Ποια είναι η πιθανότητα να περάσω την εξεταστική? - Made by Alex Bayes

Το παραπάνω PDF σχεδιάστηκε με σκοπό να καλύψει πλήρως όλη την απαραίτητη ύλη για το μάθημα πιθανοτήτων του τμήματός μας, προσφέροντας μια προσιτή και εύκολα κατανοητή παρουσίαση της θεωρίας.

Επιπλέον, στο τέλος της παρουσίασης παρέχω ένα σύνολο από παλιά θέματα εξετάσεων με λύσεις και επεξηγήσεις, ιδανικά για εξάσκηση.

Ποια είναι η πιθανότητα να περάσεις την εξεταστική? Η απάντηση σε αυτό το καίριο ερώτημα δεν θα σου δοθεί απλά, αλλά θα την υπολογίσεις ο ίδιος στο τέλος της παρουσίασης, με βάση τις γνώσεις που θα έχεις αποκτήσει.

Ενότητα 1: Αβεβαιότητα και Τύχη

Όταν δεν γνωρίζουμε όλες τις σχετικές πληροφορίες ή δεν έχουμε πλήρη κατανόηση του συστήματος ή των παραμέτρων που επηρεάζουν ένα γεγονός, τότε προκύπτει αβεβαιότητα. (Υπάρχουν πολλοί παράγοντες που μπορεί να επηρεάσουν το αν θα περάσεις την εξέταση, επομένως, υπάρχει αβεβαιότητα)

Η πιθανότητα χρησιμοποιείται για να ποσοτικοποιήσεις την αβεβαιότητα. ⭐

Εάν έχουμε αβεβαιότητα για ένα γεγονός, η πιθανότητα εκφράζει τη βεβαιότητα που έχουμε για την εμφάνιση αυτού του γεγονότος. Η πιθανότητα 0 υποδηλώνει ότι το γεγονός είναι αδύνατο να συμβεί, ενώ η πιθανότητα 1 υποδηλώνει βεβαιότητα ότι το γεγονός θα συμβεί. Ενδιάμεσες τιμές δείχνουν διάφορα επίπεδα αβεβαιότητας. (Αυτή την στιγμή δεν γνωρίζεις πιθανότητες, οπότε η πιθανότητα να περάσεις την εξέταση θεωρώ πως είναι μικρή)

Νέες πληροφορίες μπορούν να επηρεάσουν τις πιθανότητες και να μας δώσουν καλύτερη εικόνα για την αβεβαιότητα. Εάν είχαμε πλήρη και απόλυτη γνώση για όλες τις παραμέτρους που επηρεάζουν την εμφάνιση ενός γεγονότος, τότε θα μπορούσαμε να προβλέψουμε με βεβαιότητα το αν θα συμβεί η όχι. (Καθώς διαβάζεις το συγκεκριμένο PDF, θεωρώ πως η πιθανότητα να περάσεις αυξάνεται. Δεν θα φτάσει ποτέ το 1, γιατί πάντα θα υπάρχει αβεβαιότητα σχετικά με το τι μπορεί εν τέλη να συμβεί στην εξέταση. Αν μπορούσες να δεις το μέλλον, θα ήξερες με βεβαιότητα το αν θα περάσεις)

Θεωρούμε ότι ένα αποτέλεσμα είναι τυχαίο όταν υπάρχει πλήρης αβεβαιότητα για αυτό. Δηλαδή, δεν έχουμε καμία πληροφορία για αυτό και όλες οι παράμετροι που θα μπορούσαν να επηρεάσουν το αποτέλεσμα είναι άγνωστες ή αδύνατο να ληφθούν υπόψη. Επομένως, δεν μπορούμε να αξιολογήσουμε αν κάποιο αποτέλεσμα είναι πιο ή λιγότερο πιθανό από κάποιο άλλο. Σε αυτή την περίπτωση, όλα τα δυνατά αποτελέσματα θεωρούνται ισοπίθανα, δηλαδή, αποδίδουμε την ίδια πιθανότητα σε καθένα από αυτά.

Για παράδειγμα, αν ρίξουμε ένα κέρμα και δεν είμαστε σε θέση να υπολογίσουμε τη δύναμη της ρίψης, την επίδραση του αέρα, της βαρύτητας και της επιφάνειας όπου θα πέσει το κέρμα, τότε το αποτέλεσμα της ρίψης θεωρείται τυχαίο. Έτσι, κάθε πλευρά του κέρματος έχει πιθανότητα $1/2 = 0.5$ να εμφανιστεί. Είμαστε σίγουροι όμως πως το κέρμα είναι δίκαιο?

Όταν αντιμετωπίζουμε τέτοια γεγονότα, για τα οποία δεν έχουμε γνώση των παραμέτρων που τα επηρεάζουν ή των μηχανισμών που τα καθορίζουν, μπορεί αρχικά να τα θεωρούμε πλήρως τυχαία και αβέβαια. Ωστόσο, μέσω πειραμάτων και παρατήρησης, μπορεί να συγκεντρώσουμε δεδομένα που αποκαλύπτουν αν κάποια αποτελέσματα συμβαίνουν πιο συχνά από άλλα. Αυτό είναι ένα είδος νέας πληροφορίας και μπορεί να αλλάξει τις πιθανότητες.

Η χρήση πειραμάτων για την εκτίμηση πιθανοτήτων είναι βασικό θέμα στη στατιστική, η οποία είναι στενά συνδεδεμένη με τη θεωρία πιθανοτήτων.

Παράδειγμα: Κατα μέσο όρο, ποσοστό επιτυχίας στην εξέταση είναι περίπου 30%. Δεν μπορούμε να ξέρουμε πόσοι από όσους έγραψαν ήταν καλά προετοιμασμένοι, αλλά μπορούμε να εκτιμήσουμε πως είναι περίπου το 40%. Μπορούμε επίσης να θεωρήσουμε πως αν κάποιος πέρασε τότε ήταν σίγουρα καλά προετοιμασμένος, καθώς δεν υπάρχει δυνατότητα αντιγραφής. Τέλος, έστω πως πιστεύουμε ότι με πιθανότητα 90% είσαι καλά προετοιμασμένος.

Το παραπάνω μοντέλο μπορεί να χρησιμοποιηθεί για μια εκτίμηση της πιθανότητας σου να περάσεις την εξέταση. (*Θα δούμε πως μπορείς να κάνεις τους υπολογισμούς αργότερα στην παρουσίαση*)

Ερώτηση: μας δίνει το *παραπάνω* μοντέλο μια καλή εκτίμηση? 🧐

Υπάρχει θεωρούμε μια αβεβαιότητα σχετικά με το αν θα περάσεις εφόσον είσαι καλά προετοιμασμένος. Κάποιος που είναι σίγουρος πως είναι καλά προετοιμασμένος λέμε πως έχει πιθανότητα 75% να περάσει. Αυτό οφείλεται σε παράγοντες που δεν μπορούμε να προβλέψουμε, όπως το άγχος, η δυσκολία της εξέτασης και άλλες συνθήκες.

Ένα πρόβλημα είναι ότι, έχοντας μόνο δύο κατηγορίες ατόμων (καλά προετοιμασμένους και όχι καλά προετοιμασμένους), δεν λαμβάνουμε υπόψη την δυνατότητα να αυξήσεις την πιθανότητα σου πιο ψηλά από 75% αν είσαι πολύ καλά προετοιμασμένος και όχι απλά καλά προετοιμασμένος.

Ακόμα, η εκτίμηση σχετικά με το ποσοστό των καλά προετοιμασμένων ατόμων (40%) δεν μπορεί να επιβεβαιωθεί, είναι απλά μια εμπειρική εκτίμηση η οποία μπορεί να διαφέρει πολύ από την πραγματικότητα.

Τέλος, το πόσο πιθανό θεωρείς ότι είναι να είσαι καλά προετοιμασμένος είναι κάτι καθαρά υποκειμενικό το οποίο δεν μπορείς να υπολογίσεις με σιγουριά.

Παρόλα αυτά, η προσέγγιση μας δεν είναι τελείως άχρηστη. Μας βοηθά να ποσοτικοποιήσουμε την αβεβαιότητα και να κατανοήσουμε καλύτερα τις πιθανότητες επιτυχίας που έχουμε. Παρέχει ένα σημείο εκκίνησης, που μπορούμε να βελτιώσουμε στο μέλλον, με τη συλλογή περισσότερων δεδομένων και την ενσωμάτωση περισσότερων παραμέτρων στην ανάλυση.

Στην συνέχεια θα προτείνω και άλλα μοντέλα για τον υπολογισμό της πιθανότητας που έχεις να περάσεις την εξέταση. Θα μπορούσαμε ακόμα, να λάβουμε υπόψη και την μορφή του διαγωνίσματος, που θα είναι πιθανών πολλαπλής επιλογής με αρνητική βαθμολογία σε λάθος επιλογές που κάνεις.

Ερώτηση: Τελικά, υπάρχει πραγματική τύχη στο σύμπαν? Αν όχι, τότε έχουμε ελεύθερη βούληση? Θα μπορούσε κάποιος να προβλέψει την επόμενη σκέψη σου αν είχε πλήρη γνώση για το πώς λειτουργεί ο εγκέφαλός σου? Είναι δυνατόν ό,τι συνέβη και ό,τι θα συμβεί στο μέλλον να ήταν ήδη προδιαγεγραμμένο? Θα περάσεις τελικά την εξέταση ή όχι? 🤔

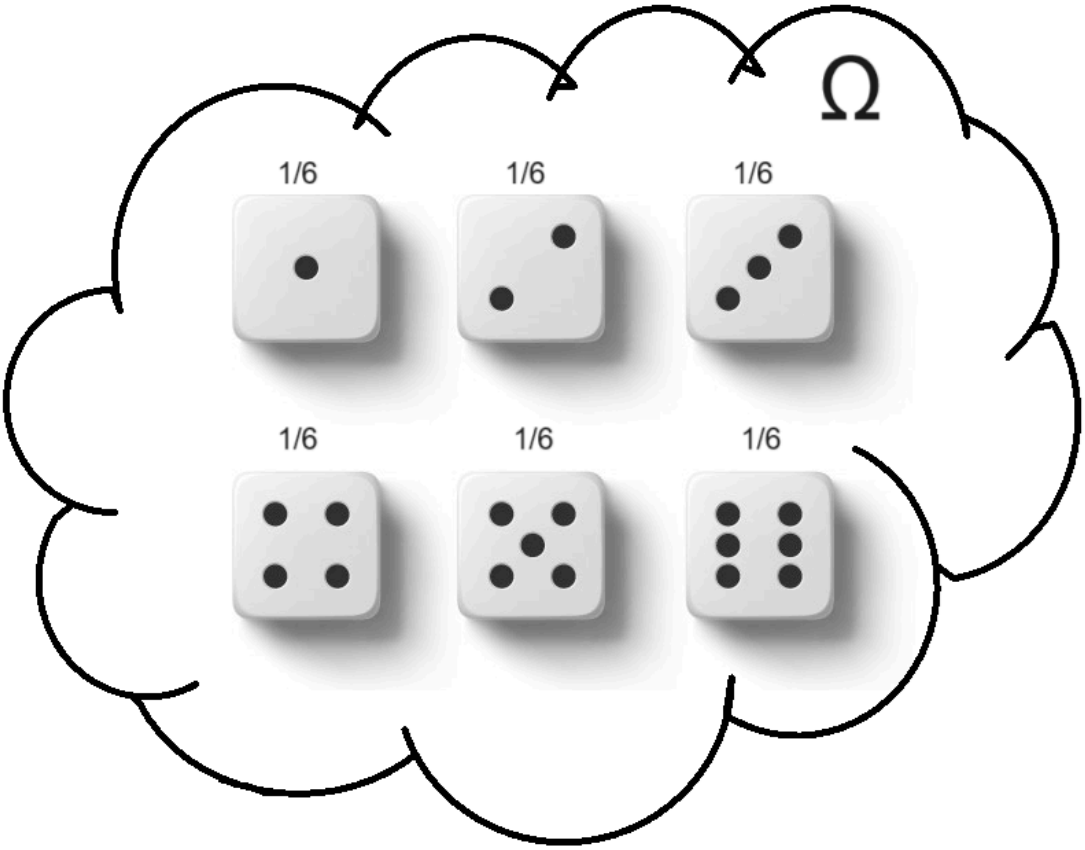
Στην κβαντική μηχανική, τα σωματίδια φαίνεται να συμπεριφέρονται ως κύματα πιθανοτήτων όταν δεν τα παρατηρούμε, γεγονός που υποδηλώνει την ύπαρξη μιας εγγενούς αβεβαιότητας στη φύση. Ο Albert Einstein διαφώνησε με την ιδέα ότι η τύχη είναι ένα βασικό χαρακτηριστικό του σύμπαντος. Αντίθετα, πίστευε ότι τα φαινόμενα που παρατηρούμε μπορούν να εξηγηθούν από άγνωστες μεταβλητές που δεν έχουμε ακόμα κατανοήσει.

Κάποια ενδιαφέρον βίντεο που προτείνω:
- What is Random?
- Dr Quantum   Double Slit Experiment
- Schrödinger's cat: A thought experiment in quantum mechanics - Chad …

Ενότητα 2: Οπτικοποίηση Πιθανοτήτων

Στην ακόλουθη ενότητα, σκοπός μου είναι να δείξω διάφορους τρόπους με τους οποίους μπορείς να οπτικοποιήσεις τις πιθανότητες. Κάτι το οποίο είναι χρήσιμο γιατί μέσα από την οπτικοποίηση των πιθανοτήτων ως γεωμετρικά στοιχεία θα μπορέσεις αργότερα να κατανοήσεις πολύ πιο εύκολα διάφορους τύπους και θεωρήματα τα οποία διαφορετικά ίσως φαινόντουσαν ακατανόητα.

Ας αρχίσουμε με ένα απλό παράδειγμα όπου έχουμε ένα ζάρι. Θεωρούμε τα πιθανά γεγονότα που μπορεί να προκύψουν εφόσον πετάξουμε αυτό το ζάρι. Το σύνολο αυτών των γεγονότων το ονομάζουμε Ω και πρόκειται για το σύνολο όλων των πιθανών αποτελεσμάτων του ζαριού. Στην συνέχεια, κατανέμουμε πιθανότητα σε κάθε ένα από τα πιθανά ενδεχόμενα. Σημαντικό είναι, το άθροισμα των πιθανοτήτων τους να είναι ίσο με 1, ούτε μεγαλύτερο ούτε μικρότερο. Καθώς το ζάρι είναι δίκαιο, θεωρούμε πως η κάθε πλευρά του έχει ίση πιθανότητα, και οπότε κάθε πλευρά θα έχει πιθανότητα 1/6 να εμφανιστεί.



Ερώτηση: ποια είναι η πιθανότητα το ζάρι να έρθει 1 η 2?

