stats545.2 Multivariate confidence intervals

Gunnar Stefansson

2. desember 2016

Copyright This work is licensed under the Creative Commons Attribution-ShareAlike License. To view a copy of this license, visit http://creativecommons.org/licenses/by-sa/1.0/ or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

Efnisyfirlit

1 Tests of hypotheses including multiple comparisons in the linear model

1.1 On distributions

$\underbrace{\mathbf{y}}_{n\times 1} \sim n(\underbrace{\mathbf{X}}_{n\times p}\underbrace{\beta}_{p\times 1}, \sigma^2 \underbrace{\mathbf{I}}_{n\times n})$

Theorem: $\hat{\psi} \sim n\left(\psi, \Sigma_{\hat{\psi}}\right)$, $\frac{||\mathbf{y}-\mathbf{X}\hat{\beta}||^2}{\sigma^2} \sim \chi^2_{n-r}$ and these two quantities are independent.

1.1.1 Details

Let $\underbrace{\mathbf{y}}_{n\times 1} \sim n(\underbrace{\mathbf{X}}_{n\times p}\underbrace{\beta}_{p\times 1}, \sigma^2 \underbrace{\mathbf{I}}_{n\times n})$
and assume $rank(\mathbf{X}) = r \leq p$.
The interest will be in obtaining some joint confidence statement on a vector, $\psi = (\psi_1, \ldots, \psi_q)$, where each $\psi_i = \mathbf{c}_i'\beta$ is an estimable function. Write $\hat{\psi} = (\hat{\psi}_1, \ldots, \hat{\psi}_q)$ for the least squares estimates with $\hat{\psi}_i = \mathbf{c}_i\hat{\beta}$ where $\hat{\beta}$ is any LS estimate and one can therefore also write $\hat{\psi}_i = \mathbf{a}_i\mathbf{y}$ for unique $a_i \in sp(\mathbf{X})$.
The above can be written more concisely as $\psi = \mathbf{C}\beta$ using obvious definitions. It follows that

$$\hat{\psi} = \mathbf{A}\mathbf{y} = \mathbf{C}\hat{\beta} \sim n(\mathbf{C}\beta, \sigma^2\mathbf{A}\mathbf{A}')$$

and the variance-covariance matrix of the estimates will be denoted

$$V[\hat{\psi}] = \Sigma_{\hat{\psi}}$$

which leads to the following theorem.

Theorem 1.1. $\hat{\psi} \sim n\left(\psi, \Sigma_{\hat{\psi}}\right)$, $\frac{||\mathbf{y}-\mathbf{X}\hat{\beta}||^2}{\sigma^2} \sim \chi^2_{n-r}$ and these two quantities are independent.

1.1.2 Handout

Proof: Let $\{\xi_1, \ldots, \xi_n\}$ be an orthonormal basis for $\mathbb{R}^n$ such that $\{\xi_1, \ldots, \xi_r\}$ form a basis for $sp(\mathbf{X})$ and let $\hat{\zeta}_1, \ldots, \hat{\zeta}_n$ be the coordinates of $\mathbf{y}$ in this basis, so that $\hat{\zeta}_i = \xi_i \cdot \mathbf{y}$. Also define $\zeta_i = E[\hat{\zeta}_i]$. It is established that $\hat{\zeta}_i$ are independent, Gaussian with common variance σ^2.
Write $\mathbf{z} = \left(\hat{\zeta}_1, \ldots, \hat{\zeta}_n\right)'$, $\mathbf{P} = [\xi_1, \ldots, \xi_n]'$ and note that $\mathbf{P}' = [\xi_1 \ldots \xi_n]$. It is then clear that the rows of $\mathbf{P}'$ are independent so $\mathbf{P}'$ is invertible (as is $\mathbf{P}$). Clearly, $\mathbf{P}\mathbf{P}' = \mathbf{I}$ so $\mathbf{P}'\mathbf{P} = \mathbf{I}$. Further, $\mathbf{z} = \mathbf{P}\mathbf{y}$ and therefore $\mathbf{y} = \mathbf{P}'\mathbf{z}$.
As elsewhere, write the LS estimates of the estimable functions in the form $\hat{\psi}_i = \mathbf{a}_i'\mathbf{y}$ where $\mathbf{a}_i \in V = sp\{\xi_1, \ldots, \xi_r\}$ so that $\hat{\psi}_i = \mathbf{a}_i'\mathbf{P}'\mathbf{z}$. It follows that $\mathbf{a}_i'\mathbf{P}' = \left[\mathbf{a}_i'\xi_1 \vdots \ldots \vdots \mathbf{a}_i'\xi_n\right]$ and of these various inner products, $\mathbf{a}_i'\xi_j = 0$ if $j > r$ (since $\mathbf{a}_i \in V$) from which it is seen that

