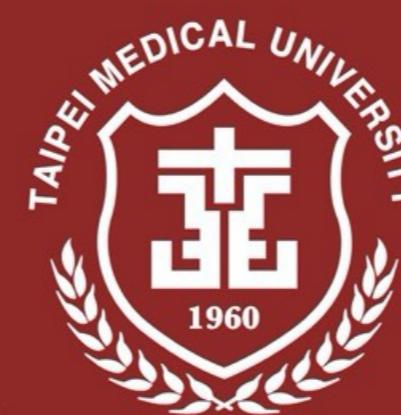


# Psychol. Statistics using R



臺北醫學大學  
TAIPEI MEDICAL UNIVERSITY

## Regression 2

Changwei W. Wu, Ph.D.

---

Graduate Institute of Mind, Brain and Consciousness  
Research Center of Brain and Consciousness  
Taipei Medical University

# Statistics

## 1. Nonlinear Least Squares

- Model fitting of nonlinear relationship

## 2. Partial / Semipartial Correlation

- Comparing multiple correlations

## 3. Causal Mediation Analysis

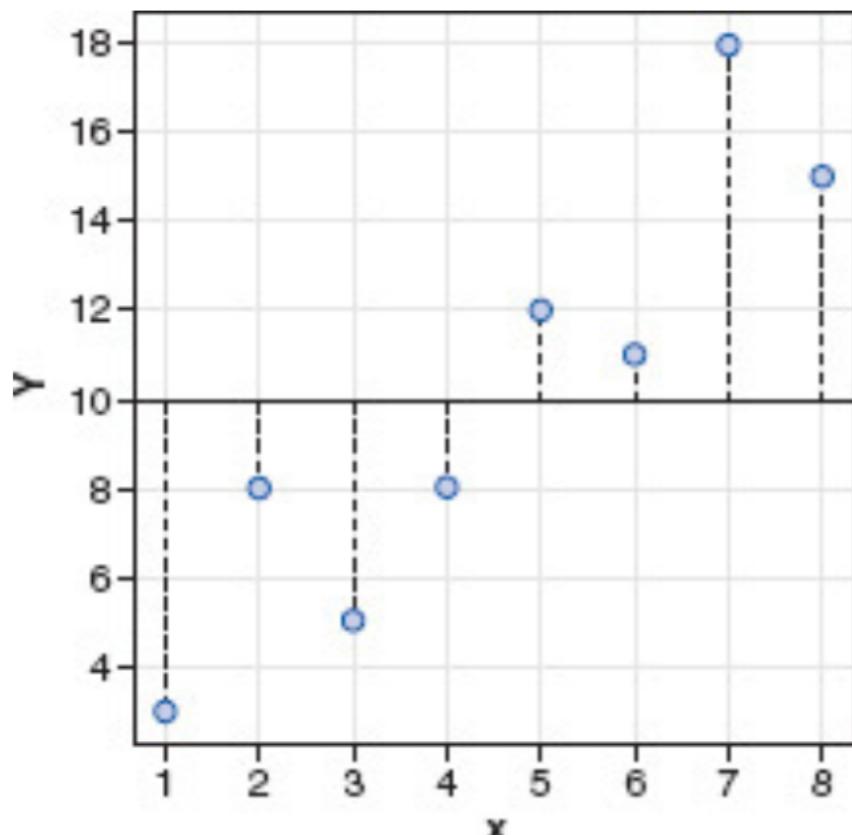
- Indirect effect from the mediator

Theories

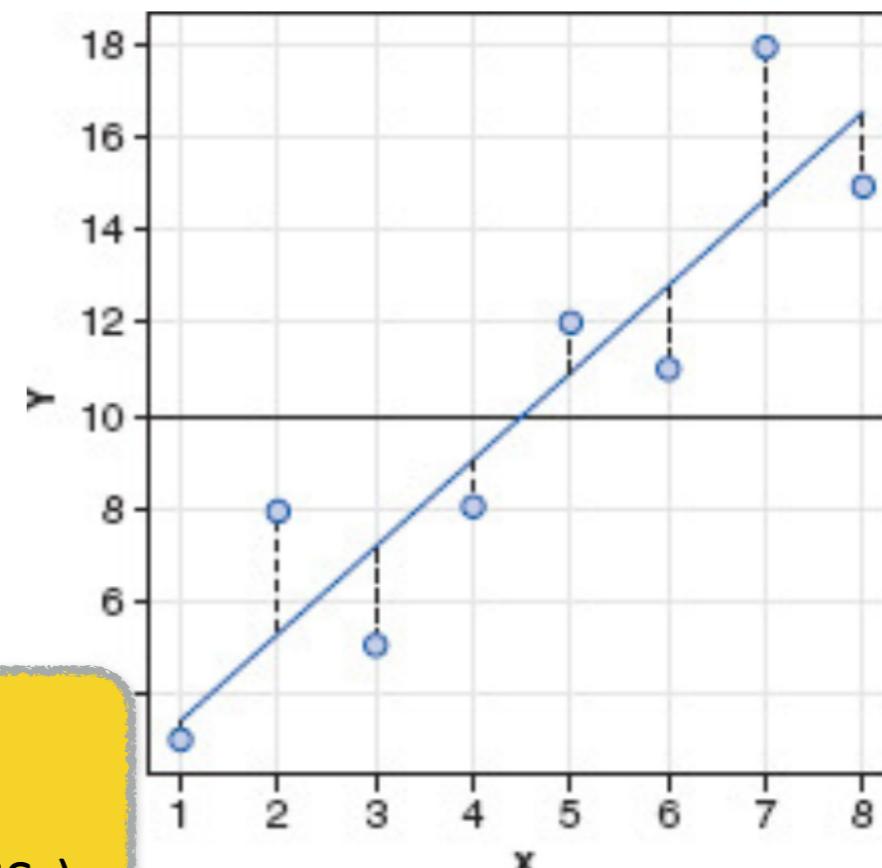
Practice

Assignment

# Regression



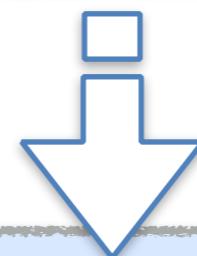
$SS_T$  uses the differences between the observed data and the mean value of Y



$SS_T$

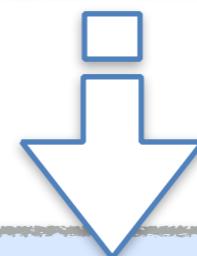
Total variance in the data ( $SS_Y$ )

$SS_R$  uses the differences between the observed data and the regression line



$SS_R$

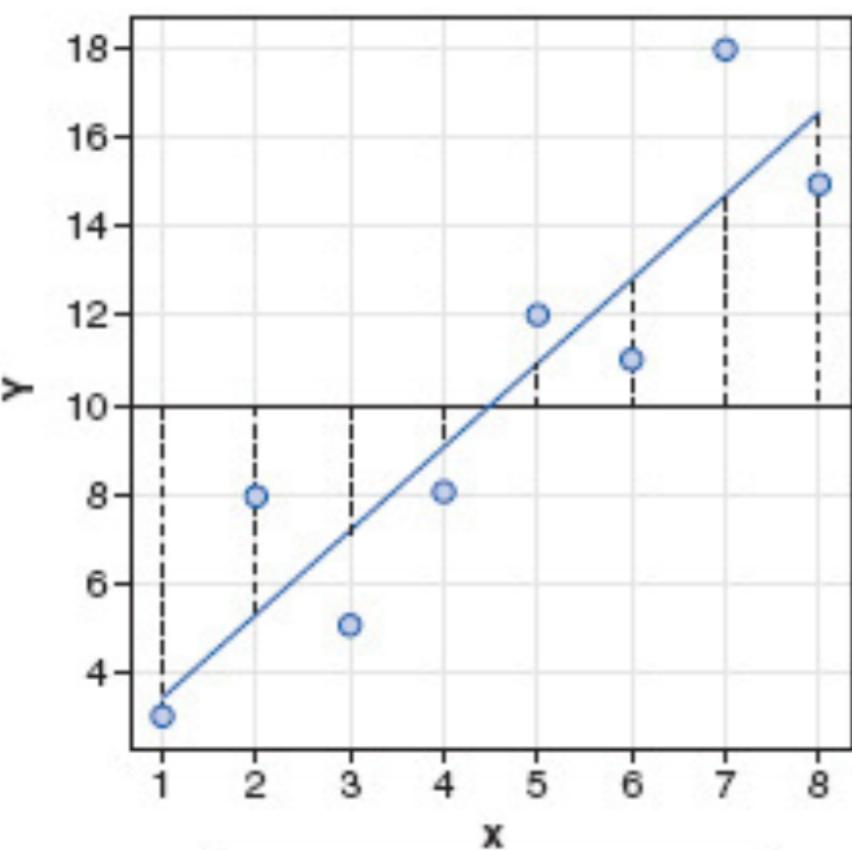
Variance explained by regression model



$SS_E$

Unexplained variance (Residue)

$$Y = X\beta + \epsilon$$



$SS_M$  uses the differences between the mean value of Y and the regression line

.1. The Sum of Squares (total) is

$$\sum(Y_i - \bar{Y})^2 = \sum(\hat{Y}_i - \bar{Y})^2 + \sum(Y_i - \hat{Y}_i)^2$$

Sum of Squares total $SS_{Total}$	=	Sum of Squares due to regression $SS_R$	+	Sum of Squares residual or error $SS_E$ .
---	---	---	---	---

# Summary of Regression

- Example: Tannin-induced proline-rich salivary proteins (PRPs) diminish the anti-nutritional effects of dietary polyphenolics in rats, so the growth in size is associated with tannin concentration.

Outcome measure (Y): Growth in rat size

$$Y = X\beta + \epsilon$$

Predictor (X): dietary Tannin concentration

① Hypothesis

$$\left. \begin{array}{l} H_0: \beta = 0 \\ H_a: \beta \neq 0 \end{array} \right\}$$

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	11.7556	1.0408	11.295	9.54e-06 ***
tannin	-1.2167	0.2186	-5.565	0.000846 ***

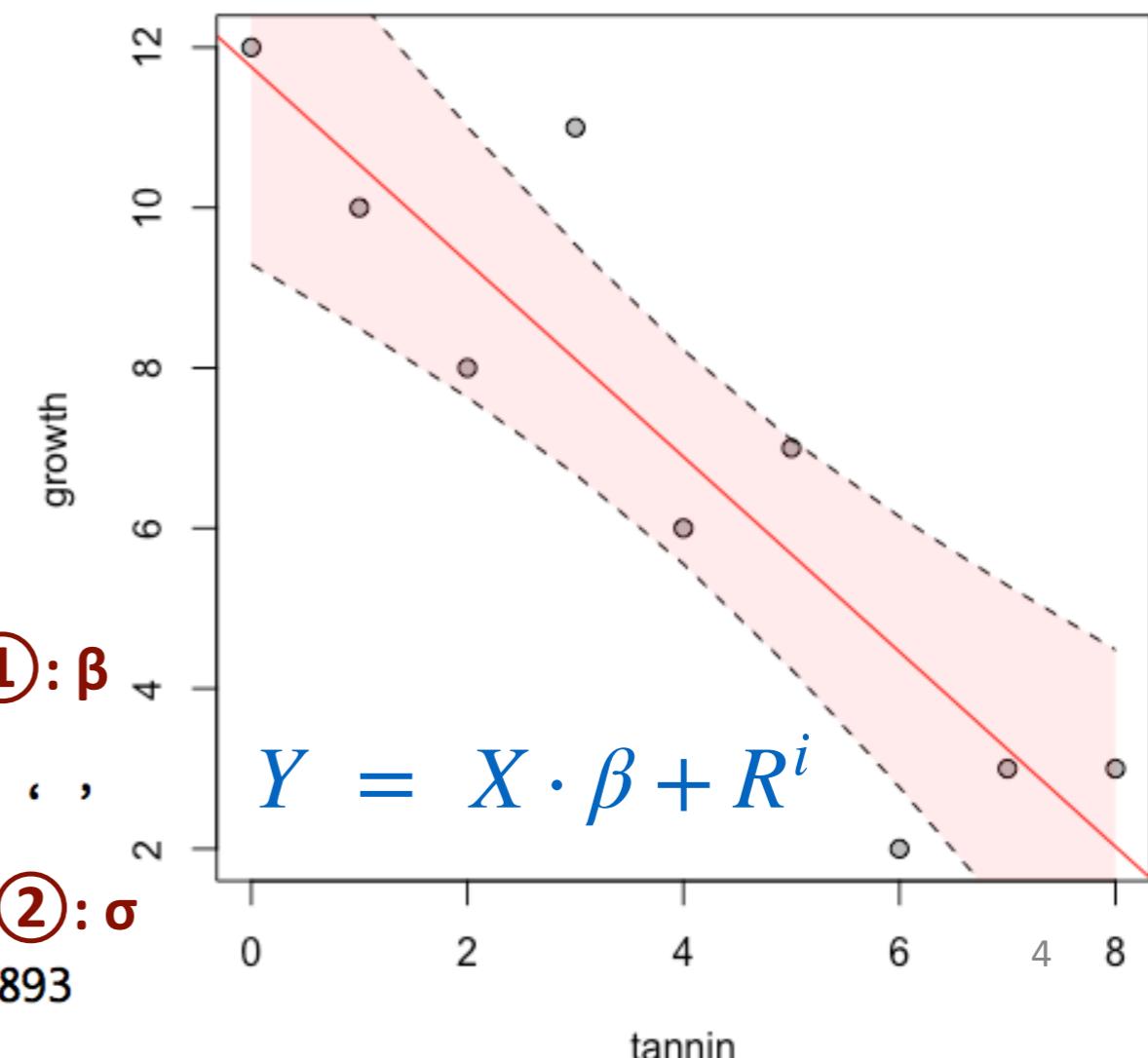
---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’

Residual standard error: 1.693 on 7 degrees of freedom

Multiple R-squared: 0.8157, Adjusted R-squared: 0.7893

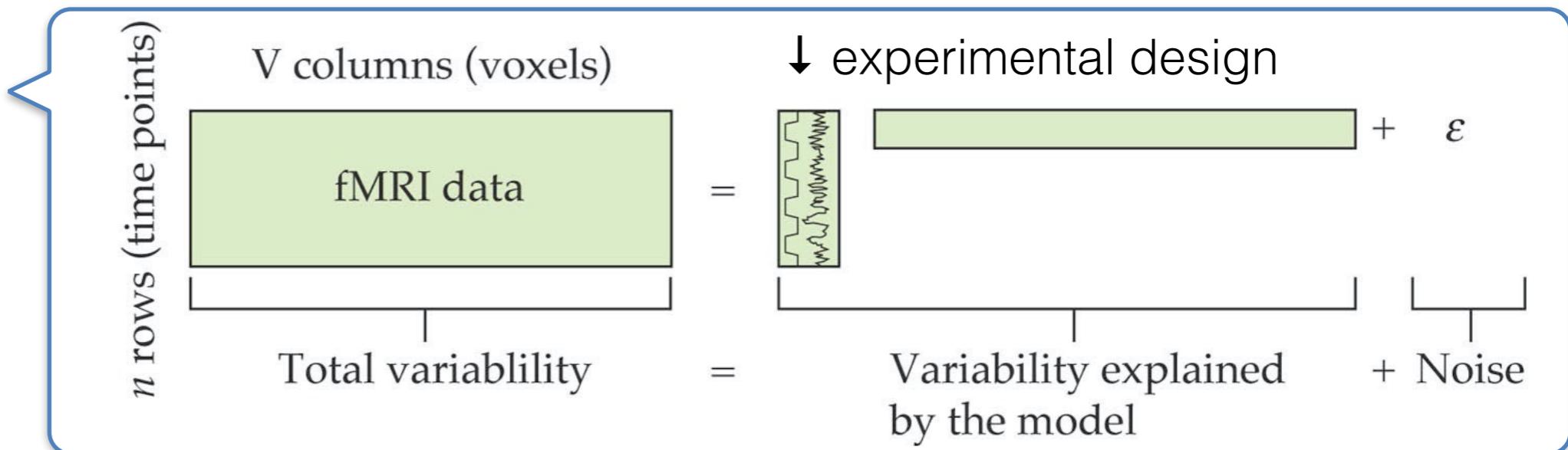
F-statistic: 30.97 on 1 and 7 DF, p-value: 0.0008461



# Testing Effect Size ( $\beta$ )

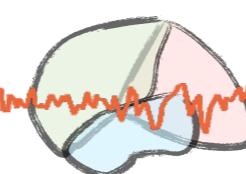
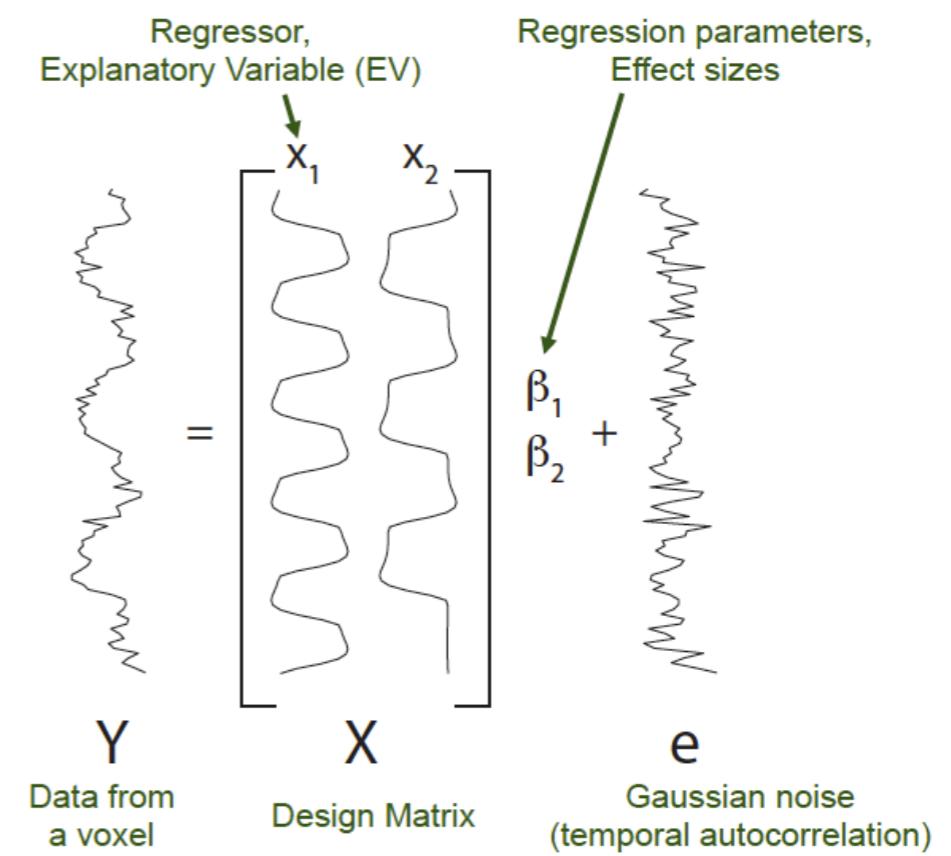
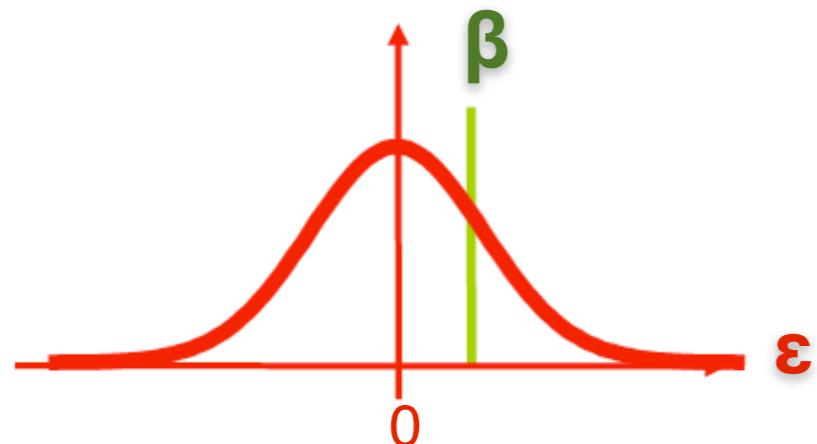
$$Y = X\beta + \epsilon$$

$$\begin{aligned} H_0 : c'\beta &= 0 \\ H_1 : c'\beta &\neq 0 \end{aligned}$$



- Every regressor has its own  $\beta$  estimate.

