การวัดค่าความเบ้และความโด่งด้วยโปรแกรมสำเร็จรูปชนิดต่าง ๆ

Measurement of Skewness and Kurtosis obtained by various package programs

วราฤทธิ์ พานิชกิจโกศลกุล

ภาควิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี

มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต ปทุมธานี 12121

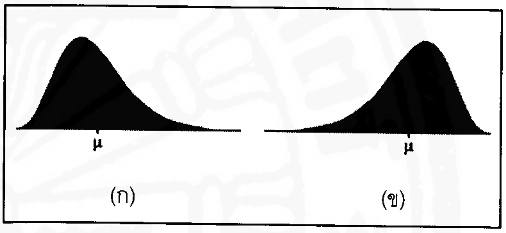
การวิเคราะห์ข้อมูลในงานวิจัยมีหลายระดับ เริ่มตั้งแต่การวิเคราะห์ข้อมูลเบื้องต้นและการวิเคราะห์ข้อมูลขั้นสูง การวิเคราะห์ข้อมูลเบื้องต้นประกอบด้วย การแจกแจงความถี่ การหาค่าร้อยละ การหาค่าสัดส่วน การหาค่าอัตราส่วน การหาค่ากลาง การหาค่าวัดการกระจาย และการหาค่าวัดรูปร่างของข้อมูล ซึ่งการหาค่าวัดรูปร่างของข้อมูลเชิงปริมาณ ได้แก่ การหาค่าวัดความเบ้ (Measures of Skewness) และการหาค่าวัดความโด่ง (Measures of Kurtosis) ของข้อมูล ค่าวัดความเบ้เป็นค่าที่ใช้วัดลักษณะของเส้นโค้งหรือข้อมูลว่าเบ้หรือไม่ ส่วนค่าวัดความโด่งเป็นค่าที่ใช้วัดความโด่งของเส้นโค้งหรือข้อมูล ซึ่งค่าวัดความเบ้และความโด่งนิยมวัดในรูปของสัมประสิทธิ์ ซึ่งเป็นค่าที่ไม่มีหน่วย

1. ค่าวัดความเบ้ (Measures of Skewness)

ข้อมูลที่มีการแจกแจงสมมาตรนั้น ค่าเฉลี่ยเลขคณิต มัธยฐาน และฐานนิยม จะมีค่าเท่ากัน ส่วนข้อมูลที่มีการแจกแจงไม่สมมาตร คือเบ้ไปทางใดทางหนึ่ง ถ้าข้อมูลส่วนใหญ่มีค่าน้อยหรือข้อมูลส่วนใหญ่อยู่ด้านซ้ายมือของการแจกแจง ลักษณะเช่นนี้เรียกว่า เบ้ขวา มีค่าวัดความเบ้เป็นบวก (Positively skewed) ดังภาพที่ 1 (ก) แต่ถ้าข้อมูลส่วนใหญ่มีค่ามาก หรือข้อมูลส่วนใหญ่อยู่ด้านขวามือของการแจกแจง จะเรียกว่า เบ้ซ้าย มีค่าวัดความเบ้เป็นลบ (Negatively skewed) ดังภาพที่ 1 (ข) การหาค่าวัดความเบ้มีหลายวิธี ในที่นี้ใช้วิธีโมเมนต์ (Moment method) ดังนี้ [1]

$$\alpha_3 = \frac{\mu_3}{\mu_2^{3/2}} \qquad (1)$$

เมื่อ μ_i คือ โมเมนต์ประชากรที่ i รอบค่าเฉลี่ย มีค่าเท่ากับ $E(X-\mu)^i$



ภาพที่ 1 (ก) ข้อมูลมีการแจกแจงแบบเบ้ขวา
(ข) ข้อมูลมีการแจกแจงแบบเบ้ซ้าย

ค่า α_3 เป็นค่าพารามิเตอร์ของประชากร ซึ่งค่าประมาณของ α_3 คือ

$$\hat{\alpha}_3 = \frac{m_3}{m_2^{3/2}} \qquad (2)$$

เมื่อ m_i คือ โมเมนต์ตัวอย่างที่ i รอบค่าเฉลี่ย มีค่าเท่ากับ

$$m_i = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^i}{n}$$ เมื่อ n คือ ขนาดตัวอย่าง

