x)\right)$$

$$T(\underline{X}) = \left(\sum_{i=1}^{n} t_1(x_i), \sum_{i=1}^{n} t_2(x_i), \ldots, \sum_{i=1}^{n} t_k(x_i)\right)$$ IS A

COMPLETE STATISTIC AS LONG AS $(w_1(\theta), w_2(\theta), \ldots w_k(\theta))$ CONTAINS AN OPEN SET IN $\mathbb{R}^k$

EX: (6.22) LET $X_1, X_2 \ldots X_n$ BE A RANDOM SAMPLE WITH pdf $f(x|\theta) = \theta x^{\theta-1}$, $0<x<1$, $\theta>0$

FIND A COMPLETE SUFFICIENT STATISTIC FOR θ

NOTICE THAT

$$f(x|\theta) = \theta e^{(\theta-1)\ln x} \quad 0<x<1,\ \theta>0$$

$$= e^{-\ln x}\,\theta\, e^{\theta \ln x}$$

THIS IS IN THE EXP. FAMILY WHERE $h(x) = e^{-\ln x}$, $c(\theta)=\theta$, $t_1(x) = \ln x$ $\Rightarrow$ $T = \sum_{i=1}^{n} \ln(x_i)$ IS SUFFICIENT AND COMPLETE

$w_1(\theta) = \theta$ WHICH CONTAINS THE OPEN SET $(0,\infty)$

SEE EX 6.15 FOR A COUNTEREXAMPLE WHERE $(w_1(\theta), w_2(\theta))$ DOES NOT CONTAIN AN OPEN SET IN $\mathbb{R}^2$.

MAIN RESULT WITH COMPLETENESS AND MINIMAL SUFFICIENT STATISTIC IS BASU'S THEOREM. PAGE 287

IF $T(\underline{X})$ IS MINIMAL SUFFICIENT AND COMPLETE AND $S(\underline{X})$ IS ANCILLARY $\Rightarrow$ $T(\underline{X})$ IS INDEPENDENT OF $S(\underline{X})$

Ex:

$X_1, X_2 \ldots X_n$ are iid $N(\mu, \sigma^2)$, ASSUME σ^2 KNOWN.
SINCE THE NORMAL MODEL BELONGS TO THE EXP. FAMILY
$T = \bar{X}$ IS MINIMAL, SUFFICIENT AND COMPLETE.
ALSO, WE KNOW $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ $\Rightarrow$ THE DIST OF
S^2 DOES NOT DEPEND ON μ. S^2 IS ANCILLARY!
By BASU'S THEOREM $\bar{X}$ IS INDEPENDENT OF S^2
(COMPARE TO WHAT WE DID IN CHAP 5)

FINAL EX:

SUP. $X_1, X_2 \ldots X_n$ ARE iid OBSERVATIONS WITH

$$f(x|\theta) = \frac{\theta}{(1+x)^{1+\theta}}, \quad 0 < x < \infty \quad \theta > 0$$

FIND A COMPLETE SUFFICIENT STATISTIC

SINCE $f(x|\theta) = \theta \exp(-(1+\theta)\log(1+x))$
$\Rightarrow f(x|\theta)$ BELONGS TO THE EXPONENTIAL FAMILY ($c(\theta) = \theta$,
$h(x) = 1$, $w(\theta) = -(1+\theta)$ $t(x) = \log(1+x)$).
THE PARAMETER $W(\theta)$ CONTAINS AN OPEN SET IN $\mathbb{R}$
$(-\infty, -1)$ $\Rightarrow$ $T(\underset{\sim}{x}) = \sum_{i=1}^{n} \log(1+x_i)$ IS SUFFICIENT AND
COMPLETE.

