Two-Sample Inferences

1 Review of univariate two-sample tests

Two *independent* random samples with the same population variance $\sigma_1^2 = \sigma_2^2 = \sigma^2$:

$$X_{11}, \ldots, X_{1n_1} \overset{i.i.d.}{\sim} \mathcal{N}(\mu_1, \sigma^2)$$

and

$$X_{21}, \ldots, X_{2n_2} \overset{i.i.d.}{\sim} \mathcal{N}(\mu_2, \sigma^2).$$

Let

$$x_{11}, \ldots, x_{1n_1}$$

and

$$x_{21}, \ldots, x_{2n_2}$$

be two observed samples with the summary statistics $\overline{x}_1$, $\overline{x}_2$, s_1^2 and s_2^2. The problem of two-sample test is to test the hypothesis $H_0 : \mu_1 = \mu_2$, or more generally, to find a confidence interval for $\mu_1 - \mu_2$.

The two random sample means obey the following sampling distributions

$$\overline{X}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} X_{1i} \sim \mathcal{N}\left(\mu_1, \frac{1}{n_1}\sigma^2\right), \quad \overline{X}_2 = \frac{1}{n_2}\sum_{i=1}^{n_2} X_{2i} \sim \mathcal{N}\left(\mu_2, \frac{1}{n_2}\sigma^2\right).$$

Since $\overline{X}_1$ and $\overline{X}_2$ are independent, we have

$$\overline{X}_1 - \overline{X}_2 \sim \mathcal{N}\left(\mu_1 - \mu_2, \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\sigma^2\right),$$

which implies

$$\frac{\left(\overline{X}_1 - \overline{X}_2\right) - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim \mathcal{N}(0, 1).$$

In practice, we need to estimate σ^2 by the two samples. Here we use the pooled sample variance to estimate the population variance

$$S_{pooled}^2 = \frac{n_1 - 1}{n_1 + n_2 - 2}S_1^2 + \frac{n_2 - 1}{n_1 + n_2 - 2}S_2^2.$$

Then we know S_{pooled}^2 is an unbiased estimator of σ^2. In fact,

$$\begin{aligned}\mathbb{E}\left(S_{pooled}^2\right) &= \frac{n_1 - 1}{n_1 + n_2 - 2}\,\mathbb{E}\left(S_1^2\right) + \frac{n_2 - 1}{n_1 + n_2 - 2}\,\mathbb{E}\left(S_2^2\right) \\ &= \frac{n_1 - 1}{n_1 + n_2 - 2}\sigma^2 + \frac{n_2 - 1}{n_1 + n_2 - 2}\sigma^2 = \sigma^2.\end{aligned}$$

There actually holds

$$\frac{\left(\overline{X}_1 - \overline{X}_2\right) - (\mu_1 - \mu_2)}{S_{pooled}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}.$$

Then

$$\mathbb{P}\left\{\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{S_{pooled}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\in[-t_{n_1+n_2-2}(0.025),t_{n_1+n_2-2}(0.025)]\right\}$$
$$=\mathbb{P}\left\{\mu_1-\mu_2\in\left[\overline{X}_1-\overline{X}_2-S_{pooled}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{n_1+n_2-2}(0.025),\overline{X}_1-\overline{X}_2+S_{pooled}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{n_1+n_2-2}(0.025)\right]\right\}$$
$$=0.95$$

Based on the two observed samples, the 95% confidence interval for $\mu_1-\mu_2$ is

$$\left[\overline{x}_1-\overline{x}_2-s_{pooled}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{n_1+n_2-2}(0.025),\overline{x}_1-\overline{x}_2+s_{pooled}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{n_1+n_2-2}(0.025)\right].$$

In particular, the null hypothesis $H_0:\mu_1-\mu_2=0$ is rejected at the level of α if

$$\left|\frac{\overline{x}_1-\overline{x}_2}{s_{pooled}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right|>t_{n_1+n_2-2}(\alpha/2),$$

since the type I error is

$$\mathbb{P}_{null}\left\{\left|\frac{\overline{X}_1-\overline{X}_2}{S_{pooled}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right|>t_{n_1+n_2-2}(\alpha/2)\right\}=\alpha.$$

