ข้อสอบ 100 ข้อ

	หน้า
บท 1 : บทนำ (ออกสอบ 5 ข้อ)	
- คำศัพท์ในสถิติ	2
- หาร sum, factorial → ทำแจกสไลด์ ex. ข้อสอบ	
บท 2 : รูปแบบการนำเสนอข้อมูล (ออกสอบ 5 ข้อ)	25
- การนำเสนอข้อมูลด้วยกราฟ	
- " ______ " จากตัวแปรเชิงปริมาณ	
- " ______ " ตัวแปรเชิงคุณภาพ	
บท 3 : มาตรวัดผลสรุปข้อมูล (ออกสอบ 15 ข้อ)	46
- มาตรวัดตำแหน่ง : มาตรวัดแนวโน้มเข้าสู่ศูนย์กลาง, มาตรวัดอื่นๆ	
- มาตรวัดการกระจาย : การกระจายสัมบูรณ์, การกระจายสัมพัทธ์	
บท 4 : ความน่าจะเป็น (ออกสอบ 15 ข้อ)	110
บท 5 : ส่วนเบี่ยงเบน, สัดส่วน (ออกสอบ 10 ข้อ)	182

New file

บท 6 : การประมาณค่า (ออกสอบ 15 ข้อ)	1
บท 7 : ทดสอบสมมติฐาน (ออกสอบ 15 ข้อ)	87
บท 10 : การทดถอยเชิงเดี่ยว (ออกสอบ 10 ข้อ) สหสัมพันธ์	204
บท 11 : Table ANOVA (ออกสอบ 10 ข้อ)	252

Grade (100 ข้อ)

A => 75+ (ผิดได้ 25)
B+ => 70+ (ผิดได้ 30)
B => 65+ (ผิดได้ 35)
C+ => 60+ (ผิดได้ 40)
C => 55+ (ผิดได้ 45)
D+ => 50+ (ผิดได้ 50)
D => 45+ (ผิดได้ 55)

บท 1) บทนำ (5 ข้อ)

สถิติพรรณนา => รายงาน, บรรยายภาพ, ใช้ค่าสถิติอย่างง่าย เช่น %, ค่าเฉลี่ย, บอกผลเป็นตัวเลข, แผนภูมิ, ตาราง

สถิติอนุมาน => มีการสุ่มตย., ทดสอบ, ทดลอง, ประมาณค่า, มีการใช้เทคนิคสถิติที่ซับซ้อน

ตัวแปรเชิงคุณภาพ => ค่าของตัวแปร "ไม่ใช่ตัวเลข"

ตัวแปรเชิงปริมาณ => "เป็นตัวเลข" แบ่งได้ 2 ประเภท
- ไม่ต่อเนื่อง => เป็น int (ค่าถูกตัดตอน)
- ต่อเนื่อง => เป็นช่วง (ไม่ถูกตัดตอน)

ตัวสถิติ => func ของค่าสังเกตที่ได้จาก ตย. สุ่ม / มีตย. ต่อ

ค่าเฉลี่ยตย. $(\bar{x}) = \frac{1}{n}\sum_{i=1}^{n} x_i$

ค.แปรปรวนตย. $(s^2) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$

เครื่องหมายรวมยอด

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \dots + x_n$$

Parameter => ค่าคงที่แสดงลักษณะปชก. / ลงท้ายด้วย ปชก.

ตัวพิมพ์เล็ก	ตัวพิมพ์ใหญ่	ชื่อ	ความหมาย
μ	M	มิว (mu)	ค่าเฉลี่ยประชากร
σ	Σ	ซิกมา (sigma)	ส่วนเบี่ยงเบนมาตรฐานประชากร
π	Π	พาย (pi)	สัดส่วนประชากร
ε	E	เอปไซลอน (epsilon)	ความคลาดเคลื่อนหรือความผิดพลาดเชิงสุ่ม
θ	Θ	ทีตา (theta)	พารามิเตอร์ใดๆที่เราศึกษา