$$\mathbf{a}_i'\mathbf{P}'\mathbf{z} = \left[\mathbf{a}_i'\xi_1 \vdots \ldots \vdots \mathbf{a}_i'\xi_r \vdots 0 \ldots 0\right]\left[\hat{\zeta}_1, \ldots, \hat{\zeta}_r, \hat{\zeta}_{r+1}, \ldots, \hat{\zeta}_n\right]' = \mathbf{a}_i'\xi_1\hat{\zeta}_1 + \ldots + \mathbf{a}_i'\xi_r\hat{\zeta}_r$$

i.e. the estimable functions are all formed from the first r of the $\hat{\zeta}_i$ and are all of the form

$$\hat{\psi}_i = \sum_1^r k_j \hat{\zeta}_j \tag{1}$$

for some constants $k_1, \ldots, k_r$.
This important result is quite general and basically states that anything that can be estimated can be derived from $\mathbf{y}$ through the column vectors of the $\mathbf{X}$-matrix.
On the other hand it is also known that $\mathbf{X}\hat{\beta}$ is the projection of $\mathbf{y}$ onto the space spanned by $\xi_1, \ldots, \xi_r$ and therefore the residual, $\mathbf{y} - \mathbf{X}\hat{\beta}$ is in the span of $\xi_{r+1}, \ldots, \xi_n$ and in fact

$$||\mathbf{y} - \mathbf{X}\hat{\beta}||^2 = \sum_{j=r+1}^{n} \hat{\zeta}_j^2 \tag{2}$$

All the results in the theorem follow easily from (1) and (2).

1.2 Confidence ellipsoids

$$\boxed{P_\beta \left[(\hat{\psi} - \psi)' \mathbf{B}^{-1} (\hat{\psi} - \psi) \leq q s^2 F_{q,n-r,1-\alpha} \right] = 1 - \alpha}$$

1.2.1 Details

Theorem 1.2. Under the above assumptions and definitions,

$$\frac{(\hat{\psi} - \psi)' \mathbf{B}^{-1} (\hat{\psi} - \psi) / q}{||\mathbf{y} - \mathbf{X}\hat{\beta}||^2 / (n - r)} \sim F_{q,n-r}$$

Noting that the denominator is the usual estimator, s^2 of σ^2, it follows that the following probability statement holds and can be used to obtain a confidence ellipsoid for ψ.

$$P_\beta \left[(\hat{\psi} - \psi)' \mathbf{B}^{-1} (\hat{\psi} - \psi) \leq q s^2 F_{q,n-r,1-\alpha} \right] = 1 - \alpha$$

These intervals are very general and lead to several important special cases.

1.2.2 Handout

It is of interest to derive confidence regions, $R(\mathbf{y}) \subseteq \mathbb{R}^n$ such that

$$P_\beta [\psi \in R(\mathbf{y})] = 1 - \alpha \quad \forall \beta \in \mathbb{R}^p.$$

Assume (without loss of generality) that $rank(C) = q$ and note that $q \leq p$.
Now, $\psi = \mathbf{C}\beta \in \mathbb{R}^q$ and the estimates can be written $\hat{\psi} = \mathbf{A}\mathbf{y}$ for an appropriate choice of $\mathbf{A}$ so $E\hat{\psi} = \psi$ and $V\hat{\psi} = \sigma^2 \mathbf{B}$ with $\mathbf{B} = \mathbf{A}\mathbf{A}'$. Next note that

$$\mathbf{C}\beta = \psi = E\hat{\psi} = \mathbf{A}\mathbf{X}\beta \quad \forall \beta$$

so that $\mathbf{C} = \mathbf{A}\mathbf{X}$ and hence $q = rank(\mathbf{C}) = rank(\mathbf{A}\mathbf{X}) \leq rank(\mathbf{A}) \leq q$ where the last inequality follows from $\mathbf{A}$ being a $q \times n$ matrix. But this implies that $rank(\mathbf{A}) = q$ and it is a know result from linear algebra that $rank(\mathbf{B}) = rank(\mathbf{A})$. Since $\mathbf{B}$ is a $q \times q$ matrix, it follows that $\mathbf{B}$ is nonsingular.

