STAT 543 ☺

Lec 19, F, Mar 7

sufficiency & completeness → Homework 4 posted, due M, Mar 10

- Exam 1 solutions, grading key, summary posted

STAT 5430: Summary to date
Where we have been
& where we are headed

- Completed
  - Introduction to Statistical Inference
  - Point Estimation
    - MME/MLE
  - Criteria for Evaluating Point Estimators
    - bias, variance, UMVUE, MSE
  - Elements of Decision Theory
    - Minimax, finding Bayes estimators

- Next: Sufficiency and Point Estimation
  - Sufficiency/Data Reduction
  - Factorization Theorem
  - Rao-Blackwell Theorem
  - Completeness/Lehman-Scheffe Theorem/UMVUE
  - Exponential Families

Sufficiency and Point Estimation (Chapter 6)

Sufficiency as Data Reduction

Definition: Let $X_1, \ldots, X_n$ be r.v.'s with joint pdf/pmf $f(\underset{\sim}{x}|\underset{\sim}{\theta})$, $\underset{\sim}{\theta} \in \Theta \subset \mathbb{R}^p$ and let $\underset{\sim}{S} \equiv (S_1, \ldots, S_k)$ be a vector of estimators. Then, $\underset{\sim}{S}$ is called (jointly) **sufficient** for $\underset{\sim}{\theta}$ if the conditional distribution of $(X_1, \ldots, X_n)$ given $\underset{\sim}{S}$ does *not* depend on $\underset{\sim}{\theta}$.

Example: Let $X_1, \ldots, X_n$ be iid Geometric(θ), $0 < \theta < 1$. Show that $S \equiv X_1 + \cdots + X_n$ is sufficient for θ.

Solution: conditional pmf of $(X_1, \ldots, X_n)$ given $S=s$ is

$$P_\theta(X_1 = x_1, \ldots, X_n = x_n \mid S = s) = \frac{P_\theta(X_1 = x_1, \ldots, X_n = x_n, S = s)}{P_\theta(S=s)}$$

$$= \begin{cases} \dfrac{P_\theta(X_1 = x_1, \ldots, X_n = x_n, S = s)}{P_\theta(S=s)} & \text{if } x_1 + x_2 + \cdots x_n = s \\ 0 & \text{o.w.} \end{cases}$$

↑ Neg-Binomial(n, θ)
\# of trials until "n" successes

$$= \begin{cases} \dfrac{P_\theta(X_1 = x_1, \ldots, X_n = x_n)}{\binom{s-1}{n-1}\theta^n(1-\theta)^{s-n}} & \text{if } x_1 + x_2 + \cdots x_n = s \\ 0 & \text{o.w.} \end{cases}$$

$$= \begin{cases} \dfrac{P_\theta(X_1 = x_1) \cdots P_\theta(X_n = x_n)}{\binom{s-1}{n-1}\theta^n(1-\theta)^{s-n}} & \text{if } x_1 + \cdots x_n = s \\ 0 & \text{o.w} \end{cases}$$

$$= \begin{cases} \dfrac{\prod_{i=1}^{n}\left[\theta(1-\theta)^{x_i-1}\right]}{\binom{s-1}{n-1}\theta^n(1-\theta)^{s-n}} & \text{if } x_1+\cdots+x_n=s \\ 0 & \text{o.w} \end{cases}$$

$$= \begin{cases} \dfrac{\theta^n(1-\theta)^{s-n}}{\binom{s-1}{n-1}\theta^n(1-\theta)^{s-n}} & \text{if } x_1+\cdots+x_n=s \\ 0 & \text{o.w} \end{cases}$$

$$= \begin{cases} 1/\binom{s-1}{n-1} & \text{if } x_1+\cdots+x_n=s \\ 0 & \text{o.w} \end{cases}$$

free of θ! $\Rightarrow$ S is sufficient for θ.

Sufficiency and Point Estimation

Factorization Theorem

Remarks on Sufficiency:

Recall in the definition: $p = \#$ of parameters, $k = \#$ of statistics

$k = p$: eg. last example Geometric(θ), $p = 1 = k$

$k > p$: eg. $X_1, \ldots, X_n$ iid UNIF($\theta, \theta+1$) $\Rightarrow \underset{\sim}{S} = (\min X_i, \max X_i)$ sufficent for θ, $p=1$

$k < p$: eg. $n=1$, $X_1 \sim N(\mu, \sigma^2)$, $p=2$, but X_1 is sufficent, $k=1$

Factorization Theorem: Let $X_1, \ldots, X_n$ be r.v.'s with joint pdf/pmf $f(\underset{\sim}{x}|\underset{\sim}{\theta})$, $\underset{\sim}{\theta} \in \Theta \subset \mathbb{R}^p$ and let $\underset{\sim}{S} = (S_1, \ldots, S_k)$ be a vector of estimators. Then, $\underset{\sim}{S}$ is sufficient for $\underset{\sim}{\theta}$ *if and only if* there exist functions $g(\underset{\sim}{S}, \underset{\sim}{\theta})$ and $h(\underset{\sim}{x})$ such that $h(\underset{\sim}{x})$ does NOT depend on $\underset{\sim}{\theta}$ and

$$\underbrace{f(\underset{\sim}{x}|\underset{\sim}{\theta})}_{X_1 \ldots X_n \text{ data pdf/pmf}} = g(\underset{\sim}{S}, \underset{\sim}{\theta})h(\underset{\sim}{x}) \quad \text{for all } \underset{\sim}{x} \text{ and all } \underset{\sim}{\theta}$$

$\underset{\sim}{S}$ & $\underset{\sim}{\theta}$ are "linked" inside $f(\underset{\sim}{x}|\underset{\sim}{\theta})$

