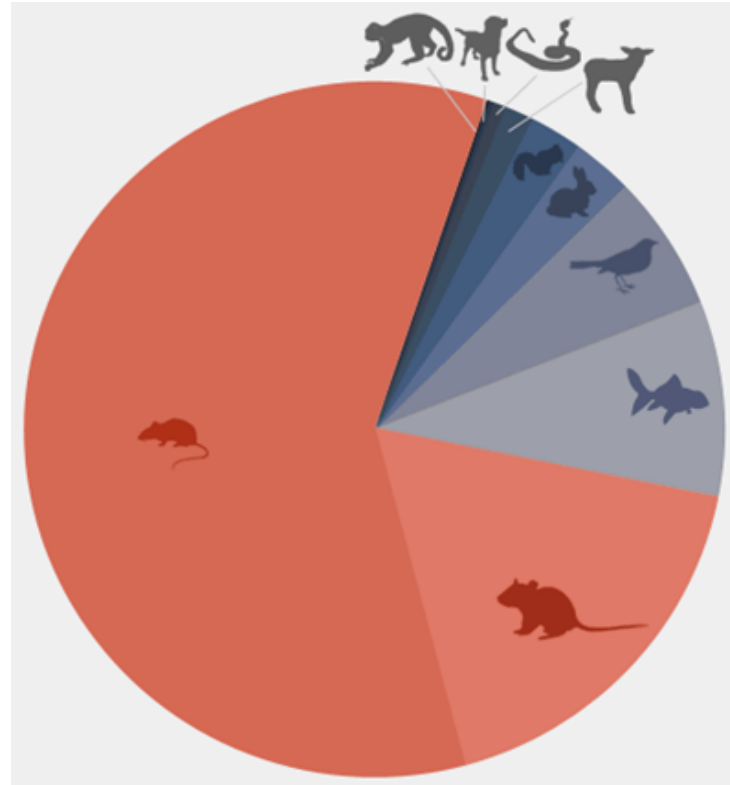
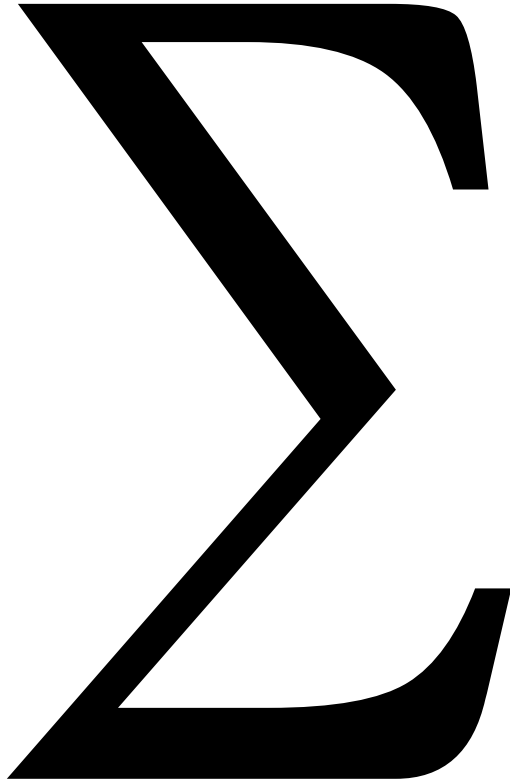


La R de Reducción: Aproximación práctica al cálculo de tamaños muestrales en experimentación con animales



Reducción en estudios diseñados para ser científicamente válidos

El uso de métodos que permitan a los investigadores:

1. Obtener niveles comparables de información a partir de menos animales
2. Obtener más información (cantidad y precisión) con el mismo número de animales

Con las consideraciones:

1. Reducción del número de animales no debe aplicarse a expensas de un mayor sufrimiento a los animales individuales
2. Si la reducción del número de animales hace imposible llegar a una conclusión válida del experimento, esto NO logra el objetivo de Reducción

Procedimientos para la Reducción:

1. Compartir datos y recursos (animales y equipos) entre grupos y organizaciones
2. Uso de tecnología como imagen en estudios longitudinales

.....

- t. Diseño de experimentos y utilización de técnicas estadísticas

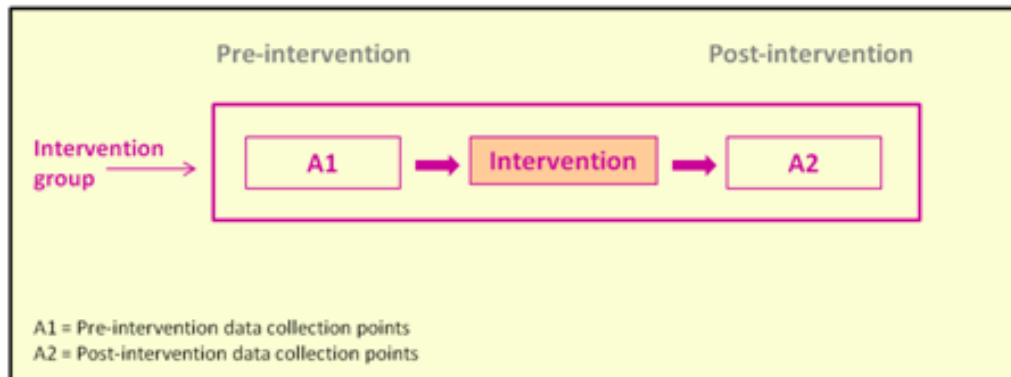
Ítems

1. Organizar las hipótesis, diseños y contrastes estadísticos
2. Utilizar los efectos correctos y definir tamaño del efecto (ES)
3. Extraer y organizar los efectos de investigaciones previas
4. Seleccionar el alpha adecuado
5. Seleccionar el nivel de potencia como elemento de Reducción
6. Calcular tamaños muestrales
7. Otros factores en el muestreo
8. Ejemplos
9. Software

