# Final Review

# Contents

- *The following sections of Casella and Berger*
  - 5.3, 5.4, 5.5.4
  - 6.1,6.2
  - 7.1, 7.2.1-7.2.2, 7.3
  - 8.1, 8.2.1-8.2.2, 8.3.1-8.3.2, 8.3.4
  - 9.1, 9.2, 9.3.1-9.3.2
  - 10.1.1-10.1.3, 10.3-10.4

# Main topics

Main topics covered

- Suffciency
- Point estimation
  - Method of moments
  - MLE
  - Bayes estimators
  - UMVUEs

# Main topics

- Testing
  - Neyman-Pearson lemma
  - Likelihood ratio tests
  - Bayesian tests
- Interval estimation
  - Inverting a test
  - Pivotal quantities
  - Bayes intervals
- Large sample inference

# Chapter 5

- Definitions (random Sample, sample size, statistic)
- Some special statistics
  - Sample mean
  - Sample variance
  - Order statistics, sample median
- Distributions of Sample mean and Sample variance
  - Student’s Theorem
- t and F distributions
- Distributions of Order Statistics

# Chapter 6

- Sufficient statistic
  - Minimal sufficiency
- Neyman's factorization theorem to find Sufficient statistic
- Ancillary statistics
  - Location, scale, or Location-scale family
- Complete Statistic
- Complete and sufficient statistic (CSS)
  - CSS independent with Ancillary statistics
- Exponential family
  - Completeness Complete and sufficient statistic in exponential family

# Chapter 7

- Point estimators (MME, MLE, Bayesian)
- Methods for finding MLES
  - Take partial derivative of log-likelihood with respect to each parameter and solve the likelihood equations.
- Invariance property of MLEs:
  - MLE of a function of $\theta$ is the function applied to MLE of $\theta$.

# Point Estimators

- Bayes estimates
  - Find posterior distribution.
  - Get posterior mean or median.
- Measures of Quality of Estimators
  - Unbiasness
  - Minimum mean-squared-error (MSE) estimator
  - UMVUE (*uniformly minimum variance unbiased estimator*)

# Quality of Estimators

- MSE
  - MSE=equals estimator variance plus estimator squared bias.
  - Finding an estimator with uniformly smaller MSE  than *every* other estimator is almost always impossible.
- Approaches for finding good estimators
  - Use an overall measure of risk: Bayes and minimax principles
  - Restrict the class of estimators: unbiased estimators, linear estimators, *. . . .*

# UE

- Score function and Fisher Information

$$S(\theta) = \frac{\partial \ln L(\theta)}{\partial \theta} = \frac{\partial \ln f(X \mid \theta)}{\partial \theta}$$

$$I(\theta) = E\left[\left(\frac{\partial \ln f(X \mid \theta)}{\partial \theta}\right)^2\right] = E\left[-\frac{\partial^2 \ln f(X \mid \theta)}{\partial \theta^2}\right]$$

- Theory of unbiased estimation (UE):
  - Cramer-Rao lower bound (CRLB)

If $T$ is a $UE$ for g($\theta$), under some regularity conditions,

$$Var(T) \geq \frac{[g'(\theta)]^2}{nI_1(\theta)}$$

# UMVUE

- Lehmann-Scheffe Theorem: A UE that is a function of a CSS is the unique UMVUE.

  If T is a *CSS* for $\theta$, then if there is $\varphi(T)$, a function of $T$, that is an UE of $g(\theta)$, then $\varphi(T)$ is UMVE of $g(\theta)$.

- Find UMVUE.
  - L-S can be used to find UMVUEs.
  - CRLB applied to best UE: Try to find an unbiased estimator whose variance achieves the lower bound

# Chapter 8

- Hypothesis testing
- *H*0 and *H*1
- Two actions: “reject *H*0” and “don't reject *H*0."
- Two types of errors
  - Type I error: reject *H*0 when *H*0 is true.
  - Type II error: accept *H*0 when *H*1 is true.
- Significance level (size of the test )
- P value
- *Power function:* $\beta(\theta)= P_{\theta}$(rejecting *H*0)

# Neyman-Pearson lemma

- Neyman-Pearson lemma $C = \{(x_{1,} ..., x_n) : \dfrac{L(\theta_0)}{L(\theta_1)} \le k\}.$
- Applying N-P lemma
  - Find MP(most powerful) test for testing simple hypotheses
  - Find UMP(uniformly most powerful): when N-P test does not depend on the alternative and has size $\alpha$, it is UMP size $\alpha$.
- UMP tests don't always exist.
  - Two sided test
- What to do if UMP test doesn't exist: Restrict unbiased class of tests and search for UMP test within that class.

# LRT

- Likelihood ratio test
  - Likelihood ratio statistic: Ratio of maximum of likelihood when θ is restricted to Θ0 to unrestricted maximum likelihood.

$$\lambda = \lambda(x_1, \cdots, x_n) = \frac{\sup_{\theta\in\Theta_0} L(\theta)}{\sup_{\theta\in\Theta} L(\theta)}$$

Reject $H_0$ in favor of $H_1$ if $\lambda \leq \lambda_0$, where $\lambda_0$ is determined by

$$\alpha = P[\lambda \leq \lambda_0 \mid H_0] = \sup_{\theta\in\Theta_0} P_\theta[\lambda \leq \lambda_0]$$

# Chapter 9

- Interval estimation
- coverage probability and confidence coeffcient
- Inverting a test to get a confidence interval
- Pivotal quantities
- Shortest length intervals

# Chapter 10

- Large sample distribution theory for point estimators (MLEs)
  - Consistency
    - Theorem: Mean squared error tending to 0 implies weak consistency.
  - Convergence in distribution (Asymptotic normality)
  - Asymptotic efficient

