

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในการศึกษาค้างนี้ ผู้วิจัยเสนอผลการวิเคราะห์ข้อมูลดังต่อไปนี้

4.1 การสร้างแบบทดสอบความถนัดทางการเรียนด้านตัวเลข

4.1.1 การทดสอบครั้งที่ 1

4.1.1.1 ค่าสถิติพื้นฐานของแบบทดสอบ

4.1.1.2 การวิเคราะห์รายชื่อจากการทดสอบครั้งที่ 1

4.1.2 การทดสอบครั้งที่ 2

4.1.2.1 ค่าสถิติพื้นฐานของแบบทดสอบ

4.1.2.2 การวิเคราะห์รายชื่อจากการทดสอบครั้งที่ 2

4.1.3 การทดสอบครั้งที่ 3 เพื่อหาคุณภาพของแบบทดสอบ

4.1.3.1 ค่าสถิติพื้นฐานของแบบทดสอบ

4.1.3.2 การวิเคราะห์รายชื่อ จากการทดสอบครั้งที่ 3

4.1.3.3 ค่าความเชื่อมั่นของแบบทดสอบ

4.1.3.4 ค่าความเที่ยงตรงของแบบทดสอบ

4.2 การสร้างเกณฑ์ปกติ

4.1 การสร้างแบบทดสอบความถนัดทางการเรียนด้านตัวเลข

4.1.1 การทดสอบครั้งที่ 1

4.1.1.1 ค่าสถิติพื้นฐานของแบบทดสอบ

ค่าสถิติพื้นฐานของแบบทดสอบประกอบด้วย จำนวนนักเรียนที่เข้าสอบ (N) จำนวนข้อของแบบทดสอบแต่ละฉบับ (K) เวลาสอบ (t) ค่าเฉลี่ย ($\bar{X}$) ค่าส่วนเบี่ยงเบนมาตรฐาน (S) และค่าความคลาดเคลื่อนมาตรฐานในการวัด (SE_{meas}) ดังแสดงในตารางที่ 1 ต่อไปนี้

ตารางที่ 1 ค่าสถิติพื้นฐานของแบบทดสอบจากการทดสอบครั้งที่ 1

แบบทดสอบ	N	K	t (นาที)	$\bar{X}$	S	SE _{meas}
ทักษะทางคณิตศาสตร์	187	50	40	19.61	8.69	2.88
อนุกรม	187	50	40	20.35	8.43	3.04
เหตุผลทางคณิตศาสตร์	187	40	40	15.46	6.30	2.82

จากตารางที่ 1 พบว่าแบบทดสอบทั้ง 3 ฉบับมีค่าเฉลี่ยต่ำกว่าครึ่งหนึ่งของคะแนนเต็ม ซึ่งแสดงว่าแบบทดสอบทั้ง 3 ฉบับค่อนข้างยาก สมควรปรับปรุงให้ง่ายขึ้น

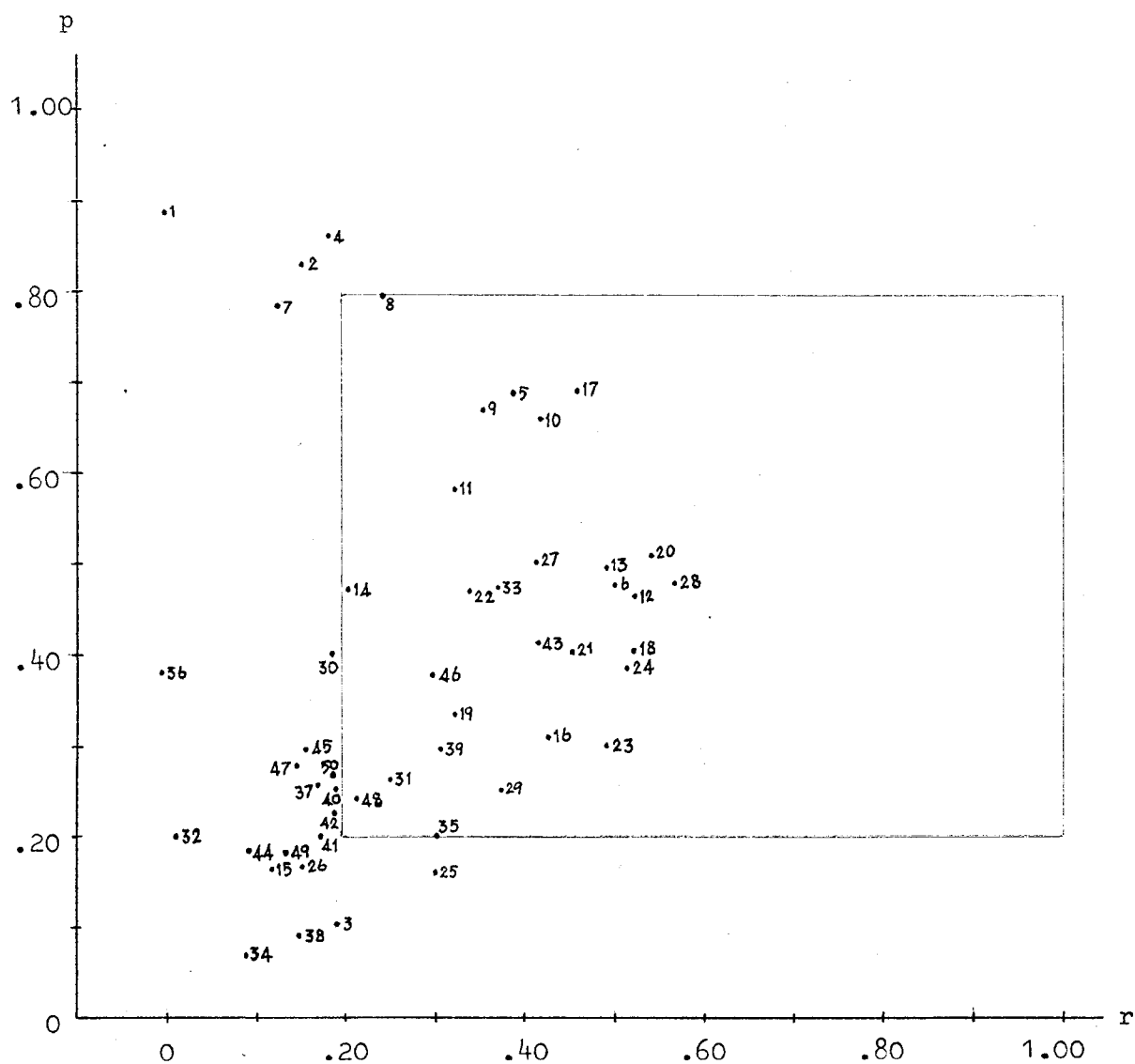
ค่าส่วนเบี่ยงเบนมาตรฐานของแบบทดสอบทั้ง 3 ฉบับ มีค่าตั้งแต่ 6.30-8.69 แบบทดสอบวัดทักษะทางคณิตศาสตร์มีค่าส่วนเบี่ยงเบนมาตรฐาน 8.69 ซึ่งสูงกว่าแบบทดสอบฉบับอื่น ๆ

ค่าความคลาดเคลื่อนมาตรฐานในการวัด มีค่าอยู่ระหว่าง 2.82 ถึง 3.04 นับว่าอยู่ในระดับปานกลางไม่สูงมากนัก พอที่จะเชื่อได้ว่าคะแนนที่ได้จากการทดสอบเป็นคะแนนที่ใกล้เคียงกับความสามารถของนักเรียนจริง ๆ

