บทที่ 3
การแปลความหมายของคะแนนที่ได้จากแบบทดสอบ

เค้าโครงเรื่อง

1. คะแนนดิบ
2. การวัดแนวโน้มเข้าสู่ส่วนกลางและการวัดการกระจาย
3. เกณฑ์ปกติชนิดต่าง ๆ ที่นิยมใช้ในแบบทดสอบทางจิตวิทยา

สาระสำคัญ

1. คะแนนดิบ คือ ผลหรือคะแนนที่ได้จากการทดสอบ

2. การวัดแนวโน้มเข้าสู่ส่วนกลาง ประกอบด้วย มัชฌิมเลขคณิต มัธยฐาน และฐานนิยม ส่วนการวัดการกระจายประกอบด้วยพิสัย ส่วนเบี่ยงเบนค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และความแปรปรวน

3. เกณฑ์ปกติคือผลที่กลุ่มตัวอย่างทำได้ ประกอบด้วยชั้นเรียนเป็นเกณฑ์ (grade norms) อายุเป็นเกณฑ์ (Age Norms) เปอร์เซ็นไทล์ และคะแนนมาตรฐาน

วัตถุประสงค์ของการเรียน

1. อธิบายและหาค่าคะแนนดิบได้
2. หาค่าการวัดแนวโน้มเข้าสู่ส่วนกลางชนิดต่าง ๆ ได้
3. หาค่าการวัดการกระจายชนิดต่าง ๆ ได้
4. อธิบายได้ว่าเกณฑ์ปกติคืออะไร มีอะไรบ้าง และแตกต่างกันอย่างไร

ในบทนี้จะกล่าวถึงสถิติที่จะเข้ามาเกี่ยวข้องกับการแปลความหมายของคะแนนที่ได้จากการทดสอบซึ่งประกอบด้วยเรื่องของคะแนนดิบ การวัดแนวโน้มเข้าสู่ส่วนกลาง การวัดการกระจายและเกณฑ์ปกติ

1. คะแนนดิบ (Raw Scores)

ในแบบทดสอบทุกประเภท หลังจากที่ทำการทดสอบแล้ว ผลหรือคะแนนที่ได้จากแบบทดสอบนั้น ๆ เรียกว่า คะแนนดิบ ก. ได้คะแนน 15 คะแนนจากการทดสอบการสะกดคำ จากประโยคนี้หมายความว่าอย่างไร และเราจะแปลความหมายของมันได้อย่างไร การที่ ก. ได้คะแนน 15 คะแนน อาจจะเป็น 15 จาก 15 หรือ 15 จาก 50 หรือ 15 จาก 20 หรือแม้แต่เราจะรู้ว่าเป็น 15 จาก 20 ซึ่งคือ 75 เปอร์เซ็นต์ ก็ไม่ได้บอกอะไรกับเราได้เลย มันจะมีความหมายก็ต่อเมื่อเรามีมาตรฐานหรือเกณฑ์อะไรที่เราสามารถที่จะเปรียบเทียบได้ ดังนั้นในแบบทดสอบทุกประเภทจะต้องมีเกณฑ์เอาไว้ให้ ซึ่งเรียกว่า **เกณฑ์ปกติ (Norms)** ในเรื่องของเกณฑ์ปกตินั้น เป็นเรื่องของผลที่ได้จากกลุ่มตัวอย่างของแบบทดสอบที่วางไว้ ซึ่งเกณฑ์ปกติคือ "ผล" ที่กลุ่มตัวอย่าง (Standardized Group) ทำได้นั่นเองในการที่จะทำให้เราสามารถทราบได้ว่า ผู้ถูกทดสอบอยู่ที่ตำแหน่งใดนั้นคือ การนำเอาคะแนนดิบที่ได้มาเปลี่ยนแปลงเป็นคะแนนตามที่เกณฑ์ปกติกล่าวไว้ โดยมีวัตถุประสงค์เพื่อให้ทราบว่า ผู้ถูกทดสอบอยู่ในตำแหน่งใดเมื่อเทียบกับกลุ่ม และทำให้เปรียบเทียบความสามารถของผู้ถูกทดสอบในแบบทดสอบที่ต่างชนิดกันได้ด้วย วิธีหนึ่งที่จะเปรียบเทียบก็คือ การเปรียบเทียบผู้ถูกทดสอบกับชั้นเรียน แล้วดูว่าผู้ถูกทดสอบนั้นได้เท่ากับชั้นใด เมื่อเทียบกับกลุ่มตัวอย่างหรือเทียบกับอายุจริงว่า ความสามารถของเขาเมื่อเทียบกับอายุจริงแล้วเป็นอย่างไรหรืออีกวิธีก็คือ ดูว่าคะแนนของเขาเมื่อเทียบกับกลุ่มตัวอย่างแล้ว เขาอยู่ต่ำกว่ากลุ่มในอัตราส่วนต่อร้อยคน หรือในรูปของต่ำกว่าค่าเฉลี่ยของกลุ่มหรืออยู่ตรงกับส่วนเบี่ยงเบนมาตรฐานเท่าใด ซึ่งที่กล่าวมานี้ก็คือเรื่องของชั้นเรียนเป็นเกณฑ์ (Grade Norms) อายุเป็นเกณฑ์ (Age Norms) **เปอร์เซ็นไทล์** และคะแนนมาตรฐานนั่นเอง แต่อย่างใดก็ตาม ก่อนอื่นเราควรที่จะศึกษาถึงสถิติบางอย่างที่เกี่ยวข้องต่อเนื่องกันเสียก่อน

1.1 การแจกแจงความถี่ (Frequency Distribution)

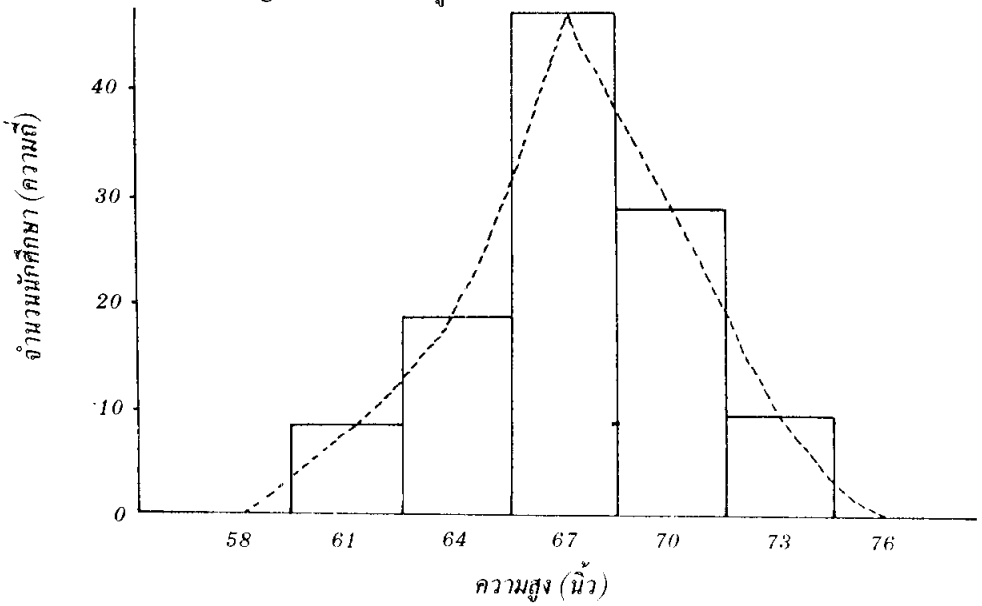
เมื่อได้คะแนนดิบต่าง ๆ แล้ว เราจะรวบรวมคะแนนดิบเหล่านี้เข้าเป็นหมวดหมู่ เพื่อที่จะสามารถมองเห็นได้ว่าคะแนนของแต่ละคนอยู่ในส่วนใดหรือชั้นใด เรียกว่า Class Frequency และจำนวนความถี่ที่แต่ละชั้นมีรวมกัน เรียกว่าเป็น Frequency Distribution

