L E C T U R E 8

§8. INFERENCES ABOUT A MEAN VECTOR

8.1 INTRODUCTION

This lecture is the first of the methodological sections of the course. We now use the concepts and results set forth in Lectures 1–7, to develop techniques for analyzing data. A large part of any analysis is concerned with inference-that is, reaching valid conclusions concern n population on the basis of information from a sample. At this point, we concentrate on inferences about a population mean vector and its component parts. Although we introduce statistical inference through initial discussions of tests of hypotheses, our ultimate aim is to present a full statistical analysis of the component means based on simultaneous confidence statements. One of the central messages of multivariate analysis is that m correlated variables must be analyzed jointly. This principle is exemplified by the methods presented in Lecture 8.

8.2 THE PLAUSIBILITY OF μ_0 AS A VALUE FOR A NORMAL POPULATION MEAN

Let us start by recalling the univariate theory for determining whether a specific value μ_0 is a plausible value for the population mean μ. From the point of view of hypothesis testing, this problem can be formulated as a test of the competing hypotheses

$$H_0: \qquad \mu = \mu_0 \qquad \text{and} \qquad H_1: \qquad \mu \neq \mu_0$$

Here H_0 is the null hypothesis and H_1 is the (two-sided) alternative hypothesis. If $X_1, X_2, \ldots, X_n$ denote a random sample from a normal population, the appropriate test statistic is

$$t = \frac{\overline{X} - \mu_0}{s/\sqrt{n}}, \quad \text{where} \quad \overline{X} = \frac{1}{n}\sum_{j=1}^{n} X_j \quad \text{and} \quad s^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \overline{X})^2$$

This test statistic has a student's t-distribution with $n-1$ degrees of freedom. We reject H_0 , that μ_0 is a plausible value of μ, if the observed $|t|$ exceeds a specified percentage point of a t-distribution with $n-1$ d.f. Rejecting H_0 when $|t|$ is large is equivalent to rejecting H_0 if its square,

$$t^2 = \frac{(\overline{X}-\mu_0)^2}{s^2/n} = n(\overline{X}-\mu_0)(s^2)^{-1}(\overline{X}-\mu_0) \tag{8.1}$$

is large. The variable t^2 in (7.1) is the square of the distance from the sample mean $\overline{X}$ to the test value μ_0. The units of distance are expressed in terms of $s/\sqrt{n}$ or estimated standard deviations of $\overline{X}$. Once $\overline{X}$ and s^2 are observed, the test becomes: Reject H_0 in favor of H_1, at significance level α, if

$$n(\overline{x}-\mu_0)\,(s^2)^{-1}\,(\overline{x}-\mu_0) > t^2_{n-1}(\alpha/2), \tag{8.2}$$

where $t_{n-1}(\alpha/2)$ denotes the upper $100(\alpha/2)$th percentile of the t-distribution with n - 1 degrees of freedom.

If H_0 is not rejected, we conclude that μ_0 is a plausible value for the normal population mean. Are there other values of μ which are also consistent with the data? The answer is yes! In fact, there is always a set of plausible values for a normal population mean. From the well-known correspondence between acceptance regions for tests of $H_0: \quad \mu = \mu_0$ versus $H_1: \quad \mu \neq \mu_0$ and confidence intervals for μ we have

$$\{\text{Do not reject} \quad H_0 : \mu = \mu_0 \quad \text{at level} \quad \alpha\} \quad \text{or} \quad \left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right| \le t_{n-1}(\alpha/2)$$

is equivalent to

$$\left\{\mu_0 \quad \text{lies in the} \quad 100(1-\alpha)\% \quad \text{confidence interval} \quad \overline{x} \pm t_{n-1}(\alpha/2)\,\frac{s}{\sqrt{n}}\right\}$$

or

$$\overline{x} - t_{n-1}(\alpha/2) \le \mu_0 \le \overline{x} + t_{n-1}(\alpha/2) \tag{8.3}$$

The confidence interval consists of all those values μ_0 that would not be rejected by the level α test of $H_0: \mu = \mu_0$.

