

ΕΝΑΛΛΑΚΤΙΚΟΙ ΤΡΟΠΟΙ ΕΚΤΙΜΗΣΗΣ ΤΗΣ ΑΒΕΒΑΙΟΤΗΤΑΣ ΤΟΥ ΣΥΣΤΗΜΑΤΙΚΟΥ ΣΦΑΛΜΑΤΟΣ ΜΕΣΩ ΕΜΒΟΛΙΑΣΜΕΝΩΝ ΔΕΙΓΜΑΤΩΝ

Μάριος Γ. Κωστάκης, Νικόλαος Σ. Θωμαΐδης, Κωνσταντίνος Η. Ευσταθίου και
Μιχαήλ Α. Κουππάρης

Εργαστήριο Αναλυτικής Χημείας, Τμήμα Χημείας, Εθνικό και Καποδιστριακό
Πανεπιστήμιο Αθηνών, Πανεπιστημιούπολη Ζωγράφου, 157 71, Αθήνα
E-mail: makostak@chem.uoa.gr

Περίληψη

Το συστηματικό σφάλμα (bias) αποτελεί μέτρο εκτίμησης της ορθότητας (trueness) μιας αναλυτικής μεθόδου. Η εκτίμηση της αβεβαιότητας του συστηματικού σφάλματος είναι μια συνιστώσα της συνολικής αβεβαιότητας που θα πρέπει να περιλαμβάνεται στη συνολική αβεβαιότητα, είτε υπάρχει, είτε όχι στατιστικά σημαντικό συστηματικό σφάλμα. Μάλιστα στη μεθοδολογία Nordtest, για την εκτίμηση της αβεβαιότητας χρησιμοποιείται, τόσο το συστηματικό σφάλμα (εφόσον αξιολογείται ως τυχαίο), όσο και η αβεβαιότητα υπολογισμού του συστηματικού σφάλματος. Μελέτες απέδειξαν ότι όταν η αβεβαιότητα του συστηματικού σφάλματος είναι μεγαλύτερη από το 10% της συνολικής αβεβαιότητας, τότε είναι σημαντική και θα πρέπει να ληφθεί υπόψη η συνεισφορά της στη συνολική αβεβαιότητα. Επομένως πρώτα θα πρέπει να υπολογίζεται η αβεβαιότητα του συστηματικού σφάλματος και ύστερα να γίνεται εκτίμηση της επίδρασής της στη συνολική. Η εκτίμηση της ορθότητας, και κατ' επέκταση της αβεβαιότητας, μπορεί να πραγματοποιηθεί με τρεις τρόπους: 1) Μέσω πιστοποιημένων υλικών αναφοράς (CRM). 2) Από διεργαστηριακές δοκιμές ικανότητας (proficiency test) ή διεργαστηριακές συγκρίσεις (collaborative trials). 3) Μέσω εμβολιασμένων δειγμάτων (spiked ή surrogate samples). Από τους τρεις τρόπους, η πιο συνηθισμένη επιλογή σε ένα εργαστήριο ρουτίνας αποτελεί ο εμβολιασμός, γιατί είναι φθηνός, άμεσα προσβάσιμος και δίνει τη δυνατότητα στον αναλυτή να δημιουργήσει δείγματα σε ένα μεγάλο εύρος συγκεντρώσεων. Ωστόσο, λόγω έλλειψης χρόνου εξισορρόπησης του δείγματος, η εκτίμηση του συστηματικού σφάλματος και της αβεβαιότητας είναι υποεκτιμημένη. Ο συνηθέστερος τρόπος εκτίμησης της αβεβαιότητας του συστηματικού σφάλματος είναι μέσω του υπολογισμού της τυπικής απόκλισης του μέσου όρου. Ο τρόπος αυτός, αν και είναι εύκολος και παρέχει ένα μέτρο της αβεβαιότητας από τα δεδομένα της επικύρωσης (ορθότητας), αποτελεί ουσιαστικά μια εκτίμηση της επαναληψιμότητας του συστηματικού σφάλματος και δεν μπορεί να χρησιμοποιηθεί σε περιπτώσεις όπου υπάρχει σημαντικό συστηματικό σφάλμα. Στη βιβλιογραφία έχουν προταθεί εναλλακτικές μέθοδοι για την εκτίμηση της αβεβαιότητας του συστηματικού σφάλματος. Αυτές οι μέθοδοι βασίζονται σε ένα λεπτομερέστερο πειραματικό σχεδιασμό εκτίμησης της αβεβαιότητας με τη βοήθεια πιο εξειδικευμένων στατιστικών μεθόδων, όπως ANOVA και Youden Plot. Στην παρούσα εργασία, γίνεται μια ανασκόπηση των διαφορετικών μεθόδων εκτίμησης του συστηματικού σφάλματος. Στη συνέχεια οι μέθοδοι αυτές εφαρμόζονται σε μια αναλυτική μέθοδο και ακολουθεί σύγκριση των αποτελεσμάτων τους.

Λέξεις-Κλειδιά: Συστηματικό Σφάλμα, Αβεβαιότητα, Εμβολιασμένα δείγματα

ALTERNATIVE WAYS FOR CALCULATION OF SYSTEMATIC ERROR

Abstract

The aim of this presentation is to study alternative ways for the calculation of systematic error. The most common laboratory practice about determination of systematic error (or bias) is spiking. This procedure is easy, not expensive and immediately available to any laboratory. The uncertainty of bias is estimated through the calculation of standard deviation of mean from the recoveries of spiking. This approach is the most common and straightforward, but it has two drawbacks. First, the standard deviation of mean treats uncertainty of bias as combined uncertainty with no deeper insight into the source of the potential bias. Second drawback is that this approach assumes that the results obtained during the study are valid for future, yet unknown, samples. An alternative way, that it has been proposed, is through nested ANOVA. In nested ANOVA, factors commonly affecting the bias, such as the concentration level or the variability of matrices, are ranked in hierarchical order. From the results of nested ANOVA, the variation of each bias component can be studied and the combined uncertainty can be calculated. In this study are presented the theory and the calculations for the determination of uncertainty of bias through nested ANOVA are presented and an example based on literature data is demonstrated.