- How to test the stat. significance?
- If  $\beta$  is non-zero, then voxel is active.
- every voxel has its  $\beta$  value →  $\beta$  maps



# Hypothesis Testing in fMRI (1st-level)

## Regression in R code

Coefficients:		$\beta$	Multiple Regression			
(Intercept)	-26.612958		Estimate	Std. Error	t value	Pr(> t )
adverts	0.084885			17.350001	-1.534	0.127
airplay	3.367425			0.006923	12.261	< 2e-16 ***
attract	11.086335			0.277771	12.123	< 2e-16 ***
---				2.437849	4.548	9.49e-06 ***
Signif. codes:	0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

The standard error of  $b$  is

$$SE(\beta) = \sqrt{\frac{MS_{err}}{SS_x}}$$

## Statistical Hypothesis:

- Null Hypothesis: Is  $c^T \beta$  Non-Zero?

$$H_0 : c^T \beta = 0$$

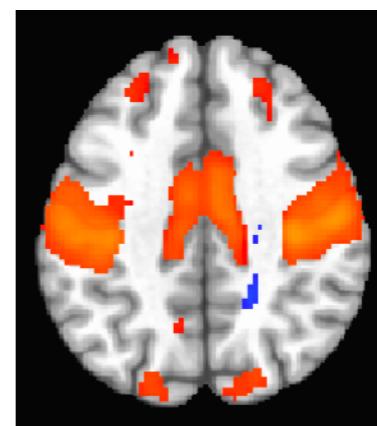
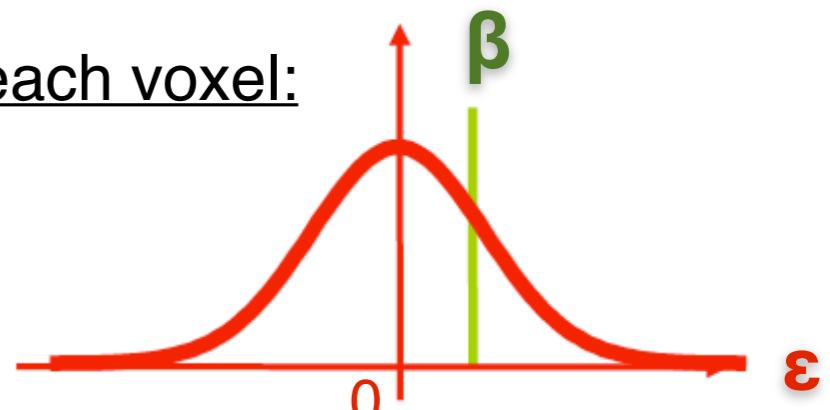
$$H_1 : c^T \beta \neq 0$$

$$t = \frac{\hat{\beta}}{\widehat{SE}(\beta)}$$

$$\begin{aligned} t &= \frac{c^T \hat{\beta} - 0}{\sqrt{Var(c^T \hat{\beta})}} = \frac{c^T \hat{\beta}}{\sqrt{Var(c^T (X^T X)^{-1} X^T Y)}} \\ &= \frac{c^T \hat{\beta}}{\sqrt{[c^T (X^T X)^{-1} X^T] \sum_{\epsilon} [c^T (X^T X)^{-1} X^T]^T}} \end{aligned}$$

$$dof = N - r$$

For each voxel:



**Fixed effect**  
for each subject  
(1<sup>st</sup>-level in SPM)  
(not generalizable)



# Confidence Intervals for Linear Regression

**FORMULA 10.5.** A  $(1 - \alpha)100\%$  confidence interval for the slope  $\beta$  of the least squares regression line is determined by

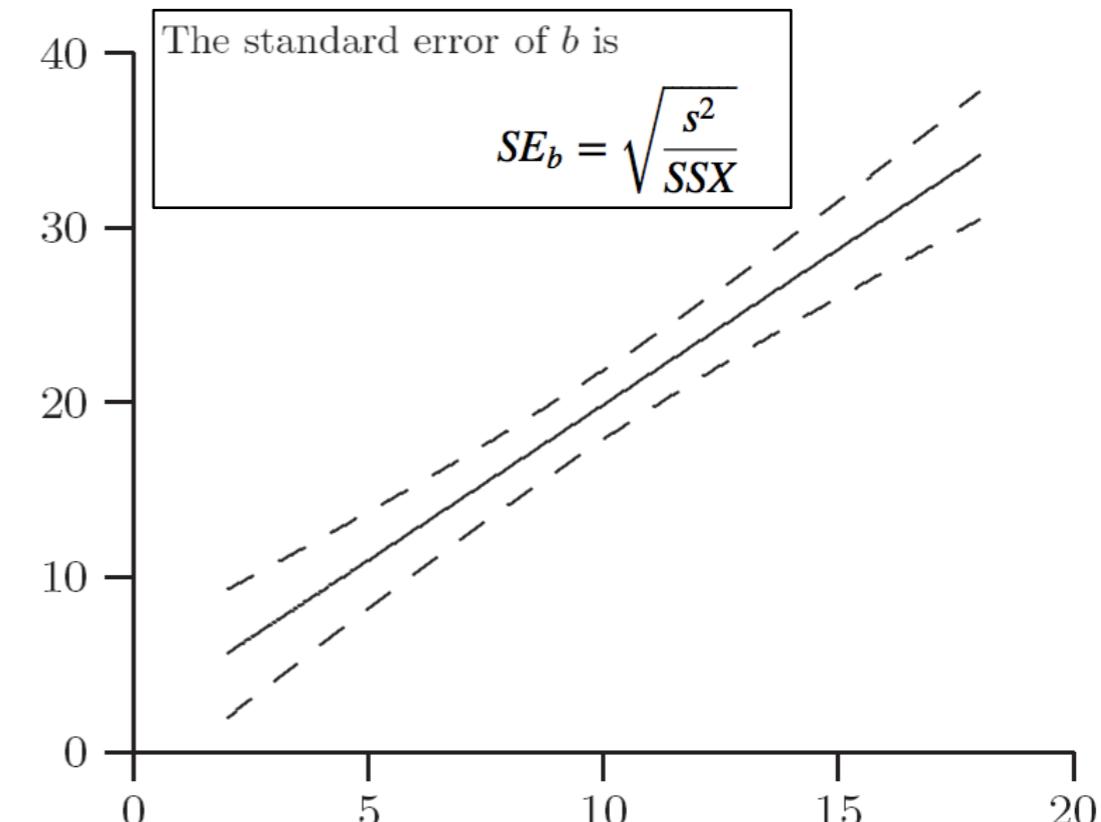
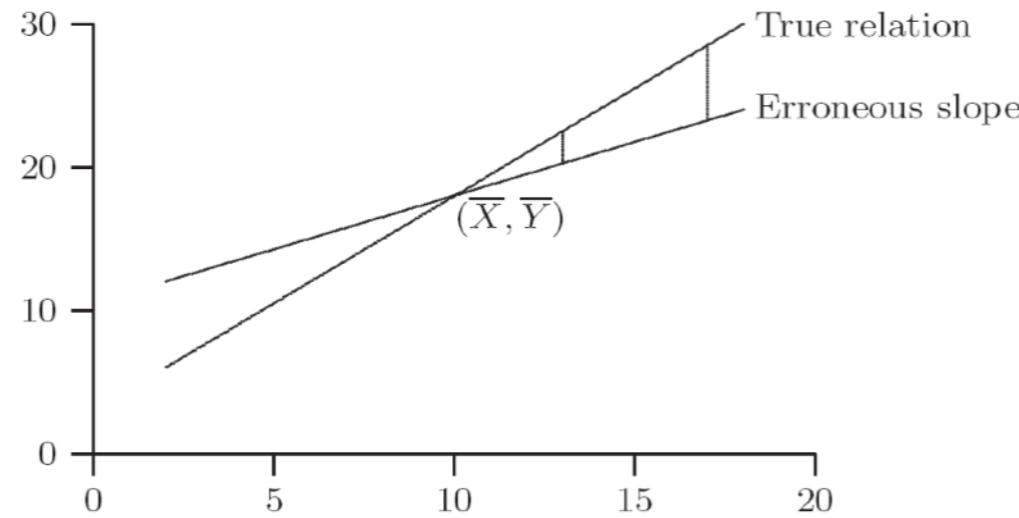
$$C \left( b - t_{1-\frac{\alpha}{2}} s_b \leq \beta \leq b + t_{1-\frac{\alpha}{2}} s_b \right) = 1 - \alpha$$

with  $\nu = n - 2$ . Its endpoints are

$$L_1 = b - t_{1-\frac{\alpha}{2}} s_b$$

and

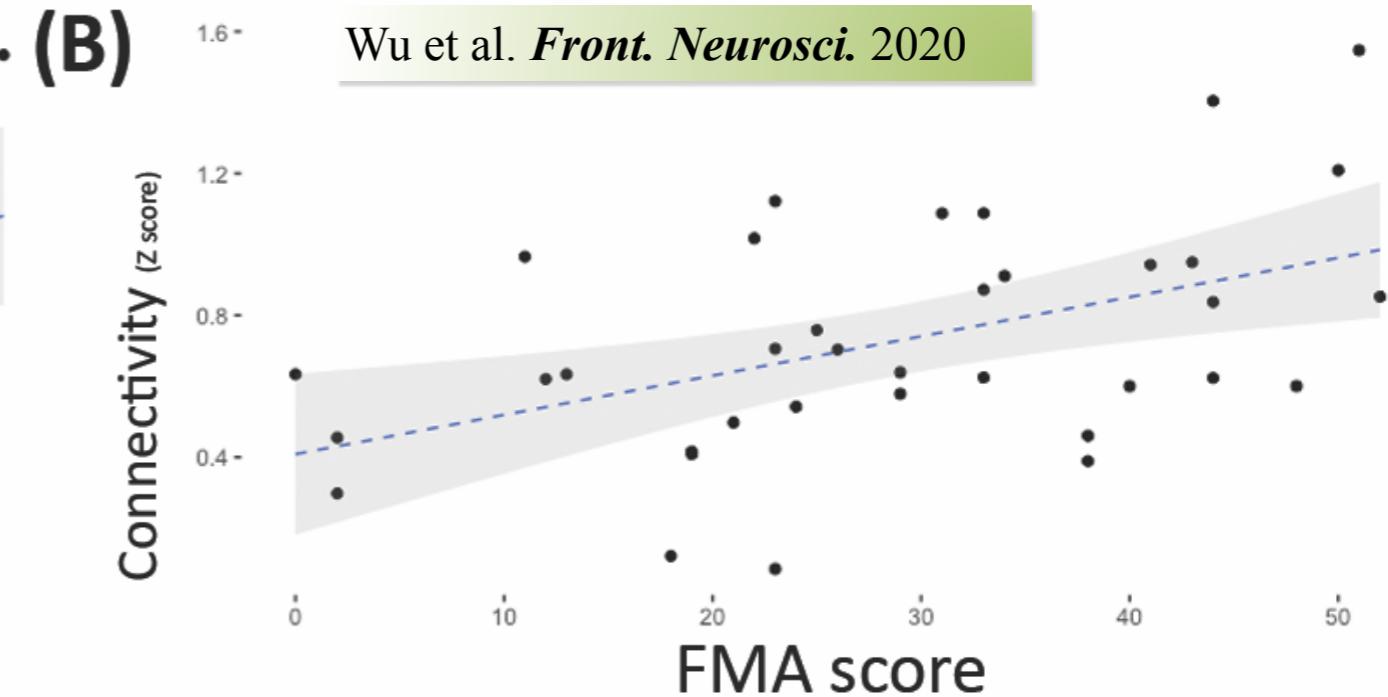
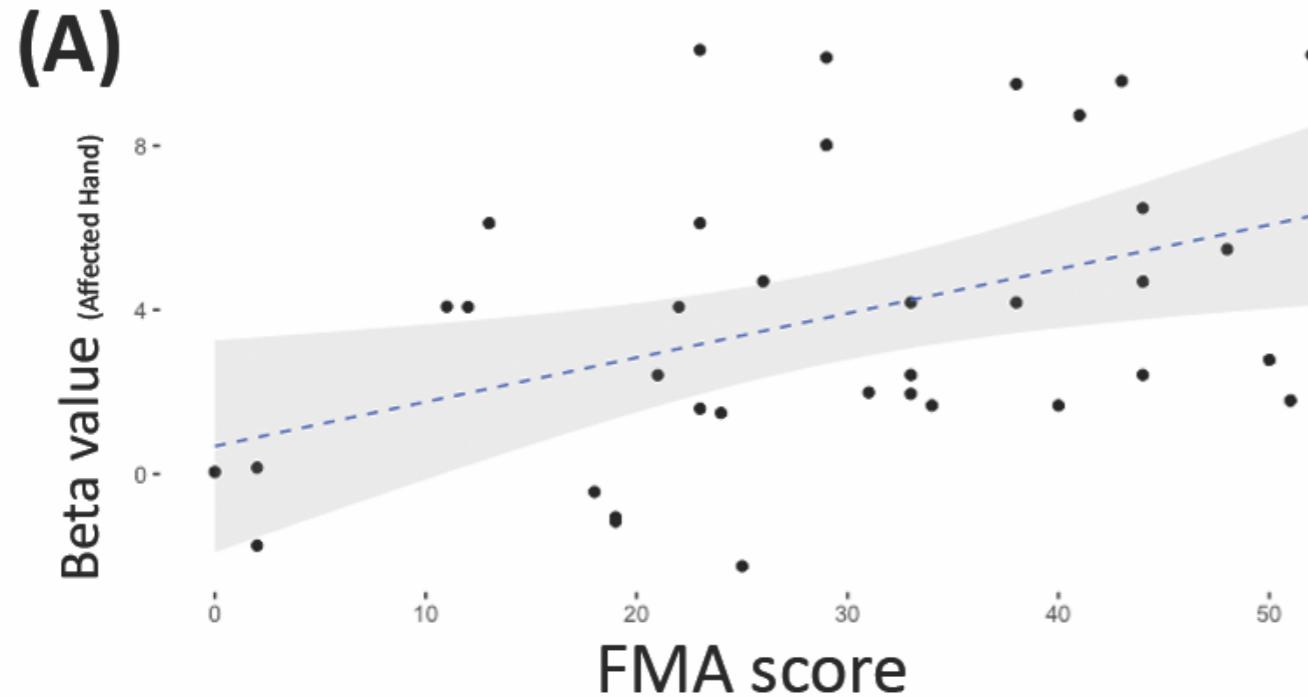
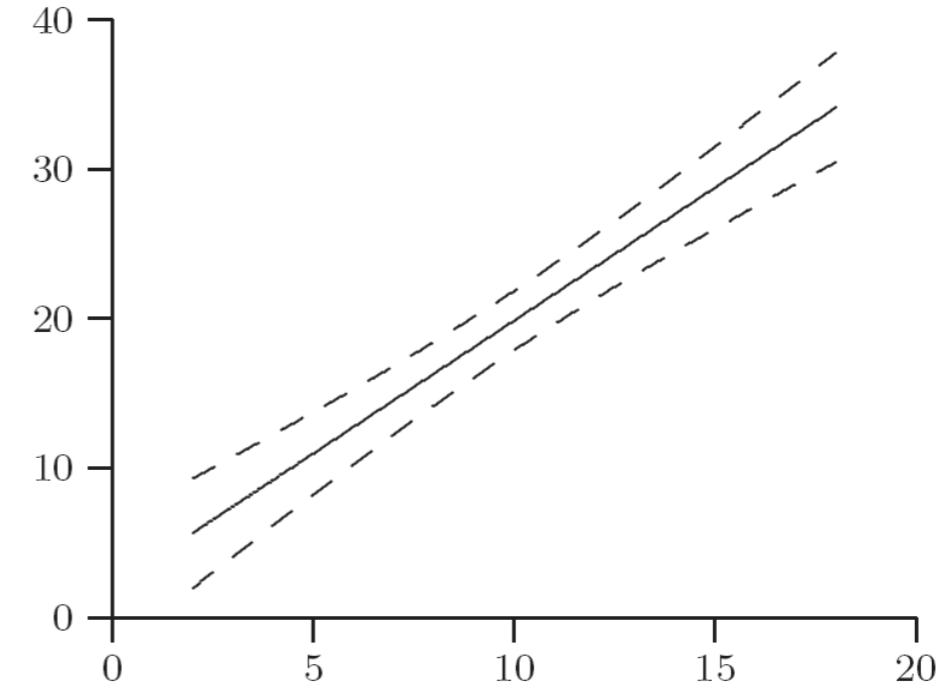
$$L_2 = b + t_{1-\frac{\alpha}{2}} s_b.$$



# Confidence Intervals for Linear Regression

The standard error of  $b$  is

$$SE_b = \sqrt{\frac{s^2}{SSX}}$$



# Confidence Intervals for Linear Regression

$s$  = mean squared error

The standard error of  $b$  is

$$SE_b = \sqrt{\frac{s^2}{SSX}}$$

from '**Unreliability**' point of view

(Error term dependent on the range of X)

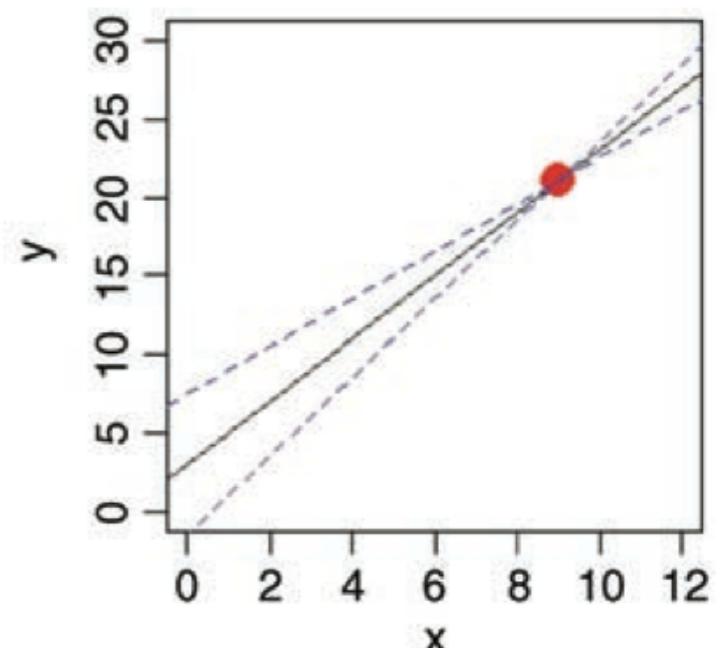
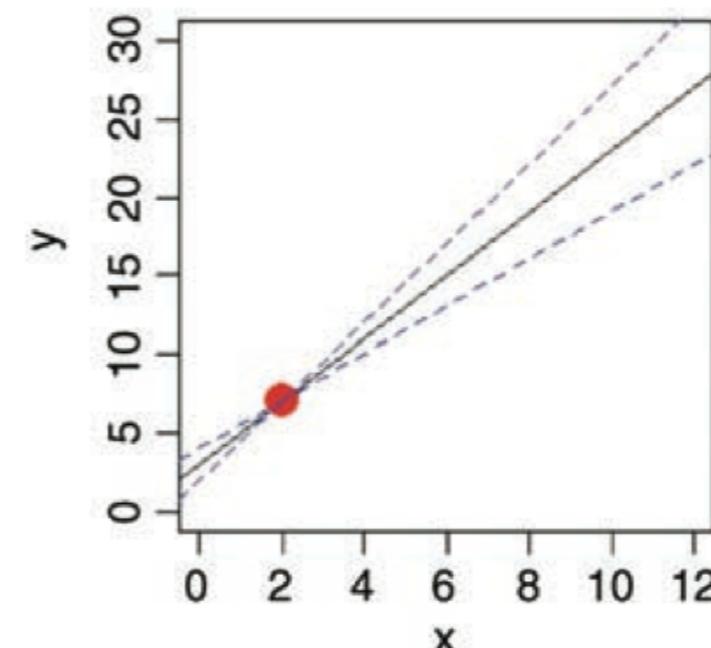
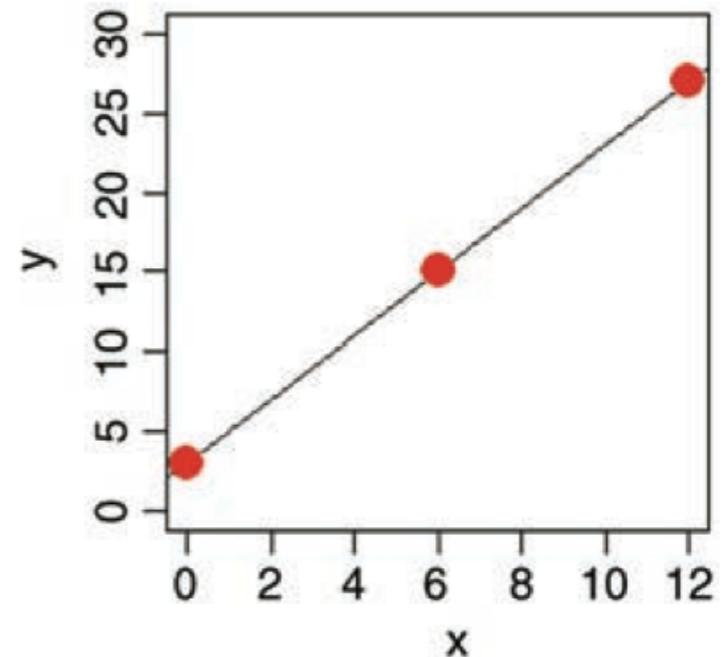
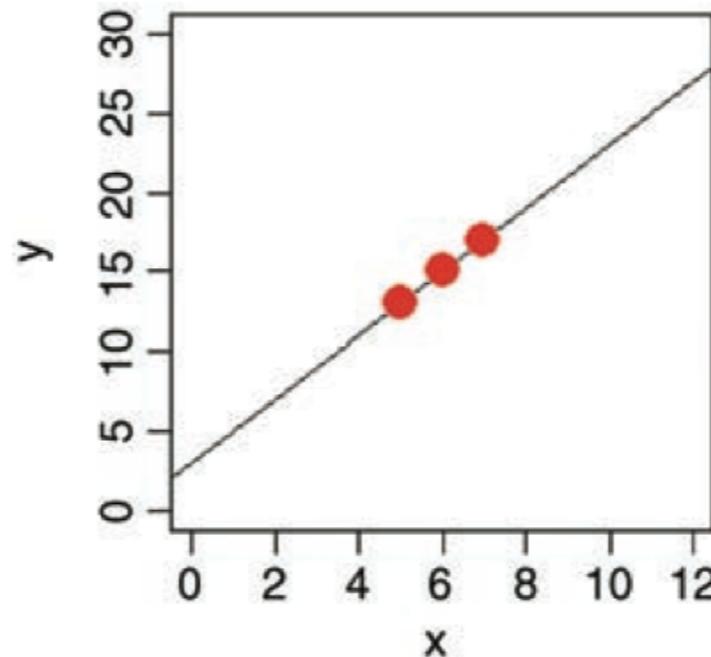
The standard error of *intercept* is

$$SE_a = \sqrt{\frac{s^2 \sum x^2}{n \times SSX}}$$

(Dependent on the mean value of X)

