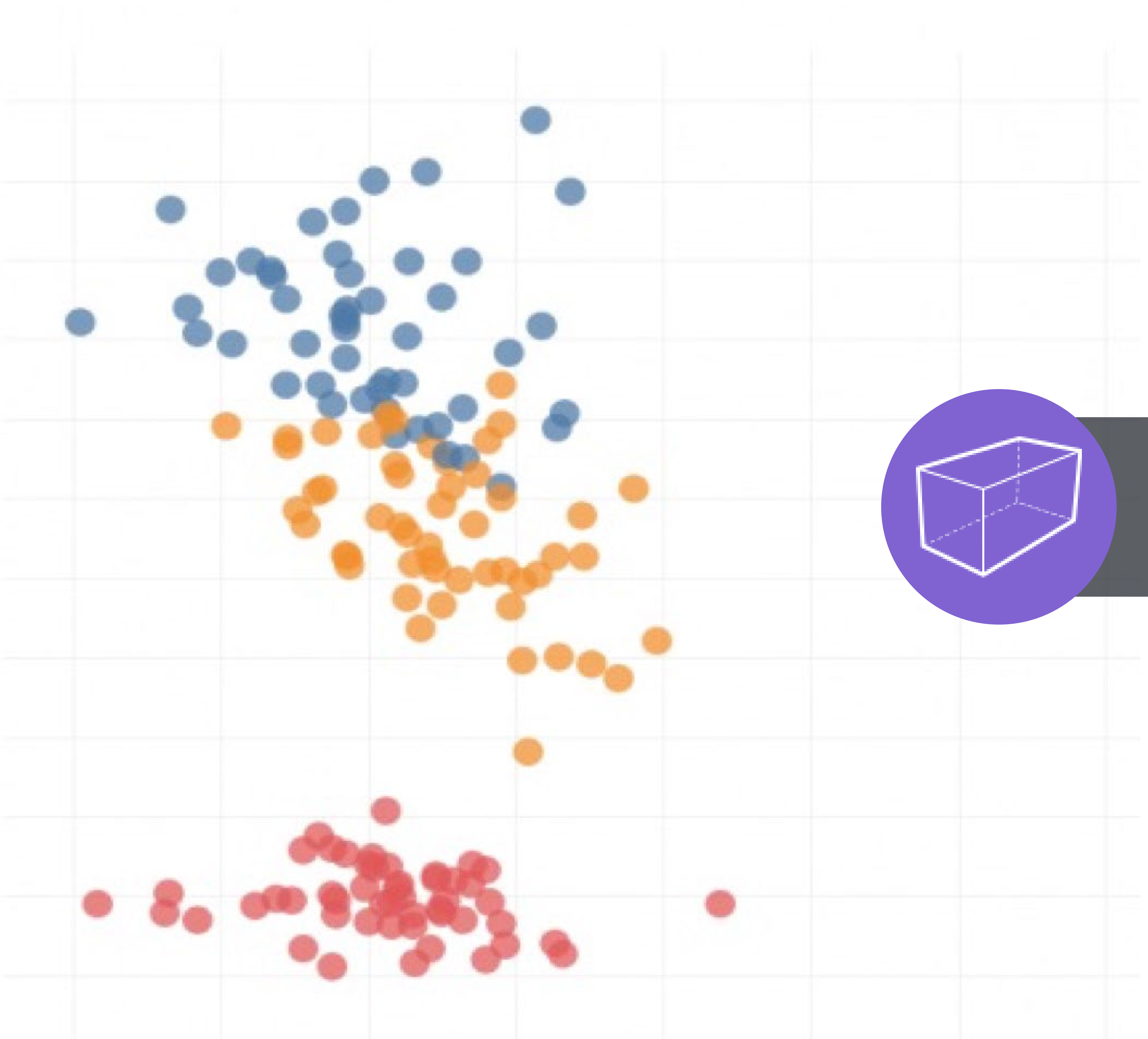
ALGES
ADVANCED LABORATORY FOR GEOSTATISTICAL SUPERCOMPUTING

ANOVA Análisis de varianza

Análisis Estadístico y Geoestadístico de Datos

Auxiliar: Fabián Soto F.
Profesor: Xavier Emery

Ingeniería de Minas
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE



Análisis de varianza

Resumen

Conceptos Básicos

- ANOVA: Análisis de Varianza → Considera variables gaussianas, independientes y de igual varianza

