บทที่ 7
การแจกแจงตัวอย่าง

7.1 คำนำ

ปัญหาที่พบเห็นกันเสมอเกี่ยวกับการสำรวจความนิยมของประชาชนของผู้สมัครรับเลือกตั้งเมื่อสมัยที่มีการเลือกตั้ง ก็คือการเลือกตัวอย่างของผู้มีสิทธิออกเสียงเลือกตั้งด้วยวิธีการต่าง ๆ กัน เพื่อประมาณผู้มีสิทธิออกเสียงเลือกตั้งว่าจะเลือกผู้สมัครรับเลือกตั้งคนใดมากน้อยแค่ไหน กระบวนการเลือกตัวอย่างนี้อ้างถึงคุณลักษณะแสดงออกของประชากรนี้เป็นเพียงตัวอย่างหนึ่งเท่านั้น สำหรับตัวอย่างอื่น ๆ พอที่จะแสดงได้ดังนี้

1. สำนักงานสถิติแห่งชาติเลือกตัวอย่างเพื่อหาข้อมูล ข่าวสารเกี่ยวกับค่าจ้าง แรงงาน รายรับรายจ่าย การศึกษาและลักษณะแสดงออกของประชากร
2. ห้างหุ้นส่วนอุตสาหกรรมเลือกตัวอย่างจากกระบวนการผลิตของเขาเพื่อตรวจสอบคุณภาพของผลิตภัณฑ์
3. องค์การสำรวจความคิดเห็นเลือกตัวอย่างเพื่อประมาณความคิดเห็นเกี่ยวกับคำสั่งที่ประกาศออกใช้โดยเฉพาะ
4. กรมเศรษฐกิจพาณิชย์ใช้กระบวนการเลือกตัวอย่างเพื่อรวบรวมข้อมูลที่เกี่ยวกับเศรษฐกิจรายได้ของผู้บริโภค
5. องค์การวิจัยตลาดใช้การเลือกตัวอย่างเพื่อขยายการตรวจสอบข้อเสนอผลิตภัณฑ์ของผู้บริโภคผลของการโฆษณา ฯลฯ
6. การเลือกตัวอย่างใช้ในการเกษตรเพื่อประมาณผลที่เก็บเกี่ยวได้ การพยากรณ์สำหรับการป่าไม้ประมาณค่าจำนวนของท่อนไม้
7. การเลือกตัวอย่างใช้เพื่อตรวจสอบหนี้ สินค้าคงคลัง ความคลาดเคลื่อน จากการจองและขายตั๋ว ฯลฯ
8. กองทัพใช้ตัวอย่างเพื่อตรวจสอบคุณภาพของอาวุธต่าง ๆ เพื่อจัดซื้อ
9. หมอและนักวิทยาศาสตร์ใช้ตัวอย่างเพื่อตรวจสอบผลยาชนิดใหม่ ๆ หรือรวบรวมข้อมูลเกี่ยวกับมะเร็งในปอดจากผลของการสูบบุหรี่ เป็นต้น

จะเห็นได้ว่าความรู้ทั้งหมดที่หามาได้จากตัวอย่างที่เลือกมาจากประชากรเป็นแนวทางไม่มากนักที่ถูกกำหนดขึ้น ในอดีตสำหรับวิธีการและทฤษฎีการเลือกตัวอย่าง อย่างไรก็ตามการพัฒนาทั่ว ๆ ไปของทฤษฎีทางสถิติที่สำคัญเกี่ยวกับวิธีการและทฤษฎีการเลือกตัวอย่างได้อย่างรวดเร็วตั้งแต่ปี 1940 ใช้ข้อมูลตัวอย่างประมาณค่าคุณลักษณะแสดงออกของประชากร(อย่างเช่น ค่ามัชฌิม ผลรวมทั้งหมดหรือสัดส่วน) และประเมินค่าประมาณเหล่านี้

7.2 ทฤษฎีการเลือกตัวอย่าง

ทฤษฎีการเลือกตัวอย่างเป็นการศึกษาถึงความสัมพันธ์ระหว่างประชากรกับตัวอย่างที่เลือกมาจากประชากร ค่าต่าง ๆ ที่ได้จากตัวอย่างนี้แหละที่จะเอาไปประมาณค่าต่าง ๆ ของประชากรที่ไม่สามารถรู้ได้ (อย่างเช่น มัชฌิมเลขคณิต ความแปรปรวน ฯลฯ ของประชากร) เรียกว่า ตัวพารามิเตอร์ ปริมาณต่าง ๆ ของตัวอย่างที่เลือกมาจากประชากร (อย่างเช่น มัชฌิมเลขคณิต ความแปรปรวน ฯลฯ ของตัวอย่าง) เรียกว่า ตัวสถิติ ตัวพารามิเตอร์เป็นปริมาณคงที่ แต่ตัวสถิติจะมีค่าเปลี่ยนจากตัวอย่างหนึ่งไปอีกตัวอย่างหนึ่ง ในที่นี้เราให้ μ เป็นมัชฌิมเลขคณิตของประชากร $\bar{x}$ เป็นมัชฌิมเลขคณิตของตัวอย่าง เราอาจสงสัยว่าปริมาณ $(\bar{x} - \mu)$ เป็นตัวพารามิเตอร์หรือตัวสถิติ เราก็มาพิจารณาดูว่าค่าของ $(\bar{x} - \mu)$ จะเปลี่ยนจากตัวอย่างไปอีกตัวอย่างหรือไม่ ถ้าหากว่าไม่เปลี่ยนก็เป็นตัวพารามิเตอร์ ถ้าหากว่าเปลี่ยนก็เป็นตัวสถิติ แต่ปริมาณ $(\bar{x} - \mu)$ เปลี่ยนจากตัวอย่างหนึ่งไปอีกตัวอย่างหนึ่ง เพราะว่า $\bar{x}$ เปลี่ยนจากตัวอย่างไปอีกตัวอย่าง

ทฤษฎีการเลือกตัวอย่างยังมีประโยชน์ในการกำหนดว่าผลต่างระหว่างสองตัวอย่างแตกต่างกันโดยบังเอิญ หรือว่าแตกต่างกันอย่างมีนัยสำคัญ อย่างเช่น ในการทดสอบเกี่ยวกับวิธีการสอนแผนใหม่ย่อมดีกว่าวิธีการสอนแผนเก่า คำตอบนี้ใช้เกี่ยวกับการทดสอบสมมติฐานว่ามีนัยสำคัญหรือไม่ ซึ่งเป็นเนื้อหาของบทที่ 9

การศึกษาการอนุมานเกี่ยวกับประชากรโดยอาศัยตัวอย่างที่เลือกมาจากประชากร การรวบรวมพร้อมทั้งแสดงถึงความถูกต้องของการอนุมานโดยใช้ทฤษฎีความน่าจะเป็นเรียกว่า การอนุมานทางสถิติ

7.3 เหตุผลในการเลือกตัวอย่าง

เราทราบแล้วว่าประชากร หมายถึง จำนวนข้อมูลทั้งหมดหรือข้อมูลที่อยู่ในข่ายการพิจารณาทั้งหมด ในทางปฏิบัติแล้วปริมาณต่าง ๆ ของประชากรหรือตัวพารามิเตอร์ เราไม่สามารถจะหาได้และงานด้านสถิติส่วนมากจึงจำเป็นต้องมีการเลือกตัวอย่างมาทำหน้าที่แทนประชากร เพราะว่า

