Introducción a la estadística Bases indispensables y uso de

Olivier Devineau olivier.devineau@fcdarwin.org.ec

Fundación Charles Darwin

Taller interno, 27-30 abril 2010

Introducción y conceptos importantes

Cosas importantes

Teoría estadística: 8:30–10:00, 10:30–12:00

• Práctica con R: 13:30–15:00, 15:30–17:00

Café: 10:00–10:30 y 15h00-15h30

Por favor, apagan los celulares

¡Preguntas bienvenidas en cualquier momento!

Agradecimientos

del taller

Logística

Agradacimientos

Diseño

experimenta

conceptos

Use material amablemente provisto por:

- Claude-Pierre Guillaume, EPHE, Montpellier, Francia
- Damien Caillaud, UT, Austin, Texas, USA
- Julien Dutheil, CNRS, Montpellier, Francia
- Vladimir Grosbois, CIRAD, Montpellier, Francia

Correcciones, comentarios y sugerencias por

• Eliana Bontti, FCD

Agradecimientos

Y también:

- Crawley, M.J. 2005. Statistics, an introduction using R. John Wiley & Sons. (con el consentimiento del autor)
- Quinn, G.P., and Keough, M.J. 2002. Experimental design and data analysis for biologists. Cambridge University Press.

Licencia

- Este documento está bajo la licencia Creative Commons: Reconocimiento - No comercial - Compartir bajo la misma licencia 3.0 Ecuador
- Para ver una copia de esta licencia, visite: http://creativecommons.org/licenses/by-nc-sa/3.0/ec/
- Código LATEX a petición

Introducc

¿Qué es?

La verdadera estadística ¿Qué puede hacer la esatística? Descriptiva / inferencial

¿Cómo empezar

Varianz

experimenta

0.

¿Qué es la estadística? Definición

- Principios y métodos para recoger, clasificar, resumir y analizar datos
- Aprender, hacer conclusiones y tomar decisiones

La verdadera estadística ¿Qué puede hacer la esatística? Descriptiva / inferencial

empezar

Varianz

evperimenta

experimenta

Otros conceptos

¿Qué es la estadística? Definición

- Principios y métodos para recoger, clasificar, resumir y analizar datos
- Aprender, hacer conclusiones y tomar decisiones

Otros

La verdadera estadística . . .

Evolución de salarios y empleados en una empresa

		Obreros	Ejecutivos	Promedio
Salario	2004	200	2000	1100
	2006	180	1800	990
Empleados	2004	1000	100	550
	2006	600	500	550

Organización del taller

¿Qué es? La verdadera

estadística ¿Qué puede hacer la esatística? Descriptiva / inferencial

inferencial ¿Cómo empezar?

experimenta

0.

La verdadera estadística . . .

Evolución de salarios y empleados en una empresa

		Obreros	Ejecutivos	Promedio
Salario	2004	200	2000	1100
	2006	180	1800	990
Empleados	2004	1000	100	550
	2006	600	500	550

Periódico Salarios bajaron en un 10%

Empresa Salario promedio por empleado aumentó de \$363.6 a \$916.3

Organización del taller

Introducció

La verdadera estadística ¿Qué puede hacer la

esatística?

Descriptiva / inferencial
¿Cómo empezar?

Varian

Diseño

experimenta

Otros

La verdadera estadística . . .

Evolución de salarios y empleados en una empresa

		Obreros	Ejecutivos	Promedio
Salario	2004	200	2000	1100
	2006	180	1800	990
Empleados	2004	1000	100	550
	2006	600	500	550

Periódico Hubo despidos en la empresa Empresa Igual número de empleados y reclutamiento

¿Qué puede hacer la esatística?

empezar?

La estadística

Puede

- Proveer criterios objetivos para probar hipótesis
- Optimizar esfuerzos

¿Qué es? La verdadera

¿Qué puede hacer la esatística?

inferencia ¿Cómo empezar?

Varianz

Diseño experimenta

Otros conceptos

Puede

- Proveer criterios objetivos para probar hipótesis
- Optimizar esfuerzos
- Evaluar razonamiento de manera crítica

VO puede

- Decir la verdad
- Compensar ausencia de controles o mala planificación
- Indicar importancia que no es probabilística

¿Qué es? La verdadera estadística

¿Qué puede hacer la esatística?

inferencia ¿Cómo empezar?

Varianz

Diseño experimental

Otros

Puede

- Proveer criterios objetivos para probar hipótesis
- Optimizar esfuerzos
- Evaluar razonamiento de manera crítica

- Decir la verdad
- Compensar ausencia de controles o mala planificación
- Indicar importancia que no es probabilística

¿Qué es? La verdadera

¿Qué puede hacer la esatística?

inferencia ¿Cómo empezar?

Varianza

Diseño experimental

Otros conceptos

Puede

- Proveer criterios objetivos para probar hipótesis
- Optimizar esfuerzos
- Evaluar razonamiento de manera crítica

- Decir la verdad
- Compensar ausencia de controles o mala planificación
- Indicar importancia que no es probabilística

¿Qué puede hacer la esatística?