- Análisis de Varianza de un Factor
- Análisis de Varianza de Dos Factores
- Análisis de Varianza de Dos Factores con Réplica
- Análisis de Varianza Múltiple

ANOVA de un Factor

- Análisis de Varianza de un Factor → $H_0: \mu_1 = \mu_2 = \cdots = \mu_k$
  $H_1: \mu_i \neq \mu_j$

Fuente de Variación	Suma de Cuadrados	Grados de Libertad	Media de Cuadrados	F
Explicada (Factor de Variación A)	$SS_1 = \sum_{i=1}^{n} n_i(\bar{X}_i - \bar{X})^2$	$k-1$	$MS_1 = \frac{SS_1}{k-1}$	$\frac{MS_1}{MS_0}$
Residual (Error)	$SS_0 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2$	$n-k$	$MS_0 = \frac{SS_0}{n-k}$	
Total	$SS = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \bar{X})^2$	$n-1$		

1. Se compara contra una distribución de Fisher con $k-1$ y $n-k$ grados de libertad.
2. Es un test unilateral.
3. $SS = SS_0 + SS_1$.

ANOVA de Dos Factores

- Análisis de Varianza de Dos Factores → Considera dos fuentes de variación

Tratamiento	B_1	B_2	...	B_c	Media
A_1	X_{11}	X_{12}	...	X_{1c}	$\bar{X}_{1.}$
A_2	X_{21}	X_{22}	...	X_{2c}	$\bar{X}_{2.}$
...					
A_f	X_{f1}	X_{f2}	...	X_{fc}	$\bar{X}_{f.}$
Media	$\bar{X}_{.1}$	$\bar{X}_{.2}$		$\bar{X}_{.c}$	$\bar{X}$

ANOVA de Dos Factores

- ¿Cómo saber si afecta la fuente A o B al valor esperado de x?

Fuente de Variación	Suma de Cuadrados	Grados de Libertad	Media de Cuadrados
A	$SS_f = \sum_{i=1}^{f} c(\bar{X}_{i.} - \bar{X})^2$	$f - 1$	$MS_f = \frac{SS_f}{f-1}$
B	$SS_c = \sum_{j=1}^{c} f(\bar{X}_{.j} - \bar{X})^2$	$c - 1$	$MS_c = \frac{SS_c}{c-1}$
Error Residual	$SS_e = \sum_{i=1}^{f}\sum_{j=1}^{c}(X_{ij} - \bar{X}_{i.})(X_{ij} - \bar{X}_{.j})$	$(f-1)(c-1)$	$MS_e = \frac{SS_e}{(f-1)(c-1)}$
Total	$SS = SS_f + SS_c + SS_e$	$\mathrm{n} - 1$	

ANOVA de Dos Factores

- H_0: No hay influencia de la fuente en el valor esperado de x
- H_1: Hay una influencia de la fuente en el valor esperado de x
    - INFLUENCIA DE A
        - Se compara el cociente $\frac{MS_f}{MS_e}$ con una variable Fisher de $(f-1)$ y $(f-1)(c-1)$ grados de libertad
        - Test de carácter unilateral
    - INFLUENCIA DE B
        - Se compara el cociente $\frac{MS_c}{MS_e}$ con una variable Fisher de $(c-1)$ y $(c-1)(f-1)$ grados de libertad
        - Test de carácter unilateral

ANOVA de Dos Factores con Réplica

Fuente de Variación	Suma de Cuadrados	Grados de Libertad	Media de Cuadrados
A	$SS_f = r\sum_{i=1}^{f} c(\bar{X}_{i..} - \bar{X})^2$	$f - 1$	$MS_f = \frac{SS_f}{f-1}$
B	$SS_c = r\sum_{j=1}^{c} f(\bar{X}_{.j.} - \bar{X})^2$	$c - 1$	$MS_c = \frac{SS_c}{c-1}$
Interacción A/B	$SS_{fc} = r\sum_{i=1}^{f}\sum_{j=1}^{c}(\bar{X}_{ij.} - \bar{X}_{i..} - \bar{X}_{.j.} + \bar{X})^2$	$(f-1)(c-1)$	$MS_{fc} = \frac{SS_{fc}}{(f-1)(c-1)}$
Error Residual	$SS_e = \sum_{i=1}^{f}\sum_{j=1}^{c}\sum_{k=1}^{r}(X_{ijk} - \bar{X}_{ij.})^2$	$fc(r-1)$	$MS_e = \frac{SS_e}{fc\,(r-1)}$
Total	$SS = SS_f + SS_c + SS_{fc} + SS_e$	$\mathrm{n} - 1$	