จากตารางที่ 3.1 *แสดงการแจกแจงความถี่ของความสูงของนักศึกษาชาย 100 คน* ***ของมหาวิทยาลัย*** *X Y Z*

ตารางที่ 3.1 *แสดงการแจกแจงความถี่ของความสูงของนักศึกษาชาย 100 คนของมหาวิทยาลัย XYZ*

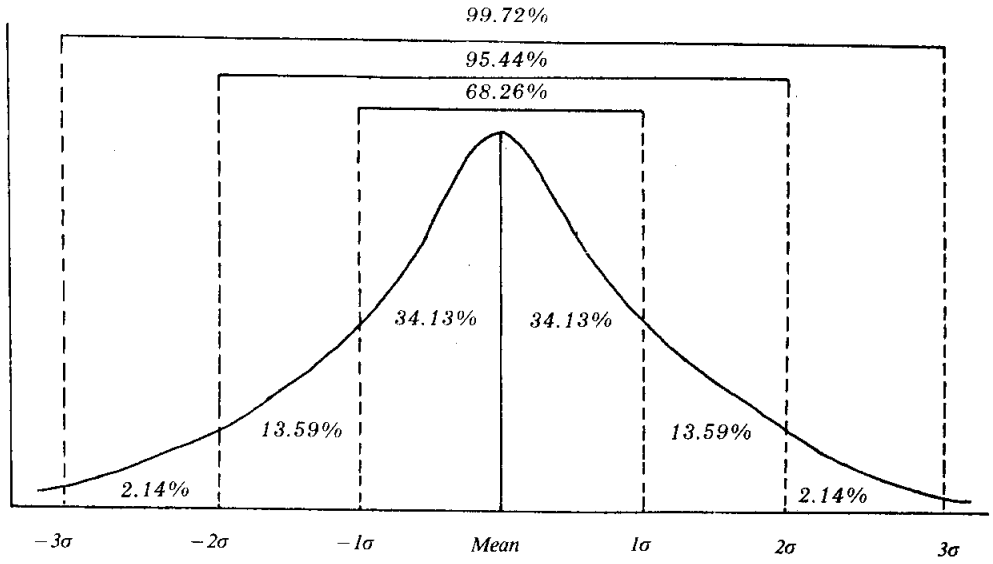
ความสูง (นิ้ว)	จำนวนนักศึกษา
60–62	5
63–65	18
66–68	42
69–71	27
72–74	8
รวม	100

ในชั้นแรกประกอบด้วย นักศึกษาที่มีความสูงระหว่าง 60–62 นิ้ว และมีนักศึกษาอยู่ 5 คน ที่มีความสูงในระดับนี้ นั่นคือ Class Frequency = 5 ซึ่งจากข้อมูลตามตารางที่ 3.1 นี้ เรียกว่า ข้อมูลกลุ่ม (Group data) นอกจากความถี่ของคะแนนจะแสดงตามตารางที่ 3.1 แล้ว ยังสามารถแสดงออกมาในรูปของ Histograms และ Frequency Polygons ได้อีกด้วย โดยมีแกนนอนแทนด้วย X และแกนตั้งแทนด้วย Y โดยให้แกนนอนเป็นความสูงของนักศึกษา ให้มีความกว้างของ histogramsเท่ากับจุดกึ่งกลางของอันตรภาคชั้น ส่วนแกนตั้ง คือ จำนวนความถี่ของนักศึกษาที่มีในแต่ละชั้น ส่วน Frequency Polygons คือเส้นที่ลากผ่านจุดกึ่งกลางของ histogram นั่นเอง ดูจากตารางภาพที่ 3.1



ตารางภาพที่ 3.1 *แสดงความถี่ของข้อมูล*

ซึ่งจากตารางภาพที่ 3.1 นี้จะเห็นว่า ความถี่ที่จุด (Plot) ออกมาจะมีลักษณะคล้าย ๆ ระฆังหรือเรียกว่าเส้นโค้งปกติ ซึ่งเส้นโค้งปกติที่สมบูรณ์ แสดงไว้ที่ตารางภาพที่ 3.2 ซึ่งเส้นโค้งชนิดนี้มีความสำคัญอย่างยิ่งต่อการคำนวณทางสถิติอื่น ๆ



ตารางภาพที่ 3.2 *แสดงเส้นโค้งปกติ*

เส้นโค้งปกติแสดงถึงจำนวนส่วนใหญ่ของข้อมูลที่ตกอยู่ในส่วนกลางและมีบางส่วนที่กระจายออกไป เส้นโค้งนี้มี 2 ด้านที่เหมือนกัน หรือเรียกว่า Symmetry โดยมีจุดสูงสุดอยู่ตรงกลาง ซึ่งลักษณะโดยทั่ว ๆ ไปของคนเราไม่ว่าจะเป็นความสูง น้ำหนักเจตคติ หรือแม้แต่ลักษณะบุคลิกภาพ จะตกอยู่ในรูปของเส้นโค้งปกติ โดยทั่วไปแล้วถ้ายิ่งมีข้อมูลมากเท่าใดก็มีแนวโน้มที่จะเป็นเส้นโค้งปกติมากขึ้นเท่านั้น

ตารางที่ 3.2 *แสดงอัตราส่วนของข้อมูลที่จะตกลงไปในช่วงระหว่างค่าส่วนเบี่ยงเบนมาตรฐานภายใต้เส้นโค้งปกติ*

ข้อจำกัดของข้อมูล	เปอร์เซ็นต์ของข้อมูล
ระหว่างค่าเฉลี่ยกับ + 1 S.D. หรือ − 1 S.D.	34.13
ระหว่างค่าเฉลี่ยกับ + 2 S.D. หรือ − 2 S.D.	47.7
ระหว่างค่าเฉลี่ย + 3 S.D. หรือ − 3 S.D.	49.9
ระหว่าง + 1 S.D. และ −1 S.D.	68.2
ระหว่าง + 2 S.D. และ −2 S.D.	95.4
ระหว่าง + 3 S.D. และ −3 S.D.	99.8

กิจกรรมการเรียนที่ 1

จับคู่ให้นักศึกษานำตัวเลขทางขวามือมาใส่หน้าข้อความทางซ้ายมือให้สัมพันธ์กัน

.....	1. คะแนนที่ได้จากการทดสอบ	1. Norms
.....	2. คะแนนเฉลี่ยของกลุ่มตัวอย่าง	2. Class Frequency
.....	3. ความถี่ของคะแนนดิบที่แต่ละชั้นมีรวมกัน	3. Normal Distribution
.....	4. จำนวนส่วนใหญ่ของข้อมูลที่ตกอยู่ใน ส่วนกลาง	4. Raw Scores
	5. การรวบรวมคะแนนดิบเข้าเป็นหมวดหมู่	5. Frequency Distribution