- Proveer criterios
- Optimizar esfuerzos

La estadística

- Compensar ausencia de controles o mala planificación

empezar?

La estadística

- Proveer criterios
- Optimizar esfuerzos

- Indicar importancia que no es probabilística

Introducció

La verdadera estadística ¿Qué puede hacer la esatística?

Descriptiva / inferencial

¿Cómo

empeza

Varianz

experiment

Owner

Otros conceptos

Primer paso para entender datos: ¡describirlos!

- Distribución normal, poisson, binomial ...
- Media, mediana
- Varianza, desviación estándar y error estándar

⇒ Estadística descriptiva informa sobre forma, centro y amplitud de los datos Industrial Confession

¿Qué es? La verdadera estadística ¿Qué puede hacer la

Descriptiva /

: Cómo

empezar?

Varianz

Diseño

Otros

Describir no es suficiente

- No es suficiente averiguar que hay variación
- ¿Variación científicamente interesante o variación natural?

Descriptiva / inferencial

¿Cómo empezar?

empezar

varianz

Diseno

experimenta

Describir no es suficiente

- No es suficiente averiguar que hay variación
- ¿Variación científicamente interesante o variación natural?

estadística ¿Qué puede hacer la esatística? Descriptiva /

inferencial

¿Cómo empezar

Varianz

Diseño

experimenta

Otros conceptos

Describir no es suficiente

- No es suficiente averiguar que hay variación
- ¿Variación científicamente interesante o variación natural?

Estadística inferencial permite:

- Distinguir entre señal y ruido
- Deducir información y llegar a conclusiones

; Cómo

empezar?

Lo más difícil es empezar

- ¿ Qué tipo de análisis?
- ¿Cómo saber que hacer? ¡habiéndolo hecho miles de veces!

Lo más difícil es empezar

- ¿ Qué tipo de análisis?
- Depende de los datos y de la pregunta inicial
- ¿Cómo saber que hacer? ¡habiéndolo hecho miles de veces!

Introducció

La verdadera estadística ¿Qué puede hacer la esatística? Descriptiva

¿Cómo empezar?

Varianz

Diseño experimenta

experimenta

Otros conceptos

Lo más difícil es empezar

- ¿Qué tipo de análisis?
- Depende de los datos y de la pregunta inicial
- ¿Cómo saber que hacer? ¡habiéndolo hecho miles de veces!

¿Estadística paramétrica o no?

Paramétrica.

- Intervalos regulares
- Hipótesis de distribución normal
- Media y error/desviación estándar

No paramétrica

- Cualquier tipo de escala
- No hipótesis de distribución (independencia)
- Mediana y desviación mediana

- ¿Cuál es la variable dependiente?
- ; De qué tipo es? ; Medida continua, número, proporción,
- ; Cuáles son las variables independientes?
- ; Son continuas? ; Categóricas? ; Ambos?

- ¿Cuál es la variable dependiente?
- ¿De qué tipo es? ¿Medida continua, número, proporción, categoría?
- ; Cuáles son las variables independientes?
- ; Son continuas? ; Categóricas? ; Ambos?

¿Cómo empezar?

Varianza

Diseño

experimenta

Otros conceptos

- ¿Cuál es la variable dependiente?
- ¿De qué tipo es? ¿Medida continua, número, proporción, categoría?
- ¿Cuáles son las variables independientes?
- ¿Son continuas? ¿Categóricas? ¿Ambos?

¿Cómo empezar?

Varianz

Diseño

experimenta

Otros conceptos

- ¿Cuál es la variable dependiente?
- ¿De qué tipo es? ¿Medida continua, número, proporción, categoría?
- ¿Cuáles son las variables independientes?
- ¿Son continuas? ¿Categóricas? ¿Ambos?

Organización del taller

Introducció ¿Qué es? La verdadera estadística ¿Qué puede hacer la

Descriptiva / inferencial ¿Cómo empezar?

varianz

experimental

experimental

¿Qué análisis? Guía de decisión

1) Variables independientes

Todas continuas

Regresión

Todas categóricas

Anova

Ambas continuas y categóricas

ncova

Descriptiva /

; Cómo empezar?

¿Qué análisis? Guía de decisión

1) Variables independientes

Todas continuas

Todas categóricas

Ambas continuas y categóricas

Regresión

Anova

Organización del taller

Introducción
¿Qué es?
La verdadera
estadística
¿Qué puede
hacer la
esatística?
Descriptiva /

¿Cómo empezar?

Vanian-

V di laliz

evperimenta

скреппенса

Otros

¿Qué análisis? Guía de decisión

1) Variables independientes

Todas continuas

Regresión

Todas categóricas

Anova

• Ambas continuas y categóricas

Ancova

Descriptiva / ; Cómo

empezar?