1. ประหยัดเวลา การรวบรวมข้อมูลจากตัวอย่างย่อมใช้เวลาน้อยกว่าการรวบรวมข้อมูลทั้งหมดของประชากร เพราะว่าตัวอย่างประกอบด้วยข้อมูลเพียงส่วนหนึ่งของประชากร หรือในกรณีที่เราต้องการข้อมูลข่าวสารเร็ว ๆ หรือมีเวลาจำกัด

2. ประหยัดค่าใช้จ่าย ค่าใช้จ่ายในการรวบรวมข้อมูลจำนวนมากย่อมมากกว่าค่าใช้จ่ายในการรวบรวมข้อมูลเพียงส่วนหนึ่ง หรือในการศึกษาประชากรทั้งหมดนั้นมันคุ้มกับผลที่ได้หรือไม่ หรือกรณีที่งบมีไม่เพียงพอสำหรับค่าใช้จ่ายในการรวบรวมข้อมูลหรือส่งคนไปสัมภาษณ์ประชากรทั้งหมด

3. ในบางเรื่องเราไม่อาจอาศัยประชากรทั้งหมดได้ อย่างเช่นการทดสอบอายุการใช้งานของหลอดไฟฟ้า ถ้าหากว่าเราทดสอบทุกหลอดก็คงไม่เหลือไว้สำหรับขายหรือการชิมรสแกงสักหม้อหนึ่ง หากชิมจนหมดหม้อก็คงจะไม่เหลือไว้ขาย หรือแม่ค้าขายผลไม้ แม่ค้าจะเอาผลไม้ผลหนึ่งหรือในกรณีผลโตก็อาจเอาเพียงส่วนหนึ่งของผลให้ชิมว่ามีรสชาติอย่างไร ผลไม้ผลหนึ่งหรือส่วนหนึ่งก็เป็นตัวอย่าง

4. ความถูกต้องได้ดีกว่า เช่น การสอบถามทัศนคติต่อการเรียนทางวิทยุหรือทางโทรทัศน์ของนักศึกษามหาวิทยาลัยรามคำแหง โดยการส่งผู้สัมภาษณ์ที่ชำนาญงานไปสอบถามนักศึกษาที่เลือกเป็นตัวอย่าง เราจะได้คำตอบสำหรับคำถามที่ต้องการทราบถูกต้องแม่นยำมากกว่าที่จะส่งผู้สัมภาษณ์ที่ไม่ชำนาญงานจำนวนมาก ๆ ไปสอบถามทุกคน เพราะเราสามารถควบคุมได้ทั่วถึง

5. ศึกษาได้กว้างขวางกว่า เนื่องจากว่าศึกษาจากข้อมูลไม่มากมาย จึงสามารถสอบถามรายละเอียดได้กว้างขวางกว่า

การเลือกตัวอย่างนั้นใช้อยู่แพร่หลายในงานด้านควบคุมคุณภาพสินค้า การวิจัยทางสังคมประเภทต่าง ๆ การตรวจบัญชี การทดลองทางการเกษตร ตลอดจนการทดลองทางวิทยาศาสตร์อีกมากมาย

7.4 ตัวอย่างสุ่มและหมายเลขสุ่ม

เพื่อที่จะสรุปทฤษฎีการเลือกตัวอย่างและการอนุมานทางสถิติให้สมบูรณ์ยิ่งขึ้น ตัวอย่างจะต้องได้รับเลือกมาเป็นตัวแทนของประชากรและตัวอย่างนั้นจะต้องประกอบด้วยสมาชิกที่มีโอกาสได้รับเลือกเท่า ๆ กันจากประชากร อย่างเช่น เขียนชื่อนักศึกษาทั้งหมดหรือกำหนดหมายเลขสมาชิกของประชากรลงในแผ่นกระดาษแต่ละแผ่นใส่ลงในกล่องแล้วคนให้เข้ากัน และหยิบแผ่นกระดาษ จากกล่อง เราอาจใช้ตารางหมายเลขสุ่มแทนก็ได้หรือเครื่องคอมพิวเตอร์เลือกใช้ได้ตามขนาดที่ **ต้องการ วิธีการเลือกแบบนี้เรียกว่า การสุ่มตัวอย่าง สำหรับการศึกษาวิธีของการสุ่มตัวอย่างหรือการเลือกตัวอย่างและปัญหาที่เกี่ยวข้องที่ยกขึ้นมา เรียกว่า การออกแบบทดลอง**

7.5 วิธีการเลือกตัวอย่างแบบต่าง ๆ

วิธีการเลือกทั่ว ๆ ไปอาจแบ่งได้ 2 แบบ คือ

1. การเลือกแบบเจาะจง การเลือกแบบนี้โดยอาศัยการตัดสินใจของผู้เลือกว่าจะเอาหน่วยใดของประชากรเป็นตัวอย่าง

2. การเลือกแบบสุ่ม การเลือกแบบนี้ทำให้ทุกหน่วยในประชากรมีโอกาสได้รับเลือกมาเป็นตัวอย่างเท่า ๆ กัน และวิธีการเลือกแบบสุ่มนี้อาจจะมีแบ่งย่อยออกไปได้อีกหลายแบบ อย่างเช่น การสุ่มตัวอย่างแบบขั้นภูมิ แบบกลุ่ม แบบระบบ เป็นต้น สำหรับตัวอย่างที่ได้มาจากการสุ่มแบบแทนที่คือ สมาชิกที่ถูกเลือกแล้วถูกใส่กลับไปในประชากรและมีโอกาสถูกเลือกอีก ประชากรที่สมาชิกได้รับเลือกแบบหยิบแล้วใส่คืนนี้ เรียกว่า ประชากรแบบไม่จำกัด สำหรับการสุ่มแบบไม่แทนที่คือ สมาชิกที่ถูกเลือกแล้วไม่มีโอกาสถูกเลือกอีกประชากรที่สมาชิกได้รับเลือกแบบหยิบแล้วไม่ใส่คืนนี้ เรียกว่า ประชากรแบบจำกัด ประชากรเป็นได้ทั้งแบบจำกัดหรือไม่จำกัด ในที่นี้จำนวนไม่จำกัดไม่ได้หมายความว่าจำนวนข้อมูลที่ประกอบเป็นประชากรนั้นจะต้องมีจำนวนมาถึงอนันต์ อาจเป็นเท่าใดก็ได้สุดแล้วแต่เหตุการณ์หรือกรณีไป พอจะยกตัวอย่าง เราหยิบลูกบอลล์ 10 ลูก ตามลำดับจาก 100 ลูก ในภาชนะใบหนึ่งโดยหยิบครั้งละลูกแล้วไม่ใส่กลับลงไปในภาชนะก่อนหยิบลูกต่อไปเรื่อย ๆ แบบนี้จนครบ 10 ลูก วิธีการเลือกแบบนี้เป็นการเลือกจากประชากรแบบจำกัด แต่ถ้าเราโยนเหรียญอันหนึ่ง 50 ครั้ง และนับจำนวนที่จะออกหัว วิธีการเลือกแบบนี้เป็นการเลือกจากประชากรแบบไม่จำกัด ในทางทฤษฎีประชากรแบบจำกัดถูกเลือกตัวอย่างโดยวิธีเลือกแล้วใส่คืนถือให้เป็นประชากรแบบไม่จำกัด เนื่องจากจำนวนใดจำนวนหนึ่งของตัวอย่างสามารถที่จะถูกเลือกไม่ได้ใช้ประชากรทั้งหมด สำหรับในทางปฏิบัติการเลือกตัวอย่างจากประชากรแบบจำกัดที่มีจำนวนมากก็สามารถที่จะพิจารณาเหมือนกับการเลือกตัวอย่างจากประชากรแบบไม่จำกัด

