

ผลการวิเคราะห์ข้อมูล

ผลการวิเคราะห์ข้อมูลตามแบบแผนที่กำหนดในบทที่ ๓ และดำเนินการตามระเบียบวิธีการ
เทียบมาตรา แล้วเปรียบเทียบผลของการเทียบมาตราจากการวิเคราะห์ความคลาดเคลื่อนมาตรฐาน
ของการเทียบมาตรา และวิเคราะห์กลุ่มสอบทานผล ทั้งเสนอตามลำดับ ดังนี้

ผลการจัดเตรียมแบบสอบเทียบมาตรา และกลุ่มตัวอย่างผู้สอบ

ผลการดำเนินการเทียบมาตรา

ผลการวิเคราะห์ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรา

ผลของการวิเคราะห์กลุ่มสอบทานผล

ผลการจัดเตรียมแบบสอบเทียบมาตรา และกลุ่มตัวอย่างผู้สอบ

๑. ผลการจัดเตรียมแบบสอบเทียบมาตรา

๑.๑ แบบสอบ

แบบสอบที่นำมาใช้ในการวิจัยครั้งนี้มี ๒ กรณี กรณีแรก คือ แบบสอบคัดเลือก
เป็นแบบสอบวัดความสามารถทั่วไปซึ่งเป็นแบบสอบชนิดเลือกตอบที่มีตัวเลือกสี่ตัว จำนวน ๑๐๐ ข้อ
วัดองค์ประกอบความสามารถรวม ๕ องค์ประกอบ ใช้เวลาในการทดสอบ ๒ ชั่วโมง ใช้ทดสอบผู้
สมัครเข้าทำงานประจำปี พ.ศ. ๒๕๒๓ จำนวน ๑๕๘๕ คน

ส่วนกรณีที่สอง คือ แบบสอบวัดผลสัมฤทธิ์ เป็นแบบสอบวัดผลสัมฤทธิ์ตามหลักสูตร
ระดับปริญญาตรีวิชา สถิติวิจัยและวัดผลการศึกษา ซึ่งเป็นแบบสอบชนิดเลือกตอบที่มีตัวเลือกห้าตัว
จำนวน ๑๐๐ ข้อ เนื้อหาที่ทำการวัดแบ่งเป็น ๓ ส่วน ครอบคลุมลักษณะวิชา ใช้เวลาในการทดสอบ
๓ ชั่วโมง ใช้ทดสอบเมื่อสิ้นสุดภาคการเรียนเมื่อปีการศึกษา ๒๕๒๓ มีผู้ลงทะเบียนเรียน และเข้า
รับการทดสอบทั้งสิ้น ๑๒๓๕๓ คน จากการวิเคราะห์ข้อมูลประชากรทั้งหมด ได้ค่าสถิติพื้นฐานของแบบสอบ

ทั้งแสดงในตารางที่ ๑ ดังนี้

ตารางที่ ๑ ค่าสถิติพื้นฐานของแบบสอบถามคัดเลือก และแบบสอบถามวัดผลสัมฤทธิ์

ค่าสถิติ	\bar{X}	SD	ρ	MIN	MAX	K	r_{tt}^*	SEmeas
<u>แบบสอบถามคัดเลือก</u>								
แบบสอบถามทั้งหมด	60.603	9.3.09	0	46.0	900	900	.274	6.666
ค่าเฉลี่ยรายข้อ	0.606	0.930	0.930	0.436	0.736			
ความแปรปรวนรายข้อ	0.226	0.026	0.997	0.250				
<u>แบบสอบถามวัดผลสัมฤทธิ์</u>								
แบบสอบถามทั้งหมด	69.92	9.9.92b	9.0	46.0	900	900	.262	6.366
ค่าเฉลี่ยรายข้อ	0.699	0.999	0.96b	0.436				
ความแปรปรวนรายข้อ	0.22b	0.02b	0.97a	0.250				

* สัมประสิทธิ์ α

๑.๒ แบบสอบถามเทียบมาตรา

การจัดเตรียมแบบสอบถามเทียบมาตราเพื่อใช้ในการวิจัยครั้งนี้ ได้ดำเนินการสุ่ม
 ความเห็นก่อนเพื่อให้ได้แบบสอบถาม X และแบบสอบถาม Y โดยให้แบบสอบถามทั้งสองชุดทำหน้าที่วัดเนื้อหา
 หรือความสามารถเชิงจิตวิทยาที่ใกล้เคียงกันดังกล่าวไว้ในบทที่ ๓ ผลการเตรียมแบบสอบถามเทียบมาตรา
 และแบบสอบถามรวมของกรณีแบบสอบถามคัดเลือก แสดงในตารางที่ ๒ (ก) ส่วนกรณีแบบสอบถามวัดผลสัมฤทธิ์
 แสดงในตารางที่ ๒ (ข)

ตารางที่ ๒ (ก) ค่าสถิติพื้นฐาน และความเที่ยงของแบบสอบเทียบมาตรา และแบบสอบรวม กรณีแบบสอบคัดเลือก

ค่าสถิติ	K	\bar{Y}	SD	MAX	MIN	$\sum \frac{S_g^2}{S_x^2}$	$r_{tt}(\alpha)$	หมายเหตุ
แบบสอบเทียบมาตรา								
ชุด X	๗๕	๑๕.๗๘๕	๖.๑๐๕	๓๐.๐	๑.๐	.๒๑๓	.๘๐๕	*
Item mean (\bar{Y}_g)		๐.๕๐๕	๐.๑๑๖	๐.๖๕๓	๐.๓๑๒			
Item variance (S_g^2)		๐.๒๓๔	๐.๐๓๓	๐.๒๕๐	๐.๑๖๗			
Inter-item (r_{gt}) correlation		๐.๑๐๖	๐.๐๓๓	๐.๕๐๓	-๐.๐๕๕			
ชุด Y	๗๕	๑๕.๗๐๕	๕.๕๖๓	๓๐.๐	๒.๐	.๒๖๖	.๗๕๖	**
Item mean (\bar{Y}_g)		๐.๕๑๔	๐.๑๒๕	๐.๗๖๕	๐.๑๕๕			
Item variance (S_g^2)		๐.๒๓๗	๐.๐๓๕	๐.๒๕๐	๐.๑๗๕			
Inter-item (r_{gt}) correlation		๐.๐๗๕	๐.๐๖๕	๐.๕๑๕	-๐.๑๐๕			
แบบสอบรวม								
V60	๗๖	๔.๓๗๖	๓.๖๕๗	๒๐.๐	๐.๐	.๓๕๔	.๖๕๖	***
Item mean (\bar{Y}_{V60g})		๐.๓๕๖	๐.๑๑๓	๐.๖๓๕	๐.๑๕๗			
Item variance (S_g^2)		๐.๒๓๓	๐.๐๓๓	๐.๒๕๐	๐.๑๗๓			
Inter-item (r_{gt}) correlation		๐.๐๕๕	๐.๐๕๓	๐.๕๐๖	-๐.๐๖๕			
V40	๑๕	๖.๖๑๔	๔.๑๖๗	๑๕.๐	๐.๐	.๓๖๓	.๖๓๗	***
Item mean (\bar{Y}_{V40g})		๐.๕๑๕	๐.๐๗๕	๐.๖๕๓	๐.๓๕๕			
Item variance (S_g^2)		๐.๒๓๓	๐.๐๑๕	๐.๒๕๐	๐.๑๕๗			
Inter-item (r_{gt}) correlation		๐.๑๕๖	๐.๑	๐.๕๐๖	-๐.๐๓๓			
V20	๗	๓.๓๑๔	๑.๕๖๕	๗.๐	๐.๐	.๕๓๖	.๖๖๓	***
Item mean (\bar{Y}_{V20g})		๐.๕๗๖	๐.๑๒๓	๐.๖๕๓	๐.๓๖๐			
Item variance (S_g^2)		๐.๒๓๖	๐.๐๑๐	๐.๒๕๐	๐.๒๓๕			
Inter-item (r_{gt}) correlation		๐.๒๓๕	๐.๐๗๗	๐.๓๗๗	๐.๑๐๕			
แบบสอบเทียบมาตราที่รวมแบบสอบรวม								
XV60	๕๖	๒๓.๖๗๕	๕.๑๕๕	๕๕.๐	๓.๐	.๔๖๖	*	
YV60	๕๖	๒๓.๒๕๖	๔.๖๖๕	๕๑.๐	๑.๐	.๕๐๕	**	
XV40	๕๕	๒๐.๕๗๕	๔.๗๖๕	๕๓.๐	๒.๐	.๔๖๕	*	
YV40	๕๕	๒๑.๐๖๕	๔.๑๕๕	๕๕.๐	๑.๐	.๕๕๕	**	
XV20	๕๓	๑๗.๗๐๕	๓.๖๖๖	๔๕.๐	๑.๐	.๕๕๓	*	
YV20	๕๓	๑๘.๑๕๖	๓.๐๕๕	๔๕.๐	๒.๐	.๕๓๖	**	

หมายเหตุ หมายเหตุที่ใช้วิเคราะห์ค่าสถิติ: * กลุ่มตัวอย่าง α , ** กลุ่มตัวอย่าง β , *** กลุ่มตัวอย่างรวม α และ β

ตารางที่ ๒ (ข) ค่าสถิติพื้นฐาน และความเที่ยงของแบบสอบถามเปรียบเทียบมารา และแบบสอบถาม กรณีสอบวัดคุณสมบัติ

ค่าสถิติ	K	\bar{X}	SD	MAX	MIN	$\leq \frac{S_g^2}{S_x^2}$	$r_{tt}(\alpha)$	หมายเหตุ
แบบสอบถามมารา								
ชุด X	๗๕	๑๗.๗๖๕	๕.๕๗๓	๓๒.๐	๑.๐	.๓๗๘	.๖๕๑	*
Item mean (\bar{X}_g)		๐.๕๕๖	๐.๒๓๑	๐.๘๗๐	๐.๑๑๓			
Item variance (S_g^2)		๐.๒๓๒	๐.๐๕๕	๐.๒๕๐	๐.๑๑๐			
Inter-item (r_{gt}) correlation		๐.๐๕๖	๐.๐๕๕	๐.๑๗๘	-๐.๐๖๕			
ชุด Y	๗๕	๑๖.๐๐๕	๕.๘๓๓	๓๕.๐	๐.๐	.๓๗๑	.๗๕๘	**
Item mean (\bar{Y}_g)		๐.๕๕๗	๐.๒๑๕	๐.๘๑๗	๐.๐๕๓			
Item variance (S_g^2)		๐.๒๐๑	๐.๐๓๗	๐.๒๕๘	๐.๐๕๓			
Inter-item (r_{gt}) correlation		๐.๐๖๖	๐.๐๕๓	๐.๑๓๕	-๐.๐๖๑			
แบบสอบถามรวม								
V60	๖๑	๕.๒๖๑	๓.๐๕๕	๒๐.๐	๐.๐	.๕๕๘	.๕๗๖	***
Item mean (\bar{X}_{V60g})		๐.๕๕๑	๐.๒๓๓	๐.๘๐๕	๐.๐๓๕			
Item variance (S_g^2)		๐.๑๕๘	๐.๐๕๕	๐.๒๕๕	๐.๐๓๕			
Inter-item (r_{gt}) correlation		๐.๐๕๗	๐.๐๕๑	๐.๒๗๕	-๐.๐๕๕			
V40	๑๕	๖.๕๑๕	๒.๓๗๕	๑๕.๐	๐.๐	.๕๓๕	.๕๕๘	***
Item mean (\bar{X}_{V40g})		๐.๕๕๘	๐.๒๑๖	๐.๘๐๕	๐.๐๓๕			
Item variance (S_g^2)		๐.๒๐๕	๐.๐๕๕	๐.๒๕๕	๐.๐๓๕			
Inter-item (r_{gt}) correlation		๐.๐๖๖	๐.๐๕๓	๐.๑๕๐	-๐.๐๑๑			
V20	๗	๒.๕๐๕	๑.๕๓๓	๗.๐	๐.๐	.๖๕๓	.๕๐๕	***
Item mean (\bar{X}_{V20g})		๐.๕๑๕	๐.๑๖๕	๐.๖๕๑	๐.๑๕๕			
Item variance (S_g^2)		๐.๒๑๕	๐.๐๓๗	๐.๒๕๕	๐.๑๕๕			
Inter-item (r_{gt}) correlation		๐.๐๕๕	๐.๐๓๖	๐.๑๕๐	๐.๐๓๗			
แบบสอบถามมาราที่รวมแบบสอบถามรวม								
XV60	๕๖	๒๖.๕๖๕	๖.๘๘๓	๕๕.๐	๘.๐	๐.๕๖๖	*	
YV60	๕๖	๒๕.๕๑๓	๗.๕๑๕	๕๖.๐	๕.๐	๐.๕๖๕	**	
XV40	๕๕	๒๓.๖๕๓	๖.๑๕๕	๕๑.๐	๗.๐	๐.๕๕๓	*	
YV40	๕๕	๒๒.๕๑๓	๖.๗๖๖	๕๑.๐	๖.๐	๐.๕๖๕	**	
XV20	๕๖	๒๐.๑๕๕	๕.๕๑๑	๓๕.๐	๓.๐	๐.๕๓๓	*	
YV20	๕๖	๑๘.๕๕๗	๕.๕๕๕	๓๕.๐	๑.๐	๐.๕๕๕	**	

หมายเหตุ ข้อมูลที่ใช้วิเคราะห์ค่าสถิติ: * กลุ่มตัวอย่าง α , ** กลุ่มตัวอย่าง β , *** กลุ่มตัวอย่างรวม α และ β

จากตารางที่ ๒ (ก) ค่าสถิติที่เกี่ยวข้องกับแบบสอบชุด X ชุด Y ตลอดจนแบบสอบรวมอีก ๓ ขนาด ของกรณีแบบสอบคัดเลือก ได้จากการใช้ข้อมูลของกลุ่มตัวอย่าง α และ β ตามแผนของการออกแบบการเก็บข้อมูล พบว่า ค่าเฉลี่ยของแบบสอบเทียบมาตราทั้งสองชุด (\bar{X}_x, \bar{X}_y) และค่าเฉลี่ยของรายข้อของชุด X และชุด Y $(\bar{X}_{xg}, \bar{X}_{yg})$ มีค่าใกล้เคียงกัน จึงอาจกล่าวได้ว่า แบบสอบทั้งสองชุดมีความยากพอ ๆ กัน แต่สัดส่วนระหว่างผลรวมความแปรปรวนรายข้อกับความแปรปรวนของแบบสอบทั้งสองชุด $(\sum s_g^2 / s_x^2)$ ของชุด X มีค่าน้อยกว่าของชุด Y ซึ่งมีผลโดยตรงต่อค่าความเที่ยง คือ ความเที่ยงของชุด X สูงกว่าความเที่ยงชุด Y เหตุการณ์เช่นนี้จัดว่าเป็นสิ่งปกติเมื่อแบบสอบสองชุดต่างก็ประกอบด้วยข้อสอบคนละข้อ และทวนความเป็นปกติเช่นนี้ที่ทำให้จำเป็นของอาศัยเทคนิคการเทียบมาตราเพื่อให้คะแนนแบบสอบสองชุดเปรียบเทียบกันได้โดยตรง

พิจารณาผลการเตรียมแบบสอบรวม ๓ ขนาด จากการให้ฉบับที่สั้นที่สุด ๗ ข้อ (V20) เป็นขั้วเรีของฉบับ ๑๔ ข้อ (V40) และต่างก็เป็นขั้วเรีของฉบับ ๒๑ ข้อ (V60) พบว่า ข้อสอบในส่วนที่เพิ่มขึ้นจากชุดที่น้อยกว่ามีลักษณะแตกต่างกันทั้งค่าเฉลี่ย และค่าอำนาจจำแนก (ภาพแสดงการแจกแจงของข้อสอบรายข้อในชุดของแบบสอบรวม แสดงในภาคผนวก ก. แผนภาพที่ ๑๖) จึงมีผลทำให้ค่าความเที่ยงไม่ได้เพิ่ม และลดความสัมพันธ์ของจำนวนข้อสอบในทฤษฎีว่าด้วยการทดสอบแบบดั้งเดิม (Classical test theory) ผลจากการสุ่มข้อสอบในแต่ละขนาดของแบบสอบรวม พบว่า V40 มีค่าความเที่ยงสูงกว่า V60 และ V20 มีค่าความเที่ยงต่ำที่สุด

หลังจากการจัดข้อสอบในชุด X Y และแบบสอบรวมแล้ว ได้ทำการจับคู่ระหว่างแบบสอบเทียบมาตรา และแบบสอบรวมเพื่อใช้ในการวิจัยต่อไป มีทั้งหมด ๓ คู่ คือ XV60-YV60 XV40-YV40 และ XV20-YV20 ค่าความเที่ยงของแบบสอบทั้ง ๓ คู่ มีค่าสูงสุด .๘๗ และต่ำสุด .๘๒ สำหรับคู่ที่มีแบบสอบรวม V40 ยังคงพบว่า มีค่าความเที่ยงสูงกว่าคู่ที่มีแบบสอบรวม V60 เล็กน้อย ซึ่งเป็นผลสืบเนื่องจากแบบสอบรวมทั้งสามมาแล้วข้างต้น

ตารางที่ ๒ (ข) ค่าสถิติบรรยายแบบสอบต่าง ๆ ของกรณีแบบสอบวัดผลสัมฤทธิ์
ได้มาในทำนองเดียวกับตารางที่ ๒ (ก) ของกรณีแบบสอบคัดเลือก พบว่า ค่าเฉลี่ยของแบบสอบ
เทียบมาตรฐาน X (\bar{X}_X) มีค่าน้อยกว่า Y (\bar{X}_Y) เล็กน้อย สัดส่วนของผลรวมของความ
แปรปรวนรายข้อกับความแปรปรวนทั้งหมด ($\leq s_g^2 / s_x^2$) ของ X มีค่าสูงกว่าของ Y
ทำให้ค่าความเที่ยง X น้อยกว่า Y

สำหรับแบบสอบรวมสามขนาด ได้ค่าเฉลี่ยของค่าเฉลี่ยรายข้อใกล้เคียงกัน
คือ ประมาณ ๐.๔ ค่าความเที่ยงของทั้งสามขนาด คือ .๔๔ .๕๓ และ .๖๕ ตามลำดับขนาดของ
แบบสอบรวม V20 V40 และ V60 ตามลำดับ (ข้อสอบในแต่ละขนาดมีการกระจายคังแผน
ภาพที่ ๑๗ ในภาคผนวก ก.)