2. การวัดแนวโน้มเข้าสู่ส่วนกลาง (Central Tendency) และการวัดการกระจาย (Measure of Variabitity)

2.1 การวัดแนวโน้มเข้าสู่ส่วนกลาง (Central Tendency) ซึ่งเป็นที่รู้จักก็คือ

2.1.1 มัชฌิมเลขคณิต (mean, X) โดยหาได้จากผลรวมของคะแนนของข้อมูลทั้งหลาย หารด้วย จานวนข้อมูล

$$\bar{X} = \frac{x_1 + x_2 + x_3 + x_4 + \ldots\ldots\ldots x_N}{N}$$

$$\frac{\sum_{j=1}^{n} X}{N} = \frac{\Sigma X}{N}$$

มัชฌิมเลขคณิต ($\bar{X}$) ของข้อมูลต่อไปนี้ 8, 3, 5, 12, 10 คือ

$$\bar{X} = \frac{8+3+5+12+10}{5} = \frac{38}{5} = 7.6$$

2.1.2 ฐานนิยม (mode) หรือข้อมูลที่มีความถี่สูงสุด ซึ่งใน Frequency Distrbution ฐานนิยม คือ ค่ากึ่งกลางของอันตรภาคชั้นที่มีความถี่มากที่สุด ซึ่งจากตารางที่ 3.1 ฐานนิยม คือ ค่าระหว่าง 66–68 ซึ่งได้แก่ 67 ซึ่งจะได้แก่ค่าสูงสุดของเส้นโค้ง จากภาพที่ 3.1 นั่นเอง

2.1.3 มัธยฐาน หรือ median ซึ่งได้แก่คะแนนที่อยู่ตรงกลางของกลุ่ม หลังจากที่เราได้มีการเรียงลำดับคะแนนจากน้อยไปหามาก หรือมากไปหาน้อยเรียบร้อยแล้ว เช่น

ตัวอย่างที่ 1 กลุ่มของข้อมูล 3, 4, 4, 5, 6, 8, 8, 10, 12

มัธยฐาน = 6

ตัวอย่างที่ 2 กลุ่มของข้อมูล 5, 5, 7, 9, 11, 12, 15, 18,

มัธยฐาน = $\frac{9+11}{2}$ = 10

2.2 การวัดการกระจาย (Measures of Variability) หรือการกระจายแต่ละบุคคลที่กระจายออกจากส่วนกลาง

2.2.1 พิสัย (Range) ซึ่งมีค่าความแตกต่างระหว่างคะแนนสูงสุดและต่ำสุดแต่นับได้ว่าการวัดโดยใช้พิสัยเป็นการวัดที่หยาบที่สุด เพราะจากคะแนนเพียง 2 ตัวเท่านั้น แต่อย่างไรก็ตาม การวัดการกระจายที่นิยมใช้กันมากที่สุด คือ การวัดโดยดูจากความแตกต่างระหว่างคะแนนของแต่ละบุคคล กับมัชฌิมเลขคณิตของกลุ่ม ซึ่งดูจากตารางที่ 3.2 จะทำให้สามารถมองเห็นความแตกต่างของการใช้การวัดการกระจายที่ต่าง ๆ กันออกไป

ตารางที่ 3.3 *แสดงความแตกต่างของการใช้การวัดการกระจายที่ต่างกัน*

X คะแนน	**(x)** ความแตกต่างจาก $\bar{X}$	**(x^2)** ความแตกต่างกำลังสอง
48	+8	64
47	+7	49
43	+3	9
41	+1	1
41	+1	1
40	0	0
38	−2	4
36	−4	16
34	−6	36
32	−8	64

$\Sigma X = 400 \quad \Sigma|x| = 40 \quad x^2 = 244$

$\bar{X} = \frac{\Sigma X}{N} = \frac{400}{10} \quad 40$

Mean deviation $= \frac{\Sigma|X|}{N} = \frac{40}{10} = 4$

Variance $= \sigma^2 = \frac{\Sigma X^2}{N} = \frac{244}{10} = 24.40$

S.D. $\sigma = \sqrt{\frac{\Sigma X^2}{N}} = \sqrt{\frac{244}{10}} = 4.9$

2.2.2 Mean Deviation หรือ Average Deviation

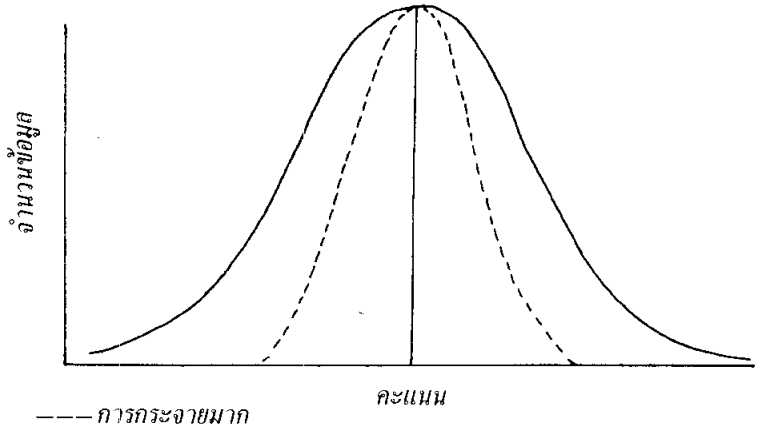
คะแนนดิบแทนตัว X และ x ใช้แทนค่าผลต่างระหว่างคะแนนดิบ (X) กับค่ามัชฌิมเลขคณิตของกลุ่ม ($\bar{X}$) และเครื่องหมาย Σ (อ่านว่า Sum of) คือผลบวกของคะแนนทั้งหมดจากแถวที่ 1 มัชฌิมคณิต ($\bar{X}$) = 40 มัธยฐาน = 40.5 และฐานนิยม คือ 41

แถวที่ 2 แสดงถึงความแตกต่างของคะแนนแต่ละคะแนนที่ไปจาก ($\bar{X}$) ซึ่งในแถวที่ 2 นี้ผลบวกจะ = 0 เพราะผลต่างทั้ง + และ – รอบ ๆ $\bar{X}$ จะเท่ากันเสมอ ถ้าเราไม่สนใจกับเครื่องหมายของความแตกต่าง เรียกว่า mean deviation หรือ Average Deviation สัญลักษณ์ของ $|x|$ หมายถึงค่าความแตกต่างทุกตัวที่นำเอามาใช้นั้นไม่คิดเครื่องหมาย ซึ่ง M.D หรือ A.V นี้ไม่เหมาะที่จะนำมาใช้ในการศึกษาต่อไป เพราะการที่ไม่สนใจกับเครื่องหมายนี้เอง