7.6 การแจกแจงตัวอย่าง

พิจารณาตัวอย่างที่อาจเป็นไปได้ทั้งหมดขนาด n ซึ่งสามารถเลือกมาจากประชากรที่กำหนดให้ (จะเป็นแบบเลือกแล้วใส่คืนหรือแบบเลือกแล้วไม่ใส่คืน) สำหรับแต่ละตัวอย่างเราสามารถคำนวณตัวสถิติ อย่างเช่น มัชฌิมเลขคณิต ส่วนเบี่ยงเบนมาตรฐาน ฯลฯ ซึ่งจะแปรไปจากตัวอย่างหนึ่งไปอีกตัวอย่างหนึ่ง ในวิธีการเช่นนี้เราสามารถแจกแจงตัวสถิติที่คำนวณมาจากตัวอย่างได้ เรียกว่า การแจกแจงตัวอย่าง

ถ้าหากว่าตัวสถิตินั้นเป็นมัชฌิมเลขคณิตตัวอย่าง เรียกการแจกแจงนี้ว่า การแจกแจงตัวอย่างของมัชฌิมเลขคณิต ในทำนองเดียวกันเราสามารถแจกแจงตัวอย่างของส่วนเบี่ยงเบนมาตรฐาน ความแปรปรวน มัธยฐาน สัดส่วน ฯลฯ ได้

แต่ละการแจกแจงตัวอย่าง เราสามารถคำนวณมัชฌิมเลขคณิต ส่วนเบี่ยงเบนมาตรฐาน ฯลฯ ดังนั้นเราจึงสามารถกล่าวได้ว่ามัชฌิมเลขคณิตและส่วนเบี่ยงเบนมาตรฐานของการแจกแจงตัวอย่างของมัชฌิมเลขคณิต ฯลฯ

7.7 นิยามทฤษฎีขีดจำกัดส่วนกลาง

กำหนดให้การแจกแจงมัชฌิมเลขคณิตตัวอย่าง $(\bar{x})$ ทั้งหมดที่ได้เลือกมาจากประชากรขนาด N จะเข้าใกล้การแจกแจงปกติ ขณะที่ขนาดตัวอย่าง (n) เข้าใกล้อนันต์

ถ้าหากว่าการแจกแจงประชากรมีมัชฌิมเลขคณิต μ และส่วนเบี่ยงเบนมาตรฐาน σ แล้วการแจกแจงมัชฌิมเลขคณิต ตัวอย่าง $(\bar{x})$ จะมีมัชฌิมเลขคณิต μ ส่วนเบี่ยงเบนมาตรฐาน $\sigma/\sqrt{n}$ (ในกรณีประชากรเป็นแบบไม่จำกัด) ภายใต้เงื่อนไขเหล่านี้

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

มีการแจกแจงปกติมาตรฐาน n (0, 1) โดยประมาณ ขณะ n เข้าใกล้อนันต์

7.8 การแจกแจงตัวอย่างของมัชฌิมเลขคณิต

สมมติว่าเลือกตัวอย่างสุ่มขนาด (n) ที่อาจเป็นไปได้ทั้งหมดแบบเลือกแล้วใส่คืนจากประชากรแบบจำกัดขนาด N(N > n) ถ้าหากว่าเราให้ $\mu_{\bar{x}}$ และ $\sigma_{\bar{x}}$ เป็นมัชฌิมเลขคณิตและส่วนเบี่ยงเบนมาตรฐานของการแจกแจงตัวอย่างของมัชฌิมเลขคณิตตามลำดับ พร้อมทั้งมัชฌิมเลขคณิตและส่วนเบี่ยงเบนมาตรฐานของประชากรเป็น μ และ σ ตามลำดับแล้ว

$$\mu_{\bar{x}} = \mu, \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \qquad \text{.......(7.8.1)}$$

ถ้าหากว่าประชากรเป็นแบบไม่จำกัดหรือการเลือกตัวอย่างเป็นแบบเลือกแล้วใส่คืน ผลลัพธ์ข้างต้นก็ลดลงเหลือเป็น

$$\mu_{\bar{x}} = \mu, \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad \text{..........(7.8.2)}$$

สำหรับตัวอย่างขนาดใหญ่ $(n > 30)$ การแจกแจงตัวอย่างของมัชฌิมจะมีการแจกแจงปกติโดยประมาณมีมัชฌิม $\mu_{\bar{x}}$ และส่วนเบี่ยงเบนมาตรฐาน $\sigma_{\bar{x}}$ สำหรับผลลัพธ์ของประชากรแบบไม่จำกัดเป็นกรณีเฉพาะของทฤษฎีขีดจำกัดส่วนกลางของทฤษฎีความน่าจะเป็น ซึ่งได้พิสูจน์ความถูกต้องมากขึ้น ขณะที่ n มีค่ามากขึ้นหรือจะกล่าวได้ว่า การแจกแจงตัวอย่างเป็น asymptotically normal

ในกรณีประชากรมีการแจกแจงปกติ การแจกแจงตัวอย่างของมัชฌิมเลขคณิตย่อมมีการแจกแจงปกติด้วย แม้ว่าค่าของ n จะมีค่าน้อยกว่า 30 (เมื่อทราบ σ) และส่วนเบี่ยงเบนมาตรฐานหรือความคลาดเคลื่อนมาตรฐาน $\sigma_{\bar{x}}$ ที่ได้เลือกมาจากประชากรแบบจำกัด ($\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$) หรือแบบไม่จำกัด ($\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$) จะไม่แตกต่างกันก็ต่อเมื่อขนาดตัวอย่าง n มีค่าน้อยกว่าห้าเปอร์เซ็นต์ของขนาดประชากร N ดังนั้นการใช้ $\frac{N-n}{N-1}$ (เรียกว่า correction factor) ก็ต่อเมื่อขนาดตัวอย่าง $n > (.05)\,N$

ตัวอย่าง (แสดงการแจกแจงตัวอย่างของ $\bar{x}$)
ประชากรประกอบด้วยค่า 2, 4, 6 เลือกตัวอย่างขนาด n = 2 จากประชากรนี้แบบเลือกแล้วใส่คืนเข้าไปในประชากร

วิธีทำ มัชฌิมเลขคณิตประชากร μ และความแปรปรวนของประชากร σ^2 คือ

$$\mu = \frac{2+4+6}{3} = 4$$

$$\sigma^2 = \frac{(2-4)^2 + (4-4)^2 + (6-4)^2}{3} = 8/3$$

ถ้าหากว่าเราเลือกตัวอย่างขนาด 2 แบบเลือกแล้วใส่คืนจะได้ตัวอย่างที่เป็นไปได้ทั้งหมด 3 × 3 = 9 คำนวณค่ามัชฌิมเลขคณิตตัวอย่าง $\bar{x}$ ของแต่ละตัวอย่างได้ดังนี้

ตัวอย่าง	ค่าเฉลี่ย	$\bar{x}$	f	$f\bar{x}$	$f(\bar{x} - E(\bar{X}))^2$
2,2	2	2	1	2	4
2,4	3	3	2	6	2
2,6	4	4	3	12	0
4,2	3	5	2	10	2
4,4	4	6	1	6	4
4,6	5		9	36	12
6,2	4				
6,4	5				
6,6	6				

