

การตัดเกรด

ผศ. สมนึก ภัททิยธนี *

การวัดคุณภาพด้านต่าง ๆ ของนักเรียนไม่ว่าจะใช้เครื่องมือชนิดใดก็ตาม ส่วนใหญ่จะปรากฏผลออกมาเป็นตัวเลข โดยเฉพาะเมื่อใช้แบบทดสอบ ตัวเลขที่ใช้แทนปริมาณความมากน้อยของคุณภาพของสิ่งที่ต้องการจะวัดนี้ เรียกว่า คะแนน โดยทั่วไปจะแบ่งคะแนนออกเป็น 2 ประเภท ได้แก่ คะแนนดิบและคะแนนแปลงรูป

1. คะแนนดิบ (Raw Score) เป็นคะแนนที่ปรากฏออกมาจากการสอบโดยตรงไม่สามารถตีความหมายที่แน่ชัดได้ว่ามีสภาพการเรียนรู้มากน้อยเท่าไร จึงจัดว่าเป็นตัวเลขลอย ๆ ที่ไม่มีความหมาย เช่น นายแดงสอบได้ 20 คะแนน แม้จะเอาไปเทียบกับคะแนนเดิมก็ให้ความหมายไม่ชัดเจน

2. คะแนนแปลงรูป (Derived Score) เป็นคะแนนที่ได้จากการนำคะแนนดิบไปเปลี่ยนให้เป็นคะแนนที่มีความหมายดีขึ้นกว่าเดิม คือ ทำให้สามารถบอกสภาพการเรียนรู้ของเด็กได้แน่ชัดขึ้นว่า เขาเก่งหรืออ่อนในวิชาใด เพียงใด คะแนนแปลงรูปนี้อาจจำแนกเป็นประเภทย่อย ๆ ได้หลายแบบ ซึ่งต่างก็ตีความได้ชัดเจนแตกต่างกัน ได้แก่

2.1 คะแนนเปอร์เซ็นต์ ได้แก่ การนำคะแนนที่ได้ไปเปรียบเทียบกับคะแนนเดิมโดยเปลี่ยนคะแนนเดิมให้มีค่าเป็น 100 เช่น สอบได้ 20 คะแนน จากคะแนนเดิม 50 แสดงว่า ถ้าคะแนนเดิม 100 จะสอบได้ 40 คะแนนหรือ เรียกว่า 40 %

2.2 คะแนนอันดับที่ ได้แก่ การนำคะแนนดิบที่ได้ ไปกำหนดตำแหน่งโดยเรียงลำดับคะแนนที่ได้จากจำนวนคนทั้งหมด แล้วมุ่งพิจารณาอันดับที่ได้เป็นสำคัญ เช่น สอบได้ 20 คะแนน เมื่อเรียงลำดับคะแนนแล้วอยู่ใน อันดับที่ 10 จากคนเข้าสอบทั้งหมด 30 คน หรือคิดเทียบอันดับที่ได้นี้จากคนเข้าสอบทั้งหมดเป็น 100 คน ก็ได้ เรียกว่าตำแหน่งร้อยละ (Percentile Rank) วิธีนี้ให้ความสำคัญแก่คะแนนน้อยกว่าอันดับที่สอบได้

2.3 คะแนนมาตรฐาน ได้แก่ การนำคะแนนดิบที่ได้ ไปเทียบกับคะแนนเฉลี่ย ($\bar{X}$) ของกลุ่มโดยพิจารณาว่า มากกว่าหรือน้อยกว่าคะแนนเฉลี่ยของกลุ่มเท่าไรคะแนนมาตรฐานที่นิยมใช้ ได้แก่

* ผู้ช่วยศาสตราจารย์ ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ มหาสารคาม

คะแนนมาตรฐาน Z คะแนนมาตรฐาน T

2.4 คะแนนมาตรฐานเก่า ได้แก่การนำคะแนนดิบที่ได้ไปเทียบกับคะแนนในกลุ่มย่อย เมื่อแบ่งกลุ่มของผู้เข้าสอบทั้งหมดออกเป็น 9 กลุ่มย่อยตามอัตราส่วนที่กำหนดไว้

แนวความคิดเกี่ยวกับคะแนนดิบ

1. คะแนนดิบไม่สามารถบอกได้ว่าใครเก่งหรืออ่อนโดยตัวเองหรือไม่มีความหมายใด ๆ เช่น นายแดงได้ 20 คะแนน

2. คะแนนดิบจากแบบทดสอบต่างฉบับกัน จะนำมาเปรียบเทียบกันย่อมไม่ได้ เช่น นาย ก. สอบวิชาภาษาไทย 40 คะแนน จากคะแนนเต็ม 60 คะแนน และสอบวิชาสังคมได้ 30 คะแนน จากคะแนนเต็ม 60 คะแนน เช่นกัน จะสรุปเขาเก่งวิชาภาษาไทยมากกว่าวิชาสังคมย่อมไม่ได้

3. การนำคะแนนดิบของแต่ละวิชามารวมกันโดยตรงย่อมไม่ถูกต้อง เพราะค่าของคะแนนแต่ละวิชามีขนาดไม่เท่ากัน

โดยสรุปแล้วคะแนนดิบยังมีความบกพร่อง 2 ประการ คือ เป็นคะแนนที่ไม่สามารถแปลความหมายได้อย่างชัดเจนด้วยตัวเองว่า นักเรียนมีความเก่งอ่อนเพียงใดและคะแนนดิบเป็นคะแนนที่ไม่สามารถจะนำมาบวกหรือลบกันได้โดยตรง เหมือนตัวเลขในทางคณิตศาสตร์ ดังนั้นคะแนนดิบของการสอบแต่ละครั้งหรือแต่ละวิชาจึงไม่ควรนำมารวมกัน

แนวความคิดเกี่ยวกับคะแนนมาตรฐาน

คะแนนมาตรฐานเป็นคะแนนแปลงรูปจากคะแนนดิบ ที่มีความหมายคลุมเครือให้มีความหมายชัดเจนขึ้น ทั้งนี้โดยทำให้คะแนนมาตรฐานมีจุดเปรียบเทียบความเก่งอ่อนที่เชื่อถือได้และทำให้มีหน่วยเท่ากัน จะมีผลทำให้สามารถนำคะแนนมาตรฐานมาบวกลบกันได้

คะแนนมาตรฐาน (Standard Score)

ถ้าต้องการเปรียบเทียบความสามารถทางการเรียนหลาย ๆ วิชาของนักเรียนคนหนึ่ง หรือความสามารถทางการเรียนของบุคคลหลายคนจากการสอบหลาย ๆ วิชา จะพิจารณาจากคะแนนดิบเท่านั้นไม่ได้ เนื่องจากข้อสอบต่างวิชากันย่อมมีความยากง่ายต่างกัน คะแนนดิบจากวิชาต่างกันจึงมีขนาด (Scale) ไม่เท่ากัน นอกจากนั้นการกระจายของคะแนนก็ต่างกันด้วย การนำคะแนนดิบมารวมกันหรือเปรียบเทียบกันจึงไม่ถูกต้อง ถ้าจะนำคะแนนมารวมหรือเปรียบเทียบกันก็ควรแปลงคะแนนดิบเหล่านั้นให้เป็นคะแนนมาตรฐานเสียก่อน ซึ่งเป็นวิธีที่มีคุณภาพสูงวิธีหนึ่ง

คะแนนมาตรฐาน คือ คะแนนที่แปลงรูปมาจากคะแนนดิบเพื่อทำให้เป็นคะแนนที่มีจุดเปรียบเทียบและขนาด (Scale) ที่เท่ากัน โดยใช้คะแนนเฉลี่ยของกลุ่มมาเป็นจุดหลักในการเปรียบเทียบ และหน่วย