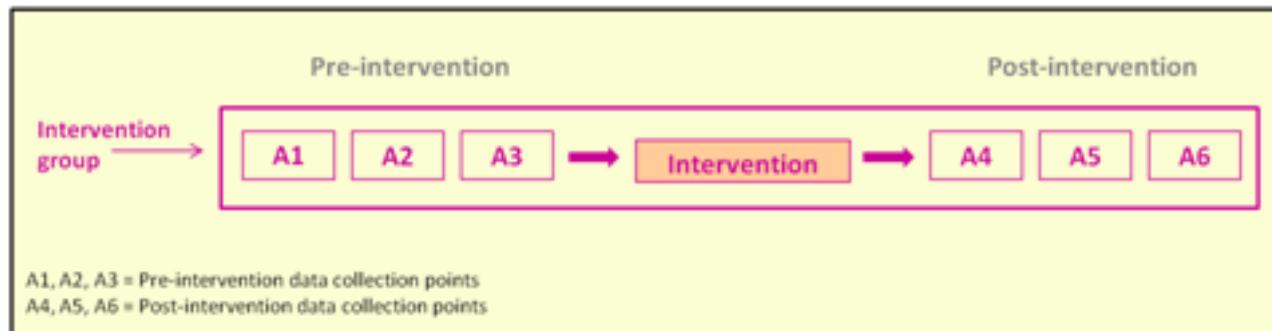
1. Organizar las hipótesis, diseños y contrastes estadísticos

Diseño no experimental o de intervención: no hay grupos de comparación

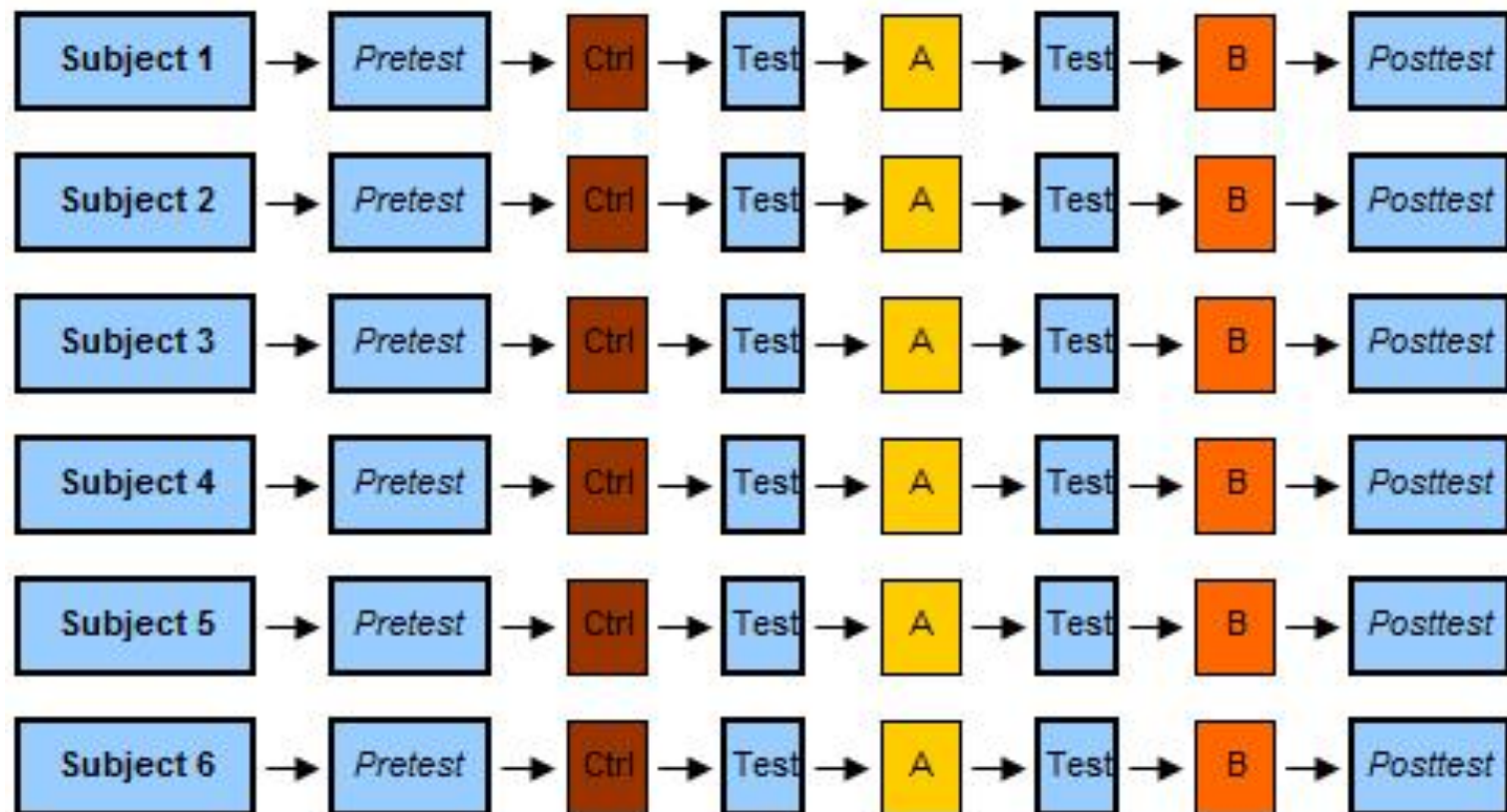
Non-Experimental Pre-Test/Post-Test Design



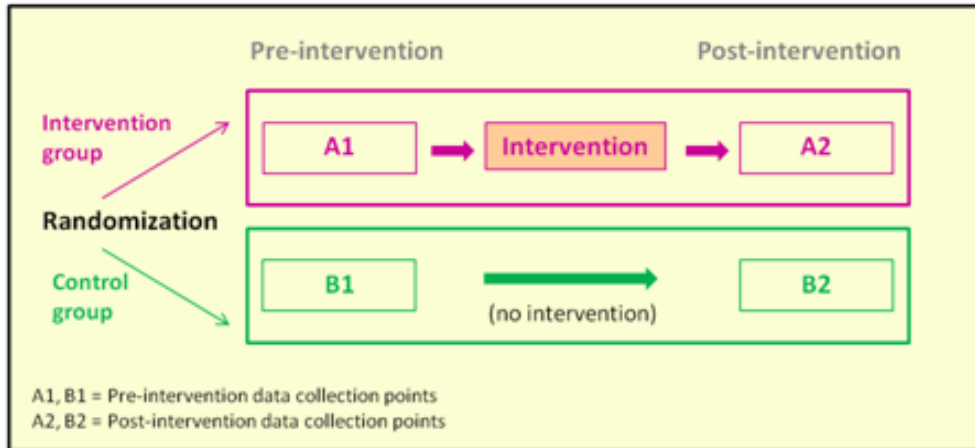
Non-Experimental Time Series Design



WITHIN SUBJECT DESIGN (REPEATED MEASURES)