Keywords: Bias, uncertainty, spiking samples

1. Εισαγωγή

Κατά τον έλεγχο της αξιοπιστίας μιας αναλυτικής μεθόδου, βασικό χαρακτηριστικό ποιότητας είναι η εκτίμηση της εγγύτητας μεταξύ της πειραματικώς μετρούμενης τιμής και της αληθούς τιμής (ή τιμής αναφοράς). Ο παραπάνω ορισμός αναφέρεται στην ορθότητα της μεθόδου (trueness). Η ορθότητα δεν είναι ποσοτικό μέγεθος, οπότε δεν μπορεί να εκφραστεί αριθμητικά. Ωστόσο, μέγεθος έκφρασης της ορθότητας αποτελεί το συστηματικό σφάλμα (bias) μιας μεθόδου. Το συστηματικό σφάλμα μπορεί να εκφραστεί με δύο τρόπους, είτε ως διαφορά μεταξύ της πειραματικής τιμής και της τιμής αναφοράς (Εξίσωση 1), είτε ως ο λόγος της πειραματικής τιμής και της τιμής αναφοράς (Εξίσωση 2). Ο τελευταίος τρόπος υπολογισμού είναι γνωστός με τον όρο ανάκτηση. Πολύ συχνά στην καθημερινή πρακτική οι όροι ορθότητα, συστηματικό σφάλμα και ανάκτηση χρησιμοποιούνται ως συνώνυμα (Barwick & Prichard 2011, Linsinger 2008).

$$b = x_{mean} - x_{ref} \quad (1)$$

$$b = \frac{x_{mean}}{x_{ref}} \quad (2)$$

Για να πραγματοποιηθεί η εκτίμηση της ορθότητας σωστά και κατ' επέκταση ο υπολογισμός του συστηματικού σφάλματος, θα πρέπει να ισχύουν οι παρακάτω προϋποθέσεις (Linsinger 2008):

- Η μήτρα στην οποία θα γίνει η εκτίμηση θα πρέπει να προσομοιάζει όσο το δυνατόν στα πραγματικά δείγματα.
- Η εκτίμηση θα πρέπει να γίνει σε όλο το εύρος συγκεντρώσεων της μεθόδου.
- Η τιμή αναφοράς θα πρέπει να είναι αξιόπιστη και η αβεβαιότητά της όσο το δυνατόν πιο μικρή.

Με βάση τις παραπάνω προϋποθέσεις, η εκτίμηση της ορθότητας μπορεί να βασιστεί σε (Magnusson and Ellison 2008, Thompson, et al. 1999, Linsinger 2008):

- Πιστοποιημένα Υλικά Αναφοράς (certified reference materials, CRM).
- Πρότυπες μέθοδοι αναφοράς.
- Διεργαστηριακές συγκρίσεις ή δοκιμασίες ικανότητας (Interlaboratory comparisons ή Collaborative trials ή Proficiency tests).
- Εμβολιασμός (spiking ή surrogates).

Ο εμβολιασμός αποτελεί την πιο φθηνή και άμεσα διαθέσιμη λύση για ένα αναλυτικό εργαστήριο. Ο εμβολιασμός μπορεί να γίνει, είτε σε λευκό (blank, placebo) δείγμα, είτε με γνωστή προσθήκη σε δείγμα που περιέχει γνωστή συγκέντρωση. Ο πρώτος τρόπος είναι ο καλύτερος για την εκτίμηση της ορθότητας, επειδή γίνεται εμβολιασμός απευθείας στη μήτρα και φαίνεται απευθείας η επίδραση του αναλύτη, ωστόσο είναι πολύ δύσκολο να παρασκευαστεί μήτρα που να προσομοιάζει τόσο πολύ στο πραγματικό δείγμα. Ο δεύτερος τρόπος αποτελεί τη συνηθισμένη πρακτική που ακολουθεί ένα αναλυτικό εργαστήριο. Δίνει τη δυνατότητα να παρασκευαστούν δείγματα που να προσομοιάζουν πολύ καλά στο δείγμα με χαμηλό κόστος. Τα κύρια μειονεκτήματα είναι ότι η μορφή του αναλύτη στο πρότυπο που εμβολιάζεται μπορεί να διαφέρει από τη μορφή του αναλύτη στη μήτρα,

επομένως απαιτείται κάποιος χρόνος εξισορρόπησης ώστε να δεσμευτεί το πρότυπο από συστατικά της μήτρας (cured samples), κάτι που πολλές φορές δεν είναι εφικτό λόγω της φύσης της μήτρας.

Η διαθεσιμότητα του αναλύτη που εμβολιάζεται σε σχέση με τον ενυπάρχοντα αναλύτη του δείγματος, έχει δείξει από μελέτες που έχουν γίνει με ισοτοπικώς επισημασμένα πρότυπα, ότι υπερεκτιμάει κατά ένα ποσοστό την ορθότητα, υποεκτιμώντας παράλληλα το συστηματικό σφάλμα. Για τους παραπάνω λόγους η αβεβαιότητα της τιμής αναφοράς είναι μεγαλύτερη σε σχέση με τους προηγούμενους τρόπους.