ที่ใช้ในการวัดก็คือ จำนวนเท่าของส่วนเบี่ยงเบนมาตรฐาน ตัวอย่างของคะแนนมาตรฐานได้แก่ คะแนนซี (Z Score) ซึ่งเป็นคะแนนมาตรฐานที่มีค่าเฉลี่ย ($\bar{X}$) เท่ากับ 0 และมีค่าความเบี่ยงเบนมาตรฐาน (S) เท่ากับ 1 ตัวอย่าง วรรณสอบวิชาภาษาอังกฤษ และวิชาวิทยาศาสตร์ ได้ 45 คะแนนเท่ากันอยากรู้ว่าเขาเก่ง 2 วิชาเท่ากันหรือไม่ ถ้าข้อมูลต่าง ๆ เป็นดังนี้

วิชา	คะแนนเดิม	คะแนนที่ได้	ค่าเฉลี่ย	ความเบี่ยงเบนมาตรฐาน
ภาษาอังกฤษ	60	45	38	3
วิทยาศาสตร์	60	45	48	6

ถ้าพิจารณาจากคะแนนดิบเท่านั้นแล้วสรุปว่าวรรณเก่งวิชาภาษาอังกฤษเท่ากับวิทยาศาสตร์ ย่อมไม่ถูกต้อง เนื่องจากคะแนนของ 2 วิชามีขนาดต่างกัน เช่นเดียวกับการที่จะเปรียบเทียบเงินสกุลต่าง ๆ จะต้องทำให้เป็นสกุลเดียวกัน ในทำนองเดียวกันการเปรียบเทียบผลสอบจะต้องแปลงคะแนนดิบให้เป็นคะแนนมาตรฐาน (คะแนนซี) เสียก่อนแล้วจึงจะเปรียบเทียบกันได้ โดยใช้สูตร ดังนี้

$$\text{คะแนนมาตรฐาน (Z)} = \frac{\text{คะแนนดิบ} - \text{คะแนนเฉลี่ย}}{\text{ความเบี่ยงเบนมาตรฐาน}} = \frac{X - \bar{X}}{S}$$

$$\text{วิชาภาษาอังกฤษ (Z)} = \frac{45 - 38}{3} = \frac{7}{3} = 2.33$$

$$\text{วิชาวิทยาศาสตร์ (Z)} = \frac{45 - 48}{6} = \frac{-3}{6} = -0.50$$

สรุปว่า วรรณเก่งภาษาอังกฤษมากกว่าวิทยาศาสตร์ (เพราะได้คะแนนมาตรฐานวิชาภาษาอังกฤษมากกว่า)

จากตัวอย่างจะเห็นว่า คะแนนซี (Z - Score) มีหน่วยที่ใหญ่จึงมีค่าเป็นทศนิยมและมีค่าติดลบ เพราะโดยธรรมชาติของคะแนนซี จะมีค่าอยู่ระหว่าง - 4.00 ถึง 4.00 ทำให้เกิดความยุ่งยากในการแปลความหมายให้คนอื่นเข้าใจ จากจุดอ่อนของคะแนนซีดังกล่าว นักวัดผลจึงเสนอให้ใช้คะแนนมาตรฐานอื่นแทน ที่นิยมใช้คือ คะแนนที (T - Score) ซึ่งมีสูตรดังนี้

$$T = 10Z + 50$$

(คะแนน T จะมีค่าเฉลี่ยเท่ากับ 50 และความเบี่ยงเบนมาตรฐานเท่ากับ 10)

คะแนนซี และคะแนนที ที่กล่าวถึงนี้เป็นคะแนนมาตรฐานที่มีลักษณะการแจกแจงเหมือนกับ การแจกแจงของคะแนนดิบทุกประการ กล่าวคือ ถ้าคะแนนดิบมีการแจกแจงแบบเบ้ (Skewed) การแจกแจงของคะแนนมาตรฐานก็จะเบ้ด้วย และเนื่องจากเป็นการแปลงรูปในลักษณะของสมการเส้นตรง วิธีการในทางการวัดผลไม่ค่อยนิยมใช้มากนัก เพราะถ้ามีการวัดหลาย ๆ วิธี หลาย ๆ กลุ่ม และการแจกแจงของแต่ละกลุ่มมีลักษณะแตกต่างกันมาก (เช่น เบ้ซ้าย เบ้ขวา) การเปรียบเทียบในแต่ละกลุ่มอาจเกิดความคลาดเคลื่อนมาก ดังนั้นนักวัดผลจึงนิยมแปลงคะแนนดิบให้เป็นคะแนนมาตรฐานที่มีการแจกแจงเป็นโค้งปกติ (Normal Curve)

เหมือนกันหมด เพื่อให้สามารถเปรียบเทียบกันได้ ซึ่งได้แก่คะแนนที่ - ปกติ (Normalized T - Score) มีวิธีแปลงดังนี้

ขั้นที่ 1 ทำการเรียงคะแนนที่สอบได้แล้วลงรอยขีด (Tally)

ขั้นที่ 2 หาค่า f และ cf

ขั้นที่ 3 หาค่า $cf + \frac{1}{2}f$

หมายเหตุ จะหาค่า $cf + \frac{1}{2}f$ ของชั้นใด ต้องใช้ค่า cf ของชั้นล่างถัดลงมา และใช้ค่า f ของชั้นนั้นเสมอ

ขั้นที่ 4 เอาค่า $cf + \frac{1}{2}f$ ไปคูณด้วย $\frac{100}{N}$ ค่าที่ได้เรียกว่าตำแหน่งเปอร์เซ็นต์ไทล์ (Percentile Rank = PR)

ขั้นที่ 5 นำค่า PR ในขั้นที่ 4 ไปเทียบค่า T ปกติ จากตารางสำเร็จรูปต่อไปนี้

การใช้ตารางอันดับที่เปอร์เซ็นต์ไทล์ของคะแนน T ปกติ

(Percentile Rank of A Normalized T - Score)

T	0	1	2	3	4	5	6	7	8	9
1	-	-	0.01	0.01	0.02	0.02	0.03	0.05	0.07	0.10
2	0.13	0.19	0.26	0.35	0.47	0.62	0.82	1.07	1.39	1.79
3	2.28	2.87	3.56	4.46	5.18	6.58	8.08	9.68	11.51	13.57
4	15.87	18.41	21.19	24.20	27.43	30.85	34.36	38.21	42.07	46.02
5	50.00	53.98	57.93	61.79	65.54	69.15	72.57	74.80	78.81	81.59
6	84.13	86.43	88.49	90.32	91.92	93.32	94.52	95.54	96.41	97.13
7	97.72	98.21	98.61	98.93	99.18	99.38	99.53	99.65	99.74	99.81
8	99.89	99.80	99.93	99.95	99.97	99.98	99.98	99.99	-	-

ค่าของคะแนน T ตามแนวดิ่ง (แถวซ้ายมือ) แสดงหลักสิบ และตามแนวนอน (แถวบน) แสดงหลักหน่วย

วิธีเทียบ

- 1.ให้นำค่าตำแหน่งเปอร์เซ็นต์ไทล์ที่คำนวณได้ มาเทียบกับค่าเปอร์เซ็นต์ไทล์ที่อยู่ในตารางนี้ซึ่งมีทศนิยม 2 ตำแหน่ง
2. หากค่าตำแหน่งเปอร์เซ็นต์ไทล์ที่คำนวณได้ไม่ตรงกับค่าใด ๆ ในตารางนี้ ให้เลือกเอาค่าในตารางนี้ที่ ใกล้มากที่สุด ไม่ว่าจะใกล้กับค่าที่น้อยกว่าหรือมากกว่าก็ตาม
- 3.ให้อ่านคะแนน T กลับกลับ จากแนวดิ่ง (แถวซ้ายมือ) และรวมกับหลักหน่วยจากแนวนอน (แถวบน) เช่น ถ้าตำแหน่งเปอร์เซ็นต์ไทล์มีค่า 91.92 ก็จะได้คะแนน $T = 64$ หรือตำแหน่งเปอร์เซ็นต์ไทล์มีค่า 37.50 จะได้คะแนน $T = 47$ เป็นต้น