The standard error of *predicted*  $\hat{y}$  is

$$SE_{\hat{y}} = \sqrt{s^2 \left[ \frac{1}{n} + \frac{(x - \bar{x})^2}{SSX} \right]}$$



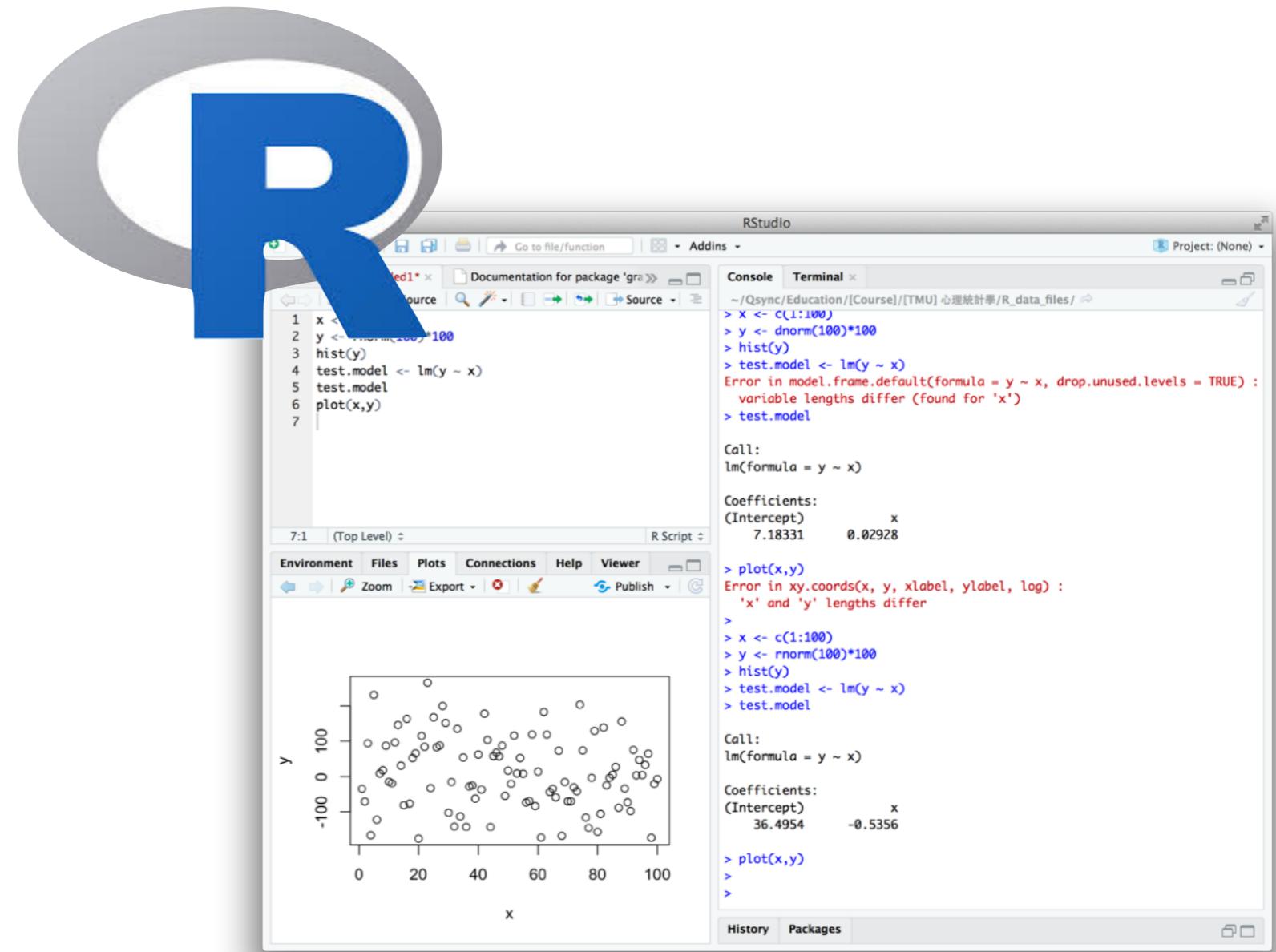
• Use 'confint' in R



臺北醫學大學  
TAIPEI MEDICAL UNIVERSITY

nls  
predict

# ① NONLINEAR LEAST SQUARES



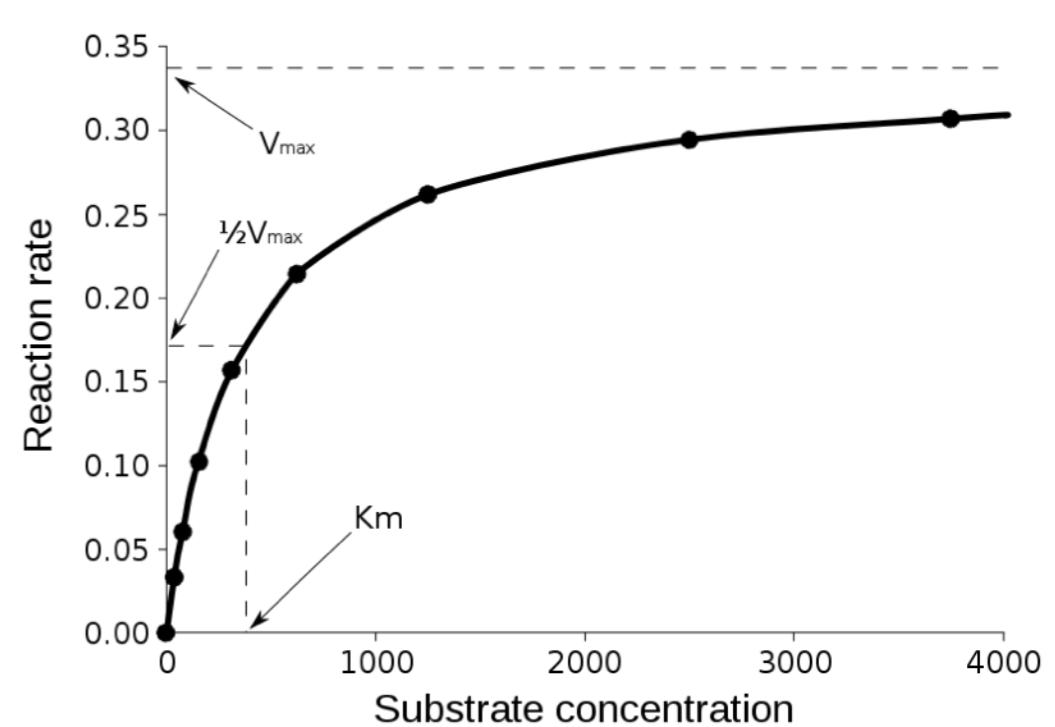
臺北醫學大學  
TAIPEI MEDICAL UNIVERSITY

# Nonlinear Model Fitting

- **Nonlinear regression** is a form of regression analysis in which observational data are modeled by a function which is a nonlinear combination of the model parameters.
- Given a specified equation using nonlinear least squares (*nls*).

Ex: asymptotic exponential:  $y = a - be^{-cx}$

- **Generalized Additive Models (GAM)**
  - When there is no particular equation to describe the relation.
  - fit non-parametric smoothers to the data without models.



[https://en.wikipedia.org/wiki/Nonlinear\\_regression](https://en.wikipedia.org/wiki/Nonlinear_regression)

**DEMO**

# Age vs. Bone Length

- **Background:** The bone length of deer jaws is increasing with age. However, the relationship is not linear, but an asymptotic exponential. What are the best parameters to fit the dataset?

$$y = a - be^{-cx}$$

Outcome measure: bone length

Predictor: age

Load: jaws.csv

①  
Hypothesis

**H<sub>0</sub>: data is not asymptotic exponential**  
**H<sub>a</sub>: data fit asymptotic exponential**

- Data import:

▶ *deer <- read.csv("jaws.csv", header = TRUE)*

## DEMO

# Age vs. Bone Length

- Assumption check:

► *Here we assume the data fulfill all assumptions for Regression.*

②  
Assumption

General form:

► ***nls(outcome ~ predictor, data=dataframe, start)***

③  
Testing

- Parameters:

- *outcome*: the variable to be predicted (dependent variable).
- *predictor(s)*: the controlled variable (independent variable).
- *dataframe*: the name of data frame you stored your indices.
- *start*: a named numeric vector of initial estimates.



臺北醫學大學  
TAIPEI MEDICAL UNIVERSITY

# Age vs. Bone Length

- Practice:

- ▶ `model1 <- nls(bone~a-b*exp(-c*age),  
start=list(a=120,b=110,c=0.064))`
- ▶ `summary(model1)`

convergence of nonlinear models can depend on good starting values critically.

- Results:

Formula: bone ~ a - b \* exp(-c \* age)

Parameters:

	Estimate	Std. Error	t value	Pr(> t )
a	115.2528	2.9139	39.55	< 2e-16 ***
b	118.6875	7.8925	15.04	< 2e-16 ***
c	0.1235	0.0171	7.22	2.44e-09 ***
---				

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1

Residual standard error: 13.21 on 51 degrees of freedom

Number of iterations to convergence: 5

Achieved convergence tolerance: 2.391e-06

## DEMO

# Age vs. Bone Length

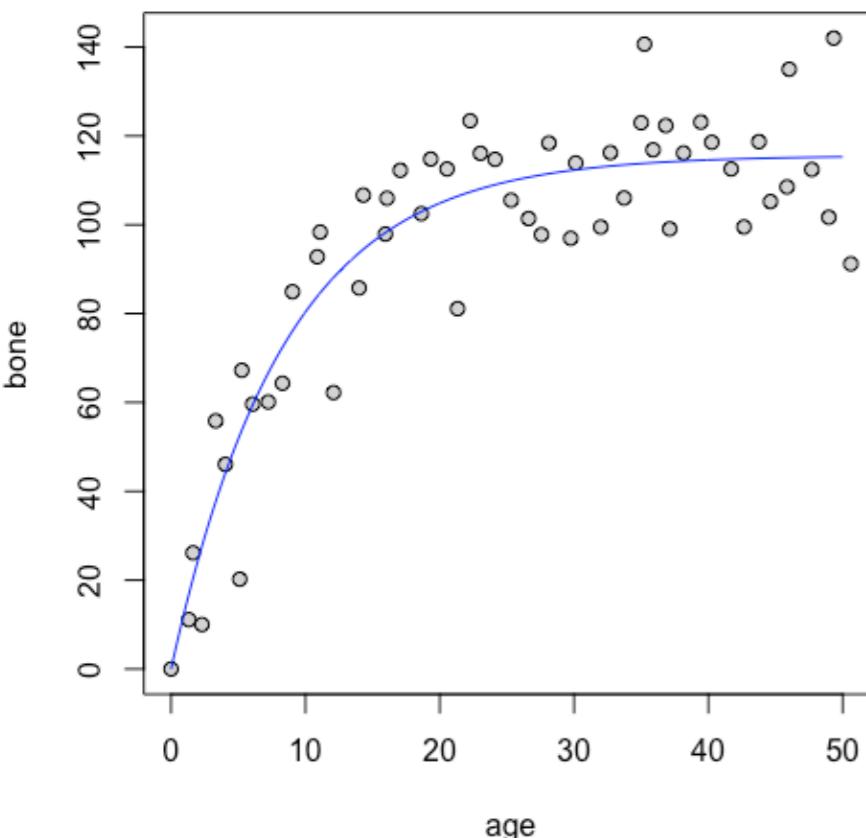
③ Testing

- Practice:

- ▶ `model2 <- nls(bone~a*(1-exp(-c*age)),  
start=list(a=120,c=0.064))`
- ▶ `summary(model2)`
- ▶ `nlstools::confint2(model2)`

- which model is better?

- Results:



Formula:  $\text{bone} \sim a * (1 - \exp(-c * \text{age}))$

Parameters:

	Estimate	Std. Error	t value	Pr(> t )
a	115.58056	2.84365	40.645	< 2e-16 ***
c	0.11882	0.01233	9.635	3.69e-13 ***
---				

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13.1 on 52 degrees of freedom

Number of iterations to convergence: 5

Achieved convergence tolerance: 1.369e-06

# Age vs. Jaw Length

④

Effect Size

- Effect size:

Parameters:

```
Estimate Std. Error t value Pr(>|t|)  
a 115.58056 2.84365 40.645 < 2e-16 ***  
c 0.11882 0.01233 9.635 3.69e-13 ***  
---  
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Residual standard error: 13.1 on 52 degrees of freedom

- How to calculate effect size for NLS?

$$r^2 = \frac{SS_R}{SS_{Total}}$$



# Coefficient of Determination

- **R<sup>2</sup>**

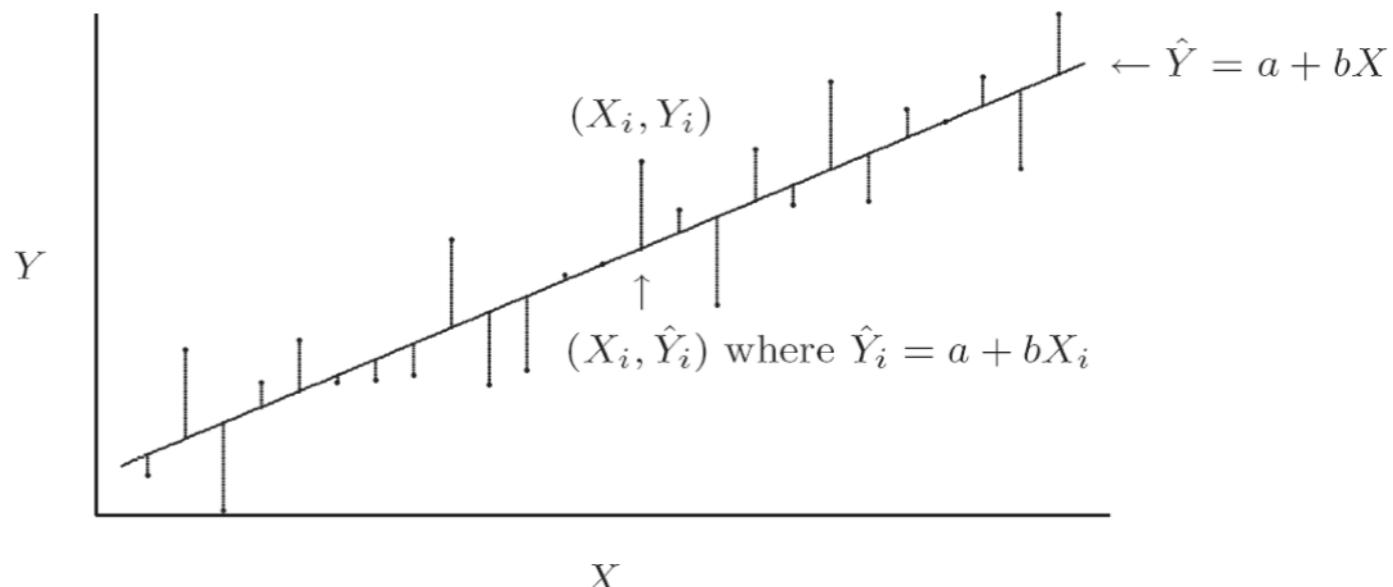
- R<sup>2</sup> lies between 0 and +1.
- The proportion of the variance in the dependent variable (Y) that is predictable from (explained by) the independent variable (X).

The Sum of Squares (total) is

$$\sum(Y_i - \bar{Y})^2 = \sum(\hat{Y}_i - \bar{Y})^2 + \sum(Y_i - \hat{Y}_i)^2$$

Sum of Squares total SS <sub>Total</sub>	=	Sum of Squares due to regression SS <sub>R</sub>	+	Sum of Squares residual or error SS <sub>E</sub> .
--	---	--	---	--

$$r^2 = \frac{SS_R}{SS_{Total}} = \frac{(SS_{XY})^2}{SS_X \cdot SS_Y}$$



## Practice

# Age vs. Jaw Length

- How to calculate effect size for NLS?

Compare it with the **NULL model**

```
null.model <- lm(bone ~ 1)  
summary.aov(null.model)  
summary(model2)
```

```
> summary.aov(null.model)  
Df Sum Sq Mean Sq F value Pr(>F)  
Residuals 53 59008 1113  
total variance (SST)
```

Parameters:

```
Estimate Std. Error t value Pr(>|t|)  
a 115.58056 2.84365 40.645 < 2e-16 ***  
c 0.11882 0.01233 9.635 3.69e-13 ***  
---
```

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

→ RSE = 13.1 SD with (52 dof)  
→ MSE = (13.1)<sup>2</sup>  
→ SSE = 52 \* (13.1)<sup>2</sup>  
= 8923.72

Residual standard error: 13.1 on 52 degrees of freedom

information of error after modeling

→ SSR = SST-SSE  
→ R<sup>2</sup> = SSR/SST

$$\text{Variance}(s^2) = \frac{\sum (x_i - \bar{x})^2}{N-1} = \frac{\sum (x_i - \bar{x})(x_i - \bar{x})}{N-1}$$



## DEMO

# Age vs. Jaw Length

④

Effect Size

- Effect size:

$$r^2 = \frac{SS_R}{SS_{Total}}$$

Parameters:

	Estimate	Std. Error	t value	Pr(> t )
a	115.58056	2.84365	40.645	< 2e-16 ***
c	0.11882	0.01233	9.635	3.69e-13 ***
---				

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13.1 on 52 degrees of freedom

→ SSR = SST-SSE = 59008-8923.72

→ R<sup>2</sup> = SSR/SST = 84.9%

⑤

Decision

- Reporting decision:

The asymptotic exponential model  $y= a(1-\exp^{-bx})$  fits the bone-length data well ( $n=54$ ,  $p < 0.001$ ), explaining 84.9% of the total variation with  $a = 115.58 \pm 2.84$  and  $b = 0.12 \pm 0.01$ .



臺北醫學大學  
TAIPEI MEDICAL UNIVERSITY

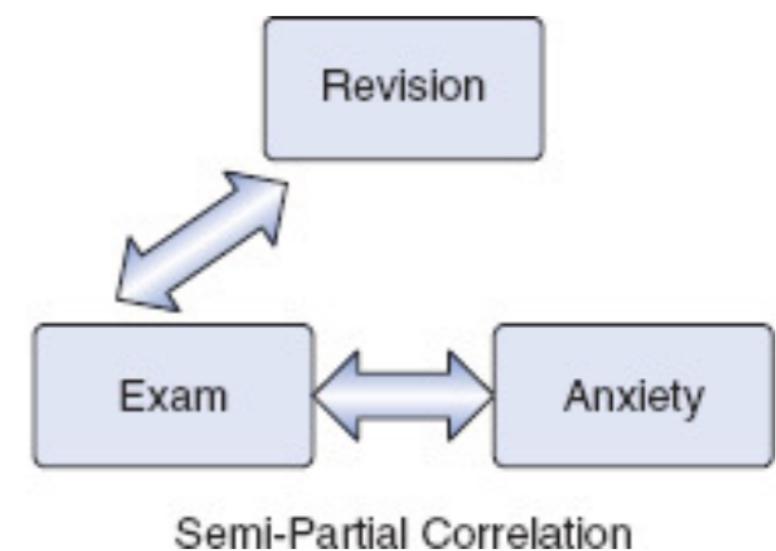
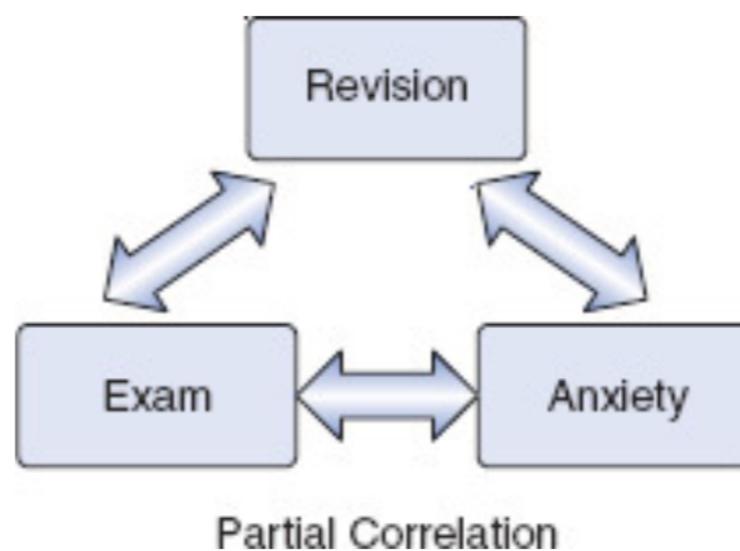
The Sum of Squares (total) is

$$\sum(Y_i - \bar{Y})^2 = \sum(\hat{Y}_i - \bar{Y})^2 + \sum(Y_i - \hat{Y}_i)^2$$

$$\begin{array}{ccl} \text{Sum of Squares} & = & \text{Sum of Squares} & + & \text{Sum of Squares} \\ \text{total} & & \text{due to regression} & & \text{residual or error} \\ SS_{Total} & & SS_R & & SS_E. \end{array}$$

Excluding contributions of the 3rd variable

## ② PARTIAL CORRELATION



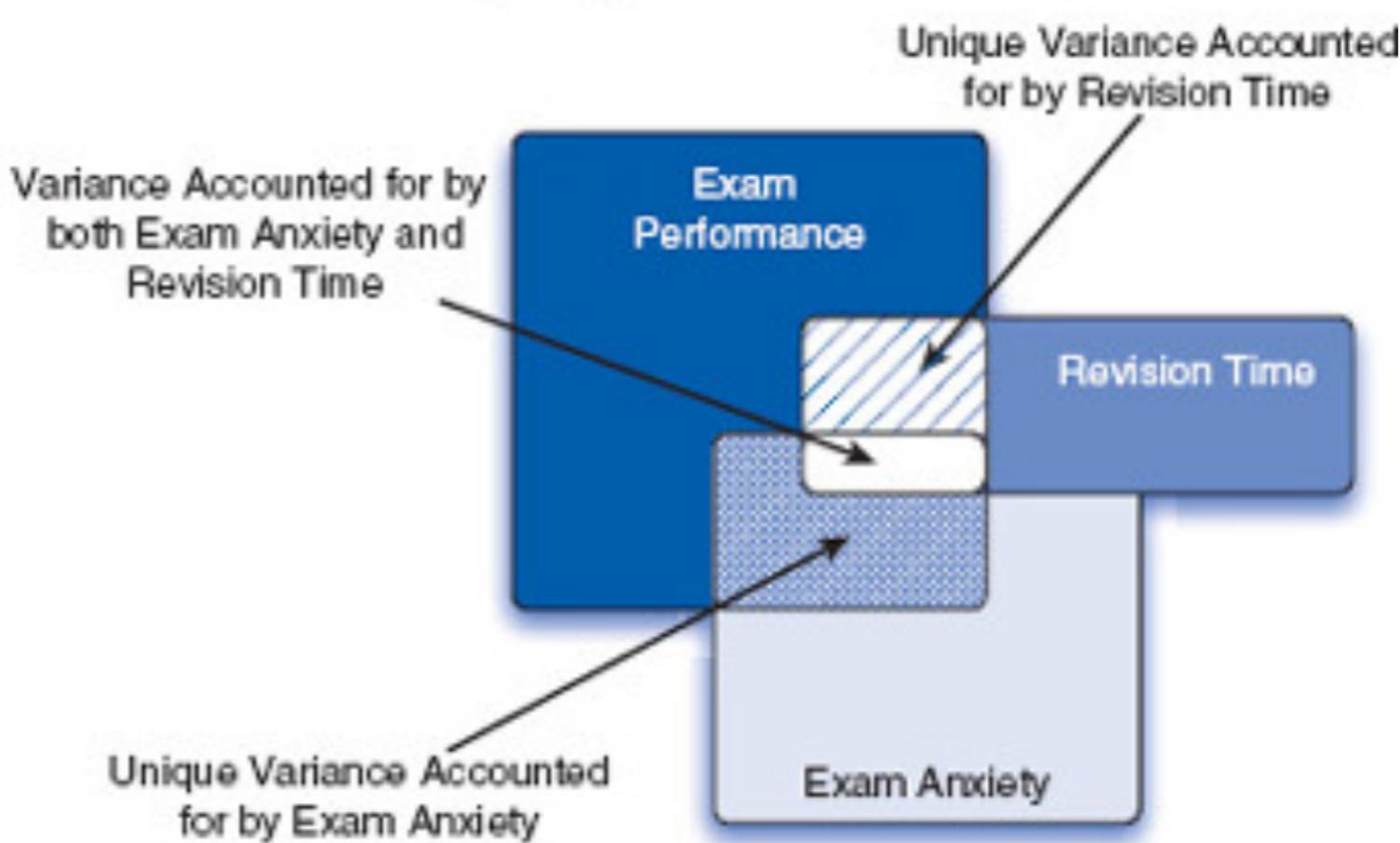
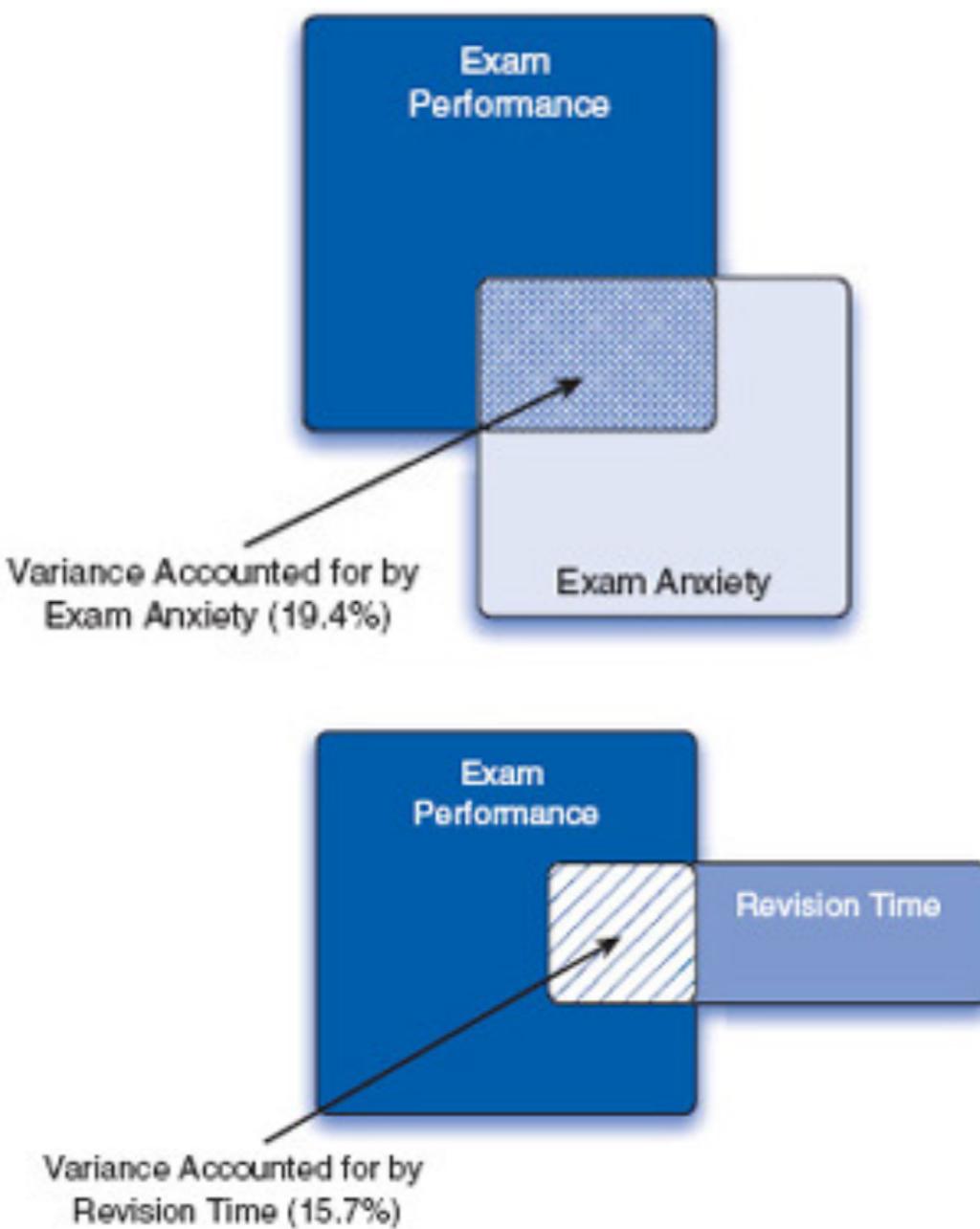
臺北醫學大學  
TAIPEI MEDICAL UNIVERSITY

> R2

	Exam	Anxiety	Revise
Exam	100.00000	19.44752	15.73873
Anxiety	19.44752	100.00000	50.30345
Revise	15.73873	50.30345	100.00000

# Partial Correlation

- The variance of **x** variable (exam score) is not uniquely explained by **y** (anxiety), which can also be accounted for by the 3<sup>rd</sup> variable (revision time).
- When you want to control/remove its effect.



- What if we want to check Exam & Anxiety with controlling Revise time?

**DEMO**

# Exam — Anxiety — Revise

- **Background:** A psychologist was interested in the effects of exam stress and revision on exam performance. She had devised a questionnaire to assess state anxiety relating to exams. This scale produced a measure of anxiety scored out of 100.
- **Anxiety** was measured before an exam, and the **percentage mark** of each student on the exam was used to assess exam performances.
- She also measured the **time spent revising** (of minutes).

Load:  
**ExamAnxiety.dat**

**Measure 1: Anxiety before exams**

**Measure 2: Percentage mark**

**Measure 3: Revise time**

①  
Hypothesis

**H<sub>0</sub>: Anxiety does not relates with Exam mark** (control Revise).  
**H<sub>a</sub>: Anxiety is linearly correlated with Exam mark** (control Revise).

# Partial Correlation

What if we want to check Exam & Anxiety with controlling Revise time?

**Partial correlation:**  $r_{ab \cdot c} = \frac{r_{ab} - r_{ac}r_{bc}}{\sqrt{1 - r_{ac}^2} \cdot \sqrt{1 - r_{bc}^2}}$

**Semi-partial correlation:**  $r_{a(b \cdot c)} = \frac{r_{ab} - r_{ac}r_{bc}}{\sqrt{1 - r_{bc}^2}}$

③  
Testing

General form: *ppcor*

- ▶ *pcor(df)\$estimate*
- ▶ *pcor.test(x, y, z, method = c("pearson", "kendall", "spearman"))*
- ▶ *spcor.test(x, y, z, method)*

- Parameters:
  - *x, y* : the variables to undergo correlation.
  - *z*: usually the variable to be regressed out.
  - *method=(pearson / spearman / kendall)*: choosing method types



# Partial Correlation

$$r_{ab \cdot c} = \frac{r_{ab} - r_{ac}r_{bc}}{\sqrt{1 - r_{ac}^2} \cdot \sqrt{1 - r_{bc}^2}}$$

What if we want to check Exam & Anxiety with controlling Revise time?

③  
Testing

- Method (1/3): Equation from the original correlation map.

```
> examData2 <- examData %>% select(Exam, Anxiety, Revise)  
> cor(examData2) %>% round(3)
```

	Exam	Anxiety	Revise
Exam	1.000	-0.441	0.397
Anxiety	-0.441	1.000	-0.709
Revise	0.397	-0.709	1.000

$$r_{EA \cdot R} = \frac{-0.441 - (0.397 \cdot -0.709)}{\sqrt{1 - 0.397^2} \cdot \sqrt{1 - r_{-0.709^2}}} = -0.247$$

- Method (2/3): “ppcor” package.

```
ppcor::pcor(examData2)$estimate
```

	Exam	Anxiety	Revise
Exam	1.000	-0.247	0.133
Anxiety	-0.247	1.000	-0.649
Revise	0.133	-0.649	1.000

```
> ppcor::pcor.test(examData2$Exam, examData2$Anxiety, examData2$Revise)  
estimate p.value statistic n gp Method  
1 -0.2466658 0.01244581 -2.545307 103 1 pearson
```



# Partial Correlation

$$r_{ab \cdot c} = \frac{r_{ab} - r_{ac}r_{bc}}{\sqrt{1 - r_{ac}^2} \cdot \sqrt{1 - r_{bc}^2}}$$

What if we want to check Exam & Anxiety with controlling Revise time?

③  
Testing

- Method (3/3): using Regression to “*regress out*” the effect of Revise time; then use the residuals (error) to conduct the correlation analysis.

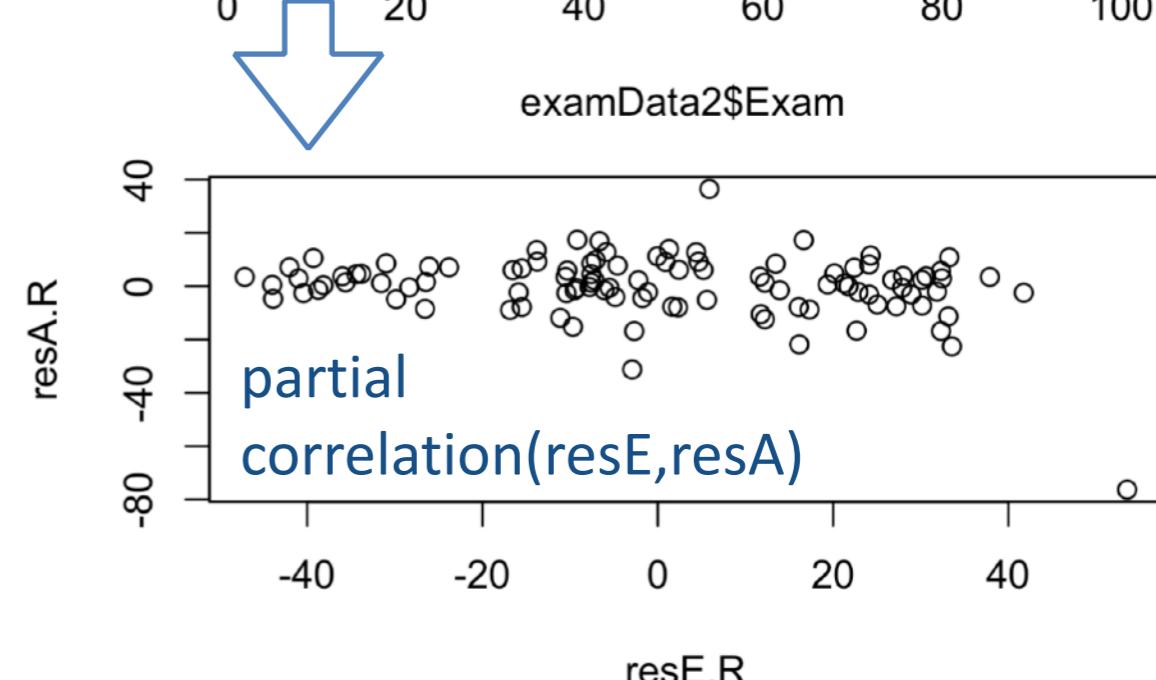
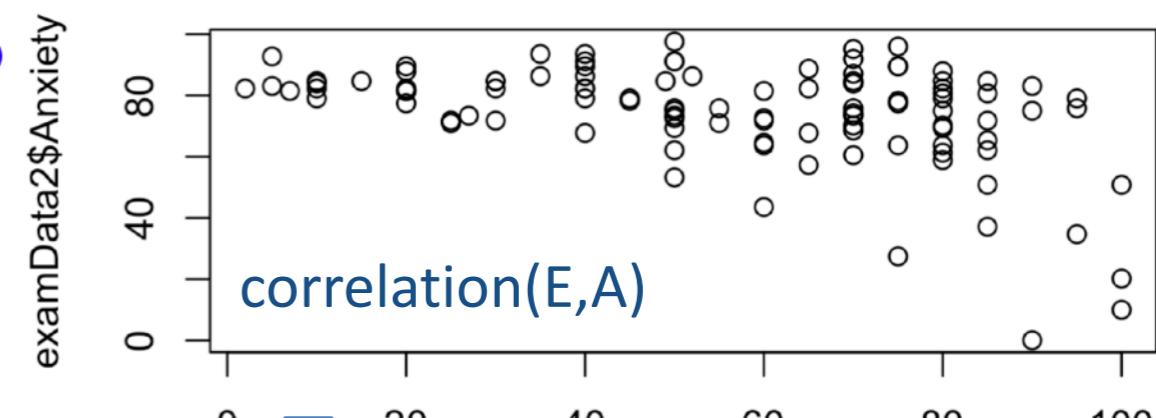
```
resE.R <- residuals(lm(Exam ~ Revise, data = examData2))
resA.R <- residuals(lm(Anxiety ~ Revise, data = examData2))
```