2. Αβεβαιότητα συστηματικού σφάλματος

Ο συνήθης τρόπος για την εκτίμηση της αβεβαιότητας του συστηματικού σφάλματος, βασίζεται στην εφαρμογή του νόμου της διάδοσης των σφαλμάτων, όπως αναφέρατε στον GUM (Guide to the Expression of Uncertainty in Measurement) και φαίνεται στην Εξίσωση 3:

$$u_b = \sqrt{\frac{s_{mean}^2}{n} + u_{ref}^2} \quad (3)$$

Όπου:

s_{mean} , η τυπική απόκλιση των n μετρήσεων του εργαστηρίου.

u_{ref} , η αβεβαιότητα της τιμής αναφοράς, ανάλογα με τον τρόπο που έγινε η εκτίμηση: α) η αβεβαιότητα του πιστοποιητικού του CRM, β) η αβεβαιότητα των εργαστηρίων για διεργαστηριακές συγκρίσεις και γ) είναι η αβεβαιότητα της συγκέντρωσης του προτύπου που εμβολιάστηκε στο δείγμα.

Εναλλακτικό τρόπο για την εκτίμηση της αβεβαιότητας παρέχει το Nordtest, σύμφωνα με το οποίο εισάγεται ένας ακόμη παράγοντας στον υφιστάμενο τρόπο υπολογισμού και αυτός είναι η διαφορά της πειραματικής από την πραγματική τιμή και συμβολίζεται ως *bias*. Οπότε η εξίσωση (3) τροποποιείται ως εξής (Magnusson et al 2012):

$$u_b = \sqrt{RMS_{bias}^2 + \frac{s_{mean}^2}{n} + u_{ref}^2} \quad (4)$$

όπου: $RMS_{bias} = \sqrt{\frac{\sum (bias)_i^2}{n}}$

Η εκτίμηση της αβεβαιότητας μέσω του υπολογισμού του τυπικού σφάλματος του μέσου όρου, παρόλο τον απλό τρόπο υπολογισμού έχει δύο μειονεκτήματα (Linsinger, 2008):

- Αντιμετωπίζει την αβεβαιότητα της ανάκτησης ως συνδυασμό αβεβαιοτήτων, δίχως κάποια διεξοδικότερη αναζήτηση της πηγής του σφάλματος.
- Θεωρεί ότι τα αποτελέσματα που παράχθηκαν από τα δείγματα κατά τη διάρκεια της μελέτης είναι έγκυρα και για τα άγνωστα, μελλοντικά δείγματα.

3. Ανάλυση διακυμάνσεων (ANOVA)

Ένας εναλλακτικός τρόπος για την εκτίμηση της αβεβαιότητας που έχει προταθεί στην βιβλιογραφία, αφορά έναν πιο ολοκληρωμένο πειραματικό σχεδιασμό για την αξιολόγηση της ορθότητας με την χρήση μιας παραλλαγής της ANOVA. Ο συγκεκριμένος τρόπος αφορά την εκτίμηση των πιθανότερων παραγόντων που επηρεάζουν την ορθότητα, την ιεράρχησή τους και την ανάλυση των διακυμάνσεων μεταξύ των παραγόντων, η επεξεργασία αυτή είναι γνωστή ως Ιεραρχική Ανάλυση Διακυμάνσεων (nested ή hierarchical ANOVA).

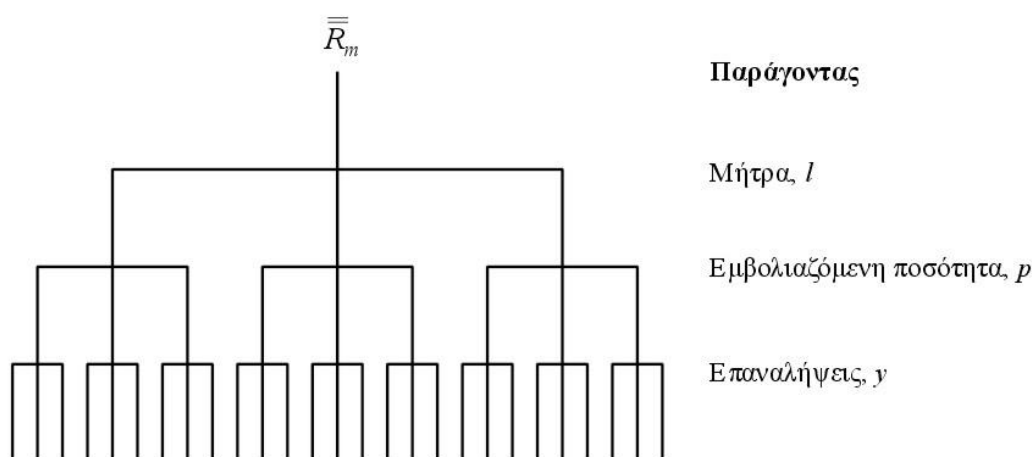
Με βάση την παραπάνω υπόθεση, η αληθής (πραγματική) ανάκτηση υπολογίζεται από την ολική μέση ανάκτηση \bar{R}_m , τη διακύμανση της ανάκτησης που οφείλεται στην τιμή της συγκέντρωσης που εμβολιάζεται στο δείγμα (ΔR_C) και τη διακύμανση της ανάκτησης που οφείλεται στη διαφορά της μήτρας (ΔR_M), όπως φαίνεται και στην Εξίσωση 5 (Boti et al 2007, Wang et al 2007):

