บทที่ 9
การทดสอบสมมติฐาน
(Tests of Hypotheses)

9.1 คำนำ

คำที่ใช้กันบ่อยที่สุดในวิชาสถิติก็คือ "การตัดสินใจ" และนี่ก็เป็นสาเหตุจริงอย่างหนึ่งที่ว่าสถิติจะเพิ่มบทบาทในการสร้างหรือการวิเคราะห์ของเกณฑ์ซึ่งขึ้นอยู่กับการตัดสินใจเป็นพื้นฐาน อย่างเช่นเราต้องการจะลงทุนเกี่ยวกับสินทรัพย์หรือเกี่ยวกับพันธบัตรรัฐบาล การโฆษณาทางหนังสือพิมพ์หรือทางโทรทัศน์ จะซื้อเครื่องจักรใหม่หรือซ่อมเครื่องเก่า เป็นต้น เราจะต้องเจอกับเหตุการณ์อย่างนี้เสมอ ในทางตรงข้ามเราก็ต้องเลือกวิธีใดวิธีหนึ่งที่ให้กำไรมากที่สุด ไม่ต้องสงสัยเลยว่าเราต้องตัดสินใจทุก ๆ วันไม่ว่าธุรกิจอุตสาหกรรมหรือการวางแผนเศรษฐกิจ ในทางวิทยาศาสตร์และในชีวิตประจำวัน เราต้องเสี่ยงกับการตัดสินใจผิด งานในด้านสถิติก็คือหาค่าของภัยนั้น (such risks) และถ้าเป็นไปได้ก็อาจตั้งเกณฑ์ (criteria) ที่จะลดโอกาสในการตัดสินใจผิดให้น้อยที่สุด

การทดสอบสมมติฐานทางสถิติก็เป็นการตัดสินใจหรือประเมินความสำคัญของผลที่ได้จากการสุ่มตัวอย่างหรือการหาว่าผลที่ได้นั้นจะเชื่อถือได้มากน้อยเพียงไร ถ้าไม่มีการประเมินความสำคัญของผลที่ได้ ก็อาจทำให้มีการตัดสินใจหรือสรุปอย่างผิด ๆ เกิดขึ้น เพราะผู้ที่ทำการสุ่มให้ความไว้วางใจแก่ผลที่ได้จากการสุ่มมากไป

9.2 ข้อความของสมมติฐาน

สมมติฐานแบ่งออกเป็นสองส่วน สมมติฐาน Ho ที่เราต้องการทดสอบ และ alternative hypothesis Ha สมมติฐานมักจะประกอบด้วยข้อความเกี่ยวกับตัวพารามิเตอร์เสมอ (หนึ่งตัวพารามิเตอร์หรือมากกว่า) ไม่เคยประกอบด้วยข้อความเกี่ยวกับตัวสถิติ

การทดสอบสมมติฐาน Ho แสดงออกในรูปของค่าเดี่ยว ๆ มากกว่าเป็นช่วง หากสมมติฐานแสดงออกเป็นค่าเดี่ยว เราก็ใช้การแจกแจงตัวอย่างเดียวของมัชฌิมเลขคณิตและการสำรวจผลที่ติดตามมาของการแจกแจงเฉพาะนี้ก็สามารถสำรวจได้ หากสมมติฐานแสดงออกเป็นช่วงเราก็

ใช้การแจกแจงตัวอย่างหลาย ๆ ตัว แต่ละตัวก็มีค่าต่าง ๆ กัน ปัญหาของการสำรวจก็ค่อนข้างจะยุ่งมากมาก

การตัดสินใจยอมรับหรือปฏิเสธสมมติฐานก็มีการอ้างอิงถึง Ho เสมอไม่เคยอ้างอิงถึง Ha

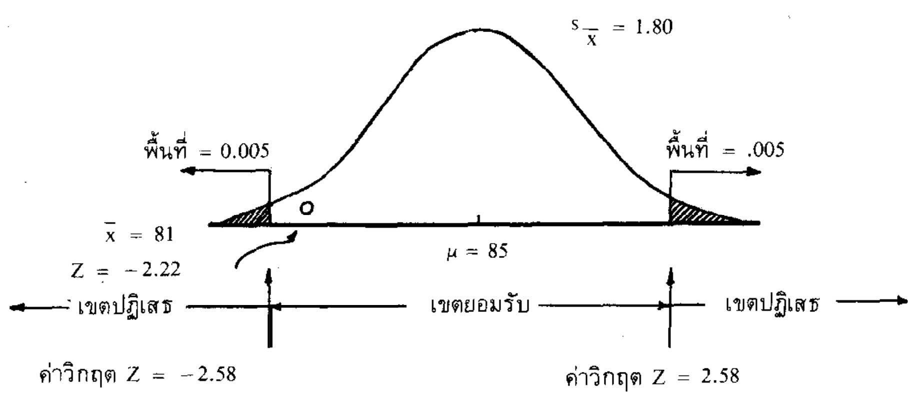
Ho เป็น "null hypothesis" ไม่มีนัยสำคัญพิเศษมาเกี่ยวพันกับเทอมนี้ เป็นเพียงสมมติฐานธรรมดาเพื่อที่จะทดสอบเท่านั้น

alternative hypothesis บรรยายถึงเงื่อนไขซึ่งหากหลักฐานของตัวอย่างเราขัดแย้งกับ Ho พอที่จะนำไปสู่การปฏิเสธ Ho เราจะปฏิเสธ Ho

9.3 การเลือก Ha การทดสอบข้างเดียวกับสองข้าง

หากเราให้ Ha : $\mu \neq 85$ เราก็กล่าวได้ว่า alternative hypothesis ไม่อยู่ในทิศทางใดทิศทางหนึ่ง (nondirectional) คืออนุญาตให้ผู้ตรวจสอบปฏิเสธสมมติฐาน (null hypothesis) หากหลักฐานพอที่จะแสดงว่า μ มีค่ามากกว่าหรือน้อยกว่าค่าของสมมติฐาน เพราะว่ามัชฌิมเลขคณิตของประชากรแตกต่างไปจากค่ามาตรฐานในทิศทางใดทิศทางหนึ่งที่ต้องการจะทราบ

การแจกแจงตัวอย่างของมัชฌิมเลขคณิต



รูปที่ 9.1 การทดสอบสมมติฐาน $\mu = 85$ ระดับนัยสำคัญ .01

บางครั้ง alternative hypothesis ที่มีทิศทางเหมาะสมดีกว่าต่อปัญหาของเรา ดังตัวอย่าง เราอาจใช้ Ha : $\mu < 85$ เราจะปฏิเสธ Ho : $\mu = 85$ ในกรณีที่หลักฐานแสดงว่า μ มีค่าน้อยกว่า 85 เราควรวางเขตวิกฤตในข้างซ้ายของการแจกแจงตัวอย่างดังแสดงในรูปที่ 9.2 ซ้ายมือแสดงถึงการทดสอบ หาก $\alpha = .05$ ค่าวิกฤตของ Z คือ -1.645 (ในกรณีที่ $\bar{X} = 81, s_x = 18, n = 100$) เราคำนวณค่าของ Z ได้ -2.22 ตกในเขตวิกฤต หากเราใช้ Ha : $\mu > 85$ เขตวิกฤตควรจะวางในข้างขวาทั้งหมดของการแจกแจงดังแสดงในรูปที่ 9.2 ขวามือ ในเหตุการณ์นี้ มัชฌิมเลขคณิตของตัวอย่างตกในเขตยอมรับ การตัดสินใจของเราต้องยอมรับ Ho

alternative Hypothesis ที่มีทิศทางจะมีความเหมาะสมเมื่อไร Ho ผิด เป็นการศึกษาค่าจริงของ μ แตกต่างไปจากค่าที่สมมติขึ้นในทิศทางเฉพาะ ดังตัวอย่างเช่น หากเราได้ริเริ่มสืบสวนโดยคิดว่าเราควรจะกระทำบางอย่าง หากมัชฌิมเลขคณิตของประชากรของค่าที่หกมีค่าต่ำกว่าค่ามาตรฐาน alternative hypothesis ที่เหมาะสมควรจะเป็น Ha : $\mu < 85$ สังเกตว่า หาก μ จริง ตกอยู่เหนือ μ สม จริง ๆ มันไม่น่าจะเป็นที่มัชฌิมเลขคณิตของตัวอย่างที่คำนวณได้นำไปสู่การปฏิเสธ Ho ในแง่ความสนใจของเรา นี้ไม่เป็นสิ่งสำคัญ เพราะว่าเราสนใจการศึกษาว่าประชากรของเราต่ำกว่ามาตรฐานหรือไม่เท่านั้น

เพราะฉะนั้น สิ่งหนึ่งที่ต้องเลือกระหว่าง alternative ไม่มีทิศทางกับมีทิศทางการเลือกสรรควรจะถูกกำหนดได้โดยเหตุผลที่ได้ให้ยกขึ้นมาศึกษาและควรจะทำก่อนที่ได้รวบรวมข้อมูล

เมื่อไรที่ข้อความของ alternative hypothesis ไม่มีทิศทาง ผลของการทดสอบเป็น การทดสอบสองข้าง เพราะว่า Ho จะถูกปฏิเสธหากคำนวณมัชฌิมเลขคณิตของตัวอย่างวางอยู่ในตำแหน่งข้างใดข้างหนึ่งของการแจกแจงตัวอย่าง ในทำนองเดียวกัน alternative hypothesis ที่มีทิศทางก็จะนำไปสู่การ ทดสอบข้างเดียว

บางครั้งการทดสอบข้างเดียวมีความเหมาะสมกว่า ดังตัวอย่างของสภาวะการทดสอบข้างเดียวแสดงได้เป็น

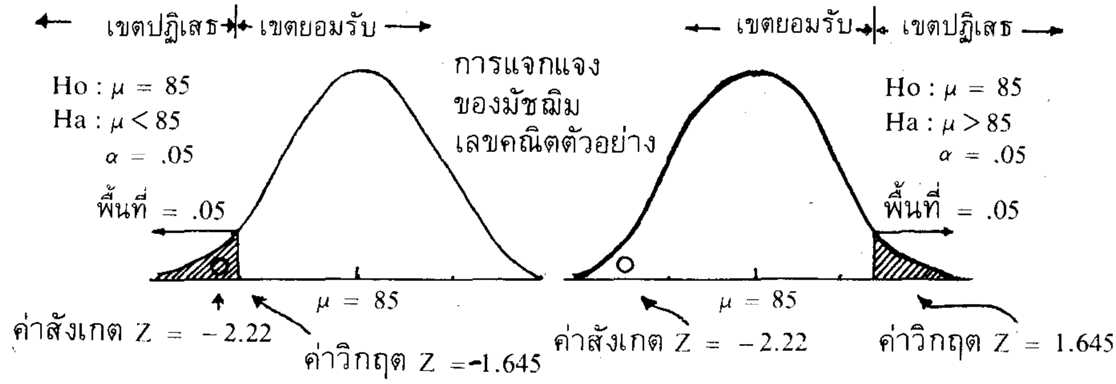
1. ผู้ผลิตสินค้าคนหนึ่งประสงค์ที่จะทดสอบอายุของหลอดไฟที่ได้ผลิตโดยกระบวนการใหม่ เขาต้องการใช้กระบวนการใหม่หากอายุเฉลี่ยของหลอดไฟมีอายุเกิน 1500 ชั่วโมงเท่านั้น

2. การทดสอบความเหมาะสมโดยธรรมชาติของเด็กอนุบาลหากการแสดงโดยเฉลี่ยด้อยกว่ามาตรฐาน ก็จำเป็นที่จะต้องจัดให้มีรายการฝึกหัดโดยธรรมชาติชนิดพิเศษ

3. ข้ออ้างว่าเมื่อไรฮอร์โมนเฉพาะที่ดูดซึมเข้าไปแล้ว จะทำให้เส้นผมงอกมากขึ้น

4. ข้อเสนอวิธีการสอนชนิดใหม่ควรจะจัดให้มีหากสามารถแสดงได้ว่าผลสำเร็จการเรียน

ภายใต้วิธีการชนิดใหม่ดีกว่าวิธีการมาตรฐาน



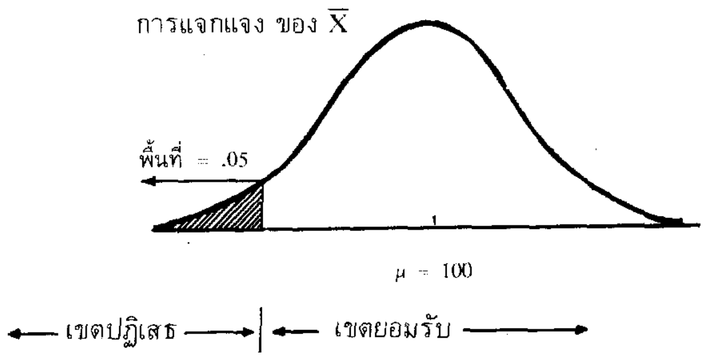
รูปที่ 9.2 ที่ตั้งของเขตที่ยอมรับและปฏิเสธเมื่อ Ha *มีทิศทาง*

ในแต่ละสภาวะเหล่านี้ พบว่ามีความแตกต่างในทิศทางเฉพาะหรือไม่ โดยเฉพาะในตัวอย่างข้อที่ 2 คำถามมีอยู่ว่า เด็ก ๆ เหล่านี้ด้อยกว่ามาตรฐานหรือไม่เมื่อเทียบกับความเหมาะสมโดยธรรมชาติ ที่อาจทดสอบโดยตั้งสมมติฐานว่าคะแนนความเหมาะสมโดยธรรมชาติเฉลี่ยประชากรของเด็กอนุบาลเท่ากับค่ามาตรฐานเทียบกับ alternative ว่าน้อยกว่าค่านี้ สังเกตว่าหากความเหมาะสมโดยธรรมชาติของเด็ก ๆ เหล่านี้เท่ากับหรือมากกว่าค่ามาตรฐาน การกระทำก็ไม่จำเป็น การจัดให้มีรายการฝึกหัดแบบพิเศษขึ้นหากการกระทำของเด็กเหล่านั้นด้อยกว่ามาตรฐานเท่านั้น ในทำนองเดียวกัน กรณีของสมมติฐานที่เกี่ยวกับฮอร์โมนของเส้นผม หากพบว่าเส้นผมงอกปกติหรือน้อยกว่าปกติ หลักฐานก็ไม่ได้ช่วยสมมติฐานของการวิจัย ในตัวอย่างนี้ อาจเป็นการทดสอบสมมติฐานของเส้นผมปกติเทียบกับ alternative ว่า เส้นผมมากกว่า

โดยทั่ว ๆ ไป alternative hypothesis ที่มีทิศทางมีความเหมาะสมเมื่อไรที่ไม่มีความแตกต่างในทางปฏิบัติในความหมายระหว่างค้นพบว่าสมมติฐานเป็นความแตกต่างหาค่าได้ในทิศทางตรงกันข้ามกับข้อความในทิศทางของ alternative hypothesis เหตุผลสำหรับเสนอแนะที่ปรากฏในรูปที่ 9.3 รูปนี้แสดงการออกแบบสำหรับการตัดสินใจเมื่อการทดสอบ Ho : $\mu = 100$ เทียบกับ Ha : $\mu < 100$ เราพบว่า เขตวิกฤตทั้งหมดอยู่ทางซ้ายมือ ค่าของ $\bar{X}$ ตกเหนือ 100 ไม่สามารถนำไปสู่การปฏิเสธ null hypothesis หากข้อความที่ไม่สามารถยอมรับนั้นหมายความว่าทิศทางของ alternative ไม่เหมาะสมก็ควรใช้การทดสอบสองข้าง

การตัดสินใจใช้ alternative ข้างเดียวต้องเป็นไปตามเหตุผลของคำถามซึ่งมีอยู่ในตัวของมันดังแสดงข้างต้น เวลาของการตัดสินใจขึ้นอยู่กับลักษณะของ alternative hypothesis ที่ได้ศึกษามาตั้งแต่

เริ่มแรก ก่อนการเก็บรวบรวมข้อมูลไม่ควรทำด้วยการสังเกตผลลัพธ์ของตัวอย่างและแล้วตั้งเขตวิกฤตในข้างของการแจกแจงตัวอย่างตรงไปทางผลลัพธ์ของตัวอย่างโอนเอียงดังตัวอย่าง หากเราใช้ระดับนัยสำคัญ 5% และทำตามกระบวนการคลาดเคลื่อนนี้อย่างมีระบบในกาลระยะยาว เรากำลังดำเนินทดสอบสองข้างจริง ๆ ที่ระดับนัยสำคัญ 10% อย่างเดียวกันการทดสอบก็ไม่เป็นที่พอใจที่จะจัดกับข้างเดียวของเราในทิศทางซึ่งเราคิดว่าผลลัพธ์อาจไปข้างนั้น หากมัชฌิมเลขคณิตของตัวอย่างปรากฏในทิศทางตรงข้ามก็เปิดการทดสอบสองข้าง หากดำเนินการทดสอบในลักษณะนี้การใช้ $\alpha = .05$ ก็จะ equivalent กับการทดสอบสองข้างที่ $\alpha = .075$ พร้อมด้วยพื้นที่ 0.05 ข้างหนึ่งอีกข้างหนึ่ง .025 แต่ส่วนที่มีพื้นที่มากกว่าวางข้างที่ปราศจากเหตุผลที่ได้สังกัปไว้ก่อนและเป็นการดีกว่ามากที่จะตัดสินล่วงหน้าจนกว่าจะพบว่าข้อขัดแย้งชนิดไหนที่มีความสำคัญจึงเลือก Ha ตามนั้น



รูปที่ 9.3 การทดสอบ

Ho : $\mu = 100$

เทียบกับ Ha : $\mu < 100$

9.4 เกณฑ์สำหรับยอมรับหรือปฏิเสธ Ho

การตัดสินใจยอมรับหรือปฏิเสธสมมติฐาน Ho ขึ้นอยู่กับเกณฑ์ที่ได้ใช้คือระดับนัยสำคัญ (α) นัยสำคัญหมายถึงอะไร ค่าของมัชฌิมเลขคณิตที่คำนวณได้อาจจะแยกไปจากค่าที่คาดหวังของสมมติฐาน เมื่อไรสมมติฐานเป็นจริง นักวิจัยมักจะพูดว่าผลลัพธ์ของการทดสอบมีนัยสำคัญที่ระดับ 0.05 นี้หมายความว่าหากใช้ $\alpha = .05$ เป็นเกณฑ์ตัดสินตัวสถิติจะตกในเขตวิกฤตและเชื่อว่า Ho ผิด

ผู้ทำการทดสอบจะเลือก α อย่างไร ตำแหน่งยิ่งอยู่ปลายสุดของการแจกแจงตัวอย่างความน่าจะเป็นที่จะเกิดการเบี่ยงเบนไปจากมัชฌิมเลขคณิตก็น้อยลง เมื่อไรสมมติฐานเป็นจริงและที่จุดไหนที่ความน่าจะเป็นที่จะเกิดขึ้นน้อยเหมาะสำหรับปฏิเสธสมมติฐานมากกว่ายอมรับ

ให้เราพิจารณาของผลการตั้งระดับสำคัญมีค่ามาก ๆ สมมติให้เป็น $\alpha = 0.25$ เมื่อไรเราตั้ง $\alpha = 0.25$ และสมมติฐานเป็นจริง 25% ของมัชฌิมเลขคณิตจะตกอยู่ในเขตวิกฤตเพราะฉะนั้นเมื่อ

ไรสมมติฐานเป็นจริง การทดสอบครั้งหนึ่งในสี่ครั้งจะนำไปสู่ข้อสรุปที่คลาดเคลื่อนว่าสมมติฐานมันผิด การเสี่ยงภัยของการกระทำความคลาดเคลื่อนนี้ดูเหมือนว่าสูงอย่างไม่น่าสนับสนุน

การลดภัยอันนี้ เราอาจปรับ α ให้มีระดับต่ำ สมมติว่าเป็น $\alpha = 0.001$ และการทดสอบก็ดำเนินการไป ค่ามัชฌิมเลขคณิตที่คำนวณได้ ก็แตกต่างไปมากจากค่าที่ควรจะเกิด หนึ่งในห้าร้อยครั้ง หากสมมติฐานเป็นจริงตามเกณฑ์ที่ใช้ เราก็จะไม่มีหลักฐานเพียงพอที่จะนำไปปฏิเสธสมมติฐานและก็ต้องยอมรับสมมติฐาน ในกรณีนี้ เราเสี่ยงภัยมากพอที่จะยอมรับสมมติฐานเมื่อไรสมมติฐานผิด กลับไปดูตัวอย่าง สมมติว่าท่านกำลังทดสอบความสมดุลของลูกเต๋าหนึ่งลูก ในการทอด 30 ครั้ง ปรากฏว่าลูกเต๋าออกหน้า 1 เสีย 20 ครั้ง ท่านจะพูดได้อย่างภูมิใจหรือไม่ว่า ท่านมีเหตุผลไม่เพียงพอที่จะปฏิเสธสมมติฐานเกี่ยวกับความสมดุลของลูกเต๋า

มีข้อเท็จจริงสองข้อที่ปรากฏให้เห็น หนึ่งกำหนดค่า α ขึ้นเป็นเกณฑ์สำหรับตัดสินชี้ขาดตั้งแต่โอกาสเกิดขึ้นยากเป็นสิ่งสำคัญ สอง การเลือกค่า α น้อยมากหรือใหญ่มากไม่มีประโยชน์ ผู้ทำการวิจัยนิยมใช้ระดับนัยสำคัญ 5% หรือ 1% ค่าเหล่านี้นำไปสู่เหตุผลที่เชื่อถือได้ว่า จะไม่มีการปฏิเสธสมมติฐานจนกว่ามันควรจะเป็นจริง ๆ ในเวลาเดียวกันค่าเหล่านี้ก็ไม่ได้กวดขันจนเกินไปที่จะยอมรับสมมติฐานที่ผิด

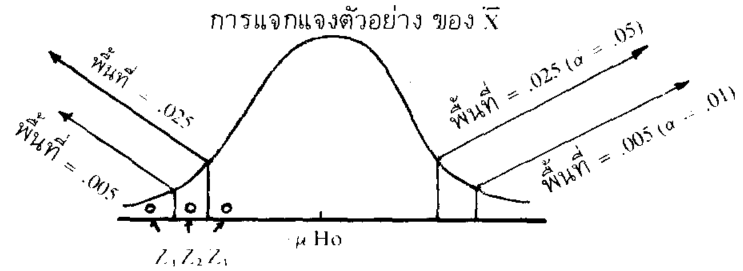
เมื่อไรดำเนินการทดสอบสองข้างและให้ $\alpha = 0.05$ ค่าวิกฤตของ Z คือ -1.96 และ $+1.96$ สำหรับการทดสอบเดียวกันแต่ให้ $\alpha = 0.01$ ค่าวิกฤตของ Z คือ -2.58 และ $+2.58$ สำหรับการทดสอบข้างเดียวขนาดของค่าวิกฤตของ Z คือ 1.64 (หรือ 1.645) เมื่อ $\alpha = 0.05$ และ 2.33 เมื่อ $\alpha = 0.01$ ความต้องการสำหรับค่าเหล่านี้เกิดขึ้นเสมอและควรจำไว้

บางครั้งนักวิจัยประเมินผลลัพธ์ของการทดสอบสมมติฐานโดยใช้เกณฑ์ทั้ง 5% และ 1% รูปที่ 9.4 แสดงสามค่าที่เป็นไปได้ของ Z จากการทดสอบสองข้าง หากค่าที่คำนวณได้เป็น Z_1 ผลลัพธ์ก็พลาดที่จะบรรลุถึงนัยสำคัญ ก็จะยอมรับ Ho หากค่าที่คำนวณได้เป็น Z_2 ผลลัพธ์ก็จะมีนัยสำคัญที่ระดับ 0.05 แต่ไม่มีนัยสำคัญที่ระดับ 0.01 หากค่าที่คำนวณได้เป็น Z_3 ก็จะปฏิเสธ Ho ไม่ว่าจะเป็นเกณฑ์ทั้ง 5% หรือ 1%

ถึงแม้ว่าเกณฑ์ 5% และ 1% มีประโยชน์สำหรับความมุ่งหมายทั่ว ๆ ไป มีบางเหตุการณ์เมื่อไรเลือกค่าอื่น ๆ ที่มีความหมายดีกว่า หากเชื่อว่าสมมติฐานก่อนทำการสรุปเราอาจใช้ α ให้มีค่าน้อยกว่า 0.01 ดังตัวอย่างเช่น การเปลี่ยนแปลงกระบวนการศึกษามีนัยปฏิเสธสมมติฐาน และถ้าการเปลี่ยนแปลงสิ้นเปลืองมาก (เวลาและเงินทอง) เราต้องแน่ใจข้อสรุปของเราก่อนการเปลี่ยนแปลง