```
> cor.test(resE.R, resA.R)
```

Pearson's product-moment correlation

```
data: resE.R and resA.R
t = -2.558, df = 101, p-value = 0.01201
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.42013496 -0.05580506
sample estimates:
```

cor  
-0.2466658



# Exam — Anxiety — Revise

⑤

Decision

- Reporting decision:

**After controlling the revise time, the Exam performance was still significantly correlated with exam anxiety,  $r = -.247$  ( $p < 0.05$ ).**

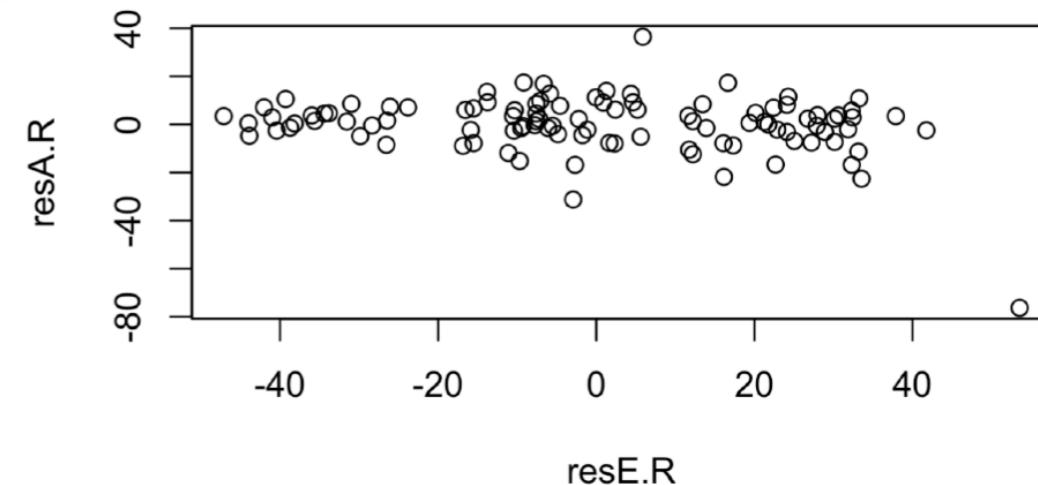
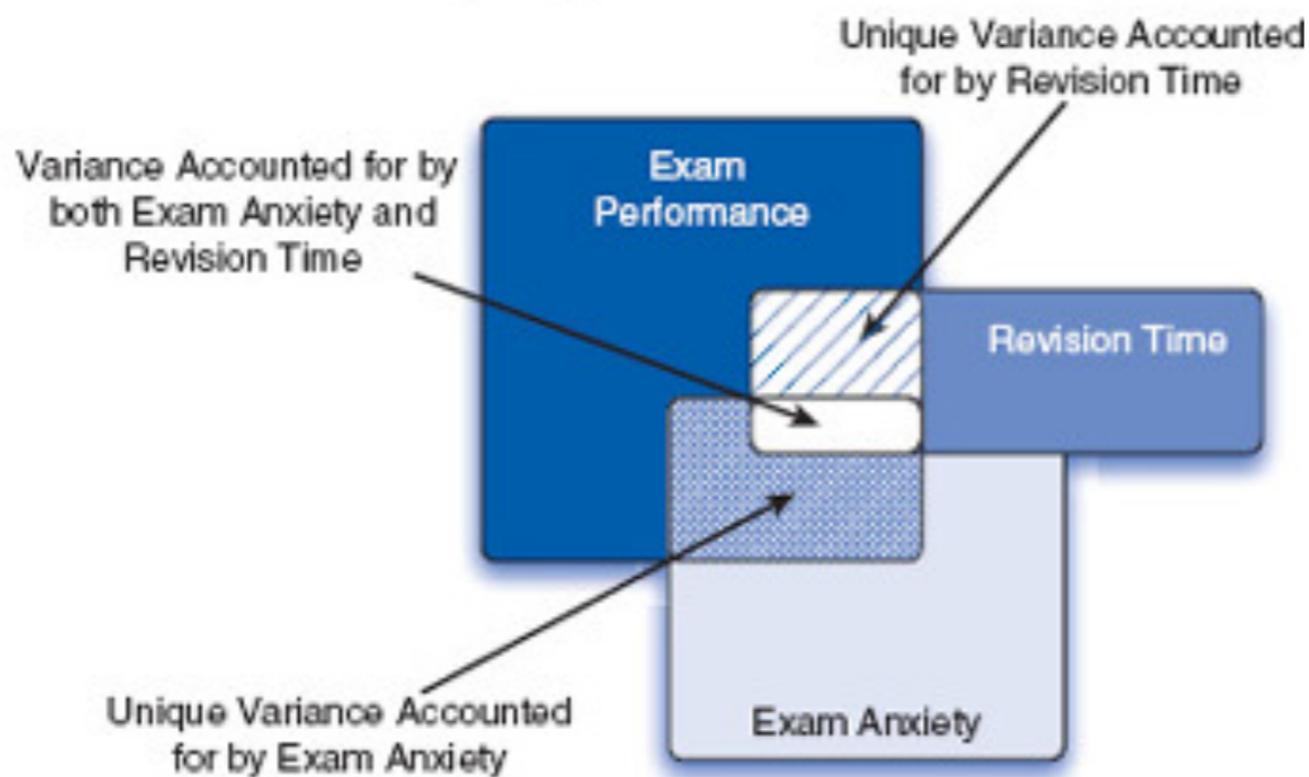
```
> cor.test(resE.R, resA.R)
```

Pearson's product-moment correlation

```
data: resE.R and resA.R
t = -2.558, df = 101, p-value = 0.01201
alternative hypothesis: true correlation is not equal to
95 percent confidence interval:
-0.42013496 -0.05580506
```

sample estimates:

cor  
-0.2466658

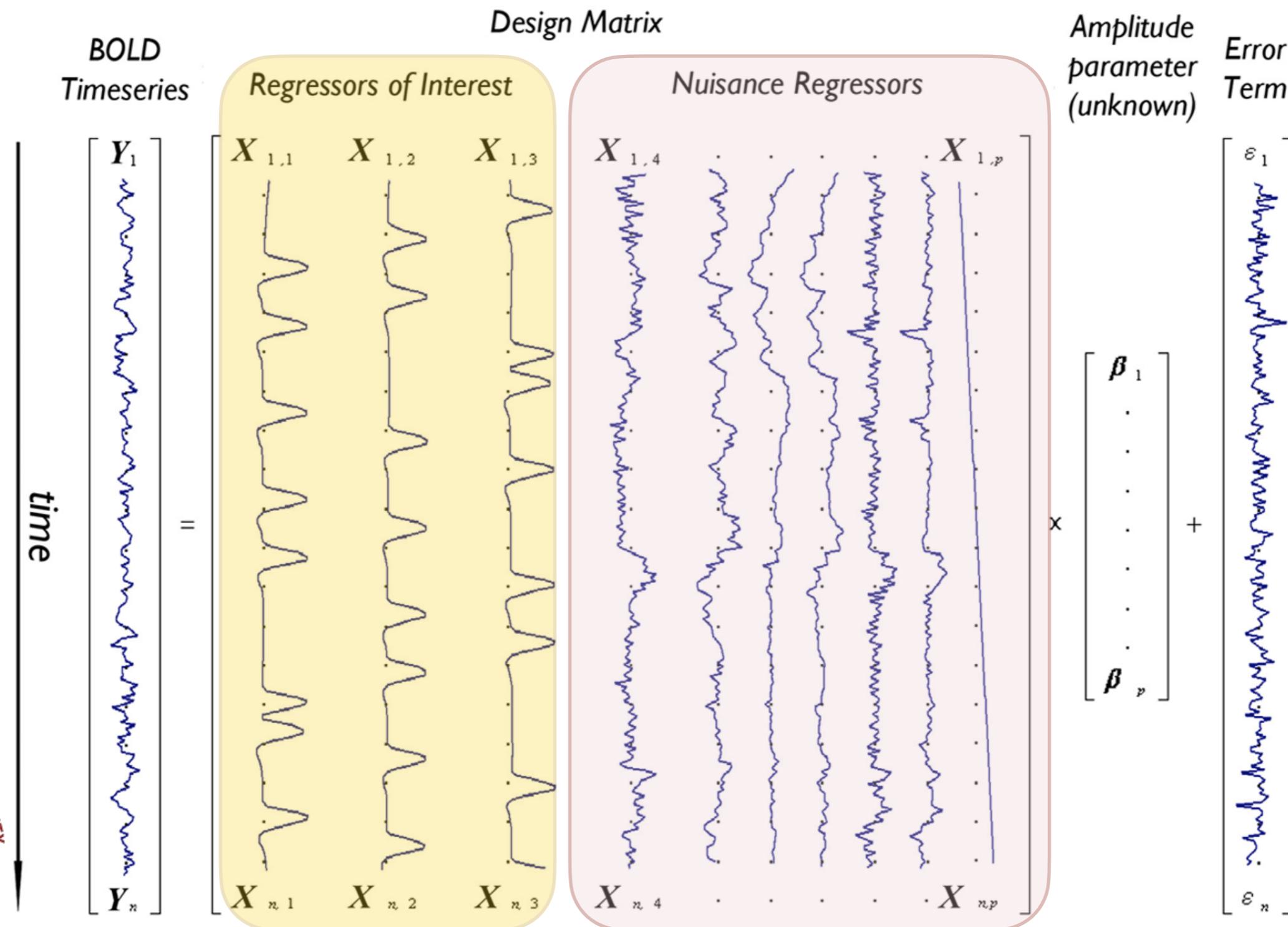


# Partial Correlation

$$r_{ab \cdot c} = \frac{r_{ab} - r_{ac}r_{bc}}{\sqrt{1 - r_{ac}^2} \cdot \sqrt{1 - r_{bc}^2}}$$

We include the “6 motion parameters” in fMRI analysis to “regress out” motion.

③  
Testing



# Semi-Partial Correlation

$$r_{a(b \cdot c)} = \frac{r_{ab} - r_{ac}r_{bc}}{\sqrt{1 - r_{bc}^2}}$$

What if we only want to control Revise time from Exam, but not Anxiety?

③  
Testing

- Method (2/3): “ppcor” package.

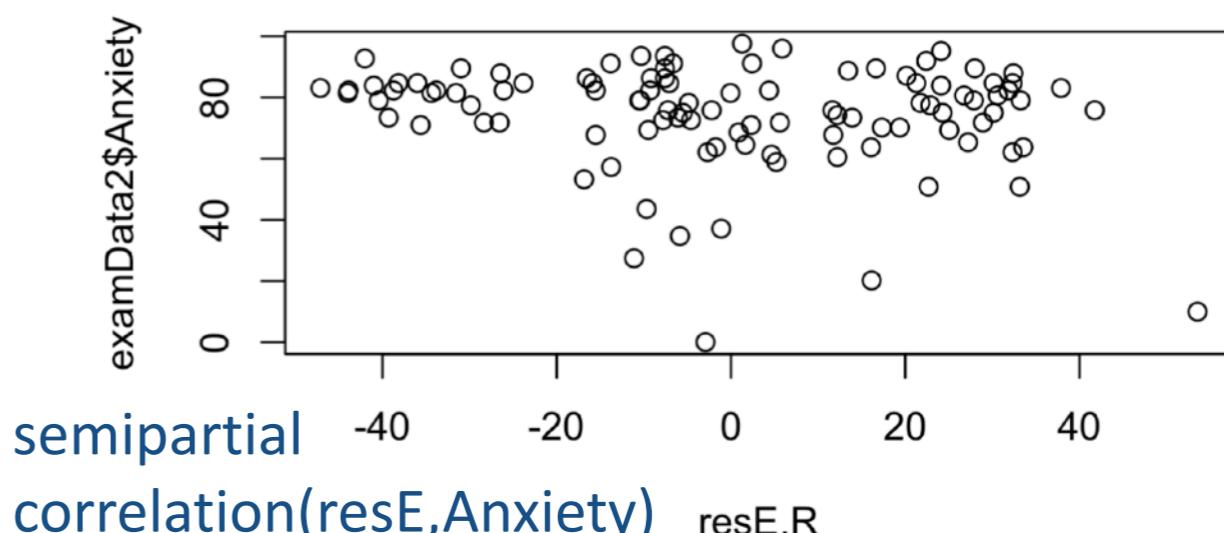
```
> ppcor::spcor.test(examData2$Anxiety, examData2$Exam, examData$Revise)
   estimate    p.value statistic   n gp Method
1 -0.173889 0.08048295 -1.765791 103  1 pearson
```

- Method (3/3): using Regression to “*regress out*” the effect of Revise time from Exam; then use the residuals to conduct the correlation analysis.

```
resE.R <- residuals(lm(Exam ~ Revise, data = examData2))
> cor.test(resE.R, examData2$Anxiety)
```

Pearson's product-moment correlation