Example: Let $X_1, \ldots, X_n$ be iid Negative-Binomial(r, θ), $0 < \theta < 1$ (known integer $r \geq 1$). Show that $S = X_1 + \cdots + X_n$ is sufficient for θ. (last time Geo(θ) $\sim$ Neg-Binom($r=1, \theta$))

Solution: joint pmf of $X_1, \ldots, X_n$ is

$$f(\underset{\sim}{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n}\left[\binom{x_i-1}{r-1}\theta^r(1-\theta)^{x_i-r} \underbrace{I\{x_i \in A_r\}}_{\text{Indicator}}\right]$$

where $A_r = \{r, r+1, r+2, \ldots\}$

$$= \underbrace{\theta^{nr}(1-\theta)^{\sum x_i - nr}}_{g(\sum_{i=1}^n x_i, \theta)} \underbrace{\left\{\prod_{i=1}^{n}\left[\binom{x_i-1}{r-1} I\{x_i \in A_r\}\right]\right\}}_{h(\underset{\sim}{x})} \quad \forall \underset{\sim}{x}, \ \forall\, 0<\theta<1$$

where $g(s, \theta) = \theta^{nr}(1-\theta)^{s-nr}$

Hence, by factorization theorem, $S = \sum_{i=1}^n X_i$ is sufficient

Sufficiency and Point Estimation

Factorization Theorem, cont'd

$\underset{\sim}{1} = (1, \dots, 1)'$

Example: Suppose $(X_1, \dots, X_n) \sim MVN(\mu \cdot \underset{\sim}{1}, \sigma^2 \cdot A)$ where $\mu \in \mathbb{R}$, $\sigma^2 > 0$ and A is a known $n \times n$ positive definite matrix. Find a sufficient statistic for (μ, σ^2).

Solution: joint pdf of $(X_1, \dots, X_n)$ is

$$f(\underset{\sim}{x} \mid \mu, \sigma^2) = \frac{1}{(\sigma^2 2\pi)^{n/2}} \frac{1}{[\det(A)]^{1/2}} \exp\left[-\frac{1}{2\sigma^2}(\underset{\sim}{x} - \mu \underset{\sim}{1})' A^{-1} (\underset{\sim}{x} - \mu \underset{\sim}{1})\right]$$

$$= \underbrace{\frac{1}{\sigma^n} \exp\left[\frac{-1}{2\sigma^2}\left[\underset{\sim}{x}' A^{-1} \underset{\sim}{x} + 2\mu \underset{\sim}{x}' A^{-1} \underset{\sim}{1} + \mu^2 \underset{\sim}{1}' A^{-1} \underset{\sim}{1}\right]\right]}_{g(\underset{\sim}{x}' A^{-1} \underset{\sim}{x},\ \underset{\sim}{x}' A^{-1} \underset{\sim}{1},\ \mu, \sigma^2)} \underbrace{\frac{1}{(2\pi)^{n/2}} \frac{1}{[\det(A)]^{1/2}} I(\underset{\sim}{x} \in \mathbb{R}^n)}_{h(\underset{\sim}{x})}$$

Hence, by Factorization Theorem,

$\underset{\sim}{S} = (\underset{\sim}{x}' A^{-1} \underset{\sim}{x},\ \underset{\sim}{x}' A^{-1} \underset{\sim}{1})$ are sufficient for (μ, σ^2)

Remarks:

1. The choice of $g(\underset{\sim}{S}, \underset{\sim}{\theta})$ and $h(\underset{\sim}{x})$ is not unique.

2. Any 1-to-1 function of a sufficient statistic is also sufficient.

   Example: In last example, suppose $A = I_{n \times n}$.

Then, $\underset{\sim}{S} = (\underset{\sim}{x}' A^{-1} \underset{\sim}{x},\ \underset{\sim}{x}' A^{-1} \underset{\sim}{1}) = \left(\sum_{i=1}^n X_i^2, \sum_{i=1}^n X_i\right)$

Note $\underset{\sim}{T} = \left(\frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2, \bar{X}_n\right)$ is a 1-to-1 function of $\underset{\sim}{S}$ $\Rightarrow$ $\underset{\sim}{T}$ is sufficient for (μ, σ^2)

Sufficiency and Point Estimation

Minimal Sufficiency

Question: Suppose $\utilde{S} \equiv (S_1, \ldots, S_k)$ is sufficient for $\utilde{\theta}$ and S_0 is another arbitrary statistic. Is $\utilde{S}^* \equiv (S_0, S_1, \ldots, S_k)$ is sufficient for $\utilde{\theta}$? Yes!

proof: Since $\utilde{S}$ is sufficient,

$f(\utilde{x}|\utilde{\theta}) = g(\utilde{S}, \utilde{\theta})h(\utilde{x})$ holds by Factorization Theorem

But $\utilde{S} = (S_1, .., S_k) = d(\utilde{S}^*)$ is a function of $\utilde{S}^*$

So, $f(\utilde{x}|\utilde{\theta}) = g(\utilde{S}, \utilde{\theta})h(\utilde{x})$

$$= g(d(\utilde{S}^*), \utilde{\theta})h(\utilde{x}) = g_1(\utilde{S}^*, \utilde{\theta})h(\utilde{x})$$

$\therefore$ $\utilde{S}^*$ is sufficient Fact. theorem.

Definition: A vector of statistics $\utilde{S}$ is called **minimally sufficient** if

1. $\utilde{S}$ is sufficient for $\utilde{\theta}$, and

2. for any other vector $\utilde{T}$ of sufficient statistics for $\utilde{\theta}$, $\utilde{S}$ is a function of $\utilde{T}$.

(Later: we can check "minimally sufficient" using "completeness with sufficiency.")

