Inference about a mean vector and hypothesis testing

Woojoo Lee

Basic framework for hypothesis testing

Hypothesis testing for a normal mean:

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

Recalling the univariate case.

$$t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$$
$$\text{where } \bar{x} = \frac{1}{n}\sum_i x_i \text{ and } s^2 = \frac{1}{n-1}\sum_i (x_i - \bar{x})^2$$

→ t-statistic follows a student t-distribution with $n-1$ df.

→ reject H_0 if observed $|t|$ exceeds $t_{n-1}(\alpha/2)$.

→ The same conclusion will be followed if $t^2 \geq t^2_{n-1}(\alpha/2)$.

→ The $100(1-\alpha)\%$ confidence interval for μ is given by $\left\{\mu \middle| \left(\frac{\bar{x}-\mu}{s/\sqrt{n}}\right)^2 \leq t^2_{n-1}(\alpha/2)\right\}$

Hotelling's T^2

Hypothesis testing for a normal population mean vector :

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

$$t^2 = (\bar{x} - \mu_0)(s^2/n)^{-1}(\bar{x} - \mu_0) \rightarrow T^2 = (\bar{\mathbf{x}} - \mu_0)^T(S/n)^{-1}(\bar{\mathbf{x}} - \mu_0)$$

Gaussian- (scaled) chi-square - Gaussian

$\rightarrow$ reject H_0 if observed T^2 is sufficiently large.

Q) How large is large ?

$$T^2 \sim \frac{(n-1)p}{n-p} F_{p.n-p} \text{ (if } n > p)$$

$\to$ reject H_0 if observed T^2 exceeds $\frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)$
because $P(T^2 > \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)) = \alpha$.

Example: Table 5.1 (Sweat data)

```
> data
   Sweat rate Sodium Potassium
1         3.7   48.5       9.3
2         5.7   65.1       8.0
3         3.8   47.2      10.9
4         3.2   53.2      12.0
5         3.1   55.5       9.7
6         4.6   36.1       7.9
7         2.4   24.8      14.0
8         7.2   33.1       7.6
9         6.7   47.4       8.5
10        5.4   54.1      11.3
11        3.9   36.9      12.7
12        4.5   58.8      12.3
13        3.5   27.8       9.8
14        4.5   40.2       8.4
```

$H_0 : \mu_0 = (4, 50, 10)^T$ vs $H_1 :$ NotH_0

What is our main assumption ?

Invariant property of T^2: $\mathbf{y} = C\mathbf{x} + d$ where C is non-singular.

Compute T^2 by using $\mathbf{y}$.

Confidence regions

$$\text{confidence interval(univ)} \rightarrow \text{confidence region (multi)}$$

Note that the region $R(X)$ is said to be a $100(1-\alpha)\%$ confidence region if

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

For the mean of a p-dim'l Gaussian distribution,

$$P(n(\bar{\mathbf{x}}-\mu)^T S^{-1}(\bar{\mathbf{x}}-\mu) \leq \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)) = 1-\alpha$$

$\rightarrow$ A $100(1-\alpha)\%$ confidence region for the mean of a p-dim'l Gaussian distribution is the ellipsoid determined by all μ such that

$$n(\bar{\mathbf{x}}-\mu)^T S^{-1}(\bar{\mathbf{x}}-\mu) \leq \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)$$

Ex: Radiation data (Table 4.1 and 4.5)

```
      radation(close) radiation(open)
[1,]        0.6223330       0.7400828
[2,]        0.5477226       0.5477226
[3,]        0.6513556       0.7400828
[4,]        0.5623413       0.5623413
[5,]        0.4728708       0.5623413
[6,]        0.5885662       0.5885662
```

```
> mean
                     [,1]
radation(close) 0.5642575
radiation(open) 0.6029812
> cov
                radation(close) radiation(open)
radation(close)      0.01435023      0.01171547
radiation(open)      0.01171547      0.01454530
```

```
> res<-eigen(cov)
> res$values
[1] 0.026163638 0.002731895
> res$vectors
          [,1]       [,2]
[1,] 0.7041574 -0.7100439
[2,] 0.7100439  0.7041574
```

Simultaneous comparisons of component means

What if we are interested in μ_i or $\mu_i - \mu_j$?

In general, how can we make the confidence statement about $\mathbf{a}^T\mu$?

When $\mathbf{x}_i \sim N_p(\mu, \Sigma)$ and fix $\mathbf{a}$,

$$\frac{\sqrt{n}(\mathbf{a}^T\bar{\mathbf{x}} - \mathbf{a}^T\mu)}{\sqrt{\mathbf{a}^T S\mathbf{a}}} \sim t_{n-1}$$

Then, $100(1-\alpha)\%$ confidence interval for $\mathbf{a}^T\mu$ is given by

$$\left[\mathbf{a}^T\bar{\mathbf{x}} - t_{n-1}(\alpha/2)\frac{\sqrt{\mathbf{a}^T S\mathbf{a}}}{\sqrt{n}},\ \mathbf{a}^T\bar{\mathbf{x}} + t_{n-1}(\alpha/2)\frac{\sqrt{\mathbf{a}^T S\mathbf{a}}}{\sqrt{n}}\right].$$

Clearly, we could make several confidence statements about the components of μ.

$\rightarrow$ We may adopt the attitude that all of the separate confidence statements should hold simultaneously.

