

Statistical inference

ศ.ดร.นพ.พงศ์เทพ วิวรรณนะเดช

LL.B., M.D., Ph.D.

ภาควิชาเวชศาสตร์ชุมชน คณะแพทยศาสตร์

มหาวิทยาลัยเชียงใหม่ เชียงใหม่ 50200

E-mail: pwiwatanadate@gmail.com

ชนิดของการวัด

(Types of Measurement)

☀ ข้อมูลเชิงคุณภาพ (Qualitative Data)

☀ ข้อมูลเชิงปริมาณ (Quantitative Data)

ข้อมูลเชิงคุณภาพ (Qualitative Data)

Categorical or Nominal Measurement

- ★ เช่น กลุ่มเลือด จังหวัดเกิด
- ★ การเปรียบเทียบกัน เช่น ดีกว่า ใหญ่กว่า ไม่มีความหมาย
- ★ ส่วนใหญ่จะพรรณนาในรูปความถี่หรือร้อยละ
- ★ การวัดที่ผลมีได้เพียง 2 อย่าง (เช่น ตาย-ไม่ตาย ชอบ-ไม่ชอบ เพศชาย-หญิง) เรียกว่า Binomial Measurement

ข้อมูลเชิงคุณภาพ (Qualitative Data) (ต่อ)

Ordinal Measurement

- * เช่น สภาพของผู้ป่วย (ป่วยมาก ปานกลาง ไม่ป่วย)
- * เรียงลำดับของแต่ละบุคคลได้ แต่ไม่สามารถเปรียบเทียบเชิงปริมาณได้

ข้อมูลเชิงปริมาณ (Quantitative Data)

- † ข้อมูลที่ได้ต้องเป็นตัวเลขเสมอและสามารถให้ความละเอียดได้ไม่มีที่สิ้นสุด
- † สามารถหาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานได้
- † การเปรียบเทียบเชิงปริมาณสามารถทำได้เสมอและมีความหมาย

ชนิดของตัวแปร (Types of Variables)

* ตัวแปรตาม (Dependent Variable-Effect)

* ตัวแปรอิสระ (Independent Variable-Cause)

ชนิดของสถิติ (Types of Statistics)

◆ สถิติเชิงพรรณนา (Descriptive Statistics)

- ✓ mean, median, mode, standard deviation

◆ สถิติเชิงอนุมาน (Analytical or Inferential Statistics)