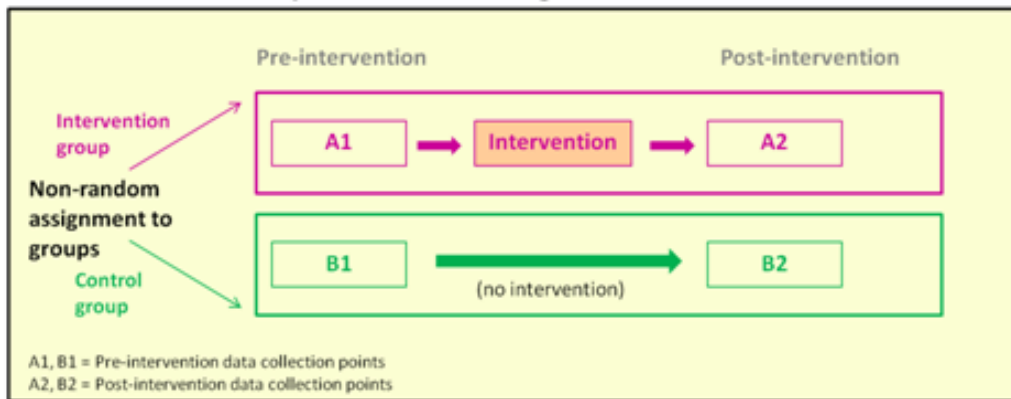


Classical Design of Randomized Experiments



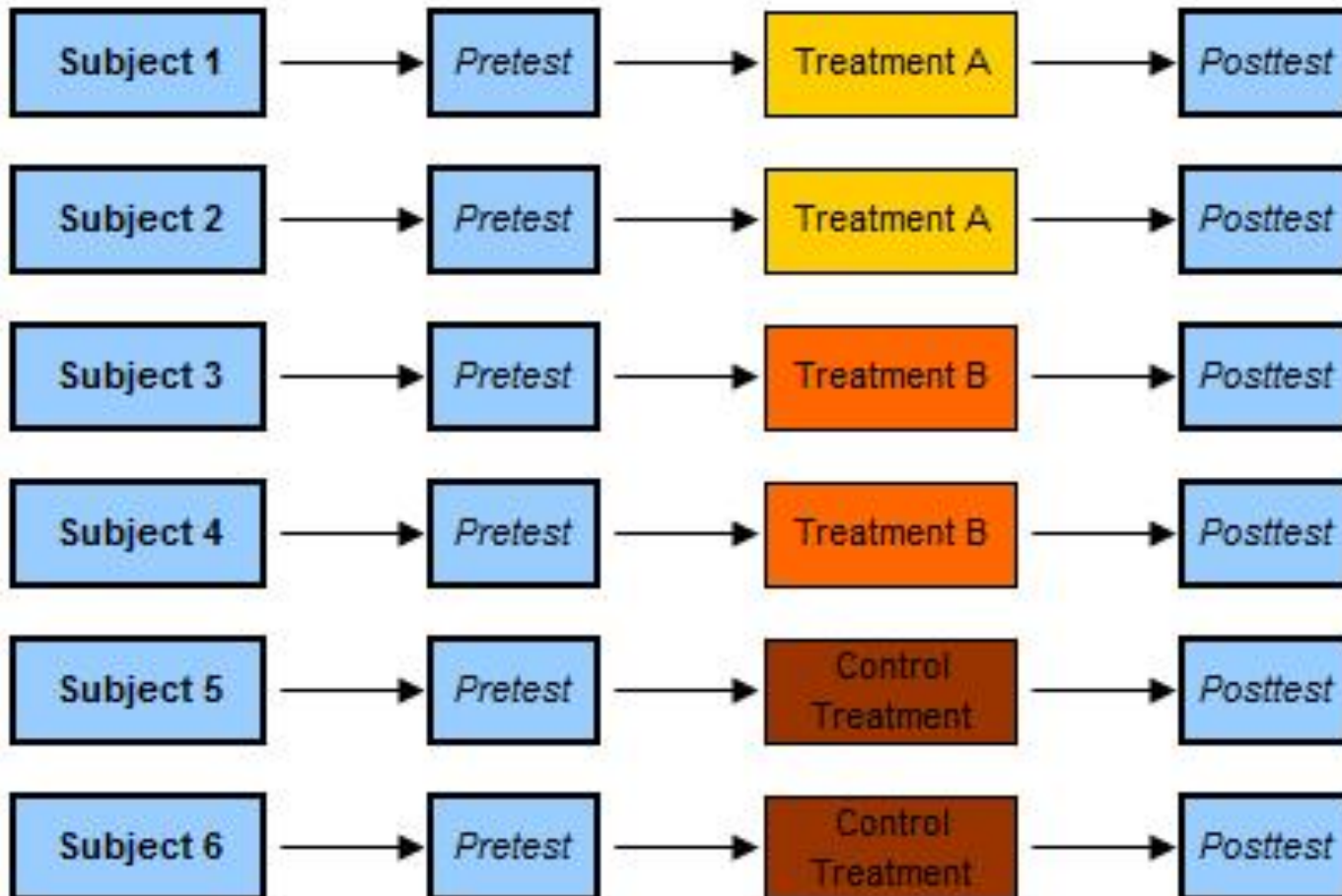
Diseños completamente aleatorizados: “gold standard”