2.2.3 ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

การวัดการกระจายที่นิยมใช้กันมากที่สุดคือ เรื่องของส่วนเบี่ยงเบนมาตรฐาน (S.D หรือ σ) โดยที่คิดเครื่องหมายและใช้วิธียกกำลัง 2 ค่าความแตกต่างในแถวที่ 2 ทั้งหมด และผลบวกของผลต่างยกกำลังสองหารด้วยจำนวนข้อมูล (N) เรียกว่า Variance หรือ mean square deviation แทนด้วยสัญลักษณ์ σ^2 ซึ่งความแปรปรวน (Variance) นี้มีประโยชน์อย่างยิ่งในการศึกษาความแตกต่างหลาย ๆ ชนิดของแต่ละบุคคลในการทำแบบทดสอบ แต่อย่างไรก็ตาม ในที่นี้เราจะให้ความสนใจกับเรื่องของส่วนเบี่ยงเบนมาตรฐาน ซึ่งมาจากความแปรปรวนมากที่สุด

ตารางภาพที่ 3.3 *แสดงให้เห็นถึงข้อมูลที่มีมัชฌิมเลขคณิตเท่ากัน แต่มีการกระจายที่ต่างกัน*



——— *การกระจายมาก*

..... *การกระจายน้อย*

นอกจากนี้แล้ว S.D. ยังอธิบายคะแนนของแต่ละบุคคลในแบบทดสอบที่ต่างกันในรูปของเกณฑ์ปกติ ซึ่งเราจะพูดถึงในตอนคะแนนมาตรฐาน (standard scores) ซึ่งการอธิบาย S.D จะชัดเจนเมื่ออธิบายในเรื่องเส้นโค้งปกติ จากภาพที่ 3.2 จะเห็นเส้นฐานที่มีระยะทาง 1, 2, 3, เหนือ mean ตัวอย่างจากตาราง 3.2 $\bar{X}$ = 40 ถ้า + 1σ = 44.9 มาจาก (40+4.9) ถ้า + 2σ จะ = 49.8 มาจาก [40+(2×4.9)] และต่อไปเรื่อย ๆ เปอร์เซ็นต์ของคะแนนที่ตกอยู่ระหว่างมัชฌิมเลขคณิตและ +1 ในเส้นโค้งปกติ = 34.13% เพราะการที่เส้นโค้งนั้นเหมือนกัน 34.13% จะตกอยู่ระหว่าง +1σ และ −1σ ซึ่งจะ = 68.26% และจำนวนข้อมูลเกือบทั้งหมด (99.72%) ตกอยู่ระหว่าง −3 และ +3 จากมัชฌิมเลขคณิต ซึ่งความสัมพันธ์อันนี้เราจะมาพูดกันอีกครั้งหนึ่งในเรื่องของคะแนนมาตรฐาน (Standard Scores) และเปอร์-เซ็นไทล์ (Percentile)

กิจกรรมการเรียนที่ 2

จากข้อมูลต่อไปนี้ให้นักศึกษาหา

1. พิสัย
2. มัชฌิมเลขคณิต
3. มัธยฐาน
4. ฐานนิยม
5. ส่วนเบี่ยงเบนมาตรฐาน

8, 9, 11, 12, 6, 7, 5, 4, 3, 5

3. เกณฑ์ปกติ (Norms)

3.1 Age Scores

คำว่า "Mental Age" ได้รับการแนะนำให้รู้จักในปี ค.ศ. 1908 ซึ่งมาจาก Binet-Simon Scales ตามความเห็นของบีเน่ท์ แบบทดสอบมีการแบ่งตามระดับความสามารถโดยที่บีเน่ท์แบ่งตามกลุ่มของอายุ เช่น ถ้าผ่านข้อนี้โดยที่ส่วนใหญ่ของกลุ่มตัวอย่าง (Standardization sample) ที่เป็นตัวแทนของเด็กอายุ 7 ปี ทำได้ นั่นคือ เด็กคนนั้นมีความสามารถอยู่ในระดับของเด็กอายุ 7 ปี ถ้าเด็กคนนั้นทำได้เท่ากับที่เด็กอายุ 8 ปี ทำได้ นั่นคือ เด็กคนนี้นั้นมีความสามารถอยู่ในระดับของเด็กอายุ 8 ปี เป็นต้น เพราะฉะนั้นคะแนนของเด็กในแบบทดสอบนี้คือระดับอายุสูงสุดที่เด็กทำได้ เช่น ถ้าเด็กอายุ 10 ปี สามารถทำได้จนถึงระดับของเด็กอายุ 12 ปี นั่นคือ อายุสมอง "Mental Age" คือ 12 ปีโดยมีอายุตามปฏิทิน (Chronological Age)

10 ปี นั่นคือ เด็กคนนั้นมีความสามารถเหนือกว่าอายุ เพราะเขามีความสามารถเท่ากับเด็กอายุ 12 ปีโดยทั่ว ๆ ไปทำได้ ในการเริ่มต้นทำแบบทดสอบของบีเน่ท์นั้น เริ่มจากการให้ผู้ถูกทดสอบทำจากข้อคำถามที่ต่ำกว่าอายุและผ่านหมด ระดับอายุที่ผู้ถูกทดสอบทำและผ่านหมดเรียกว่า อายุฐาน "Basal Age" และนำเอาคะแนนจากที่ผู้ทดสอบทำผ่านอย่างถูกต้องหมดรวมกับคะแนนที่ผู้ถูกทดสอบทำผ่าน โดยคิดเป็นจำนวนเดือน มารวมเข้าด้วยกัน ผู้ทดสอบจะหยุดการทดสอบเมื่อผู้ถูกทดสอบทำแบบทดสอบนั้นไม่ได้เลย ระดับอายุสูงสุดที่เด็กทำไม่ได้เรียกว่า อายุเพดาน "Ceiling Age" อย่างไรก็ตาม เราควรจะทราบว่า หน่วยของ M.A. จะไม่หยุดอยู่คงที่ แต่มีแนวโน้มที่จะลดลงตามการเพิ่มขึ้นของอายุ เช่น ในเด็ก 4 ขวบ ที่มีอายุสมองช้ากว่าอายุจริงอยู่ 1 ปี จะมีอายุสมองที่ช้ากว่าปกติประมาณ 3 ปี เมื่ออายุจริง = 12 ปี ทั้งนี้เป็นเพราะ 1 ปี ของการเจริญเติบโตของอายุสมองระหว่าง 3–4 ปี จะเท่ากับ 3 ปีของการเจริญเติบโตของอายุสมองจาก 9–12 ปี ที่เป็นเช่นนี้เพราะการเจริญเติบโตของสติปัญญาจะพัฒนาอย่างรวดเร็วและในระยะแรก ๆ ของชีวิตและจะลดลงเมื่อคนเราเจริญเติบโตเต็มที่

เสตอร์น และเครอสเมน ได้แสดงให้เห็นถึง I.Q. ว่าเป็นอันตราส่วนของอายุสมอง (MA) กับอายุจริง (CA) และคูณด้วย 100 เพราะหลีกเลี่ยงทศนิยม

$$\text{I.Q.} = \frac{\text{MA}}{\text{CA}} \times 100$$

ถ้าเด็กคนนี้มี MA = CA นั่นคือ I.Q. = 100 ซึ่งเท่ากับระดับเฉลี่ย ถ้าต่ำกว่า 100 เป็นการแสดงถึงช้าหรือด้อยกว่าปกติ และถ้าสูงกว่า 100 แสดงถึงความสามารถที่สูงกว่าปกติ