```
data: resE.R and examData2$Anxiety
t = -1.7746, df = 101, p-value = 0.07898
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.35545205  0.02031946
sample estimates:
cor
-0.173889
```



# Comparing Correlations

- What if we want to compare 2 correlations?

③  
Testing

General form: *psych*

- ▶ *Independent-sample case*— `r.test(n1, r1, r2, n2, pooled=T, twotailed=T)`
- ▶ *Dependent-sample case*— `r.test(n, r12, r34, r13, r24)`

$$z_{\text{Difference}} = \frac{z_{r_1} - z_{r_2}}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}}$$

- Tests the significance of the difference between **two independent correlations**.
- Exp: ExamAnxiety between Male & Female.  
`psych::r.test(52, -.506, -.381, n2=51).`
- the difference **between two dependent correlations** with different variables
- Exp: Exam-Anxiety stronger than Exam-Revision?



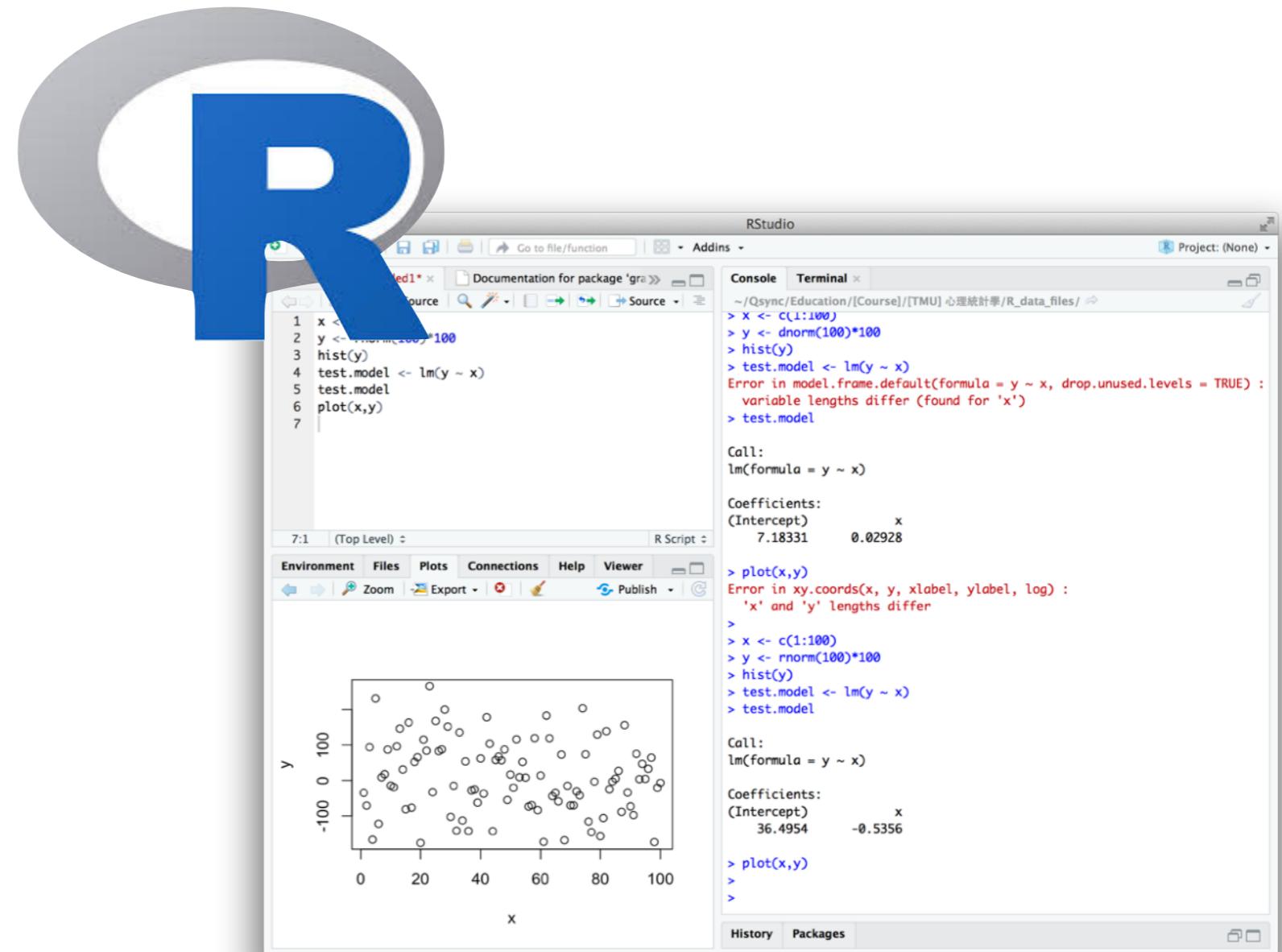
臺北醫學大學  
TAIPEI MEDICAL UNIVERSITY

<https://www.personality-project.org/r/psych/help/r.test.html>

<http://comparingcorrelations.org/>

Im  
mediate  
sem

# ③ MEDIATION ANALYSIS

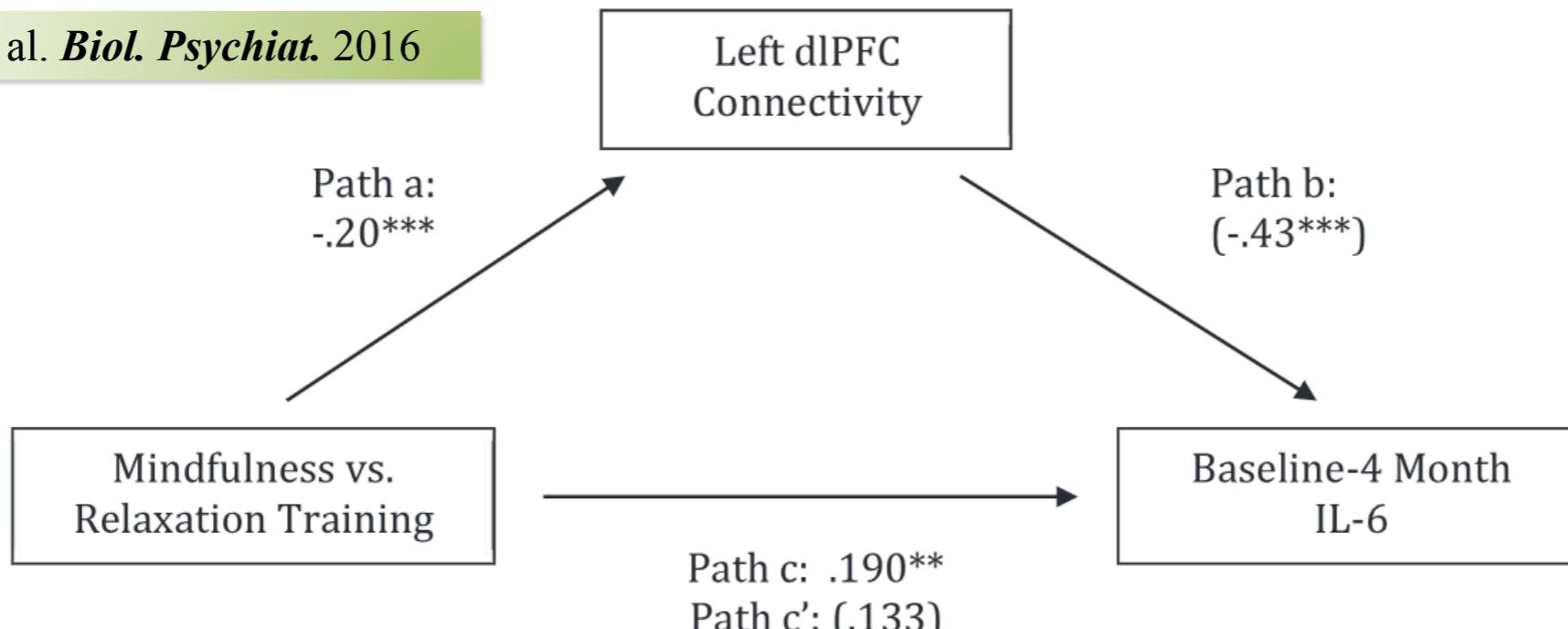


臺北醫學大學  
TAIPEI MEDICAL UNIVERSITY

# Mediation Analysis

- Correlation / Regression does not provide **Causality** !
  - Predictor (**X**) and Outcome (**Y**) are exchangeable in general.
- **Q: What mechanism enables X to cause Y ?**
  - *Causal Mediation Analysis* (be careful with the term “causal”)
  - Example: *Mindfulness* reduces the *inflammation*, which was intervened by the *brain function*.

Creswell et al. *Biol. Psychiat.* 2016

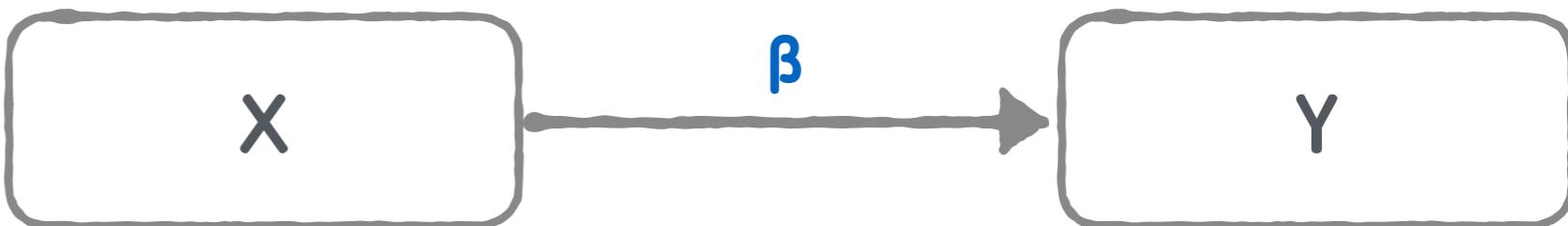


The alterations in brain connectivity statistically mediated mindfulness training improvements in IL-6 at 4-month followup.

**Figure 2.** Interleukin (IL)-6 mediation analyses. Increases in left dorsolateral prefrontal cortex (dlPFC) connectivity (Montreal Neurological Institute: -22, 52, 10) significantly mediate **(A)** the time  $\times$  treatment interaction on circulating (log transformed) IL-6. Increases in right dlPFC (Montreal Neurological Institute: 26, 42, 38) marginally significantly mediate **(B)** IL-6 effects. Numbers represent  $b$  coefficients from mixed-effect linear models, with parentheses representing  $b$  coefficients when the main effect and time  $\times$  treatment condition interaction terms and dlPFC connectivity parameter estimates are entered in a mixed-effect linear model simultaneously. \* $p = .06$ ; \*\* $p = .05$ ; \*\*\* $p < .05$ .

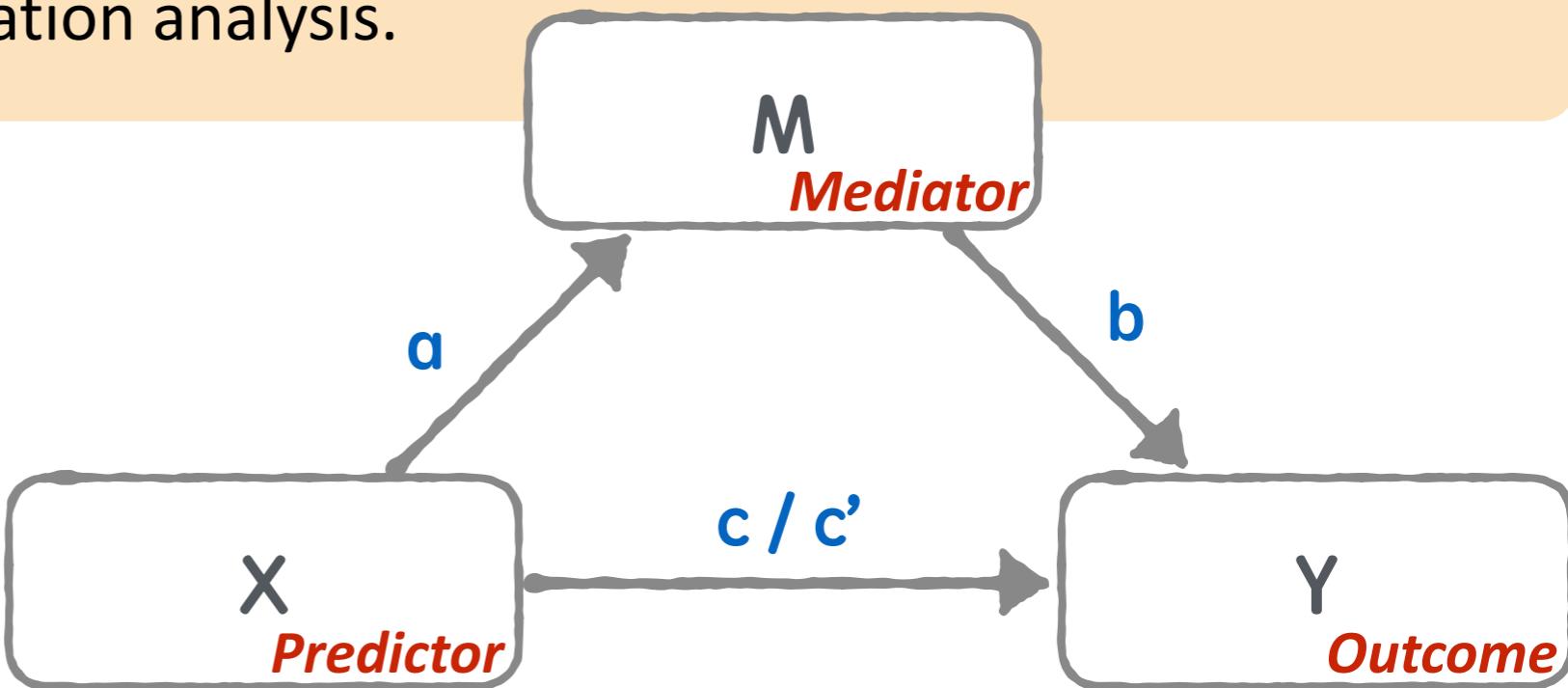
# Mediation Analysis

Original Linear Regression:



- **Mediation Analysis** is used to investigate whether one (or more than one) variable (M) transmits/intervenes the effects of a predictor variable (X) on an outcome variable (Y).
- Usually present in a path diagram.
- Both multiple regression and structural equation modeling (SEM, or path analysis) can be used to perform mediation analysis.

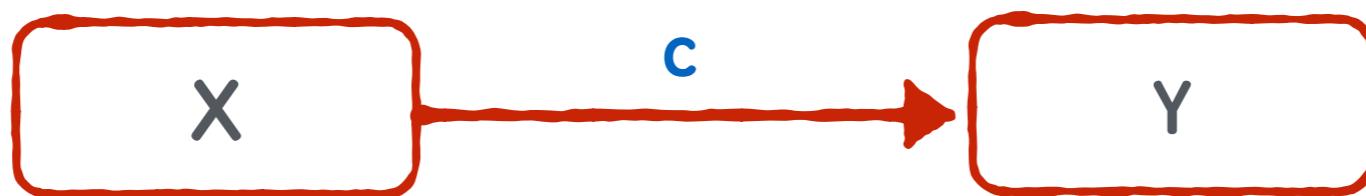
Mediation Analysis:



# Mediation: 3 Linear Models

$$Y = (\text{intercept}_1) + c(X) + e_1$$

- a, b, c, c' are used in this context to represent the regression/path coefficients.
- e represents the residual error.

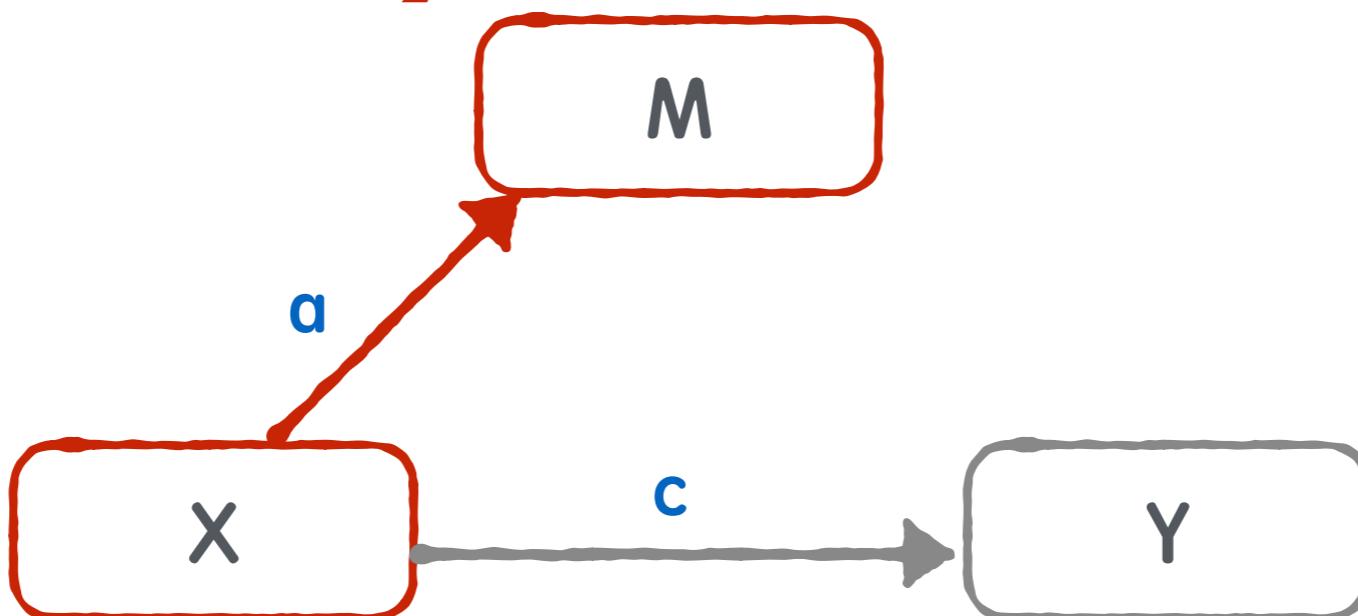


- Total effect ( $c$ )

# Mediation: 3 Linear Models

$$Y = (\text{intercept}_1) + c(X) + e_1$$

$$M = (\text{intercept}_2) + a(X) + e_2$$



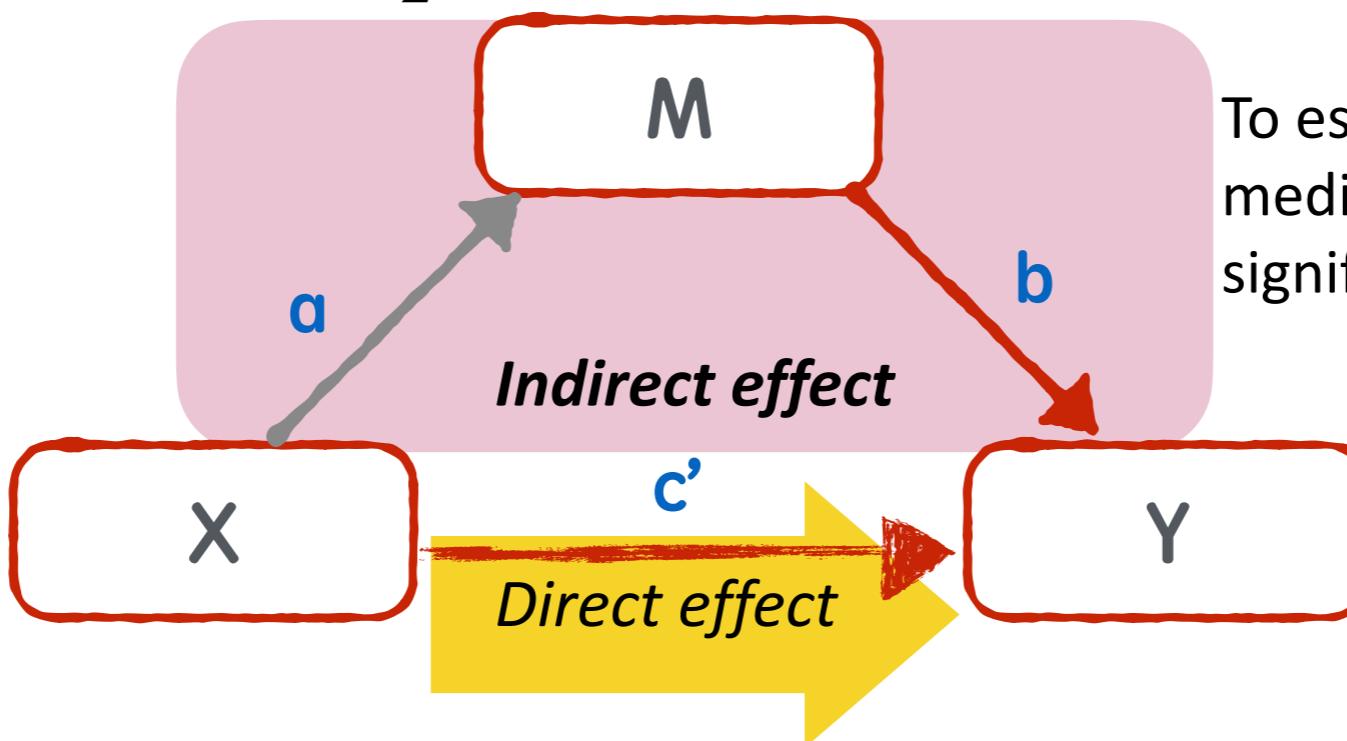
- Total effect ( $c$ )
- Path A ( $a$ )

# Mediation: 3 Linear Models

$$Y = (\text{intercept}_1) + c(X) + e_1$$

$$Y = (\text{intercept}_3) + c'(X) + b(M) + e_3$$

$$M = (\text{intercept}_2) + a(X) + e_2$$



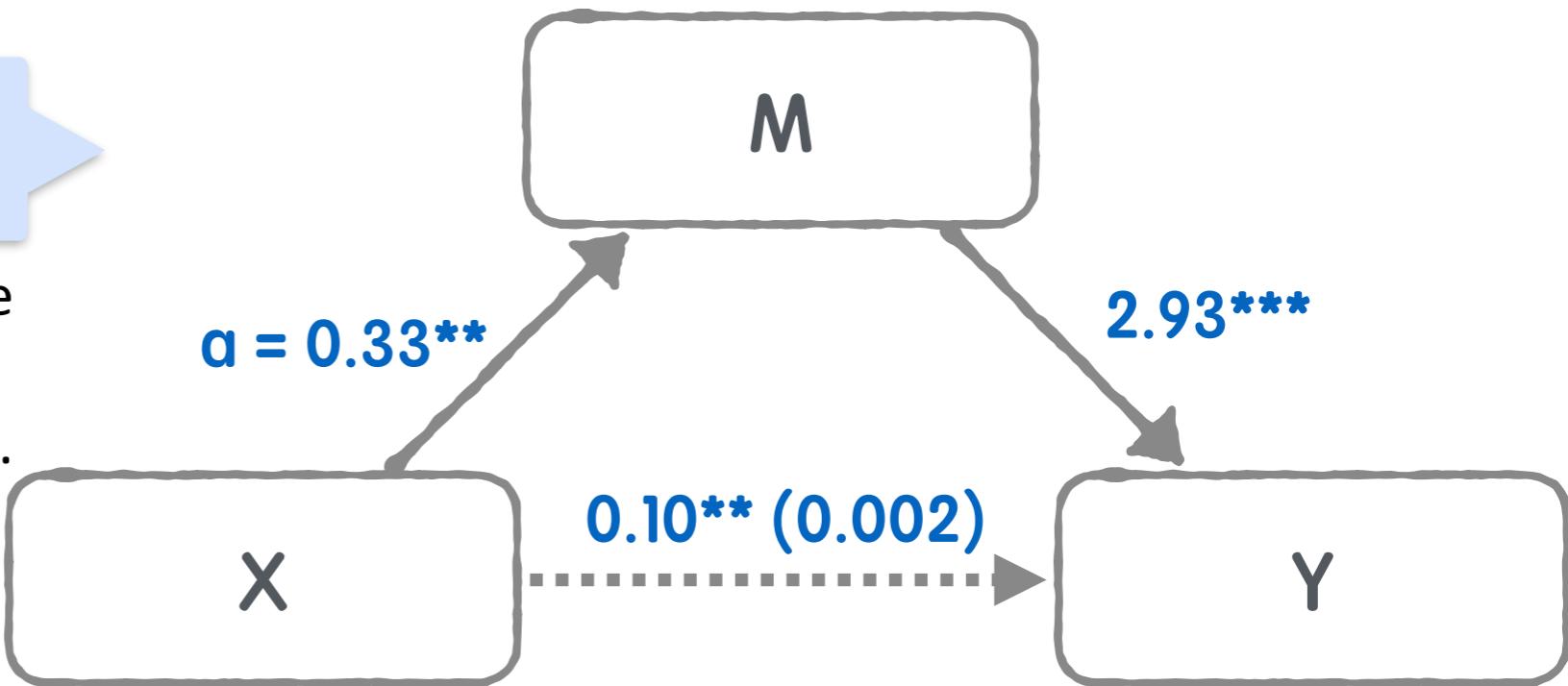
To estimate the presence of mediation, we test the significance of “**Indirect effect**”.

- Total effect ( $c$ )
- Direct effect ( $c'$ )
- Indirect effect ( $a*b$ )

# 2 Types of Mediation

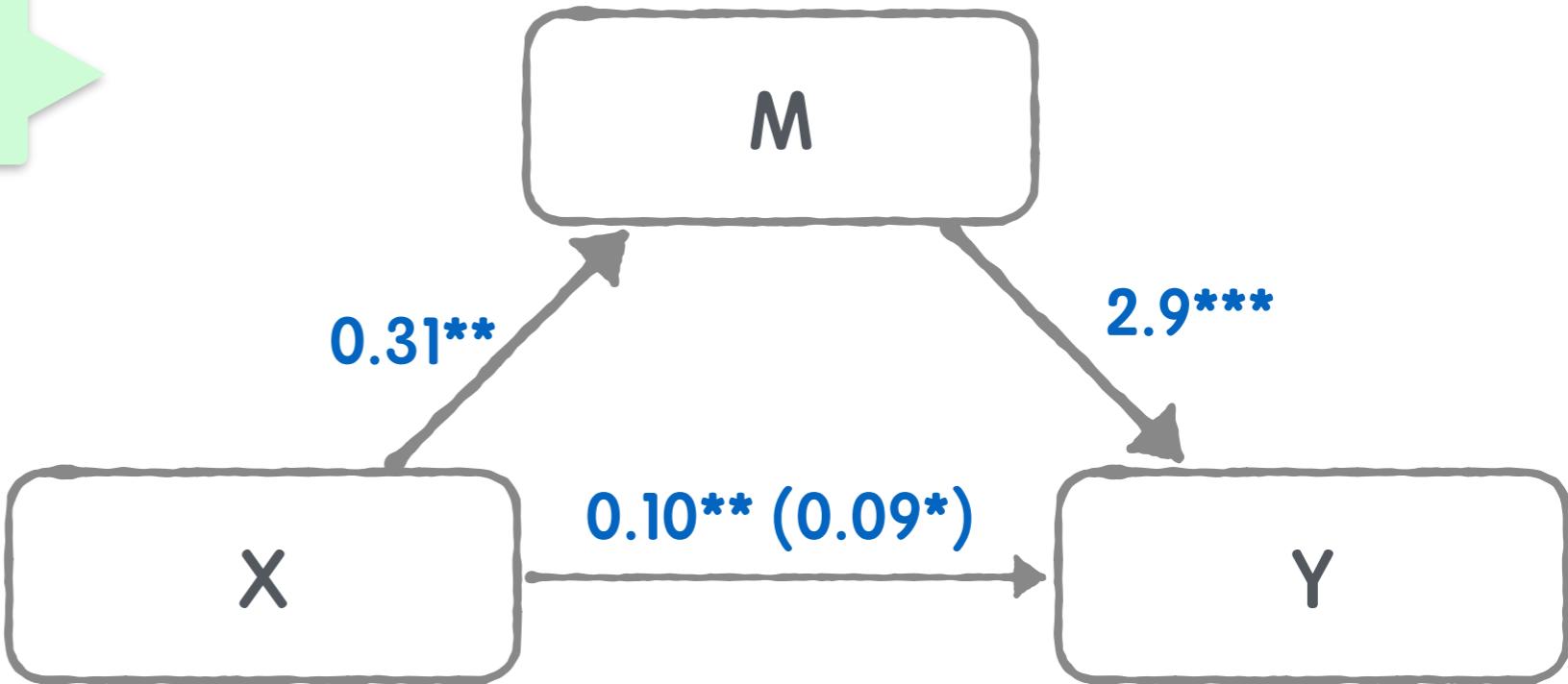
## Complete Mediation

The direct relation ( $c'$ ) between the predictor ( $X$ ) and the outcome variable ( $Y$ ) is zero — Full mediation.



## Partial Mediation

The direct relation ( $c'$ ) between  $X$  and  $Y$  is not zero — the mediator only explains this relation partially.



# Estimation of Indirect Effect

- **Product of Coefficients**: indirect effect is estimated by computing the product of the coefficients for path  $a$  and  $b$ , which represent the relation between X and M and the relation between M and Y, respectively.
  - Significance of indirect effect — dividing the product of coefficients by its standard error and the resulting ratio is compared to a standard normal distribution.