¿Qué análisis? Guía de decisión 2) Variable dependiente

Continua

Regresión normal, Anova, Ancova

Número

Tiempo hasta la muerte

Organización

Introducció

La verdader estadística ¿Qué puede hacer la esatística?

¿Cómo

empezar?

Varianz

Diseño experimental

Otros

¿Qué análisis? Guía de decisión

2) Variable dependiente

Continua

Regresión normal, Anova, Ancova

Proporción

Regresión logística

Número

Regresión log-lineal

Rinaria

Análisis logístico binario

• Tiempo hasta la muerte

Análisis de sobrevivencia

; Cómo

empezar?

¿ Qué análisis? Guía de decisión 2) Variable dependiente

Continua

Proporción

Número

Tiempo hasta la muerte

Regresión normal, Anova, Ancova

Regresión logística

Regresión log-lineal

Organización del taller

Introducció

La verdader: estadística ¿Qué puede hacer la esatística?

¿Cómo empezar?

Varianz

D. ...

experimental

Otros

¿Qué análisis? Guía de decisión 2) Variable dependiente

Continua

Proporción

Número

Binaria

• Tiempo hasta la muerte

Regresión normal, Anova, Ancova

Regresión logística

Regresión log-lineal

Análisis logístico binario

Análisis de sobrevivencia

Introducción a la estadística

Organización del taller

Introducció

La verdadera estadística ¿Qué puede hacer la esatística? Descriptiva

¿Cómo empezar?

Varianz

Diseño

ехреппента

Otros conceptos

¿Qué análisis? Guía de decisión 2) Variable dependiente

Continua

Regresión normal, Anova, Ancova

Proporción

Regresión logística

Número

Regresión log-lineal

Binaria

Análisis logístico binario

• Tiempo hasta la muerte

Análisis de sobrevivencia

Por qué la estadística? ¡Porque Todo varia!

Mucha variabilidad temporal, espacial y entre individuos:

- Genética
- Factores ambientales
- Azar
- Errores de observación y medida

¿Como medir la variabilidad?

Organización del taller

Introducción

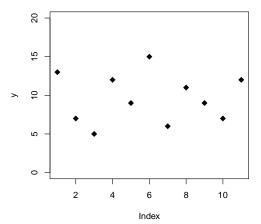
Todo varia

Concepto Grados de

Definición Otros aspect

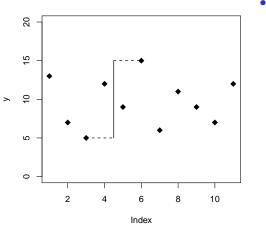
Diseño experiment:

Otros



Concepto Grados de

¿Como medir la variabilidad?



• Rango: [5, 15]

¿Como medir la variabilidad?

Organización del taller

Introducción

Todo varia

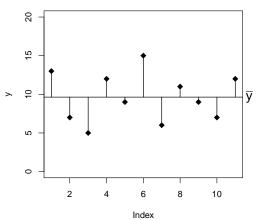
Concepto Grados de

Definición

Diseño

experiment

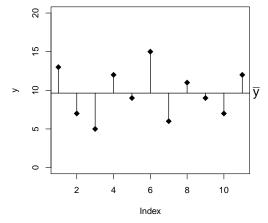
Otros conceptos



- Rango: [5, 15]
- Media y desviaciones de la media

Concepto Grados de

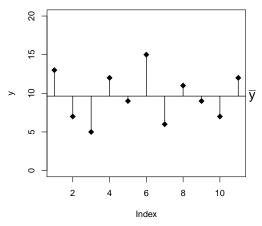
¿Como medir la variabilidad?



- Rango: [5, 15]
- Media y desviaciones de la media
- Residuales

Concepto Grados de

¿Como medir la variabilidad?

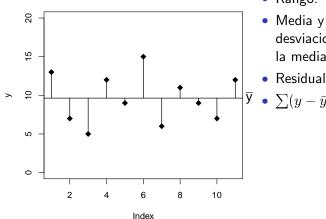


- Rango: [5, 15]
- Media y desviaciones de la media
- Residuales
- $\sum (y \bar{y})$

¿Como medir la variabilidad?

Todo varia Concepto

Grados de

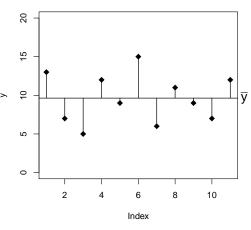


- Rango: [5, 15]
- desviaciones de la media
- Residuales

•
$$\sum (y - \bar{y}) = 0$$

Todo varia Concepto

¿Como medir la variabilidad?



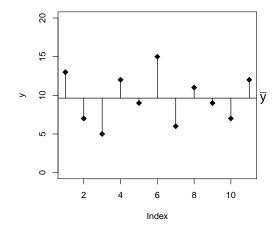
- Rango: [5, 15]
- Media y desviaciones de la media
- Residuales

•
$$\sum (y - \bar{y}) = 0$$

•
$$SS = \sum (y - \bar{y})^2$$

Concepto

¿Como medir la variabilidad?



