

การวัดการกระจายในงานวิจัยทางชีววิทยา : ค่าเบี่ยงเบนมาตรฐาน ความคลาดเคลื่อนมาตรฐาน และสัมประสิทธิ์การแปรผัน Measure of Dispersion in Biological Research: Standard Deviation, Standard Error and Coefficient of Variation

พิพัฒน์ สมภาร*

สาขาวิชาเทคโนโลยีการเกษตร คณะวิทยาศาสตร์และเทคโนโลยี

มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต ตำบลคลองหนึ่ง อำเภอคลองหลวง จังหวัดปทุมธานี 12120

Pipat Somparn*

Department of Agricultural Technology, Faculty of Science and Technology,
Thammasat University, Rangsit Centre, Khlong Nueng, Khlong Luang, Pathum Thani 12120

บทคัดย่อ

หากนักวิจัยสังเกตและวัดค่าจากปรากฏการณ์ทางชีววิทยาซ้ำ ๆ พบว่าค่าสังเกตที่ได้จะไม่เท่ากันทุกค่า ส่วนหนึ่งมีสาเหตุมาจากความผันแปรที่เกิดขึ้นเองโดยธรรมชาติ ค่าสถิติที่ใช้วัดความแปรผันของค่าสังเกตที่เป็นตัวแปรเชิงปริมาณ คือ การวัดการกระจาย ซึ่งมีหลายวิธี และแต่ละวิธีก็มีลักษณะเฉพาะที่แตกต่างกันออกไป เมื่อต้องการรายงานข้อมูลในบทความวิจัยทางชีววิทยา ผู้เขียนบทความวิจัยมักใช้สถิติพรรณนาในการอธิบายตัวอย่างที่ศึกษา ผู้เขียนบทความวิจัยนิยมใช้ค่าเบี่ยงเบนมาตรฐาน (SD) และความคลาดเคลื่อนมาตรฐาน (SE) อธิบายความผันแปรของตัวอย่าง อย่างไรก็ตาม บ่อยครั้งที่เกิดความสับสนในการใช้ค่าทั้งสอง และนำไปสู่การตีความหมายผิดพลาด ในขณะที่ค่าเบี่ยงเบนมาตรฐานใช้วัดการกระจายของค่าสังเกต แต่ค่าความคลาดเคลื่อนมาตรฐานใช้วัดความแม่นยำของค่าเฉลี่ยตัวอย่างในการประมาณค่าเฉลี่ยของประชากร นอกเหนือจากค่าเบี่ยงเบนมาตรฐานและความคลาดเคลื่อนมาตรฐาน บทความนี้ได้อธิบายสัมประสิทธิ์การแปรผัน (CV) ไว้ด้วย

คำสำคัญ : ชีววิทยา; สถิติ; การกระจาย; ค่าเบี่ยงเบนมาตรฐาน; ความคลาดเคลื่อนมาตรฐาน

Abstract

If researchers repeatedly observe and quantify a particular biological phenomenon, the measurements will rarely be identical. Part of the variability is due to an inherent variation in biological material being measured. One characteristic of a set of observations measured on a quantitative variable is known as measures of dispersion. There are a number of measures of

dispersion of the data, each of which has different attributes. When reporting data in biological research papers, authors often use descriptive statistical methods to describe their study sample. Authors often use the standard deviation (SD) and standard error (SE) to describe the variability of their sample. However, the two are frequently confused with the consequence that the wrong measure is used to describe the variability of interest; may lead to a misinterpretation of the data. Whereas the SD is a measure of the scatter of the observations, the SE is a measure of the precision of the sample mean as an estimate of the population mean. In addition to the SD and SE, coefficient of variation (CV) is also described in this paper.

Keywords: biology; statistic; dispersion; standard deviation; standard error

การวัดตัวแปรทางชีววิทยา หน่วยทดลองแต่ละหน่วย เช่น ต้นไม้แต่ละต้นหรือสัตว์แต่ละตัว ถึงแม้ว่าจะได้รับการปฏิบัติดูแลเหมือนกัน กัน แต่ค่าที่วัดได้จะมีความแตกต่างกัน แสดงให้เห็นว่ามีความแปรผันเกิดขึ้นระหว่างค่าสังเกตที่วัดได้แต่ละค่า ดังนั้นในการนำเสนอผลการทดลอง นอกเหนือจากค่าเฉลี่ยแล้ว เพื่อแสดงให้เห็นความผันแปร (variability) ของผลการทดลอง นักวิจัยส่วนใหญ่จะนำเสนอค่าสถิติที่แสดงถึงการกระจายของข้อมูลร่วมกับค่าเฉลี่ย สามารถแสดงในรูปพิสัย (range) ความแปรปรวน (variance, S^2) ค่าเบี่ยงเบนมาตรฐาน (standard deviation, SD หรือ S) หรือความคลาดเคลื่อนมาตรฐาน (standard error, SE) ช่วงความเชื่อมั่น (confidence interval, CI) และอื่น ๆ

