



ารทดสอบนัยสำคัญทางสถิติ Statistical Significance Test

- ผู้ช่วยศาสตราจารย์ บุญหญิง สมร่าง
- สาขาวิชาสถิติประยุกต์
- คณะวิทยาศาสตร์และเทคโนโลยี
- มหาวิทยาลัยหอการค้าไทย
-
- **Assistant Professor Boonying Somrang**
- Department of Applied Statistics
- School of Science and Technology
- University of the Thai Chamber of Commerce
- E-mail: boonying__som@utcc.ac.th

บทคัดย่อ

บทความนี้มีวัตถุประสงค์เพื่อให้ผู้วิจัยได้มีความเข้าใจเกี่ยวกับกระบวนการหาคำตอบโดยการทดสอบนัยสำคัญทางสถิติ ปัจจุบันการทดสอบจะพิจารณาจากค่า P-value ซึ่งเป็นค่าความน่าจะเป็นของค่าสถิติทดสอบภายใต้เงื่อนไขว่าสมมติฐานหลัก (H_0) เป็นจริง ถ้าค่า P-value มีค่าน้อย แสดงว่ามีโอกาสน้อยที่ค่าสถิติจากตัวอย่างจะมาจากประชากรที่มีลักษณะตามที่กำหนดไว้ที่ H_0 จึงปฏิเสธ H_0 ดังนั้น การตัดสินใจไม่ว่าจะปฏิเสธ H_0 หรือไม่ปฏิเสธ H_0 มีความเสี่ยงที่จะเกิดความผิดพลาดขึ้น เพราะการตัดสินใจนั้นพิจารณาจากค่าโอกาส การทดสอบนัยสำคัญจึงมิใช่เป็นการบอกกว่าสมมติฐานนั้นจริงหรือเท็จ แต่เป็นการบอกค่าโอกาสที่สมมติฐานนั้นจะเป็นจริงหรือเท็จจากค่าสถิติที่ได้จากตัวอย่าง

คำสำคัญ: นัยสำคัญ P-value สถิติทดสอบ

Abstract

This article is intended to provide researchers with a better understanding of the process of using and interpreting the results from statistical significance tests. The current test is based on the P-value, which is the probability that the null hypothesis (H_0) of the

test statistics is true. If the P-value is small, it is unlikely that sampling statistics come from the population defined by H_0 and so the null hypothesis is rejected. The decision whether to reject the null hypothesis or not is subject to error because that decision is determined from chance or probability. Hence, the significance of the test is not to say that the hypothesis is true or false, but to say that the chance or probability that the hypothesis is true or false, based on the sampling statistics.

Keywords: Significance, P-Value, Test Statistics

บทนำ

การทดสอบนัยสำคัญทางสถิติ เป็นกระบวนการหาคำตอบในศาสตร์แขนงต่าง ๆ ที่มีแบบแผน มีระเบียบวิธีอย่างเป็นขั้นเป็นตอน ส่วนใหญ่ผู้ที่ต้องการหาคำตอบโดยการทดสอบนัยสำคัญทางสถิติ มักจะทำตามขั้นตอนที่กำหนดไว้ คือ 1. ตั้งสมมติฐาน 2. กำหนดระดับนัยสำคัญ 3. เลือกตัวสถิติทดสอบและคำนวณค่า 4. เปิดตารางหาค่าจุดวิกฤติ และบอกเกณฑ์การตัดสินใจ 5. สรุปผลการตัดสินใจ โดยที่ผู้วิจัยยังไม่ค่อยเข้าใจว่า ในแต่ละขั้นตอนทำไปเพื่ออะไร ผลที่ได้มีความหมายอย่างไร เพียงแต่ทำตามขั้นตอนเหล่านี้เท่านั้น ซึ่งอาจจะยังผลให้การสรุปผลในขั้นตอนสุดท้าย หรือการอภิปรายผลการวิจัยนั้นไม่ชัดเจน หรือมีโอกาสผิดพลาดได้ จะส่งผลเสียหายต่องานวิจัยที่อุตสาหกรรมตั้งใจ ทุ่มเท เสียเวลาทำมา หรืออาจจะทำให้ไม่สามารถนำงานวิจัยนั้นไปใช้ประโยชน์ได้เต็มที่ และปัจจุบันการประมวลผลมักจะอาศัยโปรแกรมสำเร็จรูปทางสถิติมาช่วยผู้วิจัยเพียงแค่งำหนดสมมติฐาน กำหนดระดับนัยสำคัญ และสรุปผลการทดสอบ โดยพิจารณาจากค่า P-value ที่โปรแกรมคำนวณให้เท่านั้น โดยไม่จำเป็นต้องพิจารณาจากค่าสถิติทดสอบ และไม่ต้องเปิดตารางหาค่าจุดวิกฤติ ผู้วิจัยเองก็อาจจะยังไม่ทราบค่า P-value คือค่าอะไร บทความนี้จะช่วย