แบบสอบแต่ละคู่ที่นำมาใช้วิเคราะห์ในการวิจัย ได้จากการจับคู่ระหว่างแบบ
สอบเทียบมาตรา และแบบสอบรวมเช่นเดียวกับกรณีแบบสอบคัดเลือก

๒. ผลการวิเคราะห์ข้อมูลกลุ่มตัวอย่างผู้สอบ

๒.๑ กรณีแบบสอบคัดเลือก กลุ่มตัวอย่างผู้สอบได้จากการสุ่มอย่างง่ายจากประชากร
๑๕๕๕ คน โดยให้กลุ่ม α มีจำนวน ๑๕๐๐ คน และกลุ่ม β มีจำนวน ๑๕๐๐ คน เช่นเดียวกัน
โดยไม่มีหน่วยตัวอย่างซ้ำกัน แต่ละหน่วยตัวอย่างมีข้อมูลคะแนนรายข้อจากการทดสอบ จำนวน ๑๐๐
ข้อ ภายหลังจากการออกแบบให้กลุ่มตัวอย่าง α ทำแบบสอบชุด X และแบบสอบรวม V ส่วนกลุ่มตัวอย่าง β
ทำแบบสอบชุด Y และแบบสอบรวม V การประมวลข้อมูลคะแนนของแต่ละหน่วยตัวอย่าง ให้มีคะแนนตาม
แบบสอบเทียบมาตราชุดต่าง ๆ และวิเคราะห์ข้อมูลของกลุ่มตัวอย่างด้วยโปรแกรม SPSSx ผลปรากฏ
ในตารางที่ ๓

จุฬาลงกรณ์มหาวิทยาลัย

กรณีแบบสอชคักเลือก กลุ่ม α และกลุ่ม β มีคะแนนต่ำสุดในแต่ละแบบสอชคักตั้งแต่ ๐ จนถึง ๔ ไม่มีคะแนน ๐ ปรากฏอยู่ แสดงว่า ผลการสุ่มใช้ได้ และเหมาะที่จะวิเคราะห์ผล การตอบตามทฤษฎีการตอบข้อสอบซึ่งมีข้อกำหนดว่า ทฤษฎีนี้จะไม่สามารถหาค่าประมาณพารามิเตอร์ หากมีผู้ตอบที่ไม่ได้คะแนนจากการสอชคัก (Lord 1977) พิจารณาในด้านการแจกแจงของกลุ่ม และกลุ่ม ในแต่ละคู่ของแบบสอชคักเทียบมาตรา พบว่า มีความใกล้เคียงกัน โดยที่กลุ่ม α ทำคะแนน ได้สูงกว่ากลุ่ม β ทุกคู่ของแบบสอชคักประมาณ ๐.๕ คะแนน และส่วนเบี่ยงเบนมาตรฐานกลุ่ม α มีค่าน้อย กว่า β ประมาณ ๐.๕ คะแนน เช่นกัน ผลจากการสุ่มตัวอย่างผู้สอชคัก และแบบสอชคักเทียบมาตราตามที่ ปรากฏนี้ คล้ายกับสภาพการทดสอบที่จัดขึ้น ๒ ครั้ง กับกลุ่มผู้สอชคักที่มีลักษณะคล้ายคลึงกัน โดยใช้ แบบสอชคักที่วัดความสามารถอย่างเดียวกัน ต่างกันเพียงว่า รายชื่อในแบบสอชคักที่สร้างขึ้นมาต่างครั้ง ซึ่งมีผลทำให้แบบสอชคักครั้งแรกมีความยากมากกว่าแบบสอชคักชุดหลัง เหตุการณ์เช่นนี้ เป็นสิ่งที่มักพบเสมอ ในทางปฏิบัติ และเป็นสิ่งที่ทำให้การเทียบมาตรามีความจำเป็นต่อการรายงานผลการสอชคัก เพื่อให้ คะแนนสอชคักต่างครั้งสามารถเปรียบเทียบกันได้



ศูนย์วิจัยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ ๓ สถิติบรรยายกลุ่มตัวอย่าง α และ β ในการทอบแบบสอบถามต่าง ๆ

ค่าสถิติ	กลุ่มตัวอย่าง α			กลุ่มตัวอย่าง β		
	XV60	XV40	XV20	YV60	YV40	YV20
กรณีแบบสอบถามคัดเลือก						
MEAN	๒๓.๖๗๘	๒๐.๕๓๕	๑๗.๗๐๘	๒๓.๒๕๑	๒๑.๐๖๘	๑๘.๑๘๖
SD	๕.๑๘๘	๔.๗๖๘	๓.๖๖๖	๔.๖๒๘	๔.๑๘๘	๓.๐๘๘
SKEWNESS	-๐.๒๓๘	-๐.๒๘๓	-๐.๒๘๐	-๐.๑๘๘	-๐.๒๐๖	-๐.๑๗๗
KURTOSIS	-๑.๑๓๐	-๑.๑๘๕	-๑.๑๒๘	-๑.๐๘๘	-๑.๑๐๖	-๑.๐๖๘
MAX	๔๕.๐	๔๓.๐	๓๘.๐	๕๑.๐	๔๕.๐	๓๘.๐
MIN	๓.๐	๒.๐	๑.๐	๑.๐	๑.๐	๒.๐
MODE	๒๕.๐	๒๗.๐	๒๑.๐	๒๗.๐	๒๖.๐	๒๓.๐
กรณีแบบสอบถามวัดผลสัมฤทธิ์						
MEAN	๒๖.๕๒๕	๒๓.๖๘๓	๒๐.๑๘๘	๒๕.๘๑๓	๒๓.๕๑๓	๑๘.๕๘๗
SD	๖.๘๘๓	๖.๑๘๕	๕.๕๑๑	๗.๘๑๘	๖.๗๑๓	๕.๕๘๘
SKEWNESS	๐.๓๖๖	๐.๓๗๒	๐.๒๗๘	๐.๕๕๘	๐.๖๐๗	๐.๖๐๓
KURTOSIS	-๐.๑๑๗	-๐.๑๒๕	-๐.๑๑๘	-๐.๑๒๐	-๐.๐๒๕	-๐.๐๖๑
MAX	๔๕.๐	๔๑.๐	๓๕.๐	๕๓.๐	๔๑.๐	๓๕.๐
MIN	๔.๐	๓.๐	๓.๐	๕.๐	๓.๐	๒.๐
MODE	๒๕.๐๐๐	๒๓.๐๐๐	๑๘.๐๐๐	๒๖.๐๐๐	๑๕.๐๐๐	๑๖.๐๐๐
MEDIAN	๒๖.๐๐๐	๒๓.๐๐๐	๒๐.๐๐๐	๒๕.๐๐๐	๒๓.๐๐๐	๑๘.๐๐๐

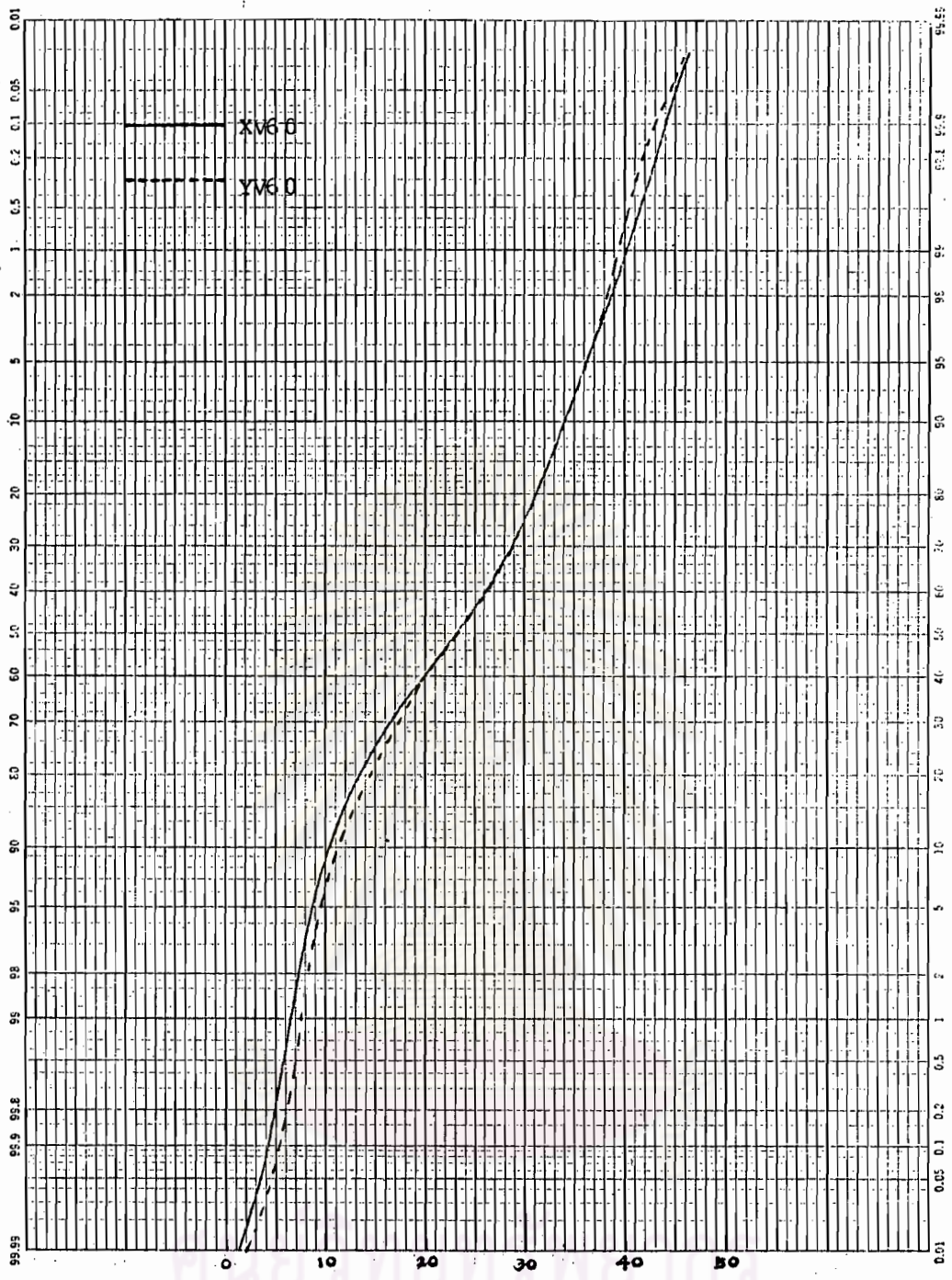
๒.๒ กรณีแบบสอมวัคผลสัมฤทธิ์ กลุ่มตัวอย่างผู้สอบได้จากการสุ่มอย่างง่ายจาก ประชากร ๑๒๕๓ คน โดยให้กลุ่ม α และกลุ่ม β กลุ่มละ ๑๕๐๐ คน การดำเนินการออกแบบ และประมวลผลข้อมูล เป็นทำนองเดียวกับกรณีแบบสอมคักเลือก ค่าสถิติพื้นฐานที่บรรยายกลุ่มตัวอย่าง แสดงในตารางที่ ๓ พบว่า คะแนนค่าสุคในกลุ่ม α และกลุ่ม β ในแต่ละแบบสอมมีค่ามากกว่า ๐ แสดงว่า ผลการสุ่มใช้ได้ และเหมาะสมสำหรับการประมาณค่าพารามิเตอร์ความสามารถผู้สอบด้วย ทฤษฎีการทอมซอสอม เช่นเดียวกับกรณีแบบสอมคักเลือก พิจารณาในค่านการแจกแจงของกลุ่ม α และกลุ่ม β พบว่า โดยทั่วไปกลุ่มตัวอย่าง α ที่ทำแบบสอม X ได้คะแนนสูงกว่ากลุ่ม β ที่ทำแบบสอม Y อยู่ประมาณ ๑.๑ คะแนน แต่ส่วนเบี่ยงเบนมาตรฐานของกลุ่ม α มีค่าต่ำกว่ากลุ่ม β ประมาณ ๐.๕ ซึ่งหมายความว่า โดยเฉลี่ยแบบสอมชุด X เป็นชุดที่ง่ายกว่าแบบสอมชุด Y เล็กน้อย

ผลการดำเนินการเทียบมาตรา

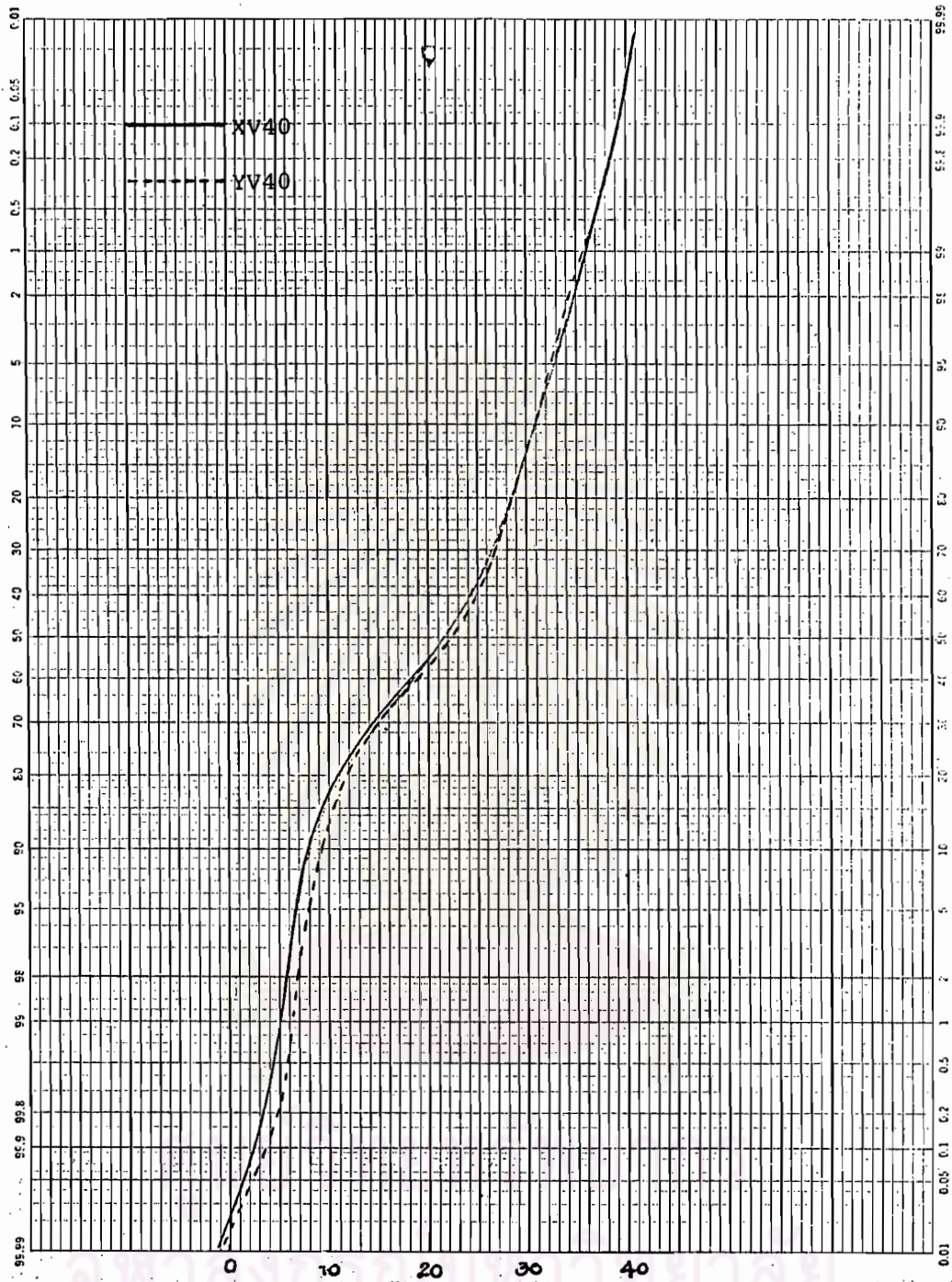
๑. ผลการเทียบมาตรารูปแบบอิกวิเปอร์เซนไทล์

การเทียบมาตรารูปแบบอิกวิเปอร์เซนไทล์ ได้ดำเนินการตามขั้นตอนที่กล่าวไว้ใน บทที่ ๓ คือ เริ่มตั้งแต่ทำตารางประมาณการแจกแจงความถี่ของกลุ่มตัวอย่างรวม แล้วนำมาหา ค่าแห่ง เปอร์ เซนไทล์ของแต่ละชั้นคะแนน ใช้โปรแกรมฟอร์แทรนช่วยในการคำนวณ (ตัวอย่าง โปรแกรมแสดงในภาคผนวก ค. โปรแกรมที่ ๑) จากนั้นจึงนำมาเขียนกราฟแสดงความสัมพันธ์ของ คะแนนแบบสอมชุด X กับค่าแห่ง เปอร์ เซนไทล์ และคะแนนแบบสอมชุด Y กับค่าแห่ง เปอร์ เซนไทล์ เขียนลงในกระดาษกราฟชนิดที่เป็น Arithmetic probability graph ปรับเส้นกราฟ ของคะแนนชุด X และชุด Y ให้เรียบด้วยมือ ผลการดำเนินการตามที่กล่าวมานี้ แสดงได้เป็นแผน ภาพเส้นกราฟของแบบสอมคู่ของแบบสอมแต่ละกรณี ดังนี้