Classical Quasi-Experimental Design

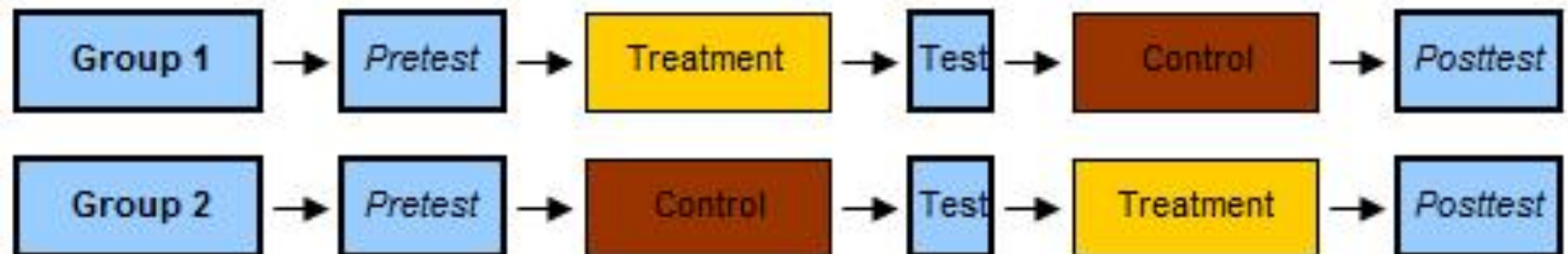


No completamente aleatorizados pero controlados

BETWEEN SUBJECTS DESIGN

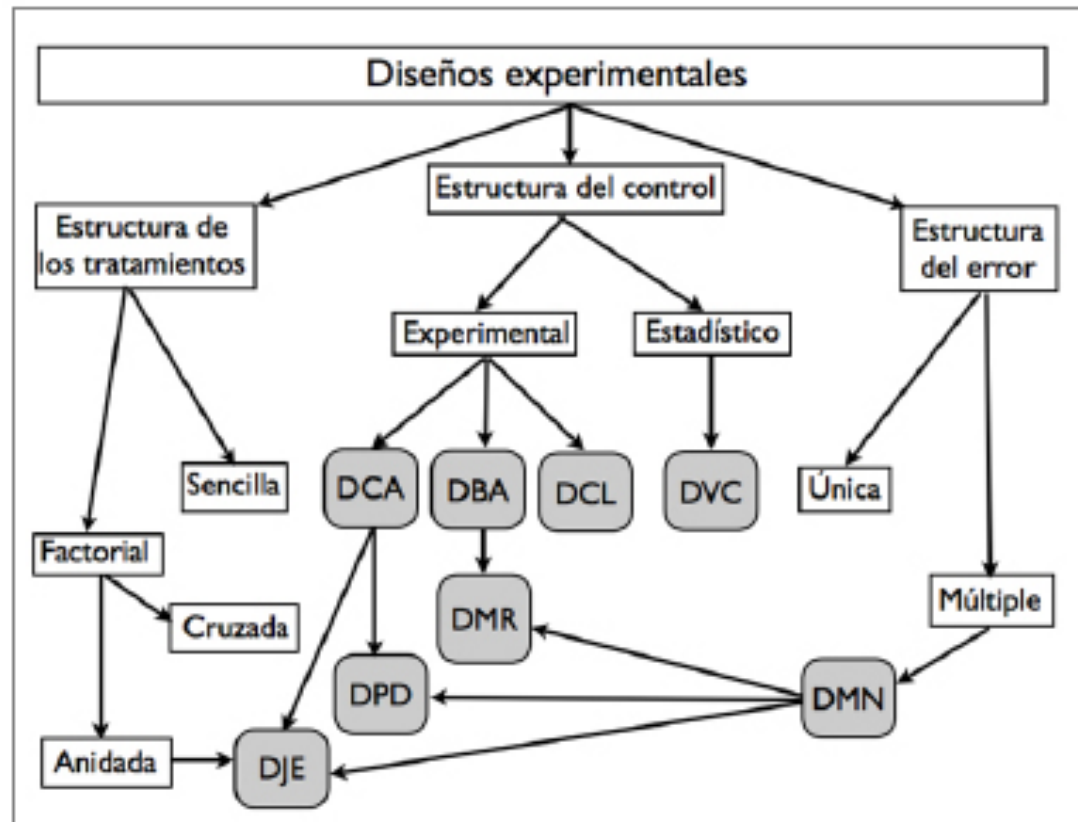


CROSSOVER - REPEATED MEASURES DESIGN



or





DCA: Diseño completamente aleatorio. DBA: Diseño de bloques aleatorios. DCL: Diseño de cuadrado latino. DVC: Diseño con variables concomitantes. DMR: Diseño de medidas repetidas. DPD: Diseño de parcela dividida. DJE: Diseño jerárquico. DMN: Diseño multinivel.

Transversal o Longitudinal

Contraste de hipótesis estadísticas

Prueba o test estadístico, usado para tomar una decisión acerca de la hipótesis sobre la evidencia que proporciona la muestra

Esta hipótesis se llama hipótesis nula H_0 y se contrasta frente a una hipótesis alternativa H_1

$$\begin{cases} H_0: \text{parámetro } \theta \text{ toma uno o varios valores} \\ H_1: \text{parámetro } \theta \text{ toma otro u otros valores} \end{cases}$$

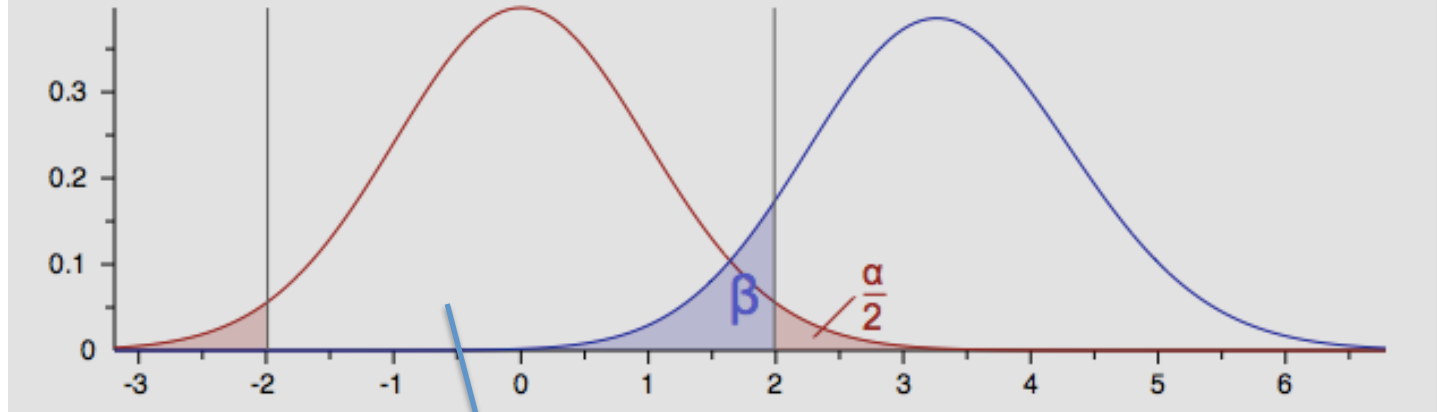
Tipos de errores en el contraste estadístico de hipótesis