ตัวพิมพ์เล็ก	ตัวพิมพ์ใหญ่	ชื่อ	ความหมาย
β	B	บีตา (beta)	สัมประสิทธิ์การถดถอยประชากรหรือความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 2
ρ	P	โร (rho)	สัมประสิทธิ์สหสัมพันธ์ประชากร
α	A	แอลฟา (alpha)	ความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 หรือระดับนัยสำคัญหรือขนาดของการทดสอบ

แบ่งชนิดข้อมูลตามแหล่ง:

1. ปฐมภูมิ => ได้ข้อมูลจากผู้ให้โดยตรง
2. ทุติยภูมิ => ได้จากหน่วยงานที่รับผิดชอบเก็บข้อมูลนั้นๆ

แบ่งชนิดข้อมูลตามระยะเวลาจัดเก็บ:

1. อนุกรมเวลา => เก็บแบบต่อเนื่องเอาไปใช้ในอนาคตได้
2. เฉพาะเวลา => เก็บเฉพาะเวลาเอาไปใช้ในอนาคตไม่ได้

แบ่งชนิดข้อมูลตามค่าข้อมูล:

1. ข้อมูลเชิงคุณภาพ => "ไม่ใช่ตัวเลข" เช่น ศาสนา เพศ
2. ข้อมูลเชิงปริมาณ => "เป็นตัวเลข" เช่น อายุ รายได้

บท 2) รูปแบบนำเสนอข้อมูล (5 ข้อ)

1. การนำเสนอข้อมูลโดยใช้รูปภาพ => แผนภูมิเส้น, แท่ง, O, รูป
2. การนำเสนอข้อมูลที่ได้จากตัวแปรเชิงปริมาณ => มี 8 แบบ
   - Arrays => เรียง น→ม, ม→น, ใช้ได้กับข้อมูล < 20
   - Dot Plots => ไม่เกิน 20, นำเสนอโดยplotข้อมูลแต่ละค่าลงบนเส้น
   - Stem-and-leaf => > 20, เอาไว้ใช้ดูข้อมูลสมมาตรไหม
   - Cumulative Distributions => เอา Arrays มาจัดอันดับ
   - การแจกแจงค.ถี่ => ข้อมูลจน.มาก สร้างตารางแจกแจง
   - Histogram => นำเสนอโดยใช้รูปกราฟแท่งแสดงค.ถี่
   - Frequency Polygon => นำเสนอโดยใช้รูปกราฟเส้นแสดงค.ถี่
   - Ogive => นำเสนอการแจกแจงค.ถี่สะสมมาทำกราฟเส้น
3. การนำเสนอข้อมูลที่ได้จากตัวแปรเชิงคุณภาพ => มี 2 แบบ
   - การแจกแจงคุณภาพ
   - การนำเสนอข้อมูลโดยใช้รูปภาพ

แนวข้อสอบบท 1

100%

ตัวอย่างของสถิติพรรณนา ได้แก่ [illegible]
- ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบวิชา ST 203 เท่ากับ 48 และ 24 คะแนน ตามลำดับ
- ตารางแสดงจำนวนนักศึกษาที่สมัครใหม่ของมหาวิทยาลัยรามคำแหงจำแนกตามคณะ ประจำปี 2545

ตัวอย่างของสถิติอนุมาน ได้แก่
- [illegible] 20,000 กิโลเมตร
- การประมาณจำนวนคนไข้ที่มาขอรับการบริการรักษาจากโรงพยาบาลแห่งหนึ่งต่อวันมีจำนวนอยู่ระหว่าง 350 – 450 คน

การศึกษาข้อมูลของปชก. เป็นบางส่วนเพื่อนำผลสรุปไปอ้างอิงปชก. ที่สนใจ (ใช้สุ่มตย.)

