การวิเคราะห์ความแปรปรวน(Analysis of variance)

ชื่อเทคนิคการวิเคราะห์ทางสถิติ : การวิเคราะห์ความแปรปรวน(Analysis of Variance) หรือที่นิยมเรียกกันว่า F-test

ประโยชน์ในเชิงสถิติ: ใช้ในการทดสอบค่าเฉลี่ยของประชากรตั้งแต่สามกลุ่มขึ้นไปว่าเท่ากันหรือไม่

ตารางสถิติที่ใช้ประกอบ: ตารางค่า F (***ไม่จำเป็นต้องใช้ตารางหากวิเคราะห์ด้วยโปรแกรมสำเร็จรูปทางสถิติ***)

ลักษณะของการทดสอบทางสถิติ: เป็นการทดสอบแบบหางเดียวและหางขวา (one-tailed and right-tailed test)

ลักษณะสมมติฐานทางสถิติที่ใช้ : H_o : $\mu 1 = \mu 2 = \mu 3$ (ค่าเฉลี่ยของประชากรมีค่าเท่ากัน) สมมติมีเพียง 3 ประชากร

$H_A : \mu_1 \neq \mu_2 \neq \mu_3$ (ค่าเฉลี่ยของประชากรไม่เท่ากัน)

ประโยชน์ในการประยุกต์ใช้งาน : ใช้ในหลายๆกรณีตัวอย่างเช่น

1)วัดว่าวิธีการสอนต่างๆกันมีประสิทธิภาพแตกต่างกันหรือไม่

2)วิธีการส่งเสริมการขายรูปแบบต่างๆกันมีผลต่อยอดขายหรือไม่

3) วัดความคิดเห็นหรือทัศนคติของคนหลายกลุ่มว่าเหมือนกันหรือไม่

4) ประเภทของกะที่ทำงานมีผลต่อผลิตภาพของพนักงานหรือไม่

5) วิธีการเรียกเก็บหนี้มีผลต่อจำนวนวันที่หนี้คงค้างอยู่หรือไม่

6) ประตูของห้างสรรพสินค้าแต่ละประตูมีคนใช้เข้าร้านแตกต่างกันหรือไม่

รูปแบบ: คำนวณหาค่าสถิติ F ซึ่งคำนวณจากค่าความแปรปรวนระหว่างกลุ่ม (between-group sample variance) หารด้วยค่าความแปรปรวนภายในกลุ่ม(within-group sample variance) ถ้ากำหนดว่าประชากรหรือกลุ่มไม่มีความแตกต่างกัน หรือใกล้เคียงกันมากจนอาจถือว่าเป็นกลุ่มเดียวกัน ค่าความแปรปรวนทั้งสองจะมีค่าใกล้เคียงกันและค่าสถิติ F จะใกล้เคียงหรือเท่ากับหนึ่ง

ในกรณีที่เป็นการวิเคราะห์ความแปรปรวนทางเดียว สูตรคำนวณค่าสถิติ F สามารถแสดงได้ดังนี้

$$F = \frac{\sigma_B^2}{\sigma_W^2} = \frac{\frac{\Sigma n_i (\bar{x}_i - \bar{X}_{GM})}{k-1}}{\Sigma \left[\frac{n_j - 1}{n_T - k} \right] s^2_j}$$

