

# 2022-2023 学年多元统计分析期末测试

命题:杨建峰 (回忆:Mathzwj)

一.(1)最简单的题是最容易被忘记的.

(2)设  $\mathbf{X}_{(n \times p)}$  为来自总体  $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  的样本阵,  $A$  与  $B$  分别为  $m \times n$  与  $p \times q$  阶阵,  $\mathbf{Y} = \mathbf{A}\mathbf{X}\mathbf{B}$ . 则  $\mathbf{Y}^V$  的分布为\_\_\_\_\_.

(3)设  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)'$  为来自总体  $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  的样本阵,  $n > p$ ,  $\boldsymbol{\Sigma} > \mathbf{0}$ . 记  $\bar{\mathbf{x}}$  与  $\mathbf{S}_n$  分别为样本均值与样本协方差阵, 则  $(n-1)\mathbf{S}_n$  的分布为\_\_\_\_\_,  $n\bar{\mathbf{x}}'\mathbf{S}_n^{-1}\bar{\mathbf{x}}$  的分布为\_\_\_\_\_.

(4)设  $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ , 有分块形式  $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}$ ,  $\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}$ ,  $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \geq \mathbf{0}$ ,  $\boldsymbol{\Sigma}_{22} > \mathbf{0}$ , 其中  $\mathbf{x}_1$  是  $r$  维的. 则  $\mathbf{x}_1 | \mathbf{x}_2$  的分布为\_\_\_\_\_.

二. 写出一随机向量服从多元正态分布  $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  的四种定义.

三. 设  $\boldsymbol{\Sigma} = k\boldsymbol{\Sigma}_0$  并且  $\boldsymbol{\mu}$  不受约束, 证明:  $k$  的 MLE 为  $\hat{k} = p^{-1}\text{tr}(\boldsymbol{\Sigma}_0^{-1}\mathbf{S})$  且  $E(\hat{k}) = \frac{n-1}{n}k$ .

四. 设  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)'$  为来自总体  $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  的样本,  $\boldsymbol{\Sigma} > \mathbf{0}$  且未知,  $p = 2k$ . 设  $\boldsymbol{\mu}$  的前  $k$  个分量为  $\boldsymbol{\mu}_{(1)}$ , 后  $k$  个分量为  $\boldsymbol{\mu}_{(2)}$ , 试导出  $H_0: \boldsymbol{\mu}_{(1)} = \boldsymbol{\mu}_{(2)}$  的 LRT 统计量和分布.

五. 对于主成分分析, 推导在  $R^n$  空间中拟合和在  $R^p$  空间中拟合的关系.

六. 写出数据阵  $\mathbf{X}_{(n \times p)}$  具有  $k$  因子模型的定义, 叙述主成分分析与因子分析的区别与联系.

七. 说明聚类分析与判别分析的差别以及如何使用移动中心聚类法进行聚类分析.

说明: 回忆的时候距离考试已过去半年以上, 部分题目已遗忘或者记忆不清楚.