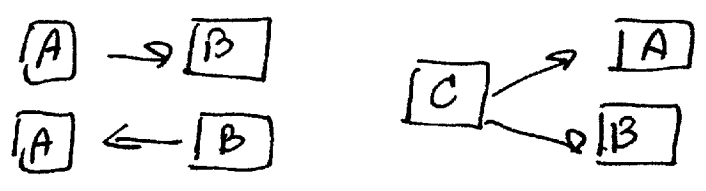


ความเป็นเหตุเป็นผล (Causality)

ความเป็นเหตุเป็นผลเป็นสิ่งที่แยกไม่ออกจากศาสตร์ทางปรัชญา และสถิติ ความเป็นเหตุเป็นผลมีลักษณะ = 0 อยู่ 3 ประการ คือ

- (1) ความสัมพันธ์
- (2) กัดทาวความสัมพันธ์
- (3) ความเชื่อเอกเทศ

ความสัมพันธ์และกัดทาวความสัมพันธ์มีจำกัดเฉพาะรูปแบบเช่น



ความสัมพันธ์ก่อนและหลังนั้น เป็นอย่างไรขึ้นอยู่กับ

- (1) ทฤษฎี (Theory)
- (2) Previous Research
- (3) Common sense

- ดังนั้น ความสัมพันธ์ที่เกิดขึ้นมา เป็นเหตุเป็นผลตามที่มีพื้นฐานมาจากทฤษฎี หรืองานวิจัยที่มีอยู่ก่อนหน้า คัดค้านจากงานนี้ นอกจากนั้น ความสัมพันธ์ที่เกิดขึ้นมาอาจเป็นเอกเทศในบางกรณี

ถ้าเมื่อเปรียบเทียบกับอิทธิพลต่อ หรือเปรียบเทียบ The 1981 ینگ จากตัวแปรอื่น นี่ก็คือ เป็นความสัมพันธ์ที่ = ความสูง 100% ขึ้นที่อาจมีผลต่อตัวแปรอื่นด้วย

บทบาทของทฤษฎีในการกำเนตความเป็นเหตุเป็นผล

- ตัวอย่างเช่น - ศึกษากับราคาหลักทรัพย์ - การเปลี่ยนแปลงกับจำนวนชั่วโมงที่อ่านหนังสือ

- ความรู้ของตัวแปรอื่นกับ GDP ของประเทศ

ความสัมพันธ์ของตัวแปรอื่นที่แปรตามบ้างกับ 0.10001 มีความสัมพันธ์กับตัวแปรอื่นทางสถิติ

๙๑=พบว่า ความสัมพันธ์ที่เราไม่พบเหตุเป็นผลกัน
กรณีความสัมพันธ์เช่นนี้ = เรียกว่า "Spurious relationships"
การสรุปว่า "๑๒ มีผลต่อ ๑๓" จึงเป็นกรณีเหตุผลผิดพลาด
หรือ ผิดตรรก (fallacy)

กรณีภาพทฤษฎี: หรือ ย่อช่วงเวลาในอดีต ประเด็น ก่อเกิดกับ
ข้อมูล: ข้อสรุปที่เราได้ทราบมา ต่อให้ประวัติ: โดยที่ ที่: ๑๒ ๑๓
มาพิจารณา แนวโน้มที่: ทราบแล้ว โมเดลที่: แนว: ๑๒ ๑๓

โมเดล (Model) หรือ แนวโน้ม

- คำทวน 1) "คุณคิดว่า ๑๒ เปลี่ยนในตอนนี้ เป็นเท่าไร"
- ๑๒ ๑๓
- ๒) "คุณเอาจากข้อมูล ๑๒ ๑๓"