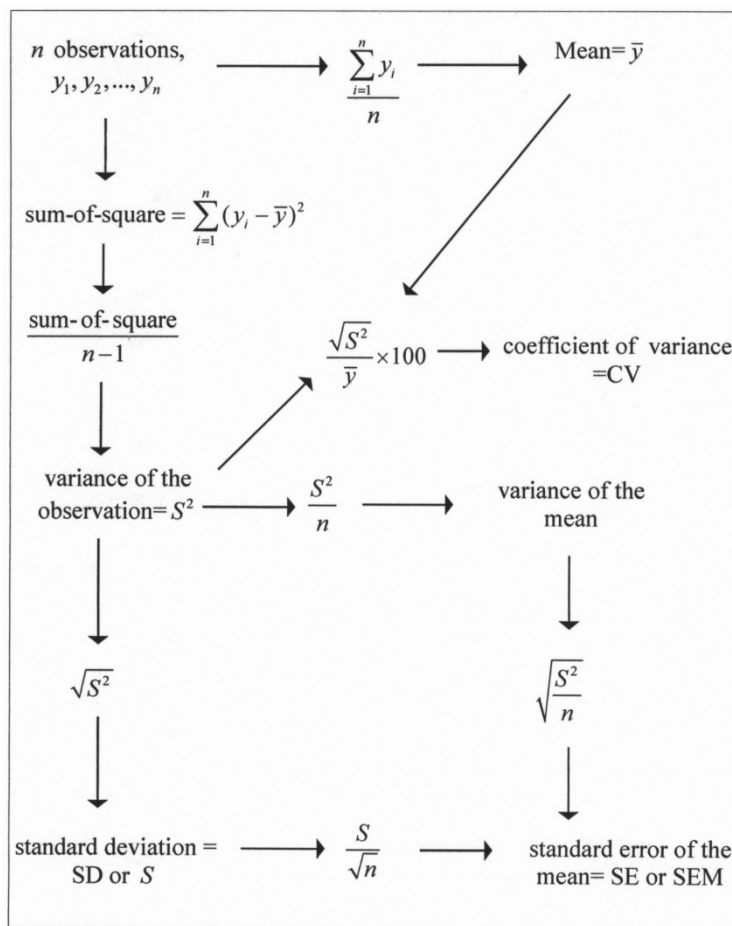
สำหรับค่าพิสัยแสดงเป็นค่าสูงสุดและต่ำสุด หรือผลต่างระหว่างค่าทั้งสอง ส่วนความแปรปรวนเป็นการนำผลรวมของผลต่างระหว่างค่าสังเกตแต่ละค่า (ซ้ำ) กับค่าเฉลี่ยมายกกำลังสอง ซึ่งค่าที่ได้เรียกว่าผลรวมกำลังสอง (sum-of-square) และหารด้วยจำนวนค่าสังเกต (ซ้ำ) ทั้งหมดหรือขนาดตัวอย่าง (sample size, n) ลบด้วยหนึ่งหรือ n-1 จากวิธีการคำนวณค่าความแปรปรวนดังกล่าวจะเห็นว่าหน่วยของความแปรปรวนจะถูกยกกำลังสอง เช่น หากหน่วยเดิม

ของค่าสังเกตเป็นกิโลกรัม หน่วยของความแปรปรวนจะเป็นกิโลกรัม² ซึ่งเป็นค่าที่ผิดธรรมชาติและยากต่อการแปลความหมาย (หากนำไปเปรียบเทียบกับค่าเฉลี่ย) ดังนั้นหากนำค่าความแปรปรวนไปถอดรากที่สอง หน่วยวัดดังกล่าวก็จะกลับไปสู่หน่วยดั้งเดิมของค่าสังเกต ค่ารากที่สองที่ได้นี้เรียกว่าค่าเบี่ยงเบนมาตรฐาน ความหมายอย่างคร่าว ๆ ของค่าเบี่ยงเบนมาตรฐานคือ ประมาณร้อยละ 68 ของข้อมูลจะตกอยู่ในช่วงค่าเฉลี่ย $\pm$ SD และประมาณร้อยละ 95 ของข้อมูลจะตกอยู่ในช่วงค่าเฉลี่ย $\pm$ 2SD ในทางทฤษฎีค่าเบี่ยงเบนมาตรฐานจะเปลี่ยนแปลงค่อนข้างน้อยถึงแม้ว่าตัวอย่างจะมีขนาดเล็กหรือใหญ่ก็ตาม ดังนั้นนักสถิติจึงใช้ค่าเบี่ยงเบนมาตรฐานเป็นค่าประมาณของความผันแปรของประชากร (σ) [1]

ส่วนความคลาดเคลื่อนมาตรฐานเป็นค่าที่บ่งบอกถึงความผันแปรของค่าเฉลี่ย อาทิ ถ้าสุ่มตัวอย่างจากประชากรเดียวกัน 10 ครั้ง (ครั้งละ n ค่าสังเกต) จะคำนวณได้ค่าเฉลี่ยตัวอย่าง (sample mean) ทั้งหมด 10 ค่า และค่าเฉลี่ยที่ได้แต่ละค่าอาจมีค่าแตกต่างกัน ดังนั้นค่าความคลาดเคลื่อนมาตรฐานจึงเป็นค่าที่ใช้วัดความผันแปรของค่าเฉลี่ยในครั้งนั้นว่าแตกต่างจากค่าเฉลี่ยจากการสุ่มครั้งอื่น ๆ (หากมีการสุ่ม) มากน้อยเพียงใด หรือกล่าวอีกนัยหนึ่งคือความ

