STAT 611-600

Theory of Statistics - Inference
Lecture 3: Minimal Sufficiency

Tiandong Wang

Department of Statistics
Texas A&M University

Sufficient Statistics: Transformation

- If T is sufficient for θ and $T = c(U)$, a mathematical function of some other statistic U, then U is also sufficient.
- If T is sufficient for θ, and $U = r(T)$ with r being **one-to-one**, then U is also sufficient.
- Remark: When one statistic is a function of the other statistic and vice versa, then they carry exactly the same amount of information.

Transformed Sufficient Statistics: Examples

Examples:

- If $\sum_{i=1}^{n} X_i$ is sufficient, so is $\bar{X}$
- If $(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2)$ are sufficient, so are $(\bar{X}, S^2)$
- If $\sum_{i=1}^{n} X_i$ is sufficient, then $(\sum_{i=1}^{m} X_i, \sum_{i=m+1}^{n} X_i)$ are sufficient, and so are $(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2)$
- Examples of non-sufficiency
  Ex. X_1, X_2 iid $Poi(\lambda)$. $T = X_1 - X_2$ is not sufficient
  Ex. $X_1, \cdots, X_n$ iid with pmf $f(x;\theta)$. $T = (X_1, \cdots, X_{n-1})$ is not sufficient.

Minimal Sufficient Statistics

It is seen that different sufficient statistics are possible. **Which one is the best?**
(Of course, equivalent statistics need not be separately considered)
Naturally, the one with maximum possible reduction.

- For $N(\theta, 1)$, $\bar{X}$ is a better sufficient statistic for θ than $(\bar{X}, S^2)$.

Definition

T is a *minimal sufficient statistic* if, given any other sufficient statistic T', there is a function $c(\cdot)$ such that $T = c(T')$.

Minimal Sufficient Statistics

- Minimal sufficient statistic has the smallest dimension among possible sufficient statistics. Often the dimension is equal to the number of free parameters (exceptions do exist).
- Partition interpretation of minimal sufficient statistics
  - Any sufficient statistic introduces a partition on the sample space.
  - The partition of a minimal sufficient statistic is the *coarsest.*

Minimal Sufficient Statistics

How to Check Minimal Sufficiency?

Theorem

A statistic T is minimal sufficient if the following property holds: For any two sample points $\boldsymbol{x}$ and $\boldsymbol{y}$, $f(\boldsymbol{x};\theta)/f(\boldsymbol{y};\theta)$ does not depend on θ (i.e. $f(\boldsymbol{x};\theta)/f(\boldsymbol{y};\theta)$ is a constant function of θ) if and only if $T(\boldsymbol{x}) = T(\boldsymbol{y})$

Examples.

- X_i iid Uniform$[\theta, \theta + 1]$.
- X_i iid N(μ, σ^2).

Minimal Sufficient Statistics: Exponential families

For iid observations from an exponential family with the probability density $f(x;\theta) = c(\theta)h(\mathbf{x})\exp\{\sum_{j=1}^{k} w_j(\theta)t_j(\mathbf{x})\}$, where θ is d dimensional, if $w_j(\theta)$'s are linearly independent, then the statistic

$$T(\mathbf{X}) = (\sum_{i=1}^{n} t_1(\mathbf{X}_i), \cdots, \sum_{i=1}^{n} t_k(\mathbf{X}_i))$$

is minimal sufficient for θ.

- In particular, for full-rank exponential family (i.e. $d = k$), T is always minimal sufficient. For the distribution with multiple parameters $\theta = (\theta_1, \cdots, \theta_s)$, the minimal sufficient statistic is also a vector $T = (T_1(\mathbf{X}), \cdots, T_r(\mathbf{X}))$.
- Often $r = s$, but not always.
- For curved exponential families, we can have $r > s$. Example: X_i iid $N(\mu, \mu^2)$.

Remarks:

- Minimal sufficient statistic is not unique. Any two are in one-to-one correspondence, so are equivalent.
- If $T(\mathbf{X})$ is minimal sufficient for $\mathcal{P}_0$, and $T(\mathbf{X})$ is sufficient for $\mathcal{P}$, and $\mathcal{P}_0 \subset \mathcal{P}$, then $T(\mathbf{X})$ is minimal sufficient for $\mathcal{P}$.

  Examples: $\mathcal{P} = \{\text{all distributions with pdf}\}$,
  $\mathcal{P}_0 = \{\text{Cauchy distribution}\}$, $T(\mathbf{X}) = (X_{(1)}, \cdots, X_{(n)})$

Ancillary Statistic

Definition

A statistic T is *ancillary* if its distribution does not depend on θ.

- Induced family is completely known, contains no information about θ
- Opposite of sufficiency.

Ancillary Statistic: Location family

General Results for Location Family with pdf $f(x-\theta)$

- T is a location invariant statistic, i.e., $T(x_1+b,\cdots,x_n+b)=T(x_1,\cdots,x_n)$. Then T is ancillary.
- Sample std S is ancillary (and so are other estimates of scale).
- Examples: iid $N(\theta,1)$. Show S^2 is ancillary.
- Examples: Location family, iid pdf $f(x-\theta)$. Consider $R=X_{(n)}-X_{(1)}$.

Ancillary Statistic: Location-scale family

General Results for Location-scale Family with pdf $\frac{1}{\sigma}f(\frac{x-\mu}{\sigma})$

- T is a location-scale invariant statistic, i.e., $T(ax_1+b,\cdots,ax_n+b)=T(x_1,\cdots,x_n)$. Then T is ancillary.
- If T_1 and T_2 are two location-scale invariant statistics (see above), then T_1/T_2 is ancillary.
- Example: X_1, X_2 Ind $N(0,\sigma^2)$. Show X_1/X_2 is ancillary.
- Example: For iid $N(\mu,\sigma^2)$, $T=((X_1-\bar{X})/S,\cdots,(X_n-\bar{X})/S)$ is ancillary.