Καθώς όλες οι πλευρές είναι ισοπίθανες. Θα μπορούσες πολύ εύκολα να συμπεράνεις ότι η πιθανότητα είναι 2/6 εφόσον σε 2 από τα 6 πιθανα ενδεχόμενα θα έχει έρθει 1 η 2. Γενικότερα όμως, ο τρόπος που θα έκανες τον συγκεκριμένο υπολογισμό, θα ήταν προσθέτοντας τις πιθανότητες των στοιχείων τα οποία ικανοποιούν την συνθήκη σου. Δηλαδή 1/6 + 1/6 = 2/6 ⭐

Για τον λόγο αυτό, σε πολλές ασκήσεις θα μας χρειαστεί να γνωρίζουμε λίγο θεωρία συνόλων. Καθώς θα απαριθμούμε ενδεχόμενα που τηρούν κάποια κριτήρια και θα αθροίζουμε τις πιθανότητές τους. Δεν πειράζει αν δεν θυμάσαι καλά σύνολα και απαρίθμηση, αφού θα τα θυμηθούμε αργότερα σε ασκήσεις.

Παράδειγμα: έστω υποσύνολα A και B του Ω, και θες να υπολογίσεις την πιθανότητα να συμβεί γεγονός που ανήκει στο A η στο B.

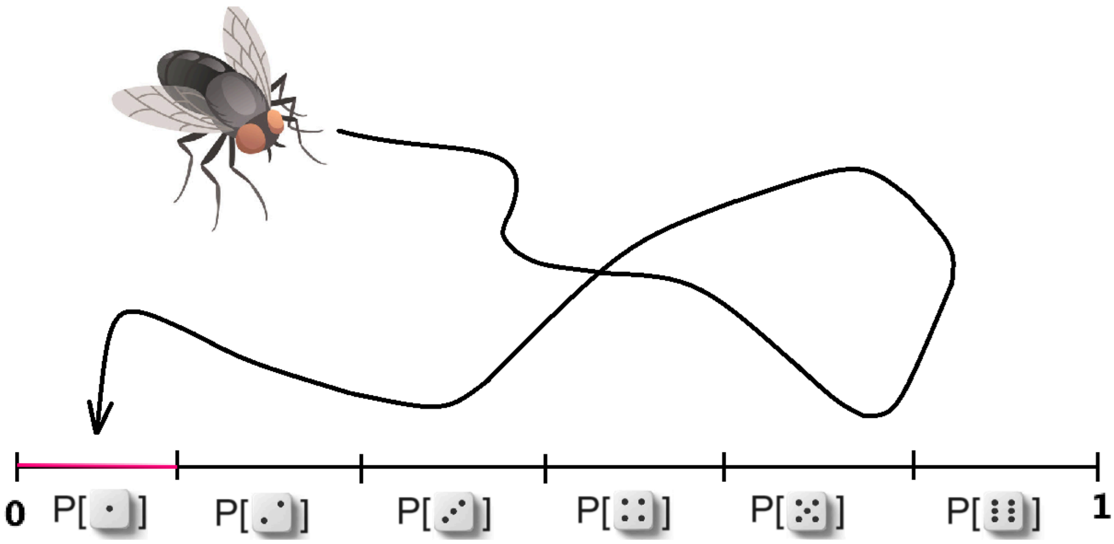
Η πιθανότητα να συμβεί γεγονός που ανήκει στο A, συμβολίζεται με P[A] και ισούται όπως είπα και πριν με το άθροισμα των πιθανοτήτων των στοιχείων του. Παρόμοια η πιθανότητα να συμβεί γεγονός που ανήκει στο B συμβολίζεται με P[B] και είναι ίση με το άθροισμα των πιθανοτήτων των δικών του στοιχείων. Η πιθανότητα να συμβεί γεγονος που ανήκει είτε στο A είτε στο B, συμβολίζεται ως $P[A \cup B]$, και ισούται με $P[A] + P[B] - P[A \cap B]$. ⭐

Η αφαίρεση με $[A \cap B]$ γίνεται διότι μπορεί το A με το B να έχουν κοινά στοιχεία, και οπότε να έχουμε αθροίσει τις πιθανότητες τους δύο φορές.

Ερώτηση: Ποια είναι η πιθανότητα αν ριξω το ζάρι 2 φορές να έρθει την πρώτη φορά 1 και την δεύτερη 3? Αλλάζει η πιθανότητα αν δεν μας ενδιαφέρει ποιο από τα δύο ήρθε πρώτο, αλλά απλά θέλουμε να έρθουν 1 και 3 με οποιαδήποτε σειρά?

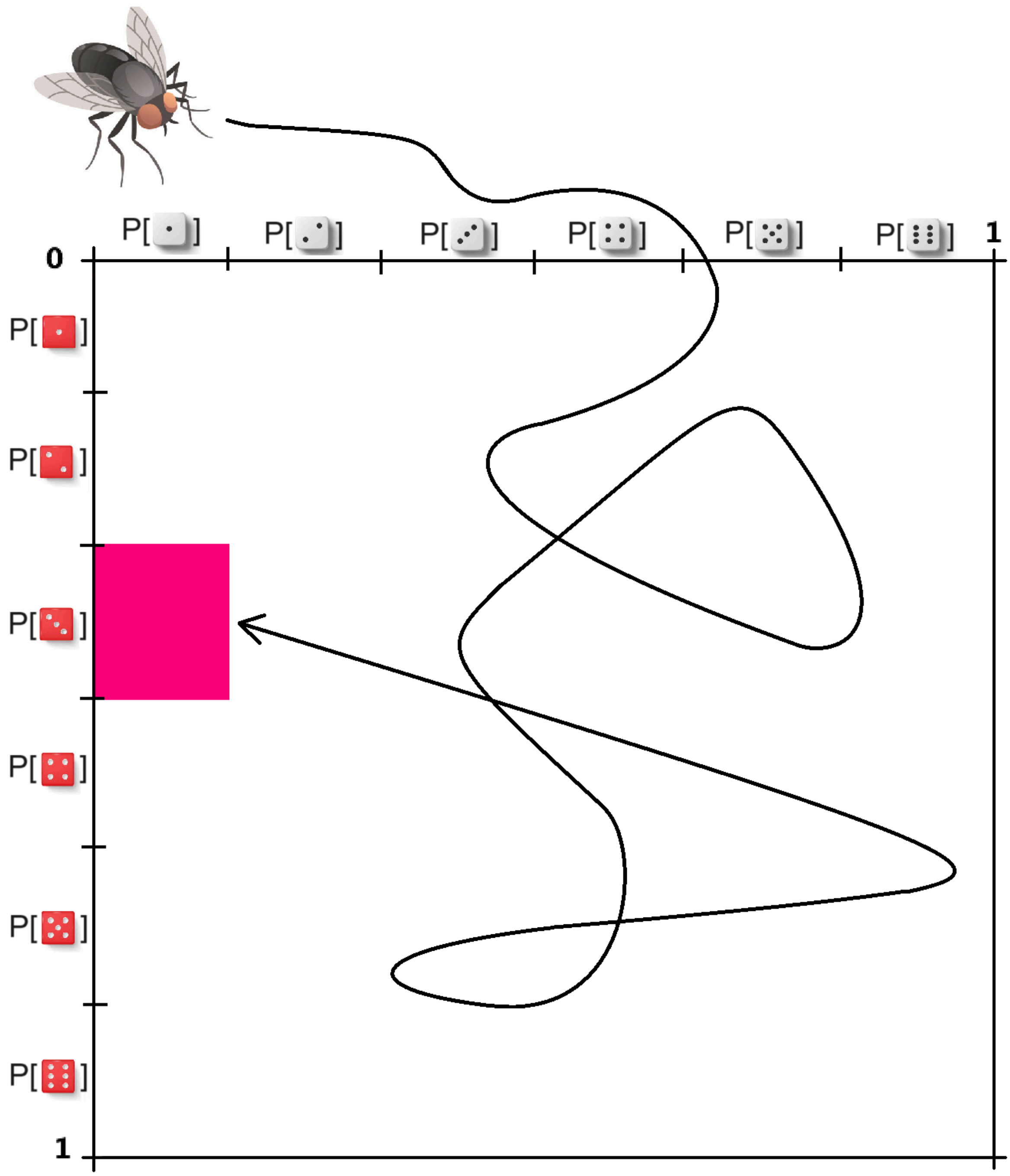
Για να απαντήσουμε τα παραπάνω ερωτήματα, θα ήθελα να προτείνω έναν ενδιαφέρον τρόπο σκέψης, με τον οποίο μπορεί κάποιος να οπτικοποιήσει τις πιθανότητες ως γεωμετρικά στοιχεία στον χώρο

Παίρνουμε τον άξονα των πραγματικών αριθμών από το 0 έως το 1 και αντιστοιχούμε μέρη του στα διάφορα πιθανά ενδεχόμενα που μπορούν να συμβούν. Το πόσο μέρος θα δοθεί σε κάθε ενδεχόμενο είναι ανάλογο της πιθανότητας που έχει να συμβεί. Στη συνέχεια, επιλέγουμε ένα σημείο του άξονα τελείως τυχαία. Οπότε, αν ένα ενδεχόμενο του είχαμε δώσει για παράδειγμα 1/6 του άξονα, τότε θα είχε πιθανότητα 1/6 να επιλεγεί.



Στο παράδειγμα πάνω έχω δώσει σε κάθε ενδεχόμενο ένα συνεχές μέρος της γραμμής με μήκος ίσο με 1/6. Αυτό που έχει σημασία όμως είναι το συνολικό μήκος του άξονα που αντιστοιχεί σε ένα ενδεχόμενο, ανεξάρτητα από το αν αυτό αποτελείται από συνεχόμενα ή διάσπαρτα τμήματα.

Αν θεωρήσουμε και τη δεύτερη ρίψη του ζαριού, περνάμε από τη μία διάσταση στις δύο διαστάσεις. Ο οριζόντιος άξονας θα χρησιμοποιηθεί για τα πιθανά αποτελέσματα της πρώτης ρίψης του ζαριού και ο κάθετος άξονας για τα πιθανά αποτελέσματα της δεύτερης ρίψης. Επιλέγοντας ένα τυχαίο σημείο στον οριζόντιο άξονα και ένα τυχαίο σημείο στον κάθετο άξονα, είναι ισοδύναμο με το να επιλέγουμε ένα τυχαίο σημείο στο δισδιάστατο επίπεδο.