Situación verdadera → Decisión tomada ↓	H_0 HIPÓTESIS CIERTA	H_0 HIPÓTESIS FALSA
RECHAZAR H_0	<i>Error tipo I, α</i> <i>$P(\text{rechazar } H_0/H_0 \text{ cierta})$</i>	<i>Acierto</i> <i>$P(\text{rechazar } H_0/H_0 \text{ falsa})$</i>
ACEPTAR H_0	<i>Acierto</i>	<i>Error tipo II, β</i> <i>$P(\text{aceptar } H_0/H_0 \text{ falsa})$</i>

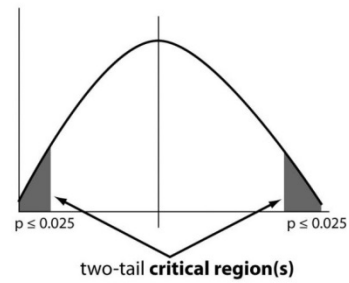
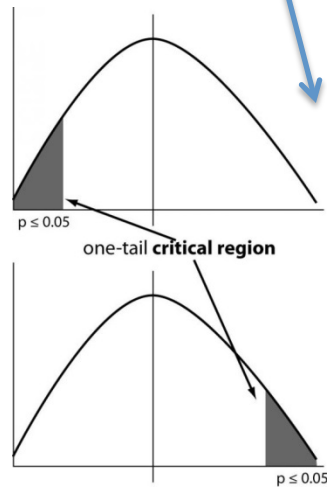
α no es estrictamente el complementario de β

POTENCIA: Probabilidad de detectar un efecto que realmente existe

$$\alpha, \text{ Power} = 100(1 - \beta) \%$$



$$Z_{1-\alpha} < Z_{1-\frac{\alpha}{2}}$$



2. Utilizar los efectos correctos y definir tamaño del efecto (ES)

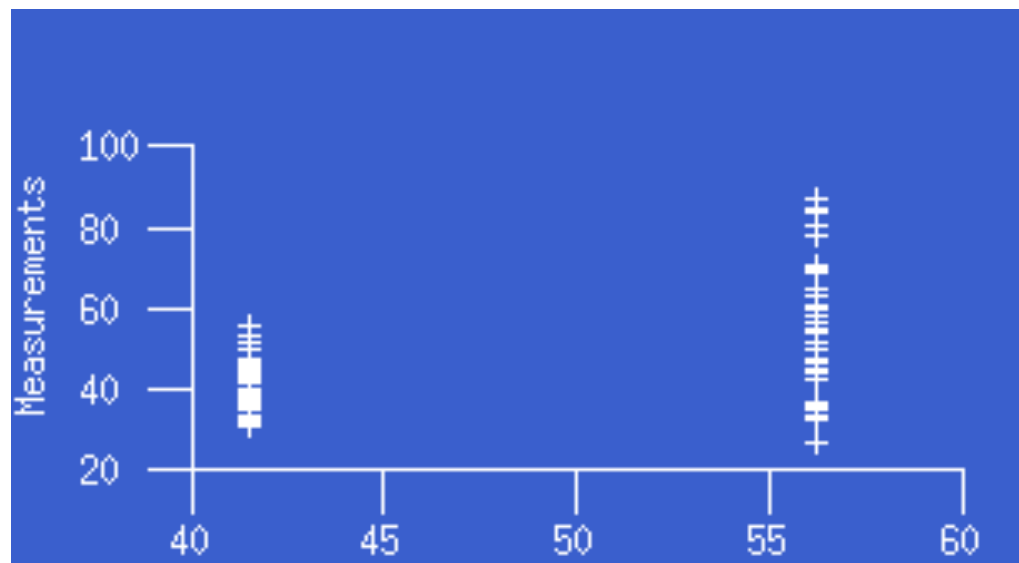
Es una medida de la diferencia entre H_0 y H_1

Su elección depende:

- ✓ El contexto teórico de la investigación
- ✓ Investigaciones similares previas
- ✓ Relación coste/beneficio en términos de resultados científicos

$$d = \frac{(\mu_1 - \mu_2)}{\sigma}$$

$$d_z = \frac{|\mu_{x-y}|}{\sigma_{x-y}}$$



$$f = \frac{\sqrt{\sum_{j=1}^k n_j (\mu_j - \bar{\mu})^2}}{N \sigma}$$

	<i>Frequencies</i>	
	Success	Failure
Treatment Group	a	b
Control Group	c	d

Odds-Ratio

$$w = \frac{ad}{bc}$$

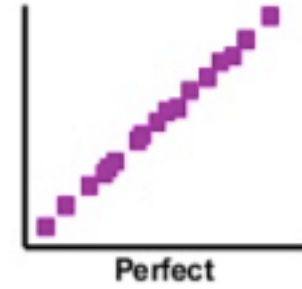
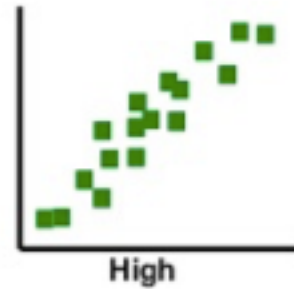
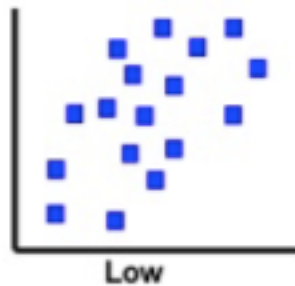
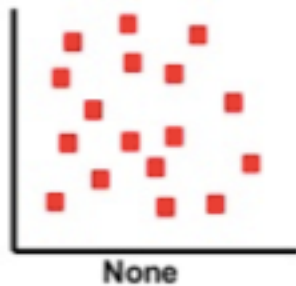
The Risk Ratio