ตัวประมาณ m_i เป็นตัวประมาณโมเมนต์ที่คงเส้นคงวา (Consistent estimates) แต่เป็นตัวประมาณโมเมนต์ที่เอนเอียง (Biased estimates) คราเมอร์ [2] แสดงตัวประมาณโมเมนต์ที่ไม่เอนเอียงของ μ_2, μ_3 และ μ_4 ดังนี้

$$u_2 = \frac{n}{n-1}m_2 = c_2 m_2 \qquad (3)$$

$$u_3 = \frac{n^2}{(n-1)(n-2)} m_3$$

$$= c_3 m_3 \quad (4)$$

$$u_4 = \frac{n(n^2-2n+3)}{(n-1)(n-2)(n-3)} m_4 - \frac{3n(2n-3)}{(n-1)(n-2)(n-3)} m_2^2$$

$$= c_4 m_4 - c_5 m_2^2 \quad (5)$$

เมื่อ u_i เป็นตัวประมาณโมเมนต์ที่ไม่เอนเอียงของ μ_i , $i = 2, 3, 4$

ดังนั้นเมื่อนำตัวประมาณโมเมนต์ที่ไม่เอนเอียง u_i ไปแทนค่าลงในสมการที่ (1) จะได้

$$\hat{\alpha}_3 = \frac{u_3}{u_2^{3/2}} \quad (6)$$

โดยทั่วไปแล้ว $\hat{\alpha}_3$ เป็นตัวประมาณที่เอนเอียง ถึงแม้ว่าจะนำค่าประมาณโมเมนต์ที่ไม่เอนเอียง u_i มาแทนค่าแล้วก็ตาม โจเรสค็อก [3] ได้แสดงตัวประมาณที่ไม่เอนเอียงของ α_3 เท่ากับ

$$\hat{\hat{\alpha}}_3 = \frac{\sqrt{n(n-1)}}{n-2} \cdot \frac{m_3}{m_2^{3/2}}$$

$$= \frac{n\sum_{i=1}^{n}(X_i - \bar{X})^3}{(n-1)(n-2)S^3} \quad (7)$$

2. ค่าวัดความโด่ง (Measures of Kurtosis)

ความโด่ง หมายถึง ความสูงของเส้นโค้งการแจกแจงว่าสูงโด่ง หรือแบนราบ ความโด่งของเส้นโค้งการแจกแจงมี 3 ลักษณะคือ [4]

1) การแจกแจงที่มีลักษณะเส้นโค้งปกติ เรียกว่า Mesokurtic

2) การแจกแจงที่มีลักษณะแบน เรียกว่า Platykurtic

3) การแจกแจงที่มีลักษณะสูง เรียกว่า Leptokurtic

การหาค่าวัดความโด่งมีหลายวิธี ในที่นี้ใช้วิธีโมเมนต์ดังนี้

$$\alpha_4 = \frac{\mu_4}{\mu_2^2} - 3 \quad (8)$$

เมื่อ μ_i คือ โมเมนต์ประชากรที่ i รอบค่าเฉลี่ย มีค่าเท่ากับ $E(X-\mu)^i$

ค่า α_4 เป็นค่าพารามิเตอร์ของประชากร ซึ่งค่าประมาณของ α_4 คือ

$$\hat{\alpha}_4 = \frac{m_4}{m_2^2} - 3 \quad (9)$$

เมื่อ m_i คือ โมเมนต์ตัวอย่างที่ i รอบค่าเฉลี่ย มีค่าเท่ากับ

$$m_i = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^i}{n}$$ เมื่อ n คือ ขนาดตัวอย่าง