โดยที่ F = ค่าสถิติ F

σ_B^2 = ค่าความแปรปรวนระหว่างกลุ่ม

σ_W^2 = ค่าความแปรปรวนภายในกลุ่ม

$\bar{x}_i$ = ค่าเฉลี่ยจากกลุ่มตัวอย่างที่ j

$\bar{X}_{GM}$ = Grand mean

S^2_i = ค่าความแปรปรวนจากกลุ่มตัวอย่างที่ j

n_i = ขนาดของกลุ่มตัวอย่างที่ j

k = จำนวนกลุ่มประชากร

n_T = จำนวนตัวอย่างจากทุกกลุ่มรวม

ตัวอย่างที่ 1:ผู้จัดการฝ่ายพัฒนาบุคคลของบริษัทผลิตสินค้าอย่างหนึ่งได้ทดลองใช้วิธีการในการฝึกพนักงานในการผลิตสินค้า 3 วิธีด้วยกัน และต้องการดูว่า ผลิตภาพของพนักงานที่ได้รับการฝึกนั้นแตกต่างกันหรือไม่ ตารางต่อไปนี้แสดงให้เห็นผลิตภาพของพนักงานที่ผ่านการฝึกทั้ง 3 วิธี

วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
15	22	18
18	27	24
19	18	19
22	21	16
11	17	22
		15

ค่าความแปรปรวนคำนวนจากค่าเฉลี่ยของกลุ่มตัวอย่าง= 20

ค่าความแปรปรวนคำนวณจากค่าความแปรปรวนในแต่ละกลุ่มตัวอย่าง=14.769

ดังนั้นค่าสถิติ F = ค่าความแปรปรวนคำนวณจากค่าเฉลี่ยของกลุ่มตัวอย่าง =20/14.769 =1.354

ค่าความแปรปรวนคำนวณจากค่าความแปรปรวนในแต่ละกลุ่มตัวอย่าง

ถ้ากำหนดให้ α =0.05 degree of freedom= 2(จำนวนวิธี-1) และ 13(จำนวนตัวอย่างทั้งหมด-จำนวนวิธี) ค่าวิกฤติ F=3.81

เนื่องจากค่าสถิติ F ที่คำนวณได้น้อยกว่าค่าวิกฤติ F ดังนั้นหมายความว่าค่าความแปรปรวนที่ได้จากการคำนวณทั้ง 2 แบบมีค่าไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ กล่าวอีกนัยหนึ่งก็คือ วิธีการในการฝึกพนักงานไม่มีผลต่อผลิตภาพของพนักงาน

ตัวอย่างที่ 2:ฝ่ายจัดการเครือข่ายร้านอาหารจานด่วนแห่งหนึ่งซึ่งมีสาขาอยู่ 4 แห่งได้รับการร้องเรียนจากลูกค้าว่า ให้บริการค่อนข้างช้าไม่ทันใจ จึงตัดสินใจที่จะทำการศึกษาว่ามีความแตกต่างในเวลา(เป็นนาที) ในการให้บริการแก่ลูกค้า และได้ทำการสุ่มตัวอย่างเวลาในการบริการแก่ลูกค้าในร้านทั้ง 4 ปรากฏผลดังนี้

ร้านที่ 1	ร้านที่ 2	ร้านที่ 3	ร้านที่ 4
4	3	2	3
5	3.5	3.5	4
6.5	4.5	5	5.5
4.5	4	6.5	2.5
5	5.5	7	2.5

ค่าความแปรปรวนคำนวณจากค่าเฉลี่ยของกลุ่มตัวอย่าง= 2.3500

ค่าความแปรปรวนคำนวณจากค่าความแปรปรวนในแต่ละกลุ่มตัวอย่าง=1.9375

ดังนั้นค่าสถิติ F = ค่าความแปรปรวนคำนวณจากค่าเฉลี่ยของกลุ่มตัวอย่าง = 2.3500 /1.9375 =1.213

ค่าความแปรปรวนคำนวณจากค่าความแปรปรวนในแต่ละกลุ่มตัวอย่าง

ถ้ากำหนดให้ α =0.05 degree of freedom= 3(จำนวนวิธี-1) และ 16(จำนวนตัวอย่างทั้งหมด-จำนวนวิธี) ค่าวิกฤติ F=3.24

เนื่องจากค่าสถิติ F ที่คำนวณได้น้อยกว่าค่าวิกฤติ F ดังนั้นหมายความว่าค่าความแปรปรวนที่ได้จากการคำนวณทั้ง 2 แบบมีค่าไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ กล่าวอีกนัยหนึ่งก็คือ ร้านอาหารทั้ง 4 ร้านใช้เวลาในการบริการลูกค้าไม่แตกต่างกัน ดังนั้นสิ่งที่ลูกค้าclaimมาจึงไม่มีมูล

