



Ακαδημία Εμπορικού Ναυτικού Ασπροπύργου
Σχολή Μηχανικών

Πτυχιακή εργασία
Οι πιθανότητες στην κοινωνία και στις κληρώσεις



Σπουδαστής: **Μαχρούς Ιωσήφ** (ΑΜ 8020)

Επιβλέπων Καθηγητής: **Στέφανος Ι. Καρναβάς, Μαθηματικός (Μ.Εδ.),**
Επίκουρος Καθηγητής

Ακαδημαϊκό έτος: **2020 – 2021**

Ημερομηνία ανάθεσης: 08. 05.2020

Ημερομηνία κατάθεσης:06.2021

Ημερομηνία εξέτασης:06.2021

A/A	Ονοματεπώνυμο	Χαρακτηρισμός	Υπογραφή
1	Στέφανος Ι. Καρναβάς Μαθηματικός (Μ.Εδ.) Επίκουρος Καθηγητής	Διακεκριμένη	
2	Αντώνιος Χαράλαμπος Ναυπηγός Μηχανολόγος Μηχανικός, Καθηγητής		
3	Χρηστίδου Αθανασία Ηλεκτρολόγος Μηχανικός Επίκουρη Καθηγήτρια		
Τελικός χαρακτηρισμός			

Περιεχόμενα

1. Κληρώσεις και λαχειοφόρες αγορές.....	4
2. Κληρώσεις με λίγους συμμετέχοντες και κληρώσεις με κάλπικο νόμισμα.....	4
3. Κληρώσεις για τρία ή περισσότερα άτομα.....	5
4. Κληρώσεις με πολλούς συμμετέχοντες.....	6
5. Μια καλά σχεδιασμένη επίσημη κλήρωση.....	7
6. Κακοσχεδιασμένες λαχειοφόρες αγορές.....	9
7. Η επιχείρηση της τύχης.....	9
8. Λαχειοφόρες αγορές, μαθηματική προσδοκία.....	10
9. Το ισπανικό χριστουγεννιάτικο λαχείο.....	10
10. Μία χριστουγεννιάτικη ιστορία.....	11
11. Πάντα πέφτει κάπου αλλού.....	11
12. Τα lotto.....	11
13. Το δέλεαρ των λώττο και το «στοίχημα του Pascal».....	12
14. Πόσους αριθμούς πρέπει να επιλέξουμε.....	12
15. Τα πλεονεκτήματα του να είναι κανείς «κανονικός». Μεγάλοι αριθμοί.....	13
16. Πολύ πιο εύκολο.....	15
17. Χρυσό θεώρημα του Bernoulli.....	15
18. Θεώρημα του Bernoulli.....	17
19. Δεν υπάρχει κακό που διαρκεί για πάντα, ή μήπως υπάρχει;.....	18
20. Μερικά λόγια για τη στατιστική.....	19
21. Καμπύλη του Gauss και κανονικότητα.....	19
22. Η κανονική καμπύλη.....	20
23. Ο νόμος των σφαλμάτων.....	21
24. Η υπόθεση των στοιχειωδών σφαλμάτων.....	22
25. Κεντρικό οριακό θεώρημα (ΚΟΘ).....	22
26. Ο πίνακας του Galton.....	23
27. John Graunt (1620–1674).....	24
28. Οι πιθανότητες στην κοινωνία. Πίνακες ζωής.....	24
29. Ορισμένοι τύποι.....	26
30. Πίνακες ζωής στην Ισπανία.....	26
31. Ασφάλειες.....	26
32. Ηλικία συνταξιοδότησης και συντάξεις.....	28
33. Άλλες εφαρμογές.....	28
34. Πιθανότητες και στατιστική στην ιατρική πρακτική.....	29
35. Πιθανότητες και DNA.....	29

1. Κληρώσεις και λαχειοφόρες αγορές

Ο σχεδιασμός δίκαιων κληρώσεων και λαχειοφόρων αγορών (στις οποίες όλοι οι συμμετέχοντες έχουν τις ίδιες πιθανότητες να κερδίσουν) είναι λίγο πιο περίπλοκος από ό τι φαίνεται. Μπορούμε να θεωρήσουμε ότι είναι κάτι που δε μας επηρεάζει, εκτός αν είμαστε, παίκτες, αλλά τότε δε θα έχουμε ακριβή αντίληψη της πραγματικότητας. Στην πραγματικότητα, σε όλη τη ζωή μας συμμετέχουμε σε πολλές κληρώσεις, μερικές φορές χωρίς να το καταλαβαίνουμε, όπως αυτές που γίνονται για την επιλογή **μελών εκλογικών τμημάτων** ή **ενόρκων** στις δίκες, η κατανομή σχολικών θέσεων στα μέρη και στις ηλικίες που είναι σπάνιες (στις πόλεις και στις ηλικίες 0 έως 6 ετών), η διανομή κοινωνικών κατοικιών, η συμμετοχή στις εξεταστικές επιτροπές διαγωνισμών (αν κάποιος είναι δημόσιος υπάλληλος) ο τόπος ή η σειρά συμμετοχής σε αυτές (αν είναι υποψήφιος). Υπήρξαν ιστορικά παραδείγματα, μερικά διαβόητα, για άδικες κληρώσεις και κακοσχεδιασμένες λαχειοφόρες αγορές (από τα οποία θα δούμε μερικά παραδείγματα), σε όλες τις περιπτώσεις χωρίς προμελετημένο τρόπο, νομίζοντας ότι έκαναν το σωστό, γεγονός που δείχνει ότι ακόμη και στον τομέα αυτόν προκύπτουν δυσκολίες σύλλληψης και εκτέλεσης. Γι' αυτό θ' αρχίσουμε με διάφορες καταστάσεις κληρώσεων για να κάνουμε σκέψεις που θα μας επιτρέψουν να είμαστε καλύτερα εξοπλισμένοι για να τις καταλάβουμε.

Η λέξη «λοταρία» για τη λαχειοφόρο αγορά προέρχεται από το lotto, ιταλική λέξη για να περιγράψει μια παρτίδα και ταυτόχρονα τον προορισμό. Ωστόσο, υπάρχουν αναφορές για λαχειοφόρες αγορές στην Παλαιά Διαθήκη ακόμη και στην Κίνα, όπου η μέθοδος αυτή χρησιμοποιήθηκε για τη χρηματοδότηση της κατασκευής του Σινικού Τείχους. Στην Ευρώπη, η ιστορία της λαχειοφόρου αγοράς ξεκινά το 1498 στην Πορτογαλία, που την επινόησε για να βοηθήσει τους ανήμπορους και για να ανταποκριθεί στις νομισματικές ανάγκες της χώρας. Το 1727 οι κάτω χώρες ίδρυσαν μία από παλαιότερας λαχειοφόρες αγορές στον κόσμο, αφού εξακολουθεί να λειτουργεί ακόμα. Αναζητούσαν να αποκαταστήσουν το δημόσιο ταμείο για τη χρηματοδότηση των πολέμων τους και για κατασκευή δημόσιων έργων. Στην Ισπανία η 1^η λαχειοφόρος αγορά ιδρύθηκε με μονοπωλιακό χαρακτήρα, το 1763, κατά τη βασιλεία του Καρόλου Γ'. Η σημερινή εθνική λοταρία γεννήθηκε στις **25.12.1811**, κατά τη διάρκεια του πολέμου της ανεξαρτησίας ως «ένα μέσο για την αύξηση των εσόδων του δημόσιου ταμείου χωρίς την επιβάρυνση των φορολογουμένων» και ονομάστηκε από τον κόσμο «σύγχρονη λοταρία» για να τη διαφοροποιήσουν από την «αρχική λοταρία».

Τα παιδιά και η λαχειοφόρος αγορά είναι αδιαχώριστα, επειδή η αθωότητα τους ως εγγύηση αμεροληψίας, ήταν η λύση στο πρόβλημα της εμπιστοσύνης για την εντιμότητα των κληρώσεων, για το λόγο αυτό οι μαθητές του κολεγίου Σαν Ιλντελφόνσο είναι αυτοί που ανακοινώνουν τα έπαθλα από το 1771.

2. Κληρώσεις με λίγους συμμετέχοντες και κληρώσεις με κάλπικο νόμισμα

Οι κληρώσεις με νόμισμα μπορεί να μην είναι πολύ συχνές, αλλά με αυτές επιχειρούμε να πετύχουμε μια καλύτερη αντίληψη των προβλημάτων που συνδέονται με τους σχεδιασμούς κληρώσεων. Έστω ότι έχουμε ένα νόμισμα και με αυτό θέλουμε να κάνουμε μία δίκαιη και αμερόληπτη κλήρωση για 2 άτομα. Τι να κάνουμε; Η πρώτη δυνατότητα είναι να θεωρήσουμε ότι το νόμισμα είναι κανονικό, δηλαδή ότι η πιθανότητα να έρθουν κορόνα ή γράμματα είναι ακριβώς ίδια, οπότε η λύση είναι απλή, ρίχνουμε κορόνα–γράμματα. Όμως ότι το νόμισμα είναι κανονικό, είναι κάτι που δε μπορούμε να γνωρίζουμε εκ των προτέρων.

Ωστόσο, με λίγη σκέψη μπορούμε να πετύχουμε μια κλήρωση που είναι πραγματικά δίκαιη, ανεξαρτήτως του νομίσματος. Έστω ότι η πιθανότητα να βγει κορόνα (C) είναι p (την οποία δε γνωρίζουμε και θα ήταν 0,5 αν το νόμισμα ήταν κανονικό), έτσι ώστε η πιθανότητα για γράμματα (X), το αντίθετο γεγονός με το προηγούμενο, θα είναι $(1-p)$. Ποια είναι η πιθανότητα να έχουμε CX (κορόνα–γράμματα) με αυτή τη σειρά; Αφού οι ρίψεις είναι ανεξάρτητες, το αποτέλεσμα της μίας δεν επηρεάζει το αποτέλεσμα της άλλης, είναι το

γινόμενο των πιθανοτήτων $prob(CX) = p(1-p)$. Και η πιθανότητα να έρθει XC με αυτή τη σειρά; Για τον ίδιο λόγο όπως προηγουμένως $prob(XC) = (1-p)p = prob(CX)$.

Άρα, έχουμε τον τρόπο για να κάνουμε μια δίκαιη κλήρωση μεταξύ δύο ατόμων με οποιαδήποτε νόμισμα. Ρίχνουμε το νόμισμα δύο συνεχόμενες φορές, με XC κερδίζει ένας από τους παίχτες με CX ο άλλος. Το μόνο μειονέκτημα είναι ότι οι ρίψεις στις οποίες έρχεται XX ή CC είναι ακατάλληλες, δε μπορούμε να τις λάβουμε υπόψη και πρέπει να συνεχίσουμε έως μια φορά να έχουμε CX ή XC. Αγνοώντας αυτές τις περιπτώσεις δεν αλλάζει ο δίκαιος χαρακτήρας της κλήρωσης. Άρα, οποιοδήποτε νόμισμα είναι καλό για να κάνουμε μια δίκαιη κλήρωση μεταξύ δύο ατόμων. Κάτι που δεν ισχύει μόνο για τα κέρματα και μπορεί να εφαρμοστεί σε οποιοδήποτε άλλο μέσο με δύο πιθανά αποτελέσματα, έστω και αν έχουν πολύ διαφορετικές πιθανότητες.

3. Κληρώσεις για τρία ή περισσότερα άτομα

Έστω τώρα ότι τα άτομα ανάμεσα στα οποία πρέπει να κάνουμε την κλήρωση είναι 3 και κάποιος προτείνει την εξής διαδικασία, ετοιμάζουμε μια τσάντα (ή αδιαφανές κουτί) με 3 μπάλες, 1 άσπρη και 2 μαύρες (όλες με ίδιο σχήμα και υφή ώστε να μη μπορούν να διακριθούν με την αφή). Οι 3 τραβούν με τη σειρά μια μπάλα που δεν επιστρέφουν στην τσάντα. Κερδίζει αυτός που βγάζει την άσπρη μπάλα. Είναι δίκαιο; Αν όχι, ποιος έχει το πλεονέκτημα, ο 1^{ος}, ο 2^{ος} ή ο 3^{ος} που τραβήξει;

Αν ήμασταν σε αυτή την κατάσταση και μπορούσαμε να επιλέξουμε σε ποια θέση θα προτιμούσαμε να τραβήξουμε την μπάλα; Για απλότητα μπορούμε να υποθέσουμε ότι οι μπάλες είναι αριθμημένες, 0, 1 και 2 και η άσπρη μπάλα είναι ο αριθμός 0. Ας δούμε τι συμβαίνει υπολογίζοντας την πιθανότητα να κερδίσει καθένας από τους 3 συμμετέχοντες.

Η σειρά τραβήγματος των 3 μπαλών είναι μία από τις έξι ακόλουθες, που αντιστοιχούν στις 6 δυνατές διατάξεις που μπορούν να προκύψουν 012, 021, 102, 120, 201, 210. Μια απλή οπτική παρατήρησή οδηγεί δυο συμπεράσματα ότι τις δυο από τις έξι φορές η μπάλα 0 θα βγει 1^η, σε άλλες δυο θα βγει 2^η και στις δύο υπόλοιπες 3^η. Με άλλα λόγια είναι αδιάφορη η σειρά με την οποία θα τραβήξουν τη μπάλα. Χρησιμοποιώντας τον κανόνα Laplace, είναι προφανές ότι η πιθανότητα να βγει η άσπρη μπάλα πρώτη είναι 2/6 (2 ευνοϊκές περιπτώσεις, 012 και 021, σε 6 δυνατές περιπτώσεις, που αντιστοιχούν σε οποιαδήποτε από τις προηγούμενες διατάξεις) και επίσης είναι 2/6 η πιθανότητα να βγει στη 2^η ή 3^η θέση. Ας δούμε τι θα συμβεί αν συλλογιστούμε με όρους δεσμευμένης πιθανότητας. Αφού υπάρχουν 3 μπάλες, η πιθανότητα να κερδίσει ο 1^{ος} είναι $prob(1^o) = 1/3$.

Για κερδίσει ο 2^{ος}, ο 1^{ος} πρέπει να έχει βγάλει μια μαύρη μπάλα και ο 2^{ος} την άσπρη. Με άλλα λόγια ο 2^{ος} πρέπει να έχει την δυνατότητα να τραβήξει (η πιθανότητα να συμβεί αυτό είναι 2/3, η δεσμευμένη πιθανότητα ότι ο 1^{ος} δεν έχει κερδίσει) και επιπλέον, να βγάλει την άσπρη (όταν υπάρχουν μόνο μια μαύρη και μια άσπρη). Συνολικά, $prob(2^o) = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}$.

Τέλος, $prob(3^o) = 1 - \frac{1}{3} - \frac{1}{3} = \frac{1}{3}$. Αν αντί για 3 είχε 100 ή N συμμετέχοντες, θα

βάζαμε 100 (ή N) μπάλες και παρόμοιοι συλλογισμοί μας οδηγούν να βεβαιώσουμε ότι η πιθανότητα να βγει η μπάλα σε μια συγκεκριμένη θέση είναι 1/100 (ή 1/N). Μία αντίστοιχη κατάσταση παρουσιάζεται όταν ο δάσκαλος θέλει να κληρώσει ένα βραβείο μεταξύ των N μαθητών και γι' αυτό γράφει έναν αριθμό μεταξύ 1 και N σε χαρτί και ζητά από κάθε μαθητή «με αλφαβητική σειρά» έναν αριθμό μεταξύ 1 και N, δίνοντας το βραβείο σε όποιον μαντέψει τον αριθμό που έγραψε. Και πάλι η σειρά (αλφαβητική στην προκειμένη περίπτωση) δεν έχει σημασία.

4. Κληρώσεις με πολλούς συμμετέχοντες

Για να μπορούν να αναλυθούν οι δυσκολίες των κληρώσεων με πολλούς συμμετέχοντες θα αναφερθούμε σε δύο ιστορικά γεγονότα που και τα δύο σχετίζονται με στρατούς. Και στις δύο περιπτώσεις οργανώθηκαν με την εντύπωση ότι ήταν δίκαιες, ότι όλοι οι συμμετέχοντες είχαν τις ίδιες πιθανότητες να επιλέγουν ή να αποκλεισθούν. Στις ΗΠΑ το 1970, με τον πόλεμο του Βιετνάμ στο απόγειό του, την ανάγκη για στρατιώτες και μια ατμόσφαιρα απόρριψής του, οργανώθηκε μια στρατιωτική κλήρωση για να επιλέξουν τους νέους που έπρεπε να αποσταλούν. Εδώ το ενδιαφέρον ήταν στο να μην επιλεγεί κάποιος. Για να οργανώσουν την κλήρωση τοποθέτησαν τις 366 πιθανές ημερομηνίες ενός έτους (από 01.01 έως 31.12, συμπεριλαμβανομένης της 29.02) κάθε μία σε μια κάψουλα και όλες τραβήχτηκαν στην τύχη. Η πρώτη που βγήκε ήταν η 14.09, μετά η 24.04 και με τον τρόπο αυτό στρατολογήθηκαν στρατιώτες, που γεννήθηκαν μεταξύ 1941 και 1952, ακολουθώντας τη σειρά των ημερομηνιών που εμφανίστηκαν στην κλήρωση (πρώτα αυτοί της 14.09, στη συνέχεια αυτοί της 24.04). Πώς σας φαίνεται η διαδικασία; Είναι δίκαιη;

Στην Ισπανία το 1997, όπου η στρατιωτική θητεία είναι υποχρεωτική για τους άνδρες μιας συγκεκριμένης ηλικίας, οι στρατευσιμοι ήταν 365.342. αλλά οι ανάγκες του στρατού ήταν λιγότερες, έτσι ώστε 16.442 από αυτούς δεν θα πάνε στρατό. Στην περίπτωση αυτή η επιλογή στην κλήρωση ήταν κάτι επιθυμητό. Οργανώθηκε με τον ακόλουθο τρόπο, πρώτα με τυχαίο τρόπο δόθηκε ένας αριθμός σε καθέναν από τους 365.342 στρατευσίμους. Για να γίνει η κλήρωση τοποθετήθηκαν 6 δοχεία με μπάλες για καθένα από τα 6 ψηφία του αριθμού που θα επιλεγόταν και από τον οποίο θα αποδεσμεύονταν 16.442 αγόρια που είχαν αριθμούς συνεχόμενους με αυτόν που βγήκε (και αν έφταναν στο τέλος θα ξεκινούσαν ξανά από τον αριθμό 1 μέχρι να ολοκληρωθεί η απαιτούμενη ποσότητα). Σε όλα τα άλλα δοχεία είχε 10 μπάλες αριθμημένες από το 0 έως το 9. Είχε προβλεφθεί ότι εάν στο 2^ο δοχείο (με τις δεκάδες χιλιάδες) έβγαине 1 μπάλα μεγαλύτερη από το 6 (όπως συνέβη στην πραγματικότητα) θα επαναλαμβανόταν το τράβηγμα στο δοχείο αυτό. Πώς σας φαίνεται η κλήρωση στην περίπτωση αυτή; Ήταν δίκαιη;

Πριν προχωρήσουμε για να εξετάσουμε την υπόλοιπη ιστορία των δύο αυτών στρατιωτικών κληρώσεων, θα δούμε μια άλλη που χρησιμοποιήθηκε για αρκετό καιρό για την κατανομή θέσεων σε ιδρύματα για τα οποία υπήρχε μεγαλύτερη ζήτηση παρά προσφορά, πχ. ορισμένες σχολές ξένων γλωσσών. Αποδιδόταν ένας αριθμός με σειρά προεγγραφής και έβγαине με κλήρωση ένας από αυτούς τους αριθμούς. Αν υπήρχαν 50 θέσεις, απονέμονταν στον αριθμό αυτόν και στους υπόλοιπους 49. Τι είναι αυτό που το κάνει μη σωστό; Ας φανταστούμε ότι θέλω να εγγραφώ. Αν πείσω 30 φίλους να γραφούν πριν από μένα και οι αριθμοί για την κλήρωση δίνονται με σειρά εγγραφής, θα παίξω με πλεονέκτημα επειδή ξέρω ότι όλοι οι φίλοι μου θα παραιτηθούν από τη θέση τους. Στην περίπτωση αυτή δεν παίξω μόνο με τον αριθμό μου, αλλά με τον δικό μου και τους 30 προηγούμενους των φίλων μου. Κάποιος μπορεί να σκεφτεί ότι είναι πολύ πολύπλοκος τρόπος για να αυξηθεί η πιθανότητα, αλλά η αλήθεια είναι ότι χρησιμοποιήθηκε για αρκετό χρονικό διάστημα, έτσι ώστε η διαδικασία έχει πλέον αλλάξει και οι αριθμοί δεν αποδίδονται με τρόπο που σχετίζεται με την προεγγραφή.

