

Ζάρια, μεσημβρινοί, γονίδια και γεωγραφίες. Μια σύντομη ιστορία της στατιστικής

Δημήτριος Ρωσσικόπουλος

Καθηγητής ΑΠΘ, Τομέας Γεωδαισίας και Τοπογραφίας, Πολυτεχνική Σχολή, ΤΑΤΜ - ΑΠΘ.
rossi@auth.gr

Περίληψη: Πως συνδέονται τα τυχερά παιχνίδια με τις λίστες θανάτου του μεσαίωνα; Οι λίστες θανάτου με τη διοίκηση των κρατών και με το έργο των γεωδαιτών και των αστρονόμων; Οι γεωδαιτικές μετρήσεις με τα γονίδια και τα πειράματα των βιολόγων; Το κλειδί για τη σύνδεσή τους είναι “ο νόμος των σφαλμάτων”. Στην εργασία αυτή περιγράφεται ένα ταξίδι στις ιστορίες της στατιστικής από τα καφενεία του μεσαίωνα ως τα βιομετρικά εργαστήρια των αρχών του 20ου αιώνα. Και κάπου στα μισά της διαδρομής, στις αρχές του 19ου αιώνα, με το πρωτοποριακό έργο των αστρονόμων και των γεωδαιτών πάνω στον βέλτιστο συνδυασμό των μετρήσεων, τη διατύπωση του νόμου των σφαλμάτων και την ανάπτυξη της μεθόδου των ελαχίστων τετραγώνων, η έννοια της στατιστικής διευρύνθηκε ώστε να συμπεριλάβει τη συλλογή, την ταξινόμηση και την ανάλυση παρατηρήσεων κάθε είδους, όπως είναι οι μετρήσεις μεγεθών, η συλλογή δειγματικών τιμών, τα αποτελέσματα πειραμάτων κλπ. Σήμερα, οι τεχνικές της στατιστικής εφαρμόζονται στις δημόσιες υπηρεσίες και στη διοίκηση των κρατών, στην πολιτική και στις επιχειρήσεις κάθε μορφής, στις κοινωνικές και οικονομικές επιστήμες, στις ιατρικές επιστήμες και στο σύνολο γενικά των επιστημών που βασίζονται στο πείραμα και στην παρατήρηση, καθώς και στην καθημερινή ζωή των ανθρώπων.

Λέξεις κλειδιά: Ιστορία της στατιστικής, Ιστορία των στατιστικών ελέγχων, Μέθοδος των ελαχίστων τετραγώνων.

1. Εισαγωγή

Ο όρος *statistik* χρησιμοποιήθηκε για πρώτη φορά το 1749 από τον γερμανό λόγιο Gottfried Achenwall για τη νεοϊδρυθείσα στατιστική υπηρεσία του πρωσικού κρατιδίου, που ιδρύθηκε σύμφωνα με τα φιλοσοφικά πρότυπα του Αριστοτέλη, όπως αυτά διατυπώνονται στο έργο του *Πολιτείες*. Περιελάμβανε τη συλλογή, την περιγραφή και την ανάλυση πληθυσμιακών δεδομένων του κρατιδίου, κάτι που έχει μικρή σχέση με την έννοια της σημερινής στατιστικής. Από τον Αριστοτέλη ο όρος *πολιτεία* εισήλθε επίσης στη λατινική γλώσσα με τη φράση *statisticum collegium* (συμβούλιο της πολιτείας), από την οποία προήλθε με τη σειρά της και η λατινική λέξη *status* που πολλοί τη θεωρούν ως ρίζα του γερμανικού όρου *statistik* και του όρου *Staatwissenschaft*, που δηλώνει τις επιστήμες του κράτους. Αλλά και ο όρος *status*, που φαίνεται να προκύπτει ως μετάφραση της λέξης *πολιτεία*, πιθανότατα

ετυμολογείται από το αρχαίο ελληνικό ρήμα *ίστημι* και του εξ αυτού παραγώγου ρήματος *στατίζω* που σημαίνει τοποθετώ, ταξινομώ, συμπεραίνω.

Κατά τον 18ο αιώνα, με τον όρο "στατιστική" εννοείται η συστηματική συλλογή δημογραφικών και οικονομικών δεδομένων από τις κυβερνήσεις των κρατών. Για τουλάχιστον δύο χιλιετίες, τα στοιχεία αυτά ήταν κυρίως απογραφές και πινακοποιημένες τιμές μεγεθών που σχετίζονται με ανθρώπινους και υλικούς πόρους, ώστε να φορολογηθούν ή να τεθούν σε στρατιωτική χρήση. Στοιχειώδεις τέτοιες απογραφές εμφανίζονται από τους μυθικούς χρόνους και φαίνεται να τις έχουν πραγματοποιήσει αρχαίοι λαοί όπως οι Κινέζοι, οι Αιγύπτιοι, οι Βαβυλώνιοι, οι Πέρσες, οι Έλληνες και οι Ρωμαίοι. Η πιο συστηματική όμως συλλογή πληθυσμιακών και οικονομικών δεδομένων άρχισε στη Βενετία και την Φλωρεντία κατά τη διάρκεια της Αναγέννησης και γρήγορα επεκτάθηκε σ' όλα τα Βασίλεια της Ευρώπης. Στα επόμενα χρόνια οι μεγάλοι ρυθμοί θνησιμότητας εξαιτίας των επιδημικών ασθενειών, των πολέμων και των λιμοκτονιών έδωσαν ιδιαίτερη ώθηση στη καταγραφή των αιτιών και των απωλειών. Έτσι, το 1348 ξεκίνησαν οι καταγραφές θανάτων από την πανώλη, της μαστίγας που διήρκεσε τέσσερις αιώνες, ώστε να βοηθηθούν οι αξιωματούχοι στη δημιουργία ενός συστήματος προειδοποίησης σχετικά με την εκδήλωση και την εξάπλωση της φοβερής αυτής ασθένειας.

Η συνδυαστική, ένας κλάδος των μαθηματικών που αποτέλεσε σημαντικό εργαλείο στην ανάπτυξη της θεωρίας των πιθανοτήτων ως μελέτη των συνδυασμών και των μεταθέσεων αντικειμένων ή γεγονότων, έχει και αυτή πολύ μακριά ιστορία. Φαίνεται να ξεκινά με τα μαγικά τετράγωνα που αναφέρονται σ' ένα παλιό κινέζικο βιβλίο του 2200 π.Χ., το *Βιβλίο των αλλαγών*, και στις ιερές *Βέδες* και σε άλλα έργα ινδών μαθηματικών, όπως του Mahāgavirācaṅga, περίπου το 850, ή τη *Χαριτωμένη* του Bhāscara, περίπου το 1150, όπου αναφέρονται μερικοί στοιχειώδεις κανόνες συνδυασμών και μεταθέσεων. Οι αρχαίοι λαοί διερευνούσαν τα ζητήματα αυτά για μυστικιστικούς και κοσμολογικούς λόγους, ενώ για το σύνολο σχεδόν των μεταγενέστερων συγγραφέων της Δύσης το κίνητρο ήταν η πρόβλεψη πιθανών εκβάσεων των τυχερών παιχνιδιών, καθώς μέσω των συνδυασμών και των μεταθέσεων γίνεται η απαρίθμηση πιθανοτήτων.

Το *τρίγωνο του Pascal*, ή πιο σωστά το *αριθμητικό τρίγωνο*, είναι μια τριγωνική διάταξη των διωνυμικών συντελεστών και εφαρμόζει κατ' εξοχή σε ζητήματα που τέθηκαν στον μεγάλο μαθηματικό για τα τυχερά παιχνίδια. Συναντάται για πρώτη φορά κατά τον 13 αιώνα στη διδασκαλία του Πέρση μαθηματικού γνωστού ως Nasir ed Dim και στο κινέζικο σύγγραμμα του 1303, *Πολύτιμο κάτοπτρο των τεσσάρων στοιχείων*. Το πρώτο ερώτημα που τέθηκε στον Blaise Pascal είναι γνωστό ως *πρόβλημα των πόντων* και διατυπώνεται ως εξής: Πως πρέπει να μοιραστεί στους παίκτες το κατατιθέμενο ποσό της πόστας όταν διακόπτεται ένα παιχνίδι χωρίς να ολοκληρωθεί ο αριθμός των παρτίδων; Το δεύτερο ερώτημα, γνωστό ως *πρόβλημα του κύβου* συνίσταται στον προσδιορισμό της πιθανότητας να πετύχει κανείς με δύο ζάρια δύο εξάρεις, με τρία ζάρια τρεις εξάρεις, κλπ. Χάρis στα ερωτήματα αυτά

άρχισε να καλλιεργείται ζωηρό ενδιαφέρον για τη νέα κατηγορία ερευνών που οδήγησε στην ανάπτυξη ενός νέου κλάδου των μαθηματικών, του λογισμού των πιθανοτήτων.