Hence

$$\hat{\psi} \sim n\left(\psi, \sigma^2\mathbf{B}\right).$$

Now, for any ν-dimensional multivariate normal random vector $\mathbf{Z}$ with variance-covariance matrix $\Sigma_{\mathbf{Z}}$ and mean vector $\mu_{\mathbf{Z}}$, it will be considered known that

$$(\mathbf{Z}-\mu_{\mathbf{Z}})'\Sigma_{\mathbf{Z}}^{-1}(\mathbf{Z}-\mu_{\mathbf{Z}}) \sim \chi^2_{\nu}.$$

This result easily follows from decomposing $\Sigma_{\mathbf{Z}}^{-1}$ into $\mathbf{LL}'$ where $\mathbf{L}$ is a lower triangular matrix and defining $\mathbf{U} = \mathbf{L}(\mathbf{Z}-\mu_{\mathbf{Z}})$. Then the components of $\mathbf{U}$ will be i.i.d. $n(0,1)$.
It is therefore seen that

$$(\hat{\psi}-\psi)'\left(\sigma^2\mathbf{B}\right)^{-1}(\hat{\psi}-\psi) \sim \chi^2_q. \tag{3}$$

From above we know that this is independent of $\frac{||\mathbf{y}-\mathbf{X}\hat{\beta}||^2}{\sigma^2} \sim \chi^2_{n-r}$
from which we obtain the above theorem.

1.3 Confidence interval for a single estimable function

For a single $\psi = \mathbf{c}'\beta\mathbf{a}'\mathbf{y}$,

$$\hat{\sigma}^2_{\hat{\psi}} = \mathbf{a}'\mathbf{a}s^2$$

and
A confidence interval for ψ: can be based on

$$(\hat{\psi}-\psi)^2 \leq \mathbf{a}'\mathbf{a}s^2F_{1,n-r,1-\alpha}$$

or on

$$P\left[\psi \in \left[\hat{\psi}-t_{n-r,1-\alpha/2}\sqrt{\mathbf{a}'\mathbf{a}}s, \hat{\psi}+t_{n-r,1-\alpha/2}\sqrt{\mathbf{a}'\mathbf{a}}s\right]\right] = 1-\alpha$$

1.3.1 Details

Consider a single ($q=1$) confidence interval for a general estimable function. Write $\psi = \mathbf{c}'\beta$ and note that $rank(\mathbf{c}) = 1$ if $\mathbf{c} \neq \mathbf{0}$. Our estimator for ψ is $\hat{\psi} = \mathbf{c}'\hat{\beta}$ and can be written $\psi = \mathbf{a}'\mathbf{y}$ for an appropriate $\mathbf{a}$.
It follows that the variance of $\hat{\psi}$ is

$$\hat{\sigma}^2_{\hat{\psi}} = \mathbf{a}'\mathbf{a}s^2$$

and a confidence interval for ψ can be based on

$$(\hat{\psi}-\psi)^2 \leq \mathbf{a}'\mathbf{a}s^2F_{1,n-r,1-\alpha}$$

or on the following corresponding probability statement:

$$P\left[\psi \in \left[\hat{\psi}-t_{n-r,1-\alpha/2}\sqrt{\mathbf{a}'\mathbf{a}}s, \hat{\psi}+t_{n-r,1-\alpha/2}\sqrt{\mathbf{a}'\mathbf{a}}s\right]\right] = 1-\alpha$$

1.3.2 Examples

When it comes to computing a confidence interval for a single estimable function, we have seen that we can simply compute the values using an interval of the form

$$\left[\hat{\psi}-t_{n-r,1-\alpha/2}\sqrt{\mathbf{a}'\mathbf{a}}s, \hat{\psi}+t_{n-r,1-\alpha/2}\sqrt{\mathbf{a}'\mathbf{a}}s\right].$$

There are a few tricks to this.