ให้ผู้วิจัยได้เข้าใจในบางประเด็นของการทดสอบนัยสำคัญทางสถิติได้ดียิ่งขึ้น เช่น ทราบว่าค่า P-value คือ ค่าอะไร เกี่ยวข้องกับค่าสถิติทดสอบอย่างไร ช่วยให้ทราบว่าการทดสอบนัยสำคัญทางสถิติ มีใช้วิธีพิสูจน์ว่าสมมติฐานนั้นเป็นจริงหรือเป็นเท็จ (บัณฑิต ถิ่นคำพร, 2553) แต่เป็นเพียงวิธีการที่จะบ่งบอกถึงโอกาสที่สมมติฐานนั้นจะเป็นจริงหรือเป็นเท็จจากค่าสถิติที่ได้จากตัวอย่าง และการตัดสินใจทุกครั้งมีความเสี่ยงที่จะเกิดความผิดพลาดได้เสมอ

การทดสอบนัยสำคัญทางสถิติ

การทดสอบนัยสำคัญทางสถิติเป็นการทดสอบเกี่ยวกับลักษณะของประชากรหรือค่าพารามิเตอร์ โดยใช้ข้อมูลจากตัวอย่าง ในการทดสอบผู้วิจัยต้องกำหนดสมมติฐานของการทดสอบ สมมติฐานของการทดสอบจะมี 2 ชนิด คือ สมมติฐานหลัก (Null Hypothesis: H_0) และสมมติฐานทางเลือก (Alternative Hypothesis: H_a) ข้อความในสมมติฐานทั้งสองเป็นข้อความที่แย้งกันอยู่ โดยสมมติฐานที่ถูกทดสอบ คือ สมมติฐานหลัก (H_0) การทดสอบนัยสำคัญจะเป็นการพิจารณาค่าความน่าจะเป็นหรือโอกาสที่จะได้ผลลัพธ์ตามค่าสถิติจากตัวอย่าง (ค่าสถิติทดสอบคือ ค่าสถิติจากตัวอย่างนั่นเอง) ภายใต้เงื่อนไขว่าสมมติฐานหลัก (H_0) เป็นจริง

ค่า P-value

ค่าความน่าจะเป็นที่คำนวณมาจากเงื่อนไขของสมมติฐานหลัก (H_0) เป็นจริงนี้ เรียกว่า ค่า P-value ถ้า P-value มีค่ามาก จะสรุปว่าการทดสอบนั้นไม่มีนัยสำคัญ จึงไม่ปฏิเสธ H_0 แต่ถ้า P-value มีค่าน้อย จะสรุปได้ว่าการทดสอบนั้นมีนัยสำคัญ หมายถึงค่าสถิติจากตัวอย่าง หรือค่าสถิติทดสอบนี้แสดงให้เห็นว่ามีโอกาสน้อยที่จะเป็นไปตามเงื่อนไขของ H_0 ดังนั้น ผู้วิจัยจึงปฏิเสธ H_0 เกณฑ์การตัดสินใจว่า P-value มีค่ามากหรือน้อย พิจารณาที่ค่าระดับนัยสำคัญ (Level of Significance) ใช้สัญลักษณ์แทนด้วย α ระดับนัยสำคัญ (α) ที่นิยมกำหนดคือที่ระดับ 0.05 0.01 และ 0.1 ถ้า P-value $> \alpha$ ถือว่ามีค่ามาก จะสรุปว่าการทดสอบไม่มีนัยสำคัญ ถ้า P-value $\leq \alpha$ ถือว่ามีค่าน้อย จึงสรุปว่าการทดสอบมีนัยสำคัญ