Before the sample is selected, the $100(1-\alpha)\%$ confidence interval in (8.3) is a random interval because the endpoints depend upon the random variables $\overline{X}$ and s. The probability that the interval contains μ is $1-\alpha$; among large numbers of such independent intervals, approximately $100(1-\alpha)\%$ of them will contain μ.

Consider now the problem of determining whether a given $m\times 1$ vector μ_0 is a plausible value for the mean of a multivariate normal distribution. We will proceed by analogy to the univariate development just presented.

A natural generalization of the squared distance in (8.1) is its multivariate analog

$$T^2 = (\overline{\mathbf{X}} - \mu_0)' \left(\frac{1}{n}\mathbf{S}\right)^{-1} (\overline{\mathbf{X}} - \mu_0) = n(\overline{\mathbf{X}} - \mu_0)' \, \mathbf{S}^{-1}(\overline{\mathbf{X}} - \mu_0) \tag{8.4}$$

where

$$\underset{(m\times 1)}{\overline{\mathbf{X}}} = \frac{1}{n}\sum_{j=1}^{n}\mathbf{X}_j, \qquad \underset{(m\times m)}{\mathbf{S}} = \frac{1}{n-1}\sum_{j=1}^{n}(\mathbf{X}_j - \overline{\mathbf{X}})(\mathbf{X}_j - \overline{\mathbf{X}})', \quad \text{and} \quad \underset{(m\times 1)}{\mu_0} = \begin{bmatrix} \mu_{10} \\ \mu_{20} \\ . \\ . \\ . \\ \mu_{m0} \end{bmatrix}$$

The statistic T^2 is called Hotelling's T^2 in honor of Harold Hotelling, a pioneer in multivariate analysis, who first obtained its sampling distribution. Here $\frac{1}{n}\mathbf{S}$ is the estimated covariance matrix of $\mathbf{X}$ (see Theorem 6.1).

If the observed statistical distance T^2 is too large - that is, if $\overline{\mathbf{x}}$ is "too far" from μ_0 – the hypothesis $H_0 : \mu = \mu_0$ is rejected. It turns out that special tables of T^2 percentage points are not required for formal tests of hypotheses. This is true because

$$T^2 \quad \text{is distributed as} \quad \frac{(n-1)m}{n-m}F_{m,n-m} \tag{8.5}$$

where $F_{m,n-m}$ denotes a random variable with an F-distribution with m and $n-m$ degrees of freedom.

To summarize, we have the following:

Let $\mathbf{X}_1$, $\mathbf{X}_2$, ..., $\mathbf{X}_n$ be a random sample from an $N_m(\mu, \Sigma)$ population. Then with $\overline{\mathbf{X}} = \frac{1}{n}\sum_{j=1}^{n}\mathbf{X}_j$, and

$$\mathbf{S} = \frac{1}{n-1}\sum_{j=1}^{n}(\mathbf{X}_j - \overline{\mathbf{X}})(\mathbf{X}_j - \overline{\mathbf{X}})',$$

$$\alpha = P\left[T^2 > \frac{(n-1)m}{n-m}F_{m,n-m}(\alpha)\right] =$$

$$P\left[n(\overline{\mathbf{X}}-\mu_0)'\,\mathbf{S}^{-1}(\overline{\mathbf{X}}-\mu_0) > \frac{(n-1)m}{n-m}F_{m,n-m}(\alpha)\right]. \tag{8.5}$$

Statement (8.5) leads immediately to a test of the hypothesis $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$. At the α level of significance, we reject H_0 in favor of H_1 if the observed

$$T^2 = n(\overline{\mathbf{X}}-\mu_0)'\,\mathbf{S}^{-1}(\overline{\mathbf{X}}-\mu_0) > \frac{(n-1)m}{n-m}F_{m,n-m}(\alpha)$$

It is informative to discuss the nature of the T^2-distribution briefly and its correspondence with the univariate test statistic.