Επίσης έχει αλλάξει για τον ίδιο λόγο το είδος της κλήρωσης κατά τη διανομή των κοινωνικών κατοικιών, που γινόταν με τον ίδιο τρόπο (αν και στην προκειμένη περίπτωση είχε μικρότερη σημασία γιατί αν παρουσιαζόσουν για να αρνηθείς στη συνέχεια κινδύνευες να μπεις σε μαύρη λίστα). Ας επιστρέφουμε στην κλήρωση του Βιετνάμ. Το αποτέλεσμα αποκάλυψε ότι εκείνοι που γεννήθηκαν στους τελευταίους μήνες του έτους ήταν πολύ περισσότεροι από εκείνους που γεννήθηκαν στους υπόλοιπους μήνες. Δεδομένου ότι η ημερομηνία γέννησης θεωρείται τυχαία (αν και αυτό θα μπορούσε να ανοίξει μια άλλη συζήτηση), ήταν αναμενόμενο ότι η κατανομή των ημερομηνιών θα ήταν επίσης τυχαία. Τι απέτυχε; Οι κάψουλες με τις ημερομηνίες γέννησης μπήκαν στο δοχείο αρχίζοντας από την

1^η Ιανουαρίου και έπειτα, με αυστηρή σειρά ημερομηνιών μέχρι να φτάσουν στις 31 Δεκεμβρίου. Χωρίς να τις αναμείξουν ιδιαίτερα άρχισαν να τραβούν, έτσι αυτές του τέλους του έτους που ήταν επάνω, βγήκαν σε πολύ υψηλά ποσοστά από το αναμενόμενο. Πώς θα μπορούσε να είχε γίνει με πιο απλό και δίκαιο τρόπο; Για παράδειγμα, να τραβήξουν μόνο μία ημερομηνία στην τύχη, μετά να γυρίσουν κατάλληλα το δοχείο και να επιλέξουν τους απαιτούμενους άνδρες που γεννήθηκαν την ημερομηνία αυτή γρήγορο και πιο δίκαιο.

Στην περίπτωση της Ισπανίας, το γεγονός ότι στο 1^ο δοχείο υπήρχε η ίδια πιθανότητα να βγει 0 ή 1, αφού υπήρχαν πέντε μπάλες για το κάθε ένα, έκανε ότι η πιθανότητα να βγει ένας αριθμός εξαρτιόταν από τον αριθμό που είχε. Έτσι, αγνοώντας το θέμα της ενδεχόμενης επανάληψης του αριθμού που βγήκε στο 2^ο δοχείο, για τους αριθμούς μεταξύ 1 και 99.999, η πιθανότητα να βγουν είναι $p(P) = \frac{1}{2} \frac{1}{99.999} = 0,000005$ ενώ για τους υπόλοιπους αριθμούς

είναι $p(S) = \frac{1}{2} \frac{1}{65.343} = 0,00000765$.

Δηλαδή, οι αριθμοί της τελευταίας ομάδας είχαν πιθανότητα να εκλεγούν σημαντικά υψηλότερη από τους πρώτους, περίπου 50% παραπάνω. Ο αριθμός που βγήκε στην κλήρωση ήταν το 155.611. Στην περίπτωση του πρώτου 5, χρειάστηκε να επαναληφθεί το τράβηγμα, επειδή αρχικά βγήκε ένα 8 (επειδή είναι μεγαλύτερο από 6). Απαλλάχθηκαν από τη στρατιωτική υπηρεσία από τον αριθμό αυτόν μέχρι το τέλος και στη συνέχεια από το 1 έως να ολοκληρωθούν οι 16.442 απαλλαγμένες. Ο πολιτικός υπεύθυνος για την οργάνωσή του και αναπληρωτής γραμματέας άμυνας εκείνη τη χρονική περίοδο, μπροστά στις ερωτήσεις των δημοσιογράφων για τις πιθανότητες και τα υπόλοιπα δεν είχε κανέναν ενδοιασμό να απαντήσει «εγώ δεν καταλαβαίνω από συλλογισμούς πιθανοτήτων γιατί είμαι των γραμμάτων». Η κλήρωση δεν επαναλήφθηκε, επειδή, σύμφωνα με πληροφορίες από το στρατό, «είχε διατηρηθεί η ισότητα ευκαιριών για όλους τους εμπλεκόμενους επειδή η ανάθεση αριθμών σε καθένα είχε γίνει με τυχαίο τρόπο».

Η τελευταία αυτή δήλωση είναι ορθή, αφού για τον υπολογισμό της πραγματικής πιθανότητας να βγει ο αριθμός ενός νέου στην κλήρωση, οι προηγούμενες πιθανότητες θα πρέπει να πολλαπλασιαστούν με την πιθανότητα ανάθεσης ενός αριθμού μεταξύ 1 και 99.999 ή μεταξύ 100.000 και 165.343. Το γεγονός αυτό δείχνει ότι οι πιθανότητες αποδεικνύονται δύσκολες να κατανοηθούν από πολλούς ανθρώπους και παρά την άγνοια που έδειξαν οι υπεύθυνοι για την κλήρωση, αυτή ήταν σωστή. Για ανάλογη αιτία δεν είναι δίκαιη μια κλήρωση για να αποτελέσεις μέλος κάποιου δικαστηρίου, κριτικής επιτροπής, προεδρείου ή παρομοίων στην οποία βγαίνουν τα δυο πρώτα γράμματα του πρώτου επώνυμου. Σε αυτή την περίπτωση όλα τα επώνυμα εξίσου πιθανά, και κάποιος, με επώνυμο για παράδειγμα, «Παράσχος», αφού η μόνη δυνατότητα για να επιλεγεί είναι να βγει «ΠΑ», έχοντας μπροστά του πολύ κοινά επώνυμα, πως «Παπαδόπουλος», δεν θα βγει σχεδόν ποτέ.

5. Μια καλά σχεδιασμένη επίσημη κλήρωση

Η κυβέρνηση της Αραγονίας ζήτησε από το τμήμα στατιστικών μεθόδων του Πανεπιστημίου της Σαραγόσα το σχεδιασμό κλήρωσης ώστε η διάθεση των κοινωνικών κατοικιών να πραγματοποιηθεί με δίκαιη διαδικασία μεταξύ όλων των αιτούντων. Όταν προσφορές για κατοικίες που πρέπει να διατεθούν έχουν μεγάλο αριθμό αιτούντων, η δυνατότητα διάθεσης ενός δοχείου με όλους τους συμμετέχοντες αριθμούς στην κλήρωση μπορεί να αποδειχθεί δύσκολο έργο να διεξαχθεί πρακτικά για διάφορους λόγους, μεταξύ των οποίων το δικαίωμα των συμμετεχόντων να ελέγξουν ότι οι αριθμοί τους εισάγοντα πραγματικά στο δοχείο. Γι' αυτό, στις περιπτώσεις αυτές συνιστάται η κλήρωση να πραγματοποιηθεί με τη διαδικασία πολλών δοχείων, ενός δοχείου για κάθε ψηφίο. Κάθε δοχείο πρέπει να περιέχει 10 μπάλες με τα ψηφία 0 έως 9. Στη συνέχεια και με σκοπό την απλοποίηση της παρουσίας, υποθέτουμε ότι υπάρχουν πάνω από 10.000 συμμετέχοντες,

δηλαδή, χρειάζονται πέντε δοχεία, το 1^ο από τα οποία αντιστοιχεί στις μονάδες, το 2^ο στις δεκάδες, το 3^ο στις εκατοντάδες, το 4^ο στις χιλιάδες και το 5^ο στις δεκάδες χιλιάδες. Αν ο αριθμός συμμετεχόντων είναι από 1.000 ως 10.000, όλα όσο θα εκτεθούν παραμένουν έγκυρα, αλλά θα υπάρχουν μόνο 4 δοχεία, έτσι όλα όσα ειπωθούν για το 5^ο δοχείο θα ισχύουν για το 4^ο. Ομοίως, αν η κλήρωση είναι μεταξύ περισσότερων από 100 άτομα και με μέγιστο 1.000 θα χρειαστούν μόνο 3 δοχεία και το ρόλο του 5^{ου} δοχείου θα παίζει το 3^ο.

Είναι σαφές ότι για κληρώσεις με 100 ή λιγότερους συμμετέχοντες ακολουθείται η διαδικασία του ενός δοχείου. Αν αποφασιστεί να χρησιμοποιηθούν πολλά δοχεία απαιτούνται μόνο 2 και αυτό των δεκάδων θα παίζει τον ρόλο που στη συνέχεια ανατίθεται στο 5^ο δοχείο. Είναι απαραίτητο να γίνει αυτή η τροποποίηση, επειδή στη διαδικασία πολλών δοχείων η εμφάνιση του 0 είναι το ίδιο πιθανή με οποιοσδήποτε άλλου αριθμού. Τώρα πρέπει να εξετάσουμε τις δυνατές καταστάσεις. Θυμίζουμε ότι σκεφτόμαστε για μια κλήρωση με 5 δοχεία (περισσότεροι από 10.000 συμμετέχοντες στην κλήρωση). Έστω ότι συμμετέχουν σε αυτή ένας αριθμός ατόμων που είναι πολλαπλάσιο του 10.000, για παράδειγμα 30.000, δηλαδή οι συμμετέχοντες έχουν αριθμούς μεταξύ 0 και 29.999. Στην περίπτωση αυτή, στο 5^ο δοχείο πρέπει να μπουν μπάλες αριθμημένες από 0 έως 2. Έτσι, για να σχηματιστεί ο αριθμός που κερδίζει πρέπει να γίνει 1 τράβηγμα από κάθε δοχείο για να σχηματιστεί ένας από τους αριθμούς που συμμετέχουν.

Αν ο αριθμός συμμετεχόντων στην κλήρωση δεν είναι πολλαπλάσιο του 10.000, για παράδειγμα 53.427, στο δοχείο που αντιστοιχεί στις δεκάδες χιλιάδες, το 5^ο δοχείο, θα εισαχθούν μπάλες από 0 μέχρι 5 και θα συνεχίσουμε όπως στην προηγούμενη περίπτωση, τραβώντας 1 μπάλα από κάθε δοχείο για να σχηματιστεί ένας αριθμός μεταξύ 0 και 59.999. Αν ο αριθμός που προκύπτει αντιστοιχεί σε αριθμό που δεν υπάρχει, πρέπει να επιστρέφουν όλες οι μπάλες που βγήκαν στα αντίστοιχα δοχεία τους, να μην είχε γίνει το 1^ο τράβηγμα και θα πραγματοποιηθεί ένα 2^ο από κάθε δοχείο για να σχηματιστεί ο νέος αριθμός.

Η διαδικασία θα επαναληφθεί αν είναι απαραίτητο, έως ότου ο αριθμός που προκύπτει να είναι έγκυρος. Για να δούμε ότι η διαδικασία είναι δίκαιη για όλους τους συμμετέχοντες στην κλήρωση θα χρησιμοποιήσουμε ένα συγκεκριμένο παράδειγμα για μεγαλύτερη ευκολία έκθεσης και χωρίς να εκχωρήσουμε το μηδέν σε κανένα. Έστω ότι οι συμμετέχοντες έχουν αριθμούς που κυμαίνονται από το 1 έως το 53.427. Η διαδικασία που περιγράφεται απαιτεί τη χρήση 5 δοχείων (μονάδες, δεκάδες, εκατοντάδες, χιλιάδες και δεκάδες χιλιάδες), την εισαγωγή μπαλών από 0 έως 9 στα 4 πρώτα, και στο 5^ο, μπάλες από 0 έως 5. Γίνεται ένα τράβηγμα σε κάθε δοχείο για να σχηματιστεί ο αριθμός, αν είναι μεταξύ 1 και 53.427, είναι ο αριθμός που «κερδίζει».

Σε αντίθετη περίπτωση, αν ο αριθμός που λαμβάνεται είναι το 0 (00000) ή μεγαλύτερος από 53.427, όλες οι μπάλες επιστρέφουν στα αντίστοιχα δοχεία τους και γίνεται νέο τράβηγμα των 5, επαναλαμβάνοντας αυτή τη διαδικασία μέχρι να λειφθεί αριθμός μεταξύ 1 και 53.427. Για να είναι αυτή η διαδικασία είναι δίκαιη πρέπει να ελέγξουμε ότι όλοι οι αριθμοί μεταξύ 1 και 53.427 έχουν τις ίδιες πιθανότητες (δηλαδή, $1/53.427$) να κερδίσουν. Θα εξετάσουμε ένα συγκεκριμένο αριθμό, πχ. το 12.525 και θα υπολογίσουμε την πιθανότητα να κερδίσει. Πρώτα, πρέπει να παρατηρηθούμε ότι βγάζοντας μια μπάλα από κάθε δοχείο, όλοι οι αριθμοί μεταξύ 00000 και 59.999 έχουν πιθανότητα να βγουν

$\frac{1}{10} \frac{1}{10} \frac{1}{10} \frac{1}{10} \frac{1}{6} = \frac{1}{60.000}$ αφού η πιθανότητα να βγει μια συγκεκριμένη μπάλα από κάθε δοχείο είναι 1 διαιρούμενο με τον αριθμό των μπαλών. Έτσι, η πιθανότητα να ληφθεί ο αριθμός 12.525 με το πρώτο τράβηγμα είναι $\frac{1}{60.000}$. Ωστόσο, μπορεί με το 1^ο τράβηγμα να

βγει ένας αριθμός που δεν είναι έγκυρος και πρέπει να πραγματοποιηθεί ένα άλλο τράβηγμα στο οποίο μπορεί επίσης να βγει ο αριθμός 12.525. Για να συμβεί αυτό, το 1^ο τράβηγμα στο οποίο να μην είναι έγκυρο, δηλαδή πρέπει να βγει είτε το 00000 είτε ένας αριθμός,

μεγαλύτερος από 53.427, συνολικά, υπάρχουν 6.572 (οι αριθμοί μεταξύ 53.428 και 59.999) συν 1 (το 00000) μη έγκυροι, και έτσι η πιθανότητα ότι στο 1^ο τράβηγμα βγαίνει ένας μη έγκυρος αριθμός είναι $\frac{\text{μη έγκυροι αριθμοί}}{\text{σύνολο αριθμών}} = \frac{6.573}{60.000}$. Άρα, η πιθανότητα να βγει ο αριθμός

12.525 σε αυτό το 2^ο τράβηγμα είναι η πιθανότητα ότι το 1^ο δεν είναι έγκυρο $\frac{6.573}{60.000}$, πολλαπλασιασμένη επί την πιθανότητα ότι στο 2^ο τράβηγμα θα βγει το 12.525, που είναι 1/60.000. Έτσι, αυτή η πιθανότητα είναι $\frac{6.573}{(60.000)^2}$. Επαναλαμβάνοντας αυτό τον

υπολογισμό και για την κατάσταση στην οποία τα πρώτα n τραβήγματα δεν είναι έγκυρα και ο αριθμός που λαμβάνεται στο τράβηγμα $(n+1)$ είναι το 12.525 και αθροίζοντας όλες τις πιθανότητες, βρίσκουμε ότι τελικά η πιθανότητα να κερδίσει ο αριθμός 12.525 είναι

$$P(12.525 \text{ νικηφόρα}) = \frac{1}{60.000} + \frac{1}{60.000} \frac{6.573}{60.000} + \frac{1}{60.000} \left(\frac{6.573}{60.000} \right)^2 + \dots + \frac{1}{60.000} \left(\frac{6.573}{60.000} \right)^n + \dots$$

$$= \frac{1}{60.000} \frac{1}{1 - \frac{6.573}{60.000}} = \frac{1}{60.000} \frac{60.000}{53.427} = \frac{1}{53.427}$$

Άρα, η διαδικασία είναι δίκαιη.

6. Κακοσχεδιασμένες λαχειοφόρες αγορές

Το κράτος (ή ο διοργανωτής) κερδίζει πόντο εκτός από την περίπτωση κακοσχεδιασμένων λαχειοφόρων αγορών. Υπάρχουν μερικά παραδείγματα λαχειοφόρων αγορών, όπως αυτή του 1728, όταν η πόλη του Παρισιού οργάνωσε μια μηνιαία λαχειοφόρο αγορά προς άντληση κεφαλαίων για την αντιμετώπιση των υποχρεώσεων της. Ο διάσημος φιλόσοφος **Βολτέρος** (1694–1778), παρά τους συλλογισμούς του για το τυχαίο, του τύπου η τύχη είναι μια λέξη δίχως νόημα, τίποτα δε μπορεί να υπάρξει χωρίς αιτία, που κάνει τις ντετερμινιστικές ομολογίες πίστης τυπικές του Διαφωτισμού, με τη βοήθεια ενός φίλου μαθηματικού συνειδητοποίησε ότι βρισκόταν μπροστά σε μια ονειρική κατάσταση: αγοράζονται όλους τους αριθμούς κέρδιζε σίγουρα. Έτσι κι έκανε, δημιουργώντας μία εταιρία με έναν αριθμό φίλων, γεγονός που συνέβαλε στην οικονομική του κατάσταση την οποία δεν παρέλειπε να αυξάνει με τα λογοτεχνικά έργα του και διάφορες επιχειρήσεις.

7. Η επιχείρηση της τύχης

Το χρηματικά ποσά που παίζουν οι χρήστες στους διάφορους τρόπους τυχερών παιχνιδιών (συμπεριλαμβανομένων των λαχειοφόρων αγορών και των καζίνο, μπίνγκο και άλλων αιθουσών τυχερών παιχνιδιών) μπορεί να αντιπροσωπεύουν σε ορισμένες δυτικές χώρες, τόσα υψηλά ποσοστά επί του ΑΕΠ όπως το 3%. Αυτός ο όγκος δαπανών μπορεί να διατηρηθεί ακόμη και σε περιόδους οικονομικών κρίσεων όπως αποδείχθηκε τα τελευταία χρόνια. Όσον αφορά τις άλλες κοινωνικές πτυχές των κληρώσεων με σημαντικά βραβεία, δεν παίζονται μόνο τα χρήματα που θα κερδηθούν, αλλά μια ολόκληρη σειρά από προλήψεις και προκαταλήψεις, η τύχη του παίχτη, τα περίεργα προαισθήματα με τους αριθμούς, τα προειδοποιητικά όνειρα ή τα μέρη της αγοράς των λαχνών (είναι καλύτερα αν έχει συμβεί σε αυτό μια καταστροφή, επειδή λόγω κάποιου είδους θεϊκής ή εξωγήινης ανταπόδοσης, η πιθανότητα κέρδους είναι μεγαλύτερη). Ένα παράδειγμα είναι ενός νικητή του 1^{ου} βραβείου του λαχείου με αριθμό που τέλειωνε στο 48 ο οποίος εξήγησε γιατί το είχε επιλέξει: «Ονειρεύτηκα το επτά, επτά νύχτες στη συνέχεια και αφού επτά επί απτά κάνει σαράντα οκτώ, αγόρασα ένα 48». Πώς να κάνεις να πιστέψει αυτός ο οραματιστής ότι $7 \cdot 7 = 49$.

8. Λαχειοφόρες αγορές, μαθηματική προσδοκία

Στις κοινωνίες υπάρχει πάντα αντίσταση στην πληρωμή φόρων, εκτός από έναν που καταβάλλεται οικειοθελώς και με ευχαρίστηση, τις επίσημες λαχειοφόρες αγορές. Επειδή σε αυτές δίχως καμιά αμφιβολία αυτός που κερδίζει πάντα είναι το κράτος. Γι' αυτό υπάρχουν επιγραμματικές φράσεις για το θέμα, όπως του Άγγλου συγγραφέα Henry Fielding (1707–1754) «Μια λαχειοφόρος αγορά είναι ένας φόρος σε όλους τους ανόητους της δημιουργίας και δόξα στον ουρανό, επιβάλλεται με ευκολία, η ευπιστία είναι πάντα στη μόδα».

Η ανάλυση των τυχερών παιχνιδιών έχει αναμφισβήτητο μαθηματικό ενδιαφέρον και κοινωνικό. Λαμβάνοντας υπόψη τον μεγάλο αριθμό ανθρώπων και τις οικονομικές ποσότητες που διακινούνται, είναι ένα σημαντικό θέμα σε όλες τις χώρες και η Ισπανία ειδικότερα είναι μια παγκόσμια δύναμη στα παιχνίδια, με δεδομένα που το επιβεβαιώνουν.