- Rango: [5, 15]
- Media y desviaciones de la media
- Residuales

•
$$SS = \sum (y - \bar{y})^2$$

 Suma de los cuadrados (sum of squares)

Concepto

Grados de

Una mejor medida de la variabilidad

•
$$SS = \sum (y - \bar{y})^2$$
, $n = 11$

experimenta

Otros conceptos

Una mejor medida de la variabilidad

- $SS = \sum (y \bar{y})^2$, n = 11

Concepto

Una mejor medida de la variabilidad

- $SS = \sum (y \bar{y})^2$, n = 11
- i Que pasa con SS si se agrega un punto?
- SS aumenta por cada nuevo punto

Concepto

Una mejor medida de la variabilidad

- $SS = \sum (y \bar{y})^2$, n = 11
- i Que pasa con SS si se agrega un punto?
- SS aumenta por cada nuevo punto

•
$$MS = \frac{\sum (y - \bar{y})^2}{n}$$

Concepto

Una mejor medida de la variabilidad

•
$$SS = \sum (y - \bar{y})^2$$
, $n = 11$

- *j* Que pasa con *SS* si se agrega un punto?
- SS aumenta por cada nuevo punto

•
$$MS = \frac{\sum (y - \bar{y})^2}{n}$$

• Desviación cuadrática media (Mean square deviation MS)

Grados de libertad

Organización del taller

Introducción

meroducció

Todo varia Concepto

Grados de libertad

Otros aspects

D: ~

experiment

Otros

Otros conceptos • Muestra de 5 números: $\bar{y}=4$

Grados de libertad

Introducción

Varianza

Todo varia Concepto Grados de libertad

Definición Otros aspecto

Diseño

experiment

Otros conceptos

2				
---	--	--	--	--

Introducción

IIIIIoduccioi

Todo varia Concepto Grados de

Definición

Otros aspecto

experimenta

Otros conceptos

2 7			
-----	--	--	--

Introducción

Todo varia Concepto Grados de

Definición
Otros aspecto

Diseño

experimenta

Otros

2	7	4		
---	---	---	--	--

Introducción

Todo varia Concepto Grados de

Definición
Otros aspecto

Otros aspecto

experimenta

experiment

Otros conceptos

2 7	4	0	
-----	---	---	--

Introducción

Todo varia Concepto Grados de libertad

Definición Otros aspecto

Diseño

experimenta

Otros

2	7	4	0	7
---	---	---	---	---

Grados de libertad

Organización del taller

Introducción

Varianza Todo varia Concepto Grados de libertad Definición

Diseño

experimenta

Otros conceptos • Muestra de 5 números: $\bar{y}=4$, $\sum y=20$

2	7	4	0	7
---	---	---	---	---

ullet Total libertad en la selección de números 1-4

Introducción

\/------

Todo varia Concepto Grados de libertad

Definición Otros aspecto

Diseño experimenta

Otros

Otros conceptos

2	7	4	0	7
---	---	---	---	---

- Total libertad en la selección de números 1-4
 - \Rightarrow 4 grados de libertad (degrees of freedom d.f.)

Introducción

Todo varia Concepto Grados de libertad Definición

Diseño experiment:

Otros

Otros conceptos

2	7	4	0	7
---	---	---	---	---

- Total libertad en la selección de números 1-4 $\Rightarrow 4$ grados de libertad (degrees of freedom d.f.)
- df = n p

2	7	4	0	7
---	---	---	---	---

- Total libertad en la selección de números 1 − 4
 ⇒ 4 grados de libertad (degrees of freedom d.f.)
- df = n p
- n= número de muestras, p= número de parámetros estimados por el modelo

Medida de la variabilidad

Organización del taller

Introducción

Varianza

Todo varia Concepto Grados de

Definición

Otros aspes

Dicoño

experimenta

Otuna

Otros conceptos • $MS = \frac{\sum (y - \bar{y})^2}{n}$

Varianza (1) Medida de la variabilidad

•
$$MS = \frac{\sum (y-\bar{y})^2}{n}$$

• No se puede calcular MS antes de conocer \bar{y}

Definición

Varianza (1) Medida de la variabilidad

•
$$MS = \frac{\sum (y-\bar{y})^2}{n}$$

- No se puede calcular MS antes de conocer \bar{y}
- *i* De donde se obtiene \bar{y} ?

