الفصل التاسع: المعالجة الإحصائية لنتائج التحليل The Treatment of Analytical Data

يقصد بالأخطاء اختلاف القيمة المقاسة او المحسوبة عن القيمة الحقيقية فإذا أعيد إجراء قياس معين لعدة مرات فان القيم ستكون مختلفة وان كان الاختلاف صغيراً.

مصادر الأخطاء Sources of errors

يمكن تقسيم الأخطاء التحليلية الى نوعين رئيسيين هما: -

اولاً - الأخطاء المحددة *Determinate errors*

وهي الأخطاء التي يمكن تحديدها وتعيينها وتسمى الأخطاء النظامية *Systematic errors* ومن انواعها:

1- **أخطاء آلية *Instrumental errors*** وهي أخطاء ناتجة عن عدم قراءة المقياس بشكل دقيق او استعمال الميزان او أدوات غير معايرة.

2- **أخطاء ناتجة عن طريقة التحليل *Methodic errors*** مثل التفاعلات غير التامة او غير القياسية وتلوث الراسب او ذوبانية الراسب او أخطاء ناتجة عن التجفيف والحرق غير التامين او وزن الجفنة في درجات حرارية مختلفة عند المقارنة.

3- **أخطاء تشغيلية *Operative errors*** وهي أخطاء ناتجة عن قلة خبرة المحلل لو المشغل للأجهزة مثل إجراء التحليل في أواني غير مغطاة مما يؤدي الى تلوث النموذج وكثرة غسل الراسب مما يؤدي الى ذوبانه وهكذا.

4- **الأخطاء الشخصية *Personal errors*** وهي الأخطاء المقترنة بالمحلل وليس الخبرة مثل عدم القدرة على تمييز الألوان المختلفة او تحديد القراءة بالضبط او اختيار نتيجة دون أخرى وبشكل عام ناتجة عن التفاوت النسبي بين الأشخاص.

ثانياً - الأخطاء غير المحددة Indeterminate errors

وهي الأخطاء التي لا يمكن تحديد مصدرها او إعادتها او السيطرة عليها ولذلك تعرف أحيانا بالأخطاء العشوائية *Random errors*. ولا تخلو نتائج التحليل الكيميائي على عدد من هذه الأخطاء غير المحددة والتي تكون قيمتها صغيرة غالباً في النتائج المقبولة وإذا تجاوزت حدود معينة ترفض النتائج ولكل طريقة تحليل حدود من الأخطاء المقبولة وهذا ما يعرف الخطأ النسبي المئوي ومنه نحسب الاسترداد (الاسترجاع) المئوي. وهناك طرق إحصائية لحساب هذه الأخطاء ومعالجتها ومن أسبابها اختلاف القياس وتفاوت درجات الحرارة وتذبذب التيار الكهربائي والضوضاء الالكترونية في الأجهزة.

ان معالجة الأخطاء المحددة وكذلك مصدر الخطأ غير المحدد يقلل من نسبة الخطأ في النتائج مثل ضبط كمية او حجم العينة ومعايرة الأدوات المستخدمة ومقايسة الأجهزة وضبطها واختيار الطرائق القياسية المعتمدة والمحلل الكيميائي ذو الخبرة المقبولة في التحليل وكفاءة ودقة الأجهزة وغيرها كلها وسائل تؤدي الى رفع دقة وضبط النتائج، ويمكن الاستعانة بالأساليب التالية لرفع كفاءة التحليل: -

1- تحليل عينات قياسية.
2- إجراء التحليل بطرق مختلفة ثم مقارنة النتائج.
3- إجراء التحليل بواسطة محللين مختلفين.
4- استعمال عينات (خلب).
5- تحليل عينات مختلفة الأوزان.

تتبع الأخطاء غير المحددة قوانين الاحتمالات وعلى هذا الأساس تستعمل بعض القواعد والقوانين الإحصائية لتقييم النتائج المختبرية ومعرفة دقة النتائج . فإذا عبرنا عن عدد القياسات بالرمز N وللقيم المختبرية $X_1 , X_2 , X_3 \ldots\ldots\ldots , X_N$ وللوسط الحسابي $\bar{x}$ فيمكن توضيح الاصطلاحات التالية في المجال الإحصائي :-