ตัวอย่าง สมมตินักเรียนเข้าสอบ 25 คน ได้คะแนนสูงสุด 23 คะแนน ต่ำสุด 12 คะแนน ส่วนคะแนนอื่น ๆ แตกต่างกันไป สามารถแปลงเป็นคะแนน T ปกติ ได้ดังนี้

คะแนน	Tally	f	cf	$cf+1/2f$	$(cf+1/2f) \frac{100}{N}$	T ปกติ
23	/	1	25	24.5	98	71
22	//	2	24	23	92	64
21	//	2	22	21	84	60
20	///	3	20	18.5	74	57
19	////	4	17	15	60	53
18	//	2	13	12	48	49
17	////	4	11	9	36	47
16	//	2	7	6	24	43
15	-	0	5	5	20	42
14	//	2	5	4	16	40
13	//	2	3	2	8	36
12	/	1	1	0.5	2	29

การให้เกรดในระบบอิงเกณฑ์ - อิงกลุ่ม

ในการประเมินผลในเรื่องใด ๆ ก็ตาม จะมีความหมายก็ต่อเมื่อนำค่าหรือคะแนนที่วัดได้นั้นไปเปรียบเทียบกับสิ่งหนึ่ง ที่เรายอมรับกันในแต่ละยุคแต่ละสมัย ในปัจจุบันเรานิยมนำไปเปรียบเทียบกับเกณฑ์หรือเปรียบเทียบกับกลุ่มผู้เรียน ดังนี้

1. การเปรียบเทียบกับเกณฑ์ เรียกว่าเป็นการให้เกรดในระบบอิงเกณฑ์ (Criterion Referenced) วิธีการนี้จะนำคะแนนที่ผู้เรียนทำได้ไปเปรียบเทียบกับเกณฑ์ หรือจุดประสงค์ที่ตั้งไว้ การประเมินแบบนี้ส่วนมากจะออกมาในรูปผ่านหรือไม่ผ่าน รู้หรือไม่รู้ หรือออกมาในรูปของเกรด A หรือ B หรือ C หรือ D หรือ E แต่เกรดที่ได้จากการประเมินแบบนี้ จะบอกได้ว่าคนที่ได้เกรด A สามารถทำอะไรได้บ้าง เช่น สามารถผ่านจุดประสงค์หลัก จุดประสงค์รองทุกข้อ หรือสามารถทำคะแนนได้ถึงร้อยละ 90 ขึ้นไป

2. การเปรียบเทียบกับกลุ่ม เรียกว่า เป็นการให้เกรดในระบบอิงกลุ่ม (Norm Referenced) วิธีการนี้จะนำคะแนนที่สอบได้ของผู้เรียนแต่ละคนไปเปรียบเทียบกับคะแนนของกลุ่มในวิชาเดียวกันเสียก่อน แล้วจึงตัดสินให้เกรด ดังนั้นการประเมินแบบนี้ ส่วนมากจะรายงานออกมาในรูปอันดับที่ (Rank) ตำแหน่ง (Percentile) หรือออกมาในรูปเกรด A หรือ B หรือ C หรือ D หรือ F แต่เกรดในการประเมินแบบนี้ จะบอกได้ในลักษณะที่ว่าเขาเก่งกว่าคนอื่นเท่าไร ไม่สามารถบอกว่า เกรด A นั้นเขามีความรู้ในเรื่องใดบ้าง จำนวนเท่าใด

การตัดเกรด

การตัดเกรดเป็นการสรุปผลการเรียนขั้นสุดท้ายเพื่อประเมินและกำหนดระดับความสามารถในการเรียนของนักเรียนว่า ผ่าน - ไม่ผ่าน หรือ เก่ง - อ่อน ระดับใด การตัดเกรดจึงเป็นการประเมินผลจากการสอบการวัดในวิชานั้น ๆ เพื่อสรุปออกมาเป็นระดับคะแนน (Grade) ซึ่งผู้สอนจะต้องพิจารณาอย่างรอบคอบเพราะการให้เกรดมีผลกระทบต่ออนาคตของผู้เรียนโดยตรง ความถูกต้องและเหมาะสมของการให้เกรดขึ้นอยู่กับองค์ประกอบ 3 ประการ คือ

1. ผลการวัด (Measurement) การวัดที่ดีจะต้องให้ผลการวัดที่ถูกต้องแม่นยำ เทียบตรงครอบคลุมและเชื่อถือได้

2. เกณฑ์การพิจารณา (Criteria) ต้องเป็นมาตรฐานที่ใช้เป็นหลักเปรียบเทียบหรือเป็นคุณลักษณะที่ตั้งไว้เป็นเป้าหมาย หรือมุ่งหวังที่จะให้เกิดแก่ผู้เรียนและใช้เป็นเครื่องตัดสินชี้ขาดระดับความสามารถของผู้เรียน

3. วิจารณ์และคุณธรรมต่าง ๆ (Value Judgement) เนื่องจากผลการวัดที่ได้เป็นเพียงข้อมูลส่วนหนึ่งเกี่ยวกับตัวเด็กเท่านั้น การประเมินผลที่เที่ยงตรงจำเป็นต้องอาศัยดุลยพินิจหรือการพิจารณาอย่างรอบคอบถี่ถ้วนของผู้สอนประกอบด้วยโดยพยายามให้เป็นธรรม ขจัดความลำเอียงหรืออคติส่วนตัวออกไปและควรคำนึงถึงความเปลี่ยนแปลงของงานของเด็กในด้านอื่น ๆ ประกอบด้วย

แบบของการตัดเกรด ในที่นี้แบ่งออกเป็น 10 แบบดังนี้

แบบที่ 1 ตัดสินให้ผ่าน - ไม่ผ่าน

หลักการสำคัญคือ กำหนดเกณฑ์ขั้นต่ำในรูปของคะแนนดิบหรือเปอร์เซ็นต์ ผลที่ได้จากการประเมินในระบบนี้จะมีแค่ 1 ค่าเท่านั้น ซึ่งการให้เกรดวิธีนี้ใช้มากในการประเมินแบบอิงเกณฑ์ โดยผู้สอน

ต้องกำหนดเกณฑ์ขั้นต่ำ (Minimum Requirement) เอาไว้ว่า ผู้สอนต้องมีความสามารถ (ตามจุดมุ่งหมาย) อะไรบ้าง จากนั้นจึงทำการตรวจสอบว่าใครผ่านหรือไม่ผ่านตามเกณฑ์ที่กำหนดไว้

การตัดสินใจให้เกรดเป็น ดี - ผ่าน - ตก ก็อนุมานให้อยู่ในลักษณะนี้

แบบที่ 2 จัดกลุ่มตามธรรมชาติ

หลักการสำคัญคือ ถ้ามีช่วงห่างระหว่างคะแนนดิบเป็นจุดแบ่งระดับคะแนน และไม่สามารถกำหนดจำนวนระดับคะแนนล่วงหน้า การใช้วิธีนี้จะต้องมีช่วงห่างระหว่างกลุ่มของคะแนนชัดเจนพอที่จะกำหนดเป็นจุดตัดของแต่ละเกรดได้ หรือถ้าไม่มีช่วงห่างของคะแนนก็พิจารณาจากความถี่ที่ต่างกันอย่างชัดเจน เกรดแต่ละเกรดควรจะมีความแตกต่างกันอย่างน้อย 1 ความคลาดเคลื่อนมาตรฐาน (Standard error of measurement) ทั้งนี้เพื่อขจัดความผิดพลาดในการให้เกรดอันเนื่องมาจากความคลาดเคลื่อนของคะแนนที่ได้จากการสอบวัด ซึ่งการคำนวณหาความคลาดเคลื่อนมาตรฐานทำได้โดยใช้สูตร

$$SE_{meas} = \sqrt{1 - r_{tt}}$$

เมื่อ SE_{meas} คือ ค่าความคลาดเคลื่อนมาตรฐาน (Standard error of measurement)

