Biostat 602 Winter 2017

Exam Aid

Basic Terminology

Model $\mathcal{P} = \{f_{\mathbf{X}}(\mathbf{x}|\theta), \theta \in \Omega\}$, which can be a family of pdf's or pmf's

Random Variables $\mathbf{X} = (X_1, \cdots, X_n)$ that can be generated from $f_{\mathbf{X}}(\mathbf{x}|\theta)$.

Data $\mathbf{x} = (x_1, \cdots, x_n)$ that is generated from $f_{\mathbf{X}}(\mathbf{x}|\theta)$.

Joint pdf/pmf Joint pdf/pmf of a random sample is the product of pdf/pmf for every single observation.

Statistic A function of data or random variables $T(\mathbf{x})$ or $T(\mathbf{X})$.

Sample Space A set of possible values of random variables $\mathcal{X}$.

Partition $A_t = \{\mathbf{x} : T(\mathbf{x}) = t\} \subseteq \mathcal{X}$.

Data Reduction Partition of sample space in terms of particular statistic.

Exponential Family

Definition 3.4.1: The random variable X belongs to an exponential family of distributions, if its pdf/pmf can be written in the form

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

where

- $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_d), \ d \leq k$,
- $w_j(\theta), \ j \in \{1, \cdots, k\}$ and $c(\boldsymbol{\theta}) \geq 0$ are real valued functions of $\boldsymbol{\theta}$ alone,
- $t_j(x)$ and $h(x) \geq 0$ only involve data,
- Support of X, i.e. the set $A = \{x : f(x|\boldsymbol{\theta}) > 0\}$ does not depend on $\boldsymbol{\theta}$.

Sufficiency Principle

Sufficient Statistic

Concept The statistic contains all information about θ

Definition 6.2.1 $f_{\mathbf{X}}(\mathbf{x}|T(\mathbf{X}))$ does not depend on θ

Theorem 6.2.2 $f_{\mathbf{X}}(\mathbf{x}|\theta)/q(T(\mathbf{X})|\theta)$ does not depend on θ
$\Longrightarrow T(\mathbf{X})$ is sufficient.

Theorem 6.2.6 (Factorization) $f_{\mathbf{X}}(\mathbf{x}|\theta) = h(\mathbf{x})g(T(\mathbf{X})|\theta)$
$\Longleftrightarrow T(\mathbf{X})$ is sufficient.

Theorem 6.2.10 (Exponential Family) $(\sum_{i=1}^{n} t_1(X_i), \cdots, \sum_{i=1}^{n} t_k(X_i))$ is sufficient

Minimal Sufficient Statistic

Concept Sufficient statistic that achieves the maximum data reduction, or coarsest partition of the sample space.

Definition 6.2.11 T is sufficient and it is a function of every sufficient statistic.

Theorem 6.2.13 $T(\mathbf{X})$ is minimal sufficient if the following is true:
$f_{\mathbf{X}}(\mathbf{x}|\theta)/f_{\mathbf{X}}(\mathbf{y}|\theta)$ is constant as a function of $\theta \Longleftrightarrow T(\mathbf{x}) = T(\mathbf{y})$

Theorem 6.2.28 Any complete sufficient statistic is also minimal sufficient.

Ancillary Statistic

Concept A statistic that does not have any information about θ.

Definition 6.2.16 Its distribution is constant to θ.

Location Family A pdf $g(x|\theta)$ belongs to location family if $g(x|\theta) = \{f(x-\theta) : \theta \in R\}$, where f is free of θ. If the transformation $Y = X - \theta$ yields a random variable that has a distribution free of θ then X belongs to location family.

Scale Family A pdf $g(x|\theta)$ belongs to scale family if $g(x|\sigma) = \left\{\frac{1}{\sigma} f\left(\frac{x}{\sigma}\right) : \sigma > 0\right\}$, where f is free of σ. If the transformation $Y = X/\sigma$ yields a random variable that has a distribution free of σ then X belongs to scale family.

Theorem 6.2.24 (Basu) A complete sufficient statistic is independent of every ancillary statistic.

Complete Statistic

Concept Any non-zero function of the statistics cannot be ancillary (there is no unnecessary part).

Definition 6.2.21 T is complete if $E[g(T)|\theta] = 0 \Longrightarrow g(T) = 0$ almost surely across all θ.

Theorem 6.2.24 (Basu) A complete and (minimal) sufficient statistic is independent of every ancillary statistics.