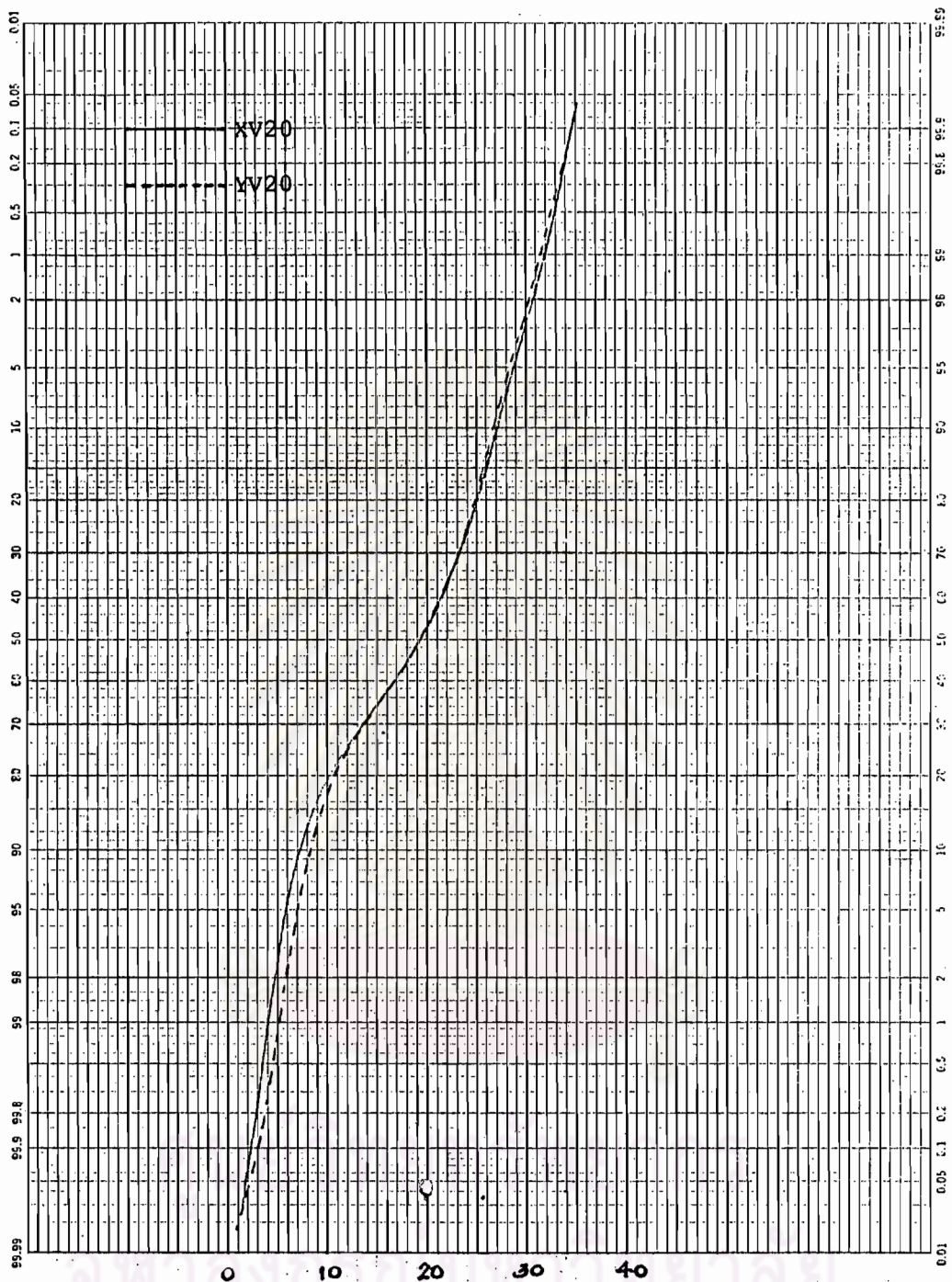
๑.๑ กรณีแบบสอมคักเลือก ผลจากการเขียนกราฟแสดงความสัมพันธ์ของค่าแห่ง เปอร์ เซนไทล์กับคะแนนแบบสอมชุด X กับชุด Y แสดงไว้ ดังนี้ ชุด XV60 YV60 แสดงในแผน ภาพที่ ๔ (ก) ชุด XV40 YV40 แสดงในแผนภาพที่ ๔ (ข) ชุด XV20 YV20 แสดงใน แผนภาพที่ ๔ (ค) จากภาพเส้นกราฟของแบบสอมชุด X และชุด Y ในแต่ละคู่ อ่านคะแนนสมมูลของ ทั้งสองทีละค่า โดยกำหนดการอ่านจากชุด X ไปหาคะแนนชุด Y ที่ค่าแห่ง เปอร์ เซนไทล์เดียวกัน



แผนภาพที่ ๔ (ก) กราฟความถี่สะสมสัมพันธ์ของแบบสขย XV60 และ YV60
กรณีแบบสอบคัทเลือก



แผนภาพที่ ๘ (ข) กราฟความถี่สะสมสัมพันธ์ของแบบสอบ XV40 และ YV40
กรณีแบบสอบคัดเลือก



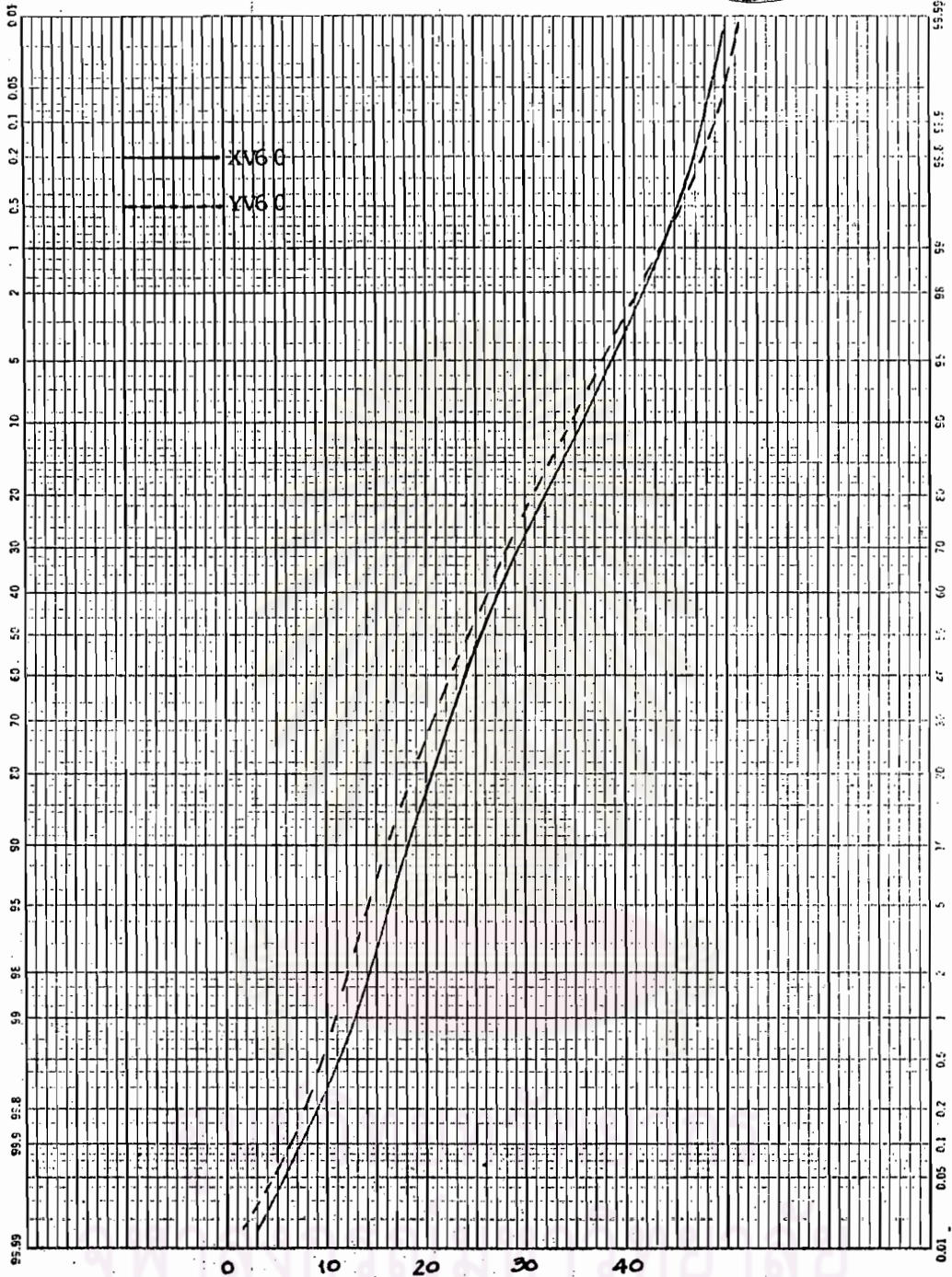
Probability Scale - 90 Divisions

แผนภาพที่ ๔ (ค) กราฟความถี่สะสมสัมพัทธ์ของแบบสอบ XV20 และ XV20
กรณีแบบสอบคัดเลือก

๑.๒ กรณีแบบสอบวัดผลสัมฤทธิ์ การดำเนินการ เป็นไปตามที่เกี่ยวกับกรณีแบบสอบ
คัดเลือก และผลของการเขียนเส้นกราฟแสดงในแผนภาพที่ ๕ (ก) ถึงแผนภาพที่ ๕ (ค) การจัดทำ
ตารางคะแนนสมมูลด้วยการอ่านค่าคะแนนจากแบบสอบชุด X ไปหาคะแนนชุด X ที่ตำแหน่งเปอร์เซ็นต์ไทล์
เดียวกัน ผลของการจัดทำตารางคะแนนสมมูลแสดงไว้ในตารางที่ ๒๑ ถึงตารางที่ ๒๓ ในภาคผนวก ข.



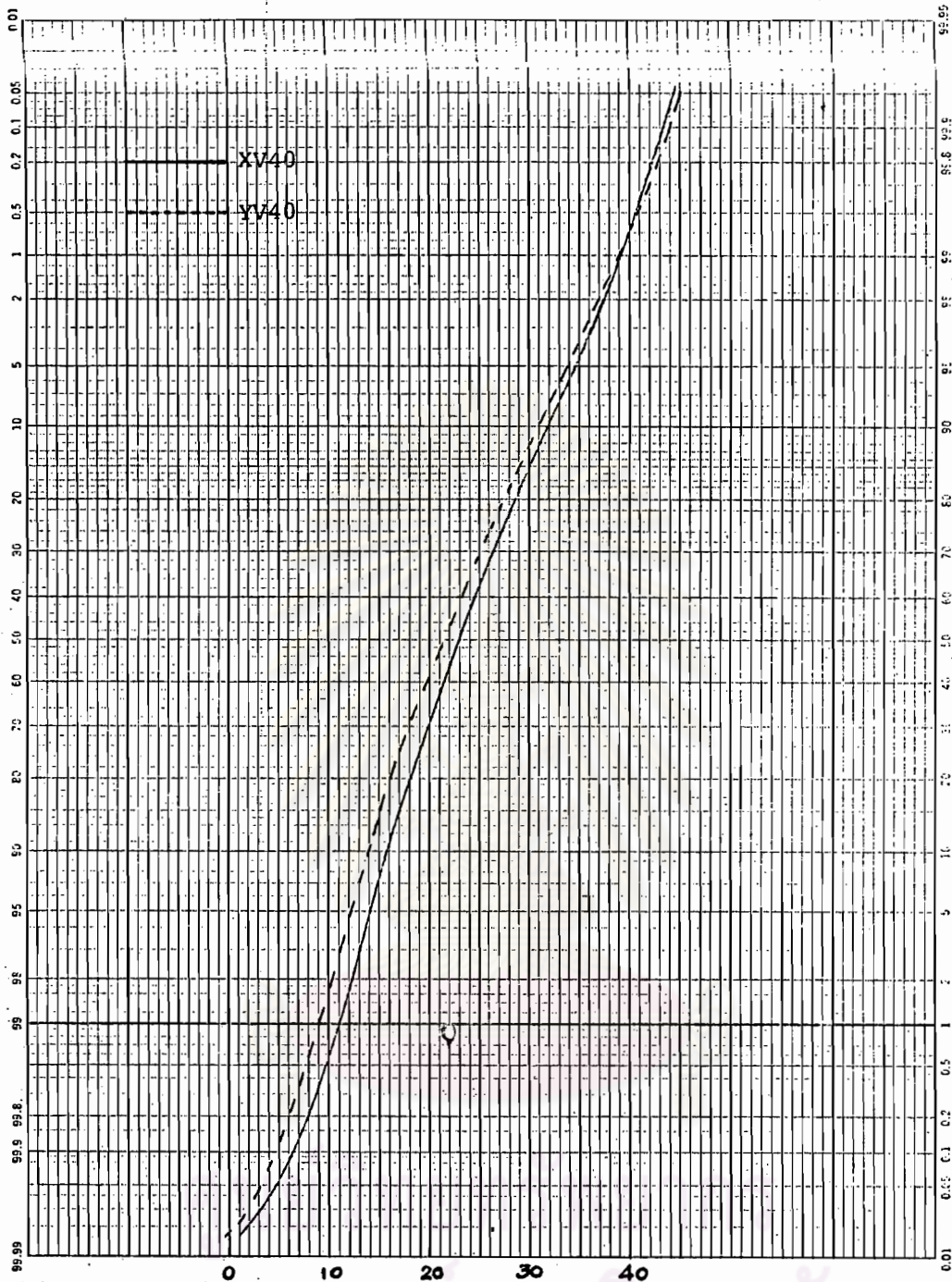
ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



Probability Scale - 99 Divisions

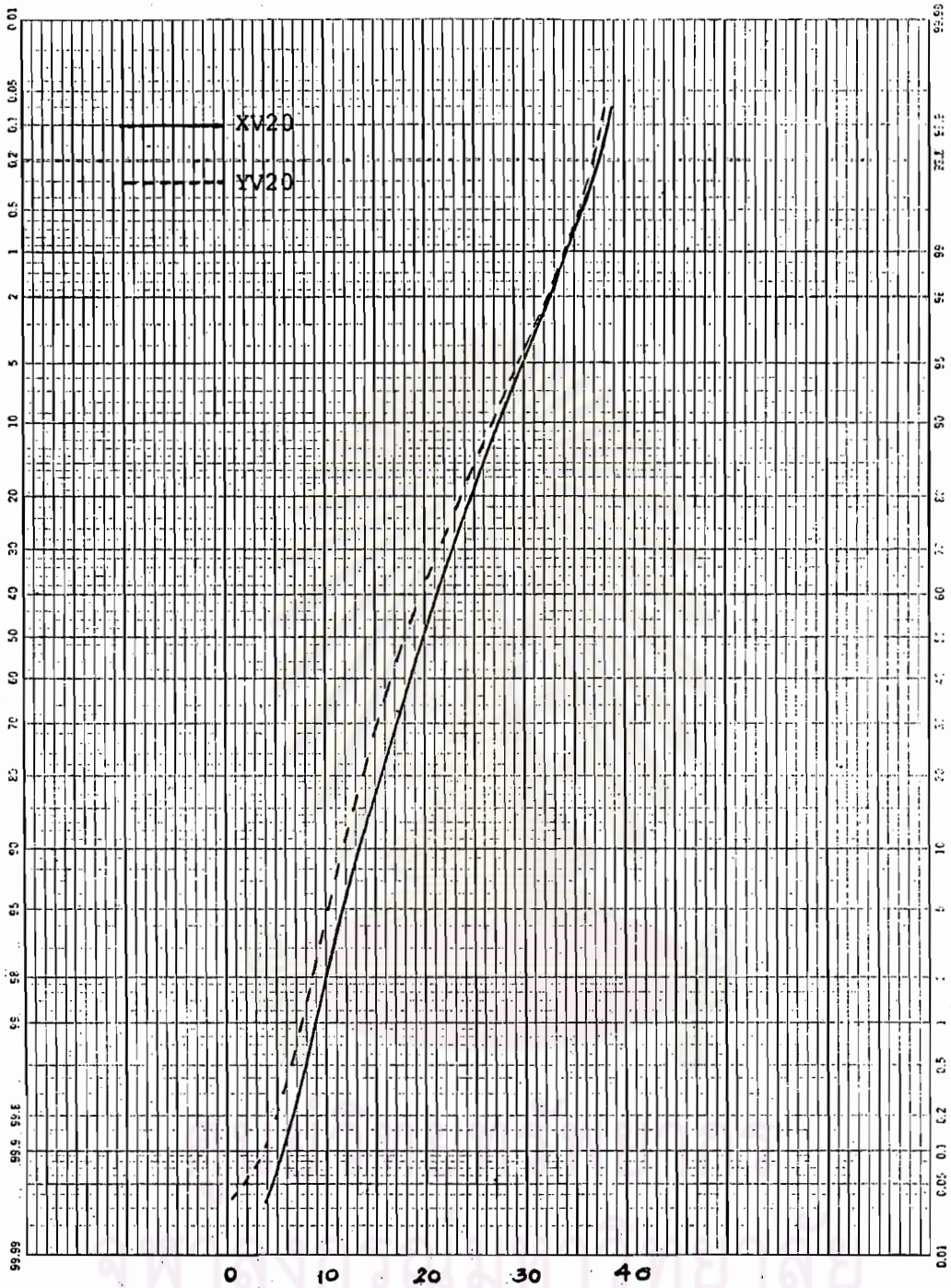
แผนภาพที่ ๕ (ก)

กราฟความถี่สะสมสัมพันธ์ของแบบสอบ XV60 และ YV60
กรณีแบบสอบวัดผลสัมฤทธิ์



Probability Scale x 90 Divisions

แผนภาพที่ ๕ (ข) กราฟความถี่สะสมสัมพันธ์ของแบบสอบ XV40 และ YV40
กรณีแบบสอบวัดผลสัมฤทธิ์



Probability Scale a 90 Divisions

แผนภาพที่ ๕ (ก) กราฟความถี่สะสมสัมพัทธ์ของแบบสอบ XV20 และ YV20
กรณีแบบสอบวัดผลสัมฤทธิ์

๒. รูปแบบการเทียบมาตรฐานเชิงเส้นตรง

การคำนวณค่าประชากรจากกลุ่มตัวอย่าง α และ β ด้วยวิธีเป็นไปไ้สูงสุด ใ้ผล
 คึงตารางที่ ๔ ส่วนค่าประมาณค่าความชัน (A) และค่าคงที่ (B) ของสมการเส้นตรง ใ้ผลคึง
 แสดงในตารางที่ ๕

ตารางที่ ๔ ค่าประมาณประชากรด้วยวิธีเป็นไปไ้สูงสุดจากกลุ่ม α และกลุ่ม β

สถิติ	กรณีแบบสอบค้เลือก			กรณีแบบสอบวัดผลสัมฤทธิ์		
	V60	V40	V20	V60	V40	V20
$\hat{\mu}_v$	๘.๓๓๓	๖.๑๖๘	๓.๓๑๘	๘.๐๐๐	๖.๘๓๓	๖.๘๖๖
$\hat{\sigma}_v^2$	๑๓.๓๓๖	๑๐.๐๓๐	๓.๘๓๘	๘.๖๖๖	๘.๖๖๘	๖.๖๖๖
$\hat{\mu}_x$	๖๖.๖๖๖	๖๐.๖๖๐	๖๘.๖๖๐	๖๖.๖๖๖	๖๖.๖๖๖	๖๖.๖๖๖
$\hat{\sigma}_x^2$	๘๘.๖๖๘	๖๖.๖๖๐	๖๖.๖๖๖	๖๖.๖๖๖	๖๖.๖๖๖	๖๖.๖๖๖
$\hat{\mu}_y$	๖๖.๖๖๖	๖๖.๖๖๐	๖๖.๖๖๖	๖๖.๖๖๖	๖๖.๖๖๖	๖๖.๖๖๖
$\hat{\sigma}_y$	๖๖.๖๖๖	๖๖.๖๖๖	๖๖.๖๖๖	๖๖.๖๖๖	๖๖.๖๖๖	๖๖.๖๖๖

ศูนย์วิทยทรัพยากร
 จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ ๕ ค่า A และ B ของสมการเทียบมาตรฐานเชิงเส้นตรง กรณีแบบสอบคัดเลือก และแบบสอบ
วัดผลสัมฤทธิ์

แบบสอบรวม	กรณีแบบสอบคัดเลือก		กรณีแบบสอบวัดผลสัมฤทธิ์	
	A	B	A	B
V60	๑.๐๖๘๘	→ ๑.๙๓๒๔	๐.๙๔๖๖	๒.๘๐๔๔
V40	๑.๐๗๔๔	→ ๑.๙๒๕๖	๐.๙๒๕๖	๒.๘๑๒๘
V20	๑.๐๘๕๓	→ ๑.๙๑๔๗	๐.๙๑๔๗	๒.๗๖๖๖

จากตารางที่ ๔ และ ๕ ค่าประมาณเหล่านี้เมื่อนำไปแทนลงในสมการเส้นตรงของ
การเทียบมาตรฐาน (สมการที่ ๑ ในบทที่ ๓) ผลที่ได้ คือ สมการเทียบมาตรฐานเชิงเส้นตรงของแต่ละ
คู่ของแบบสอบที่กองการ กังนั

๒.๑ กรณีแบบสอบคัดเลือก

$$x^*(V60) = (1.06885)y + (-1.93249) \text{ ---- (๑)}$$

$$x^*(V40) = (1.07449)y + (-1.99323) \text{ ---- (๒)}$$

$$x^*(V20) = (1.08543)y + (-1.21386) \text{ ---- (๓)}$$

สมการที่ ๑ ถึง ๓ ไ้้นำไปสร้างตารางคะแนนสมมูล แปลงคะแนนจาก
แบบสอบชุด Y ไปยังคะแนนของแบบสอบชุด X โดยแทนค่า Y ทั้งที่ ๑ จนถึงคะแนนเต็มของแต่ละชุด
ผลการเทียบคะแนนแสดงในตารางที่ ๑๘ ถึง ๒๐

๒.๒ กรณีแบบสอบวัดผลสัมฤทธิ์

$$x^*(v60) = (0.9466)y + 2.804 \text{ ----- (๔)}$$

$$x^*(v40) = (0.9381)y + 2.8159 \text{ ----- (๕)}$$

$$x^*(v20) = (10.9298)y + 2.7662 \text{ ----- (๖)}$$

สมการที่ ๔ ถึง ๖ ให้นำไปสร้างตารางคะแนนสมมูล เช่นเดียวกับกรณีแบบสอบคัดเลือก ผลปรากฏในตารางที่ ๒๑ ถึง ๒๓ ในภาคผนวก ข.