Όπως θυμίζουν οι διαφημίσεις, το ισπανικό χριστουγεννιάτικο λαχείο είναι αυτό που μοιράζει τα περισσότερα βραβεία σε όλο τον κόσμο και έτσι αναφέρεται στα διεθνή μέσα ενημέρωσης με ένα μήνυμα τουλάχιστον με έμμεσο τρόπο, ότι αυτό σημαίνει ότι η πιθανότητα απόκτησης βραβείων είναι μεγαλύτερη. Δεδομένου ότι σε πολλά παιχνίδια όπως η λαχειοφόρος αγορά υπάρχει η δυνατότητα να αποκτηθούν διάφορα βραβεία, μια σημαντική έννοια είναι της μαθηματικής ελπίδας του κέρδους, η οποία παρέχει πληροφορίες σχετικά με το μέσο κέρδος που θα μπορούσαμε να αποκτήσουμε. Για τον υπολογισμό του καθορίζονται αρχικά οι πιθανότητες κάθε βραβείου με κλασική μέθοδο, ο λόγος μεταξύ του αριθμού των ευνοϊκών περιπτώσεων να πάρουμε ένα βραβείο και του συνόλου των πιθανών περιπτώσεων.

Για παράδειγμα, σε μια κλήρωση με 100 λαχνούς ή αριθμούς, ένα 1^ο βραβείο και 2 δεύτερα, η πιθανότητα να κερδίσουμε το 1^ο βραβείο παίζοντας ένα λαχνό είναι $1/100 = 0,01$ (ή 1%) και ένα δεύτερο, $2/100 = 0,02$ (2%). Η πιθανότητα να μην κερδίσουμε είναι $97/100 = 0,97$ (97%). Η μαθηματική προσδοκία του κέρδους σε ένα παιχνίδι είναι το άθροισμα των γινομένων των δυνατών κερδών επί τις πιθανότητες απόκτησής τους μείον τα χρήματα που καταβλήθηκαν για τη συμμετοχή. Στην προηγούμενη λοταρία, αν κάθε λαχνός κοστίζει 5€, το πρώτο βραβείο είναι 100€ και τα δύο δεύτερα 40€, η ελπίδα του κέρδους είναι $100 \frac{1}{100} + 40 \frac{2}{100} - 5 = 1 + 0,8 - 5 = -3,2$. Αν η ελπίδα είναι ένας αρνητικός αριθμός (όπως στην περίπτωση αυτή) το παιχνίδι είναι δυσμενές για τον παίκτη. Αν ήταν θετικός (κάτι που δε συμβαίνει συχνά, εκτός από λάθος σχεδιασμό), θα ήταν ευνοϊκό και αν ήταν μηδέν, θα ήταν ένα αμερόληπτο ή δίκαιο παιχνίδι για τα δύο μέρη.

9. Το ισπανικό χριστουγεννιάτικο λαχείο

Στην Ισπανία παίζουν το χριστουγεννιάτικο λαχείο κυρίως λόγω εθίμου. Είναι σχεδόν μια κοινωνική υποχρέωση που συνδέεται με τα μαντολάτα, τα δώρα και τις γιορτές και συνήθως παίζεται στην εργασία, με τους φίλους ή την οικογένεια, παρόλο που αυτό που κάνουμε είναι μια εθελοντική καταβολή φόρων. Κοιτάζοντας οποιονδήποτε λαχνό βλέπουμε ποιά είναι τα βραβεία στους 85.000 αριθμούς του χριστουγεννιάτικου λαχείου, που κυμαίνονται από το 00000 έως το 84.999. Η πιθανότητα να πέσει το μέγιστο βραβείο, παίζοντας με έναν αριθμό, είναι $1/85.000 = 0,00001176$, το ίδιο όπως να πέσει το δεύτερο ή το τρίτο (τόσο μικρή όπως αναφέρεται στο ένθετο «Μια χριστουγεννιάτικη ιστορία»).

Η πιθανότητα να πέσει ένα από τα δύο τέταρτα ή τα οκτώ πέμπτα είναι $2/85.000$ και $8/85.000$, αντίστοιχα. Ότι μπορεί κάποιος να παρηγορηθεί με μικροκέρδη, με 1.774 αριθμούς, είναι 1.774 φορές μεγαλύτερη από ότι με το πρώτο βραβείο. Προσθέτοντας όλα τα βραβεία διαπιστώνεται ότι συνολικά υπάρχουν 13.354 (με εκατοντάδες, καταλήξεις και άλλες δυνατότητες), έτσι ώστε η πιθανότητα ένας παίκτης να κερδίσει κάποιο με ένα μόνο λαχνό του ενός δέκατου είναι $13.354/85.000 = 0,156870588$ σχεδόν $1/6$, δηλαδή κάτι λιγότερο από 16%. Διατυπωμένο από την άποψη του παίκτη, αν ένα άτομο παίζει τακτικά 25 διαφορετικούς αριθμούς κάθε χρόνο, θα λάβει περίπου τέσσερα ετήσια βραβεία, τα μισά από

τα οποία θα είναι απλές επιστροφές. Πολλαπλασιάζοντας τα βραβεία της λαχειοφόρου αγοράς με την πιθανότητα απόκτησής τους και αθροίζοντας φαίνεται ότι η μαθηματική προσδοκία κέρδους, για ένα δέκατο (που κοστίζει 20€) είναι -6 € (30% της τιμής του λαχνού). Με άλλα λόγια το παιχνίδι είναι δυσμενές για τους παίκτες φυσικά και κερδοφόρο για το κράτος.

10. Μία χριστουγεννιάτικη ιστορία

Ο Χουάν έχει ένα φίλο, το φύλο του οποίου δε λέει που ζει στη Θιουδάδ Ρεάλ και έχει πολύ καιρό δει (και δε γνωρίζει ούτε την τωρινή διεύθυνσή του) και κατά την τελευταία συνάντησή τους υποσχέθηκε να του στείλει αυτό το βιβλίο που τους αρέσει τόσο. Πριν από λίγες μέρες συνάντησε έναν άλλο φίλο, τον Λουίς, που του λέει ότι θα ταξιδέψει στη Θιουδάδ Ρεάλ. Ο Χουάν θυμάται την ανεκπλήρωτη υπόσχεση και δίνει στον Λουίς το βιβλίο για το φίλο του. Για να μην το χαλάσει ο Λουίς το βάζει μέσα σε ένα μεγάλο φάκελο και φεύγει για τη Θιουδάδ Ρεάλ. Όταν φτάνει εκεί αφήνει το αυτοκίνητο σε μια οδό όπου καταφέρνει να βρει θέση, βγαίνει και στο πρώτο άτομο που βλέπει του δίνει τον φάκελο «Ορίστε το βιβλίο που σου στέλνει ο φίλος σου Χουάν από τη Σαραγόσα».

Και αυτό το άτομο που συνάντησε στην τύχη του απαντά «Τι ωραία που το θυμήθηκε, το περιμένω εδώ και χρόνια. Όσο και αν μας διαβεβαιώνουν ότι αυτή η ιστορία είναι αληθινή, φαίνεται απίστευτο να συμβεί ότι αυτό ήταν ακριβώς το άτομο που έψαχνε, έτσι δεν είναι; Η πιθανότητα να συμβεί είναι λίγο μεγαλύτερη από όση έχει οποιοσδήποτε από μας που έχουμε αγοράσει νέον αριθμό, λαχείου να του πέσει το πρώτο νούμερο του χριστουγεννιάτικου λαχείου. Πράγματι, ο αριθμός των κατοίκων της Θιουδάδ Ρεάλ είναι 74.213, σύμφωνα με τη δημοτική απογραφή της 01.01.2009 και ο αριθμός λαχνών του χριστουγεννιάτικου λαχείου του 2009 ήταν 85.000.

11. Πάντα πέφτει κάπου αλλού

Το γεγονός ότι το χριστουγεννιάτικο λαχείο είναι δημοφιλές και πολλοί άνθρωποι συμμετέχουν στην κλήρωση, οδηγεί σε στρατηγικές με τις οποίες επιχειρείται να βελτιωθεί η πολύ μικρή πιθανότητα να πέσει κάποιο σημαντικό βραβείο. Ένας λάτρης των τυχερών παιχνιδιών που ζούσε σε μια μικρή πόλη ζητούσε πάντα λαχεία από συγγενή που ζούσε σε άλλη πιο μεγάλη πόλη λέγοντάς, που «αγόρασε μου εσύ επειδή πάντα πέφτει κάπου αλλού».

Βελτίωνε έτσι τις πιθανότητες να κερδίσει; Προφανώς όχι, όμως είναι μια διάχυτη αίσθησή που κάνει πολλούς ανθρώπους να αγοράζουν το λαχείο σε κάποια διάσημη περιοχή.

Αν πολλοί άνθρωποι αγοράζουν αριθμούς σε ένα συγκεκριμένο μέρος θα αυξηθεί η πιθανότητα να πέσει σε αυτό το μέρος, δεδομένου ότι θα πουλήσει περισσότερους αριθμούς, αλλά δε θα αυξηθεί η πιθανότητα να πέσει ένας συγκεκριμένος αριθμός που αγοράστηκε σε αυτό. Η αισιοδοξία μπορεί να είναι ακόμη μεγαλύτερη, όπως στην περίπτωση ενός ατόμου που για να μεγιστοποιήσει τις πιθανότητές του να κερδίσει, έλεγε «Τον αριθμό που παίζω τα Χριστούγεννα τον αγοράζω όταν αρχίζουν να πουλάνε τα λαχεία, έτσι σίγουρα δεν έχουν πουλήσει ακόμα το πρώτο νούμερο». Εδώ μπορούμε να θυμηθούμε την περίπτωση κατά την οποία έβγαινε μια μπάλα από μία τσάντα με πολλές μπάλες και μία μόνο ήταν νικήτρια, δεν είχε σημασία η σειρά με την οποία θα έβγαινε. Ας σκεφτούμε ακόμα κι αν είναι αληθινός ο συλλογισμός με τον οποίο ανοίγαμε την παράγραφο (είναι το εύκολο να πέσει κάπου αλλού επειδή παίζονται πολλοί περισσότεροι αριθμοί) εάν υπάρχει μια διαδικασία με την οποία είναι το πιο εύκολο να μας πέσει το χριστουγεννιάτικο λαχείο. Προφανώς δεν υπάρχει.

12. Τα lotto

Οι λαχειοφόρες αγορές με το όνομα lotto, στις οποίες πρέπει να επιλεγούν m αριθμοί μεταξύ 1 και N, είναι πολύ δημοφιλείς σε πολλές χώρες. Στην Ισπανία υπάρχουν τρεις, η Loteria Primitiva και το λόττο 6/49 στην Καταλονία (στις οποίες, πρέπει να επιλεγούν 6

αριθμοί μεταξύ 1 και 49) και το 7/39 της ΟΝΟΒ. Στο Ηνωμένο Βασίλειο υπάρχει επίσης το 6/49. Στην Ελβετία, το 6/45, στη Νέα Ζηλανδία, το 6/40 και στη Σουηδία υπάρχουν δύο δυνατότητες, 7/35 ή 6/48. Σε όλες τις περιπτώσεις, με μικρές παραλλαγές, υπάρχει βραβείο αν βρεθούν όλοι ή διάφοροι από τους αριθμούς που κληρώνονται. Στην περίπτωση του ισπανικού λόττο πρέπει να βρεθούν 6, 5+ ένας άλλος αριθμός που κληρώνεται ξεχωριστά (ο συμπληρωματικός) 5, 4 ή 3. Θα εστιάσουμε την ανάλυσή στο τελευταίο, το 1^ο βραβείο είναι 6 αριθμοί που κληρώνονται μεταξύ 1 και 49. Το πλήθος των διαφορετικών τρόπων με τους οποίους μπορούν να εξαχθούν 6 αριθμοί από 49 είναι

$$\binom{49}{6} = \frac{49!}{6! 43!} = \frac{44 \cdot 45 \cdot 46 \cdot 47 \cdot 48 \cdot 49}{6!} = 13.983.816, \text{ σχεδόν } 14 \text{ εκατομμύρια.}$$

Έτσι η πιθανότητα να πετύχουμε τους 6 κάνοντας ένα μόνο στοίχημά είναι

$$p = \frac{1}{13.983.816} \cong 0,000007\% \text{ ένας πραγματικά μικρός αριθμός. Σχετικά με τις πιθανότητες των άλλων βραβείων, σε κάθε περίπτωση είναι } p = \frac{\text{πιθανές περιπτώσεις}}{13.983.816} \text{ για αυτό πρέπει να υπολογιστούν οι δυνατές περιπτώσεις για κάθε αριθμό επιτυχιών.}$$

Μπορεί να μη βρεθεί κανένας από τους 6 αριθμούς και με κάθε «αποτυχία» μπορεί να ληφθεί οποιοσδήποτε από τους άλλους $49-6=43$ αριθμούς, επομένως, υπάρχουν $6 \cdot 43 = 258$ δυνατότητες να υπάρξουν 5 επιτυχίες, αλλά εδώ περιλαμβάνονται οι περιπτώσεις στις οποίες συμπίπτει ο συμπληρωματικός (που είναι 6) και εκείνες όπου δε συμπίπτει. Με παρόμοιους συλλογισμούς, οι ευνοϊκές, περιπτώσεις για 4 επιτυχίες είναι 13.545. Στην περίπτωση 3 επιτυχιών είναι 246820. Έτσι οι πιθανότητες επιτυχίας είναι:

$$\begin{aligned} 5+C \quad p &= \frac{6}{13.983.816} = \frac{1}{2.330.636}, & 5 \quad p &= \frac{252}{13.983.816} = \frac{1}{55.491} \\ 4 \quad p &= \frac{13.545}{13.983.816} = \frac{1}{1.032}, & 3 \quad p &= \frac{246.820}{13.983.816} = \frac{1}{57} \end{aligned}$$

Αθροίζοντας όλα παίρνουμε ότι η πιθανότητα βραβείου είναι 1,86%. Όπως παρατηρούμε είναι όλοι πολύ μικροί αριθμοί. Αν είμαστε σταθεροί είναι εύκολο, αφού το έτος έχει 52 εβδομάδες, περίπου μία φορά ανά έτος (αν παίζουμε μόνο μία ημέρα ή δύο αν στοιχηματίζουμε τις δύο ημέρες που έχει κλήρωση) να έχουμε ένα δελτίο με τρεις επιτυχίες. Οι υπόλοιπες πιθανότητες είναι τόσο μικρές που δε μπορούμε να αναμένουμε βραβεία με μεγαλύτερη συχνότητα και ακόμη και για τις 4 επιτυχίες πρέπει να περιμένουμε χρόνια.

13. Το δέλεαρ των λόττο και το «στοίχημα του Pascal»

Ποιος είναι ο λόγος για τη δημοτικότητα των λόττο; Ότι με ασυνείδητο τρόπο, γίνεται ένα σκεπτικό παρόμοιο με το στοίχημα του Pascal. Η μαθηματική προσδοκία είναι ίδια όπως σε πολλές άλλες λαχειοφόρες αγορές, αφού εξαρτάται από το ποσοστό των εσόδων που αφιερώνεται στα βραβεία. Αυτό που συμβαίνει είναι όταν τα βραβεία διανέμονται με διαφορετικό τρόπο και μπορεί να φτάσουν να γίνουν πολύ μεγαλύτερα. Υπάρχει πολύ μικρή πιθανότητα να πέσει ένα μεγάλο βραβείο, αλλά αν αυτό συμβεί θα είμαστε πραγματικά πλούσιοι από τη μία μέρα στην άλλη. Η κοινωνική απόφαση είναι ότι αξίζει το ρίσκο και έτσι παίζουμε.

14. Πόσους αριθμούς πρέπει να επιλέξουμε

Γιατί στην Ισπανία είναι 6/49 και στην Ελβετία 6/45; Έχει να κάνει με τον αριθμό κάτοικων της χώρας στην οποία προορίζεται, για να εκτιμηθεί ο αριθμός των πιθανών παικτών και επομένως η δυνατότητα να υπάρχουν επιτυχίες σε μεγάλα βραβεία. Είναι

καλύτερο να δούμε μια πραγματική περίπτωση. Στην Καταλονία, με περίπου 7.000.000 κατοίκους, μπορεί να παίξει κανείς το λόττο 6/49, όμως αφού ο αριθμός των δυνατών στοιχημάτων είναι σχεδόν 14.000.000 πολύ συχνά δεν υπάρχουν νικητές με 6 ούτε με 5 συν το συμπληρωματικό. Έτσι, στις 10 διαδοχικές κληρώσεις από 19.09.2009 έως 21.10.2009, σε 8 από αυτές δεν υπήρξε κανείς νικητής με 6 ή 5+ συμπληρωματικό. Μόνο στην κλήρωση της 07.10.2009 υπήρξε ένας νικητής σε κάθε κατηγορία αυτός με 6 κέρδισε 2.453.000 €, και με 5+συμπληρωματικό, 50.643 €. Στην κλήρωση της 14.10.2009 δεν υπήρξε κανείς νικητής με 6 και μόνο ένας με 5+ συμπληρωματικό, ο οποίος κέρδισε 11.600 €. Ένα παιχνίδι που βασίζει | την έλξη του σε πολύ σημαντικά βραβεία χάνει έτσι μεγάλο μέρος από αυτή.

Αντίθετα, αυτή η λαχειοφόρος αγορά είναι κατάλληλη για όλη την Ισπανία, επειδή με την αύξησή του πληθυσμού στον οποίο απευθύνεται, αυξάνει ο μέσος αριθμός των παικτών. Από το 2004 αμφισβητεί τη δημοτικότητα των λόττο κάθε χώρας ένας άλλος τρόπος στοιχήματος ακόμα πιο πολύτοκος, το Euro millions, που διεξάγεται από κοινού σε πολλές ευρωπαϊκές χώρες. Πρέπει για επιλεγεί μια ομάδα 5 αριθμών από το 1 έως το 50 και δύο αστέρια αριθμημένα από το 1 ως το 9. Επομένως ο αριθμός των διαφόρων στοιχημάτων που μπορούν να γίνουν με συλλογισμό όπως οι προηγούμενοι, είναι:

$$C_{50}^5 \cdot C_9^2 = \frac{50 \cdot 49 \cdot 48 \cdot 47 \cdot 46}{5!} \frac{9 \cdot 8}{2!} = 2.228.760 = 76.275.360$$

Το δυνητικά στοιχήματα ξεπερνούν αυτά του λόττο περισσότερο από 5 φορές, επομένως η πιθανότητα της μέγιστης επιτυχίας είναι 5 φορές μικρότερη. Ωστόσο, τα δυνητικά βραβεία είναι σημαντικά.

15. Τα πλεονεκτήματα του να είναι κανείς «κανονικός». Μεγάλοι αριθμοί

Όταν ρίχνοντας το ζάρι ρωτήσουμε κάποιον ποια η πιθανότητα να έρθει 4, είναι πολύ πιθανό να μας απαντήσει πως είναι 1/6, αν και λέγοντάς το δεν έχει πολύ σαφή ιδέα για το τι σημαίνει η πιθανότητα, για αυτό είναι ενδιαφέρον να τον ρωτήσουμε τι εννοεί λέγοντας ότι η πιθανότητα είναι «ένα έκτο». Ακόμα και αν δεν ήταν σε θέση να μας απαντήσει στην 1^η ερώτηση και του πούμε ότι η πιθανότητα είναι 1/6, αυτή η 2^η ερώτησή εξακολουθεί να έχει νόημα. Οι απαντήσεις σε αυτή τη 2^η ερώτηση μπορεί να ποικίλλουν σε μεγάλο βαθμό ανάλογα με τις γνώσεις αυτού που απαντά και ακόμη ανάλογα με τις προτιμήσεις του. Έτσι, μπορούμε να συναντήσουμε συχνά αυτές τις 3 απαντήσεις:

1. Το ζάρι έχει 6 πλευρές και μπορεί να έρθει οποιαδήποτε από αυτές,
2. Αν ρίξουμε το ζάρι πολλές φορές, περίπου στο 1/6 των φορές θα έρθει 4,
3. Εάν ποντάρουμε σε ένα παιχνίδι για να κερδίσουμε ή να χάσουμε αν βγει ή όχι το 4, τα στοιχήματα πρέπει να είναι 5 προς 1 υπέρ του να έρθει τέσσερα.