HOW TO SHOW THAT SOME STATISTIC T IS NOT COMPLETE
FOR A GIVEN θ, A SPECIFIC $g(\cdot)$. SHOW THAT $E g(T) = 0$
$\not\Rightarrow$ $g(T) = 0$

<u>Ex.</u> LET $X_1, X_2 \ldots X_n$ BE iid RVs WITH A $U(\theta, \theta+1)$
WE SAW THAT $T = (X_{(1)}, X_{(n)})$ IS MINIMAL SUFFICIENT
FOR θ

$$F(x) = \int_{\theta}^{x} 1\,du = x - \theta$$

I WILL TRY TO FIND SOME FUNCTION OF $T=(X_{(1)},X_{(n)})$,
FOR EXAMPLE, $g(T)=aX_{(n)}+bX_{(1)}+c$ SUCH THAT $E(g(T))=0$
$\not\Rightarrow g(T)=0$

TO EVALUATE $E(g(T))$, I NEED $E(X_{(1)})$ AND $E(X_{(n)})$

$P[X_{(1)}\le x] = 1-P(X_{(1)}\geq x) = 1-\prod_{i=1}^{n}P(X_i \geq x)$

$= 1-(1-(x-\theta))^n$

$f_{X_{(1)}}(x) = P[X_{(1)}\le x]' = n(1-(x-\theta))^{n-1}$ FOR $\theta\le x\le\theta+1$

ALSO, $f_{X_{(n)}}(x) = n(x-\theta)^{n-1}$; $\theta\le x\le\theta+1$ $\quad u=1-(x-\theta)$

(USING THEO 5.4.4 ON PAGE 229) $\quad x=1-u+\theta$

$\Rightarrow$

$$E(X_{(1)}) = \int_{\theta}^{\theta+1} nx\,(1-(x-\theta))^{n-1}dx = \theta - \left(\frac{n}{n+1}\right)+1$$

$\int_0^1 n(1-u+\theta)u^{n-1}du = \theta-\left(\frac{n}{n+1}\right)+1$

$$E(X_{(n)}) = \int_{\theta}^{\theta+1} nx\,(x-\theta)^{n-1}dx = \theta+\frac{n}{(n+1)}$$

$u=x-\theta$

$\int_0^1 n(u+\theta)u^{n-1}du = \frac{nu^{n+1}}{n+1}+\frac{n\theta u^n}{n}\Big|_0^1 = \frac{n}{n+1}+\theta$

IF WE MAKE $g(T) = X_{(n)}-X_{(1)}-\frac{2n}{n+1}+1$

$\Rightarrow E(g(T))=0$ BUT $g(T)$ IS NOT IDENTICALLY.

THE LIKELIHOOD FUNCTION

LET $f(\underline{x}|\theta)$ BE THE JOINT pdf OR pmf OF $\underline{X}$

IF $\underline{X}=\underline{x}$ IS OBSERVED, THE FUNCTION OF θ DEFINED

BY

$$L(\theta|\underline{x}) = f(\underline{x}|\theta) \quad \forall\theta\in\Theta$$

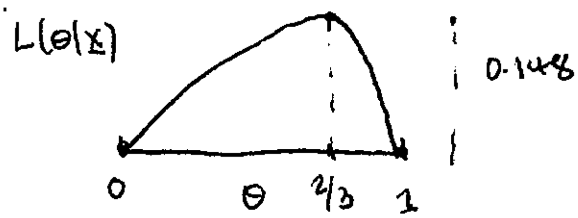
IS THE LIKELIHOOD FUNCTION OF θ.

EX: X_1, X_2, X_3 ~~iid~~ iid RVs. BERNOULLI(θ)

SUPPOSE $\sum_{i=1}^{3} x_i = 2 \Rightarrow L(\theta|\underline{x}) = f(\underline{x}|\theta) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}$

$\Rightarrow L(\theta|\underline{x}) = \theta^2(1-\theta)^{3-2} = \theta^2(1-\theta)^1$ FOR $0 \leq \theta \leq 1$