คลาดเคลื่อนมาตรฐานเป็นการวัดการกระจายของค่าเฉลี่ยแต่ละค่านั้นเอง บางครั้งจึงเรียกว่าความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย (standard error of the mean, SEM) ในขณะที่ค่าเบี่ยงเบนมาตรฐานเป็นการวัดการกระจายของค่าสังเกตแต่ละค่า ซึ่งความคลาดเคลื่อนมาตรฐานคำนวณได้จากการนำค่า

เบี่ยงเบนมาตรฐานมาหารด้วยรากที่สองของขนาดตัวอย่าง หรืออีกวิธีหนึ่งนำเอาความแปรปรวนมาหารด้วยขนาดตัวอย่าง ต่อจากนั้นจึงนำไปถอดรากที่สองวิธีการคำนวณและความสัมพันธ์ระหว่างความแปรปรวน ค่าเบี่ยงเบนมาตรฐานและความคลาดเคลื่อนมาตรฐาน แสดงดังรูปที่ 1

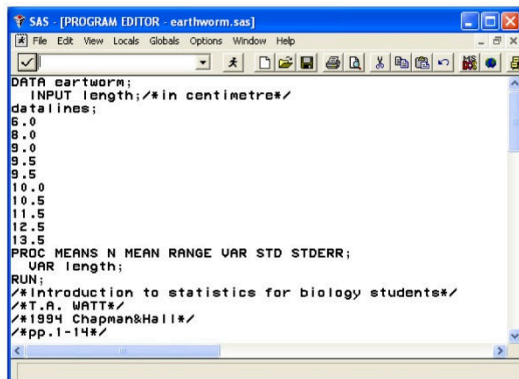


รูปที่ 1 ความสัมพันธ์ระหว่างการวัดการกระจายแบบต่าง ๆ เมื่อ y คือ ค่าสังเกตแต่ละค่า และ n คือ ขนาดตัวอย่าง

ยกตัวอย่าง สุ่มวัดความยาวไส้เดือน จำนวน 10 ตัว (หน่วยเป็นเซนติเมตร) ได้ผลดังนี้ 11.5, 10.0, 9.5, 8.0, 12.5, 13.5, 9.5, 10.5, 9.0 และ 6.0 รูปที่ 2 และ 3 แสดงชุดคำสั่งและผลลัพธ์ในการคำนวณค่าเฉลี่ยและค่าวัดการกระจาย โดยใช้ MEANS

procedure ในโปรแกรมสำเร็จรูปทางสถิติ statistical analysis system (SAS) (รูปที่ 2) จากผลลัพธ์ในรูปที่ 3 ค่าพิสัย (range) ความแปรปรวน (variance) ค่าเบี่ยงเบนมาตรฐาน (standard deviation, SD) และความคลาดเคลื่อนมาตรฐาน (standard error) เท่ากับ

7.50, 4.72, 2.17 (หรือ $\sqrt{4.72}$) และ 0.69 (หรือ $\frac{2.17}{\sqrt{10}}$ หรือ $\sqrt{\frac{4.72}{10}}$) ตามลำดับ ในการรายงานค่าความแปรปรวนนิยมให้มีจำนวนตำแหน่งของทศนิยมมากกว่าค่า (ดั้งเดิม) หนึ่งตำแหน่งเสมอ [2]



รูปที่ 2 ชุดคำสั่ง SAS ที่ใช้ในการคำนวณค่าเฉลี่ยและการกระจายของตัวอย่างความยาวไส้เดือน

ในบางครั้งสามารถแสดงค่าเบี่ยงเบนมาตรฐานเทียบเป็นร้อยละของค่าเฉลี่ย ค่าดังกล่าวเรียกว่าสัมประสิทธิ์การแปรผัน (coefficient of variation, CV) ซึ่งไม่มีหน่วยทำให้สามารถนำมาเปรียบเทียบ

ความผันแปรระหว่างการทดลองที่มีลักษณะคล้ายกันได้ แต่ไม่เป็นที่นิยมมากนัก เนื่องจากมีสมบัติทางทฤษฎีที่ค่อนข้างซับซ้อน [2] จากข้อมูลตัวอย่างข้างต้นค่าสัมประสิทธิ์การแปรผันเท่ากับ $2.1730675/10 \sim 0.2173$ หรือร้อยละ 21.73 ข้อสังเกตประการหนึ่งคือค่าสัมประสิทธิ์การแปรผันจะไม่ได้รับผลกระทบจากการคูณเพื่อแปลงมาตรวัดจากตัวอย่างข้างต้น ถ้าข้อมูลความยาวของไส้เดือนแสดงหน่วยการวัดเป็นนิ้ว (คูณทั้ง 10 ค่าด้วย 0.3937) แทนที่จะเป็นเซนติเมตร ดังนั้นค่าเฉลี่ย (3.937 นิ้ว) และค่าเบี่ยงเบนมาตรฐาน (0.8555 นิ้ว) ของตัวอย่างจะมีหน่วยเป็นนิ้วเช่นกัน แต่ค่าสัมประสิทธิ์การแปรผันของความยาวของไส้เดือนจะไม่เปลี่ยนแปลง (21.73) เนื่องจากไม่ได้รับอิทธิพลจากการแปลงมาตรวัด (นิ้วเป็นเซนติเมตร หรือเซนติเมตรเป็นนิ้ว) ดังนั้นสัมประสิทธิ์การแปรผันจึงมักถูกใช้เปรียบเทียบการกระจายของตัวแปรสองตัวหรือมากกว่าซึ่งถูกวัดด้วยหน่วยที่แตกต่างกัน อาทิ ปริมาณน้ำนมที่มีมาตรวัดเป็นกิโลกรัมกับปอนด์ หรือเปรียบเทียบลักษณะที่ต่างกัน อาทิ ส่วนสูง (เซนติเมตร) กับน้ำหนัก (กิโลกรัม)