เป็นการรวบรวม ข้อเท็จจริงในรูปของภาพแผนภาพ [illegible] (มีข้อมูลอยู่แล้ว ไม่ต้องไปสุ่ม มาศึกษา)

5. จงพิจารณาว่าข้อความต่อไปนี้จัดเป็นสถิติพรรณนาหรือสถิติอนุมาน
   ก. การประมาณจำนวนคนว่างงานใน 2 ปีข้างหน้า สถิติอนุมาน
   ข. ยอดจำหน่ายรถยนต์ในไตรมาสที่ 3 ของปี 2544 สถิติพรรณนา
   ค. การสอบถามความนิยมของผู้บริโภคที่มีต่อน้ำยาล้างจานสูตรใหม่ สถิติอนุมาน
   ง. จำนวนอุบัติเหตุจราจรทางบกในปี 2541 สถิติพรรณนา
   จ. กราฟแท่งแสดงจำนวนนักศึกษาสมัครใหม่ระหว่างปี 2536 – 2543 สถิติพรรณนา
   ฉ. รามคำแหงโพล สถิติอนุมาน
   ช. จากการทดสอบหลอดไฟที่ผลิตได้พบว่ามีอายุการใช้งานเฉลี่ยมากกว่า 2000 ชั่วโมง อนุมาน
   ซ. ปริมาณการใช้ไฟฟ้าของประเทศในช่วงฤดูร้อนเพิ่มสูงขึ้น 10% จากปีก่อน พรรณนา
   ฌ. การพยากรณ์ [illegible] ใช้ในประเทศอีก 5 ปีข้างหน้า อนุมาน
6. จงใช้ตัวเลือกต่อไปนี้เพื่อตอบคำถาม
   1. ตัวแปรเชิงคุณภาพ 2. ตัวแปรเชิงปริมาณ 3. ตัวอย่างสุ่ม
   4. พารามิเตอร์ 5. ประชากร 6. ตัวสถิติ
   ก. มัชฌิมเลขคณิตประชากรจัดเป็น 4. พารามิเตอร์
   ข. เขต 6 เขตในกรุงเทพฯที่ถูกสุ่มเลือกขึ้นมาจัดเป็น 3. ตัวอย่างสุ่ม
   ค. คะแนนสอบวิชา ST 203 จัดเป็น 2. ตัวแปรเชิงปริมาณ
   ง. จำนวนนักศึกษาทั้งหมดของมหาวิทยาลัยรามคำแหงจัดเป็น 5. ประชากร
   จ. การนับถือศาสนาจัดเป็น 1. ตัวแปรเชิงคุณภาพ
   ฉ. จำนวนรถยนต์ที่วิ่งผ่าน [illegible] 2. ตัวแปรเชิงปริมาณ
   ช. การวัดส่วนสูงของทารกแรกเกิดของโรงพยาบาลแห่งหนึ่งจัดเป็น 2. ตัวแปรเชิงปริมาณ
   ซ. ความแปรปรวนของตัวอย่างจัดเป็น 6. ตัวสถิติ
   ฌ. เพศของนักศึกษาจัดเป็น 1. ตัวแปรเชิงคุณภาพ
   ญ. จำนวนนักศึกษาที่สอบผ่านวิชา ST 203 ในภาคเรียนหนึ่งๆจัดเป็น 2. ตัวแปรเชิงปริมาณ