ค่า P-value และค่าสถิติทดสอบ

ค่า P-value คือ ค่าความน่าจะเป็นที่จะได้ค่าสถิติทดสอบภายใต้เงื่อนไขว่า H_0 เป็นจริง แต่เดิมในการทดสอบนัยสำคัญเราต้องพิจารณาว่าค่าสถิติทดสอบนั้นตกอยู่ในบริเวณวิกฤติหรือไม่ ถ้าตกอยู่ในบริเวณวิกฤติก็จะปฏิเสธ H_0 ดังนั้น เราจึงต้องเปิดตารางค่าจุดวิกฤติ เพื่อจะได้ทราบบริเวณวิกฤติ แต่ปัจจุบันส่วนใหญ่เราจะอาศัยโปรแกรมสำเร็จรูปคำนวณค่า P-value ซึ่งเป็นค่าความน่าจะเป็นของค่าสถิติทดสอบภายใต้เงื่อนไขว่า H_0 เป็นจริง จึงสามารถนำค่า P-value ไปเปรียบเทียบกับค่าระดับนัยสำคัญ (α) ซึ่งเป็นค่าความน่าจะเป็นของบริเวณวิกฤติได้เลย ถ้าค่า P-value $\leq \alpha$ แสดงว่าค่าสถิติทดสอบตกอยู่ในบริเวณวิกฤติ เราจึงปฏิเสธ H_0

การสรุปผลการทดสอบนัยสำคัญโดยพิจารณาจากค่า P-value

การทดสอบนัยสำคัญ มิใช่การพิสูจน์สมมติฐานว่าเป็นจริงหรือเป็นเท็จ เนื่องจากทางสถิติไม่สามารถทำได้ เพราะถ้าจะพิสูจน์ได้ต้องอาศัยข้อมูลจากประชากรทั้งหมด แต่การทดสอบนัยสำคัญ คือ การพิจารณาว่าถ้า H_0 เป็นจริงแล้ว โอกาสที่ค่าสถิติจากตัวอย่างจะมีผลลัพธ์ต่าง ๆ มีโอกาสมากน้อยแค่ไหน (เนื่องจากค่าสถิติจากตัวอย่างมีการกระจายโดยสุ่มหรือที่เราเรียกว่าการแจกแจงของค่าสถิติตัวอย่าง) ดังนั้น เมื่อค่าสถิติจากตัวอย่างที่เราเลือกมามีผลลัพธ์เป็นเช่นนั้นแล้ว มีโอกาสมากน้อยแค่ไหนที่จะเป็นเช่นนั้น ถ้าโอกาสน้อยผู้วิจัยก็จะปฏิเสธ H_0 เพราะนั่นหมายถึงความเสี่ยงที่ H_0 จะไม่เป็นจริงมีมาก แต่ในทางตรงกันข้าม ภายใต้เงื่อนไขว่า H_0 เป็นจริงแล้ว ถ้าโอกาสที่ค่าสถิติจากตัวอย่างจะมีผลลัพธ์เป็นเช่นนั้นมีสูง ผู้วิจัยก็คงไม่สามารถปฏิเสธ H_0 ตัวอย่างเช่น ถ้าสมมติประชากรของการศึกษา คือ ผู้ติดตามชมละครโทรทัศน์เป็นประจำ และถ้าผู้วิจัยตั้ง H_0 ว่าอายุเฉลี่ยของผู้ติดตามชมละครโทรทัศน์เป็นประจำเท่ากับ 35 ปี แต่จากการสุ่มตัวอย่างผู้ติดตามชมละครโทรทัศน์เป็นประจำได้ค่าอายุเฉลี่ยเท่ากับ 45 ปี ถ้าสมมติว่า H_0 เป็นจริง คือ อายุเฉลี่ยของผู้ติดตามชมละครโทรทัศน์เป็นประจำเท่ากับ 35 ปี จริง ถ้ามามีโอกาสเป็นไปได้หรือไม่ที่ผลการศึกษากจะเป็นเช่นนั้น คำตอบคือมีโอกาสที่จะเป็นไปได้ เพราะค่าสถิติจากตัวอย่างมีการกระจายอย่างสุ่ม แต่จะมีโอกาสมากน้อยแค่ไหนต้องทำการคำนวณค่า P-value ถ้าสมมติว่าได้ค่า P-value เท่ากับ 0.012 หมายความว่า ถ้า H_0 เป็นจริง คือ ผู้ติดตามชมละครโทรทัศน์เป็นประจำมีอายุเฉลี่ยเท่ากับ 35 ปี โอกาสที่อายุเฉลี่ยจากตัวอย่างสุ่มจะมีค่าเป็น 45 ปี เท่ากับ