Varianza (1)

Medida de la variabilidad

Organización del taller

Introducció

Varianza

Concepto Grados o

Definición

Otros aspect

Diseño

experimenta

Otros conceptos • $MS = \frac{\sum (y-\bar{y})^2}{n}$

- ullet No se puede calcular MS antes de conocer $ar{y}$
- ¿De donde se obtiene \bar{y} ?
- ullet g es un parámetro estimado de los datos

Definición

Varianza (1) Medida de la variabilidad

•
$$MS = \frac{\sum (y-\bar{y})^2}{n}$$

- No se puede calcular MS antes de conocer \bar{y}
- *i* De donde se obtiene \bar{y} ?
- \bar{y} es un parámetro estimado de los datos
- Se pierde un grado de libertad

Todo varia Grados de

Definición

Varianza (2)

Formalización y definición

Medida cuantitativa de la variabilidad:

Varianza (2) Formalización y definición

Medida cuantitativa de la variabilidad:

$$Varianza = \frac{Suma de cuadrados}{Grados de libertad} = \frac{SS}{df}$$

Varianza (2) Formalización y definición

Medida cuantitativa de la variabilidad:

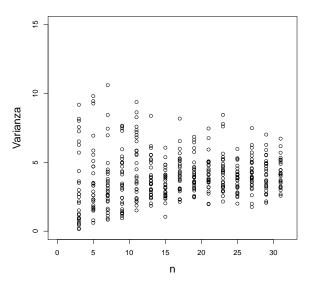
$$Varianza = \frac{Suma de cuadrados}{Grados de libertad} = \frac{SS}{df}$$

$$s^2 = \frac{\sum (y - \bar{y})^2}{n - 1}$$

Otros aspectos

Varianza y tamaño de muestra

Media: 10, Varianza: 4



Todo varia Grados de

Otros aspectos

Una medida de fiabilidad

• ¿Fiabilidad de estimaciones cuando $s^2 \nearrow$?

Todo varia Grados de

Otros aspectos

- ¿Fiabilidad de estimaciones cuando $s^2 \nearrow$?
- Fiabilidad $\propto s^2$

Todo varia

Otros aspectos

- ¿Fiabilidad de estimaciones cuando $s^2 \nearrow$?
- Fiabilidad $\propto s^2$
- ¿Y qué tal del tamaño de la muestra?

Todo varia

Otros aspectos

- ¿Fiabilidad de estimaciones cuando $s^2 \nearrow$?
- Fiabilidad $\propto s^2$
- ¿Y qué tal del tamaño de la muestra?
- Fiabilidad $\propto \frac{s^2}{n}$

Otros aspectos

- ¿Fiabilidad de estimaciones cuando $s^2 \nearrow ?$
- Fiabilidad $\propto s^2$
- ¿Y qué tal del tamaño de la muestra?
- Fiabilidad $\propto \frac{s^2}{n}$
- Qué son las unidades?

Una medida de fiabilidad

¡Error estándar de la media!

- ¿Fiabilidad de estimaciones cuando $s^2 \nearrow$?
- Fiabilidad $\propto s^2$
- ¿Y qué tal del tamaño de la muestra?
- Fiabilidad $\propto \frac{s^2}{n}$
- Qué son las unidades?
- $SE_{\bar{y}} = \sqrt{\frac{s^2}{n}}$

Todo varia Grados de

Otros aspectos

Intervalos de confianza

Muestreo repetido → rango de valores

Todo varia

Otros aspectos

- Muestreo repetido → rango de valores
- Intervalo de confianza ∝ Fiabilidad

Todo varia

Otros aspectos

- Muestreo repetido → rango de valores
- Intervalo de confianza ∝ Fiabilidad
- Distribución t de Student

Otros aspectos

- Muestreo repetido → rango de valores
- Intervalo de confianza ∝ Fiabilidad
- Distribución t de Student
- Nivel de confianza α y grados de libertad df

Otros aspectos

- Muestreo repetido → rango de valores
- Intervalo de confianza ∝ Fiabilidad
- Distribución t de Student
- Nivel de confianza α y grados de libertad df
- Número de errores estándar que se espera

Introducció

Varianza

Concepto Grados d libertad

Otros aspectos

Diseño

experiment

Otros

Muestreo repetido → rango de valores

- Distribución t de Student
- ullet Nivel de confianza lpha y grados de libertad df
- Número de errores estándar que se espera
- $CI_{95\%} = \bar{y} \pm t_{\alpha, df} \sqrt{\frac{s^2}{n}}$

Introducción a la estadística

Organización del taller

Introducción

Diseño

Diseño experimental

Seudoreplicación Aleatorización Controles

Otros

Diseño experimental

Conceptos claves

Replicación: aumenta fiabilidad

- Si replican y randomizan correctamente, ¡no hay problema!
- Diseño inadecuado \longrightarrow buenos resultados

Introducción a la estadística

Organización del taller

Introducción

Diseño

Diseño experimental

Seudoreplicación Aleatorización Controles

Otros

Diseño experimental

Conceptos claves

Replicación: aumenta fiabilidad

- Si replican y randomizan correctamente, ¡no hay problema!

Introducción a la estadística

Organización del taller

Introducción

Diseño

Diseño experimental

Seudoreplicación Aleatorización Controles

Otros

Diseño experimental

Conceptos claves

Replicación: aumenta fiabilidad

- Si replican y randomizan correctamente, ¡no hay problema!