๓. รูปแบบการเทียบมาตรฐานอิงทฤษฎีการตอบข้อสอบ (IRT)

จากการเตรียมแบบสอบเทียบมาตรฐานในแต่ละกรณี ได้แบบสอบชุด X จำนวน ๓๕ ข้อ และชุด Y จำนวน ๓๕ ข้อ แบบสอบรวมขนาดยาวที่สุด คือ ๒๐ ข้อ เมื่อนำรวมจำนวนข้อสอบที่ใช้ในการวิจัยสูงสุด คือ ๕๐ ข้อ การประมาณค่าพารามิเตอร์ของข้อสอบตามรูปแบบการตอบข้อสอบโลจิสติกสามพารามิเตอร์ด้วยโปรแกรมสำเร็จรูป LOGIST5 จึงนำข้อสอบทั้ง ๕๐ ข้อ เข้าสู่การประมาณค่าพร้อมกันหมด ๕๐ ข้อ ด้วยคำสั่งครั้งเดียว ซึ่งมีผลทำให้ค่าประมาณของข้อสอบทุกคนในแต่ละกรณีของแบบสอบมีหน่วยขมามาตราเดียวกัน ผลการวิเคราะห์สรุปในตารางที่ ๒ ทั้งนี้ (รายละเอียดอยู่ในภาคผนวก ก. ตารางที่ ๑๕ และตารางที่ ๑๖)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ ๖ สรุปค่าประมาณพารามิเตอร์ข้อสอบของแบบสอบคัดเลือก และแบบสอบวัดผลสัมฤทธิ์

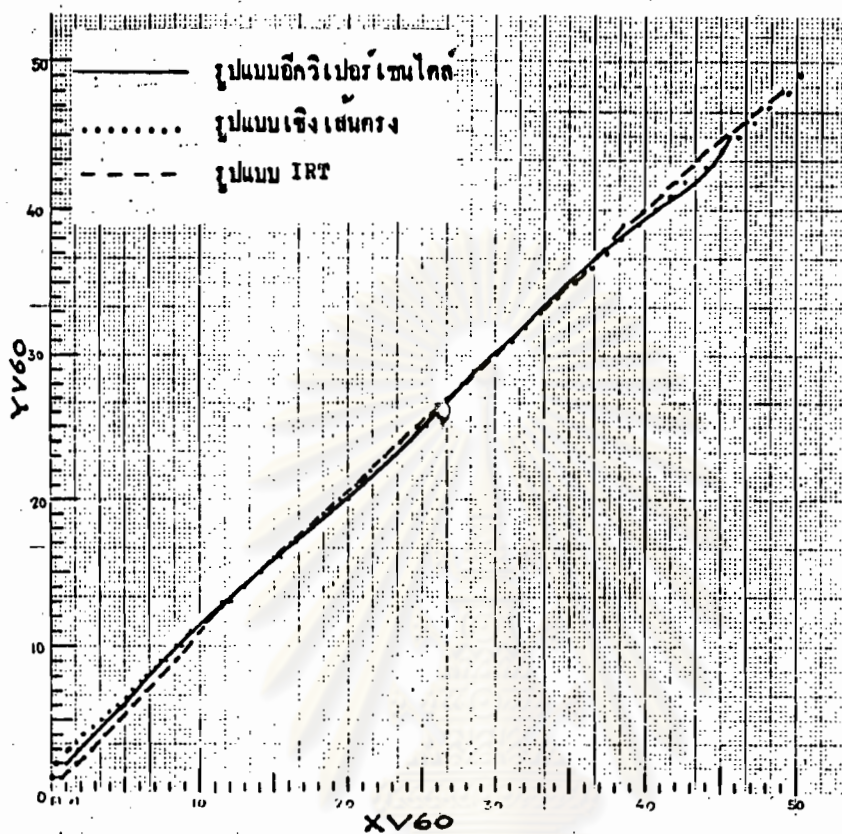
แบบสอบ	แบบสอบคัดเลือก						แบบสอบวัดผลสัมฤทธิ์					
	b		a		c		b		a		c	
	MAX	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX	MIN
XV60	๔.๕๓	-๐.๕๓	๒.๐๐	๐.๒๘	๐.๕๘	๐.๐๑	๓.๘๘	-๒.๕๓	๑.๑๕	๐.๓	๐.๓๓	๐.๐๑
YV60	๔.๕๓	-๒.๕๘	๒.๐๐	๐.๑๓	๐.๕๓	๐.๐๐	๔.๓๓	-๒.๒๘	๑.๒๕	๐.๑๓	๐.๓๘	๐.๐๐
XV40	๔.๕๓	-๐.๕๓	๑.๒๘	๐.๒๘	๐.๕๘	๐.๐๑	๓.๘๘	-๒.๕๓	๔.๑๕	๐.๑๓	๐.๓๘	๐.๐๑
YV40	๔.๕๓	-๒.๕๘	๑.๘๓	๐.๑๓	๐.๕๓	๐.๐๐	๔.๓๓	-๒.๒๘	๑.๒๕	๐.๑๓	๐.๓๘	๐.๐๐
XV20	๔.๕๓	-๐.๕๓	๑.๓๘	๐.๒๘	๐.๕๘	๐.๐๑	๓.๘๘	-๒.๕๓	๑.๑๕	๐.๑๓	๐.๓๘	๐.๐๑
YV20	๔.๕๓	-๒.๕๘	๑.๘๓	๐.๑๓	๐.๕๘	๐.๐๐	๔.๓๓	-๒.๕๓	๑.๒๕	๐.๑๓	๐.๓๘	๐.๐๐

ค่าประมาณเหล่านี้ได้นำไปหาค่าประมาณคะแนนจริงตามจำนวนข้อที่ตอบได้ของแต่ละ
ระดับความสามารถ ทั้งสมการที่ ๓ และ ๔ ในบทที่ ๓ คะแนนจริง E และ n ในระดับความสามารถ
θ เกี่ยวกัน คือ คะแนนสมมูลกัน สำหรับระดับคะแนนที่อยู่ต่ำกว่าระดับการเอาได้ใช้สมการเชิงเส้นตรง
(สมการที่ ๔ ในบทที่ ๓) คำนวณคะแนนสมมูล ผลจากการเทียบคะแนนสมมูลนี้ ได้สร้างขึ้นเป็น
ตารางแปลงคะแนนในการเทียบมาตรฐานรูปแบบ IRT แสดงไว้ในภาคผนวก ข. ตารางที่ ๑๘
ถึงตารางที่ ๒๓

๔. สรุปผลการดำเนินการเทียบมาตรา

การเทียบมาตราในแต่ละกรณีของแบบสอบซึ่งดำเนินการตามระเบียบวิธีการเทียบมาตราสามรูปแบบภายใต้การออกแบบการใช้แบบสอบร่วมภายใต้สามชานุกรมเป็น ๔ วิธี แต่ละวิธี ได้ผลสุดท้ายเป็นการวางแปลงคะแนนจากมาตราของแบบสอบชุด Y ซึ่งสมมุติให้เป็นแบบสอบชุดใหม่ เทียบไปสู่คะแนนในมาตราของแบบสอบชุด X ซึ่งสมมุติให้เป็นแบบสอบชุดเก่า การวางคะแนนแปลงนกละคู่ของแบบสอบใช้เป็นการอ้างอิงของการเทียบมาตราเพื่อรายงานผลการสอบตามมาตราที่กองการ จากตารางคะแนนแปลงที่สร้างขึ้นในการวิจัยครั้งนี้ (ในภาคผนวก ข. ตารางที่ ๑๔ ถึงตารางที่ ๒๓) เมื่อนำมาแสดงเป็นภาพกราฟในเชิงเปรียบเทียบ ปรากฏว่า สามารถมองเห็นส่วนที่สอดคล้องกันทั้งสามรูปแบบของการเทียบมาตรา ซึ่งส่วนใหญ่เป็นช่วงคะแนนที่อยู่กลาง ๆ ถึงแผนภาพที่ ๖ ซึ่งเป็นกรณีแบบสอบคัดเลือก และแผนภาพที่ ๗ ซึ่งเป็นกรณีของแบบสอบวัดผลสัมฤทธิ์ สำหรับแบบสอบคู่ที่มีความยาวมาก คือ XV60 YV60 ของทั้งสองกรณี (แผนภาพที่ ๖ (ก) และแผนภาพที่ ๗ (ก)) ความแตกต่างของเส้นคะแนนแปลงของมาตราทั้งสามแตกต่างกันเฉพาะส่วนปลายที่เป็นคะแนนสูงสุด และต่ำสุด และถ้าเปรียบเทียบแผนภาพที่ ๖ (ก) และแผนภาพที่ ๗ (ก) พบว่า ความเบี่ยงเบนของวิธีการเทียบต่างมาตราในกรณีแบบวัดผลสัมฤทธิ์มีมากกว่ากรณีแรก สำหรับแบบสอบคู่ที่สั้นที่สุด คือ XV20 YV20 ความเบี่ยงเบนของเส้นกราฟซึ่งเป็นผลการใช้รูปแบบที่แตกต่างกันเห็นได้อย่างชัดเจน ทั้งกรณีของแบบสอบคัดเลือก และแบบสอบวัดผลสัมฤทธิ์ ดังนั้น การพิจารณาจากภาพสามารถเห็นความแตกต่างของแต่ละวิธีการเทียบมาตราในเชิงประจักษ์ได้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



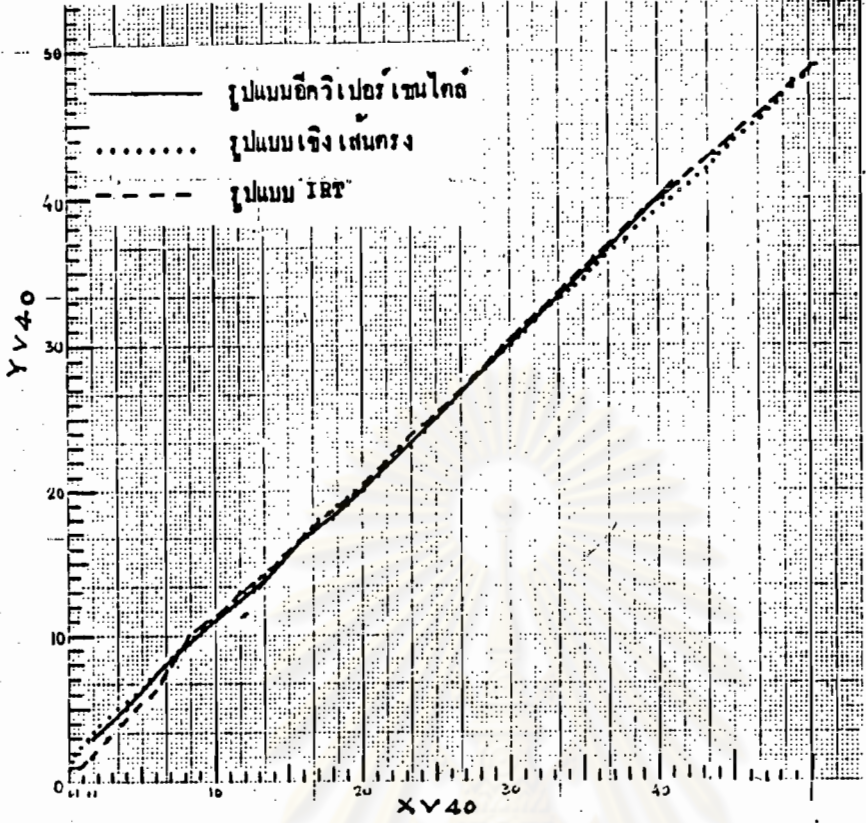
(ก)

แผนภาพที่ ๖ กราฟแสดงการเปรียบเทียบคะแนนสมมูล กรณีแบบสอบคัดเลือก
จากการเทียบมาตรฐานของแบบสอบ ๓ คู่

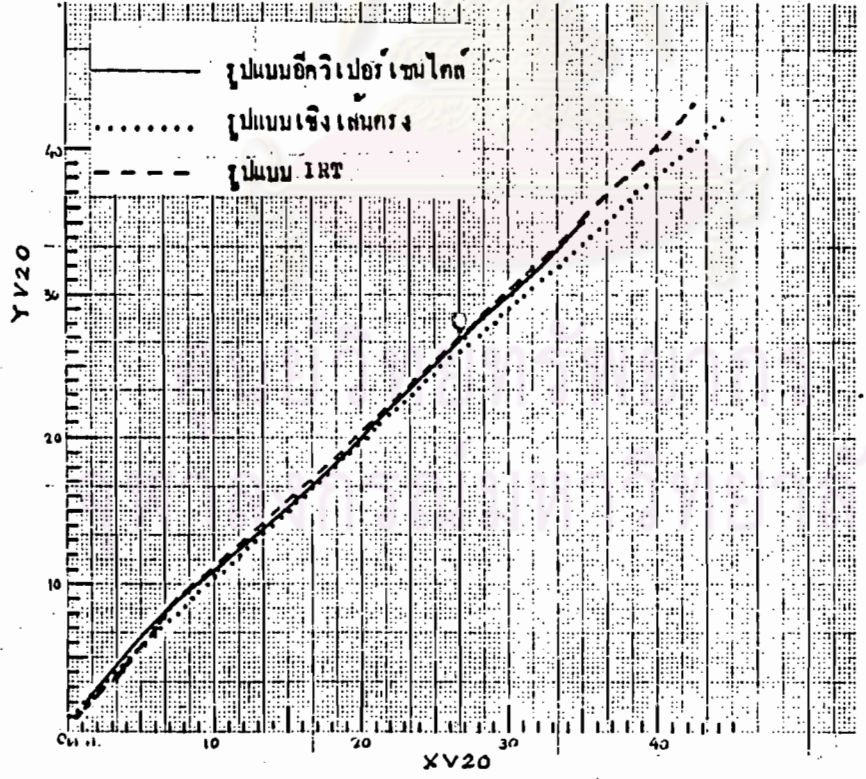
(ก) YV60 กับ XV60

(ข) YV40 กับ XV40 และ

(ค) YV20 กับ XV20

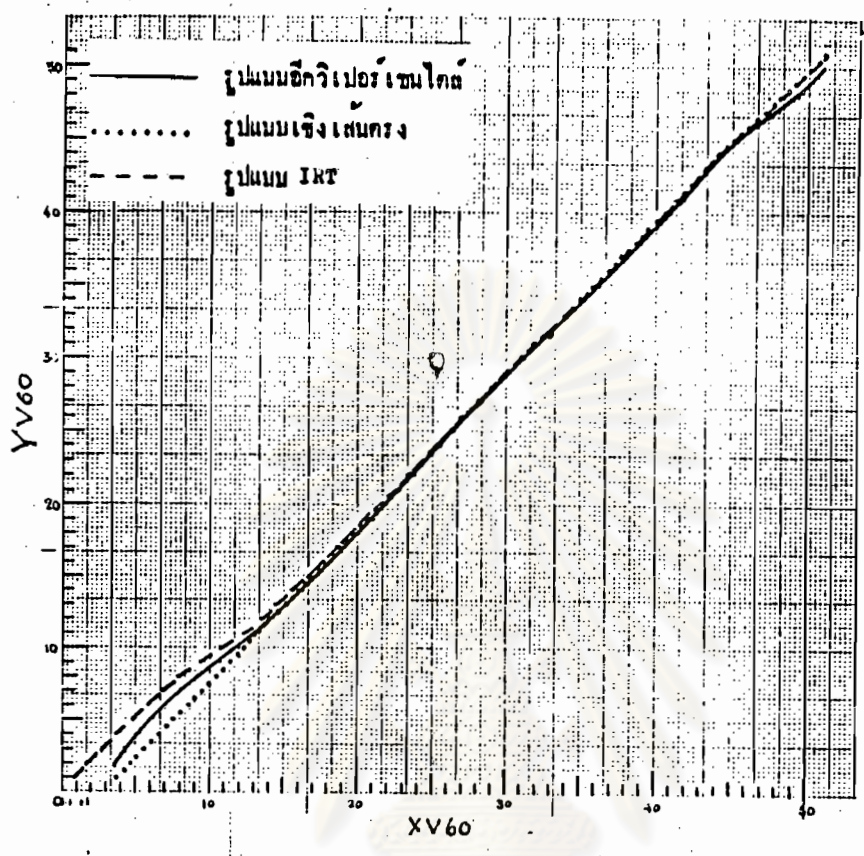


(ก)



(ข)

แผนภาพที่ ๖ (ต่อ)