N	Mean	Range	Variance	Std Dev	Std Error
10	10.0000000	7.5000000	4.7222222	2.1730675	0.6871843

รูปที่ 3 ผลลัพธ์ที่ได้จากการคำนวณค่าเฉลี่ยและการกระจายของตัวอย่างความยาวไส้เดือน

ค่าสัมประสิทธิ์การแปรผันเป็นคุณลักษณะเฉพาะของตัวแปรนั้นหรือมีค่าค่อนข้างคงที่ เช่น อัตราการเจริญเติบโตของสัตว์เลี้ยงลูกด้วยนมทุกชนิดมีแนวโน้มค่าสัมประสิทธิ์การแปรผันเท่ากับร้อยละ 12

ลักษณะทางการสืบพันธุ์ เช่น ขนาดครอก (litter size) และอัตราความสมบูรณ์พันธุ์มีความผันแปรมาก โดยมีค่าสัมประสิทธิ์การแปรผันอยู่ในช่วงร้อยละ 20-40 ผลผลิตน้ำนมและผลผลิตไข่ไก่ที่เก็บข้อมูลหลาย ๆ

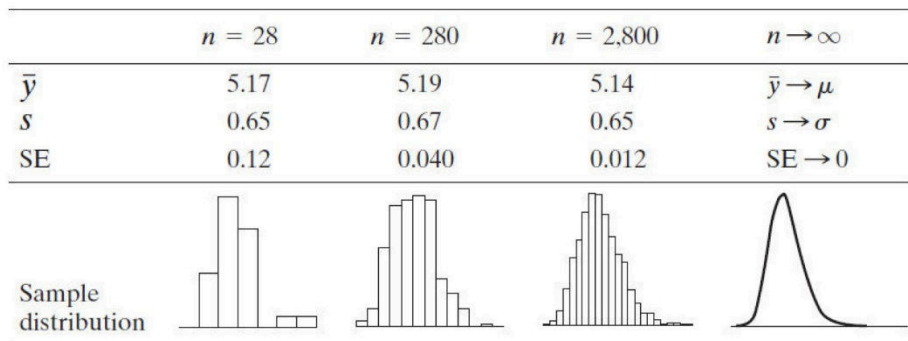
สัปดาห์ มีค่าสัมประสิทธิ์การแปรผันประมาณร้อยละ 20-25 ความยาวของกระดูกแข้ง (tibia) หรือฟันตัด (incisor) เมื่อปรับความถูกต้องด้วยน้ำหนักตัวแล้วมีค่าสัมประสิทธิ์การแปรผันประมาณร้อยละ 6 หรือสมบัติทางเคมีของดินมีค่าสัมประสิทธิ์การแปรผันต่ำกว่าสมบัติทางกายภาพ หรือความสูงของต้นข้าวโพด (ร้อยละ 2) มีค่าสัมประสิทธิ์การแปรผันต่ำกว่าความสูงของฝักวัดจากระดับดิน (ร้อยละ 7) และผลผลิตเมล็ดต่อไร่ (ร้อยละ 21) ตามลำดับ ดังนั้นนักวิจัยที่มีประสบการณ์หรือผู้เชี่ยวชาญในสาขานั้น ๆ จะทราบตัวลักษณะต่าง ๆ ที่ศึกษาของสิ่งมีชีวิตลักษณะใดมีความผันแปรมากน้อยกว่ากัน ในทางปฏิบัติ เราอาจประเมินค่าสัมประสิทธิ์การแปรผันจากค่าอัตราพันธุกรรม (heritability, h^2) ของลักษณะนั้น ๆ กล่าวคือ ลักษณะใดที่มีค่าอัตราพันธุกรรมสูง ข้อมูลจะมีความผันแปรน้อยกว่าลักษณะที่มีค่าอัตราพันธุกรรมต่ำ อาทิ ในโคเนื้อ การวัดส่วนสูงจะให้ค่าสัมประสิทธิ์การแปรผันต่ำกว่าการวัดลักษณะทางการสืบพันธุ์ (ช่วงห่างของการตกไข่) [3,4]

