INFERENCES ABOUT THE MEAN VECTOR

Reference: Johnson & Wichern (2007) *Applied Multivariate Statistical Analysis* Chapter 5.

1 Hypotheses Testing for $\boldsymbol{\mu}$

Problem:

Given a random sample $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ from a p variate normal (or approximately normal) population with mean $\boldsymbol{\mu}$ and covariance matrix Σ where Σ is unknown. Examine whether $\boldsymbol{\mu}$ is significantly different from a given value, $\boldsymbol{\mu}_0$ at a specified level of significance, say α.

Note: Since $\boldsymbol{X}_j \sim N_p(\boldsymbol{\mu}, \Sigma)$, sample mean $\overline{\boldsymbol{X}}_n \sim N_p\left(\boldsymbol{\mu}, \frac{1}{n}\Sigma\right)$.

Hypotheses: $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ against $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$.

Test Statistic: $T^2 = n\left(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}_0\right)^T \mathcal{S}_n^{-1}\left(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}_0\right)$.

Note that T^2 has a Hoteling's T^2 distribution and hence, $\dfrac{n-p}{p(n-1)}T^2$ has a F-distribution with p and $n-p$ degrees of freedom.

Decision: Reject H_0 if

$$\frac{n-p}{p(n-1)}T^2 > F_{p,n-p}(\alpha)$$

where $F_{p,n-p}(\alpha)$ is the upper 100α percentile of the $F_{p,n-p}$ distribution.

2 Confidence Region for $\boldsymbol{\mu}$

Since the distribution of $\frac{n-p}{p(n-1)}T^2$ is $F_{p,n-p}$,

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

Hence, $\mathcal{P}\left[T^2 \leq \frac{p(n-1)}{n-p}F_{p,n-p}(\alpha)\right] = 1-\alpha$ and

$$\mathcal{P}\left[n\left(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}\right)^T \mathcal{S}_n^{-1}\left(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}\right) \leq \frac{p(n-1)}{n-p}F_{p,n-p}(\alpha)\right] = 1-\alpha$$

Thus $\overline{\boldsymbol{X}}_n$ will be within $\left[\frac{p(n-1)}{n-p}F_{p,n-p}(\alpha)\right]^{1/2}$ of $\boldsymbol{\mu}$ with probability $1-\alpha$ provided distance is defined in terms of $(\mathcal{S}_n/n)^{1/2}$. Therefore, a $100(1-\alpha)\%$ confidence region for $\boldsymbol{\mu}$ is

$$n\left(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}\right)^T \mathcal{S}_n^{-1}\left(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}\right) \leq \frac{p(n-1)}{n-p}F_{p,n-p}(\alpha).$$

If $\overline{\boldsymbol{x}}_n$ is the sample average of the observed values, then

$$n\left(\overline{\boldsymbol{x}}_n - \boldsymbol{\mu}\right)^T \mathcal{S}_n^{-1}\left(\overline{\boldsymbol{x}}_n - \boldsymbol{\mu}\right) \leq \frac{p(n-1)}{n-p}F_{p,n-p}(\alpha).$$

This is an ellipsoid with center at $\overline{\boldsymbol{x}}_n$. The axes and relative lengths are given by eigenvectors $\boldsymbol{e}_i$ and eigenvalues λ_i of $\mathcal{S}_n$.

Let $c^2 = \frac{p(n-1)}{n-p}F_{p,n-p}(\alpha)$ then length of the axis along $\boldsymbol{e}_i$ is $2\sqrt{\frac{\lambda_i}{n}}c$. That is the axes of the p-variate confidence ellipsoid are

$$\overline{\boldsymbol{x}}_n \pm 2\sqrt{\frac{\lambda_i}{n}}c\ \boldsymbol{e}_i \quad \text{for } i = 1, 2, \ldots, p.$$

3 Simultaneous Confidence Interval (T^2 - Interval)

Let $\boldsymbol{a}$ be a $p \times 1$ constant vector, then for ever $\boldsymbol{a}$

$$\boldsymbol{a}^T \overline{\boldsymbol{X}}_n \pm \sqrt{\frac{p(n-1)}{n-p} F_{p,n-p}(\alpha) \ \frac{\boldsymbol{a}^T \mathcal{S}_n \boldsymbol{a}}{n}}$$

is a $100(1-\alpha)\%$ confidence interval for $\boldsymbol{a}^T \boldsymbol{\mu}$.

The successive choices of $\boldsymbol{a}^T = (1, 0, \ldots, 0)$, $\boldsymbol{a}^T = (0, 1, \ldots, 0)$, and so on through $\boldsymbol{a}^T = (0, 0, \ldots, 1)$ gives the confidence interval for $\mu_1, \mu_2, \ldots, \mu_p$ respectively.

$$\begin{aligned}
\overline{X}_{1n} \quad & \pm\sqrt{\tfrac{p(n-1)}{n-p} F_{p,n-p}(\alpha)} \ \sqrt{\tfrac{S_{11}}{n}}, \\
\overline{X}_{2n} \quad & \pm\sqrt{\tfrac{p(n-1)}{n-p} F_{p,n-p}(\alpha)} \ \sqrt{\tfrac{S_{22}}{n}}, \\
& \vdots \\
\overline{X}_{pn} \quad & \pm\sqrt{\tfrac{p(n-1)}{n-p} F_{p,n-p}(\alpha)} \ \sqrt{\tfrac{S_{pp}}{n}}.
\end{aligned}$$

where $\overline{\boldsymbol{X}}_n^T = (\overline{X}_{1n}, \overline{X}_{2n}, \ldots, \overline{X}_{pn})$. Thus confidence interval for μ_i, $(i = 1, 2, \ldots, p)$ is given by

$$\left(\overline{X}_{in} - \sqrt{\frac{p(n-1)}{n-p} F_{p,n-p}(\alpha)} \ \sqrt{\frac{S_{ii}}{n}}, \quad \overline{X}_{in} + \sqrt{\frac{p(n-1)}{n-p} F_{p,n-p}(\alpha)} \ \sqrt{\frac{S_{ii}}{n}} \right).$$