First of all, since $\mathbf{a}'\mathbf{y} = \mathbf{c}'\beta$, the variance can be obtained either from $V[\hat{\psi}] = \mathbf{a}'\mathbf{a}\sigma$ as is done above, or by using the alternative formulation

$$V[\beta] = \Sigma_{\hat{\beta}} = \sigma^2 \left(X'X\right)^{-1}$$

which gives

$$V[\hat{\psi}] = \sigma^2 \mathbf{c}' \left(X'X\right)^{-1} \mathbf{c}$$

and the corresponding confidence interval for $\psi =$:

$$\left[\mathbf{c}'\hat{\beta} - t_{n-r,1-\alpha/2}\sqrt{\mathbf{c}'\left(X'X\right)^{-1}\mathbf{c}}s, \mathbf{c}'\hat{\beta} + t_{n-r,1-\alpha/2}\sqrt{\mathbf{c}'\left(X'X\right)^{-1}\mathbf{c}s}\right].$$

In many cases it is trivial to compute $V[\hat{\psi}]$ since the estimates are classical and well known. For example there is no need to complicate the issue when looking at a contrast of the form

$$\overline{y}_{1.} - 2\overline{y}_{2.} + \overline{y}_{3.}$$

in the one-way layout with equal sample sizes J for each i. Here we see trivially that the variance of $\hat{\psi}$ is simply $\sigma^2(4/J)$ and the confidence interval becomes correspondingly trivial to compute.

1.4 Testing hypotheses for multiple estimable functions

$$H_0 : \psi_1 = \psi_2 = \ldots = \psi_q = 0 \text{ vs } H_a : \text{ not } H_0$$

Reject H_0 if

$$\hat{\psi}'\mathbf{B}^{-1}\hat{\psi} > qs^2 F_{q,n-r,1-\alpha}$$

1.4.1 Details

As another example, consider testing the hypothesis that several (linearly independent) estimable functions are zero, i.e. test

$$H_0 : \psi_1 = \psi_2 = \ldots = \psi_q = 0 \text{ vs } H_a : \text{ not } H_0$$

The simplest method to test this hypothesis is to reject H_0 if ψ is not in the confidence set, i.e.: **Reject H_0 if**

$$\hat{\psi}'\mathbf{B}^{-1}\hat{\psi} > qs^2 F_{q,n-r,1-\alpha}$$

1.5 Multiple comparisons

1.5.1 Details

The confidence ellipsoids are of course multiple comparisons in the sense that they provide information about the entire vector of estimable functions under consideration. However it is usually of greater interest to draw conclusions on the individual estimable functions, but the inference should be simultaneous. To this end, the confidence ellipsoids are used as a basis and the intervals are simply deduced from the ellipsoids as follows.
Theorem:

$$P\left[\hat{\psi}_i - \sqrt{qF_{q,n-r,1-\alpha}}\hat{\sigma}_{\hat{\psi}_i} < \psi_i < \hat{\psi}_i + \sqrt{qF_{q,n-r,1-\alpha}}\hat{\sigma}_{\hat{\psi}_i} \quad i = 1,\ldots,q\right] \geq 1-\alpha$$

Corollary: Let $L := \{\psi = \sum_1^q h_i\psi_i : h_1,\ldots,h_q \in \mathbb{R}\}$. Then

$$P\left[\hat{\psi} - \sqrt{qF_{q,n-r,1-\alpha}}\hat{\sigma}_{\hat{\psi}} < \psi_i < \hat{\psi}_i + \sqrt{qF_{q,n-r,1-\alpha}}\hat{\sigma}_{\hat{\psi}} \quad \forall\psi \in L\right] = 1-\alpha$$

1.5.2 Handout

Several interesting, useful and important methods can be derived from these confidence sets. These sets are attributed to Scheffe and are called the S-sets or S-methods of obtaining simultaneous confidence statements.