$$R = \bar{R}_m + \Delta R_C + \Delta R_M \quad (5)$$

Οι δύο τελευταίοι παράγοντες είναι άγνωστοι για τα μελλοντικά δείγματα, ωστόσο μπορεί να γίνει εκτίμησή τους μέσω του παρόντος πειραματικού σχεδιασμού. Έτσι, ο υπολογισμός της αβεβαιότητας του συστηματικού σφάλματος δίνεται από την παρακάτω εξίσωση (Boti et al 2007):

$$u(R) = \sqrt{u(\bar{R}_m)^2 + u(\Delta R_C)^2 + u(\Delta R_M)^2} \quad (6)$$

Για να γίνει εκτίμηση της αβεβαιότητας με βάση την Εξίσωση 6 θα πρέπει να μελετηθεί η επίδραση του διαφορετικού μητρικού υλικού, η επίδραση της συγκέντρωσης σε όλο το εύρος συγκεντρώσεων της μεθόδου και η διακύμανση των μετρήσεων της ίδιας ημέρας ή/και διαφορετικών ημερών (επαναληψιμότητα και αναπαραγωγιμότητα). Όπως φαίνεται και στο Σχήμα 1, αυτό μπορεί να πραγματοποιηθεί χωρίς περίπλοκους σχεδιασμούς, τοποθετώντας ιεραρχικά τους παράγοντες που θα μελετηθούν.



Σχήμα 1. Ο πειραματικός σχεδιασμός για την εκτίμηση της ανάκτησης για l διαφορετικές μήτρες, για p διαφορετικές συγκεντρώσεις σε κάθε μήτρα και για y επαναλήψεις σε κάθε επίπεδο συγκεντρώσεων κάθε μήτρας (όπως παρατίθεται (Boti et al 2007)).

Στα υψηλότερα επίπεδα τοποθετούνται οι παράγοντες που έχουν τη μεγαλύτερη επίδραση στο συστηματικό σφάλμα, ενώ στα χαμηλότερα οι παράγοντες με τη μικρότερη επίδραση. Στο Σχήμα 1 έχει τοποθετηθεί η μήτρα στο υψηλότερο επίπεδο, ωστόσο τις περισσότερες φορές στην υψηλότερη κλίμακα τοποθετείται η συγκέντρωση.

Ένα μειονέκτημα του παραπάνω σχεδιασμού είναι ο αριθμός των πειραμάτων. Αν θεωρήσουμε ότι τα ελάχιστα επίπεδα συγκεντρώσεων είναι 3, ως ελάχιστο αριθμό διαφορετικών μητρών τις 2 και ότι για κάθε επίπεδο συγκεντρώσεων θα πρέπει να γίνουν τουλάχιστον 2 επαναλήψεις, τότε προκύπτουν 12 πειράματα, γεγονός που μπορεί να αποτελεί ανασταλτικό παράγοντα, ιδίως όταν δεν υπάρχουν πολλές φορές φθηνά διαθέσιμα δείγματα και πρότυπα ή δεν υπάρχει ο κατάλληλος αυτοματισμός, ώστε να πραγματοποιηθούν τα αντίστοιχα πειράματα εύκολα. Τέλος, η πραγματοποίηση της παραπάνω διεργασίας ανάλυσης διακυμάνσεων δεν μπορεί να γίνει με τα συνήθη υπολογιστικά προγράμματα (Excel), αλλά απαιτεί κάποιο πιο εξειδικευμένο πρόγραμμα όπως MINITAB, SAS ή MATLAB.

Έπειτα από την ολοκλήρωση των πειραμάτων και την εξαγωγή των αποτελεσμάτων, τα δεδομένα εισάγονται στο υπολογιστικό πρόγραμμα, το οποίο επιστρέφει ένα πίνακα που είναι κοινός για κάθε ANOVA. Στον πίνακα των αποτελεσμάτων δίνονται οι παράγοντες κατά ιεραρχική σειρά, οι βαθμοί ελευθερίας για κάθε παράγοντα, το άθροισμα των τετραγώνων (Sum of Squares), η διακύμανση (Mean of Squares), η πειραματική τιμή του F-test και η αντίστοιχη P-value. Από τα αποτελέσματα και τα αρχικά δεδομένα υπολογίζονται οι παρακάτω παράγοντες με την εξής διαδικασία (Boti et al 2007, Dehouck et al 2003, Wang et al 2005):

Υπολογισμός της αβεβαιότητας της ολικής ανάκτησης $u(\bar{R}_m)$

Υπολογίζεται με βάση την εξίσωση

$$u(\bar{R}_m) = \sqrt{\frac{\sum_{i=1}^l u(\bar{R}_i)^2}{l^2}} \quad (7)$$

Όπου:

$u(\bar{R}_i)$: η αβεβαιότητα της μέσης ανάκτησης κάθε ομάδας. Με τον όρο μέση ανάκτηση εννοείται η μέση ανάκτηση των ομάδων, που προκύπτουν από τον διαχωρισμό σε πρώτο ιεραρχικό επίπεδο, δηλαδή αν σημαντικότερος παράγοντας είναι η μήτρα, τότε έχουμε τη μέση ανάκτηση για κάθε μήτρα.

l : ο αριθμός των ομάδων στο πρώτο επίπεδο ιεράρχησης.