0.012 การจะสรุปว่าค่า P-value ดังกล่าวมีค่ามากหรือน้อยก็เทียบกับค่าระดับนัยสำคัญ (α) สมมติว่ากรณีนี้กำหนดระดับนัยสำคัญที่ 0.05 จะเห็นว่าค่า P-value < α สรุปได้ว่าผลการศึกษามากกว่ามีโอกาสน้อยที่ตัวอย่างนี้จะมาจากประชากรที่มีอายุเฉลี่ยเท่ากับ 35 ปี

ความเสี่ยงของการเกิดความผิดพลาดจากการตัดสินใจ

การสรุปผลการทดสอบว่าจะปฏิเสธ H_0 หรือไม่ปฏิเสธ H_0 นั้น พิจารณาจากค่าความน่าจะเป็นของค่าสถิติทดสอบ (ค่า P-value) ดังนั้น การตัดสินใจไม่ว่าจะปฏิเสธหรือไม่ปฏิเสธ H_0 ผู้วิจัยมีโอกาสที่จะตัดสินใจผิดพลาดได้ นั่นคือ ผู้วิจัยอาจปฏิเสธ H_0 ทั้งๆ ที่ H_0 นั้นถูกต้อง ความผิดพลาดดังกล่าวเรียกว่า ความผิดพลาดประเภทที่ 1 (Type I Error) จากตัวอย่างที่ยกมาข้างต้น จะเห็นว่าถึงแม้จากค่า P-value จะบอกว่า มีโอกาสน้อยที่ผลจากตัวอย่างสุ่มดังกล่าวจะมาจากประชากรที่มีอายุเฉลี่ยเท่ากับ 35 ปี แต่ก็ยังมีโอกาสอยู่ ดังนั้น ถ้าสมมติในความเป็นจริงคือ อายุเฉลี่ยของประชากรเท่ากับ 35 ปีจริง แต่ค่า P-value เท่ากับ 0.012 ซึ่งน้อยกว่าระดับนัยสำคัญ 0.05 ผู้วิจัยจึงสรุปว่าผลการทดสอบมีนัยสำคัญและปฏิเสธ H_0 ที่กล่าวว่า อายุเฉลี่ยของผู้ติดตามชมละครโทรทัศน์เป็นประจำเท่ากับ 35 ปี ดังนั้น การตัดสินใจดังกล่าวจึงเกิดความผิดพลาด ความเสี่ยงที่จะตัดสินใจผิดพลาดจะเท่ากับค่าระดับนัยสำคัญ (α) ที่กำหนดนั่นเอง เนื่องจากที่เมื่อใดที่ค่า P-value น้อยกว่าระดับนัยสำคัญ ก็จะปฏิเสธ H_0 ทั้ง ๆ ที่ H_0 นั้นอาจจะเป็นจริง

ความเสี่ยงของการเกิดความผิดพลาดยังมีอีกประเภท คือ ความผิดพลาดประเภทที่ 2 (Type II

Error) ซึ่งเป็นความผิดพลาดที่เกิดจากการที่ผู้วิจัยไม่ปฏิเสธ H_0 ทั้ง ๆ ที่ H_0 นั้นไม่ถูกต้อง ค่าความเสี่ยงของการเกิดความผิดพลาดประเภทที่ 2 ใช้สัญลักษณ์แทนด้วย β ในการทดสอบนัยสำคัญค่า β ไม่สามารถกำหนดได้โดยตรง แต่เนื่องจากค่า α และ β มีความสัมพันธ์ในทิศทางที่ตรงข้ามกัน นั่นคือ ถ้า α มีค่ามาก β จะมีค่าน้อย ดังนั้น การกำหนดค่า β ให้มีค่าน้อยหรือมาก จึงทำได้โดยการกำหนดผ่านทางค่า α อย่างไรก็ตาม เราสามารถลดความเสี่ยงของการเกิดความผิดพลาดทั้งสองชนิดได้โดยการเพิ่มขนาดตัวอย่าง