2 Multivariate two-sample test by Hotelling's T^2

Question Two independent p-variate random samples with the same population covariance $\mathbf{\Sigma}_1=\mathbf{\Sigma}_2=\mathbf{\Sigma}$:

$$\vec{X}_1\sim\mathcal{N}_p(\vec{\mu}_1,\mathbf{\Sigma}_1):\vec{X}_{11},\ldots,\vec{X}_{1n_1},$$

and

$$\vec{X}_2\sim\mathcal{N}_p(\vec{\mu}_2,\mathbf{\Sigma}_2):\vec{X}_{21},\ldots,\vec{X}_{2n_2}.$$

Let

$$\vec{x}_{11},\ldots,\vec{x}_{1n_1}$$

and

$$\vec{x}_{21},\ldots,\vec{x}_{2n_2}$$

be two observed samples with the summary statistics $\overline{\vec{x}}_1$, $\overline{\vec{x}}_2$, $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$. Test $H_0:\vec{\mu}_1=\vec{\mu}_2$.

2.1 Hotelling's T^2

We have shown that

$$\overline{\vec{X}}_1\sim\mathcal{N}_p\left(\vec{\mu}_1,\frac{1}{n_1}\mathbf{\Sigma}_1\right),\quad\overline{\vec{X}}_2\sim\mathcal{N}_p\left(\vec{\mu}_2,\frac{1}{n_2}\mathbf{\Sigma}_2\right).$$

Since $\overline{\vec{X}}_1$ and $\overline{\vec{X}}_2$ are independent, we have

$$\overline{\vec{X}}_1-\overline{\vec{X}}_2\sim\mathcal{N}_p\left(\vec{\mu}_1-\vec{\mu}_2,\frac{1}{n_1}\mathbf{\Sigma}_1+\frac{1}{n_2}\mathbf{\Sigma}_2\right).$$

Moreover, since $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$, we have

$$\overline{\vec{X}}_1 - \overline{\vec{X}}_2 \sim \mathcal{N}_p \left(\vec{\mu}_1 - \vec{\mu}_2, \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \boldsymbol{\Sigma} \right).$$

This result implies

$$\left(\left(\overline{\vec{X}}_1 - \overline{\vec{X}}_2 \right) - (\vec{\mu}_1 - \vec{\mu}_2) \right)^\top \left(\left(\frac{1}{n_1} + \frac{1}{n_2} \right) \boldsymbol{\Sigma} \right)^{-1} \left(\left(\overline{\vec{X}}_1 - \overline{\vec{X}}_2 \right) - (\vec{\mu}_1 - \vec{\mu}_2) \right) \sim \chi^2_p.$$

In practice, we need to estimate $\boldsymbol{\Sigma}$ by the two samples. Here we use the pooled sample covariance to estimate the population covariance

$$\boldsymbol{S}_{pooled} = \frac{n_1 - 1}{n_1 + n_2 - 2} \boldsymbol{S}_1 + \frac{n_2 - 1}{n_1 + n_2 - 2} \boldsymbol{S}_2.$$

Then we know $\boldsymbol{S}_{pooled}$ is an unbiased estimator of $\boldsymbol{\Sigma}$. In fact,

$$\begin{aligned} \mathbb{E}\left(\boldsymbol{S}_{pooled}\right) &= \frac{n_1 - 1}{n_1 + n_2 - 2} \mathbb{E}\left(\boldsymbol{S}_1\right) + \frac{n_2 - 1}{n_1 + n_2 - 2} \mathbb{E}\left(\boldsymbol{S}_2\right) \\ &= \frac{n_1 - 1}{n_1 + n_2 - 2} \boldsymbol{\Sigma} + \frac{n_2 - 1}{n_1 + n_2 - 2} \boldsymbol{\Sigma} = \boldsymbol{\Sigma}. \end{aligned}$$