Η $u(\bar{R}_i)$ μπορεί να υπολογιστεί με δύο τρόπους:

1ος τρόπος: Ως η τυπική απόκλιση του μέσου όρου της κάθε ομάδας:

$$u(\bar{R}_i) = \sqrt{\frac{\sum_{j=1}^p (\bar{R}_{ij} - \bar{R}_i)^2}{p(p-1)}} \quad (8)$$

όπου:

\bar{R}_{ij} , ο μέσος όρος κάθε ομάδας στο δεύτερο επίπεδο ιεράρχησης.

p , ο αριθμός των ομάδων στο δεύτερο επίπεδο ιεράρχησης.

2ος τρόπος: Μέσα από την αναπαραγωγιμότητα (ή ενδιάμεση πιστότητα):

$$u(R_t) = \frac{RSD_t}{\sqrt{nr}} \quad (9)$$

όπου:

RSD_i , η τυπική απόκλιση της αναπαραγωγιμότητας.

n , ο αριθμός διαφορετικών ημερών ανάλυσης (ή αναλυτών).

Ανάλογα του πώς έχει πραγματοποιηθεί το πείραμα, δηλ. σε συνθήκες αναπαραγωγιμότητας ή επαναληψιμότητας.

r , ο αριθμός επαναλήψεων για κάθε ημέρα ανάλυσης ή για κάθε αναλύτη.

Έλεγχος σημαντικότητας

Στο σημείο αυτό που έχουμε την ολική αβεβαιότητα μπορεί να ελεγχθεί αν η ανάκτηση έχει στατιστικά σημαντική διαφορά από το 1 με βάση των παρακάτω τύπο:

$$\frac{\left| \overline{R_m} - 1 \right|}{u(R_m)} \leq k \quad (10)$$

Όπου: k , συντελεστής κάλυψης για (1- α) στάθμη εμπιστοσύνης. Αν ο αριθμός των μετρήσεων είναι μικρός ($n < 10$) τότε αντί τον παράγοντα κάλυψης, χρησιμοποιείται η τιμή της κατανομής t-student για την αντίστοιχη στάθμη εμπιστοσύνης και τους δραστικούς βαθμούς ελευθερίας.

Υπολογισμός αβεβαιοτήτων συγκέντρωσης και μάζας

Ο υπολογισμός των αβεβαιοτήτων συγκέντρωσης και μάζας θα γίνει μέσω των διακυμάνσεων που προκύπτουν από την ANOVA. Θεωρώντας ότι η σειρά ιεράρχησης είναι συγκέντρωση > μήτρα > επαναλήψεις, τότε ο υπολογισμός γίνεται με βάση τις παρακάτω εξισώσεις.

$$u(C)^2 = \frac{MS_C - MS_M}{rp} \quad (11)$$

$$u(M)^2 = \frac{MS_M - MS_r}{r} \quad (12)$$

Παρατηρήσεις:

- i. Αν η σειρά ιεράρχησης αλλάξει τότε οι παραπάνω εξισώσεις αντιστρέφονται.
- ii. Στην περίπτωση το πείραμα έχει γίνει σε συνθήκες αναπαραγωγιμότητας (δηλ. είτε σε διαφορετικές μέρες, είτε από διαφορετικούς αναλυτές) τότε έχουμε τέσσερα επίπεδα ιεράρχησης αντί τριών και μπορεί από εκεί να γίνει εκτίμηση της αβεβαιότητας του τυχαίου σφάλματος από το συμπέρασμα που εξάγεται από την ANOVA. Οι εξισώσεις παρατίθενται παρακάτω:

$$u(D)^2 = \frac{MS_D - MS_r}{r} \quad (13)$$

$$u(r)^2 = MS_r \quad (14)$$

όπου:

$u(D)$, η αβεβαιότητα που προκύπτει από τις διαφορετικές μέρες.

$u(r)$, η αβεβαιότητα από κάθε επανάληψη.

Οπότε η αβεβαιότητα του τυχαίου σφάλματος (αναπαραγωγιμότητα) $u(R)$ προκύπτει από την παρακάτω εξίσωση:

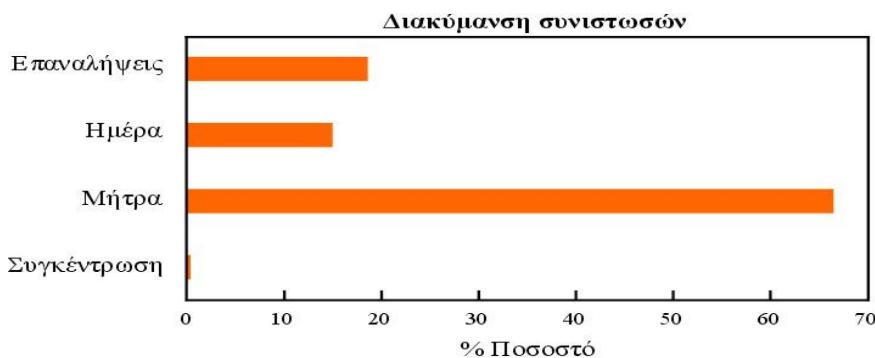
$$u(R)^2 = u(D)^2 + u(r)^2 \quad (15)$$

Στη συνέχεια, από τις Εξισώσεις 7 και 9 υπολογίζεται η αβεβαιότητα της συνολικής ανάκτησης.

