Permanova

Técnicas de Análisis Multivariante

Universitat d'Alacant
Universidad de Alicante

Departamento de Ciencias del Mar y
Biología Aplicada

Introducción

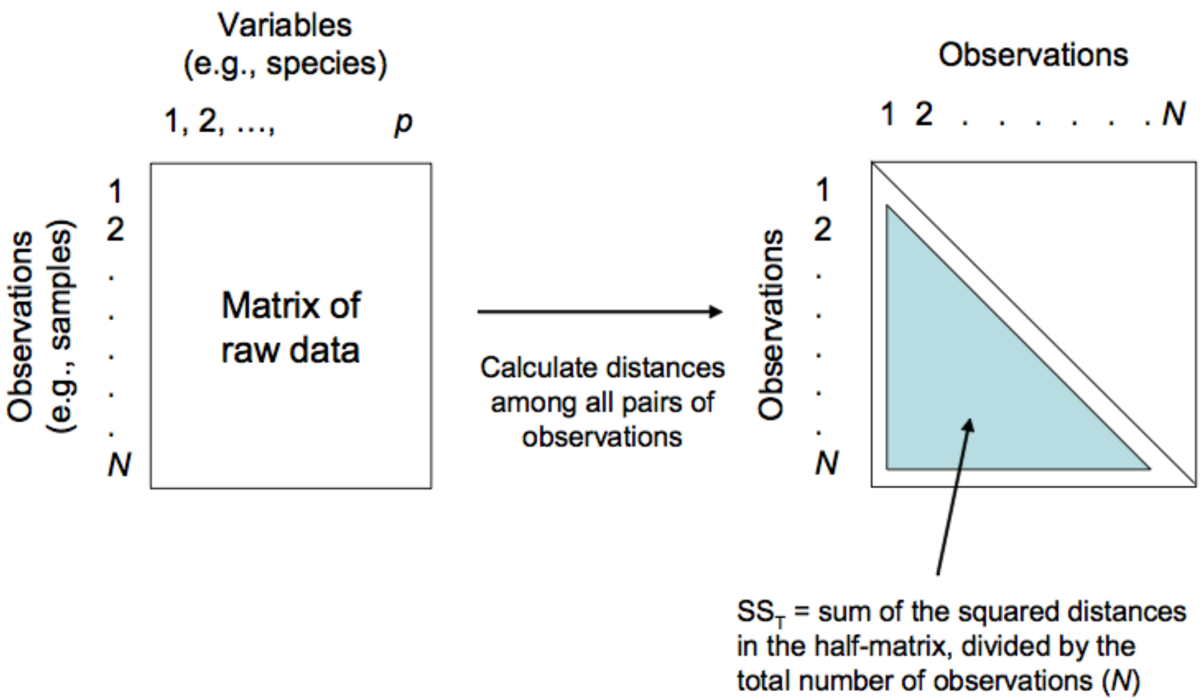
Algoritmo para medir la respuesta simultánea de una o varias variables a uno o varios factores en un diseño experimental ANOVA, basado en cualquier medida de distancias y usando análisis de permutaciones

Técnica de análisis multivariante (sobre medidas de distancia) con varios factores (balanceados) (MANOVA)

Aplica análisis de permutaciones sobre las matrices de distancia (PERmutaciones)

