Things that make you say hmmm…

My version of why we care about complete, sufficient statistics

Sam MaWhinney
Department Seminar
10/19/11

Outline (yes, really)

- Motivating Example (Flip coin 3 times)
- Sufficient statistics
- Minimal sufficient statistics
- Ancillary and conditionally sufficient statistics
- Complete statistics
- Relationship between Sufficiency and Completeness

Flip a coin 3 times

Sample Space

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

- Pr(H) = p and Pr(T) = (1-p)

P(HHH) = p^3
P(HHT) = P(HTH) = P(THH) = $p^2(1-p)$
P(TTH) = P(THT) = P(HTT) = $p(1-p)^2$
P(TTT) = $(1-p)^3$

Flip a coin 3 times

Sample Space

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

- Random Variable

X_i = 1 if Head
X_i = 0 if Tail (i = 1,2,3)

$T(\mathbf{X}) = \Sigma X_i$ = Number of Heads

$T(\mathbf{X}) \sim$ Binomial(3,p)

Flip a coin 3 times

Sample Space

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

- $T(\mathbf{X}) = \Sigma X_i$ = Number of Heads
- $T(\mathbf{X}) \sim$ Binomial(3,p)

$Pr(T(\mathbf{X})=3) = p^3$
$Pr(T(\mathbf{X})=2) = 3p^2(1-p)$
$Pr(T(\mathbf{X})=1) = 3p(1-p)^2$
$Pr(T(\mathbf{X})=0) = (1-p)^3$

Flip a coin 3 times

Sample Space

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

- $T(\mathbf{X}) = \Sigma Xi$ = Number of Heads
- $T(\mathbf{X}) \sim$ Binomial(3,p)

$Pr(T(\mathbf{X})=3) = p^3$
$Pr(T(\mathbf{X})=2) = 3p^2(1-p)$
$Pr(T(\mathbf{X})=1) = 3p(1-p)^2$
$Pr(T(\mathbf{X})=0) = (1-p)^3$

Flip a coin 3 times

Sample Space

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

- $T(\mathbf{X}) = \Sigma Xi$ = Number of Heads
- $T(\mathbf{X}) \sim$ Binomial(3,p)

$Pr(T(\mathbf{X})=3) = p^3$
$Pr(T(\mathbf{X})=2) = 3p^2(1-p)$
$Pr(T(\mathbf{X})=1) = 3p(1-p)^2$
$Pr(T(\mathbf{X})=0) = (1-p)^3$

Flip a coin 3 times

Sample Space

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

- $T(\mathbf{X}) = \Sigma Xi$ = Number of Heads
- $T(\mathbf{X}) \sim$ Binomial(3,p)

$Pr(T(\mathbf{X})=3) = p^3$
$Pr(T(\mathbf{X})=2) = 3p^2(1-p)$
$Pr(T(\mathbf{X})=1) = 3p(1-p)^2$
$Pr(T(\mathbf{X})=0) = (1-p)^3$

Flip a coin 3 times

Sample Space

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

- T(**X**) = Σ Xi = Number of Heads
- T(**X**) ~ Binomial(3,p)

Pr(T(**X**)=3) = p^3
Pr(T(**X**)=2) = $3p^2(1-p)$
Pr(T(**X**)=1) = $3p(1-p)^2$
Pr(T(**X**)=0) = $(1-p)^3$

Hmmm…..

- Now assume that we know T(**X**)=2 and want to find

Pr(HHT | T(**X**)=2) =

Pr((HHT) and T(**X**)=2) / Pr(T(**X**)=2) =

Pr(HHT) / Pr(T(**X**)=2) =

$p^2(1-p) / 3p^2(1-p) = 1/3$

Similarly Pr(HTH) = Pr(THH) = 1/3

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

Hmmm…

The dist'n of any outcome in S, given T(**X**) is independent of p.

Sample	T(X)	P(**X**|T(**X**)=3)	P(**X**|T(**X**)=2)	P(**X**|T(**X**)=1)	P(**X**|T(**X**)=0)
HHH	3	1	0	0	0
HHT	2	0	1/3	0	0
HTH	2	0	1/3	0	0
THH	2	0	1/3	0	0
TTH	1	0	0	1/3	0
THT	1	0	0	1/3	0
HTT	1	0	0	1/3	0
TTT	0	0	0	0	1

Hmmm…

The dist'n of any outcome in S, given T(**X**) is independent of p.

Sample	T(X)	P(**X**|T(**X**)=3)	P(**X**|T(**X**)=2)	P(**X**|T(**X**)=1)	P(**X**|T(**X**)=0)
HHH	3	1	0	0	0
HHT	2	0	1/3	0	0
HTH	2	0	1/3	0	0
THH	2	0	1/3	0	0
TTH	1	0	0	1/3	0
THT	1	0	0	1/3	0
HTT	1	0	0	1/3	0
TTT	0	0	0	0	1

Hmmm… The dist'n of any outcome in S, given T(**X**) is independent of p.

Sample	T(**X**)	P(**X**|T(**X**)=3)	P(**X**|T(**X**)=2)	P(**X**|T(**X**)=1)	P(**X**|T(**X**)=0)
HHH	3	1	0	0	0
HHT	2	0	1/3	0	0
HTH	2	0	1/3	0	0
THH	2	0	1/3	0	0
TTH	1	0	0	1/3	0
THT	1	0	0	1/3	0
HTT	1	0	0	1/3	0
TTT	0	0	0	0	1

Hmmm… The dist'n of any outcome in S, given T(**X**) is independent of p.