ANOVA de Dos Factores con Réplica

- 1º se analiza si hay INTERACCIÓN entre A y B
  - H_0: No hay interacción entre factores A y B
  - H_1: Existe interacción entre los factores A y B
    - Se compara el cociente $\frac{MS_{fc}}{MS_e}$ con una variable Fisher de $(f-1)(c-1)$ y $fc(r-1)$ grados de libertad

ANOVA de Dos Factores con Réplica

- 2º si existe una INTERACCIÓN entre los factores A y B, entonces:
  - H_0: No hay influencia de la fuente en el valor esperado de x.
  - H_1: Hay una influencia de la fuente en el valor esperado de x.
    - INFLUENCIA DE A
      - Se compara el cociente $\frac{MS_f}{MS_e}$ con una variable Fisher de $(f-1)$ y $fc(r-1)$ grados de libertad.
    - INFLUENCIA DE B
      - Se compara el cociente $\frac{MS_c}{MS_e}$ con una variable Fisher de $(c-1)$ y $fc(r-1)$ grados de libertad.

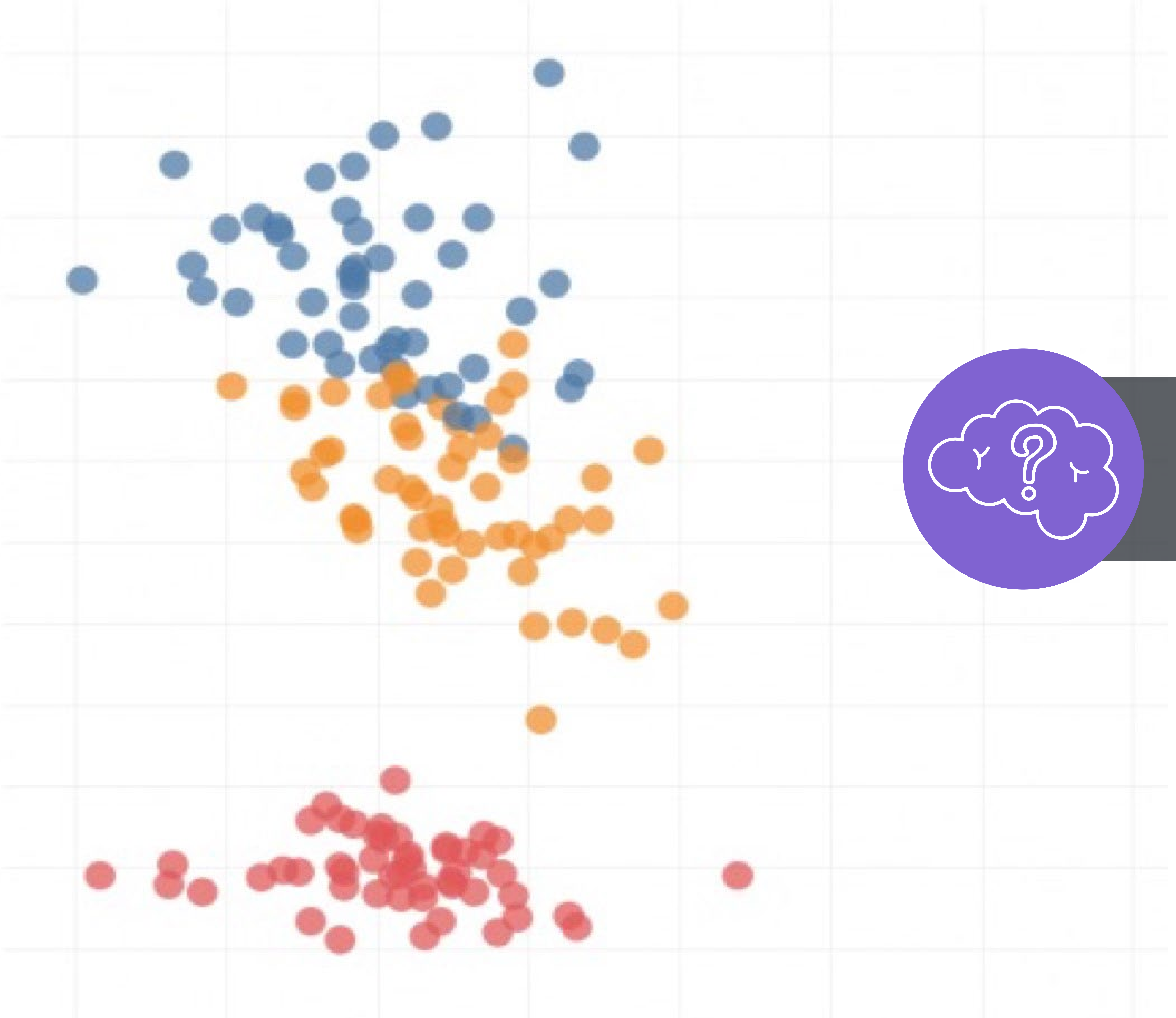
ANOVA de Dos Factores con Réplica

- 3º si no existe una INTERACCIÓN entre los factores A y B, entonces:
  - Se puede combinar la "INTERACCIÓN" con la variabilidad residual, calculando un error combinado

- Suma de Cuadrados Combinada:
  - $SS_{combinada} = SS_{fc} + SS_e$
- Grados de Libertad Combinado:
  - $GL_{combinados} = (f-1)(c-1) + fc(r-1).$
- Media de cuadrados combinada:
  - $MS_{combinada} = SS_{combinada}/GL_{combinados}$

ANOVA de Dos Factores con Réplica

- 3º si no existe una INTERACCIÓN entre los factores A y B, entonces:
  - Una vez calculado se realiza el siguiente análisis
    - INFLUENCIA DE A