ตัวประมาณ m_i เป็นตัวประมาณที่คงเส้นคงวา แต่เป็นตัวประมาณที่เอนเอียง เมื่อนำตัวประมาณที่ไม่เอนเอียง u_i ไปแทนค่าลงในสมการที่ (9) จะได้

$$\hat{\alpha}_4 = \frac{u_4}{u_2^2} - 3 \quad (10)$$

โดยทั่วไปแล้ว $\hat{\alpha}_4$ เป็นตัวประมาณที่เอนเอียง ถึงแม้ว่าจะนำค่าประมาณโมเมนต์ที่ไม่เอนเอียง u_i มาแทนค่าแล้วก็ตาม โจเรสค็อก [3] ได้แสดงตัวประมาณที่ไม่เอนเอียงของ α_4 เท่ากับ

$$\hat{\hat{\alpha}}_4 = \frac{n-1}{(n-2)(n-3)}\left[(n+1)\frac{m_4}{m_2^2} - 3(n-1)\right]$$

$$= \frac{\frac{n(n+1)}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^4 - 3\left[\sum_{i=1}^{n}(X_i-\bar{X})^2\right]^2}{(n-2)(n-3)S^4}$$

(11)

3. การหาค่าวัดความเบ้และความโด่งด้วยโปรแกรมสำเร็จรูปชนิดต่างๆ

ในปัจจุบันนักวิจัยนิยมใช้โปรแกรมสำเร็จรูปในการประมวลผล ค่าวัดความเบ้และความโด่งที่คำนวณจากโปรแกรมสำเร็จรูปชนิดต่าง ๆ นั้นอาจมีค่าเท่ากันหรือไม่เท่ากันก็ได้ สมมติว่าข้อมูลชุดหนึ่งมีจำนวน 30 ค่า ดังนี้

30 34 39 40 47 49 50 50 51 52

52 53 54 54 54 54 55 56 57 58

58 59 60 61 63 64 64 69 70 71

ในที่นี้จะคำนวณค่าวัดความเบ้และความโด่งของข้อมูล โดยใช้สูตรที่ (7) และ (11)

จากข้อมูลข้างต้น คำนวณค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน ค่าวัดความเบ้ และค่าวัดความโด่ง ได้ดังนี้

$\overline{X} = 54.267$, $S = 9.695$

ค่าวัดความเบ้ เท่ากับ

$$\hat{\alpha}_3 = \frac{n\sum_{i=1}^{n}(X_i - \overline{X})^3}{(n-1)(n-2)\,S^3}$$

$$= \frac{30(-14{,}803.3)}{(30-1)(30-2)(9.695)^3}$$

$$= -0.600$$

ค่าวัดความโด่ง เท่ากับ

$$\hat{\alpha}_4 = \frac{\frac{n(n+1)}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^4 - 3\left[\sum_{i=1}^{n}(X_i - \overline{X})^2\right]^2}{(n-2)(n-3)\,S^4}$$

$$= \frac{\frac{30(30+1)}{30-1}.(830{,}270.958) - 3(2{,}725.867)^2}{(30-2)(30-3)(9.695)^4}$$

$$= 0.649$$

ส่วนขั้นตอนหรือคำสั่งในการหาค่าวัดความเบ้และความโด่งของแต่ละโปรแกรม มีรายละเอียดดังนี้

1) การวัดความเบ้และความโด่งด้วยโปรแกรม SPSS Version 11.0 โดยใช้คำสั่ง DESCRIPTIVES ดังนี้ [5]

```
DESCRIPTIVES
VARIABLES=X
/STATISTICS=KURTOSIS SKEWNESS.
```

ค่าวัดความเบ้ เท่ากับ –0.600 และค่าวัดความโด่ง เท่ากับ 0.649 แสดงดังภาพที่ 2

Descriptive Statistics

	N	Skewness		Kurtosis	
	Statistic	Statistic	Std. Error	Statistic	Std. Error
X	30	-.600	.427	.649	.833
Valid N (listwise)	30				

ภาพที่ 2 ค่าวัดความเบ้และความโด่งที่คำนวณจากโปรแกรม SPSS

2) การวัดความเบ้และความโด่งด้วยโปรแกรม SAS Release 8.2 โดยใช้คำสั่ง PROC UNIVARIATE ดังนี้ [6]

```
DATA;
  INPUT X@@;
  CARDS;
30 34 39 40 47 49 50 50 51 52 52 53 54 54 54
54 55 56 57 58 58 59 60 61 63 64 64 69 70 71
;
PROC UNIVARIATE;
VAR X;
RUN;
```

