

STAT 3690 Lecture 09

zhiyanggeezhou.github.io

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca)

Feb 11, 2022

Sampling distributions of $\bar{\mathbf{X}}$ and \mathbf{S} (J&W Sec 4.4)

- Recall the univariate case: $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$, $n > p$
 - $S^2 \perp\!\!\!\perp \bar{X}$
 - $\sqrt{n}(\bar{X} - \mu)/\sigma \sim N(0, 1)$
 - $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$, where $S^2 = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X})^2$
 - $\sqrt{n}(\bar{X} - \mu)/S \sim t(n-1)$
 - The multivariate case
 - $\mathbf{S} \perp\!\!\!\perp \bar{\mathbf{X}}$, i.e., $\hat{\Sigma}_{\text{ML}} \perp\!\!\!\perp \hat{\mu}_{\text{ML}}$
 - $\sqrt{n}\Sigma^{-1/2}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim MVN_p(\mathbf{0}, \mathbf{I})$
 - $(n-1)\mathbf{S} = n\hat{\Sigma}_{\text{ML}} \sim W_p(n-1, \Sigma)$
 - $n(\bar{\mathbf{X}} - \boldsymbol{\mu})^\top \mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim \text{Hotelling's } T^2(p, n-1)$
-

- Wishart distribution
 - Def: $W_p(\Sigma, n)$ is the distribution of $\sum_{i=1}^n \mathbf{Y}_i \mathbf{Y}_i^\top$ with $\mathbf{Y}_1, \dots, \mathbf{Y}_n \stackrel{\text{iid}}{\sim} MVN_p(\mathbf{0}, \Sigma)$
 - A generalization of χ^2 -distribution: $W_p(\Sigma, n) = \chi^2(n)$ if $p = \Sigma = 1$
 - Properties
 - $\mathbf{A}\mathbf{A}^\top > 0$ and $\mathbf{W} \sim W_p(\Sigma, n) \Rightarrow \mathbf{A}\mathbf{W}\mathbf{A}^\top \sim W_p(\mathbf{A}\Sigma\mathbf{A}^\top, n)$
 - $\mathbf{W}_i \stackrel{\text{iid}}{\sim} W_p(\Sigma, n_i) \Rightarrow \mathbf{W}_1 + \mathbf{W}_2 \sim W_p(\Sigma, n_1 + n_2)$
 - $\mathbf{W}_1 \perp\!\!\!\perp \mathbf{W}_2$, $\mathbf{W}_1 + \mathbf{W}_2 \sim W_p(\Sigma, n)$ and $\mathbf{W}_1 \sim W_p(\Sigma, n_1) \Rightarrow \mathbf{W}_2 \sim W_p(\Sigma, n - n_1)$
 - $\mathbf{W} \sim W_p(\Sigma, n)$ and $\mathbf{a} \in \mathbb{R}^p \Rightarrow$

$$\frac{\mathbf{a}^\top \mathbf{W} \mathbf{a}}{\mathbf{a}^\top \Sigma \mathbf{a}} \sim \chi^2(n)$$

- $\mathbf{W} \sim W_p(\Sigma, n)$, $\mathbf{a} \in \mathbb{R}^p$ and $n \geq p \Rightarrow$

$$\frac{\mathbf{a}^\top \Sigma^{-1} \mathbf{a}}{\mathbf{a}^\top \mathbf{W}^{-1} \mathbf{a}} \sim \chi^2(n - p + 1)$$

- $\mathbf{W} \sim W_p(\Sigma, n) \Rightarrow$

$$\text{tr}(\Sigma^{-1} \mathbf{W}) \sim \chi^2(np)$$

-
- Hotelling's T^2 distribution
 - A generalization of (Student's) t -distribution
 - If $\mathbf{X} \sim MVN_p(\mathbf{0}, \mathbf{I})$ and $\mathbf{W} \sim W_p(\mathbf{I}, n)$, then

$$\mathbf{X}^\top \mathbf{W}^{-1} \mathbf{X} \sim T^2(p, n)$$

- $Y \sim T^2(p, n) \Leftrightarrow \frac{n-p+1}{np} Y \sim F(p, n-p+1)$
-

- Wilk's lambda distribution
 - Wilks's lambda is to Hotelling's T^2 as F distribution is to Student's t in univariate statistics.
 - Given independent $\mathbf{W}_1 \sim W_p(\Sigma, n_1)$ and $\mathbf{W}_2 \sim W_p(\Sigma, n_2)$ with $n_1 \geq p$,

$$\Lambda = \frac{\det(\mathbf{W}_1)}{\det(\mathbf{W}_1 + \mathbf{W}_2)} = \frac{1}{\det(\mathbf{I} + \mathbf{W}_1^{-1}\mathbf{W}_2)} \sim \Lambda(p, n_1, n_2)$$