2. Οι νόμοι της τύχης

Η στατιστική, όπως ορίστηκε κατά τον 18ο αιώνα, θα ξεφύγει από τον περιγραφικό χαρακτήρα της και θα θεμελιωθεί πάνω σε μια αυστηρή μαθηματική θεωρία με την ανάπτυξη ενός νέου σχετικά κλάδου των μαθηματικών, του λογισμού των πιθανοτήτων, οι απαρχές του οποίου βρίσκονται στα πανδοχεία, στα καφενεία και στα σαλόνια της αριστοκρατίας της Ευρώπης του 15ου και του 16ου αιώνα, όπου μανιώδεις και παθιασμένοι τζογαδόροι προσπαθούσαν να αυξήσουν τις πιθανότητες της επιτυχίας τους σε κάποιο παιχνίδι με χαρτιά και ζάρια καθώς σεβαστά ποσά άλλαζαν χέρια. Με τη συμμετοχή τους σε εκατοντάδες παιχνίδια και το αρρωστημένα έντονο ενδιαφέρον τους, αυτοί οι επαγγελματίες και συστηματικοί παίκτες αντιλαμβάνονταν την έννοια των εμπειρικών πιθανοτήτων ως συχνότητα εμφάνισης μιας πλευράς ενός ζαριού ή ενός χαρτιού της τράπουλας και γρήγορα συνειδητοποίησαν πως αν έπειθαν του μαθηματικούς να ασχοληθούν με τον υπολογισμό τους, τα τυχερά παιχνίδια θα απέβαιναν περισσότερο κερδοφόρα, ή τουλάχιστον λιγότερο ζημιογόνα. Από την ενασχόληση αυτή δεν κατάφεραν να ξεφύγουν ούτε οι πιο διάσημοι μαθηματικοί και επιστήμονες, όπως φαίνεται στα αποσπάσματα από το έργο του Γαλιλαίου *Σκέψεις για τα παιχνίδια με ζάρια*¹, όπου παρατίθενται μερικοί υπολογισμοί για τον καθορισμό ευνοϊκών περιπτώσεων σε παιχνίδι με τρεις κύβους. Οι υπολογισμοί αυτοί του εξασφάλισαν μια θέση στην προϊστορία ενός νέου κεφαλαίου των μαθηματικών, του λογισμού των πιθανοτήτων. Όμως το πρώτο ίσως αποφασιστικό βήμα στην ανάπτυξη του λογισμού των πιθανοτήτων έγινε με το *Βιβλίο για παιχνίδια με κύβους*² που γράφτηκε το 1526 από έναν φανατικό λάτρη των επιτραπέζιων παιχνιδιών και του ζαριού, τον μεγάλο μαθηματικό και επιφανή γιατρό του Μιλάνου Giralomo Cardano. Το πιο σημαντικό είναι το σχόλιό του για το παιχνίδι των ζαριών, όπου για πρώτη φορά διατυπώνεται η έννοια της πιθανότητας³. Δηλώνει ο Cardano ότι αν το ζάρι είναι αμερόληπτο οι έξι έδρες του έχουν την ίδια πιθανότητα εμφάνισης. Χρησιμοποιεί δε τον λόγο του πλήθους των ευνοϊκών προς τις μη ευνοϊκές εκβάσεις ενός παιχνιδιού για να εκφράσει την πιθανότητα εμφάνισης των ευνοϊκών ή μη ευνοϊκών περιπτώσεων. Τον

1 Galileo Galilei, *Sopra le scoperte dei dadi*, ... Εκδόθηκε μεταφρασμένο στα Αγγλικά από τον E. H. Thorne ως παράρτημα στο βιβλίο της F. N. David, *Games, Gods and Gambling*.

2 Giralomo Cardano (1526). *Liber de Ludo Aleae*. Μεταφράστηκε στα Αγγλικά από τον Sidney Henry Gould και περιέχεται στο *The Book on Games of Chance* (New York, 1961).

3 Από τον de Moivre χρησιμοποιήθηκε για πρώτη φορά ο λατινικός όρος *probabilitate* το 1711 ενώ η αγγλική της απόδοση *probability* χρησιμοποιήθηκε στα μαθηματικά για πρώτη φορά το 1718.

λόγο αυτόν τον ορίζει ως *odds*, όρος που χρησιμοποιείται μέχρι σήμερα στα τυχερά παιχνίδια και στον ιππόδρομο.

Περίπου εκατό χρόνια αργότερα, ένας διάσημος παίκτης τυχερών παιχνιδιών, ο γάλλος ευγενής Chevalier de Meré, αποφάσισε να συμβουλευτεί τον περίφημο μαθηματικό και φιλόσοφο Blaise Pascal, στέλνοντάς του στο Παρίσι ερωτήσεις για τα τυχερά παιχνίδια. Ο Pascal άρχισε να αλληλογραφεί με φίλους του μαθηματικούς, κυρίως με τον Pierre Fermat στην Τουλούζη⁴. Η αλληλογραφία αυτή αποτέλεσε την αρχή της θεωρίας πιθανοτήτων αν και στις εργασίες των δύο γάλλων μαθηματικών το ενδιαφέρον επικεντρωνόταν στα τυχερά παιχνίδια και στον τζόγο και όχι στην ανάπτυξη μιας μαθηματικής θεωρίας. Αυτά τα προβλήματα τα αντιμετώπιζαν συνήθως ως απλή εφαρμογή του συνδυαστικού λογισμού, με λόγους χωρίς να αναφέρεται ο όρος πιθανότητα.

Όταν ο μεγάλος ωρολογοποιός Christian Huygens επισκέφτηκε τη Γαλλία ενημερώθηκε για τα προβλήματα που τέθηκαν από τον γάλλο ευγενή και απασχολούσαν τον Pascal και τον Fermat. Επιστρέφοντας στην Ολλανδία άρχισε αμέσως να συγγράφει μια σύντομη πραγματεία που εκδόθηκε το 1657 με τίτλο *Περί της λογικής στα παιχνίδια με κύβους*⁵. Την πραγματεία αυτή την περιέλαβε στο πρώτο μέρος του έργου του *Τέχνη της υπόθεσης*⁶ ο Jacob Bernoulli, γενάρχης της δυναστείας των Bernoulli, εμπλουτισμένη με σημαντικά σχόλια. Το *Ars conjectandi* εκδόθηκε το 1713, οκτώ χρόνια μετά το θάνατο του Bernoulli, με την επιμέλεια του ανιψιού του Nicolai. Στο δεύτερο μέρος της εργασίας αυτής, έχοντας ο Bernoulli σταθερά στραμμένη την προσοχή του στον υπολογισμό των πιθανοτήτων, πραγματεύεται τον συνδυαστικό λογισμό και δίνει τους τύπους με τα αθροίσματα των διαδοχικών δυνάμεων των ακεραίων αριθμών με τους συντελεστές που παρέμειναν στην ιστορία των μαθηματικών γνωστοί ως *αριθμοί Bernoulli*. Για να αποδείξει δε τη χρησιμότητα όλων αυτών στη θεωρία των πιθανοτήτων, παρουσιάζει ως εφαρμογή τη λύση ενός ενδιαφέροντος προβλήματος που δανείζεται από την ιατρική. Στη συνέχεια του *Ars conjectandi* ο Bernoulli στρέφεται σε εφαρμογές σημαντικότερες από τα τυχερά παιχνίδια, όπως προαναγγέλλει ο τίτλος του: *Διδάσκοντας τη χρήση και τις εφαρμογές των παραπάνω θεωριών σε ζητήματα πολιτικά, κοινωνικά και οικο-*

4 Blaise Pascal (1623-1662), *Lettres à Fermat* (29/7/1654, 24/8/1654, 27/10/1654), *Traité du Triangle arithmétique* (1654, éd. posthume 1665). Pierre de Fermat (1601-1665), *Lettres à Pascal* (29/8/1654, 25/9/1654). Η αλληλογραφία εκδόθηκε μεταφρασμένη στα Αγγλικά από τον Maxine Merrington ως παράρτημα στο βιβλίο της F. N. David, *Games, Gods and Gambling*.

5 Christiaan Huygens (1657). *De ratiociniis in Ludo aleae*. In: F van Schooten (ed.): *Exercitium Mathematicorum*. Εκδόθηκε στα αγγλικά από τον John Arbuthnot (1692). *Of the Laws of Chance*. Ήταν το πρώτο αγγλικό βιβλίο πάνω στις πιθανότητες.

6 Jacob Bernoulli. *Ars conjectandi*. Εκδόθηκε στα λατινικά το 1713 από τους εκδότες Thurneysen Brothers Press, Basel. Το μνημειώδες αυτό έργο μεταφράστηκε στα αγγλικά από την Edith Dudley Sylla: *The Art of Conjecturing, together with Letter to a Friend on Sets in Court Tennis* (Johns Hopkins University Press, 2005).

νομικά. Δυστυχώς δεν πρόλαβε να τελειώσει το έργο του. Στις επιστολές του προς τον Leibniz βρέθηκε διατυπωμένος ο νόμος των μεγάλων αριθμών, ένα από τα σημαντικότερα θεωρήματα της θεωρίας πιθανοτήτων, τον οποίο δεν πρόλαβε να δημοσιεύσει, αλλά τον είχε διατυπώσει και αποδείξει τουλάχιστον δέκα χρόνια πριν. Πρόκειται για την πρώτη προσπάθεια διατύπωσης του *κεντρικού οριακού θεωρήματος*, όπου ο Bernoulli έδειξε ότι η πιθανότητα να συμβεί ένα γεγονός μπορούσε να υπολογιστεί από τις παρατηρούμενες συχνότητες εμφάνισης του γεγονότος αυτού και προσπάθησε να βρει το ανώτατο όριο αριθμού των παρατηρήσεων που ήταν απαραίτητες ώστε να είναι κανείς “βέβαιος” ότι έχει εκτιμήσει σωστά τις πιθανότητες.

Το πρώτο αυτό στάδιο της εξέλιξης της θεωρίας των πιθανοτήτων ολοκληρώνεται με το έργο *Δοκίμιο ανάλυσης των τυχερών παιχνιδιών*⁷ του ιερωμένου Pierre Raymond de Montmort που εκδόθηκε το 1708 και το έργο *Περί μετρήσεως της τύχης, ή, μια μέθοδος υπολογισμού των πιθανοτήτων στα τυχερά παιχνίδια*⁸ του Abraham de Moivre που εκδόθηκε το 1711. Δεν θα ασχοληθούμε με το έργο του αιδεσιμότατου Montmort το οποίο, είναι μεν σημαντικό στην ιστορία των μαθηματικών, όμως στο *Δοκίμιο ανάλυσης των τυχερών παιχνιδιών* δεν έχει να επιδείξει κάτι σημαντικό για τη συνέχεια της ιστορίας μας, παρά μόνο για όσους ασχολούνται με τις εφαρμογές των πιθανοτήτων στα τυχερά παιχνίδια.