Furthermore, we have the following sampling distribution result:

$$\begin{aligned} &\left(\left(\overline{\vec{X}}_1 - \overline{\vec{X}}_2 \right) - (\vec{\mu}_1 - \vec{\mu}_2) \right)^\top \left(\left(\frac{1}{n_1} + \frac{1}{n_2} \right) \boldsymbol{S}_{pooled} \right)^{-1} \left(\left(\overline{\vec{X}}_1 - \overline{\vec{X}}_2 \right) - (\vec{\mu}_1 - \vec{\mu}_2) \right) \\ &\sim \frac{(n_1 + n_2 - 2)p}{n_1 + n_2 - 1 - p} F_{p, n_1 + n_2 - 1 - p}. \end{aligned}$$

This sampling distribution result implies the following Hotelling's T^2 for two-sample test $H_0 : \vec{\mu}_1 - \vec{\mu}_2 = \vec{\delta}_0$

$$T^2 = \left((\bar{\vec{x}}_1 - \bar{\vec{x}}_2) - \vec{\delta}_0 \right)^\top \left(\left(\frac{1}{n_1} + \frac{1}{n_2} \right) \boldsymbol{S}_{pooled} \right)^{-1} \left((\bar{\vec{x}}_1 - \bar{\vec{x}}_2) - \vec{\delta}_0 \right)$$

We reject $H_0 : \vec{\mu}_1 - \vec{\mu}_2 = \vec{\delta}_0$ at the level of α if

$$T^2 > \frac{(n_1 + n_2 - 2)p}{n_1 + n_2 - 1 - p} F_{p, n_1 + n_2 - 1 - p}(\alpha),$$

with the Type I error control

$$\mathbb{P}_{null} \left(T^2 > \frac{(n_1 + n_2 - 2)p}{n_1 + n_2 - 1 - p} F_{p, n_1 + n_2 - 1 - p}(\alpha) \right) = \alpha.$$

2.2 Confidence ellipse for the mean difference

The $100(1 - \alpha)\%$ confidence ellipse of $\vec{\delta} = \vec{\mu}_1 - \vec{\mu}_2$ is

$$\left((\bar{\vec{x}}_1 - \bar{\vec{x}}_2) - (\vec{\mu}_1 - \vec{\mu}_2) \right)^\top \left(\left(\frac{1}{n_1} + \frac{1}{n_2} \right) \boldsymbol{S}_{pooled} \right)^{-1} \left((\bar{\vec{x}}_1 - \bar{\vec{x}}_2) - (\vec{\mu}_1 - \vec{\mu}_2) \right) \leq \frac{(n_1 + n_2 - 2)p}{n_1 + n_2 - 1 - p} F_{p, n_1 + n_2 - 1 - p}(\alpha).$$

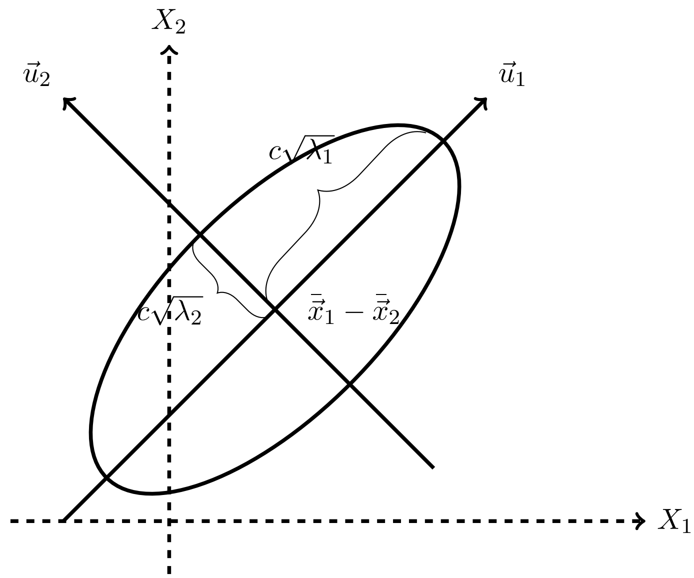
When $n_1 \gg p$ and $n_2 \gg p$, the $100(1 - \alpha)\%$ confidence ellipse of $\vec{\delta} = \vec{\mu}_1 - \vec{\mu}_2$ is

