$t = \frac{b - \beta}{S_b}$

ความไม่แน่นอนของการวัดจากกราฟมาตรฐาน

อนุสิทธิ์ สุขม่วง

บทคัดย่อ

ค่าความไม่แน่นอนของการวัดเป็นพารามิเตอร์ที่แสดงช่วงของค่าที่วัดได้ เป็นปริมาณที่จำเป็นในการรายงานผลการวัด เนื่องจากเป็นค่าที่ต้องใช้ในการตัดสินใจหรือตัดสินผลการวัดด้วยเช่นกัน ดังนั้นในการคำนวณ ค่าความไม่แน่นอนของการวัด จะต้องนำความไม่แน่นอนจากแหล่งต่างๆ มาคำนวณให้ครบถ้วน และใช้หลักการทางสถิติให้ถูกต้อง ความไม่แน่นอนจากกราฟมาตรฐาน (calibration curve) เป็นแหล่งหนึ่งที่สำคัญ แนวทางในการคำนวณค่าความไม่แน่นอนจากกราฟมาตรฐาน เริ่มต้นจากการสร้างสมการของกราฟมาตรฐาน ทดสอบความเหมาะสมของสมการที่จะนำไปใช้วัด การหาความคลาดเคลื่อนของสมการ ความคลาดเคลื่อนของความชันและความคลาดเคลื่อนจุดตัดแกน Y การคำนวณค่าความไม่แน่นอนจากกราฟมาตรฐานมีสองแบบ คือ คำนวณค่าความคลาดเคลื่อนของ Y จากค่า X และ คำนวณค่าความคลาดเคลื่อนของ X จากค่า Y ขึ้นอยู่กับการใช้งาน ในการคำนวณมีค่าต่างๆ มากมาย สามารถนำโปรแกรมสำเร็จรูป Excel มาใช้ เพื่อความสะดวกถูกต้อง และรวดเร็ว

คำนำ

เครื่องมือวิเคราะห์ทางเคมีที่ใช้วัดผลการวิเคราะห์จะแสดงผลลัพธ์เป็นสัญญาณในรูปแบบฟังก์ชันทางคณิตศาสตร์ต่างๆ กัน เช่น แบบเส้นตรง แบบเส้นโค้ง หรือแบบลอการิทึม ขึ้นอยู่กับธรรมชาติของสิ่งที่ถูกวัด และหลักการทางทฤษฎีของการวัด กราฟมาตรฐานเป็นกราฟที่แสดงความสัมพันธ์ระหว่างสัญญาณจากเครื่องมือวัดกับความเข้มข้น หรือปริมาณของสารมาตรฐาน การสร้างกราฟมาตรฐานที่เหมาะสมทำให้สามารถนำไปใช้วัดปริมาณสารในตัวอย่างที่ต้องการได้อย่างถูกต้อง ใช้ได้ และมีความน่าเชื่อถือ ค่าความไม่แน่นอนจากกราฟมาตรฐานเป็นแหล่งของความไม่แน่นอนหนึ่ง ที่นำไปคำนวณค่าความไม่แน่นอนของการวัด ในบทความนี้จะกล่าวถึงกราฟมาตรฐานที่เป็นเส้นตรง

การสร้างสมการของกราฟมาตรฐาน ที่เป็นสมการเชิงเส้น

$$y = a + bx \quad (1)$$

a คือค่า y-intercept ของเส้นตรง

b คือค่า slope ของเส้นตรง

โดยที่

$$b = \frac{S_{xy}}{S_{xx}} \quad (2)$$

$$S_{xx} = \sum (x_i - \bar{x})^2$$

$$S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) \quad (3)$$

$$a = \bar{y} - b\bar{x}$$

ความคลาดเคลื่อนในการประมาณค่า

ในกระบวนการวัดทางวิทยาศาสตร์จะมีความไม่แน่นอนของผลการวัดเกิดขึ้นได้จากหลายๆ ปัจจัย การวัดค่าโดยใช้กราฟมาตรฐาน ค่าที่ได้จากตัวอย่างเป็นค่าสถิติที่จะนำไปประมาณค่าของประชากรทั้งหมด ความคลาดเคลื่อนย่อมเกิดขึ้น ทั้งจากสมการที่สร้างจากข้อมูลตัวอย่าง การใช้ค่าความชัน และค่าตัดแกน Y การคำนวณค่าความคลาดเคลื่อนของค่าความชัน และค่าตัดแกน Y จะใช้ค่า Residual Standard deviation, S_R ซึ่งวัดการกระจายของข้อมูลรอบเส้นถดถอย สำหรับค่า x แต่ละค่า และวัดการกระจายในแนวแกน y เป็นค่าเริ่มต้น ดังสมการ

