

## P Value มีค่าน้อยความแตกต่าง ระหว่างกลุ่มยิ่งมากจริงหรือ

อรุณ จิรวัดน์กุล วท.บ. (อาชีวอนามัย), วท.ม. (ชีวสถิติ), M.Sc. (Clinical Epidemiology)

การแปลผลว่า ถ้าค่า p value มีค่าน้อย ความแตกต่างระหว่างกลุ่มยิ่งมีต่างกันมาก เป็นการแปลผล p value ที่ไม่ถูกต้องซึ่งเคยเขียนลงบทความมุมมอง โดยได้อธิบายว่าค่า p value จะเล็กหรือใหญ่สัมพันธ์โดยตรงกับทั้งความต่างและขนาดตัวอย่าง ขนาดความต่างที่เท่ากัน ถ้าเพิ่มขนาดตัวอย่างก็ทำให้ p value มีค่าเล็กลงได้ จึงไม่สามารถระบุขนาดความต่างระหว่างกลุ่มด้วยค่า p value

ในการทำความเข้าใจความผิดพลาดในประเด็นนี้ถ้า นักวิจัยเข้าใจหลักการทดสอบสมมติฐาน จะช่วยให้เข้าใจว่าทำไมขนาดของค่า p value จึงไม่สามารถนำมาสรุปขนาดความแตกต่างระหว่างกลุ่มได้

ในการทดสอบสมมติฐานเปรียบเทียบค่าเฉลี่ย 2 กลุ่ม สูตรที่ใช้ในการทดสอบคือ

$$t = (\bar{X}_1 - \bar{X}_2) / \sqrt{s_p^2 / n_1 + s_p^2 / n_2}$$

จากสูตรข้างบนขนาดของค่าสถิติ t จะขึ้นอยู่กับขนาดความต่างของค่าเฉลี่ยระหว่างกลุ่ม  $(\bar{X}_1 - \bar{X}_2)$ หารด้วย standard error  $(\sqrt{s_p^2 / n_1 + s_p^2 / n_2})$  ในประชากรศึกษา ส่วนเบี่ยงเบนมาตรฐานจะมีค่าคงที่ทำให้ค่า  $s_p^2$  ที่คำนวณจากค่าส่วนเบี่ยงเบนมาตรฐานทั้งสองกลุ่มจะมีค่าคงที่ด้วย ดังนั้น ขนาดของ standard error จะขึ้นอยู่กับขนาดตัวอย่าง  $n_1$  และ  $n_2$

ถ้าตัวอย่างที่ศึกษามีขนาดใหญ่ standard error จะมีขนาดเล็ก เมื่อตัวหามีขนาดเล็กค่า สถิติ t จะมีขนาดใหญ่

เมื่อนำค่า t ไปแปลงเป็นค่า p value จะได้ p value ที่มีขนาดเล็ก

ในทางกลับกันถ้าตัวอย่างที่นำมาศึกษามีขนาดเล็ก standard error ที่ได้จะมีขนาดใหญ่ เมื่อตัวหามีขนาดใหญ่ ค่าสถิติ t จะมีขนาดเล็ก เมื่อนำค่า t ไปแปลงเป็นค่า p value จะได้ p value ที่มีขนาดใหญ่

ตัวอย่างในการทดสอบสมมติฐานเพื่อเปรียบเทียบค่าเฉลี่ยสองประชากร จากตัวอย่างที่ศึกษามีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานดังนี้  $\bar{X}_1 = 23, s_1 = 3.7, n_1 = 27, \bar{X}_2 = 25.2, s_2 = 4.2$  และ  $n_2 = 27$  ข้อมูลจากตัวอย่าง นำมาคำนวณค่า  $(\bar{X}_1 - \bar{X}_2)$  ได้เท่ากับ 2.2 และ  $s_p^2$  ได้เท่ากับ 15.67 คำนวณค่า standard error ได้ 1.077 ค่า  $t = 2.04$  และ p value 0.023 ถ้านำข้อมูลจากตัวอย่างที่ศึกษามาคำนวณค่า p value เมื่อเพิ่มหรือลดขนาดตัวอย่างได้ผลดังแสดงในตารางที่ 1

จากตารางที่ 1 ในกรณีศึกษาในประชากรเดิมที่มี

ตารางที่ 1 การเปลี่ยนแปลงค่า p value เมื่อมีการเปลี่ยนแปลงขนาดตัวอย่าง

n	se	t	p value
20	1.251	1.76	0.043
27	1.077	2.04	0.023
35	0.946	2.33	0.011
40	0.885	2.49	0.007