เมื่อนำค่ามัชฌิมเลขคณิตของตัวอย่างทั้ง 9 ตัวอย่างมาแจงตามความถี่จะได้การแจกแจงตัวอย่างของค่ามัชฌิมเลขคณิตดังข้างต้น

$$\mu_{\bar{X}} = \frac{\Sigma f\bar{x}}{\Sigma f} = \frac{36}{9} = 4 = \mu$$

$$\sigma^2_{\bar{X}} = \frac{\Sigma f(\bar{x} - \mu_{\bar{X}})^2}{\Sigma f} = \frac{12}{9} = \frac{4}{3}$$

$$= \frac{8/3}{2} = \frac{\sigma^2}{n}$$

นี้แสดงว่าการแจกแจงตัวอย่างของ $\bar{x}$ มีค่าส่วนเฉลี่ยเลขคณิต μ ความแปรปรวน $\sigma^2/_n$

กรณีเลือกตัวอย่างแบบเลือกแล้วไม่ใส่คืนจะได้ตัวอย่างที่เป็นไปได้ทั้งหมด 6 ค่ามัชฌิมเลขคณิต $\bar{x}$ ของแต่ละตัวอย่างเป็น

ตัวอย่าง	$\bar{x}$
2, 4	3
2, 6	4
4, 2	3
4, 6	5
6, 2	4
6, 4	5

x	f	$f\bar{x}$	$f(\bar{x} - \mu_{\bar{X}})^2$
3	2	6	2
4	2	8	0
5	2	10	2
	6	24	4

$$\mu_{\bar{X}} = \frac{\Sigma f\bar{x}}{\Sigma f} = \frac{24}{6} = 4 = \mu$$

$$\sigma^2_{\bar{X}} = \frac{\Sigma f(\bar{x} - \mu_{\bar{X}})^2}{\Sigma f} = \frac{4}{6}$$

$$= \frac{8/3}{2}\left(\frac{3-2}{3-1}\right) = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$$

การแจกแจงตัวอย่างของ $\bar{x}$ มีค่ามัชฌิมเลขคณิต μ ความแปรปรวน $\frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$

ตัวอย่าง สมมติว่าความสูงของนักศึกษาชาย 3,000 คน ของมหาวิทยาลัยแห่งหนึ่งมีการแจกแจงปกติ ซึ่งมีมัชฌิมเลขคณิต 68.0 นิ้ว กับส่วนเบี่ยงเบนมาตรฐาน 3.0 นิ้ว ถ้าเลือก 80 ตัวอย่างประกอบด้วยตัวอย่างละ 25 คน จงคำนวณหามัชฌิมเลขคณิต $\mu_{\bar{X}}$ และส่วนเบี่ยงเบนมาตรฐาน $\sigma_{\bar{X}}$ ของการแจกแจงตัวอย่างของมัชฌิมเลขคณิต ตัวอย่างที่หวังว่าจะได้ในเมื่อการเลือกโดยวิธี

1. เลือกแล้วใส่คืน

2. เลือกแล้วไม่ใส่คืน

3. จะมีกี่ตัวอย่างที่คาดหวังจะหาได้ระหว่างมัชฌิมเลขคณิตตัวอย่าง $(\bar{x})$ 66.8 กับ 68.3 นิ้ว

4. จะมีกี่ตัวอย่างที่ท่านคาดหวังจะหาได้เมื่อมัชฌิมเลขคณิตตัวอย่างน้อยกว่า 66.4 นิ้ว

วิธีทำ

1 . $\mu_{\bar{X}} = \mu = 68.0$ นิ้ว

$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = 3/\sqrt{25} = 0.6$ นิ้ว

2. $\mu_{\bar{X}} = \mu = 68.0$ นิ้ว

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{3}{\sqrt{25}}\sqrt{\frac{3000-25}{3000-1}}$$

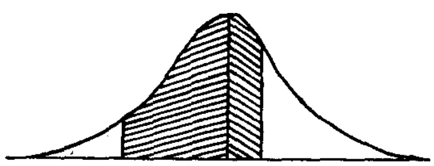
ซึ่งมีค่าน้อยกว่า 0.6 นิ้ว น้อยมากและสามารถพิจารณาให้เหมือนกับการเลือกตัวอย่างแบบเลือกแล้วใส่คืน เพราะว่า $n < .05N$

ดังนั้น เราจึงคาดหวังการแจกแจงตัวอย่างของการทดลองของมัชฌิมเลขคณิตเข้าใกล้ปกติ มีมัชฌิมเลขคณิต 68.0 นิ้ว และส่วนเบี่ยงเบนมาตรฐาน 0.6 นิ้ว

3. ค่ามัชฌิมของตัวอย่างในหน่วยมาตรฐานกำหนดได้โดย

$$Z = \frac{\bar{x} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{\bar{x} - 68.0}{0.6}$$

$$P(66.8 < \bar{x} < 68.3) = P\left(\frac{66.8 - 68.0}{0.6} < \frac{\bar{x} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} < \frac{68.3 - 68.0}{0.6}\right)$$



$= P(-2.0 < z < 0.5)$

= พื้นที่ใต้โค้งปกติระหว่าง z = -2.0 กับ z = 0.5

= (พื้นที่ระหว่าง z = -2 กับ Z = 0)

+ (พื้นที่ระหว่าง z = 0 กับ z = 0.5)

$= 0.4712 + 0.1915 = 0.6687$

ดังนั้น ตัวอย่างที่คาดหวัง $= (80)(0.6687) = 53$

4. $P(\bar{x} < 66.4) = P\left(\frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} < \frac{66.4 - 68.0}{0.6}\right)$

$= P(Z < -2.67)$

= พื้นที่ใต้โค้งปกติทางซ้ายของ $Z = -2.67$

= (พื้นที่ทางซ้ายของ $Z = 0$) - (พื้นที่ระหว่าง $Z = -2.67$ กับ $Z=0$)

= $0.5 - 0.4962 = 0.0038$

ดังนั้น ตัวอย่างที่คาดหวัง = (80)(0.0038) = **0.304** หรือศูนย์

7.9 การแจกแจงตัวอย่างของสัดส่วน

สัดส่วนคนส่วนมากมักจะนึกถึงตัวเลขที่เป็นเศษส่วนหรือร้อยละ แต่ถ้าเรามองให้ลึกซึ้งเข้าไปอีก เราจะพบว่าสัดส่วนก็คือค่ามัชฌิมนั่นเอง ดังตัวอย่างสมมติว่าเราโยนเหรียญ 3 อัน ถ้าออกหัวมีค่าเท่ากับ 1 ออกก้อยมีค่าเท่ากับ 0 สำหรับเหรียญแต่ละอันในการโยนนี้ออกหัว 1 ก้อย 2 มัชฌิมของหัวที่ออกไปในการโยนเหรียญ 3 อัน ก็คือ $\frac{1 + 0 + 0}{3} = \frac{1}{3}$ นั่นก็คือ สัดส่วนของหัวก็เท่ากับ $\frac{1}{3}$ หรือมีตัวอย่างที่ได้มาแบบสุ่มของแม่บ้าน 500 คน 320 คน ชอบใช้สบู่ชนิด A ขณะเดียวกับ 180 คน ชอบใช้สบู่ชนิด B ดังนั้น สัดส่วนของแม่บ้านที่ชอบใช้สบู่ชนิด A

$$\frac{x}{n} = \frac{320}{500} = 0.64$$

ผลที่ได้จากการแจกแจงตัวอย่างของค่ามัชฌิมก็นำมาใช้ได้ในการแจกแจงตัวอย่างของสัดส่วน