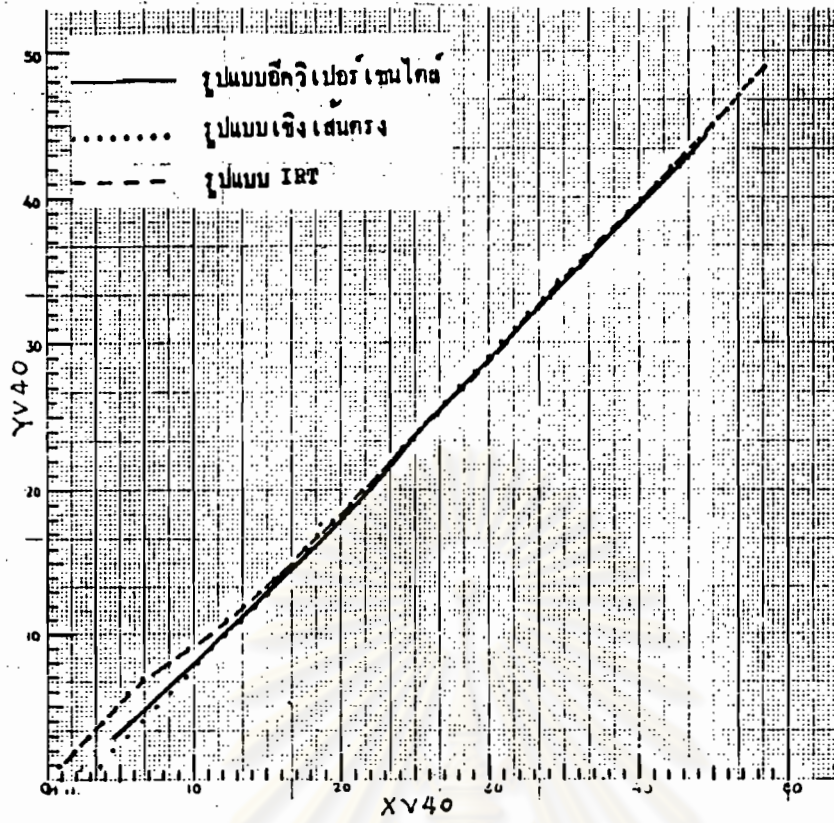


(ก)

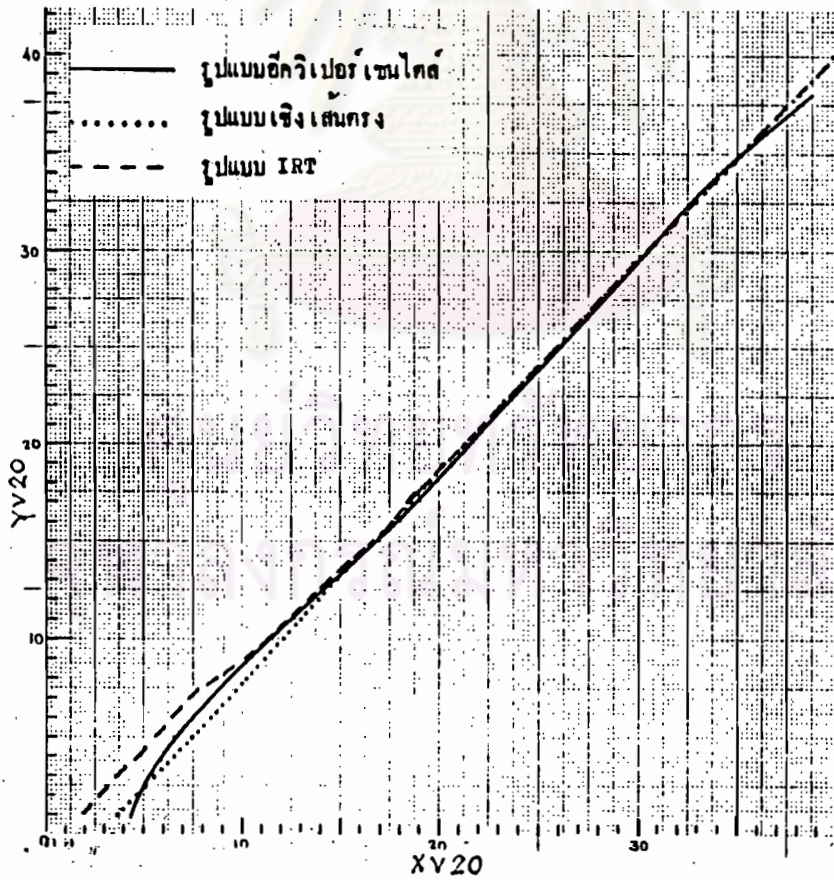
แผนภาพที่ ๗ กราฟแสดงการเปรียบเทียบคะแนนสมมูล กรณีแบบสอบวัดผลสัมฤทธิ์
จากการเทียบมาตรฐานของแบบสอบ ๓ คู่

- (ก) YV60 กับ XV60
- (ข) YV40 กับ XV40 และ
- (ค) YV20 กับ XV20

ศูนย์วิจัยและพัฒนา
จุฬาลงกรณ์มหาวิทยาลัย



(ก)



(ค)

แผนภาพที่ ๗ (ต่อ)

ผลการวิเคราะห์ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรา

การวิเคราะห์ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรา (SEE) แต่ละรูปแบบมีวิธีการที่ต่างกัน ซึ่งขึ้นอยู่กับวิธีการวิเคราะห์ความแปรปรวนแอสซิมโทติกของคะแนนแปลง (x^*) ตามคะแนน y ที่กำหนด การศึกษาครั้งนี้จึงได้กำหนดการรายงานค่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตราจากการประเมินความคลาดเคลื่อนที่คะแนน ซึ่งแปลงเป็นคะแนนมาตรฐาน $z(y)$ จาก ๐.๐ ถึง + ๓.๐ โดยเพิ่มขั้นบันได ๐.๕ รวม ๗ ขั้น เหมือนกันทุกวิธีการของการเทียบมาตรา ผลการวิเคราะห์รายรูปแบบมีดังนี้

๑. ความคลาดเคลื่อนมาตรฐานของการเทียบมาตราแบบอควิปเปอร์เซนไทล์

ความคลาดเคลื่อนมาตรฐานของการเทียบมาตราแบบอควิปเปอร์เซนไทล์ประมาณได้ จากสูตรที่ ๑๐ (บทที่ ๓) ค่า SEE ที่ประเมินจากจุด $z(y)$ ที่ต่างกัน ค่าที่ได้ต่างกันด้วย ลักษณะของ SEE เป็นค่าต่อเนื่องให้ภาพเป็นเส้นกราฟสมมาตรโดยมีค่าต่ำสุดที่ $z(y)$ เป็น ๐ การเสนอในรายงานจึงเสนอเฉพาะค่า $z(y)$ ทั้งแก่ ๐.๐ ถึง + ๓.๐ เริ่มตั้งแต่ $z(y)$ มีค่าเป็น ๐.๐ แล้วเพิ่มขั้นบันได ๐.๕ รวม ๗ ขั้น ผลการวิเคราะห์แสดงในตารางที่ ๗ ดังนี้

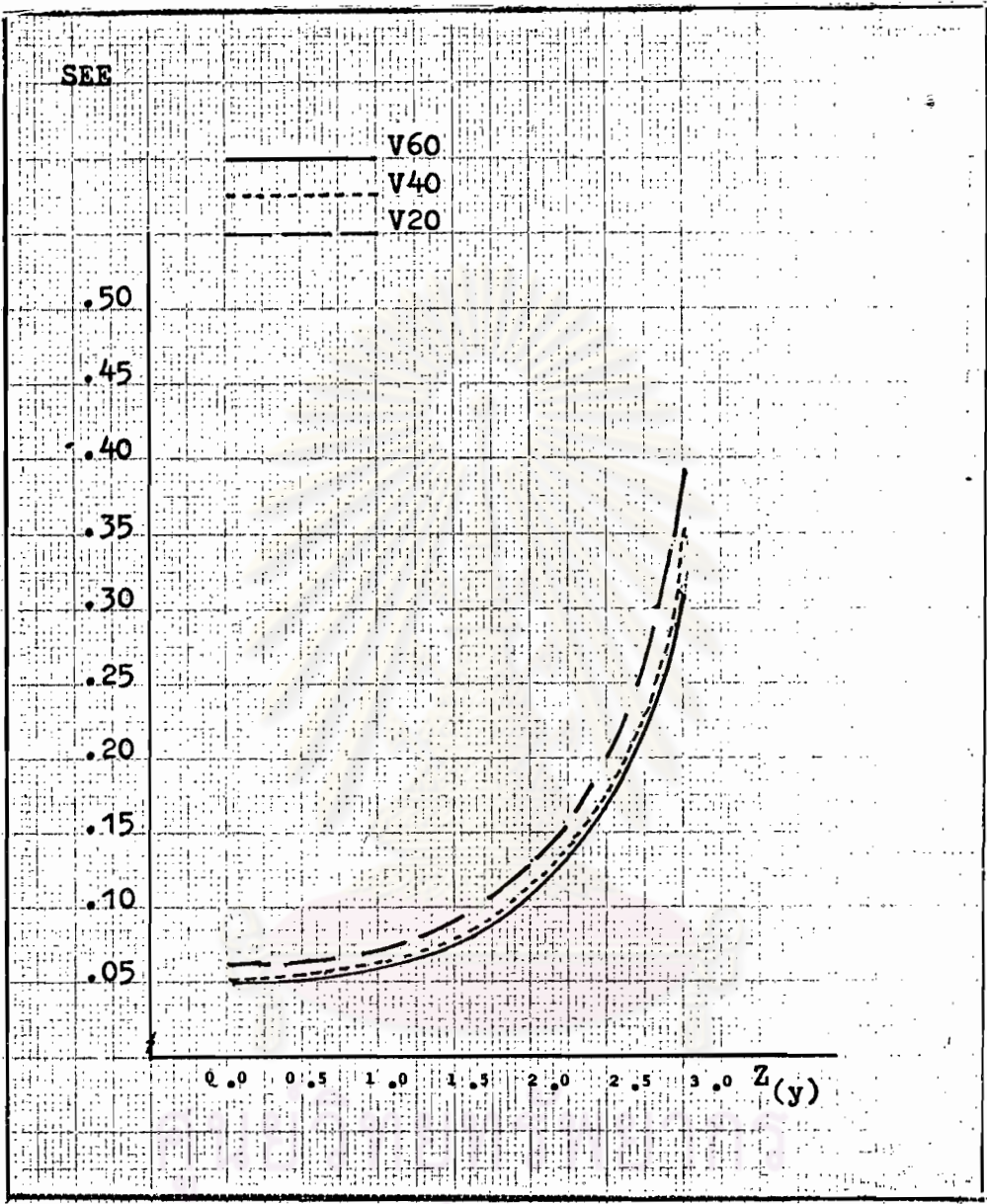
ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ ๗ ค่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐานรูปแบบอควิเปอร์ เซนโกล์

$Z(y)$	กรณีแบบสอบคัดเลือก			กรณีแบบสอบวัดผลสัมฤทธิ์		
	V60	V40	V20	V60	V40	V20
๐.๐	.๐๕๐๐	.๐๕๒๒	.๐๕๕๓	.๐๖๖๕	.๐๗๓๘	.๐๘๓๐
๐.๕	.๐๕๒๓	.๐๕๔๖	.๐๖๒๕	.๐๖๕๖	.๐๗๓๓	.๐๘๖๕
๑.๐	.๐๖๐๒	.๐๖๒๘	.๐๗๑๕	.๐๘๐๕	.๐๘๕๐	.๑๐๐๓
๑.๕	.๐๗๓๖	.๐๘๐๓	.๐๙๑๕	.๑๐๒๓	.๑๑๓๖	.๑๒๗๕
๒.๐	.๑๑๐๕	.๑๑๕๒	.๑๒๓๕	.๑๔๖๕	.๑๖๓๓	.๑๘๓๕
๒.๕	.๑๗๑๑	.๑๘๓๓	.๒๑๓๘	.๒๓๗๕	.๒๖๕๕	.๒๙๕๕
๓.๐	.๒๖๕๕	.๒๙๑๑	.๓๒๐๕	.๔๓๗๕	.๔๘๓๓	.๕๔๓๖

๑.๑ กรณีแบบสอบคัดเลือก

จากตารางที่ ๗ ค่า SEE ที่ใช้แบบสอบรวม V60 มีพิสัยอยู่ระหว่าง .๐๕๐๐ ถึง .๒๖๕๕ V40 อยู่ระหว่าง .๐๕๒๒ ถึง .๒๙๑๑ และ V20 อยู่ระหว่าง .๐๕๕๓ ถึง .๓๒๐๕ เมื่อพิจารณาจากจุด $Z(y)$ ที่ประเมินจุดเดียวกัน ค่า SEE ที่ได้จากวิธีที่ใช้แบบสอบรวมที่ยาวกว่า ไหลค่าน้อยกว่า เช่น $Z(y)$ เท่ากับ ๐.๐ พบว่า SEE ต่ำสุด คือ จาก V60 และสูงสุดได้จาก V20 SEE มีลักษณะทำนองเดียวกันนี้ตลอดการเทียบมาตรฐาน ซึ่งแสดงให้เห็นถึงแบบแผนของ SEE อย่าง ชัดเจน ดังปรากฏเป็นภาพกราฟในแผนภาพที่ ๕ เมื่อพิจารณาการเพิ่มของ SEE เมื่อจุดที่ประเมิน $Z(y)$ เพิ่มขึ้นในแต่ละวิธีการเทียบมาตรฐาน ซึ่งทั้ง ๓ วิธี มีแบบแผนเดียวกัน พบว่า การเพิ่มของ SEE จะเพิ่มขึ้นทีละน้อย ยกเว้นเมื่อ $Z(y)$ ที่ ๒.๕ และ $Z(y)$ ที่ ๓.๐ ค่า SEE เพิ่มขึ้นมาก อย่างชัดเจน



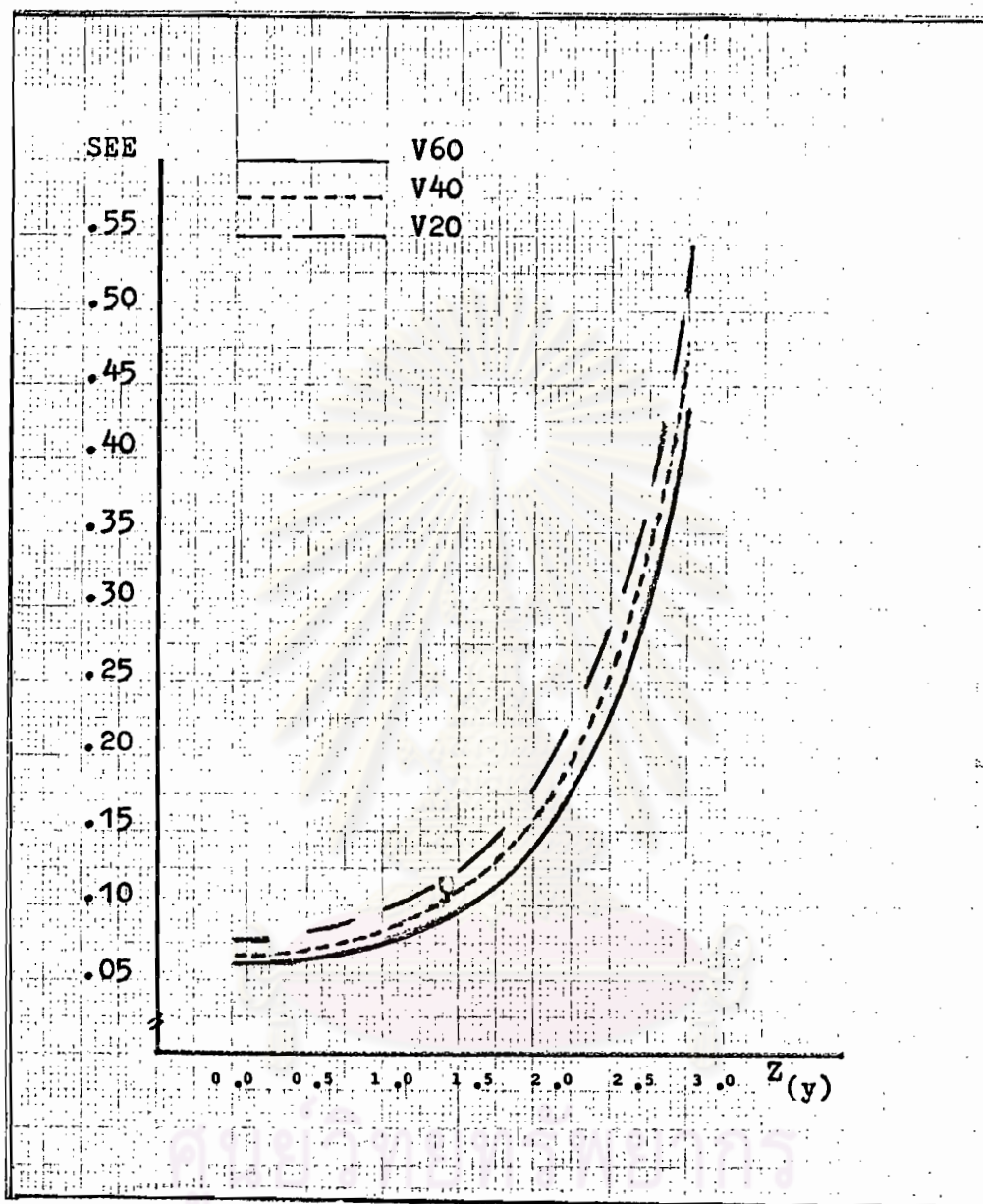
แผนภาพที่ ๔ กราฟเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐาน
 รูปแบบอควิปเปอร์เซนไทล์ กรณีแบบสอบคัทเลือก

๑.๒ กรณีแบบสอบวัดผลสัมฤทธิ์

จากตารางที่ ๘ ค่า SEE ที่ใช้แบบสอบรวม V60 มีพิสัยอยู่ระหว่าง .๐๖๖๕ ถึง .๑๑๑๕ V40 อยู่ระหว่าง .๐๖๖๕ ถึง .๑๑๑๕ และ V20 อยู่ระหว่าง .๐๘๓๐ ถึง .๑๑๑๕ เมื่อเปรียบเทียบ SEE ที่จุดประเมินเดียวกัน จุดต่อจุด ค่า SEE จากวิธีที่ใช้แบบสอบรวมที่ยาวกว่า ให้ค่าน้อยกว่า เช่น ที่ $Z_{(y)}$ เป็น ๐.๐ ค่า SEE จากวิธีทั้ง ๓ นำมาเรียงลำดับจากน้อยไปหา มากได้ ดังนี้ V60 V40 และ V20 และที่จุด $Z_{(y)}$ ที่ประเมินอื่น ๆ ค่า SEE มีปริมาณที่นำมา จัดอันดับในทำนองเดียวกัน ซึ่งสามารถแสดงเป็นภาพเสนกราฟให้เห็นแบบแผนการเพิ่มและลดของค่า SEE ในแต่ละวิธี และแต่ละจุดที่ประเมิน ดังแสดงในแผนภาพที่ ๕ เมื่อพิจารณาการเพิ่มของ SEE เมื่อจุดที่ประเมินเพิ่มขึ้น พบว่า ลักษณะการเพิ่มอย่างมีแบบแผนทำนองเดียวกับกรณีแบบสอบคัดเลือก แต่ค่าที่เพิ่มขึ้นของ SEE เมื่อจุดประเมินเป็นคะแนน $Z_{(y)}$ สูง ๆ มีอัตราการเพิ่มที่สูงกว่ากรณีแบบ สอบคัดเลือก แสดงว่า ผลการเทียบมาตราในกรณีแบบสอบวัดผลสัมฤทธิ์มีความแม่นยำน้อยกว่ากรณีของ แบบสอบคัดเลือกทั้งสามขนาดความยาวของแบบสอบรวมที่ใช้ โดยเฉพาะอย่างยิ่งในกรณีของแบบสอบ วัดผลสัมฤทธิ์การเทียบมาตราที่ระดับคะแนนสูงมากหรือต่ำมากมีความแม่นยำในการเทียบที่น้อยกว่าการ เทียบที่จุดอื่น