$$w = \frac{a/(a+b)}{c/(c+d)}$$

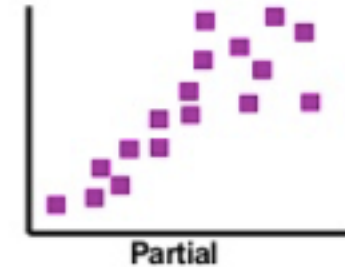
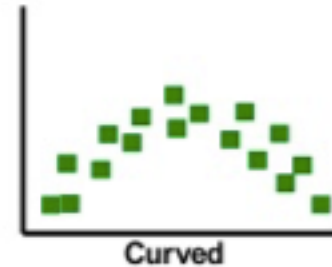
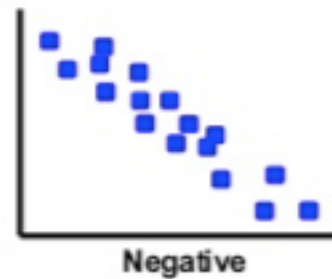
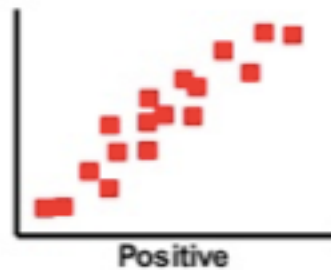
3. Extraer los efectos de investigaciones previas

Correlación $r = \rho$

Degrees of correlation:



Types of correlation:



Tamaño de los efectos: aproximando la hipótesis

Cohen (1977,1988) recomienda los siguientes niveles para tamaños de efectos:

	Index	small	medium	large
<u>t-Test on Means</u>	d	0.20	0.50	0.80
<u>t-Test on Correlations</u>	r	0.10	0.30	0.50
<u>F-Test (ANOVA)</u>	f	0.10	0.25	0.40
<u>F-Test (MCR)</u>	f ²	0.02	0.15	0.35
<u>Chi-Square Test</u>	w	0.10	0.30	0.50

- small effect: ¿clínicamente relevante?
- large effect: ¿demasiado obvio?

4. Seleccionar el alpha adecuado

✓ Error tipo I, α *P(rechazar H_0 / H_0 cierta)*

p-valor=0.002(**) < 0.01 $\longrightarrow$ Rechazar H_0 con 1% error

p<0.05	*
p<0.01	**
p<0.001	***

✓ No es válida cualquier otra consideración en función de que el p-valor sea grande o pequeño

✓ Si no se **rechaza** H_0 **calcular** error de tipo II, β

Selección de alpha

- ✓ Corrección de Bonferroni
- ✓ Alpha ajustado de Hochberg

Test	p	Standard alpha	Bonferroni	Hochberg alpha
	.031	.05 (sig!)	.0125 (not sig!)	.025 (not sig!)
	.02	.05 (sig!)	.0125 (not sig!)	.0167 (sig!)
	.011	.05 (sig!)	.0125 (sig!)	.0125 (sig!)
	.038	.05 (sig!)	.0125 (not sig!)	.05 (sig!)

5. Seleccionar el nivel de potencia como elemento de Reducción

- .80 situación generalmente admitida admitida
- .90 decisiones críticas

a) Si no se controla en el muestreo previo puede ocurrir que los datos sean insuficientes para el contraste

Ejemplo: Un estudio simple, dos grupos Tratamiento y Placebo
El Tratamiento es realmente efectivo. El estudio se repite muchas veces

Una potencia de 0.8 indica que el 80% de las veces que repitamos el experimento obtendremos diferencias estadísticamente significativas pero el 20% no tiene por qué ser así

b) No se pueden interpretar resultados NO significativos sin control de la potencia



No publicar los resultados
Repetir los experimentos
Eleva los tamaños muestrales....

Trabajando con la potencia de los contrastes, análisis:



a) *A priori*. Calcular N dados α, β, ES



b) *Post hoc*. Calcular $1 - \beta$ dado α, N, ES



c) De *compromiso*. Calcular $\alpha, 1 - \beta$ dado $q = \beta / \alpha, N, ES$

6. Calcular tamaños muestrales

$$N > \left(\frac{Z_{1-\frac{\alpha}{2}} \sigma}{d} \right)^2$$

Relación entre tamaño de muestra y error tipo I

• Nivel de confianza: $1 - \alpha \uparrow \rightarrow Z_{1-\frac{\alpha}{2}} \uparrow \rightarrow N \uparrow$