"ความพยายามในจิตตาคือ ๑๒ ๑๓ ๑๔ ๑๕ ๑๖ ๑๗ ๑๘ ๑๙ ๒๐ ๒๑ ๒๒ ๒๓ ๒๔ ๒๕ ๒๖ ๒๗ ๒๘ ๒๙ ๓๐ ๓๑ ๓๒ ๓๓ ๓๔ ๓๕ ๓๖ ๓๗ ๓๘ ๓๙ ๔๐ ๔๑ ๔๒ ๔๓ ๔๔ ๔๕ ๔๖ ๔๗ ๔๘ ๔๙ ๕๐ ๕๑ ๕๒ ๕๓ ๕๔ ๕๕ ๕๖ ๕๗ ๕๘ ๕๙ ๖๐ ๖๑ ๖๒ ๖๓ ๖๔ ๖๕ ๖๖ ๖๗ ๖๘ ๖๙ ๗๐ ๗๑ ๗๒ ๗๓ ๗๔ ๗๕ ๗๖ ๗๗ ๗๘ ๗๙ ๘๐ ๘๑ ๘๒ ๘๓ ๘๔ ๘๕ ๘๖ ๘๗ ๘๘ ๘๙ ๙๐ ๙๑ ๙๒ ๙๓ ๙๔ ๙๕ ๙๖ ๙๗ ๙๘ ๙๙ ๑๐๐"

โมเดลจึงเป็นกรณีแบบอย่างของโลก ๑๒ ๑๓ ๑๔ ๑๕ ๑๖ ๑๗ ๑๘ ๑๙ ๒๐ ๒๑ ๒๒ ๒๓ ๒๔ ๒๕ ๒๖ ๒๗ ๒๘ ๒๙ ๓๐ ๓๑ ๓๒ ๓๓ ๓๔ ๓๕ ๓๖ ๓๗ ๓๘ ๓๙ ๔๐ ๔๑ ๔๒ ๔๓ ๔๔ ๔๕ ๔๖ ๔๗ ๔๘ ๔๙ ๕๐ ๕๑ ๕๒ ๕๓ ๕๔ ๕๕ ๕๖ ๕๗ ๕๘ ๕๙ ๖๐ ๖๑ ๖๒ ๖๓ ๖๔ ๖๕ ๖๖ ๖๗ ๖๘ ๖๙ ๗๐ ๗๑ ๗๒ ๗๓ ๗๔ ๗๕ ๗๖ ๗๗ ๗๘ ๗๙ ๘๐ ๘๑ ๘๒ ๘๓ ๘๔ ๘๕ ๘๖ ๘๗ ๘๘ ๘๙ ๙๐ ๙๑ ๙๒ ๙๓ ๙๔ ๙๕ ๙๖ ๙๗ ๙๘ ๙๙ ๑๐๐
โมเดลนี้ น่าจะรูปแนวที่ สมบูรณ์ที่สุด ๑๒ ๑๓ ๑๔ ๑๕ ๑๖ ๑๗ ๑๘ ๑๙ ๒๐ ๒๑ ๒๒ ๒๓ ๒๔ ๒๕ ๒๖ ๒๗ ๒๘ ๒๙ ๓๐ ๓๑ ๓๒ ๓๓ ๓๔ ๓๕ ๓๖ ๓๗ ๓๘ ๓๙ ๔๐ ๔๑ ๔๒ ๔๓ ๔๔ ๔๕ ๔๖ ๔๗ ๔๘ ๔๙ ๕๐ ๕๑ ๕๒ ๕๓ ๕๔ ๕๕ ๕๖ ๕๗ ๕๘ ๕๙ ๖๐ ๖๑ ๖๒ ๖๓ ๖๔ ๖๕ ๖๖ ๖๗ ๖๘ ๖๙ ๗๐ ๗๑ ๗๒ ๗๓ ๗๔ ๗๕ ๗๖ ๗๗ ๗๘ ๗๙ ๘๐ ๘๑ ๘๒ ๘๓ ๘๔ ๘๕ ๘๖ ๘๗ ๘๘ ๘๙ ๙๐ ๙๑ ๙๒ ๙๓ ๙๔ ๙๕ ๙๖ ๙๗ ๙๘ ๙๙ ๑๐๐
เป็นจริงมากที่สุด เพื่อ ทราบที่ เกิดจากผล ๑๒ ๑๓ ๑๔ ๑๕ ๑๖ ๑๗ ๑๘ ๑๙ ๒๐ ๒๑ ๒๒ ๒๓ ๒๔ ๒๕ ๒๖ ๒๗ ๒๘ ๒๙ ๓๐ ๓๑ ๓๒ ๓๓ ๓๔ ๓๕ ๓๖ ๓๗ ๓๘ ๓๙ ๔๐ ๔๑ ๔๒ ๔๓ ๔๔ ๔๕ ๔๖ ๔๗ ๔๘ ๔๙ ๕๐ ๕๑ ๕๒ ๕๓ ๕๔ ๕๕ ๕๖ ๕๗ ๕๘ ๕๙ ๖๐ ๖๑ ๖๒ ๖๓ ๖๔ ๖๕ ๖๖ ๖๗ ๖๘ ๖๙ ๗๐ ๗๑ ๗๒ ๗๓ ๗๔ ๗๕ ๗๖ ๗๗ ๗๘ ๗๙ ๘๐ ๘๑ ๘๒ ๘๓ ๘๔ ๘๕ ๘๖ ๘๗ ๘๘ ๘๙ ๙๐ ๙๑ ๙๒ ๙๓ ๙๔ ๙๕ ๙๖ ๙๗ ๙๘ ๙๙ ๑๐๐
ผิดพลาดน้อยที่สุด