ให้ p เป็นสัดส่วนของประชากรหรือค่ามัชฌิมของประชากร

$\frac{x}{n}$ เป็นสัดส่วนของตัวอย่างหรือค่ามัชฌิมของสัดส่วน

ดังนั้น มัชฌิมของประชากรและส่วนเบี่ยงเบนมาตรฐานของประชากร ก็คือ

$$\mu = p; \ \sigma = \sqrt{p(1-p)} \qquad \text{..........(7.9.1)}$$

การแจกแจงตัวอย่างของสัดส่วนซึ่งมีมัชฌิม μ_p หรือ $E(\frac{x}{n})$ และส่วนเบี่ยงเบนมาตรฐาน σ_p หรือ $\sqrt{V(\frac{x}{n})}$ ก็กำหนดให้

$$\mu_p = p \ ; \ \sigma_p = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{pq}{n}} \qquad \text{.........(7.9.2)}$$

ในเมื่อ $q = 1 - p$, n คือขนาดตัวอย่าง ประชากรมีการแจกแจงทวินาม สำหรับค่าของ $n > 30$ การแจกแจงตัวอย่างก็มีลักษณะใกล้เคียงกับการแจกแจงปกติ

ข้อสังเกต

สมการ (7.9.2) คำนวณหามาได้โดยการหารมัชฌิมเลขคณิตและส่วนเบี่ยงเบนมาตรฐาน (np และ $\sqrt{npq}$ ของการแจกแจงทวินามด้วย n (ดู บทที่ 6) และประชากรเป็นแบบจำกัด แต่วิธีการเลือกเป็นแบบเลือกแล้วใส่คืน ในกรณีวิธีการเลือกเป็นแบบเลือกแล้วไม่ใส่คืน ค่าส่วนเบี่ยงเบนมาตรฐาน σ_p ก็ใช้สูตร

$$\sigma_p = \sqrt{\frac{pq}{n}} \sqrt{\frac{N-n}{N-1}} \qquad \text{.......(7.9.3)}$$

จากทฤษฎีขีดจำกัดส่วนกลางเมื่อขนาดตัวอย่าง n เข้าใกล้อนันต์

$$Z = \frac{\frac{x}{n} - p}{\sqrt{\frac{pq}{n}}}$$

ก็จะมีการแจกแจงปกติมาตรฐาน คือ มีมัชฌิมเลขคณิตศูนย์และส่วนเบี่ยงเบนมาตรฐานหนึ่ง

ตัวอย่าง

เครื่องจักรผลิตสกรูพบว่ามีชำรุด 2% จงคำนวณหาความน่าจะเป็นที่ผลิตสกรู 400 ตัว จะมีสกรูชำรุด

1. 3% หรือมากกว่า
2. 2% หรือน้อยกว่า

วิธีทำ

$$p = 0.02;\ \sqrt{V(x/n)} = \sqrt{pq/n} = \sqrt{(.02)(.98)/400}$$

$$= .14/20 = .007$$

1. การแปลงตัวแปรไม่ต่อเนื่องให้เป็นตัวแปรต่อเนื่อง จะต้องบวกหรือลบด้วย $1/2n = 1/800 = .00125, \frac{x}{n} \geqslant .03$

$$Z = \frac{(\frac{x}{n} \pm \frac{1}{2n}) - p}{\sqrt{\frac{pq}{n}}}$$

$$p\left[(\frac{x}{n} - \frac{1}{2n}) \geqslant .03 - .00125\right]$$

$$= P\left[\frac{(\frac{x}{n} - \frac{1}{2n}) - p}{\sqrt{\frac{pq}{n}}} \geqslant \frac{(.03 - .00125) - .02}{.007}\right]$$

$= P[Z \geqslant 1.25]$

= พื้นที่ใต้โค้งปกติทางขวาของ $Z = 1.25$

= (พื้นที่ทางขวาของ $Z = 0$) – (พื้นที่ระหว่าง $Z = 0$ กับ $Z = 1.25$)

$= 0.5 - .3944 = .1056$

ความน่าจะเป็นที่ผลิตสกรู 400 ตัว จะมีสกรูชำรุด 3% หรือมากกว่าเป็น 0.1056

2.

$$P\left[\frac{x}{n} \leqslant .02\right] = P\left[\frac{(\frac{x}{n} + \frac{1}{2n}) - p}{\sqrt{\frac{pq}{n}}} \leqslant \frac{.02 + .00125 - .02}{.009}\right]$$

$= P[Z \leqslant 0.18]$

= พื้นที่ภายใต้โค้งปกติทางซ้ายของ $Z = 0.18$

$= \quad$ (พื้นที่ทางซ้ายของ $Z = 0$) + (พื้นที่ระหว่าง $Z = 0$ กับ $Z = 0.18$)

$= \quad 0.5000 + .0714 = .5714$

ความน่าจะเป็นที่ผลิตสกรู 400 ตัว จะมีสกรูชำรุด 2% หรือน้อยกว่าเป็น 0.5714

7.10 การแจกแจงตัวอย่างของผลต่างและผลบวกของตัวสถิติ

สมมติว่ามีสองประชากรที่กำหนดให้ สำหรับแต่ละตัวอย่างขนาด n_1 เลือกจากประชากรที่หนึ่งคำนวณค่าสถิติ s_1 ซึ่งให้การแจกแจงตัวอย่างมีมัชฌิมเลขคณิตและส่วนเบี่ยงเบนมาตรฐานเป็น μ_{s_1} กับ σ_{s_1} ตามลำดับ ในทำนองเดียวกันแต่ละตัวอย่างขนาด n_2 เลือกจากประชากรที่สองคำนวณค่าสถิติได้ s_2 ซึ่งมีการแจกแจงตัวอย่างมีมัชฌิมเลขคณิตกับส่วนเบี่ยงเบนมาตรฐานเป็น μ_{s_2} กับ σ_{s_2} จากส่วนประกอบที่อาจเป็นไปได้ทั้งหมดของตัวอย่างเหล่านี้ของสองประชากร เราสามารถคำนวณหาการแจกแจงของผลต่าง ๆ $s_1 - s_2$ ซึ่งเรียกว่าการแจกแจงตัวอย่างของผลต่างของตัวสถิติมีมัชฌิมเลขคณิตและส่วนเบี่ยงเบนมาตรฐานแสดงได้เป็น $\mu_{s_1-s_2}$ กับ $\sigma_{s_1-s_2}$ ตามลำดับ นั่นคือ

$$\mu_{s_1-s_2} = \mu_{s_1} - \mu_{s_2} \text{ และ } \sigma_{s_1-s_2} = \sqrt{\sigma_{s_1}^2 + \sigma_{s_2}^2} \qquad \text{.........(7.10.1)}$$