ค่าวัดความเบ้ เท่ากับ –0.600 และค่าวัดความโด่ง เท่ากับ 0.649 แสดงดังภาพที่ 3

```
                 The UNIVARIATE Procedure
                      Variable:  X

                         Moments

N                          30    Sum Weights                 30
Mean               54.2666667    Sum Observations          1628
Std Deviation       9.6951226    Variance            93.9954023
Skewness           -0.6001541    Kurtosis            0.64899658
Uncorrected SS          91072    Corrected SS        2725.86667
Coeff Variation     17.865705    Std Error Mean      1.77007912
```

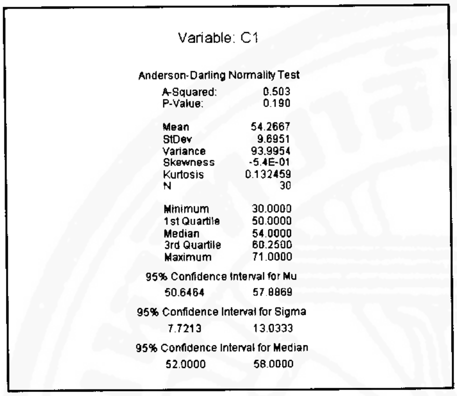
ภาพที่ 3 ค่าวัดความเบ้และความโด่งที่คำนวณจากโปรแกรม SAS

3) การวัดความเบ้และความโด่งด้วยโปรแกรม MINITAB Release 11.12 มีขั้นตอนดังนี้ [7]

เลือกเมนูตามขั้นตอนดังต่อไปนี้ Stat → Basic Statistics → Descriptive Statistics เมื่อปรากฏกล่องข้อความ ให้เลือกตัวแปรใส่ในช่อง Variables จากนั้นคลิกที่ปุ่ม Graphs… เพื่อให้แสดงค่าสถิติขั้นสูง และกราฟ โดยคลิกให้

ปรากฏเครื่องหมายถูกหน้าช่อง Graphical Summary หลังจากนั้นคลิกที่ปุ่ม OK

ค่าวัดความเบ้ เท่ากับ –0.540 และค่าวัดความโด่งเท่ากับ 0.132 แสดงดังภาพที่ 4



ภาพที่ 4 ค่าวัดความเบ้และความโด่งที่คำนวณจากโปรแกรม MINITAB

4)การวัดความเบ้และความโด่งด้วยโปรแกรม Maple 6 โดยใช้คำสั่ง describe ดังนี้ [8]

```
with(stats):
describe[skewness]([30,34,39,40,47,49,50,50,51,52
,52,53,54,54,54,54,55,56,57,58,58,59,60,61,63,64,64
,69,70,71]);
describe[kurtosis]([30,34,39,40,47,49,50,50,51,52,
52,53,54,54,54,54,55,56,57,58,58,59,60,61,63,64,64,
69,70,71]);
```

ค่าวัดความเบ้ เท่ากับ –0.570 และค่าวัดความโด่งเท่ากับ 3.352

5) การวัดความเบ้และความโด่งด้วยโปรแกรม Microsoft Excel โดยใช้คำสั่ง SKEW และ KURT ดังนี้ [9]

```
=SKEW(A1:A30)
=KURT(A1:A30)
```

ค่าวัดความเบ้ เท่ากับ –0.600 และค่าวัดความโด่งเท่ากับ 0.649

จะเห็นได้ว่าค่าวัดความเบ้และความโด่งของข้อมูลที่คำนวณจากโปรแกรมสำเร็จชนิดต่างๆ บางโปรแกรมก็ให้ค่าเท่ากัน บางโปรแกรมก็ให้ค่าต่างกัน สรุปได้ดังนี้

ตารางที่ 1 ค่าวัดความเบ้และค่าวัดความโด่งของข้อมูลข้างต้นจำแนกตามโปรแกรมที่ใช้

โปรแกรม	ค่าวัดความเบ้	ค่าวัดความโด่ง
SPSS, SAS, Microsoft Excel	–0.600	0.649
MINITAB	–0.540	0.132
Maple	–0.570	3.352

ค่าวัดความเบ้และความโด่งของข้อมูลที่คำนวณจากโปแกรม SPSS, SAS และ Microsoft Excel มีค่าเท่ากัน คือ ค่าวัดความเบ้และความโด่ง เท่ากับ –0.600 และ 0.649 ตามลำดับ ซึ่งค่าที่คำนวณจากโปรแกรมทั้งสามตรงกับค่าที่คำนวณจากสูตร นั่นคือ เป็นค่าวัดความเบ้และความโด่งที่ไม่เอนเอียง