• Varianza muestral: $\sigma \uparrow \rightarrow N \uparrow$

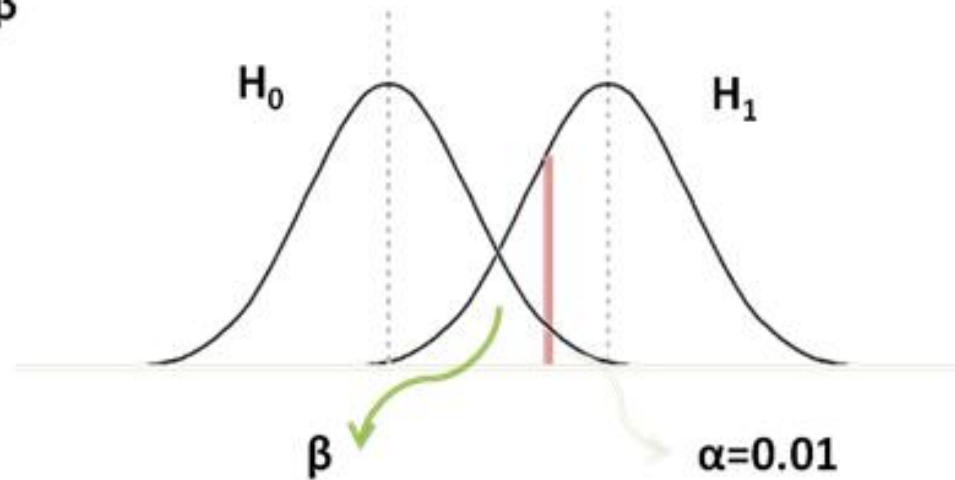
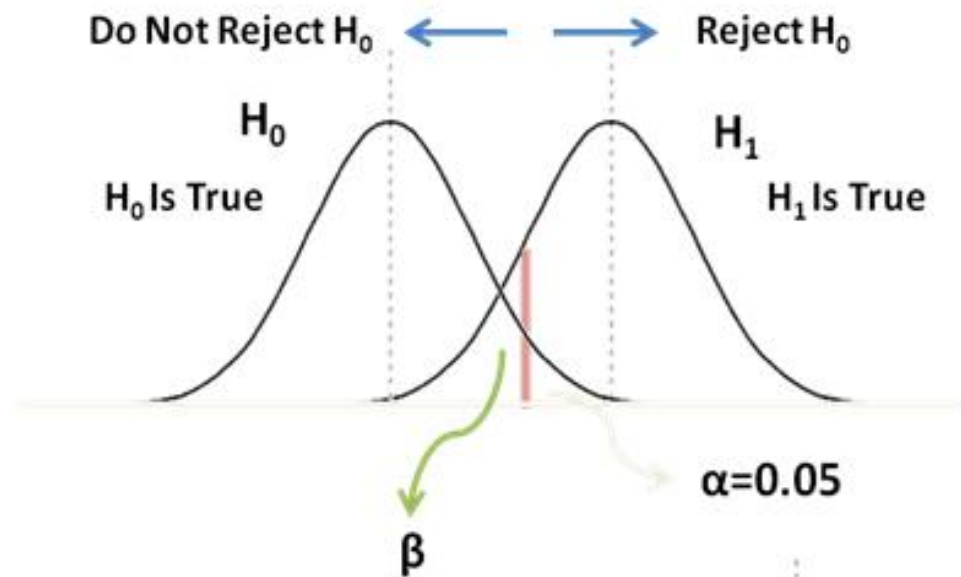
• Tamaño del efecto: $d \downarrow \rightarrow N \uparrow$

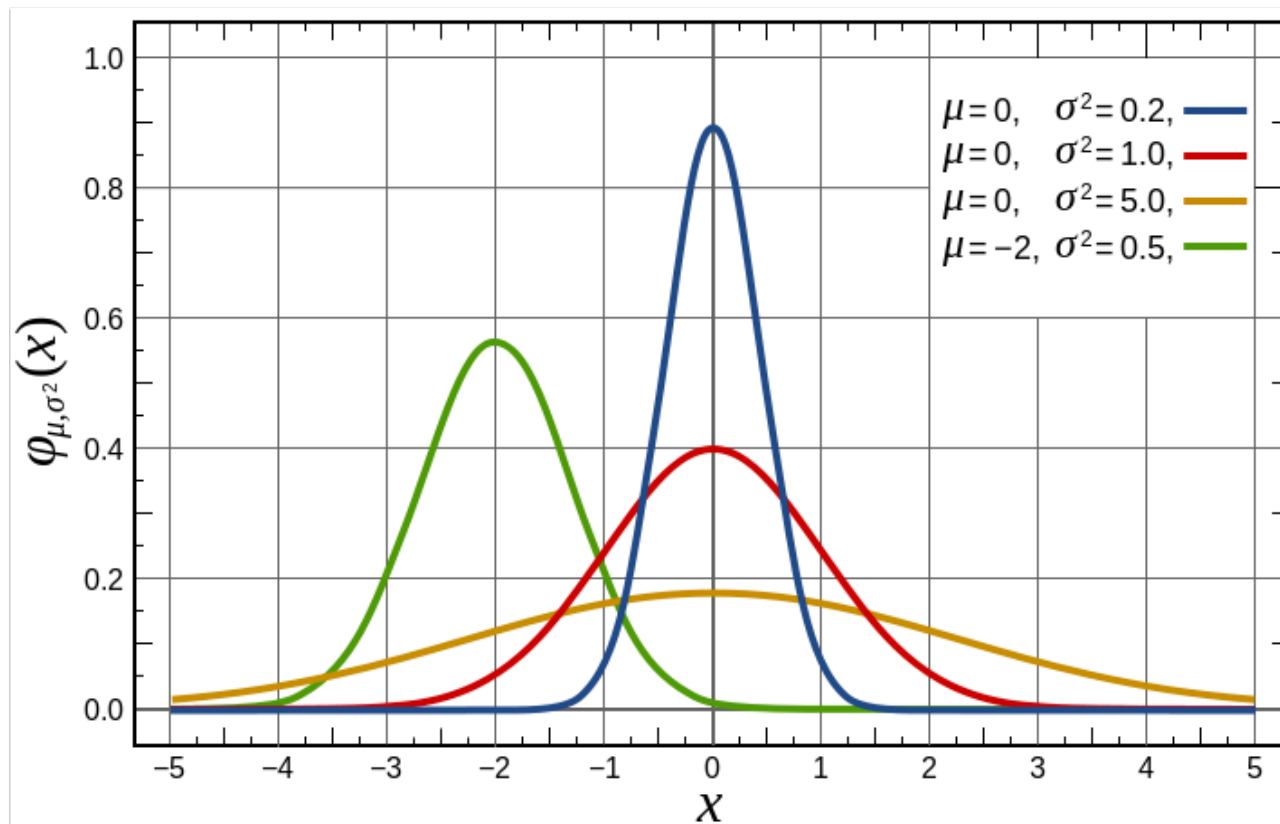
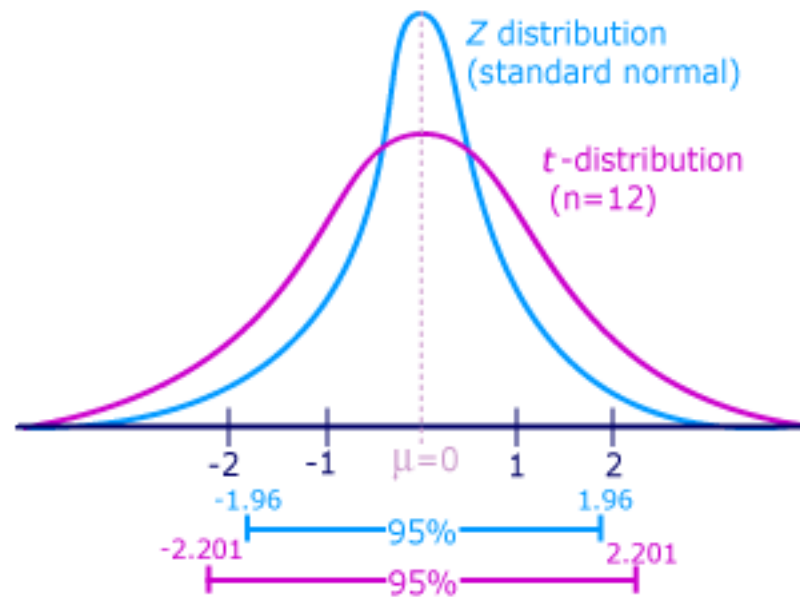
• Dirección del contraste: una o dos colas: $\text{dos colas} \rightarrow N \uparrow$
Para una cola debe reemplazarse el valor Z: $Z_{1-\alpha} < Z_{1-\frac{\alpha}{2}}$