ผู้วิจัยอาจมีปัญหาว่าควรกำหนดค่าความเสี่ยงทั้ง 2 ประเภทอย่างไร โดยทั่วไปจะพิจารณาจากผลกระทบที่จะได้รับการเกิดความผิดพลาดทั้ง 2 ประเภท ถ้าความเสี่ยงของการเกิดความผิดพลาดประเภทที่ 1 จะส่งผลกระทบต่อผู้ใช้ผลการวิจัยที่รุนแรงกว่า ก็ควรที่จะกำหนดให้ α มีค่าน้อย ถ้าความเสี่ยงของการเกิดความผิดพลาดประเภทที่ 2 จะส่งผลกระทบต่อผู้ใช้ผลการวิจัยที่รุนแรงกว่า ก็ควรที่จะกำหนดให้ α มีค่ามาก ดังตัวอย่างเช่น การวิจัยที่กำหนดว่า

H_0 : วิธีการรักษาแบบใหม่ไม่ได้ให้ผลการรักษาที่ดีกว่าวิธีแบบเก่า

H_a : วิธีการรักษาแบบใหม่ให้ผลการรักษาที่ดีกว่าวิธีแบบเก่า

สมมติว่าในความเป็นจริง คือ H_0 ถูกต้อง แต่ผลการศึกษาจากตัวอย่างได้ค่า P-value มีค่าน้อย (ซึ่งมีโอกาสเกิดขึ้นเท่ากับค่าระดับนัยสำคัญที่กำหนด) ผู้วิจัยจึงปฏิเสธ H_0 ไปสรุปว่าวิธีการรักษาแบบใหม่ให้ผลการรักษาที่ดีกว่าวิธีแบบเก่า การตัดสินใจดังกล่าวจะส่งผลต่อการรักษาผู้ป่วย แทนที่ผู้ป่วยจะได้รับการรักษาด้วยวิธีเก่า ซึ่งมีการใช้รักษามานานและเป็น

ที่ยอมรับถึงผลการรักษา แต่อาจจะมีต้นทุนที่สูงกว่า ถ้าผลการทดสอบออกมาดังกล่าวอาจมีการตัดสินใจเลือกใช้วิธีใหม่ซึ่งมีต้นทุนต่ำกว่า แต่ผลการรักษาไม่ดีเท่าวิธีเดิม ดังนั้น ผลกระทบที่เกิดขึ้นอาจส่งผลต่อชีวิตของผู้ป่วย ดังนั้น งานวิจัยลักษณะดังกล่าวจึงควรกำหนดให้โอกาสที่จะปฏิเสธ H_0 ทั้ง ๆ ที่ H_0 เป็นจริง หรือความเสี่ยงที่จะเกิดความผิดพลาดประเภทที่ 1 (α) มีค่าน้อย หรืออย่างเช่นงานวิจัยที่กำหนด

H_0 : ยาตัวใหม่ที่คิดค้นขึ้นไม่มีอันตรายต่อผู้ป่วย

H_a : ยาตัวใหม่ที่คิดค้นขึ้นมีอันตรายต่อผู้ป่วย

ถ้าสมมติว่าในความเป็นจริง คือ ยาตัวใหม่นั้น ถึงแม้จะมีประสิทธิภาพต่อผลการรักษาที่ดีขึ้นแต่จะมีอันตรายต่อผู้ป่วยมากกว่า แต่ผลการทดสอบ คือ ค่า P-value มีค่าสูง ผู้วิจัยจึงไม่ปฏิเสธ H_0 สรุปว่า ยาตัวดังกล่าวไม่มีอันตรายต่อผู้ป่วย การตัดสินใจดังกล่าวจะส่งผลต่อชีวิตผู้ป่วย ดังนั้น งานวิจัยลักษณะดังกล่าวจึงควรกำหนดให้ค่าความเสี่ยงที่จะเกิดความผิดพลาดประเภทที่ 2 (β) มีค่าน้อย