บ่อยครั้งที่ระดับของความผันแปรในการทดลองมีความสัมพันธ์กับค่าเฉลี่ยของตัวแปรตาม อาทิ ในสัตว์บางชนิด น้ำหนักแรกเกิดมีค่าเบี่ยงเบนมาตรฐานต่ำกว่าน้ำหนักเต็มวัย (อีกนัยหนึ่ง คือน้ำหนักตัวมาก ค่าเบี่ยงเบนมาตรฐานมาก) ด้วยเหตุผลดังกล่าวค่าสัมประสิทธิ์การแปรผันจึงเป็นค่าที่ค่อนข้างคงที่มากกว่าค่าเบี่ยงเบนมาตรฐาน ซึ่งค่าสัมประสิทธิ์การแปรผันจะมีค่าเดียวสำหรับทุกกลุ่มในการทดลอง ถ้าในแต่ละทรีตเมนต์มีค่าเบี่ยงเบนมาตรฐานเป็นสัดส่วนโดยตรงกับค่าเฉลี่ยของตัวแปรตามในการทดลองนั้น อย่างไรก็ตาม หากระดับของความผันแปรไม่สัมพันธ์กับระดับของค่าเฉลี่ย ในกรณีหลังนี้ค่าสัมประสิทธิ์การแปรผันจะไม่มี ความหมายในการแปลผล [5]

กรณีของค่าเบี่ยงเบนมาตรฐานและความคลาดเคลื่อนมาตรฐาน บ่อยครั้งที่นักวิจัยสับสนในการนำไปประยุกต์ใช้ ต่อไปนี้เป็นตัวอย่างเพื่อเปรียบเทียบความแตกต่างระหว่างการวัดการกระจายทั้งสองแบบ นักพันธุศาสตร์ซึ่งน้ำหนักลูกแกะเพศเมียแรกเกิดจำนวน 28 ตัว โดยลูกแกะทั้งหมดคลอดในเดือนเมษายน และเป็นแกะพันธุ์เดียวกัน คือ Rambouillet นอกจากนี้แกะทุกตัวคลอดเพียงตัวเดียว มีใช้ฝาแฝด น้ำหนักลูกแกะแต่ละตัวแสดงในตารางที่ 1 จากข้อมูลดังกล่าวคำนวณค่าเฉลี่ยเท่ากับ 5.17 กิโลกรัม ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.65 กิโลกรัม และค่าความคลาดเคลื่อนมาตรฐานเท่ากับ 0.12 กิโลกรัม ค่าเบี่ยงเบนมาตรฐานอธิบายความผันแปรของน้ำหนักตัวลูกแกะตัวหนึ่งจากลูกแกะตัวอื่น ๆ ในขณะที่ค่าความคลาดเคลื่อนมาตรฐานอธิบายความไม่น่าเชื่อถือ (อันเนื่องมาจากความผิดพลาดโดยสุ่ม) ของค่าเฉลี่ยตัวอย่างซึ่งใช้ประมาณค่าเฉลี่ยประชากร อีกวิธีหนึ่งที่สามารถทำให้มองเห็นภาพความแตกต่างของค่าเบี่ยงเบนมาตรฐานและความคลาดเคลื่อนมาตรฐานชัดเจนขึ้น คือ การพิจารณาตัวอย่างที่มีขนาดแตกต่างกัน จากรูปที่ 4 เมื่อขนาดตัวอย่างใหญ่ขึ้น ค่าเฉลี่ยตัวอย่าง ($\bar{y}$) และค่าเบี่ยงเบนมาตรฐานตัวอย่าง (s) มีแนวโน้มเข้าใกล้ค่าเฉลี่ยประชากร (μ) และค่าเบี่ยงเบนมาตรฐานประชากร (σ) มากขึ้น ตามลำดับ ตรงกันข้ามค่าความคลาดเคลื่อนมาตรฐานกลับมีแนวโน้มลดลง เมื่อขนาดตัวอย่างใหญ่ขึ้น อีกนัยหนึ่งคือเมื่อ n มีค่าใหญ่มาก ๆ ความคลาดเคลื่อนมาตรฐานจะเล็กมาก ๆ ดังนั้นการประมาณค่าเฉลี่ยประชากรด้วยค่าเฉลี่ยตัวอย่างจะแม่นยำมากขึ้นตามลำดับ ซึ่งจากตัวอย่างแกะ สมมุติสุ่มตัวอย่างครั้งที่ 2 และ 3 ให้มีจำนวนเท่ากับ 280 และ 2,800 ตัว ตามลำดับ จะเห็นว่าเมื่อขนาดตัวอย่างใหญ่ขึ้น ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของตัวอย่างจะเข้าใกล้ค่าเฉลี่ยและค่า

เบี่ยงเบนมาตรฐานของประชากรตามลำดับ ส่วนความคลาดเคลื่อนมาตรฐานจะมีแนวโน้มเข้าใกล้ 0 หรือหมายความว่าหากค่าความคลาดเคลื่อนมาตรฐาน

เท่ากับ 0 ค่าเฉลี่ยตัวอย่างจะมีค่าเท่ากับค่าเฉลี่ยประชากรนั่นเอง ข้อสังเกตค่าความคลาดเคลื่อนมาตรฐานจะมีค่าต่ำกว่าค่าเบี่ยงเบนมาตรฐานเสมอ [5]



รูปที่ 4 การสุ่มตัวอย่างที่มีขนาดต่าง ๆ กันของน้ำหนักแรกเกิดแกะแพคเมียร์พันธุ์ Rambouillet
ที่มา : Samuels และคณะ [1]