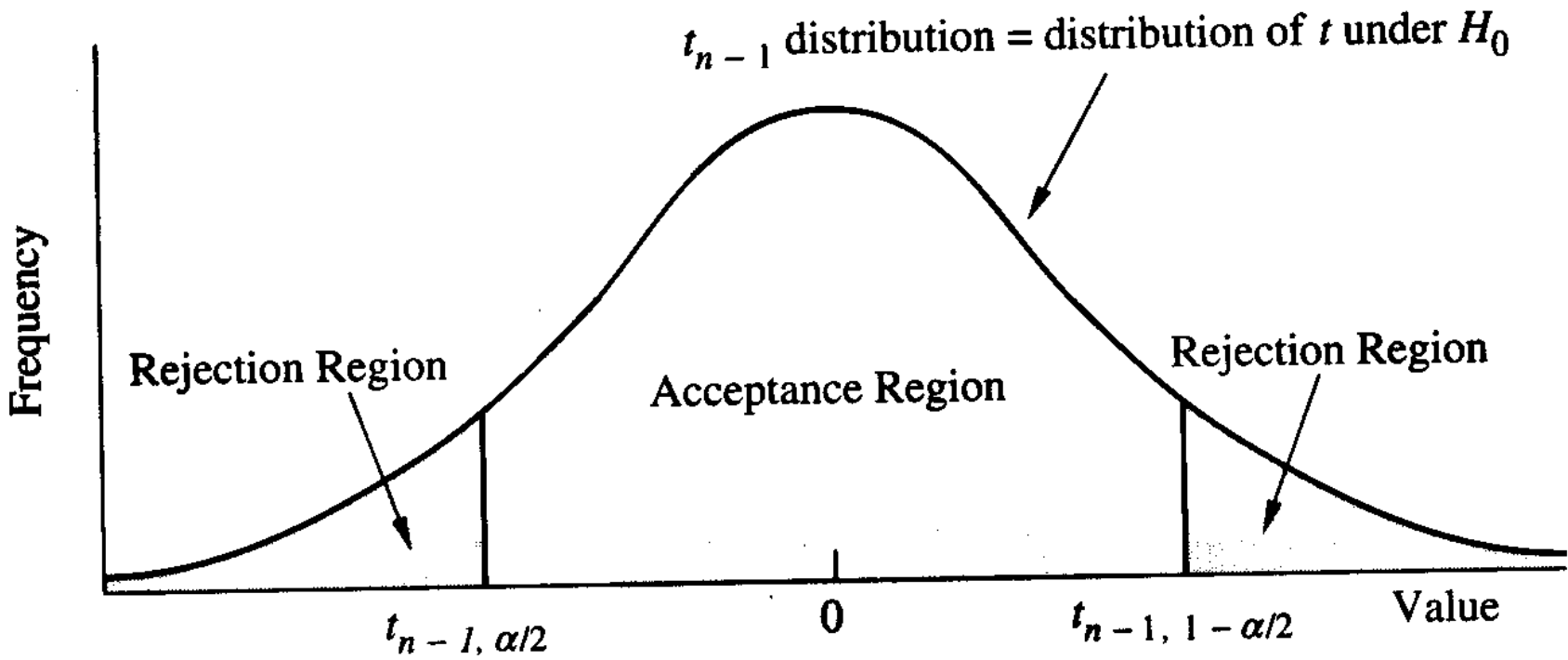
- ✓ Parametric

- ✓ Nonparametric

สถิติกลุ่มพาราเมตริก

- ข้อมูลที่ทำการวิเคราะห์จะต้องเป็นข้อมูลเชิงปริมาณและทำหน้าที่เป็นตัวแปรตาม (ยกเว้น chi-square test)
- ข้อมูลควรมีการกระจายแบบปกติ (Normal Distribution)
- ข้อดีของสถิติกลุ่มนี้คือวิเคราะห์ง่ายและให้อำนาจการบอกนัยสำคัญสูง
- แม้ข้อมูลจะมีการกระจายต่างไปจากแบบปกติก็ยังสามารถใช้ได้หากมีจำนวนตัวอย่างมากพอ (> 30 ตัวอย่าง)





สถิติกลุ่มพาราเมตริก (ต่อ)

ตัวอย่าง

- ★ Z-Test
- ★ Student's t -Test
- ★ Analysis of Variance (One-Way, Multi-Way)
- ★ Chi-Square Test
- ★ Regression

สถิติเชิงอนุมานกลุ่มนอนพาราเมตริก

- ※ ใช้ได้กับการกระจายทุกชนิด
- ※ การวิเคราะห์ทำไม่ได้หากมีตัวแปรในการศึกษาจำนวนมาก
- ※ เมื่อขนาดตัวอย่างมีมาก การคำนวณจะยุ่งยากและใช้เวลามาก
- ※ อำนวยในการบอกนัยสำคัญน้อยกว่าแบบพาราเมตริก
- ※ ตัวอย่าง
 - ✓ Sign Test
 - ✓ Wilcoxon Signed Rank Sum Test
 - ✓ Mann-Whitney U Test
 - ✓ Kruskal-Wallis Test
 - ✓ Friedman's Test

สถิติเชิงอนุมานกลุ่มพาราเมตริก

Mean Inference

★ *One-Sample Tests*

- ✓ Z-Test (ทราบ population variance)
- ✓ Student's *t*-Test (ไม่ทราบ population variance)

สถิติเชิงอนุมานกลุ่มพารามेटริก (ต่อ)

Mean Inference (ต่อ)

★ *Two-Sample Tests*

- ✓ Paired t -Test
- ✓ t -Test for Independent Samples with Equal Variances
- ✓ t -Test for Independent Samples with Unequal Variances

สถิติเชิงอนุมานกลุ่มพารามетริก (ต่อ)

Variance Inference

- ✦ One-Sample χ^2 Test
- ✦ Two-Sample F Test for Equality of Two Variances
- ✦ Multisample:
 - ✓ One-Way Analysis of Variance [*ค่า variances จะต้องเท่ากัน (Levene's Test)*]
 - ✓ Multi-Way Analysis of Variance

สถิติเชิงอนุมานกลุ่มพาราเมตริก (ต่อ)