Οι ιδέες του Bernoulli αναπτύχθηκαν περαιτέρω από τον αντάξιό του μαθηματικό Abraham de Moivre, ο οποίος με το οριακό θεώρημα των πιθανοτήτων που διατύπωσε, όρισε για πρώτη φορά τον νόμο των σφαλμάτων, που αργότερα εξελίχθηκε ως κανονική κατανομή, ως το όριο της “μελλοντικής” διωνυμικής κατανομής. Κατέληξε δε σε πολύ λογικότερες τιμές από τον Bernoulli για τις δοκιμές που ήταν απαραίτητες στην πειραματική προσέγγιση των πραγματικών πιθανοτήτων. Το βιβλίο του *De mensura sortis ...*, ή οι μετέπειτα αγγλικές εκδόσεις του, *The Doctrine of Chances ...*, αποτελείται από προβλήματα που αναφέρονται σε δημοφιλή παιχνίδια της εποχής και ανάμεσά τους και εκείνα που αναφέρονται στο γενικότερο κύκλο προβλημάτων, γνωστό ως διάρκεια του παιχνιδιού. Ο de Moivre περιέλαβε το έργο του αυτό ως υπόμνημα, αναθεωρημένο κάθε φορά, στις πολλές εκδόσεις του *Αναλυτικά ανάμικτα περί σειρών και τετραγωνισμών*⁹, όπου συμπεριλαμβάνονται και οι εργασίες του πάνω στον συνδυαστικό λογισμό και στις αντίστρο-

7 Pierre Raymond de Montmort (1708). *Essai d'Analyse sur les Jeux d'hasard*. Paris: Jackue Quil-lau.

8 Abraham de Moivre (1711). De mensura sortis, seu, Probabilitate eventuum in Ludis a caso fortuito pendentibus. *Philosophical Transactions*, Vol. 27, No. 329, pp. 213-214. Μεταφράστηκε στα αγγλικά από τον B. McClintock (1984): On the measurement of chance, or, on the probability of events in games depending upon fortuitous chance (*Intern. Statist. Rev.* 52, 237-262).

9 Abraham de Moivre (1730): *Miscellanea analytica de seriebus et quadraturis*. London: J. Tonson and J. Watts.

φες εξισώσεις, θέματα που προέκυψαν από την ενασχόλησή του με τη θεωρία των πιθανοτήτων. Οι σελίδες του έργου του που αναφέρονται στις *ασφάλειες ζωής*, όπου περιγράφει τις εφαρμογές των πιθανοτήτων στις δικές του μελέτες για την τιμολογιακή πολιτική των ασφαλιστικών συμβολαίων, αποδεικνύουν ότι, όπως και ο Bernoulli, προέβλεπε τη χρησιμότητα και τη δυνατότητα εφαρμογής της θεωρίας πιθανοτήτων σε προβλήματα κοινωνικού ενδιαφέροντος και όχι μόνο στα τραπέζια των τυχερών παιχνιδιών. Συνεχιστής του έργου του de Moivre θεωρείται ο άγγλος διαπρεπής μαθηματικός Thomas Simpson στον οποίο ανήκει το έργο *Η φύση και οι νόμοι της τύχης* που έγραψε το 1740 και δύο χρόνια αργότερα, η συνέχεια του έργου αυτού *Θεωρία ανατοκισμού και χρεολυσίας*¹⁰.

3. Η Πολιτική Αριθμητική

Περίπου την ίδια χρονική περίοδο με τη θεωρία των τυχερών παιχνιδιών κατά το δεύτερο μισό του 17ου αιώνα, γεννήθηκε η αγγλική Πολιτική Αριθμητική, ένα σύστημα καταγραφής και επεξεργασίας δεδομένων που σχετίζονται με κοινωνικά ζητήματα. Ο όρος καθιερώθηκε από τον πολυπράγμονα μαθηματικό και έναν από τους πρώτους οικονομομήτρες, τον William Petty. Πνευματική κληρονομιά του ήταν το αριστούργημα *Πολιτική Αριθμητική* που εμφανίστηκε το 1690, τρία χρόνια μετά από το θάνατό του¹¹. Μήνυμα του Petty είναι να καταγράφονται όλες οι φυσικές, κοινωνικές και πολιτικές συνιστώσες ενός κράτους με αριθμούς, προκειμένου να έχουν μια σταθερή βάση για τη λήψη αποφάσεων και να αποφεύγονται έτσι πολιτικές αντιπαραθέσεις. Πιθανότατα ο Petty επηρεάστηκε από τον άγγλο φίλο του John Graunt, ο οποίος είχε συγγράψει το 1662 το διάσημο βιβλίο του *Φυσικές και Πολιτικές Παρατηρήσεις* που βασιζόνταν στα αποτελέσματα της ανάλυσης των δελτίων θνησιμότητας του Λονδίνου και όπου προσπάθησε να εφαρμόσει εμπειρικές πιθανότητες πέρα από τα τυχερά παιχνίδια¹². Το βιβλίο αυτό χάρισε στον Graunt τη φήμη του ιδρυτή της δημογραφίας, της επιστήμης που ασχολείται με τη μελέτη του πληθυσμού ώστε να διαπιστωθεί το σύνολο των κατοίκων που τον αποτελούν, η σύνθεσή του, οι μεταβολές του και όλα τα φαινόμενα που προκαλούν τις μεταβολές αυτές.

Η εργασία του Graunt δεν άφησε αδιάφορους τους μαθηματικούς της εποχής του.

10 Thomas Simpson (1740): *The nature and laws of Chances*. London, Printed by Edward Cave.

– (1742): *The doctrine of Annuities and Reversions*. London, Printed for J. Nourse.

11 William Petty (1676), *Political Arithmetic* (εκδόθηκε για πρώτη φορά το 1690). Βλ. επίσης, McCormick, Ted (2009): *William Petty and the Ambitions of Political Arithmetic*. Oxford University Press.

12 William Petty. *Natural and political observations, mentioned in a following index, and made upon the bills of mortality by John Graunt, citizen of London; with reference to the government* (1662).

Το 1693 ο αστρονόμος Edmond Halley, γνωστός για τη μεγάλη του συνεισφορά στην υπόθεση του γεωγραφικού μήκους, δημοσίευσε ένα άρθρο σχετικά με τις ασφάλειες ζωής, όπου παρουσίαζε μια ανάλυση του για τις ηλικίες θανάτου με βάση τα στατιστικά στοιχεία του γερμανού κληρικού Caspar Neumann σχετικά με τη θνησιμότητα της πόλη του Breslau. Το άρθρο αυτό υποδείκνυε στη κυβέρνηση να πουλάει ασφάλειες ζωής σε τιμή βασισμένη στην ηλικία του υποψήφιου ασφαλισμένου. Ο πίνακας ζωής για την πόλη του Breslau, θεωρείται πλέον σταθμός στην ιστορία της δημογραφίας. Λίγα χρόνια πρωτότερα, ο Huygens το 1669 και ο Johan de Witt το 1671 εφάρμοσαν τον λογισμό των πιθανοτήτων στον υπολογισμό μιας μέσης διάρκειας ζωής των ανθρώπων, με σκοπό την εκτίμηση της τιμής των ασφαλειών από τους πίνακες θνησιμότητας.

Κύρια ιδέα του Graunt ήταν να υπολογίζονται οι μέσοι όροι ως αντιπροσωπευτικές τιμές των πληθυσμών, κάτι που συνηθίζεται στην ανάλυση δεδομένων μέχρι τις μέρες μας. Στην πραγματικότητα, αυτή η ιδέα δεν είναι μόνο κεντρική για τη στατιστική, αλλά και για όλες τις φυσικές επιστήμες, για κάθε επιστημονικό κλάδο που στηρίζεται στο πείραμα και στην ανάλυση των αποτελεσμάτων του πειράματος. Βέβαια, ο Graunt δεν είχε επίγνωση της σχέσης μεταξύ του “μεγέθους του δείγματος” και των “σταθερών συμπεριφορών” του και ούτε γνώριζε την έννοια της “πιθανότητας”. Παρ’ όλα αυτά, όρισε τη γενική στρατηγική που πρέπει να ακολουθεί κανείς όταν ασχολείται με τυχαίες θα λέγαμε σήμερα δειγματοληψίες. Η στρατηγική αυτή βασίζεται σε “μέσους όρους” που στοχεύουν στην ανίχνευση συμπεριφορών που διατυπώνονται π.χ. ως “νόμοι των σφαλμάτων” ή, όπως είναι γνωστές σήμερα, ως “στατιστικές κατανομές”. Λίγα χρόνια αργότερα προς το τέλος του 17ου αιώνα ο James Bernoulli συνέδεσε την έννοια της πιθανότητας με την έννοια της τύχης προσπαθώντας να τη μετασχηματίσει σε μια αυστηρή μαθηματική θεωρία. Ωστόσο, έμεινε ένα κενό στην ανάπτυξη της σχετικής θεωρίας ώστε να μπορούν να συνδεθούν οι πιθανότητες των τυχερών παιχνιδιών με τις λίστες του Graunt και του Petty. Το κλειδί για τη σύνδεση αυτή και την αρχή της μαθηματικής τεκμηρίωσης της Πολιτικής Αριθμητικής ήταν οι *αντίστροφες πιθανότητες*.