  - **Delta Method (Sobel Test)**: the approach for estimating standard error for product of coefficients and comparing to standard normal distribution to determine  $p$ -value — tends to show bias in small samples (e.g.,  $n < 50$ ) and assumes sampling distribution of indirect effect is normal, which is not often tenable.
  - **Resampling Method**: the approach, such as percentile bootstrapping, in which the normal distribution is not presumed but resampling still tends to show bias in small samples (e.g.,  $n < 20$ ).



## DEMO

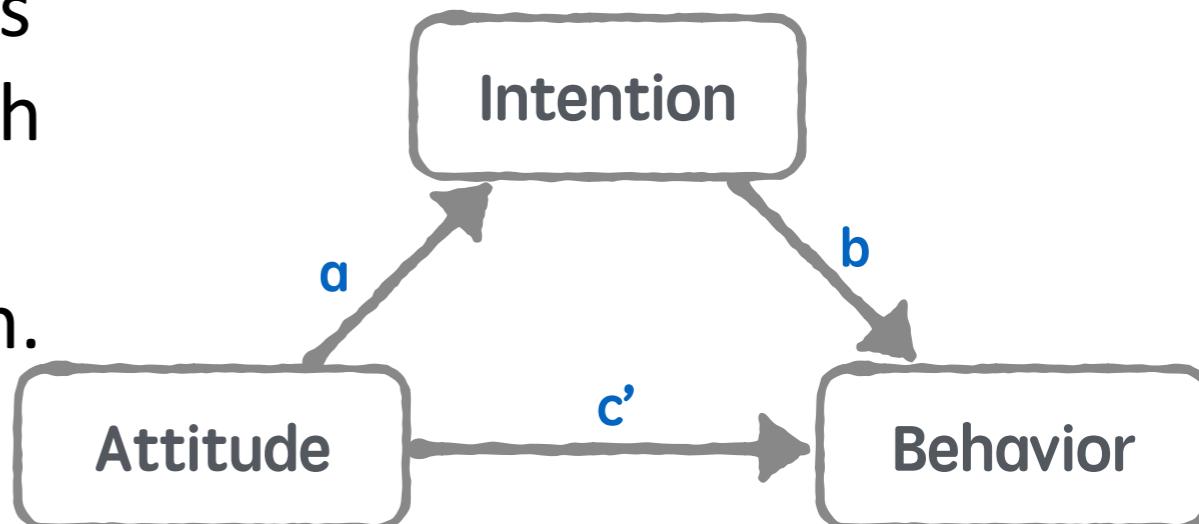
# Attitude, Intention, Behavior

- **Background:** Psychologists would like to test the relationship between the attitude, intention and planned behavior. They regard the intention as the mediator and formulated the path diagram like this. Please use the collected data to verify the mediation.

**Outcome measure: Behavior**

**Predictor: Attitude**

**Mediator: Intention**



- **Set up Hypothesis:**

①  
Hypothesis

$H_0: \text{mediation} = 0$   
 $H_a: \text{mediation} \neq 0$

- **Data import:**

► `reg <- read.csv("PlannedBehavior.csv", header = TRUE)`



臺北醫學大學  
TAIPEI MEDICAL UNIVERSITY

Load:  
PlannedBehavior.csv

## DEMO

# Attitude, Intention, Behavior

- Regression (1): total effect

▶ *c\_path = lm(behavior~attitude,  
data=Mediation\_data)*

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	2.33552	0.22214	10.514	< 2e-16 ***
attitude	0.24126	0.06688	3.608	0.000392 ***

Total effect = **c = 0.241\*\*\***

- Regression (2): Path A

▶ *a\_path = lm(intention~attitude,  
data=Mediation\_data)*

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1.46083	0.19393	7.533	1.76e-12 ***
attitude	0.48405	0.05838	8.291	1.73e-14 ***

**a = 0.484\*\*\***

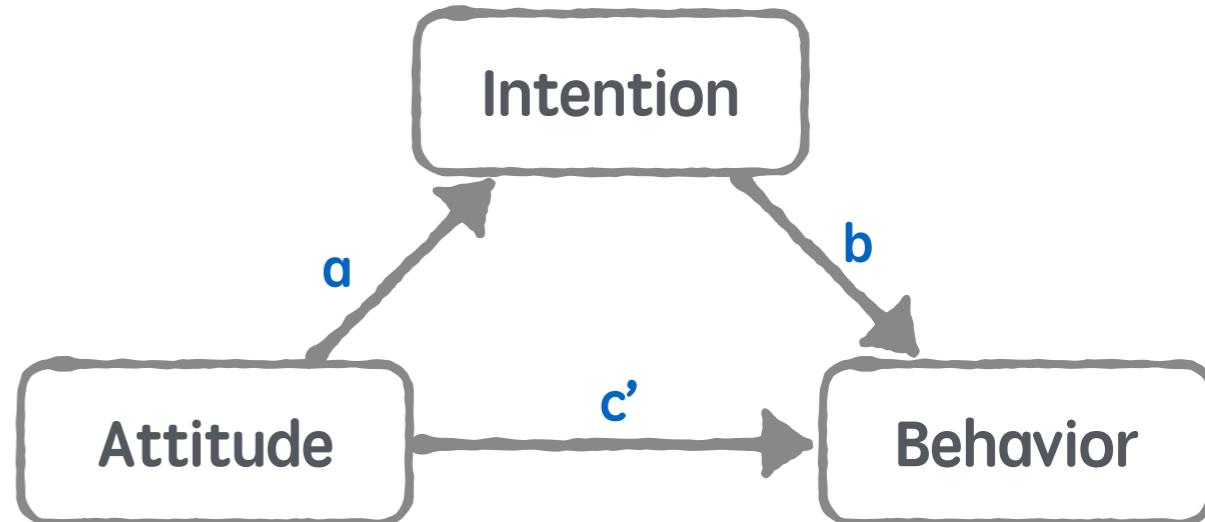
- Regression (3): Path B & C'

▶ *b\_path = lm(behavior~intention+attitude,  
data=Mediation\_data)*

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1.69617	0.23358	7.262	8.79e-12 ***
intention	0.43767	0.07561	5.788	2.78e-08 ***
attitude	0.02941	0.07196	0.409	0.683

**b = 0.438\*\*\***

**c' = 0.029 (n.s.)**



## DEMO

# Attitude, Intention, Behavior

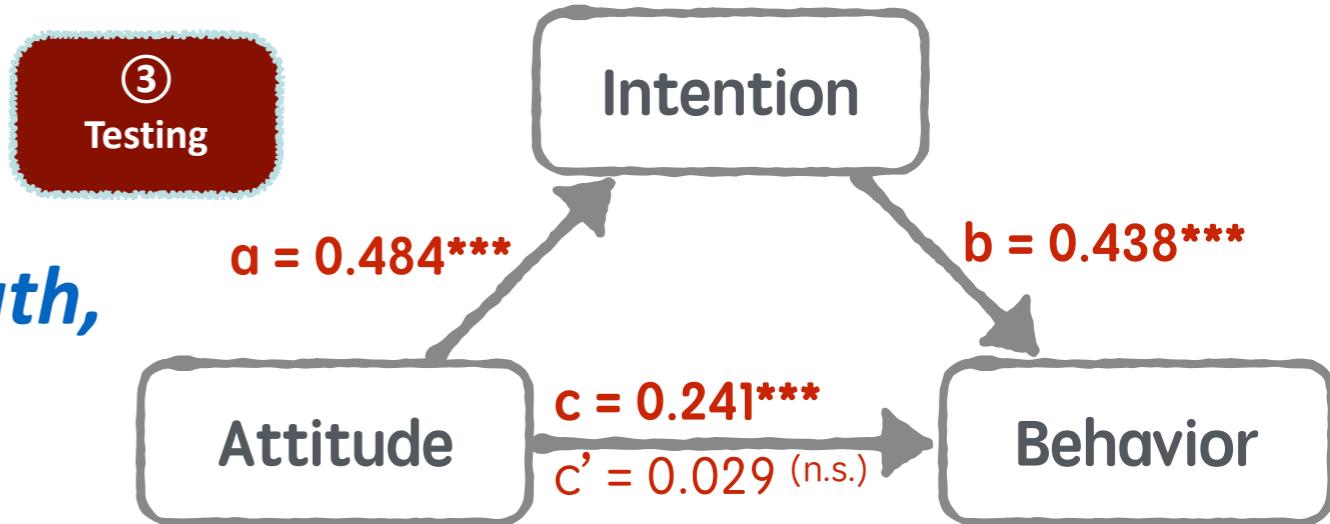
- Test the indirect effect:

- ▶ *library(mediation)*
- ▶ *results = mediate(a\_path, b\_path, sims=5000, treat='attitude', mediator='intention', boot=T)*
- ▶ *summary(results)*

- Results:

	Estimate	95% CI Lower	95% CI Upper	p-value	
ACME	0.2119	0.1324	0.30	<2e-16	***
ADE	0.0294	-0.0943	0.16	0.64	
Total Effect	0.2413	0.1163	0.37	<2e-16	***
Prop. Mediated	0.8781	0.5192	1.69	<2e-16	***
---					
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Sample Size Used: 199



- Total effect ( $c$ )
- Average direct effect ( $c'$ )
- Average causal mediation effect ( $a*b$ )
- Proportion mediated:  $ACME/Total\ effect.$

**DEMO**

# Attitude, Intention, Behavior

④

Effect Size

- **Effect size:**

$$\text{proportion mediated} = 1 - \left(\frac{c'}{c}\right) = \frac{ab}{ab + c'}$$

Proportion mediated is unstable unless sample size is at least 500, so this is optional for the report.

Proportion mediated: **ACME/Total effect.**

Prop. Mediated 0.8781

0.5192

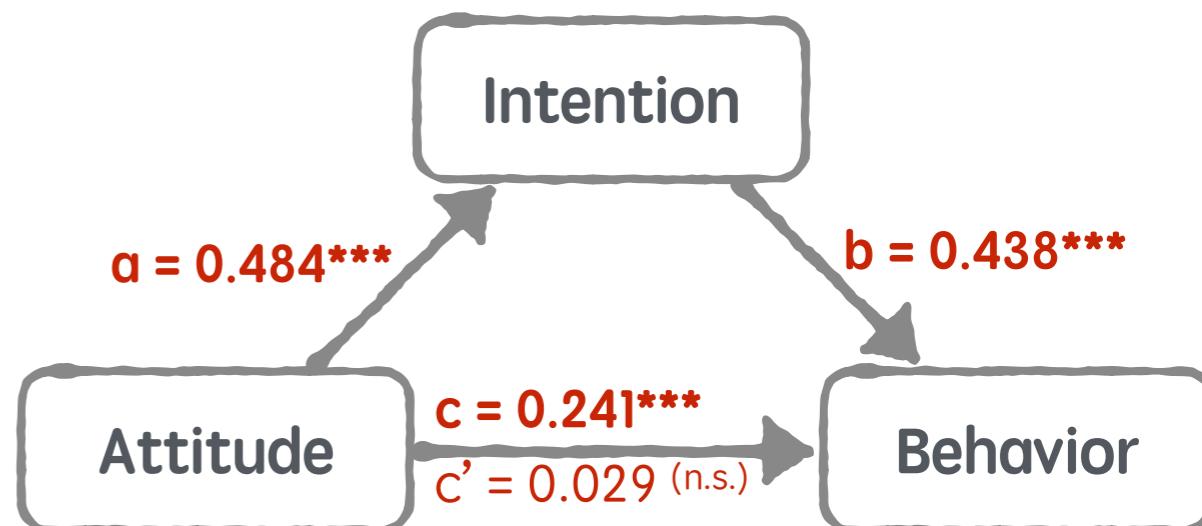
1.69

⑤

Decision

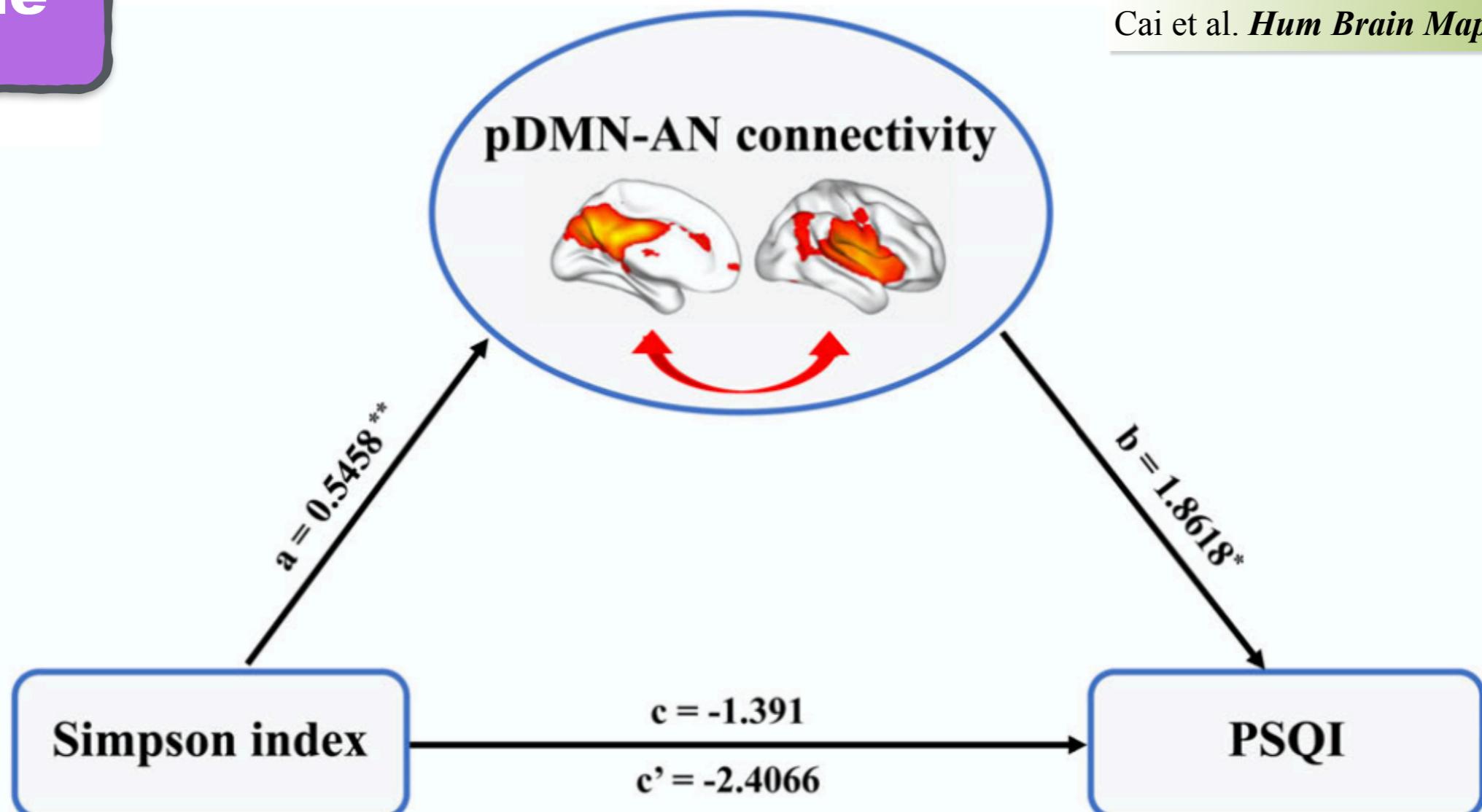
- **Reporting decision:**

**Based on the 199 observations, the Intention acts as a complete mediator on the relationship between Attitude and Planned behavior (indirect effect = 0.212, C.I.[0.132, 0.30],  $p < 0.001$ ).**



## Example

Cai et al. *Hum Brain Mapp* 2021



- Test whether the relation between microbiota and sleep quality was intervened by brain connectivity.
- Indirect effect was significant, but the direct relation was not obvious.

# Discussion

## Mediation Analysis

David P. MacKinnon, Amanda J. Fairchild,  
and Matthew S. Fritz

Department of Psychology, Arizona State University, Tempe, Arizona 85287-1104;  
email: david.mackinnon@asu.edu, amanda.fairchild@asu.edu, matt.fritz@asu.edu

### 1. Nonlinear Least Squares

- Model fitting of nonlinear relationships

### 2. Partial / Semipartial Correlation

- Comparing multiple correlations

### 3. Causal Mediation Analysis

- Indirect effect from the mediator

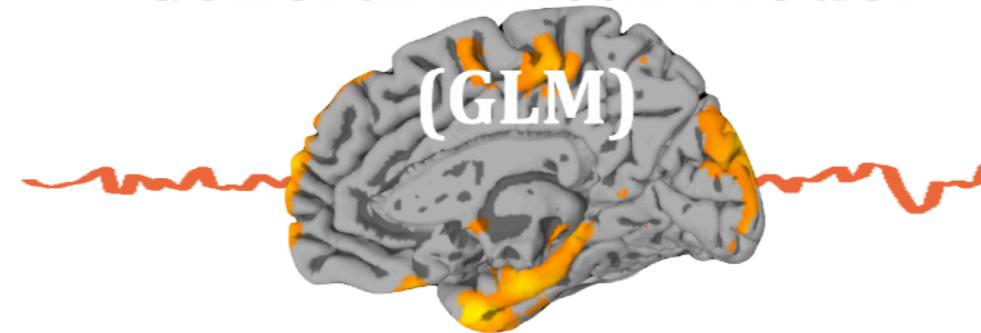
Regression assumptions  
Residual error  
Random & Fixed effects

# GENERAL LINEAR MODEL IN FMRI

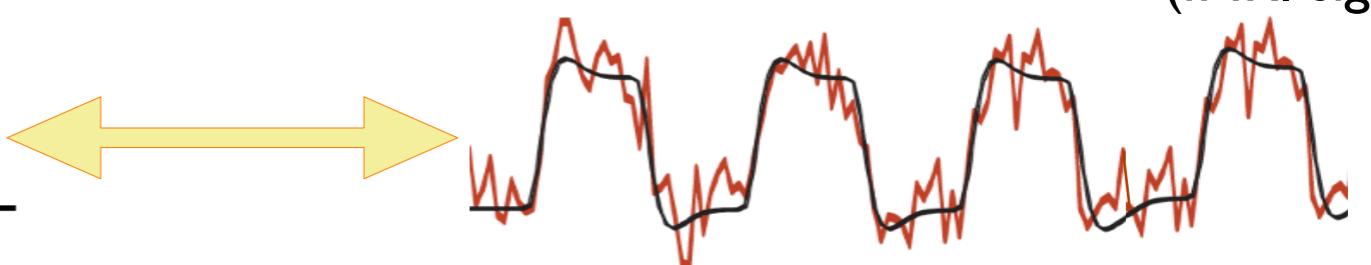
## Functional MRI

Week 10

General Linear Model



Paradigm  
(Experimental design)

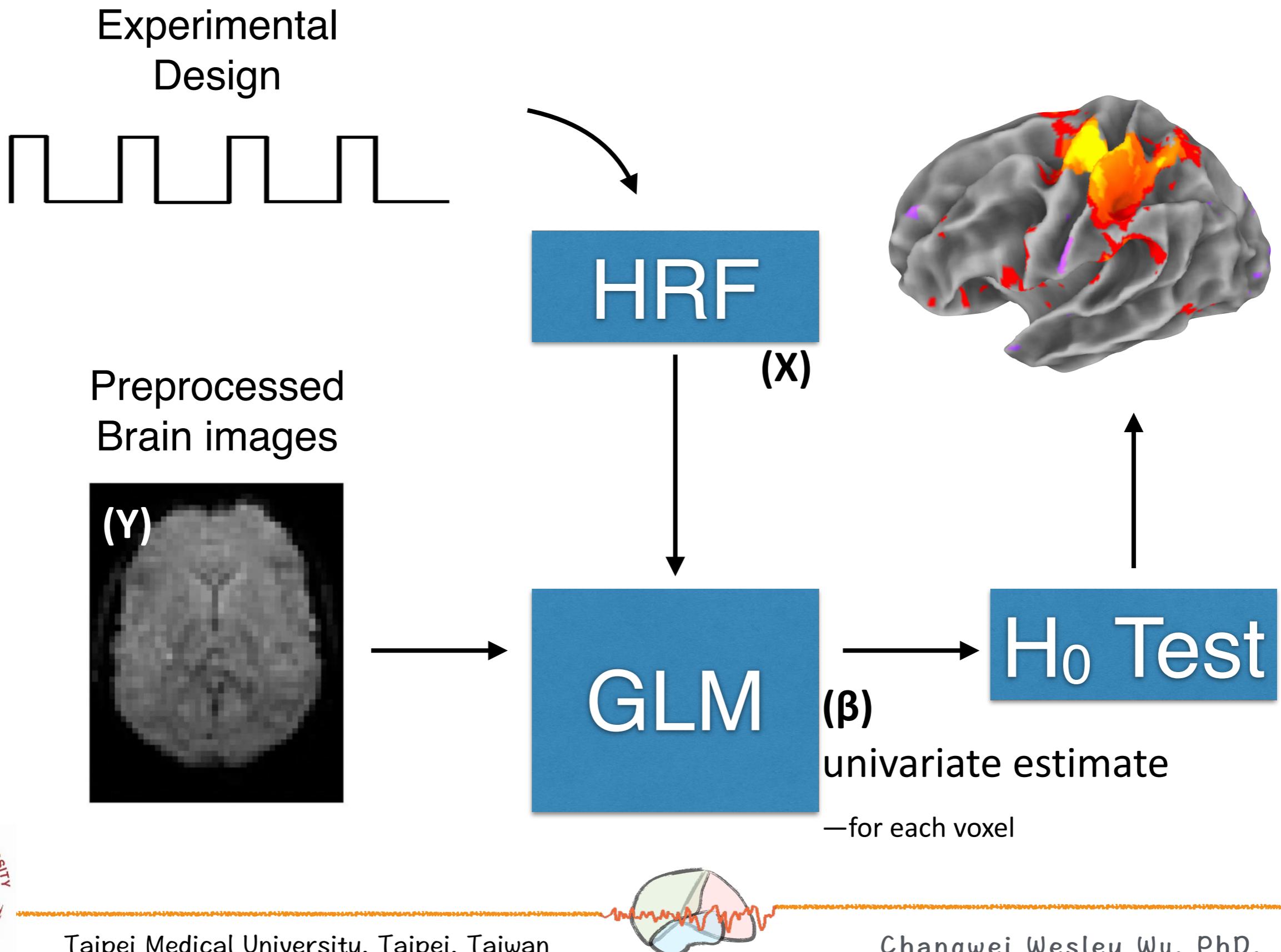


BOLD Response  
(fMRI signal)



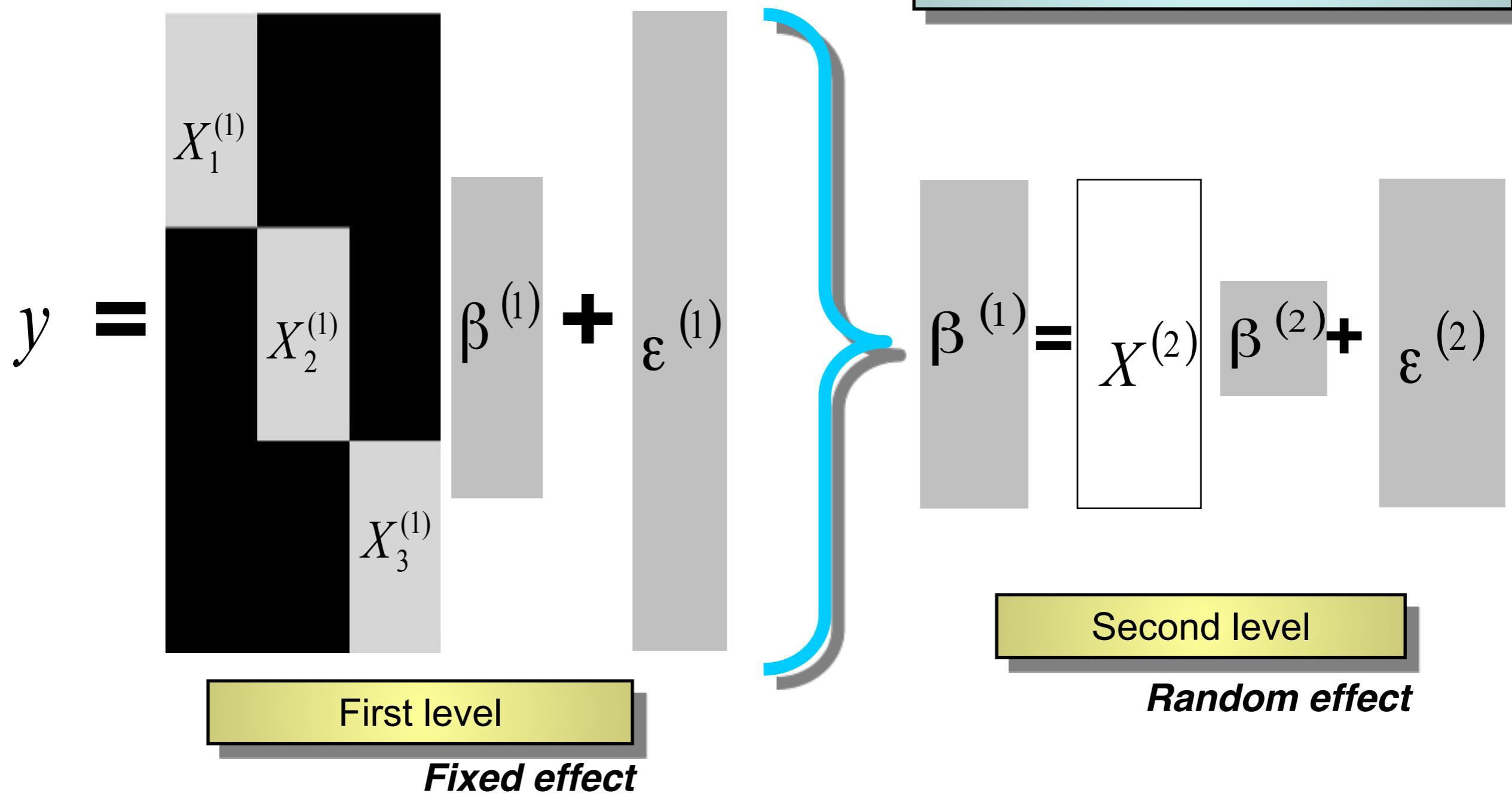
臺北醫學大學  
TAIPEI MEDICAL UNIVERSITY