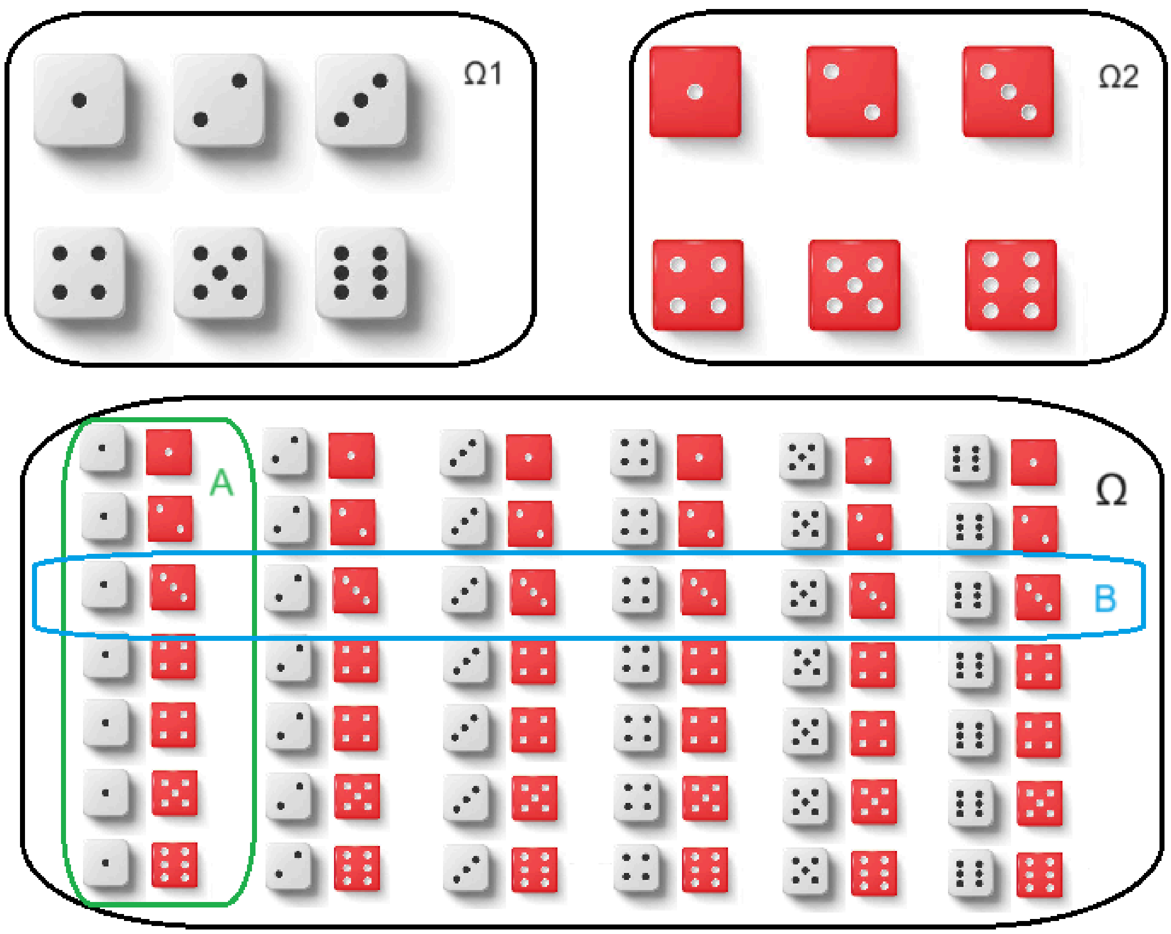


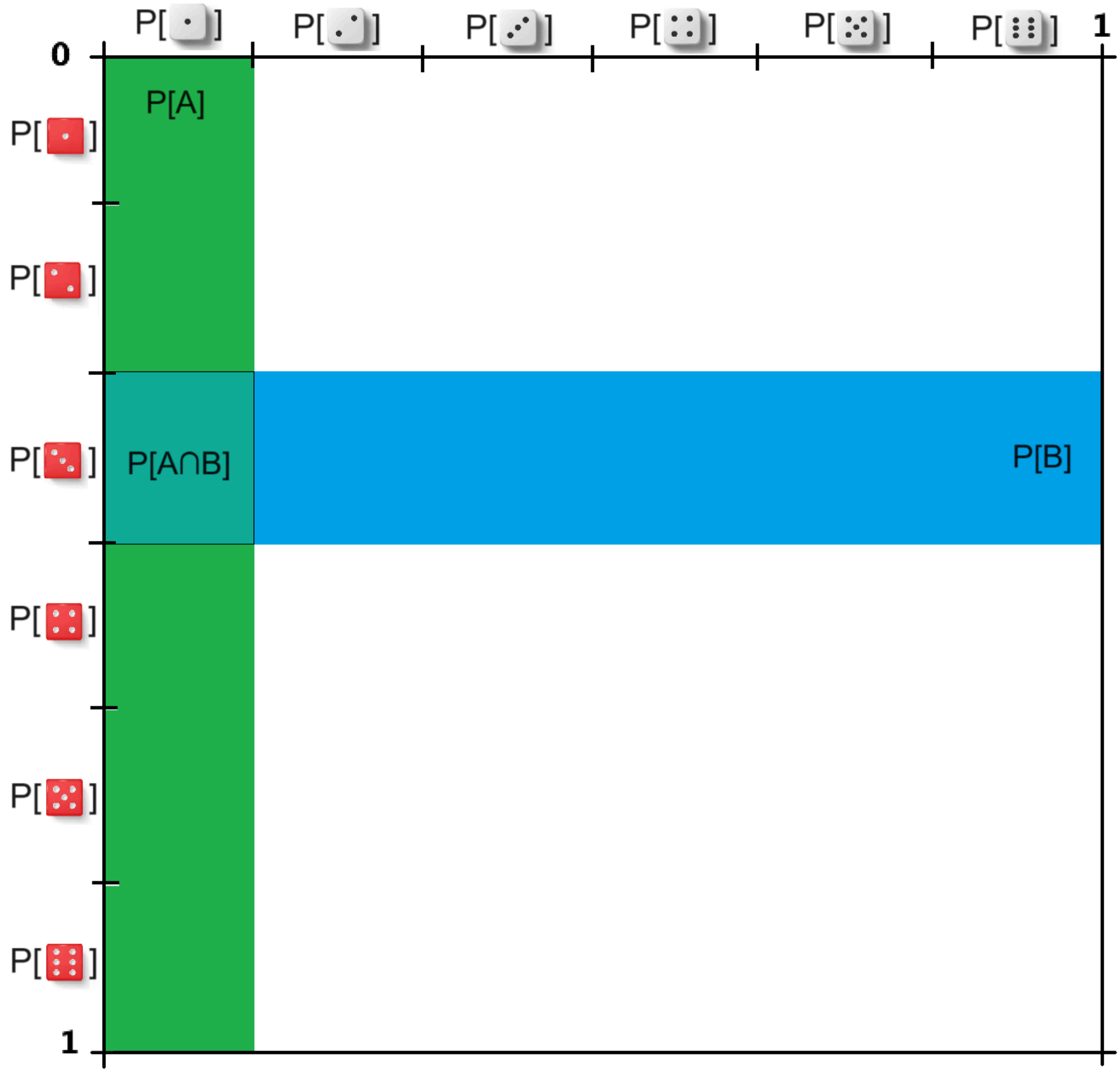
Κάθε σημείο στο δισδιάστατο επίπεδο έχει ίση πιθανότητα να επιλεγεί. Επομένως, η πιθανότητα το πρώτο ζάρι να έρθει 1 και το δεύτερο να έρθει 3 είναι ανάλογη του εμβαδόν της επιφάνειας που αντιστοιχεί στο ζάρι 1 για τον οριζόντιο άξονα και στο ζάρι 3 για τον κάθετο άξονα. Αυτό το εμβαδόν υπολογίζεται εύκολα και είναι ίσο με 1/6 * 1/6 = 1/36.

Αν δεν μας ενδιαφέρει ποιο από τα δύο ζάρια έφερε το 1 ή το 3, αλλά θέλουμε απλά να εμφανιστούν οι αριθμοί 1 και 3 με οποιαδήποτε σειρά, τότε θα πρέπει να προσθέσουμε και την επιφάνεια που αντιστοιχεί στο ζάρι 3 για τον οριζόντιο άξονα και στο ζάρι 1 για τον κάθετο άξονα. Συνεπώς, η πιθανότητα θα είναι το άθροισμα των δύο εμβαδόν 1/36 + 1/36 = 2/36 = 1/18.

Γενικότερα, όταν έχεις ένα γεγονός A με πιθανότητα P[A] και ένα γεγονός B με πιθανότητα P[B], αν τα δύο γεγονότα είναι ανεξάρτητα, δηλαδή, το να συμβεί το ένα δεν επηρεάζει την πιθανότητα του να συμβεί το άλλο, τότε η πιθανότητα να συμβούν και τα δύο γεγονότα μαζί συμβολίζεται ως P[A ∩ B] και είναι ίση με το γινόμενο P[A] * P[B]. ⭐

Παράδειγμα: A = “το πρώτο ζάρι ήρθε 1” και B = “το δεύτερο ζάρι ήρθε 3”



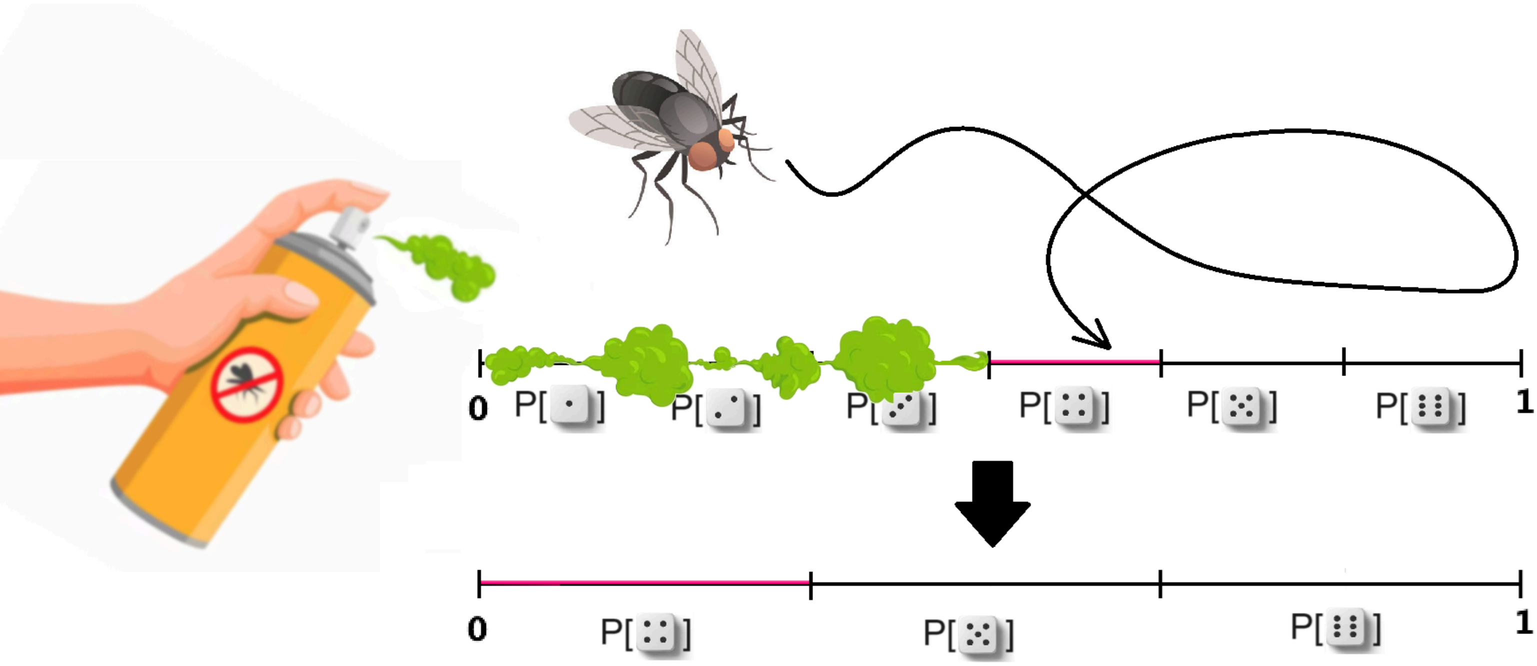


Αν Ω1 είναι το σύνολο όλων των πιθανών αποτελεσμάτων της πρώτης ρίψης, και Ω2 είναι το σύνολο όλων των πιθανών αποτελεσμάτων της δεύτερης ρίψης, τότε το σύνολο Ω ορίζεται ως το καρτεσιανό γινόμενο Ω1 και Ω2 και περιέχει όλους τους πιθανούς συνδυασμούς των δύο ρίψεων.

Όταν ορίζω γεγονός A = “το πρώτο ζάρι ήρθε 1”, στην πραγματικότητα αναφέρομαι σε ένα υποσύνολο του Ω, που περιέχει όλα τα ενδεχόμενα στα οποία το πρώτο ζάρι ήρθε 1. Η πιθανότητα να συμβεί ενδεχόμενο που να ανήκει στο A συμβολίζεται ως P[A] και είναι ίση με 1/6 καθώς τόσο είναι η πιθανότητα να έρθει 1 στην πρώτη ρίψη. Αντίστοιχα όταν ορίζω B = “το δεύτερο ζάρι ήρθε 3” αναφέρομαι σε ένα υποσύνολο του Ω που περιέχει όλα τα ενδεχόμενα στα οποία το δεύτερο ζάρι ήρθε 3, P[B] = 1/6 καθώς τόσο είναι η πιθανότητα να έρθει 3 στην δεύτερη ρίψη. P[A∩B] = P[A] * P[B] = 1/36

Πλέον γνωρίζεις τον κανόνα του γινομένου για ανεξάρτητα γεγονότα 🏆

Συνεχίζοντας, έστω πως τώρα κάνουμε το εξής πείραμα με έναν φίλο. Ρίχνω ένα ζάρι και κλείνω τα μάτια μου. Χωρίς να γνωρίζω το αποτέλεσμα, μαθαίνω από τον φίλο μου πως το ζάρι δεν έχει έρθει 1, 2 η 3. Πως θα μεταβληθούν οι πιθανότητες των πλευρών των οποίων δεν έχω ακόμα απορρίψει?



Εφόσον γνωρίζουμε ότι το ζάρι δεν έχει έρθει 1, 2 η 3, τα μόνα σημεία στα οποία μπορεί να έχει προσγειωθεί η “μύγα της τύχης” είναι μέσα στα κομμάτια που αντιστοιχούν στις πλευρές 3, 4 και 5. Η πιθανότητα να επιλεγεί σημείο μέσα σε ένα από αυτά τα κομμάτια, είναι ανάλογη του ποσοστού του απομένοντα χώρου το οποίο καλύπτει το κομμάτι.

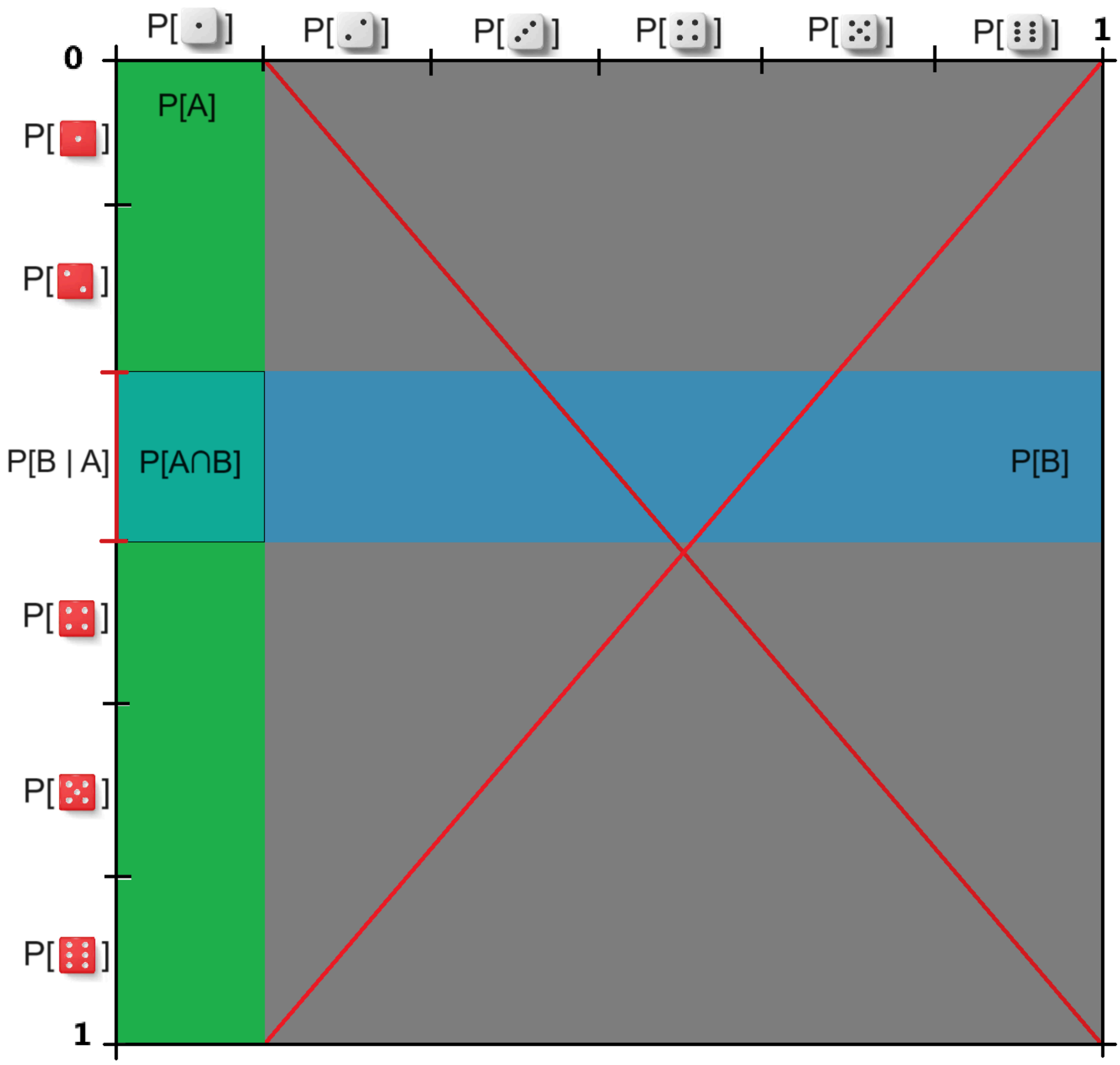
Έχοντας απορρίψει συνολική πιθανότητα ίση με 3/6, ο χώρος που απομένει θα ισούται με 1 - 3/6 = 1/2. Κάθε μία από τις πλευρές που μένουν, τους έχουμε αντιστοιχίσει μήκος του άξονα ίσο με 1/6. Επομένως, η πιθανότητα μιας πλευράς να επιλεγεί θα μετατραπεί σε (1/6) / (1/2) = 1/3 (καθώς κάνουμε διαίρεση μήκος αντιστοιχισμένο στην πλευρά δια μήκος απομένοντα χώρου)

Γενικότερα, όταν αποκτούμε πληροφορίες που αποκλείουν ένα σύνολο ενδεχομένων με συνολική πιθανότητα C. Για κάποιο ενδεχόμενο A από αυτά που μένουν, το οποίο αρχικά είχε πιθανότητα P[A], γίνεται αναθεώρηση της πιθανότητας του σε $P'[A] \;=\; \frac{P[A]}{1-C}$ ⭐

Η παραπάνω μέθοδος εφαρμόζεται και για το δισδιάστατο επίπεδο. Για παράδειγμα, ρίχνοντας δύο ζάρια, αν μάθουμε το αποτέλεσμα της μιας ρίψης, απορρίπτουμε μέρος του δισδιάστατου επιπέδου, και κάνουμε αναθεώρηση των πιθανοτήτων για τα ενδεχόμενα που απομένουν.

Για δύο κανονικές ρίψεις, μπορούμε να δούμε ότι, μαθαίνοντας το αποτέλεσμα της μίας ρίψης, οι πιθανότητες της άλλης ρίψης εξακολουθούν να καλύπτουν το 1/6 του απομένοντα χώρου. Αυτό δείχνει πως οι δύο ρίψεις είναι ανεξάρτητες. Το αν δύο γεγονότα είναι ανεξάρτητα ή όχι, είναι αποτέλεσμα του πως έχουμε κατανείμει χώρο του επιπέδου στα διάφορα δυνατά ενδεχόμενα.