1- **الوسط الحسابي *Average , Mean*** : هو المجموع الجبري للقيم مقسوماً على عددها

$$\bar{x} = \frac{\sum x_1 + x_2 + x_3 \ldots\ldots\ldots x_n}{N}$$

2- **المنوال *Mode*** هو القيمة الأكثر تكرار في مجموعة القيم

3- **المدى *Range*** هو الفرق بين اصغر قيمة واكبر قيمة

4- **الانحراف *Deviation*** هو الفرق بين القيمة المحصلة x_i والوسط الحسابي $\bar{x}$ دون النظر الى الاشارة ويمز له d حيث :

$$d = x_i - \bar{x}$$

5- **متوسط الانحراف** ***Average deviation*** هو متوسط مجموع الانحراف لقيم المفردات عن وسطها الحسابي $\bar{d}$

$$\bar{d} = \frac{\sum d}{N} = \frac{\sum(x_1 - \bar{x}) + (x_2 - \bar{x}) \ldots\ldots (x_n - \bar{x})}{N}$$

6- **الانحراف القياسي** ***Standard deviation*** هو القيمة الموجبة للجذر التربيعي لمجموع مربعات الانحراف لقيم المفردات عن وسطها الحسابي مقسوما على (N – 1) اذا كان عدد المفردات صغيرا ويرمز له S او يكون عددها كبيرا ويكون الرقم مقسوماً على N ويرمز له في هذه الحالة σ

$$S = \sqrt{\frac{\sum(x_i - \bar{x})^2}{N-1}}$$

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{N}}$$

7- **التباين** ***Variance*** هو مربع الانحراف القياسي ويرمز له عادة S^2 ويعتبر من المقاييس المهمة في الحسابات الإحصائية, ويستعمل لمعرفة مدى بعثرة القراءات

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{N-1}$$

8- **معامل التباين :** ويستخدم لمعرفة التباين لوصف توافق النتائج التحليلية

$$C_V = \frac{S}{\bar{x}} \times 100$$

9- **الانحراف القياسي النسبي المئوي** ***Relative S.D.***

$$RSD\ \% = \frac{S}{\bar{x}} \times 100$$

طرائق التعبير عن الدقة وعن التوافق

أولا - الدقة *Accuracy*

مقياس لمدى تقارب النتائج مع القيمة الحقيقية وكلما كان الفرق بين القيم المقاسة والقيمة الحقيقية قليلاً زادت دقة النتائج. ويعبر عنها بعدة طرق:

1- **الخطأ المطلق** ***Absolute errors*** هو الاختلاف بين القيمة الحقيقية والقيمة المقاسة مع الأخذ بالإشارة بنظر الاعتبار

$$E = X_i - X_t$$

2- **الخطأ النسبي** ***Relative errors*** هو قيمة الخطأ المطلق مقسوماً على القيمة الحقيقة مضروبا بالمئة

$$E_{rel} = \frac{X_i - X_t}{X_t} \times 100$$

القيمة الحقيقية ***True Value*** (X_t) هي القيمة التي تمثل حقيقة مقدار الشيء المحلل وإذا لم تكن معروفة او يصعب الحصول عليها تستخدم القيمة الأكثر احتمالاً او المعدل (القيمة النظرية).

القيمة المقاسة (الملاحظة) هي القيمة التي يتم الحصول عليها بالملاحظة او بالتجربة فإذا كان عدد القيم المقاسة N أصبحت القيم المقاسة $(x_1, x_2, x_3, x_4 \ldots\ldots\ldots\ldots x_N)$

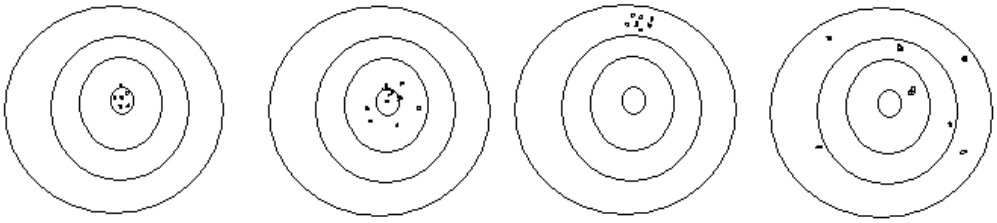
ثانياً – التوافق (الضبط) Precision

هو مقياس لمدى تقارب النتائج مع بعضها البعض بغض النظر عن قيمتها الحقيقية وكلما كان الفرق بين القيم العددية للنتائج صغيراً كلما كان التوافق اعلى، ويعبر عن التوافق بعدة طرائق: -

1- توافق مجموعة متسلسلة من النتائج لقياس اساس معدل الانحراف $\bar{d} = \frac{\sum d}{N}$

2- توافق مجموعة متسلسلة من النتائج لقياس معدل الانحراف القياسي النسبي $RSD = \frac{S}{\bar{x}}$

3- توافق مجموعة متسلسلة من النتائج لقياس معدل الانحراف القياسي النسبي المئوي $RSD\ \% = \frac{S}{\bar{x}} \times 100$.