Introducción a la estadística

Organización del taller

IIILIOGUCCIOI

Diseño experimental

Replicación Seudoreplicación Aleatorización Controles

Otros

Diseño experimental

Conceptos claves

Replicación: aumenta fiabilidad

- Si replican y randomizan correctamente, ¡no hay problema!
- Diseño inadecuado buenos resultados

Replicación Seudoreplicación

Replicación

- Permite aumentar la fiabilidad y cuantificar la variabilidad dentro de un tratamiento

Organización del taller

Introduccion

Diseño

experiment

Replicación Seudoreplicación Aleatorización

Controles Inferiencia

- Permite aumentar la fiabilidad y cuantificar la variabilidad dentro de un tratamiento
- Medidas repetidas deben:
 - Ser independientes (individuos distintos)
 - No formar una serie temporal
 - No estar agrupadas juntas en un lugar
 - Tener escala espacial adecuada

Introducción

Varianza

Diseño experiment

Replicación

Seudoreplicación Aleatorización Controles

Inferiencia

- Permite aumentar la fiabilidad y cuantificar la variabilidad dentro de un tratamiento
- Medidas repetidas deben:
 - Ser independientes (individuos distintos)
 - No formar una serie temporal
 - No estar agrupadas juntas en un lugar
 - Tener escala espacial adecuada

Introducción

Varianza

Diseño experiment

Replicación

Seudoreplicación Aleatorización

Inferiencia

- Permite aumentar la fiabilidad y cuantificar la variabilidad dentro de un tratamiento
- Medidas repetidas deben:
 - Ser independientes (individuos distintos)
 - No formar una serie temporal
 - No estar agrupadas juntas en un lugar
 - Tener escala espacial adecuada

Replicación

- Permite aumentar la fiabilidad y cuantificar la variabilidad dentro de un tratamiento
- Medidas repetidas deben:
 - Ser independientes (individuos distintos)
 - No formar una serie temporal
 - No estar agrupadas juntas en un lugar

Replicación

- Permite aumentar la fiabilidad y cuantificar la variabilidad dentro de un tratamiento
- Medidas repetidas deben:
 - Ser independientes (individuos distintos)
 - No formar una serie temporal
 - No estar agrupadas juntas en un lugar
 - Tener escala espacial adecuada

Seudoreplicación

Replicación (2)

 Idealmente: una réplica de cada tratamiento debe estar agrupada en un bloque y cada tratamiento debe estar repetido en varios bloques

Seudoreplicación

Organización del taller

Introducción

Vaniana.

Diseño experimenta

Replicación Seudoreplicación Aleatorización Controles

Otros

- Tantas como sea posible ©
- ¿Cómo saber?
 - ⇒ Indicación sobre varianza base y magnitud de la respuesta al tratamiento
- Método práctico (en general): ≥ 30

Seudoreplicación

- Tantas como sea posible ©
- ¿Cómo saber? Estudios pilotos y experiencia.

Replicación Seudoreplicación

- Tantas como sea posible ©
- ¿Cómo saber? Estudios pilotos y experiencia

Controles

- Tantas como sea posible ©
- ¿Cómo saber? Estudios pilotos y experiencia
 - ⇒ Indicación sobre varianza base y magnitud de la respuesta al tratamiento

- Tantas como sea posible ©
- ¿Cómo saber? Estudios pilotos y experiencia ⇒ Indicación sobre varianza base y magnitud de la respuesta al tratamiento
- Método práctico (en general): ≥ 30

Seudoreplicación

Poder y réplicas

• Poder: probabilidad de rechazar H_0 cuando es falsa

Replicación Seudoreplicación

- Poder: probabilidad de rechazar H_0 cuando es falsa
- ; Cuantas réplicas para detectar un efecto δ con 80%probabilidad de no cometer un error?

Replicación Seudoreplicación

- Poder: probabilidad de rechazar H_0 cuando es falsa
- ¿Cuantas réplicas para detectar un efecto δ con 80%probabilidad de no cometer un error?
- Experiencia y/o estudio piloto

Seudoreplicación

- Poder: probabilidad de rechazar H_0 cuando es falsa
- ¿Cuantas réplicas para detectar un efecto δ con 80%probabilidad de no cometer un error?
- Experiencia y/o estudio piloto
 - \Rightarrow Primera estimación del efecto δ y de la varianza s^2

- Poder: probabilidad de rechazar H_0 cuando es falsa
- ¿Cuantas réplicas para detectar un efecto δ con 80%probabilidad de no cometer un error?
- Experiencia y/o estudio piloto
 - \Rightarrow Primera estimación del efecto δ y de la varianza s^2

$$n \approx \frac{8 * s^2}{\delta^2}$$

Seudoreplicación

Seudoreplicación

Condición importante: independencia de los errores

- Medidas repetidas del mismo individuo → seudoreplicación temporal
- Varias medidas del mismo lugar → seudoreplicación spacial
- ¿Cuántos grados de libertad?

Organización del taller

Introducción

Diseño experimenta

Seudoreplicación Aleatorización

Controles Inferiencia

Otros conceptos

¿Qué hacer con seudoreplicación?