Sample	T(**X**)	P(**X**|T(**X**)=3)	P(**X**|T(**X**)=2)	P(**X**|T(**X**)=1)	P(**X**|T(**X**)=0)
HHH	3	1	0	0	0
HHT	2	0	1/3	0	0
HTH	2	0	1/3	0	0
THH	2	0	1/3	0	0
TTH	1	0	0	1/3	0
THT	1	0	0	1/3	0
HTT	1	0	0	1/3	0
TTT	0	0	0	0	1

Hmmm… The dist'n of any outcome in S, given T(**X**) is independent of p.

Sample	T(**X**)	P(**X**|T(**X**)=3)	P(**X**|T(**X**)=2)	P(**X**|T(**X**)=1)	P(**X**|T(**X**)=0)
HHH	3	1	0	0	0
HHT	2	0	1/3	0	0
HTH	2	0	1/3	0	0
THH	2	0	1/3	0	0
TTH	1	0	0	1/3	0
THT	1	0	0	1/3	0
HTT	1	0	0	1/3	0
TTT	0	0	0	0	1

Sufficient Statistic

- The dist'n of **X** given T(**X**) does not involve the parameter p.
- True whenever we condition on a sufficient statistic
  - T is for sufficient
- Def'n of Sufficient (C&B):

  *A statistic T(**X**) is a sufficient statistic for θ if the conditional dist'n of the sample **X** given the value T(**X**) does not depend on θ.*

Sufficiency Principle (C&B):

If $T(\mathbf{X})$ is a sufficient statistic for θ then any inference about θ should depend on the sample $\mathbf{X}$ only through the value $T(\mathbf{X})$.

That is, if $\mathbf{x}$ and $\mathbf{y}$ are two sample points such that $T(\mathbf{x})=T(\mathbf{y})$, then the inference about θ should be the same whether $\mathbf{X}=\mathbf{x}$ or $\mathbf{Y}=\mathbf{y}$ is observed.

	x	*y*	
HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

Hmmm…. Sufficiency and Data Partition

We make our inference based on how we partition our data.

The sufficient statistic, ΣX_i, defines the partition of S.

Any sample in the same partition makes the same inference.

$\Sigma X_i = 3$	$\Sigma X_i = 2$		
HHH	HHT	HTH	THH
TTH	THT	HTT	TTT
	$\Sigma X_i = 1$		$\Sigma X_i = 0$

Sufficient Statistics

- Now we can base our inferences on T(**X**) (T is for sufficient).
- ideally (although not necessarily) T(**X**) is of smaller dimension than **X**
  - and ideally the same dimension as θ….
- How do we find sufficient statistics?

Finding Sufficient Statistics

Theorem (C&B, short hand):

If $\mathbf{X} \sim p(\mathbf{x}\,|\,\theta)$ and $T(\mathbf{X}) \sim q(\mathbf{t}\,|\,\theta)$

Then for every $\mathbf{x}$ in the sample space:

$T(\mathbf{X})$ sufficient $\rightarrow$ *$p(\mathbf{x}\,|\,\theta) / q(T(\mathbf{X})\,|\,\theta)$ constant wrt θ*

$T(\mathbf{X})$ sufficient $\leftarrow$ *$p(\mathbf{x}\,|\,\theta) / q(T(\mathbf{X})\,|\,\theta)$ constant wrt θ*

Coin flip example

- $p(\mathbf{x}|\theta) = p^{\Sigma x}(1-p)^{3-\Sigma x}$
- $q(\mathbf{t}|\theta) = \binom{3}{\Sigma x}\, p^{\Sigma x}(1-p)^{3-\Sigma x}$
- $p(\mathbf{x}|\theta) / q(\mathbf{t}|\theta) = 1/\binom{3}{\Sigma x}$
- Déjà vu (Hmmm…) : Pr(HHT | T(**X**)=2) = **…**
  Pr(HHT) / Pr(T(**X**)=2) =
  $p(X=\mathbf{x}|\theta) / q(T=\mathbf{t}|\theta) = p^2(1-p) / 3p^2(1-p) = 1/3$

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

Finding Sufficient Statistics, T(**X**)

- In previous theorem, we need T(**X**) (data partition) and dist'n of T(**X**)… what if no obvious choice for T(**X**)?
- Factorization Thm (C&B) (short hand):

  *Sample **X**~ $f(\mathbf{x}|\theta)$, i.e. $f(\mathbf{x}|\theta)$ is the joint dist'n*

  *T(**X**) sufficient* $\rightarrow$ $f(\mathbf{x}|\theta)=g(T(\mathbf{x})|\theta)h(\mathbf{x})$

  *T(**X**) sufficient* $\leftarrow$ $f(\mathbf{x}|\theta)=g(T(\mathbf{x})|\theta)h(\mathbf{x})$

  *Functions $g(t|\theta)$ and $h(\mathbf{x})$ exist for all **x** and all parameter points θ.*

Coin flip example

*T(**X**) sufficient* $\rightarrow$ $f(\mathbf{x}|\theta)=g(T(\mathbf{x})|\theta)h(\mathbf{x})$

*T(**X**) sufficient* $\leftarrow$ $f(\mathbf{x}|\theta)=g(T(\mathbf{x})|\theta)h(\mathbf{x})$

$f(\mathbf{x}|\theta) = p^{\Sigma x}(1-p)^{3-\Sigma x}$

$g(T(\mathbf{x})|\theta) = p^{\Sigma x}(1-p)^{3-\Sigma x}$

$h(x) = 1$

$T(\mathbf{x}) = \Sigma x$ = sufficient

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

Easiest* way of finding Sufficient Statistics

Exponential Family Thm (C&B)
(short hand and, for simplicity, assuming **one** parameter):

$X_1 \ldots X_n \sim f(x|\theta)$, iid

$f(x|\theta) = h(x)\, c(\theta) \exp(w(\theta)\, t(x))$ *(note: for an individual X)*

Then

$T(X) = \sum_{j=1}^{n} t(x_j)$

is a sufficient statistic for θ.