S คือ ค่าความเบี่ยงเบนมาตรฐาน (Standard deviation)

r_{tt} คือ ค่าความเชื่อมั่น (Reliability) ของคะแนนชุดนั้น

ตัวอย่าง คะแนนดิบจากการสอบวิชาหนึ่ง มีการแจกแจงความถี่ดังนี้

คะแนนดิบ	ความถี่	
87	1	
85	1	
84	2	
83	1	กลุ่มที่ 1
80	2	
79	1	กลุ่มที่ 2
78	3	
77	1	
74	4	
73	5	
72	9	

71	10	
70	7	
69	6	กลุ่มที่ 3
68	5	
67	2	
66	1	
65	1	
60	2	
59	1	กลุ่มที่ 4
58	-	
57	-	
56	1	

จากตารางการแจกแจงความถี่ข้างต้นจะเห็นว่า มีช่วงห่างของกลุ่มคะแนนอยู่ 4 ช่วง ซึ่งอาจจะให้คะแนนในกลุ่มแรกได้เกรด A กลุ่มที่ 2 ได้เกรด B และ C,D ตามลำดับ หรืออาจจะตัด 3 เกรด โดยรวมคะแนนกลุ่มที่ 2 และ 3 เข้าด้วยกัน สิ่งนี้จะต้องพิจารณาเหตุผลอื่น ๆ ประกอบด้วย เช่น ความยากง่ายของข้อสอบ เป็นต้น

หรือถ้าจะพิจารณาจากความคลาดเคลื่อนมาตรฐานเป็นเกณฑ์จะได้ดังนี้

$$\text{หาค่า } S \text{ ของคะแนนชุดนี้ ตามสูตร } S = \sqrt{\frac{N \sum X^2 - (\sum X)^2}{N(N-1)}}$$

ได้ค่า $S = 5.75$

$$\text{หาค่า } r_{tt} \text{ (ความเชื่อมั่นของคะแนน) ตามสูตร } r_{tt} = \frac{n}{n-1} \left\{ 1 - \frac{\bar{X}(n-\bar{X})}{ns^2} \right\}$$

$$\begin{aligned} \text{สมมติได้ค่า } r_{tt} &= .76 \\ \text{แทนค่าในสูตร } SE_{meas} &= S \sqrt{1 - r_{tt}} \\ &= 5.75 \sqrt{1 - .76} \\ &= 5.75 \sqrt{.24} \\ &= 5.75 (\pm .49) \\ &= \pm 2.82 \\ &= \pm 3 \end{aligned}$$

หมายความว่า แต่ละช่วงเกรดคะแนนควรจะต่างกันอย่างน้อย ± 3

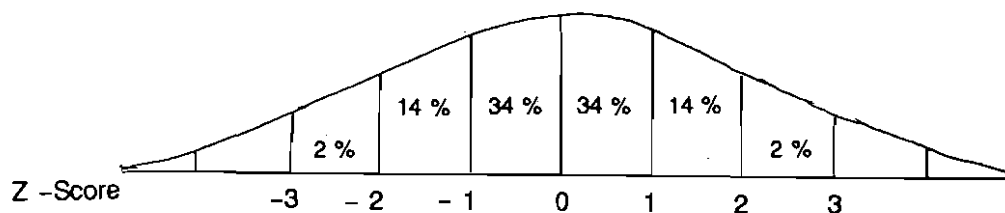
ในกรณีที่ไม่มีช่วงห่างของคะแนน การตัดเกรดโดยวิธีนี้คงจะทำได้ยาก แต่ถ้าคะแนนชุดนั้นมีความถี่แตกต่างกันอย่างชัดเจน ผู้ประเมินอาจจะพิจารณาจุดแบ่งเกรดตรงจุดที่มีความแตกต่างของความถี่อย่างชัดเจน

แบบที่ 3 ใช้เกณฑ์ที่คาดหวังหรือแบบตั้งเกณฑ์ไว้ตายตัว

หลักการสำคัญ คือ กำหนดเกณฑ์ในรูปของคะแนนดิบ หรือเปอร์เซ็นต์ขึ้นมาก่อน เช่น ได้ 90 % ขึ้นไป ได้เกรด A ได้ 70 - 89 % ได้เกรด B เป็นต้น ระบบนี้มีจุดอ่อนตรงที่ว่าเปอร์เซ็นต์หรือคะแนนของผู้เรียนขึ้นอยู่กับความยากง่ายของข้อสอบ โดยผู้สอบมิได้ใช้วิจารณญาณของตนเองเกี่ยวข้องด้วย เช่น ถ้าผู้สอนสังเกตพบว่านักเรียนกลุ่มนี้ค่อนข้างเก่งและมีผลการเรียนอยู่ในขั้นที่น่าพอใจมาก แต่ข้อสอบที่ให้นักเรียนทำนั้นยากเกินไป เด็กกลุ่มนี้ก็อาจจะได้ F หลายคน ซึ่งไม่สอดคล้องกับการประเมินของผู้สอน

แบบที่ 4 กำหนดเปอร์เซ็นต์ตามการแจกแจงของโค้งปกติ

หลักการสำคัญ คือ เมื่อจำนวนผู้เข้าสอบมาก คะแนนมักจะกระจายอยู่ในรูปของโค้งปกติและผู้สอนกำหนดระดับคะแนนขึ้นมาก่อน เนื่องจากโค้งปกติได้แบ่งพื้นที่ใต้โค้งออกเป็นอัตราส่วนดังนี้



ดังนั้นครูต้องกำหนดระดับ คะแนนขึ้นมาก่อน เช่น ต้องการแบ่งเป็น 5 ระดับ อาจจะแบ่งดังนี้

ระดับคะแนน	4	มี 7 %	มีจำนวน	6 คน
ระดับคะแนน	3	มี 24 %	มีจำนวน	19 คน
ระดับคะแนน	2	มี 38 %	มีจำนวน	30 คน
ระดับคะแนน	1	มี 24 %	มีจำนวน	19 คน
ระดับคะแนน	0	มี 7 %	มีจำนวน	6 คน

การกำหนดเปอร์เซ็นต์ในแต่ละระดับขึ้นอยู่กับดุลยพินิจของครูผู้สอน เช่น เคทเทิล (Cattle 1950) ได้กำหนดระดับคะแนนเป็นดังนี้

ระดับคะแนน	4	มี 10 %	ระดับคะแนน	1	มี 20 %
ระดับคะแนน	3	มี 20 %	ระดับคะแนน	0	มี 10 %
ระดับคะแนน	2	มี 40 %			

แบบที่ 5 วิธีของดักลาส (Douglas)

หลักการสำคัญ คือ จำนวนระดับคะแนน จะเป็นเท่าไร ขึ้นอยู่กับอัตราส่วนระหว่างคะแนนต่ำสุดกับคะแนนสูงสุด ดังตัวอย่างข้างล่าง มีผู้เข้าสอบ 90 คน

คะแนน	ความถี่	คะแนน	ความถี่	คะแนน	ความถี่
58	1	40	2	28	3
56	1	39	2	27	2
55	2	38	1	26	4
51	1	37	5	25	6
50	1	36	5	24	3
49	2	35	2	23	1
47	3	34	7	22	1
46	1	33	3	21	2
45	3	32	3	20	1
44	2	31	5	18	1
42	6	30	1	17	2
41	2	29	3		