      - Se compara el cociente $\frac{MS_f}{MS_{combinado}}$ con una variable Fisher de $(f-1)$ y $(fcr-f-c+1)$ grados de libertad
    - INFLUENCIA DE B
      - Se compara el cociente $\frac{MS_c}{MS_{combinado}}$ con una variable Fisher de $(c-1)$ y $(fcr-f-c+1)$ grados de libertad



ALGES
ADVANCED LABORATORY FOR GEOSTATISTICAL SUPERCOMPUTING

Problemas

ANOVA

Problema 1:

- Se desea saber si existe una diferencia significativa en las leyes medias analizadas por los diferentes laboratorios. Si hay diferencia significativa, ¿Podría identificar el laboratorio diferente?
- Archivo: *“Round Robin.xls”*

Problema 1:

- Se desea saber si existe una diferencia significativa en las leyes medias analizadas por los diferentes laboratorios. Si hay diferencia significativa, ¿Podría identificar el laboratorio diferente?
- Archivo: *"Round Robin.xls"*

Problema 2:

- Construir la tabla de análisis de varianza para los siguientes datos.
- ¿Existen diferencias significativas entre bloques y/o entre tratamientos?

	Bloque 1	Bloque 2	Bloque 3	Bloque 4	Bloque 5
Tratamiento 1	6,325	7,020	5,896	6,619	6,123
Tratamiento 2	7,186	7,361	6,132	7,106	6,250
Tratamiento 3	6,643	6,511	5,928	6,532	5,938

Problema 3:

- Se busca determinar si el uso de un nuevo acero aumenta la velocidad de perforación. Para ello, se hicieron 4 pruebas con cada tipo de acero (actual / nuevo) en dos tipos de roca (dura y blanda).
- ¿Valdrá la pena cambiar el acero actual por el acero nuevo?

[m/min]	Roca Dura	Roca Blanda
Acero Actual	40.2 41.3 39.7 40.5	42.2 42.5 44.0 43.8
Acero Nuevo	41.1 40.8 40.4 40.9	42.9 42.2 43.9 43.7