*Assuming exponential family, important information for qualifying exams.

Coin flip example

$f(x|\theta)= h(x)\, c(\theta) \exp(w(\theta) t(x))$

$T(X) = \Sigma t(x_j)$ is sufficient

$f(x|\theta) = p^x(1-p)^{1-x}$ (one x, Bernoulli dist'n)

$= (p/(1-p))^x (1-p)^1$

$= (1-p) \exp\{x \log (p/(1-p))\}$

$t(x)=x$, $h(x) =1$,

$c(\theta)=(1-p)$, $w(\theta)=\log(p/(1-p))$

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

Coin flip example

$f(x|\theta)= h(x)\, c(\theta) \exp(w(\theta) t(x))$

$T(X) = \Sigma t(x_j)$ is sufficient

$f(x|\theta) = p^x(1-p)^{1-x}$ (one x, Bernoulli dist'n)

$= (p/(1-p))^x (1-p)^1$

$= (1-p) \exp\{x \log (p/(1-p))\}$

$t(x)=x$, $h(x) =1$,

$c(\theta)=(1-p)$, $w(\theta)=\log(p/(1-p))$

$T(\mathbf{x}) = \Sigma x =$ sufficient

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

Minimal Sufficient Statistics

- We can use the factorization theorem to argue that the sample is sufficient.

  $f(\mathbf{x}|\theta)=g(T(\mathbf{x})|\theta)h(\mathbf{x})$

  *Define $h(\mathbf{x})=1$ **and** $T(\mathbf{x})=\mathbf{x}$*

- For a continuous outcome, we would partion the data by the order statistics

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

Minimal Sufficient Statistics

Any one-to-one function of a sufficient statistic is also sufficient –defines the same partition.

Σx_i defines same partition as $(\Sigma x_i)^2$ or $\Sigma x_i/n$ or Σx_i^2

$(\Sigma X_i)^2 = 9$ $(\Sigma X_i)^2 = 4$

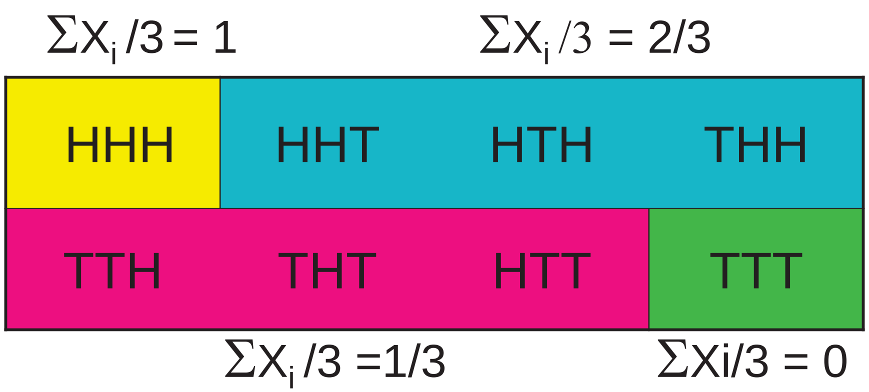
HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

$(\Sigma X_i)^2 = 1$ $(\Sigma X_i)^2 = 0$

Minimal Sufficient Statistics

Any one-to-one function of a sufficient statistic is also sufficient –defines the same partition.

Σx_i defines same partition as $(\Sigma x_i)^2$ or $\Sigma x_i/n$ or Σx_i^2



Minimal Sufficient Statistics

- Numerous sufficient statistics for a problem
- Coin flip:

  (X_1, X_2, X_3) (the complete data) is sufficient

  Could also show: $(X_1, X_2 + X_3)$ is sufficient….

  Any 1-to-1 transformation of ΣX_i is sufficient

  Note ΣX_i is a function of other sufficient stats…

  Hmmmm….

Minimal Sufficient Statistics

- Goal: Most data reduction with no loss of information.
- Def'n (C&B)

  A sufficient statistics $T(\mathbf{X})$ is called a minimal sufficient statistic if, for any other sufficient statistic $T'(\mathbf{X})$, $T(\mathbf{x})$ is a function of $T'(\mathbf{x})$.

Minimal Sufficient Statistics

- Goal: Most data reduction with no loss of information.
- Practical Theorem for identifying minimal sufficient statistics (C&B, short hand):

  Sample $\mathbf{X} \sim f(\mathbf{x}|\theta)$, i.e. $f(\mathbf{x}|\theta)$ is the joint dist'n

  $T(\mathbf{X})$ minimal sufficient $\rightarrow$ $f(\mathbf{x}|\theta) / f(\mathbf{y}|\theta)$ constant wrt θ if $T(\mathbf{x}) = T(\mathbf{y})$.

  $T(\mathbf{X})$ minimal sufficient $\leftarrow$ $f(\mathbf{x}|\theta) / f(\mathbf{y}|\theta)$ constant wrt θ if $T(\mathbf{x}) = T(\mathbf{y})$.