Ο αιδεσιμότατος Thomas Bayes και στη συνέχεια ο Pierre Simon Laplace επιχείρησαν να καλύψουν το κενό αυτό θέτοντας το πρόβλημα της αντίστροφης πιθανότητας, πρόβλημα που στη συνέχεια θα παίξει σημαντικό ρόλο στην ιστορία της εφαρμογής των μεθόδων που σήμερα ονομάζουμε *θεωρία πιθανοτήτων* και *στατιστική*, στις διάφορες επιστήμες. Η μετά θάνατον του Bayes δημοσιευμένη το 1764 εργασία του με τίτλο *Πραγματεία για την επίλυση ενός προβλήματος στη θεωρία της τύχης*¹³, όπου περιέχονταν το θεώρημα των δεσμευμένων πιθανοτήτων, ενώ αγνοή-

13 Bayes T. (1763). An Essay towards solving a Problem in the Doctrine of Chances". *Phil. Trans.*, 53, 370 - 418. Με την έκφραση "doctrine of chances", η οποία εισήχθη από τον Abraham de Moivre ως τίτλος του βιβλίου του, εννοείται η θεωρία πιθανοτήτων.

θηκε από τους σύγχρονους του, επηρέασε σημαντικά την ανάπτυξη της σύγχρονης στατιστικής. Ο Pierre-Simon Laplace γενίκευσε το θεώρημα του Bayes και το χρησιμοποίησε στη λύση διαφόρων προβλημάτων στην ουράνια μηχανική, στη στατιστική ιατρική, στη νομολογία και σε άλλες εφαρμογές, καθιερώνοντας τον κλάδο της Μπεϋζιανής στατιστικής. Έθεσε το θέμα της αντίστροφης πιθανότητας ορίζοντας διάφορες συναρτήσεις κατανομής, ως αντίστροφες των συναρτήσεων που περιγράφουν τη θεωρητική κατανομή των σφαλμάτων¹⁴.

Πως όμως συναντώνται και συνενώνονται τελικά οι τεχνικές καταγραφής της αγγλικής *political arithmetic* και τα εργαλεία διακυβέρνησης του κράτους της γερμανικής *statistik* με τις θεωρίες για τα τυχερά παιχνίδια που ξεκίνησαν από τον Pascal στα μέσα του 17ου αιώνα και ολοκληρώθηκαν ως νόμοι των σφαλμάτων και ως αντίστροφες πιθανότητες στη διάρκεια περίπου των εκατό επόμενων χρόνων;

Ο σκοτσέζος ευγενής και πολιτικός, Sir John Sinclair, γνώρισε το 1786 στη Γερμανία τις μεθόδους της πολιτικής έρευνας που χρησιμοποιούσαν οι γερμανικές υπηρεσίες και θέλησε να χρησιμοποιήσει αυτές τις "στατιστικές" μεθόδους, όπως τις ονόμασε αντιγράφοντας τον γερμανικό όρο *statistik*, για τη μέτρηση του ύψους της ευημερίας των κατοίκων ενός έθνους ώστε να βρει τρόπους για τη βελτίωσή της. Έστειλε λοιπόν ειδικά διαμορφωμένο ερωτηματολόγιο σε όλους τους υπουργούς και δημόσιους άρχοντες της χώρας του. Τα αποτελέσματα της εργασίας του ήταν ζωτικής σημασίας για την κατανόηση της κοινωνικής κατάστασης στη Σκωτία, στις παραμονές μάλιστα της βιομηχανικής επανάστασης. Αν και χαρακτηρίστηκε από πολλούς ως κλοπή η χρήση του όρου στατιστική, η λέξη ενσωματώθηκε στην αγγλική γλώσσα. Ο μεγάλος κοινωνικός αντίκτυπος της εργασίας του και ο θόρυβος που δημιουργήθηκε στα τέλη του 18ου αιώνα αποκατέστησε τον χαμένο ενθουσιασμό των μαθηματικών και ξεκίνησε κατά τη διάρκεια του 19ου αιώνα μια καταπληκτική ανάπτυξη της στατιστικής ως ένα νέο κεφάλαιο πλέον των μαθηματικών.

Όμως το πραγματικό κίνητρο για την εφαρμογή μιας αυστηρής μαθηματικής θεωρίας στην ανάλυση των δημογραφικών δεδομένων και στην εφαρμογή της στις διοικητικές πρακτικές ενός κράτους, προήλθε από μία άλλη, απρόσμενη κατεύθυνση. Τα δύο κορυφαία επιστημονικά προβλήματα της εποχής εκείνης, η εύρεση του γεωγραφικού μήκους στα ωκεάνια ταξίδια και ο ορισμός του σχήματος της γης, επηρέασαν και δρομολόγησαν τις εξελίξεις και προς αυτή την κατεύθυνση. Οι αστρονόμοι και κυρίως οι γεωδαίτες, αν θα μπορούσαμε να τους διαχωρίσουμε την εποχή αυτή, έδωσαν τη λύση αναπτύσσοντας τη μέθοδο των ελαχίστων τετραγώνων για την ανάλυση των μετρήσεών τους.

14 Laplace, Pierre Simon (1774). *Memoire sur la Probabilite des Causes par les Evenements. l'Academie Royale des Sciences*, 6, 621 - 656. Μεταφράστηκε στα αγγλικά από τον S. M. Stigler το 1986 ως "Memoir on the Probability of the Causes of Events" (*Statistical Science*, 1(3), 359 - 378).

4. Μετρώντας τα αστέρια και τη γη

Η μέθοδος των ελαχίστων τετραγώνων διατυπώθηκε ταυτόχρονα και ανεξάρτητα από δύο μαθηματικούς, αστρονόμους και γεωδαιίτες: τον Adrien-Marie Legendre και τον Carl Friedrich Gauss. Και οι δύο προσπαθούσαν να προσδιορίσουν τις τροχιές ουρανίων σωμάτων από διακριτές παρατηρήσεις με σφάλματα αλλά και να συνδυάσουν με βέλτιστο τρόπο τις γεωδαιτικές παρατηρήσεις για τον υπολογισμό των τόξων πάνω στους γήινους μεσημβρινούς ώστε να οριστεί το σχήμα της γης. Η δημοσίευση της μεθόδου το 1809 από τον Gauss, υπήρξε αφορμή για ισχυρές φιλονικίες και διεκδικήσεις της πρωτοτυπίας της από τους δύο αυτούς επιστήμονες.

Είναι δύσκολο να παρακολουθήσει κανείς την εξέλιξη αυτής της μεθόδου ανάλυσης των μετρήσεων πριν από τους δύο αυτούς κορυφαίους μαθηματικούς. Το πρόβλημα των σφαλμάτων των παρατηρήσεων ήταν γνωστό από την εποχή του Κέπλερ και του Γαλιλαίου, η μέθοδος όμως των ελαχίστων τετραγώνων συνδέθηκε με τα δύο κορυφαία προβλήματα που ταλάνιζαν τους επιστήμονες τον 18ο αιώνα: με το πρόβλημα των τριών σωμάτων ώστε να περιγραφεί η κίνηση της σελήνης και με το πρόβλημα του προσδιορισμού του σχήματος της γης. Η ιδέα του συνδυασμού των παρατηρήσεων άρχισε να αναπτύσσεται από το αστρονομικό έργο του μαθηματικού και χαρτογράφου Tobias Mayer και η ιδέα ενός αντικειμενικού κριτηρίου για τη διόρθωση των παρατηρήσεων γεννήθηκε από το γεωδαιτικό έργο του Roger Boscovich. Ακολούθησε στη συνέχεια η εξέλιξη του "νόμου των σφαλμάτων" ή, όπως ειπώθηκε αργότερα, "της κανονικής κατανομής των σφαλμάτων", μια θεωρία που περιγράφει τη συνολική συμπεριφορά των σφαλμάτων και επιτρέπει την εισαγωγή της στους υπολογισμούς των παρατηρήσεων. Μια θεωρία που πάρα πολλοί προσθέσανε κάτι, αλλά που μελετήθηκε συστηματικά από τον Gauss.

Ένα από τα πρώτα έργα πάνω στο θέμα του προσδιορισμού των καταλληλότερων τρόπων συνδυασμού των μετρήσεων είναι ο τόμος *Ανάμικτα έργα του λογισμού των σφαλμάτων σε διάφορους κλάδους*, μέρος του έργου *Αρμονία Μέτρων*¹⁵ που εκδόθηκε το 1722 από τον Robert Smith, αλλά πρόκειται για τα χειρόγραφα του εξαδέλφου του και προκατόχου του στην έδρα του πανεπιστημίου του Cambridge, αστρονόμου Roger Cotes. Ο Thomas Simpson, το 1757, ως συνεχιστής του έργου του Cotes, προσπάθησε, σε άρθρο του που περιέχεται στη συλλογή *Ποικίλα άρθρα ...*¹⁶ να δείξει τα πλεονεκτήματα που ωφελείται η αστρονομία αν δεχθεί ως τιμή ενός αγνώστου μεγέθους το μέσο όρο των τιμών που λαμβάνονται από πολλές μετρήσεις του μεγέθους αυτού. Στο άρθρο αυτό διατυπώνεται για πρώτη φορά το αξίωμα ότι *θετικά και αρνητικά σφάλματα έχουν την ίδια πιθανότητα να εμφανι-*

15 Roger Cotes (1722). *Opera Miscellanea Aestinatio Errorum in Mixta Mathesi*. Περιέχεται στο *Harmonia mensurarum*, (επιμέλεια έκδοσης Robert Smith). Cambridge.

16 Thomas Simpson, (1757). *Miscellaneous Tracts on Some Curious and very Interesting Subjects in Mechanics, Physical-Astronomy, and Speculative Mathematics*. London: J. Nourse.