ตารางด้านล่างแสดงผลที่ได้จากโปรแกรมสำเร็จรูปทางสถิติ

Oneway

```
[DataSet0]  C:\Users\danai_pat\Documents\ex-anova-fast food.sav
```

ANOVA

X

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	7.050	3	2.350	1.213	.337
Within Groups	31.000	16	1.938		
Total	38.050	19			

การต่อยอดเทคนิคนี้: ในกรณีที่ผู้วิเคราะห์คาดว่าอาจมีปัจจัยตัวที่สองที่มีผลต่อค่าที่คำนวณได้ ผู้วิเคราะห์อาจใช้เทคนิคเดียวกัน แต่จะเป็นเทคนิควัดความแปรปรวนที่เกิดจากปัจจัยสองปัจจัย(two-way factor analysis of variance) ตลอดจนผลเชื่อมโยงระหว่างทั้งสองปัจจัย ดังตัวอย่างต่อไปนี้

ตัวอย่างที่ 3: ผู้ผลิตสินค้าแห่งหนึ่งกำลังต้องการศึกษาว่า การจัดวางสินค้าผลิตภัณฑ์ของตนในห้างสรรพสินค้าตลอดจนสีของบรรจุภัณฑ์จะมีผลต่อยอดขายของสินค้าหรือไม่ ตารางต่อไปนี้แสดงให้เห็นยอดขายรวมของสินค้าทั้งประเทศใน 3 สัปดาห์ที่ผ่านมาแยกตามการจัดวางของและสีของบรรจุภัณฑ์ (หน่วย: 10,000 ชิ้น)

	สีแดง	สีส้ม	สีฟ้า
ระดับพื้นล่างสุด	0.01 0.02 0.01	4.5 5.3 5.2	0.4 0.5 0.4
ระดับเอว	0.04 0.05 0.03	3.1 2.5 3.6	0.2 0.4 0.4
ระดับสายตา	0.01 0.02 0.02	2.4 1.3 1.5	0.5 0.4 0.6
ระดับบนสุด	0.01 0.00 0.01	0.8 1.1 1.0	0.3 0.2 0.4

ค่าความแปรปรวนคำนวณจากค่าเฉลี่ยของกลุ่มตัวอย่างตามลักษณะการจัดวางสินค้า=3.23

ค่าความแปรปรวนคำนวณจากค่าเฉลี่ยของกลุ่มตัวอย่างตามสีของบรรจุภัณฑ์ =25.14

ค่าความแปรปรวนอันเกิดจากความเชื่อมโยงของการจัดวางสินค้าและสีของบรรจุภัณฑ์ =3.08

ค่าความแปรปรวนคำนวณจากค่าความแปรปรวนในแต่ละกลุ่มตัวอย่าง=0.07

ดังนั้นค่าสถิติ F ที่ใช้วัดผลของการจัดวางสินค้า = 3.23/0.07 =43.2

ค่าสถิติ F ที่ใช้วัดผลของสีของบรรจุภัณฑ์ = 25.14/0.07 = =336.4

ค่าสถิติ F ที่ใช้วัดผลของความเชื่อมโยงของการจัดวางสินค้าผนวกเข้ากับสีของบรรจุภัณฑ์ = 3.08/0.07 =41.2

ถ้ากำหนดว่าค่า α =0.05 และที่ degree of freedom ของเศษ(จำนวนแถว-1)และส่วน(จำนวนแถวX จำนวนสดมภ์X ขนาดของตัวอย่างหักลบหนึ่ง)เท่ากับ 3,24 ตามลำดับ ค่าวิกฤติ F = 3.01 ซึ่งน้อยกว่าค่า F ที่คำนวณได้(43.2) ดังนั้นลักษณะการจัดวางสินค้ามีผลต่อยอดขายของสินค้า