Coin flip example

$T(\mathbf{X})$ minimal sufficient $\leftrightarrow$ *$f(\mathbf{x}|\theta) / f(\mathbf{y}|\theta)$ constant wrt θ if $T(\mathbf{x}) = T(\mathbf{y})$.*

$f(\mathbf{x}|\theta) = p^{\Sigma x}(1-p)^{3-\Sigma x}$

$f(\mathbf{y}|\theta) = p^{\Sigma y}(1-p)^{3-\Sigma y}$

$f(\mathbf{x}|\theta) / f(\mathbf{y}|\theta) = (p/(1-p))^{\Sigma x - \Sigma y}$

$f(\mathbf{x}|\theta) / f(\mathbf{y}|\theta)$ constant wrt θ if $\Sigma x = \Sigma y$

$T(\mathbf{x}) = \Sigma x$ = minimal sufficient

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

Minimal Sufficient Statistics

Any one-to-one function of a **minimal sufficient statistic** is also minimal sufficient –defines the same partition.
Σx_i defines same partition as $(\Sigma x_i)^2$ or $\Sigma x_i/n$ or Σx_i^2

$(\Sigma X_i)^2 = 9$ $(\Sigma X_i)^2 = 4$

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

$(\Sigma X_i)^2 = 1$ $(\Sigma X_i)^2 = 0$

Hmmm….. Coin flipping and dice rolling

- What if…
- Give everyone a die (dice) and the number rolled determined the number times the coin flipped

$q_n = Pr(N=n) = 1/6$; for n = 1,2,3,4,5,6

$f_{n,x}(n,x) = q_n * f(x|n)$

$= (1/6) * \prod^n p^x(1-p)^{n-x} = (1/6) * p^{\Sigma x}(1-p)^{n-\Sigma x}$

$= (1/6) * (p/(1-p))^{\Sigma x} (1-p)^n$

$f(\mathbf{x},\mathbf{n}|\theta) / f(\mathbf{y},\mathbf{m}|\theta) = \dfrac{(1/6)\,(1-p)^n\,(p/1-p)^{\Sigma^n x}}{(1/6)\,(1-p)^m\,(p/1-p)^{\Sigma^m y}}$

Need both n=m and $\Sigma x = \Sigma y$ to be constant wrt p

n and Σx are minimal sufficient (#heads not enough)

Here n is an **ancillary statistic**… (another def'n required)

Ancillary Statistics

- Same Goal: Most data reduction with no loss of info.
- Ancillary Statistics (C&B):

  A statistic $S(\mathbf{X})$ whose dist'n doesn't depend on θ is called ancillary. (S is for ancillary)

  We may have a minimal sufficient statistic, $T' = (n, \Sigma x)$ with dimension(T') > dimension(θ).

  Coin flip: $T' = (n, \Sigma x)$, $\theta = Pr(H) = p$

  We can think of Σx as 'conditionally sufficient' because it is used as a sufficient statistic in inference conditional on N=n.

  An ancillary statistic is an important component of the minimal sufficient statistic… **Hmmmm….**

Complete Statistics

- Intuition: We'd like our minimal sufficient statistic to be independent of ancillary statistics.
- Occurs with completeness….
- Completeness Def'n (C&B):

Let $f(t|\theta)$ be a family of pdfs or pmfs for a statistic $T(\mathbf{X})$. The family of probability dist'ns is called complete if

$E_\theta[g(T)] = 0$ *for all* $\theta \rightarrow P_\theta(g(T)=0)=1$ *for all* θ.

Equivalently, $T(\mathbf{X})$ is called a complete statistic.

Completeness is a property of a family of dist'ns.

Hmmm….. Or Huh?!?

Coin Flip Example: Completeness

- $E_\theta[g(T)] = 0$ *for all* $\theta \rightarrow P_\theta(g(T)=0)=1$ *for all* θ.
- T(X) ~ binomial(3,p) (assuming n=3 known)
- $q(\mathbf{t}|\theta) = \binom{3}{\Sigma x} p^{\Sigma x}(1-p)^{3-\Sigma x}$
- $E[g(T)] = g(0)*1*p^0(1-p)^{3-0} + g(1)*3*p^1(1-p)^{3-1} + g(2)*3*p^2(1-p)^{3-2} + g(3)*1*p^3(1-p)^{3-3}$

$$= (1-p)^3 [g(0)*1*(p/(1-p))^0 + g(1)*3*(p/(1-p))^1 + g(2)*3*(p/(1-p))^2 + g(3)*1*(p/(1-p))^3]$$

=? 0 → g(0)=g(1)=g(2)=g(3) = 0

Coin Flip Example: Completeness

- $E_\theta[g(T)] = 0$ *for all* $\theta \rightarrow P_\theta(g(T)=0)=1$ *for all* θ.
- $T(\mathbf{X})$ ~ binomial(3,p) (assuming n=3 known)
- $q(\mathbf{t}|\theta) = \binom{3}{\Sigma x} p^{\Sigma x}(1-p)^{3-\Sigma x} = \binom{3}{t} p^t (1-p)^{3-t}$
- $E[g(T)] = g(0)\, 1*p^0(1-p)^{3-0} + g(1)\, 3*p^1(1-p)^{3-1} + g(2)\, 3*p^2(1-p)^{3-2} + g(3)\, 1*p^3(1-p)^{3-3}$

$$= (1-p)^3 [g(0)(p/(1-p))^0 + g(1)3(p/(1-p))^1 + g(2)3(p/(1-p))^2 + g(3)(p/(1-p))^3] =? 0$$

since $(1-p)^3 \neq 0$

Coin Flip Example: Completeness

- $E_\theta[g(T)] = 0$ *for all* $\theta \rightarrow P_\theta(g(T)=0)=1$ *for all* θ.
- T(X) ~ binomial(3,p) (assuming n=3 known)
- $q(\mathbf{t}|\theta) = \binom{3}{\Sigma x} p^{\Sigma x}(1-p)^{3-\Sigma x}$
- $E[g(T)] = 0 \rightarrow g(0)\, 1*p^0(1-p)^{3-0} + g(1)\, 3*p^1(1-p)^{3-1} + g(2)\, 3*p^2(1-p)^{3-2} + g(3)\, 1*p^3(1-p)^{3-3}$

$$\rightarrow 0 \;?= g(0)*1 + g(1)*3*(p/(1-p))^1 + g(2)*3*(p/(1-p))^2 + g(3)*1*(p/(1-p))^3$$