$$\left((\bar{\vec{x}}_1 - \bar{\vec{x}}_2) - (\vec{\mu}_1 - \vec{\mu}_2) \right)^\top \left(\left(\frac{1}{n_1} + \frac{1}{n_2} \right) \boldsymbol{S}_{pooled} \right)^{-1} \left((\bar{\vec{x}}_1 - \bar{\vec{x}}_2) - (\vec{\mu}_1 - \vec{\mu}_2) \right) \leq \chi^2_p(\alpha).$$

In the case $p = 2$, by finding the spectral decomposition

$$\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \boldsymbol{S}_{pooled} = \lambda_1 \vec{u}_1 \vec{u}_1^\top + \lambda_2 \vec{u}_2 \vec{u}_2^\top,$$

and denoting $c = \sqrt{\frac{(n_1+n_2-2)p}{n_1+n_2-1-p} F_{p,n_1+n_2-1-p}(\alpha)}$, the $100(1-\alpha)\%$ confidence ellipse of $\vec{\delta} = \vec{\mu}_1 - \vec{\mu}_2$ is plotted as follows:



2.3 Equivalence under invertible linear transformations

Recall the problem of two-sample test: Two independent p-variate random samples with the same population covariance $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$:

$$\vec{X}_1 \sim \mathcal{N}_p(\vec{\mu}_1, \boldsymbol{\Sigma}_1) : \vec{X}_{11}, \ldots, \vec{X}_{1n_1},$$

and

$$\vec{X}_2 \sim \mathcal{N}_p(\vec{\mu}_2, \boldsymbol{\Sigma}_2) : \vec{X}_{21}, \ldots, \vec{X}_{2n_2}.$$

Let

$$\vec{x}_{11}, \ldots, \vec{x}_{1n_1}$$

and

$$\vec{x}_{21}, \ldots, \vec{x}_{2n_2}$$

be two observed samples with the summary statistics $\bar{\vec{x}}_1$, $\bar{\vec{x}}_2$, $\boldsymbol{S}_1^x$ and $\boldsymbol{S}_2^x$. Test $H_0 : \vec{\mu}_1 = \vec{\mu}_2$.

Suppose $\boldsymbol{C}$ is a $p \times p$ invertible matrix and $\vec{d}$ is a $p \times 1$ vector. Consider the transformation $\vec{Y}_1 = \boldsymbol{C}\vec{X}_1 + \vec{d}$ and $\vec{Y}_2 = \boldsymbol{C}\vec{X}_2 + \vec{d}$. Then we have

$$\vec{Y}_1 \sim \mathcal{N}(\boldsymbol{C}\vec{\mu}_1 + \vec{d}, \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}^\top), \quad \vec{Y}_2 \sim \mathcal{N}(\boldsymbol{C}\vec{\mu}_2 + \vec{d}, \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}^\top),$$

with independent samples $\vec{y}_{lj} = \boldsymbol{C}\vec{x}_{lj} + \vec{d}$ for all $l = 1, 2$ and $j = 1, 2, \ldots, n_l$.

Clearly, we have the equivalence of two-sample tests:

$$H_0 : \vec{\mu}_1 = \vec{\mu}_2 \Longleftrightarrow H_0 : \boldsymbol{C}\vec{\mu}_1 + \vec{d} = \boldsymbol{C}\vec{\mu}_2 + \vec{d}.$$

Now the question is whether Hotelling's T^2 gives us the same result before and after linear transformation.