3.2 เปอร์เซ็นไทล์ (Percentile)

คะแนนหรือค่าของเปอร์เซ็นไทล์เป็นการแสดงออกในรูปของร้อยละ ที่คนในกลุ่มตัวอย่างมาตรฐานได้คะแนนต่ำกว่าคะแนนดิบที่กำหนดไว้ เช่น ถ้าผู้ถูกทดสอบร้อยละ 28 ทำข้อสอบเหตุผลทางคณิตศาสตร์ได้ถูกต้องต่ำกว่า 15 ข้อ ดังนั้นคะแนนดิบจะตรงกับเปอร์-เซ็นไทล์ที่ 28 นั่นคือ เปอร์เซ็นต์ไทล์แสดงให้ทราบถึงตำแหน่งในกลุ่มจาก 100 คน แต่เปอร์-เซ็นไทล์เริ่มนับจากคะแนนต่ำสุดแทนการนับจากคะแนนสูงสุด ยิ่งตำแหน่งเปอร์เซ็นไทล์ต่ำเท่าใดก็หมายความว่าบุคคลนั้นอยู่ในตำแหน่งที่ต่ำลงเท่านั้น

เปอร์เซ็นไทล์ที่ 50 (P_{50}) มีความหมายเช่นเดียวกับมัธยฐาน ค่าเปอร์เซ็นไทล์ต่ำกว่า 50 แสดงว่า มีความสามารถต่ำกว่าค่าเฉลี่ย ถ้าเปอร์เซ็นไทล์สูงกว่า 50 แสดงว่า มีความสามารถสูงกว่าค่าเฉลี่ย ค่าเปอร์เซ็นไทล์ไม่ควรที่จะนำมาปนกับค่าร้อยละ เพราะค่าร้อยละเป็นค่าของคะแนนดิบที่ได้จากข้อที่ตอบได้ถูกต้อง แต่ค่าเปอร์เซ็นไทล์หมายถึงค่าร้อยละของ

บุคคล คะแนนดิบที่มีค่าต่ำกว่าคะแนนใด ๆ ในตัวอย่างมาตรฐานย่อมมีค่าเปอร์เซ็นไทล์เป็นศูนย์ (P_0) และถ้าคะแนนดิบที่มีค่าสูงกว่าคะแนนใด ๆ ในตัวอย่างมาตรฐานย่อมมีค่าเปอร์-เซ็นไทล์เป็น 99 (P_{99}) ตัวอย่างเช่น ผู้ถูกทดสอบคนหนึ่งทำคะแนนทดสอบได้ 70 คะแนน และคะแนน 70 คะแนนนี้อยู่ตำแหน่งเปอร์เซ็นไทล์ที่ 60 หมายความว่า ร้อยละ 60 ของกลุ่มตัวอย่างได้คะแนนต่ำกว่า 70 คะแนน วิธีหาค่าเปอร์เซ็นไทล์จากหลักสูตร

$$PR = 100 - \frac{100R - 50}{N}$$

PR = ตำแหน่งเปอร์เซ็นไทล์

R = อันดับของคะแนนหรือลำดับที่

N = จำนวนคนในกลุ่ม

3.3 คะแนนมาตรฐาน (Standard Scores)

คะแนนมาตรฐาน หรือ **Z-Scores** คือค่าที่หาได้จากความแตกต่างระหว่างคะแนนดิบกับมัชฌิมเลขคณิตของกลุ่ม และหารด้วยค่า S.D. ของกลุ่ม

$$Z = \frac{X - \bar{X}}{S.D} \text{ หรือ } \frac{x}{S.D}$$

X = คะแนนดิบ

$\bar{X}$ = มัชฌิมเลขคณิตของกลุ่ม

S.D. = ค่าเบี่ยงเบนมาตรฐานของกลุ่ม

ผู้ถูกทดสอบ	คะแนนดิบ (X)	$(X - \bar{X})$	Z
A	3	−7	−1.11
B	6	−4	−.63
C	7	−3	−.47
D	9	−1	−.16
E	15	5	.79
F	20	10	1.58
Σ	60	.00	.00
$\bar{X}$	10	.00	.00
S.D.	6.32	6.32	1.00

ตารางที่ 3.4 *แสดงวิธีคำนวณหาคะแนนมาตรฐาน*

ในการใช้คะแนนมาตรฐานนั้น หมายความว่า เราใช้คะแนนมาตรฐานในฐานะที่เป็นหน่วยของการวัด จากตัวอย่างข้างบน A อยู่ที่ −1.11 S.D. ต่ำกว่ามัชฌิมเลขคณิต ในขณะที่ F อยู่ที่ 1.58 S.D. สูงกว่ามัชฌิมเลขคณิต คะแนนมาตรฐานนิยมใช้ในการเปรียบเทียบคะแนนจากการใช้เครื่องมือทดสอบที่ต่างชนิดกัน พิจารณาคะแนนจากการทดสอบ ภาษาอังกฤษและคณิตศาสตร์ ของผู้ถูกทดสอบกลุ่มเดียวกัน และสมมติว่าค่ามัชฌิมเลขคณิตและส่วนเบี่ยงเบนมาตรฐาน (S.D.) เป็น

วิชา	$\bar{X}$	S.D.
ภาษาอังกฤษ	65	8
คณิตศาสตร์	52	12

ผู้ถูกทดสอบคนหนึ่งทำคะแนนภาษาอังกฤษได้ 57 และได้คะแนนคณิตศาสตร์ 58 เราสามารถที่จะเปรียบเทียบวิชาทั้ง 2 นี้ โดยใช้คะแนนมาตรฐาน

