

Στατιστική, μηδενική υπόθεση και τεκμήριο αθωότητας...

/ [Πεμπτούσια](#)



2fe9fc5c-f665-41b5-9947-0abe3b7c9030

Ο στατιστικολόγος

ενδιαφέρεται για τις μετρήσεις ή τα γεγονότα με βάση την ιδέα ότι απορρέουν από ένα θεμελιώδες φαινόμενο και από ένα συνδυασμό «ανεξέλεγκτων» παραγόντων διαταραχής. Πράγμα που σημαίνει ότι επαναλαμβάνοντας τις παρατηρήσεις, υπό σταθερές συνθήκες, τα διαδοχικά αποτελέσματα θα διαφέρουν «λιγότερο ή περισσότερο» μεταξύ τους.

Το πιο απλό σχήμα είναι το «γραμμικό ή προσθετικό πρότυπο». Μπορεί να γίνει πιο περίπλοκο (πολλαπλασιαστικό ή κάτι άλλο), όμως τίποτε δεν αλλάζει επί της ουσίας. Είναι λοιπόν επαρκές για να κατανοηθεί το ζήτημα του τυχαίου για τον στατιστικολόγο. Έτσι λοιπόν, μια μέτρηση m θεωρείται ως το αποτέλεσμα ενός «κύριου» φαινομένου, η οποία δίνει μια τιμή μ στην οποία προστίθεται ένα «σφάλμα» e το οποίο δημιουργείται από «ανεξέλεγκτους» παράγοντες: $m = \mu + e$ (η ποσότητα e μπορεί να είναι θετική ή αρνητική). Το θεμελιώδες ζήτημα της στατιστικής είναι η «εκτίμηση» του μ και η δυνατότητα να μπορεί να πει πότε μια μέτρηση ή μια εκτίμηση διαφέρει από άλλες που ελήφθησαν υπό παρεμφερείς συνθήκες. Το τυχαίο λοιπόν παρεμβαίνει στον όρο e . Ο όρος αυτός είναι ανάλογος του όρου σφάλμα των φυσικών, και εδώ ξαναβρίσκουμε την κατά Laplace οπτική του τυχαίου η οποία απορρέει από ένα σύνολο μη γνωστών παραγόντων.

Ας πάρουμε ένα κλασικό παράδειγμα από τον χώρο της αγρονομίας, υπογραμμίζοντας με την ευκαιρία ότι κατά μεγάλο μέρος η στατιστική αναπτύχθηκε με αφορμή πολύ συγκεκριμένα προβλήματα, πολλά από τα οποία εισέφερε αυτός ο επιστημονικός και τεχνικός τομέας⁶. Ας υποθέσουμε ότι θέλουμε να μάθουμε αν ένα χημικό προϊόν είναι λίπασμα, αν επιδρά στη φυτική παραγωγικότητα αυξάνοντας την απόδοση μιας καλλιέργειας. Για τον σκοπό αυτό μπορούμε να πραγματοποιήσουμε ένα απλό πείραμα: ένα τμήμα γης μάρτυρας, χωρίς λίπασμα, και ένα τμήμα λιπασμένο, καλλιεργούνται ταυτόχρονα. Θα λάβουμε έτσι δυο αποδόσεις m_0 και m_1 . Μπορούμε όμως να πούμε ότι το m_1 είναι

«σημαντικά διαφορετικό» από το m_0 ; Και ότι το λίπασμα έχει μια επίδραση, ακόμα και αν το m_1 είναι μεγαλύτερο από το m_0 ;

Πράγματι, αντιλαμβανόμαστε γρήγορα ότι αν πραγματοποιήσουμε πολλούς μάρτυρες θα ληφθούν διαφορετικά αποτελέσματα. Το ίδιο ισχύει και για τα λιπασμένα τμήματα γης. Εκτός από εξαιρετικές περιπτώσεις, όπως για παράδειγμα αν οι αποδόσεις των πειραματικών τμημάτων γης είναι όλες ανώτερες από τους μάρτυρες, θα είναι δύσκολο να εξαχθούν συμπεράσματα. Οι διαφορές μεταξύ μαρτύρων ή μεταξύ πειραματικών τμημάτων γης αποδίδονται σε «ανεξέλεγκτους παράγοντες» -δεδομένου ότι μόνο οι αντίστοιχες ποσότητες σπόρων και λιπασμάτων είναι γνωστές- για παράδειγμα εδάφη εκ φύσεως λιγότερο ή περισσότερο γόνιμα ή μια γενετική ετερογένεια των σπερμάτων ή ακόμα ελαφρά διαφορετικές μικροκλιματικές συνθήκες.

Για να πουν αν οι παρατηρούμενες διαφορές μπορούν ή όχι να αποδοθούν στην επεξεργασία, οι στατιστικολόγοι εκπονούν στατιστικές δοκιμασίες.