1.6 Data-snooping

Can use the S-method for data-snooping.
Normally use a large α
Better than LSD: Know explicitly the error rate

1.6.1 Details

Suppose we are interested in **searching for significance** or **data-snooping**. Normally this is not permitted since usually the hypotheses to be tested need to be specified in advance. However, the confidence sets discussed in this tutorial are all simultaneous and can therefore be searched in arbitrary detail.
Suppose Ψ is a set of estimable functions, e.g. a set spanned by q estimable functions: $\Psi = \left\{\psi = k_1\psi_1 + \ldots + k_q\psi_q\right\}$ where $\psi_i = \mathbf{c}_i'\beta$ and $\mathbf{c}_1, \ldots, \mathbf{c}_n$ are linearly independent. Then from the earlier results we can assert

$$P\left[\hat{\psi} - \sqrt{qF^*}\hat{\sigma}_{\hat{\psi}} \leq \psi \leq \hat{\psi} + \sqrt{qF^*}\hat{\sigma}_{\hat{\psi}} \quad \forall\psi \in \Psi\right] = 1 - \alpha$$

and we are therefore allowed to **search** among all estimable functions within the set to find significant effects.
The "trick" here lies in the cutoff-point, $qF^* = qF_{q,n-r,1-\alpha}$, which takes into account the dimension of the space.

2 Special cases of Scheffes confidence sets: Applications to simple linear regression

2.1 The setup

2.1.1 Handout

We will assume the model to be $y_i \sim n(\alpha + \beta x_i, \sigma^2)$, independent.
In this case the OLS estimators, $\hat{\alpha}$, $\hat{\beta}$ are well known linear combinations of the y-values. They can be written as

$$\hat{\beta} = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} \mathbf{a_1 y} \\ \mathbf{a_2 y} \end{pmatrix} = \mathbf{Ay}$$

for an appropriate choice of $\mathbf{a_1}$,$\mathbf{a_2}$ and $\mathbf{A}$.
The variance-covariance matrix has been derived elsewhere as $\sigma^2\mathbf{AA}' = \sigma^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}$.
Since σ^2 can be estimated with $s^2 = MSE$, the variances and covariances of $\hat{\alpha}$ and $\hat{\beta}$ can easily be estimated.

2.2 The intercept

2.2.1 Handout

The intercept alone is a simple linear function of the full parameter vector, i.e.

$$\psi = \alpha = (1,0)\begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

the corresponding estimate is

$$\hat{\psi} = \hat{\alpha} = (1,0)\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \mathbf{a_1 y}.$$

The variance of this particular estimable function is well known

$$\hat{\sigma}^2_{\hat{\psi}} = \dots$$

Since this is a single estimable function we have $q = 1$. If the x-values are not all the same then $\mathbf{X}$ has full rank so $r = p = 2$ and we obtain

...

the same CI as before.

2.3 The slope

$\square$

2.3.1 Handout

The slope alone is a simple linear function of the full parameter vector, i.e.

$$\psi = \beta = (0,1)\begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

the corresponding estimate is

$$\hat{\psi} = \hat{\beta} = (0,1)\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \mathbf{a_2 y}.$$

The variance of this particular estimable function is well known

$$\hat{\sigma}^2_{\hat{\psi}} = \dots$$

Since this is a single estimable function we have $q = 1$. If the x-values are not all the same then $\mathbf{X}$ has full rank so $r = p = 2$ and we obtain

...

the same CI as before.

2.4 A simultaneous confidence set for the slope and intercept

2.4.1 Handout

Recall that the vector $\psi = (\hat{\alpha}, \hat{\beta})'$ is estimable in simple linear regression if the x-values are not all the same. A simultaneous confidence set for ψ is based on the point estimate $\psi = (\hat{\alpha}, \hat{\beta})'$ and the corresponding covariance matrix and the earlier result

$$\frac{(\hat{\psi} - \psi)' \mathbf{B}^{-1} (\hat{\psi} - \psi)/q}{||\mathbf{y} - \mathbf{X}\hat{\beta}||^2/(n-r)} \sim F_{q,n-r} \qquad (4)$$

where in this case $\mathbf{B} = \mathbf{X}'\mathbf{X}$.
Equation 4 provides a confidence set,

$$\left\{ \psi : \frac{(\hat{\psi} - \psi)' \mathbf{B}^{-1} (\hat{\psi} - \psi) / q}{||\mathbf{y} - \mathbf{X}\hat{\beta}||^2 / (n - r)} \leq F_{q, n-r, 1-\alpha} \right\}, \tag{5}$$

which describes an ellipse in the (α, β)-plane.
This confidence set can be used to obtain simultaneous bound on the two parameters.