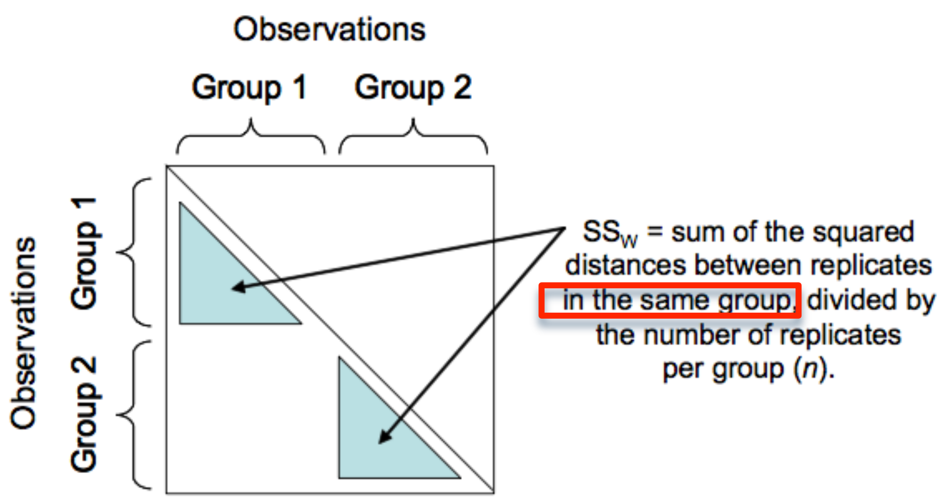
PerMANOVA

Ejemplo Unifactorial



$$SS_T = \frac{1}{N} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} d_{ij}^2$$

Ejemplo Unifactorial



$$SS_W = \frac{1}{n}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} d_{ij}^2 \varepsilon_{ij}$$

Pseudo F-ratio

$$F = \frac{SS_A/(a-1)}{SS_W/(N-a)}$$

Permanova

P-valor

La idea es calcular el p-valor usando permutaciones.

Si no hubiera diferencias entre los valores se podría permutar los valores (todos iguales)

Calcular la cantidad de permutaciones que pueden resultar con valores mas extremos

$$P = \frac{\left(No.\,of\ F^{\pi} \geq F\right)+1}{\left(Total\ no.\,of\ F^{\pi}\right)+1}$$

Permanova

P-valor depende del **Número de permutaciones**

Si sólo podemos hacer 10 permutaciones (i.e. 2 grupos x 3 réplicas)

$$(6!)/[2!(3!)^2]= 10$$

El mínimo p-valor = 0.10

Resultaría insuficiente para análisis estadístico clásico

Permanova Requisitos

Similares a los de ANOVA:

Independencia del término del **error**

Normalidad de los mismos

Homogeneidad sus varianzas (muy sensible a la heterocedasticidad entre varianzas entre niveles de factores)

Aditividad de los tratamientos (factores)

Permanova - Medidas de distancia

$$d_{ij} = \frac{\sum_{k=1}^{p} \left| y_{ik} - y_{jk} \right|}{\sum_{k=1}^{p} (y_{ik} + y_{jk})}$$

Disimilaridad de Bray-Curtis

$$d_{ij} = \sqrt{\sum_{k=1}^{p} \frac{1}{y_{\bullet k}} \left(\frac{y_{ik}}{y_{i\bullet}} - \frac{y_{jk}}{y_{j\bullet}} \right)^2}$$

Chi square metric

$$d_{ij} = \sqrt{\frac{\sum_{k=1}^{p} \left| y_{ik} - y_{jk} \right|}{\sum_{k=1}^{p} (y_{ik} + y_{jk})}}$$

RS Bray-Curtis

$$d_{ij} = \sqrt{\sum_{k=1}^{p} \left(\sqrt{\frac{y_{ik}}{y_{i\bullet}}} - \sqrt{\frac{y_{jk}}{y_{j\bullet}}} \right)^2}$$

Distancia de Hellinger

$$d_{ij} = \sqrt{\sum_{k=1}^{p} (y_{ik} - y_{jk})^2}$$

Distancia Euclídea

$$d_{ij} = \sum_{k=1}^{p} \left\{ \left| y_{ik} - y_{jk} \right| / R_k \right\}$$

Disimilaridad de Gower

Permanova - Resumen

Técnica de análisis multivariante (sobre medidas de distancia) con varios factores (balanceados) (MANOVA)

Aplica análisis de permutaciones sobre las matrices de distancia (Permutaciones) para obtener p-valor

Limitado por el número de permutaciones posibles
Número de permutaciones recomendable (entre 200 y 1000) (cuantas más, mejor :)

También funciona con otras medidas de distancia (o incluso con valores de variables) pero debemos tener en cuenta el significado del test