ضبط قليل , دقة قليلة دقة قليلة , ضبط عالي دقة عالية , ضبط قليل دقة عالية , ضبط عالي

مثال/ تم قياس حجم سائل من قبل ستة طلاب وكانت نتائجهم كما يلي :-

24.32, 24.38, 24.45, 24.29, 24.3, 24.31 cm^3 احسب الوسط الحسابي والوسيط والمدى والانحراف القياسي والتباين لهذه القياسات؟

الحل /

$$Mean = \frac{\sum X_i}{N} = \frac{146.05}{6} = 24.34\ cm^3$$

$$Range = 24.45 - 24.29 = 0.16\ \ cm^3$$

لاستخراج الوسيط يتم ترتيب الأرقام (الحجوم) اما تصاعدياً او تنازلياً 24.29, 24.30, 24.31, 24.32, 24.38, 24.45 cm^3 , , ولما كان عدد القياسات زوجيا فإن :

$$Mean = \frac{1}{2}\ (x_N + x_{N+1})$$

$$Mean = \frac{1}{2}\ (24.31 + 24.32) = 24.315\ cm^3$$

نرتب القيم في جدول رياضي ومن الجدول نستخرج قيم الانحراف ومربع الانحراف

الحجوم (سم^3)	الانحراف d (سم^3)	مربع الانحراف d2 (سم^3)^2
24.32	- 0.022	0.000489
24.38	0.038	0.001444
24.45	0.108	0.011664
24.24	- 0.052	0.002704
24.30	- 0.042	0.001764
24.31	- 0.032	0.001024

$$S = \sqrt{\frac{(x_i - \bar{x})^2}{N-1}} = \sqrt{\frac{d^2}{N-1}} = 0.019084\ cm^3$$

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{N-1} = \frac{\sum d^2}{N-1} = \frac{0.019084}{6-1} = 0.003817$$

دقّة وأهمية الوسط الحسابي

عند عدم توفر القيم الحقيقية نستعمل الوسط الحسابي كبديل عن القيمة الحقيقية وذلك بأخذ قياسات كثيرة من X الى X_N واخذ الوسط الحسابي لها $\bar{x}$ وقد تختلف هذه القيمة قليلاً او كثيراً عن القيمة الحقيقية لذلك وجدت اساليب احصائية لتقويم النتائج منها :-

1- اختبار الطالب (اختبار t) – مجال الثقة – *Student's test*

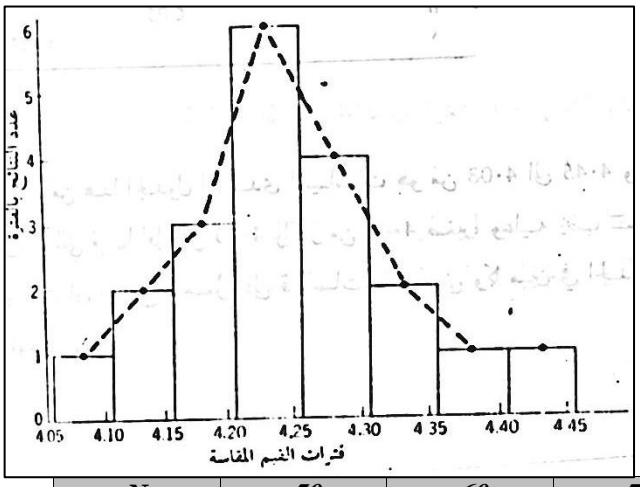
في هذا الاختبار نعتمد على المعادلة التالية :

$$\pm t = \frac{(\bar{x} - \mu)\sqrt{N}}{S}$$

حيث : $\bar{x}$: الوسط الحسابي لعدد محدود من النتائج

μ : الوسط الحسابي لعدد كبير من النتائج

S : الانحراف القياسي لمجموعة من القراءات ذات اعداد محددة N من القراءات



ويتعذر حساب قيم t من المعادلة لعدد صغير من القراءات لان قيمة S و μ لاتمثلان الحقيقة غير ان قيم t يمكن الحصول عليها من الجداول الخاصة بذلك والتي تأخذ بنظر الاعتبار التباين بين قيم $\bar{x}$ و μ , وتكتب المعادلة بشكل اكثر فائدة كما يلي :-

$$\mu = \bar{x} \pm \frac{t.S}{\sqrt{N}} = \left(\bar{x} + \frac{t.S}{\sqrt{N}} \ldots\ldots\ldots \bar{x} - \frac{t.S}{\sqrt{N}}\right)$$