$\rightarrow$ A price must be paid for the "simultaneous" confidence.

Note that

$$\max_{\mathbf{a}} \frac{n(\mathbf{a}^T(\bar{\mathbf{x}}-\mu))^2}{\mathbf{a}^T S \mathbf{a}} = n(\overline{\mathbf{x}}-\mu)^T S^{-1}(\overline{\mathbf{x}}-\mu) = T^2$$

Result. Let $\mathbf{x}_i$ be a random sample from $N_p(\mu, \Sigma)$.
Then, simultaneously for all $\mathbf{a}$, the following interval

$$\mathbf{a}^T\bar{\mathbf{x}} \pm \sqrt{\frac{p(n-1)}{n(n-p)} F_{p,n-p}(\alpha)\mathbf{a}^T S \mathbf{a}}$$

will contain $\mathbf{a}^T\mu$ with prob $1-\alpha$.
For the proof, note that

$$n(\overline{\mathbf{x}}-\mu)^T S^{-1}(\overline{\mathbf{x}}-\mu) \leq c^2$$

is equivalent to

$$\frac{n(\mathbf{a}^T(\bar{\mathbf{x}}-\mu))^2}{\mathbf{a}^T S \mathbf{a}} \leq c^2 \text{ for every } \mathbf{a}.$$

Q) Construct the $100(1-\alpha)\%$ simultaneous confidence intervals for $\mu_1, \cdots, \mu_p$.

A comparison of simultaneous confidence intervals with one at a time intervals

Consider the confidence intervals of individual means.

One-at-a-time confidence intervals are

$$\bar{x_1} - t_{n-1}(\alpha/2)\sqrt{\frac{s_{11}}{n}} \leq \mu_1 \leq \bar{x_1} + t_{n-1}(\alpha/2)\sqrt{\frac{s_{11}}{n}}$$

$$\vdots$$

$$\bar{x_p} - t_{n-1}(\alpha/2)\sqrt{\frac{s_{pp}}{n}} \leq \mu_p \leq \bar{x_p} + t_{n-1}(\alpha/2)\sqrt{\frac{s_{pp}}{n}}$$

Q)If x_i are independent, what is P(All the one-at-a-time intervals contain μ_i) ? and discuss its implication in practice.

Bonferroni method of multiple comparisons

Let C_i denote a confidence statement about μ_i with

$$P(C_i \text{ is true}) = 1 - \alpha_i$$

$$\text{P(all } C_i \text{ true)=1-P(at least one } C_i \text{ false)}$$

$$\text{P(at least one } C_i \text{ false)} \leq \textstyle\sum_i P(C_i \text{ false})$$

$$=(1 - P(C_1\text{true}))+\cdots+(1 - P(C_m\text{true}))$$

Q) Is Bonferroni correction valid regardless of the correlation structure ?

Consider the following individual intervals

$$\bar{x}_i \pm t_{n-1}(\alpha/2m)\sqrt{\frac{s_{ii}}{n}}.$$

Then,

$$\begin{aligned} & P\left(\bar{x}_i \pm t_{n-1}(\alpha/2m)\sqrt{\frac{s_{ii}}{n}} \text{ contains } \mu_i \text{ for all } i\right) \\ = \quad & \text{Apply the Bonferroni method here !} \end{aligned}$$

Ex: Radiation data (Table 4.1 and 4.5)

It would be interesting to compare the length of Bonferroni interval(LB) with the length of T^2-interval (LT).

$$\frac{LB}{LT} = \frac{t_{n-1}(\alpha/2p)}{\sqrt{(n-1)pF_{p,n-p}(\alpha)/(n-p)}}.$$

Remark) See Table 5.4. This tells that the Bonferroni method provides shorter intervals when $m = p$.

Large sample inference on mean vector

When the sample size is large, tests of hypotheses and confidence regions for μ can be constructed without the Gaussian assumption.

Suppose that $\mathbf{x}_1, \cdots, \mathbf{x}_n$ be a random sample with mean μ and positive definite covariance matrix Σ.

All large-sample inferences about μ are based on a χ^2-distn (in this book).

$$P(n(\bar{\mathbf{x}} - \mu)^T S^{-1}(\bar{\mathbf{x}} - \mu) \leq \chi_p^2(\alpha)) \approx 1 - \alpha$$

Thus, $H_0 : \mu = \mu_0$ is rejected at a level of significance approximately α if

$$n(\bar{\mathbf{x}} - \mu)^T S^{-1}(\bar{\mathbf{x}} - \mu) > \chi_p^2(\alpha)$$

Remark) Compare $(n-1)pF_{p,n-p}(\alpha)/(n-p)$ and $\chi_p^2(\alpha)$.

Result. Let $\mathbf{x}_i$ be a random sample with mean μ and covariance Σ.
If $n >> p$, then

$$\mathbf{a}^T\bar{\mathbf{x}} \pm \sqrt{\chi_p^2(\alpha)}\sqrt{\frac{\mathbf{a}^T S\mathbf{a}}{n}}$$

will contain $\mathbf{a}^T\mu$, for every $\mathbf{a}$ with prob $1-\alpha$ approximately.

Remark) When n is large, compare the above result with
the Bonferroni simultaneous confidence intervals for individuals means.

Remark) It is a good practice to check the normality of the observations. Although small-to-moderate departures from normality do not cause substantial difficulties for large sample size, extreme deviations could cause problems.