4. Παράδειγμα εφαρμογής

Η τυλοσίνη (tylosin) είναι αντιβιοτικό που ανήκει στην κατηγορία των μακρολιδίων (macrolides). Τα μακρολίδια είναι ομάδα αντιβιοτικών που έχουν αντιβακτηριακές ιδιότητες και δρουν κατά των κατά Gram θετικών και ορισμένων κατά Gram αρνητικών βακτηριδίων. Χρησιμοποιούνται ευρέως στην ιατρική και την κτηνιατρική πρακτική. Λανθασμένη χρήση τους θα έχει ως αποτέλεσμα να βρεθούν υπολείμματα στους βρώσιμους ιστούς ή στα τρόφιμα, γεγονός που είναι επικίνδυνο για τους καταναλωτές λόγω της αλλεργικής αντίδρασης που μπορούν να προκαλέσουν οι ίδιες ή οι μεταβολίτες τους με τον ανθρώπινο οργανισμό (Wang et al 2005). Ο προσδιορισμός της τυλοσίνης έγινε σε 4 επίπεδα συγκεντρώσεων ($l = 4$) για 3 διαφορετικές μήτρες ($p = 3$), δύο διαφορετικές μέρες για κάθε επίπεδο και μήτρα ($n = 2$) και 3 επαναλήψεις για κάθε μέρα ($r = 3$). Τα δεδομένα που παρουσιάζονται είναι από τη βιβλιογραφία και τα πρωτογενή δεδομένα παρουσιάζονται στο Παράρτημα (Wang et al 2012).

Στο Σχήμα 2 φαίνεται διαγραμματικά η διακύμανση της κάθε συνιστώσας, όπως προέκυψε από το λογισμικό, θέτοντας τις συνιστώσες του συστηματικού σφάλματος ως εξής: συγκέντρωση > μήτρα > διαφορετικές μέρες > επαναλήψεις.



Σχήμα 2. Η ποσοστιαία διακύμανση κάθε επιπέδου (σειρά ιεράρχησης: συγκέντρωση > μήτρα > διαφορετικές ημέρες > επαναλήψεις).

Από το παραπάνω σχήμα παρατηρούμε ότι η μεγαλύτερη διακύμανση οφείλεται στο υλικό της μήτρας, ακολουθούν η επαναληψιμότητα και η μεταξύ ημερών διακύμανση, ενώ η συγκέντρωση έχει μηδενική διακύμανση. Όσον αφορά τη συνιστώσα της συγκέντρωσης από το πρόγραμμα προέκυψε ότι η ποσοστιαία διακύμανση της συγκέντρωσης είναι αρνητική. Το φαινόμενο αυτό συμβαίνει σε περιπτώσεις όπου η διακύμανση του ενός παράγοντα είναι μικρότερη από αυτού που βρίσκεται χαμηλότερα ιεραρχικά, άρα το *MS* του ενός είναι μικρότερο του άλλου, οπότε και η εκτιμώμενη διακύμανση που προκύπτει είναι αρνητική (Vander Heyden, et al, 1999). Η σειρά ιεράρχησης που ακολουθήθηκε είναι αυτή που προτείνεται στις περισσότερες μελέτες. Ωστόσο, το ίδιο πρόβλημα αντιμετωπίζεται στις περισσότερες των περιπτώσεων, γι' αυτό και έχει προταθεί να τοποθετείται το μητρικό υλικό στο υψηλότερο ιεραρχικό επίπεδο και μετά να ακολουθεί η συγκέντρωση (Boti, Sakkas, & Albanis, 2007). Όπως είναι φανερό από το Σχήμα 3, το ίδιο πρόβλημα αντιμετωπίζεται και στην περίπτωση αυτή. Οπότε θα γίνει η ίδια μετατροπή στη σειρά ιεράρχησης, αν και λόγω της συμμετρίας του πειραματικού σχεδιασμού δεν θα πρέπει να υπάρχουν σημαντικές διαφορές μεταξύ των δύο αποτελεσμάτων (Vander Heyden et al 1999). Μετά από την παραπάνω τροποποίηση στην κατάταξη των παραγόντων προκύπτει αντίστοιχα ο Πίνακας 1 και το Σχήμα 3:

Πίνακας 1. Συμπέρασμα που εξάγεται από την εφαρμογή της nested ANOVA στο MINITAB (σειρά ιεράρχησης: μήτρα>συγκέντρωση>διαφορετικές μέρες>επαναλήψεις).

Source	<i>d.f.</i> ^α	Level	<i>SS</i> ^β	<i>MS</i> ^γ	<i>F</i>	<i>P</i>
Matrix	2	<i>l</i> = 3	0,1553	0,0776	77,467	0,000
Concentration	9	<i>p</i> = 4	0,0090	0,0010	0,364	0,931
Day	12	<i>n</i> = 2	0,0330	0,0028	3,426	0,001
Error	48	<i>r</i> = 3	0,0385	0,0008		
Total	71		0,2359			

α: degrees of freedom (βαθμοί ελευθερίας), β: Sum of Squares, γ: Mean of Squares



Σχήμα 3. Η ποσοστιαία διακύμανση κάθε επιπέδου (σειρά ιεράρχησης: μήτρα>συγκέντρωση>διαφορετικές μέρες>επαναλήψεις).

Έτσι με βάση τα παραπάνω δεδομένα και τις εξισώσεις που παρατέθηκαν προκύπτουν τα παρακάτω:

Πίνακας 2. Πίνακας αποτελεσμάτων για την εκτίμηση των συνιστωσών και της συνολικής αβεβαιότητας του παραδείγματος.

Εκτιμώμενη παράμετρος	Αποτελέσματα*	Αναφορά
$u(\overline{R}_m)$	0,001866	Εξίσωση 7
$u(\overline{R}_{l=1})$	0,004240	Εξίσωση 8
$u(\overline{R}_{l=2})$	0,003002	Εξίσωση 8
$u(\overline{R}_{l=3})$	0,002081	Εξίσωση 8
$u(\overline{R}_m)^2$	0,000003	
$u(C)^2$	0,00000	Εξίσωση 11
$u(M)^2$	0,003192	Εξίσωση 12
$u(R)$	0,0567	Εξίσωση 6

*τα πρωτογενή δεδομένα υπολογίστηκαν ως ανακτήσεις και επομένως μιλάμε για σχετικές αβεβαιότητες.