เราอาจตั้งค่า $\alpha = 0.10$ หรือบางที่ 0.20 สภาวะเช่นนั้นอาจเกิดขึ้นได้ในระยะต้นของรูปการทดสอบ เมื่อไรการค้นพบค่าที่เป็นไปได้มีความสำคัญกว่าการขจัดอุบายที่ไม่เกิดผล

อะไรก็ตามที่ได้ใช้เป็นระดับนัยสำคัญ การตัดสินใจควรจะทำล่วงหน้า หากเรามองผลลัพธ์ของข้อมูลตัวอย่างก่อนถึงการตัดสินใจนี้ มีการตั้งระดับนัยสำคัญที่จุดหนึ่ง ซึ่งให้ข้อสรุปอยู่ในเส้นของสังกัปอะไรก็ได้ที่เราอาจมี ในเหตุการณ์หนึ่งเหตุการณ์ใด การตั้งระดับนัยสำคัญควรจะเป็นไปตามเหตุผลของการศึกษา ไม่ได้มาจากผลลัพธ์เฉพาะที่คำนวณหาได้ในตัวอย่าง



รูปที่ 9.4 สามผลลัพธ์ที่อาจเป็นไปได้ของการทดสอบสองข้างเมื่อ $\alpha = 0.05$ กับ $\alpha = 0.01$

9.5 การตัดสินใจทางสถิติ

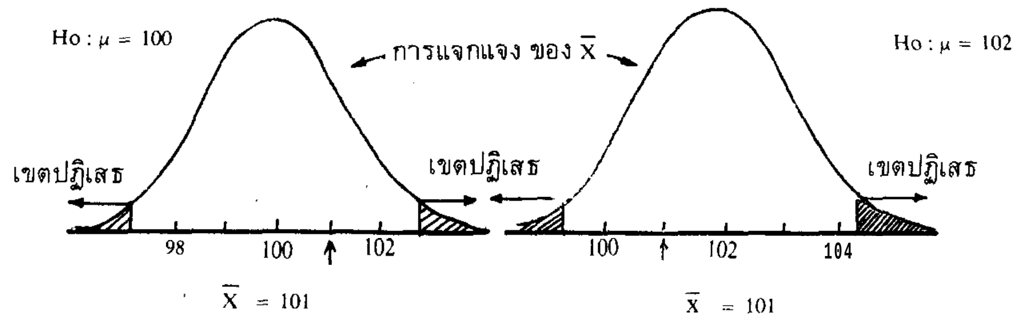
การตัดสินใจยอมรับ Ho หรือปฏิเสธ Ho ที่ผลลัพธ์ของการทดสอบสมมติฐาน เมื่อการทดสอบสมมติฐานที่ $\mu = 100$ เทียบกับ alternative hypothesis Ha : $\mu \neq 100$ หากปฏิเสธสมมติฐาน Ho หมายความว่าเราไม่เชื่อมัชฌิมเลขคณิตของประชากรเป็น 100 นอกจากนั้น ความน่าจะเป็นที่จะคำนวณมัชฌิมเลขคณิตของตัวอย่างที่ได้เกิดขึ้นชนิดนั้นมีค่าน้อยกว่าเมื่อสมมติฐานเป็นจริง ความเชื่อมั่นของเราในการตัดสินใจถูก มีมากกว่าที่จะปฏิเสธสมมติฐาน

ส่วนการยอมรับสมมติฐานไม่ได้หมายความว่า เราเชื่อสมมติฐานเป็นจริง การตัดสินใจนี้สะท้อนไปถึงหลักความจริงที่ว่าเราไม่มีหลักฐานเพียงพอที่จะปฏิเสธสมมติฐาน (สมมติฐานได้รับการสนับสนุน) มาพิจารณาตัวอย่างข้างต้นเมื่อสมมติฐานเป็น $\mu = 100$ หากมัชฌิมเลขคณิตของตัวอย่างเป็น $\bar{X} = 101$ ดูรูปที่ 9.5 ซ้ายมือ การตัดสินใจของเราจะยอมรับสมมติฐาน แต่สมมติว่าสมมติฐานเป็น $\mu = 102$ และ $\bar{x} = 101$ เราก็ยังยอมรับสมมติฐานที่ $\mu = 102$ เหมือนเดิม ดังรูป 9.5 ขวามือ จึงขอสรุปสั้น ๆ ดังนี้ ปฏิเสธสมมติฐานหมายความว่าไม่มีเหตุผลเพียงพอที่จะเชื่อว่าสมมติฐานเป็นจริงแต่หากยอมรับสมมติฐานหมายความว่า เราเชื่อสมมติฐานได้เป็นจริง นั้นไม่ได้หมายความว่าสมมติฐานต้องเป็นจริงหรือ แม้ว่าสมมติฐานน่าเป็นจริงก็ควรจำยอมรับด้วย สำหรับสมมติฐานอื่น ๆ หากได้ทดสอบโดยใช้ข้อมูลตัวอย่างเดียวกัน

การทดสอบสมมติฐานที่ $\mu = 100$ เราคำนวณ Z ได้เป็น

$$Z = \frac{\bar{X} - \mu_{สม}}{\frac{s_X}{\sqrt{n}}}$$

หาก Z มีค่ามากพอก็จะปฏิเสธสมมติฐาน ขนาดของ Z ขึ้นอยู่กับทั้งตัวตั้งและตัวหาร หากขนาดตัวอย่างมีขนาดใหญ่ตัวหาร $s_X/\sqrt{n}$ จะมีค่าน้อย ในกรณีนี้ ความสัมพันธ์ระหว่างผลต่าง $\bar{X}$ กับ $\mu_{สม}$ มีค่าน้อยอาจให้ค่า Z มากพอที่จะนำเราไปปฏิเสธสมมติฐานเรียกว่า มีนัยสำคัญทางสถิติ



รูปที่ 9.5 การทดสอบสมมติฐานที่ $\mu = 100$ หรือที่ $\mu = 102$ *เมื่อ* $\bar{x} = 101$

แต่ความแตกต่างระหว่าง $\mu_{จริง}$ กับ $\mu_{สม}$ น้อยมากไม่มีความสำคัญ หากใช้ตัวอย่างขนาดเล็กมาก ความคลาดเคลื่อนมาตรฐานจะมีค่ามาก และหากสมมติฐานผิดก็เป็นการยากที่จะค้นคว้า เว้นแต่ความแตกต่างระหว่าง $\mu_{จริง}$ กับ $\mu_{สม}$ มีค่ามาก ข้อเสนอแนะนี้มีความสำคัญมากเพื่อพิจารณาออกแบบการทดลองเกี่ยวกับการเลือกใช้ขนาดตัวอย่าง

9.6 ความคลาดเคลื่อนในการทดสอบสมมติฐาน

ลักษณะของข้อความที่จะตั้งเป็นสมมติฐานมีสองชนิดคือ สมมติฐาน Ho เป็นจริงหรือสมมติฐาน Ho ผิด ในทำนองเดียวกัน การตัดสินใจที่เป็นไปได้ก็มีสองชนิด ยอมรับสมมติฐานหรือปฏิเสธ นำเข้ามารวมกันก็จะมีสี่ข้อความที่เป็นไปได้ ดังตารางที่ 9.1 หากสมมติฐานเป็นจริงและเรายอมรับ หรือสมมติฐานผิดและเราปฏิเสธ เป็นการตัดสินใจที่ถูก สำหรับข้อความอื่น ๆ ในตารางอยู่ในความคลาดเคลื่อน ความคลาดเคลื่อนชนิดที่หนึ่งหรือที่สองดังตัวอย่างแสดงการกระทำความคลาดเคลื่อนชนิดที่หนึ่งและที่สอง อย่างเช่น หากเราหยุดกระบวนการผลิตถึงแม้ว่ากระบวนการผลิตนั้นยังดำเนินการอย่างเหมาะสมอยู่ หมายความว่าเราได้กระทำความคลาดเคลื่อนชนิดที่หนึ่ง หากไม่หยุดกระบวนการผลิต ถึงแม้ว่ามีบางสิ่งบางอย่างของกระบวนการผลิตเกิดเสียขึ้น หมายความว่าเราได้กระทำความคลาดเคลื่อนชนิดที่สอง มีอีกตัวอย่างหนึ่งแสดงการกระทำความคลาดเคลื่อนชนิดที่หนึ่งและที่สอง

อย่างเช่น นาย ก. กำลังดื่มกาแฟอยู่กับเพื่อนของเขา เขาทั้งสองตกลงกันที่จะโยนเหรียญเพื่อตัดสินใจว่าใครจะเป็นผู้จ่ายค่ากาแฟ นาย ก. พนันกับเพื่อนของเขาแสดงว่า เขาไว้ใจเพื่อนของเขา นาย ก. ก็ตั้งสมมติฐานว่าเพื่อนของเขาซื่อตรงและไม่โกงเขา หากวันต่อไปก็พนันกันด้วย การโยนเหรียญและนาย ก. ก็เสียพนันทั้งสองครั้ง นาย ก. ก็ยังไม่สงสัยเพื่อนของเขา แต่ถ้านาย ก. เสียพนัน 1,000 ครั้ง เขาอาจสงสัยว่าเพื่อนของเขากำลังโกงเขาและไม่ยอมรับสมมติฐานว่าเพื่อนของเขาซื่อตรง หากสมมติฐานดั้งเดิมเป็นจริง ก็ไม่น่าที่นาย ก. ควรจะเสียพนัน 1,000 ครั้ง เพราะว่านาย ก. เสียพนัน 1,000 ครั้งติดต่อกัน นาย ก. จึงไม่ยอมรับสมมติฐาน

การที่นาย ก. ไม่ยอมรับสมมติฐานอาจเป็นการถูกต้อง เพื่อนของเขาอาจไม่มีความซื่อตรงจริง ๆ ในขณะเดียวกัน สมมติฐานดั้งเดิมอาจถูกก็เป็นได้ นั่นเป็นแต่เพียงในทางทฤษฎีสำหรับเพื่อนของเขาชนะ 1,000 ครั้งติดต่อกัน โดยปราศจากการโกง ถ้าเพื่อนของเขามีความซื่อตรงจริง ๆ แต่นาย ก. ตัดสินใจจากหลักฐานว่าเพื่อนของเขาไม่ซื่อตรง นาย ก. ก็กำลังกระทำความคลาดเคลื่อนชนิดที่หนึ่ง คือไม่ยอมรับสมมติฐานเมื่อสมมติฐานนั้นเป็นจริง

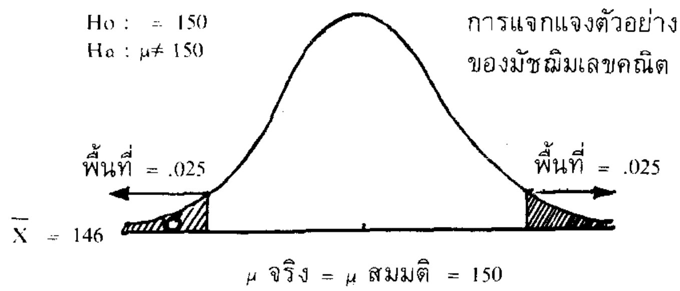
หากเพื่อนของนาย ก. ชนะ 6 ครั้งในการโยนเหรียญ 10 ครั้งแน่นอน นาย ก. จะไม่สงสัยถึงความซื่อตรงของเพื่อนของเขา หรือ นาย ก. ยอมรับสมมติฐานที่เพื่อนของเขามีความซื่อตรง ผลจากการสรุปอาจถูกต้อง แต่เพื่อนของเขาไม่ซื่อตรงก็อาจเป็นไปได้ การที่เขาเสียพนันนั้นเป็นไปโดยบังเอิญแต่ก็พยายามที่จะเอาชนะอยู่ตลอดเวลาด้วยความไม่ซื่อตรง ถ้าเพื่อนของนาย ก. ไม่มีความซื่อตรงจริง ๆ แต่นาย ก.ก็เชื่อต่อไปว่า เพื่อนของเขาซื่อตรง นาย ก. ก็กำลังกระทำความคลาดเคลื่อนชนิดที่สอง คือยอมรับสมมติฐานเมื่อสมมติฐานนั้นผิด

จากเหตุการณ์ที่กล่าวพอสรุปดังในตาราง

		ลักษณะของข้อความ	
		Ho ผิด	Ho จริง
การตัดสินใจ	ยอมรับ Ho	ความคลาดเคลื่อนชนิดที่ II	การตัดสินใจที่ถูก
	ปฏิเสธ Ho	การตัดสินใจที่ถูก	ความคลาดเคลื่อนชนิดที่ I

ตารางที่ 9.1 ชนิดของความคลาดเคลื่อนในการทดสอบ

ความคลาดเคลื่อนชนิดที่ I คือปฏิเสธสมมติฐาน Ho เมื่อสมมติฐาน Ho เป็นจริง พิจารณารูปการแจกแจงตัวอย่างดังในรูปที่ 9.6 แสดงถึงการทดสอบสมมติฐานที่ $\mu = 150$ เทียบกับ alternative hypotheses $\mu \neq 150$ เมื่อใช้ระดับนัยสำคัญ 5% สมมติว่า $\bar{X} = 146$ และการทดสอบได้ดำเนินการทำให้เราปฏิเสธสมมติฐาน Ho เหตุผลในการปฏิเสธ Ho คือว่าหากสมมติฐานเป็นจริง มัชฌิมเลขคณิตของตัวอย่างเบี่ยงเบนไปควรจะเกิดขึ้นน้อยกว่า 5% ของครั้ง เพราะฉะนั้นจึงดูเหมือนมีเหตุผลมากกว่าที่จะเชื่อว่ามัชฌิมเลขคณิตของตัวอย่างแตกต่างไปจากที่กำหนดใน Ho อย่างไรก็ตามมัชฌิมเลขคณิตของตัวอย่างนี้สามารถเป็นค่าเบี่ยงเบน เหล่านั้นที่คำนวณได้ตลอดจนความไม่แน่นอนของการสุ่มตัวอย่างจากประชากรที่สมมติฐานเป็นจริงจากค่าจำกัดความเขตวิกฤตพิสูจน์องศาของความเบี่ยงเบนนั้นเมื่อสมมติฐานเป็นจริง 5% ของมัชฌิมเลขคณิตจะบรรลุถึงหรือเกินเขตเหล่านั้น เพราะฉะนั้นเมื่อไรดำเนินการทดสอบตามเกณฑ์นี้และสมมติฐานเป็นจริงแล้ว 5% ของมัชฌิมเลขคณิตของตัวอย่างจะนำเราไปสู่การสรุปที่คลาดเคลื่อน ปฏิเสธสมมติฐาน



รูปที่ 9.6 การทดสอบสมมติฐานที่เป็นจริงแต่ $\bar{X}$ นำไปสู่ความคลาดเคลื่อนชนิดที่หนึ่ง

เมื่อพบกับสมมติฐานเป็นจริง เพราะฉะนั้นเราต้องเสี่ยงต่อการสรุปที่ผิด การกำหนดขนาดของการเสี่ยงด้วยความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่ I (α)

α = P (ความคลาดเคลื่อนชนิดที่ I)

= P (ปฏิเสธ Ho|Ho เป็นจริง)

สัญลักษณ์ α คือความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่ I หรือระดับนัยสำคัญมีความหมายเหมือนกัน ระดับนัยสำคัญ 5% กำหนดความน่าจะเป็นของการปฏิเสธสมมติฐานเมื่อสมมติฐานเป็นจริงคือ 0.05

ข้อสังเกต ความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่ I คำนวณหาได้ในสภาวะที่สมมติฐานเป็นจริงเท่านั้น หากสมมติฐานผิดไม่อาจกระทำความคลาดเคลื่อนชนิดนี้

ความคลาดเคลื่อนชนิดที่ II คือการยอมรับสมมติฐาน Ho เมื่อสมมติฐาน Ho ผิด สมมติว่า

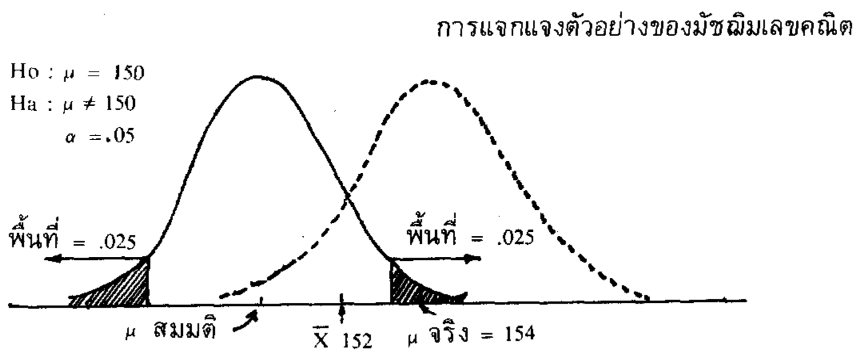
การทดสอบสมมติฐานที่ $\mu = 150$ เทียบกับ alternative hypothesis ที่ $\mu \neq 150$ ใช้ระดับนัยสำคัญ 5% เลือกตัวอย่างหนึ่งหาค่า $\bar{x}$ ได้ 152 มัชฌิมเลขคณิตของประชากรจริง ๆ คือ 154 ดูรูป 9.7 แสดงสมการแจกแจงตัวอย่าง การแจกแจงจริงแสดงด้วยเส้นประ 150 เป็นมัชฌิมเลขคณิตตามสมมติฐาน มัชฌิมเลขคณิตของตัวอย่างเป็นของการแจกแจงตัวอย่างจริง แต่ต้องการทดสอบสมมติฐานที่ $\mu = 150$ การประเมินมัชฌิมเลขคณิตจะต้องเป็นไปตามตำแหน่งของมัชฌิมเลขคณิตในการแจกแจงตัวอย่างแสดงด้วยเส้นทึบเปรียบเทียบกับการแจกแจงนั้น จะไม่มีการเบี่ยงเบนไปมากนัก (เบี่ยงเบนไปจากมัชฌิมเลขคณิต 150) ที่จะปฏิเสธ Ho เพราะฉะนั้นการตัดสินใจก็จะยอมรับสมมติฐานที่ $\mu = 150$ ซึ่งเป็นการตัดสินใจที่คลาดเคลื่อน เราก็ได้กระทำความคลาดเคลื่อนชนิดที่ II

ความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่ II กำหนดได้เป็น

β = P (ความคลาดเคลื่อนชนิดที่ II)

= P (ยอมรับ Ho|Ho ผิด)

ใช้อักษรกรีก β แสดงความน่าจะเป็นนี้ สังเกตว่าความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่ II หาค่าได้ในสภาวะที่สมมติฐานผิดเท่านั้น หากสมมติฐานถูกก็ไม่อาจกระทำความคลาดเคลื่อนชนิดนี้



รูปที่ 9.7 การทดสอบ Ho : Ho ผิดแต่ $\bar{X}$ นำไปสู่ความคลาดเคลื่อนชนิดที่สอง

9.7 ความคลาดเคลื่อนชนิดที่ I และความคลาดเคลื่อนชนิดที่ II

เราได้ศึกษาสองความคลาดเคลื่อนที่อาจเกิดขึ้นในการทดสอบสมมติฐาน ความคลาดเคลื่อนชนิดที่ I กับความคลาดเคลื่อนชนิดที่ II ความน่าจะเป็นของการกระทำความคลาดเคลื่อนแต่ละชนิดนิยามได้ดังนี้

ความคลาดเคลื่อนชนิดที่ I

α = P (ปฏิเสธ Ho|Ho เป็นจริง)

ความคลาดเคลื่อนชนิดที่ II

β = P (ยอมรับ Ho|Ho ผิด)

ในการทดสอบสมมติฐาน ความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่ I มองเห็นได้เพราะว่าผู้ทดสอบต้องเลือกค่า α ที่เขาประสงค์จะใช้ ส่วนการเสี่ยงภัยของความคลาดเคลื่อนชนิดที่ II ได้รับการพิจารณาโดยตรงน้อยมาก เป็นเพราะขนาดตัวอย่างใหญ่เพียงพอหรือการพิจารณาความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่ II พร้อมทั้งอิทธิพลของตัวประกอบที่มีผลต่อความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ II จะแสดงในหัวข้อต่อไป

9.8 อานุภาพของการทดสอบ (Power of Test)

การตัดสินใจยอมรับสมมติฐานที่ผิดเป็นพื้นที่ส่วนตรงข้ามกับส่วนที่ปฏิเสธสมมติฐานที่ผิด ความน่าจะเป็นของการปฏิเสธสมมติฐานที่ผิดคือ $(1-\beta)$

$(1-\beta)$ = P (ปฏิเสธ Ho|Ho ผิด)

$(1-\beta)$ เป็นความน่าจะเป็นของการอ้างความแตกต่างว่ามีนัยสำคัญหรือไม่เมื่อความแตกต่างจริงคำนวณหาค่าได้จริง ๆ ความน่าจะเป็นของการกระทำ $(1-\beta)$ เรียกว่าอานุภาพของการทดสอบระหว่างหลาย ๆ วิธีการของทดสอบวิธีการหนึ่งที่มีอานุภาพมากที่สุดคือว่าการเสนอความน่าจะเป็นที่มีค่ามากที่สุดที่จะเอาไปปฏิเสธ Ho เมื่อไรสมควรที่จะปฏิเสธสมมติฐาน

ในหัวข้อต่อไป เราจะตรวจสอบหลาย ๆ เงื่อนไขที่มีผลต่อ β เนื่องจากว่า β กับอานุภาพของการทดสอบเป็น Complementary กันหากลด β จะไปเพิ่มอานุภาพของการทดสอบ

9.9 ความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่ I และชนิดที่ II

สมมติว่าสมมติฐานต่อไปนี้ต้องการที่จะตรวจสอบ

Ho : $\mu = 80$

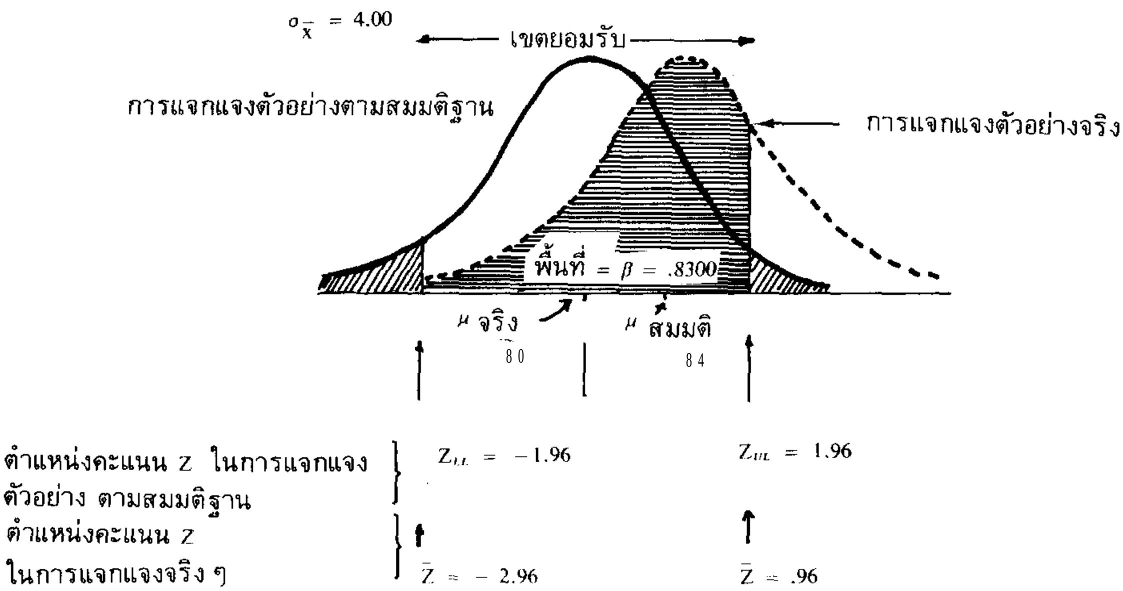
Ha : $\mu \neq 80$

หากขนาดตัวอย่างเท่ากัน 25 $\alpha = 0.05$ $\sigma = 20$ แล้ว $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 4.00$ ต้องการทดสอบสมมติฐานเราก็จำเป็นที่จะต้องทราบการแจกแจงตัวอย่างของมัชฌิมเลขคณิตเมื่อสมมติฐานเป็นจริง การแจกแจงตัวอย่างนี้แสดงด้วยเส้นทึบของรูปที่ 9.8 หากคำนวณมัชฌิมเลขคณิตของตัวอย่างตกอยู่ภายในขีดจำกัด $Z = \pm 1.96$ จะยอมรับ Ho

สมมติว่าสมมติฐานผิด ความจริงเป็น $\mu_{\text{จริง}} = 84$ ในการสุ่มตัวอย่าง ได้ค่าของ $\bar{X}$ จริง ๆ จะดำเนินตามการแจกแจงตัวอย่างจริงแสดงด้วยเส้นประในรูปที่ 9.8 หากมัชฌิมเลขคณิตของ