ส่วนโปรแกรม MINITAB ให้ค่าวัดความเบ้และความโด่ง เท่ากับ -0.540 และ 0.132 ตามลำดับ และโปรแกรม Maple ให้ค่าวัดความเบ้และความโด่ง เท่ากับ –0.570 และ 3.352 ตามลำดับ ซึ่งค่าวัดความเบ้และความโด่งที่คำนวณจากโปรแกรม MINITAB และ Maple มีค่าแตกต่างจากค่าที่คำนวณจากโปรแกรม SPSS, SAS และ Microsoft Excel

4. สูตรที่ใช้ในการคำนวณค่าวัดความเบ้และความโด่งของโปรแกรมสำเร็จรูปชนิดต่างๆ

สูตรที่ใช้ในการคำนวณค่าวัดความเบ้และความโด่งของโปรแกรม Maple เป็นสูตรที่คำนวณโดยใช้วิธีโมเมนต์ ซึ่งเป็นตัวประมาณที่เอนเอียง คำนวณจาก

$$\hat{\alpha}_3 = \frac{m_3}{m_2^{3/2}}$$

$$= \frac{\dfrac{\sum_{i=1}^{n}(X_i-\bar{X})^3}{n}}{\dfrac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n}\cdot\sqrt{\dfrac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n}}}$$

$$= \frac{\sum_{i=1}^{n}(X_i-\bar{X})^3}{n(S^*)^3}$$

$$\hat{\alpha}_4 = \frac{m_4}{m_2^2}$$

$$= \frac{\sum_{i=1}^{n}(X_i-\bar{X})^4}{n} \times \frac{n^2}{\left[\sum_{i=1}^{n}(X_i-\bar{X})^2\right]^2}$$

$$= \frac{\sum_{i=1}^{n}(X_i-\bar{X})^4}{n(S^*)^4}$$

โดยที่ $S^* = \sqrt{\dfrac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n}}$

ส่วนสูตรที่ใช้ในการคำนวณค่าวัดความเบ้ของโปรแกรม MINITAB มีวิธีการหาเช่นเดียวกับโปรแกรม Maple แต่เนื่องจาก S^{*2} เป็นตัวประมาณที่เอนเอียงของ σ^2 ดังนั้นจึงใช้ S^2 นิยามโดย $S^2 = \dfrac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n-1}$ ซึ่งตัวประมาณที่ ไม่เอนเอียงของ σ^2 แทน และสูตรที่ใช้ในการคำนวณค่าวัดความโด่งของโปรแกรม MINITAB ก็มีวิธีการเช่นเดียวกับค่าวัดความเบ้ แต่มีการปรับค่าด้วย –3 เช่นเดียวกับโปรแกรม SPSS, SAS และ Microsoft Excel สูตรต่างๆ สามารถสรุปได้ใน ตารางที่ 2

ในการพิจารณาว่าข้อมูลมีลักษณะสมมาตรหรือไม่นั้น สามารถพิจารณาได้จากค่าวัดความเบ้ ดังนี้ ถ้าค่าวัดความเบ้เท่ากับ 0 แสดงว่า ข้อมูลมีลักษณะสมมาตร ถ้าค่าวัดความเบ้มีค่าน้อยกว่า 0 แสดงว่า ข้อมูลมีลักษณะเบ้ซ้าย แต่ถ้าค่าวัดความเบ้มีค่ามากกว่า 0 แสดงว่า ข้อมูลมีลักษณะเบ้ขวา