ในทำนองเดียวกัน ที่ degree of freedom ของเศษ(จำนวนสดมภ์-1)และส่วน(จำนวนแถวXจำนวนสดมภ์ X ขนาดของตัวอย่างหักลบหนึ่ง)เท่ากับ2,24 ตามลำดับ ค่าวิกฤติ F=3.40 ซึ่งน้อยกว่าค่า F ที่คำนวณได้(336.4) ดังนั้นสีของบรรจุภัณฑ์มีผลต่อยอดขายของสินค้า

และเช่นเดียวกัน ที่ degree of freedomของเศษ (จำนวนแถวลบหนึ่งXจำนวนสดมภ์ลบหนึ่ง) และส่วนเท่ากับ 6,24 ตามลำดับ ค่าวิกฤติ F=2.51 ซึ่งน้อยกว่าค่า F ที่คำนวณได้ (41.2) ดังนั้นความเชื่อมโยงของทั้งลักษณะการจัดวางสินค้าและสีของบรรจุภัณฑ์มีผลต่อยอดขาย

ข้อสังเกต: การวิเคราะห์แบบนี้ตอบได้แต่เพียงว่าค่าเฉลี่ยของผลที่วัดจากวิธีการต่างๆแตกต่างกันหรือไม่ แต่ไม่ได้ตอบคำถามว่าวิธีการไหนดีที่สุด

Contribution this issue: ดร. ดนัย ปัตตพงศ์*

* ผู้แต่ง/เจ้าของ Statistics Talks(70+ issues) เอกสารวิชาการด้านการวิจัยและสถิติประยุกต์ที่นำเสนอเครื่องมือทางสถิติเพื่อการวิจัยในระดับปริญญาโท เอก ตลอดจนตัวอย่างของการนำไปประยุกต์ใช้ในชีวิตจริง

* ผู้พัฒนา FINALYSIS... a dedicated software สำหรับให้บริการนักธุรกิจที่ต้องการวิเคราะห์ความเป็นไปได้ทางการเงินของโครงการพัฒนาอสังหาริมทรัพย์(บ้านจัดสรร/จัดสรรที่ดินเพื่อการอุตสาหกรรม/อาคารชุด/อาคารสำนักงานให้เช่า) โรงแรม โรงพยาบาลเอกชน ห้างสรรพสินค้า โรงงานน้ำตาล โรงงานกระดาษ โรงไฟฟ้าชีวมวล ฯลฯได้เห็นตัวเลขก่อนโครงการเกิดหลีกเลี่ยงความผิดพลาดเป็นร้อยเป็นพันล้านหากเกิดการลงทุนจริง (กำหนด DEBUT 1 เมษายน 2569)

* ผู้แต่งหนังสือ" การวิเคราะห์ความเป็นไปได้ทางการเงินและการจัดวงเงินเครดิตของโครงการลงทุน " ประกอบด้วยตัวอย่างของธุรกิจจริงที่ไม่เปิดเผยชื่อนับ 100 บริษัท ครอบคลุมอุตสาหกรรม 24 อุตสาหกรรม

* Co-developer ซอฟแวร์ en◉gex◉cel® สำหรับใช้ทดสอบ/เรียนรู้ศัพท์ (ประกอบด้วยแบบฝึกหัดและเฉลยกว่า 90 บทครอบคลุมศัพท์ระดับ SAT/IELTS/TOEFL กว่า 12,000 คำ) และไวยากรณ์อังกฤษ (ประกอบด้วยแบบฝึกหัดและเฉลยกว่า 160 บทหรือกว่า 10,000 ข้อครอบคลุมเนื้อหาระดับอุดมศึกษาและTOEFL) มาพร้อมกับไฟล์เสียง/ไฟล์ข้อมูล/ฯลฯ อีกมาก (กำหนด DEBUT 1 เมษายน 2569)