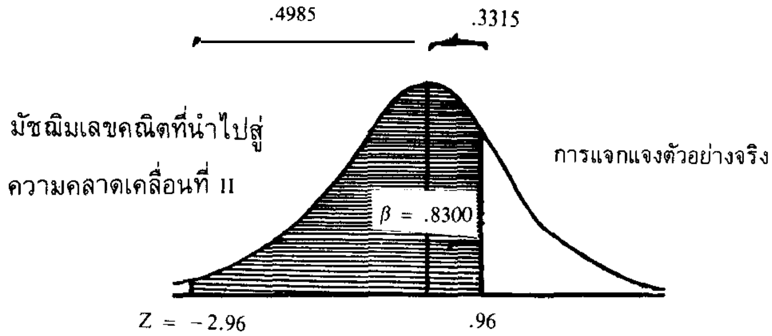
ตัวอย่างที่คำนวณได้เป็นค่าหนึ่งของค่าเหล่านั้นแสดงในส่วนแลเงาของการแจกแจง การตัดสินใจยอมรับ Ho : จะเป็นการกระทำความคลาดเคลื่อนชนิดที่ II ความน่าจะเป็นของการกระทำความคลาดเคลื่อนนี้จะเป็นเท่าไร พื้นที่ทั้งหมดภายใต้เส้นโค้ง (เส้นประ) จะเท่ากับพื้นที่แลเงาภายใต้เส้นโค้งนั้น

การคำนวณหาพื้นที่ในส่วนแลเงาของเส้นโค้งขอบเขตของมันต้องแสดงในรูปของคะแนน Z ขอบเขตที่อยู่ในคะแนน Z ที่เกี่ยวพันกับการแจกแจงตัวอย่างตามสมมติฐาน $Z_{LL} = -1.96$ กับ $Z_{UL} = +1.96$ เส้นทางต่อข้อแก้ปัญหาต้องแสดงความแตกต่างระหว่าง $\mu_{\text{สม}}$ กับ $\mu_{\text{จริง}}$ และเปลี่ยนเป็นหน่วยคะแนน Z $(\mu_{\text{สม}} - \mu_{\text{จริง}}) / \sigma_{\bar{X}} = 4.00/4.00 = 1.00$ ขอบเขตล่าง 1.96 หน่วยของคะแนน Z ต่ำกว่า $\mu_{\text{สม}}$ คือ (1.96 + 1.00) หน่วยของคะแนน Z ต่ำกว่า $\mu_{\text{จริง}}$ เพราะฉะนั้นที่ตั้งของการแจกแจงจริงคือ $Z_{LL} = -2.96$ ขอบเขตบน 1.96 หน่วยของคะแนน Z เหนือ $\mu_{\text{สม}}$ คือ (1.96 – 1.00) หน่วยของคะแนน Z เหนือ $\mu_{\text{จริง}}$ เพราะฉะนั้นที่ตั้งของการแจกแจงจริงคือ $Z_{UL} = +0.96$ การคำนวณหาพื้นที่ระหว่างขอบเขตเหล่านี้อาจทำสำเร็จลงได้ด้วยวิธีการแจกแจงปกติ



รูปที่ 9.8 ความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ II เมื่อทดสอบสมมติฐานที่ $\mu = \mu_{\text{สม}}$

พื้นที่ระหว่าง Z_{ll} กับมัชฌิมเลขคณิตคือ 0.4985 และพื้นที่ระหว่างมัชฌิมเลขคณิตกับ Z_{ul} คือ 0.3315 ตามตารางของพื้นที่ภายใต้เส้นโค้งปกติ ผลบวกของสองพื้นที่เหล่านี้คือ 0.8300 ความน่าจะเป็นของการคำนวณมัชฌิมเลขคณิตตกอยู่ภายในสองขอบเขตคือ β ขั้นสุดท้ายของข้อแก้ปัญหาแสดงได้ดังรูปที่ 9.9



รูปที่ 9.9 ความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ II

หากยังไม่สามารถกำหนดค่า $\mu_{\text{จริง}}$ ก็ไม่สามารถคำนวณค่า β ได้ ถึงแม้ไม่ทราบค่านี้ ความสามารถที่จะคำนวณ β ตามค่าแตกต่างของ $\mu_{\text{จริง}}$ ทำให้น่าจะเห็นผลที่เกิดขึ้นของการทดสอบสมมติฐานภายใต้เหตุการณ์ต่าง ๆ ด้วยเหตุนี้จึงต้องวางแผนการทดสอบตามเหตุการณ์นั้น การคำนวณค่า β ดังแสดงข้างต้นเป็นการแสดงการทดสอบสมมติฐานเกี่ยวกับมัชฌิมเลขคณิตตัวเดียว แนวความคิดนี้สามารถประยุกต์กับการทดสอบสมมติฐานอื่น ๆ ดังตัวอย่างเช่น สามารถนำไปใช้สำหรับทดสอบความแตกต่างระหว่างสองมัชฌิมเลขคณิต ในตัวอย่างของเรา เราทราบส่วนเบี่ยงเบนมาตรฐานของประชากร σ เมื่อไรที่ต้องประมาณค่า σ ด้วย s ที่คำนวณได้จากตัวอย่าง เส้นโค้งปกติก็เป็นตัวแบบโดยประมาณเท่านั้น ค่า β จะมากขึ้นขณะที่ขนาดตัวอย่างลดลง ตัวแบบอื่น ๆ อาจเป็นประโยชน์สำหรับเหตุการณ์เหล่านี้ แต่ในการแสดงเหตุผลพื้นฐานอยู่ในที่ที่ไม่สามารถกำหนดวิธีปฏิบัติให้เป็นระบบได้

ตัวอย่างที่ 9.1 เป็นที่ทราบกันว่ามีวัคซีนชนิดหนึ่งจะมีผลเพียง 25% ภายหลังระยะเวลา 2 ปีเท่านั้น เพื่อที่จะกำหนดวัคซีนชนิดใหม่และแพงกว่าว่ามีประสิทธิภาพในการป้องกันไวรัสชนิดเดียวกันสำหรับในกาลระยะยาว สุ่มเลือกบุคคลที่ฉีดวัคซีนชนิดใหม่มา 20 คน ถ้าหากว่าน้อยกว่า 9 คนของจำนวน 20 คน ติดเชื้อไวรัสภายในระยะเวลา 2 ปี วัคซีนใหม่จะได้รับการพิจารณาว่ามีประสิทธิภาพดีกว่าที่ใช้อยู่ในปัจจุบัน เมื่อไรที่ 9 คน หรือมากกว่าเกินระยะเวลา 2 ปี โดยปราศจากการติดเชื้อไวรัส ถ้าหากว่า x เป็นจำนวนบุคคลผู้ซึ่งสุขภาพสมบูรณ์สำหรับอย่างน้อย 2 ปี ความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่หนึ่ง (α) คำนวณหาได้

$$\alpha = P\text{ (ความคลาดเคลื่อนชนิดที่หนึ่ง)}$$
$$= P\ (X \geq 9 \mid p = 1/4)$$

$$= \sum_{x=9}^{20} \binom{20}{x} (1/4)^x (3/4)^{20-x}$$

$$= 1 - \sum_{x=0}^{8} \binom{20}{x} (1/4)^x (3/4)^{20-x}$$

$$= 1 - 0.959$$

$$= 0.0409$$

เรากำลังทดสอบสมมติฐาน $p = 1/4$ ที่ $\alpha = 0.0409$ ซึ่งมีค่าน้อยมาก เพราะฉะนั้นจึงไม่เหมาะที่จะทำความคลาดเคลื่อนชนิดที่หนึ่ง โดยมากไม่เคยเป็น 9 คนหรือมากกว่าที่พ้นอันตรายไวรัสสำหรับระยะเวลา 2 ปี ในการใช้วัคซีนชนิดใหม่ซึ่งเทียบกับบุคคลหนึ่งในตลาด

ความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่สอง β ไม่อาจคำนวณได้เว้นแต่ว่าเรามี alternative hypothesis เฉพาะ หากเราทดสอบสมมติฐานว่า $p = 1/4$ เทียบกับ alternative hypothesis ที่ $p = 1/2$ แล้วเราสามารถคำนวณความน่าจะเป็นของการยอมรับ Ho เมื่อ Ho ผิด เราคำนวณหาความน่าจะเป็นที่น้อยกว่า 9 คนในกลุ่มเกินระยะเวลา 2 ปี เมื่อ $p = 1/2$ ในกรณีนี้

$$\beta = P\ (\text{ความคลาดเคลื่อนชนิดที่สอง})$$

$$= P\ (X < 9/p = (1/2)$$

$$= \sum_{x=0}^{8} \binom{20}{x} (1/2)^x (1/2)^{20-x}$$

$$= 0.2517$$

นี้เป็นความน่าจะเป็นค่อนข้างสูงในการดำเนินการทดสอบที่เลว แต่เหมาะที่เราจะปฏิเสธวัคซีนชนิดใหม่ เมื่อไรที่วัคซีนชนิดใหม่มีประสิทธิภาพเหนือกว่าวัคซีนที่ใช้อยู่ปัจจุบัน เราต้องการทั้งความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่ I กับที่ II ให้มีค่าน้อย วิธีป้องกันความคลาดเคลื่อนชนิดที่ II คือเมื่อค่า p เป็นอย่างน้อย 0.7 หาก $p = 0.7$ การดำเนินการทดสอบได้

$$\beta = P\ (\text{ความคลาดเคลื่อนชนิดที่สอง})$$

$$= P\ (X < 9 \mid p = 0.7)$$

$$= \sum_{x=0}^{8} \binom{20}{x} (0.7)^x (0.3)^{20-x}$$

$$= 0.0015 \qquad = 0.0051$$

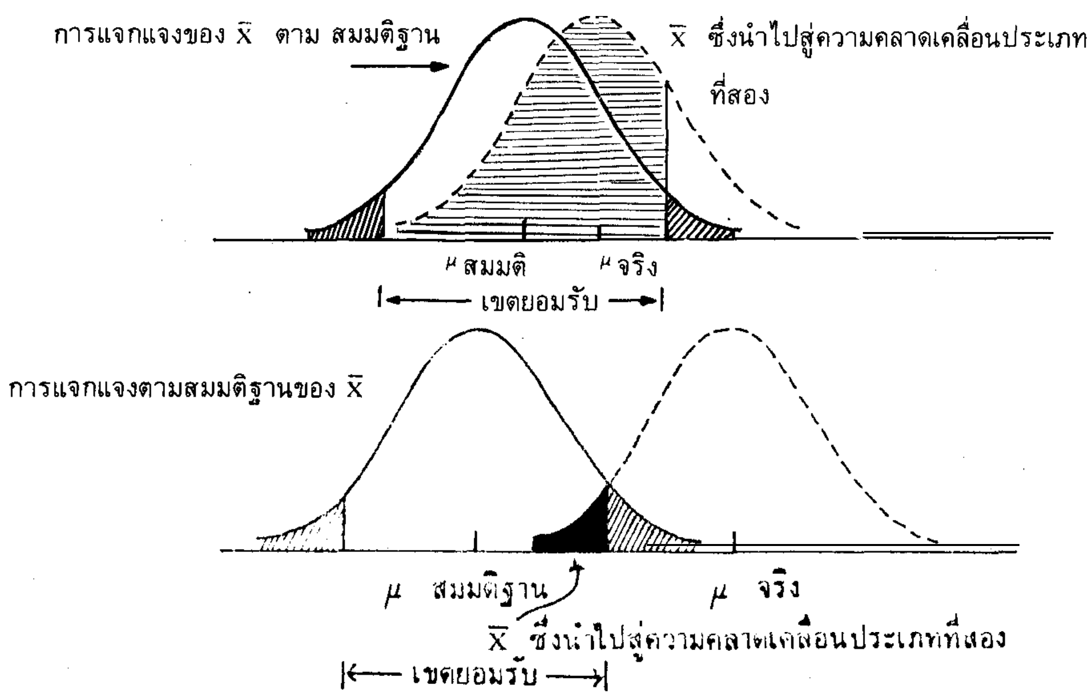
ความน่าจะเป็นของความคลาดเคลื่อนชนิดที่สองมีค่าน้อย จึงไม่ควรที่จะปฏิเสธวัคซีนชนิดใหม่ เมื่อมีผล 70% ภายหลังระยะเวลา 2 ปี

9.10 ตัวประกอบที่มีผลต่อความคลาดเคลื่อนชนิดที่ II

9.10.1 ระยะทางระหว่างมัชฌิมเลขคณิตจริงกับมัชฌิมเลขคณิตตามสมมติฐาน

ความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ II มีความสัมพันธ์ต่อระยะทางระหว่างมัชฌิมเลขคณิตตามสมมติฐานกับมัชฌิมเลขคณิตจริง หากระยะทางใกล้กันมาก $\bar{X}$ คำนวณได้จากการแจกแจงตัวอย่างจริงจะเหลื่อมพื้นที่ยอมรับมาก เพราะฉะนั้นหลาย ๆ ค่าของตัวอย่างจะนำไปสู่การยอมรับที่ผิดของสมมติฐาน Ho รูปบนของรูปที่ 9.10 แสดงถึงสภาวะนี้มัชฌิมเลขคณิตที่คำนวณได้ตลอดการแจกแจงสุ่มจะดำเนินไปตามการแจกแจงที่แสดงด้วยเส้นประ พื้นที่เกินครึ่งหนึ่งของทั้งหมดจะนำไปสู่การยอมรับที่ผิดของสมมติฐานที่ $\mu = \mu_{สม}$

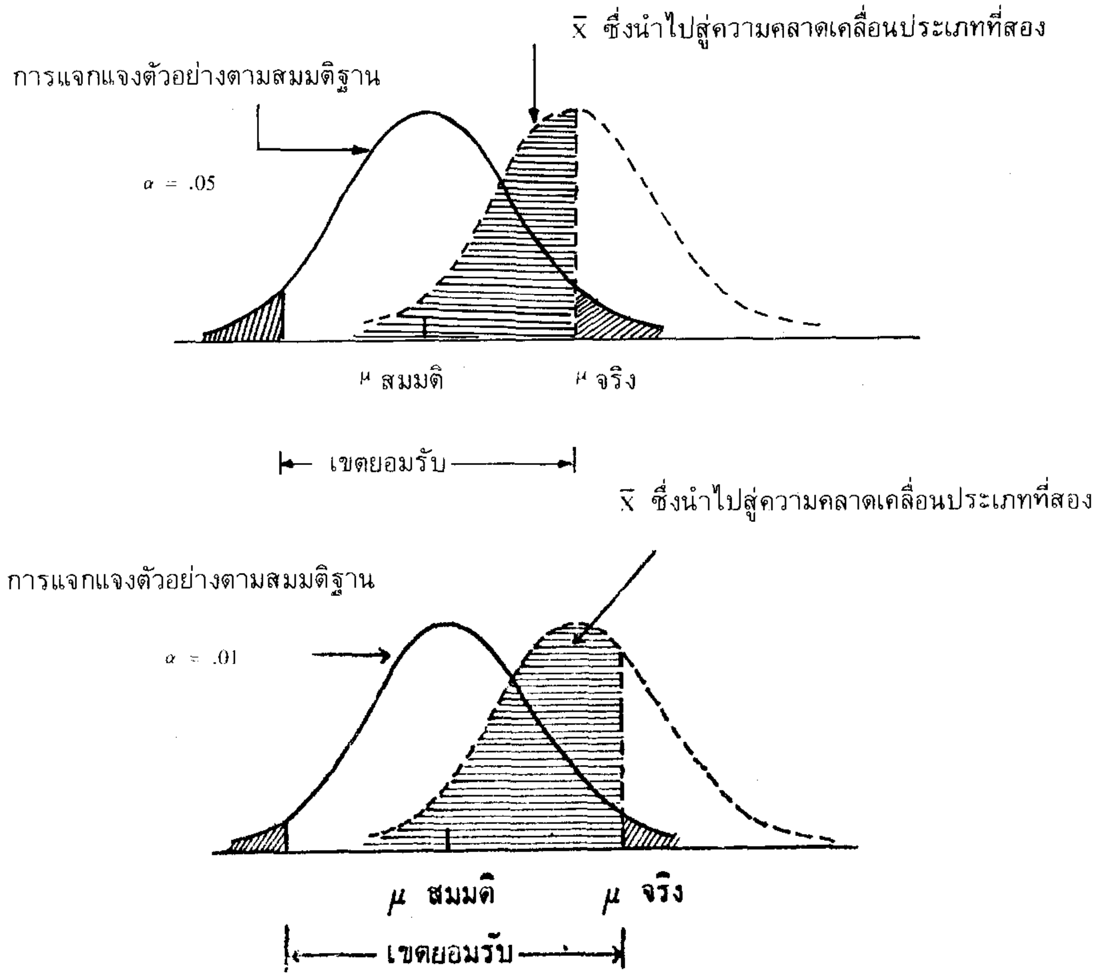
ส่วนรูปข้างล่างของรูปที่ 9.10 มัชฌิมเลขคณิตจริงกับมัชฌิมเลขคณิตตามสมมติฐานมีความแตกต่างกันมาก ในที่นี้จะมีสองสามตัวอย่างที่นำไปสู่การยอมรับ Ho โดยทั่ว ๆ ไปเมื่อไรสมมติฐานผิดระยะทางระหว่าง $\mu_{จริง}$ กับ $\mu_{สม}$ มากขึ้นความน่าจะเป็นของการยอมรับสมมติฐานผิดน้อยลง



รูปที่ 9.10 ความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ II เป็นฟังก์ชันของขนาดของผลต่างระหว่าง μ_{hyp} กับ μ_{true}

9.10.2 การเลือกระดับนัยสำคัญ (α)

β มีความสัมพันธ์ต่อการเลือก α ในรูปที่ 9.11 รูปบนแสดงความน่าจะเป็นของการยอมรับ Ho เมื่อ Ho ผิดและ $\alpha = 0.05$ ส่วนการแสดงรูปข้างล่างในสภาวะเหมือนกัน นอกจากว่า $\alpha = 0.01$ ทั้งสองรูป มัชฌิมเลขคณิตของตัวอย่างแทนได้ด้วยพื้นที่แลเงาในการแจกแจงแสดงด้วยเส้นประเป็นพื้นที่ส่วนที่ยอมรับ สัดส่วนของความถี่ของมัชฌิมเลขคณิตนั้นเมื่อไร $\alpha = .01$ มีค่ามากกว่าเมื่อไร $\alpha = 0.05$ เพราะฉะนั้นจึงมีโอกาสหามัชฌิมเลขคณิตของตัวอย่างซึ่งจะนำไปสู่การยอมรับ Ho ผิดมากกว่าเมื่อไร α มีค่าน้อยกว่า โดยทั่ว ๆ ไป การลดการเสี่ยงภัยของความคลาดเคลื่อนชนิดที่ I จะเพิ่มการเสี่ยงภัยของการกระทำความคลาดเคลื่อนชนิดที่ II



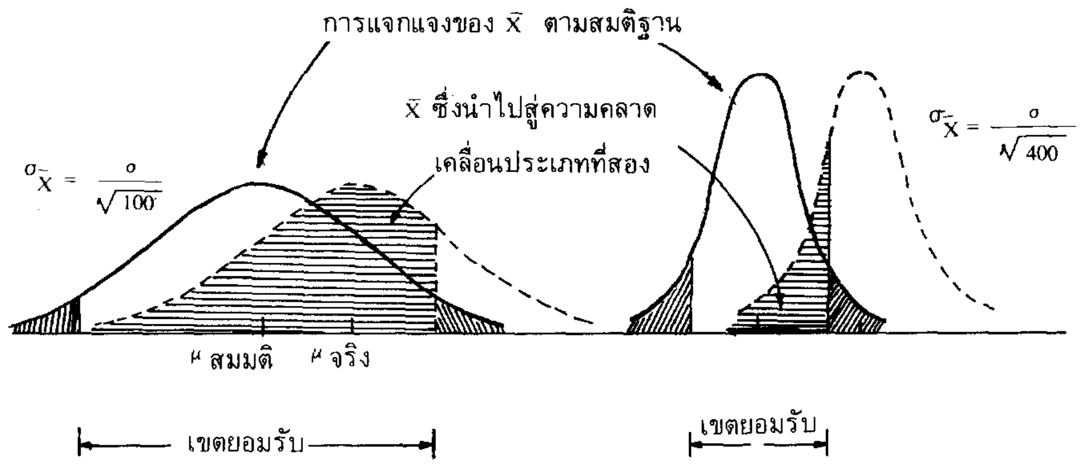
รูปที่ 9.11 ความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ II เป็นฟังก์ชันของ α

ความสัมพันธ์ระหว่าง α กับ β สะท้อนถึงการเปลี่ยนแปลงในทางปฏิบัติของการวิจัยมา 40 กว่าปีแล้ว ในสมัยก่อน เขาต้องการค่า $Z \geqslant \pm 3$ ก่อนเปิดเผยการปฏิเสธสมมติฐาน ค่า Z นี้มีข้อแย้งกันมากเกินไปที่ค่านี้ควรจะเกิดขึ้นน้อยกว่าสามครั้งในหนึ่งพันครั้งของการแจกแจงปกติ ในทางปฏิบัติทุกวันนี้ การปฏิเสธสมมติฐานเมื่อขัดแย้งกันเกิดขึ้นมากเกินไปที่ควรจะเกิดขึ้น 5% ของครั้งหรือน้อยกว่า ($Z = \pm 1.96$) หรือ 1% ของครั้ง หรือน้อยกว่า ($Z = \pm 2.58$) เป็นเรื่องธรรมดาและเป็นที่พอใจมากกว่า

การพิจารณาเลือก α เบื้องต้นควรเป็นไปตามเหตุผลของการทดลองเป็นสำคัญ การเลือก α ให้มีค่าน้อยที่สุดก็จะมีผลต่อ β การควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ II ให้พอเหมาะสามารถทำได้โดยการเลือกขนาดของตัวอย่างซึ่งจะกล่าวในหัวข้อต่อไป

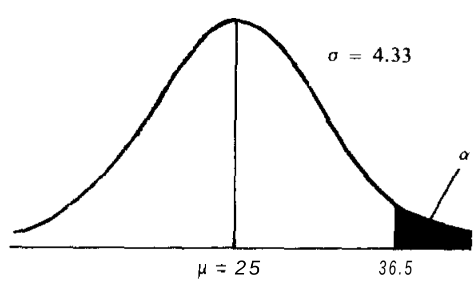
9.10.3 ขนาดตัวอย่าง

ความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่ II มีความสัมพันธ์กับขนาดตัวอย่าง พิจารณาการทดสอบสมมติฐานที่ $\mu = \mu_{สม}$ เมื่อขนาดตัวอย่างเป็น 100 ความคลาดเคลื่อนมาตรฐานของมัชฌิมเลขคณิตจะเป็น $\sigma_{\bar{X}} = \sigma/\sqrt{100}$ หากขนาดตัวอย่างเป็น 400 ความคลาดเคลื่อนมาตรฐานของมัชฌิมเลขคณิตควรจะเป็น $\sigma_{\bar{X}} = \sigma/\sqrt{400}$ เท่ากับครึ่งหนึ่งเท่านั้น รูปที่ 9.12 แสดงถึงการแจกแจงตัวอย่างเมื่อตัวประกอบอื่น ๆ เหมือนกันหมดนอกจากขนาดตัวอย่างเท่านั้นที่ไม่เหมือนกันโดยทั่ว ๆ ไป ขนาดตัวอย่างโตขึ้นส่วนเบี่ยงเบนมาตรฐาน $\sigma_{\bar{X}}$ ของการแจกแจงตัวอย่างของ $\bar{x}$ จะมีค่าน้อยลง เปรียบเทียบสองการแจกแจงแสดงว่าเมื่อขนาดตัวอย่างใหญ่ขึ้น การเหลื่อมระหว่างสองการแจกแจงจะน้อยลง เพราะฉะนั้น การเสี่ยงภัยจะน้อยลงในการเลือกตัวอย่างหนึ่งที่จะนำไปสู่การยอมรับสมมติฐานที่ผิด นั่นคือ ขนาดตัวอย่างใหญ่ขึ้น ความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่ II น้อยลง



รูปที่ 9.12 ความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ II เป็นฟังก์ชั่นของขนาดตัวอย่าง

ตัวอย่างที่ 9.2 จากตัวอย่างที่ 9.1 หากสุ่มตัวอย่างเท่ากับ 100 และมีอยู่ 37 คน หรือมากกว่าเกินระยะ 2 ปี โดยปราศจากการติดเชื้อไวรัส เราปฏิเสธสมมติฐานที่ $P = \frac{1}{4}$ และยอมรับ alternative hypothesis $P > \frac{1}{4}$ ค่าวิกฤต คือ 36.5 ค่าที่อาจเป็นไปได้ทั้งหมดเหนือค่า 36.5 เป็นองค์ประกอบของพื้นที่วิกฤต ค่าที่อาจเป็นไปได้ทั้งหมดใต้ค่า 36.5 เป็นพื้นที่ยอมรับ