การทดสอบนัยสำคัญเรื่องทั่ว ๆ ไป มักจะกำหนดระดับนัยสำคัญที่ระดับ 0.05 แต่ผู้วิจัยอาจกำหนดให้มีค่าสูงหรือต่ำกว่านี้ได้ แล้วแต่ความเหมาะสมดังที่กล่าว เช่น กำหนดเป็น 0.1 ถ้าการเกิดความผิดพลาดประเภทที่ 2 จะส่งผลกระทบต่อผู้ใช้ผลการวิจัยที่รุนแรงกว่า หรือกำหนดเป็น 0.01 หรือ 0.005 ถ้าการเกิดความผิดพลาดประเภทที่ 1 จะส่งผลกระทบต่อผู้ใช้ผลการวิจัยที่รุนแรงกว่า

การสรุปผลการทดสอบนัยสำคัญทางสถิติ

ตามที่กล่าวไว้ตั้งแต่ต้นว่าการทดสอบนัยสำคัญ สมมติฐานของการทดสอบทางสถิติจะมี 2 ชนิด คือ สมมติฐานหลัก (Null Hypothesis: H_0) และสมมติฐานทางเลือก (Alternative Hypothesis: H_a)

แต่สมมติฐานที่ถูกทดสอบ คือ H_0 ข้อความที่กำหนดที่ H_0 เป็นข้อความที่แสดงถึงความไม่แตกต่างกัน (=) หรืออาจเป็นข้อความที่แสดงถึงความแตกต่างก็ได้ แต่ต้องมีเครื่องหมายแสดงความไม่แตกต่างอยู่ด้วย (\leq หรือ \geq) หรือเป็นข้อความที่แสดงถึงความไม่มีความสัมพันธ์กัน เช่น

H_0 : อายุเฉลี่ยของผู้ติดตามชมละครโทรทัศน์เป็นประจำไม่ต่ำกว่า 35 ปี

$$(\mu \geq 35)$$

หรือ

H_0 : อายุเฉลี่ยของผู้ติดตามชมละครโทรทัศน์ช่อง 3 เป็นประจำ และผู้ติดตามชมละครโทรทัศน์ช่อง 7 เป็นประจำไม่แตกต่างกัน

$$(\mu_1 = \mu_2)$$

หรือ

H_0 : ระดับรายได้ไม่มีความสัมพันธ์กับการตัดสินใจซื้อโทรศัพท์สมาร์ทโฟน

ในการทดสอบนัยสำคัญทางสถิติ สมมติฐานที่ถูกทดสอบ คือ H_0 ดังนั้น การตัดสินใจจึงมีอยู่ 2 ประการ คือ ปฏิเสธ H_0 หรือไม่ปฏิเสธ H_0 การสรุปผลการทดสอบกรณีไม่ปฏิเสธ H_0 ผู้วิจัยอาจสรุปว่าไม่มีความแตกต่างกัน หรือ ไม่มีความแตกต่างกันทางสถิติ (ถ้าเป็นความแตกต่างกันทางคณิตศาสตร์ จะเป็นเช่นนี้ คือ สมมติ $\bar{X}_1 = 12$ และ $\bar{X}_2 = 14$ ลักษณะนี้ \bar{X}_1 และ \bar{X}_2 จะแตกต่างกันทางคณิตศาสตร์ แต่อาจไม่แตกต่างกันทางสถิติ) กรณีทดสอบความสัมพันธ์ก็สรุปว่า ไม่มีความสัมพันธ์กันตามข้อความที่กำหนดที่ H_0

ถ้าผลการทดสอบมีนัยสำคัญจะปฏิเสธ H_0 ถ้าปฏิเสธ H_0 จะเท่ากับยอมรับ H_a โดยปริยาย การสรุปใช้ข้อความที่กำหนดไว้ที่ H_a แล้วต่อด้วยคำว่า อย่างมีนัยสำคัญที่ระดับนัยสำคัญที่กำหนด เพื่อ