Στον Πίνακα 2 ο υπολογισμός της ολικής ανάκτησης έγινε μέσω του υπολογισμού της τυπικής απόκλισης του μέσου όρου (Εξίσωση 8). Στον Πίνακα 3 που ακολουθεί φαίνονται τα αποτελέσματα και οι υπολογισμοί της αβεβαιότητας της ολικής ανάκτησης μέσω της αναπαραγωγιμότητας.

Πίνακας 3. Αποτελέσματα υπολογισμού αβεβαιότητας ολικής ανάκτησης $u(\overline{R}_m)$ μέσω αναπαραγωγιμότητας.

Εκτιμώμενη Παράμετρος	$l = 1$	$l = 2$	$l = 3$	$l = 4$	Αναφορά
MS_D	0,0035	0,0036	0,0008	0,0031	
MS_i	0,0006	0,0010	0,0012	0,0004	
$u(D)^2$	0,0010	0,0009	0,0000	0,0009	Εξίσωση 13
$u(r)^2$	0,0006	0,0010	0,0012	0,0004	Εξίσωση 14
$u(R)^2$	0,0016	0,0019	0,0012	0,0013	Εξίσωση 15
$u(\overline{R}_i)^2$	0,0001	0,0001	0,0001	0,0001	Εξίσωση 9
$u(\overline{R}_m)^2$	0,0000206				Εξίσωση 7
$u(\overline{R}_m)$	0,0045				

Από το παραπάνω αποτέλεσμα και την εξίσωση (6) υπολογίζεται η αβεβαιότητα της συνολικής ανάκτησης μέσω αναπαραγωγιμότητας. Παρακάτω, παρατίθενται τα αποτελέσματα με τις δύο διαφορετικές μεθόδους υπολογισμού:

Πίνακας 4. Αβεβαιότητα συνολικής ανάκτησης μέσω τυπικής απόκλισης μέσου όρου και μέσω αναπαραγωγιμότητας.

	Αβεβαιότητα συνολικής ανάκτησης
Μέσω τυπικής απόκλισης μέσου όρου	0,0565
Μέσω αναπαραγωγιμότητας	0,0567

Στον Πίνακα 4 φαίνεται ότι η διαφορά μεταξύ των δύο αβεβαιοτήτων είναι πολύ μικρή και είναι αναμενόμενο, καθώς ο παράγοντας που επιδρά περισσότερο στη συνολική αβεβαιότητα του συστηματικού σφάλματος είναι η διακύμανση του μητρικού υλικού, ενώ η αβεβαιότητα της συνολικής ανάκτησης είναι αμελητέα. Τέλος από την Εξίσωση 10 υπολογίζεται η πειραματική τιμή του t-test (και με τους δύο τρόπους υπολογισμού της αβεβαιότητας), ώστε να εκτιμηθεί αν υπάρχει στατιστικά σημαντική διαφορά από το 1 οπότε και αν απαιτείται διόρθωση του αποτελέσματος. Η σύγκριση γίνεται μέσω του 1,96 που είναι ο συντελεστής κάλυψης για στάθμη εμπιστοσύνης 95%. Στον Πίνακα 5 φαίνονται τα αποτελέσματα.

Πίνακας 5. Αποτελέσματα t-test με βάση την εξίσωση (12) για τους δύο τρόπους υπολογισμού της συνολικής ανάκτησης.

	t-test
Μέσω τυπικής απόκλισης μέσου όρου	8,43
Μέσω αναπαραγωγιμότητας	3,46

Και με τους δύο τρόπους το συμπέρασμα που εξάγεται από το t-test είναι ότι απορρίπτεται η μηδενική υπόθεση, οπότε θεωρητικά θα πρέπει στα αποτελέσματα να εφαρμοστεί κάποιος συντελεστής διόρθωσης. Ωστόσο, όπως φαίνεται, τα αποτελέσματα είναι στην περιοχή ανακτήσεων 80-120% και μάλιστα πολύ κοντά στο 100%, ενώ η αβεβαιότητα είναι πολύ μικρή. Επομένως η μηδενική υπόθεση απορρίπτεται λόγω της μικρής συνολικής αβεβαιότητας. Οπότε δεν θεωρείται απαραίτητο να εισαχθεί κάποιος συντελεστής διόρθωσης.

5. Συμπεράσματα

Καταλήγοντας, παρουσιάστηκε ένα εναλλακτικός τρόπος για την εκτίμηση της αβεβαιότητας του συστηματικού σφάλματος. Με βάση την παραπάνω διαδικασία μπορεί να προκύψει ότι:

- Αποτελεί ένα τρόπο για την εκτίμηση της αβεβαιότητας με μια πιο βαθύτερη εικόνα για τους παράγοντες που δημιουργούν το συστηματικό σφάλμα, κάτι που πολλές φορές είναι αναγκαίο και κυρίως με τις σύγχρονες τεχνικές ανάλυσης (π.χ. φασματομετρία μαζών).
- Πέρα από τον υπολογισμό της αβεβαιότητας, ο παραπάνω πειραματικός σχεδιασμός μπορεί να εφαρμοστεί, ώστε να γίνει συνολικά η εκτίμηση των αναλυτικών παραμετρών μιας μεθόδου συμπεριλαμβάνοντας διαφορετικές μήτρες, διαφορετικά επίπεδα συγκεντρώσεων και της αναπαραγωγιμότητας συνολικά, σε αντίθεση με την κοινή πρακτική όπου αυτά θα πρέπει να εκτιμηθούν με διαφορετικές μελέτες, για κάθε μητρικό υλικό.