Categorical Data Inference

◆ Chi-Square Test

- ✓ ใช้วิเคราะห์ทีละ 2 ตัวแปร และต้องวัดแบบเชิงคุณภาพทั้งคู่
- ✓ ให้ค่า p -value ที่เป็นค่าประมาณเสมอ
- ✓ ใช้ไม่ได้ถ้า 20% ของเซลล์มี expected values < 5 หรือ expected values < 1 แม้เพียงเซลล์เดียว

◆ Fisher's Exact Test

- ✓ ให้ค่าที่ถูกต้อง
- ✓ ใช้เวลาในการคำนวณมาก

การทดสอบนัยสำคัญทางสถิติ

- ✿ ในศาลจะถือว่า “จำเลยบริสุทธิ์” จนกว่า “จะพิสูจน์โดยปราศจากข้อสงสัยว่ามีความผิดจริง”
- ✿ ในทางสถิติสัดส่วนที่ยอมให้สรุปผิดได้มักอยู่ที่ 5% หรือ 1%

การทดสอบนัยสำคัญทางสถิติ

- * ความผิดพลาดที่เกิดจากการสรุปว่า “มีนัยสำคัญทางสถิติ” โดยที่ในความเป็นจริงเกิดโดยบังเอิญ เปรียบเสมือนการตัดสินว่าผู้บริสุทธิ์เป็นผู้มีความผิด ความผิดพลาดชนิดนี้เรียกว่า *type I error*
- * สำหรับ *type II error* คือการไม่สามารถบอกนัยสำคัญทางสถิติได้ ทั้งๆ ที่ “มีนัยสำคัญทางสถิติ” เปรียบเสมือนการยกฟ้องผู้กระทำความผิดในทางศาล

การทดสอบสมมติฐาน

สมมติฐาน คือข้อความเกี่ยวกับ

- ❁ การกระจาย (เช่น “การกระจายเป็นแบบปกติ”), หรือ
- ❁ แสดงค่าของพารามิเตอร์ (เช่น “ $\mu = 10$ ”), หรือ
- ❁ ความสัมพันธ์ระหว่างพารามิเตอร์ (เช่น “ $\mu_1 = \mu_2$ ” —หมายความว่าค่าเฉลี่ยของประชากรมีค่าเท่ากัน)

การทดสอบสมมติฐาน

- ❁ การทดสอบสมมติฐานคือ กระบวนการตัดสินใจบนพื้นฐานของข้อมูล มากกว่าการใช้ความเห็นส่วนบุคคลเพียงการพิจารณาจากตัวเลขเท่านั้น
- ❁ โดยความเป็นจริงต่างบุคคลย่อมให้ความเห็นที่แตกต่างกันหากเพียงแค่มองตัวเลข
- ❁ ในทางตรงข้ามการทดสอบสมมติฐานเป็นกระบวนการตัดสินใจที่เป็นมาตรฐาน ที่ให้ผลเสมอต้นเสมอปลายไม่ว่าใครจะเป็นผู้ทดสอบ

การทดสอบสมมติฐาน

- ❁ สมมติฐานที่จะทดสอบเรียกว่า *null hypothesis* และแสดงโดย H_0 , ที่จะบ่งบอกว่า ไม่มีความแตกต่าง หรือไม่มีความสัมพันธ์ระหว่างพารามิเตอร์ เปรียบเสมือนการให้หลักประกันว่า ผู้ถูกกล่าวหาเป็นผู้บริสุทธิ์จนกว่าจะพิสูจน์ว่ามีความผิดจริง
- ❁ การทดสอบสมมติฐานจึงเป็นกระบวนการตัดสินใจ โดยพิจารณาจากชุดของข้อมูลและภายใต้ข้อสมมติฐานของ H_0 ก็จะนำไปสู่การตัดสินใจว่าจะปฏิเสธ H_0 หรือไม่

การทดสอบสมมติฐาน

- ❁ ในทางตรงข้าม alternative hypothesis แสดงด้วย H_A คือ สมมติฐานที่ตรงข้ามกับ null hypothesis H_0 เปรียบเสมือน ข้อกล่าวหาแก่จำเลย ภายใต้สมมติฐาน H_A จะถือว่ามีความแตกต่างกันจริง เช่น $\bar{x}_1 \neq \bar{x}_2$ ไม่ใช่เกิดโดยบังเอิญ เพราะ $\mu_1 \neq \mu_2$
- ❁ เราจะปฏิเสธ null hypothesis ถ้ามีหลักฐานเพียงพอที่แสดงให้เห็นว่า alternative hypothesis ถูกต้อง