Theorem 6.2.25 (Exponential Family) If the parameter space $\Theta = \{(w_1(\theta), \dots, w_k(\theta) : \theta \in \Omega\}$ contains an open subset of $\mathbb{R}^k$, then $(\sum_{i=1}^n t_1(X_i), \cdots, \sum_{i=1}^n t_k(X_i))$ is complete

Theorem 6.2.28 Any complete sufficient statistic is also minimal sufficient.

Summary of Point Estimation

Basics

Point Estimator Any function $W(\mathbf{X})$ of a sample, or any statistic.

Likelihood Function pdf/pmf as a function of θ given data, instead of a function of data given θ, i.e. $L(\theta|\mathbf{x}) = f_{\mathbf{X}}(\mathbf{x}|\theta)$

Score Function $u(\theta|X) = \frac{\partial}{\partial\theta} \log f_{\mathbf{X}}(\mathbf{x}|\theta) = \frac{\partial}{\partial\theta} \log L(\theta|\mathbf{x})$

Fisher Information $I_n(\theta) = \mathrm{E}\left[\{u(\theta|X)\}^2\right], \quad I_n(\theta) = nI(\theta)$ in the case of a random sample.

Estimators

Method-of Moments Estimator Estimator obtained by equating sample moments (function of $\mathbf{X}$) to their corresponding population moments (function of θ).

Maximum Likelihood Estimator (MLE) $\hat{\theta}$ is MLE if $L(\hat{\theta}(\mathbf{x})|\mathbf{x}) \geq L(\theta|\mathbf{x}) \quad , \forall\theta \in \Omega$ where $\hat{\theta}(\mathbf{x}) \in \Omega$

Criteria for evaluating Estimators

Bias: $\text{Bias}(\hat{\theta}) = \text{E}(\hat{\theta}) - \theta$

MSE: $\text{MSE}(\hat{\theta}) = \text{E}[(\hat{\theta} - \theta)]^2 = \left[\text{Bias}(\hat{\theta})\right]^2 + \text{Var}(\hat{\theta})$

Unbiased Estimator: An estimator $W(\mathbf{X})$ is unbiased for $\tau(\theta)$, if it satisfies

$$\text{E}[W(\mathbf{X})] = \tau(\theta)$$

Best Unbiased Estimator (UMVUE): Uniformly Minimum Variance Unbiased Estimator (UMVUE) for $\tau(\theta)$ is $W(\mathbf{X})$ such that

$$(i)\quad \text{E}[W(\mathbf{X})] \quad = \quad \tau(\theta)$$

and (ii) $\text{Var}[W(\mathbf{X})] \leq \text{Var}[W^*(\mathbf{X})]$ for any unbiased estimator $W^*(\mathbf{X})$ of $\tau(\theta)$.

Strategies for finding the best unbiased estimator

Cramer-Rao Theorem If $\text{E}[W(\mathbf{X})|\theta] = \tau(\theta)$, assuming regulatory conditions hold,

$$\text{Var}[W(\mathbf{X})|\theta] \geq \frac{[\tau'(\theta)]^2}{E\left[\{\frac{\partial}{\partial\theta}\log f_{\mathbf{X}}(\mathbf{X}|\theta)\}^2|\theta\right]} = \frac{[\tau'(\theta)]^2}{I_n(\theta)}$$

Lemma 7.3.11 $I(\theta) = \text{E}\left[\{u(\theta|X)\}^2\right] = -\text{E}\left[\frac{\partial^2}{\partial\theta^2}\log f_X(X|\theta)\right]$

Corollary 7.3.15 $W(\mathbf{X})$ is UMVUE that attains Cramer-Rao bound if and only if

$$\frac{\partial}{\partial\theta}\log L(\theta|\mathbf{X}) = u_n(\theta|\mathbf{X}) = a(\theta)[W(\mathbf{X}) - \tau(\theta)]$$

Exponential Family For one-parameter exponential family, $W(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^n t(X_i)$ is the best unbiased estimator for its expected value.

Theorem 7.3.19 (Uniqueness of UMVUE) UMVUE is unique.

Theorem 7.3.20 (UMVUE and UE of zeros) Best unbiased estimator $\Longleftrightarrow$ Uncorrelated with any unbiased estimators of zero.

Theorem 7.3.23 (Lehmann Scheffè) Theorem 7.3.23 - Any function of a complete sufficient statistic T that is unbiased for $\tau(\theta)$ is UMVUE for $\tau(\theta)$.