The T^2 applied to the two original samples is denoted as

$$T_x^2 = \left(\left(\overline{\vec{x}}_1 - \overline{\vec{x}}_2\right)\right)^\top \left(\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \boldsymbol{S}^x_{pooled}\right)^{-1} \left(\left(\overline{\vec{x}}_1 - \overline{\vec{x}}_2\right)\right)$$

where

$$\boldsymbol{S}^x_{pooled} = \frac{n_1 - 1}{n_1 + n_2 - 2}\boldsymbol{S}^x_1 + \frac{n_2 - 1}{n_1 + n_2 - 2}\boldsymbol{S}^x_2.$$

Similarly, the T^2 applied to the two samples after transformation is denoted as

$$T_y^2 = \left(\left(\overline{\vec{y}}_1 - \overline{\vec{y}}_2\right)\right)^\top \left(\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \boldsymbol{S}^y_{pooled}\right)^{-1} \left(\left(\overline{\vec{y}}_1 - \overline{\vec{y}}_2\right)\right)$$

where

$$\boldsymbol{S}^y_{pooled} = \frac{n_1 - 1}{n_1 + n_2 - 2}\boldsymbol{S}^y_1 + \frac{n_2 - 1}{n_1 + n_2 - 2}\boldsymbol{S}^y_2.$$

The linear transformations give

$$\overline{\vec{y}}_1 = \boldsymbol{C}\overline{\vec{x}}_1, \quad \overline{\vec{y}}_2 = \boldsymbol{C}\overline{\vec{x}}_2, \quad \boldsymbol{S}^y_1 = \boldsymbol{C}\boldsymbol{S}^x_1\boldsymbol{C}^\top, \quad \boldsymbol{S}^y_2 = \boldsymbol{C}\boldsymbol{S}^x_2\boldsymbol{C}^\top.$$

Then

$$\begin{aligned}
\boldsymbol{S}^y_{pooled} &= \frac{n_1 - 1}{n_1 + n_2 - 2}\boldsymbol{S}^y_1 + \frac{n_2 - 1}{n_1 + n_2 - 2}\boldsymbol{S}^y_2. \\
&= \frac{n_1 - 1}{n_1 + n_2 - 2}\boldsymbol{C}\boldsymbol{S}^x_1\boldsymbol{C}^\top + \frac{n_2 - 1}{n_1 + n_2 - 2}\boldsymbol{C}\boldsymbol{S}^x_2\boldsymbol{C}^\top \\
&= \boldsymbol{C}\left(\frac{n_1 - 1}{n_1 + n_2 - 2}\boldsymbol{S}^x_1 + \frac{n_2 - 1}{n_1 + n_2 - 2}\boldsymbol{S}^x_2\right)\boldsymbol{C}^\top.
\end{aligned}$$

Then we can establish the equivalence between the Hotelling's T^2's:

$$\begin{aligned}
T_y^2 &= \left(\left(\overline{\vec{y}}_1 - \overline{\vec{y}}_2\right)\right)^\top \left(\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \boldsymbol{S}^y_{pooled}\right)^{-1} \left(\left(\overline{\vec{y}}_1 - \overline{\vec{y}}_2\right)\right) \\
&= \left(\boldsymbol{C}\left(\overline{\vec{x}}_1 - \overline{\vec{x}}_2\right)\right)^\top \left(\boldsymbol{C}\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \boldsymbol{S}^x_{pooled}\boldsymbol{C}^\top\right)^{-1} \left(\boldsymbol{C}\left(\overline{\vec{x}}_1 - \overline{\vec{x}}_2\right)\right) \\
&= \left(\left(\overline{\vec{x}}_1 - \overline{\vec{x}}_2\right)\right)^\top \boldsymbol{C}^\top(\boldsymbol{C}^\top)^{-1} \left(\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \boldsymbol{S}^x_{pooled}\right)^{-1} \boldsymbol{C}^{-1}\boldsymbol{C}\left(\left(\overline{\vec{x}}_1 - \overline{\vec{x}}_2\right)\right) \qquad (\because \boldsymbol{C} \text{ is invertible}) \\
&= \left(\left(\overline{\vec{x}}_1 - \overline{\vec{x}}_2\right)\right)^\top \left(\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \boldsymbol{S}^x_{pooled}\right)^{-1} \left(\left(\overline{\vec{x}}_1 - \overline{\vec{x}}_2\right)\right) \\
&= T_x^2.
\end{aligned}$$