Coin Flip Example: Completeness

- *$E_\theta[g(T)]= 0$ for all $\theta \rightarrow P_\theta(g(T)=0)=1$ for all θ.*
- T(X) ~ binomial(3,p) (assuming n=3 known)
- $q(\mathbf{t}|\theta) = \binom{3}{\Sigma x} p^{\Sigma x}(1-p)^{3-\Sigma x}$
- $E[g(T)] = g(0)\ 1*p^0(1-p)^{3-0} + g(1)\ 3*p^1(1-p)^{3-1} + g(2)\ 3*p^2(1-p)^{3-2} + g(3)\ 1*p^3(1-p)^{3-3}$

$$\rightarrow 0 \ ?= g(0)*1 + g(1)*3*(p/(1-p))^1 + g(2)*3*(p/(1-p))^2 + g(3)*1*(p/(1-p))^3$$

≠ 0

Coin Flip Example: Completeness

- *$E_\theta[g(T)]= 0$ for all $\theta \rightarrow P_\theta(g(T)=0)=1$ for all θ.*
- T(X) ~ binomial(3,p) (assuming n=3 known)
- $q(\mathbf{t}|\theta) = \binom{3}{\Sigma x} p^{\Sigma x}(1-p)^{3-\Sigma x}$
- $E[g(T)] = g(0)\ 1*p^0(1-p)^{3-0} + g(1)\ 3*p^1(1-p)^{3-1} + g(2)\ 3*p^2(1-p)^{3-2} + g(3)\ 1*p^3(1-p)^{3-3}$

$$\rightarrow 0 \ ?= g(0)*1 + g(1)*3*(p/(1-p))^1 + g(2)*3*(p/(1-p))^2 + g(3)*1*(p/(1-p))^3$$

≠ 0 for all p

Coin Flip Example: Completeness

- *$E_\theta[g(T)]= 0$ for all $\theta \rightarrow P_\theta(g(T)=0)=1$ for all θ.*
- T(X)= Σx ~ binomial(3,p) (assuming n=3 known)
- $q(\mathbf{t}|\theta) = \binom{3}{\Sigma x} p^{\Sigma x}(1-p)^{3-\Sigma x}$
- $E[g(T)] = g(0)\ 1*p^0(1-p)^{3-0} + g(1)\ 3*p^1(1-p)^{3-1} + g(2)\ 3*p^2(1-p)^{3-2} + g(3)\ 1*p^3(1-p)^{3-3}$

$$\rightarrow 0 \ ?= g(0)*1 + g(1)*3*(p/(1-p))^1 + g(2)*3*(p/(1-p))^2 + g(3)*1*(p/(1-p))^3$$

g(0)=g(1)=g(2)=g(3)=0

Coin Flip Example: Completeness

- *$E_\theta[g(T)]= 0$ for all $\theta \rightarrow P_\theta(g(T)=0)=1$ for all θ.*
- T(X)= Σx ~ binomial(3,p) (assuming n=3 known)
- $q(\mathbf{t}|\theta) = \binom{3}{\Sigma x} p^{\Sigma x}(1-p)^{3-\Sigma x}$
- $E[g(T)] = g(0)\ 1*p^0(1-p)^{3-0} + g(1)\ 3*p^1(1-p)^{3-1} + g(2)\ 3*p^2(1-p)^{3-2} + g(3)\ 1*p^3(1-p)^{3-3}$

$$\rightarrow 0 \ ?= g(0)*1 + g(1)*3*(p/(1-p))^1 + g(2)*3*(p/(1-p))^2 + g(3)*1*(p/(1-p))^3$$

g(0)=g(1)=g(2)=g(3)=0

Σx_i is complete

Coin Flip Example: Completeness

- *$E_\theta[g(T)] = 0$ for all $\theta \rightarrow P_\theta(g(T)=0)=1$ for all θ.*
- $T(X) = \Sigma x \sim$ binomial(3,p) (assuming n=3 known)
- $q(\mathbf{t} | \theta) = \binom{3}{\Sigma x} p^{\Sigma x}(1-p)^{3-\Sigma x}$
- $E[g(T)] = \quad g(0)\, 1^*p^0(1-p)^{3-0} + g(1)\, 3^*p^1(1-p)^{3-1}$
  $\quad + g(2)\, 3^*p^2(1-p)^{3-2} + g(3)\, 1^*p^3(1-p)^{3-3}$

$\rightarrow 0$?= $g(0)^*\, 1$
$+ g(1)^*\, 3 * (p/(1-p))^1$
$+ g(2)^*\, 3 * (p/(1-p))^2$ $\quad g(0)=g(1)=g(2)=g(3)=0$
$+ g(3)^*\, 1 * (p/(1-p))^3$

Σx_i is complete
(and sufficient and minimal sufficient…)

Completeness and Ancillary Statistics

- Basu's Theorem (C&B):

  If $T(\mathbf{X})$ is complete and minimal sufficient statistic;
  *then $T(\mathbf{X})$ is independent of **every** ancillary statistic.*

  Note: if a dist'n is complete, then there are no unbiased estimators of zero, except zero itself…..