ตัวอย่างที่ได้รับเลือกมานี้มีความอิสระกัน

ถ้าหากว่า s_1 กับ s_2 คือมัชฌิมเลขคณิตตัวอย่างจากสองประชากรแสดงได้เป็น $\bar{x}$ กับ $\bar{y}$ แล้วการแจกแจงตัวอย่างของผลต่างของมัชฌิมเลขคณิตตัวอย่างสำหรับประชากรแบบไม่จำกัดพร้อมด้วยมัชฌิมเลขคณิตและส่วนเบี่ยงเบนมาตรฐาน μ_1, σ_1 กับ μ_2, σ_2 ตามลำดับนั่นคือ

$$\mu_{\bar{x}-\bar{y}} = \mu_{\bar{x}} - \mu_{\bar{x}} = \mu_1 - \mu_2$$

$$\sigma_{\bar{x}-\bar{y}} = \sqrt{\sigma_{\bar{x}}^2 + \sigma_{\bar{y}}^2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad \text{..........(7.10.2)}$$

จากผลลัพธ์นี้สามารถหาการแจกแจงตัวอย่างของผลต่างของสัดส่วนจากสองประชากรที่มีการแจกแจงทวินาม พร้อมด้วยพารามิเตอร์ p_1, q_1 กับ p_2, q_2 ตามลำดับ ในกรณีนี้ s_1 กับ s_2 คือสัดส่วนของความสำเร็จ $\frac{x_1}{n_1}$ กับ $\frac{x_2}{n_2}$ สมการ (7.10.1) ให้ผล

$$\mu_{p_1-p_2} = E\left(\frac{x_1}{n_1} - \frac{x_2}{n_2}\right) = p_1 - p_2 \qquad \text{..........(7.10.3)}$$

$$\sigma_{p_1-p_2} = \sqrt{V\left(\frac{x_1}{n_1} - \frac{x_2}{n_2}\right)} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

ถ้าหากว่า n_1 และ n_2 ใหญ่พอ ($n_1, n_2 > 30$) การแจกแจงตัวอย่างของผลต่างมัชฌิมหรือสัดส่วนจะมีการแจกแจงเข้าใกล้ปกติ

สำหรับการแจกแจงตัวอย่างของผลบวกของสองตัวสถิติจะมีมัชฌิมและส่วนเบี่ยงเบนมาตรฐานของการแจกแจงนี้กำหนดได้โดย

$$\mu_{s_1+s_2} = \mu_{s_1} + \mu_{s_2} \text{ และ } \sigma_{s_1+s_2} = \sqrt{\sigma_{s_1}^2 + \sigma_{s_2}^2} \qquad \text{..........(7.10.6)}$$

จะเห็นได้ว่าถ้า n_1 และ n_2 มีความอิสระกัน ค่าส่วนเบี่ยงเบนมาตรฐานของการแจกแจงตัวอย่างของผลต่างหรือผลบวกของสองตัวสถิติจะเท่ากัน เว้นแต่ว่า n_1 และ n_2 ไม่มีความอิสระเท่านั้น

ตัวอย่าง หลอดไฟฟ้าที่ผลิตจากบริษัท ก. มีอายุเฉลี่ย 1400 ชั่วโมง พร้อมด้วยส่วนเบี่ยงเบนมาตรฐาน 200 ชั่วโมง ขณะที่หลอดไฟฟ้าที่ผลิตจากบริษัท ข. มีอายุเฉลี่ย 1200 ชั่วโมง พร้อมด้วยส่วนเบี่ยงเบนมาตรฐาน 100 ชั่วโมง ถ้าหากว่าสุ่มตัวอย่างขนาด 125 หลอด จากแต่ละบริษัทเพื่อทดสอบ จงหาความน่าจะเป็นที่อายุเฉลี่ยของหลอดไฟจากบริษัท ก. มากกว่าอายุเฉลี่ยของหลอดไฟบริษัท ข.

1. อย่างน้อย 160 ชั่วโมง　　　　2. มากกว่า 250 ชั่วโมง

วิธีทำ ให้ $\bar{x}_ก$ และ $\bar{x}_ข$ เป็นอายุเฉลี่ยของตัวอย่าง ก. และ ข. ตามลำดับแล้ว

$$\mu_ก - \mu_ข = 1400 - 1200 = 200 \text{ ชั่วโมง}$$

$$\sigma_{\bar{x}_ก-\bar{x}_ข} = \sqrt{\frac{\sigma_ก^2}{n_ก} + \frac{\sigma_ข^2}{n_ข}} = \sqrt{\frac{(200)^2}{125} + \frac{(100)^2}{125}}$$

$$= 20 \text{ ชั่วโมง}$$

1. $P[(\bar{x}_ก - \bar{x}_ข) \geqslant 160] = P\left[\frac{(\bar{x}_ก - \bar{x}_ข) - (\mu_ก - \mu_ข)}{\sigma_{\bar{x}_ก - \bar{x}_ข}} \geqslant \frac{160 - 200}{20}\right]$

$$= P(Z \geqslant -2)$$

= พื้นที่ภายใต้โค้งปกติทางขวาของ $Z = -2$

= (พื้นที่ระหว่าง $Z = -2$ กับ $Z = 0$) +
(พื้นที่ทางขวาของ $Z = 0$)

$= 0.4772 + 0.5000$

$= .9772$

ความน่าจะเป็นที่อายุเฉลี่ยของหลอดไฟจากบริษัท ก. มากกว่าอายุเฉลี่ยของหลอดไฟบริษัท ข. อย่างน้อย 160 ชั่วโมงเป็น 0.9772

2. $P\left[(\bar{x}_ก - \bar{x}_ข) > 250\right] = P\left[\frac{(\bar{x}_ก - \bar{x}_ข) - (\mu_ก - \mu_ข)}{\sigma_{\bar{x}_ก - \bar{x}_ข}} > \frac{250 - 200}{20}\right]$

$= P(Z > 2.5)$

= พื้นที่ใต้โค้งปกติทางขวาของ $Z = 2.5$

= (พื้นที่ทางขวาของ $Z = 0$) –
(พื้นที่ระหว่าง $Z = 0$ กับ $Z = 2.5$)

$= 0.5000 - 0.4938$

$= 0.0062$

ความน่าจะเป็นที่อายุเฉลี่ยของหลอดไฟจากบริษัท ก. มากกว่าอายุเฉลี่ยของหลอดไฟบริษัท ข. มากกว่า 250 ชั่วโมง เป็น 0.0062

ตัวอย่าง นาย ก. และ นาย ข. เล่นเกมโยนเหรียญกันคนละ 50 ครั้ง นาย ก. จะชนะเกมได้ ถ้าหากว่าเขาโยนเหรียญปรากฏเป็นหัว 5 ครั้ง หรือมากกว่านาย ข. โยน นอกจากนั้น นาย ข. ชนะ จงหาความน่าจะเป็นที่ นาย ก. ชนะ

วิธีทำ $\frac{x_ก}{n_ก}$ และ $\frac{x_ข}{n_ข}$ เป็นสัดส่วนที่ปรากฏเป็นหัวจากการโยนของนาย ก. และ นาย ข. ถ้าหากว่าเหรียญนั้นสมดุล ความน่าจะเป็น p ที่ปรากฏเป็นหัวคือ $1/2$

$$\mu_{p_1-p_2} = E\left(\frac{x_ก}{n_ก} - \frac{x_ข}{n_ข}\right) = p_ก - p_ข = (1/2) - (1/2) = 0$$

$$\sigma_{p_1-p_2} = \sqrt{V\left(\frac{x_\text{ก}}{n_\text{ก}} - \frac{x_\text{ข}}{n_\text{ข}}\right)} = \sqrt{\frac{p_\text{ก}\, q_\text{ก}}{n_\text{ก}} + \frac{p_\text{ข}\, q_\text{ข}}{n_\text{ข}}} = \sqrt{\frac{(\frac{1}{2})(\frac{1}{2})}{50} + \frac{(\frac{1}{2})(\frac{1}{2})}{50}}$$

$$= \sqrt{\frac{2(\frac{1}{2})(\frac{1}{2})}{50}} = 0.10$$

ตัวแปรมาตรฐานสำหรับผลต่างของสัดส่วนคือ

$$Z = \frac{\left(\frac{x_\text{ก}}{n_\text{ก}} - \frac{x_\text{ข}}{n_\text{ข}}\right) - (p_\text{ก} - p_\text{ข})}{\sqrt{\frac{p_\text{ก}\, q_\text{ก}}{n_\text{ก}} + \frac{p_\text{ข}\, q_\text{ข}}{n_\text{ข}}}}$$