Residual Standard deviation $$s_R = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n - 2}} \quad (4)$$

Standard deviation of slope $s_b = \frac{s_R}{\sqrt{S_{xx}}}$ (5)

Standard deviation of intercept $s_a = s_R\sqrt{\frac{\sum x_i^2}{nS_{xx}}}$ (6)

วิธีการวัดค่า เมื่อนำสมการความสัมพันธ์ที่ได้ไปใช้วัดค่า มี 2 ลักษณะที่ต่างกัน

1. Direct Calibration เป้าหมายคือการวัดค่า y สิ่งที่ต้องการหาคือความคลาดเคลื่อนของ y, จากปริมาณนำเข้า x

จากสมการความสัมพันธ์ y = a + bx

ให้ K เป็นจำนวนสารละลายมาตรฐานที่สร้างกราฟ และ จาก Rules of uncertainty propagation

ได้ $S_{yo} = S_R\sqrt{\frac{1}{K} + \frac{(x-\bar{x})^2}{S_{xx}}}$, (7)

2. Indirect Calibration เป้าหมายคือการวัดค่า x สิ่งที่ต้องการหาคือความคลาดเคลื่อนของ x, S_{x_o} จากปริมาณนำเข้า y

จากสมการความสัมพันธ์ x = (y-a)/b

และ จาก Rules of uncertainty propagation

- กรณี วัดค่า x ครั้งเดียว

$$S_{x_o} = \frac{S_R}{b}\sqrt{1 + \frac{1}{K} + \frac{(x-\bar{x})^2}{S_{xx}}} \quad (8)$$

- กรณี วัดค่า x จำนวน m ครั้ง และใช้ค่าเฉลี่ยของ x

$$S_{x_o} = \frac{S_R}{b}\sqrt{\frac{1}{m} + \frac{1}{K} + \frac{(x-\bar{x})^2}{S_{xx}}} \quad (9)$$

ช่วงความเชื่อมั่นของ x ใด ๆ

$$x = x_o \pm t_{(n-2)}S_R \quad (10)$$

ทดสอบความเหมาะสมของสมการ

การทดสอบว่าสมการเชิงเส้นที่ได้มีความเหมาะสมและสามารถนำไปประเมินค่าของประชากรได้นั้น ต้องทดสอบความเป็นเส้นตรง โดยใช้ **t-test** และทดสอบความเหมาะสมทั้งความใช้ได้ และความน่าเชื่อถือโดยใช้ **F-test**

1. ทดสอบโดย **t-test**

H_0 : $\beta = 0$

H_1 : $\beta \neq 0$

ค่าสถิติที่ใช้ในการทดสอบ คือ

$t = \frac{b-\beta}{S_b}$, โดยที่ β เป็นค่า Slope ของประชากร

df = n-2 , t_{crt} = 2.571

$t_{cal} > t_{crt}$ การทดสอบมีนัยสำคัญ แสดงว่าสมการเชิงเส้นนี้มีความสัมพันธ์เป็นเส้นตรง

2. ทดสอบโดย **F-test**

ค่าสถิติที่ใช้ในการทดสอบ คือ

$F = \frac{MSR}{MSE}$ ซึ่งเป็นค่าอัตราส่วนของความแปรปรวนของ **regression** เปรียบเทียบกับความแปรปรวนของ **error**

ให้ SST : Sum of squares total

SSR : Sum of squares regression

SSE : Sum of squares error

SST = SSR + SSE

$SST = \sum_{i=1}^{n}(y_i - \bar{y})^2,\ SSR = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2,\ SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

ในแต่ละ sum of squares เมื่อหารด้วย df ของแต่ละตัว จะได้ค่าความแปรปรวน ดังนี้

MST = SST/n-1 , MSR = SSR/k ,

MSE = SSE/n-k-1

ปฏิเสธสมมติฐานเมื่อ ค่า F ที่คำนวณได้มีค่ามากกว่าค่าวิกฤติ

ตัวอย่าง การหาปริมาณ **Fluorescein** โดยใช้กราฟมาตรฐานที่วัดสารละลายมาตรฐาน **Fluorescein** โดยใช้เครื่อง **Fluorescence spectrometer** ได้ผล ดังนี้

Fluorescence intensities:	2.1	5.0	9.0	12.6	17.3	21.0	24.7
Concentration, pg/mL	0	2	4	6	8	10	12