ตารางที่ 1 น้ำหนักแรกเกิดลูกแกะแพคเมียร์พันธุ์ Rambouillet จำนวน 28 ตัว

น้ำหนักแรกเกิด (กิโลกรัม)						
4.3	5.2	6.2	6.7	5.3	4.9	4.7
5.5	5.3	4.0	4.9	5.2	4.9	5.3
5.4	5.5	3.6	5.8	5.6	5.0	5.2
5.8	6.1	4.9	4.5	4.8	5.4	4.7

ที่มา : Samuels และคณะ [1]

การตัดสินใจว่าจะเลือกนำเสนอค่าเบี่ยงเบนมาตรฐานของตัวอย่างหรือความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ยขึ้นอยู่กับเจตนาของนักวิจัยว่าต้องการสื่อให้ผู้อ่านได้รับข้อมูลอย่างไร ถ้าการสรุปผลการทดลองของนักวิจัยอยู่บนพื้นฐานของการใช้ข้อมูลเพียงชุดเดียวเป็นตัวแทนของประชากรทั้งหมด ในกรณีนี้ควรรายงานเป็นความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย แต่ถ้าการสรุปผลการทดลองจำกัดเฉพาะตัวอย่างที่นักวิจัยมีอยู่เท่านั้น กรณีหลังนี้การรายงานค่าเบี่ยงเบนมาตรฐานของตัวอย่างจะเป็นสิ่งที่บ่งบอก

ถึงความซื่อสัตย์ของนักวิจัยในการนำเสนอข้อมูล เนื่องจากความคลาดเคลื่อนมาตรฐานมีค่าต่ำกว่าค่าเบี่ยงเบนมาตรฐานเสมอ ดังนั้นการใช้ความคลาดเคลื่อนมาตรฐานรายงานความผันแปร จึงทำให้ผู้อ่านคิดว่าการวัดข้อมูลมีความเที่ยงมาก อนึ่งผู้เชี่ยวชาญทางสถิติหลายท่าน เช่น Gotelli และ Ellison [6]; Olsen [7] และ Curran-Everett และ Benos [8] สนับสนุนให้รายงานค่าเบี่ยงเบนมาตรฐาน ซึ่งเป็นค่าที่มีความเที่ยงและสะท้อนให้เห็นความผันแปรของค่าสังเกตแต่ละค่าในตัวอย่างชุดนั้นอย่างแท้จริง อย่างไรก็ตาม หากนักวิจัยแทรกขนาดตัวอย่างเข้าไปในเนื้อหา (text) รูปภาพ คำอธิบายรูปภาพหรือตาราง ผู้อ่านสามารถคำนวณค่าเบี่ยงเบนมาตรฐานหรือความคลาดเคลื่อนมาตรฐานกลับไปกลับมาได้ (ดูรูปที่ 1 ประกอบ)

ถึงแม้จะมีความชัดเจนในทางสถิติว่าหากข้อมูลมีการกระจายแบบปกติ ค่าสถิติของตัวอย่างที่ควรใช้อธิบายพารามิเตอร์ของประชากรคือค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐาน แต่บ่อยครั้งที่นักวิจัย (หรือ

ผู้ทรงคุณวุฒิ/ผู้ประเมินของวารสาร) เลือกใช้ (หรือนำให้นักวิจัยใช้) ความคลาดเคลื่อนมาตรฐานอธิบายความผันแปรของตัวอย่าง จากการรวบรวมข้อมูลของ Nagele [9] พบว่าร้อยละ 23 (198/860 บทความ) ของบทความวิจัยในวารสารทางด้านวิสัญญี (anesthesia) รวม 4 วารสาร ที่ถูกเผยแพร่ในปี พ.ศ. 2544 ใช้ความคลาดเคลื่อนมาตรฐานไม่เหมาะสม โดยใช้ความคลาดเคลื่อนมาตรฐานเป็นค่าสถิติพรรณนาอธิบายความผันแปรของตัวอย่างที่ศึกษา สำหรับข้อมูลที่ถูกวิเคราะห์โดยการทดสอบไม่อิงพารามิเตอร์ (non-parametric test หรือ distribution-free test) อาทิ การทดสอบแมนน์-วิตนีย์ (Mann-Whitney test) การรายงานผลการทดลองที่เหมาะสม คือ มัธยฐาน (median) ร่วมกับค่าต่ำสุด ค่าสูงสุด เปอร์เซ็นไทล์ที่ 25 และ 75 ตามลำดับ [7]