ในวิชาภาษาอังกฤษ $\frac{X-\bar{X}}{S.D.} = \frac{57-65}{8} = -1.0$

ในวิชาคณิตศาสตร์ $\frac{X-\bar{X}}{S.D.} = \frac{58-52}{12} = 0.5$

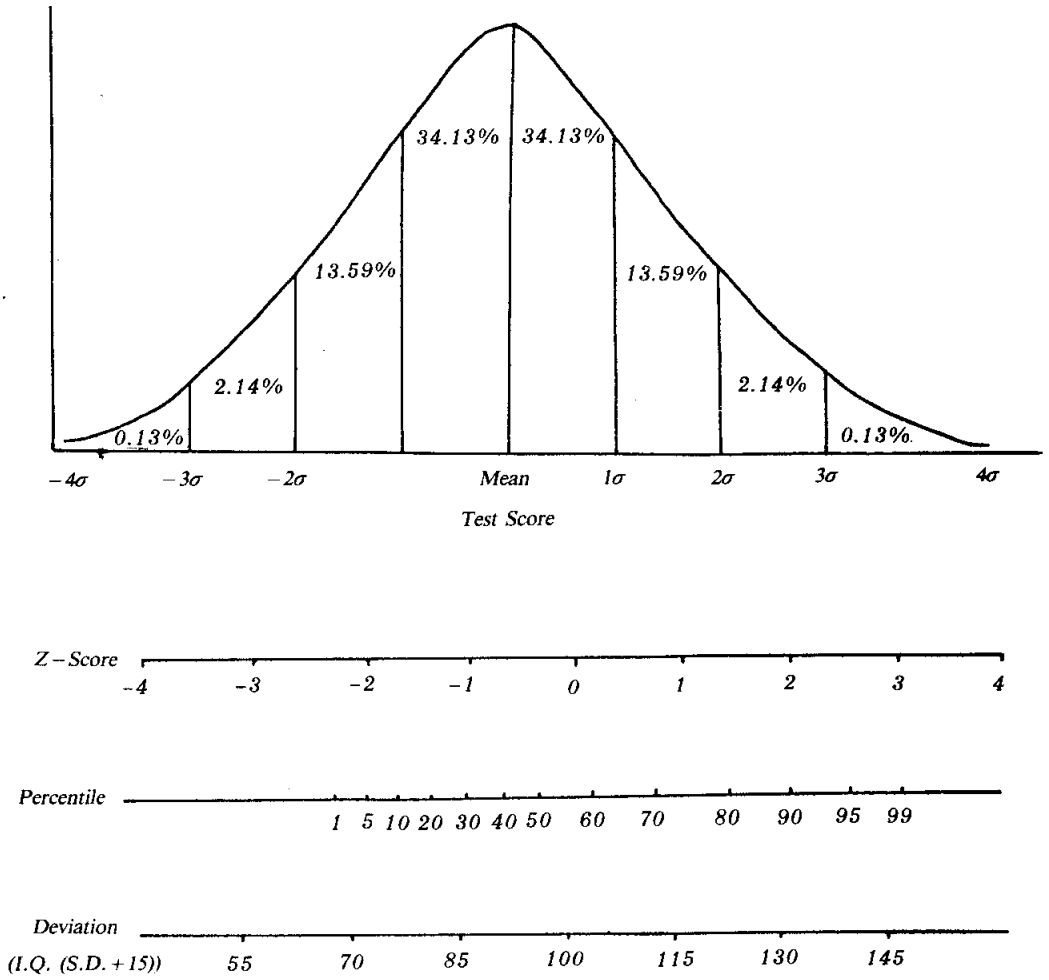
นั่นคือ ในวิชาภาษาอังกฤษ ผู้ถูกทดสอบอยู่ต่ำกว่ามัชฌิมเลขคณิตของกลุ่มอยู่ 1 S.D. ในขณะที่ในวิชาคณิตศาสตร์ ผู้ถูกทดสอบจะอยู่เหนือกว่ามัชฌิมเลขคณิตของกลุ่มอยู่ .5 จากผลอันนี้แสดงว่า ผู้ถูกทดสอบนี้ทำภาษาอังกฤษได้ต่ำกว่ากลุ่ม แต่ทำคณิตศาสตร์ได้ดีกว่ากลุ่ม นอกจากจะใช้คะแนนมาตรฐานดูความสามารถของผู้ถูกทดสอบในต่างแบบทดสอบแล้ว ยังสามารถใช้ศึกษาดูความแตกต่างในความสามารถของบุคคลในกลุ่มเดียวกันได้อีกด้วย เช่น ก. และ ข. ทำแบบทดสอบชุดหนึ่ง ผลปรากฏว่า ก. ได้คะแนน 65 ข. ได้คะแนน 58 มัชฌิมเลขคณิต ($\bar{X}$) ของกลุ่ม = 60, S.D. = 5

คะแนนของ ก. (X_1) = 65 คะแนนของ ข. (X_2) = 58

$Z_1 = \frac{65-60}{5}$ $Z_2 = \frac{58-60}{5}$

$= +1.0$ $= -0.40$

เพราะฉะนั้น คะแนนของ ก. และ ข. เมื่อเทียบกับกลุ่มแล้ว ก. อยู่ในตำแหน่งหรือมีความสามารถที่ดีกว่า โดยเฉลี่ยของกลุ่ม ในขณะที่ ข. อยู่ในตำแหน่งหรือมีความสามารถด้อยกว่ากลุ่มนั่นเอง



ตารางภาพที่ 3.4 *แสดงความสัมพันธ์ระหว่างลักษณะคะแนนชนิดต่าง ๆ ที่มีการแจกแจงปกติ*

3.4 เกณฑ์ปกติทางการศึกษา (Educational Norms)

โดยปกติส่วนใหญ่ การแปลคะแนนในแบบทดสอบความสัมฤทธิผลในความหมายของ "Grade Norms" ซึ่งการอธิบายผลของแบบทดสอบในลักษณะนี้ เป็นการอธิบายความสามารถในฐานะที่เทียบกับระดับความสามารถในการเรียนในชั้นต่าง ๆ เช่น ความสามารถของผู้ถูกทดสอบในการสะกดตัวเท่ากับเด็กที่เรียนอยู่ในชั้น 7 (grade 7) ความสามารถในการอ่านของผู้ถูกทดสอบเท่ากับเด็กที่เรียนอยู่ในชั้น 8 ความสามารถทางคณิตศาสตร์ของผู้ถูกทดสอบเท่ากับเด็กที่เรียนอยู่ในชั้น 5 เพราะฉะนั้น จะเห็นว่าความสามารถของผู้ถูกทดสอบจะใช้เปรียบเทียบเช่นเดียวกับเรื่องของอายุสมอง

Grade Norm ได้มาจากการคำนวณหามัชฌิมเลขคณิตของคะแนนดิบของเด็กในแต่

ละชั้นเรียน ถ้าคะแนนเฉลี่ยจากคะแนนดิบของนักเรียนในกลุ่มตัวอย่าง (Standardization Sample) ในชั้น 4 ในวิชาคณิตศาสตร์ได้เท่ากับ 23 นั่นคือ คะแนน 23 มีความหมายเท่ากับความสามารถของเด็กชั้น 4 อย่างไรก็ตาม ในทางทฤษฎีแล้วเราสามารถที่จะเทียบบัญญัติไตรยางค์หาเศษของชั้นเรียนแต่ละชั้นได้ เพราะใน 1 ปีการศึกษานั้นมีเวลาถึง 10 เดือน เพราะฉะนั้นถ้าทำการทดสอบ ผลสอบโดยเฉลี่ยของตอนเริ่มต้นของชั้น 4 ผลที่ได้คือ 4.0 และถ้าทำตอนกลาง ๆ ปี ผลเฉลี่ยที่ได้อาจเป็น 4.5 และอื่น ๆ ได้เช่นกัน

ตารางที่ 3.5 *แสดง grade norms ความเข้าใจในการอ่านแบบทดสอบ Iowa-Every-Pupil ของนักเรียนชั้น 3 และชั้น 5 (Spitzer, 1947)*

ชั้น 3 เปอร์เซ็นไทล์	เทียบเท่าชั้น	ชั้น 5 เปอร์เซ็นไทล์	เทียบเท่าชั้น
99	8-2	99	9-8
90	6-0	90	8-4
75	4-9	75	7-2
50	3-9	50	5-9
25	3-3	25	4-8
10	2-6	10	4-0
1	1-9	1	2-9

ระดับชั้นเรียนที่เขียนว่า 8-2 หมายความว่ามีความรู้เท่ากับเด็กที่เรียนชั้น 8 มา 2 เดือน จาก 10 เดือนของปีการศึกษา