σθούν και ορίζεται ως νόμος των σφαλμάτων μια εξίσωση γραμμικής μορφής που οδηγεί σ' ένα ισοσκελές τρίγωνο.

Ο Pierre Simon Laplace, το 1774, ήταν ο πρώτος που εφάρμοσε τις αρχές της θεωρίας των πιθανοτήτων, την οποία ο ίδιος χαρακτήριζε ως κοινή λογική που μπορεί να εκφραστεί με τη γλώσσα των μαθηματικών, στη συζήτηση για τα σφάλματα των παρατηρήσεων. Για να περιγράψει τον νόμο των σφαλμάτων πρότεινε μια εκθετική παράσταση και τρία χρόνια αργότερα, το 1777, μια λογαριθμική. Προσπάθησε να επεξεργαστεί μαθηματικά το άθροισμα των απόλυτων τιμών ως κριτήριο βέλτιστης εκτίμησης, η πολυπλοκότητα όμως των υπολογισμών τον αποθάρρυνε και τα παράτησε. Την ίδια χρονιά δημοσιεύτηκε ένα νέο έργο του Daniel Bernoulli που αναφερόταν στον συνδυασμό των αποτελεσμάτων περισσότερων μετρήσεων. Ο Bernoulli δεν συμφωνούσε με την απόρριψη ορισμένων μετρήσεων, επειδή αυτές φαίνονται αρχικά ότι δε συμφωνούν με τις υπόλοιπες, μια τακτική που ακολουθούσαν οι αστρονόμοι εκείνης της εποχής.

Την 1^η Ιανουαρίου του 1801 αναγγέλθηκε από τον αστρονόμο Giuseppe Piazzi η ανακάλυψη ενός νέου πλανήτη, που ονομάστηκε από τον ίδιο *Φερδινάνδειος Δήμητρα*. Όταν έδειξε το νέο πλανήτη στους συναδέλφους του, ο ιταλός αστρονόμος, τους έθεσε ταυτόχρονα και το πρόβλημα του προσδιορισμού της τροχιάς του. Ο Gauss έλυσε το πρόβλημα εφαρμόζοντας τη νέα μέθοδο υπολογισμού των πλανητικών τροχιών, την ιδέα της οποίας είχε, όπως ο ίδιος ισχυρίστηκε, από το 1795. Τη μέθοδο αυτή εφάρμοσε και στην περίπτωση της Παλλάδος, ενός άλλου πλανήτη που ανακαλύφθηκε το 1802 από τον αστρονόμο Όλβερς και τέλος τη δημοσίευσε το 1809 στο βιβλίο του *Θεωρία περί της κινήσεως των ουρανίων σωμάτων*¹⁷. Η τιμή όμως της πρώτης δημοσίευσης της μεθόδου ανήκει στον Adrien Marie Legendre, ο οποίος ήδη την είχε παρουσιάσει το 1805 στο παράρτημα του βιβλίου του *Νέα μέθοδος προσδιορισμού των τροχιών των κομητών*¹⁸ με τον τίτλο “*Sur la méthode des moindres quarrés*”, η μέθοδος των ελαχίστων τετραγώνων, το όνομα που επρόκειτο να μείνει στο μέλλον. Πάντως, σε αντίθεση με τον Legendre, ο Gauss παρουσίασε τα ελάχιστα τετράγωνα ως τη μέθοδο που ελαχιστοποιεί τα σφάλματα της εκτίμησης με αυστηρή θεωρητική τεκμηρίωση, βασισμένη στη θεωρία των πιθανοτήτων και στην κατανομή που ακολουθούν τα σφάλματα των παρατηρήσεων. Έδωσε μάλιστα τη μαθηματική μορφή στην κατανομή που ονομάστηκε

17 Gauss, Carl Fredrich (1809). *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium*. Hamburg, Friedrich Perthes and I. H. Besser. Εκδόθηκε σε μετάφραση του C. H. Davis (1963): *Theory of Motions of the Heavenly Bodies Moving about the Sun in Conic Sections*. Dover.

18 Legendre, Adrien Marie (1805). *Nouvelle Méthode pour la détermination des Orbites des Comètes*. Paris: Courcier. Το παράρτημα που αφορά στη μέθοδο των ελαχίστων τετραγώνων περιέχεται μεταφρασμένο στα αγγλικά στο βιβλίο D. E. Smith (1959). *A Source Book in Mathematics*. Dover.

το 1870 κανονική από τον γερμανό στατιστικό Wilhelm Lexis, ολοκληρώνοντας στην κατεύθυνση αυτή τις εργασίες του Laplace. Έδειξε ουσιαστικά ο Gauss ότι αν τα σφάλματα ακολουθούν την κανονική κατανομή, τότε το κριτήριο των ελαχίστων τετραγώνων δίνει εκτιμήσεις ίδιες με τη μέθοδο της μέγιστης πιθανοφάνειας.

Οι αστρονόμοι και οι γεωδαίτες είδαν τη χρησιμότητα της μεθόδου, ειδικά επειδή τα σφάλματα στις παρατηρήσεις τους ήταν ο κανόνας, καθώς τα προκαλούσε όχι μόνον η ατέλεια των οργάνων αλλά και παράγοντες του περιβάλλοντος των μετρήσεων που δυσκόλευαν τον παρατηρητή, όπως π.χ. η διάθλαση της ατμόσφαιρας. Με τη διατύπωση του κεντρικού οριακού θεωρήματος το 1810 από τον Laplace, το ζήτημα της κατανομής των σφαλμάτων στη μέθοδο των ελαχίστων τετραγώνων έχασε τη σημασία του. Το 1812 ο Laplace δημοσίευσε τη μεγάλη του πραγματεία *Αναλυτική θεωρία των πιθανοτήτων*¹⁹, μία σύνθεση όλων των μέχρι τότε εξελίξεων, που για μία ολόκληρη γενιά παρέμεινε το κυρίως βιβλίο για το θέμα των ελαχίστων τετραγώνων. Το 1821, 1823 και 1826 ο Carl Friedrich Gauss δημοσίευσε τρεις εργασίες του σε μορφή απομνημονευμάτων, με κοινό τίτλο *Θεωρία συνδυασμού των παρατηρήσεων (με) ελάχιστες επιδράσεις σφαλμάτων*²⁰. Στις εργασίες του αυτές επιστρέφει και διατυπώνει τη θεωρία των ελαχίστων τετραγώνων ως εκτίμηση ελάχιστης μεταβλητότητας, χωρίς την υπόθεση της κανονικής κατανομής των μετρήσεων. Κατά την περίοδο αυτή ο διάσημος επιστήμονας ασχολήθηκε αποκλειστικά με γεωδαιτικές εργασίες. Ανάμεσα στα άλλα του έργα της εποχής εκείνης είναι και η εφαρμογή της μεθόδου των ελαχίστων τετραγώνων στον υπολογισμό του τριγωνομετρικού δικτύου της περιοχής του Αμβούργου καθώς και εργασίες σχετικές με το σχήμα της γης. Εξάλλου, η τρίτη εργασία του που δημοσιεύτηκε το 1826 με υπότιτλο *Supplementum*, είναι σε μεγάλο βαθμό αυτόνομο έργο, αφιερωμένο στην εφαρμογή των ελαχίστων τετραγώνων στη γεωδαισία. Ήδη όμως από το 1815 τα ελάχιστα τετράγωνα καθιερώθηκαν ως το επίσημο εργαλείο των γεωδαιτών και των αστρονόμων στη Γαλλία, την Ιταλία, την Πρωσία και λίγο αργότερα την Αγγλία.

5. Η κοινωνική φυσική και η βιομετρική παράδοση

Η νέα επιστήμη των πιθανοτήτων και της στατιστικής χρησιμοποιήθηκε καταρχάς κυρίως στη γεωδαισία, για να ελέγξουν οι γεωδαίτες τα σφάλματα των μετρήσεών τους με τη μέθοδο των ελαχίστων τετραγώνων. Ο βέλγος μαθηματικός και αστρο-

19 Laplace, Pierre Simon (1812). *Theorie Analytique des Probabilites*. Courcier, Paris. Reprinted as "Oeuvres Completes de Laplace", 7, 1878-1912. Paris: Gauthier-Villars.

20 Carl Friedrich Gauss (1821, 1823, 1826). *Combinatioonis Observationum Erroribus Minimis Obnoxiae*. Εκδόθηκε σε μετάφραση του G. W. Stewart το 1995: *Theory of the Combination of Observations. Least Subject to Errors* (Society of Industrial and Applied Mathematics).

νόμος Adolphe Quetelet ήταν άριστος γνώστης της πολυπλοκότητας των κοινωνικών φαινομένων και των πολλών μεταβλητών που απαιτούνται να μετρηθούν, ώστε να περιγραφούν τα φαινόμενα αυτά. Με τον τρόπο που συσσωρεύονται γύρω από τις πραγματικές τιμές οι επηρεασμένες με σφάλματα μετρήσεις, έτσι ακριβώς δημιούργησε την έννοια του μέσου ανθρώπου, γύρω από τον οποίον κατανέμονται τα χαρακτηριστικά των πραγματικών ανθρώπων. Στο αριστούργημά του *Για τον άνθρωπο και την ανάπτυξη των ικανοτήτων του, ή Δοκίμιο της κοινωνικής φυσικής*²¹ που κυκλοφόρησε το 1835, ονόμασε σφάλμα, σε αντιστοιχία με τα σφάλματα των παρατηρήσεων, την απόκλιση από αυτή τη θεωρητική κατανομή, ανοίγοντας το δρόμο στον κυρίαρχο ρόλο της κανονικής κατανομής. Διατύπωσε έτσι τους στατιστικούς νόμους που διέπουν φαινόμενα όπως οι ρυθμοί γεννήσεων, θανάτων, γάμων, αυτοκτονιών και εγκληματικότητας, συνδέοντας τις τιμές αυτών των μεταβλητών με άλλους κοινωνικούς παράγοντες και δημιουργώντας τον κλάδο της κοινωνικής στατιστικής, ο ίδιος την αποκάλεσε *κοινωνική φυσική*.