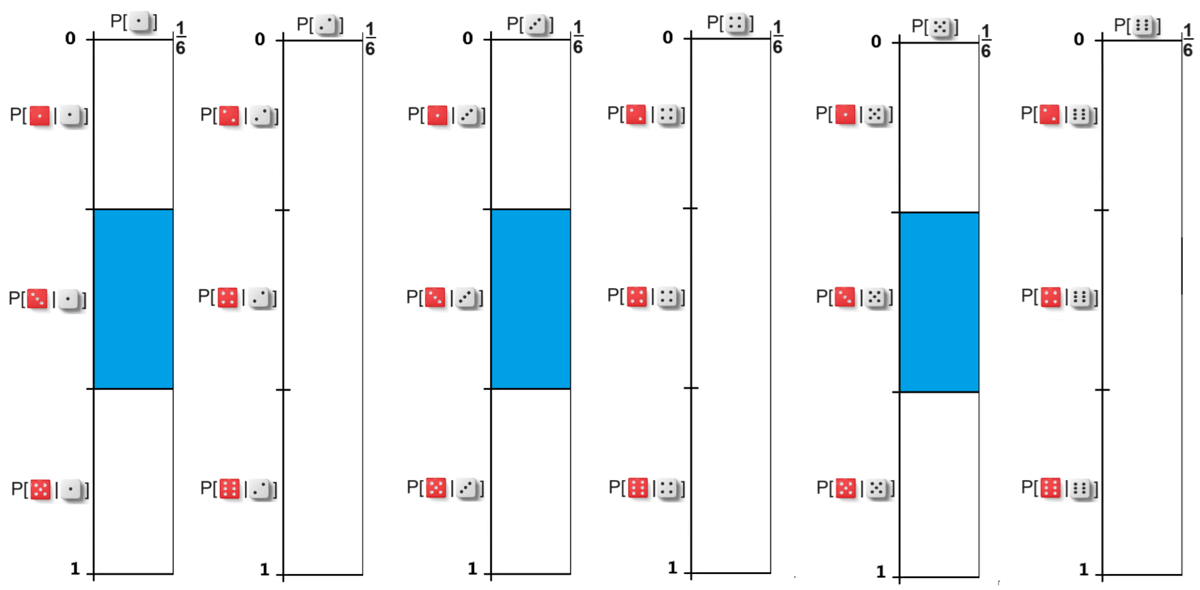
Παράδειγμα: Έστω A = “το πρώτο ζάρι ήρθε 1” και B = “το δεύτερο ζάρι ήρθε 3”. Η πιθανότητα το δεύτερο ζάρι να ήρθε 3 δεδομένου ότι το πρώτο ζάρι ήρθε 1 συμβολίζεται ως P[B | A] και ισούται με $\frac{P[A \cap B]}{P[A]} = \frac{1/36}{1/6} = 6/36 = 1/6$



Μπορείς να σκεφτείς πως θα έμοιαζε το σχήμα για γεγονότα τα οποία είναι εξαρτημένα? Ο κάθετος άξονας που ποσοτικοποιεί πιθανότητες της δεύτερης ρίψης δεν βολεύει. Θέλουμε οι πιθανότητες της δεύτερης ρίψης να προσαρμόζονται κατάλληλα ανάλογα το αποτέλεσμα της πρώτης ρίψης.

Παράδειγμα: Ρίχνω 2 ζάρια, και έχοντας τα μάτια μου κλειστά, μαθαίνω από τον φίλο μου πως το άθροισμα των αποτελεσμάτων είναι άρτιος αριθμός. Πλέον οι πιθανότητες των πλευρών μιας ρίψης θα πρέπει να προσαρμόζονται μαθαινοντας το αποτέλεσμα της άλλης ριψης.

Έστω πως μαθαίνουμε ότι η πρώτη ρίψη ήρθε 1. Γνωρίζουμε πως σίγουρα το δεύτερο ζάρι δεν θα έχει έρθει 2, 4 η 5 καθώς το άθροισμα πρέπει να είναι άρτιος αριθμός. Έχοντας απορρίψει τις παραπάνω πλευρές, θα γίνει αναθεώρηση των πιθανοτήτων των υπόλοιπων πλευρών (1, 3 η 5) σε 1/3 σύμφωνα με την μέθοδο που είδαμε προηγουμένως.

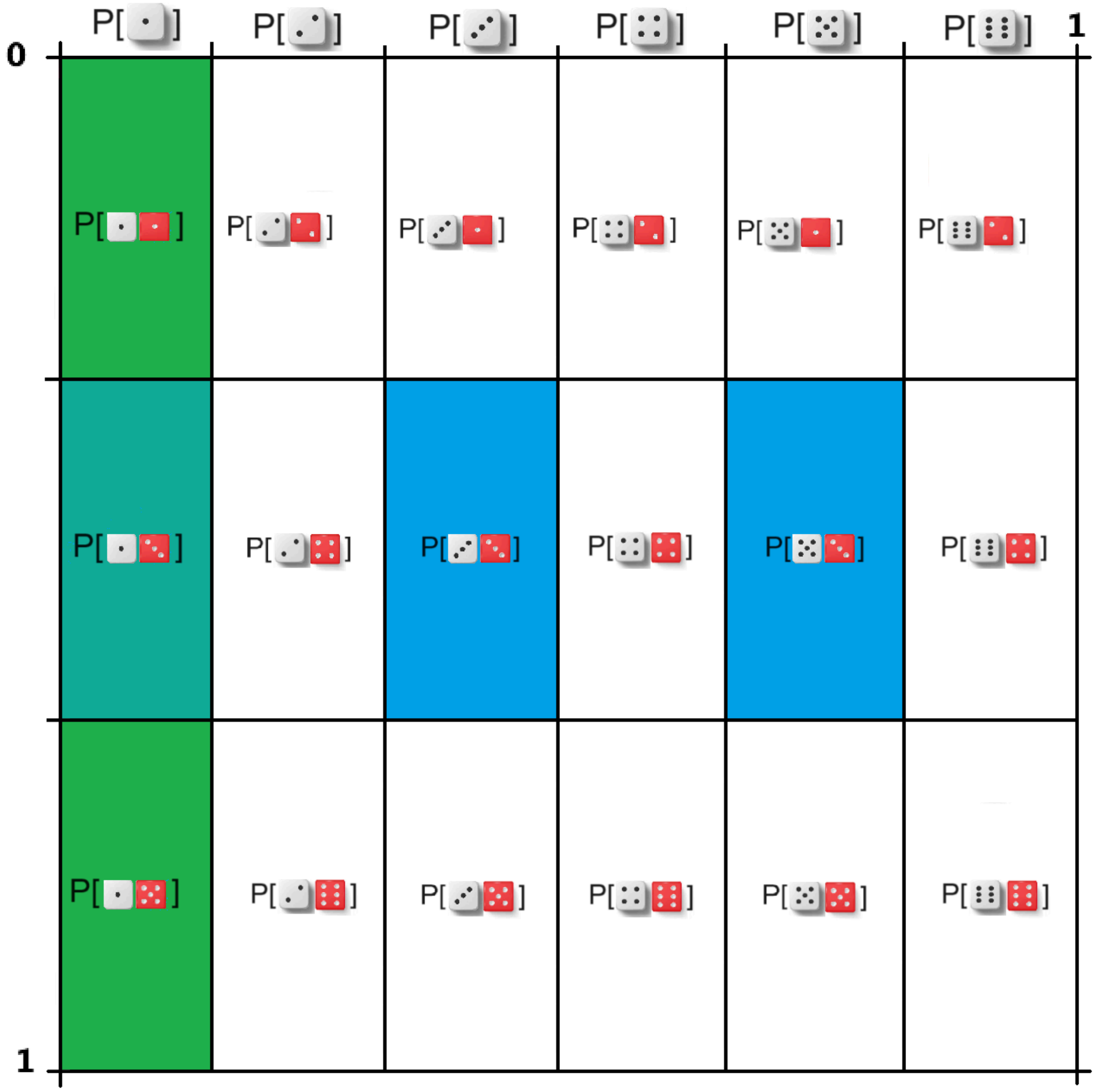


Κατανέμουμε τον χώρο, έτσι ώστε, απορρίπτοντας το μέρος του άξονα που δεν αντιστοιχεί σε ενδεχόμενα όπου η πρώτη ρίψη ήρθε 1, τα ενδεχόμενα όπου η δεύτερη ρίψη έχει έρθει 1, 3 η 5 να καλύπτουν το 1/3 του απομένοντα χώρου. Ανάλογα προσαρμόζουμε και για άλλα αποτελέσματα πρώτης ρίψης.

Οι μπλε επιφάνειες στην εικόνα είναι όλα τα κομμάτια τα οποία αντιστοιχούν σε ενδεχόμενα όπου η δεύτερη ρίψη έχει έρθει 3. Αν αθροίσεις τα εμβαδά τους, υπολογίζεις την πιθανότητα το δεύτερο ζάρι να έρθει 3, η οποία είναι 1/6.

Γενικότερα, για γεγονότα Α και Β, η πιθανότητα να συμβεί το Α συμβολίζεται ως P[A] και είναι ίση με την πιθανότητα να συμβεί το Α και το Β συν την πιθανότητα να συμβεί το Α και όχι το Β, δηλάδη P[A] = P[A∩B] + P[A∩B’] ⭐

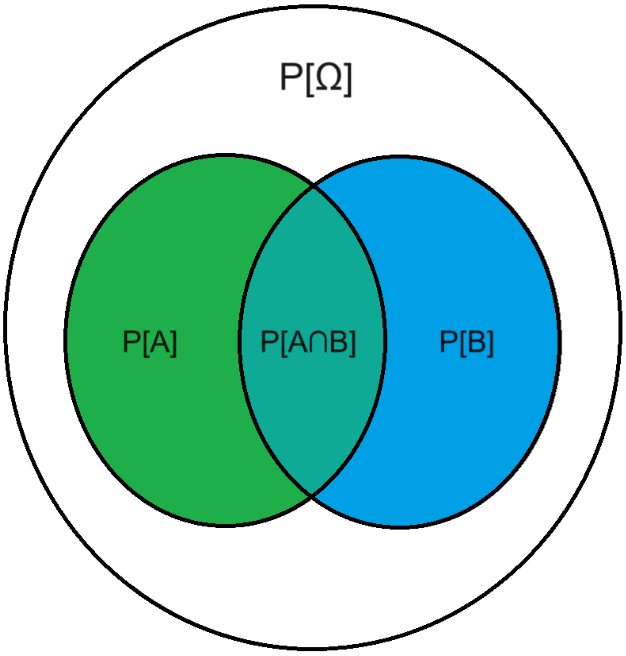
Πως υπολογίζουμε την πιθανότητα να συμβεί το Α και το Β? Ο κανόνας του γινομένου που είχαμε δει πιο πριν ισχύει μόνο για ανεξάρτητα γεγονότα.



Η πιθανότητα να συμβεί το Α και το Β συμβολίζεται ως P[A∩B] και ειναι ιση με το εμβαδόν της επιφάνειας που αντιστοιχεί σε ενδεχόμενα όπου συμβαίνουν τα Α και Β μαζί, το οποίο εμβαδόν εύκολα υπολογίζεται ως P[A] * P[B | A] ⭐

Παρατήρησε ότι, αν τα δύο γεγονότα είναι ανεξάρτητα, τότε P[B | A] = P[B] και οπότε P[A∩B] = P[A] * P[B] όπως είχαμε δει και προηγουμένως.

Η εξάρτηση μεταξύ των δύο ενδεχομένων δεν είναι μονόδρομη, αλλά είναι αμφίδρομη. Χρησιμοποιώντας το πάνω σχήμα, οι τύποι που έχουμε δει ως τώρα ισχύουν και προς τις δύο μεριές. Για να γίνει το παραπάνω πιο φανερό, θα ήθελα να προτείνω έναν επιπλέον τρόπο οπτικοποίησης. Για Α = “το πρώτο ζάρι ήρθε 1” και Β = “το δεύτερο ζάρι ήρθε 3”, κάποιος χώρος του πεδίου αντιστοιχεί στο P[A], κάποιος στο P[B] και έχουν κάποιον κοινό χώρο P[A∩B]. Το πως αυτό μοιάζει ακριβώς πάνω στο επίπεδο δεν μας ενδιαφέρει.



Στην πραγματικότητα, το μόνο που μας ενδιαφέρει είναι πόσο χώρο καταλαμβάνει ένα ενδεχόμενο στο επίπεδο, όχι η θέση αυτού του χώρου, αν είναι συνεχής ή διασπασμένος σε διάφορα μέρη, ή το σχήμα του.

Ως P[A | B] έχουμε ορίσει το ποσοστό του P[B] το οποίο καλύπτεται από το P[A∩B]. Το οποίο υπολογίζεται ως P[A∩B] / P[B] προφανώς. Αντίστοιχα, το P[B | A] θα είναι το ποσοστό του P[A] το οποίο καλύπτεται από το P[A∩B] το οποίο υπολογίζεται ως P[A∩B] / P[A].

Το P[A∩B] μπορεί να υπολογιστεί είτε ως P[A] * P[B | A] είτε ως P[B] * P[A | B], εφόσον, αν το P[A | B] για παράδειγμα είναι το ποσοστό του P[B] το οποίο καλύπτεται από το P[A∩B], αν πολλαπλασιάσεις αυτό το ποσοστό με P[B] προφανώς παίρνεις την ποσότητα P[A∩B].

Τέλος, επίσης εύκολα από θεωρία συνόλων είναι φανερό ότι P[A] = P[A∩B] + P[A∩B’] και P[B] = P[A∩B’] + P[A’∩B] όπου A’ = Ω - A και B’ = Ω - B. Το P[A∩ B’] μπορεί να υπολογιστεί ως P[A] * P[B’ | A] η ως P[B’] * P[A | B’]. Αν γνωρίζουμε το P[B | A] τότες P[B’ | A] = 1 - P[B | A]. Και άλλα πολλά…

Όλες αυτές οι σχέσεις βγαίνουν εύκολα από το σχήμα πάνω, για αυτό και η γεωμετρική αναπαράσταση βοηθάει. Απλά φτιάχνεις το σχέδιο με βάση τα δεδομένα που σου δίνει η άσκηση και μπορείς αμέσως να καταλάβεις ποια άλλα δεδομένα μπορείς να υπολογίσεις από αυτά.

Από όσα έχουμε δει προκύπτει και το γνωστό θεώρημα του Bayes, το οποίο χωρίς την οπτικοποίηση που κάναμε ίσως φαινόταν λίγο δυσνόητο.

Θεώρημα του Bayes: $P[B \mid A] = \frac{P[A \cap B]}{P[A]} = \frac{P[B] * P[A \mid B]}{P[A \cap B] + P[A \cap B']}$ ⭐

Ας δούμε λίγες ασκήσεις για εξάσκηση, χωρίς ζάρια αυτή την φορά! ⚔️

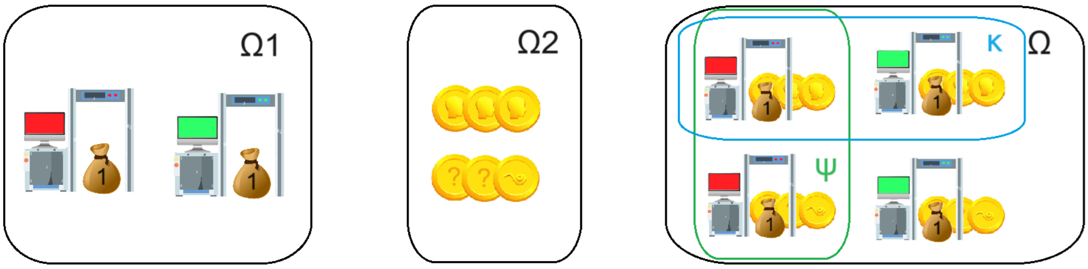
Άσκηση1: Μετά από πολύ προσπάθεια, έχοντας αντιμετωπίσει πολλά εμπόδια, επιτέλους φτάνεις στο μυθικό ιπτάμενο νησί, πετώντας με τον πιστό σου κόκκινο δράκο Αλδουιν.