เก็บแบบต่อเนื่องเอาไปใช้ในอนาคตได้

เก็บเฉพาะเวลาเอาไปใช้ในอนาคตไม่ได้

7. จงพิจารณาว่าข้อมูลต่อไปนี้จัดเป็นข้อมูลอนุกรมเวลาหรือข้อมูลเฉพาะระยะเวลาใดเวลาหนึ่ง
   ก. ยอดจำหน่ายรถยนต์ในประเทศไทยระหว่างปี 2540 - 2544 อนุกรม
   ข. การสำรวจค่าใช้จ่ายในช่วงปีใหม่ของประชาชนในกทม. เฉพาะเวลา
   ค. จำนวนผู้ป่วยที่ติดเชื้อเอดส์ปี 2530 – ปัจจุบัน อนุกรม
   ง. การวัดการเต้นของหัวใจจากเครื่องวัดทุกๆ ชั่วโมง อนุกรม
   จ. การวัดความ [illegible] เฉพาะเวลา
8. $x_1 = 1$, $x_2 = 3$, $x_3 = 4$, $x_4 = 6$; $y_1 = 2$, $y_2 = -5$, $y_3 = 7$, $y_4 = 8$; $f_1 = 3$, $f_2 = 8$, $f_3 = 10$, $f_4 = -1$
   จงหา $\sum_{i=1}^{4} x_i$, $\sum_{i=1}^{4} x_i^2$, $\sum_{i=1}^{4} f_i x_i$, $\sum_{i=1}^{4} f_i y_i^2$, $\sum_{i=1}^{4} f_i(x_i+y_i)$ และ $\sum_{i=1}^{4}(x_i-2)$
   1. $1+3+4+6 = 14$ #
   2. $1+9+16+36 = 62$ #
   3. $(3)(1)+(8)(3)+(10)(4)+(-1)(6) = 3+24+40-6 = 61$ #
   4. [illegible] $= 658$ #
   5. [illegible] $= 89$ #
   6. [illegible] $= 6$ #
9. จงหาค่าของ $3!+6!$, $6!-3!$, $3!6!$ และ $\frac{6!}{3!}$
   $3!+6! = (3 \cdot 2)+(6 \cdot 5 \cdot 4 \cdot 3 \cdot 2) = 6 + 720 = 726$ #
   $6!-3! = 720-6 = 714$ #
   $3!6! = 4320$ #
   $\frac{6!}{3!} = \frac{6 \cdot 5 \cdot 4 \cdot \cancel{3 \cdot 2}}{\cancel{3 \cdot 2}} = 6 \cdot 5 \cdot 4 = 120$ #

100%

1. ถ้าข้อมูลชุดหนึ่งมีขีดจำกัดของชั้นดังนี้ 31 – 35 , 36 – 40 , 41 – 45 และ 46 –50 แล้ว
   (1) จงหาความกว้างของชั้นข้อมูล
   (2) จงหาขอบเขตของชั้นข้อมูล
   (3) จงหาจุดกึ่งกลางชั้น

ช่วง 31-35
1. ค.กว้างของชั้นข้อมูล
   $35-31+1 = 5$
   $35.5-30.5 = 5$ #
2. ขอบเขตของชั้นข้อมูล
   ขอบเขตล่าง $= 31-0.5 = 30.5$ #
   ขอบเขตบน $= 35+0.5 = 35.5$ #
3. จุดกึ่งกลางของชั้นข้อมูล
   $\frac{(35+31)}{2} = \frac{66}{2} = 33$ #

ช่วง 36-40
1. ค.กว้างของชั้นข้อมูล
   $40-36+1 = 5$
   $40.5-35.5 = 5$ #
2. ขอบเขตของชั้นข้อมูล
   ขอบเขตล่าง $= 36-0.5 = 35.5$ #
   ขอบเขตบน $= 40+0.5 = 40.5$ #
3. จุดกึ่งกลางของชั้นข้อมูล
   $\frac{(40+36)}{2} = \frac{76}{2} = 38$ #

ช่วง 41-45
1. ค.กว้างของชั้นข้อมูล
   $45-41+1 = 5$
   $45.5-40.5 = 5$ #
2. ขอบเขตของชั้นข้อมูล
   ขอบเขตล่าง $= 41-0.5 = 40.5$ #
   ขอบเขตบน $= 45+0.5 = 45.5$ #
3. จุดกึ่งกลางของชั้นข้อมูล
   $\frac{(45+41)}{2} = \frac{86}{2} = 43$ #

บท 3) มาตรวัดสรุปผล (15 ข้อ)