Example 8.1. (Evaluating T^2.) Let the data matrix for a random sample of size $n = 3$ from a bivariate normal population be

$$\mathbf{X} = \begin{bmatrix} 6 & 9 \\ 10 & 6 \\ 8 & 3 \end{bmatrix}$$

Evaluate the observed T^2 for $\mu_0' = [9, 5]$. What is the sampling distribution of T^2 in this case? We find

$$\overline{\mathbf{x}} = \begin{bmatrix} \overline{x}_1 \\ \overline{x}_2 \end{bmatrix} = \begin{bmatrix} \frac{6+10+8}{3} \\ \frac{9+6+3}{3} \end{bmatrix} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$$

and

$$s_{11} = \frac{(6-8)^2 + (10-8)^2 + (8-8)^2}{2} = 4$$

$$s_{12} = \frac{(6-8)(9-6) + (10-8)(6-6) + (8-8)(3-6)}{2} = -3$$

$$s_{22} = \frac{(9-6)^2 + (6-6)^2 + (3-6)^2}{2} = 9$$

so

$$\mathbf{S} = \begin{bmatrix} 4 & -3 \\ -3 & 9 \end{bmatrix}$$

Thus,

$$\mathbf{S}^{-1} = \frac{1}{(4)(9)-(-3)(-3)}\begin{bmatrix} 9 & 3 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{9} \\ \frac{1}{9} & \frac{4}{27} \end{bmatrix}$$

and, from (8.4),

$$\mathbf{T}^2 = 3\,[\,8-9, \quad 6-5\,]\begin{bmatrix} \frac{1}{3} & \frac{1}{9} \\ \frac{1}{9} & \frac{4}{27} \end{bmatrix}\begin{bmatrix} 8-9 \\ 6-5 \end{bmatrix} = 3\,[\,-1, \quad 1\,]\begin{bmatrix} -\frac{2}{9} \\ \frac{1}{27} \end{bmatrix} = \frac{7}{9}$$

Before the sample is selected, T^2 has the distribution of a

$$\frac{(3-1)2}{(3-2)}F_{2,3-2} = 4F_{2,1}$$

random variable.

The next example illustrates a test of the hypothesis $H_0 : \mu = \mu_0$ using data collected as part of a search for new diagnostic techniques at the university of Wisconsin Medical School.

Example 8.2 (Testing a multivariate mean vector with T^2.) Perspiration from 20 healthy females was analyzed. Three components, X_1 =sweat rate, X_2 = sodium content, and X_3 = potassium content, were measured. Test the hypothesis $H_0 : \mu' = [4,\ 50,\quad 10]$ against $H_0 : \mu' \neq [4,\ 50,\quad 10]$ at level of significance $\alpha = 0.1$.

Computer calculations provide

$$\overline{\mathbf{x}} = \begin{bmatrix} 4.640 \\ 45.400 \\ 9.965 \end{bmatrix}$$

,

$$\mathbf{S} = \begin{bmatrix} 2.879 & 10.010 & -1.810 \\ 10.010 & 199.788 & -5.640 \\ -1.810 & -5.640 & 3.628 \end{bmatrix}$$

and

$$\mathbf{S}^{-1} = \begin{bmatrix} .586 & -.022 & .258 \\ -.022 & .006 & -.002 \\ .258 & -.002 & .402 \end{bmatrix}$$

We evaluate

$$T^2 = 20[4.64-4,\ 45.4-50,\ 9.965-10] \begin{bmatrix} .586 & -.022 & .258 \\ -.022 & .006 & -.002 \\ .258 & -.002 & .402 \end{bmatrix} \begin{bmatrix} 4.64-4 \\ 45.4-50 \\ 9.965-10 \end{bmatrix} = 9.74$$

Comparing the observed $T^2 = 9.74$ with the critical value

$$\frac{(n-1)m}{n-m}F_{m,n-m}(0.1) = 3.353 \cdot 2.44 = 8.18$$

we see that $T^2 = 9.74 > 8.18$, and consequently, we reject H_0 at 10% level of significance.