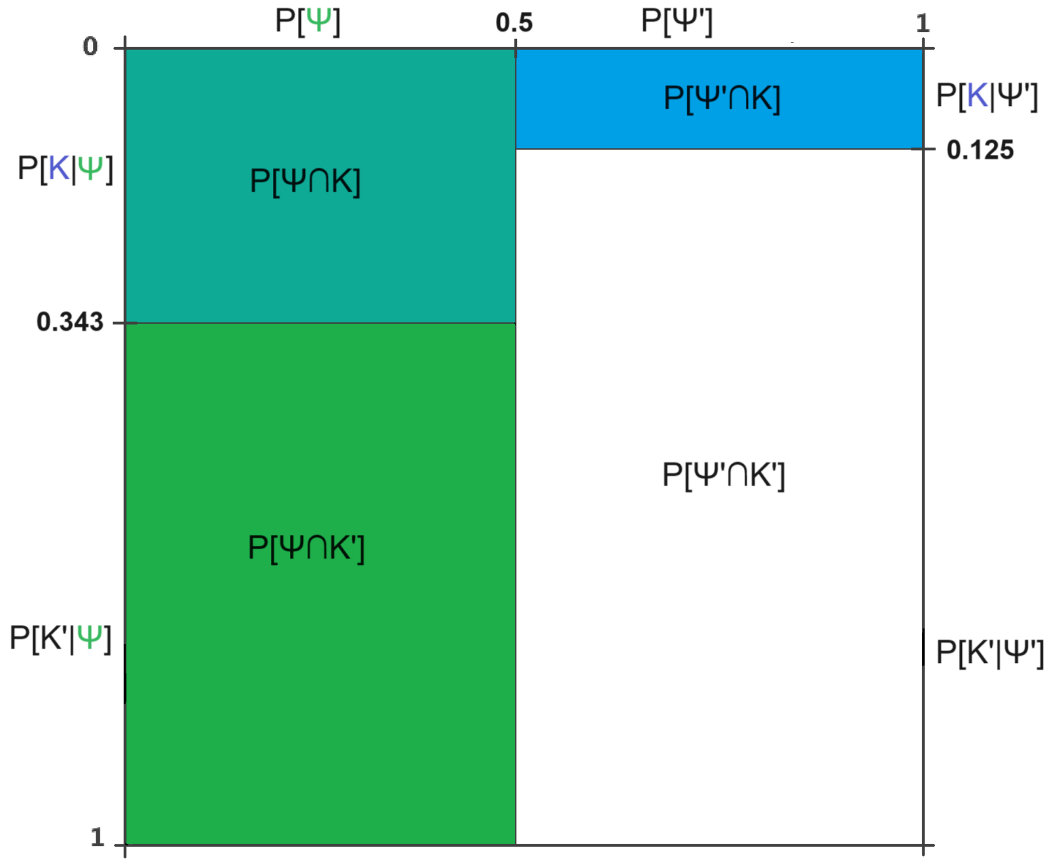
Ο σοφός γέρος του χωριού σε είχε ενημερώσει πως στο νησί θα βρεις δύο σακιά γεμάτα με χρυσά νομίσματα μεγάλης αξίας. Ένα από τα δύο όμως θα περιέχει ψεύτικα νομίσματα. Ο Αλδουιν είναι μικρός, και μπορεί να κουβαλήσει το πολύ ένα σακί μαζί με εσένα. Τα ψεύτικα νομίσματα είναι πολύ δύσκολο να τα ξεχωρίσεις από τα κανονικά. Ο γέροντας όμως, έχοντας αντιμετωπίσει παρόμοιες καταστάσεις στις δικές του περιπέτειες στο παρελθόν, σου έχει αποκαλύψει μια κρίσιμη διαφορά.

Ενώ τα πραγματικά νομίσματα, αν τα πετάξεις, έχουν 50% πιθανότητα να έρθουν κορώνα ή γράμματα, τα ψεύτικα, επειδή είναι πιο βαριά η πλευρά της κορώνας, έχουν 70% πιθανότητα να έρθουν κορώνα.

Διαλέγεις το πρώτο από τα σακιά. Παίρνεις ένα νόμισμα από μέσα και το πετάς 3 φορές. Και τις 3 φορές έρχεται κορόνα. Αποφασίζεις πως με μεγάλη πιθανότητα το συγκεκριμένο σακί είναι το σακί με τα ψεύτικα νομίσματα, οπότε κρατάς το άλλο. Πόσο βέβαιος είσαι όμως για την επιλογή σου?

Λύση: Έστω Ψ = "το σακί περιέχει ψεύτικα νομίσματα" και Κ = "ήρθε κορόνα 3 φορές". Η άρνηση του Ψ, δηλαδή Ψ', είναι "το σακί δεν περιέχει ψεύτικα νομίσματα", και η άρνηση του Κ, δηλαδή Κ', είναι "δεν ήρθε κορόνα 3 φορές".





Η πιθανότητα να περιέχει το σακί ψεύτικα νομίσματα είναι P[Ψ] = 0.5, αφού υπάρχει ίση πιθανότητα το σακί να περιέχει είτε πραγματικά είτε ψεύτικα νομίσματα. Η πιθανότητα να μην περιέχει το σακί ψεύτικα νομίσματα, P[Ψ'], θα είναι επομένως επίσης 0.5.

Αν το σακί περιέχει πραγματικά νομίσματα (Ψ'), η πιθανότητα να έρθει κορόνα 3 φορές σε συνεχόμενες ρίψεις, P[K|Ψ'], υπολογίζεται με βάση τον κανόνα του γινομένου ανεξάρτητων γεγονότων ως 0.5 * 0.5 * 0.5 = 0.125, αφού η πιθανότητα να έρθει κορόνα σε κάθε ρίψη είναι 0.5.

Αντίστοιχα, αν το σακί περιέχει ψεύτικα νομίσματα (Ψ), η πιθανότητα να έρθει κορόνα σε κάθε ρίψη είναι 0.7. Επομένως, η πιθανότητα να έρθει κορόνα 3 φορές συνεχόμενα, P[K|Ψ], υπολογίζεται ως 0.7 * 0.7 * 0.7 = 0.343.

Εμας μας ενδιαφερει να υπολογίσουμε ποια είναι η πιθανότητα το σακί να περιέχει ψεύτικα νομίσματα δεδομένου ότι ήρθε κορόνα και στις 3 ρίψεις. Αυτή η πιθανότητα εκφράζει το πόσο βέβαιοι είμαστε ότι επιλέξαμε το σωστό σακί στο τέλος.

Σύμφωνα με τον κανόνα του Bayes, το παραπάνω μπορεί να υπολογιστεί ως P[Ψ|K] = P[Ψ∩K] / P[K]

Μπορούμε να υπολογίσουμε P[K] = P[Ψ∩K] + P[Ψ'∩K] όπου P[Ψ∩K] = P[Ψ] * P[K|Ψ] = 0.5 * 0.343 = 0.1715 και P[Ψ'∩K] = P[Ψ'] * P[K|Ψ'] = 0.5 * 0.125 = 0.0625. Επομένως $P[\Psi|K] = \frac{0.1715}{0.1715 + 0.0625} = 0.7329059829$

Συμπέρασμα: Είσαι περίπου 73% βέβαιος πως έχεις επιλέξει το σωστό σακί.

Άσκηση2: Στην Αττική, το 95% των εγκλημάτων λαμβάνει χώρα κατά τη διάρκεια της νύχτας, ενώ το 54% από αυτά διαπράττεται εντός των ορίων της Αθήνας. Παρατηρούμε ότι μόλις το 2% των εγκλημάτων συμβαίνει την ημέρα μέσα στην Αθήνα.

Βάσει αυτών των πληροφοριών, ποιο είναι το ποσοστό των εγκλημάτων που συμβαίνει τη νύχτα εντός της Αθήνας; Και ποιο είναι το αντίστοιχο ποσοστό για τα εγκλήματα που διαπράττονται τη νύχτα εκτός της Αθήνας;

	Μέσα στην Αθήνα	Έξω από την Αθήνα
Βράδυ	a	b
Μέρα	c	d

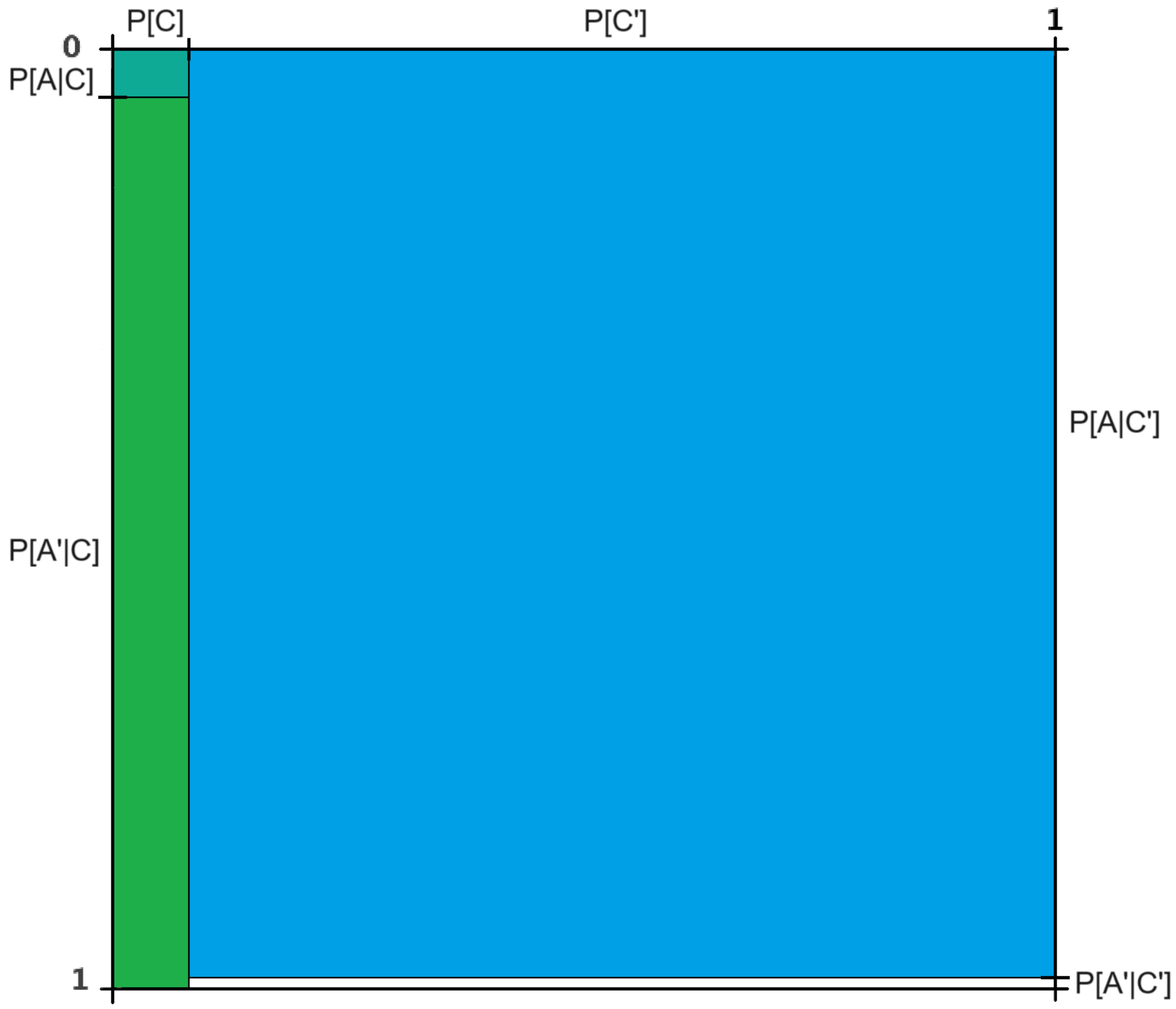
Λύση: Γνωρίζοντας πως το 95% λαμβάνει χώρα το βράδυ, θα πρέπει να ισχύει η σχέση a + b = 0.95. Εφόσον το 54% συμβαίνει μέσα στην Αθήνα θα πρέπει επίσης a + c = 0.54. Τέλος, εφόσον το 2% συμβαίνει την μέρα μέσα στην Αθήνα αυτό σημαίνει πως c = 0.02

Εφόσον c = 0.02 μπορούμε να υπολογίσουμε a = 0.54 - 0.02 = 0.52. Και οπότε b = 0.95 - 0.52 = 0.43. Τέλος, θα πρέπει a + b + c + d = 1 και οπότε d = 1 - 0.52 - 0.43 - 0.02 = 0.03.

Άσκηση3: Η πιθανότητα να έχει κάποιος COVID είναι 8%, ενώ η πιθανότητα να μην έχει είναι 92%. Κάνοντας test για να μάθεις αν έχεις COVID, το αποτέλεσμα εξαρτάται από το εάν έχεις ή όχι τον ιό. Συγκεκριμένα, εάν έχεις COVID, το test θα βγει θετικό με πιθανότητα 95% και αρνητικό με πιθανότητα 5%. Αντίθετα, εάν δεν έχεις τον ιό, υπάρχει 1% πιθανότητα το τεστ να δείξει εσφαλμένα θετικό και 99% πιθανότητα να είναι αρνητικό.