จงคำนวณหาค่าความไม่แน่นอนของ x_0 เมื่อวัดค่า intensities ได้ 2.9 pg/mL

วิธีทำ 1. หาความชัน (b) และจุดตัดแกน Y (a) โดยใช้สมการ (2) และ (3) ได้ค่า

b = 1.93 , a = 1.52

2. คำนวณหาความคลาดเคลื่อนของค่า x ในแนวแกน y (S_R) ค่าความคลาดเคลื่อนของความชัน (S_b) และความคลาดเคลื่อนของจุดตัดแกน y (S_a)

$S_R = \sqrt{0.9368/5} = \sqrt{0.18736} = 0.4329$ จากสมการ (4)

$S_b = 0.4329/\sqrt{112} = 0.4329/10.58 = 0.0409$

จากสมการ (5)

$S_a = 0.4329/\sqrt{\frac{364}{7\times112}} = 0.2950$ จากสมการ (6)

เมื่อ y_0 = 2.9,

x_0 = (2.9-1.52)/1.93 = 0.716 pg/mL

จากสมการ (1)

S_{xx} = 112.0, K = 7, วัดค่าครั้งเดียว

$$S_{x0} = \frac{0.4329}{1.93}\sqrt{1+\frac{1}{7}+\left(\frac{2.9-13.1}{112}\right)^2}$$

= 0.264 จากสมการ (8)

3. คำนวณหาช่วงความเชื่อมั่น 95% จากสมการ (10)

$t_{0.025}$ = 2.571 , df = 7-2 =5

= 0.716 ± 2.571*0.264

= 0.716 ± 0.68 pg/mL

ความเข้มข้นของ Fluorescein ที่วัดจากกราฟมาตรฐานมีค่า เท่ากับ 0.716 ± 0.68 pg/mL ที่ระดับความเชื่อมั่น 95%

การวัด Fluorescein ที่ความเข้มข้น 0.716 pg/mL มีค่าความไม่แน่นอนจากกราฟมาตรฐาน ±0.264 pg/mL เป็นข้อมูลที่นำไปใช้เป็นค่าความไม่แน่นอนของการวัด Fluorescein

เนื่องจากในการคำนวณค่าจะประกอบด้วยตัวเลขจำนวนมาก เพื่อความสะดวก และป้องกันความผิดพลาด สามารถใช้สูตร และฟังชั่นใน **Excel** คำนวณค่าได้ ดังภาพที่ 1

									Predicted y	Residual²	SSR
	x_i	y_i	$x_i-\bar{x}$	$(x_i-\bar{x})^2$	$y_i-\bar{y}$	$(y_i-\bar{y})^2$	x^2	$(x_i-\bar{x})(y_i-\bar{y})$	$\hat{y}=a+bx$	$(y-\hat{y})^2$	$(\hat{y}-\bar{y})^2$
	0	2.1	-6.0	36.0	-11.0	121.00	0.00	66.00	1.518	0.339	134.146
	2	5	-4.0	16.0	-8.1	65.61	4.00	32.40	5.379	0.143	59.620
	4	9	-2.0	4.0	-4.1	16.81	16.00	8.20	9.239	0.057	14.905
	6	12.6	0.0	0.0	-0.5	0.25	36.00	0.00	13.100	0.250	0.000
	8	17.3	2.0	4.0	4.2	17.64	64.00	8.40	16.961	0.115	14.905
	10	21	4.0	16.0	7.9	62.41	100.00	31.60	20.821	0.032	59.620
	12	24.7	6.0	36.0	11.6	134.56	144.00	69.60	24.682	0.000	134.146
sum=	42.0	91.7	0.0	112.0	0.0	418.3	364.0	216.2	0.9368	0.9368	417.343
average=	6.0	13.1									
K=	7	7									
m=	1		S_R=	0.4328							
Slope(b)=	1.9304		S_b=	0.0409							
Intercept (a)=	1.5179		S_a=	0.2949							

ภาพที่ 1 แสดงแผ่นคำนวณโดยใช้โปรแกรม Excel

นอกจากนั้น ยังสามารถใช้ Data analysis ของ Excel โดยเลือกจากเมนู Tools -> Data analysis -> regression คำนวณค่า เพื่อประมวลผลต่อไปได้ ดังภาพที่ 2