ความต่างระหว่างค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานเท่าเดิม จะพบว่าเมื่อเพิ่มขนาดตัวอย่างให้ใหญ่ขึ้นจะทำให้ค่า p value มีขนาดเล็กลง

ความแตกต่างของค่าเฉลี่ยเท่าเดิมเมื่อเพิ่มขนาดตัวอย่างในการศึกษาทำให้ standard error มีขนาดเล็กลง (ภาพที่ 1) ลักษณะของ sampling distribution of sample mean จะมีการกระจายน้อยลง ทำให้ส่วนที่ทับซ้อนกันของการแจกแจงทั้งสองน้อยลง มีผลทำให้ขนาดของ p value เล็กลง

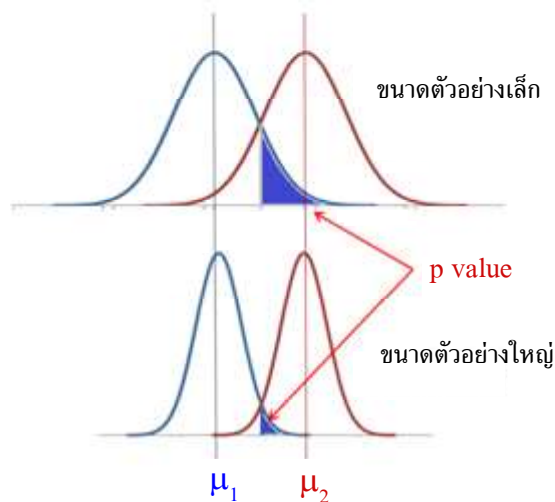
จากตัวเลขการคำนวณในตารางที่ 1 และลักษณะการแจกแจงในภาพที่ 1 สรุปได้ว่าขนาดความต่างระหว่างกลุ่มเท่าเดิม ถ้ามีการเพิ่มขนาดตัวอย่างจะมีผลทำให้ค่าของ p value เล็กลง ดังนั้นจึงไม่สามารถนำค่า p value ไปสรุปขนาดความต่างระหว่างกลุ่มได้

การทดสอบสมมุติฐานที่ได้ค่า p value เล็กมากเท่าไรยิ่งจะช่วยเพิ่มความมั่นใจว่าโอกาสที่เกิดความผิดพลาดจากการปฏิเสธสมมุติฐาน จะมีค่าน้อย

ในกรณีผลการทดสอบพบว่าต่างกันอย่างไรไม่มีนัยสำคัญ ถ้ามีการเพิ่มขนาดตัวอย่างให้มากพอ จะสามารถเปลี่ยนผลการทดสอบเป็นพบความแตกต่างกันอย่างมีนัยสำคัญได้

ดังนั้นในการพิจารณาสรุปผลวิจัยจะใช้ความต่างอย่างมีนัยสำคัญอย่างเดียวไม่ได้ เมื่อพบความต่างอย่างมีนัยสำคัญแล้ว จะต้องพิจารณาต่อว่า ขนาดความต่างที่พบนั้นว่ามีความสำคัญต่อนำไปใช้งาน (clinical importance หรือ public health importance) หรือไม่ สถิติที่ใช้พิจารณา

ภาพที่ 1 แสดง sampling distribution of sample mean เมื่อขนาดตัวอย่างใหญ่ขึ้น



ขนาดความต่างที่ควรนำมาพิจารณาร่วมกับค่า p value ได้แก่ 95% ช่วงเชื่อมั่น Risk reduction, Absolute risk reduction เป็นต้น

### สรุป

ขนาดของ p value ไม่สามารถนำมาสรุปเป็นขนาดความต่างระหว่างกลุ่ม ขนาดของ p value จะเล็กลงตามขนาดตัวอย่างที่เพิ่มขึ้น p value ยังมีขนาดเล็กยิ่งช่วยเพิ่มความมั่นใจในการปฏิเสธสมมุติฐาน เมื่อผลการทดสอบพบว่ามีความแตกต่างกันอย่างมีนัยสำคัญ นักวิจัยต้องพิจารณาว่า ขนาดความต่างที่พบนั้นมีความสำคัญต่อการนำไปใช้งานหรือไม่