Βασιζόμενοι σε αυτά τα δεδομένα, ποια είναι η πιθανότητα να έχω πραγματικά COVID αν το test έχει δείξει ότι δεν έχω τον ιό;

Λύση: έστω A = “το test βγήκε αρνητικό” και C = “έχω covid”, θέλουμε να υπολογίσουμε το P[C|A] ενώ γνωρίζουμε πως P[C] = 0.08, P[C’] = 0.92, P[A|C] = 0.05, P[A’|C] = 0.95, P[A|C’] = 0.99 και P[A’|C’] = 0.01



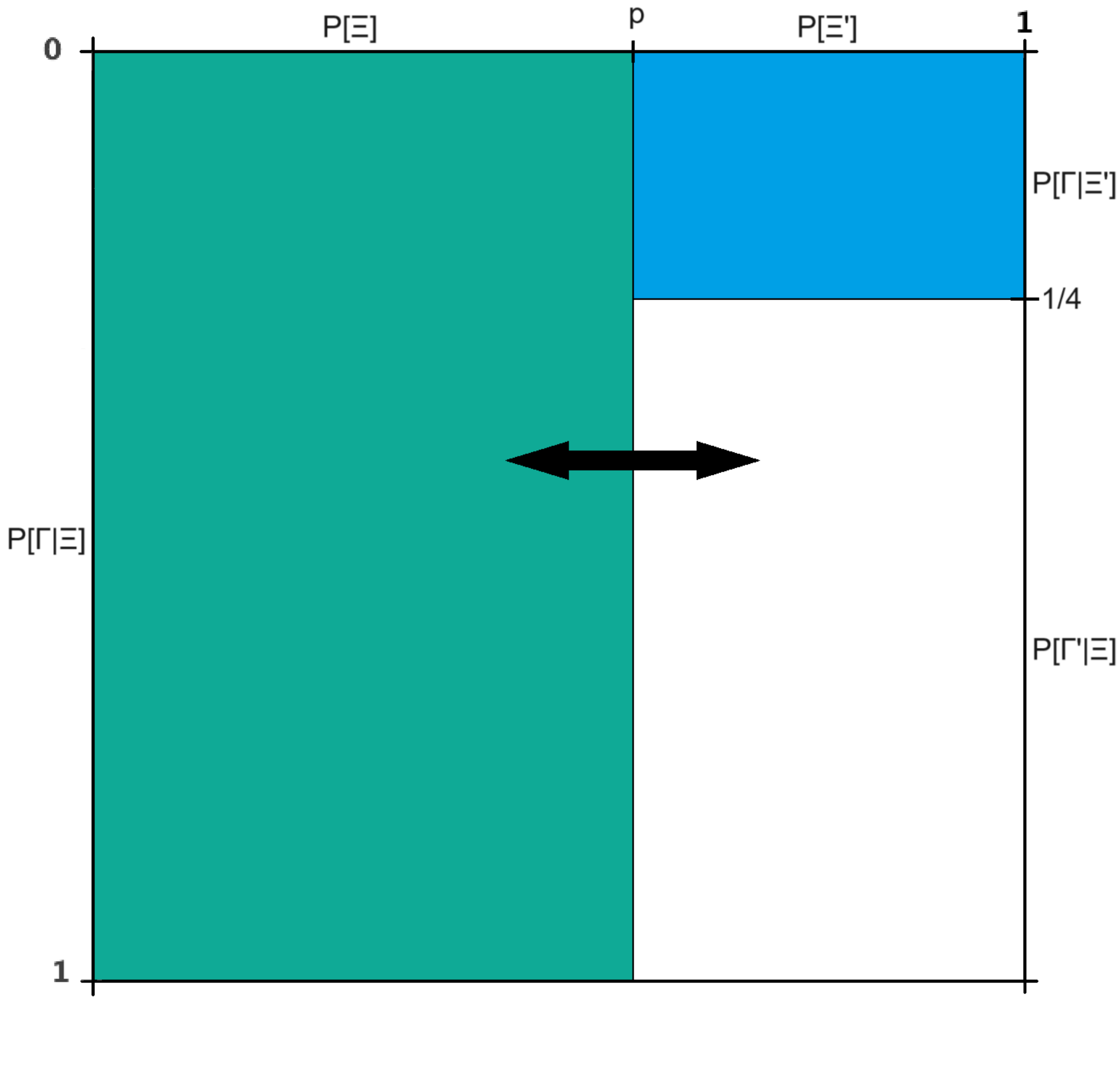
$$P[C|A] = \frac{P[C\cap A]}{P[A]} = \frac{P[C] * P[A|C]}{P[C] * P[A|C] + P[C'] * P[A|C']} = \frac{0.08 * 0.05}{0.08 * 0.05 + 0.92 * 0.99} = 0.00437254044$$

Συμπέρασμα: δεδομένου ότι το test βγήκε αρνητικό, η πιθανότητα να έχω COVID είναι περίπου ίση με 0.437%

Άσκηση4: Εάν ο Αχλιόπτας μας δώσει ένα διαγώνισμα με ερωτήσεις πολλαπλών επιλογών, όπου κάθε ερώτηση περιλαμβάνει τέσσερις επιλογές, και πιστεύει πως η πιθανότητα να γνωρίζουμε τη σωστή απάντηση σε ένα ερώτημα είναι p, ενώ είναι δεδομένο ότι αν γνωρίζουμε τη σωστή απάντηση θα την επιλέξουμε οπωσδήποτε, τότε ποια πιστεύει πως είναι η πιθανότητα να γνωρίζουμε τη σωστή απάντηση σε ένα ερώτημα αν απαντήσαμε σωστά?

Λύση: Έστω πως Ξ = “ξέρεις την σωστή απάντηση” και Γ = “έγραψες την σωστή απάντηση”. Από εκφώνηση γνωρίζουμε πως P[Ξ] = p, P[Ξ’] = 1 - p, P[Γ|Ξ] = 1 και αν θεωρήσουμε πως όταν δεν γνωρίζεις απαντάς τελείως τυχαία τότε εφόσον υπάρχουν 4 επιλογές P[Γ|Ξ’] = 1/4

Μπορούμε να υπολογίσουμε $P[\Xi|\Gamma] = \frac{P[\Xi\cap\Gamma]}{P[\Gamma]} = \frac{p}{p + (1-p) * (1/4)}$



Θεώρημα: Στις ασκήσεις 5 και 6, θα εφαρμόσω ένα θεμελιώδη θεώρημα των πιθανοτήτων: όταν επιλέγεται ένα τυχαίο στοιχείο από ένα σύνολο N, η πιθανότητα να επιλεγεί στοιχείο που να ανήκει σε ένα υποσύνολο M του N ισούται με |M| / |N|. ⭐

Η πιθανότητα κάθε στοιχείου να επιλεγεί είναι ίσα διαιρεμένη μεταξύ όλων των στοιχείων του N. Επομένως, κάθε στοιχείο έχει πιθανότητα ιση με 1/|N|. Το να επιλεγεί στοιχείο που να ανήκει στο M θα έχει πιθανότητα ίση με το άθροισμα των πιθανοτήτων όλων των στοιχείων του M, το οποίο αθροίζει σε |M| / |N|

Παράδειγμα: Αν 40 από τα 100 άτομα σε ένα δωμάτιο καπνίζουν, η πιθανότητα να διαλέξεις τυχαία έναν καπνιστή είναι 40/100 ή 0.4.

Άσκηση5: Έχουμε συνολικά 150 υπολογιστές. Από αυτούς, 120 συνδέονται στο διαδίκτυο και 45 διαθέτουν εκτυπωτή, ενώ 30 υπολογιστές διαθέτουν και τα δύο, δηλαδή τόσο εκτυπωτή όσο και σύνδεση στο διαδίκτυο. Όταν επιλέγουμε έναν υπολογιστή τυχαία και διαπιστώνουμε ότι διαθέτει εκτυπωτή, ποια είναι η πιθανότητα αυτός ο υπολογιστής να έχει επίσης σύνδεση στο διαδίκτυο?

Λύση: Μπορεί να λυθεί με θεώρημα του Bayes όπως και οι προηγούμενες, αλλά δεν είναι απαραίτητο να μπούμε στον κόπο. Εύκολα θα μπορούσες να συμπεράνεις πως η πιθανότητα είναι 30 / 45 καθώς από τους 45 υπολογιστές που έχουν εκτυπωτή οι 30 έχουν πρόσβαση στο διαδίκτυο.

Αν θες να το δοκιμάσεις με Bayes, όρισε προτάσεις I = “έχει ίντερνετ” και E = “έχει εκτυπωτή”, θα ισχύει πως $P[I] = 120/150$, $P[E] = 45/150$, $P[I \cap E] = 30/150$, $P[I|E] = P[I \cap E] / P[E] = (30/150) / (45/150) = 30 / 45$

Άσκηση6: Στο αμφιθέατρο είμαστε 100 άτομα, ο Αχλιόπτας μας προκαλεί να σηκώσουμε το χέρι εάν η μητέρα μας γεννήθηκε σε άρτιο έτος ή εάν έχουμε πειραματιστεί με ναρκωτικά στο παρελθόν. Συνολικά, 70 άτομα σηκώνουν το χέρι. Κάνοντας μια εκτίμηση, πόσοι πιστεύουμε πως έχουν κάνει ναρκωτικά?

Λύση: Έστω |T| = “άτομα μέσα στο αμφιθέατρο” = 100, |X| = “άτομα που σήκωσαν το χέρι τους” = 70, |M| = “άτομα που έχουν μητέρα που γεννήθηκε άρτιο έτος” και |N| = “άτομα που έχουν κάνει ναρκωτικά”. Τότες με βάση την εκφώνηση θα ισχύει πως $|X| = |M \cup N| = |M| + |N| - |M \cap N|$

Μπορούμε να υποθέσουμε ότι $|M| = |T| / 2 = 50$ καθώς το ότι βρισκόμαστε στο αμφιθέατρο δεν εξαρτάται από το πότε γεννήθηκε η μητέρα μας. Περιμένουμε πως περίπου οι μισοί θα έχουμε μητέρα που γεννήθηκε άρτιο έτος.

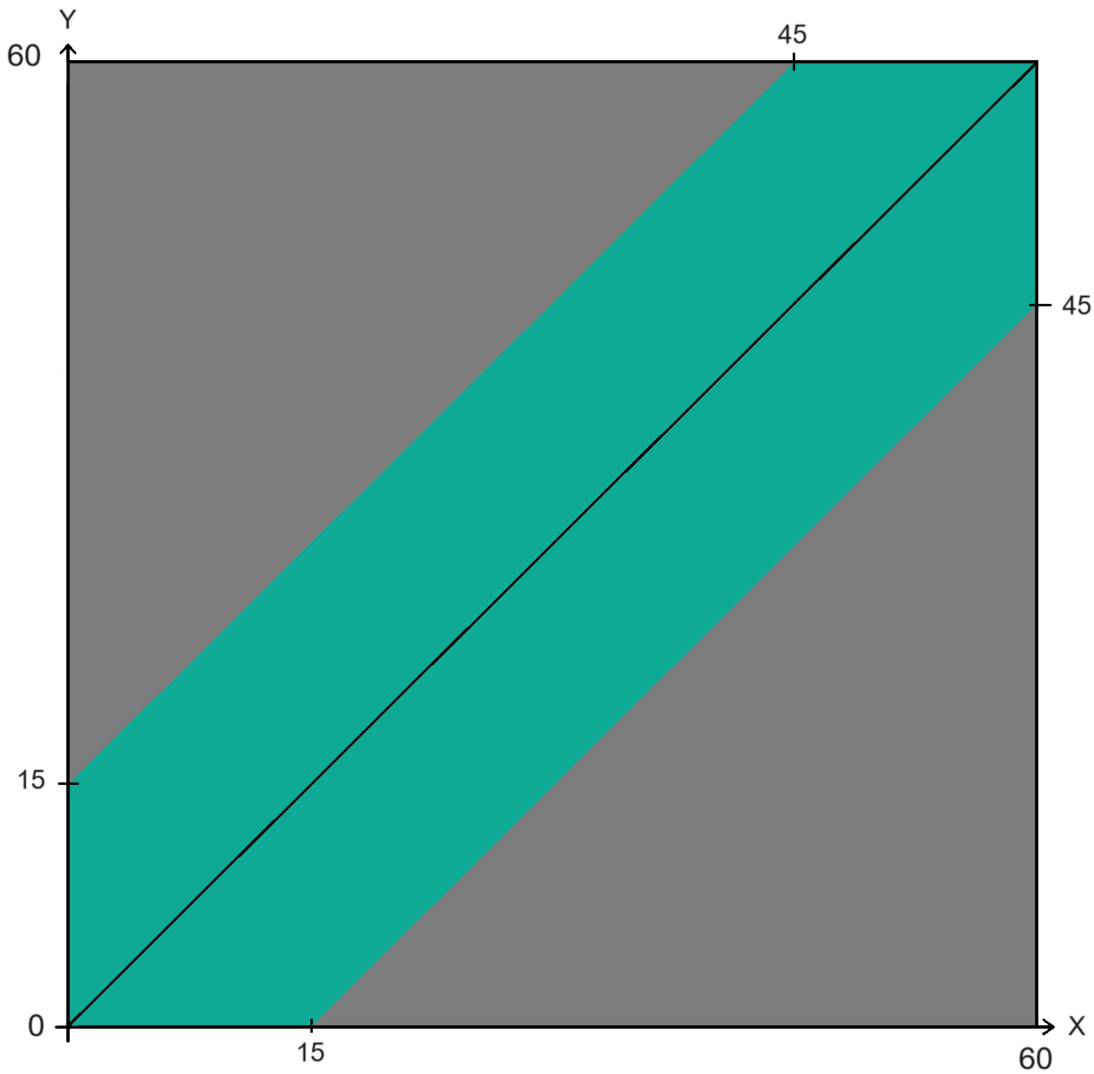
Ακόμα, μπορούμε να υποθέσουμε πως $|M \cap N| = |N| / 2$ καθώς το ότι κάποιος έχει κάνει ναρκωτικά δεν εξαρτάται από το πότε γεννήθηκε η μητέρα του. Περιμένουμε πως περίπου οι μισοί από όσους έχουν κάνει ναρκωτικά θα έχουν μητέρα που γεννήθηκε άρτιο έτος. Με βάση τα παραπάνω:

$|X| = |M| + |N| - |M \cap N| \Leftrightarrow 70 = 50 + |N| - |N| / 2 \Leftrightarrow |N| = 40$

Ερώτηση: Αν σήκωσε κάποιος το χέρι, ποια πιστεύουμε πως είναι η πιθανότητα να έχει κάνει ναρκωτικά?

Το παραπάνω εύκολα μπορεί να υπολογιστεί ως $|N| / |X| = 40/70 = 4/7$

Άσκηση7: Δύο φίλοι έχουν κανονίσει να συναντηθούν σε ένα μπαρ κάποια στιγμή μεταξύ 9 και 10 το βράδυ. Κάθε ένας από αυτούς θα φτάσει σε ένα τυχαίο χρονικό σημείο εντός αυτής της ώρας, ανεξάρτητα από τον άλλον. Η ερώτηση είναι: Ποια είναι η πιθανότητα οι δύο φίλοι να μην περιμένουν ο ένας τον άλλον για περισσότερο από 15 λεπτά;



Το γράφημα πάνω παρουσιάζει όλα τα πιθανά ζευγάρια ώρας άφιξης των δύο φίλων. Ο άξονας X απεικονίζει την ώρα άφιξης του πρώτου φίλου και ο άξονας Y του δεύτερου. Σε αυτό το σύστημα, το σύνολο όλων των σημείων (Ω) είναι άπειρο, καθώς οι ώρες άφιξης είναι συνεχείς και όχι διακριτές.

Η διαγώνιος του γραφήματος περιλαμβάνει όλα τα σημεία όπου οι δύο φίλοι φτάνουν ταυτόχρονα. Περιοχές κοντά στη διαγώνιο, όπου η απόλυτη διαφορά |X - Y| είναι μικρότερη ή ίση με 15, απεικονίζουν τις περιπτώσεις όπου κανένας από τους δύο δεν χρειάζεται να περιμένει τον άλλο για περισσότερο από 15 λεπτά.

Για να υπολογίσουμε την πιθανότητα να προκύψει ζεύγος το οποίο να ανήκει μέσα σε αυτή την ευρύτερη διαγώνια περιοχή, συγκρίνουμε το εμβαδόν της με το συνολικό εμβαδόν του επιπέδου (60 * 60). Αν το εμβαδόν αυτής της περιοχής αποτελούσε, για παράδειγμα, το 50% του επιπέδου, τότε η πιθανότητα για ένα τέτοιο σενάριο θα ήταν 0.5.