รูป ก. ความน่าจะเป็นของความคลาดเคลื่อนชนิดที่หนึ่ง

เพื่อที่จะคำนวณความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่หนึ่ง เราจะใช้เส้นโค้งปกติพร้อมด้วย

$$\mu = np = 100\left(\frac{1}{4}\right) = 25; \sigma = \sqrt{npq} = \sqrt{100\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)} = 4.33$$

จากรูป

$$\alpha = P\ (\text{ความคลาดเคลื่อนชนิดที่หนึ่ง})$$
$$= P\ (X > 36.5 \mid H0\ \text{เป็นจริง})$$

ค่า Z ที่สมนัยกับ X = 36.5 คือ

$$Z = \frac{36.5 - 25}{4.33} = 2.656$$

เพราะฉะนั้น

$$\alpha = P\ (Z > 2.656)$$
$$= 1 - P\ (Z \leqslant 2.656)$$
$$= 1 - 0.9961$$
$$= 0.0039$$

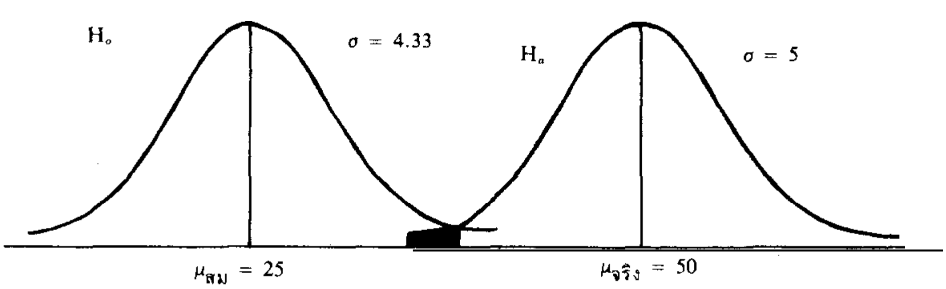
หาก Ho ผิดและค่าจริงเป็น $p = \frac{1}{2}$ เราสามารถคำนวณความน่าจะเป็นของความคลาดเคลื่อนชนิดที่สองโดยการประมาณเส้นโค้งปกติพร้อมด้วย

$\mu = np = (100)\ (\frac{1}{2}) = 50;\ \sigma = \sqrt{npq} = \sqrt{(100)\ (\frac{1}{2})\ (\frac{1}{2})} = 5$ ความน่าจะเป็นของการยอมรับ Ho ผิด กำหนดได้โดยพื้นที่แลเงาในรูป ข. ด้วยเหตุนี้

$$\beta = P\ (\text{ความคลาดเคลื่อนชนิดที่สอง})$$
$$= P\ (X \leqslant 36.5 \mid Ha\ \text{เป็นจริง})$$

ค่า Z ที่สมนัยกับ X = 36.5 คือ

$$Z = \frac{36.5 - 50}{5} = -2.7$$



รูป ข. ความน่าจะเป็นของความคลาดเคลื่อนชนิดที่สอง

เพราะฉะนั้น $\beta = P(z < -2.7)$

$= 0.0035$

ความคลาดเคลื่อนชนิดที่หนึ่งกับชนิดที่สองเกิดขึ้นได้ยาก หากการทดลองประกอบด้วยตัวอย่างขนาด 100 คน

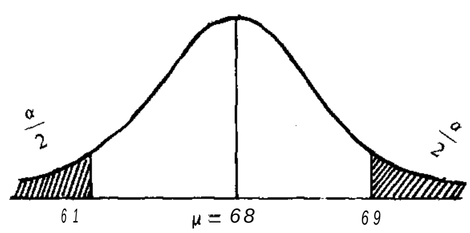
ความคิดรวบยอดที่ได้กล่าวข้างต้นสามารถมองเห็นด้วยกราฟ เมื่อประชากรเป็นแบบต่อเนื่อง พิจารณาสมมติฐานที่ความสูงเฉลี่ยของนักศึกษาในวิทยาลัยแห่งหนึ่งเป็น 68 นิ้ว เทียบกับ alternative hypothesis ที่ความสูงเฉลี่ยของนักศึกษาไม่เท่ากับ 68 นิ้ว นั่นคือ

$H_o : \mu = 68$

$H_a : \mu \neq 68$ ($\mu < 68$ หรือ $\mu > 68$)

สมมติว่าส่วนเบี่ยงเบนมาตรฐานของประชากร $\sigma = 3.6$ ขนาดตัวอย่าง $n = 36$ $\sigma_{\bar{X}} = \sigma / \sqrt{n} = 3.6 / 6 = 0.6$ พื้นที่วิกฤตแสดงได้โดยพื้นที่แลเงาในรูปโดยการเลือก $\bar{X} < 67$ กับ $\bar{X} > 69$ เพราะฉะนั้น พื้นที่ยอมรับสมมติฐานคือ $67 < \bar{X} < 69$ ด้วยเหตุนี้ $\bar{X}$ ของเราตกภายในพื้นที่วิกฤตก็ปฏิเสธ Ho

ความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่หนึ่งหรือระดับนัยสำคัญของการทดสอบของเราเท่ากับผลบวกของพื้นที่ที่แลเงาในแต่ละข้างของการแจกแจงในรูป ค. เพราะฉะนั้น



รูป ค. พื้นที่วิกฤตสำหรับการทดสอบ $\mu = 68$ เทียบกับ $\mu \neq 68$

$\alpha = P(\bar{X}_1 < 67 \mid Ho$ เป็นจริง$) + P(\bar{X}_2 > 69 \mid Ho$ เป็นจริง$)$ ค่า Z ที่สมนัยกับ $\bar{X}_1 = 67$ กับ $\bar{X}_2 = 69$ เมื่อ Ho เป็นจริงคือ

$$Z_1 = \frac{67 - 68}{0.6} = -1.67 \; ; Z_2 = \frac{69 - 68}{0.6} = 1.67$$

เพราะฉะนั้น

$$\begin{aligned} \alpha &= P(Z < -1.67) + P(Z > 1.67) \\ &= 2P(Z < -1.67) \\ &= 0.0950 \end{aligned}$$

ดังนั้น 9.5% ของพื้นที่ทั้งหมดที่ควรนำเราไปปฏิเสธ $\mu = 68$ นิ้ว เมื่อ μ เป็นจริง เพื่อที่จะลด α เราจะต้องเลือกขนาดตัวอย่างเพิ่มขึ้นหรือขยายพื้นที่ยอมรับสมมติฐานให้เราเพิ่มขนาดตัวอย่างเป็น n = 64 แล้ว

$$\sigma_{\bar{X}} = 3.6/8 = 0.45$$

$$Z_1 = \frac{67 - 68}{0.45} = -2.22$$

$$Z_2 = \frac{69 - 68}{0.45} = 2.22$$

ดังนั้น

$$\begin{aligned} \alpha &= P(Z < -2.22) + P(Z > 2.22) \\ &= 2P(Z < -2.22) \\ &= 0.0264 \end{aligned}$$

การลด α ไม่เป็นการเพียงพอที่จะประกันวิธีการดำเนินการทดลองที่ดี เราต้องประเมิน β สำหรับ Ha ที่ต่าง ๆ ที่เรารู้สึกว่าควรยอมรับถ้าเป็นจริง

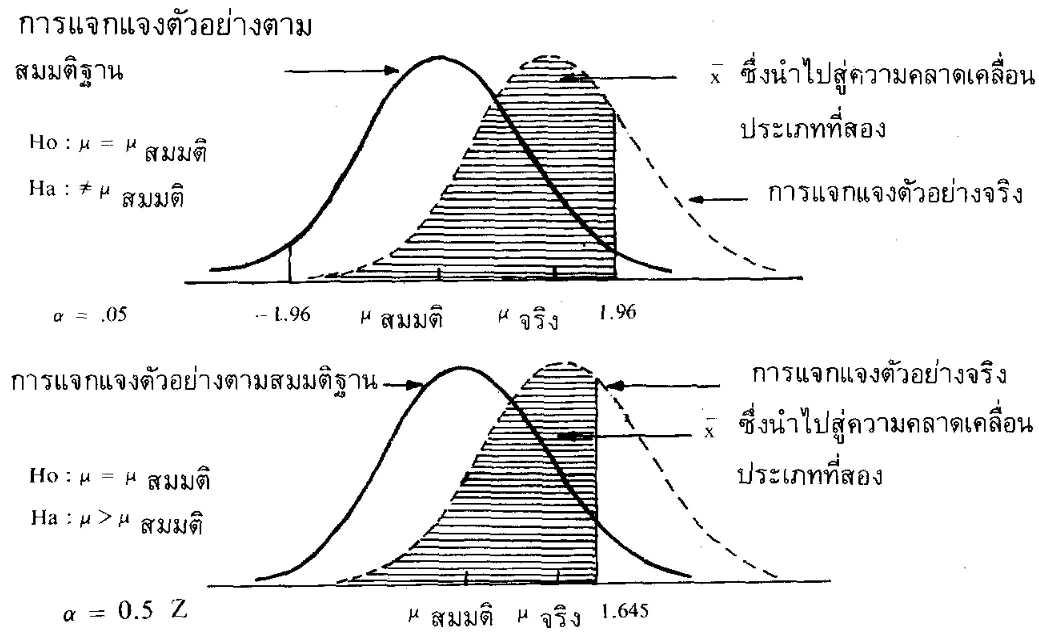
9.10.4 การเปลี่ยนแปลงจากการชั่ง ตวง วัด

เราพบว่าการเพิ่มขนาดตัวอย่างหรือลดขนาดของ σ จะลดการเสี่ยงภัยของความคลาดเคลื่อนชนิดที่ II หากมองดูอย่างผิวเผินก็อาจคิดว่าเหลือวิสัยที่ผู้ตรวจสอบจะควบคุมได้ แท้จริงแล้วโอกาสนี้ยังเปิดกว้างอยู่เสมอ แต่การวางแผนล่วงหน้าเป็นสิ่งจำเป็นอย่างยิ่ง

ส่วนเบี่ยงเบนมาตรฐานของเซ็ทจากการวัดสะท้อนถึงการเปลี่ยนแปลงเนื่องมาจากตัวประกอบที่เราประสงค์จะศึกษา แต่ยังสะท้อนถึงการเปลี่ยนแปลงเนื่องมาจากปัจจัยที่ไม่เกี่ยวข้อง การเปลี่ยนแปลงที่ไม่เกี่ยวข้องของปัจจัยมีแนวโน้มที่จะเพิ่ม σ มากกว่าสิ่งอื่น ๆ เพราะฉะนั้นความพยายามที่จะขจัดปัจจัยนั้นจะมีแนวโน้มลด σ และแล้วจะลด β ด้วย ดังตัวอย่างเช่น การทดสอบในสาขาวิทยาศาสตร์ที่เกี่ยวกับพฤติกรรม ซึ่งมีข้อยุ่งยากในบางสิ่งบางอย่างที่

เปลี่ยนแปลงได้ไปในทางที่เลว การปรับปรุงเครื่องมือที่ใช้วัดจะมีผลเชื่อถือได้ในการลดค่า σ โดยให้ตัวประกอบอื่น ๆ คงที่

9.10.5 การทดสอบข้างเดียวเทียบกับการทดสอบสองข้าง การเสี่ยงภัยของความคลาดเคลื่อนชนิดที่สองจะมีต่อการเลือก alternative hypothesis ในการทดสอบข้างเดียวของ Ho : $\mu = \mu_{\text{สม}}$ เทียบกับ alternative hypothesis Ha : $\mu > \mu_{\text{จริง}}$ หาก Ho ผิดและ $\mu_{\text{จริง}} > \mu_{\text{สม}}$ ดูรูปที่ 9.13 แสดงการทดสอบ Ho เมื่อ $\alpha = 0.05$ ทั้งสองรูปใช้แทนเงื่อนไขการทดสอบเหมือนกัน เว้นแต่ลักษณะของ alternative hypotheis เมื่อไร Ha ไม่มีทิศทาง $z = \pm 1.96$ และเมื่อไร Ha มีทิศทาง $z = 1.645$ เปรียบเทียบสัดส่วนของมัชฌิมเลขคณิตของตัวอย่างนำไปสู่การยอมรับ Ho ที่ผิด ถึงจะน้อยกว่าการทดสอบข้างเดียว โดยที่ตัวประกอบอื่น ๆ คงที่ ความน่าจะเป็นของการกระทำความคลาดเคลื่อนชนิดที่ II ของการทดสอบข้างเดียวน้อยกว่าการทดสอบสองข้างจากหลักความจริงที่ว่าการทดสอบข้างเดียวมีอานุภาพมากกว่า (อย่างเช่น รับภาระการเสี่ยงภัยของความคลาดเคลื่อนชนิดที่สองน้อยกว่า)



รูปที่ 9.13 ความคลาดเคลื่อนชนิดที่สองในการทดสอบข้างเดียว และการทดสอบสองข้างเมื่อ $\alpha = 0.05$

9.11 **การทดสอบเกี่ยวกับมัชฌิมเลขคณิต** (ตัวอย่างขนาดใหญ่)

การทดสอบเกี่ยวกับมัชฌิมเลขคณิตสำหรับประชากรหนึ่งแบ่งตามขนาดของตัวอย่างที่ใช้ในการทดสอบในที่นี้เราพิจารณา $n > 30$ ซึ่งอนุญาตให้ใช้ทฤษฎีขีดจำกัดส่วนกลางและคะแนน Z ส่วนในหัวข้อต่อไปในเมื่อ $n \leq 30$ และประชากรมีการแจกแจงปกติ เราใช้ตัวสถิติ t สำหรับ $n > 30$ เป็นที่ยอมรับกันอย่างกว้างขวางว่า s^2 เป็นตัวค่าประมาณที่สมเหตุสมผลของ σ^2 และเป็นผลเปรียบเทียบเหมือนกับทราบ σ^2 สำหรับตัวอย่างสุ่มของ $n \leq 30$ ไม่มีข้ออ้างเช่นนั้นแต่จำเป็นจะต้องใช้รูปใหม่

กระบวนการทดสอบและการตีความผลที่ได้ของการทดสอบเกี่ยวกับมัชฌิมเลขคณิต alternative hypothesis มีสามชนิดด้วยกัน (1) น้อยกว่า (<) (2) ไม่เท่ากัน (≠) และ (3) มากกว่า ซึ่งจะแสดงในแต่ละตัวอย่างต่อไป

9.12 วิธีการดำเนินงานทั่ว ๆ ไปสำหรับการทดสอบสมมติฐาน

เหตุผลและการดำเนินงานทั่ว ๆ สำหรับการทดสอบสมมติฐานทางสถิติทั้งหมด ไม่ว่าจะเป็นมัชฌิมเลขคณิต ความถี่หรือคุณลักษณะอื่น ๆ ของประชากรย่อมมีความสำคัญเหมือนกันทั้งในกรณีสองตัวอย่างหรือหลาย ๆ ตัวอย่างตามลำดับขั้นดังนี้

1. ตั้งสมมติฐานเกี่ยวกับตัวพารามิเตอร์ของประชากร (อย่างเช่น มัชฌิมเลขคณิตของประชากร)
2. เลือกตัวอย่างสุ่มหนึ่งจากประชากรและคำนวณค่าของตัวสถิติ
3. ตรวจสอบคุณลักษณะของการแจกแจงตัวอย่างสุ่มของตัวสถิติภายใต้การพิจารณาเพื่อศึกษาผลลัพธ์ของตัวอย่าง (sample outcomes) อะไรควรจะเกิดขึ้นถ้าสมมติฐานเป็นจริง
4. ยอมรับสมมติฐาน หากผลลัพธ์ของตัวอย่างอยู่ในขีดของผลลัพธ์ที่ได้คาดหวัง ถ้าสมมติฐานเป็นจริงและปฏิเสธสำหรับผลลัพธ์อื่น ๆ

9.13 การประมาณค่าความคลาดเคลื่อนมาตรฐานของมัชฌิมเลขคณิตเมื่อไรที่ไม่ทราบ σ

ในการทดสอบสมมติฐาน มีอยู่สิ่งหนึ่งที่จำเป็นจะต้องคำนวณหาคือ ส่วนเบี่ยงเบนมาตรฐานของการแจกแจงตัวอย่างสุ่มของมัชฌิมเลขคณิตซึ่งเราได้เรียนมาแล้วในบทก่อน ๆ ปริมาณดังนี้เรียกว่าความคลาดเคลื่อนมาตรฐานของมัชฌิมเลขคณิต เขียนได้เป็น

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

หากเราทราบค่าของ σ ก็เป็นการดีไป แต่โดยส่วนมากแล้วไม่ทราบส่วนเบี่ยงเบนมาตรฐานของประชากร ในเหตุการณ์เช่นนั้น ก็จำเป็นต้องประมาณค่าจากตัวอย่างและตัวสถิติตัวนั้นคือส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง s ซึ่งเป็นตัวค่าประมาณที่ดีของ σ และมีสูตรการคำนวณได้เป็น

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{(n - 1)}}$$

เราแทนค่า σ ด้วย s และเปลี่ยนสัญลักษณ์สำหรับความคลาดเคลื่อนมาตรฐานของมัชฌิมเลขคณิตจาก $\sigma_{\bar{X}}$ เป็น $s_{\bar{X}}$ ดังสูตร

$$s_{\bar{X}} = \frac{s}{\sqrt{n}}$$

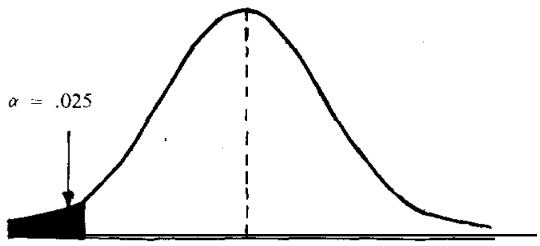
การแทนค่าประมาณสำหรับค่าจริงของ σ จะนำไปสู่ความคลาดเคลื่อนในการดำเนินการโดยทั่ว ๆ ไป ความคลาดเคลื่อนนี้จะหมดไปในกรณีตัวอย่างที่มีขนาดใหญ่ $(n > 30)$ ส่วนในกรณีตัวอย่างขนาดเล็กก็จะใช้การทดสอบแบบ t ซึ่งจะกล่าวในหัวข้อต่อไป

ตัวอย่างที่ 9.3 มหาวิทยาลัยแห่งหนึ่งต้องหาวิธีการตัดเวลาเรียนโดยไม่ให้เกิดความยุ่งยากต่อกระบวนการเรียนในเวลาการทดลองมีข้อเสนอแนะหนึ่งต้องการลดเวลาเลิกระหว่างชั้นเรียนปัจจุบันห้านาที โดยได้เลือกตัวอย่างสุ่มหนึ่งประกอบด้วยนักศึกษา 36 คน ใช้เวลาเปลี่ยนชั้นเรียนเฉลี่ย 4.85 นาที ส่วนเบี่ยงเบนมาตรฐาน 0.5 นาที ทดสอบว่าเวลาเฉลี่ยควรจะลดหรือไม่ที่ระดับนัยสำคัญ $\alpha = 0.025$

วิธีทำ ขั้นแรกตั้งเขตวิกฤตและทดสอบ μ = เวลาเฉลี่ยจริง (ที่ต้องการระหว่างการเปลี่ยนชั้นเรียน) การทดสอบก็เป็นการทดสอบข้างด้านซ้ายมือเนื่องจากว่าเป็นปัญหาเกี่ยวกับการลดเวลา

1. Ho : $\mu = 5.0$ นาที
   Ha : $\mu < 5.0$ นี้ต้องการเขตวิกฤตด้านเดียว
2. สำหรับ $\alpha = 0.025$, $n = 36$ ค่าทดสอบที่เหมาะสมคือคะแนน Z

ทดสอบ ปฏิเสธ Ho หากคำนวณ $z < -1.96$
ไม่สามารถปฏิเสธ Ho หากคำนวณ $z > -1.96$
สงวนการตัดสินหากคำนวณ $z = -1.96$



เขตสำหรับปฏิเสธ Ho   เขตสำหรับยอมรับ Ho

3. ใช้เวลาการทดลอง

$$\text{คำนวณ } Z = \frac{4.85 - 5.00}{(.5)/\sqrt{36}} = \frac{-.15}{(1/12)} = -.15 \times 12 = -1.80$$

4. การตัดสินใจ เราไม่สามารถปฏิเสธ Ho สำหรับ $\alpha = 0.025$ (เนื่องจากว่า $-1.80 > -1.96$)

นักศึกษาเหล่านี้ไม่สามารถใช้เวลาเปลี่ยนชั้นเรียนน้อยกว่า 5.0 นาที ไม่ควรลดเวลาเลิกระหว่างชั้นเรียน ในขั้นที่ 2 หากคำนวณค่า Z ได้ -1.96 โดยประมาณก็ไม่ควรจะทำการตัดสินใจ ถึงแม้ว่าจะเป็นกฎที่หยาบมากก็สงวนการตัดสินใจไว้เมื่อไร คำนวณค่า $|Z| \leqslant 0.05$ โดยทั่ว ๆ ไปไม่มีโอกาสให้สำหรับคลาดเคลื่อนอย่างน้อยที่สุด 5/100 จากสาเหตุของความไม่แน่นอนอยู่แล้ว หากไม่สิ้นเปลืองเท่าไร เราอาจเลือกตัวอย่างอื่น ๆ เพื่อทดสอบขึ้นใหม่ กระบวนการควรจะซ้ำจนกระทั่งบรรลุถึงการตัดสินใจที่มั่นคง

ตัวอย่าง 9.4 จงทดสอบว่าสลักเกลียวที่ใช้ประกอบเครื่องรถยนต์ควรจะมีความยาวเฉลี่ย 1.500 นิ้ว หากตัวอย่างสุ่มหนึ่งประกอบด้วยสลักเกลียว 900 ตัว ได้มาจากการผลิตเครื่องจักรหนึ่ง ให้ค่า $\bar{x} = 1.504$ นิ้ว พร้อมด้วย $s = .050$ นิ้ว เครื่องจักรนี้อยู่ภายในการตัดสินที่ยอมรับหรือ

วิธีทำ ในที่นี้ μ = ความยาวเฉลี่ยของสลักเกลียวที่ผลิตด้วยเครื่องจักรนี้ ปัญหานี้จะต้องพิจารณาถึงเหตุผลว่าไม่ควรยอมรับสลักเกลียวยาวเกินไปหรือสั้นเกินไป เพราะจะเป็นเหตุให้เครื่องเสียได้ จึงต้องทดสอบสองข้าง

1. Ho : $\mu = 1.500$

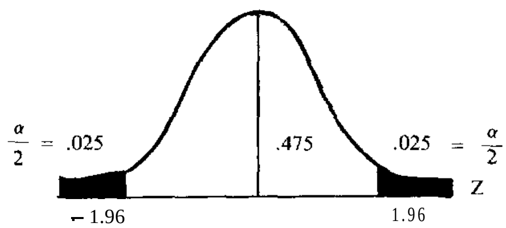
   Ha : $\mu \neq 1.500$

2. a = 0.05, n = 900

ทดสอบ ปฏิเสธ Ho หากคำนวณ $z < -1.96$ หรือ $Z > 1.96$ ไม่สามารถปฏิเสธ Ho หาก $-1.96 < z < 1.96$

3. $z = \dfrac{1.504 - 1.500}{.050 / \sqrt{900}} = \dfrac{.004}{.050 / \sqrt{900}} = 2.40$

4. ตัดสินใจ ปฏิเสธ Ho ที่ a = .05



พื้นที่ปฏิเสธ Ho	พื้นที่ยอมรับ Ho	พื้นที่ปฏิเสธ Ho

สรุป ความยาวเฉลี่ยของสลักเกลียวไม่เป็น 1.500 นิ้ว alternative hypothesis ไม่ได้กำหนดว่าสลักเกลียวยาวเกินไปหรือสั้นเกินไป การคำนวณหาทิศทางของการตัดสินใจเป็นสิ่งจำเป็นจากค่าประมาณของช่วงความเชื่อมั่น $1 - \alpha = .95$ คือ

$$1.504 - 1.96\frac{.05}{30} < \mu < 1.504 + 1.96\frac{.05}{30}$$

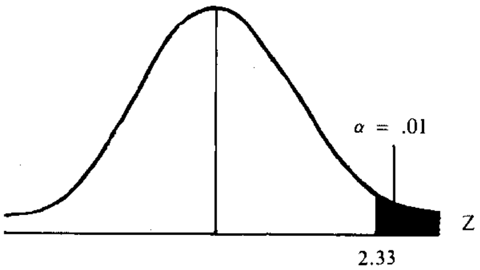
$$1.501 < \mu < 1.507 \text{ นิ้ว}$$

สังเกตว่าค่าน้อยที่สุดของช่วงนี้ 1.501 นิ้ว (เกินค่า 1.500 นิ้ว) เราก็สรุปได้ว่าเครื่องจักรผลิตสลักเกลียวโดยเฉลี่ยยาวเกินความต้องการ