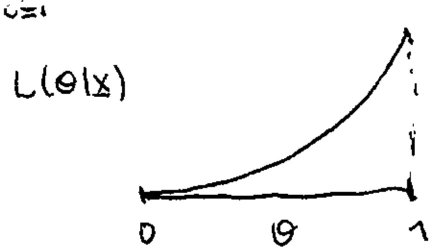
PLOT:



THE CURVE IS GIVING INFORMATION ON THE VALUES OF θ THAT ARE MORE "PLAUSIBLE". (θ AROUND $2/3$) UNDER THE SAMPLE

IF $\sum_{i=1}^{3} x_i = 3 \Rightarrow L(\theta|\underline{x}) = \theta^3, \; 0 < \theta < 1$

PLOT:

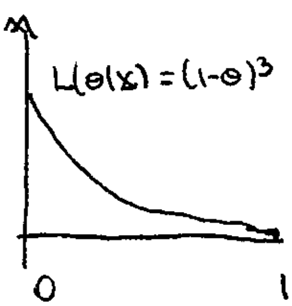


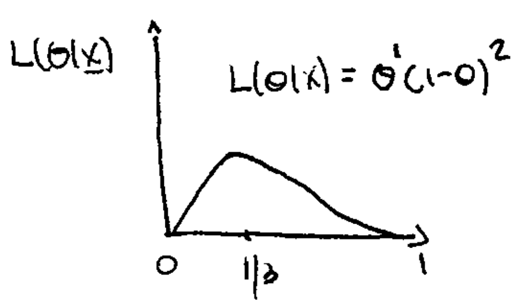
THE MORE PLAUSIBLE VALUES ARE AROUND 1

NOTICE THAT, $L(\theta|\underline{x})$ IS NOT A pdf ON Θ. FOR THE LAST EX. $\int_0^1 L(\theta|\underline{x})d\theta = \int_0^1 \theta^3 d\theta = \frac{\theta^4}{4} = \frac{1}{4} \neq 1$

THE OTHER TWO CASES $\sum_{i=1}^{3} x_i = 0 \Rightarrow L(\theta|\underline{x}) = (1-\theta)^3$

$\sum_{i=1}^{3} x_i = 1$





LIKELIHOOD PRINCIPLE

IF WE HAVE TWO EXPERIMENTS E1, E2 AND $\underline{X}1$ AND $\underline{X}2$ ARE SAMPLE POINTS FROM E1, E2 RESPECTIVELY SUCH THAT :

$$L(\theta|\underline{X}_2) = C(\underline{X}_1,\underline{X}_2)\, L(\theta|\underline{X}_1)$$

THEN, THE CONCLUSIONS DRAWN ABOUT θ SHOULD BE IDENTICAL.

Ex.

SUPPOSE THAT E1 IS SUCH THAT X_1 FOLLOWS A BINOMIAL DIST WITH $n=10$ AND θ BEING THE PROB OF SUCCESSES, i.e.

$$f(x_1|\theta) = \binom{10}{x_1}\theta^{x_1}(1-\theta)^{10-x_1}, \quad 0 \le x_1 \le 10$$

ALSO, SUPPOSE THAT E2 IS SUCH THAT X2 FOLLOWS A NEGATIVE BINOMIAL DIST WITH $r=7$ AND THE SAME PROB. OF SUCCESS θ, i.e.

$$f(x_2|\theta) = \binom{7+x_2-1}{x_2}\theta^7(1-\theta)^{x_2}; \quad x_2=0,1,\ldots$$

NOW SUPPOSE WE OBSERVE $X1=7$ AND $X2=3$

THEN, UNDER EXP 1

$$L(\theta|x_1=7) = \binom{10}{7}\theta^7(1-\theta)^3$$

UNDER EXP 2

$$L(\theta|x_2=3) = \binom{9}{3}\theta^7(1-\theta)^3$$

SO

$$L(\theta|x_1=7) = \frac{\binom{10}{7}}{\binom{9}{3}} L(\theta|x_2=3) \Rightarrow \text{CONCL. ON } \theta \text{ MUST BE THE SAME.}$$