Easiest* way of finding Complete Statistics

Exponential Family Thm (C&B)
(short hand and, for simplicity, assuming **one** parameter):

$X_1 \ldots X_n \sim f(x|\theta)$, iid

$f(x|\theta) = h(x)\, c(\theta) \exp(w(\theta) t(x))$ *(note: for an individual X)*

Then

$$T(X) = \sum_{j=1}^{n} t(x_j)$$

is a complete statistic as long as the parameter space Θ contains an open subset in $\mathbb{R}^k$.

*Assuming exponential family, important information for qualifying exams.

Hmmm…. Relationship between Sufficiency and Completeness

- Sufficiency / completeness move in opposite directions*

Let $T(\mathbf{X}) = h(U(\mathbf{X}))$

T and U both statistics (ftns of data). Since T is ft'n of U:
$U(\mathbf{X})$ is a reduction of the data
$T(\mathbf{X}) = h(U(\mathbf{X}))$ is a further reduction.

T sufficient $\rightarrow$ U sufficient

T complete $\leftarrow$ U complete

Recall: A sufficient statistics $T(\mathbf{X})$ is called a minimal sufficient statistic if, for any other sufficient statistic $T'(\mathbf{X})$, $T(\mathbf{x})$ is a function of $T'(\mathbf{x})$.

* Louise Ryan class notes….

Relationship between Sufficiency and Completeness

$T(\mathbf{X}) = h(U(\mathbf{X}))$

$U(\mathbf{X})$: reduction of the data

$T(\mathbf{X}) = h(U(\mathbf{X}))$:further reduction

T sufficient → U sufficient

T complete ← U complete

Sufficient	**X** = full data vector	
Sufficient	U(**X**) a reduction of the data	At some point lose completeness
Minimal Sufficient Further reduction, lose sufficiency	T(**X**) = h(U(**X**)) a further reduction	Complete ?
	further reduction	Complete
	Constant, most extreme reduction	Complete

Relationship between Sufficiency and Completeness

$T(\mathbf{X}) = h(U(\mathbf{X}))$

$U(\mathbf{X})$: reduction of the data

$T(\mathbf{X}) = h(U(\mathbf{X}))$:further reduction

T sufficient → U sufficient

T complete ← U complete

Sufficient	HHH TTH	HHT THT	HTH HTT	THH TTT	Finest partition (8) (X_1, X_2, X_3)
Sufficient	HHH TTH	HHT THT	HTH HTT	THH TTT	U(**X**) = Reduction of data (6 partitions)
Minimal Sufficient / Complete	HHH TTH	HHT THT	HTH HTT	THH TTT	T(**X**) Grouped by ΣX_i # Heads (4 partitions)
Complete	HHH TTH	HHT THT	HTH HTT	THH TTT	$T_1(\mathbf{X})$ = further reduction. (3 partitions)
Complete	HHH TTH	HHT THT	HTH HTT	THH TTT	Coarsest partition (1) $T_2(\mathbf{X})$=Constant for all **X**

Relationship between Sufficiency and Completeness

- Define Random Variable consistent with Finest partition.
- $T_F(\mathbf{x}) = I_{F1}(\mathbf{x}), I_{F2}(\mathbf{x}), I_{F3}(\mathbf{x}), I_{F4}(\mathbf{x}), I_{F5}(\mathbf{x}), I_{F6}(\mathbf{x}), I_{F7}(\mathbf{x}), I_{F8}(\mathbf{x})$

  Where $I_{Fi}(\mathbf{x})$ is an indicator function:

  $I_{Fi}(\mathbf{x})=1$ if in partition i,

  $I_{Fi}(\mathbf{x})= 0$ else i=1,2,…8
- $f(\mathbf{x}|p)= (p^3)^{I_{F1}(\mathbf{x})} * (p^2(1-p))^{I_{F2}(\mathbf{x})+I_{F3}(\mathbf{x})+I_{F4}(\mathbf{x})} *$

  $* (p(1-p)^2)^{I_{F5}(\mathbf{x})+I_{F6}(\mathbf{x})+I_{F7}(\mathbf{x})} * (1-p^3)^{I_{F8}(\mathbf{x})}$

  By Factorization theorem: $T_F(\mathbf{x})$ is sufficient.

1 HHH	2 HHT	3 HTH	4 THH
5 TTH	6 THT	7 HTT	8 TTT

Relationship between Sufficiency and Completeness

- Is $T_F(\mathbf{x})$ complete?

  $E[g(T_F(\mathbf{x}))= \Sigma P(T) g(T) = (p^3) * [g(I_{F1})]$

  $+ (p^2(1-p))* [g(I_{F2})+g(I_{F3})+g(I_{F4})]$

  $+ (p(1-p)^2)* [g(I_{F5})+g(I_{F6})+g(I_{F7})]$

  $+ (1-p)^3 * [g(I_{F8})]$ =? 0 for all p

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

Relationship between Sufficiency and Completeness

- Is $T_F(\mathbf{x})$ complete?

$E[g(T_F(\mathbf{x}))= \Sigma P(T)\, g(T) = (p^3) * [g(I_{F1})]$
$+ (p^2(1-p))* [g(I_{F2})+g(I_{F3})+g(I_{F4})]$
$+ (p(1-p)^2)* [g(I_{F5})+g(I_{F6})+g(I_{F7})]$
$+ (1-p)^3 * [g(I_{F8})] =? 0$ for all p ?

$[g(I_{F1})]=0$

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

Relationship between Sufficiency and Completeness

- Is $T_F(\mathbf{x})$ complete?

$E[g(T_F(\mathbf{x}))= \Sigma P(T)\, g(T) = (p^3) * [g(I_{F1})]$
$+ (p^2(1-p))* [g(I_{F2})+g(I_{F3})+g(I_{F4})]$
$+ (p(1-p)^2)* [g(I_{F5})+g(I_{F6})+g(I_{F7})]$
$+ (1-p)^3 * [g(I_{F8})] =? 0$ for all p ?