ในการวิจัยสำรวจมีอยู่หลาย ๆ กรณีที่เลือก Ha ครั้งแรกโดยให้ $\mu \neq$ หากปฏิเสธ Ho ก็ทำการทดสอบลักษณะแตกต่างออกไปอีก เช่น ข้างเดียว และต้องใช้ข้อมูลอย่างใหม่ในแต่ละการทดสอบ ด้วยเหตุนี้ในตัวอย่างสุดท้าย ก็อาจเลือกตัวอย่างอื่น ๆ และให้ Ha : $\mu > 1.50$ นิ้ว

ตัวอย่าง 9.5 ผู้ผลิตอาหารสัตว์หนึ่งกำลังพิจารณาประกาศให้รางวัลต่อพนักงาน หรือตัวแทนจำหน่ายที่ขายได้มากกว่า สุ่มเลือกตัวแทนจำหน่ายมาสี่สิบเก้าคน ลำดับการขายต่อการประกาศรางวัลเฉลี่ย 1,550 บาท ต่อตัวแทน พร้อมด้วยส่วนเบี่ยงเบนมาตรฐาน 249 บาท ระหว่างระยะ

ทดลองขายของตัวแทนเดียวกันปรากฏว่าขายได้เฉลี่ย 1,647 บาท (ภายหลังการปรับการขยายของราคาและต้นทุนของรางวัล) จากเกณฑ์ของการขายเพิ่มขึ้นนี้พอที่จะคุ้มค่ารางวัลหรือไม่ใช้ $\alpha = 0.01$ สำหรับ $\mu =$ ขายได้เฉลี่ยของตัวแทน



พื้นที่ยอมรับ Ho | พื้นที่ปฏิเสธ Ho

1. Ho : $\mu = 1,550$ บาท

   Ha : $\mu > 1,550$ บาท การทดสอบข้างขวา

2. $\alpha = 0.01$ n = 49

การทดสอบ ปฏิเสธ Ho หากคำนวณ $z > 2.33$

3. คำนวณค่า $z = \dfrac{1,647 - 1,550}{294 / \sqrt{49}} = 2.31$

4. การตัดสินใจ สงวนการตัดสินขณะ 2.31 – 2.33 ≤ .05 ข่าวสารนี้ไม่อนุญาตการตัดสินใจได้อย่างชัดเจนที่ระดับนัยสำคัญ $\alpha = 0.01$ ควรจะได้ข่าวสารมากกว่านี้ภายใต้ระบบการทดลองและการดำเนินการทดสอบอื่น ๆ

แต่ละตัวอย่างในหัวข้อนี้แสดงถึงแบบต่าง ๆ ของ alternative hypothesis ซึ่งเป็นโครงสร้างการดำเนินการทดสอบสมมติฐาน

9.14 การทดลองแบบทวินามและการทดสอบเกี่ยวกับสัดส่วน

การทดสอบเกี่ยวกับสัดส่วนก็มีวิธีดำเนินการเหมือนการทดสอบมัชฌิมเลขคณิต แต่ขนาดตัวอย่างต้องมีขนาดใหญ่พอ ตัวสถิติที่ใช้ทดสอบก็คือ z

สำหรับการทดลองแบบทวินามพร้อมด้วยตัวอย่างใหญ่พอ ($np \geq 5$, $nq \geq 5$) การแจกแจงตัวอย่าง

$$Z = \frac{\frac{x}{n} - p}{\sqrt{\frac{pq}{n}}} = \frac{x - np}{\sqrt{np\,(1-p)}}$$

ก็จะมีการแจกแจงปกติมีมัชฌิมเลขคณิตศูนย์ส่วนเบี่ยงเบนมาตรฐานหนึ่ง

การใช้แบบ z ในการทดสอบค่าของ p ที่กำหนดขึ้นภายใต้สมมติฐาน ดังตัวอย่าง เช่น Ho : p = .4 ค่า .4 ใช้แทนที่ p ในฟอร์มของ z

ในหัวข้อนี้มีหลาย ๆ ตัวอย่างแสดงการทดสอบสัดส่วนจริงของความสำเร็จ p ในการทดสอบแบบทวินาม การทดสอบเกี่ยวกับความเห็นของผู้ออกเสียง เปอร์เซ็นต์ของส่วนที่เสียของกระบวนการผลิต และสินค้าที่ชอบของผู้บริโภคเกี่ยวกับสินค้าในตลาด ดังตัวอย่างต่อไปนี้

ตัวอย่างที่ 9.6 บริษัทหนึ่งได้ผลิตกลอนประตูเสียไป 8% แต่ละชั่วโมงสุ่มเลือกตัวอย่างหนึ่งประกอบด้วยกลอนประตู 100 ตัว เพื่อควบคุมคุณภาพ ถ้าหากว่าระหว่างเวลาบ่ายสามโมงสุ่มเลือกตัวอย่างหนึ่งปรากฏว่าเสีย 15% อยากทราบว่ากระบวนการผลิตควรจะหยุดเพื่อปรับปรุงใหม่ที่ $\alpha = 0.02$

วิธีทำ ให้ p = เปอร์เซ็นต์ของกลอนประตูที่ผลิตจริง เนื่องจากว่าเรากำลังป้องกันผลิตภัณฑ์ไม่ให้เสียมากเกินไป alternative hypothesis จึงใช้ "มากกว่า"

1. Ho : p = .08

   Ha : p > .08

2. a = .02, n = 100

   ทดสอบปฏิเสธ Ho หากคำนวณ Z > 2.05

3. คำนวณ $z = \frac{15 - 8}{\sqrt{100 \times .08 \times .92}} = 2.58$

4. ตัดสินใจ ปฏิเสธ Ho ที่ $\alpha = .02$

สรุป เปอร์เซ็นต์ของกลอนประตูเสีย เพิ่มขึ้นอย่างมีนัยสำคัญ ตามเกณฑ์นี้กระบวนการผลิตควรจะหยุดและปรับปรุงใหม่

ตัวอย่างที่ 9.7 ในการสำรวจตัวอย่างหนึ่งที่ได้ดำเนินการที่มหาวิทยาลัยสหศึกษา สุ่มถามนักศึกษาหญิง 400 คน หากเขาพอใจไม่ให้มีการบังคับชั่วโมงสำหรับนักศึกษาหญิง มีนักศึกษาหญิง 186 คน พอใจไม่ให้มีการบังคับ มีหลักฐานเพียงพอที่จะแสดงว่านักศึกษาหญิงที่พอใจไม่ให้มีการบังคับชั่วโมงน้อยกว่านักศึกษาส่วนใหญ่ในมหาวิทยาลัยนี้หรือ

วิธีทำ เนื่องจากว่าระดับการเสี่ยงภัย α ไม่ได้กำหนด เราก็มีสิทธิที่จะเลือกของเราเอง เราให้ $\alpha = 0.05$ สำหรับ p = เปอร์เซ็นต์ของนักศึกษาหญิงที่มหาวิทยาลัยนี้พอใจที่ไม่ให้มีการบังคับเกี่ยวกับชั่วโมงของเขา

1. $\text{Ho} : p = \frac{1}{2}$

   $\text{Ha} : p < \frac{1}{2}$

2. สำหรับ $\alpha = 0.05, n = 400$

   ทดสอบ ปฏิเสธ Ho หากคำนวณ $z < -1.64$

3. คำนวณค่า Z = $\dfrac{\mathbf{186} - 400\,(\frac{1}{2})}{\sqrt{400\,(\frac{1}{2})\,(\frac{1}{2})}} = -1.40$

4. การตัดสินใจ เราไม่สามารถปฏิเสธ Ho สำหรับ $\alpha = .05$

สรุป เปอร์เซ็นต์ของนักศึกษาหญิงที่มหาวิทยาลัยนี้ผู้ซึ่งพอใจไม่ให้มีการบังคับเกี่ยวกับชั่วโมงสำหรับนักศึกษาหญิงน้อยกว่า 50% ความคิดเห็นของเพศหญิงแบ่งออกเป็นสองส่วนเท่า ๆ กัน เกี่ยวกับคำสั่งนี้หรือเพศหญิงส่วนใหญ่พอใจไม่ให้มีการบังคับ

ตัวอย่างที่ 9.8 แผนกการวิจัยตลาดของบริษัทเครื่องดื่มต้องการกำหนดว่าเด็กวัยรุ่นสามารถที่จะพิสูจน์ว่าโคล่าของเขาแตกต่างไปจากตราอื่น ๆ ด้วยการชิม สุ่มเลือกเด็กวัยรุ่นสองร้อยคน แต่ละคนให้แก้วที่บรรจุโคล่าสี่ใบ มีอยู่ใบหนึ่งที่บรรจุโคล่าของบริษัท ภายหลังการชิมทั้งสี่แก้ว มีอยู่ 64 คน ที่สามารถพิสูจน์โคล่าของบริษัทได้ถูกต้อง ผลลัพธ์นี้มีค่ามากกว่าค่าที่คาดหวังอย่างมีนัยสำคัญหรือใช้ $\alpha = 0.025$

วิธีทำ ในการตัดสินใจว่าจะทดสอบสมมติฐานอะไรนั้น เราเริ่มต้นด้วยการสมมติโอกาสที่ได้รับเลือกแต่ละแก้วในสี่แก้วเท่า ๆ กัน นั่นคือ หากแต่ละคนไม่สามารถพิสูจน์โคล่านี้ด้วยการชิมจริง ๆ แล้วโอกาส $\frac{1}{4}$ สำหรับการพิสูจน์ของเขาก็เป็นได้ด้วยการเดา ความสามารถที่ทราบโคล่าของบริษัท ด้วยการชิมจากตราต่าง ๆ ควรแสดงได้จากการพิสูจน์ทางสถิติมากกว่า $\frac{1}{4}$ ด้วยเหตุนี้สำหรับ p = เปอร์เซ็นต์ที่แท้จริงของเด็กวัยรุ่นผู้ซึ่งสามารถพิสูจน์โคล่านี้จากสามตราอื่น ๆ ด้วยการชิมจริง และ $\frac{x}{n} = \frac{64}{200} = 0.32$

1. $Ho : p = \frac{1}{4}$

   $Ha : p > \frac{1}{4}$

การทดสอบปฏิเสธ Ho หากคำนวณ $Z > 1.96$ $(\alpha = 0.025)$

2. คำนวณ $Z = \dfrac{(64/200) - (1/4)}{\sqrt{(.25 \times .75) / 200}} = \dfrac{0.07}{0.03} = 2.33$

3. การตัดสินใจ ปฏิเสธ Ho สำหรับ $\alpha = 0.025$

สรุป โดยทั่ว ๆ ไป เด็กวัยรุ่นมีความสามารถบางอย่างในการพิสูจน์โคล่านี้ด้วยการชิมของตัวเองอยู่แล้ว มากกว่า 25% ของเด็กวัยรุ่นทั้งหมดมีความสามารถที่จะพิสูจน์โคล่านี้ในการเปรียบเทียบกับสามตราอื่น ๆ

สรุปขั้นการทดสอบสมมติฐานเกี่ยวกับมัชฌิมเลขคณิตและสัดส่วน

ในการอนุมานผลทางสถิติขึ้นอยู่กับการเลือกตัวอย่าง เพราะว่าการเลือกประชากรทั้งหมดจะไม่เป็นการประหยัดและไม่จำเป็น หากมีความระมัดระวังตั้งแต่เริ่มแรกผลที่ได้มาจากตัวอย่างค่อนข้างดีมาก การสุ่มตัวอย่างเป็นกระบวนการหนึ่งที่ให้ตัวค่าประมาณที่จะเอาไปเป็นตัวแทนประชากร นอกจากนั้นแบบความน่าจะเป็นที่ได้จากการสุ่มตัวอย่างก็ยังสามารถนำไปใช้สำหรับการอนุมานเกี่ยวกับมัชฌิมเลขคณิต เปอร์เซ็นต์และอื่น ๆ ของประชากร เราสามารถใช้ตารางสุ่มหรือตัวเลขสุ่มด้วย computer เพื่อเลือกเป็นตัวอย่างก็ได้

การแจกแจงความน่าจะเป็นของค่าที่เป็นไปได้ทั้งหมดสำหรับตัวสถิติหนึ่งเป็นตัวแปรเชิงสุ่มหนึ่ง เรียกว่าการแจกแจงตัวอย่างของตัวแปร ตัวสถิติที่เราใช้มากที่สุดคือมัชฌิมเลขคณิตของตัวอย่าง สำหรับตัวอย่างขนาดใหญ่ $(n > 30)$ วิธีปฏิบัติปัญหาความน่าจะเป็นเกี่ยวกับค่าของมัชฌิมเลขคณิตของตัวอย่างโดยการใช้ทฤษฎีขีดจำกัดส่วนกลางและกฎนี้ก็ใช้กันอย่างกว้างขวางเนื่องจากว่าการ

แจกแจงตัวอย่างของมัชฌิมเลขคณิตมีการแจกแจงปกติจึงใช้คะแนน Z ในการตอบปัญหาความน่าจะเป็นเกี่ยวกับค่าของมัชฌิมเลขคณิตของตัวอย่างสำหรับ $n \leqslant 30$ ขณะที่ x มีการแจกแจงปกติและทราบ σ^2 กฎนี้เป็นจริง

การอนุมานทางสถิติแบ่งออกเป็นสองหัวข้อใหญ่คือ (1) การประมาณค่าซึ่งเกี่ยวกับการประมาณค่าของตัวพารามิเตอร์ (2) การทดสอบในการตัดสินสภาวะที่เกี่ยวกับปริมาณการเสี่ยงภัย การศึกษาตัวพารามิเตอร์มีมัชฌิมเลขคณิตและสัดส่วนของประชากร

การทดสอบสมมติฐานเป็นวิธีการทางวิทยาศาสตร์ประยุกต์ขั้นตอนของการทดสอบมีดังนี้

1. พิจารณาปัญหา
2. กำหนด alternative hypothesises ในลักษณะที่เป็นจริง
3. สร้างกระบวนการทดสอบ วิธีการสำหรับการตัดสินใจจากหลักฐานของตัวอย่างซึ่งเป็นทางเลือกที่สมเหตุสมผล
4. เลือกตัวอย่างเพื่อใช้ทดสอบ
5. ทำการตัดสินใจหรือแก้ไขสมมติฐานเพื่อทำการทดสอบต่อไป

การทดสอบของเราเกี่ยวกับค่าจ้างสำหรับตัวพารามิเตอร์ μ กับ p อย่างไรก็ตาม หลักการทดสอบก็ไม่ได้บังคับต่อสองกรณีเหล่านี้ แต่ส่วนมากเป็นไปตามกระบวนทดสอบในบทนี้

ทิศทางของการทดสอบกำหนดได้ใน alternative hypothesis "ไม่เท่ากับ" ของ alternative เป็นเขตวิกฤตทั้งสองข้างสำหรับปฏิเสธ Ho

พิจารณาสองการเสี่ยงภัยการทดสอบ α วัดโอกาสการปฏิเสธสมมติฐานโดยไม่ถูกต้อง β วัดโอกาสเกี่ยวกับการยอมรับสมมติฐานที่ผิด การทดสอบที่ดีต้องการรักษาค่าการเสี่ยงภัยทั้งสองแบบให้น้อยหรือเข้าใกล้ศูนย์ แต่วิธีการที่ต้องการตัวอย่างใหญ่พอซึ่งอาจสิ้นค่าใช้จ่ายมาก ดูลที่สมเหตุสมผลคือต้องดัดแปลงทั้งการเสี่ยงภัย (α และ β) กับต้นทุนที่สะท้อนถึงขนาดตัวอย่าง

9.15 การทดสอบเกี่ยวกับมัชฌิมเลขคณิต (ตัวอย่างขนาดเล็ก)

เราได้พิจารณาการอนุมานเกี่ยวกับมัชฌิมเลขคณิตและค่าของเปอร์เซ็นต์สำหรับตัวอย่างขนาดใหญ่ เรากลับมายังเทคนิคตัวอย่างขนาดเล็กเพื่อทดสอบเกี่ยวกับมัชฌิมเลขคณิต ก่อนอื่นเราทราบว่าความคลาดเคลื่อนมาตรฐานต้องคำนวณมาจากตัวพารามิเตอร์ของประชากร อย่างเช่น $\sigma_{\bar{X}} = \sigma / \sqrt{n}$ ซึ่งต้องการทราบส่วนเบี่ยงเบนมาตรฐานของประชากรในทางปฏิบัติโดยมากไม่ทราบตัวพารามิเตอร์ที่ต้องการและต้องประมาณค่าความคลาดเคลื่อนมาตรฐานจากข้อมูลของตัวอย่าง

ความคลาดเคลื่อนมาตรฐานนั้นเขียนได้เป็น $s_{\bar{X}} = s/\sqrt{n}$ และจากสูตรนี้จะเห็นได้ว่า ปริมาณของความคลาดเคลื่อนเป็นฟังก์ชั่นของขนาดตัวอย่าง ตัวอย่างขนาดใหญ่กว่าย่อมมีความคลาดเคลื่อนน้อยกว่า เมื่อไรขนาดตัวอย่างมีขนาดเล็ก การทดสอบแบบ Z ก็ต้องเปลี่ยนไปเป็นอย่างอื่น นักศึกษาอาจเข้าใจผิดคิดว่าปัญหาการทดสอบแบบ Z ขึ้นอยู่กับขนาดตัวอย่าง เราควรจะทำความเข้าใจตั้งแต่นี้ต่อไป ประเด็นพื้นฐานคือว่าสูตรสำหรับความคลาดเคลื่อนมาตรฐานขึ้นอยู่กับตัวพารามิเตอร์ของประชากรหรือตัวค่าประกอบของตัวอย่างของตัวพารามิเตอร์เหล่านั้น หากความคลาดเคลื่อนมาตรฐานขึ้นอยู่กับตัวพารามิเตอร์ เราก็ใช้กระบวนการแบบ Z โดยไม่คำนึงถึงขนาดตัวอย่าง หากความคลาดเคลื่อนมาตรฐานขึ้นอยู่กับตัวค่าประมาณของตัวพารามิเตอร์ของประชากรและขนาดตัวอย่างต้องขนาดน้อยกว่าหรือเท่ากับ 30 เราก็ใช้กระบวนการแบบ t ที่จะกล่าวต่อไปนี้

9.16 คุณลักษณะของการแจกแจงแบบ t

การแจกแจงแบบ t มีลักษณะคล้ายคลึงการแจกแจงปกติ แบบของการแจกแจงแบบ t เขียนได้เป็น

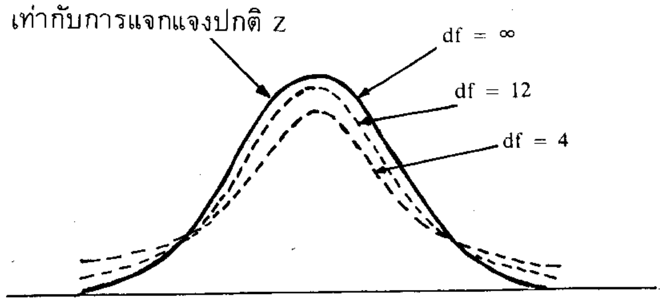
$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

และมีการคำนวณอะไรอีกหลาย ๆ อย่างที่คล้ายกับกระบวนการแบบ Z เมื่อไรที่ตัวอย่างมีขนาดใหญ่ ค่าของ $s_{\bar{X}}$ จะเข้าใกล้ค่าของ $\sigma_{\bar{X}}$ และ t ก็จะคล้ายกับ Z มากที่สุด การแจกแจงของ t ก็เข้าใกล้ปกติ

การแจกแจงแบบ t ไม่ได้เป็นการแจกแจงเดี่ยว ๆ แต่เป็นการแจกแจงไปในแบบครอบครัว และรูปการแจกแจงก็ขึ้นอยู่กับขนาดตัวอย่างหรือจำนวนองศาแห่งความอิสระ (degrees of freedom) เมื่อไรจำนวนองศาแห่งความอิสระ เป็นจำนวนอนันต์ (d.f. $= \infty$) การแจกแจงแบบ t ก็เหมือนกับการแจกแจงแบบ Z ขณะจำนวนองศาแห่งความอิสระลดลง คุณลักษณะของการแจกแจงแบบ t ก็จะเริ่มแยกออกจากการแจกแจงแบบ Z รูปที่ 9.14 แสดงถึงวัตถุประสงค์อันนี้ โดยเฉพาะ การแจกแจงแบบปกติกับการแจกแจงแบบ t สำหรับ df = 12 กับ df = 4

การแจกแจงแบบปกติกับการแจกแจงแบบ t

สิ่งที่มีลักษณะคล้ายกัน	**สิ่งที่มีลักษณะต่างกัน**
1. มัชฌิมเป็นศูนย์	1. เป็น leptokurtic มากกว่าการแจกแจงปกติ
2. เป็นรูปสมมาตร	2. มีส่วนเบี่ยงเบนมาตรฐานมากกว่าการแจกแจงปกติ
3. มีฐานนิยมเดียวเท่านั้น	3. ขึ้นอยู่กับองศาแห่งความอิสระ



รูปที่ 9.14 การแจกแจงแบบ t สำหรับสามระดับขององศาแห่งความอิสระ

คุณลักษณะที่แตกต่างกันของการแจกแจงแบบ t มีผลที่สำคัญในการอนุมานทางสถิติ ดังตัวอย่าง ถ้าหากว่าเราทดสอบสมมติฐานที่ $\alpha = .05$ ค่าวิกฤติของ z เป็น ± 1.96 สำหรับการทดสอบสองด้าน แต่ขนาดของ t สำหรับซึ่งความน่าจะเป็นเป็น .05 ของการคำนวณค่าเบี่ยงเบนมากกว่า 1.96 นี้ก็สามารถศึกษาได้จากรูปที่ 9.14 หากของการแจกแจงแบบ t ลาดลงตามแกนนอนน้อยกว่าการแจกแจงปกติ

t จะต้องมากกว่า z เท่าไร จึงจะสมนัยกันกับเหตุการณ์ที่เหมือนกันนาน ๆ สักครั้งที่ขึ้นอยู่กับขนาดตัวอย่าง ตารางข้างล่างนี้แสดงเหตุการณ์นี้ สำหรับการเลือกค่าของ df ขนาดวิกฤติของ t ที่สมนัยกับการทดสอบสองด้านนำไปสู่เกณฑ์การตัดสินใจ $\alpha = .05$ ขนาดของ t สำหรับการแจกแจงแบบ t ความน่าจะเป็นเป็น .05 ค่าส่วนเบี่ยงเบนที่น้อยที่สุดคำนวณหาได้

df	5	10	25	50	100	500	∞
t.05	2.571	2.228	2.060	2.008	1.984	1.965	1.96

ข้อสังเกต

1. จำนวนองศาแห่งความอิสระเป็นอนันต์ ขนาดวิกฤตของ t เหมือนกับขนาดวิกฤตของ Z : 1.96

2. จำนวนองศาแห่งความอิสระเล็กลง ขนาดวิกฤตของ t โตขึ้น

3. พฤติการณ์ของ t จะคล้ายคลึงกับ z มากที่สุดจนกระทั่งจำนวนองศาแห่งความอิสระตั้งแต่ 100 ขึ้นไป

9.17 องศาแห่งความอิสระและการแจกแจงแบบ t

คำว่า "องศาแห่งความอิสระ" หมายถึงความอิสระของค่าสังเกตที่จะแปรไป โดยทั่ว ๆ ไป จำนวนองศาแห่งความอิสระจะสัมพันธ์กับจำนวนค่าสังเกต ซึ่งมีความอิสระที่จะแปรไปได้อย่างสมบูรณ์ สิ่งหนึ่งที่อาจสมมติขึ้นนี้ก็เหมือนกับจำนวนค่าสังเกตในตัวอย่าง แต่มีเงื่อนไขสอดคล้องกับข้อบังคับเพื่อให้จำนวนองศาแห่งความอิสระน้อยลง เราจะแสดงแนวทางศึกษาองศาแห่งความอิสระเกี่ยวกับการคำนวณ s_x

สมมติว่าตัวอย่างหนึ่งประกอบด้วยสามข้อมูลของค่าสังเกต 33, 35, 40 ถ้าหากว่าคำนวณค่า s_x ได้จากสูตร

$s_x = \sqrt{\Sigma(X - \overline{X})^2/(n-1)}$ เราเริ่มด้วยการหามัชฌิม $\overline{X} = 36$ และขั้นต่อไปสร้างคะแนนเบี่ยงเบนได้ด้วยการลบแต่ละค่าสังเกตด้วย มัชฌิม ดังแสดงได้ดังนี้ ($\Sigma x = 108$; $\overline{X} = 36$)

X	$(X - \overline{X})$
33	$(33 - 36) = -3$
35	$(35 - 36) = -1$
40	$(40 - 36) = +4$