ซึ่งจะเห็นว่า

เซลล์ B17 คือ ค่า a เซลล์ B7 คือ ค่า S_R

เซลล์ B18 คือ ค่า b เซลล์ C17 คือ ค่า S_a

เซลล์ B18 คือ ค่า S_b

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	*Regression Statistics*								
4	Multiple R	0.99888							
5	R Square	0.99776							
6	Adjusted R Square	0.99731							
7	Standard Error	0.43285							
8	Observations	7							
9									
10	ANOVA								
11		*df*	*SS*	*MS*	*F*	*Significance F*			
12	Regression	1	417.3432	417.3432	2227.5276	0.0000			
13	Residual	5	0.9368	0.1874					
14	Total	6	418.28						
15									
16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
17	Intercept	1.518	0.295	5.146	0.004	0.760	2.276	0.760	2.276
18	xi	1.930	0.041	47.197	0.000	1.825	2.035	1.825	2.035

ภาพที่ 2 แสดงผลการประมวลผล จาก Data analysis ในโปรแกรม Excel

ผลการทดสอบความเหมาะสมของสมการ

1. ทดสอบความเป็นเส้นตรง โดย t-test

ค่าสถิติที่ใช้ในการทดสอบ คือ t โดยที่

df = n-2 , t_{crt} = 2.571

$$t_{cal} = \frac{1.9303-0}{0.0409}$$

t_{cal} = 47.1966

$t_{cal} > t_{crt}$ การทดสอบมีนัยสำคัญ แสดงว่าสมการเชิงเส้นนี้มีความสัมพันธ์เป็นเส้นตรง

2. ทดสอบความเหมาะสมทั้ง ความใช้ได้ และความน่าเชื่อถือ ทดสอบโดย F-test

ผลจากการคำนวณในภาพที่ 2 เซลล์ E12 ได้ค่า F_{cal} = 2,227.528 ซึ่งมากกว่าค่าวิกฤติ F_{crt} = 6.608 หรือ *p*-value ในเซลล์ F12 มีค่าน้อยกว่า 0.025 แสดงว่าสมการเชิงเส้นนี้มีความใช้ได้ และมีความน่าเชื่อถือ เหมาะสมที่จะใช้ในการประเมินค่าต่อไป

จากตัวอย่าง แสดงให้เห็นว่าสามารถคำนวณค่าความไม่แน่นอนของการวัดจากปัจจัยความไม่แน่นอนอันเนื่องมาจากกราฟมาตรฐานได้ โดยใช้หลักทางสถิติและโปรแกรมสำเร็จรูปช่วยในการคำนวณ ทำให้เกิดความสะดวกรวดเร็วในการปฏิบัติงาน และสามารถตรวจสอบถึงความเหมาะสมที่จะนำไปใช้งานได้ โดยเหมาะสมทั้งความใช้ได้ และ ความน่าเชื่อถือ

ผู้ที่สนใจเรื่องการคำนวณค่าความไม่แน่นอนของการวัด สามารถเข้ารับการฝึกอบรมจากหลักสูตร การคำนวณค่าทางสถิติสำหรับงานวิเคราะห์ทดสอบ และหลักสูตรความไม่แน่นอนของการวัดทางเคมี ซึ่งจัดโดยสำนักพัฒนาศักยภาพนักวิทยาศาสตร์ห้องปฏิบัติการ กรมวิทยาศาสตร์บริการ

เอกสารอ้างอิง

EURACHEM/ CITAC. Guide: Quantifying uncertainty in analytical measurement. [Online] [cite dated 10 January 2009] Availablefrom Internet : http://www.chem.utoronto.ca/coursenotes/analsi/StatsTutorial/ErrRegr.html

European Federation of National Associations of Measurement, Testing and Analytical Laboratories (EUROLAB). Guide to the evaluation of measurement uncertainty for quantitative tests results. Eurolab documents;Technical reports August, No. 1/2006 [Online] [cite dated 10 January 2009] Available from Internet : http://www.eurolab.org/docs/technical%20report/EL_11_01_06_387%20Technical%20report%20%20Guide_Measurement_uncertainty.pd

Miller, James N. ; and Miller, Jane C. **Statistics and chemometrics for analytical chemistry.** 4th edition, **Dorchester, Dorset : Dorset,** 1999.

กรมวิทยาศาสตร์บริการ. **เอกสารประกอบการฝึกอบรมหลักสูตรการคำนวณค่าทางสถิติสำหรับงานวิเคราะห์ทดสอบ.** 15-16 มกราคม 2552, กรุงเทพมหานคร : สถานศึกษาเคมีปฏิบัติ, 2552.

สำนักงานสถิติแห่งชาติ. **เอกสารประกอบการบรรยาย การวิเคราะห์ความถดถอยและสหสัมพันธ์** กรุงเทพมหานคร: สำนักงานสถิติแห่งชาติ, 2551.