Η βιομετρική παράδοση ιδρύθηκε κυρίως από τον άγγλο φυσιοδίφη, εξερευνητή και ανθρωπολόγο Sir Francis Galton, εξάδελφο του Δαρβίνου, ο οποίος χρησιμοποίησε στατιστικές μεθόδους στη μελέτη των κληρονομικών χαρακτηριστικών, με σκοπό τη βελτίωση της φυσικής και διανοητικής διάπλασης του ανθρώπινου είδους μέσω μιας επιλεκτικής αναπαραγωγής. Αυτός ήταν ο κύριος στόχος της ευγονικής, τον όρο αυτό χρησιμοποίησε και καθιέρωσε ο ίδιος ο Galton, και η στατιστική του παρείχε τα εργαλεία για τη δημιουργία των μαθηματικών προτύπων που θα του επέτρεπαν την ανάλυση των στοιχείων της κληρονομικότητας. Ο Galton εφάρμοσε την κανονική κατανομή όχι τόσο ως “νόμο των σφαλμάτων”, αλλά ως “νόμο των αποκλίσεων” από τις φυσικές δεξιότητες, που του επιτρέπει την κατάταξη των ανθρώπων, όπου το ιδεώδες δεν είναι ο μέσος άνθρωπος του Quetelet, αλλά ο μεγαλοφυής²².

Στο βιομετρικό εργαστήριο του Galton και μέσα από τις έρευνές του σχετικά με την κληρονομικότητα, ξεπήδησαν οι έννοιες της παλινδρόμησης και της συσχέτισης. Στο εργαστήριο αυτό κατέγραφε και πινακοποιούσε τα ύψη, τα βάρη, καθώς και άλλα χαρακτηριστικά των μελών της κάθε οικογένειας που έπαιρνε μέρος στο πείραμά του. Προσπαθούσε να βρει κάποιον μαθηματικό τύπο ώστε να προβλέπει τα χαρακτηριστικά των παιδιών μετρώντας τα χαρακτηριστικά των γονέων. Αναλύοντας τους πίνακές του παρατήρησε πως τα αγόρια με τους πολύ ψηλούς πατέρες έτειναν να είναι κοντότερα από αυτούς, ενώ τα αγόρια με κοντούς πατέρες, υψηλότερα. Τα ύψη των παιδιών «παλινδρομούσαν» προς το μέσο ύψος των πατέ-

21 Quetelet, Adolphe (1835). *Sur l'homme et le développement de ses facultés ou Essai de physique sociale*. Paris, Bachelier, Imprimeur-Libraire, Quai des Augustins, No. 55.

22 Από τις βιογραφίες που εκδόθηκαν για τον Galton η σημαντικότερη είναι του Karl Pearson, *The life, letters and labours of Francis Galton* (3 τόμοι: 1914, 1924, 1930). Cambridge University Press.

ρων του δείγματός του. Ως ένας αυστηρός και απaráβατος νόμος της φύσης που αναγκάζει τα ύψη των ανθρώπων να παραμένουν σταθερά κατά μέσο όρο. Το φαινόμενο αυτό, που αργότερα ονομάστηκε *regression* (παλινδρόμηση), ο Galton το ονόμασε *reversion* (αναστροφή). Στη συνέχεια ο Galton έδωσε ένα αριθμητικό μέτρο για να εκφράσει το βαθμό συσχέτισης δύο μεταβλητών, το ύψος του παιδιού για παράδειγμα με το μέσο ύψος των πατέρων, και να αποδώσει το πόσο καλά μπορεί να περιγραφεί η σχέση μεταξύ των μεταβλητών αυτών. Το ονόμασε *coefficient of correlation* (συντελεστή συσχέτιση) και το παρουσίασε ως ένα εναλλακτικό εργαλείο της παλινδρόμησης. Λίγο αργότερα ο άγγλος μαθηματικός Karl Pearson διατύπωσε την ιδέα της παλινδρόμησης και της συσχέτισης του Galton στην πιο ολοκληρωμένη τους μορφή και τις μετέτρεψε σε στατιστικά εργαλεία χρήσιμα σε κάθε επιστήμη που βασίζεται στην παρατήρηση.

Το ενδιαφέρον του Karl Pearson για τις εφαρμογές της στατιστικής στα βιολογικά προβλήματα της κληρονομικότητας και της εξέλιξης το κίνησαν οι εξελικτικές θεωρίες του Galton και η φιλία του με τον βιολόγο Raphael Weldon, ο οποίος τους έφερε σε επαφή. Από το 1893 έως το 1912 ο Pearson, ένας από τους θεμελιωτές της σύγχρονης στατιστικής, έγραψε μία σειρά από 18 μελέτες τις οποίες εξέδωσε με τίτλο *Μαθηματικές συμβολές στη θεωρία της εξέλιξης*²³. Στην έκδοση αυτή περιέβαλλε πολλές από τις σημαντικότερες εργασίες του για την παλινδρόμηση και τη συσχέτιση αλλά και το περίφημο κριτήριο *χι τετράγωνο* που αναπτύχθηκε το 1900 και επέλυε το πρόβλημα του κατά πόσο μία θεωρητική κατανομή προσαρμόζεται επαρκώς στην κατανομή των δεδομένων που περιέγραφε, μια διαδικασία γνωστή ως *έλεγχος καλής προσαρμογής*.

Στο γύρισμα του νέου αιώνα, το 1901, οι Pearson, Galton και Weldon ένωσαν τις προσπάθειές τους για να εκδώσουν ένα νέο επιστημονικό περιοδικό που θα εφαρμόζε τις ιδέες του Pearson στη βιολογία. Το περιοδικό *Biometrika*, το οποίο χρηματοδοτήθηκε από τον Galton, έγινε το κυρίαρχο έντυπο στο χώρο της στατιστικής. Τα τελευταία χρόνια της ζωής του ο Galton χρηματοδότησε και τη δημιουργία μίας έδρας Ευγονικής στο University College του Λονδίνου, η οποία διατηρείται μέχρι σήμερα ως έδρα γενετικής. Πρώτος καθηγητής της έδρας ήταν ο Karl Pearson και στη συνέχεια, μετά τη συνταξιοδότησή του, ο μαθηματικός και βιολόγος Ronald Aylmer Fisher, αντίπαλος του Pearson στο πεδίο των επιστημονικών διενέξεων.

23 Pearson, K., *Mathematical Contributions to the Theory of Evolution*. Οι εργασίες αυτές βρίσκονται αναδημοσιευμένες στο βιβλίο: Pearson, K. (1956). *Karl Pearson's Early Statistical Papers*. Cambridge University Press.

6. Το χρονικό της σημαντικότητας

Ένα από τα μεγάλα επιτεύγματα του Karl Pearson ήταν η δημοσίευση του ελέγχου της καλής προσαρμογής στην εργασία του Σχετικά με το κριτήριο ώστε ένα δεδομένο σύστημα αποκλίσεων από τις πιθανές τιμές στην περίπτωση ενός συστήματος συσχετισμένων μεταβλητών είναι τέτοιο ώστε να μπορεί λογικά να υποθεθεί ότι έχει προκύψει από τυχαία δειγματοληψία.²⁴ Ο έλεγχος αυτός ήταν ο πρόδρομος ενός σημαντικού κεφαλαίου της σύγχρονης στατιστικής ανάλυσης, που ονομάζεται έλεγχος υποθέσεων (hypothesis testing), ή έλεγχος σημαντικότητας (significance testing). Συγκρίνοντας τις παρατηρούμενες με τις προβλεπόμενες τιμές των συχνοτήτων εμφάνισης, ο Pearson υπολόγισε ένα στατιστικό κριτήριο που ακολουθεί την ίδια κατανομή, ανεξάρτητα από τον τύπο των δεδομένων που χρησιμοποιήθηκαν. Μπορούσε δηλαδή να συνοψίσει τις κατανομές πιθανοτήτων αυτής της στατιστικής ποσότητας σε μία και να χρησιμοποιήσει τους ίδιους πίνακες για κάθε έλεγχο. Η κατανομή αυτή ανήκε σε μια ομάδα λοξών κατανομών που τις είχε ήδη ορίσει από το 1876 ο γερμανός γεωδαίτης Friedrich Robert Helmert χρησιμοποιώντας το ελληνικό γράμμα, ως κατανομές χ τετράγωνο²⁵. Ονόμασε λοιπόν τον στατιστικό του έλεγχο ο Pearson chi square goodness of fit test, που αποδίδεται στα ελληνικά ως έλεγχος καλής προσαρμογής χ τετράγωνο. Ο έλεγχος έχει μία μόνο παράμετρο, την οποία ο Ronald Fisher σε μια εργασία του το 1922 ονόμασε degrees of freedom, όρος που χρησιμοποιείται από τότε στην ανάλυση δεδομένων για να οριστεί η διαθέσιμη πλεονάζουσα πληροφορία. Η λέξη παράμετρος αποτελεί αντιδάνειο της αγγλικής parameter, η οποία προέρχεται από τις ελληνικές λέξεις “παράμετρηση” και είναι δημιουργήμα των άγγλων στατιστικών αυτής της περιόδου. Προσδιορίζεται με αυτήν ένας αριθμός, διαφορετικός από την τιμή μιας μέτρησης, ή που δεν προκύπτει από κάποια διαδικασία μέτρησης αλλά υπολογίζεται από κάποιες ιδιότητες των μετρήσεων. Για παράδειγμα οι λοξές κατανομές που αναπτύχθηκαν ανεξάρτητα από τους Helmert και Pearson, ορίζονται από τέσσερις παραμέτρους: τη μέση τιμή των τιμών των παρατηρήσεων, την τυπική της απόκλιση, τη συμμετρία και την κύρτωση.