Relación entre tamaño de muestra, error tipo I y error tipo II

$$N > \left(\frac{\left(Z_{1-\frac{\alpha}{2}} + Z_{1-\beta} \right) \sigma}{d} \right)^2$$

- Potencia del test: $1 - \beta \uparrow \rightarrow N \uparrow$





7. Otros factores en el muestreo

- ✓ **Variable respuesta:** utilizar medidas altas en sensibilidad y bajas en error de medida
- ✓ **Diseño experimental:** algunos modelos como medidas repetidas aportan más potencia
- ✓ **Grupos:** el N se reduce bajando el número de contrastes
- ✓ **Procedimiento y modelo estadístico:** p.e. el estudio de interacciones requiere más N para alcanzar suficiente potencia
- ✓ **Modificar la respuesta:** reducir extremos, transformar variables. Categorizar variables lleva a pérdida de potencia
- ✓ **Propósito del estudio:** replicar trabajos anteriores implica un mayor número de muestra. En general se desestima el error de muestreo
- ✓ **Missing data:** en general reducen potencia, incluso en caso de imputación

8. Ejemplos

Comparamos 10 grupos con un tamaño de efecto medio $f=0.25$

¿Cuántos sujetos necesitamos?

1a

Input:

- i. Tipo de análisis: a priori (establecer N)
- ii. Tipo de Test: F-test (ANOVA)
- iii. Alpha: 0.05
- iv. Potencia(1-beta): 0.95
- v. Tamaño efecto f : 0.25

Output:

Total tamaño de muestra: 390

Potencia actual: 0.9525

Supongamos que sólo se dispone de $N=200$

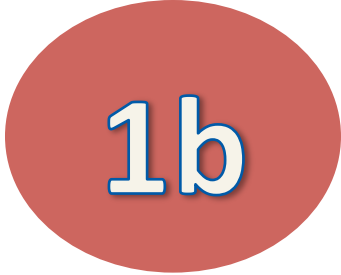
Input:

- i. Tipo de análisis: a posteriori
- ii. Tipo de Test: F-test (ANOVA)
- iii. Tamaño de la muestra: 200
- iv. Tamaño del efecto: 0.25
- v. Ratio beta/alpha: 1

Output:

Alpha: 0.1592

Potencia actual: 0.8408



1b

Supongamos dos poblaciones A y B para compararse sobre la variable aleatoria x de desviación estándar σ en ambas poblaciones

2a

¿Cuántos sujetos necesitamos?

Input:

- i. Tipo de análisis: a priori (establecer N)
- ii. Tipo de Test: t-Test , muestras independientes, dos colas
- iii. Alpha: 0.05
- iv. Potencia(1-beta): 0.90
- v. Tamaño efecto f : dado $\mu_1=10$ $\mu_2=12$ y $\sigma=4$

$$d = \frac{(\mu_1 - \mu_2)}{\sigma} = 0.5$$

Output:

Total tamaño de muestra: 172

Potencia actual: 0.9032

Supongamos que sólo se dispone de $N=140$

2b

Input:

- i. Tipo de análisis: de compromiso
- ii. Tipo de Test: t-Test , muestras independientes, dos colas
- iii. Tamaño de la muestra: $n_1=70$, $n_2=70$
- iv. Tamaño del efecto: 0.50
- v. Ratio beta/alpha: 2

Output:

Alpha: 0.0670

Potencia actual: 0.8408

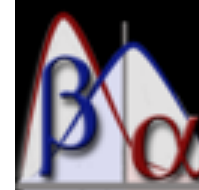
9. Software

En general requieren el mismo tipo de información para diseñar y calcular los muestreos.

- Libres

Gpower <http://www.gpower.hhu.de>

Optimal Design <https://sites.google.com/site/optimaldesignsoftware/home>
(repeated measures, longitudinal y multilevel designs)



- Licenciados

PASS <http://www.ncss.com/index.htm>

SPSS Sample Power, módulo independiente de SPSS. IBM

SAS tiene proc power

Stata tiene el sampsi command, incluyendo fpower, powerreg y aipe



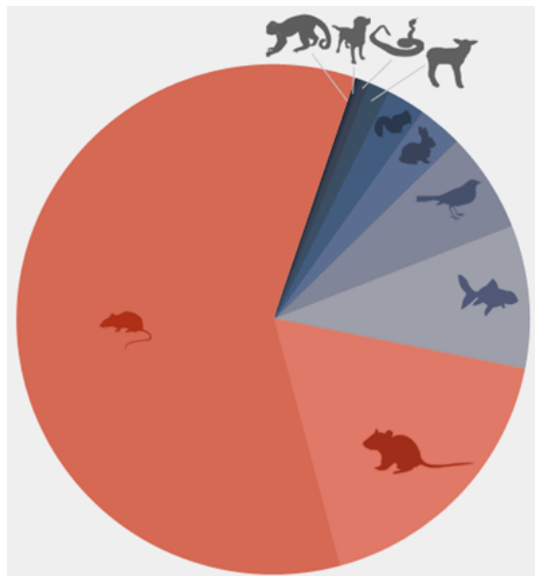
Mplus modelos avanzados y complicados, simulaciones de Monte Carlo

<http://www.statmodel.com/power.shtml>

<http://www.statmodel.com/ugexcerpts.shtml>

Mplus

$$\sum_{i=1}^N$$



$$N \downarrow$$

$$\alpha \downarrow \quad 1 - \beta \uparrow$$