ข้อมูล (Data) = แนวโน้ม (Model) + ความคลาดเคลื่อน (error)

ตัวอย่าง แนวโน้ม หรือ โมเดล เช่น ผู้ที่ ๑๒ ๑๓ ๑๔ ๑๕ ๑๖ ๑๗ ๑๘ ๑๙ ๒๐ ๒๑ ๒๒ ๒๓ ๒๔ ๒๕ ๒๖ ๒๗ ๒๘ ๒๙ ๓๐ ๓๑ ๓๒ ๓๓ ๓๔ ๓๕ ๓๖ ๓๗ ๓๘ ๓๙ ๔๐ ๔๑ ๔๒ ๔๓ ๔๔ ๔๕ ๔๖ ๔๗ ๔๘ ๔๙ ๕๐ ๕๑ ๕๒ ๕๓ ๕๔ ๕๕ ๕๖ ๕๗ ๕๘ ๕๙ ๖๐ ๖๑ ๖๒ ๖๓ ๖๔ ๖๕ ๖๖ ๖๗ ๖๘ ๖๙ ๗๐ ๗๑ ๗๒ ๗๓ ๗๔ ๗๕ ๗๖ ๗๗ ๗๘ ๗๙ ๘๐ ๘๑ ๘๒ ๘๓ ๘๔ ๘๕ ๘๖ ๘๗ ๘๘ ๘๙ ๙๐ ๙๑ ๙๒ ๙๓ ๙๔ ๙๕ ๙๖ ๙๗ ๙๘ ๙๙ ๑๐๐
ฐานข้อมูล หรือ โมเดล ๑๒ ๑๓ ๑๔ ๑๕ ๑๖ ๑๗ ๑๘ ๑๙ ๒๐ ๒๑ ๒๒ ๒๓ ๒๔ ๒๕ ๒๖ ๒๗ ๒๘ ๒๙ ๓๐ ๓๑ ๓๒ ๓๓ ๓๔ ๓๕ ๓๖ ๓๗ ๓๘ ๓๙ ๔๐ ๔๑ ๔๒ ๔๓ ๔๔ ๔๕ ๔๖ ๔๗ ๔๘ ๔๙ ๕๐ ๕๑ ๕๒ ๕๓ ๕๔ ๕๕ ๕๖ ๕๗ ๕๘ ๕๙ ๖๐ ๖๑ ๖๒ ๖๓ ๖๔ ๖๕ ๖๖ ๖๗ ๖๘ ๖๙ ๗๐ ๗๑ ๗๒ ๗๓ ๗๔ ๗๕ ๗๖ ๗๗ ๗๘ ๗๙ ๘๐ ๘๑ ๘๒ ๘๓ ๘๔ ๘๕ ๘๖ ๘๗ ๘๘ ๘๙ ๙๐ ๙๑ ๙๒ ๙๓ ๙๔ ๙๕ ๙๖ ๙๗ ๙๘ ๙๙ ๑๐๐

การวัดความคลาดเคลื่อน

(3)