Το 1908 ο William Sealy Gosset, ένας βιολόγος που δούλευε στη ζυθοποιία Guinness στο Δουβλίνο και επεξεργαζόταν τεχνικές ποιοτικού ελέγχου της παραγωγής βασισμένες σε μικρό αριθμό δειγμάτων, δημοσίευσε στο περιοδικό

24 Pearson, K. (1900). On the Criterion that a given System of Deviations from the Probable in the Case of a Correlated System of Variables is such that it can be reasonably supposed to have arisen from Random Sampling. *Philosophical Magazine Series 5*, 50 (302): 157 - 175.

25 Helmert, F. R. (1876). Ueber die Wahrscheinlichkeit der Potenzsummen der Beobachtungsfehler und über einige damit im Zusammenhange stehende Fragen. *Zeitschrift für Mathematik und Physik*, 21, 102 - 219.

Biometrika μία εργασία με τίτλο *Το πιθανόν σφάλμα του μέσου*²⁶. Στην εργασία αυτή εισήγαγε την κατανομή t για μικρά δείγματα και τη δημοσίευσε, όπως και τις επόμενες εργασίες του, με το ψευδώνυμο “student” με σεβασμό στους μεγάλους δασκάλους της στατιστικής, το δε κριτήριο t ονομάζεται από τότε “κριτήριο student”. Ο όρος *studentized residuals*, *ομαλοποιημένα υπόλοιπα* όπως αποδίδεται στα ελληνικά, δόθηκε προς τιμήν του μεγάλου και σεμνού στατιστικού Gosset, διότι η ιδέα της διαίρεσης του υπολοίπου προς την εκτίμηση της τυπικής του απόκλισης είναι κεντρική στη λογική του ελέγχου του. Το *εσωτερικά και εξωτερικά ομαλοποιημένο υπόλοιπο* αποτελεί επίσης σημαντική τεχνική για την ανίχνευση των ακραίων τους τιμών, μια τεχνική που στις γεωδαιτικές της εφαρμογές ονομάζεται *σάρωση δεδομένων*.

Ορόσημο στην ιστορία της στατιστικής θεωρείται η σημαντικότερη ίσως εργασία του Ronald Aylmer Fisher, *Η Συσχέτιση μεταξύ συγγενών στην υπόθεση της Μεντελικής Κληρονομικότητας*²⁷, όπου εισήγαγε την τεχνική της ανάλυσης της μεταβλητότητας για τον έλεγχο και τον διαχωρισμό των “σημαντικών” αποτελεσμάτων από τα τυχαία σφάλματα. Στην εργασία αυτή χρησιμοποιήθηκαν για πρώτη φορά οι όροι *variance* (*μεταβλητότητα*) και *analysis of variance* (*ανάλυση μεταβλητότητας*). Τέθηκε επίσης το θέμα της εκτίμησης των συνιστωσών της μεταβλητότητας αναφοράς, τις οποίες ο Fisher ονομάζει *components of variation*²⁸. Το πρόβλημα αυτό όμως περιγράφηκε για πρώτη φορά από τον γερμανό γεωδαίτη Friedrich Robert Helmert στη δεύτερη έκδοση του βιβλίου του *Η συνόρθωση των παρατηρήσεων με τη μέθοδο των ελαχίστων τετραγώνων*²⁹ το 1907, όπου και ανέπτυξε μια μέθοδο υπολογισμού των βαρών ασυσχέτιστων μεταξύ τους γεωδαιτικών παρατηρήσεων βασισμένη στη μέθοδο της μέγιστης πιθανοφάνειας. Η πρωτοποριακή αυτή εργασία του Helmert δυστυχώς αγνοήθηκε από τους στατιστικούς της εποχής εκείνης, επειδή είχαν στραμμένη την προσοχή τους στην εξέλιξη και στην κληρονομικότητα και στα επιτεύγματα της στατιστικής που προέκυπταν από τις μελέτες των κουνελιών του Galton και των μοσχομπίζελων του Mendel, και στο έργο πάνω στις μελέτες αυτές των Pearson, Gosset και Fischer και όχι στην ανάλυση των γεωδαιτικών μετρήσεων.

Ένα μεγάλο μέρος της δουλειάς των Pearson και Gosset επικαλύφθηκε από την αντίστοιχη δουλειά του Fisher. Στο άρθρο του *Περί της κατανομής που ακολουθούν οι συναρτήσεις σφάλματος γνωστών στατιστικών ποσοτήτων* που συνέγραψε το

26 “Student” (1908). The Probable Error of a Mean. *Biometrika*, 6, 1 - 25.

27 Fisher, R.A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Royal Society Edinburgh*, 52, 399 - 433.

28 Ο όρος *components of variance* εμφανίστηκε για πρώτη φορά το 1939 στην εργασία του H. Daniels, The estimation of components of variance (*J. Roy. Stat. Soc. Suppl.* 6, 186 - 197).

29 Helmert, F. R. (1907). *Die Ausgleichsrechnung nach der Methode der kleinsten Quadrate*. Leipzig und Berlin, Druck und Verlag Von B. G. Teubner.

1924, παρουσίασε τον έλεγχο chi-τετράγωνο του Pearson και τον έλεγχο t του William Gosset στο ίδιο πλαίσιο μαζί με την κανονική κατανομή και την ανάλυση της μεταβλητότητας, χρησιμοποιώντας τη γενικότερη δική του z-κατανομή, που χρησιμοποιείται από τότε ως κατανομή F .³⁰ Ένα χρόνο αργότερα, η ενασχόλησή του με τον σχεδιασμό πειραμάτων και με τις μελέτες του για την ανάλυση της μεταβλητότητας είχε ως αποτέλεσμα την έκδοση του πρώτου του βιβλίου, *Στατιστικές Μέθοδοι για ερευνητές*.³¹ Αυτό το βιβλίο τυπώθηκε σε πολλές εκδόσεις και μεταφράσεις στα επόμενα χρόνια, και έγινε το πρότυπο έργο αναφοράς για τους επιστήμονες πολλών κλάδων. Το 1935, ακολούθησε *Ο σχεδιασμός των πειραμάτων*, που επίσης άσκησε μεγάλη επιρροή.³² Μεταξύ άλλων συνεισφορών, το βιβλίο εισήγαγε την έννοια της μηδενικής υπόθεσης στο πλαίσιο του διάσημου πειράματος *The Lady Tasting Tea*.

Ο Fisher εκλαΐκωσε τον "έλεγχο σημαντικότητας" για να βοηθήσει τους ερευνητές να αποφασίσουν αν θα τροποποιήσουν το επόμενο πείραμά τους ή αν ισχύει η μηδενική υπόθεση. Για την εφαρμογή του, ορίζεται η μηδενική υπόθεση, μια επιλεγμένη πιθανότητα για τη σωστή μηδενική υπόθεση γνωστή αργότερα ως επίπεδο σημαντικότητας, και ένα σύνολο δειγματικών τιμών. Ο έλεγχος βασίζεται στην υπολογισμένη πιθανότητα από την κατανομή που ακολουθεί η στατιστική ποσότητα που χρησιμοποιείται και από τις τιμές του δείγματος, και στη σύγκρισή της με την επιλεγμένη της τιμή. Αν η πιθανότητα αυτή, ορίζεται ως τιμή p , είναι μεγαλύτερη από την επιλεγμένη της τιμή, τότε υπάρχουν ισχυρές ενδείξεις ότι η μηδενική υπόθεση ισχύει. Αν και την τιμή αυτή τη χρησιμοποίησε ως κριτήριο ο Karl Pearson στον έλεγχο *χι τετράγωνο* και συνέταξε τους πρώτους πίνακες για τον υπολογισμό της, πρώτος τη χρησιμοποίησε σε στατιστικό έλεγχο ο Pierre Simon Laplace στη δεκαετία του 1770, κατά την ανάλυση των στατιστικών στοιχείων από σχεδόν μισό εκατομμύριο γεννήσεις. Οι στατιστικές τότε έδειξαν υπέρβαση των αγοριών σε σχέση με τα κορίτσια. Κατέληξε στον υπολογισμό μια τιμής p , που η υπέρβασή της οδηγούσε σε αποδεκτό αποτέλεσμα.

Σημαντικό μέρος της ιστορίας της στατιστικής καταλαμβάνει η διαμάχη μεταξύ των Pearson και Fisher, η οποία κράτησε πολλά χρόνια, ακόμα και μετά τον θάνατο του Pearson, και αφορούσε, όπως προκύπτει από την αλληλογραφία μεταξύ των τριών επιστημόνων (Pearson, Gosset και Fisher), στη διαφορετική φιλοσοφική θεώρηση της στατιστικής ως επιστήμης. Η Πηρσόνεια εκδοχή, όπως αναπτύχθηκε και παρουσιάστηκε από τον Egon Pearson, τον υιό, και τον πολωνό μαθηματικό Jerzy Neyman, θεωρεί τη στατιστική ως μια επιστήμη λήψης αποφάσεων, ενώ ο

30 Fisher, R.A. (1924). On a Distribution Yielding the Error Functions of Several Well Known Statistics *Proceedings of the International Congress of Mathematics, Toronto*, 2: 805 - 813.