4.1.1.2 การวิเคราะห์รายชื่อจากการทดสอบครั้งที่ 1

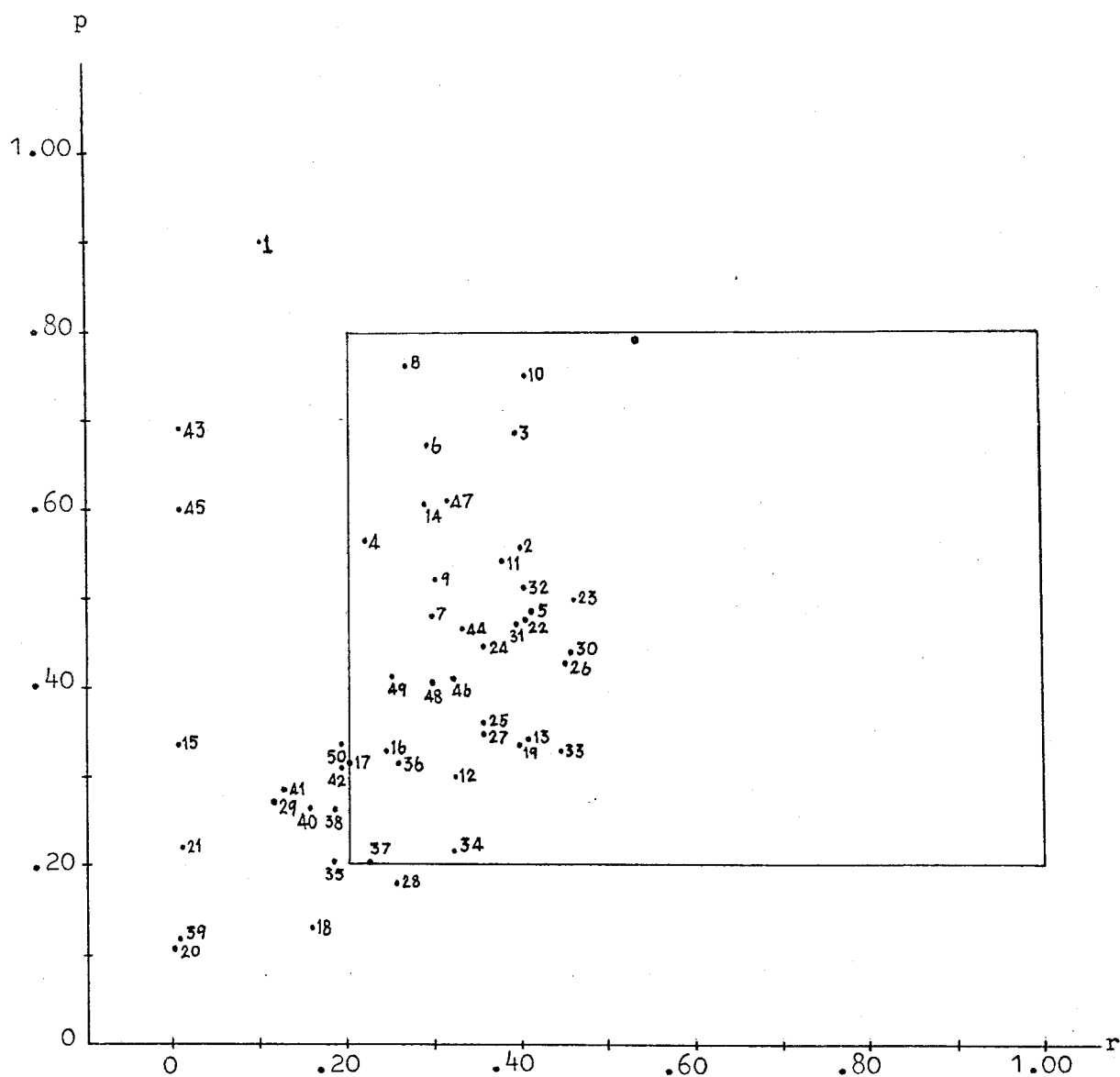
ผู้วิจัยนำผลจากการทดลองครั้งที่ 1 ของแบบทดสอบทั้ง 3 ฉบับมาคำนวณค่าความยากง่าย และค่าอำนาจจำแนกเพื่อคัดเลือกข้อสอบที่มีคุณภาพอยู่ในเกณฑ์ที่ต้องการโดยใช้ ไมโครคอมพิวเตอร์ ผลการวิเคราะห์รายชื่อของแบบทดสอบทั้ง 3 ฉบับ แสดงเป็นกราฟได้ดังภาพที่ 2, 3 และ 4 ดังต่อไปนี้

ภาพที่ 2 กราฟแสดงคุณภาพข้อสอบรายข้อของแบบทดสอบวัดทักษะทางคณิตศาสตร์
จากการทดสอบครั้งที่ 1



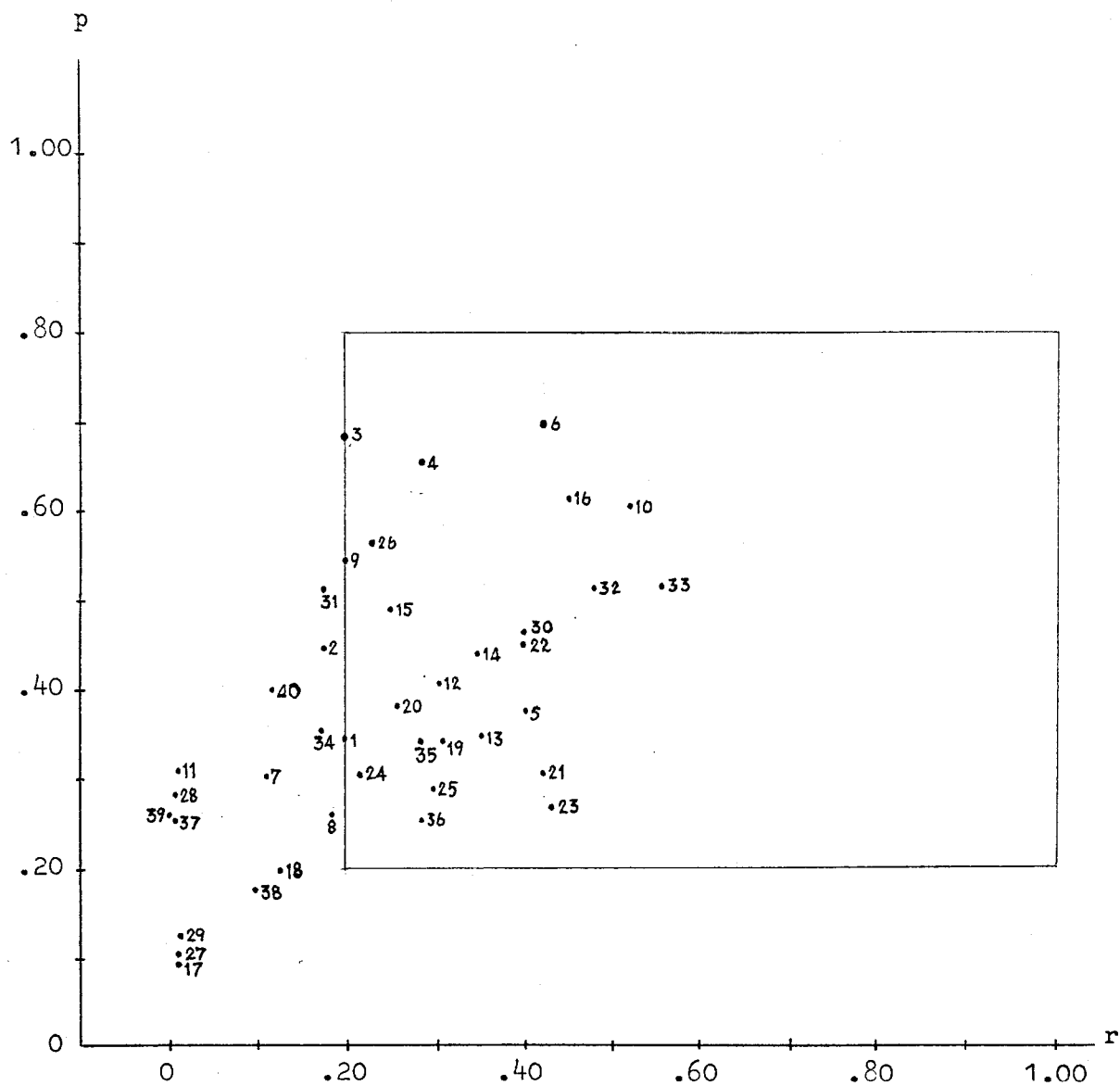
จากภาพที่ 2 พบว่าข้อสอบที่ต้องปรับปรุงแก้ไขคือข้อสอบที่มีค่าความยากง่าย (p) ต่ำกว่า .20 หรือสูงกว่า .80 และค่าอำนาจจำแนก (r) ต่ำกว่า .20 (ยาใจ, 2527) ซึ่งมีจำนวน 22 ข้อ จากทั้งหมด 50 ข้อ (รายละเอียดของค่า p และ r แสดงไว้ในตารางที่ 12 ภาคผนวก ก.)

ภาพที่ 3 กราฟแสดงคุณภาพข้อสอบรายข้อของแบบทดสอบอนุกรม จากการทดสอบ
ครั้งที่ 1



จากภาพที่ 3 พบว่าข้อสอบข้อที่ต้องปรับปรุงแก้ไขมีจำนวน 16 ข้อ จากทั้งหมด 50 ข้อ
(รายละเอียดของค่า p และ r แสดงไว้ในตารางที่ 12 ภาคผนวก ก.)