ตัวแปรต่อเนื่องของ 5 ครั้งหรือมากกว่า คือ 4.5 หรือมากกว่า ดังนั้นผลต่างของสัดส่วนควรเป็น $\frac{4.5}{50} = 0.09$ หรือมากกว่า

$$P\left[\left(\frac{x_\text{ก}}{n_\text{ก}} - \frac{x_\text{ข}}{n_\text{ข}}\right) - \frac{1}{2n} \geqslant 0.09\right]$$

$$= P\left[\frac{\left(\frac{x_\text{ก}}{n_\text{ก}} - \frac{x_\text{ข}}{n_\text{ข}}\right) - \frac{1}{2n} - (p_\text{ก} - p_\text{ข})}{\sqrt{\frac{p_\text{ก}\, q_\text{ก}}{n_\text{ก}} + \frac{p_\text{ข}\, q_\text{ข}}{n_\text{ข}}}} \geqslant \frac{.09 - 0}{.10}\right]$$

$= P\,(Z \geqslant 0.9)$

$=$ พื้นที่ใต้โค้งปกติทางขวาของ $Z = 0.9$

$=$ (พื้นที่ทางขวาของ $Z = 0$) $-$

(พื้นที่ระหว่าง $Z = 0$ กับ $Z = .9$)

$= 0.5000 - .3159$

$= 0.1841$

ความน่าจะเป็นที่ นาย ก. ชนะเป็น 0.1841

7.11 ความคลาดเคลื่อนมาตรฐาน

ส่วนเบี่ยงเบนมาตรฐานของการแจกแจงตัวอย่างของตัวสถิติส่วนมากเรียกว่าความคลาดเคลื่อนมาตรฐาน ในตารางที่ 7.1 เรามีรายชื่อความคลาดเคลื่อนมาตรฐานของการแจกแจงตัวอย่างสำหรับตัวสถิติต่าง ๆ ภายใต้เงื่อนไขของการสุ่มตัวอย่างจากประชากรแบบไม่จำกัดหรือการเลือกตัวอย่างแบบเลือกแล้วใส่คืนจากประชากรแบบจำกัด พร้อมทั้งหมายเหตุกำหนดเงื่อนไขภายใต้ผลลัพธ์ที่มีเหตุผลและข้อความที่ตรงกับเนื้อหาอื่น ๆ

ปริมาณ μ, σ, p, μ_r และ $\bar{x}$, s, $\frac{x}{n}$, m_r คือ มัชฌิมเลขคณิต ส่วนเบี่ยงเบนมาตรฐาน สัดส่วน โมเมนต์ที่ r รอบมัชฌิมเลขคณิตของประชากร และตัวอย่างตามลำดับ

ถ้าหากว่าขนาดตัวอย่าง n ใหญ่พอ การแจกแจงตัวอย่างเข้าใกล้ปกติ เมื่อไร $n < 30$ เราเรียกตัวอย่างว่าตัวอย่างขนาดเล็ก กรณีที่ไม่ทราบตัวพารามิเตอร์ σ, p, μ เราใช้ตัวสถิติที่สมนัยกัน s, $\frac{x}{n}$, $\bar{x}$ มาประมาณค่าได้

ตารางที่ 7.1

ความคลาดเคลื่อนมาตรฐานสำหรับบางการแจกแจงตัวอย่าง

การแจกแจงตัวอย่าง	ความคลาดเคลื่อนมาตรฐาน	หมายเหตุ
มัชฌิมเลขคณิต	$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$	นี้เป็นจริงสำหรับตัวอย่างขนาดใหญ่หรือเล็ก การแจกแจง-ตัวอย่างของมัชฌิมเข้าใกล้ปกติเมื่อ $n > 30$ แม้ว่าประชากรไม่ปกติ $\mu_{\bar{X}} = \mu$, มัชฌิมเลขคณิตของประชากรทุก ๆ กรณี
สัดส่วน	$\sigma_p = \sqrt{\frac{pq}{n}}$	$\mu_p = p$ ทุกกรณี
ส่วนเบี่ยงเบนมาตรฐาน	(1) $\sigma_s = \frac{\sigma}{\sqrt{2n}}$ (2) $\sigma_s = \sqrt{\frac{\mu_4 - \mu_2^2}{4n\mu_2}}$	สำหรับ $n \geqslant 100$ การแจกแจงตัวอย่างของ s เข้าใกล้ปกติ σ_s ตามข้อ (1) เท่านั้น

การแจกแจงตัวอย่าง	ความคลาดเคลื่อนมาตรฐาน	หมายเหตุ
		ถ้าหากว่าประชากรปกติหรือเข้าใกล้ปกติ ถ้าประชากรไม่ปกติให้ใช้ (2) สังเกตว่าข้อ (2) ลดรูปเป็นข้อ (1) เมื่อ $\mu_2 = \sigma^2$ และ $\mu_4 = 3\sigma^4$ กรณีซึ่งเป็นจริงสำหรับประชากรปกติ $n \geqslant 100$; $\mu_s = \sigma$ (เกือบเท่ากัน)
มัธยฐาน	$\sigma_{med} = \sigma\sqrt{\frac{\pi}{2n}}$	สำหรับ $n > 30$ การแจกแจงตัวอย่างของมัธยฐานเข้าใกล้ปกติมาก ผลลัพธ์ที่กำหนดเป็นจริง ถ้าประชากร เป็นปกติเท่านั้น $\mu_{med} = \mu$
ความแปรปรวน	(1) $\sigma_s^2 = \sigma^2\sqrt{\frac{2}{n}}$ (2) $\sigma_s^2 = \sqrt{\frac{\mu_4 - \mu_2^2}{n}}$	เงื่อนไขก็เหมือนกับส่วนเบี่ยงเบนมาตรฐาน สังเกตว่าข้อ (2) จะลดรูปเป็นข้อ (1) ในกรณีประชากรเป็นปกติ $\mu_s^2 = \frac{\sigma^2(n-1)}{n}$ ซึ่งจะเข้าใกล้ σ^2 มาก เมื่อ n มีค่ามาก

แบบฝึกหัดที่ 7

1) ประชากรประกอบด้วยเลข 5 จำนวน 2,3,6,8,11 เลือกตัวอย่างให้มีขนาดเท่ากับ 2 โดยวิธี without replacement จากประชากรนี้จงหา

ก. มัชฌิมเลขคณิตของประชากร

ข. ส่วนเบี่ยงเบนมาตรฐานของประชากร

ค. มัชฌิมเลขคณิตของมัชฌิมเลขคณิตของตัวอย่าง

ง. ส่วนเบี่ยงเบนมาตรฐานของมัชฌิมเลขคณิตของตัวอย่าง

(ก. 6.0, ข. 3.29, ค. 6.0, ง. 2.32)

2) จากปัญหาข้อที่ 1 เลือกตัวอย่างให้มีขนาดเท่ากับ 2 โดยวิธี without replacement จากประชากร จงหา

ก. มัชฌิมเลขคณิตของมัชฌิมของตัวอย่าง (6.0)

ข. ส่วนเบี่ยงเบนมาตรฐานของมัชฌิมเลขคณิตของตัวอย่าง (2.01)

3) เลือกตัวอย่างที่มีขนาดเท่ากับ 2 โดยวิธี with replacement จากประชากร 1,2,3,4,5 จงหามัชฌิมเลขคณิตของแต่ละตัวอย่าง ทำตารางความถี่ของมัชฌิมเลขคณิตของแต่ละตัวอย่างเขียน histogram สำหรับประชากรและสำหรับการแจกแจงของมัชฌิมเลขคณิตของตัวอย่างและเปรียบเทียบรูปร่างของ histogram.