$X_1 \ldots X_n \xrightarrow{\text{data reduction}}$ $\utilde{T}$ is sufficient for $\utilde{\theta}$ (more)

$\downarrow$ data reduction $g(\utilde{T}) = \utilde{S}$ ($\utilde{S}$ is a function of $\utilde{T}$)

minimally sufficient $\utilde{S}$ (less)

e.g. In last MVN example, $A = I_{n\times n}$, it turns out that $\utilde{S} = (\sum_{i=1}^n X_i^2, \sum_{i=1}^n X_i)$ is minimally sufficient for (μ, σ^2).

Later: we can show "minimal sufficiency" using "completeness"

Sufficiency and Point Estimation

Remarks on Sufficiency

1. If $X_1, \ldots, X_n$ is a random sample (iid) from pdf/pmf $f(x|\underset{\sim}{\theta})$, $\underset{\sim}{\theta} \in \Theta$, then the order statistics $X_{(1)}, \ldots, X_{(n)}$ are sufficient for $\underset{\sim}{\theta}$.

   proof: By the factorization theorem, $X_{(1)}, \ldots, X_{(n)}$ are sufficient for $\underset{\sim}{\theta}$ because we can write

$$\begin{aligned}\text{the joint pdf/pmf } f(\underset{\sim}{x}|\underset{\sim}{\theta}) &= \prod_{i=1}^{n} f(x_i|\underset{\sim}{\theta}) = \prod_{i=1}^{n} f(x_{(i)}|\underset{\sim}{\theta}) \\ &= \underbrace{g(x_{(1)}, \ldots, x_{(n)}, \underset{\sim}{\theta})}_{\prod_{i=1}^{n} f(x_{(i)}|\underset{\sim}{\theta})}\, \underbrace{h(\underset{\sim}{x})}_{1}, \quad \text{for all } \underset{\sim}{x}, \underset{\sim}{\theta}\end{aligned}$$

2. If $\underset{\sim}{S} = (S_1, S_2, \ldots, S_k)$ is sufficient for real-valued $\theta \in \Theta \subset \mathbb{R}$, then any Bayes estimator is a function of $\underset{\sim}{S}$.

$$f_{\theta|\underset{\sim}{x}}(\theta) \propto f(\underset{\sim}{x}|\theta)\pi(\theta)$$
$$\propto g(\underset{\sim}{S}, \theta)\pi(\theta)$$

   Example: From homework, consider $X_1, \ldots, X_n$ iid Bernoulli(θ), $0 < \theta < 1$; loss $L(t, \theta) = \frac{(t-\theta)^2}{\theta(1-\theta)}$; and uniform(0,1) prior $\pi(\theta)$.

   Then the Bayes estimator is $T_0 = \bar{X}_n$, which is sufficient for θ (by factorization theorem).

3. If $\underset{\sim}{S} = (S_1, S_2, \ldots, S_k)$ is sufficient for $\underset{\sim}{\theta} \in \Theta \subset \mathbb{R}^p$ and $\hat{\theta}$ is the unique MLE of $\underset{\sim}{\theta}$, then $\hat{\theta}$ is a function of $\underset{\sim}{S}$.

$f(\underset{\sim}{x}|\theta) = g(\underset{\sim}{S}, \theta)h(\underset{\sim}{x})$ by Fact theorem

Sufficiency and Point Estimation

Rao-Blackwell Theorem & Sufficiency

Rao-Blackwell Theorem. Let $f(\underset{\sim}{x}|\underset{\sim}{\theta}) = f(x_1, \ldots, x_n|\underset{\sim}{\theta})$ be the joint pdf/pmf of $(X_1, \ldots, X_n)$ and $\underset{\sim}{S} = (S_1, S_2, \ldots, S_k)$ be sufficient for $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_p) \in \Theta \subset \mathbb{R}^p$.

Also let T be any UE of a real-valued $\gamma(\underset{\sim}{\theta})$ and $T^* = \mathrm{E}(T|\underset{\sim}{S})$ *(this conditional expectation does not depend on* $\underset{\sim}{\theta}$*, since* $\underset{\sim}{S}$ *is sufficient, and so is a statistic).*

new estimator (just depend on $\underset{\sim}{S}$ through data & NOT on $\underset{\sim}{\theta}$)

Then,

1. T^* is a function of $\underset{\sim}{S}$ and an UE of $\gamma(\underset{\sim}{\theta})$.

2. $\mathrm{Var}_{\underset{\sim}{\theta}}(T^*) \leq \mathrm{Var}_{\underset{\sim}{\theta}}(T)$, for all $\underset{\sim}{\theta} \in \Theta$.

3. If $\mathrm{Var}_{\underset{\sim}{\theta}_0}(T^*) = \mathrm{Var}_{\underset{\sim}{\theta}_0}(T)$ holds for some $\underset{\sim}{\theta}_0 \in \Theta$, then $P_{\underset{\sim}{\theta}_0}(T = T^*) = 1$.

Idea: T is UE of $\gamma(\theta)$ $\xrightarrow{\text{condition on sufficent } \underset{\sim}{S}}$ new $T^* = E(T|\underset{\sim}{S})$

"Rao-Blackwellization" or we say we "Rao-Blackwellize" T using $\underset{\sim}{S}$

Remarks

- Given an UE T of $\gamma(\underset{\sim}{\theta})$, the theorem shows how to obtain an UE T^* that is at least as good as T in terms of variance (in fact, better than T unless $T = T^*$ with probability 1 for all $\underset{\sim}{\theta}$). That is, you can "Rao-Blackwellize" an UE T by conditioning on a sufficient statistic $\underset{\sim}{S}$.

$T^* = E(T|\underset{\sim}{S})$

(ASIDE: $T = g(\underset{\sim}{S}) \Rightarrow T^* = E(T|\underset{\sim}{S}) = E(g(\underset{\sim}{S})|\underset{\sim}{S}) = g(\underset{\sim}{S})$)

- For finding an UMVUE of $\gamma(\underset{\sim}{\theta})$ we may restrict attention to the class of estimators that are functions of a sufficient statistic.