องศาแห่งความอิสระสำหรับปัญหานี้ ค่าของ $\overline{X}$ เป็นค่าคงที่ คะแนนตัวแรก 33 สามารถมีค่าแตกต่างกัน คือ อิสระต่อการแปร ค่าสังเกตตัวที่สองก็เช่นเดียวกัน 35 แต่ถ้าหากว่าเราประสงค์ที่จะกำหนดค่าหนึ่งค่าใดกับค่าสังเกตสองตัวแรกแล้ว เราจะไม่มีความอิสระที่จะกำหนดค่าของค่าสังเกตตัวที่สาม

โปรดจำไว้ว่ามัชฌิมเลขคณิตของสามค่าสังเกตเป็น 36 ถ้าหากว่ากำหนดสองค่าสังเกตแรกแล้ว ค่าสังเกตตัวที่สามต้องเป็นค่าที่เราต้องการเพื่อว่าให้ผลรวมของค่าสามสังเกตจะเท่ากับ 108 เมื่อไรหารด้วยสามแล้วก็จะให้ $\overline{X} = 36$ จำนวนองศาแห่งความอิสระในปัญหานี้คือ 2

ถ้าหากว่ามีสี่ค่าสังเกต ก็จะมีสามค่าสังเกตจากสี่ตัว มีความอิสระกันที่จะผันแปรโดยทั่ว ๆ ไป จำนวนองศาแห่งความอิสระจะน้อยกว่าขนาดตัวอย่างไปหนึ่ง ตัวอย่างเช่น $df = n - 1$ จำนวนองศาแห่งความอิสระไม่จำเป็นจะต้องเป็น $n - 1$ เสมอไป ขึ้นอยู่กับชนิดของปัญหาแต่ละชนิด

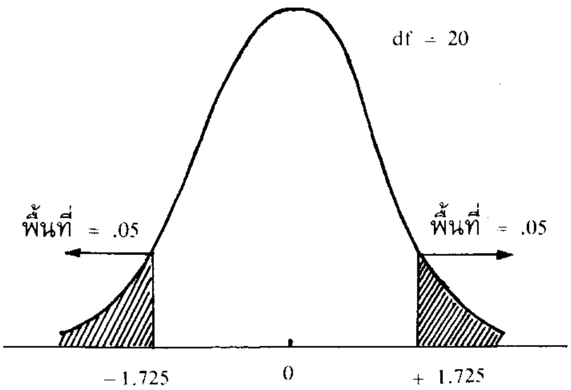
9.18 การใช้การแจกแจงแบบ t

การแจกแจงทางทฤษฎีของ t ปรากฏอยู่ในตารางภาคผนวกในการใช้ตารางนี้ สิ่งสำคัญที่จะต้องจำคือว่าพื้นที่ภายใต้การแจกแจงที่กำหนดให้ของ t เป็น 1.00 เหมือนตารางเส้นโค้งปกติ ตารางไม่ได้แสดงค่าลบกับค่าบวกของ t ไว้ เนื่องจากว่าพื้นที่ตกเหนือค่า t เป็นบวกเหมือนกับพื้นที่ตกได้ค่า t เป็นลบ นี้แสดงได้ดังรูปที่ 9.15

แต่ละแถวของตารางระบุตามองศาแห่งความอิสระเฉพาะ ค่าของ t แสดงถึงพื้นที่จะวางอยู่ข้างบนของตาราง ลักษณะการใช้ตารางอาจจะต้องแสดงด้วยตัวอย่าง ดังพิจารณาปัญหาต่อไปนี้

ปัญหา เมื่อไร df = 20 ค่า t ที่มากที่สุดจะเป็นเท่าไร สำหรับ 5% ของครั้ง

วิธีทำ ดูตาราง t และหาแถวที่สมนัยกับ df = 20 เพื่อหาคอลัมน์ที่สมนัยกับพื้นที่ในด้านหนึ่ง = .05 ที่จุดตัดของแถวกับคอลัมน์นี้อ่าน t = 1.725 คล้ายกับการหาตำแหน่งเปอร์เซ็นต์ไทล์ที่จุด 95 ของการแจกแจงแบบ t ค่าของมันเป็นบวก t = + 1.725 ดังรูปที่ 9.15



รูปที่ 9.15 ค่าของ t สำหรับ df = 20 ซึ่งสมนัยกับพื้นที่ในแต่ละด้านของการแจกแจงแบบ t = .05

ถ้าหากว่าเราทดสอบสมมติฐานที่ $\mu = 100$ เทียบกับ alternative ด้านเดียว $\mu > 100$ และใช้เกณฑ์การตัดสินใจ $\alpha = .05$ ค่าวิกฤตของ t เมื่อ df = 20 ถ้าเป็นค่าวิกฤตของเส้นโค้งปกติ z ก็จะเป็น z = 1.645

ปัญหา เมื่อ df = 20 การทดสอบสองด้านที่ α = .05 ค่าของ t ที่แตกต่างกันของเขตปฏิเสธจากเขตยอมรับเป็นเท่าไร

วิธีทำ เราต้องหาค่าของ t ที่ 95% อยู่กึ่งกลางภายในเขตเหล่านี้ จะมีค่าของ t อยู่ต่ำกว่า 2.5% ของค่า t ที่วางอยู่ กับค่าของt 2.5% อยู่สูงกว่าค่า t ที่วางอยู่ ขนาดของ t อ่านได้ตามพื้นที่ในด้านหนึ่ง = .025 สำหรับ df = 20 ค่าของมันคือ t = + 2.086

9.19 การทดสอบสมมติฐานเกี่ยวกับหนึ่งมัชฌิมเลขคณิต (การแจกแจงแบบ t)

สมมติมีคำกล่าวกันว่าความสูงของผู้ใหญ่ (ชาย) มีความสูงกว่าที่เคยเป็น เราประสงค์เพื่อที่จะทดสอบสมมติฐานนี้จากการสำรวจทั่วประเทศ 20 ปีก่อน พบว่าความสูงเฉลี่ยของชายอายุ 21 ปี 69.5 นิ้ว ต้องการศึกษาคำถามนี้ สุ่มเลือกชายมา 25 คน เพื่อวัดความสูง แปลงคำถามของการวิจัยเป็นสมมติฐานทางสถิติ การตัดสินใจทำการทดสอบสมมติฐานที่ระดับนัยสำคัญ 5%

$$Ho : \mu = 69.5 \qquad \alpha = .05$$

$$Ha : \mu > 69.5$$

เราคำนวณหามัชฌิมเลขคณิตและส่วนเบี่ยงเบนมาตรฐานของการวัด 25 ครั้ง ความคลาดเคลื่อนมาตรฐานของมัชฌิมเลขคณิต ดังแสดงข้างล่างนี้

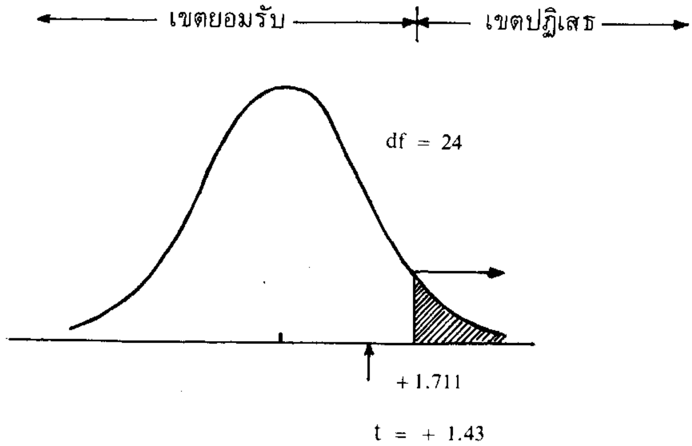
$$n = 25, \overline{X} = 70.4, s_x = 3.15, s_{\bar{X}} = \frac{s_x}{\sqrt{n}} = \frac{3.15}{\sqrt{25}} = .63$$

เพื่อทดสอบสมมติฐาน เราคำนวณ t

$$t = \frac{\bar{X} - \mu}{s_{\bar{X}}} = \frac{70.4 - 69.5}{.63} = +1.43$$

จำนวนองศาแห่งความอิสระที่สัมพันธ์กับ t คือ df = n − 1 = 24

การแจกแจงของ t สำหรับองศาแห่งความอิสระแสดงในรูปที่ 9.16 เพราะว่า Ha ระบุว่า μ มากกว่า 69.5 เขตวิกฤตต้องวางด้านขวาของการแจกแจง ตามเกณฑ์การตัดสินใจที่ใช้ ควรปฏิเสธ Ho ถ้าหากว่าได้มัชฌิมเลขคณิตของตัวอย่างนำไปสู่ค่าของ t เบี่ยงเบนมากที่ซึ่งความน่าจะเป็นของการเกิดขึ้นของมันควรเป็น .05 ถ้าหากว่าสมมติฐานเป็นจริง ค่าวิกฤต t ที่ซึ่ง 5% ของการแจกแจงแบบ t ด้านขวามือจากตารางสำหรับ df = 24 คือ t = + 1.711 ดังรูป 9.16 เนื่องจากว่ามัชฌิมเลขคณิตของตัวอย่าง 70.4 เป็นข้อแย้งกันไม่เพียงพอที่จะนำไปสู่ค่า t = + 1.711 หรือมากกว่า การตัดสินใจของเราจึงต้องยอมรับ Ho



รูปที่ 9.16 การทดสอบสมมติฐานเกี่ยวกับ μ เทียบกับ alternative *ที่มีทิศทาง* $\alpha = .05$

จากตัวอย่างข้างต้น เราสมมติว่า $s_{\bar{x}}$ คำนวณเสร็จเรียบร้อยแล้ว เมื่อไรปัญหาเริ่มจากข้อมูลดิบ จำเป็นต้องหา $s_{\bar{x}}$ วิธีการคำนวณก็ได้กล่าวไว้ในหัวข้อ 9.13

9.20 การทดสอบสมมติฐานเกี่ยวกับผลต่างระหว่างสองมัชฌิมเลขคณิตที่อิสระกัน

การทดสอบผลต่างระหว่างสองมัชฌิมเลขคณิตที่แตกต่างกัน กระบวนการตัวอย่างขนาดใหญ่ ปัญหาของการทดสอบสมมติฐานสำหรับสองมัชฌิมเลขคณิตเท่ากันอาจช่วยในการหาคำตอบของปัญหาได้ อย่างเช่น สุ่มเลือกตัวอย่างมาสองตัวอย่างขนาดตัวอย่างละ 50 ตัวอย่างแรกให้อาหารหนูอย่างปกติ ขณะที่ตัวอย่างสองให้อาหารหนูผสมวิตามิน ภายหลังชั่วระยะเวลาหนึ่ง เอาหนูแต่ละตัวมาชั่งน้ำหนัก คำนวณหาน้ำหนักเฉลี่ยแต่ละตัวอย่าง เราอาจสมมติว่าน้ำหนักเฉลี่ยของตัวอย่างที่ให้อาหารผสมวิตามินเป็น .47 ปอนด์ อีกตัวอย่างหนึ่งเป็น .43 ปอนด์ อย่างไรก็ตามเราไม่สามารถที่จะศึกษาเปรียบเทียบมัชฌิมเลขคณิตของตัวอย่างเฉพาะได้ เพราะว่าสองตัวอย่างให้มัชฌิมเลขคณิตแตกต่างกันแม้ว่าสองตัวอย่างได้รับปฏิบัติเหมือนกัน

ปัญหาที่สำคัญไม่ได้เกี่ยวกับตัวอย่างแต่เกี่ยวกับประชากรที่ตัวอย่างได้มา เป็นไปได้ไม่มีมัชฌิมเลขคณิตของประชากรภายใต้การให้อาหารปกติเหมือนกับมัชฌิมเลขคณิตของประชากรภายใต้การให้อาหารปกติผสมวิตามิน มัชฌิมเลขคณิตของประชากรทั้งสองควรจะเหมือนกันถ้าหากว่าวิตามินไม่มีผล ปัญหานี้อาจแสดงออกได้เป็นสมมติฐานทางสถิติเพื่อที่จะทดสอบดังนี้

$$Ho : \mu_v - \mu = 0$$

$$Ha : \mu_v - \mu \neq 0$$

Ha อาจมีทิศทางหรือไม่มีทิศทางก็ได้ ถ้าหากว่าความสนใจของเราพบว่าวิตามินมีส่วนทำให้เพิ่มน้ำหนัก Ha : $\mu_v - \mu > 0$ ก็เป็น alternative ที่เหมาะสม อย่างไรก็ตามเราต้องการสันนิษฐานที่จะทราบว่าผลอาจอยู่ในแง่ดีหรือแง่เสียอย่างใดอย่างหนึ่ง จึงต้องใช้ทดสอบสองด้าน การทดสอบก็ใช้แบบเดียวกับการทดสอบหนึ่งมัชฌิมเลขคณิตประชากร

9.21 การแจกแจงตัวอย่างสุ่มของผลต่างระหว่างสองมัชฌิมเลขคณิตตัวอย่าง

สมมติว่าสุ่มเลือกตัวอย่างหนึ่งจากประชากร X และสุ่มเลือกอีกตัวอย่างหนึ่งจากประชากร Y คำนวณหามัชฌิมเลขคณิตแต่ละตัวอย่างและหาผลต่างระหว่างสองมัชฌิมเลขคณิตเหล่านี้ ใส่ตัวอย่างเหล่านี้กลับเข้าไปในประชากรแล้วทำการสุ่มแบบเดียวกันเพื่อคำนวณหาสองมัชฌิมเลขคณิตและผลต่างระหว่างมัชฌิมเลขคณิต ถ้าหากว่าทำการทดสอบซ้ำ ๆ กัน โดยไม่จำกัดจำนวนครั้งค่าผลต่าง ($\bar{X} - \bar{Y}$) ก็จะสร้างการแจกแจงตัวอย่างสุ่มของผลต่างระหว่างสองมัชฌิมเลขคณิตตัวอย่าง

สมมติว่าประชากร X ประกอบด้วยสามข้อมูล 3, 5, 7 และประชากร Y ประกอบด้วยสามข้อมูลเดียวกัน 3, 5, 7 ถ้าหากว่าเลือกตัวอย่างแบบหยิบแล้วใส่คืนจากแต่ละประชากรก็จะได้ตัวอย่างที่อาจเป็นไปได้ 9 ตัวอย่าง (ของแต่ละประชากร) และมี 9 มัชฌิมเลขคณิตที่เป็นไปได้ (ถึงแม้ว่าจะมีบางมัชฌิมเลขคณิตมีค่าเหมือนกันก็ต้องใส่ลงในตาราง)

ตาราง

ตัวอย่างและมัชฌิมเลขคณิตของตัวอย่างที่เป็นไปได้สำหรับตัวอย่างขนาดสอง (แบบหยิบแล้วใส่คืน)

ประชากร X 3, 5, 7 $\mu_x = \mu_y = 5$

ประชากร Y 3, 5, 7 $\sigma_x = \sigma_y = 1.633$

ตัวอย่างจากประชากร X	$\bar{X}$	ตัวอย่างจากประชากร Y	$\bar{Y}$
3,3	3	3,3	3
3,5	4	3,5	4
3,7	5	3,7	5
5,3	4	5,3	4
5,5	5	5,5	5
5,7	6	5,7	6
7,3	5	7,3	5
7,5	6	7,5	6
7,7	7	7,7	7

ความต้องการของเราคือผลต่างระหว่างมัชฌิมเลขคณิตตัวอย่าง $\overline{X}$ กับมัชฌิมเลขคณิตตัวอย่าง $\overline{Y}$ เป็นเท่าไร เราต้องพิจารณาว่ามีเก้ามัชฌิมเลขคณิตตัวอย่างใน X กับที่แต่ละตัวอาจจับคู่กับตัวใดตัวหนึ่งของเก้าตัวมัชฌิมเลขคณิต ตัวอย่างใน Y ดังนั้นจึงต้องมีถึง 81 คู่ที่เป็นไปได้ของมัชฌิมเลขคณิตของ X กับมัชฌิมเลขคณิตของ Y มัชฌิมเลขคณิตตัวอย่างที่เป็นไปได้แสดงอยู่ที่ด้านข้างของตารางค่าผลต่าง 81 ค่าระหว่างแต่ละคู่ $(\overline{X} - \overline{Y})$ ก็แสดงอยู่ในด้านใน ถ้าหากว่าสุ่มเลือกตัวอย่างอิสระกัน แต่ละค่าของผลต่างก็จะปรากฏเท่ากัน ถ้าหากว่าเรารวบรวมค่าผลต่างเหล่านี้ในรูปของการแจกแจงความถี่ ผลก็จะได้ดังตารางต่อไป

มีหลาย ๆ ลักษณะของการแจกแจงความถี่ หนึ่ง มัชฌิมเลขคณิตของการแจกแจงระหว่างผลต่าง μ_x กับ μ_y เป็นศูนย์ สอง การแจกแจงเป็นเส้นโค้งปกติโดยประมาณดังรูป 9.17 เส้นโค้งปกติสามารถปรับให้เข้ากับการแจกแจงนี้และปรากฏว่าความถี่ที่ตนหวังของเส้นโค้งปกติใกล้เคียงกับความถี่จริงของมัชฌิมเลขคณิตเฉพาะ (แสดงได้โดยความสูงของแท่งในแผนภูมิ)

ตาราง

ผลต่างระหว่างคู่ของมัชฌิมเลขคณิตตัวอย่าง $(\overline{X} - \overline{Y})$ สำหรับคู่ตัวอย่าง X กับ Y ที่เป็นไปได้

				$\overline{Y}$					
$\overline{X}$	3	4	4	5	5	5	6	6	7
7	4	3	3	2	2	2	1	1	0
6	3	2	2	1	1	1	0	0	−1
6	3	2	2	1	1	1	0	0	−1
5	2	1	1	0	0	0	−1	−1	−2
5	2	1	1	0	0	0	−1	−1	−2
5	2	1	1	0	0	0	−1	−1	−2
4	1	0	0	−1	−1	−1	−2	−2	−3
4	1	0	0	−1	−1	−1	−2	−2	−3
3	0	−1	−1	−2	−2	−2	−3	−3	−4

ถ้าหากว่าการแจกแจงตัวอย่างของมัชฌิมเลขคณิตสำหรับ X กับ Y เป็นปกติโดยประมาณ การแจกแจงของผลต่างระหว่างคู่ของมัชฌิมเลขคณิตที่เลือกมาโดยสุ่มจะเป็นปกติโดยประมาณด้วย ในสภาวะของเราการแจกแจงของประชากร X กับ Y มีการแจกแจงคล้ายสี่เหลี่ยมผืนผ้า

แต่การแจกแจงของผลต่างระหว่างคู่ของมัชฌิมเลขคณิตตัวอย่างคล้ายเส้นโค้งปกติ ดังนั้นผลของการทำนายสำหรับมัชฌิมเลขคณิตเดี่ยว ๆ ด้วยทฤษฎีขีดจำกัดส่วนกลางมีผลสำคัญเกี่ยวกับการแจกแจงของผลต่างระหว่างมัชฌิมเลขคณิตตัวอย่างด้วย เพราะฉะนั้น เส้นโค้งปกติยังคงรักษาตัวแบบที่สำคัญสำหรับการแจกแจงตัวอย่างของผลต่างระหว่างมัชฌิมเลขคณิตแม้ว่าการแจกแจงประชากรผิดไปจากปกติบ้าง

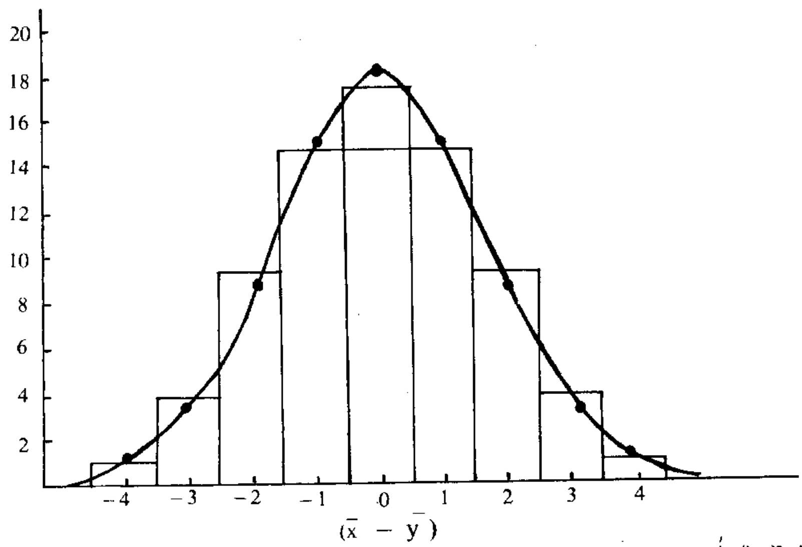
ตาราง

การแจกแจงของผลต่างระหว่างคู่ของมัชฌิมเลขคณิตตัวอย่างที่เป็นไปได้

$(\bar{X} - \bar{Y})$	f
4	1
3	4
2	10
1	16
0	19
−1	16
−2	10
−3	4
−4	1
	81

$\mu_{\bar{x} - \bar{y}} = 0$

$\sigma_{\bar{x} - \bar{y}} = 1.633$



รูปที่ 9.17 การแจกแจงของผลต่างระหว่างคู่ของมัชฌิมของตัวอย่างที่เป็นไปได้

9.22 มัชฌิมเลขคณิตตัวอย่างที่ไม่อิสระกันและอิสระกัน

ส่วนเบี่ยงเบนมาตรฐานของการแจกแจงตัวอย่างสุ่มของผลต่างระหว่างมัชฌิมเลขคณิตที่ขึ้นอยู่กับว่าคู่ของมัชฌิมเลขคณิตตัวอย่างที่มีความอิสระกันหรือไม่มีความอิสระกัน ถึงแม้ว่ามัชฌิมเลขคณิตของการแจกแจงตัวอย่างเป็น $\mu_x - \mu_y$ ในแต่ละกรณี มีวิธีการอยู่สามวิธีการที่จะนำเราไปศึกษาว่ามัชฌิมเลขคณิตของตัวอย่างมีความอิสระกันหรือไม่มีความอิสระกัน หนึ่ง ถ้าหากว่าเลือกสองตัวอย่างโดยสุ่มจากสองประชากรตามลำดับตัวอย่างจะมีความอิสระกัน สอง ถ้าหากว่าปัญหาเหมือนกันใช้สำหรับเงื่อนไขทั้งสองของการศึกษา สาม ถ้าหากว่าเอาสองตัวแปรมาจับคู่กันเพื่อดำเนินการตัวอย่างก็จะไม่มีความอิสระกัน

ส่วนเบี่ยงเบนมาตรฐานของการแจกแจงตัวอย่างของผลต่างระหว่างสองมัชฌิมเลขคณิตเรียกว่าความคลาดเคลื่อนมาตรฐานของผลต่างระหว่างสองมัชฌิมเลขคณิตเราใช้สัญลักษณ์เป็น $\sigma_{\bar{X}-\bar{Y}}$ ถ้าหากว่าการแจกแจงประกอบด้วยผลต่างระหว่างมัชฌิมเลขคณิตของตัวอย่างที่ไม่มีความอิสระกันก็จะมีส่วนเบี่ยงเบนมาตรฐานเขียนได้เป็น

$$\sigma_{\bar{X}-\bar{Y}} = \sqrt{\sigma^2_{\bar{X}} + \sigma^2_{\bar{Y}} - 2\rho_{xy}\,\sigma_{\bar{X}}\,\sigma_{\bar{Y}}}$$

ในเมื่อ

$\sigma_{\bar{X}}$ = ความคลาดเคลื่อนมาตรฐานของมัชฌิมเลขคณิต $\bar{X}$

$\sigma_{\bar{Y}}$ = ความคลาดเคลื่อนมาตรฐานของมัชฌิมเลขคณิต $\bar{Y}$

ρ_{xy} = สัมประสิทธิสหสัมพันธ์ของประชากร X กับ Y ถ้าหากว่า X กับ Y มีความอิสระกับ $\rho_{xy} = 0$ เทอม $2\rho_{xy}\sigma_{\bar{X}}\sigma_{\bar{Y}}$ ก็ไม่ปรากฏ ความคลาดเคลื่อนมาตรฐานของผลต่างระหว่างสองมัชฌิมเลขคณิตที่อิสระกันเขียนได้เป็น

$$\sigma_{\bar{X}-\bar{Y}} = \sqrt{\sigma^2_{\bar{X}} + \sigma^2_{\bar{Y}}}$$

เนื่องจากว่าค่าของประชากรที่ต้องการไม่ทราบจึงจำเป็นต้องใช้การประมาณค่าปริมาณเหล่านี้ สำหรับตัวอย่างไม่มีความอิสระกันค่าประมาณของ $\sigma_{\bar{X}-\bar{Y}}$ เขียนได้เป็น

$$s_{\bar{X}-\bar{Y}} = \sqrt{s^2_{\bar{X}} + s^2_{\bar{Y}} - 2r_{xy}\,s_{\bar{X}}s_{\bar{Y}}}$$