ศูนย์วิจัยและพัฒนา
จุฬาลงกรณ์มหาวิทยาลัย



แผนภาพที่ ๕

กราฟเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐาน
รูปแบบอควิเปอร์ เซนโตล กรณีสอบวัดผลสัมฤทธิ์

๒. ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐานรูปแบบเชิงเส้นตรง

ประเมินความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐาน ได้เสนอการประเมินที่
คะแนนมาตรฐานของ y คือ ที่ $Z(y)$ เท่ากับ ๐ ๐.๕ ๑.๐ ๑.๕ ๒.๐ ๒.๕ และ ๓.๐ เช่นเดียวกับ
รูปแบบอิกวิเปอร์เซนไทล์ ผลการวิเคราะห์ความคลาดเคลื่อนของการเทียบมาตรฐานของแบบสอบทั้งสอง
กรณี แสดงในตารางที่ ๔

ตารางที่ ๔ ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐานรูปแบบเชิงเส้นตรง

$Z(y)$	กรณีแบบสอบคัดเลือก			กรณีแบบสอบวัดผลสัมฤทธิ์		
	V60	V40	V20	V60	V40	V20
๐.๐	.๑๕๕๕	.๑๖๐๐	.๑๕๖๕	.๑๔๔๔	.๒๑๘๖	.๒๕๕๑
๐.๕	.๑๗๒๓	.๑๗๗๖	.๑๗๖๕	.๒๐๘๓	.๒๔๐๑	.๒๘๒๓
๑.๐	.๒๑๕๕	.๒๒๐๗	.๒๒๑๗	.๒๕๗๕	.๒๕๕๑	.๓๔๒๖
๑.๕	.๒๗๖๔	.๒๗๘๖	.๒๗๖๖	.๓๒๖๒	.๓๖๕๐	.๔๒๕๕
๒.๐	.๓๓๕๕	.๓๓๗๗	.๓๓๖๖	.๓๙๕๐	.๔๕๒๖	.๕๑๗๘
๒.๕	.๔๐๒๑	.๔๐๒๓	.๔๐๕๕	.๔๕๖๖	.๕๕๕๕	.๖๑๗๘
๓.๐	.๔๗๖๓	.๔๘๖๖	.๕๗๕๕	.๕๖๐๐	.๖๓๗๗	.๗๒๐๗

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

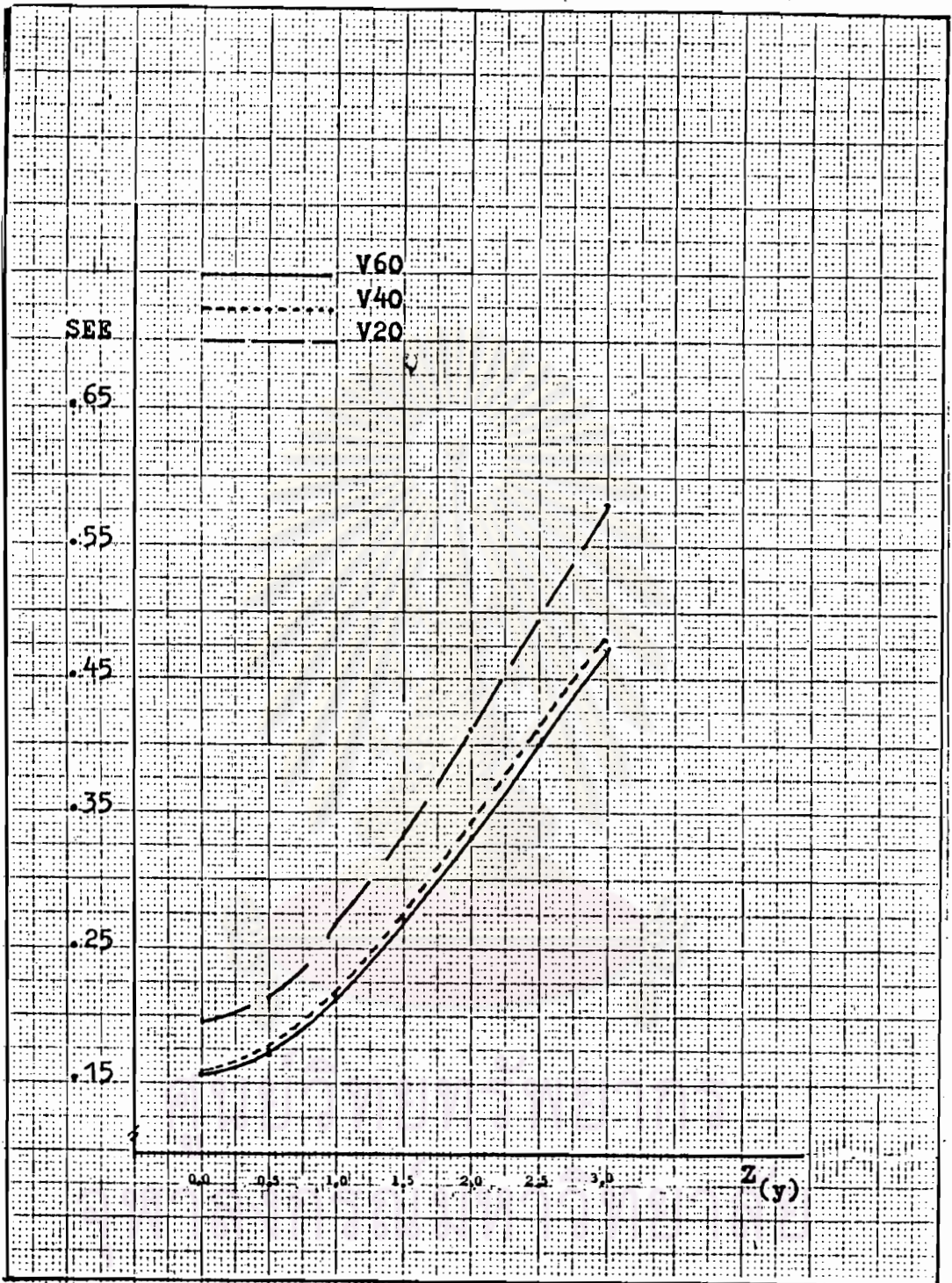
๒.๑ กรณีแบบสอบคัดเลือก

ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐานรูปแบบเชิงเส้นตรง มีลักษณะสมมาตร โดยมีค่า SEE ค่าสุดที่ $Z(y)$ เท่ากับ ๐.๐ จากตารางที่ ๔ ค่า SEE ที่ประเมิน ณ แต่ละจุดของ $Z(y)$ มีค่าเพิ่มขึ้นเมื่อจุดที่ประเมินเพิ่มขึ้น พิสัยของ SEE ที่ใช้ V60 อยู่ระหว่าง .๑๕๕๕ ถึง .๔๓๑๓ V40 อยู่ระหว่าง .๑๖๐๐ ถึง .๔๔๓๖ และ V20 อยู่ระหว่าง .๑๕๖๕ ถึง .๔๓๕๕ เมื่อพิจารณาว่าค่า SEE ที่เกิดขึ้นระหว่างวิธี ณ จุดประเมินใด ๆ พบว่า วิธีที่ใช้แบบสอบรวมที่ยาวกว่า ให้ค่า SEE ที่น้อยกว่าเสมอ ความแตกต่างระหว่างวิธีที่ใช้ V60 และ V40 มีไม่มากเท่ากับความแตกต่างระหว่าง V40 กับ V20 ซึ่งแสดงให้เห็นว่า ในการเทียบมาตรฐานรูปแบบเชิงเส้นตรงที่ใช้แบบสอบรวมยาวกว่าให้ผลการเทียบที่แม่นยำกว่าก็จริง แต่ในกรณีการเพิ่มความยาวแบบสอบรวม V60 เป็น V40 ให้ผลแตกต่างที่ไม่ชัดเจน ผลการวิเคราะห์แสดงเป็นกราฟในแผนภาพที่ ๑๐

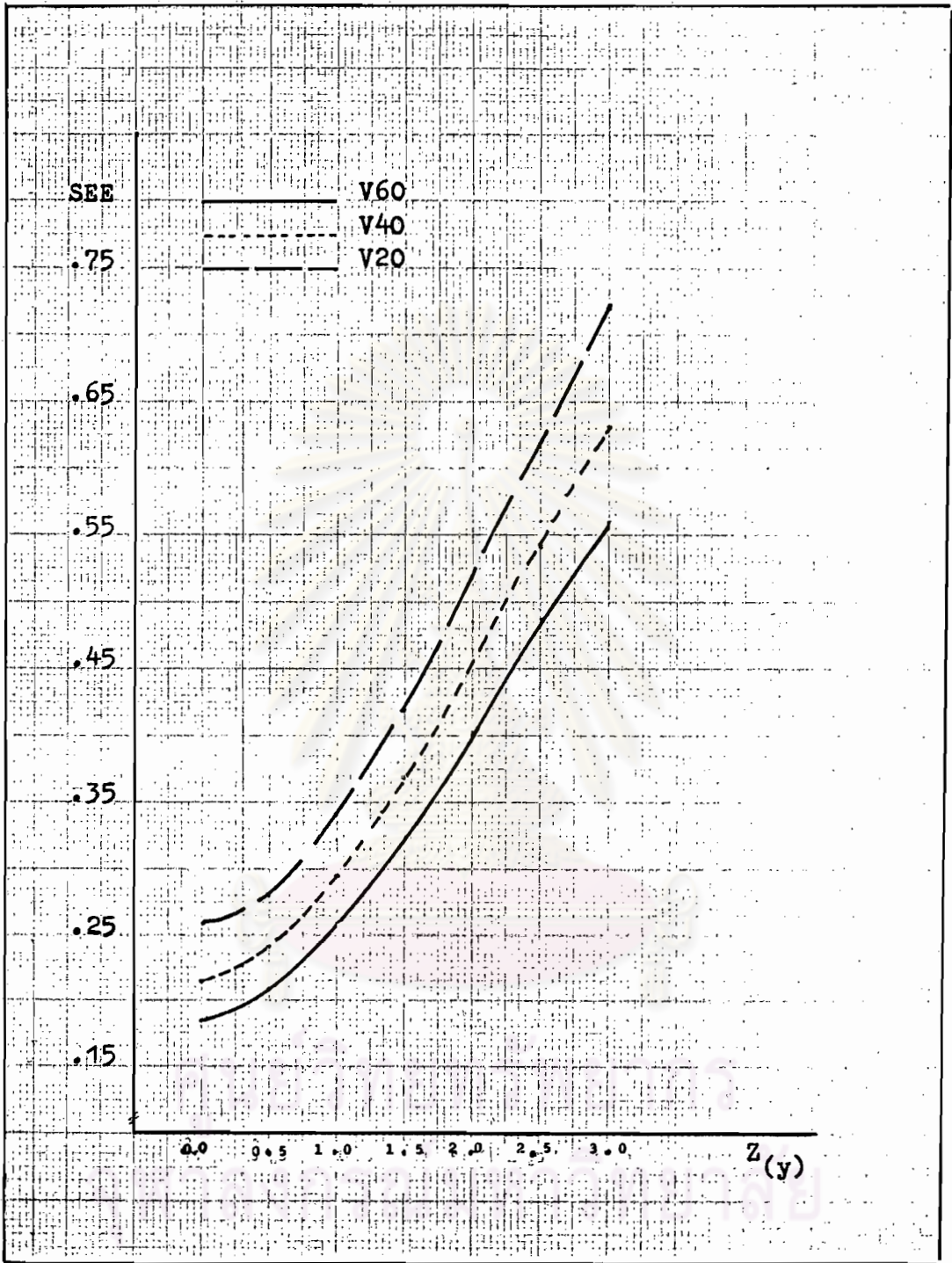
๒.๒ กรณีแบบสอบวัดผลสัมฤทธิ์

ลักษณะของความคลาดเคลื่อนมาตรฐานของกรณีแบบสอบวัดผลสัมฤทธิ์ มีแบบแผนทำนองเดียวกับกรณีของแบบสอบคัดเลือก คือ สมมาตร และมีการเพิ่มความจุดที่ประเมิน พิสัยของ SEE ที่ใช้ V60 อยู่ระหว่าง .๑๔๔๔ ถึง .๕๖๐๐ V40 อยู่ระหว่าง .๑๓๔๖ ถึง .๖๓๓๓ และ V20 อยู่ระหว่าง .๑๕๕๑ ถึง .๕๒๐๗ การเพิ่มขึ้นของค่า SEE จากวิธีหนึ่งไปอีกวิธีหนึ่งมีลักษณะของการเพิ่มที่พอ ๆ กัน ดังแสดงในแผนภาพที่ ๑๑

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



แผนภาพที่ ๑๐ กราฟเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐาน
รูปแบบเชิงเส้นตรง กรณีแบบสอบกคเลือก



แผนภาพที่ ๑๑

กราฟเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐาน
รูปแบบเชิงเส้นตรง กรณีแบบสอบวัดผลสัมฤทธิ์

๓. ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐานรูปแบบ IRT

วิเคราะห์ความแปรปรวนแอสซิมโทติกของคะแนนแปลงทัวส์ทรที่ ๑๔ (ในบทที่ ๓) และหาค่าความคลาดเคลื่อนของการเทียบมาตรฐานทัวส์ทรที่ ๑๔ (ในบทที่ ๓) กับกลุ่มตัวอย่างเทียบมาตรฐานทั้งสองกรณีของแบบสอบ โดยวิเคราะห์หาคู่ของการเทียบมาตรฐาน เสนอผลการประเมินความคลาดเคลื่อนที่ระบับความสามารถ (๑) ที่ต้องการตามจุดประเมินของคะแนนมาตรฐานของคะแนนจริงที่อยู่ภายในช่วง ± ๓.๐ ผลการวิเคราะห์ได้ค่า SEE แสดงในตารางที่ ๕ ดังต่อไปนี้

ตารางที่ ๕ ความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐานรูปแบบ IRT

$Z(y)$	กรณีแบบสอบคัดเลือก			กรณีแบบสอบวัดผลสัมฤทธิ์		
	V60	V40	V20	V60	V40	V20
-๒.๐	*	*	*	.๐๖๒๘	.๐๖๕๔	.๐๖๖๕
-๑.๕	.๐๒๐๐๓	.๐๒๑๔	.๐๒๒๕	.๐๓๓๑	.๐๓๕๒	.๐๓๖๖
-๑.๐	.๐๓๐๘	.๐๓๑๘	.๐๓๒๘	.๐๓๖๑	.๐๓๕๕	.๐๓๓๕
-๐.๕	.๐๓๖๖	.๐๓๓๒	.๐๓๒๕	.๐๓๒๐	.๐๓๓๕	.๐๓๕๒
๐.๐	.๐๓๕๑	.๐๓๕๑	.๐๓๐๘	.๐๒๕๕	.๐๓๓๒	.๐๓๒๕
๐.๕	.๐๒๘๐	.๐๒๓๖	.๐๒๕๕	.๐๒๘๕	.๐๒๕๓	.๐๓๑๒
๑.๐	.๐๒๖๒	.๐๒๖๕	.๐๒๖๕	.๐๒๓๑	.๐๒๘๒	.๐๒๕๖
๑.๕	.๐๒๓๘	.๐๒๓๓	.๐๒๓๖	.๐๒๖๕	.๐๒๓๓	.๐๒๘๓
๒.๐	.๐๒๐๘	.๐๒๐๓	.๐๒๐๖	.๐๒๕๕	.๐๒๖๒	.๐๒๓๒
๒.๕	.๐๑๘๓	.๐๑๘๐	.๐๑๘๓	.๐๒๕๕	.๐๒๕๒	.๐๒๖๕
๓.๐	.๐๑๕๓	.๐๑๖๑	.๐๑๖๓	.๐๒๒๓	.๐๒๓๘	.๐๒๔๘

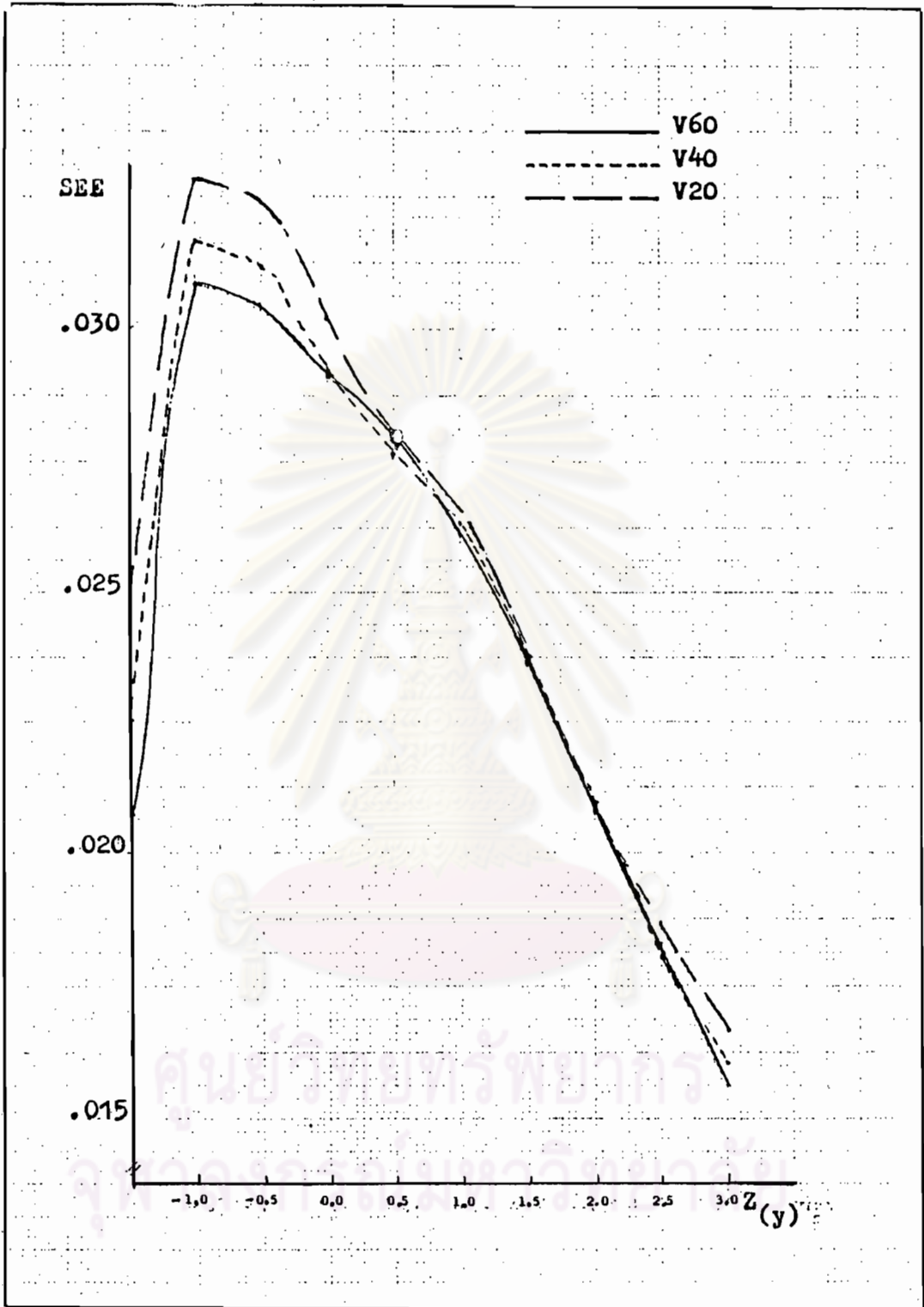
* ทำการระบับการเคา

จากตารางที่ ๑๑ การวิเคราะห์ค่าความคลาดเคลื่อนมาตรฐานของการเทียบมาตรา
ของทุกคู่ของแบบสอบเริ่มที่คะแนนต่ำสุดที่เป็นไปได้ตามรูปแบบของทฤษฎีการตอบข้อสอบสามพารามิเตอร์
ซึ่งหมายถึง คะแนนที่อยู่สูงกว่าค่าเฉลี่ยการกระจายข้อของแบบสอบทั้งคู่ ลักษณะทั่วไปของ SEE
พบว่า มีลักษณะไม่สมมาตร ไม่เป็นเส้นตรง มีแนวโน้มของการลดลงของค่า SEE เมื่อประเมินที่
ระดับความสามารถที่สูงขึ้น