- วิธีทำ 1. เมื่อตรวจให้คะแนนเรียบร้อย จึงเรียงคะแนนจากสูงสุดจนถึงต่ำสุด
2. หาความถี่ของแต่ละคะแนน
3. หาพิสัย = คะแนนสูงสุด - คะแนนต่ำสุด
- $$\text{พิสัย} = 58 - 17 = 41 \text{ คะแนน}$$
4. หาคะแนนต่ำสุดด้วยคะแนนสูงสุด ผลหารที่ได้จะเป็นทศนิยม ซึ่งนำไปเปรียบเทียบกับตารางเทียบจำนวนระดับคะแนน ข้างล่างว่า จะมีกี่ระดับคะแนน ดังนี้
- $$\frac{\text{คะแนนต่ำสุด}}{\text{คะแนนสูงสุด}} = \frac{17}{58} = 0.29$$

$0.29 < 0.49$ จึงมี 5 ระดับ คือ 4,3,2,1,0

5. จากตารางเทียบจำนวนระดับคะแนนในหน้าถัดไป ถ้ามี 5 ระดับคะแนน ตัวเลขที่ไปหารพิสัยเพื่อทราบช่วงแต่ละระดับคะแนนเป็น 6 นำ 6 ไปหารพิสัย $= \frac{41}{6} = 6.8 = 7$ ดังนั้นระดับคะแนนแต่ละระดับจะห่างกัน 7 คะแนน ยกเว้นระดับคะแนนที่ 2 ซึ่งอยู่ตรงกลางและมีคนมากกว่าระดับคะแนนอื่น จึงให้มีช่วงห่าง 14 คะแนน (เพราะการสอบ กันนักเรียน จำนวนมากคะแนนมักจะอยู่ในรูปโค้งปกติ)

ฉะนั้น จำนวนคนในแต่ละระดับคะแนน เป็นดังนี้

ระดับคะแนน 4 คะแนนระหว่าง 52 - 58 คะแนน มีจำนวน 4 คน
 ระดับคะแนน 3 คะแนนระหว่าง 45 - 51 คะแนน มีจำนวน 11 คน
 ระดับคะแนน 2 คะแนนระหว่าง 31 - 44 คะแนน มีจำนวน 45 คน
 ระดับคะแนน 1 คะแนนระหว่าง 24 - 30 คะแนน มีจำนวน 22 คน
 ระดับคะแนน 0 คะแนนระหว่าง 17 - 23 คะแนน มีจำนวน 8 คน

ตารางเทียบจำนวนระดับคะแนน

ผลหารระหว่างคะแนน ต่ำสุดกับสูงสุด	จำนวนระดับคะแนน	ตัวเลขที่ไปหารพิสัย เพื่อทราบช่วง
0.95 ขึ้นไป	1	2
0.90 - 0.94	2	3
0.70 - 0.89	3	4
0.50 - 0.69	4	5
0.49 ลงมา	5	6

แบบที่ 8 วิธีของสตุต (Dewey B. Stuit)

หลักการสำคัญ คือ ต้องพิจารณาผู้เรียนทั้งหมดว่ามีความสามารถอยู่ในกลุ่มใด (จาก 7 กลุ่ม) และต้องอาศัยค่ามัธยฐานกับค่าความเบี่ยงเบนมาตรฐาน วิธีการนี้เป็นการแก้ปัญหาคำถามการตัดเกรดซึ่งเนื่องมาจากความสามารถของผู้เรียนแต่ละกลุ่มภาคเรียนไม่ตัดเทียมกัน แต่ผู้สอนกำหนดเกณฑ์การให้ระดับคะแนนไว้ล่วงหน้าหรือเหมือนกันในทุกภาคเรียน ถ้าเป็นเช่นนี้ กลุ่มผู้เรียนที่มีความสามารถสูง และมีจำนวนมากก็อาจจะมีโอกาสได้ระดับคะแนนสูง ๆ จำนวนน้อย หรือ พิจารณาในแง่ตรงกันข้าม กลุ่มที่มีความสามารถต่ำ ก็มีโอกาได้คะแนนสูง ๆ จำนวนมาก ดังนั้น เพื่อแก้ปัญหาดังกล่าว สตุต จึงเสนอว่าขั้นแรกต้องประเมินกลุ่มผู้เรียนก่อนว่า มีความสามารถโดยเฉลี่ยอยู่ในระดับใด แล้วให้เกรดตามความสามารถของกลุ่ม สิ่งที่จะบอกความสามารถของกลุ่มผู้เรียนได้ดีและสะดวกที่สุดคือค่าเฉลี่ยของคะแนนเฉลี่ยสะสม (Grade point average) ของกลุ่ม แต่ถ้าไม่เคยเรียนมาเลย ก็ต้องรอไประยะหนึ่งจนผู้สอนประเมินความสามารถของกลุ่มได้แล้วก็ไปเทียบกับตารางสำเร็จรูปเพื่อหาตัวคงที่หรือแฟคเตอร์ที่จะนำไปใช้ในการคำนวณเพื่อหาจุดต่ำสุดของเกรด A ดังนี้

ตารางพิจารณาระดับความสามารถ

ระดับความสามารถ	ค่าเฉลี่ยของคะแนน เฉลี่ยสะสม	จุดต่ำสุดของเกรดสูงสุดที่จะให้ (Lower Limit Factor)
ดีเลิศ	2.8	0.7
ดีมาก	2.6	0.9
ดี	2.4	1.1
ดีพอใช้	2.2	1.3
ปานกลาง	2.0	1.5
อ่อน	1.8	1.7
อ่อนมาก	1.6	1.9

ขั้นตอนการตัดเกรดแบบนี้

1. ประเมินคร่าว ๆ ว่า กลุ่มผู้เรียนมีความสามารถระดับใด (ดีเลิศ ดีมาก ฯลฯ ถ้าประเมินไม่ได้ให้ใช้ G.P.A เฉลี่ย เช่น ได้มากกว่า 2.60 ก็เป็นกลุ่มดีเลิศ เป็นต้น)
2. คำนวณหาค่ามัธยฐาน (Median) ของคะแนนชุดนั้น
3. คำนวณหาค่าความเบี่ยงเบนมาตรฐานของคะแนน จากสูตร

$$S = \sqrt{\frac{N \sum x^2 - (\sum x)^2}{N(N-1)}}$$

หรืออาจใช้สูตรวิธีลัด ดังนี้

$$S = \frac{\text{ผลรวมคะแนนสูงสุด } \frac{N}{6} \text{ คะแนน} - \text{ผลรวมคะแนนต่ำสุด } \frac{N}{6} \text{ คะแนน}}{N/2}$$

เมื่อ N คือ จำนวนคนทั้งหมดในกลุ่ม

4. คำนวณจุดต่ำสุดของเกรด A โดยนำค่าแฟคเตอร์ที่ได้จากตารางสำเร็จรูปไปคูณกับค่าความเบี่ยงเบนมาตรฐาน แล้วนำผลคูณที่ได้ไปบวกกับค่ามัธยฐาน ผลลัพธ์ที่ได้จะเป็นคะแนนต่ำสุดของเกรด
คะแนนต่ำสุดของเกรด B จะหาได้จากนำค่าความเบี่ยงเบนมาตรฐานไปลบออกจาก
คะแนนต่ำสุดของเกรด A สำหรับคะแนนต่ำสุดของเกรด C และ D ก็ดำเนินการเช่นเดียวกัน

ตัวอย่าง ผลการสอบวิชาเคมี นักศึกษาจำนวน 50 คน ซึ่ง G.P.A.เฉลี่ยของกลุ่มเท่ากับ 2.35 มีการแจกแจงความถี่ของคะแนนดังนี้

คะแนนดิบ	จำนวนคน	
33	1	
32	2	A = 3 คน
30	5	
29	5	
27	8	B = 18 คน
25	9	
24	7	
22	5	C = 21 คน
16	5	D = 5 คน
15	2	
12	1	F = 3 คน