$T^* = E(T|\underset{\sim}{S})$ function of $\underset{\sim}{S}$

Sufficiency and Point Estimation

Rao-Blackwell Theorem: Illustration

$n=2$

Example: Suppose X_1, X_2 are iid Exponential(θ). Note $T = X_1$ is an UE of θ and $\text{Var}_\theta(T) = \text{Var}_\theta(X_1) = \theta^2$.

$E_\theta T = E_\theta X_1 = \theta$

Also note that $S = X_1 + X_2$ is sufficient for θ by factorization theorem & S is GAMMA(2,θ)-distributed.

Verify that

1. $T^* = \text{E}_\theta(T|S) = \text{E}_\theta(X_1|S)$ is a function of S;
2. T^* doesn't depend on θ;
3. T^* is unbiased for θ;
4. and compare $\text{Var}_\theta(T)$ and $\text{Var}_\theta(T^*)$

Solution: Given $S = s > 0$, first find the conditional pdf of $X_1|S = s$ as

$$f(x_1|S=s) = \frac{f_{X_1,S}(x_1, s|\theta)}{f_S(s|\theta)} = \frac{f_{X_1,X_2}(x_1, x_2 = s - x_1|\theta)}{f_S(s|\theta)}$$

joint pdf of (X_1, S)

$$= \begin{cases} \dfrac{\theta^{-2}e^{-x_1/\theta}e^{-(s-x_1)/\theta}}{\theta^{-2}se^{-s/\theta}} = s^{-1} & \text{if } 0 < x_1 < s \\ 0 & \text{otherwise} \end{cases}$$

Gamma pdf

So, given $S = s > 0$, the conditional distribution of X_1 is UNIF$(0,S)$

Hence, the conditional expectation is $\text{E}_\theta(X_1|S = s) = \frac{0+S}{2} = S/2$

Now, treating S as a random variable, we have $T^* = \text{E}_\theta(X_1|S) = \frac{S}{2} = \frac{X_1+X_2}{2} = \bar{X}_2$

① T^* is function of S

② T^* doesn't depend on θ

③ T^* is UE of θ $(E_\theta(\bar{X}_2) = \theta)$

④ $Var_\theta(\bar{X}_2) = \theta^2/2 < Var_\theta(X_1) = \theta^2$

Sufficiency and Point Estimation

Completeness

side note: to be used with sufficiency

- "Completeness" is a statistical property for use in conjunction with sufficiency.

- "Completeness" has a rather technical definition: a statistic $\underset{\sim}{T}$ is complete if the only function $u(\underset{\sim}{T})$ of $\underset{\sim}{T}$ that can be an UE of zero is $u(\underset{\sim}{T}) = 0$ wp1

Note: $u(\underset{\sim}{T}) \equiv 0$ (always zero) has expectation zero (i.e. $E_{\theta} u(\underset{\sim}{T}) = 0$).

It is true that $u(T) \equiv 0$ is the ONLY function of $\underset{\sim}{T}$ that is U.E of zero? If so, $\underset{\sim}{T}$ is complete

"Completeness" Definition: Let $f(\underset{\sim}{x}|\underset{\sim}{\theta}) = f(x_1, \ldots, x_n|\underset{\sim}{\theta})$, $\underset{\sim}{\theta} \in \Theta \subset \mathbb{R}^p$, be the *joint* pdf/pmf of $(X_1, \ldots, X_n)$ and let $f_{\underset{\sim}{T}}(\underset{\sim}{t}|\underset{\sim}{\theta})$ denote the pdf/pmf of a vector of statistics $\underset{\sim}{T}$. Then,

sampling distribution of $\underset{\sim}{T}$

(i) $\underset{\sim}{T}$ (and/or the family $\mathcal{F}_{\underset{\sim}{T}} \equiv \{f_{\underset{\sim}{T}}(\underset{\sim}{t}|\underset{\sim}{\theta}) : \underset{\sim}{\theta} \in \Theta\}$) is called **complete** if, for any real-valued function $u(\underset{\sim}{T})$,

$$\text{whenever } \underbrace{\mathrm{E}_{\underset{\sim}{\theta}}\, u(\underset{\sim}{T}) = 0}_{\text{says "}u(T)\text{ is U.E of zero"}} \text{ holds for all } \underset{\sim}{\theta} \in \Theta \text{ then } P_{\underset{\sim}{\theta}}\Big[u(\underset{\sim}{T}) = 0\Big] = 1 \text{ for } \underset{\sim}{\theta} \in \Theta \quad (3)$$

then, $u(\underset{\sim}{T})$ must zero w.p. 1

(ii) $\underset{\sim}{T}$ is called **bounded complete** if (4) holds for all *bounded* functions $u(\cdot)$.

*Completeness of $\underset{\sim}{T}$ depends on $f_{\underset{\sim}{T}}(t|\theta)$ & "richness/size of parameter space Θ"

The more parameter $\underset{\sim}{\theta}$ in Θ, the more constraints are on $u(T)$ in $E_{\theta}(u(T)) = 0$, $\forall \underset{\sim}{\theta} \in \Theta$

e.g. $T \equiv X_1 \sim N(\theta, 1)$, $\theta \in \Theta$

① $\Theta = \{1\}$

Note: $u(X_1) = X_1 - 1$

$E_\theta u(X_1) = E_\theta(X_1 - 1) = 0$, $\forall \theta \in \Theta = \{1\}$

But, $u(X_1)$ is NOT zero w.p. 1 ← $N(0,1) \sim X_1 - 1$

($P_\theta(X_1 - 1 = 0) = P(Z = 0) = 0$)

So, $T \equiv X_1$ is NOT complete.