Συνολικά, το επίπεδο έχει εμβαδόν 60 * 60. Το εμβαδόν της διαγώνιας περιοχής μπορεί να βρεθεί αν αφαιρέσουμε από το 60 * 60 την περιοχή των δύο γκρι τριγώνων. Τα τρίγωνα είναι συμμετρικά, οπότε αρκεί να βρούμε το εμβαδόν του ενός, καθώς το άλλο θα έχει το ίδιο ακριβώς εμβαδόν.

Θυμήσου ότι το εμβαδόν ενός τριγώνου μπορεί να υπολογιστεί ως:

Εμβαδόν τριγώνου = $\frac{βάση * ύψος}{2}$

Το πάνω τρίγωνο είναι ένα ανάποδο ορθογώνιο τρίγωνο, με βάση ίση με 45 και ύψος ίσο με 45. Επομένως, το εμβαδόν του υπολογίζεται ότι ισούται με 45^2 / 2 = 1012.5.

Το ίδιο εμβαδόν θα έχει και το κάτω τρίγωνο. Έτσι, μαζί και τα δύο τρίγωνα θα έχουν εμβαδόν 1012.5 * 2 = 2025.

Για να βρούμε το εμβαδόν της διαγώνιας λωρίδας, αφαιρούμε το 2025 από το 60 * 60, και έτσι βρίσκουμε εμβαδόν 60^2 − 2025 =157560^2 = 1575.

Επομένως, η πιθανότητα θα είναι 1575 / (60 * 60) = 0.4375.

Εφαρμογή Θεωρίας Πιθανοτήτων σε Κοινωνικά Φαινόμενα:

Οι πεποιθήσεις μας μπορούν να εκφραστούν μαθηματικά ως πιθανότητες.

Για παράδειγμα, αν δω έναν άνδρα που φοράει κουκούλα, μπορεί να υπολογίσω μια πιθανότητα 70% ότι είναι κλέφτης, βασισμένος σε προηγούμενες εμπειρίες μου ή ιστορίες που έχω ακούσει.

Το φαινόμενο του "confirmation bias" αναφέρεται στην τάση μας όχι μόνο να αγνοούμε πληροφορίες που αντικρούουν τις ήδη καθιερωμένες πεποιθήσεις μας, αλλά και να αναζητούμε ενεργά πληροφορίες που επιβεβαιώνουν αυτές τις πεποιθήσεις. Αν, για παράδειγμα, διαβάσω μια μελέτη που δείχνει ότι οι κλέφτες συνήθως δεν φορούν κουκούλα (ώστε να μην φαίνονται απειλητικοί και να σε πλησιάζουν πιο εύκολα), το μυαλό μου μπορεί να υποβαθμίσει αυτή την πληροφορία επειδή δεν συνάδει με όσα ήδη πιστεύω. Επιπλέον, μπορεί να εστιάσω σε άλλες πληροφορίες ή ιστορίες που ενισχύουν την αρχική μου εκτίμηση. Αυτός ο συνδυασμός μπορεί να οδηγήσει σε υπερεκτίμηση της πιθανότητας που δίνω στον άντρα να είναι κλέφτης.

Το φαινόμενο αυτό είναι κοινό και συχνά συγχέεται με τον ρατσισμό, παρόλο που διαφέρει σημαντικά. Ο ρατσισμός εμφανίζεται όταν αρνούμαστε να εφαρμόσουμε τον κανόνα του Bayes για να αναθεωρήσουμε την πιθανότητα, παρά τα πρόσθετα δεδομένα που έχουμε στη διάθεσή μας. Για παράδειγμα, αν ο άνδρας με την κουκούλα βρίσκεται σε δημόσιο χώρο, φαίνεται φιλικός ή είναι κάποιος που γνωρίζω προσωπικά, αλλά εγώ εξακολουθώ να πιστεύω ότι υπάρχει 70% πιθανότητα να είναι κλέφτης μόνο λόγω της κουκούλας, αυτό θεωρείται ρατσισμός. Αυτή η άρνηση συχνά πηγάζει από συναισθήματα όπως του μίσους, της ζήλιας, του φόβου για το διαφορετικό, της ανάγκης να νιώθεις υπεράνω των άλλων, ή διάφορων προκαταλήψεων που πιστεύεις φανατικά.

Και τα δύο φαινόμενα είναι αρνητικά, αλλά το confirmation bias είναι μια διαδικασία που συχνά δεν ελέγχουμε συνειδητά και επομένως δεν μας καθιστά απαραίτητα κακούς ανθρώπους. Παρόλα αυτά, η αναγνώριση και η προσπάθεια περιορισμού του μπορούν να μας βοηθήσουν να λαμβάνουμε πιο αντικειμενικές και δίκαιες αποφάσεις.

Κάποια ενδιαφέρον βίντεο που προτείνω:

- Bayes theorem, the geometry of changing beliefs
- The quick proof of Bayes' theorem
- How to systematically approach truth - Bayes' rule
- What Is Confirmation Bias?

Ενότητα 3: Επανάληψη σε Απαρίθμηση

Η ακόλουθη ενότητα είναι μια σύντομη επανάληψη σε θεωρία απαρίθμησης

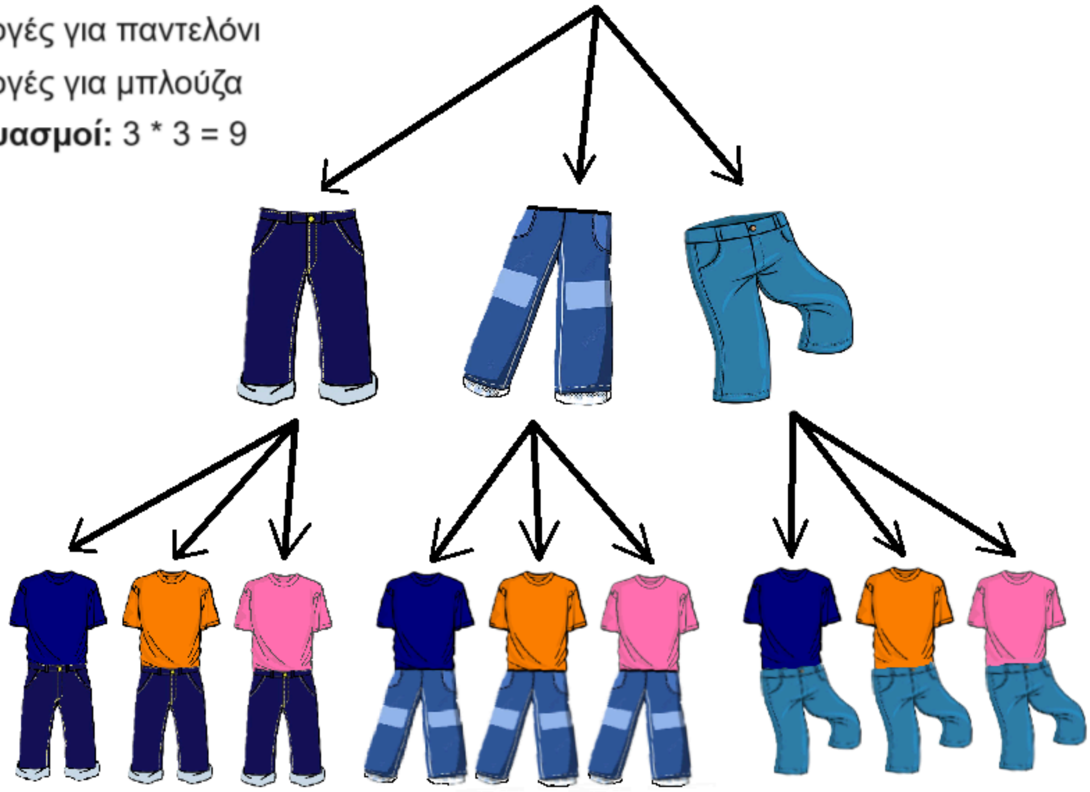
Αρχή Πολλαπλασιασμού: Αν έχουμε k βήματα και κάθε βήμα i έχει ni επιλογές, τότε συνολικά υπάρχουν n1 * n2 * ⋯ * nk δυνατοί συνδυασμοί⭐

Παράδειγμα: Επιλογή ρούχων όπου επιλέγουμε 1 από m παντελόνια και 1 από n μπλούζες, δίνει m * n δυνατούς συνδυασμούς.

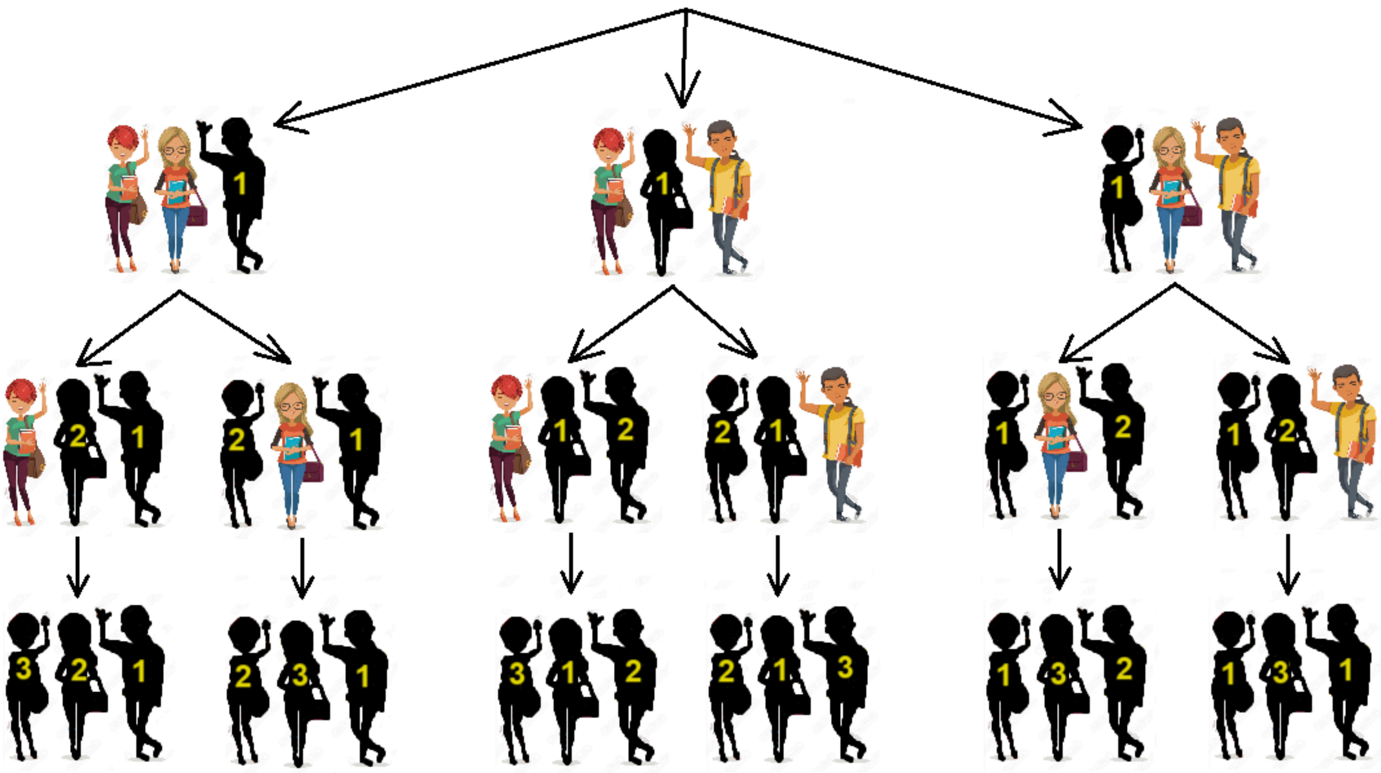
Βήμα1: 3 επιλογές για παντελόνι

Βήμα2: 3 επιλογές για μπλούζα

Δυνατοί συνδυασμοί: 3 * 3 = 9



Ερώτηση: Με πόσους τρόπους μπορώ να διατάξω 3 ανθρώπους?



Βήμα1: 3 επιλογές για το ποιος θα είναι πρώτος
Βήμα2: 2 επιλογές για το ποιος θα είναι δεύτερος
Βήμα3: 1 επιλογή για το ποιος θα είναι τρίτος
Δυνατοί συνδυασμοί: 3 * 2 * 1 = 6

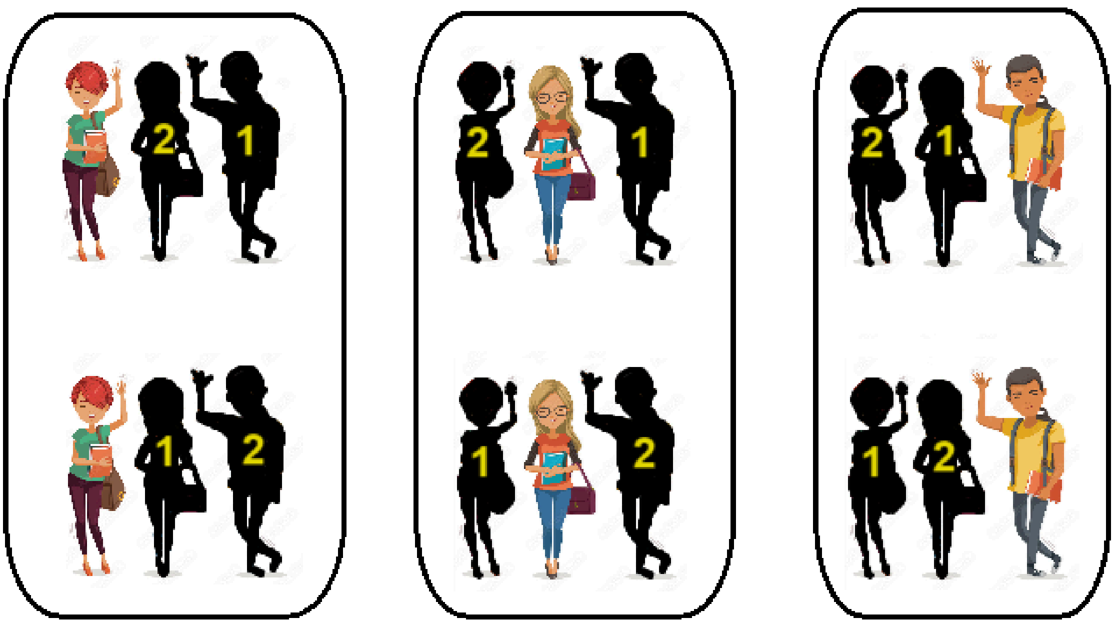
Γενικότερα, οι τρόποι με τους οποίους μπορείς να διατάξεις n στοιχεία είναι n! (το παραγοντικό του n) το οποίο ισούται με n * (n-1) * (n-2) * …. * 1 ⭐

Παρατήρηση: Οι τρόποι με τους οποίους μπορείς να διατάξεις n στοιχεία είναι ίσοι με τους τρόπους με τους οποίους μπορείς να επιλέξεις n στοιχεία όταν η σειρά των επιλεγμένων στοιχείων έχει σημασία.

Αν θέλαμε να μετρήσουμε τους τρόπους με τους οποίους μπορούμε να διατάξουμε 2 από τους 3 ανθρώπους, θα μπορούσαμε απλά να σταματήσουμε στο βήμα 2 αντί για το βήμα 3.