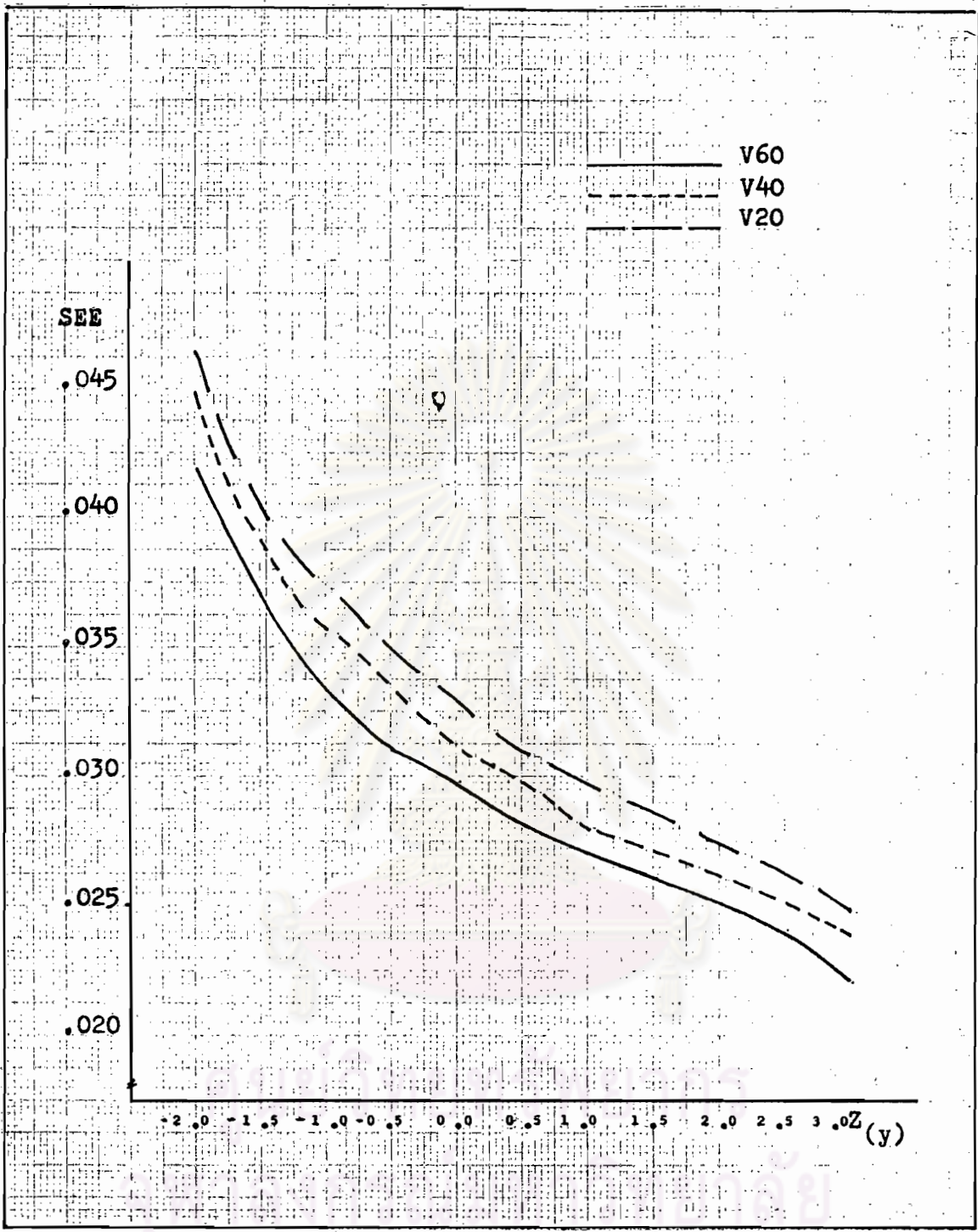
๓.๑ กรณีแบบสอบคัดเลือก พบว่า ก่อนที่ค่า SEE จะเริ่มลดลงที่ $Z(y)$ เท่ากับ
-๑.๑๐ นั้น SEE มีค่าน้อย แล้วเพิ่มขึ้นถึงจุด $Z(y)$ เท่ากับ -๑.๑๐ จากนั้นค่า SEE จะเริ่มลดลง
เป็นลำดับอย่างสม่ำเสมอ คือ ลดลงประมาณ .๐๐๒ จากจุดหนึ่งไปยังอีกจุดหนึ่ง ไม่มีลักษณะเป็น
เส้นตรง ดังแสดงในแผนภาพที่ ๑๒

๓.๒ กรณีแบบสอบวัดผลสัมฤทธิ์ พบว่า จากจุดประเมินที่ $Z(y)$ เท่ากับ -๒.๑๐
ค่า SEE จะค่อย ๆ ลดลงอย่างสม่ำเสมอด้วยค่าค่อนข้างน้อย เช่น ในการเทียบที่ใช้ V60 มีพิสัย
ของการลดอยู่ระหว่าง ๐.๐๐๐ ถึง .๐๐๒ V20 ตั้งแต่ .๐๐๑ ถึง .๐๐๒ ส่วน V40 มีการลด
เหมือนกันโดยตลอด คือ .๐๐๑ ดังแสดงในแผนภาพที่ ๑๓

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



แผนภาพที่ ๑๒ กราฟเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐาน
รูปแบบ IRT กรณีแบบสอบคัดเลือก



แผนภาพที่ ๑๓ กราฟเปรียบเทียบความคลาดเคลื่อนมาตรฐานของการเทียบมาตรฐานรูปแบบ IRT กรณีแบบสองวัตถุดิบสัมฤทธิ์

๘. ประสิทธิภาพสัมพัทธ์ (RE)

ประสิทธิภาพสัมพัทธ์ของวิธีการเทียบมาตรา (RE) ที่ใช้แบบสอบรวมขนาดต่างกัน ในแต่ละรูปแบบของการเทียบมาตรา ใช้เป็นมาตรการในการเปรียบเทียบความแม่นยำของการเทียบมาตราภายในแต่ละรูปแบบ จากการวิเคราะห์เชิงทฤษฎีที่พบว่า วิธีการเทียบมาตราที่ใช้แบบสอบรวมที่ยาวกว่าจะให้ผลที่มีความแม่นยำกว่า การสร้างสูตรเพื่อคำนวณประสิทธิภาพสัมพัทธ์จึงได้ใช้ความคลาดเคลื่อนมาตรฐานอันเกิดจากแบบสอบรวม V60 เป็นตัวเทียบในการหาอัตราส่วนร้อยละของความคลาดเคลื่อนมาตรฐานอันเกิดจาก V40 และ V20 ค่า RE ที่คำนวณได้เป็นดังนี้เปรียบเทียบประสิทธิภาพของวิธีนั้น ๆ เมื่อเทียบกับวิธีที่ใช้ V60 ผลการวิเคราะห์ RE ของวิธีการเทียบมาตราที่ทำการวิจัย แสดงในตารางที่ ๑๐



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ ๑๐ ประสิทธิภาพสัมพัทธ์ของวิธีเทียบมาตรฐานที่ใช้แบบสรบรวม V40 และ V20 เมื่อเทียบกับ V60

Z(y)	อีคิวเปอร์เซนไทล์		เชิงเส้นตรง		IRT	
	V40	V20	V40	V20	V40	V20
กรณีแบบสอบคัดเลือก						
๐.๐	๕๕.๗๔	๔๓.๗๕	๕๖.๘๗	๗๕.๐๘	๑๐๐.๐๐	๕๕.๗๒
๐.๕	๕๕.๗๔	๔๓.๘๑	๕๗.๑๗	๗๕.๖๓	๑๐๑.๕๐	๕๕.๖๓
๑.๐	๕๕.๘๕	๔๓.๙๒	๕๗.๖๘	๘๐.๒๒	๕๕.๖๘	๕๗.๗๖
๑.๕	๕๕.๘๕	๔๓.๙๘	๕๗.๗๓	๘๐.๖๕	๑๐๐.๕๒	๑๐๐.๕๕
๒.๐	๕๕.๙๒	๔๓.๙๗	๕๗.๘๕	๘๐.๘๗	๑๐๐.๕๘	๑๐๐.๕๗
๒.๕	๕๕.๖๒	๔๓.๙๖	๕๗.๘๓	๘๑.๒๖	๑๐๑.๖๖	๕๗.๘๖
๓.๐	๕๖.๒๕	๔๔.๒๕	๕๗.๓๑	๘๑.๓๕	๕๕.๖๑	๕๕.๐๑
เฉลี่ย	๕๕.๘๗	๔๔.๘๓	๕๗.๑๕	๘๐.๕๕	๑๐๐.๖๖	๕๗.๕๘
ส่วนเบี่ยงเบนมาตรฐาน	๐.๑๕	๐.๑๘	๐.๑๕	๐.๑๕	๑.๐๑	๒.๕๐
กรณีแบบสอบวัดผลสัมฤทธิ์						
๐.๐	๕๖.๐๑	๕๐.๑๒	๕๖.๓๗	๗๒.๘๗	๕๕.๘๓	๕๐.๘๘
๐.๕	๕๐.๐๓	๕๐.๐๕	๕๖.๗๑	๗๓.๗๕	๕๕.๕๕	๕๑.๓๕
๑.๐	๕๐.๓๓	๕๐.๒๕	๕๗.๓๕	๗๕.๒๘	๕๖.๐๕	๕๑.๕๕
๑.๕	๕๐.๐๕	๕๐.๐๒	๕๗.๘๖	๗๖.๓๗	๕๖.๗๐	๕๑.๕๘
๒.๐	๕๐.๐๖	๕๐.๐๕	๕๘.๑๖	๗๗.๐๕	๕๖.๗๕	๕๓.๓๘
๒.๕	๕๐.๓๑	๕๐.๐๐	๕๘.๘๐	๗๘.๗๕	๕๖.๘๒	๕๒.๕๒
๓.๐	๕๐.๖๕	๕๐.๖๖	๕๘.๕๒	๗๗.๗๐	๕๕.๓๗	๕๑.๕๓
เฉลี่ย	๕๐.๒๐	๕๐.๑๗	๕๗.๘๑	๗๕.๕๗	๕๖.๒๕	๕๑.๘๖
ส่วนเบี่ยงเบนมาตรฐาน	๐.๒๓	๐.๒๓	๑.๐๖	๑.๕๖	๐.๕๕	๐.๗๖

จากตารางที่ ๑๐ พิจารณาค่าเฉลี่ยของ RE ได้ข้อสรุปโดยส่วนรวมว่า วิธีการเทียบ
มาตราที่ใช้แบบสอบร่วมที่ยาวกว่า ให้ผลที่มีความแม่นยำมากกว่า ซึ่งเป็นผลที่สอดคล้องกับสมมุติฐาน
การวิจัยข้อที่ ๑ เมื่อพิจารณาในแต่ละรูปแบบที่ใช้ ได้ผลการวิเคราะห์ ดังนี้

รูปแบบการเทียบมาตราแบบคั้งเดิมสองรูปแบบ คือ อิกวิเปอร์เซนไทล์กับเชิงเส้นตรง
ให้ผลสอดคล้องกับทั้งสองกรณีของแบบสอบ กล่าวคือ เมื่อใช้วิธีที่มี V60 เป็นฐานของการเปรียบเทียบ
เทียบความแม่นยำแล้ว วิธีที่มี V40 มีผลการเทียบมาตราที่แม่นยำเป็นอันดับสอง และวิธีที่ใช้ V20
มีความแม่นยำเป็นอันดับสาม ถ้าหากพิจารณาในรายกรณีของแบบสอบ พบว่า กรณีแบบสอบคัดเลือก
วิธีที่ใช้ V40 มีประสิทธิภาพสูงใกล้เคียงกับวิธีที่ใช้ V60 ถึงร้อยละ ๕๕ ในรูปแบบอิกวิเปอร์-
เซนไทล์ และร้อยละ ๕๗ ในรูปแบบเชิงเส้นตรง ส่วนในกรณีแบบสอบวัดผลสัมฤทธิ์ วิธีที่ใช้ V40
มีประสิทธิภาพเทียบเท่ากับ V60 เป็นร้อยละ ๕๐ ในรูปแบบอิกวิเปอร์เซนไทล์ และร้อยละ ๕๗ ใน
รูปแบบเชิงเส้นตรง

รูปแบบการเทียบมาตรา IRT พิจารณาจากค่าเฉลี่ยของ RE พบว่า ความแม่นยำ
ที่ได้จากแบบสอบร่วมที่ยาวกว่ามีค่าสูงกว่าที่ใช้แบบสอบร่วมที่สั้นกว่า แต่มีข้อค้นพบที่แตกต่างจากรูปแบบ
การเทียบมาตราที่กล่าวมาแล้ว ๒ รูปแบบแรก ตรงที่ว่า ค่า RE อันเนื่องมาจากการใช้รูปแบบ IRT
นั้น สูงกว่าเมื่อใช้รูปแบบคั้งเดิม ถ้าหากพิจารณาในรายกรณีของแบบสอบ พบว่า ประสิทธิภาพของ
การเทียบมาตราด้วย V40 และ V20 มีลักษณะเป็นทำนองเดียวกันกับการใช้รูปแบบคั้งเดิม กล่าวคือ
ในกรณีแบบสอบคัดเลือก วิธีที่ใช้ V40 มีประสิทธิภาพสูงมากถึงร้อยละ ๑๐๐ ในขณะที่กรณีแบบสอบ
วัดผลสัมฤทธิ์มีประสิทธิภาพร้อยละ ๕๖ เมื่อเทียบกับการใช้ V60 ส่วนการใช้ V20 ในกรณีแบบสอบ
คัดเลือกมีประสิทธิภาพร้อยละ ๕๗ และกรณีแบบสอบวัดผลสัมฤทธิ์คือเป็นร้อยละ ๕๖ เมื่อเทียบกับ V60

ผลการวิเคราะห์กลุ่มตัวอย่างสอบทานผล

๑. ลักษณะของกลุ่มตัวอย่างสอบทานผล

จากการสุ่มตัวอย่างอย่างง่ายจากประชากรเกี่ยวกับกลุ่มตัวอย่างเทียบมาตรา α
และ β ได้พบลักษณะของกลุ่มตัวอย่างสอบทานผล ดังแสดงในตารางที่ ๑๑ ดังนี้

ตารางที่ ๑๑ สถิติบรรยายกลุ่มตัวอย่างสอบทานผล

แบบสอบ	กรณีแบบสอบคัดเลือก				กรณีแบบสอบวัดผลสัมฤทธิ์			
	MEAN	SD	MAX	MIN	MEAN	SD	MAX	MIN
XV60	๒๒.๓๑๖	๔.๒๐๕	๔๔	๓	๒๖.๕๕๑	๖.๕๕๕	๕๒	๒
YV60	๒๓.๐๓๓	๔.๕๔๕	๔๕	๔	๒๕.๑๓๓	๗.๒๕๖	๕๑	๒
XV40	๒๐.๕๕๓	๔.๘๐๖	๔๑	๒	๒๓.๗๕๓	๖.๒๘๑	๔๖	๑
YV40	๒๐.๘๖๕	๔.๑๑๗	๔๑	๓	๒๒.๓๗๓	๖.๕๗๑	๔๕	๑
XV20	๑๗.๗๐๕	๗.๗๐๓	๓๖	๑	๒๐.๒๔๗	๕.๕๖๑	๓๕	๑
YV20	๑๘.๐๑๕	๗.๐๑๘	๓๕	๒	๑๘.๘๖๕	๕.๘๓๘	๓๕	๑

ตารางที่ ๑๑ กลุ่มตัวอย่างสอบทานผลขนาด ๑๕๐๐ คน ทุกคนมีผลการสอบจากแบบสอบ ชุด X ทั้ง ๓ ชุด คือ XV60 XV40 และ XV20 ขณะเดียวกันมีผลการสอบของชุด Y ค่าย คือ YV60 YV40 และ YV20 ทั้งนี้เป็นผลจากการประมวลข้อมูลคะแนนสอบตามผลการสอบรายข้อในแบบสอบที่กำหนดขึ้นจากการจัดชุดของแบบสอบที่กล่าวมาแล้วข้างต้น

จุฬาลงกรณ์มหาวิทยาลัย

๒. ผลการคำนวณค่าดัชนี C

การวิเคราะห์ความเพียงพอของวิธีการเทียบมาตรา ใช้การเปรียบเทียบดัชนีความแตกต่างระหว่างวิธีการคำนวณค่าดัชนีความแตกต่าง C ได้เริ่มจากการใช้ตารางเทียบคะแนนสมมูลที่สร้างขึ้น (ตารางที่ ๑๔ ถึง ๒๓ ในภาคผนวก ข.) แปลงคะแนนในมาตราของแบบสอบถาม Y ให้เป็นคะแนน x^* ที่สมนัยกัน แล้วหาค่าความแตกต่างระหว่างคะแนนแปลง x^* และคะแนนเกณฑ์ x ของแต่ละรายบุคคล ใช้ผลต่างดังกล่าวมาคำนวณค่าดัชนีความแตกต่าง C ตามสูตรที่ ๑๖ (บทที่ ๓) ผลการวิเคราะห์แสดงในตารางที่ ๑๒