- Promediar seudoreplicación y hacer análisis sobre medias
- Hacer análisis separados por cada período de tiempo
- Usar análisis de series de tiempo o modelos de efectos mixtos

Introducción

Varianza

Diseño experimenta

Replicación Seudoreplicació

Aleatorización Controles

Controles Inferiencia

- ¿Cómo seleccionar un árbol al azar en una selva?
- ¿Hojas accesibles?
- ¿Cerca del laboratorio?
- ¿Parece sano?
- ¡Sin insectos?
- ⇒ ¡Sesgo en la fotosíntesis!

Introducción

Diseño

experimenta

Seudoreplicación Aleatorización

Controles

Inferiencia

- ¿Cómo seleccionar un árbol al azar en una selva?
- ¿Hojas accesibles?
- ¿Cerca del laboratorio?
- ¿Parece sano?
- ¿Sin insectos?
- ⇒ ¡Sesgo en la fotosíntesis!

Introduccion

Diseño experiment

experimenta Replicación

Seudoreplicación Aleatorización

Controles Inferiencia

- ¿Cómo seleccionar un árbol al azar en una selva?
- ¿Hojas accesibles?
- ¿Cerca del laboratorio?
- ¿Parece sano?
- ¿Sin insectos?
- ⇒ ¡Sesgo en la fotosíntesis!

Introducción

Diseño

experiment Replicación

Seudoreplicación Aleatorización

Controles Inferiencia

- ¿Cómo seleccionar un árbol al azar en una selva?
- ¿Hojas accesibles?
- ¿Cerca del laboratorio?
- ¿Parece sano?
- ¿Sin insectos?
- ⇒ ¡Sesgo en la fotosíntesis!

Aleatorización

Organización del taller

Introducción

Diseño experiment

Replicación Seudoreplicación Aleatorización

Controles

Inferiencia

- ¿Cómo seleccionar un árbol al azar en una selva?
- ¿Hojas accesibles?
- ¿Cerca del laboratorio?
- ¿Parece sano?
- ¿Sin insectos?
- ⇒ ¡Sesgo en la fotosíntesis

Aleatorización

Organización del taller

Introduccion

Diseño

Replicación Seudoreplicación Aleatorización

Controles Inferiencia

- ¿Cómo seleccionar un árbol al azar en una selva?
- ¿Hojas accesibles?
- ¿Cerca del laboratorio?
- ¿Parece sano?
- ¿Sin insectos?
- ⇒ ¡Sesgo en la fotosíntesis!

Selección aleatoria de un árbol

Organización

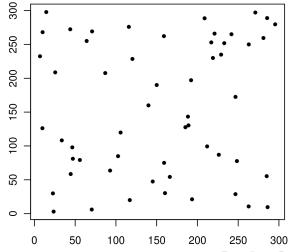
Introducción

Diseño

experiment: Replicación

Seudoreplicación Aleatorización

Controles



Selección aleatoria de un árbol

Organización del taller

Introducción

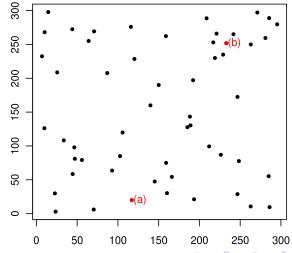
Diseño

Diseno experiment

Replicación Seudoreplicación Aleatorización

Controles

Otros



Selección aleatoria de un árbol

Organización del taller

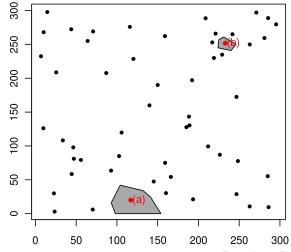
Introducción

Disco

Diseno experiment

Replicación Seudoreplicación Aleatorización

Controles



Organización del taller

Introducción

Diseño

Replicación Seudoreplicación Aleatorización

Controles

Inferienci

Otros

• No controles, no conclusiones

Replicación Seudoreplicación Inferiencia

¿Cuánto tiempo?

- Idealmente: determinar duración por adelantado

Seudoreplicación Inferiencia

¿Cuánto tiempo?

- Idealmente: determinar duración por adelantado
- NO seguir experimento hasta que se obtenga un "buen" resultado

Seudoreplicación

Inferiencia

Inferencia fuerte

- Formular una hipótesis clara
- Diseñar un test aceptable
- Sin replicación, aleatorización y controles, no hay progreso

Modelaje probabilidad

- Datos: lo que pasó

probabilidad

- Datos: lo que pasó
- Descripción → patrones → mecanismos

Modelaje probabilidad

- Datos: lo que pasó
- Descripción → patrones → mecanismos
- Modelo para explicar y predecir

- Datos: lo que pasó
- Descripción → patrones → mecanismos
- Modelo para explicar y predecir
- Varios (muchos) modelos están ajustados a los datos

- Datos: lo que pasó
- Descripción → patrones → mecanismos
- Modelo para explicar y predecir
- Varios (muchos) modelos están ajustados a los datos
- → Modelo mínimo y adecuado

probabilidad

Modelaje estadístico

Mínimo: Suficientemente simple

probabilidad

Modelaje estadístico

Mínimo: Suficientemente simple

Adecuado: ¿Por qué usar modelo que no describe los

datos?