② $\Theta \equiv \mathbb{R}$. Then, $T \equiv X_1$ is complete (later)

Sufficiency and Point Estimation

Completeness

Example. Suppose $X_1, \ldots, X_n$ are iid Poisson(θ), $\theta > 0$. Show that $T \equiv \sum_{i=1}^n X_i$ is complete.

Solution: $T \sim \text{Poisson}(n\theta)$, $\theta > 0$.

Now for some $u(\cdot)$, suppose it holds that $E_\theta u(T) = 0$, $\forall \theta > 0$.

$$\Leftrightarrow \sum_{t=0}^{\infty} u(t) e^{-n\theta} \frac{(n\theta)^t}{t!} = 0, \quad \forall \theta > 0$$

$$\Leftrightarrow \sum_{t=0}^{\infty} u(t) \frac{(n\theta)^t}{t!} = 0, \quad \forall \theta > 0.$$

$$\Leftrightarrow \left[u(0) + u(1)(n\theta) + u(2)\frac{(n\theta)^2}{2!} + \cdots \right] = 0, \quad \forall \theta > 0$$

let $\theta \downarrow 0$, get $u(0) = 0$.

$$\Leftrightarrow \left[u(1) + u(2)\frac{(n\theta)}{2!} + \cdots \right] = 0, \quad \forall \theta > 0$$

let $\theta \downarrow 0$, get $u(1) = 0$.

So, get for any $t \geq 0$, $u(t) = 0$

$\Rightarrow P_\theta(u(T) = 0) = 1$ for all $\theta > 0$

$\therefore$ T is complete!

(Later, we'll see easier ways to check completeness)

Sufficiency and Point Estimation

Remarks on Completeness

1. If T is complete, then T is boundedly complete; the converse is false.

connection between sufficiency & completeness

2. If T is sufficient and boundedly complete, then T is minimal sufficient.

   So by Remark 1 above, if T is sufficient and complete, then T is minimal sufficient.

3. Suppose T is complete and $h_1(T), h_2(T)$ are two estimators of $\gamma(\underset{\sim}{\theta})$

$$
\begin{aligned}
& \text{if } \mathrm{E}_{\underset{\sim}{\theta}} h_1(T) = \gamma(\underset{\sim}{\theta}) = \mathrm{E}_{\underset{\sim}{\theta}} h_2(T), \ \text{ for all } \underset{\sim}{\theta} \in \Theta \\
\Rightarrow \ & \mathrm{E}_{\underset{\sim}{\theta}} u(T) = 0, \ \text{ for all } \underset{\sim}{\theta} \in \Theta, \quad \text{where } u(T) = h_1(T) - h_2(T) \\
\Rightarrow \ & P_{\underset{\sim}{\theta}}\big(u(T) = 0\big) = 1, \ \text{ for all } \underset{\sim}{\theta} \in \Theta \\
\Rightarrow \ & P_{\underset{\sim}{\theta}}\big(h_1(T) = h_2(T)\big) = 1, \text{for all } \underset{\sim}{\theta} \in \Theta
\end{aligned}
$$

   Hence, there can be at most one (i.e., unique) UE of a parametric function $\gamma(\underset{\sim}{\theta})$ that is a function of a complete statistic.

4. Let $T \equiv h(X_1, \ldots, X_n)$ be an UE of $\gamma(\theta)$ & suppose $\underset{\sim}{S}$ is sufficient.

Recall:

$T \xrightarrow[\text{Rao-Blackwellize}]{\text{sufficient } \underset{\sim}{S}} T^* = E(T|\underset{\sim}{S})$ is U.E of $\gamma(\theta)$ & $Var_\theta(T^*) \le Var_\theta(T)$

$T \xrightarrow[\text{L-S Theorem (next)}]{\text{sufficient \& complete } \underset{\sim}{S}} T^* = E(T|\underset{\sim}{S})$ is UMVUE of $\gamma(\theta)$!

Sufficiency and Point Estimation

Lehmann-Scheffe Theorem: Completeness + Sufficiency + UE = UMVUE

Lehmann-Scheffe Theorem. Let $f(\underset{\sim}{x}|\underset{\sim}{\theta}) = f(x_1, \ldots, x_n|\underset{\sim}{\theta})$ be the *joint* pdf/pmf of $(X_1, \ldots, X_n)$, $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_p) \in \Theta \subset \mathbb{R}^p$. Let $\underset{\sim}{S} = (S_1, S_2, \ldots, S_k)$ be a complete and sufficient statistic. If $T^* \equiv T(\underset{\sim}{S})$ is an UE of $\gamma(\underset{\sim}{\theta})$ and is a function of $\underset{\sim}{S}$, then T^* is the UMVUE of $\gamma(\underset{\sim}{\theta})$.

Proof. Let T be any UE of $\gamma(\underset{\sim}{\theta})$. We must show $\text{Var}_{\underset{\sim}{\theta}}(T^*) \leq \text{Var}_{\underset{\sim}{\theta}}(T)$

Define $T_1 = E(T|\underset{\sim}{S})$. Since $\underset{\sim}{S}$ is sufficient, by the Rao-Blackwell theorem, we know

T_1 is a function of $\underset{\sim}{S}$ & U.E. of $\gamma(\theta)$ &

$\text{Var}_\theta(T_1) \leq \text{Var}_\theta(T)$, $\forall \theta$

Now T_1 & T^* are functions of $\underset{\sim}{S}$ (complete) & both are U.E of $\gamma(\theta)$

Since $\underset{\sim}{S}$ is complete, we know

$P_\theta(T_1 = T^*) = 1$, $\forall \theta$

$\Rightarrow \text{Var}_\theta(T^*) = \text{Var}_\theta(T_1) \leq \text{Var}_\theta(T)$, $\forall \theta$