ภาพที่ 4 กราฟแสดงคุณภาพข้อสอบรายข้อของแบบทดสอบวัดเหตุผลทางคณิตศาสตร์
จากการทดสอบครั้งที่ 1



จากภาพที่ 4 พบว่าข้อสอบข้อที่ต้องปรับปรุงแก้ไขมีจำนวน 15 ข้อ จากทั้งหมด 40 ข้อ
(รายละเอียดของค่า p และ r แสดงไว้ในตารางที่ 12 ภาคผนวก ก.)

เมื่อนำค่าความยากง่ายและค่าอำนาจจำแนกไปหาค่าเฉลี่ยโดยค่าความยากง่ายเฉลี่ยใช้ค่าความยากง่ายมาตรฐาน (Δ) มาเฉลี่ยแล้วเปลี่ยนกลับเป็นค่าความยากง่าย ส่วนค่าอำนาจจำแนกเฉลี่ยนี้แปลงค่า r ให้เป็นคะแนนมาตรฐาน (Z) โดยวิธีการของฟิชเชอร์ (Fisher's Z Transformation) แล้วหา Z เฉลี่ยจากนั้นจึงเปลี่ยนกลับเป็นค่า r (โกวิท, 2527) จึงได้ค่าความยากง่ายเฉลี่ยและค่าอำนาจจำแนกเฉลี่ย ดังแสดงในตารางที่ 2 ต่อไปนี้

ตารางที่ 2 ค่าความยากง่ายเฉลี่ยและค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบทั้ง 3 ฉบับ
จากการทดสอบครั้งที่ 1

แบบทดสอบ	ค่าความยากง่าย			ค่าอำนาจจำแนก		
	ต่ำสุด	สูงสุด	เฉลี่ย	ต่ำสุด	สูงสุด	เฉลี่ย
ทักษะทางคณิตศาสตร์	.08	.89	.38	.00	.56	.24
อนุกรม	.11	.90	.40	.00	.46	.24
เหตุผลทางคณิตศาสตร์	.10	.68	.37	.02	.56	.22

4.1.2 การทดสอบครั้งที่ 2

4.1.2.1 ค่าสถิติพื้นฐานของแบบทดสอบประกอบด้วย จำนวนนักเรียนที่เข้าสอบ (N) จำนวนข้อของแบบทดสอบแต่ละฉบับ (K) เวลาสอบ (t) ค่าเฉลี่ย ($\bar{X}$) ค่าส่วนเบี่ยงเบนมาตรฐาน (S) และค่าความคลาดเคลื่อนมาตรฐานในการวัด (SE_{meas}) ดังแสดงในตารางที่ 3 ต่อไปนี้

ตารางที่ 3 ค่าสถิติพื้นฐานของแบบทดสอบจากการทดสอบครั้งที่ 2

แบบทดสอบ	N	K	t (นาที)	$\bar{X}$	S	SE _{meas}
ทักษะทางคณิตศาสตร์	215	50	40	20.44	9.95	2.99
อนุกรม	215	50	40	19.22	8.06	3.12
เหตุผลทางคณิตศาสตร์	215	40	40	15.85	6.01	2.88

เมื่อได้ปรับปรุงแก้ไขข้อสอบบางข้อของแบบทดสอบทั้ง 3 ฉบับแล้ว จากตารางที่ 3 พบว่าแบบทดสอบวัดทักษะทางคณิตศาสตร์ และแบบทดสอบวัดเหตุผลทางคณิตศาสตร์มีค่าเฉลี่ยสูงขึ้นจากเดิม แสดงว่าแบบทดสอบทั้ง 2 ฉบับนี้ เมื่อได้ปรับปรุงแก้ไขแล้วทำให้ง่ายขึ้น หรืออาจจะเป็นเพราะกลุ่มตัวอย่างมีความสามารถมากกว่ากลุ่มตัวอย่างที่ใช้ในการทดสอบครั้งที่ 1 แต่แบบทดสอบอนุกรมมีค่าเฉลี่ยลดลงเพียงเล็กน้อย แสดงว่าแบบทดสอบฉบับนี้ เมื่อได้ปรับปรุงแก้ไขแล้วทำให้ยากขึ้นหรืออาจจะเป็นเพราะกลุ่มตัวอย่างที่เปลี่ยนไป

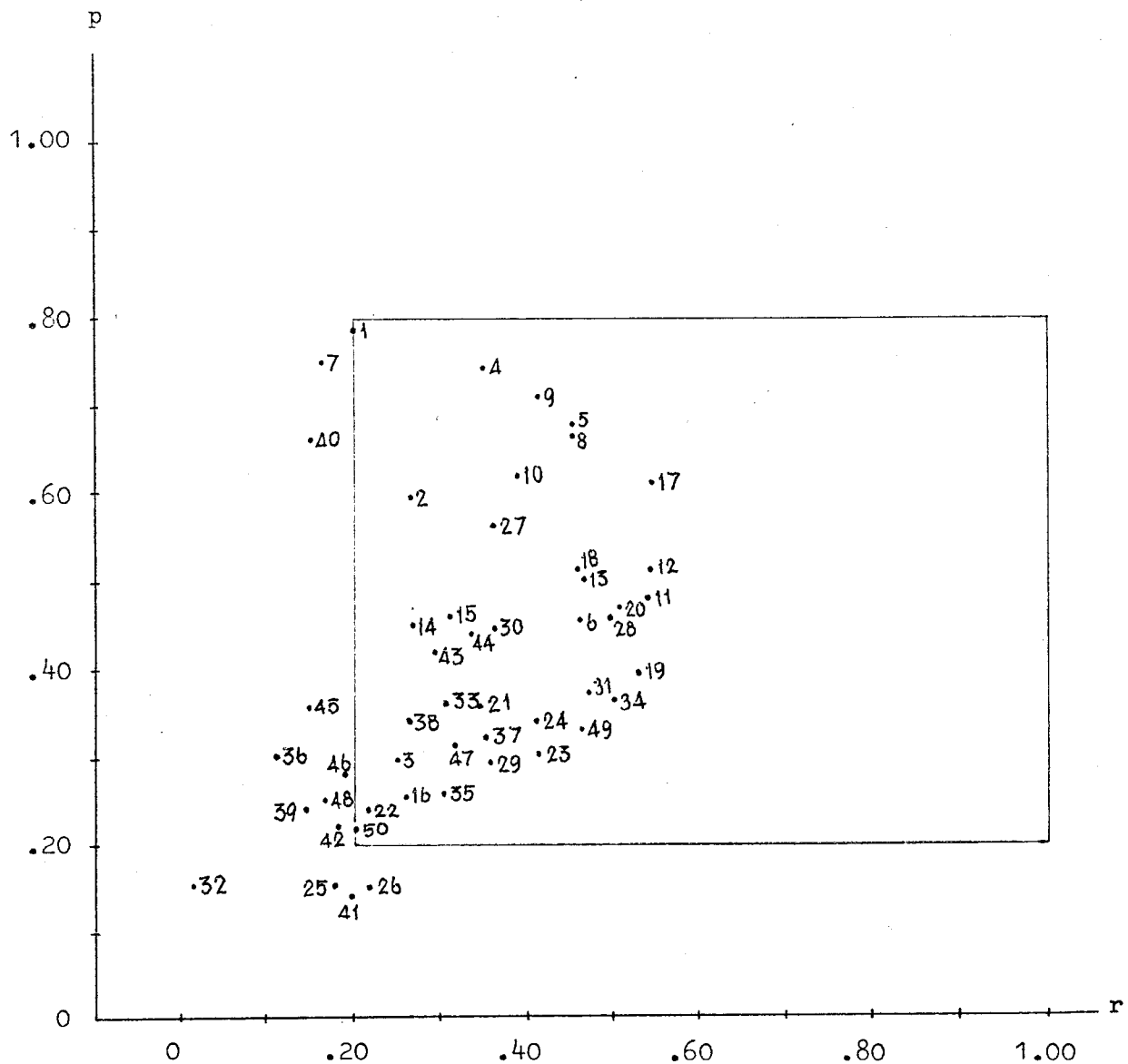
ค่าส่วนเบี่ยงเบนมาตรฐานของแบบทดสอบทั้ง 3 ฉบับ พบว่าแบบทดสอบวัดทักษะทางคณิตศาสตร์มีค่าเพิ่มขึ้น แต่แบบทดสอบอนุกรมและแบบทดสอบวัดเหตุผลทางคณิตศาสตร์ลดลงอาจจะเป็นเพราะกลุ่มตัวอย่างที่เปลี่ยนไป