Η βασική αρχή συνίσταται στη δημιουργία ενός μαθηματικού προτύπου της διασποράς των αποτελεσμάτων, υπό την αποκαλούμενη «μηδενική υπόθεση», δηλαδή ότι οι διαφορές θεωρούνται «μη σημαντικές», οφειλόμενες μόνο στο τυχαίο, σε έναν «θόρυβο» που επάγεται από την πολλαπλότητα των επιδράσεων των ανεξέλεγκτων παραγόντων. Οι διαφορές αυτές ερμηνεύονται ως σφάλματα μέτρησης.

Το πρότυπο επιτρέπει την αποτίμηση της πιθανότητας του παρατηρούμενου αποτελέσματος. Αν αυτή η πιθανότητα είναι μικρή, για παράδειγμα μικρότερη του 0,05, ελάχιστο όριο που ορίζεται a priori, η μηδενική υπόθεση απορρίπτεται και συμπεραίνουμε ότι υπάρχει σημαντική διαφορά μεταξύ των υπό μελέτη τιμών και άρα ότι η χρήση λιπάσματος έχει επίδραση.

Θα σημειώσουμε με την ευκαιρία αυτή ότι, ακόμα και εκτός του τομέα της στατιστικής, πολλοί από τους συλλογισμούς μας και τα πρότυπα «απόφασης» που ακολουθούμε βασίζονται στο ανάλογο των προτύπων της στατιστικής που δημιουργήθηκαν με βάση αυτή τη μηδενική υπόθεση. Τα πρότυπα αυτά είναι ισχυρά έναντι της απόρριψης αυτής της υπόθεσης.

Αυτό συμβαίνει για παράδειγμα στις δικαστικές έρευνες όπου το ανακρινόμενο πρόσωπο θεωρείται a priori αθώο (μηδενική υπόθεση). Ο εισαγγελέας θα απορρίψει αυτή την υπόθεση αν τουλάχιστον ένα γεγονός τη διαψεύδει.

Στην πραγματικότητα, το τυχαίο του στατιστικολόγου είναι ένα είδος «καλάθου αχρήστων» στον οποίο τοποθετούνται όλα όσα δεν γνωρίζουμε ή όσα αποφασίσαμε

να αγνοήσουμε και τα θεωρήσαμε «δεύτερης κατηγορίας», σε σχέση με το κύριο φαινόμενο το οποίο μελετάμε. Αυτή η «δεύτερη κατηγορία» λειτουργεί με ένα φαινομενικά διαλείποντα, τυχαίο τρόπο, ως ένα είδος θορύβου.

Με δεδομένα τα προαναφερθέντα, αν και η άποψη αυτή περί του τυχαίου είναι ακόμα ευρέως αποδεκτή από τη μεγάλη πλειοψηφία, οι στατιστικολόγοι ενδιαφέρονται επίσης για τις λεγόμενες «στοχαστικές» διεργασίες, δηλαδή τα γεγονότα που διαδραματίζονται στον χρόνο και τον χώρο.

Πρόκειται για παράδειγμα για την περίπτωση των διακυμάνσεων των τιμών του Χρηματιστηρίου ή για την κατανομή των ατόμων σε έναν γεωγραφικό χώρο. Οι μέθοδοι που διαμορφώνουν επιτρέπουν να δοκιμαστούν για παράδειγμα οι υποθέσεις ανεξαρτησίας των ημερήσιων τάσεων των τιμών ή η ομοιομορφία της κατανομής στον χώρο των ατόμων (μηδενικές υποθέσεις). Επιπλέον μπορούν να συνεισφέρουν καθοριστικά στην εκτίμηση των παραμέτρων που αφορούν διεργασίες αυτού του τύπου ή ακόμα αιτιοκρατικά πρότυπα (θα δούμε ένα τέτοιο παράδειγμα στο κεφάλαιο 4).

Σημείωση 6: Ο Sir Ronald Fisher, χωρίς αμφιβολία από τους πιο διάσημους στατιστικολόγους του 20ου αιώνα, πραγματοποίησε ένα μεγάλο μέρος των εργασιών του με βάση τέτοια παραδείγματα. Υπήρξε επίσης γενετιστής των πληθυσμών και από τους πρώτους δημιουργούς της συνθετικής θεωρίας της εξέλιξης (είδε κεφάλαιο 2, τμήμα V). Εργάστηκε για μεγάλο χρονικό διάστημα στον αγρονομικό σταθμό του Rothamsted στη Μεγάλη Βρετανία.

Παρατήρηση: το παρόν άρθρο αποτελεί απόσπασμα του βιβλίου: Η ΑΝΑΓΚΑΙΟΤΗΤΑ ΤΟΥ ΤΥΧΑΙΟΥ

Συγγραφέας: ALAIN PAVÉ, ISBN 978-960-531-282-4

Σχήμα: 17Χ24, Σελ. : 192, Τιμή: 21,30 ευρώ (με ΦΠΑ 6.5%)

Εκδόσεις: ΔΙΑΥΛΟΣ (www.diavlos-books.gr)