- Resort to approximations for computation: $\{(p - n_2 + 1)/2 - n_1\} \ln \Lambda(p, n_1, n_2) \approx \chi^2(n_2 p)$

Hypothesis testing

- Model: $\mathbf{X}_1, \dots, \mathbf{X}_n \sim f_{\theta^*} \in \{f_{\theta} : \theta \in \Theta\}$
 - $\theta^* \in \mathbb{R}^p$: the p -vector of parameters of interest, fixed and unknown
 - $\Theta \subset \mathbb{R}^p$: the parameter space
- Hypotheses $H_0 : \theta^* \in \Theta_0$ v.s. $H_1 : \theta^* \in \Theta_1$
 - $\Theta_0 \cap \Theta_1 = \emptyset$
 - $\Theta_0 \cup \Theta_1 = \Theta$
- Rejection/critical region R
 - Reject H_0 if observations $\in R$
- Level α : $\sup_{\theta \in \Theta_0} \Pr_{\theta}(\text{observations} \in R) \leq \alpha$
 - When $\theta^* \in \Theta_0$, $\Pr(\text{type I error}) = \Pr_{\theta^*}(\text{observations} \in R) \leq \sup_{\theta \in \Theta_0} \Pr_{\theta}(\text{observations} \in R) \leq \alpha$
 - Type I error: H_0 is incorrectly rejected
 - Type II error: H_0 is incorrectly accepted
- p -value: a specific rejection region
 - Reject H_0 if p -value $\in [0, \alpha]$
 - Not always well-defined
- Necessary components in reporting a testing result
 1. Hypotheses
 2. Name of approach
 3. Value of test statistic
 4. Rejection region/ p -value
 5. Conclusion: e.g., at the α level, we reject/do not reject H_0 , i.e., we believe XXX

Likelihood ratio test (LRT)

- Minimize the type II error rate subject to a capped type I error rate (under certain classical circumstances)
- Test statistic

$$\lambda(\mathbf{X}_1, \dots, \mathbf{X}_n) = \frac{\sup_{\theta \in \Theta_0} L(\theta; \mathbf{X}_1, \dots, \mathbf{X}_n)}{\sup_{\theta \in \Theta} L(\theta; \mathbf{X}_1, \dots, \mathbf{X}_n)} = \frac{L(\hat{\theta}_0; \mathbf{X}_1, \dots, \mathbf{X}_n)}{L(\hat{\theta}; \mathbf{X}_1, \dots, \mathbf{X}_n)}$$

- $\hat{\theta}_0$: ML estimator for $\theta \in \Theta_0$
- $\hat{\theta}$: ML estimator for $\theta \in \Theta$

- Rejection region $R = \{\mathbf{x}_1, \dots, \mathbf{x}_n : \lambda(\mathbf{x}_1, \dots, \mathbf{x}_n) \leq c\}$
 - \mathbf{x}_i is the realization of \mathbf{X}_i
 - c is chosen such that

$$\sup_{\theta \in \Theta_0} \Pr_{\theta}(\lambda(\mathbf{X}_1, \dots, \mathbf{X}_n) \leq c) = \alpha.$$

- * Have to know the null distribution of $\lambda(\mathbf{X}_1, \dots, \mathbf{X}_n)$, i.e., the distribution of $\lambda(\mathbf{X}_1, \dots, \mathbf{X}_n)$ with $\theta \in \Theta_0$

- p -value

$$p(\mathbf{x}_1, \dots, \mathbf{x}_n) = \sup_{\theta \in \Theta_0} \Pr_{\theta}\{\lambda(\mathbf{X}_1, \dots, \mathbf{X}_n) \leq \lambda(\mathbf{x}_1, \dots, \mathbf{x}_n)\}$$

- Null distribution of $\lambda(\mathbf{X}_1, \dots, \mathbf{X}_n)$

- Use the accurate distribution of $\lambda(\mathbf{X}_1, \dots, \mathbf{X}_n)$ if it is known; otherwise see below for an approximation.
- As $n \rightarrow \infty$,

$$-2 \ln \lambda(\mathbf{X}_1, \dots, \mathbf{X}_n) \approx \chi^2(\nu)$$

- * ν : the difference in numbers of free parameters between H_0 and H_1
- * Leading to an (asymptotic) rejection region $\{\mathbf{x}_1, \dots, \mathbf{x}_n : -2 \ln \lambda(\mathbf{x}_1, \dots, \mathbf{x}_n) \geq \chi_{\nu, 1-\alpha}^2\}$
 - $\chi_{\nu, 1-\alpha}^2$ is the $(1 - \alpha)$ - quantile of $\chi^2(\nu)$.