Further taking $\boldsymbol{a}^T = (0, 0, \ldots, a_i, 0, \ldots, 0, a_k, 0, \ldots, 0)$ with $a_i = -a_k = 1$, we can obtain a $(1-\alpha)\%$ confidence interval for $\mu_i - \mu_k$ as follows:

$$\Bigg(\big(\overline{X}_{in} - \overline{X}_{kn}\big) - \sqrt{\tfrac{p(n-1)}{n-p} F_{p,n-p}(\alpha)} \ \sqrt{\tfrac{S_{ii} - 2S_{ik} + S_{kk}}{n}},$$
$$\big(\overline{X}_{in} - \overline{X}_{kn}\big) + \sqrt{\tfrac{p(n-1)}{n-p} F_{p,n-p}(\alpha)} \ \sqrt{\tfrac{S_{ii} - 2S_{ik} + S_{kk}}{n}} \Bigg).$$

Note: Since $\dfrac{p(n-1)}{n-p} F_{p,n-p} = T^2$, the above confidence intervals are called T^2 intervals.

4 Bonferroni Intervals

Bonferroni confidence intervals can be computed for any number $m(\leq p)$ of individuals means. The length of a Bonferroni interval is shorter than the length of the corresponding T^2 interval and they are equal only if $m = p = 1$.

For any $m(\leq p)$, $100(1-\alpha)\%$ Bonferroni intervals for μ_i $(i = 1, 2, \ldots, m)$ are given by

$$\left(\overline{X}_{in} - t_{n-1}\left(\frac{\alpha}{2m}\right)\sqrt{\frac{S_{ii}}{n}}, \quad \overline{X}_{in} + t_{n-1}\left(\frac{\alpha}{2m}\right)\sqrt{\frac{S_{ii}}{n}}\right).$$

Note:

Length of Bonferroni interval $= 2t_{n-1}\left(\frac{\alpha}{2m}\right)\sqrt{\frac{S_{ii}}{n}}$ and

Length of T^2 interval $= 2\sqrt{\frac{p(n-1)}{n-p}F_{p,n-p}(\alpha)}\sqrt{\frac{S_{ii}}{n}}$.

Thus, the ratio

$$\frac{\text{Length of Bonferroni interval}}{\text{Length of } T^2 \text{ interval}} = \frac{t_{n-1}\left(\frac{\alpha}{2m}\right)}{\sqrt{\frac{p(n-1)}{n-p}F_{p,n-p}(\alpha)}}$$

This does not dependent on random quantities, $\overline{\boldsymbol{X}}_n$ and $\mathcal{S}_n$ and using numerical computation we can show that this ratio is less than 1 for all $m \leq p$ and $p > 1$.

5 Large sample Theory

When the sample size is large, test of hypotheses and confidence intervals for $\boldsymbol{\mu}$ can be constructed without the normality assumption on the population.

Consider a random sample $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ from a p-variate population (not necessarily normal) with mean $\boldsymbol{\mu}$ and positive definite covariance matrix Σ. Further, n is large relative to p.

It can be proved that distribution of T^2, for this case, is approximately chi-square with p degrees of freedom. That is,

$$n\left(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}\right)^T \mathcal{S}_n^{-1}\left(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}\right) \quad \approx \quad \chi^2_p.$$

Note that when n is large relative to p

(a) $\mathcal{S}_n \approx \Sigma$ (using law of large numbers)

(b) and hence, the distribution of T^2 is approximately χ^2_p.

(c) Further, it can be proved that, as $n \to \infty$

$$\frac{p(n-p)}{n-p} F_{p,n-p} \to \chi^2_p$$

6 Hypotheses Testing for $\boldsymbol{\mu}$

Consider the hypotheses:

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 \quad \text{against} \quad H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0.$$

Since n is large relative to p, test statistic

$$T^2 = n\left(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}_0\right)^T \mathcal{S}_n^{-1}\left(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}_0\right) \quad \approx \quad \chi_p^2.$$

Hence reject H_0 if $T^2 > \chi_p^2(\alpha)$ where $\chi_p^2(\alpha)$ is the upper 100α percentile of the chi-square distribution with p degrees of freedom.

7 Confidence Region for $\boldsymbol{\mu}$

A $100(1-\alpha)\%$ confidence interval for $\boldsymbol{\mu}$ is

$$n\left(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}\right)^T \mathcal{S}_n^{-1}\left(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}\right) \leq \chi_p^2(\alpha)$$

Note that this is an ellipsoid and the axes of the ellipsoid can calculate using the eigenvalues and eigenvectors of $\mathcal{S}$. See Week 4 Lecture Note for details.

8 Simultaneous Confidence Intervals

A $100(1-\alpha)\%$ confidence interval for μ_i for $i = 1, 2, \ldots, p$ is given by

$$\left(\overline{X}_{in} - \sqrt{\chi_p^2(\alpha)}\,\sqrt{\frac{S_{ii}}{n}}, \quad \overline{X}_{in} + \sqrt{\chi_p^2(\alpha)}\,\sqrt{\frac{S_{ii}}{n}}\right).$$