3 Simultaneous confidence intervals

3.1 Bonferroni correction

Two random samples:

$$\vec{X}_{11}, \ldots, \vec{X}_{1n_1} \overset{i.i.d.}{\sim} \mathcal{N}_p(\vec{\mu}_1, \boldsymbol{\Sigma}_1)$$

and

$$\vec{X}_{21}, \ldots, \vec{X}_{2n_2} \overset{i.i.d.}{\sim} \mathcal{N}_p(\vec{\mu}_2, \boldsymbol{\Sigma}_2).$$

For $j = 1, \ldots, p$, let's just investigate the two samples on the j-th variate:

$$X_{11j}, \ldots, X_{1n_1j} \overset{i.i.d.}{\sim} \mathcal{N}(\mu_{1j}, \sigma^2_{1j})$$

and

$$X_{21j}, \ldots, X_{2n_2j} \overset{i.i.d.}{\sim} \mathcal{N}(\mu_{2j}, \sigma_{2j}^2)$$

Assume for each $j = 1, \ldots, p$, the variances of the j-th variate in the two samples are equal:

$$\sigma_{1j}^2 = \sigma_{2j}^2 := \sigma_j^2.$$

The summary statistics are

$$\overline{X}_{1j}, \quad S_{1j}^2, \quad \overline{X}_{2j}, \quad S_{2j}^2.$$

For an observation of these two samples, the summary statistics are

$$\overline{x}_{1j}, \quad s_{1j}^2, \quad \overline{x}_{2j}, \quad s_{2j}^2.$$

Question: How to test $H_0 : \vec{\mu}_1 - \vec{\mu}_2 = \vec{0}$? How to find simultaneous intervals for μ_j?

Recall that we have the following sampling distributions

$$\overline{X}_{1j} \sim \mathcal{N}\left(\mu_{1j}, \frac{1}{n}\sigma_{1j}^2\right), \quad \overline{X}_{2j} \sim \mathcal{N}\left(\mu_{2j}, \frac{1}{n}\sigma_{2j}^2\right), \ j = 1, \ldots, p.$$

By the assumption of equal variances: $\sigma_{1j}^2 = \sigma_{2j}^2 = \sigma_j^2$, $j = 1, \ldots, p$ and the independence of the two samples, we have

$$\overline{X}_{1j} - \overline{X}_{2j} \sim \mathcal{N}\left(\mu_{1j} - \mu_{2j}, \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\sigma_j^2\right),$$

which implies

$$\frac{(\overline{X}_{1j} - \overline{X}_{2j}) - (\mu_{1j} - \mu_{2j})}{\sigma_j\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim \mathcal{N}(0, 1).$$

Since σ_j is unknown, we need to estimate it from the data. Here we use the pooled sample variance to estimate the population variance

$$S_{pooled,j}^2 = \frac{n_1 - 1}{n_1 + n_2 - 2}S_{1j}^2 + \frac{n_2 - 1}{n_1 + n_2 - 2}S_{2j}^2.$$

By replacing σ_j with $S_{pooled,j}$, we have the following sampling distribution result

$$\frac{(\overline{X}_{1j} - \overline{X}_{2j}) - (\mu_{1j} - \mu_{2j})}{S_{pooled,j}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}.$$

Then we have the following covering probability

$$\mathbb{P}(A_j) = 1 - \alpha/p.$$

where

$$A_j := \left\{\overline{X}_{1j} - \overline{X}_{2j} - S_{pooled,j}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}t_{n_1+n_2-2}\left(\frac{\alpha}{2p}\right) \leq \mu_{1j} - \mu_{2j} \leq \overline{X}_{1j} - \overline{X}_{2j} + S_{pooled,j}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}t_{n_1+n_2-2}\left(\frac{\alpha}{2p}\right)\right\}.$$