สำหรับตัวอย่างที่มีความอิสระกัน ค่าประมาณ $\sigma_{\bar{x}-\bar{y}}$ เขียนได้เป็น

$$s_{\bar{x}-\bar{y}} = \sqrt{s^2_{\bar{x}} + s^2_{\bar{y}}}$$

เพราะว่าไม่ทราบค่าประชากร สูตรเหล่านี้จึงใช้แสดงกระบวนการเพื่อทดสอบสมมติฐานเกี่ยวกับสองมัชฌิมเลขคณิต

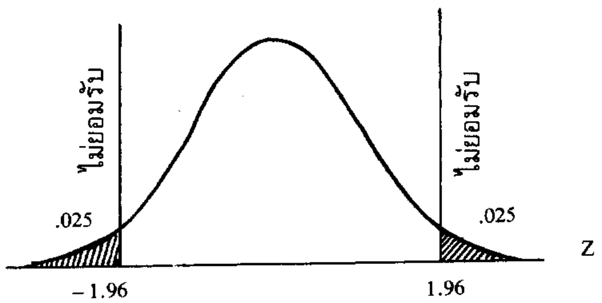
ตัวอย่าง 1 สมมติว่าเราต้องการจะเปรียบเทียบหลอดไฟฟ้าสองชนิดจากการทดลอง 100 ดวง ซึ่งทำโดยบริษัท ก ค่าเฉลี่ยอายุของหลอดไฟฟ้าเท่ากับ 952 ชั่วโมง ส่วนเบี่ยงเบนมาตรฐาน 85 ชั่วโมง ขณะเดียวกับหลอดไฟฟ้า 50 ดวง ทำโดยบริษัท B ค่าเฉลี่ยอายุเท่ากับ 987 ชั่วโมง ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 92 ชั่วโมง เราสามารถเขียนสัญลักษณ์ของการทดลองนี้ได้

$$n_x = 100,\ \bar{X} = 952,\ s_x = 85$$
$$n_y = 50,\ \bar{Y} = 987,\ s_y = 92$$

ให้ μ_x กับ μ_y เป็นค่าเฉลี่ยอายุของหลอดไฟฟ้าสองชนิดสมมติฐานที่ต้องการจะทดสอบ และ alternative hypothesis คือ

$$Ho : \mu_x = \mu_y\ (\mu_x - \mu_y = 0)$$
$$Ha : \mu_x \neq \mu_y\ (\mu_x - \mu_y \neq 0)$$

เราใช้ระดับนัยสำคัญ 0.05



รูปที่ 9.18

การแจกแจงตัวอย่างของผลต่างระหว่างสองมัชฌิมเลขคณิต ($\bar{X} - \bar{Y}$) ประมาณให้เป็นเส้นโค้งปกติดังรูป 9.18 การคำนวณผลต่างในรูปของผลคะแนน Z

$$Z = \frac{(\bar{X}_x - \bar{Y}_x) - (\mu_x - \mu_y)}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}$$

ภายใต้สมมติฐานค่าเฉลี่ยของการแจกแจงตัวอย่าง ($\bar{X} - \bar{Y}$) มีค่าเท่ากับศูนย์ แทนค่าตัวเลขสำหรับการทดลองเกี่ยวกับหลอดไฟฟ้าสองชนิดเราได้

$$Z = \frac{952 - 987}{\sqrt{\frac{85^2}{100} + \frac{92^2}{50}}} = -\frac{35}{15.5} = -2.26$$

เราจะไม่ยอมรับสมมติฐาน ถ้าหากว่า $Z < -1.96$ หรือ $Z > 1.96$ ยอมรับสมมติฐาน ถ้าหากว่า $-1.96 \leqslant Z \leqslant 1.96$

เนื่องจากว่าค่าสถิติ Z น้อยกว่า -1.96 เราก็สรุปได้ว่า ผลต่างระหว่าง 952 กับ 987 มีนัยสำคัญและไม่ยอมรับสมมติฐาน Ho มี $\alpha = .05$ (เราควรซื้อหลอดไฟฟ้าชนิดที่สอง แต่อาจมีตัวประกอบอื่น ๆ เช่น ราคา)

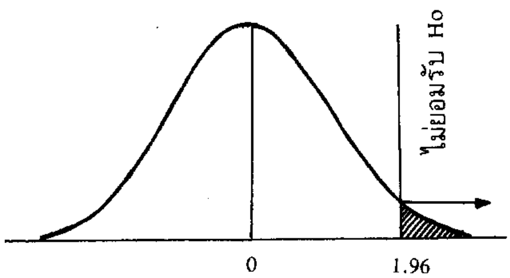
ตัวอย่าง 2 เปรียบเทียบวิธีการสอนสองชนิด สำหรับตัวอย่างสุ่มหนึ่งประกอบด้วยนักศึกษา $n_y = 64$ คน สอนด้วยวิธีการแบบเก่า คะแนนเฉลี่ยจากการทดสอบข้อสอบแบบมาตรฐานเป็น 68.8 ส่วนเบี่ยงเบนมาตรฐาน 5.2 สำหรับอีกตัวอย่างสุ่มหนึ่งประกอบด้วยนักศึกษา n_x 80 คน สอนด้วยวิธีการแบบใหม่ คะแนนเฉลี่ยจากการทดสอบข้อสอบเหมือนกันเป็น 70.5 ส่วนเบี่ยงเบนมาตรฐาน 5.6 ต้องการทดสอบสมมติฐานว่าผลของการสอนแบบใหม่มีคะแนนเฉลี่ยสูงกว่าที่ระดับนัยสำคัญ $\alpha = 0.025$

วิธีทำ นี้เป็นการทดสอบเกี่ยวกับความแตกต่างจริงในทางปฏิบัติถัวเฉลี่ย (μ_x = คะแนนเฉลี่ยจริงของวิธีการสอนแบบใหม่ μ_y = คะแนนเฉลี่ยจริงของวิธีการสอนแบบเก่า)

Ho : $\mu_x = \mu_y$ (หรือ $\mu_x - \mu_y = 0$)

Ha : $\mu_x > \mu_y$ (หรือ $\mu_x - \mu_y > 0$)

สำหรับ $\alpha = 0.025$



รูปที่ 9.19 การทดสอบสมมติฐานที่ $\mu_x - \mu_y$ *เทียบกับ* Ha : *ที่* $\mu_x - \mu_y > 0$

$$s_{\bar{x}_1-\bar{y}_2} = \sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}} = \sqrt{\frac{(5.6)^2}{80} + \frac{(5.2)^2}{64}}$$

$$= \sqrt{.392 + .4225} = .90$$

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{s_{\bar{x}-\bar{y}}} = \frac{(70.50 - 68.8) - 0}{.90}$$

$$= \frac{1.70}{.90} = 1.89$$

เราจะไม่ยอมรับ Ho ถ้าหากว่าค่า $z > 1.96$ และยอมรับ Ho ถ้าหากว่า $Z \leqslant 1.96$

เนื่องจากว่าค่าสถิติ Z น้อยกว่า 1.96 จึงยอมรับ Ho ที่ $\alpha = 0.025$

เราก็สรุปได้ว่าหลักการไม่ได้ช่วยกระบวนการที่มีประสิทธิภาพที่สูงกว่าสำหรับการศึกษาวิธีการแบบใหม่

ตัวอย่าง 3 (การทดสอบสมมติฐานของผลต่างสำหรับสองมัชฌิมเลขคณิตที่สัมพันธ์กัน)

อาจารย์คนหนึ่งต้องการประมาณค่าในความเข้าใจเพิ่มขึ้น สำหรับนักศึกษาซึ่งสำเร็จกระบวนวิชาของเขาเลือกตัวอย่างโดยสุ่มประกอบด้วยนักศึกษาสิบคนมาทดสอบก่อน แล้วมาทดสอบตอนจบกระบวนวิชาอีกครั้ง โดยใช้ข้อสอบเท่าเทียมกัน เราจะทดสอบสมมติฐานว่าคะแนนเพิ่มขึ้นมากกว่า 30 คะแนน โดยสมมติว่าการแจกแจงเป็นปกติ

นักศึกษา	ทดสอบครั้งหลัง	ทดสอบครั้งก่อน	X_d	$(X_d - \bar{X}_d)$	$(X_d - \bar{X}_d)^2$
1	51	15	36	2	4
2	63	21	42	8	64
3	55	20	35	1	1
4	68	36	32	−2	4
5	48	12	36	2	4
6	38	9	29	−5	25
7	54	17	37	3	9
8	73	42	31	−3	9
9	49	26	23	−11	1
10	57	18	39	5	25
			340	0	266

$\bar{X}_d = 34$

วิธีทำ เราใช้การทดสอบค่าความแตกต่างที่เป็นคู่พร้อมด้วย μ_d = มัชฌิมเลขคณิตของค่าที่เพิ่มขึ้นในคะแนนสอบ ใช้ $\alpha = .05$

Ho : $\mu_d = 30$ คะแนน

Ha : $\mu_d > 30$ คะแนน

$$\bar{X}_d = 34 ;\ s_d^2 = \frac{266}{10-1} = 29.56$$

$$t = \frac{\bar{X}_d - \mu d}{\sqrt{s^2_d/n}} = \frac{34 - 30}{\sqrt{29.56/10}} = \frac{4}{1.72} = 2.33$$

ปฏิเสธ Ho ถ้าหากว่า $t > 1.83$ $(t_{.05,9})$

เนื่องจากว่าค่าสถิติ $t > 1.83$ เราจึงปฏิเสธ Ho มี $\alpha = .05$ นี้แสดงว่าค่าเพิ่มขึ้นเฉลี่ยมากกว่า 30 คะแนน สำหรับการทดสอบครั้งหลัง

ตัวอย่าง 4

คู่	ปฏิกิริยาต่อการกระตุ้นแสงสีเขียว X	ปฏิกิริยาต่อการกระตุ้นแสงสีแดง Y	X^2	Y^2	XY
1	28	25	784	625	700
2	26	27	676	729	702
3	33	28	1089	784	924
4	30	31	900	961	930
5	32	29	1024	841	928
6	30	30	900	900	900
7	31	32	961	1024	992
8	18	21	324	441	378
9	22	25	484	625	550
10	24	20	576	400	480
	274	268	7718	7330	7484

$Ho : \mu_x = \mu_y \; (\mu_x - \mu_y = 0)$

$Ha : \mu_x \neq \mu_y \; (\mu_x - \mu_y \neq 0)$

$\alpha = .05$

1. $\bar{X} = \dfrac{274}{10} = 27.4 \qquad \bar{Y} = \dfrac{268}{10} = 26.8$

2. $s_x = \dfrac{1}{n}\sqrt{n\Sigma X^2 - (\Sigma X)^2} = \dfrac{1}{10}\sqrt{10(7718) - (274)^2} = \dfrac{45.9}{10} = 4.59$

$s_y = \dfrac{1}{n}\sqrt{n\Sigma Y^2 - (\Sigma Y)^2} = \dfrac{1}{10}\sqrt{10(7330) - (268)^2} = \dfrac{38.4}{10} = 3.84$

3. $R_{xy} = \dfrac{n\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{(n\Sigma X^2 - (\Sigma X)^2)\,(n\Sigma Y^2 - (\Sigma Y)^2)}} = \dfrac{10(7484) - (274)(268)}{(45.9)(38.4)}$

$= \dfrac{1408}{(45.9)\,(38.4)} = .80$

4. $s_{\bar{x}} = \dfrac{s_x}{\sqrt{n-1}} = \dfrac{4.59}{10-1} = 1.53$; $s_{\bar{y}} = \dfrac{s_y}{\sqrt{n-1}} = \dfrac{3.84}{\sqrt{10-1}} = 1.28$

5. $s_{\bar{x}-\bar{y}} = \sqrt{s^2_{\bar{x}} + s^2_{\bar{y}} - 2r_{xy}s_{\bar{x}}s_{\bar{y}}} = \sqrt{(1.53)^2 + (12.8)^2 - 2(.80)(1.53)(1.28)}$

$= \sqrt{3.98 - 3.12} = \sqrt{.86} = .93$

6. $Z = \dfrac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{s_{\bar{x}-\bar{y}}} = \dfrac{(27.4 - 26.8) - o}{.93}$

$= +.65$

เนื่องจากว่าเขตยอม Ho อยู่ระหว่าง Z = −1.96 กับ 1.96 การตัดสินใจของเราคือ ยอมรับ Ho ที่ $\alpha = .05$

เราอาจดำเนินการทดสอบได้อีกวิธีการหนึ่ง (แบบตัวอย่างที่สี่)

คู่	X	Y	D	D^2
1	28	25	3	9
2	26	27	−1	1
3	33	28	5	25
4	30	31	−1	1
5	32	29	3	9
6	30	30	0	0
7	31	32	−1	1
8	18	21	−3	9
9	22	25	−3	9
10	24	20	4	16
			6	80

การคำนวณ

$\bar{D} = \dfrac{\Sigma D}{n} = \dfrac{6}{10} = .6$

$s^2_d = \dfrac{1}{n-1}\left[\Sigma D^2 - \dfrac{(\Sigma D)^2}{n}\right]$

$= \dfrac{1}{10-1}\left[80 - \dfrac{6^2}{10}\right] = \dfrac{76.4}{9}$

$sd = 2.91$

$s_{\bar{d}} = \dfrac{sd}{\sqrt{n}} = \dfrac{2.91}{\sqrt{10}} = .92$

$Ho : \mu_d = o$

$Ha : \mu_d \neq o$

$Z = \dfrac{\bar{D} - \mu_d}{s_{\bar{d}}} = \dfrac{.6 - 0}{.92} = .65$

เนื่องจากว่าเขตยอมรับ Ho อยู่ระหว่าง Z = −1.96 กับ 1.96 การตัดสินใจของเราคือ ยอมรับ Ho ที่ $\alpha = 0.05$

ในตัวอย่างที่ 4 นี้เหมาะสมสำหรับกรณีมัชฌิมเลขคณิตที่สัมพันธ์กันของตัวอย่างขนาดใหญ่เท่านั้นจึงสามารถเปรียบเทียบกันได้ระหว่างวิธีการทั้งสอง

ข้อดีในการใช้ตัวอย่างที่มีความสัมพันธ์กัน คือต้องการลดการกระจายที่คาดหวังระหว่างสองมัชฌิมเลขคณิตของตัวอย่าง ความแปรปรวนของผลต่างของมัชฌิมเลขคณิตของสองตัวอย่างที่มีความสัมพันธ์กันมีค่าน้อยกว่าความแปรปรวนของผลต่างของสองมัชฌิมเลขคณิตที่อิสระกัน

$$(\sigma_{\bar{x}}^2 + \sigma_{\bar{y}}^2 - 2\rho_{xy}\sigma_{\bar{x}}\sigma_{\bar{y}}) < \sigma_{\bar{x}}^2 + \sigma_{\bar{y}}^2)$$

ข้อเสียในการใช้ตัวอย่างที่มีความสัมพันธ์กัน คือค่าสังเกตต้องคงอยู่ในประชากรของเรื่องที่จับคู่กันและอยู่บนพื้นฐานของคุณลักษณะแสดงออกถึงความสัมพันธ์ของค่าสังเกตที่ต้องการหาการเลือกตัวอย่างเป็นคู่โดยสุ่มจะต้องเลือกจากประชากรนี้ และจะต้องทราบ ϱ_{xy} สหสัมพันธ์ระหว่างคู่ของค่าสังเกตในหมู่ประชากร

9.23 ความแตกต่างระหว่างค่าเฉลี่ยของตัวอย่างที่มีขนาดเล็ก

ถ้าตัวอย่างของเราเลือกมาจากประชากรที่เข้าใกล้เส้นโค้งปกติ และ $\sigma_1 = \sigma_2 \ (= \sigma)$ นัยสำคัญของการทดสอบผลต่างระหว่างค่าเฉลี่ยของตัวอย่างขนาดเล็กทั้งสองจะเป็นไปตาม Student-t distribution เราใช้การแจกแจงตัวอย่างของตัวสถิติ

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad \text{.........(9.23.1)}$$

ให้ $\sigma_1 = \sigma_2 = \sigma$ สูตร (19.23.2) ก็จะได้เป็น

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \qquad \text{.........(9.23.2)}$$

ถ้าเราประมาณค่า σ ในสูตรนี้ด้วยตัวสถิติ

$$\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \qquad \text{.........(9.23.3)}$$

(เราก็จะได้ตัวหารของ (9.23.1) ภายในรู้ทของ (9.23.3) เป็นส่วนเฉลี่ยถ่วงน้ำหนักของ s_1^2 กับ s_2^2 และเนื่องจาก $n_1 - 1$ ของส่วนเบี่ยงเบนจากค่าเฉลี่ยไม่ขึ้นแก่กันใน s_1^2 และ $n_2 - 1$ ของส่วนเบี่ยงเบนจากค่าเฉลี่ยไม่ขึ้นแก่กันใน s_2^2 เรามีส่วนเบี่ยงเบนที่ไม่ขึ้นแก่กันทั้งหมด $(n_1 - 1) + (n_2 - 1)$

$= n_1 + n_2 - 2$ จากค่าเฉลี่ยซึ่งเป็นตัวค่าประมาณพื้นฐานของเราของ σ ที่มีองศาแห่งความอิสระจำนวนมาก) เป็น student-t distribution ที่มีองศาแห่งความอิสระ $n_1 + n_2 - 2$ เราก็สามารถทดสอบสมมติฐาน

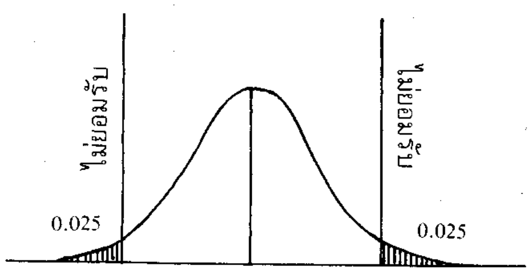
Hypothesis (Ho) : $\mu_1 = \mu_2$

เทียบ

Alternative (H_a) : $\mu_1 \neq \mu_2$

ด้วยเกณฑ์ดังแสดงในรูป 9.20 ใช้ระดับนัยสำคัญ 0.05 เกณฑ์ของเราอ่านได้

ไม่ยอมรับสมมติฐานถ้า $t < -t_{.025}$ หรือ $t > t_{.025}$ ยอมรับสมมติฐาน ถ้า $-t_{.025} \leqslant t \leqslant t_{.025}$ ในเมื่อ t คำนวณได้จากสูตร (9.23.1) และองศาแห่งความอิสระเท่ากับ $n_1 + n_2 - 2$



รูปที่ 9.20

เพื่อแสดงกระบวนการนี้ สมมติว่าเราต้องการศึกษาเพื่อเปรียบเทียบจำนวนนิโคตินของบุหรี่สองชนิด บุหรี่ชนิดเกล็ดทอง 10 มวน มีจำนวนนิโคตินเฉลี่ย 23.1 มิลลิกรัม ส่วนเบี่ยงเบนมาตรฐาน 1.5 มิลลิกรัม ขณะเดียวกัน 8 มวน ของบุหรี่สามิต มีจำนวนนิโคตินเฉลี่ย 22.7 มิลลิกรัม ส่วนเบี่ยงเบนมาตรฐาน 1.7 มิลลิกรัม (เราสมมติว่าตัวอย่างทั้งสองนี้เป็นตัวอย่างสุ่มจากประชากรเชิงปกติและเนื่องจากส่วนเบี่ยงเบนมาตรฐานของตัวอย่างเข้าใกล้ตัวอย่างทั้งสองที่สุ่มมาจากประชากรที่มีส่วนเบี่ยงเบนมาตรฐานเท่ากัน) ใช้ระดับนัยสำคัญ 0.05

แทนค่า $n_1 = 10$, $\bar{x}_1 = 23.1$, $s_1 = 1.5$, $n_2 = 8$, $\bar{x}_2 = 22.7$ และ $s_2 = 1.7$ ใน (9.11.1) เราคำนวณได้

$$t = \frac{23.1 - 22.7}{\sqrt{\frac{9(1.5)^2 + 7(1.7)^2}{16}\left(\frac{1}{10} + \frac{1}{8}\right)}} = 0.53$$

เนื่องจาก $t_{.025} = 2.12$ สำหรับ $n_1 + n_2 - 2 = 10 + 8 - 2 = 16$ degrees of freedom (ดูตาราง t) เราพบว่าผลต่างระหว่างค่าเฉลี่ยของนิโคตินของบุหรี่สองชนิดไม่มีนัยสำคัญที่ $\alpha = .05$

ตัวอย่าง 2 พ่อค้าแลกรถเก่ามีความสนใจเกี่ยวกับการเปรียบเทียบอายุของรถที่ทำจากบริษัท ก กับ บริษัท ข ที่จะนำมาแลกเปลี่ยน จากประสบการณ์ของเขาทำให้เขาเชื่อว่า โดยทั่ว ๆ ไป รถที่ทำจากบริษัท ก มาทำการแลกเปลี่ยนกับรถที่ทำจากบริษัท ข จะมีอายุเก่ามากกว่าหนึ่งปี จะทดสอบความเชื่อของเขาโดยใช้ $\alpha = .025$ และข้อมูลกำหนดได้จากตาราง

อายุของรถในการแลกเปลี่ยน

รถที่ทำจากบริษัท ก.				รถที่ทำจากบริษัท ข.			
3	3	4	5	2	3	3	4
5	5	6	6	4	4	4	5
7	8	8		4	4	4	
				5	6	4	

วิธีทำ $\mu_ก$ = อายุเฉลี่ยของรถทำจากบริษัท ก ที่เอามาแลกเปลี่ยน

$\mu_ข$ = อายุเฉลี่ยของรถทำจากบริษัท ข ที่เอามาแลกเปลี่ยน

Ho : $\mu_ก - \mu_ข = 1$

Ha : $\mu_ก - \mu_ข > 1$ $(\mu_ก > \mu_ข + 1)$

$$t = \frac{(\bar{x}_ก - \bar{x}_ข) - (\mu_ก + \mu_ข)}{\sqrt{s_p^2\left[\left(\frac{1}{n_ก}\right) + \left(\frac{1}{n_ข}\right)\right]}}$$

$$\bar{X}_ก = \frac{\Sigma x}{n_ก} = \frac{3 + 5 + 7 + \dots\dots + 5 + 6}{11}$$

$$= \frac{60}{11} = 5.45$$

$$\bar{X}_{ข} = \frac{2 + 4 + \ldots\ldots + 4 + 5}{14}$$

$$= \frac{56}{14} = 4$$

$$s_p^2 = \frac{\Sigma(X_{ก i} - \bar{X}_{ก})^2 + \Sigma(X_{ข i} - \bar{X}_{ข})^2}{n_{ก} + n_{ข} - 2}$$

ในเมื่อ $\Sigma(X_{ก i} - \bar{X}_{ก})^2 = (3 - 5.45)^2 + (5 - 5.45)^2 + \ldots.. + (6 - 5.45)^2$

$$= 30.525$$

$$\Sigma(X_{ข i} - \bar{X}_{ข})^2 = (2 - 4)^2 + (4 - 4)^2 + \ldots.. + (5 - 4)^2$$

$$= 8$$

$$s_p^2 = \frac{30.525 + 8}{11 + 14 - 2} = \frac{38.525}{23}$$

$$= 1.6838$$

$$t = \frac{(5.45 - 4.00) - 1}{\sqrt{1.68\left(\frac{1}{11} + \frac{1}{14}\right)}} = \frac{.45}{.522}$$

$$= .861$$

เราปฏิเสธ Ho ถ้าหากว่า $t > 2.069$ ($t_{.025:23}$)

เราไม่สามารถปฏิเสธ Ho ที่ $\alpha = 0.025$

โดยทั่ว ๆ ไปเจ้าของรถที่ทำจากบริษัท ก ที่ได้นำรถไปแลกเปลี่ยนกับพ่อค้านี้ส่วนมากจะมีอายุเก่ากว่าเจ้าของรถที่ทำจากบริษัท ข หนึ่งปี

หมายเหตุ เพราะว่าตัวอย่างเป็นตัวอย่างขนาดเล็กท่านควรจะเข้าใจด้วยว่าภัยที่เกิดจากความคลาดเคลื่อนประเภทที่สองอาจสูงมาก ดังนั้นโอกาสที่การตัดสินใจนี้อาจผิดได้

แบบฝึกหัดที่ 9.1

1. ตัวอย่างสุ่มหนึ่งประกอบด้วย แยม 30 ขวด "แต่ละขวดมีน้ำหนักสุทธิ 12 ออนซ์" มีป้ายปิดน้ำหนักเฉลี่ยสุทธิ 11.9 ออนซ์ ส่วนเบี่ยงเบนมาตรฐาน 0.3 ออนซ์ ป้ายปิดน้ำหนักที่ขวดนี้สามารถทำให้เราไม่ยอมรับสมมติฐาน $\mu = 12$ ออนซ์ หรือ ถ้าเราทดสอบเทียบกับ alternative hypothesis $\mu < 12$ ออนซ์ ที่ระดับนัยสำคัญ 0.01 ($z = -1.82$, ยอมรับ)