$\Rightarrow T^* = h(\underset{\sim}{S})$ is UMVUE of $\gamma(\theta)$ [& so is T_1]

Remark. The R-B theorem & L-S theorem together suggest two methods for finding the UMVUE:

1. Method I: Given a parametric function $\gamma(\underset{\sim}{\theta})$, find an UE of $\gamma(\underset{\sim}{\theta})$ that is a function of a complete and sufficient statistic.

easier

$\underset{\sim}{S}$ sufficiency + complete

$T^* = h(\underset{\sim}{S})$

$E_\theta T^* = \gamma(\theta)$, $\forall \theta$

then T^* is UMVUE

2. Method II: Start with any UE T of $\gamma(\underset{\sim}{\theta})$. Then $T^* = E(T|\underset{\sim}{S})$ is the UMVUE of $\gamma(\underset{\sim}{\theta})$, if $\underset{\sim}{S}$ is complete and sufficient.

a little harder

find $T^* = E(\underset{\text{any UE}}{T}|\underset{\text{complete \& sufficient}}{\underset{\sim}{S}}) = h(\underset{\sim}{S})$

Sufficiency and Point Estimation

Lehmann-Scheffe Theorem: Illustrations

Example. Let $X_1, \ldots, X_n$ be iid Poisson(θ), $\theta > 0$. Find the UMVUE of θ.

(could here use CRLB to find UMVUE)

Solution: Check $S = \sum_{i=1}^{n} X_i$ is sufficient (check by Factorization theorem) & is also complete (later)

Use $\bar{X}_n = \frac{S}{n} \Rightarrow$ check $E_\theta(\bar{X}_n) = E_\theta(X_1) = \theta, \forall \theta$

So $\bar{X}_n$ is UE of θ & a function of complete/sufficient $S \Rightarrow \bar{X}_n$ is UMVUE of θ.

$\tilde{S}^2 = \frac{n}{n-1}[\frac{\sum x_i^2}{n} - (\bar{X}_n)^2]$

Note: $\tilde{S}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ has $E_\theta \tilde{S}^2 = Var_\theta(X_1) = \theta, \forall \theta > 0$

So $\tilde{S}^2$ is UE of θ. So, $E(\tilde{S}^2 | S) = \bar{X}_n$ by L-S theorem

Example. Let $X_1, \ldots, X_n$ be iid Bernoulli(θ), $0 < \theta < 1$. Find the UMVUE of $\gamma(\theta) = \theta^r(1-\theta)^{n-r}$, for a fixed (known) integer $1 \le r \le n$.

Solution: Check $S \equiv \sum_{i=1}^{n} X_i$ is sufficient & also complete (later)

Note $S \sim \text{Binomial}(n, \theta)$, $P_\theta(S = s) = \binom{n}{s}\theta^s(1-\theta)^{n-s}$, $0 \le s \le n$

Define $T^* = \begin{cases} \frac{1}{\binom{n}{r}} & \text{if } S = r \\ 0 & \text{o.w} \end{cases} = \frac{I[S=r]}{\binom{n}{r}}$

which is a function of S &

$E_\theta(T^*) = \frac{1}{\binom{n}{r}} E_\theta(I[S=r]) = \frac{P_\theta(S=r)}{\binom{n}{r}} = \gamma(\theta) = \theta^r(1-\theta)^{n-r}, \forall \theta$

So, T^* is UMVUE of $\tau(\theta)$ by L-S theorem.

Sufficiency and Point Estimation

Exponential Families (for Checking Sufficiency/Completeness)

Definition: A family of pdf/pmf $\{f(\underset{\sim}{x}|\underset{\sim}{\theta}) : \underset{\sim}{\theta} \in \Theta\}$, $\Theta \subset \mathbb{R}^p$, is called an **exponential family** if it can be written in the form

pmf/pdf of data

$$f(\underset{\sim}{x}|\underset{\sim}{\theta}) = \begin{cases} c(\underset{\sim}{\theta})h(\underset{\sim}{x})\exp\left[\underbrace{\sum_{i=1}^{k} q_i(\underset{\sim}{\theta})t_i(\underset{\sim}{x})}_{t_i(\underset{\sim}{x}),\ q_i(\underset{\sim}{\theta}),\ i=1,\dots,k}\right] & \underset{\sim}{x} \in A \quad \leftarrow \text{support} \\ 0 & \text{otherwise} \end{cases}$$

where

$\leftarrow$ support

$A \equiv \{\underset{\sim}{x} : f(\underset{\sim}{x}|\underset{\sim}{\theta}) > 0\}$ does NOT depend on $\underset{\sim}{\theta}$,

$c(\underset{\sim}{\theta}) > 0$ and $h(\underset{\sim}{x}) > 0$ are positive-valued functions,

and $q_i(\underset{\sim}{\theta})$, $t_i(\underset{\sim}{x})$ are real-valued functions for $i = 1, \dots, k$.

$\leftarrow$ tool to determine/find complete & sufficient statistics

Theorem: Let $\underset{\sim}{X}_1, \dots, \underset{\sim}{X}_n$ be a (possibly vector-valued) random sample (iid) from $f(\underset{\sim}{x}|\underset{\sim}{\theta})$, where $\{f(\underset{\sim}{x}|\underset{\sim}{\theta}) : \underset{\sim}{\theta} \in \Theta\}$ is an exponential family admitting a representation as above. If

$$\left\{\overbrace{[q_1(\underset{\sim}{\theta}), \dots, q_k(\underset{\sim}{\theta})]}^{k\text{-tuple}} : \underset{\sim}{\theta} \in \Theta\right\} \supset \underbrace{(a_1, b_1) \times \cdots \times (a_k, b_k)}_{\text{open set/rectangle in } \mathbb{R}^k}$$

$\leftarrow$ set of all k-tuples over $\underset{\sim}{\theta} \in \Theta$ must an "open set" in $\mathbb{R}^k$

for some $a_i < b_i$, $i = 1, \dots, k$, then

$$\underset{\sim}{S} = \left(\sum_{j=1}^{n} t_1(\underset{\sim}{X}_j), \dots, \sum_{j=1}^{n} t_k(\underset{\sim}{X}_j)\right) \quad \leftarrow k \text{ statistics}$$

is complete and sufficient.