With this fact, we actually get simultaneous confidence intervals:

$$\mathbb{P}(A_1 \cap \ldots \cap A_p) = 1 - \mathbb{P}((A_1 \cap \ldots \cap A_p)^c) = 1 - \mathbb{P}(A_1^c \cup \ldots \cup A_p^c) \geq 1 - \sum_{j=1}^{p}\mathbb{P}(A_j^c) = 1 - \alpha.$$

Then we have the $(1 - \alpha)$ simultaneous confidence intervals

$$\overline{x}_{1j} - \overline{x}_{2j} - s_{pooled,j}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}t_{n_1+n_2-2}\left(\frac{\alpha}{2p}\right) \leq \mu_{1j} - \mu_{2j} \leq \overline{x}_{1j} - \overline{x}_{2j} + s_{pooled,j}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}t_{n_1+n_2-2}\left(\frac{\alpha}{2p}\right).$$

Moreover, Bonferroni correction rejects H_0 at level α, if

$$\max_{1\leq j\leq p}\left|\frac{\overline{x}_{1j} - \overline{x}_{2j}}{s_{pooled,j}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\right| \geq t_{n_1+n_2-2}\left(\frac{\alpha}{2p}\right),$$

3.2 Simultaneous confidence intervals from confidence region

In order to derive simultaneous confidence intervals from the confidence region, we need the following inequalities:

- Cauchy-Schwarz inequality
$$(\vec{x}^\top \vec{d})^2 \le (\vec{x}^\top \vec{x})(\vec{d}^\top \vec{d}).$$

- Extended Cauchy-Schwarz inequality: For positive definite $\boldsymbol{B}$, and two vectors $\vec{x}$ and $\vec{d}$, then
$$(\vec{x}^\top \vec{d})^2 \le (\vec{x}^\top \boldsymbol{B}\vec{x})(\vec{d}^\top \boldsymbol{B}^{-1}\vec{d}).$$

  Proof.
$$(\vec{x}^\top \vec{d})^2 = (\vec{x}^\top \boldsymbol{B}^{\frac{1}{2}}\boldsymbol{B}^{-\frac{1}{2}}\vec{d})^2 = \left((\boldsymbol{B}^{\frac{1}{2}}\vec{x})^\top(\boldsymbol{B}^{-\frac{1}{2}}\vec{d})\right)^2 \le (\boldsymbol{B}^{\frac{1}{2}}\vec{x})^\top(\boldsymbol{B}^{\frac{1}{2}}\vec{x})(\boldsymbol{B}^{-\frac{1}{2}}\vec{d})^\top(\boldsymbol{B}^{-\frac{1}{2}}\vec{d}) = (\vec{x}^\top \boldsymbol{B}\vec{x})(\vec{d}^\top \boldsymbol{B}^{-1}\vec{d}).$$
  □

- Maximization lemma: Let $\boldsymbol{B}$ be positive definite and $\vec{d}$ be a given vector. Then for an arbitrary nonzero vector $\vec{x}$,
$$\max_{\vec{x}\neq\vec{0}} \frac{(\vec{x}^\top \vec{d})^2}{\vec{x}^\top \boldsymbol{B}\vec{x}} = \vec{d}^\top \boldsymbol{B}^{-1}\vec{d},$$
  with the maximum attained when $\vec{x} = c\boldsymbol{B}^{-1}\vec{d}$ for $c \neq 0$.

  Proof. By the extended Cauchy-Schwarz inequality,
$$(\vec{x}^\top \vec{d})^2 \le (\vec{x}^\top \boldsymbol{B}\vec{x})(\vec{d}^\top \boldsymbol{B}^{-1}\vec{d}).$$
  Since $\vec{x} \neq \vec{0}$ and $\boldsymbol{B}$ is positive definite, we have $\vec{x}^\top \boldsymbol{B}\vec{x} > 0$, which implies
$$\frac{(\vec{x}^\top \vec{d})^2}{\vec{x}^\top \boldsymbol{B}\vec{x}} \le \vec{d}^\top \boldsymbol{B}^{-1}\vec{d}.$$
  It is easy to verify that the bound is attained for $\vec{x} = c\boldsymbol{B}^{-1}\vec{d}$. □