2. ทดสอบสมมติฐานว่าผู้โดยสารรถส่วนตัว ซึ่งไม่ได้ใช้ในวัตถุประสงค์ทางธุรกิจได้ขับรถถัวเฉลี่ย 12,000 ไมล์ต่อปี ถ้าตัวอย่างสุ่มประกอบด้วยรถ 100 คัน ขับรถถัวเฉลี่ย 12,750 ไมล์ต่อปี ส่วนเบี่ยงเบนมาตรฐาน 2,500 ไมล์ ใช้ alternative ทั้งสองข้างและระดับนัยสำคัญ 0.05 ($z = 3$, ไม่ยอมรับ)

3. ตัวอย่างสุ่มประกอบด้วยท่อนเหล็กกล้า 8 ท่อน มีความต้านทานแรงกลเฉลี่ย 56,598 psi กับส่วนเบี่ยงเบนมาตรฐาน 564 psi ทดสอบสมมติฐานว่าความต้านทานแรงเฉลี่ยจริงของท่อนเหล็กกล้าที่ได้จากตัวอย่างนี้เท่ากับ 56,000 psi ใช้ two tailed test กับระดับนัยสำคัญ 0.05 ผลของการสรุปของท่านควรจะเหมือนกันหรือถ้าท่านใช้ระดับนัยสำคัญ 0.01 ($t = 3.00$ ไม่ยอมรับ)

4. ผลการทดสอบ (ใช้ข้อสอบชนิดเดียวกัน) ของตัวอย่างที่ได้สุ่มมาจากสองโรงเรียน ๆ ละ 50 คน ดังปรากฏผลดังนี้

   $n_1 = 50 \quad \bar{x}_1 = 89 \quad s_1 = 4$

   $n_2 = 50 \quad \bar{x}_2 = 92 \quad s_2 = 3$

   ทดสอบความแตกต่างระหว่างค่าเฉลี่ยของตัวอย่างทั้งสองที่ระดับนัยสำคัญ 0.05 ว่ามีนัยสำคัญหรือไม่ ($z = -4.24$ มีนัยสำคัญ)

5. ในการสำรวจอุปนิสัยการซอบซื้อของของผู้หญิงจ่ายตลาด 400 คน ที่ได้สุ่มเลือกมาจากซูเปอร์มาร์เกต เอ. ในนครแห่งหนึ่ง ค่าใช้จ่าย อาหารเฉลี่ยสัปดาห์ละ 40 ดอลลาร์ ส่วนเบี่ยงเบนมาตรฐาน 12 ดอลลาร์ สำหรับผู้หญิงจ่ายตลาดอีก 400 คน ที่ได้สุ่มเลือกมาจากซูเปอร์มาร์เกต บี. ในนครแห่งเดียวกัน ค่าใช้จ่ายอาหารเฉลี่ยสัปดาห์ละ 32 ดอลลาร์ ส่วนเบี่ยงเบนมาตรฐาน 15 ดอลลาร์ ทดสอบที่ระดับนัยสำคัญ 0.05 ว่าค่าใช้จ่ายอาหารเฉลี่ยรายสัปดาห์ของผู้จ่ายตลาดของประชากรทั้งสองจากตัวอย่างที่เลือกมาเท่ากันหรือไม่ ($z = 8.33$ มีนัยสำคัญ)

6. ตัวอย่างสุ่มหนึ่งประกอบด้วยร้านค้าขนาดกลาง 8 ร้าน ในพระนคร ได้ลดเปอร์เซ็นต์ 11.4, 8.3, 7.9, 10.3, 7.8, 10.7, 9.4 และ 8.6 ขณะเดียวกันอีกตัวอย่างสุ่มหนึ่งประกอบด้วย 10 ร้าน ในจังหวัดเชียงใหม่ได้ลดเปอร์เซ็นต์ 9.8, 8.9, 8.8, 10.3, 12.2, 9.5, 11.5, 11.0, 7.8 และ 10.2 ผลต่างของค่าเฉลี่ย (ระดับนัยสำคัญ 0.05) ในการลดเปอร์เซ็นต์ของสองจังหวัดนี้มีนัยสำคัญหรือไม่ ($t = -1.09$ ไม่มีนัยสำคัญ)

7. การทดลองทางเกษตรปรากฏว่า ผลจากการปลูกข้าวโพดชนิดหนึ่ง 6 แปลง ให้ผลผลิตเฉลี่ย 85.3 บุชเชลต่อเอเคอร์ ส่วนเบี่ยงเบนมาตรฐาน 5.8 บุชเชลต่อเอเคอร์ ขณะเดียวกันผลผลิตเฉลี่ยจากข้าวโพดอีกชนิดหนึ่ง 6 แปลง ได้ 92.7 บุชเชลต่อเอเคอร์ ส่วนเบี่ยงเบนมาตรฐาน 6.1 บุชเชลต่อเอเคอร์ จงแสดงว่าผลต่างระหว่างผลผลิตเฉลี่ย มีนัยสำคัญ 0.05

9.24 การทดสอบความแตกต่างระหว่างสัดส่วน

ปัญหาที่ต้องตัดสินใจบ่อย ๆ คือ ความแตกต่างของข้อมูล ระหว่างสัดส่วนของสองตัวอย่างเป็นนัยสำคัญหรือไม่ ดังตัวอย่าง จากการตรวจสอบพบว่าเสมียนคนหนึ่งบรรจุรายงานลงในแฟ้มผิด 7 ใน 200 รายงาน ในขณะเดียวกันอีกคนหนึ่งบรรจุรายงานลงในแฟ้มผิด 10 ใน 400 เราอาจต้องการที่จะตัดสินใจความแตกต่างระหว่าง $\frac{7}{200} = 0.035$ กับ $\frac{10}{400} = 0.025$ ว่าโดยเฉลี่ยเสมียนทั้งสองไม่ได้ทำผิดจำนวนเท่ากัน ในทำนองเดียวกันกระบวนการหนึ่งผลิตสินค้าเสียไป 16 ใน 400 ชิ้น ขณะเดียวกันกระบวนการอื่นก็ได้ผลิตสินค้าเสียไป 24 ใน 300 ชิ้น อยากทราบว่า ความแตกต่างระหว่าง $\frac{16}{400} = 0.04$ กับ $\frac{24}{300} = 0.08$ เป็นไปสมเหตุผลหรือไม่

ถ้า x_1 และ x_2 เป็นจำนวนของความสำเร็จ ซึ่งได้จากตัวอย่างสุ่มที่มีขนาดใหญ่ n_1 และ n_2 โดยการเลือกมาจากประชากรที่มีสัดส่วน p_1 และ p_2 การแจกแจงตัวอย่างของสถิติ

$$\frac{x_1}{n_1} - \frac{x_2}{n_2}$$

สามารถประมาณให้เป็นเส้นโค้งปกติซึ่งมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานเป็น

$$\mu = p_1 - p_2 \qquad \text{.........(9.24.2)}$$

และ

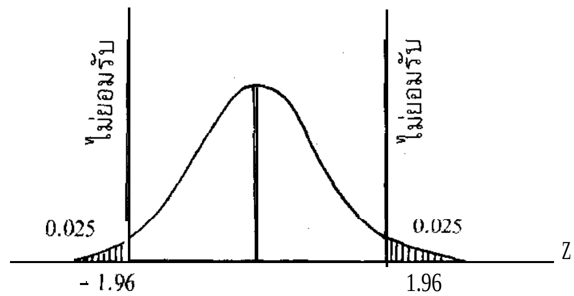
$$\sigma_{\frac{x_1}{n} - \frac{x_2}{n}} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \qquad \text{..........(9.24.1)}$$

เพื่อแสดงความหมายของการแจกแจงตัวอย่างของผลต่างระหว่าง สัดส่วนทั้งสอง สมมติว่า p_1 เป็นสัดส่วนจริงของผู้มีสิทธิเลือกตั้งชายในพระนคร ซึ่งจะเลือกหมายเลข 1 และ p_2 เป็น

สัดส่วนจริงของผู้มีสิทธิเลือกตั้งหญิง ถ้าเราส่งคนงานเป็นจำนวนมากไปสัมภาษณ์ผู้ชายที่มีสิทธิออกเสียงเลือกตั้งโดยวิธีสุ่ม มีขนาดตัวอย่างเท่ากับ n_1 และผู้มีสิทธิ์ออกเสียงเลือกตั้งหญิงมีขนาดตัวอย่างเท่ากับ n_2 ในพระนคร เราได้ตัวเลข x_1 และ x_2 ของจำนวนผู้มีสิทธิออกเสียงชายและหญิงที่เลือกผู้สมัครหมายเลข 1 ไม่เหมือนกันหมด ถ้าเราเอาค่าต่าง ๆ มาทำเป็นหมู่ซึ่งคนงานได้เก็บรวบรวมผลต่างระหว่าง x_1/n_1 กับ x_2/n_2 เราก็จะได้การแจกแจงตัวอย่างที่ทำการทดลองของตัวสถิติ $(x_1/n_1) - (x_2/n_2)$ การแจกแจงตัวอย่างที่ได้กล่าวข้างต้นก็จะสมนัยกับการแจกแจงตัวอย่างในทางทฤษฎี

เนื่องจากเราสนใจการทดสอบ $p_1 = p_2$ เราก็ตั้งสมมติฐาน $p_1 = p_2 (= p)$ และ alternative ทั้งสองข้าง $p_1 \neq p_2$ แทนค่า p สำหรับ p_1 และ p_2 ใน (9.24.1) และ (9.24.2) เราพบว่าภายใต้สมมติฐานนี้ สูตรสำหรับค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของการแจกแจงตัวอย่างของ $(x_1/n_1) - (x_2/n_2)$ เป็น

รูปที่ 9.21



$$\mu = 0 \qquad \text{..........(9.24.3)}$$

และ

$$\sigma_{\frac{x_1}{n_1} - \frac{x_2}{n_2}} = \sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \qquad \text{..........(9.24.4)}$$

เนื่องจากไม่สามารถคำนวณได้จากสูตร (9.24.4) ถ้าไม่รู้ค่า p และ p ก็ไม่รู้เสียด้วย เราประมาณค่า p ได้โดยเอาความสำเร็จของสองตัวอย่างผสมกันดังเช่น

$$\frac{x_1 + x_2}{n_1 + n_2} \qquad \text{..........(9.24.5)}$$

เราใช้การประมาณด้วยเส้นโค้งปกติ นี้เป็นเหตุผลหนึ่งที่ว่าทำไมตามที่กล่าวข้างต้นจึงใช้ตัวอย่างที่มีขนาดใหญ่

เราจะมีการทดสอบสมมติฐานเกี่ยวกับอัตราส่วนทั้งสองได้อย่างไร ดังตัวอย่างจากข้างต้นเกี่ยวกับเสมียน 5 คน ที่บรรจุรายงานลงในแฟ้มผิด ให้ p_1 และ p_2 เป็นความน่าจะเป็นที่เสมียนทั้งสองบรรจุรายงานลงในแฟ้มผิด สมมติฐานของเราก็คือ

Hypothesis (Ho) : $p_1 = p_2$

Alternative (H_a) : $p_1 \neq p_2$

เราให้ระดับนัยสำคัญเป็น 0.05 พิจารณาเส้นโค้งปกติของรูปที่ 9.21 ซึ่งใช้แทนการแจกแจงตัวอย่างของ $(x_1/n_1) - (x_2/n_2)$ เราก็สามารถตั้งเกณฑ์ (Criterion) ดังนี้

ไม่ยอมรับสมมติฐานถ้า $z < -1.96$ หรือ $z > 1.96$ ยอมรับสมมติฐาน ถ้า $-1.96 \leqslant z \leqslant 1.96$ ในเมื่อ z คำนวณได้จากสูตร

$$z = \frac{\frac{x_1}{n_1} - \frac{x_2}{n_2}}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad \text{.........(9.24.6)}$$

ในเมื่อ $p = (x_1 + x_2) / (n_1 + n_2)$

สูตรคำนวณค่า z หาได้จากการลบกันจากผลต่างระหว่าง สัดส่วนของตัวอย่างทั้งสองกับค่าเฉลี่ยของการแจกแจงตัวอย่าง กำหนดให้ใช้ (9.24.3) แล้วหารด้วยส่วนเบี่ยงเบนมาตรฐานที่กำหนดให้ใช้ (9.24.4)

แทนค่าตัวเลขโดย $x_1 = 7, n_1 = 200, x_2 = 10$ และ $n_2 = 400$ สูตร (9.24.5) และ (9.24.6) ก็จะได้

$$p = \frac{7 + 10}{200 + 400} = 0.028$$

$$z = \frac{\frac{7}{400} - \frac{10}{200}}{\sqrt{(0.028)(0.972)\left(\frac{1}{200} + \frac{1}{400}\right)}} = 0.71$$

เนื่องจากค่านี้อยู่ระหว่าง -1.96 กับ 1.96 เรายอมรับสมมติฐานที่ $p_1 = p_2$ (สมมติฐานที่อัตราส่วนของการบรรจุรายงานลงในแฟ้มผิดโดยเฉลี่ยของเสมียนสองคนเหมือนกัน) เราก็สามารถกล่าวได้ว่าผลต่างระหว่าง สัดส่วนของตัวอย่างทั้งสองไม่มีนัยสำคัญนี้หมายความว่า ผลต่างระหว่าง 0.035 กับ 0.025 ไม่มากพอที่จะสรุปว่าเสมียนคนหนึ่งจะดีกว่าอีกคนหนึ่ง และเสมียนทั้งสองก็ไม่ได้ทำงานดีเท่ากัน

(**หมายเหตุ** Hypothesis $p_1 = p_2$ เราใช้ Null hypothesis : $p_1 = p_2$ ก็ได้ ใช้อักษรย่อ Ho ส่วน Alternative hypothesis ใช้อักษรย่อ H_a ขึ้นอยู่กับความนิยม อย่างเช่น Ho : $p_1 = p_2$; Ha : $p_1 \neq p_2$)

แบบฝึกหัดที่ 9.2

1. จากข่าวพอที่จะเชื่อถือได้ว่า 30 เปอร์เซ็นต์ของนักศึกษาทั้งหมดที่เข้าศึกษาในวิทยาลัยลาออกระหว่างหรือภายหลังปีแรก ทดสอบข่าวนี้ against alternative ว่า จำนวนเปอร์เซ็นต์น้อยกว่า 30 เปอร์เซ็นต์ ถ้าตัวอย่างสุ่มที่มีจำนวนนักเรียน 500 คน ผู้ซึ่งเข้าศึกษาในวิทยาลัย 2495 มี 124 คน ลาออกระหว่างหรือภายหลังปีแรกใช้ระดับนัยสำคัญ 0.05 ($z = -2.54$ ไม่ยอมรับ)

2. บุคลากรคนหนึ่งพูดว่า 60 เปอร์เซ็นต์ของหญิงโสดทำงานเลขานุการินีจะเลิกงานภายหลังแต่งงาน 2 ปี ทดสอบสมมติฐาน against alternative ที่ $p \neq 0.60$ ถ้าหากว่าระหว่างเลขานุการินี 200 คน 112 คน เลิกงานภายหลังแต่งงาน 2 ปี ใช้ระดับนัยสำคัญ 0.05

3. องค์การวิจัยแห่งหนึ่งได้ทำการทดสอบสมมติฐานว่า ผู้สมัครรับเลือกตั้ง นาย ก. จะได้รับเสียง 55 เปอร์เซ็นต์ against alternative ว่าผู้สมัครรับเลือกตั้ง นาย ก. จะได้รับเสียงน้อยกว่า 55 เปอร์เซ็นต์ องค์การวิจัยนี้ควรจะตัดสินใจอย่างไรที่ระดับนัยสำคัญ 0.01 ถ้าหากว่าตัวอย่างสุ่มจากผู้มีสิทธิออกเสียง 1,000 คน 506 คน พอใจผู้สมัครรับเลือกตั้ง นาย ก. ($z = -2.80$ ไม่ยอมรับ)

4. ทดสอบสมมติฐานว่า อัตราการย้ายครอบครัวทุก ๆ 3 ปี ต่อครั้งหนึ่งเท่ากับ 0.50 จากการสำรวจตัวอย่างแสดงว่าระหว่าง 400 ครอบครัว 218 ครอบครัวย้ายระหว่าง 3 ปีที่ผ่านมา ใช้ระดับนัยสำคัญ 0.05

5. การโฆษณาของบริษัทผลิตรถยนต์ว่า 35 เปอร์เซ็นต์ของรถยนต์ทั้งหมดสร้างโดยบริษัทในปี 1945 ยังมีภาวะดีอยู่ จงแสดงว่าการโฆษณานี้เป็นที่ยอมรับที่ระดับนัยสำคัญ 0.01 ($H_a: p \neq 0.35$) ถ้าตัวอย่างสุ่มของจำนวนรถยนต์ 800 คัน 257 คัน ยังมีภาวะดีอยู่ในปี 1955 ($z = -1.76$, ยอมรับ)

6. ร้านขายของชำส่วนใหญ่เชื่อจากประสบการณ์ว่าอย่างน้อย 70 เปอร์เซ็นต์ของผู้หญิงไปซื้อของชอบบริการด้วยตัวเองอย่างเดียวกับร้าน supermarket ในการติดต่อเพื่อสร้างร้านใหม่ บริษัทได้สำรวจตลาดจากตัวอย่างของผู้หญิงไปซื้อของโดยการสัมภาษณ์ 600 คน 406 คน ชอบบริการด้วยตัวเอง ทดสอบสมมติฐาน $p = 0.70$ against alternative hypothesis $p < 0.70$ ที่ระดับนัยสำคัญ 0.05

7. กระบวนการผลิตสินค้าสำเร็จรูป 2 กระบวนการ ได้ผลิตสินค้าเสียไป 16 กับ 24 ชิ้น ในตัวอย่างที่มีขนาด 400 และ 300 ตามลำดับ ใช้ระดับนัยสำคัญ 0.05 เพื่อทดสอบความแตกต่างระหว่าง สัดส่วนของตัวอย่างทั้งสองของสินค้าที่เสียว่ามีนัยสำคัญหรือไม่ ($z = -2.26$; ความแตกต่างมีนัยสำคัญ)

8. จากการศึกษาตลาดในนครนิวยอร์คปรากฏผลจากการเลือกตัวอย่างแม่บ้าน 100 คน 68 คน ชอบเหล้าไวท์ชนิด A มากกว่าชนิด B จากการศึกษาอย่างเดียวกันในลอสแองเจลิส 213 คน ใน 300 แม่บ้านชอบเหล้าไวท์ชนิด A มากกว่าชนิด B ใช้ระดับนัยสำคัญ 0.01 เพื่อทดสอบความแตกต่างระหว่าง สัดส่วนทั้งสองที่ชอบเหล้าไวท์ชนิดนี้ว่ามีนัยสำคัญหรือไม่

9. ห้างร้านค้าหนึ่งได้รวบรวมข้อมูลเพื่อศึกษาบัญชีค้างจ่าย จากตัวอย่างหนึ่งประกอบด้วย 600 บัญชีรายตัว ซึ่งเปิดบัญชีโดยบุคคลที่อาศัยอยู่ในชุมชนนั้นมากกว่า 5 ปี มี 58 ราย ที่ไม่ชำระในครั้งเดียวจากอีกตัวอย่างหนึ่งประกอบด้วย 400 บัญชีรายตัว ซึ่งเปิดบัญชีโดยบุคคลที่อาศัยอยู่ในชุมชนนั้นน้อยกว่า 5 ปี มี 26 ราย ที่ไม่ชำระในครั้งเดียว ใช้ระดับนัยสำคัญ 0.05 เพื่อทดสอบความแตกต่างระหว่าง สัดส่วนของบัญชีค้างจ่ายว่ามีนัยสำคัญหรือไม่ ($z = 1.8$, ความแตกต่างไม่มีนัยสำคัญ)

10. ตัวอย่างหนึ่งประกอบด้วยใบสั่งซื้อสินค้า 500 ราย ในเขตแลกเปลี่ยนหนึ่งพบว่า 300 ราย ดำเนินงานโดยผู้หญิง ขณะเดียวกันมีอีกตัวอย่างหนึ่งประกอบด้วยใบสั่งซื้อสินค้า 500 ราย ในอีกเขตแลกเปลี่ยนหนึ่งพบว่า 250 ราย ดำเนินงานโดยผู้หญิงนี้จะเป็นหลักฐานหรือว่าสัดส่วนของลูกค้าผู้หญิงเหมือนกันในการแลกเปลี่ยนทั้งสอง สำหรับในระยะเวลาที่เกี่ยวข้อง ใช้ระดับนัยสำคัญ 0.05

11. ตัวอย่างสุ่มประกอบด้วยข้อมูล 16 จำนวน ที่ได้เลือกมาจากประชากรปกติซึ่งมีค่ามัชฌิมเท่ากับ 50 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 10 ข้อมูลของตัวอย่างมีดังนี้

62	43	60	49	72	56	45	46
37	56	41	43	36	45	56	49

ถ้าหากว่าไม่รู้ค่ามัชฌิมของประชากร ใช้ 5% ระดับนัยสำคัญ ทดสอบสมมติฐานว่า ค่ามัชฌิมของประชากรเท่ากับ

(ก) 40, (ข) 49,

(ค) 50, (ง) 51

(จ) 60 ใช้ two tailed test วิธีดำเนินการเป็นไปตามหัวข้อ (9.9) คำนวณค่า z และสรุป

เนื่องจากรู้ค่ามัชฌิมจริงของประชากรเท่ากับ 50 ก็สามารถกำหนดคำสรุปว่าถูกหรือผิด สำหรับแต่ละกรณีใน 5 กรณี ผลจากการสรุปถูกต้องหรือไม่ควรจะทดสอบแบบความคลาดเคลื่อนชนิดที่หนึ่ง หรือชนิดที่สอง

แบบฝึกหัดนี้ตั้งใจจะให้แสดงการกระทำความคลาดเคลื่อนชนิดที่สอง ถ้าค่ามัชฌิมของประชากรตามสมมติฐานมีค่าใกล้เคียงกับค่ามัชฌิมจริงของประชากร

(ก) $z = 3.90$; ไม่ยอมรับ (ข) $z = 0.30$; ยอมรับ

(ค) $z = -0.10$; ยอมรับ (ง) $z = -0.50$; ยอมรับ

(จ) $z = -4.01$; ไม่ยอมรับ

12. ตัวอย่างสุ่มหนึ่งประกอบด้วยข้อมูล 2500 ซึ่งเลือกมาโดยวิธี with replacement จากป้ายในตะกร้ามีค่ามัชฌิมของตัวอย่างเท่ากับ 49.9 โดยใช้ 1% ระดับนัยสำคัญ ทดสอบสมมติฐานทั้งห้าอย่างเดียวกับแบบฝึกหัดที่ 11

แบบฝึกหัดนี้ตั้งใจจะแสดงการลด probability ในการกระทำความคลาดเคลื่อนชนิดที่หนึ่งและชนิดที่สองในเวลาเดียวกัน โดยเพิ่มขนาดของตัวอย่าง

(ก) $z = 49.5$; ไม่ยอมรับ (ข) $z = 4.5$; ไม่ยอมรับ

(ค) $z = -0.5$; ยอมรับ (ง) $z = -5.5$; ไม่ยอมรับ

(จ) $z = -50.5$; ไม่ยอมรับ

13. ตัวอย่างสุ่มประกอบด้วยข้อมูล 25 จำนวน ซึ่งเลือกจากป้ายในตะกร้า ($\mu = 50, \sigma = 10$) ค่ามัชฌิมของตัวอย่างเท่ากับ 54 ใช้ระดับนัยสำคัญ 1% และ 5% ทดสอบสมมติฐานว่าค่ามัชฌิมของประชากรเท่ากับ (ก) 50 และ (ข) 49 มีสี่การทดลองด้วยกัน การทดสอบแต่ละครั้ง หาค่า z และสรุป เนื่องจากรู้ค่ามัชฌิมของประชากรเท่ากับ 50 จึงสามารถกำหนดคำสรุปว่าถูกหรือผิด สำหรับแต่ละกรณีในสี่กรณี ผลการสรุปถูกหรือไม่ หรือควรจะทำความคลาดเคลื่อนชนิดที่หนึ่งหรือที่สอง

แบบฝึกหัดนี้ตั้งใจจะแสดงหลักความจริงว่าการเปลี่ยนระดับนัยสำคัญโดยไม่เปลี่ยนขนาดตัวอย่างก็จะดีในบางสิ่งบางอย่างและเสียบางสิ่งบางอย่าง การใช้ 5% ระดับนัยสำคัญ ชอบที่จะนำไปสู่ความคลาดเคลื่อนชนิดที่หนึ่ง และไม่ชอบที่จะนำไปสู่ความคลาดเคลื่อนชนิดที่สองกว่าการใช้ 1% ระดับนัยสำคัญ