สำหรับการพิจารณาว่าข้อมูลมีลักษณะโด่งมากหรือน้อยนั้น สามารถพิจารณาได้จากค่าวัดความโด่ง ค่าวัดความโด่งที่คำนวณจากโปรแกรม SPSS, SAS, Microsoft Excel และ MINITAB ถ้ามีค่าเท่ากับ 0 แสดงว่า ข้อมูลมีลักษณะโด่งปกติ ถ้ามีค่าน้อยกว่า 0 แสดงว่า ข้อมูลมีลักษณะโด่งน้อย แต่ถ้ามีค่ามากกว่า 0 แสดงว่า ข้อมูลมีลักษณะโด่งมาก ส่วนค่าวัดความโด่งที่คำนวณจากโปรแกรม Maple ถ้ามีค่าเท่ากับ 3 แสดงว่า ข้อมูลมีลักษณะโด่งปกติ ถ้ามีค่าน้อยกว่า 3 แสดงว่า ข้อมูลมีลักษณะโด่งน้อย แต่ถ้ามีค่ามากกว่า 3 แสดงว่า ข้อมูลมีลักษณะโด่งมาก

5. บทสรุป

ในการคำนวณค่าวัดความเบ้และความโด่งของข้อมูลจากโปรแกรมสำเร็จรูปชนิดต่างๆ นั้นบางโปรแกรมก็ให้ค่าเท่ากัน บางโปรแกรมก็ให้ค่าแตกต่างกันออกไป ซึ่งค่าวัดความเบ้และความโด่งที่คำนวณจาก SPSS, SAS, Microsoft Excel ให้ค่าวัดซึ่งมีสมบัติที่ดีในเชิงสถิติของตัวประมาณ กล่าวคือเป็นตัวประมาณที่ไม่เอนเอียงของพารามิเตอร์ ดังนั้น การอ่านผลลัพธ์จากโปรแกรมสำเร็จรูปควรเข้าใจที่มาของค่าต่างๆ และศึกษาคู่มือการใช้โปรแกรมหรือระบบความช่วยเหลือ (Help) ให้เข้าใจก่อนเริ่มใช้โปรแกรม

ตารางที่ 2 สูตรที่ใช้ในการคำนวณค่าวัดความเบ้และความโด่งของโปรแกรมสำเร็จรูปชนิดต่างๆ

ค่าวัด	โปรแกรม		
	SPSS, SAS และ Microsoft Excel	MINITAB	Maple
ความเบ้ ($\hat{\alpha}_3$)	$\dfrac{n\sum_{i=1}^{n}(X_i-\overline{X})^3}{(n-1)(n-2)S^3}$	$\dfrac{\sum_{i=1}^{n}(X_i-\overline{X})^3}{nS^3}$	$\dfrac{\sum_{i=1}^{n}(X_i-\overline{X})^3}{n(S^*)^3}$
ความโด่ง ($\hat{\alpha}_4$)	$\dfrac{\frac{n(n+1)}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^4-3\left[\sum_{i=1}^{n}(X_i-\overline{X})^2\right]^2}{(n-2)(n-3)S^4}$	$\dfrac{\sum_{i=1}^{n}(X_i-\overline{X})^4}{nS^4}-3$	$\dfrac{\sum_{i=1}^{n}(X_i-\overline{X})^4}{n(S^*)^4}$

โดยที่ $S = \sqrt{\dfrac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{n-1}}$ และ $S^* = \sqrt{\dfrac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{n}}$

6. เอกสารอ้างอิง

[1] Larsen, R.J. and Marx M.L., An Introduction Mathematical Statistics and Its Applications., Prentice-Hall, New Jersey, 790 p, 2001.

[2] Cramer, H., Mathematical Methods of Statistics., Princeton university press, Princeton, 575 p, 1945.

[3] Joreskog, K.G., Formulas for Skewness and Kurtosis,http://www.ssicentral.com/lisrel/techdocs/kurtosis.pdf, สืบค้น 26 มิถุนายน 2548.

[4] Kirk, R.E., Statistics : An Introduction., Harcourt Brace College Publisher, Texas, 755 p, 1999.

[5] SPSS Inc., SPSS Base 10.0 Applications Guide. SPSS Inc., Chicago, 537 p, 1999.

[6] Delwiche, L.D., and Slaughter, S., The Little SAS Book., SAS Publishing, Cary, 268 p, 2001.

[7] Minitab Inc., Minitab Help : Descriptive Statistics., Minitab Inc., 1996.

[8] Heal, K.M., Hansen, M.L., and Rickard, K.M., Maple 6 Learning Guide., Waterloo Maple Inc., Waterloo, 314 p, 2000.

[9] Kirkup, L., Data Analysis with Excel., Cambridge University Press, Cambridge, 446 p, 2002.