โดยทั่วไปในการทดลองทางชีววิทยา นักวิจัยสนใจที่จะเปรียบเทียบตัวอย่างจาก 2 กลุ่ม หรือมากกว่าว่ามีความแตกต่างกันหรือไม่ เช่น นักชีววิทยาอาจต้องการเปรียบเทียบหนูสายพันธุ์ดั้งเดิม (wild type) กับหนูสายพันธุ์กลาย (mutant type) หรือแม่ไก่ที่เลี้ยงในกรงทั่วไปกับแม่ไก่ในกรงเพิ่มพูนสภาพแวดล้อม (enriched cage) ดังนั้นเพื่อตัดสินว่าความแตกต่างที่เกิดขึ้นเป็นความแตกต่างที่แท้จริง (มีนัยสำคัญ) หรือเป็นแค่เพียงความแตกต่างที่เกิดขึ้นโดยบังเอิญ ซึ่งการนำเสนอข้อมูลโดยใช้ภาพ (visual display) สามารถใช้แท่งความคลาดเคลื่อน (error bar) แทนค่าสถิติที่แสดงความผันแปรค่าใดค่าหนึ่งที่กล่าวมาข้างต้น แผนภูมิแท่งหรือแผนภูมิเส้นที่แสดงร่วมกับแท่งความคลาดเคลื่อนของค่าความคลาดเคลื่อนมาตรฐานหรือช่วงความเชื่อมั่นบ่งชี้ว่าค่าเฉลี่ยที่คาดหวังของข้อมูลทั้งหมดจากการทดลองหรือประชากรตกอยู่ในช่วงใด โดยที่ช่วงห่างระหว่างขอบบนและขอบล่างของแท่งความคลาดเคลื่อน แสดงค่าที่

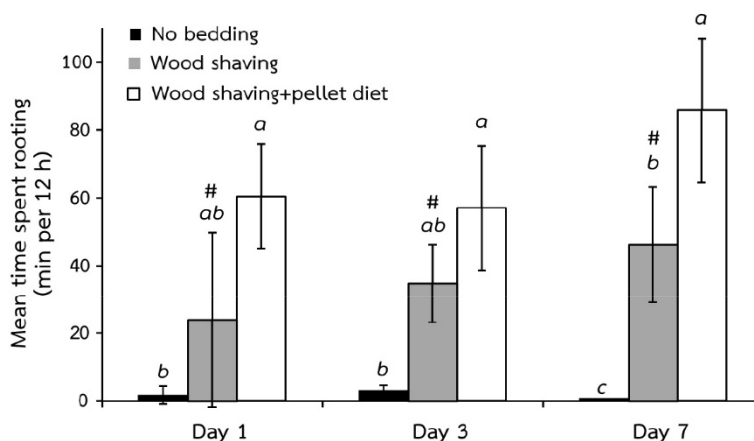
เป็นไปได้มากที่สุดของค่าเฉลี่ยประชากร

การทดสอบสมมติฐาน หากค่าพี (p-value) มีค่ามากกว่า 0.05 เราไม่สามารถสรุปว่าอิทธิพลมีความแตกต่าง (อย่างมีนัยสำคัญ) ในขณะเดียวกันเราก็ไม่สามารถสรุปได้ว่าอิทธิพลมีค่าเท่ากับศูนย์ เนื่องจากอิทธิพลอาจมีอยู่จริง แต่อิทธิพลมีขนาดเล็กมาก หรือการทดลองอาจมีจำนวนซ้ำน้อยเกินไป ไม่มากพอที่จะแสดงให้เห็นนัยสำคัญ อย่างไรก็ตาม ในการรายงานผลการทดสอบสมมติฐาน วารสารระดับนานาชาติบางเล่มอาจกำหนดให้ผู้เขียนระบุค่าพี (exact p-value) ลงไปในเนื้อหาของบทความ เช่น $P=0.15$ หรือ $P=0.002$ สำหรับหลักการคร่าว ๆ ในการแปลความหมายของแท่งความคลาดเคลื่อนคือ หากแท่งความคลาดเคลื่อนระหว่างสองกลุ่มเหลื่อมกันน้อย หรือมีช่องห่าง (gap) ระหว่างแท่งความคลาดเคลื่อน (ขอบล่างของกลุ่มหนึ่งกับขอบบนของอีกกลุ่มหนึ่ง) กว้าง การทดสอบนัยสำคัญจะให้ค่าพีต่ำ หรือหมายความว่าหลักฐานมีความหนักแน่นเพียงพอที่สามารถบ่งชี้เป็นนัยว่าความแตกต่าง (อิทธิพล) นั้นมีอยู่จริง อย่างไรก็ตาม ข้อเสนอแนะดังกล่าวเป็นแค่เพียงแนวทางในการพิจารณาเบื้องต้นจากภาพเท่านั้น เนื่องจากถึงแม้ว่าจะมีการเหลื่อมกันระหว่างแท่งความคลาดเคลื่อนของแต่ละกลุ่ม แต่ความแตกต่างของค่าเฉลี่ยระหว่างกลุ่มก็ยังมีโอกาสมีนัยสำคัญทางสถิติได้ [10] ในการสร้างแผนภาพ นักวิจัยต้องอธิบายให้ชัดเจนว่าสถิติที่เลือกใช้อธิบายความผันแปรของตัวอย่างในแผนภาพคืออะไร เพื่อป้องกันการเข้าใจผิดและนำไปสู่การแปลผลหรือการอนุมานที่ผิดพลาดของผู้อ่าน [11] รูปที่ 5 แสดงระยะเวลาที่ใช้ในการดูด (rootings) ของลูกสุกรหลังหย่านมที่เลี้ยงไว้ในคอกซึ่งมีวัสดุรองพื้นต่าง ๆ กัน โดยนักวิจัยได้อธิบายไว้ได้ภาพว่าค่าสถิติที่นำเสนอความผันแปรคือค่าเบี่ยงเบนมาตรฐาน จากภาพจะเห็นว่าค่าเบี่ยงเบนมาตรฐานของพฤติกรรมการดูดมีค่าค่อนข้าง