สูตรที่ใช้ตอนไม่มีตารางแจกแจง ค.ถี่

ค่าเฉลี่ย(μ)($\bar{X}$) = sum/n

มัธยฐาน(Md) = $(n+1)/2$

ฐานนิยม(Mo) = เลขซ้ำที่เยอะสุด

ควอร์ไทล์ที่ r (Q_r) = $r(N+1)/4$

เดไซล์ที่ r (D_r) = $r(N+1)/10$

เปอร์เซ็นไทล์ (P_r) = $r(N+1)/100$

พิสัย (R) = $X_{max} - X_{min}$

$IQR. = Q_3 - Q_1$

$Q.D. = IQR/2$

ส่วนเบี่ยงเบนเฉลี่ย(A.D) = $\frac{1}{N} sum|X_i - \mu|$

ค.แปรปรวน(σ^2) = $\frac{1}{N} sum(X_i - \mu)^2$

ค.แปรปรวนตพ.(S^2) = $\frac{1}{n-1} sum(X_i - \bar{X})^2$

ส่วนเบี่ยงเบนมาตรฐาน (σ) = $\sqrt{\sigma^2}$; $(S) = \sqrt{S^2} = (\bar{X})^2$

สัมประสิทธิ์การแปรผัน (C.V) = σ/μ

$Z = x - \mu/\sigma$; $Z = x - \bar{X}/S$

- สีเหลืองเป็นของสถิติ สูตรเหมือนกัน แค่เปลี่ยนตัว

สูตรที่ใช้ตอนมีตารางแจกแจง ค.ถี่

ค.ถี่สัมพัทธ์ = $f/\Sigma f$

ชั้นสมมุติ = ชั้นที่มี ค.ถี่มากที่สุด

A = Mid ชั้นสมมุติ

I = ค.กว้างของชั้น

L = ขอบเขตล่าง

$\mu(\bar{X}) = A + \left(\frac{\Sigma fd}{\Sigma f}\right)I$

$Md = L + \left[\frac{N/2 - \Sigma f}{f_m}\right]I$

Σf = ค.ถี่สะสมของชั้นก่อนหน้า

f_m = ค.ถี่ชั้นที่พิจารณา

$Mo = L + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right)I$

Δ_1 = ผลต่าง ค.ถี่ระหว่างชั้น Mo กับชั้นก่อนหน้า

Δ_2 = ผลต่าง ค.ถี่ระหว่างชั้น Mo กับชั้นถัดมา

$D_r, Q_r, P_r = L + \left(\frac{rN/10 - \Sigma f}{fd}\right)I$

Σf = ค.ถี่สะสมของชั้นก่อนหน้า

f_d = ค.ถี่ชั้นที่พิจารณา

$\sigma^2 = I^2(\Sigma fd^2/N - (\Sigma fd/N)^2)$

บท 4 ค.น่าจะเป็น (15 ข้อ)

วิธีการกำหนดค.น่าจะเป็นของเหตุการณ์ มี 3 วิธี

1. วิธีค.ถี่สัมพัทธ์ => ต้องเกิดการทดลองจริงๆ

$$P(A) = \frac{\text{จน.ครั้งที่เหตุการณ์เกิดขึ้น}}{\text{จน.ครั้งทั้งหมดของทดลองสุ่ม}}$$

2. วิธีแบบฉบับ => ไม่จำเป็นต้องทดลอง

$$P(A) = \frac{\text{จน.result ของการทดลอง}}{\text{จน.result ที่ possible ของการทดลอง}}$$

3. วิธีอัตนัย => วิธีกำหนด prob. ให้ situation โดยอาศัยผู้รู้หรือผู้มีประสบการณ์

การคำนวณ prob.

บวก prob. => $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

บวก prob. เหตุการณ์ไม่เกิดร่วม ($P(A \cap B) = 0$)

=> $P(A \cup B) = P(A) + P(B)$

prob. มีเงื่อนไข => $P(A/B) = P(A \cap B)/P(B)$

คูณ prob. => $P(A \cap B) = P(B) \cdot P(A/B)$

อิสระ => $P(A \cap B) = P(A) \cdot P(B)$

เหตุการณ์เติมเต็ม => $P(\bar{A}) = 1 - P(A)$

ตัวแปรสุ่มไม่ต่อเนื่อง

การแจกแจงปัวซง => มักใช้นับจน.situation ที่เกิดขึ้นในช่วงเวลาหรือพื้นที่หนึ่ง เช่น จน.อุบัติเหตุ มีลักษณะเป็นตัวแปรเต็มบวก $X = 0, 1, 2, \ldots$