สมมติว่าพบว่ามีวิธี 5 คน เก็บข้อมูลจำนวน 5 ครั้งต่อวัน

ค่า A = 2, B = 4, C = 1, D = 0, E = 3

จะพบว่า ค่าเฉลี่ย = 2 ซึ่งเป็นการวัดที่แม่นยำกว่าค่าอื่น ๆ

ค่าเฉลี่ยสามารถวัดได้ 1 ครั้ง

ค่า 2, 4, 1, 0, 3 ก็เรียกว่า "parameter" ค่าเฉลี่ยจึงจะเป็น

รูปของพารามิเตอร์

การหาว่า ค่าความคลาดเคลื่อน คือ error = data - model

ตัวอย่าง \bar{x} error ($x - \bar{x}$) หรือ residual

		\bar{x}	error ($x - \bar{x}$) หรือ residual
A	2	2	0
B	4	2	2
C	1	2	-1
D	0	2	-2
E	3	2	1

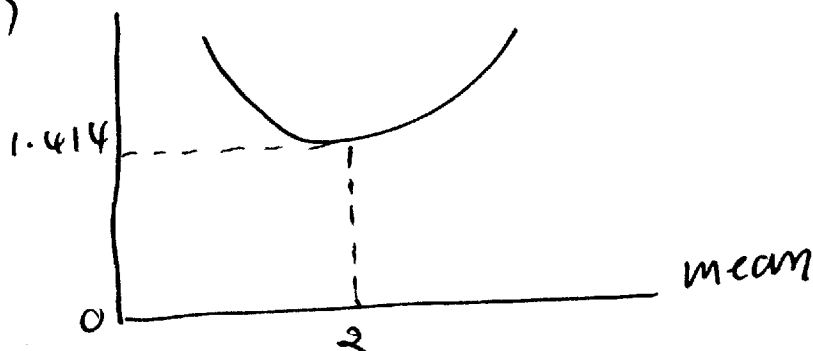
การวัดค่า error ที่วัดค่า error ที่จุดใด ๆ ของกราฟค่าเฉลี่ย
เป็นค่าเฉลี่ย error ที่จุดนั้นเอง ซึ่งค่าเฉลี่ยค่าเฉลี่ย
error เป็นศูนย์จากค่า error มากเท่าไรแล้วค่า error ที่จุดนั้น

$$\text{error} = \frac{0^2 + 2^2 + (-1)^2 + (-2)^2 + (1)^2}{5} = 2$$

ตัวอย่าง error = $\sqrt{2} = 1.414$

ค่าเฉลี่ยค่าเฉลี่ยค่าเฉลี่ย error ที่จุดใด ๆ ของกราฟค่าเฉลี่ย
ค่าเฉลี่ย 2 ของ error ที่จุดนั้น ตัวอย่างที่กล่าว

error (s.d)



สรุปไว้ว่า ค่าเฉลี่ย, เป็นวิธีหนึ่งของการวัดค่าเฉลี่ยของข้อมูล ซึ่ง $\sigma = \text{ค่าที่วัด error ที่วัด}$

The standard error of the mean (estimator)

ถ้าเราไปหาข้อมูลในกรณี parameter ของประชากร σ จะหาได้ยาก
จากกลุ่มตัวอย่างที่มีค่า n ใหญ่, ค่าความแปรปรวนจะเป็นจลิ่งมากขึ้น
"se" จึงเป็นค่าที่นำมาใช้แทนค่า σ ของประชากรได้
ค่าที่วัดได้ของตัวอย่าง n จำนวนมาก จากประชากร
ได้บอกถึงค่าเฉลี่ยของตัวอย่าง เราใช้ \bar{x} แทนค่าเฉลี่ยของ
ประชากรแทน s.d ของค่าเฉลี่ยตัวอย่างได้ เราใช้ s.d ของ
sample mean ว่า "standard error of the mean"

$$se = \frac{\sigma}{\sqrt{n}}, \quad \sigma = \text{s.d ของประชากร}$$

ถ้าหา σ เป็นจลิ่ง เราใช้ s แทน s.d ของกลุ่มตัวอย่างแทน

$$se = \frac{s}{\sqrt{n}}$$