Η 1^η από τις απαντήσεις μπορεί να ανταποκρίνεται στη διαίσθηση μας και δεν απαιτεί ειδικές μαθηματικές γνώσεις, Αντίθετα, η 3^η είναι εντελώς αποδεκτή από άτομα συνηθισμένα να στοιχηματίζουν σε διάφορα παιχνίδια και μπορεί να ανταποκρίνεται περισσότερο στις εμπειρίες τους ως παικτών, παρά στις γνώσεις τους για τις πιθανότητες. Η ποικιλία παιχνιδιών είναι τόσο μεγάλη και το ενδιαφέρον για αυτά τόσο διαδεδομένο που πολλοί άνθρωποι θα υπερασπιστούν με πραγματικό πάθος αυτή την απάντηση έως την πιο λογική. Αυτό που μπορούμε να διαπιστώσουμε χωρίς μεγάλη δυσκολία είναι ότι για μεγάλο αριθμό ατόμων με λίγες ή καθόλου γνώσεις πιθανοτήτων η 2^η από τις απαντήσεις είναι αυτή που έχει λιγότερους υπερασπιστές. Κι όμως είναι αυτή που μπορεί να αποδειχθεί μαθηματικά. Μερικά δευτερόλεπτα συλλογισμού επιτρέπουν να συμπεράνουμε ότι αυτή η δήλωση είναι πιο πολύπλοκη απ' ότι μπορεί να φαίνεται αρχικά και περικλείει αρκετά λεπτές μαθηματικές έννοιες, Πράγματι, καθένας μπορεί να υποστηρίξει ότι με απόλυτους όρους, αυτή η δήλωση δε θα επιβεβαιωθεί όσες φορές και αν ρίξουμε το ζάρι αν το κάνουμε έναν αριθμό φορές που δεν είναι πολλαπλάσιο του 6. Επομένως, τι είναι αυτό που δηλώνει

πραγματικά η 2^η απάντηση; Δηλώνει ότι αν ρίχνουμε επ' αόριστον ή τουλάχιστον για ένα μεγάλο αριθμό φορών ένα ζάρι, το ποσοστό των φορών που θα έρθει ένα τεσσάρι (ή οποιοδήποτε άλλο ένα από τα 6 πιθανά νούμερα) θα προσεγγίσει για όσο θέλουμε, το 1/6.

Αυτός ο τύπος αποτελεσμάτων ονομάζονται στις πιθανότητες **νόμοι των μεγάλων αριθμών** (τώρα είναι πιο σαφές από πού προέρχεται το όνομα των νόμων). Εάν παραμένουμε ακόμα σκεπτικοί σχετικά με την ακρίβεια αυτών των νόμων, θα δούμε μια άλλη κατάσταση που μπορεί να βοηθήσει να εξαλείψουμε τις αμφιβολίες μας. Η ρουλέτα είναι ένα από τα πιο γνωστά παιχνίδια. Είναι ίσως ένα από αυτά που παίζονται περισσότερο σε όλο τον κόσμο όλες τις ημέρες και τα βραβεία μπορεί να ανέλθουν σε σημαντικά ποσά.

Ουσιαστικά, η **ευρωπαϊκή ρουλέτα** είναι ένας κύλινδρος στο εσωτερικό του οποίου υπάρχει ένας περιστρεφόμενος δίσκος 37 θέσεων, εναλλάξ μαύρες κόκκινες αριθμημένες από το 1 ως 36 συν το 0, που είναι σε θέση διαφορετικού χρώματος (πράσινο) και πρέπει να διατηρεί μια ευαίσθητη και λεπτή ισορροπία μεταξύ όλων των θέσεων στις οποίες τοποθετούνται οι αριθμοί. Στόχος του παιχνιδιού είναι να προβλεφθεί σε ποιον αριθμό ή χρώμα του περιστρεφόμενου δίσκου θα πέσει η μικρή μπάλα που ρίχνει ο κρουπιέρης. Υπάρχουν παραλλαγές αυτού του μοντέλου ρουλέτας, όπως η **αμερικανική ρουλέτα, η οποία έχει επίσης το «διπλό μηδέν»** και περιλαμβάνει μερικές μικρές διαφοροποιήσεις ως προς τα στοιχήματα και τα βραβεία. Εδώ αναφερόμαστε στην ευρωπαϊκή ρουλέτα.

Πρώτα γίνονται τα στοιχήματα, έπειτα γυρίζει με δύναμη ο τροχός και στη συνέχεια η μπάλα ρίχνεται από τον κρουπιέρη στο εξωτερικό τμήμα του τροχού, όπου παραμένει κάνοντας γύρους. Όταν ο τροχός χάνει την επαρκή ταχύτητα, η μπάλα πέφτει στις υποδοχές, όπου παραμένει αναπηδώντας από το ένα νούμερο στο άλλο, έως ότου καθίσει τελικά σε μία από τις 37 υποδοχές. Η «μαγεία» της κίνησης του τροχού έχει επηρεάσει την ανθρωπότητα σχεδόν από την εμφάνισή του. Η φαινομενική ακινησία του κέντρου, μαζί με την αύξηση της ταχύτητας καθώς απομακρυνόμαστε από αυτό και η αβεβαιότητα για το σημείο που θα σταματήσει ήταν η αιτία για τα πολλά παιχνίδια που έχουν τον τροχό ως βάση.

Συμφώνα με όλες τις ενδείξεις η δημιουργία της ρουλέτας και των κανόνων του παιχνιδιού της πολύ παρόμοιων με αυτούς που γνωρίζουμε σήμερα, οφείλεται **στον Pascal που εφηύρε μία ρουλέτα με 36 αριθμούς**. Φαίνεται ότι η επιλογή 36 αριθμών ενισχύει ακόμα περισσότερο τη μαγεία, επειδή **το άθροισμα ανιών των πρώτων 36 αριθμών δίνει τον κατεξοχήν μαγικό αριθμό 666**. Η επιλογή του αριθμού 36 είναι ίσως επειδή έχει πολλούς διαιρέτες. Αν και υπάρχουν πολλοί τρόποι πονταρίσματος στη ρουλέτα, για να απλουστευθεί η παρουσίαση θα υποθέσουμε ότι ένας παίκτης παίζει μόνο μονά/ζυγά, κόκκινο/μαύρο ή μεγάλα/μικρά (μεγάλα είναι το ποντάρισμα στα νούμερα 19 έως 36 και μικρά, 1 έως 18). Για κάθε ευρώ που ποντάρει, ο παίκτης παίρνει αν κερδίσει, αυτό το ευρώ και άλλο ένα. Ποιά η πιθανότητα να κερδίσει αυτό το επιπλέον ευρώ και ποιά να χάσει το ευρώ που στοιχημάτισε; Είναι σημαντικό να τονιστεί ότι αριθμός 0 δεν έχει χρώμα ούτε θεωρείται για τους σκοπούς του παιχνιδιού μονός ζυγός ούτε μπαίνει στο ποντάρισμα μεγάλα/μικρά. Επομένως η πιθανότητα να κερδίσει κανείς σε οποιοδήποτε από αυτά τα πονταρίσματα είναι 18/37 και να χάσει 19/37. **Το καζίνο έχει μία πιθανότητα 1/37 υπέρ του (2,70%)**. Αφού το κέρδος είναι θετικό ενός ευρώ με πιθανότητα 18/37 και αρνητικό ενός ευρώ με πιθανότητα 19/37, 1 μαθηματική προσδοκία ή μέσο κέρδος κάθε παρτίδας είναι

$$(+1) \cdot \frac{18}{37} + (-1) \cdot \frac{19}{37} = \frac{(-1)}{37} = -0,027\text{€}$$

δηλαδή, σε κάθε παρτίδα χάνουμε κατά μέσο όρο 2/7 λεπτά του €. Η κατάσταση αυτή είναι ίδια σαν να στοιχηματίζαμε σε ένα μόνο αριθμό, αφού στην περίπτωση αυτή, η πιθανότητα να κερδίσουμε είναι 1/37, να χάσουμε, 36/37 και αν ποντάρουμε ένα ευρώ και κερδίζαμε, παίρνουμε 1+ 335 €. Το μέσο κέρδος είναι επίσης:

$$35 \cdot \frac{1}{37} + (-1) \cdot \frac{36}{37} = \frac{(-1)}{37} = -0,027$$

Ας υποθέσουμε ότι παίζουμε 100 παρτίδες, ποντάροντας σε κόκκινο ή μαύρο ένα ευρώ σε κάθε μία. Στο τέλος των παρτίδων θα έχουμε χάσει $100 \cdot 0,027 = 2,7$ €; Αν κάτι είναι σαφές είναι ότι δε θα έχουμε κερδίσει ή χάσει μη ακέραιους αριθμούς. Όμως θα είμαστε πιο κοντά στο να έχουμε χάσει αυτό το ποσό ή πιο κοντά στο να έχουμε κερδίσει, για παράδειγμα, 50 €; Αυτά τα ερωτήματα είναι του ίδιου τύπου όπως αν αναρωτηθούμε αν σε ένα μεγάλο αριθμό ριξιμάτων ενός ζαριού θα εμφανιστεί το 4 στο ένα έκτο των φορών. Μπορούμε να σκεφτούμε άλλες καταστάσεις στις οποίες οι ερωτήσεις είναι παρόμοιες. Όταν λέγεται ότι ένα φάρμακο έχει αποτελεσματικότητα 80% αυτό σημαίνει ότι αν δοθεί σε μεγάλο αριθμό ασθενών θα είναι χρήσιμο μόνο για το 80% από αυτούς;

Αν το φάρμακο έχει παρενέργειες στο 1% των περιπτώσεων, σημαίνει πως αυτές οι επιπτώσεις μπορούν να επηρεάσουν το 1% των ασθενών που έλαβαν θεραπεία με αυτό; Η απάντησή σε αυτά θέματα δίνεται από τους νόμους μεγάλων αριθμών. Υπάρχουν δύο τύποι ερωτήσεων που σχετίζονται με την ιδέα των πιθανοτήτων. Από τη μία πλευρά, μπορούμε να σκεφτούμε, θα αντανακλούν και με τι τρόπο οι πιθανότητες των γεγονότων στα αποτελέσματα που θα πάρουμε; Από την άλλη πλευρά, το αντίστροφο πρόβλημα, λαμβάνοντας υπόψη τα αποτελέσματα που παρατηρήθηκαν, μπορούμε να συμπεράνουμε τις σχετικές πιθανότητές τους; Αυτά τα δύο ερωτήματα είναι διαφορετικής φύσης, αφού το σκεπτικό που ακολουθείται είναι αντίστροφο. Στην 1^η περίπτωση, από τις πιθανότητες που έχουν αποδοθεί, θέλουμε να συμπεράνουμε τα αποτελέσματα που θα πάρουμε, ενώ στη 2^η, από τις παρατηρήσεις θέλουμε να συνάγουμε τις πιθανότητες που ελέγχουν το φαινόμενο. Αυτή η 2^η προσέγγιση είναι αυτή που ακολουθεί η στατιστική.

16. Πολύ πιο εύκολο

Σήμερα η προσομοίωση με υπολογιστή επιτρέπει να γίνει εύκολα η επανάληψη πειραμάτων για όσες φορές θέλουμε. Για να ρίξετε ένα κέρμα και να δείτε τον αριθμό των κορόνων ή των γραμμμάτων και τις μεγαλύτερες σειρές κορόνων ή γραμμμάτων, μπορείτε να πάτε στη διεύθυνση http://nlvm.usu.edu/en/nav/frames_asid_305_g_3_t_5.html?from=topic_t_5.html έναν εικονικό χειριστή του πολιτειακού πανεπιστημίου της Γιούτα, όπου υπάρχουν μοντέλα για άλλες καταστάσεις.

17. Χρυσό θεώρημα του Bernoulli

Αρχίζουμε να ρίχνουμε ένα ζάρι πολλές φορές. Αν το ρίξουμε 300 φορές, θα πάρουμε 50 τεσσάρια. Τι θα συμβεί αν το ρίξουμε 3.000 φορές; Ο αναγνώστης μπορεί να θεωρήσει ότι κανείς που έχει τα λογικά του δε θα καθίσει να ρίξει χιλιάδες φορές ένα ζάρι και θα σημειώσει τα αποτελέσματά του. Και αν του πούμε ότι κάποιος μαθηματικός το έκανε, μπορεί να χρησιμεύσει για να επιβεβαιώσει τη λαϊκή ιδέα ότι αυτοί συχνά έχουν λίγο «πειραγμένο το μυαλό». Όμως ο μόνος τρόπος για να ελεγχθούν τα αποτελέσματα που αναμένονται από τη θεωρία είναι κάνοντας πειράματα, γεγονός του τα θέτει πιο κοντά στις πειραματικές επιστήμες από ότι πιστεύεται συνήθως.

Αυτό έκανε ο Γάλλος φυσιοδίφης Ζορζ Λουί Λεκλέρκ, κόμης De Buffon, ο οποίος έριξε ένα κέρμα 4.040 φορές και έφερε 2.048 κορόνες δηλαδή, αναλογία $2048/4.040 = 0,5069$ ή σε ποσοστιαίες μονάδες, 50,69% κορόνες. Μπορούν να εμφανισθούν ειδικές περιπτώσεις στις οποίες δεν υπάρχουν πολλά ενδιαφέροντα πράγματα να κάνουμε και είναι καλύτερο να βρεθούν ασχολίες που κρατούν απασχολημένο το μυαλό, για να συμβάλουν στην συναισθηματική ισορροπία. Αυτό συνέβη στον Νοτιοαφρικάνο μαθηματικό John Kerrich, ο οποίος φυλακίστηκε κατά τη διάρκεια του Β' παγκοσμίου πολέμου και στη φυλακή «ψυχαγωγούταν» ρίχνοντας ένα νόμισμα 10.000 φορές φέρνοντας 5.067 κορόνες αναλογία $5067/10.000 = 0.5067$. Σε διάφορα στάδια της διαδικασίας το ποσοστό των κορόνων απομακρυνόταν από το αναμενόμενο 50%, αλλά με την αύξηση των ριξιμάτων πλησίαζε σταδιακά την τιμή αυτή. Στα πρώτα 10 ριξίματα έφερε μόνο 4 κορόνες (40%) και στα 10

επόμενα, 6 κορόνες, και έτσι σε 20 ριζίματα είχε ακριβώς 10 κορόνες/ Μετά από 100 ριζίματα, το ποσοστό ήταν μόνο 44,96%, με 200 ανέβηκε στο 50.2% μέχρι που τελικά έφτασε, στα 10.000 ριζίματα, στις 5.067 κορόνες, 50,67%, τιμή κοντινή στο αναμενόμενο 50% για τις κορόνες.

Ο Jacob Bernoulli (1654–1705) αφιέρωσε αρκετές δεκαετίες για να μελετήσει το πρόβλημα και κατάφερε να αποδείξει μαθηματικά ότι το ποσοστό των κορόνων που λαμβανόταν με το ρίξιμο ενός νομίσματος επ' αόριστο πλησίαζε, χωρίς αμφιβολία, στο 50%. Στην περίπτωση του ζαριού, το θεώρημα του δείχνει ότι η αναλογία των φορών που έρχεται 4 προσεγγίζει το 1/6. Ο Bernoulli ονόμασε το αποτέλεσμα του «χρυσό θεώρημα», αλλά οι τωρινές εκδόσεις του αποτελέσματος είναι γνωστές ως «**νόμοι των μεγάλων αριθμών**» αδύναμοι νόμοι και ισχυροί νόμοι. Το προσδιοριστικό μεγάλων αριθμών αναφέρεται στο γεγονός ότι τα συμπεράσματα είναι βέβαια όταν επαναλαμβάνουμε το πείραμα επ' αόριστόν.

Είναι προφανές ότι δε θα επαναλάβουμε διαδικασία για πάντα, ούτε ακόμα και με τη βοήθεια ενός υπολογιστή, όσο δυνατός κι αν είναι. Μπορούμε να το κάνουμε πάρα πολλές φορές, αλλά θα είναι πάντα ένας πεπερασμένος αριθμός. Επομένως, τι εννοεί το συμπέρασμα ακόμα καλύτερα, τι χρησιμεύει ένα συμπέρασμα που δε μπορούμε να επαληθεύσουμε ποτέ; Εδώ εμφανίζεται η σημαντική μαθηματική έννοια του ορίου, επειδή αυτό που βεβαιώνουν οι νόμοι των μεγάλων αριθμών είναι ότι το ποσοστό των φορών που θα φέρνουμε κορόνα πλησιάζει, καθώς αυξάνουμε τον αριθμό των ριζιμάτων του νομίσματος (ή του ζαριού στην άλλη περίπτωση), στην πιθανότητα που δεχόμαστε πως έχει το αποτέλεσμα, 1/2 στην περίπτωση να έρθει κορόνα 1/6 στην περίπτωση να έρθει το 4.

Ο νόμος των μεγάλων αριθμών είναι ένας κανόνας που κάθε άτομο γνωρίζει με κάποιο φυσικό ένστικτο και χωρίς προηγούμενη εκπαίδευση. Μπορούμε να πούμε ότι υπάρχει μέσα μας με τον γενετικό κώδικα. Για να βρει το «χρυσό θεώρημα» στην αρχική του μορφή, ο Bernoulli κατασκεύασε ένα κουτί στο οποίο είχε 5.000 όμοιες μπάλες, 3.000 άσπρες και 2.000 μαύρες. Βγάζουμε μια μπάλα, σημειώνουμε το χρώμα της και την βάζουμε πίσω στο κουτί (για να μην τροποποιηθεί η αρχική του σύνθεση). Μετά βγάζουμε μια άλλη μπάλα και επαναλαμβάνουμε τη διαδικασία πολλές φορές. Είναι σαφές ότι οι δυνατότητες τραβήγματος μίας άσπρης μπάλας κάθε φορά είναι 56,0%, Τα ερώτημα που έθεσε ο Bernoulli ήταν με ποια ακρίβεια θα είναι το 60% και με ποια πιθανότητα θα υπάρχει τέτοια ακρίβεια; Με μια πρώτη ματιά μοιάζει με ένα λογοπαίγνιο δύσκολο να κατανοηθεί, αλλά ας συνεχίσουμε για να δούμε το βάθος του προβλήματος που έθεσε ο Bernoulli.

Αν βγάλουμε 200 μπάλες από το κουτί με άσπρες και μαύρες μπάλες, μπορούμε να έχουμε 120 άσπρες μπάλες (60%) ή 100 (50%) ή 125 (62%). Όμως τι πιθανότητες έχουμε ότι το ποσοστό των άσπρων μπαλών θα είναι μεταξύ 55% και 65%; Αν θέλουμε να γίνουμε πιο ακριβείς θα θέσουμε τις πιθανότητες να πάρουμε ένα ποσοστό άσπρων μπαλών κοντά στο 60% για παράδειγμα μεταξύ 59% και 61%. Μια τελευταία πρόσθετη ερώτηση, όμως όχι λιγότερο σημαντική, θα αυξηθούν οι πιθανότητες μας αν αντί να βγάλουμε 200 μπάλες βγάλουμε ένα εκατομμύριο; Αυτό το είδος ερωτημάτων, που αποτέλεσαν αντικείμενο εργασίας του Bernoulli, περιέχουν δύο είδη σφαλμάτων ή αβεβαιοτήτων που πρέπει να ορίσουμε επακριβώς. Από τη μία πλευρά, η απόκλιση του πραγματικού ποσοστού των μπαλών που είμαστε πρόθυμοι να δεχτούμε. Για παράδειγμα, ότι το ποσοστό που αποκτήθηκε είναι μεταξύ 50% και 61%. Από την άλλη πλευρά, δε θα μπορέσουμε ποτέ να διασφαλίσουμε με απόλυτη ακρίβεια ότι το ποσοστό μας θα είναι μεταξύ αυτών των ορίων, όμως αν επιθυμούσαμε να είναι σε πολλές από τις επαναλήψεις του πειράματος που κάνουμε, δηλαδή αυτό να συμβαίνει με ακρίβεια 95%, επομένως, στις 95% των φορών που επαναλαμβάνουμε το πείραμα της εξαγωγής των 200 μπαλών (ή του 1.000.000 μπαλών). Φαίνεται ότι δε μας ενδιαφέρει ο χρόνος που θα χρειαστούμε για τόσες επαναλήψεις.