หลักฐานทางสถิติ

- ❁ ถ้า null hypothesis เกี่ยวข้องกับค่า μ (เช่น $H_0: \mu = 10$)
ที่ที่ดีที่สุดที่จะหาข้อมูลเกี่ยวกับ μ คือการคำนวณค่า \bar{x}
- ❁ \bar{x} เรียกว่า *test statistic* ซึ่งเป็นสถิติที่ใช้วัดความแตกต่างระหว่างข้อมูลและสิ่งที่คาดหมาย ถ้า null hypothesis เป็นจริง (เช่น “ $\mu = 10$ ”)
- ❁ ถ้าความแตกต่างระหว่างข้อมูลและสิ่งที่คาดหมายจาก null hypothesis ยากที่จะอธิบายด้วยความบังเอิญ ดังนั้นจะต้องปฏิเสธ null hypothesis และยอมรับ alternative hypothesis แทน

ความผิดพลาด

ข้อเท็จจริง	การตัดสินใจ	
	ยอมรับ H_0	ปฏิเสธ H_0
H_0 เป็นความจริง	ตัดสินใจถูกต้อง	Type I error
H_0 ไม่เป็นความจริง	Type II error	ตัดสินใจถูกต้อง

ความผิดพลาด Type I และ Type II

1. *Type I*: ความผิดพลาดจากการปฏิเสธ H_0 (ที่ถูกต้อง)
 2. *Type II*: ความผิดพลาดจากการไม่ปฏิเสธ H_0 (ที่ไม่ถูกต้อง)
- ❁ โอกาสของความผิดพลาด type I แสดงด้วย α และมักเรียกว่า “ระดับนัยสำคัญ” ของการทดสอบ
 - ❁ โอกาสของความผิดพลาด type II แสดงด้วย β และ $(1 - \beta)$ เรียกว่า “อำนาจ” ของการทดสอบ
 - ❁ โดยทั่วไปเป้าหมายของการทดสอบสมมติฐาน คือการเลือกสถิติที่ทำให้ α and β เล็กที่สุดเท่าที่จะเป็นไปได้
 - ❁ ในการวิเคราะห์ทางสถิติ เรามักจะกำหนดค่า α ที่คงที่ระดับหนึ่ง เช่น 0.01 หรือ 0.05 และใช้สถิติที่ทำให้ค่า β น้อยที่สุด หรือทำให้อำนาจการบอกนัยสำคัญสูงสุดโดยที่ใช้ขนาดตัวอย่างน้อยที่สุด

ความผิดพลาด Type I และ Type II

1. *Type I*: ความผิดพลาดจากการปฏิเสธ H_0 (ที่ถูกต้อง)
 2. *Type II*: ความผิดพลาดจากการไม่ปฏิเสธ H_0 (ที่ไม่ถูกต้อง)
- ❁ สัดส่วนของความผิดพลาด type II ขึ้นกับหลายปัจจัยคือ
- ✓ ขนาดของความแตกต่าง เช่นเดียวกับ ความรุนแรงของการกระทำ ความผิดในทางศาล หากความแตกต่างยังมีมาก โอกาสของความผิดพลาด type II จะยิ่งน้อย
 - ✓ ขนาดของความแปรปรวนในข้อมูล (“noise”) เช่นเดียวกับ คุณภาพของหลักฐานที่ใช้ในการพิจารณาของศาล
 - ✓ ขนาดของตัวอย่าง เช่นเดียวกับปริมาณของหลักฐานที่ใช้ในการพิจารณาของศาล โอกาสของความผิดพลาด type II จะสูงในการศึกษาที่มีจำนวนตัวอย่างน้อย

การเปรียบเทียบความคล้ายคลึง

การทดสอบนัยสำคัญ	↔	การตัดสินใจในศาล
Null hypothesis	↔	“จำเลยบริสุทธิ์จนกว่าจะพิสูจน์ได้ว่าผิดจริง”
การออกแบบการศึกษา	↔	การสอบสวนของตำรวจ
ข้อมูล=การทดสอบทางสถิติ	↔	หลักฐาน=การแสดงต่อศาล
หลักการทางสถิติ	↔	วิธีพิจารณาคดี
การตัดสินใจในทางสถิติ	↔	การตัดสินใจของศาล
Type I error	↔	ตัดสินผู้บริสุทธิ์ว่ามีความผิด
Type II error	↔	ยกฟ้องอาชญากร