ค่าความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบทั้ง 3 ฉบับ พบว่ามีค่าใกล้เคียงกับการทดสอบครั้งที่ 1 มาก แสดงว่าคะแนนจากการทดสอบเป็นคะแนนที่ใกล้เคียงกับความสามารถของนักเรียนจริง ๆ

4.1.2.2 การวิเคราะห์รายข้อจากการทดสอบครั้งที่ 2

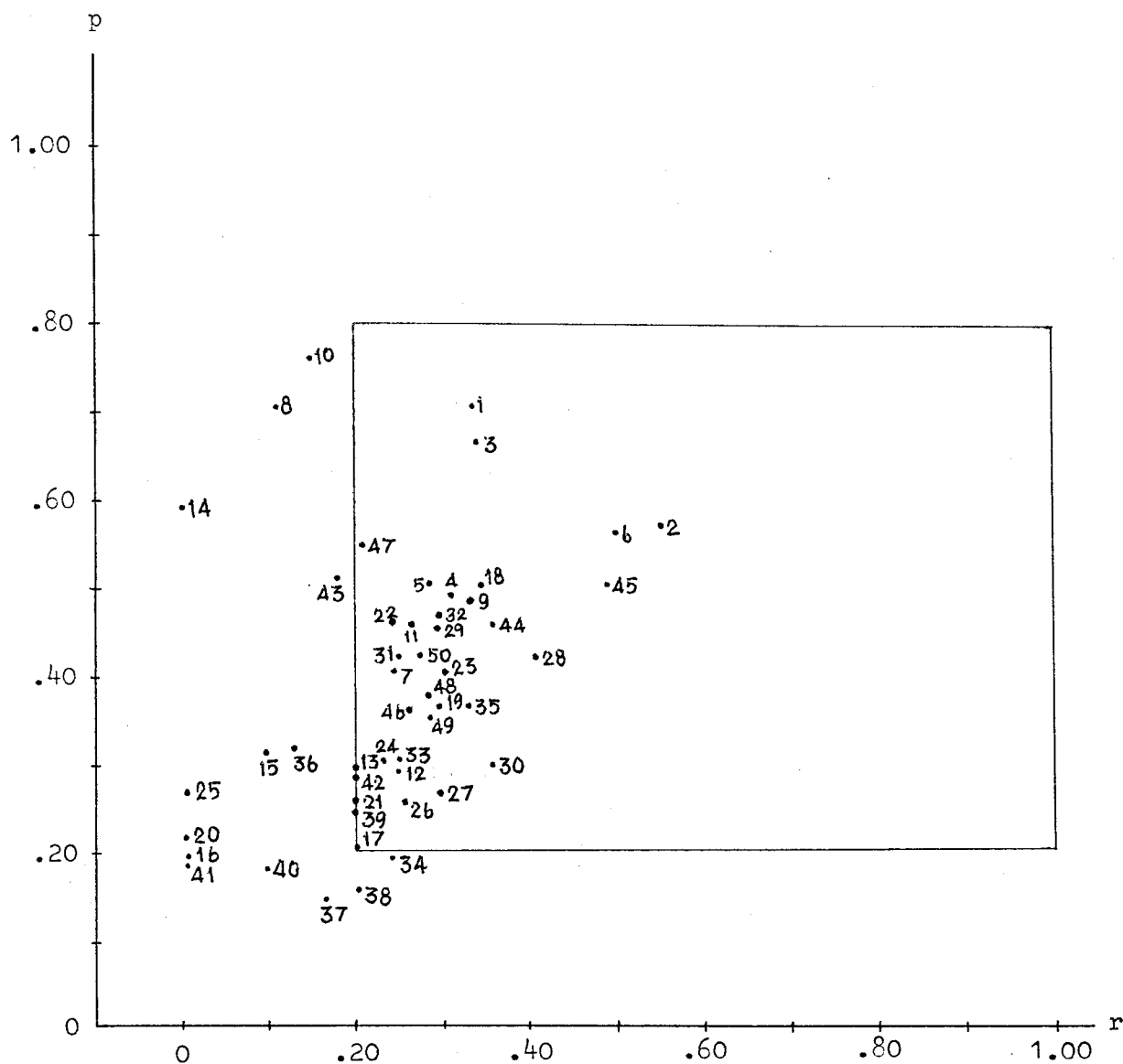
ผู้วิจัยนำผลจากการทดสอบครั้งที่ 2 ของแบบทดสอบทั้ง 3 ฉบับ มาคำนวณค่าความยากง่าย และค่าอำนาจจำแนกโดยใช้ไมโครคอมพิวเตอร์ ผลการวิเคราะห์รายข้อของแบบทดสอบทั้ง 3 ฉบับ แสดงเป็นกราฟได้ดังภาพที่ 5, 6 และ 7 ดังต่อไปนี้

ภาพที่ 5 กราฟแสดงคุณภาพข้อสอบรายข้อของแบบทดสอบวัดทักษะทางคณิตศาสตร์
จากการทดสอบครั้งที่ 2



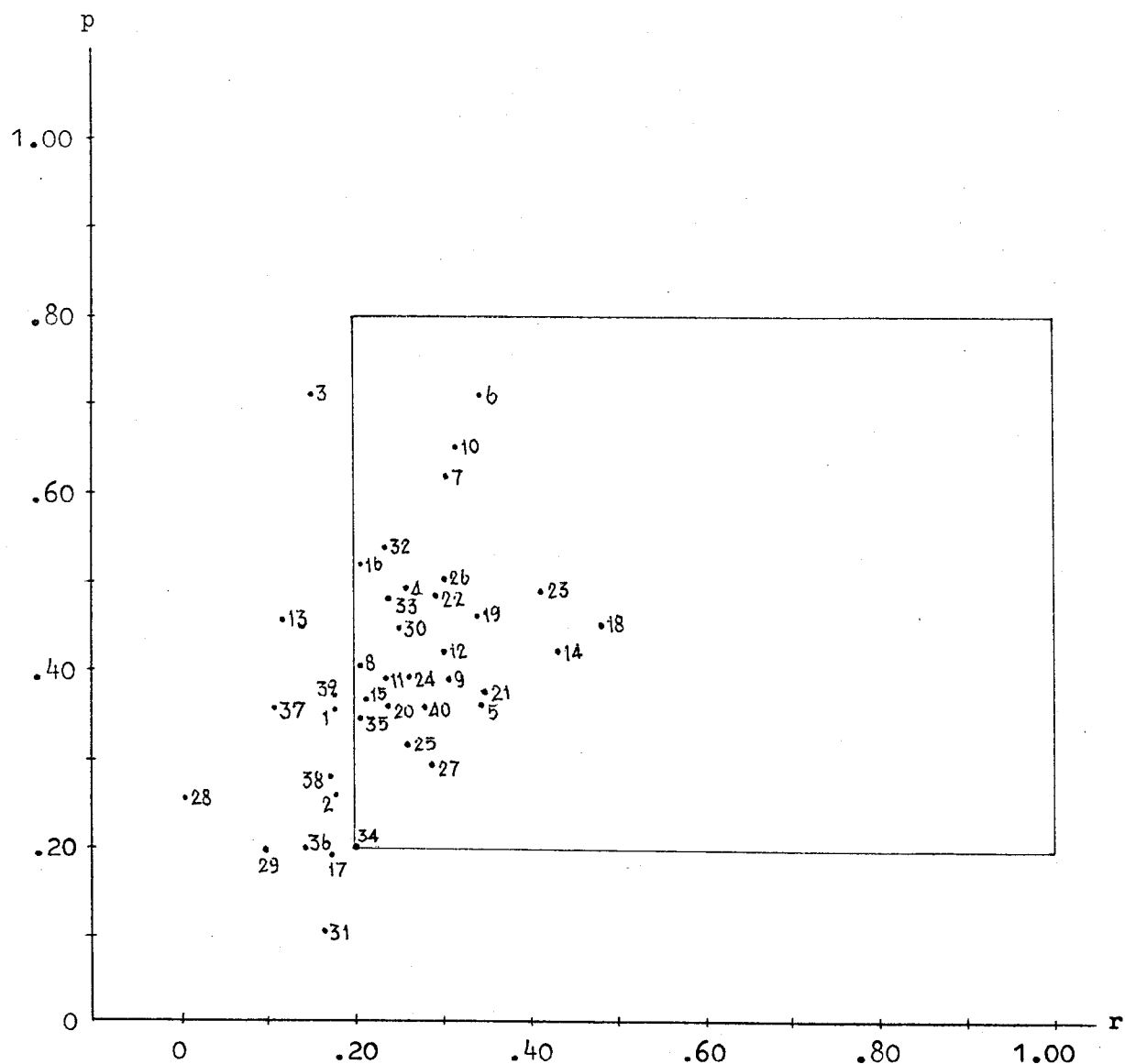
จากภาพที่ 5 พบว่าข้อสอบที่มีคุณภาพตามเกณฑ์มีจำนวน 38 ข้อ จากทั้งหมด 50 ข้อ
(รายละเอียดของค่า p และ r แสดงไว้ในตารางที่ 13 ภาคผนวก ก.)