Είναι αδύνατο να καθορίσουμε εκ των προτέρων και τα δύο σφάλματα. Αυτό του έδειξε ο Bernoulli είναι ότι αν επαναλάβουμε το πείραμα αρκετές φορές δηλαδή, αν

βγάλουμε αρκετές μπάλες, είναι δυνατό να πετύχουμε το ποσοστό των άσπρων μπαλών να πλησιάζει το 60% όσο επιθυμούμε. Ο όρος «να πλησιάζει ... τόσο» μπορεί να κυμαίνονται μεταξύ 59% και 61% , ή μεταξύ 59,9999% και 60,0001%, δηλαδή τόσο επιθυμεί κάποιος. Επιπλέον, το χρυσό θεώρημα παρέχει έναν τύπο για τον αριθμό των επαναλήψεων που απαιτούνταν για την επίτευξη αυτής της προσέγγισης. Το θεώρημα Bernoulli έχει δύο διαφορετικά μέρη. Ένα, ίσως το πιο σημαντικό, είναι ότι είναι δυνατό να επιτευχθεί η επιθυμητή ακρίβεια με οποιοδήποτε αριθμό (πεπερασμένο) δοκιμών. Το 2^ο προβλέπει τον αριθμό δοκιμών που απαιτούνται για την επίτευξη αυτής της ακρίβειας. Αυτό το 2^ο μέρος έχει μια πραγματική πρακτική εφαρμογή. Σε οποιαδήποτε μελέτη που γίνεται με ερωτήσεις μπορεί να οριστεί το επίπεδο σφάλματος που είμαστε έτοιμοι να δεχτούμε και να καθοριστεί ο αριθμός των ερωτηματολογίων που πρέπει να συμπληρωθούν για να εκτιμηθεί αυτό το σφάλμα με την επιθυμητή ακρίβεια. Έστω ότι γνωρίζουμε τη δημόσια υποστήριξη που ένα δημόσιο λειτούργημα έχει στην πόλη και θέλουμε να μάθουμε την ακρίβεια με την οποία, ρωτώντας έναν ορισμένο αριθμό πολιτών, το ποσοστό που θα πάρουμε θα αποκλίνει από το θεωρητικό κατά μια προκαθορισμένη απόκλιση. Χρησιμοποιώντας το χρυσό θεώρημα του Bernoulli μπορούμε να καθορίσουμε τον αριθμό ατόμων τα οποία πρέπει να ρωτήσουμε.

Τα αποτελέσματα του Bernoulli δεν αποδείχτηκαν τόσο πρακτικά όσο θα θέλαμε επειδή οι υπολογισμοί του περιλαμβάνουν μεγάλο αριθμό προσεγγίσεων που έχουν ως αποτέλεσμα ο αριθμός των ερωτηματολογίων να είναι υπερβολικά μεγάλος. Πρέπει να πούμε ότι μιλάμε για ένα αποτέλεσμα από τα τέλη του 17^{ου} αιώνα, που επιτεύχθηκε με τα τότε διαθέσιμα εργαλεία. Επηρεάζει επίσης η ακρίβεια που θέτουμε και ο Bernoulli εργάστηκε πάντα με αυτό που ονόμαζε «ηθική βεβαιότητα», που σήμαινε ακρίβεια 99,9%. Οι νόμοι των μεγάλων αριθμών στις σύγχρονες εκδοχές τους έχουν βελτιώσει την εκτίμηση σχετικά με τον αριθμό των δοκιμών που απαιτούνται. Ωστόσο, ακόμη και με τα σημερινά αποτελέσματα των πιθανοτήτων, αν απαιτούμε υπερβολική ακρίβεια, 99,99999%, για παράδειγμα και μικρή απόκλιση σε σχέση με το αναμενόμενο ποσοστό, θα πρέπει να ρωτήσουμε περισσότερους ανθρώπους από όσους ζουν στην πόλη. Σε κάθε περίπτωση, είναι προφανές ότι δεν πρέπει να εργαζόμαστε ποτέ ούτε με υπερβολικό σφάλμα ούτε με υπερβολική ακρίβεια.

Ο Jacob Bernoulli ήταν 50 ετών όταν πέθανε χωρίς να προφτάσει να τελειώσει το χειρόγραφο που περιλάμβανε το θεώρημά του. Οι εκδότες προσπάθησαν να βάλουν τον αδελφό του Johan να το ολοκληρώσει. Και όταν αυτός και ο ανιψιός του Nikolas αρνήθηκαν, το έργο παρέμεινε 8 χρόνια αδημοσίευτο και εκδόθηκε τελικά το 1713 ως Ars Conjectandi. Ακόμα και σήμερα εξακολουθεί να είναι μια ενδιαφέρουσα ανάγνωση. Ο Bernoulli έδειξε πώς με τον υπολογισμό μπορούσε να κάποιος εμβαθύνει στην κατανόηση των πιθανοτήτων.

Το ενδιαφέρον, το βάθος και η ομορφιά του προβλήματος που τέθηκε κι επιλύθηκε από τον Bernoulli είναι τέτοιου μεγέθους ώστε οποιαδήποτε σειρά πειραμάτων που επαναλαμβάνονται υπό τις ίδιες συνθήκες και με ανεξάρτητο τρόπο και στα οποία μπορούν να υπάρξουν δύο δυνατά αποτελέσματα, όπως το κουτί με 5.000 μπάλες, ονομάζονται στις σύγχρονες πιθανότητες «δοκιμές Bernoulli».

18. Θεώρημα του Bernoulli

Έστω ότι ως αποτέλεσμα ενός πειράματος μπορούμε να αποκτήσουμε ένα συγκεκριμένο γεγονός που θα ονομάσουμε A, η πιθανότητα του οποίου να συμβεί είναι p. Επαναλαμβάνουμε διαδοχικά το πείραμα n φορές και σημειώνουμε σε πόσες από αυτές εμφανίζεται το αποτέλεσμα A. Αν το γεγονός A εμφανίστηκε m φορές, το πηλίκο m/n αντιπροσωπεύει την αναλογία των φορών που εμφανίστηκε το A (σχετική συχνότητα εμφάνισης του A). Η διαφορά, σε απόλυτες τιμές, μεταξύ της πιθανότητας p και της σχετικής συχνότητας m/n μετρά το σφάλμα που θα κάναμε αν χρησιμοποιούσαμε τη σχετική συχνότητα έως προσέγγιση της πραγματικής πιθανότητας. Ο Bernoulli έδειξε ότι η πιθανότητα αυτής της διαφοράς μπορεί να γίνει τόσο μικρή όσο θα θέλαμε

επαναλαμβάνοντάς το πείραμα αρκετές φορές, ότι δηλαδή η πιθανότητα αυτής της διαφοράς τείνει στο μηδέν καθώς το n αυξάνεται. Με μαθηματικούς όρους αυτό εκφράζεται λέγοντας ότι αν είναι θετικός αριθμός οσοδήποτε μικρός, επαληθεύεται ότι $\lim_{n \rightarrow \infty} \left(\left| \frac{m}{n} - p \right| > \varepsilon \right) = 0$.

19. Δεν υπάρχει κακό που διαρκεί για πάντα, ή μήπως υπάρχει;

Οι νόμοι των μεγάλων αριθμών βεβαιώνουν ότι, αν παρατηρήσουμε την εμφάνιση ενός γεγονότος που έχει πιθανότητα p , η συχνότητα πλησιάζει στο p καθώς αυξάνουμε τις παρατηρήσεις. Αυτό σημαίνει ότι αν έχουμε μία σειρά παρατηρήσεων στις οποίες η συχνότητα είναι πολύ κάτω από την προβλεπόμενη τιμή, θα πρέπει να εμφανιστεί, κάποια στιγμή, μία άλλη σειρά αποτελεσμάτων που θα μας πλησιάσει στην αναμενόμενη συχνότητα (το ίδιο αν μία σειρά είναι πολύ πάνω από τη συχνότητα), ή αν συμμετέχουμε σε ένα παιχνίδι στο οποίο η πιθανότητα νίκης είναι κοντά στο $1/2$ και χάνουμε μεγάλο αριθμό παρτίδων, το ιδανικό είναι να συνεχίσουμε να παίζουμε, επειδή πρέπει να εμφανιστεί μία σειρά, που μας ευνοεί έτσι ώστε ο αριθμός των κερδισμένων παρτίδων να συμπίπτει με τις πιθανότητές μας να κερδίσουμε; Δηλαδή, δεν υπάρχει κακό που διαρκεί για πάντα. Και κάποια στιγμή θα πρέπει να αρχίσουμε να κερδίζουμε.

Ας σκεφτούμε ένα παιχνίδι στο καζίνο όπου οι μόνοι παίκτες είναι η μπάνκα και εσείς. Ένας παίκτης караδοκεί κοντά και παρακολουθεί πώς συσσωρεύετε χαμένες πατρίδες. Αν ο παρατηρητής έχει μαθηματικές γνώσεις φαίνεται λογικό ότι όταν δει πως έχουμε χάσει πολλές φορές να αποφασίσει, να συμμετάσχει στοιχηματίζοντας στην πλευρά μας. Όμως, δεν αποφασίζει να ποντάρει. Είναι λογική η συμπεριφορά του; Οι νόμοι των μεγάλων αριθμών δεν βεβαιώνουν ότι οι εμφανίσεις του γεγονότος εξισορροπούνται σε μικρές σειρές, αλλά ότι το κάνουν σε μια πάρα πολύ μεγάλη σειρά επαναλήψεων. Δεν υπάρχει τίποτα που μας επιτρέπει να βεβαιώσουμε ότι μετά από μία σειρά απωλειών θα ακολουθήσει μια άλλη με κέρδη ή το αντίστροφο. Οι νόμοι πληρούνται όταν η συχνότητα είναι πολύ κοντά στο όριο της ακολουθίας των συχνοτήτων, αλλά δεν υπάρχει λόγος να πληρούνται με οποιονδήποτε πεπερασμένο αριθμό παρατηρήσεων. Με άλλα λόγια, η τύχη δεν έχει μνήμη.

Μπορούμε να το διαπιστώσουμε εύκολα. Λέμε σε 40 άτομα να ρίξουν ένα νόμισμα 50 φορές και να σημειώσουν τα αποτελέσματα. Σε καθεμία από τις 40 σειρές οι κορόνες και τα γράμματα δεν κατανέμονται με ομοιόμορφο τρόπο ούτε υπάρχουν σειρές που εξισορροπούν τις συχνότητες. Όμως συγκεντρώνοντας τα 2.000 ριξίματα, η συχνότητα των κορόνων (ή των γραμμάτων) είναι κοντά στο $1/2$. Παρά τα παραπάνω είναι πιθανό να έχουμε δει κάποιον να παίζει σε μία αυτόματη μηχανή παιχνιδιών ο οποίος δε σταματά μέχρι να βγάλει ένα μεγάλο βραβείο. Αντιλέγει τα όσα είπαμε; Ο παίκτης ρισκάρει να χάσει πολλές παρτίδες επειδή περίμενε ότι η τύχη του θα αλλάξει. Είναι εύκολο ακόμα και αν έχει παρατηρήσει τη μηχανή βλέποντας ότι άλλοι παίκτες έκαναν πριν αποφασίσει να παίζει αυτός. Αντίθετα με τον «παρατηρητή» του καζίνο, αποφάσισε να συμμετάσχει, όταν σκέφτηκε ότι το βραβείο κόντευε «να πέσει». Σε αυτά τα παιχνίδια, αν και τα αποτελέσματα κάθε παρτίδας είναι τυχαία, οι μηχανές υποχρεούνται από το νόμο να επιστρέφουν σε βραβεία ένα μέρος αυτών που παίχτηκαν (πχ. το 70%) σε ένα μέγιστο αριθμό διαδοχικών παρτίδων, που συνήθως είναι μεγάλος (πχ. 40.000). Αυτό που συμβαίνει είναι προγραμματίζονται για να μη δίνουν τα μεγάλα βραβεία μέχρι το τέλος της σειράς, αλλά τα διανέμουν για να πληρούν τις νομικές απαιτήσεις. Κάθε οριζόμενο αριθμό παρτίδων, όχι πάντα τον ίδιο, οι μηχανές δίνουν το μεγάλο βραβείο. Εάν προσθέσουμε ότι τα αποτελέσματα ελέγχονται με πληροφορικά μέσα, δεν είναι εντελώς τυχαία, είναι κατανοητό ότι υπάρχουν «παρατηρητές» για το κυνήγι του βραβείου.

20. Μερικά λόγια για τη στατιστική

Έστω ότι επαναλαμβάνουμε ένα πείραμα για να παρατηρήσουμε την εμφάνιση ενός γεγονότος του οποίου την πιθανότητα αγνοούμε. Επαναλαμβάνουμε το πείραμά πολλές φορές και παρατηρούμε την εμφάνιση του γεγονότος. Η σχετική συχνότητα του γεγονότος είναι κατάλληλη προσέγγιση για την άγνωστη πιθανότητα; Ακόμα και αν πιστεύουμε ότι γνωρίζουμε εκ των προτέρων την πιθανότητα και επαναλάβουμε το πείραμά πολλές φορές, η σχετική συχνότητα εμφάνισης του γεγονότος μπορεί να μας κάνει να διαφοροποιήσουμε την εκ των προτέρων πιθανότητα. Αυτή η προσέγγιση, στην οποία επανασυλλέγουμε πληροφορίες μέσω των αποτελεσμάτων των πειραμάτων και από αυτές προσπαθούμε να εξάγουμε συμπεράσματα σχετικά με την πιθανότητα ενός γεγονότος ή άλλα χαρακτηριστικό του πειράματος είναι το αντικείμενο της στατιστικής. Οι πιθανότητες κάνουν παραγωγικό συλλογισμό. Η στατιστική χρησιμοποιεί έναν επαγωγικό συλλογισμό.

Οι πιθανότητες υποθέτουν ως αληθινό ένα πιθανολογικό μοντέλο και από αυτό αντλούν συμπεράσματα σχετικά με τα αποτελέσματα που θα παράγει ένα πείραμα. Η στατιστική μπορεί να θεωρήσει ένα μοντέλο για το πείραμα, αλλά αγνοώντας τα χαρακτηριστικά του εστιάζει στην παρατήρηση των αποτελεσμάτων και από αυτά, μέσω μιας επαγωγικής διαδικασίας, λαμβάνει τα χαρακτηριστικά του πειράματος, ως την πιθανότητα ότι θα συμβεί ένα γεγονός. Και στις δύο διαδικασίες, τα αποτελέσματα υπόκεινται σε κάποιο σφάλμα που τόσο η στατιστική όσο και η πιθανότητα είναι ικανές να ποσοτικοποιήσουν, καθεμιά με τους κανόνες της και σε αυτές παίζουν συνήθως σημαντικό ρόλο οι νόμοι των μεγάλων αριθμών ή άλλα αποτελέσματα όπως το κεντρικό οριακό θεώρημα.

Στο παράδειγμα του δοχείου του Bernoulli, αν γνωρίζουμε τη σύνθεσή του μπορούμε να καθορίσουμε την πιθανότητα (ή μια εκτίμησή της) ότι αν κάνουμε n τραβήγματα με επανατοποθέτηση, η αναλογία των άσπρων θα είναι τόσο κοντά στο $3/5$ όσο επιθυμούμε. Μπορούμε επίσης να καθορίσουμε το πλήθος τραβηγμάτων που πρέπει να κάνουμε ώστε η πιθανότητα ότι η διαφορά μεταξύ σχετικής συχνότητας και $3/5$ να είναι εντός ορισμένων ορίων (ωστόσο μικρό κι αν είναι), να είναι τόσο κοντά στο 1 όσο θέλουμε. Είναι η **πιθανολογική προσέγγιση**.

Αν γνωρίζουμε μόνο ότι το δοχείο υπάρχουν άσπρες και μαύρες μπάλες, όμως όχι πόσες ούτε σε ποιο ποσοστό, αρχίζουμε να βγάζουμε μπάλες σημειώνοντας το χρώμα τους και επιστρέφοντάς τις στο δοχείο. Από την παρατηρούμενη σχετική συχνότητα άσπρων και μαύρων μπορούμε να συμπεράνουμε ποια είναι η αναλογία μέσα στο δοχείο με κάποιο σφάλμα ή ανοχή. Πιο συγκεκριμένα, μπορεί να καθοριστεί η πιθανότητα ότι η αναλογία θα είναι εντός ορισμένων ορίων. Είναι η **στατιστική προσέγγιση**. Η τελευταία είναι η βάση των ερευνών που είναι τόσο κοινές στην κοινωνία. Για να μάθουμε τις προτιμήσεις υπέρ ενός υποψηφίου δημάρχου ρωτάμε πολλά άτομα (τραβάμε μπάλες), σημειώνουμε την απάντησή τους (χρώμα) και συνάγουμε με κάποιο επίπεδο εμπιστοσύνης (πιθανότητα) το ποσοστό της λαϊκής υποστήριξης (ποσοστό άσπρων μπαλών) που έχει αυτός ο υποψήφιος (δοχείο), με ένα κάποιο σφάλμα (απόκλιση).

21. Καμπύλη του Gauss και κανονικότητα

Έλεγε ο Sir Francis Galton (1822–1911) «Όσο πιο μεγάλος είναι ο αριθμός και μεγαλύτερη η φαινομενική αναρχία, τόσο πιο τέλειος είναι ο τομέας του. Είναι ο υπέρτατος νόμος του παραλογισμού. Κάθε φορά που λαμβάνεται ένα μεγάλο δείγμα χασοκών στοιχείων και διατάσσονται ανάλογα με το μέγεθός τους, εμφανίζεται μια απροσδόκητη και θαυμάσια μορφή κανονικότητας, λανθάνουσα κάτω από αυτό». Ο Bernoulli ενδιαφερόταν να αποδείξει ότι η πιθανότητα πως η απόκλιση μεταξύ σχετικής συχνότητας και πιθανότητας γίνεται τόσο μικρή όσο επιθυμούμε, προσεγγίζει το ένα. Αλλά ποτέ δεν ενδιαφέρθηκε για την ακριβή εκτίμηση αυτής της πιθανότητας, κάτι που επιτρέπει η κανονική κατανομή και το θεώρημα κεντρικού ορίου **κανονική κατανομή**

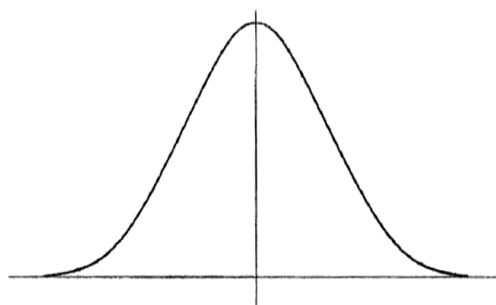
22. Η κανονική καμπύλη

Ο Abraham De Moivre (1667–1754) αποπειράθηκε να υπολογίσει το άθροισμα των όρων που εμφανίζονται στο διώνυμο του Newton $(a+b)^n$ όπου n είναι ένας μεγάλος φυσικός αριθμός. Συνέδεσε αυτό το θέμα με την πιθανότητα ενός γεγονότος, όπου n ο αριθμός επαναλήψεων του πειράματος που δημιουργεί το γεγονός. Άρχισε να μελετά την περίπτωση $(1+1)^n$, που σχετίζεται με πείραμα με μόνο δύο πιθανά αποτελέσματα που είναι εξίσου πιθανά. Ο De Moivre ήθελε να υπολογίσει την πιθανότητα σε n επαναλήψεις του πειράματος, το ζητούμενο γεγονός (που είναι εξίσου πιθανό με το αντίθετο) θα εμφανιζόταν έναν αριθμό φορές που θα ήταν περίπου $\frac{n}{2}$. Διαπίστωσε ότι μια καλή προσέγγιση αυτής της

πιθανότητας λαμβανόταν με τη βοήθεια του ολοκληρώματος $\frac{2}{\sqrt{2\pi n}} \int_0^1 e^{\frac{-x^2}{2n}} dx$ όπου

αναγνωρίζεται αυτό που σήμερα ονομάζεται **κανονική κατανομή**. Η Γκαουσιανή ή κανονική κατανομή είναι μια κατανομή πιθανοτήτων που καθορίζεται από τη συνάρτηση πυκνότητας (το συνεχές αντίστοιχο της συνάρτησης πιθανότητας των διαρκών μεταβλητών)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(\frac{x-\mu}{\sigma}\right)^2}$$
 όπου μ , σ^2 είναι η μέση τιμή και η διασπορά της. Το γράφημα της που ονομάζεται καμπύλη ή **καμπάνα του Gauss** είναι το παρακάτω.