Sufficiency and Point Estimation

Exponential Families: Illustration

Example. Let $X_1, \ldots, X_n$ be iid Gamma(α, β), $\alpha, \beta > 0$. Show that $\underset{\sim}{T} = (\sum_{i=1}^n X_i, \prod_{i=1}^n X_i)$ is complete and sufficient.

Solution: $X_1, \ldots, X_n$ iid, so consider the pdf of X_1

$$f(x|\alpha,\beta) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} & , x>0 \\ 0 & \text{o.w.} \end{cases}$$

$e^{\log x^\alpha}$

$$= \begin{cases} c(\underset{\sim}{\theta}) h(x) \exp\left[-x/\beta + \alpha \log x\right] & , x>0 \\ 0 & \text{o.w.} \end{cases}$$

$\underset{\sim}{\theta} = (\alpha, \beta)$, $c(\underset{\sim}{\theta}) = \frac{1}{\Gamma(\alpha)\beta^\alpha}$, $h(x) = x^{-1}$, $A \overset{\text{support}}{=} (0, \infty)$

$t_1(x) = x$, $q_1(\underset{\sim}{\theta}) = -1/\beta$, $t_2(x) = \log x$, $q_2(\underset{\sim}{\theta}) = \alpha$

check

$$\{ [q_1(\underset{\sim}{\theta}), q_2(\underset{\sim}{\theta})] \in \mathbb{R}^2 : \underset{\sim}{\theta} = (\alpha, \beta) \in (0,\infty) \times (0,\infty) \}$$

$$= \{ [-1/\beta, \alpha] \in \mathbb{R}^2 : \alpha, \beta > 0 \} = (-\infty, 0) \times (0, \infty)$$

$\supset (-1, 0) \times (0, 1)$

$(-10, -\pi) \times (1, 100)$

contains some open interval

By Theorem,

$$\underset{\sim}{S} = \left(\sum_{j=1}^n t_1(X_j), \sum_{j=1}^n t_2(X_j) \right) = \left(\sum_{j=1}^n X_j, \sum_{j=1}^n \log X_j \right)$$

is sufficient & complete

So, $\underset{\sim}{T} = \left(\sum_{j=1}^{n} X_j, \prod_{j=1}^{n} X_i\right)$ is one-to-one function with

$$\underset{\sim}{S} = \left(\sum_{j=1}^{n} X_j, \sum_{j=1}^{n} \log X_j\right)$$

& so

$\underset{\sim}{T}$ is complete &

sufficient

Note: looked at the problem as "n" iid X_i's (real-valued) & worked with $f(x|\theta)$ ← 1 obs.

or could have "1" vector $\underset{\sim}{x} = (x_1, \ldots, x_n)$ & worked with $f(\underset{\sim}{x}|\theta)$ ← all data

e.g. Suppose $X_1, \ldots, X_n$ are independent & $X_i \sim \text{Poisson}(i\theta)$, $\theta > 0$, $i = 1, \ldots, n$

$\underset{\sim}{X} = (X_1, \ldots, X_n)$ ← 1 vector

$$f(\underset{\sim}{x}|\theta) = \prod_{i=1}^{n} \frac{e^{-i\theta}(i\theta)^{x_i}}{x_i!} = \underbrace{e^{-\sum_{i=1}^{n} i\theta}}_{c(\theta)} \underbrace{\left(\prod_{i=1}^{n} \frac{(i)^{x_i}}{x_i!}\right)}_{h(\underset{\sim}{x})} \exp(\underbrace{\log\theta}_{q_1(\theta)} \cdot \underbrace{\sum_{i=1}^{n} x_i}_{t_1(\underset{\sim}{x})})$$

$e^{\log\theta^{\sum x_i}}$

check $\{q_1(\theta) \equiv \log\theta : \theta > 0\} = (-\infty, \infty) \supset (0,1)$

$\Rightarrow t_1(\underset{\sim}{X}) = \sum_{i=1}^{n} X_i$ is complete & sufficient by theorem (applied to 1 $\underset{\sim}{X}$)

Sufficiency and Point Estimation

Ancillary Statistics & Basu's Theorem

Definition: A statistic $\underset{\sim}{T}$ is called **ancillary** if its distribution does NOT depend on any parameters.

you can know the distribution of $\underset{\sim}{T}$ completely

... doesn't depend on unknown $\underset{\sim}{\theta}$

$\leftarrow$ location parameter

Example 1. Let $X_1, \ldots, X_n$ be iid $N(\mu, 1)$, $\mu \in \mathbb{R}$.

$T = X_1 - X_2 \sim N(0,2) \leftarrow$ ancillary stat.

$= (X_1 - \mu) - (X_2 - \mu) \qquad X_i - \mu \sim N(0,1)$ $i=1,2$

$T = \dfrac{X_1+X_2}{2} - \dfrac{X_3+X_4}{2} \sim N(0,1) \leftarrow$ ancillary statistic

(use differences with location parameter)

Example 2. Let $X_1, \ldots, X_n$ be iid Exponential(θ), $\theta > 0$.