ถึงแม้ว่าเรื่องของ grade norm จะเป็นที่นิยมใช้กัน แต่ก็คงมีข้อสังเกตอยู่หลายประการ ประการแรก เนื่องมาจากเนื้อหาวิชาที่สอนเปลี่ยนไปตามแต่ละชั้นเรียน แต่ grade norm เหมาะจะใช้กับวิชาที่มีการสอนทุกระดับชั้นเรียน ถึงแม้ว่าทุกชั้นเรียนจะมีวิชาเรียนที่คล้ายคลึงกัน แต่ระดับความก้าวหน้าทางวิชาเดียวกัน ในแต่ละชั้นเรียนต่างกันหรือต่างวิชาในชั้นเดียวกัน หรือพูดได้ง่าย ๆ ว่า หน่วยของชั้นเรียนไม่เท่ากัน และการไม่เท่ากันนี้จะปรากฏให้เห็นอย่างไม่สม่ำเสมอในวิชาที่ต่างกัน

ประการต่อไปคือ grade norm ทำให้เกิดความเข้าใจผิดจากผลที่ได้ นั่นคือ ถ้าผู้ถูกทดสอบอยู่ชั้น 4 ทำสอบวิชาเลขคณิต ได้ผลเท่ากับ 6-9 ผลนี้ไม่ได้หมายความว่า ผู้ถูกทดสอบได้เรียนรู้เลขคณิต ก่อนขึ้นปีที่ 7 มาแล้ว และประการสุดท้าย grade norm จะถูกสรุปว่า เป็นผลของการสอน เช่น ครูที่สอนชั้น 6 อาจจะคิดว่า นักเรียนในชั้นของตนน่าจะมี

ความสามารถใกล้เคียงกันเมื่อทดสอบ เมื่อใช้ grade norm แล้วข้อสงสัยนี้ย่อมจะไม่เกิดขึ้น เพราะ grade norm ย่อมทำให้เกิดพิสัยของความสัมฤทธิผลในชั้นเรียนที่ต่างกันด้วย

ตารางที่ 3.6 *แสดงประเภทของเกณฑ์ปกติสำหรับแบบทดสอบทางด้านการศึกษาและจิตวิทยา*

ประเภทของเกณฑ์ปกติ	สิ่งที่ใช้ในการเปรียบเทียบ	ประเภทของกลุ่มตัวอย่าง
Grade Norms	ผู้ถูกทดสอบกับกลุ่ม	กลุ่มที่ประสบความสำเร็จในชั้นเรียน
Age Norms	ผู้ถูกทดสอบกับกลุ่ม	กลุ่มที่ประสบความสำเร็จตามอายุ
เปอร์เซ็นไทล์	เปอร์เซ็นต์ของกลุ่มตัวอย่างที่อยู่ต่ำกว่าผู้ถูกทดสอบ	กลุ่มของชั้นเรียนหรืออายุที่ผู้ถูกทดสอบเป็นสมาชิกอยู่
คะแนนมาตรฐาน	ค่าของส่วนเบี่ยงเบนมาตรฐานที่ผู้ถูกทดสอบอยู่เหนือหรือต่ำกว่ากลุ่ม	กลุ่มของชั้นเรียนหรืออายุที่ผู้ถูกทดสอบเป็นสมาชิกอยู่

กิจกรรมการเรียนที่ 3

ชุดที่ 1 *จับคู่ให้เลือกตัวเลขทางขวามือมาใส่หน้าข้อความทางซ้ายมือ*

..... 1. ช่วงอายุในแบบทดสอบที่ผู้ถูกทดสอบทำไม่ได้เลย 1. Chronological Age
..... 2. ช่วงอายุรวมทั้งหมดที่ผู้ถูกทดสอบทำได้ 2. Basal Age
..... 3. ช่วงอายุในแบบทดสอบที่ผู้ถูกทดสอบทำได้หมด 3. Ceiling Age
..... 4. อายุจริงตามปฏิทิน 4. Mental Age

ชุดที่ 2 *ถูก-ผิดให้เขียนเครื่องหมาย ✓ ลงหน้าข้อความที่ถูก และ × หน้าข้อความที่ผิด*

..... 1. ใจเรียนอยู่ชั้น ม.4 แต่สอบเทียบชั้น ม.6 ได้แล้วเป็นเรื่องของ Age Norms
..... 2. อุ้ยสอบได้คะแนนสูงสุดในชั้นดังนั้นจึงอยู่ที่ P_{100}
..... 3. กบอายุสมองเท่ากับอายุจริง แสดงว่า กบเรียนตามชั้นเรียน
..... 4. ตุ๊กสอบได้ $Z = +1$ แสดงว่าดีกว่าติ๊กซึ่งอยู่ที่ $Z = -1$
..... 5. แป้งสอบได้คะแนนต่ำสุดในชั้น แสดงว่าแป้งอยู่ที่ P_0

สรุป

คะแนนดิบ (Raw Scores) คือผลหรือคะแนนที่ได้จากการทดสอบ ซึ่งถ้ามีข้อมูลเป็นจำนวนมากเราสามารถจะจัดกระทำข้อมูลโดยการแจกแจงความถี่ของข้อมูลนั้น ๆ และข้อมูลส่วนใหญ่จะตกอยู่ตรงกลางและมีบางส่วนกระจายออกไป ซึ่งเราเรียกการกระจายนี้ว่า เส้นโค้งปกติ (Normal Distribution) นอกจากนี้แล้วคะแนนยังสามารถอธิบายในรูปของการวัดแนวโน้มเข้าสู่ส่วนกลาง (Central Tendency) ที่เป็นที่รู้จักได้แก่ มัชฌิมเลขคณิต (mean) ฐานนิยม (mode) และมัธยฐาน (median) การวัดการกระจาย (Measures of Variability) คือการวัดที่เกี่ยวกับการกระจายของแต่ละบุคคลที่กระจายออกมาจากส่วนกลาง ประกอบด้วยพิสัย (range), mean deviation และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

คะแนนดิบที่ได้จากการทดสอบจะไม่สามารถแปลความหมายได้ถ้าแบบทดสอบนั้นไม่มีเกณฑ์ปกติหรือ Norms ไว้ให้ Norms คือค่าเฉลี่ยที่ได้จากกลุ่มตัวอย่าง Norms ที่นิยมใช้ในแบบทดสอบทางจิตวิทยามีหลายชนิด เช่น Age Norms, Grade Norms, Percentile, Standard Scores เป็นต้น

แบบฝึกหัด

ตอนที่ 1 *ให้นักศึกษาเลือกคำตอบที่ถูกต้องที่สุดเพียงคำตอบเดียว*

1. การนำเอาคะแนนดิบที่ได้มาแปลงเป็นคะแนนตามที่เกณฑ์ปกติกล่าวไว้เพื่อ
   (1) เปรียบเทียบความสามารถของผู้ถูกทดสอบกับกลุ่มตัวอย่าง
   (2) เปรียบเทียบความสามารถของผู้ถูกทดสอบในต่างวิชากัน
   (3) เปรียบเทียบความสามารถของผู้ถูกทดสอบในแบบทดสอบที่ต่างกัน
   (4) เปรียบเทียบความสามารถระหว่างกลุ่ม
   (5) เปรียบเทียบความสามารถระหว่างผู้ถูกทดสอบในกลุ่ม
2. ข้อความใดต่อไปนี้ที่ไม่ถูกต้องในเรื่องของเปอร์เซ็นไทล์
   (1) เปอร์เซ็นไทล์เป็นเรื่องของคนในกลุ่มตัวอย่างมาตรฐานที่ได้คะแนนต่ำกว่าคะแนนดิบ
   (2) ตำแหน่งเปอร์เซ็นไทล์ยิ่งต่ำลง บุคคลนั้นยิ่งมีตำแหน่งต่ำลงด้วย
   (3) ค่าเปอร์เซ็นไทล์สูงสุดคือ 100
   (4) ค่าเปอร์เซ็นไทล์ต่างกับค่าร้อยละ
   (5) P_{50} มีความหมายเช่นเดียวกับมัธยฐาน
3. จากการทดสอบความสัมฤทธิผลในการเรียนวิชาภูมิศาสตร์ ปรากฏว่า อ้วน สอบได้เท่ากับเด็กที่เรียนอยู่ในชั้น ป.6 จากข้อความดังกล่าวเป็นเรื่องของ
   (1) Z-Score (2) Percentile (3) Age-Norm (4) Grade-Norm (5) Mental Age