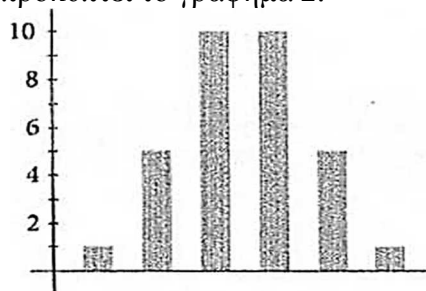
Ο Laplace βελτίωσε το αποτέλεσμα του De Moivre αποδεικνύοντας ότι έστω ένα πείραμα στο οποίο μπορεί να εμφανιστεί ένα γεγονός E με πιθανότητα p ή το συμπλήρωμά του (με πιθανότητα $1-p$). Όταν επαναλαμβάνουμε το πείραμα υπό τις ίδιες συνθήκες n φορές και υπάρχουν n εμφανίσεις του γεγονότος E , η πιθανότητα ότι η διαφορά μεταξύ m/n

και p είναι μεταξύ ορισμένων τιμών και συγκεκριμένα $-\frac{t\sqrt{2p(1-p)}}{\sqrt{n}}$ και $\frac{t\sqrt{2p(1-p)}}{\sqrt{n}}$ είναι

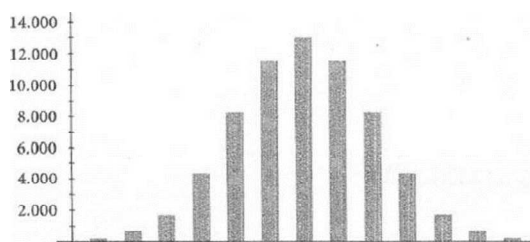
περίπου $\frac{2}{n} \int_0^1 e^{-x^2} dx + \frac{e^{-x^2}}{\sqrt{2\pi np(1-p)}}$.

Αυτό το αποτέλεσμα είναι μια πρώτη εκδοχή αυτού που είναι σήμερα γνωστό ως ΚΟΘ και είναι το κλειδί που εξηγεί τη δύναμη των προβλέψεων μεγάλου μέρους των στατιστικών εργαλείων. Πέρα από τους τύπους με τους οποίους εκφράζεται, αυτό που λέει το θεώρημα είναι ότι αν δοθεί ένα αρκετά μεγάλο δείγμα, τα χαρακτηριστικά αυτού του δείγματος θα αναπαράγουν αυτά του πληθυσμού από τον οποίο προέρχονται και ο τελευταίος θα μπορεί να μελετηθεί με μεθόδους πιθανοτήτων και στατικής. Αυτό που ήθελε να κάνει ο De Moivre ήταν να προσθέσει τους όρους των σειρών του τριγώνου Pascal που κατείχαν θέσεις πολύ χαμηλά στο τρίγωνο (n μεγάλο). Σήμερα πιστεύουμε ότι αυτός ο υπολογισμός γίνεται,

γρήγορα και με ακρίβεια, με μια αριθμομηχανή ή έναν όχι πολύ ισχυρό υπολογιστή. Όμως βλέπουμε ότι στη γραμμή 200 υπάρχουν όροι με 59 ψηφία (οι αριθμομηχανές δε χρησιμεύουν) και κατά τον 18^ο αιώνα τα εργαλεία υπολογισμού δεν ήταν τα σημερινά. Το πρόβλημα που εξετάστηκε από τον De Moivre ήταν, όπως ο ίδιος αναγνώριζε, πολύ δύσκολο. Οι περιορισμοί του υπολογισμού ήταν αυτό που είχε αναγκάσει τους Bernoulli και De Moivre να χρησιμοποιήσουν προσεγγίσεις οι οποίες βελτιώθηκαν με την ανάπτυξη νέων τεχνικών και εργαλείων υπολογισμού. Στο τρίγωνο του Pascal, αν πάρουμε τους αριθμούς μιας σειράς, όπως της 5^η και αν αναπαραστήσουμε σε ένα γράφημα κάθε αριθμό με ύψος ανάλογο με την αξία του προκύπτει το γράφημα 1. Αν κάνουμε το ίδιο στη 16^η γραμμή προκύπτει το γράφημα 2.



Γράφημα 1



Γράφημα 2

Κατεβαίνοντας στο τρίγωνο Pascal, οι επόμενες εικόνες πλησιάζουν ολοένα και περισσότερο το σχήμα της κανονικής καμπύλης, δηλαδή το άθροισμα αυτών των συντελεστών μπορεί να προσεγγιστεί από την κοινή κατανομή και αυτό ήταν επίτευγμα του De Moivre.

23. Ο νόμος των σφαλμάτων

Σε όλη την ιστορία, η παρατήρηση του ουράνιου θόλου, των πλανητών, των άστρων ή των κομητών είναι ένα από τα επιστημονικά θέματα που έχει εξάψει περισσότερο την προσοχή, επειδή η κατανόηση και η γνώση του ηλιακού συστήματος (και άλλων συστημάτων πέρα από το δικό μας) θεωρούνται απαραίτητες για να καταλάβουμε τον πλανήτη και τη ζωή μας. Τον 18^ο αιώνα, οι επιστήμονες που εργάζονταν στη φυσική, τα μαθηματικά και την ουράνια μηχανική βρίσκονταν συχνά αντιμέτωποι με το πρόβλημα των αποκλίσεων στις παρατηρήσεις. Φαίνεται ότι ήδη από τον 16^ο αιώνα διαδόθηκε η πρακτική της λήψης αρκετών μετρήσεων μιας άγνωστης ποσότητας στις ίδιες συνθήκες (ή τις πιο παρόμοιες που ήταν δυνατόν), με σκοπό να ληφθεί μια αξιόπιστη τιμή.

Έτσι, υπήρχαν διάφορες μετρήσεις και έπρεπε να αποφασιστεί σχετικά με την πραγματική τιμή αυτού του μεγέθους. Μια δυνατότητα ήταν να ληφθεί ο μέσος αριθμητικός όρος ως η «καλύτερη» τιμή, αλλά μαζί με αυτή την προσέγγιση, τον 18^ο αιώνα υπήρχε μια άλλη σχολή σκέψης που όριζε ως πραγματική τιμή αυτή μιας παρατήρησης που πραγματοποιήθηκε με «ιδιαιτέρη προσοχή». Η αμφισβήτηση συνεχίστηκε έως ότου εμφανίστηκε η κανονική κατανομή ως ένα κατάλληλο μοντέλο για τα σφάλματα. Η κανονική κατανομή εμφανίστηκε με μεγάλη δύναμη όταν ο Carl Friedrich Gauss (1777–1855) τη θεώρησε ως την κατάλληλη κατανομή για τα σφάλματα σε σχέση με τη μέθοδο των ελαχίστων τετραγώνων, που αναπτύχθηκε ως μέθοδος για την προσαρμογή μετρήσεων στην αστρονομία. Η μέθοδος βασίζεται στο εξής, μια παράμετρος για την οποία μπορούμε να λάβουμε μετρήσεις (επιρρεπείς σε σφάλματα) που επαναλαμβάνονται, εξαρτάται από μια σειρά μεγεθών που δεν είναι επιρρεπή σε σφάλματα και αναζητούμε την καλύτερη συνάρτηση που αντιπροσωπεύει την παράμετρο, με την έννοια ότι το άθροισμα των τετραγώνων των αποκλίσεων μεταξύ αυτής της συνάρτησης και των παρατηρήσεών μας για την παράμετρο να είναι ελάχιστο (τα τετράγωνα λαμβάνονται για να δοθεί η ίδια τιμή σε μία

αρνητική ή θετική απόκλιση). Υποθέτοντας ότι τα σφάλματα ακολουθούσαν το νόμο της κανονικής κατανομής, ο Gauss απέδειξε ότι η συνάρτηση που ελαχιστοποιούσε τα σφάλματα εφαρμόζοντας τη μέθοδο των ελάχιστων τετραγώνων είναι η συνάρτηση που κάνει πιο πιστευτές ή αληθοφανείς, με όρους των πιθανοτήτων, τις παρατηρήσεις δηλαδή, που αποδίδει τη μέγιστη πιθανότητα στις παρατηρήσεις. Η κανονική κατανομή αναγνωρίστηκε τελικά ως σημαντική κατανομή στη μελέτη σφαλμάτων και η μέθοδος των ελάχιστων τετραγώνων υιοθετήθηκε στην αστρονομία και τη γεωδαισία. Επιπλέον από τις αστρονομικές μετρήσεις εξακριβώθηκε η συμφωνία μεταξύ των παρατάσεων και νομού της κανονικής κατανομής.

24. Η υπόθεση των στοιχειωδών σφαλμάτων

Μια οδός για να υποστηρίξει το νόμο κανονικής κατανομής ως νόμο των σφαλμάτων προήλθε από τη θεωρία των στοιχειωδών σφαλμάτων. Συνίσταται στην υπόθεση ότι σφάλμα που γίνεται κάνοντας μια παρατήρηση ή μέτρηση μιας παραμέτρου είναι απλά το άθροισμα ενός μεγάλου αριθμού σφαλμάτων, που οφείλονται οι διάφορες ανεξάρτητες πηγές, καθένα από τα οποία είναι μικρό σε σύγκριση με το άθροισμα όλων. Με την αρχή αυτή αποδείχθηκε ότι η κατανομή του συνολικού σφάλματος ακολουθεί το νόμο κανονικής κατανομής.

Αυτό είναι το γνωστό ως ΚΟΘ. Η αρχή των στοιχειωδών σφαλμάτων έγινε ευρέως αποδεκτή στη διάρκεια του 19^{ου} αιώνα. Πολλοί επιστήμονες αφιέρωσαν τις προσπάθειές τους να αποδείξουν ότι υπό την προϋπόθεση αυτή, ο νόμος πιθανοτήτων που διέπει το συνολικό σφάλμα είναι ο νόμος κανονικής κατανομής. Μεταξύ τους ξεχωρίζει ο Γερμανός Friedrich Wilhelm Bessel (1784–1846) ο οποίος χρησιμοποίησε τη μέθοδο ελάχιστων τετραγώνων σε πολλές εφαρμογές και παρείχε εμπειρικά στοιχεία για την υποστήριξη του νόμου κανονικής κατανομής ως νομού των σφαλμάτων, όπως είχε προτείνει ο Gauss.

Ο Bessel παρουσιάζει την εξής κατάσταση *«Ένα σφάλμα παρατήρησης είναι το άθροισμα ενός μεγάλου αριθμού στοιχειωδών λαθών, που οφείλονται σε διαφορετικές και ανεξάρτητες μεταξύ τους αιτίες. Κανένα στοιχειώδες σφάλμα δεν υπερβαίνει σημαντικά τα άλλα και τα θετικά και αρνητικά σφάλματα του ίδιου μεγέθους εμφανίζονται με την ίδια συχνότητα. Επιπλέον, οι νόμοι που διέπουν τα στοιχειώδη σφάλματα δεν είναι απαραίτητα οι ίδιοι»*. Υπάρχουν δύο σημαντικές πτυχές στην υπόθεση του Bessel κανένα από τα σφάλματα δεν έχει ιδιαίτερη σημασία και τα θετικά και αρνητικά σφάλματα διανέμονται με συμμετρικό τρόπο. Όταν ένας δάσκαλος πρέπει να βαθμολογήσει τις εργασίες των μαθητών τους βάζει έναν αριθμητικό βαθμό μεταξύ 0 και 10. Όμως, είναι πραγματικά σε θέση να εξηγήσει τη διαφορά μεταξύ ενός βαθμού 8,8 και ενός άλλου 8,9; Θα ρισκάρουμε να πούμε πως όχι. Όταν βαθμολογεί, υπόκειται σε συνθήκες που θα μπορούσαν να προκαλέσουν σφάλματα.

Η αντίληψη δεν είναι ίδια όταν διαβάζει την 5^η εργασία ή όταν διαβάζει την 45^η, η κούραση συσσωρεύεται, ο τρόπος έκφρασης των μαθητών ποικίλλει, η περίοδος για να παρουσιάσει τις βαθμολογίες μπορεί να κάνει το χρόνο που αφιερώνει στον έλεγχο κάθε εξέτασης να ποικίλλει πολύ. Υπάρχουν διάφορες αιτίες που μπορούν να επηρεάσουν τη βαθμολόγηση, τόσο προς τα πάνω όσο και προς τα κάτω, διάφορες πηγές στοιχειακού σφάλματος. Θα μπορούσαμε να υποθέσουμε ότι αριθμητική βαθμολογία υπόκειται σε κάποια μεταβλητότητα που οφείλεται στη συσσώρευση των αιτιών, χωρίς αυτό να σημαίνει ότι είναι άδικη. Ίσως αυτή είναι η βάση του μύθου που λέει ότι οι δάσκαλοι βαθμολογούν ακολουθώντας το νόμο κανονικής κατανομής ή προσαρμόζοντας τις βαθμολογίες τους σε αυτόν. Είναι ένα παράδειγμα για να φανεί πώς το πρόβλημα του σφάλματος είναι ένα καθημερινό και σύγχρονο πρόβλημα.

25. Κεντρικό οριακό θεώρημα (ΚΟΘ)

Το ΚΟΘ αναφέρει ότι αν μια μέτρηση είναι το αποτέλεσμα του αθροίσματος ενός μεγάλου αριθμού παραγόντων (επιρρεπών σε σφάλματα) οι οποίοι ενεργούν ανεξάρτητα και

κανένας από αυτούς δεν κυριαρχεί στο συνολικό άθροισμα, ο νόμος πιθανοτήτων του συνολικού αθροίσματος κατανέμεται σύμφωνα με το νόμο κανονικής κατανομής, ανεξάρτητα από το νόμο πιθανοτήτων που ελέγχει τους παράγοντες. Θα αναλύσουμε μια συγκεκριμένη κατάσταση για να καταλάβουμε τι υπάρχει πίσω από αυτή τη διαβεβαίωση. Αναφορικά με το νόμισμα που έχουμε ρίξει τόσες φορές, το ΚΟΘ λέει ότι η αναλογία του αριθμού κορόνων που θα βγουν θα πλησιάζει το 50% καθώς αυξάνονται οι ρίψεις, και η πιθανότητα ότι αυτή η διαφορά είναι μεταξύ δύο συγκεκριμένων ορίων μπορεί να καθοριστεί από την κανονική κατανομή.

Για παράδειγμα, εάν ρίξουμε το νόμισμα 100 φορές, η πιθανότητα ότι η αναλογία των κορόνων είναι μεταξύ 4.8% και 5.2% είναι 0,3108. Αν το ρίξουμε 400 φορές, θα είναι 0,5762, και αν το ρίξουμε 5.000 φορές, αυξάνεται στο 0,9958. Δηλαδή, είμαστε πολύ σίγουροι ότι το ποσοστό των κορόνων θα είναι μεταξύ αυτών των τιμών που αναφέρονται. Αυτές οι πιθανότητες καθορίστηκαν χρησιμοποιώντας την κανονική κατανομή. Μια άλλη κατάσταση εμφανίζεται όταν ένας κατασκευαστής ράβδων δημητριακών γράφει στις συσκευασίες ότι καθεμιά ζυγίζει 60 g. Μερικές φορές το πραγματικό βάρος θα υπερβαίνει το προβλεπόμενο και άλλες θα είναι μικρότερο, αλλά τα βάρη θα ακολουθούν την κανονική κατανομή και λογικά, το αναγραφόμενο βάρος στη συσκευασία θα είναι το μέσο βάρος και η κανονική κατανομή επιτρέπει να καθορίσουμε την πιθανότητα ότι το βάρος μιας ράβδου είναι μεταξύ ορισμένων ορίων, επαληθεύεται η κανονικότητα των πιθανοτήτων.

Οι δύο προηγούμενες καταστάσεις επιτρέπουν να ξεχωρίσουμε τις δύο απόψεις που οδήγησαν στην αναζήτηση μαθηματικών εξηγήσεων οι οποίες δημιούργησαν το ΚΟΘ, ο προσδιορισμός ότι η μέση τιμή των σφαλμάτων των παρατηρήσεων (ιστορικά σε αστρονομία, γεωδαισία) ήταν μεταξύ ορισμένων ορίων και η ανάπτυξη και αποδοχή της υπόθεσης των στοιχειωδών σφαλμάτων ως βασικής αρχής για το ότι η κατανομή σφαλμάτων σε μια διαδικασία μέτρησης ακολουθεί το νόμο της κανονικής κατανομής.

Μια μαθηματική διατύπωση μιας απλής εκδοχής του ΚΟΘ αναφέρει «Εάν $X_1, X_2, \dots$ είναι ανεξάρτητες τυχαίες μεταβλητές μετρήσεις με την ίδια κατανομή, με πεπερασμένη μέση κοινή τιμή μ και τα τετράγωνα των οποίων έχουν επίσης πεπερασμένη μέση τιμή, η πιθανότητα το άθροισμα $X_1 + X_2 + \dots + X_n$ δεόντως κανονικοποιημένο να είναι μεταξύ των τιμών a, b προσεγγίζεται όταν το n αυξάνει επ' αόριστο, από την κανονική καμπύλη». Με ακριβείς μαθηματικούς όρους, αυτό που επαληθεύεται είναι ότι

$$\lim_{n \rightarrow \infty} P_{n \rightarrow \infty} \left[a \leq \frac{(X_1 + X_2 + \dots + X_n) - n \cdot \mu}{\sigma \sqrt{n}} \leq b \right] = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx \quad \text{όπου} \quad \sigma^2 = E(X^2) - \mu^2.$$

Το ΚΟΘ μπορεί να αναδιατυπωθεί λέγοντας ότι επιβεβαιώνεται η κανονικότητα των πιθανοτήτων ή ότι οι μεταβλητές ικανοποιούν την κανονικότητα των πιθανοτήτων.

Το ΚΟΘ ή η κανονικότητα των πιθανοτήτων επαληθεύονται ανεξάρτητα από την κατανομή που ακολουθούν οι μεταβλητές. Κατά κάποιο τρόπο αυτό που λέει το θεώρημα είναι ότι δε μπορούμε να προβλέψουμε την ατομική συμπεριφορά μιας μεταβλητής ή ατόμου, αλλά μπορούμε να προβλέψουμε τη συμπεριφορά του μέσου όρου του πληθυσμού.

Το ΚΟΘ πρόσφερε εκείνη την εποχή και προσφέρει και σήμερα, μια πιο βαθιά γνώση της συμπεριφοράς των παρατηρήσεων και της σχέσης της με τον φυσικό κόσμο. Οι ίδιες ιδέες έχουν εφαρμοστεί για τη μελέτη χαρακτηριστικών (φυσικών και κοινωνικών) σε ανθρώπινους πληθυσμούς και προς έκπληξη πολλών, διαπιστώθηκε ότι σε αρκετές περιπτώσεις η συμπεριφορά των χαρακτηριστικών των ανθρωπίνων πληθυσμών ανταποκρίνεται σε αυτό που προβλέπεται από το ΚΟΘ.

26. Ο πίνακας του Galton

Ο πίνακας που επινοήθηκε από τον Galton είναι μια διάταξη που επιτρέπει να ελεγχθεί πειραματικά η κανονικότητα των πιθανοτήτων και το ΚΟΘ. Πρόκειται για

κατακόρυφη πλάκα με μια σειρά από καρφιά που ακολουθούν το σχήμα του τριγώνου του Pascal. Κάτω από μια χοάνη από την οποία αφήνουμε να πέσουν όμοιες μπάλες βρίσκεται ένα καρφί και κάτω από αυτό άλλα δύο, ένα δεξιά και ένα αριστερά. Κάτω από αυτά βρίσκονται με τη σειρά τους άλλα καρφιά τοποθετημένα αριστερά και δεξιά αυτών που βρίσκεται στην παραπάνω σειρά, και ούτω καθεξής. Όταν μια μπάλα πέφτει μέσα από τα διαδοχικά επίπεδα, χτυπά σε κάποιο καρφί, το οποίο την κατευθύνει προς τα δεξιά ή αριστερά ενώ συνεχίζει να πέφτει. Τέλος, οι μπάλες συλλέγονται σε θήκες που βρίσκονται κάτω από καθέναν από τους χώρους που είναι κάτω από την τελευταία σειρά καρφιών.

Οι μπάλες που στοιβάζονται σχηματίζουν την κανονική καμπύλη. Η ερμηνεία της διαδικασίας είναι η εξής, αρχικά μια μπάλα αντιπροσωπεύει το σφάλμα που κάνουμε κατά τη μέτρηση ενός χαρακτηριστικού. Αυτό το σφάλμα της μέτρησης (ή τυχαία μεταβλητή) μπορεί να πάρει μόνο δύο τιμές (αριστερά και δεξιά, υπόλειμμα ή πλεόνασμα), με ίσες πιθανότητες.