จะบอกให้ทราบว่า การตัดสินใจปฏิเสธ H_0 ครั้งนี้มีโอกาสที่จะเกิดความผิดพลาดประเภทที่ 1 เท่ากับค่าระดับนัยสำคัญที่กำหนด เช่น สรุปว่าแตกต่างกันอย่างมีนัยสำคัญที่ $\alpha = 0.05$ หรือมีความสัมพันธ์กันที่ระดับนัยสำคัญ 0.05

ในการทดสอบนัยสำคัญผู้วิจัยมักคาดหวังให้ผลสรุปออกมาว่ามีนัยสำคัญ คือ ต้องการให้มีการปฏิเสธ H_0 เพื่อยอมรับ H_a แต่การจะปฏิเสธ H_0 หรือไม่ปฏิเสธ H_0 นั้น ขึ้นอยู่กับค่าสถิติจากตัวอย่าง ดังนั้น เมื่อผลการทดสอบไม่มีนัยสำคัญ ผู้วิจัยบางท่านอาจจะสรุปว่าจากหลักฐานที่มีอยู่ (ข้อมูลตัวอย่าง) ไม่เพียงพอที่จะสรุปว่าค่าเฉลี่ยแตกต่างกันอย่างมีนัยสำคัญ หรือไม่เพียงพอที่จะสรุปว่าตัวแปรมีความสัมพันธ์กันอย่างมีนัยสำคัญ

การที่ผู้วิจัยมักคาดหวังให้ผลสรุปออกมาว่ามีนัยสำคัญ คือ ต้องการให้มีการปฏิเสธ H_0 เพื่อยอมรับ H_a เนื่องจากโดยทั่วไปการทดสอบนัยสำคัญนั้น ผู้วิจัยต้องการยืนยันถึงความแตกต่าง หรือยืนยันถึงความสัมพันธ์กันระหว่างตัวแปร เพื่อจะสามารถนำผลสรุปไปใช้ประโยชน์ในการแก้ปัญหา หรือการตัดสินใจได้ การทดสอบเพื่อยืนยันว่า ไม่แตกต่างกันหรือไม่มีความสัมพันธ์กัน จะนำไปใช้ประโยชน์ในทางปฏิบัติไม่ได้ การกำหนดระดับนัยสำคัญ (α) ก็เป็นอีกประเด็นหนึ่งที่สำคัญที่จะทำให้ผลสรุปออกมาเป็นเช่นไร บางครั้งผู้วิจัยต้องการให้งานวิจัยมีความน่าเชื่อถือ โดยกำหนดระดับความเชื่อมั่นสูง คือ ตั้งระดับนัยสำคัญ (α) ให้มีค่าต่ำ ๆ การตั้ง α ให้มีค่าต่ำจะทำให้โอกาสปฏิเสธ H_0 เกิดได้ยาก ถึงแม้ค่าสถิติจากตัวอย่างอาจจะแตกต่างจากค่าที่กำหนดไว้ที่ H_0 มาก ก็ยังไม่สามารถปฏิเสธ H_0 ได้ ซึ่งไม่เป็นผลดีสำหรับผู้วิจัยที่ต้องการนำผลการทดสอบที่ยืนยันความแตกต่างไปใช้ประโยชน์ต่อไป

บทสรุป

การทดสอบนัยสำคัญไม่สามารถบอกได้ว่าสมมติฐานนั้นเป็นจริงหรือเป็นเท็จ แต่เป็นเพียงแค่บอกโอกาสที่สมมติฐานนั้นจะเป็นจริงหรือเป็นเท็จว่ามีมากน้อยแค่ไหน โดยพิจารณาจากค่า P-value ซึ่งเป็นค่าความน่าจะเป็นของค่าสถิติจากตัวอย่างภายใต้เงื่อนไขว่า H_0 เป็นจริง เมื่อเป็นเรื่องของโอกาส การตัดสินใจไม่ว่าจะปฏิเสธ H_0 หรือไม่ปฏิเสธ H_0 จึงมีความเสี่ยงที่ผู้วิจัยจะสรุปผิด ดังนั้น ในการทดสอบจึงควรต้องใช้ระเบียบวิธีวิจัยที่มีความเหมาะสมทุกขั้นตอน เพื่อจะได้ตัวอย่างที่เป็นตัวแทนที่ดีของประชากร จะช่วยลดความเสี่ยงของการตัดสินใจผิดพลาดได้ และการกำหนดระดับนัยสำคัญจะต้องคำนึงถึงผลกระทบที่อาจเกิดขึ้น และคำนึงถึงข้อเท็จจริงจากหลักฐานที่ได้มาจากตัวอย่างด้วยเป็นสำคัญมากกว่าความเชื่อส่วนตัวของผู้วิจัยเอง