สูง ซึ่งแสดงให้เห็นว่าลูกสุกรแต่ละตัวในแต่ละกลุ่มแสดงการดูดค่อนข้างผันแปรมาก

สำหรับในเนื้อหาของบทความวิจัย บ่อยครั้งที่นักวิจัยรายงานผลในรูปแบบ “ค่าเฉลี่ย $\pm$ ความผันแปร” นักวิจัยและผู้ประเมินของวารสารต้องแน่ใจว่า

ค่าทั้งสองเทอมนี้ได้ถูกอธิบายไว้อย่างชัดเจน ยกตัวอย่าง อธิบายไว้ในส่วนอุปกรณ์และวิธีการ (materials and methods) ว่า “ผลการทดลองแสดงเป็นค่าเฉลี่ย $\pm$ ค่าเบี่ยงเบนมาตรฐาน”



The effect of environmental enrichment on rooting behaviour in piglets on day 1, 3 and 7 after weaning. Error bars represent standard deviation. Significant ($P < 0.05$) treatment effects are denoted by: *abc* significant difference between treatments; # significant difference between enriched and barren pen treatments.

รูปที่ 5 ตัวอย่างการนำเสนอข้อมูลด้วยแผนภูมิแท่ง (ค่าเฉลี่ย) ร่วมกับแท่งความคลาดเคลื่อน (ค่าเบี่ยงเบนมาตรฐาน) (ที่มา : ดัดแปลงจาก พัพฒน์ และคณะ [12])

สรุป

การอธิบายการกระจายของข้อมูลเป็นความผิดพลาดที่พบบ่อยในการรายงานวิจัยหรือการทดลองทางชีววิทยา สถิติที่ใช้อธิบายการกระจายของตัวอย่างมีให้เลือกใช้หลายค่า แต่ละค่ามีข้อจำกัดหรือวัตถุประสงค์ในการใช้ที่แตกต่างกันไป แต่ค่าเบี่ยงเบนมาตรฐานและความคลาดเคลื่อนมาตรฐานเป็นสถิติที่นักวิจัยส่วนใหญ่ใช้ในการอธิบายการกระจายของตัวอย่าง ดังนั้นนักวิจัยจึงควรเลือกใช้ค่าสถิติที่แสดงการกระจายให้เหมาะสมกับบริบทงานวิจัยของตน รวมทั้งควรระมัดระวังในการนำเสนอานวิจัยที่ได้เผยแพร่ในสื่อต่างๆ ของนักวิจัยท่านอื่นมาประยุกต์ใช้อย่างรอบคอบ

กิตติกรรมประกาศ

ผู้เขียนขอขอบพระคุณ รองศาสตราจารย์ ดร. วราฤทธิ์ พานิชกิจโกศลกุล ที่ให้ความกรุณาตรวจทานบทความวิชาการนี้

รายการอ้างอิง

- [1] Samuels, M.L., Witmer, J.A. and Schaffner, A.A., 2012, *Statistics for the Life Sciences*, 4th Ed., Prentice Hall, Upper Saddle River, 654 p.
- [2] Petrie, A. and Watson, P., 1999, *Statistics for Veterinary and Animal Science*, Alden Press, Oxford, 243 p.

- [3] Morris, T.R., 1999, Experimental Design and Analysis in Animal Science, Biddles, Ltd., Guildford, 208 p.
- [4] Aaron, D.K. and Hays, V.W., 2004, How many pig? Statistical power considerations in swine nutrition experiments, J. Anim. Sci. 82 (E. Suppl.): E245-E254.
- [5] Mead, R., Curnow, R.N. and Hasted, A.M., 1993, Statistical Methods in Agriculture and Experimental Biology, Chapman & Hall, London, 415 p.
- [6] Gotelli, N.J. and Ellison, A.M., 2013, A Primer of Ecological Statistics. 2nd Ed., Sinauer Associates, Inc., Sunderland, 614 p.
- [7] Olsen, C.H., 2003, Review of the use of statistics in Infection and Immunity, Infect. Immun. 71: 6689-6692.
- [8] Curran-Everett, D. and Benos, D.J., 2004, Guidelines for reporting statistics in journals published by the American Physiological Society, Am. J. Physiol. Cell Physiol. 287: C243-C245.
- [9] Nagele, P., 2001. Misuse of standard error of the mean (SEM) when reporting variability of a sample: A critical evaluation of four anaesthesia journals, Br. J. Anaes. 90: 514-516.
- [10] Schenker, N. and Gentleman, J., 2001, On judging the significance of differences by examining the overlap between confidence intervals, Am. Stat. 55 : 182-186.
- [11] Cumming, G. and Fidler, D., 2007, Error bars in experimental biology, J. Cell Biol. 177: 7-11.
- [12] พิพัฒน์ สมภาร, นกตล ทองย้อย และนวิญญา พิมพ์า, 2560, อิทธิพลของการเพิ่มพูนสภาพแวดล้อมต่อพฤติกรรมของลูกสุกรหลังหย่านม, ว. วิทย์. เกษตร 48(2)(พิเศษ): 248-255.