Η τιμή αντιστοιχεί στη μετατόπιση που προκαλεί το καρφί που βρίσκεται κάτω από τη χοάνη όταν η μπάλα χτυπά πάνω του. Όταν κινηθεί σε μια χαμηλότερη σειρά, αυτό που συμβαίνει είναι ότι παράγεται ένα νέο σφάλμα που συσσωρεύεται στο προηγούμενο. Αν θεωρήσουμε ότι η κίνηση κάθε μπάλας σε μια σειρά αντιπροσωπεύει τη συσσώρευση σφαλμάτων, το τελικό αποτέλεσμα, που συνίσταται στο να παρατηρήσουμε τη θέση όπου αποθηκεύεται η μπάλα, αναπαριστά το άθροισμα τόσων λαθών όσων και οι γραμμές που υπάρχουν στη διάταξη. Αν τοποθετήσουμε ένα τρίγωνο του Pascal πάνω από τη διάταξη αυτή, οι αριθμοί του τριγώνου δείχνουν τον αριθμό των δρόμων που οδηγούν σε κάθε καρφί.

27. John Graunt (1620–1674)

Ήταν ιδιοκτήτης επιχείρησης ψιλικών που του παρείχε αρκετά χρήματα για μπορεί να αφιερώνεται σε χόμπι τόσο μακρινά από την επιχείρησή του όπως η μελέτη του πληθυσμού του Λονδίνου. Οι αναφορές θνησιμότητας ήταν αρχαία που συνέλεγαν από τον 16^ο αιώνα τα βαφτίσια και τις κηδείες (με την ηλικία και την προφανή αιτία θανάτου) των διαφόρων ενοριών του Λονδίνου, που επέτρεπαν στις αρχές να παρακολουθούν τα ίχνη των επιδημιών που μάστιζαν την πόλη. Από το 1603 ενός από το χειρότερα χρόνια της, το αρχείο γίνονταν εβδομαδιαία. Ο Graunt μελέτησε αυτά τα στοιχεία, οργάνωσε τις πληροφορίες με τη μορφή που τώρα ονομάζονται πίνακες ζωής ή θνησιμότητας και ανέλυσε διάφορες κοινωνικές πτυχές του πληθυσμού όπως, αντίθετα από ότι συνέβαινε στην ύπαιθρο, στην πόλη ο αριθμός των κηδειών υπερέβαινε τον αριθμό των βαπτίσεων. Είχαν ξεκινήσει οι μελέτες της δημογραφίας. Ο Graunt δημοσίευσε το 1662 το Bills of mortality (ο πλήρης τίτλος είναι Natural and political observation mentioned in a following index and made upon the bills of mortality) που έγινε δεκτό μεγάλο ενδιαφέρον από την αγγλική κοινωνία και του χάρισε στην ένταξη στη Royal Society.

28. Οι πιθανότητες στην κοινωνία. Πίνακες ζωής

Οι πίνακες ζωής ή θνησιμότητας είναι ένα στατιστικό εργαλείο που επιτρέπει τη μελέτη της επίπτωσης της θνησιμότητας σε διάφορους πληθυσμούς σε μια χρονική περίοδο. Μία από τις πιο κοινές εφαρμογές τους είναι οι ασφάλειες. Οι πρώτοι πίνακες ζωής ή θνησιμότητας οφείλονται στον Graunt ο οποίος, μαζί με τον William Petty, επεξεργάστηκαν στο 2^ο μισό του 17^{ου} αιώνα τις πρώτες στατιστικές για τον πληθυσμό, γεγονός που τους καθιστά προδρόμους της σύγχρονης δημογραφίας. Μετά από τη μελέτη των στοιχείων της θνησιμότητας, ο Graunt πήρε μία ομάδα 100 ατόμων και έδειξε πώς αυτή η αρχική ομάδα μειωνόταν (λόγω θανάτου των μελών της) με την πάροδο των ετών. Παρατηρείται ότι η βρεφική θνησιμότητα κατά τα πρώτα 6 χρόνια ήταν στο Λονδίνο του 17^{ου} αιώνα 36% και μόνο το 1% ξεπερνούσε τα 76 έτη. Οι μόνες ουσιαστικές αλλαγές στους σημερινούς πίνακες ζωής ή θνησιμότητας σε σχέση με τον πίνακα του Graunt είναι ότι σε αυτούς χρησιμοποιείται η στιγμή της γέννησης αντί για τη στιγμή της σύλληψης, το μέγεθος της

ομάδας είναι διαφορετικό και περιλαμβάνουν περισσότερες πληροφορίες.

Κατά τη στιγμή	επιβιώνουν
Της σύλληψης	100
Στο τέλος του 6 ^{ου} έτους	64
Στο τέλος του 16 ^{ου} έτους	40
Στο τέλος του 26 ^{ου} έτους	25
Στο τέλος του 36 ^{ου} έτους	16
Στο τέλος του 46 ^{ου} έτους	10
Στο τέλος του 56 ^{ου} έτους	6
Στο τέλος του 66 ^{ου} έτους	3
Στο τέλος του 76 ^{ου} έτους	1
Στο τέλος του 86 ^{ου} έτους	0

Η ιδέα μπορεί να εφαρμοστεί σε διαφορετικές καταστάσεις ή πληθυσμούς, όχι μόνο των ανθρώπων. Βασικός σκοπός είναι η μελέτη της θνησιμότητας ή με θετικούς όρους, του μέσου χρόνου της υπολειπόμενης διάρκειας ζωής ή προσδόκιμου ζωής των ατόμων ενός πληθυσμού. Οι πίνακες ζωής ή θνησιμότητας για τους ανθρώπινους πληθυσμούς καταρτίζονται τόσο σε εθνικό όσο και σε άλλα κατώτερα επίπεδα και ομαδοποιούνται σύμφωνα με γεωγραφικά, εθνογραφικά ή διοικητικά κριτήρια.

Η μεθοδολογία καταρτισμού τους, περιέχεται στα πρωτόκολλα της **Human Mortality Database (HMD)** όμως ανάλογα με τον πληθυσμό που μελετάται, μπορεί να απαιτούνται προσαρμογές. Αν στην αρχή το βασικό χαρακτηριστικό ενός πίνακα ζωής ή θνησιμότητας ήταν το προσδόκιμο ζωής, σήμερα περιλαμβάνει διάφορες βιομετρικές συναρτήσεις που επιτρέπουν την ανάλυση διαφορετικών χαρακτηριστικών του πληθυσμού που σχετίζονται με τη διάρκεια της ζωής των ατόμων. Παρακάτω περιγράφονται οι βασικές συναρτήσεις που περιλαμβάνονται συνήθως. Όλα αυτά τα χαρακτηριστικά αναφέρονται σε καθεμία από τις ηλικίες που περιλαμβάνονται στους πίνακες.

Προσδόκιμο ζωής, $EV(x)$. Το προσδόκιμο ζωής δείχνει το μέσο αριθμό ετών, μετά από την τωρινή του ηλικία x , που απομένουν σε ένα άτομο να ζήσει αν τα χαρακτηριστικά ζωής του πληθυσμού παραμένουν σταθερά. Το EV είναι ένας μέσος όρος που βασίζεται στην εμπειρία μίας υποθετικής ομάδας ατόμων του ίδιου πληθυσμού. Η διαθέσιμη τεχνολογία επιτρέπει τα στοιχεία για τον υπολογισμό αυτού του μέσου όρου να ενημερώνονται κάθε χρόνο, από τα πραγματικά δεδομένα θνησιμότητας, αυξάνοντάς έτσι την αξιοπιστία των πληροφοριών.

Πιθανότητα θανάτου, $q(x)$. Παρά το όνομα τους, οι πίνακες ζωής δείχνουν συχνά τον αριθμό των αναμενόμενων θανάτων ανά 1.000 άτομα πληθυσμού, δηλαδή την πιθανότητα πολλαπλασιασμένη επί 1.000. Αυτός είναι ο λόγος που αυτή η τιμή προσδιορίζεται ως «κίνδυνος θανάτου» αντί για μία πιθανότητα.

Θεωρητικοί θάνατοι, $d(x)$. Είναι ο αριθμός των θεωρητικών θανάτων που αντιστοιχούν σε κάθε μία από τις ηλικίες του πίνακα.

Επιζώντες, $L(x)$. Δηλώνει τον αριθμό ατόμων του πληθυσμού που είναι ζωντανοί σε μια συγκεκριμένη ηλικία.

Μέσος όρος ετών που έζησαν το τελευταίο έτος ζωής αυτοί που πεθαίνουν με συμπληρωμένα ηλικία x , $m(x)$. Είναι ο μέσος χρόνος που έζησαν μετά τη συμπλήρωση της ηλικίας «εκείνα τα άτομα του πληθυσμού που πεθαίνουν με αυτή την ηλικία,

Σταθερός πληθυσμός στην ηλικία x , $PE(x)$. Ο συνολικός αριθμός των ετών που έζησαν τα άτομα του πληθυσμού με συμπληρωμένη ηλικία x .

29. Ορισμένοι τύποι

Υπάρχουν σχέσεις που χρησιμεύουν για να κατανοήσουμε πώς γίνονται οι υπολογισμοί. Υπάρχουν μερικές προφανείς, όπως αυτή που λέει ότι ο αριθμός επιζώντων στην ηλικία $x+1$ είναι ίσος με αυτόν των επιζώντων μέχρι την αμέσως προηγούμενη ηλικία μείον τους θεωρητικούς θανάτους στην ίδια ηλικία $L(x+1) = L(x) - d(x)$. Είναι προφανές ότι η πιθανότητα αναμενόμενου θανάτου σε ηλικία x είναι ίση με την αναλογία μεταξύ θεωρητικών θανάτων και επιζώντων σε αυτή την ηλικία $q(x) = d(x) / L(x)$.

Το προσδόκιμο ζωής αναπαριστά το μέσο αριθμό των ετών ζωής που μένουν σε ένα άτομο ηλικίας x που ανήκει στην αρχική ομάδα. Η τιμή του προκύπτει από το λόγο του συνολικού χρόνου (σε έτη) που μένει να ζήσουν τα άτομα της ομάδας από τότε που θα συμπληρώσουν την ηλικία x ετών μέχρι την πλήρη εξαφάνιση τους και τον αριθμό επιζώντων της ίδιας ομάδας στην ηλικία x . Δηλαδή $EV(x) = \sum_{y=x}^{\infty} L(y) / L(x)$. Δεδομένου ότι κάθε άτομο που επιβιώνει στην ηλικία x συμβάλλει με ένα έτος στον συνολικό αριθμό των ετών που αποτελούν το σταθερό πληθυσμό και κατά μέσο όρο καθένας από αυτούς που πεθαίνουν με την ηλικία x συμβάλλει με $m(x)$ έτη, ο σταθερός πληθυσμός υπολογίζεται με την έκφραση $PE(x) = L(x+1) + m(x) \cdot d(x)$.

30. Πίνακες ζωής στην Ισπανία

Ένα μέρος του πίνακα ζωής ή θνησιμότητας των ετών 1991 και 2008 του ισπανικού πληθυσμού, ανεξάρτητα από το φύλο που αναφέρεται σε μία ομάδα 100.000 ατόμων. Εμφανίζονται μόνο οι βιομετρικές εξισώσεις που αντιστοιχούν στα πρώτα έτη, στις δεκάδες και στα τελευταία έτη. Στους πίνακες, ο κίνδυνος θανάτου εκφράζει τον αριθμό θανάτων ανά 1.000 άτομα και οι επιζώντες δηλώνουν τα άτομα που φθάνουν στην αντίστοιχη ηλικία για κάθε 100.000 άτομα. Επιπλέον, ο μέσος όρος των ετών που έζησαν το τελευταίο έτος ζωής στην ηλικιακή ομάδα των 100 και πλέον ετών, αντιστοιχεί στο μέσο όρο των ετών που έζησαν με συμπληρωμένα 100 έτη. Θα συγκρίνουμε αυτά τα δεδομένα του έτους 2008 με τα αντίστοιχα του έτους 1991. Από τη σύγκριση των πινάκων 1991 και 2008 διαπιστώνεται ότι έχει σημειωθεί μεγάλη βελτίωση στο ποσοστό βρεφικής θνησιμότητας καθώς έχει μειωθεί περισσότερο από 50% (72 θάνατοι ανά 1.000 γεννήσεις το 1991 σε σύγκριση με 34,6 το 2008). Στις μικρές ηλικίες η κατάσταση είναι παρόμοια.

Αυτό έχει μεγάλη επίπτωση στον υπολογισμό του μέσου όρου που αντιπροσωπεύει το προσδόκιμο ζωής, που πέρασε από 77,08 έτη με τη γέννηση το 1991 σε 81,24 το 2008, μία αύξηση μεγαλύτερη από 4 έτη. Ωστόσο, αν παρατηρήσουμε το προσδόκιμο ζωής με τη συμπλήρωση 65 ετών, επίσημη ηλικία συνταξιοδότησης στην Ισπανία το 2010, φαίνεται ότι εκείνη τη στιγμή το προσδόκιμο ζωής αυξήθηκε από 17,59 χρόνια το 1991 σε σχεδόν 20 το 2008, δηλαδή μια αύξηση σχεδόν 35 ετών, αρκετά χαμηλότερα από την αύξηση του προσδόκιμου ζωής κατά τα γέννηση. Πιο λεπτομερείς αναλύσεις για κάθε ηλικία είναι απαραίτητες στις δημογραφικές μελέτες και αποτελούν βασικό εργαλείο που συμβάλλει στην οργάνωση των κοινωνιών, αφού επιτρέπουν να εκτιμηθούν από τις δαπάνες για συντάξεις έως το πιθανό κόστος της υγειονομικής περίθαλψης. Για τα μελέτη αυτών και άλλων ενδιαφερόντων χαρακτηριστικών καταρτίζονται εξειδικευμένοι πίνακες όπως πίνακες ζωής ή θνησιμότητας ανά φύλο ή για «πληθυσμούς χωρίς φυσική ανικανότητα», σύμφωνα με την επίσημη ορολογία. Επίσης, το προσδόκιμο ζωής έχει ενταχθεί σε δείκτες ευημερίας ή ανθρώπινης ανάπτυξης οι οποίοι μετρούν των κατάσταση των χωρών στο διεθνές πλαίσιο.

31. Ασφάλειες

Μία από τις πιο προσοδοφόρες επιχειρήσεις στην κοινωνία είναι οι ασφαλιστικές εταιρείες. Υπάρχουν πολλοί τύποι (ασφάλισης, ζωής, αυτοκινήτου, υγείας, οικείας). Πίσω από όλες αυτές υπάρχει μία σειρά από μαθηματικούς υπολογισμούς που λαμβάνει υπόψη

τόσο το δικό μας κίνδυνο (την πιθανότητα να χρειαστεί να μας πληρώσουν) όσο και το χρόνο κατά τον οποίο θα μπορούσαμε να απαιτήσουμε αποζημίωση για αυτό που έχει ασφαλιστεί. Από τις πιο δημοφιλείς είναι οι ασφάλειες ζωής στις οποίες ένα άτομο ασφαλίζει τη ζωή του, έτσι ώστε αν πεθάνει η οικογένειά του λαμβάνει ένα ορισμένο ποσό που ανακουφίζει για την απουσία του. Ο καθορισμός της δόσης που πρέπει να καταβληθεί και του ποσού που θα ληφθεί σε περίπτωση θανάτου βασίζονται σε μια σειρά μαθηματικών-στατιστικών υπολογισμών που επιχειρούν να προβλέψουν πόσο χρόνο ακόμα θα ζήσει ο ασφαλισμένος. Είναι προφανές ότι η διάρκεια ζωής ενός μεμονωμένου ατόμου δεν είναι προβλέψιμη αλλά είναι αλήθεια ότι σε μεγάλους πληθυσμούς παρατηρούνται μία σειρά από κανονικότητες που επιτρέπουν, χρησιμοποιώντας εργαλεία πιθανοτήτων και στατιστικής να εκτιμήσουμε το μέσο χρόνο ζωής των ατόμων μίας συγκεκριμένης ομάδας.

Η ύπαρξη ασφαλίσεων είναι παλαιότερη από ότι μπορούμε να φανταστούμε και υπάρχουν τεκμηριωμένες ασφαλίσεις για τον κίνδυνο, ναυαγίων και για τα εμπορεύματα που αποστέλλονται σε πλοία στην αρχαία Ελλάδα και Ρώμη. Στο Μεσαίωνα οι προσκυνητές που πήγαιναν στους Αγίους Τόπους μπορούσαν να υπογράψουν μια «ασφάλισή διάσωσης», έτσι ώστε κατέβαλαν ένα τέλος και αν κατά τη διάρκεια του ταξιδιού τους απήγαγαν και ζητούσαν λύτρα, τα πλήρωνε ο ασφαλιστής. Ωστόσο, δεν υπήρχε το αντίστοιχο της ασφάλειας της οποίας το κεντρικό στοιχείο να είναι ο θάνατος του ασφαλισμένου. Όχι μόνο δεν υπήρχε αλλά και θεωρούταν ιεροσυλία, αφού ζωή και θάνατος διέποντας από τα σχέδια του Θεού και δε μπορούσαν να γίνουν αντικείμενο μελέτης. Η ασφάλιση της ζωής (ή του θανάτου) πήγαινε ενάντια στα σχέδια του δημιουργού. Η Αναγέννηση χαλάρωσε τις ηθικές πεποιθήσεις και βοήθησε την άνθηση των επιχειρήσεων και συνεπώς των ασφαλειών ζωής.

Ωστόσο, οι χώρες των οποίων οι ηθικοί κανόνες ήταν αυστηρότεροι εξακολουθούσαν να μην τις δέχονται και αν δεν αρκούσε η πειθώ, υπήρχε η ρητή απαγόρευση. Η Ισπανία το έκανε το 1570 και η Ολλανδία το 1598. Έναν αιώνα αργότερα, ο υπολογισμός της πιθανότητας θανάτου και του προσδόκιμου ζωής είχαν καθιερωθεί για τα καλά και κανείς δεν τους αμφισβητούσε. Οι πιο διστακτικοί αρκέστηκαν να διαπιστώνουν ότι οι στατιστικές κανονικότητες που παρατηρούνταν στους ανθρώπινους πληθυσμούς ήταν επίσης, ένα σημάδι της θείας τάξης. Ο υπολογισμός του τέλους ή ασφαλιστρου μιας ασφάλειας ακολουθεί διαδικασία παρόμοια με ενός τυχερού παιχνιδιού, οι επιχειρήσεις καθορίζουν την πιθανότητα του συνόλου των δυνητικών κινδύνων που ασφαλίζουν, πολλαπλασιάζοντας την πιθανότητα του κινδύνου (που εκτιμούν από τα δεδομένα των τελευταίων ετών) επί το μέσο κόστος του και προσθέτοντας όλους τους δυνητικούς κινδύνους. Στο ποσό αυτό προσθέτουν τα έξοδα διαχείρισης και το κέρδος που θέλουν να αποκτήσουν. Με όρους παιχνιδιών οι παίκτες είναι οι ασφαλισμένοι που παίζουν ενάντια στη μπάνκα, που είναι η ασφαλιστική εταιρεία.

Ωστόσο, στην περίπτωση αυτή δεν έχουν κανένα συμφέρον να κερδίσουν το βραβείο, που σημαίνει να έχουν κάποιο ατύχημα και να πάρουν το ποσό ασφάλισης. Για παράδειγμα, αν ταξινομήσουμε τους κινδύνους μίας ασφάλειας αυτοκινήτου σε τρεις κατηγορίες (βλ. παρακάτω πίνακα) και οι πληροφορίες που έχει η εταιρεία είναι αυτές που φαίνονται στον πίνακα, π μαθηματική προσδοκία (EM) του κόστους των ατυχημάτων θα ήταν:
 $EM = 0,003 \cdot 7.000 + 0,05 \cdot 3.000 + 0,3 \cdot 500 = 21 + 150 + 150 = 321€$ κατά συνέπιά, το κόστος B του ασφαλιστήριου συμβολαίου είναι $P = EM + \text{έξοδα διαχείρισης} + \text{επιθυμητό κέρδος}$.