ภาพที่ 6 กราฟแสดงคุณภาพข้อสอบรายข้อของแบบทดสอบอนุกรม จากการทดสอบ
ครั้งที่ 2



จากภาพที่ 6 พบว่าข้อสอบที่มีคุณภาพตามเกณฑ์มีจำนวน 36 ข้อ จากทั้งหมด 50 ข้อ
(รายละเอียดของค่า p และ r แสดงไว้ในตารางที่ 13 ภาคผนวก ก.)

ภาพที่ 7 กราฟแสดงคุณภาพข้อสอบรายข้อของแบบทดสอบวัดเหตุผลทางคณิตศาสตร์
จากการทดสอบครั้งที่ 2



จากภาพที่ 7 พบว่าข้อสอบที่มีคุณภาพตามเกณฑ์มีจำนวน 28 ข้อ จากทั้งหมด 40 ข้อ
(รายละเอียดของค่า p และ r แสดงไว้ในตารางที่ 13 ภาคผนวก ก.)

เมื่อนำค่าความยากง่ายและค่าอำนาจจำแนกไปหาค่าเฉลี่ย จึงได้ค่าความยากง่ายเฉลี่ย และค่าอำนาจจำแนกเฉลี่ย ดังแสดงในตารางที่ 4 ต่อไปนี้

ตารางที่ 4 ค่าความยากง่ายเฉลี่ยและค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบทั้ง 3 ฉบับ
จากการทดสอบครั้งที่ 2

แบบทดสอบ	ค่าความยากง่าย			ค่าอำนาจจำแนก		
	ต่ำสุด	สูงสุด	เฉลี่ย	ต่ำสุด	สูงสุด	เฉลี่ย
ทักษะทางคณิตศาสตร์	.14	.79	.40	.06	.54	.31
อนุกรม	.13	.70	.38	.04	.55	.25
เหตุผลทางคณิตศาสตร์	.10	.70	.39	.05	.48	.24

จากตารางที่ 4 พบว่า แบบทดสอบทั้ง 3 ฉบับมีค่าความยากง่ายเฉลี่ยสูงขึ้น แสดงว่าแบบทดสอบทั้ง 3 ฉบับ ยากขึ้นกว่าเดิม ส่วนค่าอำนาจจำแนกเฉลี่ยนั้น แบบทดสอบทั้ง 3 ฉบับก็มีค่าอำนาจจำแนกเฉลี่ยสูงขึ้น แสดงว่าแบบทดสอบทั้ง 3 ฉบับ จำแนกเด็กเก่ง เด็กอ่อนได้ดีขึ้น

4.1.3 การทดสอบครั้งที่ 3 เพื่อหาคุณภาพของแบบทดสอบ

การทดสอบในครั้งที่ 1 และครั้งที่ 2 ที่ผ่านไปแล้ว เป็นการทดสอบเพื่อปรับปรุงแก้ไขและคัดเลือกแบบทดสอบที่สร้างขึ้นทั้ง 3 ฉบับ โดยวิเคราะห์ข้อสอบเป็นรายข้อและพิจารณาค่าสถิติพื้นฐานประกอบ เพื่อจัดข้อสอบข้อที่เหมาะสมเข้าเป็นแบบทดสอบสำหรับการทดสอบครั้งที่ 3 ต่อไป

การทดสอบครั้งที่ 3 เป็นการทดสอบเพื่อหาคุณภาพของแบบทดสอบที่สร้างขึ้นทั้ง 3 ฉบับ โดยหาค่าความยากง่าย ค่าอำนาจจำแนก ค่าความเชื่อมั่นของแบบทดสอบ และหาค่าความเที่ยงตรงของแบบทดสอบในการวิเคราะห์ผลจากการทดสอบครั้งที่ 3 นี้ใช้คอมพิวเตอร์โปรแกรม SPSS^x ณ ศูนย์คอมพิวเตอร์ มหาวิทยาลัยขอนแก่น ซึ่งผลการวิเคราะห์จะได้เสนอเป็นลำดับต่อไป

4.1.3.1 ค่าสถิติพื้นฐานของแบบทดสอบ

ค่าสถิติพื้นฐานของแบบทดสอบประกอบด้วยจำนวนนักเรียนที่เข้าสอบ (N) จำนวนข้อของ

แบบทดสอบแต่ละฉบับ (K) เวลาสอบ (t) ค่าเฉลี่ย ($\bar{X}$) ค่าส่วนเบี่ยงเบนมาตรฐาน (S) และค่าความคลาดเคลื่อนมาตรฐานในการวัด (SE_{meas}) ดังแสดงในตารางที่ 5 ต่อไปนี้

ตารางที่ 5 ค่าสถิติพื้นฐานของแบบทดสอบ จากการทดสอบครั้งที่ 3

แบบทดสอบ	N	K	t (นาที)	$\bar{X}$	S	SE_{meas}
ทักษะทางคณิตศาสตร์	720	38	30	19.34	8.89	2.37
อนุกรม	720	36	34	17.04	7.97	2.30
เหตุผลทางคณิตศาสตร์	720	28	30	13.74	5.94	2.30

จากตารางที่ 5 พบว่าแบบทดสอบทั้ง 3 ฉบับ มีค่าเฉลี่ยประมาณครึ่งหนึ่งของคะแนนเต็ม ซึ่งสูงกว่าการทดสอบครั้งที่ 1 และครั้งที่ 2 แสดงว่าแบบทดสอบทั้ง 3 ฉบับ ได้รับปรับปรุงให้ดีขึ้น มีค่าเฉลี่ยสูงขึ้น

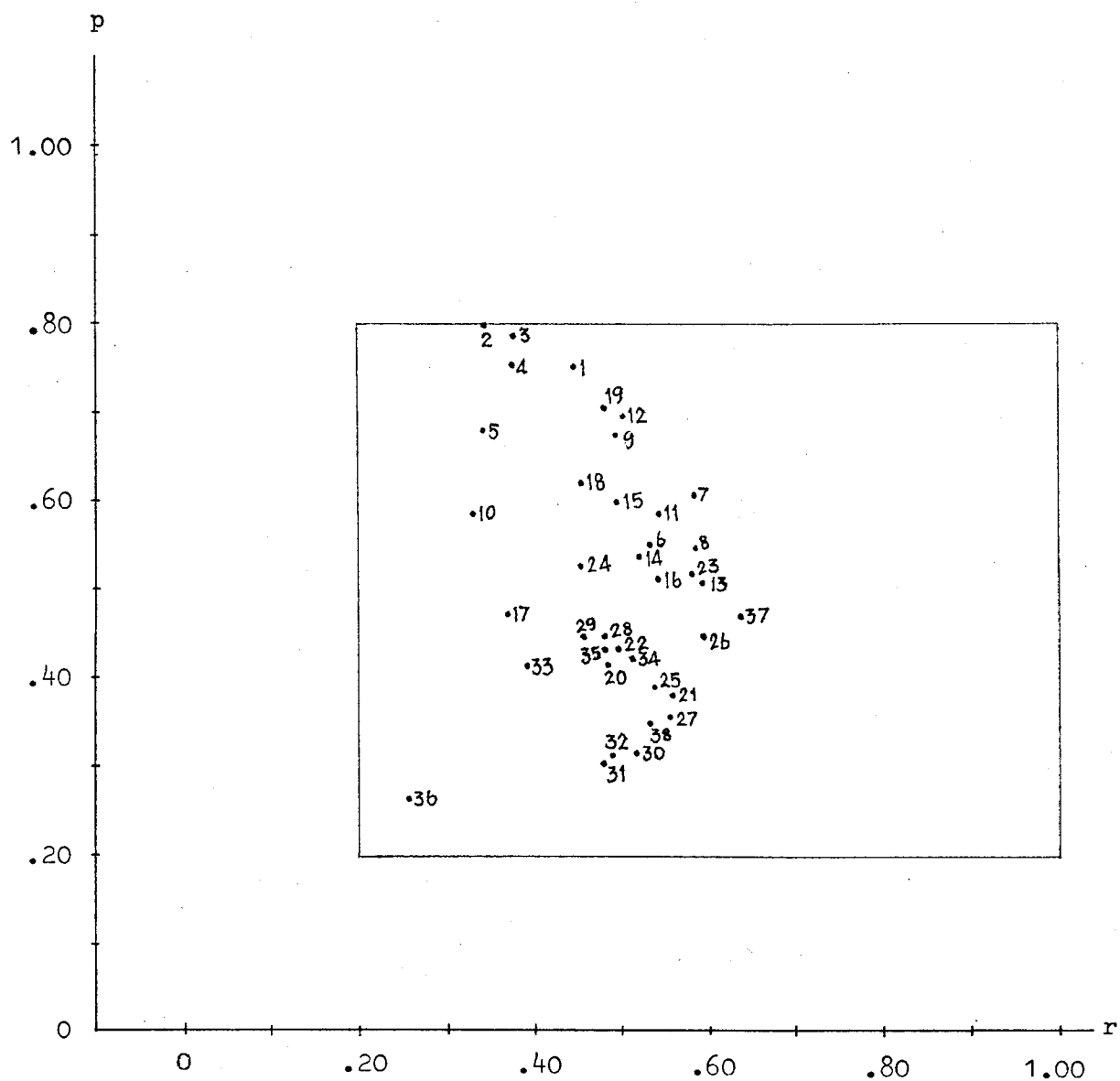
ค่าส่วนเบี่ยงเบนมาตรฐานของแบบทดสอบทั้ง 3 ฉบับ พบว่ามีค่าใกล้เคียงกับการทดสอบครั้งที่ 1 และครั้งที่ 2 แสดงว่าการกระจายของคะแนนจากการทดสอบทั้ง 3 ครั้ง ใกล้เคียงกัน