ถ้าจะคัดเกรดโดยใช้วิธีนี้

ขั้นที่ 1 เทียบ G.P.A ของกลุ่ม (เท่ากับ 2.35) กับตารางสำเร็จรูป ได้ค่าแฟคเตอร์เท่ากับ 1.1

ขั้นที่ 2 หาค่ามัธยฐานของคะแนนชุดนี้ = 25

ขั้นที่ 3 คำนวณค่าเบี่ยงเบนมาตรฐาน (วิธีลัด)

$$S = \frac{(33+32+32+---+30) - (12+15+15+---+16)}{50/2}$$

$$= \frac{247 - 122}{25}$$

$$= 5$$

ขั้นที่ 4

คำนวณจุดต่ำสุดของเกรด A และเกรดอื่น ๆ

$$\text{คะแนนต่ำสุดของเกรด A} = (1.1 \times 5) + 25 = 30.50 \approx 31$$

$$\text{คะแนนต่ำสุดของเกรด B} = 31 - 5 = 26$$

$$\text{คะแนนต่ำสุดของเกรด C} = 26 - 5 = 21$$

$$\text{คะแนนต่ำสุดของเกรด D} = 21 - 5 = 16$$

คะแนนต่ำสุดของเกรด F คือ ตั้งแต่ 15 ลงมา

หมายเหตุ บางเกรด อาจจะไม่มีคนได้ และการประเมินแต่ละครั้ง ไม่จำเป็นต้องให้ F เสมอไป ทั้งนี้ขึ้นอยู่กับดุลยพินิจของผู้ประเมิน

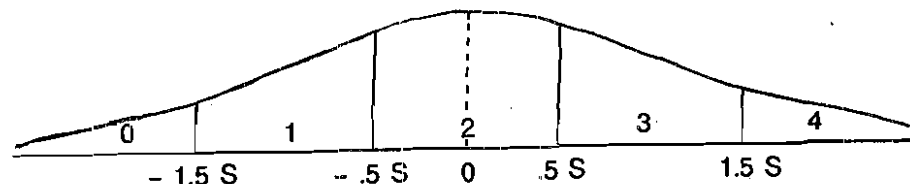
แบบที่ 7 ใช้ค่ามัธยฐาน (Median)

หลักการสำคัญ คือ อาศัยค่าของมัธยฐานกับค่าความเบี่ยงเบนมาตรฐาน ไม่ต้องคำนึงถึงความสามารถของกลุ่มผู้เรียน และการแจกแจงของข้อมูลว่าเป็นรูปโค้งปกติหรือไม่ (วิธีนี้คล้ายกับวิธีของสตุท) มีวิธีการดังนี้

1. ตรวจให้คะแนนแล้วเรียงคะแนนจากมากไปหาน้อย
2. หาคะแนน Median นับจากหัวท้ายของคะแนนจนถึงกึ่งกลาง ถ้าคะแนนเป็นจำนวนคู่ Median เท่ากับผลรวมของคะแนนคู่กลางหารด้วย 2
3. หา SD โดยประมาณ

$$S = \frac{\text{ผลรวมคะแนนสูงสุด } \frac{N}{2} \text{ คะแนน} - \text{ผลรวมคะแนนต่ำสุด } \frac{N}{2} \text{ คะแนน}}{N/2}$$

4. หาขอบเขตของแต่ละระดับคะแนน ดังนี้



- ระดับคะแนน 4 ตั้งแต่ 1.5 S ขึ้นไป
 ระดับคะแนน 3 ระหว่าง .5 S ถึง 1.5 S
 ระดับคะแนน 2 ระหว่าง - .5 S ถึง .5 S
 ระดับคะแนน 1 ระหว่าง - 1.5 S ถึง - .5 S
 ระดับคะแนน 0 ตั้งแต่ - 1.5 S ลงมา

ตัวอย่าง คะแนนวิชาสังคมศึกษา ส. ๐๓๕ นักเรียนสอบได้คะแนน ดังนี้

คะแนน	ความถี่	คะแนน	ความถี่
33	1	21	8
32	3	20	10
31	2	19	10
30	7	18	9
29	8	17	5
28	12	16	4

คะแนน	ความถี่	คะแนน	ความถี่
27	10	15	0
26	13	14	3
25	11	13	1
24	10	12	1
23	9	n =	150
22	13		

1. Median = 24

2. $\frac{N}{6} = \frac{150}{6} = \frac{25}{6}$ คน

3. หา S (โดยประมาณ)

$$S = \frac{(33+96+62+210+232+112) - (12+13+42+64+85+162+38)}{150/2}$$

$$= \frac{745 - 416}{75} = \frac{329}{75}$$

$$= 4.38$$

๕

4. เมื่อ S = 4.38

ค่า .5 S = .5 x 4.38 = 2.19

ค่า 1.5 S = 1.5 x 4.38 = 6.57

ระดับคะแนน 4 = Median + 1.5 S = 24 + 6.57 = 30.57 ตั้งแต่ 30 คะแนนขึ้นไปจำนวน 13 คน

ระดับคะแนน 3 = Median + .5 S = 24 + 2.19 = 26.19 ระหว่าง 26 ถึง 29 คะแนน

จำนวน 43 คน

ระดับคะแนน 2 = Median - .5 S = 24 - 2.19 = 21.81 ระหว่าง 21 ถึง 25 คะแนน

จำนวน 51 คน

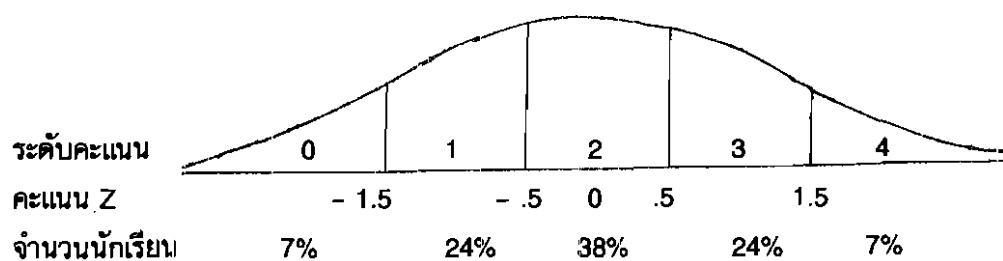
ระดับคะแนน 1 = Median - 1.5 S = 24 - 6.57 = 17.43 ระหว่าง 17 ถึง 20 คะแนน

จำนวน 34 คน

ระดับคะแนน 0 ตั้งแต่ 16 คะแนนลงมา จำนวน 9 คน

แบบที่ 8 ใช้คะแนนมาตรฐาน (Z-Score)

หลักการสำคัญ คือ จุดแบ่งระดับคะแนน (Z) ไม่มีกฎเกณฑ์ตายตัว ขึ้นอยู่กับผู้ประเมินจะมีเกณฑ์อย่างไร มีผู้เสนอเกณฑ์การแบ่งระดับคะแนนออกหลายวิธี เช่น



ตัวอย่าง มีคนเข้าสอบ 127 คน เป็นดังนี้				
คะแนน	ความถี่	คะแนน	ความถี่	
25	2	13	16	
24	1	12	10	
23	1	11	4	
22	4	10	11	
21	5	9	8	
20	5	8	5	
19	6	7	6	
18	1	6	3	
17	6	5	2	
16	13	4	1	
15	7	n = 127		
14	10			

วิธีทำ

1. หา $\sum X$ (คือผลบวกของคะแนนทุกคน) = 1745

2. $\bar{X} = 1745/127 = 13.74$

3. หา S (โดยประมาณ)