๓. การประเมินระดับการยอมรับความเพียงพอของวิธีเทียบมาตรา

จากการพิจารณาปริมาณค่า โดยใช้หลักการของปีเตอร์สันและคณะ (Petersen and others 1982: 93-94) ซึ่งกำหนดจุดกัณฑ์การยอมรับของส่วนเบี่ยงเบนมาตรฐานของการแจกแจงของคะแนนชุด x ที่เป็นเกณฑ์ของการวิเคราะห์กลุ่มสมมติฐานผลตั้งที่อธิบายไว้ในบทที่ ๓ ผลการประเมินค่า C ที่คำนวณได้ พบว่า ในกรณีของแบบสอบถามคัดเลือก ค่า C ทั้ง ๕ ค่า กระจายอยู่ในระดับน่าพอใจอย่างมาก และระดับน่าพอใจ ส่วนกรณีแบบสอบถามวัดผลสัมฤทธิ์ ค่า C ทั้ง ๕ ค่า อยู่ในระดับน่าพอใจทั้งหมด ทั้งเสนอในตารางที่ ๑๒

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ ๑๒ คัดนี้เปรียบเทียบความแตกต่าง C จากกลุ่มสอบทานผล และระดับความน่าพอใจของวิธี
เทียบมาตรา

แบบสอบ รวม	กรณีแบบสอบคัดเลือก			กรณีแบบสอบวัดผลสัมฤทธิ์		
	EQ	LI	IRT	EQ	LI	IRT
V60	.๑๖๕๕**	.๑๗๖๕**	.๑๖๕๕**	.๒๕๖๖*	.๒๕๐๓*	.๒๕๖๕*
V40	.๑๗๖๖**	.๑๕๕๕*	.๑๕๕๕**	.๓๐๐๑*	.๒๕๕๖*	.๓๐๖๕*
V20	.๒๓๖๖*	.๒๖๕๕*	.๒๕๖๖*	.๓๕๖๖*	.๓๖๖๖*	.๓๕๕๖*

** ความเพียงพออยู่ในระดับน่าพอใจอย่างมาก

* ความเพียงพออยู่ในระดับน่าพอใจ

จากตารางที่ ๑๒ แสดงให้เห็นว่า วิธีการเทียบมาตราทั้ง ๕ วิธี ในแต่ละกรณีของ
แบบสอบให้ผลการเทียบคะแนนสมมูลที่มีความเพียงพอในระดับที่น่าพอใจ ถ้าพิจารณาในแต่ละกรณีของ
แบบสอบ สามารถเห็นได้อย่างชัดเจนว่า วิธีการที่ใช้แบบสอบรวมที่มีความยาวมากกว่าให้ผลที่น่าพอใจ
มากกว่าโดยเฉพาะในกรณีแบบสอบคัดเลือก ถ้าพิจารณาผลการเทียบมาตราของกรณีแบบสอบสองกรณี
พบว่า กรณีแบบสอบคัดเลือกให้ผลการเทียบที่น่าพอใจกว่า ไม่ว่าจะเป็นการมองในประเด็นของรูปแบบ
การเทียบมาตรา หรือประเด็นของความยาวแบบสอบรวม

จุฬาลงกรณ์มหาวิทยาลัย

๔. ผลการทดสอบความมีนัยสำคัญของความแตกต่างของค่า C

จากการวางที่ ๑๒ คณิต C ที่ปรากฏในแต่ละกรณีของแบบสอบถามมีความสอดคล้องกันอย่างมีแบบแผน กล่าวคือ การเทียบมาตราที่ใช้แบบสอบถามที่ยาวกว่า ทำให้ได้คณิต C ที่มีค่าน้อยกว่า ซึ่งแสดงว่า ความแตกต่างระหว่างคะแนนแปลง กับคะแนนเกณฑ์ของแต่ละรายบุคคลที่เกิดขึ้นนั้นมีค่าน้อย การวิเคราะห์ค่า C นี้ มีจุดประสงค์เพื่อนำไปสู่การทดสอบสมมติฐาน ๒ ประการ ดังต่อไปนี้

๔.๑ ความยาวของแบบสอบถาม

พิจารณาในประเด็นความยาวของแบบสอบถาม พบว่า ผลการวิเคราะห์ได้ค่าคณิตความแตกต่าง C ที่สอดคล้องกับความคลาดเคลื่อนมาตรฐานของการเทียบมาตราอย่างมาก คือ เมื่อความยาวของแบบสอบถามเพิ่มขึ้น ค่า C มีค่าน้อยลง หมายความว่า วิธีการเทียบมาตราที่ใช้แบบสอบถามที่ยาวกว่า เป็นวิธีที่มีความเพียงพอในการเทียบคะแนนสมมูลมากกว่า ซึ่งเป็นผลที่สอดคล้องกับสมมติฐานของการวิจัยข้อที่ ๒

การทดสอบความแตกต่างของค่า C ใช้ทดสอบด้วยนันทพาราเมตริก Friedman model กับค่าอันดับของ C ค่า C จากการวางที่ ๑๒ เมื่อนำมาจัดอันดับตามความมากน้อยภายในแต่ละรูปแบบของการเทียบมาตราได้ผลดังปรากฏในการวางที่ ๑๓



ตารางที่ ๑๓ อันคัมค่าซีเปรียบเทียบความแตกต่างภายในแต่ละรูปแบบของการเทียบมาตรา

รูปแบบ	กรณีแบบสอบคัดเลือก			กรณีแบบสอบวัดผลสัมฤทธิ์		
	V60	V40	V20	V60	V40	V20
EQ	•	๒	๓	•	๒	๓
LI	•	๒	๓	•	๒	๓
IRT	•	๒	๓	•	๒	๓

จากการทดสอบความมีนัยสำคัญของค่าอันคัมของ C พบว่า อันคัมของค่า C ที่เกิดจากการใช้แบบสอบรวม V60 แตกต่างจากค่าอันคัมของค่า C ที่เกิดจากการใช้แบบสอบรวม V20 อย่างมีนัยสำคัญที่ .๐๕ ในกรณีของแบบสอบทั้งสองกรณี ส่วนการเปรียบเทียบคู่อื่นไม่พบความแตกต่างอย่างมีนัยสำคัญ แสดงว่า วิธีการเทียบมาตราที่ใช้แบบสอบรวมร้อยละ ๒๐ หรือ ๒๕ ข้อ ให้ความเพียงพอต่อการแปลงคะแนนสมมูลมากกว่าวิธีการเทียบมาตราที่ใช้แบบสอบรวมขนาดร้อยละ ๒๐ หรือ ๒๕ ข้อ

๔.๒ รูปแบบของการเทียบมาตรา

จากการวิเคราะห์กลุ่มสอบทานผลด้วยการพิจารณาจากค่า C ที่เกิดขึ้นในแต่ละรูปแบบของการเทียบมาตรา และทดสอบความมีนัยสำคัญทางสถิติด้วยการเปรียบเทียบรายคู่ที่สนใจตามแบบ Friedman เช่นเดียวกัน การทดสอบในประเด็นของแบบสอบรวม ตารางที่ ๑๔ แสดงการจัดอันคัมตามค่า C ในแต่ละกรณีของแบบสอบ ดังนี้

ตารางที่ ๑๔ อันคัมคาค์ชนิดนี้เปรียบเทียบความแตกต่างภายในแต่ละความยาวแบบสอบรวม

แบบสอบรวม	กรณีแบบสอบคัดเลือก			กรณีแบบสอบวัดผลสัมฤทธิ์		
	EQ	LI	IRT	EQ	LI	IRT
V60	•	๓	๒๑	๒	•	๓
V40	•	๓	๒	๒	•	๓
V20	•	๓	๒	๒	•	๓
อันคัมเฉลี่ย	•	๓	๒	๒	•	๓

ผลการทดสอบความมีนัยสำคัญของค่าอันคัมในกรณีแบบสอบคัดเลือก ซึ่งทำการทดสอบ ๓ คู่ คือ ระหว่างอีควิเปอร์เซนไทล์กับเชิงเส้นตรง ระหว่างอีควิเปอร์เซนไทล์กับ IRT และระหว่างเชิงเส้นตรงกับ IRT ปรากฏว่า การเปรียบเทียบระหว่างอีควิเปอร์เซนไทล์กับเชิงเส้นตรงมีความแตกต่างอย่างมีนัยสำคัญที่ .๐๕ อีกสองคู่ไม่ปรากฏว่ามีนัยสำคัญของความแตกต่าง

ผลการทดสอบความมีนัยสำคัญของอันคัมในกรณีแบบสอบวัดผลสัมฤทธิ์ ซึ่งทำการทดสอบ ๓ คู่ เช่นเดียวกับกรณีแบบสอบคัดเลือก ผลปรากฏว่า การเปรียบเทียบระหว่างเชิงเส้นตรงกับ IRT มีความแตกต่างอย่างมีนัยสำคัญที่ .๐๕ แสดงว่า วิธีการเทียบมาตราที่ใช้รูปแบบอีควิเปอร์เซนไทล์ให้ความเพียงพอต่อการแปลงคะแนนสมมูลมากกว่าวิธีเชิงเส้นตรง ในกรณีแบบสอบคัดเลือก ส่วนในกรณีแบบสอบวัดผลสัมฤทธิ์ รูปแบบเชิงเส้นตรงให้ความเพียงพอมากกว่ารูปแบบ IRT

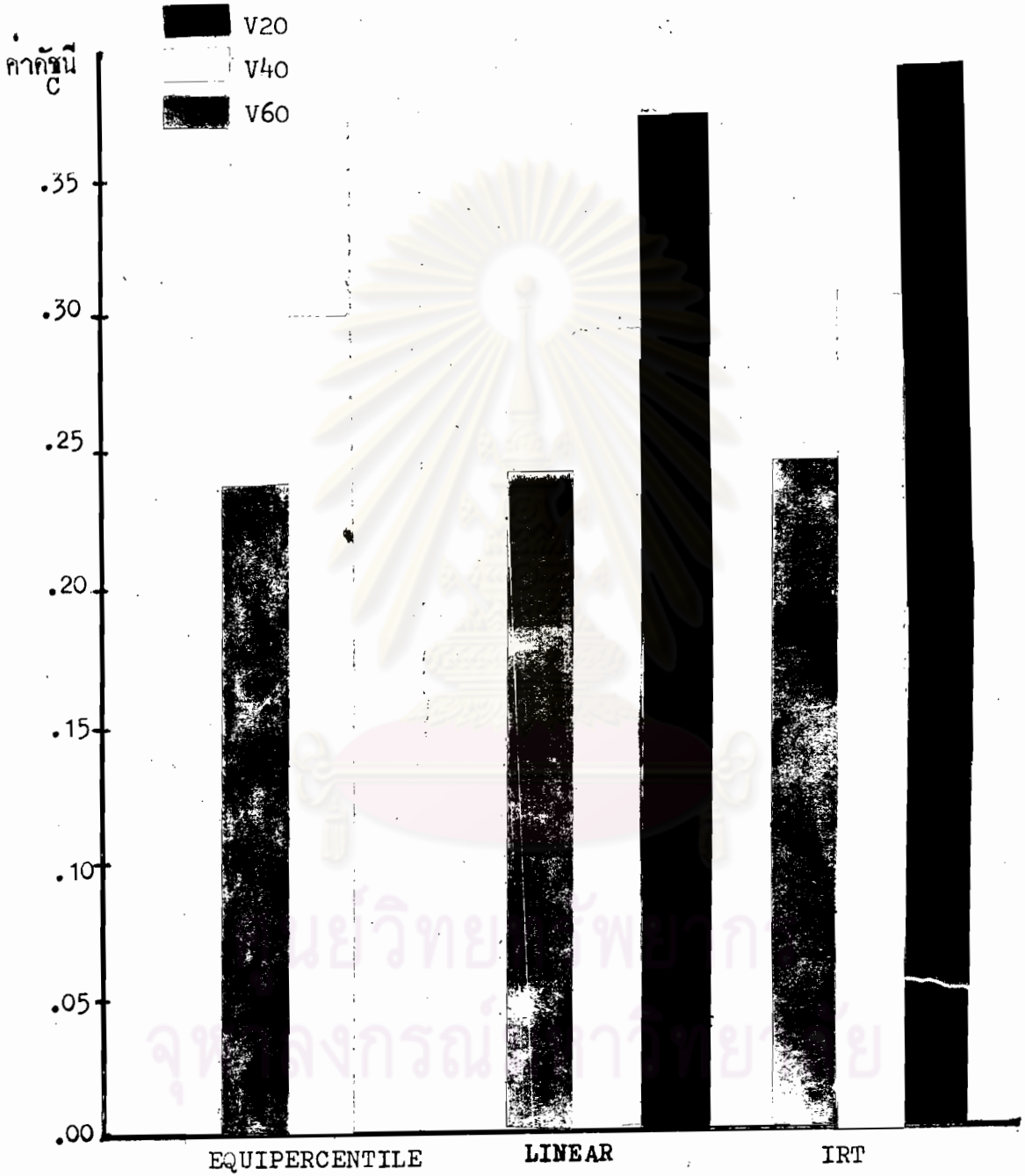
๕. สรุปผลการวิเคราะห์กลุ่มสอบทานผล

จากการวิเคราะห์กลุ่มสอบทานผลในแต่ละกรณีของแบบสอบ สรุปได้ดังนี้ ในประเด็นของความยาวของแบบสอบรวม V60 ให้ผลการเทียบที่มีความคงเส้นคงวากว่าการใช้ V20 ทั้งสองกรณีของแบบสอบ แต่ในประเด็นของรูปแบบของการเทียบมาตรา พบว่า รูปแบบอิกวิเปอร์เซนไทล์ ให้ผลการเทียบที่มีประสิทธิภาพเชิงเส้นตรงในกรณีแบบสอบคัดเลือก ส่วนในกรณีแบบสอบวัดผลสัมฤทธิ์ รูปแบบเชิงเส้นตรงให้ผลการเทียบที่มีประสิทธิภาพกว่า IRT ทั้งนี้ ข้อค้นพบที่ได้สอดคล้องกับสมมุติฐานของการวิจัยข้อที่ ๓ แต่ไม่สอดคล้องกับสมมุติฐานของการวิจัยข้อที่ ๔

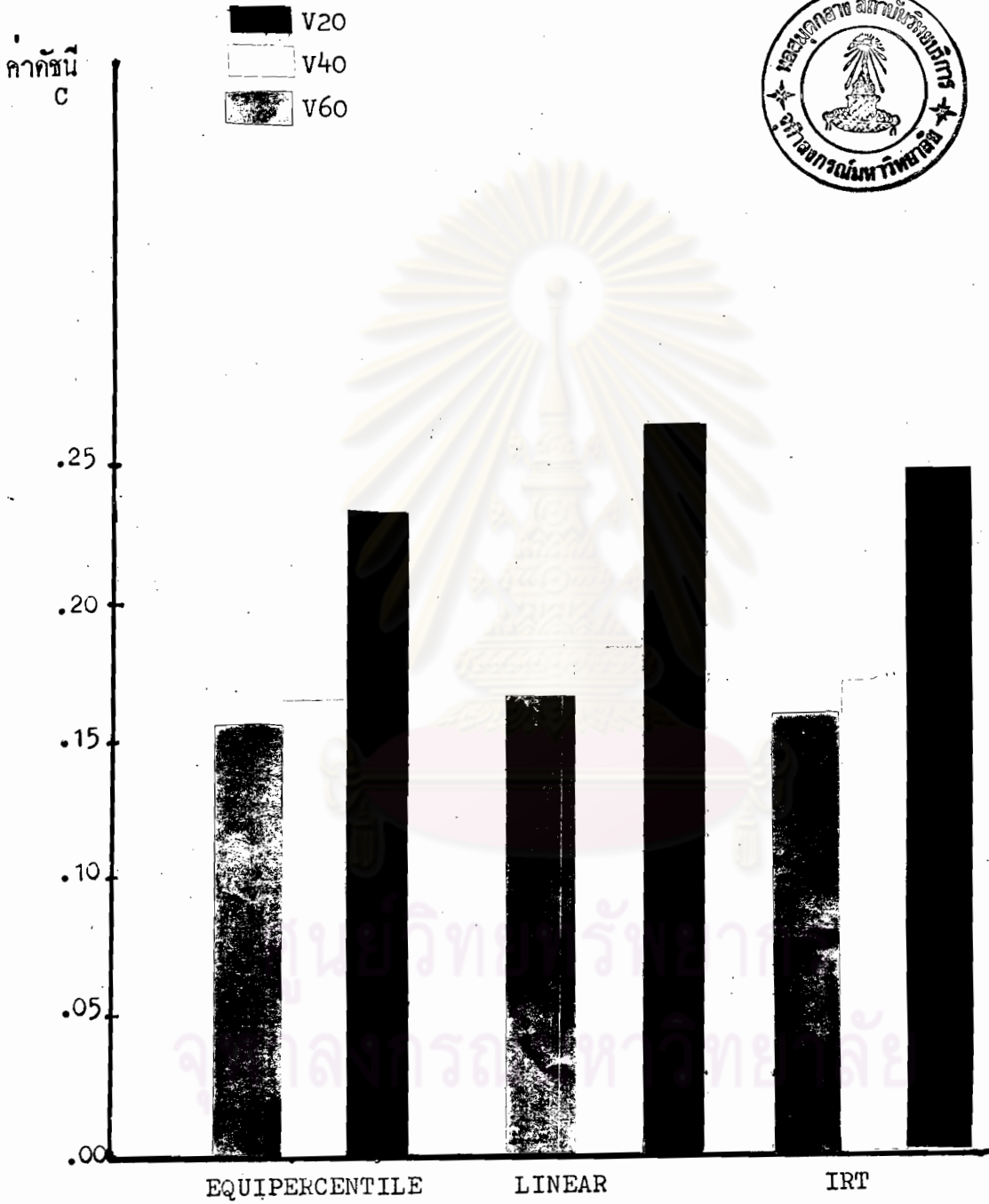
เนื่องจากค่า C ที่ใช้เป็นตัวนี้เปรียบเทียบความแตกต่างที่มีคุณสมบัติของความเป็นมาตรฐาน สามารถนำมาเปรียบเทียบกันได้โดยตรงในทุกสถานการณ์ของการทดสอบ และชนิดของแบบสอบ ดังนั้น ผลการวิเคราะห์นี้ทำให้สรุปได้อีกประเด็นว่า การเทียบมาตราที่ใช้กรณีแบบสอบคัดเลือก ให้ผลการเทียบที่เหนือกว่ากรณีแบบสอบวัดผลสัมฤทธิ์ทุกวิธี

ผลการวิเคราะห์ทั้งประเด็นของความยาว รูปแบบการเทียบมาตรา และกรณีของแบบสอบหรือสถานการณ์ของการทดสอบ สรุปเป็นแผนภูมิแท่งทั้งแผนภาพที่ ๑๔ และ ๑๕

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



แผนภาพที่ ๕ แผนภูมิแท่งแสดงค่าความต่างเปรียบเทียบความแตกต่างจากการวิเคราะห์
กลุ่มสอบทานผล กรณีแบบสอบวัดผลสัมฤทธิ์



แผนภาพที่ ๑๘ แผนภูมิแท่งแสดงค่าดัชนีเปรียบเทียบความแตกต่างจากการวิเคราะห์
กลุ่มสอบทานผล กรณีแบบสอบคัดเลือก