4) ขนาดของตัวอย่างเท่ากับ 25 และความคลาดเคลื่อนมาตรฐานของมัชฌิมเลขคณิต เป็น 2.4 ขนาดของตัวอย่างจะเป็นเท่าใดถ้าความคลาดเคลื่อนลดลง 1.5 (177)

5) สมมติว่าความสูงของนักศึกษาชาย 3,000 คน ของมหาวิทยาลัยแห่งหนึ่งเป็นการแจกแจงปกติซึ่งมีมัชฌิมเลขคณิต 68.0 นิ้วกับส่วนเบี่ยงเบนมาตรฐาน 3.0 นิ้ว ถ้าเลือก 80 ตัวอย่าง ประกอบด้วยตัวอย่างละ 25 คน จงคำนวณหามัชฌิมเลขคณิต ($\bar{x}_{\bar{X}}$) และส่วนเบี่ยงเบนมาตรฐาน ($\sigma_{\bar{X}}$) ของมัชฌิมเลขคณิตของตัวอย่างที่หวังว่าจะได้ในเมื่อการเลือกโดยวิธี

ก. with replacement ข. without replacement

(ก. 68.0 นิ้ว, 0.6 นิ้ว, ข. 68.0 นิ้ว, 0.6 นิ้ว)

6) จากปัญหาข้อที่ 5 จะมีกี่ตัวอย่างที่ท่านคาดหวังจะหาได้

ก. ระหว่างมัชฌิมเลขคณิต 66.8 นิ้ว กับ 68.3 นิ้ว (53)

ข. มัชฌิมเลขคณิตที่น้อยกว่า 66.4 นิ้ว (0)

7) ขนาดของตัวอย่างเท่ากับ 40 ที่เลือกจากประชากร ซึ่งมีมัชฌิม $\mu = 36.4$ กับส่วนเบี่ยงเบนมาตรฐาน $\sigma = 5.2$ จงหา probability ที่มัชฌิมของตัวอย่าง จะอยู่ระหว่าง 36 กับ 38(0.6623)

8) เลือกตัวอย่างขนาด 75 จากประชากรซึ่งมีมัชฌิมและส่วนเบี่ยงเบนมาตรฐาน คือ $\mu = 112$ กับ $\sigma = 10$ จงหา probability ที่มัชฌิมของตัวอย่างอยู่ระหว่าง 110.5 กับ 113.5 (.8064)

9) ลูกปืน (ball bearings) ห้าร้อยลูกมีน้ำหนักเฉลี่ย 5.02 ออนซ์ และส่วนเบี่ยงเบนมาตรฐาน 0.30 ออนซ์ จงหาความน่าจะเป็นที่ตัวอย่างสุ่มประกอบด้วยลูกปืน 100 ลูก ที่ได้เลือกจากลูกปืนกลุ่มนี้จะมีน้ำหนักรวม

ก. ระหว่าง 496 กับ 500 ออนซ์ (0.2164)

ข. มากกว่า 510 ออนซ์ (0.0015)

10) จงหาความน่าจะเป็นที่จะคลอดเด็ก 200 คนต่อไป (ก) จะเป็นเด็กชายน้อยกว่า 40% (ข) จะเป็นเด็กหญิงระหว่าง 43% ถึง 57% (ค) จะเป็นเด็กชายมากกว่า 54% สมมติว่าความน่าจะเป็นที่จะคลอดเป็นเด็กชายและหญิงมีโอกาสเท่า ๆ กัน

((ก) 0.0012 (ข) 0.9616 (ค) 0.1131)

11) จาก 1000 ตัวอย่าง โดยที่แต่ละตัวอย่างประกอบด้วยเด็ก 200 คน จงคำนวณหาจำนวนตัวอย่างที่ท่านคาดหวังว่า (ก) เป็นเด็กชายน้อยกว่า 40% (ข) เป็นเด็กหญิงระหว่าง 40% ถึง 60% (ค) เป็นเด็กหญิง 53% หรือมากกว่า ((ก) 1 (ข) 998 (ค) 215))

12) ผู้ผลิตสินค้าสำเร็จรูปคนหนึ่งส่งสินค้าออก 1000 กล่อง แต่ละกล่องบรรจุหลอดไฟฟ้า 100 หลอด ถ้าหากว่าหลอดไฟฟ้าโดยปกติเสีย 5% จงคำนวณหาจำนวนกล่องที่ท่านคาดหวังว่า (ก) หลอดไฟฟ้าดีน้อยกว่า 90 หลอด (ข) หลอดไฟฟ้าดี 98 หลอดหรือมากกว่า ((ก) 6 (ข) 125))

13) กล่องใบหนึ่งบรรจุลูกบอล 80 ลูก ซึ่งมีบอลแดง 60% และบอลขาว 40% จาก 50 ตัวอย่าง ซึ่งแต่ละตัวอย่างประกอบด้วยบอล 20 ลูก โดยเลือกแบบหยิบแล้วใส่คืนจากกล่องนี้จะมีจำนวนที่ตัวอย่างที่สามารถคาดหวังว่าประกอบด้วย (ก) จำนวนบอลแดงและขาวเท่ากัน (ข) บอลแดง 12 ลูก และบอลขาว 8 ลูก (ค) บอลแดง 8 ลูก และบอลขาว 12 ลูก (ง) บอลขาว 10 ลูกหรือมากกว่า

((ก) 6 (ข) 9 (ค) 2 (ง) 19)

14) A และ B เป็นลวดสองชนิด มีความทนต่อแรงดึงเฉลี่ย 4000 และ 4500 ปอนด์ ส่วนเบี่ยงเบนมาตรฐาน 300 และ 200 ปอนด์ตามลำดับ ถ้าหากว่าเอาลวด A มา 100 เส้น และ B มา 50 เส้น

เพื่อทดสอบ จงหาความน่าจะเป็นที่ความทนต่อแรงดึงเฉลี่ยของ B จะ (ก) มากกว่า A อย่างน้อย 600 ปอนด์ (ข) มากกว่า A อย่างน้อย 450 ปอนด์ ((ก) 0.0077 (ข) 0.8869)

15) การเลือกตั้งแสดงว่าคู่แข่งคนหนึ่งได้รับคะแนนเสียง 65% จงหาความน่าจะเป็นที่สองตัวอย่างโดยที่แต่ละตัวอย่างประกอบด้วยผู้ออกเสียง 200 คน แสดงคะแนนความแตกต่างมากกว่า 40% ในรูปของสัดส่วนที่ลงคะแนนให้คู่แข่งคนนั้น (0.0162)