$\uparrow$ scale parameter

$T = \dfrac{X_1}{\sum_{i=1}^n X_i} = \dfrac{Z_1}{\sum_{i=1}^n Z_i} \qquad Z_i = X_i/\theta \sim Exp(1)$

ancillary stat.

$T = X_1/X_2 = Z_1/Z_2$

(use ratios with scale parameter)

Basu's Theorem: Let $f(\underset{\sim}{x}|\underset{\sim}{\theta})$, $\underset{\sim}{\theta} \in \Theta \subset \mathbb{R}^p$, be the joint pdf/pmf of $X_1, \ldots, X_n$. Suppose that $\underset{\sim}{S} = (S_1, \ldots, S_k)$ is complete and sufficient, and that $\underset{\sim}{T} = (T_1, \ldots, T_m)$ in ancillary. Then, $\underset{\sim}{S}$ and $\underset{\sim}{T}$ are independent for all $\underset{\sim}{\theta}$.

(doesn't matter which $\underset{\sim}{\theta}$ generates data.... $P(T \in B)$

$\uparrow$ ancillary remove $\underset{\sim}{\theta}$

$\underset{\sim}{S}$ & $\underset{\sim}{T}$ are independent)

$P_{\underset{\sim}{\theta}}(\underset{\sim}{S} \in A, \underset{\sim}{T} \in B) = P_{\underset{\sim}{\theta}}(\underset{\sim}{S} \in A) P_{\underset{\sim}{\theta}}(\underset{\sim}{T} \in B)$

for any $\underset{\sim}{\theta}$

Sufficiency and Point Estimation

Ancillary Statistics & Basu's Theorem: Illustration

Example. Let $X_1, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$.

1. Show that $T = \dfrac{X_1 - \bar{X}_n}{S_n}$ and $(\bar{X}_n, S_n^2)$ are independent, where $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ and $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$.
2. Find the UMVUE of $P_{\mu,\sigma^2}(X_1 \le 2) = \underbrace{\Phi}_{N(0,1)\text{ cdf}}\left(\dfrac{2-\mu}{\sigma}\right)$. $= P_{\mu,\sigma^2}\left(\underbrace{\dfrac{X_1-\mu}{\sigma}}_{\sim N(0,1)} \le \dfrac{2-\mu}{\sigma}\right)$

Solution:

Step 1. Check that $(\bar{X}_n, S_n^2)$ is complete & sufficient statistic for $\underset{\sim}{\theta} = (\mu, \sigma^2)$.... use Exponential Families

Hence, by Basu's theorem, it is enough to show $T = \dfrac{X_1 - \bar{X}_n}{S_n}$ is ancillary.

Define $Z_i = \dfrac{X_i - \mu}{\sigma} \sim N(0,1)$, $i = 1, \ldots, n$

$$T = \frac{X_1 - \bar{X}_n}{S_n} = \frac{[(X_1-\mu) - (\bar{X}_n - \mu)]/\sigma}{\left\{\frac{1}{n-1}\sum_{i=1}^n \left[\frac{(X_i-\mu)}{\sigma} - \frac{(\bar{X}_n-\mu)}{\sigma}\right]^2\right\}^{1/2}}$$

$$= \frac{Z_1 - \bar{Z}_n}{\sqrt{\frac{1}{n-1}\sum_{i=1}^n (Z_i - \bar{Z}_n)^2}}$$

$\leftarrow$ distribution of T does NOT depend on $\underset{\sim}{\theta} = (\mu, \sigma^2)$

$\therefore$ T is ancillary & hence independent of $(\bar{X}_n, S_n^2)$

2. To find UMVUE of $\Phi\left(\frac{2-\mu}{\sigma}\right)$, we're going to use Method II

(find $\underline{\text{some}}$ UE of $\Phi\left(\frac{2-\mu}{\sigma}\right)$ & compute $E(\cdot \mid \underset{\sim}{S})$, $\underset{\sim}{S} \equiv (\bar{X}_n, S_n^2)$)

Note $T_1 = \begin{cases} 1 & \text{if } X_1 \le 2 \\ 0 & \text{o.w.} \end{cases}$ is UE of $\Phi\left(\frac{2-\mu}{\sigma}\right)$

since $E_{\underset{\sim}{\theta}} T_1 = P_{\underset{\sim}{\theta}}(X_1 \le 2) = \Phi\left(\frac{2-\mu}{\sigma}\right)$

So, by Lehmann-Scheffe theorem, the UMVUE of $\Phi\left(\frac{2-\mu}{\sigma}\right)$ is

$$T^{*} = E(T_1 \mid \underset{\sim}{S}), \quad \underset{\sim}{S} = (\bar{X}_n, S_n^2)$$

$$= P(X_1 \le 2 \mid \underset{\sim}{S})$$

$$= P\left(\frac{X_1 - \bar{X}_n}{S_n} \le \frac{2-\bar{X}_n}{S_n} \,\Big|\, (\bar{X}_n, S_n^2)\right)$$

$$= P\left(T \le \frac{2-\bar{X}_n}{S_n} \,\Big|\, (\bar{X}_n, S_n^2)\right)$$

$$= \int_{-\infty}^{\frac{2-\bar{X}_n}{S_n}} f_{T\mid \underset{\sim}{S}}(t)\,dt \quad \leftarrow \text{conditional pdf of } T \text{ given } \underset{\sim}{S} = (\bar{X}_n, S_n^2)$$

$$= \int_{-\infty}^{\frac{2-\bar{X}_n}{S_n}} f_T(t)\,dt \quad \leftarrow \text{marginal pdf of } T \text{ (doesn't depend on } \underset{\sim}{\theta}) \text{ using independence by Basu's theorem}$$

$$= F_T\left(\frac{2-\bar{X}_n}{S_n}\right), \text{ where } F_T(t) = P(T \le t),\ t \in \mathbb{R}$$

could compute this (simulation)