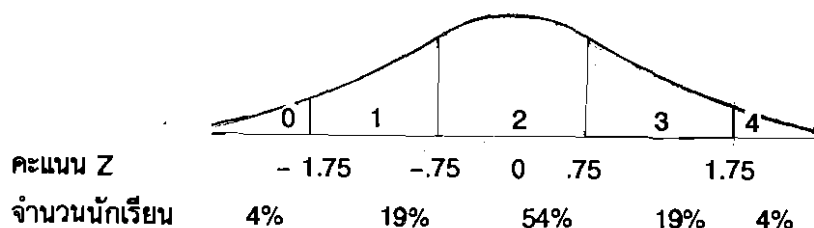
$$S = \frac{\text{ผลรวมคะแนนสูงสุด } \frac{N}{6} \text{ คะแนน} - \text{ผลรวมคะแนนต่ำสุด } \frac{N}{6} \text{ คะแนน}}{N/2}$$

$$= \frac{(50+24+23+88+105+100+57) - (4+10+18+42+40+36)}{127/2}$$

$$= \frac{447 - 150}{63.5} = 4.6$$

4. หา Z - Score ของแต่ละคน เช่น นาย ก. ได้ 15 คะแนน, $Z = \frac{X - \bar{X}}{S}$ ดังนั้น คะแนน Z ของนาย ก. $= \frac{15 - 13.74}{4.6} = .27$

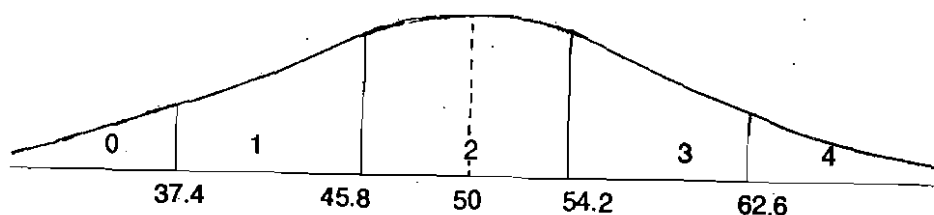
เพราะฉะนั้น นาย ก. ได้ระดับคะแนน เป็น 2
หมายเหตุ ยังมีผู้เสนอการตั้งเกณฑ์เป็นแบบอื่น ๆ เช่น



แบบที่ 9 ใช้ คะแนนมาตรฐานที่ - ปกติ (Normalized T - Score)

หลักการสำคัญ คือ คะแนนกระจายอยู่ในรูปของโค้งปกติ (Normal Curve) และต้องการเปรียบเทียบความสามารถในกลุ่ม ส่วนจำนวนระดับคะแนน ขึ้นอยู่กับดุลยพินิจของผู้ประเมิน ขั้นตอนในการให้ระดับคะแนน มีดังนี้

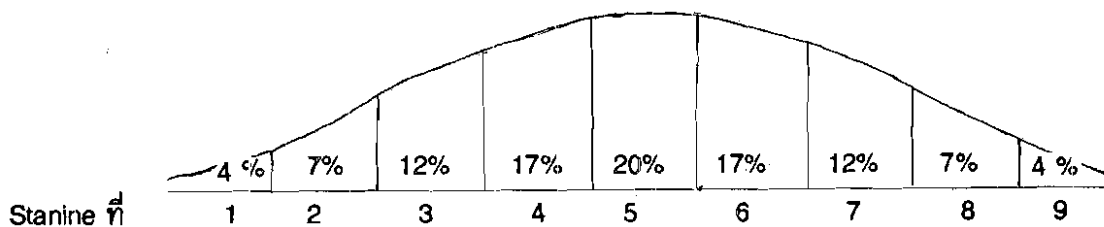
1. หาพิสัยของคะแนนที่ เช่น จากตัวอย่างในหน้า 57 จะได้ค่าพิสัยเป็น 42 (เพราะ 71-29)
2. เอาจำนวนระดับคะแนนที่ต้องการไปหารค่าพิสัย ผลลัพธ์ที่ได้คืออันตรภาคชั้นของคะแนนที่ เช่น ต้องการ 5 ระดับ ดังนั้น $42 \div 5 = 8.4$ และ $8.4 \div 2 = 4.2$ (ถ้าต้องการระดับคะแนนเป็นเลขคู่ ไม่เอา 2 หารค่าอันตรภาคชั้น)



ระดับคะแนน 4	คะแนนที่	ตั้งแต่ 63 ขึ้นไป	จำนวน	3 คน
ระดับคะแนน 3	คะแนนที่	ตั้งแต่ 55 - 62	จำนวน	5 คน
ระดับคะแนน 2	คะแนนที่	ตั้งแต่ 46 - 54	จำนวน	10 คน
ระดับคะแนน 1	คะแนนที่	ตั้งแต่ 38 - 45	จำนวน	4 คน
ระดับคะแนน 0	คะแนนที่	ตั้งแต่ 37 ลงมา	จำนวน	3 คน

แบบที่ 10 ใช้คะแนนมาตรฐานเก้า (Stanine Score)

หลักการสำคัญ คือ แบ่งคะแนนของพื้นที่ได้โค้งปกติออกเป็นอัตราส่วน เช่นเดียวกับการแปลงคะแนนดิบให้เป็นคะแนนที่ - ปกติ แต่แบ่งออกเป็น 9 กลุ่ม โดยแต่ละกลุ่มกำหนดอัตราส่วนเป็นลักษณะตายตัวโดยกำหนดให้กลุ่มที่มีคะแนนต่ำสุดเป็น Stanine ที่ 1 เรียงลำดับไปเรื่อยจนถึงกลุ่มที่มีคะแนนสูงสุดซึ่งเป็น Stanine ที่ 9 ดังนี้



ลำดับขั้นในการแปลงคะแนนดิบเป็นคะแนน Stanine ดังนี้

ขั้นที่ 1 บรรจุนับจำนวนผู้สอบในแต่ละ Stanine โดยคำนวณตามทฤษฎี จากสูตร ดังนี้

$$\text{จำนวนในแต่ละกลุ่ม} = \% \text{ ของพื้นที่ใต้โค้ง} \times \frac{N}{100}$$

ขั้นที่ 2 สร้างตารางแจกแจงความถี่ของคะแนนดิบ

ขั้นที่ 3 แบ่งจำนวนผู้สอบในตารางแจกแจงความถี่ออกเป็น 9 กลุ่ม โดยยึดหลักการดังนี้

3.1 ยึดคะแนนที่อยู่ตรงกลางของกลุ่ม (ค่ามัธยฐาน : Median) เป็นจุดเริ่มต้นในการแบ่ง ซึ่งจะตรงกับ Stanine ที่ 5 พอดี จากนั้นจึงแบ่งจำนวนคนในช่วงที่สูงและต่ำกว่าตามลำดับ

3.2 คะแนนที่เท่ากันต้องอยู่ใน Stanine เดียวกัน ในการแบ่งกลุ่ม หากมีคะแนนบางตัวซ้ำกัน และคะแนนที่ซ้ำกันนี้อยู่ระหว่างรอยต่อของคะแนน Stanine 2 ค่า ซึ่งมีผลทำให้คะแนนที่เท่ากันถูกแบ่งออกไปอยู่ใน Stanine ต่างกัน ต้องมีการปรับแต่งให้คะแนนที่เท่ากันนั้นอยู่ใน Stanine เดียวกัน โดยปรับให้อยู่ใน Stanine ที่ใกล้กับค่ามัธยฐาน (Stanine ที่ 5) เช่น ถ้าเป็นคะแนนที่คาบเกี่ยวระหว่าง Stanine ที่ 7 และ 8 ก็ควรปรับให้เป็น Stanine ที่ 7

การวัดผลการศึกษา 71

ตัวอย่าง นักเรียน 50 คน สอบวิชาเคมี ได้คะแนนดังตารางแจกแจงความถี่ ดังนี้

คะแนนดิบ	จำนวนคน	คะแนน Stanine ที่	จำนวนจริง	จำนวนคนในแต่ละ Stanine จากการ คำนวณตามทฤษฎี
33	1	9	3	(2)
32	2			
30	5	8	5	(3.5)
29	5	7	5	(6)
27	8	6	8	(8.5)
25	9	5	9	(10)
24	7	4	7	(8.5)
22	5	3	5	(6)
16	5	2	5	(3.5)
15	2	1	3	(2)
12	1			