- ข้อมูลเกี่ยวกับการนับจน.เหตุการณ์

การแจกแจงทวินาม => ข้อมูลเป็นจน.ครั้งของเหตุการณ์ที่เกิดขึ้นในจน.การทดลองที่กำหนด n ครั้ง และแต่ละการทดลองมี prob. เกิดขึ้น

ต่อเนื่อง

การแจกแจงปรกติ => เจอบ่อยในธรรมชาติและการคำนวณทางสถิติ เช่น ปริมาณไฟฟ้าต่อเดือนของแต่ละครัวเรือน

- ข้อมูลที่เป็นตัวเลขที่สามารถรับค่าใดก็ได้ ในช่วงเวลาหนึ่งๆ

หลักการนับ มี 3 วิธี

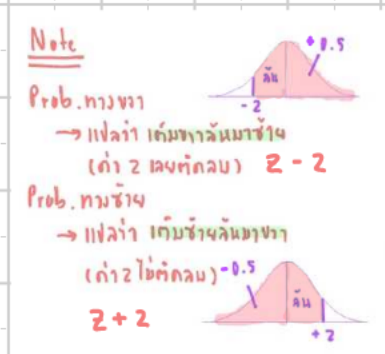
1. หลักการคูณ
2. การเรียงสับเปลี่ยน Permu (^nP_r)
   - สิ่งของ n สิ่งที่แตกต่างกัน คราวละ r สิ่ง = $\frac{n!}{(n-r)!}$
   - มีสิ่งของที่ซ้ำกัน = $\frac{n!}{r_1! \ldots r_k!}$
3. การจัดหมู่ Combi (^nC_r) = $\frac{n!}{r!(n-r)!}$

ค่าคาดหมาย

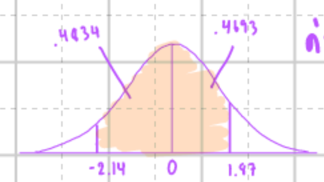
$E(X) = \mu = \sum P(X = x)$

$Var(X) = \sigma^2 = \underbrace{E(X^2)}_{P(X^2 = x)} - (E(X))^2$ → เอา result $E(X)$ ไปยกกำลัง 2

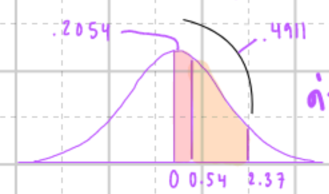
ตาราง Z



• ระหว่าง



ค่าจุด + กัน



ค่าจุด - กัน

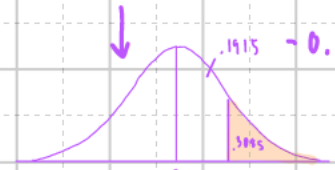
$$z = \frac{x - \mu}{\sigma}$$

Step : หา z ที่จะให้ prob. = ?

1. เอาเลขที่โจทย์ให้ -0.5 จะได้พท.
2. เอาพท.ไปเปิดตาราง z
3. ดูว่าค่า z ติด -, + มั้ย ดูเอาจากรูป

Step : หา x ~ N / P(X >, < ?)

1. แทนค่า $z = \frac{x - \mu}{\sigma}$
2. เอาค่า z ไปเปิดตาราง
3. ดูจากรูปเอาว่าจะ ±0.5
   - \+ รูปอยู่ระหว่าง 2 ฝั่ง
   - \- รูปอยู่ฝั่งเดียวกัน



บท 5 การแจกแจงการชักตย. (10 ข้อ)

แบบคืนที่ ; ปชก.จำกัด & ไม่จำกัด

$E(\bar{X}) = \mu_{\bar{x}} = \mu$

$S.D.(\bar{X}) = \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

$\bar{X}\ (z) = \frac{x-\mu}{\sigma/\sqrt{n}} \longrightarrow \sigma_{\bar{x}}$

สัดส่วนตย.