# Large sample tests

- Asymptotic distribution of likelihood ratio statistic:
  - Under $H0$, $-2 \log \lambda(X)$ converges in distribution to a random variable having a $\chi^2$ distribution as $n \to \infty$.
- LRT to assess multinomial goodness of fit the commonly used Pearson $\chi^2$ test is an approximation of the LRT
- Other large sample tests: Asymptotic normality of estimators can be used to construct large sample tests, such as Wald test, and score test

# Large sample confidence intervals

- Large sample confidence intervals
  - Inverting a large sample LRT, Wald, or score test to get a approximate confidence interval

# Chapter 11 ANOVA

- **One-way ANOVA**
  - Model assumption
  - Inferences regarding linear combinations of means
  - ANOVA F test
  - Partioning of Sums of Squares
- **Two-way ANOVA**
  - Without interactions
  - With interactions

$$T_{\mathbf{a}} = \frac{\sum_{i=1}^{k} a_i \bar{Y}_{i\bullet} - \sum_{i=1}^{k} a_i \theta_i}{\sqrt{S_p^2 \sum_{i=1}^{k} (a_i^2 / n_i)}} \sim t(N-k)$$

$$H_0: \quad \sum_{i=1}^{k} a_i \theta_i = 0$$

$$T_{\mathbf{a}}^{*} = \frac{\sum_{i=1}^{k} a_i \bar{Y}_{i\bullet}}{\sqrt{S_p^2 \sum_{i=1}^{k} (a_i^2 / n_i)}} \overset{H_0}{\sim} t(N-k)$$

$$H_0 : \theta_1 = \frac{1}{2}(\theta_2 + \theta_3) \qquad H_1 : \theta_1 \neq \frac{1}{2}(\theta_2 + \theta_3)$$

$$H_0: \quad \theta_1 = \theta_2 = \cdots = \theta_k, \qquad H_1: \quad \theta_i \neq \theta_j, \text{ for some } i, j.$$

$$SS_T = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \bar{\bar{Y}})^2 \overset{H_0}{\sim} \sigma^2 \chi^2 (N-1)$$

$$SS_W = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{i\bullet})^2 \sim \sigma^2 \chi^2 (N-k)$$

$$SS_B = \sum_{i=1}^{k} n_i \left( \bar{Y}_{i\bullet} - \bar{\bar{Y}} \right)^2 \overset{H_0}{\sim} \sigma^2 \chi^2 (k-1)$$

$$F = \frac{SS_B / (k-1)}{SS_W / (N-k)} = \frac{MS_B}{MS_W} \overset{H_0}{\sim} F(k-1, N-k)$$

- Reject $H_0$ if $F > F_{k-1, N-k, 1-\alpha}$

Table 11.2.1. *ANOVA table for oneway classification*

| Source of variation | Degrees of freedom | Sum of squares | Mean square | $F$ statistic |
|---|---|---|---|---|
| Between treatment groups | $k-1$ | SSB = $\sum n_i(\bar{y}_i - \bar{\bar{y}})^2$ | MSB = SSB/$(k-1)$ | $F = \frac{\text{MSB}}{\text{MSW}}$ |
| Within treatment groups | $N-k$ | SSW = $\sum\sum(y_{ij} - \bar{y}_{i\cdot})^2$ | MSW = SSW/$(N-k)$ | |
| Total | $N-1$ | SST = $\sum\sum(y_{ij} - \bar{\bar{y}})^2$ | | |

# Two way ANOVA
# Partitioning of Variation (without interaction)

$$SST = \sum_{j=1}^{b}\sum_{i=1}^{a}(X_{ij} - \bar{X}_{..})^2 \qquad df(T) = ab - 1$$

---

$$SSA = b\sum_{i=1}^{a}(\bar{X}_{i\bullet} - \bar{X}_{\bullet\bullet})^2 \qquad df(A) = a - 1$$

$$SSB = a\sum_{j=1}^{b}(\bar{X}_{\bullet j} - \bar{X}_{\bullet\bullet})^2 \qquad df(B) = b - 1$$

$$SSE = \sum_{j=1}^{b}\sum_{i=1}^{a}(X_{ij} - \bar{X}_{i\bullet} - \bar{X}_{\bullet j} + \bar{X}_{..})^2 \qquad df(E) = (a-1)(b-1)$$

**SST=SSA+SSB+SSE**

**dfT=dfA+dfB+dfE**

# Partitioning of Variation (with interaction)

$$SST = \sum_{j=1}^{b}\sum_{i=1}^{a}\sum_{k=1}^{n_{ij}}(X_{ijk} - \bar{X}_{\dots})^2 \qquad df(T) = abc - 1$$

---

$$SSA = \sum_{i=1}^{a} n_{i\bullet}(\bar{X}_{i\bullet\bullet} - \bar{X}_{\bullet\bullet\bullet})^2 \qquad df(A) = a - 1$$

$$SSB = \sum_{j=1}^{b} n_{\bullet j}(\bar{X}_{\bullet j\bullet} - \bar{X}_{\bullet\bullet\bullet})^2 \qquad df(B) = b - 1$$

$$SSAB = \sum_{j=1}^{b}\sum_{i=1}^{a} n_{ij}(\bar{X}_{ij\bullet} - \bar{X}_{i\bullet\bullet} - \bar{X}_{\bullet j\bullet} + \bar{X}_{\bullet\bullet\bullet})^2 \qquad df(AB) = (a-1)(b-1)$$

$$SSE = \sum_{j=1}^{b}\sum_{i=1}^{a}\sum_{k=1}^{n_{ij}}(X_{ijk} - \bar{X}_{ij\bullet})^2 \qquad df(E) = N - ab$$