สำหรับค่าความคลาดเคลื่อนมาตรฐานของการวัด ปรากฏว่ามีค่าค่อนข้างต่ำทุกฉบับ คือ มีค่าอยู่ระหว่าง 2.43 ถึง 2.72 แสดงว่าคะแนนที่ได้มาเป็นคะแนนที่เบี่ยงเบนไปจากคะแนนจริงเพียงเล็กน้อยเท่านั้น พอที่จะเชื่อถือได้ว่าเป็นคะแนนที่ได้จากความสามารถของนักเรียนในกลุ่มตัวอย่างจริง ๆ

4.1.3.2 การวิเคราะห์รายข้อ จากการทดสอบครั้งที่ 3

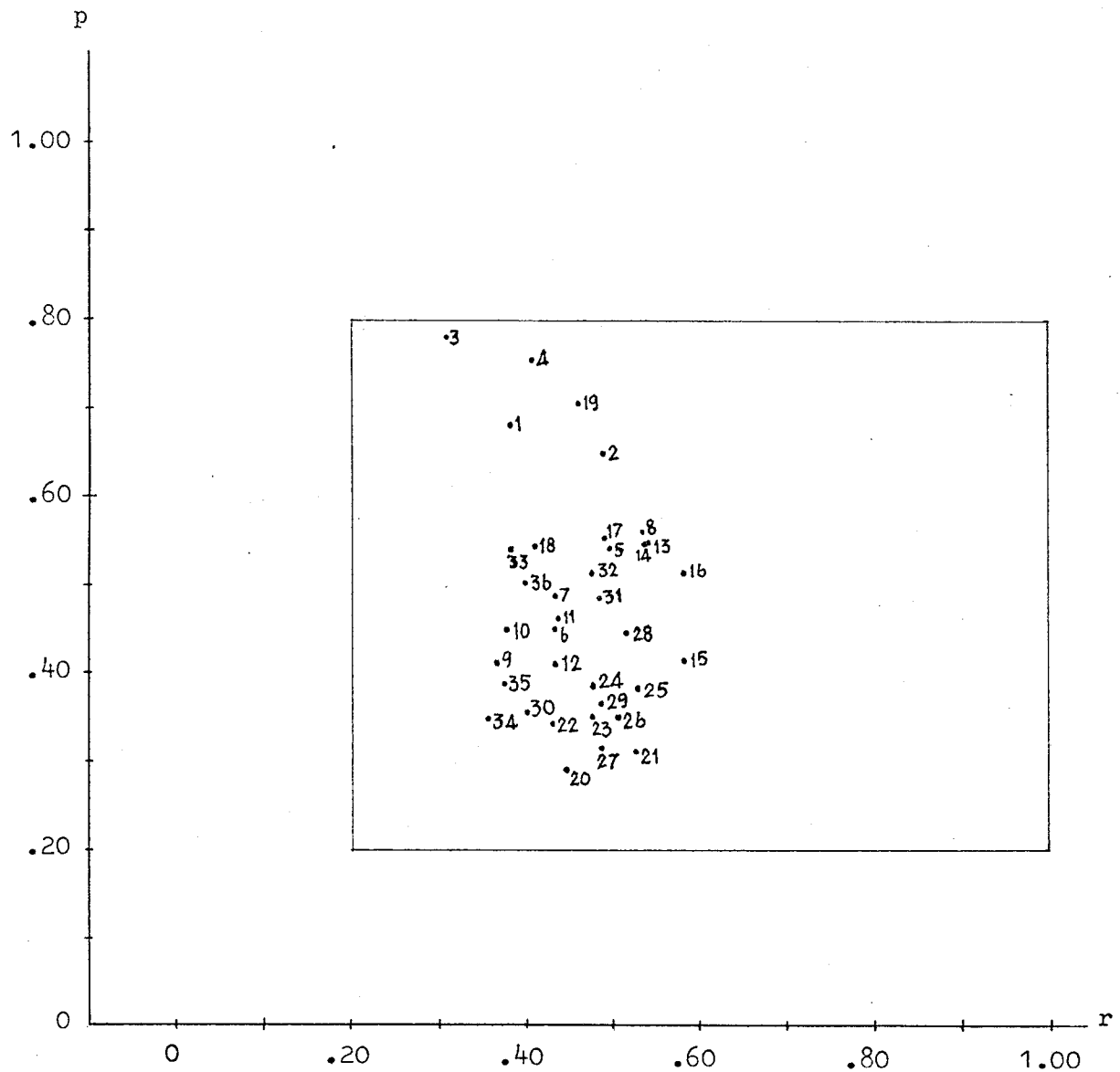
ผู้วิจัยนำผลจากการทดสอบครั้งที่ 3 ของแบบทดสอบทั้ง 3 ฉบับ มาคำนวณค่าความยากง่าย และค่าอำนาจจำแนก โดยใช้คอมพิวเตอร์โปรแกรม SPSS* ผลการวิเคราะห์รายข้อของแบบทดสอบทั้ง 3 ฉบับ แสดงเป็นกราฟได้ดังภาพที่ 8, 9 และ 10 ดังต่อไปนี้

ภาพที่ 8 กราฟแสดงคุณภาพข้อสอบรายข้อของแบบทดสอบวัดทักษะทางคณิตศาสตร์
จากการทดสอบครั้งที่ 3