Modelaje estadístico

Mínimo: Suficientemente simple

Adecuado: ¿Por qué usar modelo que no describe los

datos?

Mejor modelo: La menor proporción de varianza que no sea

explicada (desviación residual mínima)

Parsimonia probabilidad

La navaja de Occam

Principio de parsimonia

- Con varias explicaciones igualmente válidas
- Correcta: la más simple

probabilidad

La navaja de Occam

Principio de parsimonia

- Con varias explicaciones igualmente válidas
- Correcta: la más simple

La navaja de Occam

Principio de parsimonia

- Con varias explicaciones igualmente válidas
- Correcta: la más simple

- Tan pocos parámetros como sea posible
- Modelos lineales > no lineales
- Pocas condiciones > muchas
- Pocas variables > muchas
- 1 explicación simple > varias explicaciones complicadas

La navaja de Occam

Principio de parsimonia

- Con varias explicaciones igualmente válidas
- Correcta: la más simple

- Tan pocos parámetros como sea posible
- Modelos lineales > no lineales
- Pocas condiciones > muchas
- Pocas variables > muchas
- 1 explicación simple > varias explicaciones complicadas

La navaja de Occam

Principio de parsimonia

- Con varias explicaciones igualmente válidas
- Correcta: la más simple

- Tan pocos parámetros como sea posible
- Modelos lineales > no lineales
- Pocas condiciones > muchas
- Pocas variables > muchas
- 1 explicación simple > varias explicaciones complicadas

La navaja de Occam

Principio de parsimonia

- Con varias explicaciones igualmente válidas
- Correcta: la más simple

- Tan pocos parámetros como sea posible
- Modelos lineales > no lineales
- Pocas condiciones > muchas
- Pocas variables > muchas
- 1 explicación simple > varias explicaciones complicadas

La navaja de Occam

Principio de parsimonia

- Con varias explicaciones igualmente válidas
- Correcta: la más simple

- Tan pocos parámetros como sea posible
- Modelos lineales > no lineales
- Pocas condiciones > muchas
- Pocas variables > muchas
- 1 explicación simple > varias explicaciones complicadas

Parsimonia probabilidad La navaja de Einstein

Einstein: "Un modelo debe ser tan simple como posible. Pero no más simple"

Máxima probabilidad

Máximo de verosimilitud

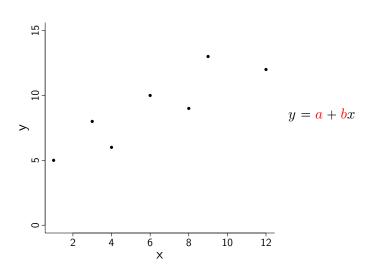
(Maximum Likelihood: ML)

- Dado los datos
- Y dado un modelo
- ¿ Qué valores de parámetros hacen a los datos observados más probables?
- ⇒ Estimadores sin sesgo que minimizan la varianza

Máxima

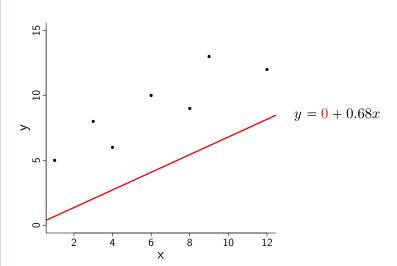
probabilidad

Máximo de verosimilitud

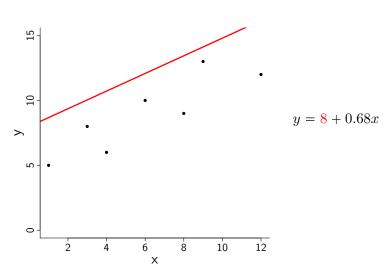


Máxima probabilidad

Máximo de verosimilitud

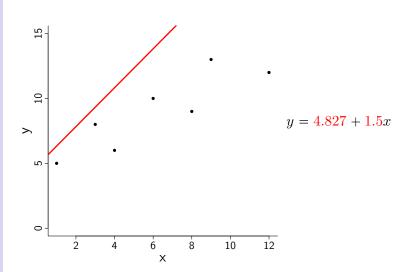


Máximo de verosimilitud



Máxima probabilidad

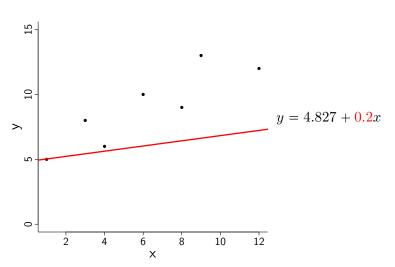
Máximo de verosimilitud



Máximo de verosimilitud

Modelaje

Máxima probabilidad



Máxima probabilidad

Máximo de verosimilitud