$E(p) = \mu_p = \pi$

$S.D.(p) = \sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$

$p\ (z) = \frac{p-\pi}{\sigma_p}$

$\bar{x}_1 - \bar{x}_2$

$E(\bar{x}_1 - \bar{x}_2) = \mu_1 - \mu_2$

$S.D(\bar{x}_1 - \bar{x}_2) = \sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

$\bar{x}_1 - \bar{x}_2\ (z) = \frac{(x_1 - x_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}$

แบบไม่คืนที่ ; ปชก.จำกัด

$E(\bar{X}) = \mu_{\bar{x}} = \mu$

$S.D.(\bar{X}) = \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$

$p_1 - p_2$

$E(p_1 - p_2) = \pi_1 - \pi_2$

$S.D(p_1 - p_2) = \sigma_{p_1 - p_2} = \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}$

$p_1 - p_2\ (z) = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sigma_{p_1 - p_2}}$

บท 6 การประมาณค่า (15 ข้อ)

วิธีหา α : ค่าเชื่อมั่น 95% (ดูที่ตาราง t)

$100(1-\alpha)\% = 95\%$	$\alpha = 0.05$
$(1-\alpha) = 0.95$	$\alpha/2 = 0.025$
$\alpha = 1-0.95$	$Z_{0.025} = 1.96$

การประมาณค่า 1 กลุ่มปชก.

1. กรณีทราบ σ, σ^2

$\Rightarrow \bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

2. กรณีไม่ทราบ σ, σ^2

$(n>30) \Rightarrow \bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$

$(n<30) \Rightarrow \bar{x} \pm t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

3. ประมาณค่าสัดส่วน

$\Rightarrow p \pm z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}$

แบบจุด $\Rightarrow \hat{\pi} = p = \frac{x}{n}$

$\sigma \Rightarrow \hat{\sigma}_p = s_p = \sqrt{\frac{p(1-p)}{n}}$

การประมาณค่า 2 กลุ่มปชก.

tip จะใช้ตาราง t ; $n<30, n_1+n_2<30$

ปชก.ทั้ง 2 กลุ่มเป็นอิสระ

1. กรณีทราบ σ, σ^2

- การแจกแจงปกติ $\Rightarrow (\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\, \sigma_{\bar{x}_1 - \bar{x}_2}$
- ไม่มีแจกแจงปกติ ขนาดใหญ่ ↑

2. กรณีไม่ทราบ σ, σ^2 แต่มีการแจกแจงปกติ

- $n_1+n_2 \geq 30 \Rightarrow (\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\, \hat{\sigma}_{\bar{x}_1 - \bar{x}_2}$; $\hat{\sigma} = s$ ใส่ s แทน σ
- $n_1+n_2 < 30 \Rightarrow (\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2, n_1+n_2-2}\, \hat{\sigma}_{\bar{x}_1 - \bar{x}_2}$ (3)

$\sigma_1^2 = \sigma_2^2 = \sigma^2$ $\quad \hat{\sigma}_{\bar{x}_1 - \bar{x}_2} = \sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ (2)

$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$ (1)

3. ประมาณค่าสัดส่วน

$\Rightarrow (p_1 - p_2) \pm Z_{\alpha/2}\, \hat{\sigma}_{p_1 - p_2}$

$\hookrightarrow \frac{x_1}{n} - \frac{x_2}{n}$

คุณสมบัติตัวประมาณค่า

1. ค. ไม่เอนเอียง $\Rightarrow$ ตัวประมาณค่ามี $\bar{x}$ และ $E(x)$ = parameter
2. ค. ต้องกัน $\Rightarrow$ ตัวประมาณค่ามีค่าเข้าใกล้ parameter prob. = 1 เมื่อ $n \rightarrow \infty$
3. ค. มีประสิทธิภาพ
4. ค. เพียงพอ