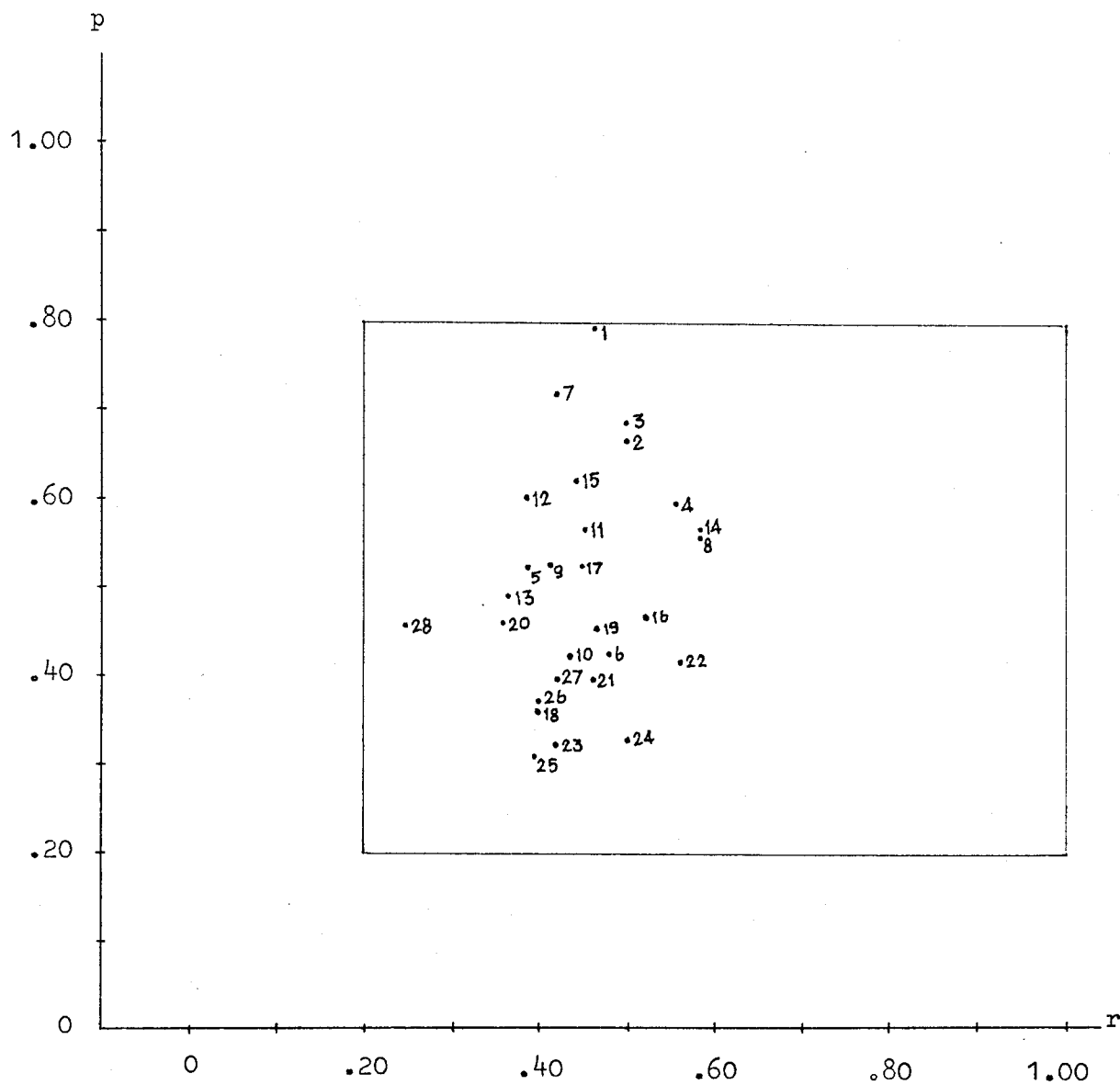
จากภาพที่ 8 พบว่าข้อสอบทุกข้อมีค่า p และ r อยู่ในเกณฑ์ดี แสดงว่าข้อสอบทุกข้อมีความยากง่ายพอเหมาะและสามารถจำแนกเด็กเก่ง เด็กอ่อนได้ดี (รายละเอียดของค่า p และ r แสดงไว้ในตารางที่ 14 ภาคผนวก ก.)

ภาพที่ 9 กราฟแสดงคุณภาพข้อสอบรายข้อของแบบทดสอบอนุกรมจากการทดสอบครั้งที่ 3



จากภาพที่ 9 พบว่าข้อสอบทุกข้อมีค่า p และ r อยู่ในเกณฑ์ดี แสดงว่าข้อสอบทุกข้อมีความยากง่ายพอเหมาะและสามารถจำแนกเด็กเก่ง เด็กอ่อนได้ดี (รายละเอียดของค่า p และ r แสดงไว้ในตารางที่ 14 ภาคผนวก ก.)

ภาพที่ 10 กราฟแสดงคุณภาพข้อสอบรายชื่อของแบบทดสอบวัดเหตุผลทางคณิตศาสตร์
จากการทดสอบครั้งที่ 3



จากภาพที่ 10 พบว่าข้อสอบทุกข้อมีค่า p และ r อยู่ในเกณฑ์ดี แสดงว่าข้อสอบทุกข้อมีความยากง่ายพอเหมาะ และสามารถจำแนกเด็กเก่ง เด็กอ่อนได้ดี (รายละเอียดของค่า p และ r แสดงไว้ในตารางที่ 14 ภาคผนวก ก.)

เมื่อนำค่าความยากง่ายและค่าอำนาจจำแนกไปหาค่าเฉลี่ย จึงได้ค่าความยากง่ายเฉลี่ย และค่าอำนาจจำแนกเฉลี่ย ดังแสดงในตารางที่ 6 ต่อไปนี้

ตารางที่ 6 ค่าความยากง่ายเฉลี่ยและค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบทั้ง 3 ฉบับ
จากการทดสอบครั้งที่ 3

แบบทดสอบ	ค่าความยากง่าย			ค่าอำนาจจำแนก		
	ต่ำสุด	สูงสุด	เฉลี่ย	ต่ำสุด	สูงสุด	เฉลี่ย
ทักษะทางคณิตศาสตร์	.27	.80	.50	.26	.60	.47
อนุกรม	.29	.78	.46	.31	.58	.44
เหตุผลทางคณิตศาสตร์	.30	.79	.48	.24	.57	.42

จากตารางที่ 6 พบว่าแบบทดสอบทั้ง 3 ฉบับ มีค่าความยากง่ายเฉลี่ยใกล้เคียงกับ .50 แสดงว่าแบบทดสอบทั้ง 3 ฉบับ มีความยากง่ายอยู่ในระดับปานกลาง คือ .40-.60 (โกวิท, 2527) และแบบทดสอบทั้ง 3 ฉบับ มีค่าความยากง่ายเฉลี่ยสูงขึ้น แสดงว่าแบบทดสอบทั้ง 3 ฉบับ ง่ายขึ้นกว่าเดิม ส่วนค่าอำนาจจำแนกเฉลี่ยนั้น แบบทดสอบทั้ง 3 ฉบับ มีค่าอำนาจจำแนกเฉลี่ยสูงขึ้นมากกว่า .40 ทั้ง 3 ฉบับ ซึ่ง Ebel (อ้างถึงใน โกวิท, 2527) กล่าวว่าค่าอำนาจจำแนกตั้งแต่ .40 ขึ้นไป เป็นค่าอำนาจจำแนกที่ดีมาก แสดงว่าแบบทดสอบทั้ง 3 ฉบับ สามารถจำแนกเด็กเก่ง เด็กอ่อนได้ดีมาก

4.1.3.3 ค่าความเชื่อมั่นของแบบทดสอบ

จากการทดสอบทั้ง 3 ครั้ง ค่าความเชื่อมั่นของแบบทดสอบฉบับต่าง ๆ ได้เปลี่ยนแปลงไปจากเดิมบ้าง บางฉบับมีค่าความเชื่อมั่นเพิ่มขึ้น บางฉบับมีค่าความเชื่อมั่นลดลง ซึ่งค่าความเชื่อมั่นของแบบทดสอบทั้ง 3 ฉบับ จากการทดสอบทั้ง 3 ครั้ง ได้แสดงในตารางที่ 7 ต่อไปนี้

ตารางที่ 7 ค่าความเชื่อมั่นของแบบทดสอบ จากการทดสอบทั้ง 3 ครั้ง

แบบทดสอบ	ทดสอบครั้งที่ 1		ทดสอบครั้งที่ 2		ทดสอบครั้งที่ 3	
	r_{tt}	K	r_{tt}	K	r_{tt}	K
ทักษะทางคณิตศาสตร์	.90**	50	.92**	50	.93**	38
อนุกรม	.88**	50	.87**	50	.92**	36
เหตุผลทางคณิตศาสตร์	.81**	40	.78**	40	.85**	28

**มีนัยสำคัญทางสถิติที่ระดับ .01

จากตารางที่ 7 พบว่า แบบทดสอบทั้ง 3 ฉบับ จากการทดสอบทั้ง 3 ครั้ง มีค่าความเชื่อมั่นสูงมากกว่า .50 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่าแบบทดสอบทั้ง 3 ฉบับ มีความเชื่อถือได้สูงเพราะคะแนนที่ได้จากการทดสอบใกล้เคียงกับคะแนนจริงของนักเรียน (โกวิท, 2527)

4.1.3.4 ค่าความเที่ยงตรงของแบบทดสอบ

ค่าความเที่ยงตรงของแบบทดสอบทั้ง 3 ฉบับ ในการศึกษาครั้งนี้ ได้หาความเที่ยงตรงตามสภาพโดยใช้สูตรของ Pearson Product Moment Correlation Coefficient โดยใช้คะแนนจากแบบทดสอบมาตรฐานความถนัดทางการเรียน ระดับมัธยมศึกษาตอนต้น ชุดเรียงอันดับ/ข ของสำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เป็นคะแนนเกณฑ์ และทดสอบความมีนัยสำคัญทางสถิติ โดยใช้ t-test ดังแสดงในตารางที่ 8 ต่อไปนี้

ตารางที่ 8 ค่าความเที่ยงตรงตามสภาพและการทดสอบความมีนัยสำคัญทางสถิติของแบบทดสอบทั้ง 3 ฉบับ