# Recap: Calculating Brain “Activation”



# Hierarchical models

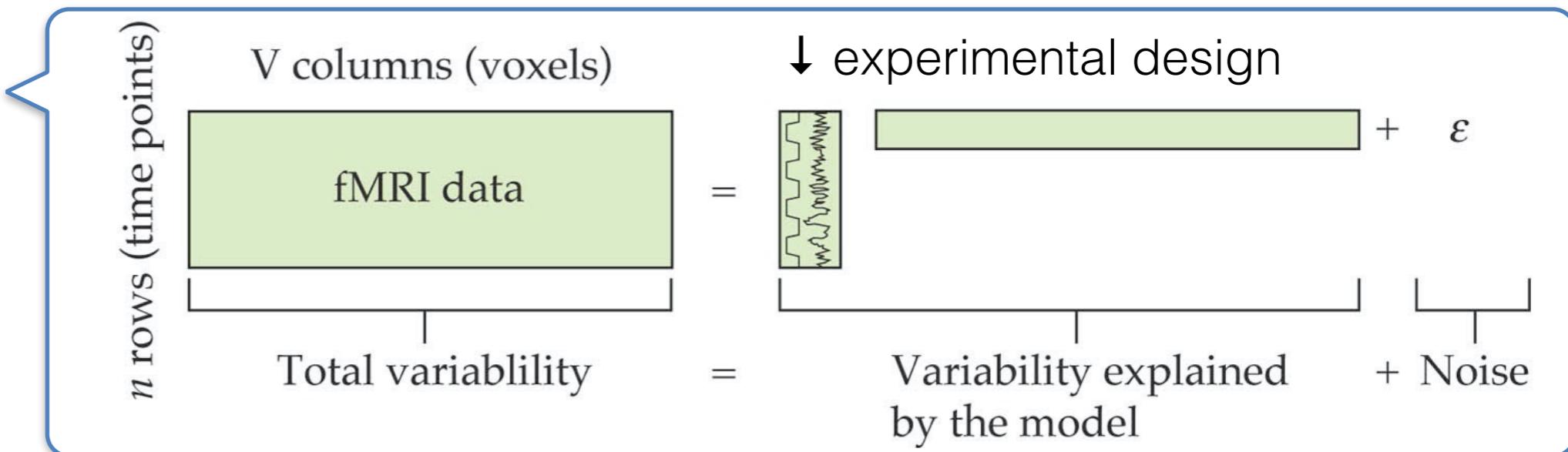
Example: Two level model



# Testing Effect Size ( $\beta$ )

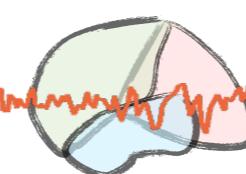
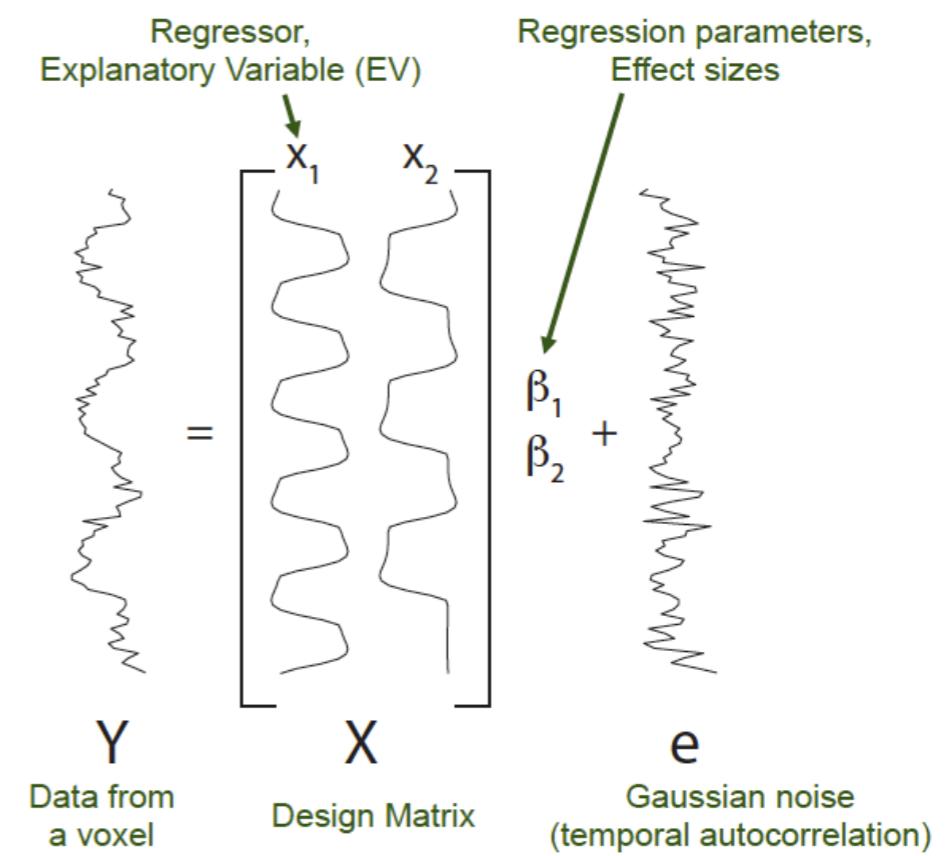
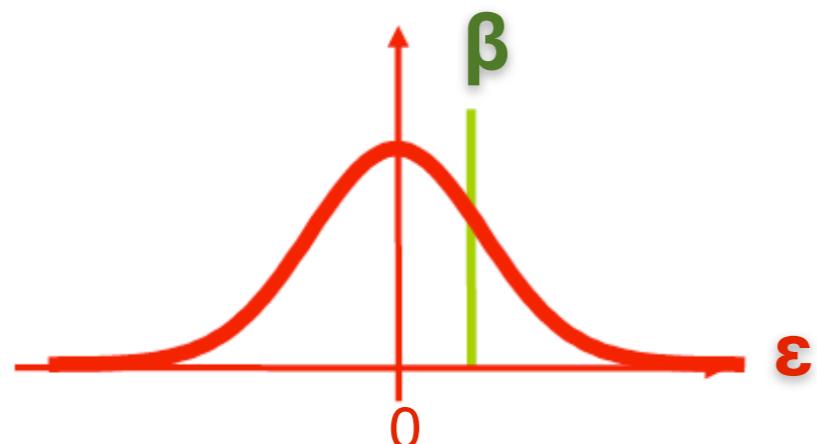
$$Y = X\beta + \epsilon$$

$$\begin{aligned} H_0 : c'\beta &= 0 \\ H_1 : c'\beta &\neq 0 \end{aligned}$$



## ► Every regressor has its own $\beta$ estimate.

- How to test the stat. significance?
- If  $\beta$  is non-zero, then voxel is active.
- every voxel has its  $\beta$  value →  $\beta$  maps



# Hypothesis Testing in fMRI (1st-level)

## Regression in R code

Coefficients:		$\beta$	Multiple Regression			
(Intercept)	-26.612958		Estimate	Std. Error	t value	Pr(> t )
adverts	0.084885			17.350001	-1.534	0.127
airplay	3.367425			0.006923	12.261	< 2e-16 ***
attract	11.086335			0.277771	12.123	< 2e-16 ***
---				2.437849	4.548	9.49e-06 ***
Signif. codes:	0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

The standard error of  $b$  is

$$SE(\beta) = \sqrt{\frac{MS_{err}}{SS_x}}$$

## Statistical Hypothesis:

- Null Hypothesis:  $H_0 : c'\beta = 0$   
Is  $c^T\beta$  Non-Zero?

$$H_0 : c'\beta = 0$$

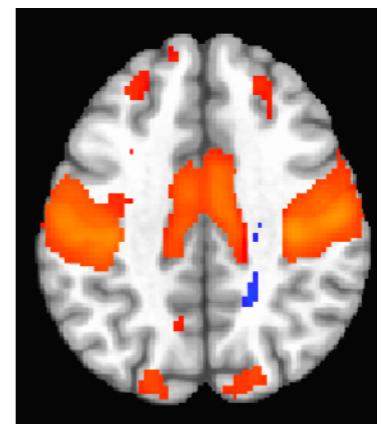
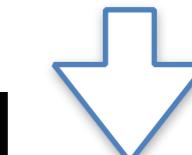
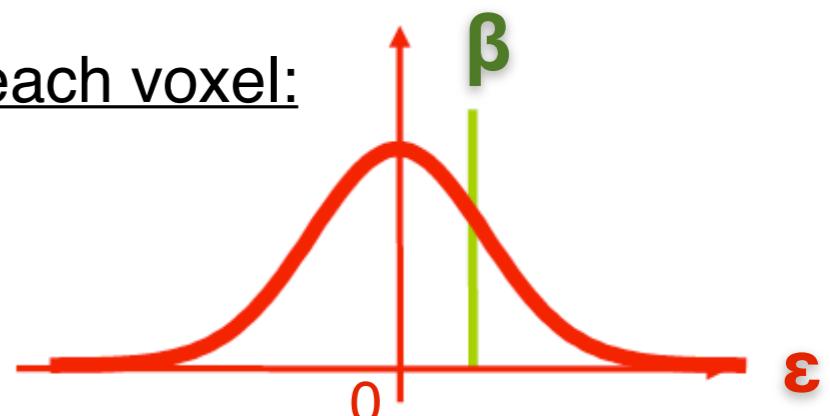
$$H_1 : c'\beta \neq 0$$

$$t = \frac{\hat{\beta}}{\widehat{SE}(\beta)}$$

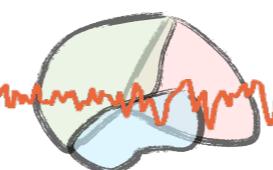
$$\begin{aligned} t &= \frac{c^T \hat{\beta} - 0}{\sqrt{\text{Var}(c^T \hat{\beta})}} = \frac{c^T \hat{\beta}}{\sqrt{\text{Var}(c^T (X^T X)^{-1} X^T Y)}} \\ &= \frac{c^T \hat{\beta}}{\sqrt{[c^T (X^T X)^{-1} X^T] \sum_{\epsilon} [c^T (X^T X)^{-1} X^T]^T}} \end{aligned}$$

$$dof = N - r$$

For each voxel:



**Fixed effect**  
for each subject  
(1<sup>st</sup>-level in SPM)  
(not generalizable)



# Gauss-Markov Theorem

**[Gauss-Markov Theorem]** Within GLM, if  $\Sigma_\epsilon = \sigma_\epsilon^2 I$ , then the method of least squares always produces minimum variance unbiased estimators of  $\beta$ .

t-statistic      
$$t = \frac{c^T \hat{\beta}}{\sqrt{[c^T (X^T X)^{-1} X^T] \Sigma_\epsilon [c^T (X^T X)^{-1} X^T]^T}}$$

- Assumption of Noise  
*(Temporal autocorrelation):*
1. Homogeneity across TRs
  2. Independence across TRs.

*Homogeneity of variances (same  $\sigma_\epsilon$ )*

$$\Sigma_\epsilon = \begin{bmatrix} \sigma_\epsilon^2 & 0 & 0 & \dots \\ 0 & \sigma_\epsilon^2 & 0 & \dots \\ 0 & 0 & \sigma_\epsilon^2 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

*Independence*

fMRI data is easily to break the assumptions  
 →  $\beta$ s are not generalizable to entire population

$\beta^{(1)}$ : First level  $\beta$  for each participant

# Gauss-Markov Theorem

## [Final Hypothesis Test]

$H_0: \mathbf{c}^T \boldsymbol{\beta} = \mathbf{0}$ , where  $\mathbf{c}$  is the COPE (e.g.,  $\mathbf{c}' = [0 \underline{0} 1 \underline{1} 0 -1 -1 0 \underline{0}]$ )

$H_1: \mathbf{c}^T \boldsymbol{\beta} \neq \mathbf{0}$

**[Proposition 1]** Consider any version of the GLM that assumes **homogeneity and independence of variance** (i.e.,  $\Sigma_{\epsilon} = \sigma_{\epsilon}^2 I$ ). Suppose  $\mathbf{c}^T$  is a row vector of constants, then under the null hypothesis  $H_0: \mathbf{c}^T \boldsymbol{\beta} = \mathbf{0}$ , the following statistic would be:

## Student's t-test (simple form, $\mathbf{c}=1$ )

**t-statistic**

$$t = \frac{\widehat{c' \beta}}{\widehat{SE(c' \beta)}}$$

- **Amplitude:** higher  $\beta \rightarrow$  high t
- **Variation:** high  $\text{std}(\epsilon) \rightarrow$  low t

$$\begin{aligned}
 \text{Var}(\underline{c}' \hat{\underline{\beta}}_G) &= \text{Var}\left[\frac{1}{M} \sum_{i=1}^M \underline{c}' (\mathbf{X}'_i \mathbf{X}_i)^{-1} \mathbf{X}'_i \underline{B}_i\right] \\
 &= \frac{1}{M^2} \sum_{i=1}^M \text{Var}[\underline{c}' (\mathbf{X}'_i \mathbf{X}_i)^{-1} \mathbf{X}'_i \underline{B}_i] \\
 &= \frac{1}{M^2} \sum_{i=1}^M \underline{c}' (\mathbf{X}'_i \mathbf{X}_i)^{-1} \mathbf{X}'_i \Sigma_i \mathbf{X}_i (\mathbf{X}'_i \mathbf{X}_i)^{-1} \underline{c} \\
 &= \frac{1}{M^2} \sum_{i=1}^M \underline{c}' (\mathbf{X}'_i \mathbf{X}_i)^{-1} \mathbf{X}'_i [\sigma_w^2 I] \mathbf{X}_i (\mathbf{X}'_i \mathbf{X}_i)^{-1} \underline{c} \\
 &= \frac{\sigma_w^2}{M^2} \sum_{i=1}^M \underline{c}' (\mathbf{X}'_i \mathbf{X}_i)^{-1} (\mathbf{X}'_i \mathbf{X}_i) (\mathbf{X}'_i \mathbf{X}_i)^{-1} \underline{c} \\
 &= \frac{\sigma_w^2}{M^2} \sum_{i=1}^M \underline{c}' (\mathbf{X}'_i \mathbf{X}_i)^{-1} \underline{c}.
 \end{aligned}$$

p.s. here  $B_i = Y_i$  (fMRI data)



# THANK YOU FOR YOUR ATTENTION

E-mail: [sleepbrain@tmu.edu.tw](mailto:sleepbrain@tmu.edu.tw)