การแบ่งระดับคะแนน อาจจะเป็นดังนี้

ระดับคะแนน เกิดจาก Stanine 9	มีจำนวน 3 คน
ระดับคะแนน เกิดจาก Stanine 8 และ 7	มีจำนวน 10 คน
ระดับคะแนน เกิดจาก Stanine ที่ 6,5 และ 4	มีจำนวน 24 คน
ระดับคะแนน เกิดจาก Stanine ที่ 3 และ 2	มีจำนวน 10 คน
ระดับคะแนน เกิดจาก Stanine ที่ 1	มีจำนวน 3 คน

การรวมคะแนน

ในการตัดสินผลการเรียนของแต่ละวิชานั้น ผู้สอนต้องใช้ผลการวัดหลาย ๆ อย่างมารวมกัน เช่น จากคะแนนภาคปฏิบัติ คะแนนรายงาน หรือโครงการ คะแนนสอบกลางภาคและคะแนนสอบปลายภาค จึงมีปัญหว่าคะแนนดิบที่ได้จากส่วนต่าง ๆ เหล่านี้ จะนำมารวมกันได้ทันทีหรือไม่

ก่อนจะตอบปัญหานี้ ขอให้พิจารณาตัวอย่างนี้ สมมติ วิชาเคมี คะแนนเต็ม 100 คะแนน

แบ่งเป็นคะแนนกลางภาค 40 คะแนน (แบ่งเป็น 2 ส่วนย่อย ส่วนละ 20 คะแนน) ปลายภาค 60 คะแนน แดง และ ดำ ได้คะแนนดังนี้

ชื่อ	คะแนนกลางภาค		คะแนนปลายภาค	รวม
	ครั้งที่ 1 (20)	ครั้งที่ 2 (20)		
แดง	18	18	40	76
ดำ	15	15	46	76

จะเห็นว่า ผู้สอนคงตัดสินว่า เก่งเท่ากัน หรือได้ระดับคะแนนเหมือนกัน ย่อมไม่ยุติธรรมแก่ผู้เข้าสอบ จึงไม่ควรนำคะแนนดิบจากหลาย ๆ ส่วนมารวมกัน

วิธีการที่ดีและยุติธรรมแก่ผู้เข้าสอบ ก็โดยการแปลงคะแนนดิบเหล่านี้ ให้เป็นคะแนนมาตรฐานรูปใดรูปหนึ่งก่อน แล้วนำมาวมกัน ดังตัวอย่างนี้ สมมติในการสอบดังกล่าวได้คำนวณหาข้อมูลเพิ่มเติมเป็นดังนี้

$$\text{สอบกลางภาคครั้งที่ 1 } \bar{X} = 15, S = 3$$

$$\text{สอบกลางภาคครั้งที่ 2 } \bar{X} = 12, S = 5$$

$$\text{สอบปลายภาค } \bar{X} = 30, S = 10$$

ถ้าแปลงคะแนนของแดงและดำเป็น คะแนน Z จะได้ดังนี้ (สูตร $Z = \frac{X - \bar{X}}{S}$)

ชื่อ	คะแนน Z กลางภาค		คะแนน Z ปลายภาค	รวม
	ครั้งที่ 1	ครั้งที่ 2		
แดง	1	1.2	1	3.2
ดำ	0	0.6	1.6	2.2

จะเห็นว่า แดง เก่งกว่า ดำ

วิธีรวมคะแนน

การรวมคะแนนที่ช่วยให้การตัดสินมีความยุติธรรมหรือมีประสิทธิภาพสูง มีหลายวิธี แต่ในที่นี้จะเสนอ 2 วิธี ดังนี้

1. การรวมโดยใช้คะแนนมาตรฐาน
2. การรวมโดยใช้ระดับคะแนน (เกรด)

1. การรวมโดยใช้คะแนนมาตรฐาน

ในการวัดผล คะแนนที่อยู่ในรูปเดียวกัน เช่น คะแนน Z หรือ T เป็นคะแนนที่มีหน่วยเท่ากัน จึงสามารถนำมาบวกลบกันได้ การรวมคะแนน Z หรือ T ก็รวมตามน้ำหนักที่ได้กำหนดไว้ เช่น ถ้ากำหนดสอบกลางภาคเป็น 40% (สอบ 2 ครั้ง สอบปลายภาค เป็น 60%)

$$Z \text{ รวม} = 2Z + 2Z + 6Z = 10Z$$

$$\text{หรือ } T \text{ รวม} = 2T + 2T + 6T = 10T$$

2. การรวมโดยใช้ระดับคะแนน (เกรด)

ในบางครั้งผู้สอนให้คะแนนนักเรียนในรูปของเกรดไว้แล้ว เช่น การจัดอันดับคุณภาพของงานภาคปฏิบัติ หรืองานอื่น ๆ ผู้สอนได้ประเมินออกมาเป็นเกรด A,B,C,D หรือ F หรือ 4,3,2,1,0 การให้เกรดของวิชานี้จึงได้จากการนำเกรดของงานแต่ละส่วนมารวมกัน การรวมโดยวิธีนี้ก็จะต้องนำน้ำหนักที่กำหนดให้กับงานหรือการสอบแต่ละครั้งมาคำนวณด้วย เช่น ถ้าการสอบครั้งที่ 1 ครั้งที่ 2 และการสอบปลายภาคมีน้ำหนักเป็น 1:1:2 ตามลำดับ นักเรียนจะได้เกรดของวิชานี้ ดังนี้

ตารางการรวมระดับคะแนน

นักเรียน	การสอบ			น้ำหนักรวม	ระดับคะแนน (น้ำหนักรวม ÷ 4)	เกรดของวิชานี้
	ครั้งที่ น้ำหนัก	I 1	II 1	III 2		
สิน		C (2x1)	C (2x1)	B (3x2)	10	C+
สม		A (4x1)	C (2x1)	B (3x2)	12	B
สอน		C	C	C	8	C
สาย		D	C	F	3	D-
เสริม		A	D	D	7	C-
สร้อย		B	C	A	13	B+

หมายเหตุ จากระดับ คะแนนที่เกิดจากการหาค่าเฉลี่ย (น้ำหนักรวม ÷ 4) ย่อมเป็นค่าทศนิยม ดังนั้นการให้เกรด อาจจำแนกรายละเอียดลงไปอีก ดังนี้

เกณฑ์การเปลี่ยนระดับคะแนนเป็นเกรด

ระดับคะแนน	เกรด
3.75 - 4.00	A
3.50 - 3.74	A ⁻
3.25 - 3.49	B ⁺
3.00 - 3.24	B
2.75 - 2.99	B ⁻

ระดับคะแนน	เกรด
2.50 - 2.74	C ⁺
2.00 - 2.49	C
1.75 - 1.99	C ⁻
1.50 - 1.75	D ⁺
1.00 - 1.49	D
0.50 - 0.99	D ⁻
0.00 - 0.49	F

หนังสืออ้างอิง

- โกวิท ประวาลพฤษษ์ และ สมศักดิ์ สันตุระเวชญ์. การตัดเกรด GRADING. กรุงเทพมหานคร : เจริญพัฒน์, 2518. 89 หน้า.
- นิไลบล นิ่มกิ่งรัตน์. วิธีการให้เกรดในการประเมินผลการเรียน. ม.ป.ท. : 2524. 58 หน้า.
- สมนึก ภัททิยธนี. การประเมินผลและการสร้างแบบทดสอบ. กทม. : ประสานการพิมพ์, 2533. 280 หน้า.