บท 7 ทดสอบสมมติฐาน

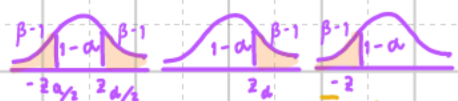
รูปแบบการทดสอบสมมติฐาน

1 กลุ่มปชก. (p ก็ใส่ π แทน)

2 ด้าน, ขวา, ซ้าย

$H_0: \mu = \mu_0 \;,\; \mu \leq \mu_0 \;,\; \mu \geq \mu_0$

$H_1: \mu \neq \mu_0 \;,\; \mu > \mu_0 \;,\; \mu < \mu_0$



- ไม่ทราบ σ^2 $(n \geq 30) \Rightarrow z = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$

$(n < 30) \Rightarrow T = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$ แทน $t_{\alpha/2,\, n-1}$

- ทราบ $\sigma^2 \Rightarrow z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

- สัดส่วนปชก. $(p) \Rightarrow z = \dfrac{x - n\pi_0}{\sqrt{n\pi_0(1-\pi_0)}}$; $p = \dfrac{x}{n}$

2 กลุ่มปชก. (p ก็ใส่ π แทน)

2 ด้าน, ขวา, ซ้าย

$H_0: \mu_1 - \mu_2 = 0 \;,\; \mu_1 - \mu_2 \leq 0 \;,\; \mu_1 - \mu_2 \geq 0$

$H_1: \mu_1 - \mu_2 \neq 0 \;,\; \mu_1 - \mu_2 > 0 \;,\; \mu_1 - \mu_2 < 0$

- ทราบ σ^2 (มีการแจกแจงปกติ) $\Rightarrow z = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

$\bar{x}_1 - \bar{x}_2$

- ไม่ทราบ σ^2 $(n_1 + n_2 \geq 30) \Rightarrow z = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

$(n_1 + n_2 < 30) \Rightarrow T = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{s_p^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$

$\sigma_1^2 = \sigma_2^2 = \sigma^2$

$s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$

แทน $z \rightarrow t_{\alpha,\, n_1+n_2-2}$

- สัดส่วนปชก. $(p_1 - p_2)$; $q = 1 - p$

$d_0 = 0 \Rightarrow z = \dfrac{p_1 - p_2}{\sqrt{pq\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$ $\quad d_0 \neq 0 \Rightarrow z = \dfrac{(p_1 - p_2) - d_0}{\sqrt{\dfrac{p_1 q_1}{n_1} + \dfrac{p_2 q_2}{n_2}}}$

$d_0 \neq \rightarrow$ คือ % ที่โจทย์บอก

Step

1. ตั้ง H_0, H_1
2. กำหนด α
3. วาดรูป
4. คำนวณค่า z (p-value)
5. สรุปผล: p-value < α Rejected H_0

p-value > α Accepted H_0

บท 11 ANOVA

สิ่งที่ต้องรู้ก่อนทำ ANOVA

① รู้ค่า k ② รู้ N ③ $F^{v_1,\, v_2}_{k-1,\, k-N,\, \alpha}$ → นท.กำหนด

ทำ ANOVA

SSB	k-1	MSB = SSB / k-1	F = MSB / MSE
SSE	N-k	MSE = SSE / N-k	
SST (total)	N-1		

บท 10 ถดถอย

การถดถอยเชิงเดี่ยว

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

$\hat{\beta}_1 = \dfrac{\Sigma(x - \bar{x})(y - \bar{y})}{\Sigma(x - \bar{x})^2}$ หรือ $\hat{\beta}_1 = \dfrac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$

$\hat{\beta}_1$ = ค.ชัน , $\hat{\beta}_0$ = จุดตัดแกน x

ดูทิศทางให้ดู $\hat{\beta}_1$ ว่า +, -

สหสัมพันธ์

ρ (rho) สัมประสิทธิ์สหสัมพันธ์ปชก.

r " ______ " ตย. $\boxed{-1 \leq r \leq 1}$

สัมประสิทธิ์การตัดสินใจ

r^2 $\boxed{0 \leq r^2 \leq 1}$ ถามหา % $\Rightarrow r^2 \times 100$