Inference about a Mean Vector

[see Johnson, chapter 5, p. 210 ff.]

The Plausibility of μ_0 as a Value for a Normal Population Mean

Assume we have a random sample $X_1, ..., X_n$ from an $\mathcal{N}_p(\mu, \Sigma)$. In this section, we want to test whether a specific value μ_0 is a plausible value for the true population mean μ.

We test the hypotheses

$$H_0 : \mu = \mu_0 \quad \text{against} \mu \neq \mu_0.$$

An appropriate test statistic is the the so-called Hotelling's T^2-statistic

$$T^2 = (\overline{X} - \mu)^T \big(\frac{1}{n} S\big)^{-1} (\overline{X} - \mu) = n(\overline{X} - \mu) S^{-1} (\overline{X} - \mu) \sim \frac{(n-1)p}{n-p} F_{p,n-p} \text{ under } H_0.$$

From this, we can conclude the following decision rule: at the α level of significance, we reject H_0 in favor of H_1 if

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

Hotelling's T^2 and Likelihood Ratio Tests

Another reasonable approach is the idea of Likelihood Ratio. To determine wether μ_0 is a plausible value of μ, the maximum of the multivariate normal Likelihood $L(\mu_0, \Sigma)$ is compared with the unrestricted maximum of $L(\mu, \Sigma)$. The resulting ratio is called the likelihood ratio statistic Λ:

$$\Lambda = \frac{\max\limits_{\Sigma} L(\mu_0, \Sigma)}{\max\limits_{\mu, \Sigma} L(\mu, \Sigma)} = \left(\frac{|\hat{\Sigma}|}{|\hat{\Sigma_0}} \right)^{n/2},$$

where $\hat{\Sigma}$ is the ML-estimate in the unrestricted case and $\hat{\Sigma_0}$ is the ML-estimate under H_0.

The Likelihood ratio test and the test based on the Hotelling's T^2 statistic are equivalent.

Theorem (p. 218)**.** *Let $X_1, X_2, ..., X_n$ be a random sample from an $\mathcal{N}_p(\mu, \Sigma)$ population. Then the test based on the Hotelling's T^2 statistics is equivalent to the likelihood ratio test of $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$ because*

$$\Lambda^{2/n} = \left(1 + \frac{T^2}{n-1} \right)^{-1}$$

The following theorem generalizes the idea of the Likelihood-Ratio-Test for the test

$$H_0 : \theta \in \Theta_0 \text{ vs. } \theta \notin \Theta_0.$$

Theorem (p. 220)**.** *When the sample size n is large, under the null hypothesis H_0,*

$$-2 \ln \Lambda = -2 \ln \left(\frac{\max\limits_{\theta \in \Theta_0} L(\theta)}{\max\limits_{\theta \in \Theta} L(\theta)} \right) \overset{a}{\sim} \chi^2_{\nu - \nu_0},$$

where $\nu - \nu_0 = (\text{dimension of } \Theta) - (\text{dimension of } \Theta_0)$.

Confidence Regions and Simultaneous Comparisons of Component Means

The idea is now to extend the concept of a univariate confidence interval to a multivariate confidence region.

Definition (Confidence region). *Let $\theta \in \Theta$ be a vector of unknown population parameters and $X = [X_1, X_2, ..., X_n]$ our data matrix. $R(X)$ is said to be a $100(1-\alpha)\%$ confidence region if*

$$\mathbb{P}(R(X) \textit{ will cover the true } \theta) = 1 - \alpha.$$

Using that $n(\overline{X} - \mu)^T S^{-1}(\overline{X} - \mu) \sim \frac{(n-1)p}{n-p} F_{p,n-p}$ we get that the ellipsoid

$$\mu : n(\overline{x} - \mu)^T S^{-1}(\overline{x} - \mu) \sim \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)$$

gives us a $100(1-\alpha)$ confidence region for μ

Theorem (simultaneous confidence interval). *Let $X_1, X_2, ..., X_n$ be a random sample from an $\mathcal{N}_p(\mu, \Sigma)$ population with Σ positive definite. Then, simultaneously for all a, the intervall*

$$\left(a^T\overline{X} - \sqrt{\frac{p(n-1)}{n(n-p)} F_{p,n-p}(\alpha) a^T S a}, a^T\overline{X} + \sqrt{\frac{p(n-1)}{n(n-p)} F_{p,n-p}(\alpha) a^T S a} \right)$$

will contain $a^T\mu$ with probability $1 - \alpha$.

Large Sample Inferences about a Population Mean Vector

So far, we always built our confidence regions under the assumption that our sample was normal. Justified by the Central Limit theorem, if the mean and the covariance exist, we can assume that our sample is normal.

Theorem (Hypothesis testing for large samples, p. 235). *Let $X_1, X_2, ..., X_n$ be a random sample from a population with mean μ and positive definite covariance matrix Σ. When $n - p$ is large, the hypthesis $H_0 : \mu = \mu_0$ is rejected in favor of $H_1 : \mu \neq \mu_0$, at a level of significance approximately α, if the observed*

$$n(\overline{x} - \mu_0)^T S{-}1(\overline{x} - \mu_0) > \chi_p^2(\alpha).$$

Theorem (Confidence intervals for large samples, p. 235). *Let $X_1, X_2, ..., X_n$ be a random sample from a population with mean μ and positive definite covariance matrix Σ. If $n - p$ is large*

$$a^T\overline{X} \pm \sqrt{\chi_p^2(\alpha)} \sqrt{\frac{a^T S a}{n}}$$

will contain $a^T\mu$ for every a with probabyility approximately $1 - \alpha$.

The question of what is a large sample size is not easy to answer. In 1 or 2 dimensions, sample sizes in the range 30 to 50 can usually be considered large.