$[g(I_{F1})]=0$
$[g(I_{F8})]=0$

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

Relationship between Sufficiency and Completeness

- Is $T_F(\mathbf{x})$ complete?

$E[g(T_F(\mathbf{x}))= \Sigma P(T)\, g(T) = (p^3) * [g(I_{F1})]$
$+ (p^2(1-p))* [g(I_{F2})+g(I_{F3})+g(I_{F4})]$
$+ (p(1-p)^2)* [g(I_{F5})+g(I_{F6})+g(I_{F7})]$
$+ (1-p)^3 * [g(I_{F8})] =? 0$ for all p ?

$[g(I_{F1})]=0$
$[g(I_{F8})]=0$
$[g(I_{F2})+g(I_{F3})+g(I_{F4})]=0$ for all p

if $g(I_{F2})=-[g(I_{F3})+g(I_{F4})]$….
many combinations =0 that don't require g()=0!!

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

Relationship between Sufficiency and Completeness

- Is $T_F(\mathbf{x})$ complete?

$E[g(T_F(\mathbf{x}))= \Sigma P(T)\, g(T) = (p^3) * [g(I_{F1})]$
$+ (p^2(1-p))* [g(I_{F2})+g(I_{F3})+g(I_{F4})]$
$+ (p(1-p)^2)* [g(I_{F5})+g(I_{F6})+g(I_{F7})]$
$+ (1-p)^3 * [g(I_{F8})] =? 0$ for all p ?

$[g(I_{F1})]=0$
$[g(I_{F8})]=0$
$[g(I_{F2})+g(I_{F3})+g(I_{F4})]$*
$[g(I_{F5})+g(I_{F6})+g(I_{F7})]$*
if $g(I_{F5})=-[g(I_{F6})+g(I_{F7})]$….
*many combinations =0 that don't require g()=0!!

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

Relationship between Sufficiency and Completeness

- Is $T_F(\mathbf{x})$ complete?
  - We have separated samples with identical information (Σx_i)

$$
\begin{aligned}
E[g(T_F(\mathbf{x})) = \Sigma\, P(T)\, g(T) = \quad & (p^3) * \ [g(I_{F1})] \\
& + (p^2(1-p)) * [g(I_{F2})+g(I_{F3})+g(I_{F4})] \\
& + (p(1-p)^2) * [g(I_{F5})+g(I_{F6})+g(I_{F7})] \\
& + (1-p)^3 \quad * [g(I_{F8})] =?\ 0 \ \text{ for all } p\ ?
\end{aligned}
$$

*many combinations =0 that don't require g()=0!!

Not complete.

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

Relationship between Sufficiency and Completeness

- Define Random Variable consistent w Coarsest partition.
- $T_C(\mathbf{x}) = I_{C1}(\mathbf{x})$

  Where $I_{C1}(\mathbf{x})$ is an indicator function:

  $I_{C1}(\mathbf{x}) = 1$ if in partition 1,

  $I_{C1}(\mathbf{x}) = 0$ else
- Is $I_{C1}(\mathbf{x})$ sufficient
- $f(\mathbf{x}|p) = (p^3 + 3p^2(1-p) + 3p(1-p)^2 + (1-p)^3)^{IC1(\mathbf{x})} = 1^{IC1(\mathbf{x})} = 1$

By factorization theorem: $T_C(\mathbf{x})$ is NOT sufficient.

Lose sufficiency when combine samples with different info.

1	HHH	HHT	HTH	THH
	TTH	THT	HTT	TTT

Relationship between Sufficiency and Completeness

- Is $I_{C1}(\mathbf{x})$ complete?
- $f(\mathbf{x}|p) = (p^3 + p^2(1-p) + p(1-p)^2 + (1-p)^3)^{IC1(\mathbf{x})} = 1^{IC1(\mathbf{x})} = 1$

$E[g(T_F(\mathbf{x})) = \Sigma\, P(T)\, g(T) = \quad (1) * \ [g(I_{C1})] \ =?\ 0 \ \text{ for all } p$

$\rightarrow [g(I_{C1})] = 0$

$I_{C1}(\mathbf{x})$ IS complete.

We haven't separated anything with identical information (Σx_i)

1	HHH	HHT	HTH	THH
	TTH	THT	HTT	TTT

Relationship between Sufficiency and Completeness

- Random Variable consistent w Minimal Sufficient partition.
- $T_{MS}(\mathbf{x}) = I_{MS1}(\mathbf{x}), I_{MS2}(\mathbf{x}), I_{MS3}(\mathbf{x}), I_{MS4}(\mathbf{x})$

  Where $I_{MSi}(\mathbf{x})$ is an indicator function:

  $I_{MSi}(\mathbf{x}) = 1$ if in partition i,

  $I_{MSi}(\mathbf{x}) = 0$ else; i=1,2,…4
- $f(\mathbf{x}|p) = (p^3)^{I_{MS1}(\mathbf{x})} * (3*p^2(1-p))^{I_{MS2}(\mathbf{x})} * \ * \ (3*p(1-p)^2)^{I_{MS3}(\mathbf{x})} * (1-p^3)^{I_{MS4}(\mathbf{x})}$

By factorization theorem: $T_{MS}(\mathbf{x})$ is sufficient.