ان قيم t في الجداول تختلف باختلاف حدود الثقة *Confidence Interval* حيث توجد جداول رياضية خاصة بذلك يتم الرجوع إليها من المصادر ويمكن وضع الجدول التالي للتوضيح :

N	*50*	*60*	*70*	*80*	*90*	*95*	*99*
2	1.000	1.000	1.376	1.963	0318	12.706	63.657
3	0.816	1.064	0.386	1.886	2.920	4.303	9.925
4	0765	0.978	1.250	1.638	2.353	3.182	5.841
5	0.741	0.944	1.190	1.533	2.132	2.776	4.604
6	0.727	0.920	1.156	1.476	2.015	2.571	4.032

واجب // متابعة واكمال الجدول من اي مصدر

وضعت المعادلة السابقة للتعبير عن الاختلاف بين $\bar{x}$ و μ ان اهم المعالجات الإحصائية للتعبير عن توافق النتائج هو استخدام منحني غاوس الطبيعي *The Gaussian curve* وهو المنحني الذي يربط تكرارية عدد القياسات بانحرافها.

منحني غاوس الطبيعي

ان احتمالية وقوع أي قراءة وقعت بصورة عشوائية ضمن هذا المدى وهو مستوى الثقة *Confidence level* وان مجال الثقة للقراءة تساوي *Confidence limit* $(C.l. = X_i \pm t.\sigma)$

مثال/ وجد محلل كيميائي ان نتيجة تحليل النسبة المئوية للحديد في إحدى خاماته هي $\bar{x} = 15.30$, $N = 4$, $S = 0.1$, احسب μ بحدود ثقة 95 و 99 ؟

الحل// $Confidence\ limit\ for\ 95\%\quad \mu = \bar{x} \pm \frac{t.S}{\sqrt{N}} = 15.30 \pm \frac{0.1 \times 2.353}{\sqrt{4}} = 15.30 \pm 1.2$

$$Confidence\ limit\ for\ 99\%\quad \mu = \bar{x} \pm \frac{t.S}{\sqrt{N}} = 15.30 \pm \frac{0.1 \times 5.841}{2} = 15.30 \pm 0.39$$

2- اختبار *f* *f - test*

يستخدم هذا الاختبار لبيان الفرق بين طريقتين مختبريتين استعملتا لتحليل نموذج معين او تقديم طريقة جديدة في التحليل بالمقارنة مع طريقة مثبتة سابقاً ويعبر عن f بالمعادلة التالية: -

$$F = \frac{S_1}{S_2}$$

حيث S_1 و S_2 التباين في الطريقة الاولى والثانية على التوالي, كلما كان قيمة F المحسوبة تقترب من قيمة F المثبتة في الجداول والتي تتضمن درجات الحرية *Degree of freedom* لقيم $N - 1$ بمقدار كبير كلما استدل على وجود تطابق كبير بين الطريقتين واذا كانت قيمة f المحسوبة اكبر مما هي عليه في الجدول يبين ان هناك اختلاف كبير بين الطريقتين.

حذف نتيجة مختبرية

عند إعادة تحليل نموذج معين لعدة مرات تظهر غالباً نتائج التحليل مختلفة اختلافً ملحوظاً عن بقيتها وان الأسباب التي أدت الى ذلك الاختلاف غير معروفة ويصعب تحديدها وهناك بعض القواعد التي يمكن استخدامها لمعرفة مدى قابلية إهمال نتيجة مختبرية او عدمها ومنها: -

1- قاعدة أربعة إضعاف معدل الانحراف

تنص هذه القاعدة على ما يلي (تستبعد القيمة المشكوك فيها إذا كان انحرافها عن الوسط الحسابي لبقية النتائج اكثر من أربعة إضعاف انحرافها عن متوسط الانحراف لبقية النتائج، اما اذا كانت اقل من أربعة إضعاف يجوز استبعادها او ضمها).

مثال / تم قياس حجم قاعدة معينة لتسحيح حجم حامض مجهول لعدة مرات وكانت النتائج كما يلي :

$11.02\ cm^3$, 11.40 , 11.12 , 10.9 بين هل بالإمكان حذف القراءة 11.40 ؟

الحل / نهمل 11.40 مؤقتاً ونستخرج الوسط الحسابي ومعدل الانحراف لبقية القراءات

المتوسط الحسابي = $11.01\ cm^3$

معدل الانحراف = $4 \times 0.076\ cm^3$

انحراف النتيجة المستبعدة عن الوسط الحسابي = $0.39\ cm^3$ أربعة إضعاف معدل الانحراف = $0.30\ cm^3$ وعليه يتطلب إهمال القراءة .