2.5 Confidence band for the regression line

2.5.1 Handout

The simultaneous confidence set for the two parameters in SLR can be used to obtain a confidence band for the regression line.
The confidence band for the regression line is a simultaneous statement on all points in the set

$$\mathcal{C} = \{\alpha + \beta x : x \in \mathbb{R}\}$$

Now, the variance of the estimates $\hat{\alpha} + \hat{\beta}x$ is well known and it is also clear that the above confidence set is a subset of

$$\mathcal{L} = \{\psi = c_1\alpha + c_2\beta : c_1, c_2 \in \mathbb{R}\}.$$

This is the set of all linear combinations of the two-dimensional parameter vector $(\alpha, \beta)'$, which is an estimable function,

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

The above set, $\mathcal{L}$, consists of all linear combinations of ψ_1 and ψ_2 and can be written as

$$\mathcal{L} = \left\{ \psi = \sum_{h=1}^{2} h_i \psi_i : h_1, h_2 \in \mathbb{R} \right\}.$$

This demonstrates that $\mathcal{L}$ is spanned by two estimable functions, $\psi_1 = \alpha$ and $\psi_2 = \beta$ and $\mathcal{L}$, as in the corollary earlier. It therefore has dimension $q = 2$ and one can use a corresponding F-cutoff to obtain simultaneous confidence bounds for the entire regression line.
To derive the actual formulae, note that a generic point on the regression line, $\psi = \psi_x = \alpha + \beta x$ (an element of $\mathcal{L}$) is predicted with $\hat{\psi} = \hat{\alpha} + \hat{\beta}x$, which has variance

$$\sigma^2_{\hat{\psi}} = V\left[(1, x)\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}\right] = \sigma^2 (1, x) \left(X'X\right)^{-1} \begin{pmatrix} 1 \\ x \end{pmatrix}$$

and as usual, this variance is estimated using

$$\hat{\sigma}^2_{\hat{\psi}} = s^2 (1, x) \left(X'X\right)^{-1} \begin{pmatrix} 1 \\ x \end{pmatrix}.$$

The confidence band for the entire regression line thus becomes

$$\hat{\alpha} + \hat{\beta}x \pm s\sqrt{2F_{2, n-2, 1-\alpha}(1, x) \left(X'X\right)^{-1} \begin{pmatrix} 1 \\ x \end{pmatrix}}.$$

Note the several "tricks"here, where we know the appropriate variances and can use them directly.

3 The Bonferroni approach to multipe comparisons

3.1 Bonferroni confidence intervals

Bonferroni intervals: Simple Always work Conservative

3.1.1 Details

In general, consider two events, A and B having the same probability, $P[A] = P[B] = \alpha'$. In the current situation, A is the event "confidence interval 1 is wrong" and B is the event "confidence interval 2 is wrong".
The probability of both confidence intervals being correct is

$$
\begin{aligned}
P[A^c \cap B^c] &= P[(A \cup B)^c] \\
&= 1 - P[A \cup B] \\
&= 1 - (P[A] + P[B] - P[A \cap B]) \\
&\geq 1 - P[A] - P[B] \\
&= 1 - 2\alpha'
\end{aligned}
$$

It follows that if two confidence statements are made, each with error rate $\alpha' = \alpha/2$, or confidence $100(1-\alpha/2)\%$, then the overall confidence is at least $100(1-\alpha)\%$, i.e. the probability of any error is reduced to α.
References ISBN: 0412982811

4 Tukeys confidence intervals

4.1 Pairwise multiple comparisons

Tukey's method for pairwise comparisons works!

4.1.1 Details

When all pairwise comparisons are of equal importance, the interest is in being able to make statements of the form

$$P[|\bar{X}_i - \bar{X}_j| \leq d_{ij} \text{ for all } i, j] \geq 1 - \alpha$$

Usually, d_{ij} is taken proportional to the common standard deviation, s and written either as $qs/\sqrt{n}$ or $ws/\sqrt{1/n_i + 1/m_j}$ in the case of unequal sample sizes.
The function TukeyHSD in R and the procedure "proc glm" in SAS (with the Tukey option) can be used for general, and valid, pairwise multiple comparisons.

4.2 Tukeys confidence intervals

□

4.2.1 Details

Tukeys confidence intervals

5 Simultaneous confidence intervals for all contrasts

5.1 Scheffes method

□

6 Comparing confidence sets

6.1 Scheffe, Tukey and Bonferroni

□