Can also show minimal sufficient …

1 HHH	2 HHT	HTH	THH
3 TTH	THT	HTT	4 TTT

Relationship between Sufficiency and Completeness

- Is $T_{MS}(\mathbf{x})$ complete?

$$E[g(T_{MS}(\mathbf{x}))= \Sigma P(T)\, g(T) = (p^3) * [g(I_{MS1})] + (3p^2(1-p)) * [g(I_{MS2})] + (3p(1-p)^2) * [g(I_{MS3})] + (1-p)^3 * [g(I_{MS4})] =?\ 0 \text{ for all } p$$

1 HHH	2 HHT HTH THH
3 TTH THT HTT	4 TTT

Relationship between Sufficiency and Completeness

- Is $T_{MS}(\mathbf{x})$ complete?

$$E[g(T_{MS}(\mathbf{x}))= \Sigma P(T)\, g(T) = (p^3) * [g(I_{MS1})] + (3p^2(1-p)) * [g(I_{MS2})] + (3p(1-p)^2) * [g(I_{MS3})] + (1-p)^3 * [g(I_{MS4})] =?\ 0 \text{ for all } p$$

$[g(I_{MS1})] = 0$

1 HHH	2 HHT HTH THH
3 TTH THT HTT	4 TTT

Relationship between Sufficiency and Completeness

- Is $T_{MS}(\mathbf{x})$ complete?

$$E[g(T_{MS}(\mathbf{x}))= \Sigma P(T)\, g(T) = (p^3) * [g(I_{MS1})] + (3p^2(1-p)) * [g(I_{MS2})] + (3p(1-p)^2) * [g(I_{MS3})] + (1-p)^3 * [g(I_{MS4})] =?\ 0 \text{ for all } p$$

$[g(I_{MS1})] = 0$

$[g(I_{MS2})] = 0$

1 HHH	2 HHT HTH THH
3 TTH THT HTT	4 TTT

Relationship between Sufficiency and Completeness

- Is $T_{MS}(\mathbf{x})$ complete?

$$E[g(T_{MS}(\mathbf{x}))= \Sigma P(T)\, g(T) = (p^3) * [g(I_{MS1})] + (3p^2(1-p)) * [g(I_{MS2})] + (3p(1-p)^2) * [g(I_{MS3})] + (1-p)^3 * [g(I_{MS4})] =?\ 0 \text{ for all } p$$

$[g(I_{MS1})] = 0$

$[g(I_{MS2})] = 0$

$[g(I_{MS3})] = 0$

1 HHH	2 HHT HTH THH
3 TTH THT HTT	4 TTT

Relationship between Sufficiency and Completeness

- Is $T_{MS}(\mathbf{x})$ complete?

$E[g(T_{MS}(\mathbf{x}))= \Sigma\, P(T)\, g(T) = (p^3) * [g(I_{MS1})]$
$+ (3p^2(1-p))* [g(I_{MS2})]$
$+ (3p(1-p)^2)* [g(I_{MS3})]$
$+ (1-p)^3 * [g(I_{MS4})] =?\ 0$ for all p

$[g(I_{MS1})] =0$
$[g(I_{MS2})] =0$
$[g(I_{MS3})] =0$
$[g(I_{MS4})] =0$

1 HHH	2 HHT	HTH	THH
3 TTH	THT	HTT	4 TTT

Relationship between Sufficiency and Completeness

- Is $T_{MS}(\mathbf{x})$ complete?

$E[g(T_{MS}(\mathbf{x}))= \Sigma\, P(T)\, g(T) = (p^3)* [g(I_{MS1})]$
$+ (3p^2(1-p))* [g(I_{MS2})]$
$+ (3p(1-p)^2)* [g(I_{MS3})]$
$+ (1-p)^3 * [g(I_{MS4})] =?\ 0$ for all p

$[g(I_{MS1})] =0$
$[g(I_{MS2})] =0$ $T_{MS}(\mathbf{x})$ is complete
$[g(I_{MS3})] =0$
$[g(I_{MS4})] =0$

1 HHH	2 HHT	HTH	THH
3 TTH	THT	HTT	4 TTT

Relationship between Sufficiency and Completeness

$T(\mathbf{X}) = h(U(\mathbf{X}))$
$U(\mathbf{X})$: reduction of the data
$T(\mathbf{X}) = h(U(\mathbf{X}))$:further reduction

T sufficient → U sufficient
T complete ← U complete

Sufficient	HHH	HHT	HTH	THH	Finest partition (8)
	TTH	THT	HTT	TTT	(X_1, X_2, X_3)
Sufficient	HHH	HHT	HTH	THH	$U(\mathbf{X})$ = Reduction of data
	TTH	THT	HTT	TTT	(6 partitions)
Minimal Sufficient / Complete	HHH	HHT	HTH	THH	$T(\mathbf{X})$ Grouped by ΣX_i # Heads
	TTH	THT	HTT	TTT	(4 partitions)
Complete	HHH	HHT	HTH	THH	$T_1(\mathbf{X})$ = further reduction.
	TTH	THT	HTT	TTT	(3 partitions)
Complete	HHH	HHT	HTH	THH	Coarsest partition (1)
	TTH	THT	HTT	TTT	$T_2(\mathbf{X})$=Constant for all $\mathbf{X}$

The Point?

HHH	HHT	HTH	THH
TTH	THT	HTT	TTT

- Sufficiency lost when **combine** samples with **different** information in the same partition
- Completeness lost when samples with **same** information are **separated** into different partitions (dimension too big)
- Sufficiency and Completeness together provides all information with minimal dimension (and indepence from ancillary stats)

  Complete, Sufficient Statistic…. **Statistical Nirvana!!**

Hmmm….

- Things to think about / add:
  - N(μ,1)
  - N(μ, μ^2)*
  - U(0,θ)
  - U(θ, θ+k)*

**exponential family….. as long as the parameter space Θ contains an open subset in $\boldsymbol{R}^k$……*