2- اختبار Q Q test

يعرف بالمعادلة التالية:

$$Q = \frac{X - Y}{R}$$

حيث X تمثل القيمة المراد استبعادها , Y اقرب قيمة لها

R المدى (اكبر قيمة عددية – اصغر قيمة عددية)

مثال / عند تحليل نماذج من الكالسايت وجدت القيم 55.95 , 56.00 , 56.04 , 56.08 , 56.23 ولقد لوحظ ان القيمة الأولى هي اكبر بكثير من بقية القيم بين إمكانية إبقاء او إهمال هذه القيمة عند مستوى ثقة 90% ؟

الحل / من الجداول للقراءة 5 تحت مستوى ثقة 90% = 0.64

$$Q = \frac{56.23 - 56.08}{56.23 - 55.95} = \frac{0.15}{0.28} = 0.54$$

لخمس قياسات وعند مستوى ثقة 90% فإن قيم Q الحرجة = 0.64 وهنا Q = 0.54 لذا فإن هذه القيمة اصغر من Q الحرجة لذا يمكن إبقاءها .

واجب // مجموعة من النتائج 4.30 , 4.20 , 4.28 , 4.45 , 4.17 بيّن إمكانية رفض او قبول النتيجة 4.45 ضمن النتائج بمستوى ثقة 90% = (0.64) ؟

الأرقام المعنوية

الرقم 3.51 له ثلاث أرقام معنوية 3 و 5 و 1, ويتم تقريبها باستخدام قواعد خاصة, فمثلاً لتقريب الرقم 3.450 = 3.45 , اما 3.459=3.46 وهكذا ... تراجع من أي مصدر.

الاسترداد المئوي

تمثل القيمة الناتجة من نسبة الخطأ مطروحة من الرقم 100 اي: -

الاسترداد المئوي = 100 – نسبة الخطأ , وهي عكس نسبة الخطأ.

الحساسية

تمثل التغير في شدة الاستجابة للأجهزة مع تغير التراكيز المستخدمة

$$S = \frac{\Delta I}{\Delta C}$$

S = الحساسية , ΔI = شدة الاستجابة , ΔC = تغير التراكيز, وعند رسم ΔI مقابل ΔC نحصل على خط مستقيم يمر بنقطة الاصل وكلما كان ميل الخط المستقيم عالي تكون الحساسية عالية والعكس صحيح.

$$\Delta I > \Delta C$$

حد الكشف

اصغر تركيز للمادة المحللة يمكن الكشف عنه بالطرق التحليلية المستعملة وبدرجة الثقة المحددة فإذا كان التركيز الأولي $C_{min} = 2 \times 10^{-5}$ g/ml فان الايون المبحوث عنه سيكشف بالطريقة المستعملة فقط تحت ظروف يكون فيها تركيزه اكبر من 2×10^{-5} g/ml .

$$\Delta L = \frac{3\,C \times S}{\bar{x}}$$

ΔL : حد الكشف C : التركيز $\bar{x}$: الوسط الحسابي S : الانحراف القياسي

المصادر

الكتاب المنهجي

- الكيمياء التحليلية، الجزء الاول، الاسس العامة للتحليل الكمي الوزني، صفاء رزوقي المرعب، 1989.
- الكيمياء التحليلية، الجزء الثاني، طرائق الفصل الحديثة، صفاء رزوقي المرعب، 1989.

كتب مساعدة

- "Fundamental of Analytical Chemistry" by Doglas A. Skooge, Donald M. West and James Holler, Eighth Edition, 1988.
- "Analytical Chemistry" by Gary D. Christian, Purnendu K. (Sandy) Dasgupta, Kevin A. Schug, 7th Edition, 2014.
- "An Introduction to Analytical Chemistry" by Doglas A. Skooge & Donald M. West, 4th Edition, 1986.
- "Analytical Chemistry" by Gary D. Christian, John Wiley and Sons, Inc. 6th Edition, 2004.
- الاساسيات النظرية للكيمياء التحليلية اللاعضوية، التحليل الكمي الوزني والحجمي، هادي كاظم عوض واخرون، 1986.
- طرائق وتقنيات حديثة في التحليل الكيميائي الآلي د. جميل موسى ضباب
- مفتاح الابداع للكيمياء التحليلية Analytical Chemistry، عمر جبر حلوة، الطبعة الاولى، 2008.