แบบทดสอบ	r_{xy}	t
ทักษะทางคณิตศาสตร์	.64	4.88**
อนุกรม	.76	10.71**
เหตุผลทางคณิตศาสตร์	.66	5.71**

** มีนัยสำคัญทางสถิติที่ระดับ .01

จากตารางที่ 8 พบว่า แบบทดสอบทั้ง 3 ฉบับมีค่าความเที่ยงตรงตามสภาพสูงมากกว่า .50 อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่า แบบทดสอบทั้ง 3 ฉบับ สามารถวัดความถนัดทางการเรียนด้านตัวเลขได้ตามความมุ่งหมาย

แต่การหาความเที่ยงตรงตามสภาพซึ่งใช้แบบทดสอบมาตรฐานความถนัดทางการเรียนระดับมัธยมศึกษาตอนต้น ชุดเรียงอันดับ/ข ของสำนักทดสอบทางการศึกษาและจิตวิทยามหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เป็นคะแนนเกณฑ์นั้น เป็นแบบทดสอบเรียงอันดับ ซึ่งมีโครงสร้างคล้ายคลึงกันกับแบบทดสอบอนุกรมที่ผู้วิจัยสร้างขึ้น อาจจะทำให้ค่าความเที่ยงตรงตามสภาพของแบบทดสอบอนุกรมมีค่าสูงแตกต่างจากฉบับอื่น ๆ จึงทำการทดสอบผลต่างของค่าสหสัมพันธ์ของแบบทดสอบในการทำนายแบบทดสอบมาตรฐานความถนัดทางการเรียนดังแสดงในตารางที่ 9 ต่อไปนี้

ตารางที่ 9 การทดสอบความแตกต่างระหว่างความเที่ยงตรงตามสภาพของแบบทดสอบความถนัดทางการเรียนด้านตัวเลข

ความสัมพันธ์ระหว่างแบบทดสอบเรียงอันดับ กับแบบทดสอบอื่น ๆ	r	t
เรียงอันดับ-อนุกรม	.76	5.82**
เรียงอันดับ-ทักษะทางคณิตศาสตร์	.64	
เรียงอันดับ-อนุกรม	.76	5.06**
เรียงอันดับ-เหตุผลทางคณิตศาสตร์	.66	

**มีนัยสำคัญทางสถิติที่ระดับ .01

จากตารางที่ 9 พบว่า ความเที่ยงตรงตามสภาพของแบบทดสอบอนุกรม มีค่าสูงกว่าความเที่ยงตรงตามสภาพของแบบทดสอบวัดทักษะทางคณิตศาสตร์และแบบทดสอบวัดเหตุผลทางคณิตศาสตร์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 แสดงว่าแบบทดสอบอนุกรมมีโครงสร้างที่คล้ายคลึงกับแบบทดสอบมาตรฐานความถนัดทางการเรียนรู้ชุดเรียงอันดับ

4.2 คะแนนเกณฑ์ปกติ

คะแนนเกณฑ์ปกติของแบบทดสอบทั้ง 3 ฉบับ หาโดยใช้คะแนนมาตรฐานในรูปของคะแนนตำแหน่งเปอร์เซ็นต์ไทล์ และคะแนนที่ปกติ ดังแสดงในตารางที่ 10 และตารางที่ 11 ต่อไปนี้

ตารางที่ 10 คะแนนเกณฑ์ปกติของแบบทดสอบทั้ง 3 ฉบับ โดยใช้คะแนนมาตรฐานในรูปของ
คะแนนตำแหน่งเปอร์เซ็นต์ไทล์

คะแนนดิบ	ตำแหน่งเปอร์เซ็นต์ไทล์ของแบบทดสอบ			คะแนนดิบ
	ทักษะทางคณิตศาสตร์	อนุกรม	เหตุผลทางคณิตศาสตร์	
38	99			38
37	99			37
36	97	99		36
35	95	99		35
34	92	99		34
33	90	98		33
32	89	96		32
31	86	94		31
30	84	92		30
29	82	91		29
28	80	89	99	28
27	78	86	99	27
26	75	83	98	26
25	72	81	97	25
24	69	78	95	24
23	66	75	92	23
22	63	72	90	22
21	59	69	86	21
20	56	66	82	20
19	52	63	79	19
18	49	59	74	18
17	44	55	70	17

ตารางที่ 10 (ต่อ)

คะแนนดิบ	ตำแหน่งเปอร์เซ็นต์ไคล์ของแบบทดสอบ			คะแนนดิบ
	ทักษะทางคณิตศาสตร์	อนุกรม	เหตุผลทางคณิตศาสตร์	
16	40	49	65	16
15	37	44	60	15
14	33	40	54	14
13	29	38	48	13
12	25	31	43	12
11	21	26	37	11
10	17	21	31	10
9	13	17	25	9
8	9	13	19	8
7	6	10	14	7
6	3	7	9	6
5	2	4	6	5
4	1	2	3	4
3	1	1	1	3
2				2
1				1

จากตารางที่ 10 พบว่า แบบทดสอบทั้ง 3 ฉบับ มีคะแนนดิบสูงสุด (38, 36 และ 28) ตรงกับคะแนนตำแหน่งเปอร์เซ็นต์ไคล์ที่ 99 และแบบทดสอบทั้ง 3 ฉบับ มีคะแนนดิบ 3 คะแนน ซึ่งตรงกับคะแนนตำแหน่งเปอร์เซ็นต์ไคล์ที่ 1 ถ้านักเรียนคนหนึ่งได้คะแนนดิบ 25 คะแนน จากการทำแบบทดสอบวัดทักษะทางคณิตศาสตร์ แบบทดสอบอนุกรม และแบบทดสอบวัดเหตุผลทางคณิตศาสตร์ ซึ่งตรงกับคะแนนตำแหน่งเปอร์เซ็นต์ไคล์ที่ 72, 81 และ 97 ตามลำดับ ซึ่งหมายความว่า มีนักเรียนที่สอบได้คะแนนต่ำกว่านักเรียนคนนี้ ร้อยละ 72 ร้อยละ 81 และร้อยละ 97 ตามลำดับ

ตารางที่ 11 คะแนนเกณฑัปกติของแบบทดสอบทั้ง 3 ฉบับ โดยใช้คะแนนมาตรฐานในรูปของคะแนนที่ปกติ

คะแนนดิบ	คะแนนปกติของแบบทดสอบ			คะแนนดิบ
	ทักษะทางคณิตศาสตร์	อนุกรม	เหตุผลทางคณิตศาสตร์	
38	82			38
37	73			37
36	69	79		36
35	66	75		35
34	64	73		34
33	63	70		33
32	62	67		32
31	61	66		31
30	60	64		30
29	60	63		29
28	59	62	80	28
27	58	61	74	27
26	57	60	71	26
25	56	59	69	25
24	55	58	66	24
23	54	57	64	23
22	53	56	63	22
21	52	55	61	21
20	52	54	59	20
19	51	53	58	19
18	50	52	57	18
17	49	51	55	17

ตารางที่ 11 (ต่อ)

คะแนนดิบ	คะแนนที่ปกติของแบบทดสอบ			คะแนนดิบ
	ทักษะทางคณิตศาสตร์	อนุกรม	เหตุผลทางคณิตศาสตร์	
16	48	50	54	16
15	47	48	53	15
14	46	47	51	14
13	44	46	50	13
12	43	45	48	12
11	42	44	47	11
10	41	42	45	10
9	39	40	43	9
8	37	39	41	8
7	34	37	39	7
6	32	35	37	6
5	30	33	34	5
4	28	30	31	4
3	25	27	28	3
2	22	17	23	2
1	18	14	18	1

จากตารางที่ 11 พบว่า แบบทดสอบวัดทักษะทางคณิตศาสตร์มีช่วงคะแนนที่ปกติอยู่ระหว่าง T18-T82 แบบทดสอบอนุกรมมีช่วงคะแนนที่ปกติอยู่ระหว่าง T14-T79 และแบบทดสอบวัดเหตุผลทางคณิตศาสตร์มีคะแนนที่ปกติอยู่ระหว่าง T18-T80 ถ้านักเรียนคนหนึ่งได้คะแนนดิบ 16 คะแนนจากการทำแบบทดสอบวัดทักษะทางคณิตศาสตร์ แบบทดสอบอนุกรม และแบบทดสอบวัดเหตุผลทางคณิตศาสตร์ ซึ่งตรงกับ T48, T50 และ T54 ตามลำดับ หมายความว่านักเรียนคนนี้ทำคะแนนได้ต่ำกว่าครึ่งหนึ่ง เท่ากับครึ่งหนึ่ง และสูงกว่าครึ่งหนึ่งตามลำดับ เมื่อเทียบกับเกณฑ์ปกติ