- Η εκτίμηση του συντελεστή διόρθωσης μέσω του υπολογισμού του t-test γίνεται για όλο τον πειραματικό σχεδιασμό και όχι για ξεχωριστά, για κάθε επίπεδο συγκεντρώσεων ή για κάθε μητρικό υλικό.

Παράλληλα θα μπορούσαν να παρατηρηθούν τα εξής μειονεκτήματα:

- Ο πειραματικός σχεδιασμός και ο υπολογισμός είναι πολυπλοκότερος σε σχέση με την κοινή αναλυτική πρακτική, οπότε απαιτεί περισσότερο χρόνο για τον προσεκτικό σχεδιασμό του πειράματος.
- Ο αριθμός των πειραμάτων που πρέπει να πραγματοποιηθούν είναι μεγαλύτερος, το οποίο αν και είναι ανασταλτικός παράγοντας ανάλογα με τη φύση του πειράματος, ωστόσο λόγω της εξέλιξης του τεχνολογικού εξοπλισμού ο φόρτος εργασίας είναι μειωμένος για τον αναλυτή σε σχέση με το παρελθόν.
- Απαιτείται λογισμικό διαφορετικό του excel, που πολλές φορές δεν είναι διαθέσιμο και εύκολο στη χρήση για τον αναλυτή.

Παράρτημα

Πρωτογενή δεδομένα του παραδείγματος από την βιβλιογραφία (Wang et al 2012)

	Ημέρα Ανάλυσης	Επίπεδο Συγκέντρωση 1	Επίπεδο Συγκέντρωση 2	Επίπεδο Συγκέντρωση 3	Επίπεδο Συγκέντρωση 4
Μήτρα 1	1	0,894	0,899	0,907	0,948
		0,924	0,937	0,908	0,913
		0,932	0,912	0,931	0,878
	2	0,978	0,949	0,933	0,935
		1,019	1,014	0,939	0,93
		0,984	0,969	0,953	0,903
Μήτρα 2	1	1,086	1,052	1,11	1,001
		1,072	1,041	1,02	1,029
		1,076	1,03	1,02	1,043
	2	1,091	0,991	0,992	1,078
		1,009	1,036	1,088	1,055
		1,053	1,09	0,999	1,076
Μήτρα 3	1	0,934	0,967	0,953	0,952
		0,978	0,931	0,936	0,942
		0,954	0,949	0,963	0,925
	2	0,971	0,97	0,981	0,994
		0,994	0,999	0,935	1,011
		0,948	1,055	0,993	1,006

6. Βιβλιογραφία

- Barwick V. J. and Pichard E. “*Eurachem Guide: Terminology in Analytical Measurement - Introduction to VIM 3*”, 2011.
- Boti V. I., Sakkas V. A. and Albanis T. A., “*Measurement uncertainty arising from trueness of analysis of two endocrine disruptors and their metabolites in environmental samples part I: Ultrasonic extraction from diverse sediment matrices*”, *Journal of Chromatography A*, 1146, pp 139-147, 2007.
- Dehouck P., Van Looy E., Haghedooren E., Deckers K., Vander Heyden Y., Adams E., Roets E. and Hoogmertens J., “*Analysis of erythromycin and benzoylperoxide in topical gels by liquid chromatography*”, *Journal of Chromatography B*, 794, pp 293-302, 2003.
- Linsinger P.J. Thomas, “*Use of recovery and bias information in analytical chemistry and estimation of its uncertainty contribution*”, *Trends in Analytical Chemistry*, 27, pp 916-923, 2008.
- Magnusson B. and Ellison L.R.S., “*Treatment of uncorrected measurement bias in uncertainty estimation for chemical measurements*”, *Analytical and Bioanalytical Chemistry*, 390, pp 201-213, 2008.
- Magnusson B., Naykk T., Hovind H., and Krysell M., “*Handbook for calculation of measurement uncertainty in environmental laboratories (NT TR 537 - Edition 3.1)*”, 2012 Nordtest.
- Maroto A., Boque R., Riu J. and Rius Xavier F., “*Should non-significant be included in the uncertainty budget?*”, *Accreditation and Quality Assurance*, 7, pp 90-94, 2002.
- Miller N. J. and Miller C. J., “*Statistics and Chemometrics for Analytical Chemistry*”, 2010, Pearson Education Limited.
- Thompson M., Ellison L. R. S., Fajgelj A., Willetts P. and Wood R., “*Harmonised Guidelines for the use of recovery information in Analytical Measurement*”, *Pure and Applied Chemistry*, 71, pp 337-348, 1999.
- Vander Heyden Y., De Braekeleer K., Zhu Y., Roets E., Hoogmartens J., De Beer J., and Massart L. D., “*Nested designs in ruggedness testing*”, 20, pp 875-887, 1999.
- Wang J., Leung D. and Butterworth F., “*Determination of five Macrolide Antibiotic Residues in Eggs using Liquid Chromatography/Electrospray Ionization Tandem Mass Spectrometry*”, *Journal of Agricultural and Food Chemistry*, 53, pp 1857-1865, 2005.
- Wang J., MacNeil D. J. and Kay F. J., “*Chemical Analysis of Antibiotic Residues in Food*”, 2012, Wiley.