Now we are ready to introduce how to derive simultaneous intervals from the confidence region. Recall that the $(1-\alpha)$ confidence region is

$$\left((\bar{\vec{x}}_1 - \bar{\vec{x}}_2) - (\vec{\mu}_1 - \vec{\mu}_2)\right)^\top \left(\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\boldsymbol{S}_{pooled}\right)^{-1} \left((\bar{\vec{x}}_1 - \bar{\vec{x}}_2) - (\vec{\mu}_1 - \vec{\mu}_2)\right) \le \frac{(n_1+n_2-2)p}{n_1+n_2-1-p} F_{p,n_1+n_2-1-p}(\alpha).$$

By the maximization lemma,

$$\max_{\vec{a}\neq\vec{0}} \frac{\left(\vec{a}^\top \left((\bar{\vec{x}}_1 - \bar{\vec{x}}_2) - (\vec{\mu}_1 - \vec{\mu}_2)\right)\right)^2}{\vec{a}^\top \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\boldsymbol{S}_{pooled}\vec{a}} = \left((\bar{\vec{x}}_1 - \bar{\vec{x}}_2) - (\vec{\mu}_1 - \vec{\mu}_2)\right)^\top \left(\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\boldsymbol{S}_{pooled}\right)^{-1} \left((\bar{\vec{x}}_1 - \bar{\vec{x}}_2) - (\vec{\mu}_1 - \vec{\mu}_2)\right).$$

Therefore, when $\vec{\mu}_1 - \vec{\mu}_2$ lies in the confidence region, for all $\vec{a}$,

$$\frac{\left(\vec{a}^\top \left((\bar{\vec{x}}_1 - \bar{\vec{x}}_2) - (\vec{\mu}_1 - \vec{\mu}_2)\right)\right)^2}{\vec{a}^\top \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\boldsymbol{S}_{pooled}\vec{a}} \le \frac{(n_1+n_2-2)p}{n_1+n_2-1-p} F_{p,n_1+n_2-1-p}(\alpha),$$

which is equivalent to

$$\begin{aligned}
&\vec{a}^\top \left(\overline{\vec{x}}_1 - \overline{\vec{x}}_2\right) - \sqrt{\frac{(n_1+n_2-2)p}{n_1+n_2-1-p} F_{p,n_1+n_2-1-p}(\alpha) \vec{a}^\top \left(\frac{1}{n_1} + \frac{1}{n_2}\right) \boldsymbol{S}_{pooled} \vec{a}} \\
&\leq \vec{a}^\top (\vec{\mu}_1 - \vec{\mu}_2) \\
&\leq \vec{a}^\top \left(\overline{\vec{x}}_1 - \overline{\vec{x}}_2\right) + \sqrt{\frac{(n_1+n_2-2)p}{n_1+n_2-1-p} F_{p,n_1+n_2-1-p}(\alpha) \vec{a}^\top \left(\frac{1}{n_1} + \frac{1}{n_2}\right) \boldsymbol{S}_{pooled} \vec{a}}.
\end{aligned}$$

In particular, by choosing $\vec{a} = [1, 0, \ldots, 0]^\top, [0, 1, 0, \ldots, 0], \ldots, [0, \ldots, 0, 1]$, we have the simultaneous intervals

$$\begin{aligned}
&\overline{x}_{1j} - \overline{x}_{2j} - s_{pooled,j} \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \frac{(n_1+n_2-2)p}{n_1+n_2-1-p} F_{p,n_1+n_2-1-p}(\alpha)} \\
&\leq \mu_{1j} - \mu_{2j} \\
&\leq \overline{x}_{1j} - \overline{x}_{2j} + s_{pooled,j} \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \frac{(n_1+n_2-2)p}{n_1+n_2-1-p} F_{p,n_1+n_2-1-p}(\alpha)}.
\end{aligned}$$