4. นักการศึกษา ได้สร้างแบบทดสอบเพื่อทดสอบความสัมฤทธิผลในการเรียนวิชาคณิต-ศาสตร์ของนักเรียนชั้น ป.4 พบว่า ผลการทดสอบนั้นไปสอดคล้องกับการทดสอบของครูประจำชั้นด้วย จากข้อความนี้ แบบทดสอบนี้มีลักษณะของ
   (1) ความเที่ยงตรง
   (2) ความน่าเชื่อถือ
   (3) เกณฑ์ปกติ
   (4) คุณค่าในการทำนาย
   (5) ความยากง่ายของแบบทดสอบ
5. คะแนนเฉลี่ยของกลุ่มตัวอย่างเป็นเรื่องของ
   (1) ความเที่ยงตรง
   (2) ความน่าเชื่อถือ
   (3) เกณฑ์ปกติ
   (4) คุณค่าในการทำนาย
   (5) ความยากง่ายของแบบทดสอบ

ข้อ 6-11 ให้นักเรียนศึกษาใช้ข้อมูลต่อไปนี้ในการตอบคำถาม

ผลการทดสอบวิชาภาษาอังกฤษ	15	20	20	30	25	S.D. = 5.09
ผลการทดสอบวิชาภาษาไทย	19	19	16	26	30	S.D. = 5.17

6. ค่า $\overline{X}$ ของผลการสอบวิชาภาษาอังกฤษและภาษาไทย คือ
   (1) 15, 14 (2) 20, 19 (3) 19, 20 (4) 22, 22 (5) 30, 30
7. ค่ามัธยฐาน (median) ของผลการสอบวิชาภาษาอังกฤษและภาษาไทยคือ
   (1) 15, 14 (2) 20, 19 (3) 19, 20 (4) 22, 22 (5) 30, 30
8. ค่าฐานนิยม (mode) ของผลการสอบวิชาภาษาอังกฤษและภาษาไทย คือ
   (1) 15, 14 (2) 20, 19 (3) 19, 20 (4) 22, 22 (5) 30, 30
9. ค่าพิสัย (range) ของผลการสอบวิชาภาษาอังกฤษและภาษาไทย คือ
   (1) 15, 14 (2) 20, 19 (3) 19, 20 (4) 22, 22 (5) 30, 30

10. นาย ก. และนาย ข. สอบวิชาภาษาอังกฤษได้ 18, 21 ตามลำดับ ถ้าอยากทราบว่า นาย ข. สอบได้ตำแหน่งดีกว่า นาย ก. มากน้อยเพียงใด เมื่อเทียบกับกลุ่ม ดูได้จาก
    (1) คะแนนดิบ
    (2) ค่าเฉลี่ย ($\bar{X}$)
    (3) ส่วนเบี่ยงเบนมาตรฐาน (S.D.)
    (4) แปลงค่าเป็นหน่วยของคะแนนมาตรฐาน
    (5) แปลงค่าให้อยู่รูปของเส้นโค้งปกติ
11. นาย ฮ. สอบวิชาภาษาอังกฤษได้ 20 และภาษาไทยได้ 23 ถ้าอยากทราบว่า นาย ฮ. สอบวิชาใดได้ตำแหน่งที่ดีกว่าเมื่อเทียบกับกลุ่ม ดูได้จาก
    (1) คะแนนดิบ
    (2) ค่าเฉลี่ย ($\bar{X}$)
    (3) ส่วนเบี่ยงเบนมาตรฐาน (S.D.)
    (4) แปลงค่าเป็นหน่วยของคะแนนมาตรฐาน
    (5) แปลงค่าให้อยู่ในรูปของเส้นโค้งปกติ
12. นักจิตวิทยาได้สร้างแบบทดสอบเพื่อวัดผลการเรียนคณิตศาสตร์ของเด็กชั้น ป.4 แบบทดสอบที่นักจิตวิทยาสร้างขึ้นนี้เป็นแบบทดสอบ
    (1) สติปัญญา (2) ความถนัด
    (3) ความสัมฤทธิผล (4) ความสามารถพิเศษ (5) ความสนใจ

ตอนที่ 2 *ถูก-ผิด ให้นักศึกษาเขียนเครื่อง ✓ ลงหน้าข้อความที่ถูกและเครื่องหมาย × ลงหน้าข้อความที่ผิด*

..... 1. คะแนนดิบ (Raw Scores) สามารถบอกถึงความสามารถของผู้ถูกทดสอบได้
..... 2. เปอร์เซ็นไทล์เป็นตัวบอกถึงความสามารถของบุคคลคิดเป็นร้อยละเช่นเดียวกับเปอร์เซ็นต์
..... 3. ผู้ถูกทดสอบได้ค่าเปอร์เซ็นไทล์ = 0 ได้
..... 4. อายุสมองเป็นเรื่องของความสามารถของสมองซึ่งสัมพันธ์กับอายุ
..... 5. อายุตามปฏิทิน (CA) คืออายุตามที่เป็นจริง
..... 6. Age Scores คือการใช้ชั้นเรียนเป็นเกณฑ์ในการวัดความสามารถ
..... 7. ผู้ถูกทดสอบได้ค่าเปอร์เซ็นไทล์เท่ากับ 100 ไม่ได้
..... 8. Grade Norm เป็นเกณฑ์ปกติที่นิยมใช้ในแบบทดสอบความสัมฤทธิผล
..... 9. Grade Norm มีผลทำให้เกิดพิสัยของสัมฤทธิผลในชั้นเรียนที่ต่างกัน
..... 10. คะแนนที่ได้จากการทดสอบทางจิตวิทยาจะมีการตัดสินว่าได้หรือตก

ตอนที่ 3 *ให้นักเรียนตอบคำถามต่อไปนี้*

1. คะแนนดิบคืออะไรและเกี่ยวข้องกับเกณฑ์ปกติ (Norms) อย่างไร

2. การวัดแนวโน้มเข้าสู่ส่วนกลางคืออะไร อธิบายพร้อมบอกตัวอย่างประกอบ

3. การวัดการกระจายมีประโยชน์อย่างไร

4. Age Norms คืออะไรและเกี่ยวข้องกับการหาค่า I.Q. อย่างไร

5. คะแนนมาตรฐานคืออะไรและมีประโยชน์อย่างไรในการทดสอบ