Τύπος ατυχήματος	Πιθανότητα	Κόστος επανόρθωσης
Γενική απώλεια	$3/1.000 = 0,003$	7.000 €
Σοβαρά	$5/100 = 0,05$	3.000 €
Μικρά	$3/10 = 0,3$	500 €

Η ασφάλιση μπορεί να θεωρηθεί ως ένα «παιχνίδι» δυσμενές για τον ασφαλισμένο

σχεδόν όλα τα χρόνια (υπό την προϋπόθεση ότι δεν παθαίνει ατύχημα), αλλά με αυτή καλύπτεται με λίγα χρήματα, από τον κίνδυνο να χρειαστεί να πληρώσει ένα υψηλό ποσό ή να αγοράσει ένα νέο αυτοκίνητο σε περίπτωση ατυχήματος, αγοράζοντας μία ασφάλιση, με την καταβολή ενός όχι πολύ υψηλού ποσού κάθε έτος (ή μόνο μία φορά στην περάτωση των ταξιδιωτικών ασφαλίσεων) ελπίζουμε να πάρουμε ως βραβείο ότι σε περίπτωση που μας συμβεί κάποιο ατύχημα, που είναι λίγο πιθανό αλλά όχι αδύνατο, να μπορούν να αντέξουμε οικονομικά. Γενικά, οι ασφάλειες μπορούν να θεωρηθούν ως μια λαχειοφόρος αγορά, ένα παιχνίδι στο οποίο μπορεί να αξίζει να στοιχηματίσουμε.

32. Ηλικία συνταξιοδότησης και συντάξεις

Οι σύγχρονες κοινωνίες με συστήματα κοινωνικής προστασίας έχουν ως ένα από τα μεγαλύτερα επιτεύγματά τους τη φροντίδα για τους ηλικιωμένους, στους οποίους χορηγούν μία σύνταξη μόλις πάψουν να εργάζονται, όταν φτάσουν σε ένα όριο ηλικίας. Κατ' επανάληψη, σε χώρες με αυτό το σύστημα κοινωνικής προστασίας (οι περισσότερες από τις ευρωπαϊκές) ανοίγει μία συζήτηση για το αν οι εργαζόμενοι πρέπει να παρατείνουν την εργασιακή τους ζωή, δηλαδή αν πρέπει να εργάζονται μέχρι ένα υψηλότερο όριο ηλικίας για να έχουν δικαίωμα στη σύνταξη. Αυτό το είδος συζητήσεων, ανεξάρτητα από τη σοβαρότητα και την αυστηρότητα με την οποία αντιμετωπίζονται, έχουν ως βασικό στοιχείο τη διάρκεια ζωής. Όσο περισσότερα χρόνια ζει ένα άτομο τόσο περισσότερα χρόνια θα λαμβάνει τη σύνταξή του, υποθέτοντας ότι δεν πεθαίνει πριν φθάσει στην ηλικία συνταξιοδότησης.

Όπως συμβαίνει με τις ασφάλειες, ο προσδιορισμός κατάλληλων πινάκων ζωής ή θνησιμότητας και ο υπολογισμός του προσδόκιμου ζωής είναι βασικά στοιχεία για την αντιμετώπιση των ζητημάτων αυτού του είδους. Τώρα, δεν απαιτούνται μόνο κατάλληλοι πίνακες, αλλά και μια σωστή ανάλυσή τους, δεδομένου ότι γενικά, δεν αρκεί να δούμε το στοιχείο για το προσδόκιμο ζωής κατά τη γέννηση, αλλά θα απαιτηθεί να αναλυθεί το προσδόκιμο ζωής στην ηλικία που μας ενδιαφέρει ή άλλες βιομετρικές εξώσεις.

33. Άλλες εφαρμογές

Υπάρχουν πολλές άλλες εφαρμογές των πινάκων ζωής ή θνησιμότητας, τόσο στο πλαίσιο των ανθρωπινων πληθυσμών όσο και αλλού. Για παράδειγμα, στη μηχανική ενός εξαρτήματος ή μηχανήματος που μπορεί να παρουσιάσει βλάβη και πρέπει να αντικατασταθεί από άλλο δεν είναι παρά ο υπολογισμός της ωφέλιμης ζωής αυτού του εξαρτήματος. Αναμφίβολα, η ύπαρξη κατάλληλων πινάκων ζωής ή θνησιμότητας είναι ζωτικής σημασίας για πολλές βιομηχανικές διαδικασίες. Αν και το παραπάνω παράδειγμα μπορεί να φαίνεται κάτι χωρίς μεγάλη επιρροή στην καθημερινή ρουτίνα μας, τίποτα δεν απέχει περισσότερο από την πραγματικότητα. Ας σκεφτούμε ότι αυτό το μηχάνημα που μπορεί να χαλάσει είναι η τηλεόραση ή το πλυντήριό, για να μην πούμε για τις λάμπες που χρησιμοποιούμε. Γιατί πράγμα μιλάμε λοιπόν; Ακριβώς για το χρόνο του περνά κατά μέσο όρο, μέχρι το πλυντήριο ή η τηλεόρασή να χρειάζονται επισκευή.

Μια στιγμή συλλογισμού μας κάνει να δούμε ότι οι κατασκευαστές γνωρίζουν τους πίνακες ζωής θνησιμότητας, των εξαρτημάτων και εμείς όχι. Τι σημαίνει αυτό; Δυστυχώς, ότι οι κατασκευαστές ορίζουν τις εγγυήσεις των προϊόντων σύμφωνα με τα προσδόκιμα ζωής που εμείς δε γνωρίζουμε. Είναι πιθανό μερικές συσκευές όπως οι τηλεοράσεις να παρουσιάσουν λιγότερες βλάβες, όμως άλλες όπως τα πλυντήρια που υπόκεινται σε έντονη διαδικασία φθοράς, μπορεί να υποστούν περισσότερες βλάβες και ποιός δεν είχε την αίσθηση ότι η βλάβη συνέβη ακριβώς τη στιγμή που έληξε η εγγύηση;

Η κατάσταση αυτή δεν οφείλεται στην τύχη, αλλά στη μελέτη που κάνουν οι εταιρείες για το προσδόκιμο ζωής των προϊόντων τους. Είναι προφανές ότι καμία εγγύηση δεν είναι μεγαλύτερη από την αναμενόμενη διάρκεια ζωής της συσκευής, αλλά ακριβώς το αντίθετο. Έτσι γίνεται καλύτερα κατανοητό τι αντιπροσωπεύει μια εγγύηση. Ευτυχώς, οι

πιθανότητες και η στατιστική είναι γνωστές για όλους όσους θέλουν να τις μελετήσουν, γεγονός που επιτρέπει στις αρχές, λαμβάνοντας υπόψη τις πληροφορίες που παρέχονται από εμπειρογνώμονες, να αποφύγουν καταχρηστικές πρακτικές και να ορίσουν μια ελάχιστη διάρκεια για την εγγύηση των προϊόντων, η οποία είναι υποχρεωτική για τους κατασκευαστές. Κάτι για το οποίο είναι επίσης σημαντικές οι ενώσεις καταναλωτών.

34. Πιθανότητες και στατιστική στην ιατρική πρακτική

Με τα αποτελέσματα μερικών κλινικών αναλύσεων, ο ιατρός συμβουλευεται μια σειρά στοιχείων τα οποία, σύμφωνα με διαδικασίες πιθανοτήτων και στατιστικής, του παρέχονται με κατανοητό τρόπο ώστε να λάβει μέτρα που θεωρεί κατάλληλα. Η τελική απόφαση είναι για τον ιατρό. Τα στοιχεία, η επεξεργασία τους και ο τρόπος με τον οποίο του παρέχονται είναι ένα εργαλείο υποστήριξης, για την απόφασή του. Όμως, μερικές από τις πληροφορίες μπορεί να έχουν αποφασιστική σημασία ώστε ο ιατρός να πάρει αποφάσεις ή να επιβάλει πρότυπα συμπεριφοράς. Για παράδειγμα, ξέρουμε ότι ορισμένες «συστάσεις», μαζί με τις αγωγές, δείχνουν ότι κάτι είναι «εκτός των κανονικών ορίων». Τι σημαίνει αυτό και πώς αποφασίζεται; Στην ιατρική ως κανονικά όρια ενός συγκεκριμένου χαρακτηριστικού εννοούνται αυτά μεταξύ των οποίων βρίσκεται η πλειοψηφία του πληθυσμού στον οποίο ανήκουμε και καθορίζονται με βάση το ΚΟΘ και την κανονική κατανομή.

Η διαδικασία για να γίνει αυτό είναι να προσδιοριστεί μεταξύ ποιών τιμών βρίσκεται η μέση τιμή με ορισμένη πιθανότητα, το πρόβλημα που ήδη μελέτησε ο Laplace. Οι τιμές μεταξύ των οποίων πρέπει να είναι ο μέσος όρος είναι αυτό που στη σημερινή στατιστική ονομάζεται διαστήματα εμπιστοσύνης. Μια παρόμοια περίπτωση είναι όταν ένας παιδίατρος λέει ότι το εκατοστημόριο στο οποίο βρίσκεται ένα παιδί από την άποψη βάρους και ύψους είναι 85% ή 95%. Τα διαστήματα εμπιστοσύνης για τις διαφορετικές ηλικίες που χρησιμοποιούν οι παιδίατροι αναφέρονται στα βάθη και ύψη μεταξύ των οποίων είναι ο πληθυσμός των παιδιών που αποτελούν τον πληθυσμό αναφοράς.

Ότι ένα παιδί είναι στο εκατοστημόριο 80 του ύψους και 75 του βάρους σημαίνει ότι το 80% των παιδιών ίδιας ηλικίας θα είναι το πολύ, ίδιου ύψους με αυτό και το ίδιο για το εκατοστημόριο 75% όσον αφορά το βάρος.

Είναι σημαντικό να έχουμε έναν καλό πληθυσμό αναφοράς αν αυτός που χρησιμοποιήθηκε για τον προσδιορισμό της κανονικότητας του ύψους και του βάρους ήταν πολύ διαφορετικός από αυτόν που αντιστοιχεί, τα αποτελέσματα δεν θα είχαν καμιά εγκυρότητα. Γι' αυτό υπάρχουν διαφορετικές ζώνες αναφοράς για αγόρια και κορίτσια, αφού το φύλο καθορίζει τον πληθυσμό της αναφοράς.

35. Πιθανότητες και DNA

Από τα μέσα της δεκαετίας του 1980 η χρήση των προφίλ δεσοξυριβονουκλεϊκού οξέος (DNA) χρησιμοποιείται σε δικαστικές υποθέσεις για την απόδειξη πατρότητας, οικογενειακής σχέσης μεταξύ διαφορετικών ατόμων, αθωότητας ή ενοχής ποινικών υπόπτων.

Η ιδέα είναι η σύγκριση DNA των παιδιών με αυτό των θεωρούμενων πατέρων στην περίπτωση των τεστ πατρότητας, ή του DNA που προέρχεται από δείγματα που συλλέχτηκαν στη σκηνή ενός εγκλήματος με αυτό που λαμβάνεται από τους υπόπτους. Αυτά τα τεστ έχουν βελτιωθεί και σήμερα αποτελούν μία βασική απόδειξη σε πολλές νομικές διαδικασίες λόγω της μεγάλης ποικιλίας μεταξύ των προφίλ DNA των διαφορετικών ατόμων ακόμα και αν ανήκουν στη ίδια εθνική ομάδα. Διάφορα γεγονότα, όπως η φήμη ορισμένων νομικών υποθέσεων στις οποίες έχει χρησιμοποιηθεί αυτός ο τύπος τεστ οι τηλεοπτικές σειρές όπως το CSI έχουν κάνει γνωστές αυτές τις τεχνικές. Αυτό που είναι λιγότερο γνωστό είναι ότι πίσω από την ανάπτυξη αυτών των προφίλ DNA βρίσκονται ισχυρά εργαλεία πιθανοτήτων, τα οποία έχουν οδηγήσει στη **δικαστική στατιστική**.

Μία από τις πρώτες δίκες στις οποίες χρησιμοποιήθηκαν τεστ DNA έγινε στην

Αγγλία το 1987 όπου ένας νέος 17 ετών κατηγορήθηκε για το βιασμό και τη δολοφονία δύο κοριτσιών, ενός το 1983 και του άλλου το 1987. Ορισμένες περιστασιακές ενδείξεις και το ιστορικό του νεαρού έμοιαζαν να τον δείχνουν ένοχο χωρίς πολλές αμφιβολίες. Ο νεαρός ζήτησε η ανάλυση του αίματος και το DNA να συγκριθούν με τα δείγματα σπέρματος που βρέθηκαν στα δύο κορίτσια. Το συμπέρασμα ήταν πειστικό τα 2 κορίτσια είχαν βιαστεί από το ίδιο άτομο αλλά δεν ήταν ο κατηγορούμενος. Στις νομικές υποθέσεις όπου επιχειρείται να προσδιοριστεί ο δράστης ή οι δράστες ενός εγκλήματος μετριέται κάποιο χαρακτηριστικό που επιτρέπει τον καθορισμό της πιθανής σχέσης του κατηγορούμενου με το έγκλημα.

Η ιδέα είναι απλή, αλλά το πρόβλημα είναι η πρακτική εφαρμογή της. Το θέμα είναι να οριστεί κάποιο χαρακτηριστικό που επιτρέπει να αναγνωριστεί με ακρίβεια ένα άτομο. Ένα από αυτά είναι τα δακτυλικά αποτυπώματα, τα οποία χρησιμοποιούνται όποτε είναι δυνατό, επειδή προσδιορίζουν πολύ καλά ένα άτομο. Ωστόσο, οποιοσδήποτε «επαγγελματίας» κακοποιός προσπαθεί να μην αφήσει δακτυλικά, αποτυπώματα και επιπλέον, μπορούν να καταστραφούν ή να αλλοιωθούν εύκολα. Το DNA είναι ένα χαρακτηριστικό που πλεονεκτεί σε σχέση με τα αποτυπώματα επειδή προσδιορίζει με σχεδόν μοναδικό τρόπο ένα άτομο (εκτός από τους ομοζυγωτούς δίδυμους που μοιράζονται το DNA και περιέργως όχι τα δακτυλικά αποτυπώματα), εκτός του ότι έχει διάφορα προτερήματα. Το πρώτο είναι ότι υπάρχει στον πυρήνα όλων των κυττάρων, έτσι ώστε αρκεί ελάχιστη ποσότητα οργανικής ύλης για τη μέτρηση του (μια τρίχα, μια σταγόνα αίματος, ένα απομεινάρι δέρματος).

Το δεύτερο είναι ότι είναι πολύ σταθερό και παραμένει αμετάβλητο για μεγάλο χρονικό διάστημα. Το μόνο μειονέκτημα είναι οι τεχνολογικοί περιορισμοί για να μετρηθεί με ακρίβεια, αφού αναλύονται μόνο μερικοί γενετικοί δείκτες και αν και το DNA είναι πρακτικά μοναδικό, υπάρχουν άνθρωποι που μοιράζονται έναν περιορισμένο αριθμό δεικτών. Δε θα μιλήσουμε εδώ για τον τρόπο ανάλυσης του DNA, αλλά θα περιγράψουμε πώς ο υπολογισμός των πιθανοτήτων συμβάλλει στη λήψη αποφάσεων σχετικά με το θεωρούμενο έγκλημα. Για να το απεικονίσουμε, θα θεωρήσουμε μία απλοποιημένη περίπτωση, λίγο ρεαλιστική αλλά χρήσιμη για τους σκοπούς μας. Έχει διαπραχτεί ένα έγκλημα και η αστυνομία συλλέγει δείγματα DNA στο χώρο που έγινε και έχει συλλάβει έναν ύποπτο, από τον οποίο έχει λάβει DNA. Έστω ότι κάθε δείγμα DNA παράγει ένα απλό δεδομένο, ανεξάρτητα από το αν στη δίκη εμφανίζονται και άλλες δυνατότητες, θα θεωρήσουμε ότι υπάρχουν μόνο δύο επιλογές ο ύποπτος είναι ένοχος (Ο) ή αθώος (Ι) και η απόφαση θα βασιστεί στα τεστ DNA. Ο σκοπός, επομένως, είναι η διάκριση μεταξύ των δύο ακόλουθων υποθέσεων:

1. Τα δείγματα προέρχονται από το ίδιο άτομο (C),
2. Τα δείγματα προέρχονται από διαφορετικά άτομα (I).

Συμβολίζουμε με Ev (ένδειξη) τη σύμπτωση των δειγμάτων που προέρχονται από το DNA που συλλέχτηκαν στον τόπο του εγκλήματος και του ύποπτου και με S τις άλλες περιπτώσεις. Το θεώρημα Bayes αναφέρει ότι για δύο γεγονότα A, B και συμβολίζοντας με $p(A/B)$ τη δεσμευόμενη πιθανότητα του γεγονότος A με δεδομένο το γεγονός B, υπό των

προϋπόθεσή ότι $p(B) \neq 0$, είναι $p(A/B) = \frac{p(B/A) \cdot p(A)}{p(B)}$. Χρησιμοποιώντας το

συμβολισμό της δεσμευμένης πιθανότητας και το θεώρημα Bayes, η αναλογία της ενοχής προς την αθωότητα δίνεται από

$$\frac{\text{prob}(C/Ev, S)}{\text{prob}(I/Ev, S)} = \frac{\text{prob}(Ev/C, S)}{\text{prob}(Ev/I, S)} \cdot \frac{\text{prob}(C/S)}{\text{prob}(I/S)}$$

Από αυτή την έκφραση φαίνεται ότι οι δύο βασικές πιθανότητες που πρέπει να καθοριστούν είναι η πιθανότητα της ένδειξης Ev αν ο ύποπτος είναι ένοχος ή αν είναι αθώος. Δεν αρκεί να εξεταστεί μόνο η πιθανότητα της ένδειξης όταν ο ύποπτος είναι αθώος και να

βγει το συμπέρασμα ότι μία μικρή τιμή αυτής της πιθανότητας είναι ισχυρή απόδειξη υπέρ της ενοχής. Η πιθανότητα της ένδειξης αν ο ύποπτος είναι ένοχος πρέπει επίσης να ληφθεί υπόψη. Αυτές οι δύο πιθανότητες δεν είναι συμπληρωματικές και θα μπορούσαν να είναι και οι δύο πολύ μικρές. Σε τέτοιες λεπτές διακρίσεις παίζουν δικηγόροι και εισαγγελείς και μία παρερμηνεία αυτών των λεπτών εννοιών των πιθανοτήτων μπορεί να σημαίνει την καταδίκη ενός αθώου ή την αθώωσή ενός ενόχου. Οι τεχνικές των δικηγόρων έχουν μικρή σχέση με τις πιθανότητες και μεγάλη με τη συμφέρουσα ερμηνεία τους.

Η προηγούμενη έκφραση δείχνει επίσης την εξάρτισης της αναλογίας ενοχής προς την αθωότητα με βάση άλλον τύπο αποδείξεων (S), που δεν πρέπει να παραμελούνται ποτέ. Ο καθορισμός αυτών των πιθανοτήτων της ένδειξης. Εν αν ο ύποπτος είναι ένοχος ή αθώος είναι ένα από τα αμφιλεγόμενα ζητήματα στις νομικές υποθέσεις όπου χρησιμοποιούνται τεστ DNA. Χρησιμοποιούνται διαδικασίες που επιτρέπουν τον καθορισμό τους με τον πιο αντικειμενικό δυνατό τρόπο, όπως η αναζήτηση πληθυσμών αναφοράς στους οποίους πραγματοποιούνται οι υπολογισμοί, χρησιμοποιώντας βάσεις δεδομένων ανά εθνικότητα ή φύλο. Όμως υπάρχουν πολλές περιπτώσεις στις οποίες το πρόβλημα δε λύνεται, καθώς ο τρόπος δημιουργίας βάσεων δεδομένων και οι υπολογισμοί που γίνονται με αυτές συνήθως δεν είναι διαφανείς. Επιπλέον, στα τεστ DNA μετρούνται συνήθως πολλοί δείκτες και έτσι πιθανότητα να έχουν εμφανιστεί όλοι μπορεί να ληφθεί με τον πολλαπλασιασμό των ατομικών πιθανοτήτων αν γίνει δεκτή η ανεξαρτησία τους, κάτι που είναι ένα άλλο σημείο αμφισβήτησης που χρησιμοποιείται στις δίκες. Τα τεστ DNA είναι πολύ σημαντικά για την αναγνώριση ατόμων σε καταστάσεις πολεμικών συγκρούσεων, καταστροφών ή ατυχημάτων.