บรรณานุกรม

- Krupee. 2009. **Significance Level 0.05 or 0.01 ... Which one !?** [ออนไลน์]. Available: <http://krupee.blogspot.com/2009/11/05-01.html> (in Thai).
- ครูพี. 2552. **ตั้งระดับนัยสำคัญ 0.05 หรือ 0.01... เท่าไหร่ดี !?** [ออนไลน์]. เข้าถึงจาก: <http://krupee.blogspot.com/2009/11/05-01.html>
- Mendenhall, William., Beaver, Robert J., and Beaver, Barbara M. 1999. **Introduction to Probability and Statistics**. 10th ed. Pacific Grove, CA: Duxbury Press.
- Ott, R. Lyman., and Longnecker, Michael. 1999. **An Introduction to Statistical Methods and Data Analysis**. 5th ed.

- Pacific Grove, CA: Duxbury Press.
- Piyapimolsith, Chatsiri. 2001. **Type I Error, Type II Error and Power of the Test** [Online]. Available: <http://www.watpon.com/Elearning/stat19.htm> (in Thai).
- ฉัตรศิริ ปิยะพิมลสิทธิ์. 2544. **Type I Error, Type II Error และอำนาจการทดสอบ** [ออนไลน์]. เข้าถึงจาก: <http://www.watpon.com/Elearning/stat19.htm>
- Sookmuang, Anusit. 2006. **Risks in Decision Making Using Hypothesis Testing** [Online]. Available: http://www.dss.go.th/dssweb/st-articles/files/lpd_5_2549_error-risk.pdf (in Thai).
- อนุสิทธิ์ สุขม่วง. 2549. **ความเสี่ยงในการตัดสินใจ โดยการทดสอบสมมติฐาน** [ออนไลน์]. เข้าถึงจาก: http://www.dss.go.th/dssweb/st-articles/les/lpd_5_2549_error-risk.pdf
- Srisaard, Boonchom. 2012. **Solving Problems of Statistics, Measurement and Research** [Online]. Available: http://www.kroobannok.com/news__file/p90458671037.pdf (in Thai).
- บุญชม ศรีสะอาด. 2556. **ไขปัญหาเกี่ยวกับสถิติ วัตถุประสงค์ และวิจัย ตอนที่ 1** [ออนไลน์]. เข้าถึงจาก: http://www.kroobannok.com/news__file/p90458671037.pdf
- Thailand. Coastal Fisheries Research and Development Bureau. 2012. **Comparison of the Difference of Statistical Mean** [Online]. Available: <http://www.coastalaqua.com/webboard/index.php?topic=1174.15> (in Thai).
- สำนักวิจัยและพัฒนาประมงชายฝั่ง. 2555. **การเปรียบเทียบความแตกต่างของค่าเฉลี่ยทางสถิติ** [ออนไลน์]. เข้าถึงจาก: <http://www.coastalaqua.com/webboard/index.php?topic=1174.15>
- Thinkumrop, Bhandit. 2010. **Statistical Significance Test and Confidence Interval** [Online]. Available: http://www.dmbn.net/mdbtemplate/mytemplate/template.php?component=view__article&qid=16 (in Thai).
- บัณฑิต ถิ่นคำพร. 2553. **การทดสอบนัยสำคัญทางสถิติและช่วงความเชื่อมั่น** [ออนไลน์]. เข้าถึงจาก: http://www.dmbn.net/mdbtemplate/mytemplate/template.php?component=view__article&qid=16



Assistant Professor Boonying Somrang received her Master of Science in Applied Statistics from the National Institute of Development Administration. She is currently a lecturer in the School of Science and Technology, University of the Thai Chamber of Commerce. Her research interest focuses on time series analysis and regression analysis.