31 Fisher, R.A. (1925). *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.

32 Fisher, R.A. (1935). *The Design of Experiments*. Edinburgh: Oliver and Boyd. Ακολούθησαν άλλες 8 εκδόσεις του βιβλίου μέχρι το 1971.

Fischer και οι κληρονόμοι της άποψής του την αποδέχονται ως μια επιστήμη ανάλυσης δεδομένων με σκοπό την απόσπαση πληροφοριών από τον φυσικό κόσμο για τη διαμόρφωση γνώσης. Η διαμάχη αυτή αποτελεί συνέχεια μιας ισχυρής αντίθεσης προς τη βιομετρική σχολή του Pearson που προήλθε από την ανακάλυψη στις αρχές του 20 αι. ενός άρθρου του Gregor Mendel, με τις θεωρίες του ιδρυτή της γενετικής επιστήμης για την κληρονομικότητα και τα αποτελέσματα των πειραμάτων του με τα μωσχομπίτζελα. Αυτή η διαμάχη με τους βιομέτρους να διεκδικούν τη στατιστική αυστηρότητα και τους μεντελιανούς την καλύτερη κατανόηση της βιολογίας, κυριάρχησε τις δύο πρώτες δεκαετίες του 20ου αιώνα, έως ότου, το 1918, ο Fisher επεξεργάστηκε στατιστικά τα αποτελέσματα του Mendel και κατάφερε να συνδέσει τις δύο προσεγγίσεις στη σύγχρονη σύνθεση της εξελικτικής βιολογίας.

Ο Πολωνός Jerzy Neyman, ένας από τους σημαντικότερους αρχιτέκτονες της σύγχρονης στατιστικής, υπήρξε καθηγητής του εργαστηρίου στατιστικής του πανεπιστημίου Berkeley της Καλιφόρνιας. Ενώ ήταν ακόμα στην Πολωνία ανέπτυξε την ιδέα της εκτίμησης του διαστήματος εμπιστοσύνης και την εφαρμογή της στον έλεγχο υποθέσεων και διατύπωσε στη συνέχεια τον *έλεγχο της μηδενικής υπόθεσης* σε συνεργασία με τον Egon Pearson, με τον οποίον γνωρίστηκε σε κάποιο ταξίδι του στο Παρίσι. Οι Neyman - Pearson θεωρούσαν ότι η διατύπωση του ελέγχου της μηδενικής υπόθεσης ήταν μια βελτιωμένη γενίκευση του ελέγχου σημαντικότητας. Ο Fisher πίστευε ότι ο έλεγχος της μηδενικής υπόθεσης δεν ήταν εφαρμοστέος στην επιστημονική έρευνα, γιατί συχνά, κατά τη διάρκεια ενός πειράματος, ανακαλύπτεται ότι οι αρχικές παραδοχές σχετικά με τη μηδενική υπόθεση είναι αμφισβητήσιμες λόγω απρόσμενων πηγών σφάλματος. Η χρήση άκαμπτων αποφάσεων του τύπου απόρριψη/αποδοχή που βασίζονται σε υποθέσεις που διατυπώθηκαν πριν από τη συλλογή των δεδομένων, μπορεί να οδηγήσει την επιστημονική έρευνα σε μαζική σύγχυση. Παρά τη σφοδρή διαμάχη με τις αντιλήψεις του Fischer, ο έλεγχος αυτός προέκυψε τελικά ως ένα υβρίδιο από τη σύνθεση όλων των απόψεων που αναπτύχθηκαν από τις δύο πλευρές στις αρχές του 20 αιώνα.

Κατά την ανάπτυξη της θεωρίας τους οι Neyman και Pearson αναγνώρισαν την ανάγκη της διατύπωσης και της εναλλακτικής υπόθεσης και όρισαν τις πιθανές λαθεμένες αποφάσεις κατά τον έλεγχο της μηδενικής υπόθεσης της σχετικής με την άγνωστη τιμή του πληθυσμού, με βάση τις εκτιμήσεις από τις παρατηρήσεις ενός δείγματος. Κάλεσαν σφάλμα πρώτου τύπου το σφάλμα απόρριψης μιας υπόθεσης που ισχύει και σφάλμα δεύτερου τύπου το σφάλμα αποδοχής μια λανθασμένης υπόθεσης. Έδωσαν σημασία στην πιθανότητα απόρριψης μιας υπόθεσης όταν είναι ψευδής. Κάλεσαν αυτή την πιθανότητα *ισχύς του ελέγχου* (power of test) και πρότειναν τον όρο *critical region* για να δηλώσουν ένα σύνολο δειγματικών στατιστικών τιμών που οδηγούν στην απόρριψη της υπόθεσης που ελέγχεται. Το "εμβάδόν" της κρίσιμης αυτής περιοχής, που την ονόμασαν *επίπεδο σημαντικότητας* (level of significance), είναι η πιθανότητα για σφάλμα πρώτου τύπου.

Ο έλεγχος των υποθέσεων και το κεφάλαιο της στατιστικής συμπερασματολογίας αποτελούν εφαρμογή της θεωρίας των πιθανοτήτων. Ο πυρήνας της ιστορικής διαφωνίας μεταξύ των Fisher και Neyman/Pearson οφείλεται καθαρά στις υποκειμενικές τους απόψεις πάνω στον ορισμό και στις φιλοσοφικές ερμηνείες των πιθανοτήτων. Οι επικρίσεις, οι διαφωνίες και τα σχόλια των στατιστικολόγων πάνω στη φιλοσοφία των πιθανοτήτων και του έλεγχου των υποθέσεων, που δε σταμάτησαν αλλά συνεχίζονται μέχρι σήμερα, συνέβαλαν στο να καταστεί η στατιστική ένα ισχυρό εργαλείο που άλλαξε ριζικά την πειραματική προσέγγιση όλων των επιστημών.

Βιβλιογραφία

- David, F. N. (1909): *Games, Gods and Gambling. A history of probability and statistical ideas*. (1962, Charles Griffin and Com., 1998 Dover Publications, Inc.).
- Delvin, K. (2010). *Φερμά - Πασκάλ. Το τελευταίο παιχνίδι*. Εκδόσεις Τραυλός. (2008, The unfinished game).
- Desrosières, A. (2005): *Η πολιτική των μεγάλων αριθμών. Ιστορία της στατιστικής αντίληψης*. Εκδόσεις Μεταίχμιο. (1993, *La politique des grands nombres. Histoire de la raison statistique*, Éditions La Découverte and Syros).
- Eisenhart, C. (1964): The Meaning of “Least” in Least Squares. *Journal of the Washington Academy of Science*, vol. 54, No. 2, 24-32.
- Everitt, B. S. (2001): Οι κανόνες της τύχης. Πιθανότητες, κίνδυνοι και στατιστική. Εκδόσεις Κάτοπτρο. (1999, *Chance Rules – An Informal Guide to Probability, Risk and Statistics*, Springer).
- Farebrother, R. W. (1999): Fitting Linear Relationships. *A History of the Calculus of Observations*. Springer.
- Folks, J. L. (1981): *Ideas of Statistics*. John Wiley and Sons.
- Hacking, I. (2006). *The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference*. Cambridge Series on Statistical & Probabilistic Mathematics.
- Koren, J. (1970): *The History of Statistics: Their Development and Progress in Many Countries*. Burt Franklin.
- Loria, G. (1971): *Ιστορία των μαθηματικών*. Έκδοση της Ελληνικής Μαθηματικής Εταιρείας.
- Owen, D. B. (1976): *On the History of Statistics and Probability*. Marcel Dekker, Inc.
- Pearson, K. (1978): *The History of Statistics in the 17th and 18th Centuries against the changing background of intellectual, scientific and religious thought*. Lectures by Karl Pearson given at the University College London during the academic sessions 1921-1933. Edited by E.S. Pearson. Charles Griffin, London.
- Plackett, R. L. (1949): A Historical Note on the Method of Least Squares. *Biometrika*, vol. 36, Issue 3/4, 458-460.

- Plackett, R. L. (1983): Karl Pearson and the chi-squared Test. *International Statistical Review*, 51, 59-72.
- Porter, T. M. (1986): *The Rise of Statistical Thinking 1820-1900*. Princeton University Press.
- Ρωσσικόπουλος, Δ. (2006): *Μέτρον Γεωμετρικόν. Η ιστορία των επιστημών της αποτόπωσης*. Εκδόσεις Ζήτη.
- Ρωσσικόπουλος, Δ. (2009): *Το γεωγραφικό μήκος. Ιστορίες με αστέρια, μεσημβρινούς και χάρτες*. Εκδόσεις Ζήτη.
- Salsburg, D. (2001): *The Lady Tasting Tea. How statistics revolutionized science in the twentieth century*. W. H. Freeman and Company.
- Seal, H. L. (1967): Studies in the History of Probability and Statistics. XV: The Historical Development of the Gauss Linear Model. *Biometrika*, vol. 54, Issue 1/2, 1-24.
- Stigler, S. M. (1981). Gauss and the Invention of Least Squares. *Ann. Statist.* 9 (3): 465 - 474.
- Stigler, S. M. (1986): *The History of Statistics. The Measurement of Uncertainty before 1900*. Harvard University Press.
- Stigler, S. M. (1999): Statistics on the Table. *The History of Statistical Concepts and Methods*. Harvard University Press.
- Todhunter, I. (1965): *A History of the Mathematical Theory of Probability*. Chelsea Publishing Company.
- Weigle, D. C. (1994): Historical Origins of Contemporary Statistical Testing Practices: How in the World Did Significance Testing Assume Its Current Place in Contemporary Analytic Practice? Paper presented at the *Annual meeting of the Southwest Educational Research Association*, San Antonio, Texas, January 27.
- Westergaard, H. (1969): *Contributions to the History of Statistics*. Augustus M. Kelley.