Γενικότερα, αν θέλουμε να βρούμε τον αριθμό των τρόπων να διατάξουμε k από n στοιχεία, αυτό μπορεί να υπολογιστεί ως $P(n,\ k)\ =\ \frac{n!}{(n-k)!}$⭐

Ο παραπάνω τύπος θα συμπεριλάβει στον υπολογισμό τους τρόπους με τους οποίους μπορούν να επιλεγούν τα ίδια στοιχεία αλλά με διαφορετική σειρά.



Κάθε τρόπος επιλογής 2 από 3 ανθρώπους μετράται δύο φορές, μία για κάθε διαφορετική διάταξη των δύο ατόμων. Αν θέλαμε να μετρήσουμε τρόπους να επιλέξουμε 2 άτομα, χωρις να μας ενδιαφέρει η σειρά, θα μπορούσαμε να διαιρέσουμε με τον αριθμό των τρόπων διάταξης 2 ατόμων, δηλαδή με 2.

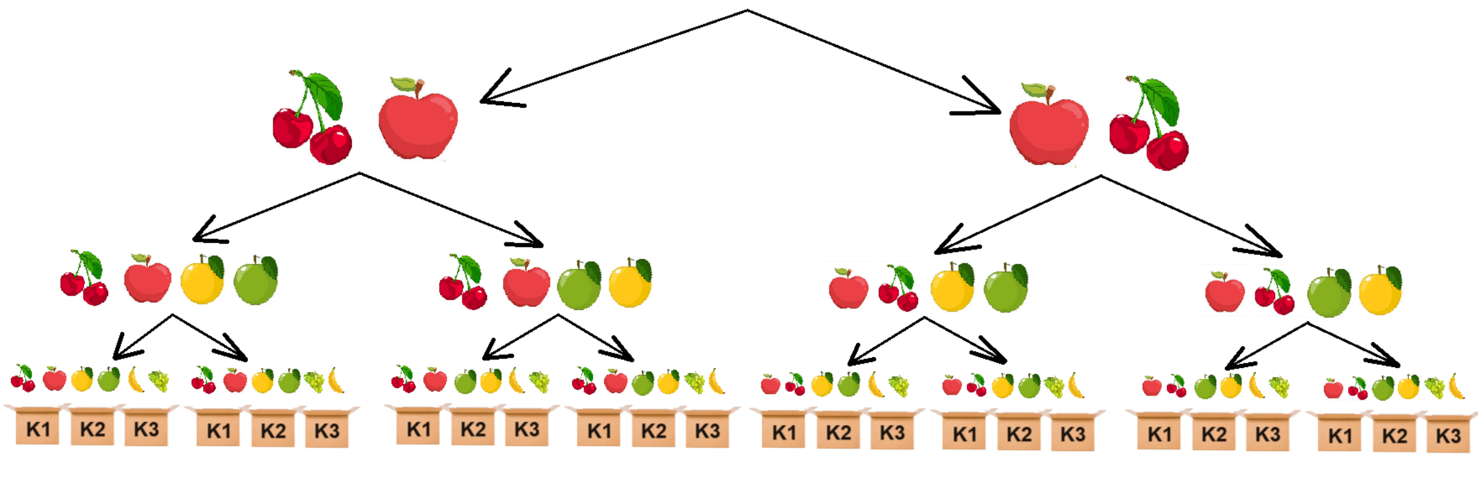
Γενικότερα, αν θέλουμε να βρούμε τον αριθμό των τρόπων να επιλέξουμε k από τα n στοιχεία, χωρίς να μας ενδιαφέρει η σειρά με την οποία θα τα επιλέξουμε, αυτό μπορεί να υπολογιστεί ως $C(n,\ k)\ =\ \frac{n!}{(n-k)!\,k!}$⭐

Κάναμε διαίρεση με k! διότι όπως και στο παράδειγμα με τους ανθρώπους, για κάθε ζευγάρι k στοιχείων ο αρχικός τύπος μετρούσε όλες τις δυνατές διατάξεις αυτών των στοιχείων, οι οποίες είναι συνολικά k! όπως έχουμε ήδη αποδείξει.

Ας δούμε τώρα πώς μπορούμε να υπολογίσουμε τους δυνατούς τρόπους να τοποθετήσουμε n στοιχεία σε m κουτιά, όπου το κάθε κουτί i μπορεί να χωρέσει ki στοιχεία και k1 + k2 + … + km = n (για να χωράνε όλα τα στοιχεία)

Μια προσέγγιση είναι να θεωρήσουμε τα κουτιά ως διατεταγμένα στη σειρά και να υπολογίσουμε όλους τους δυνατούς τρόπους διάταξης των n στοιχείων. Στην κάθε διάταξη, τα πρώτα k1 στοιχεία τοποθετούνται στο πρώτο κουτί, τα επόμενα k2 στοιχεία στο δεύτερο κουτί, και ούτω καθεξής:

Στο παράδειγμα πάνω βλέπουμε μια πιθανή διάταξη 6 φρούτων που μπαίνουν σε 3 κουτιά τα οποία χωράνε 2 φρούτα το καθένα. Θέλω να σκεφτούμε όμως το εξης, πόσες διατάξεις υπάρχουν που καταλήγουν στο ίδιο αποτέλεσμα? Για παράδειγμα, αν το μήλο ήταν αριστερά από το κεράσι, και πάλι με τον ίδιο τρόπο θα βάζαμε τα φρούτα στα κουτιά, δεν μας ενδιαφέρει η σειρά με την οποία τοποθετούνται εντός κάθε κουτιού.



Βήμα1: 2 πιθανές διατάξεις για τα στοιχεία του πρώτου κουτιου
Βήμα2: 2 πιθανές διατάξεις για τα στοιχεία του δεύτερου κουτιού
Βήμα3: 2 πιθανές διατάξεις για τα στοιχεία του τρίτου κουτιού
Δυνατοί συνδυασμοί: 2 * 2 * 2 = 8

Κάθε τρόπος τοποθέτησης των 6 φρούτων στα 3 κουτιά μετράται 8 φορές, λόγω του φαινομένου που έδειξα πάνω. Γι' αυτόν τον λόγο, από τον αριθμό των δυνατών μεταθέσεων των 6 φρούτων, που είναι 6!, πρέπει να διαιρέσουμε με 8. Το 8 προκύπτει από το γινόμενο του αριθμού των μεταθέσεων των στοιχείων του πρώτου κουτιού, επί τον αριθμό των μεταθέσεων των στοιχείων του δεύτερου κουτιού, επί τον αριθμό των μεταθέσεων των στοιχείων του τρίτου κουτιού. Τελικά υπολογίζουμε 6! / 8 = 90 τρόποι να μπουν στα κουτιά.

Γενικότερα, όταν θες να μετρήσεις τρόπους να βάλεις n στοιχεία σε m κουτιά όπου κάθε κουτί i χωράει ki στοιχεία, και k1 + k2 + … + km = n (ώστε να υπάρχει χώρος για όλα τα στοιχεία), αυτό μπορεί να υπολογιστεί με τον ακόλουθο τύπο: $C(n,\ k1,\ k2,\ k3,\ ...,\ km) \ = \ \frac{n!}{k1! * k2! * k3! * ... * km!}$ ⭐

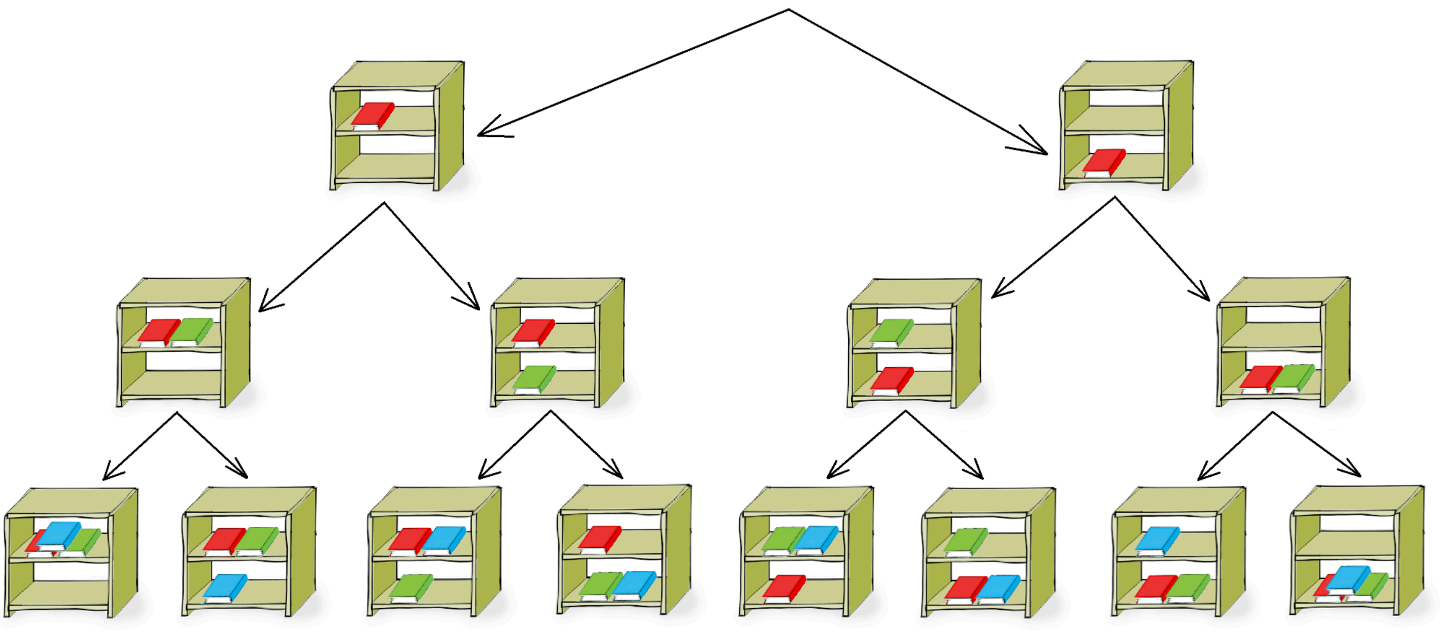
Αν θες να μετρήσεις τρόπους να βάλεις τα στοιχεία σε m κουτιά όπου όμως μπορεί να μην χωράνε όλα τα στοιχεία μέσα στα κουτιά, μια άλλη πιο χρονοβόρα, αλλά ισοδύναμη προσέγγιση, είναι η χρήση της αρχής του πολλαπλασιασμού. Έστω πως είχαμε 7 φρούτα για παράδειγμα:

Βήμα1: διάλεξε 2 από τα 7 φρούτα να μπουν στο πρώτο κουτί
Βήμα2: διάλεξε 2 από τα 5 φρούτα που μένουν να μπουν στο δεύτερο κουτί,
Βήμα3: διάλεξε 2 από τα 3 φρούτα που μένουν για να μπουν στο τρίτο κουτί.
Δυνατοί συνδυασμοί: C(7, 2) * C(5, 2) * C(3, 2)

Παρατήρηση: Η επιλογή k στοιχείων από n είναι ισοδύναμη με την τοποθέτηση k στοιχείων σε ένα κουτί k1 (που αντιπροσωπεύει τα επιλεγμένα στοιχεία) και n - k στοιχείων σε ένα άλλο κουτί k2 (που αντιπροσωπεύει τα στοιχεία που δεν επιλέχθηκαν). Επομένως, C(n, k) = C(n, k, n-k)

Από την πάνω παρατήρηση, θα ισχύει επίσης, πως το να βάλεις n στοιχεία σε m κουτιά όπου κάθε κουτί i χωράει ki στοιχεία, και k1 + k2 + … + km = r < n (δηλαδή δεν χωράνε όλα τα στοιχεία), αυτό είναι ισοδύναμο με το αν είχαμε και ένα επιπλέον φανταστικό κουτί το οποίο χωράει n - r στοιχεία και αποθηκεύει τα στοιχεία που μένουν εκτός των κουτιών k1, k2, …, km. Επομένως, τέτοιες περιπτώσεις μπορούν να υπολογιστούν ως $C(n,\ k1,\ k2,\ k3,\ ...,\ km,\ n\ -\ r)$ ⭐

Ερώτηση: με πόσους τρόπους μπορείς να βάλεις 3 βιβλία σε 2 ράφια? Όπου το κάθε ράφι θεωρούμε έχει χώρο για οποιοδήποτε αριθμό βιβλίων.



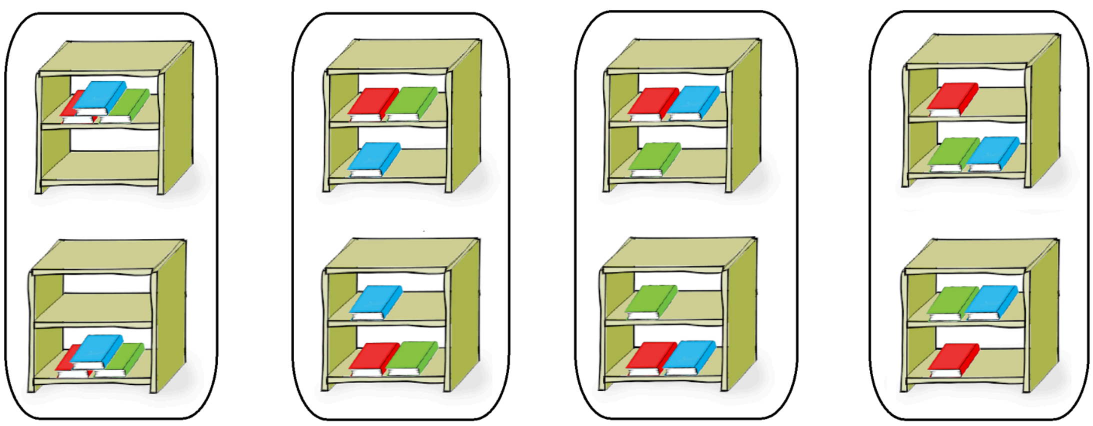
Βήμα1: το πρώτο βιβλίο έχει 2 ράφια στα οποία μπορεί να τοποθετηθεί
Βήμα2: το δεύτερο βιβλίο έχει 2 ράφια στα οποία μπορεί να τοποθετηθεί
Βήμα3: το τρίτο βιβλίο έχει 2 ράφια στα οποία μπορεί να τοποθετηθεί.
Δυνατοί συνδυασμοί: $2^3 = 8$

Γενικότερα, ο αριθμός των τρόπων να επιλέξουμε m στοιχεία από ένα σύνολο n στοιχείων, όταν επιτρέπεται η επανάληψη επιλογής κάποιου στοιχείου, και μας ενδιαφέρει η σειρά επιλογής, υπολογίζεται με τον τύπο n^m ⭐

— Από το σύνολο n = 2 ραφιών κάναμε m = 3 επιλογές

Παράδειγμα: Αν θέλουμε να παραβιάσουμε έναν κωδικό 5 ψηφίων χρησιμοποιώντας μέθοδο ωμής βίας, πόσες προσπάθειες θα χρειαστούν; Κάθε ψηφίο του κωδικού έχει 10 πιθανές επιλογές (από 0 έως 9). Επομένως, έχουμε να επιλέξουμε 5 ψηφία από ένα σύνολο 10 στοιχείων, με επανάληψη να επιτρέπεται. Εύκολα υπολογίζουμε $10^5 = 100.000$ προσπάθειες

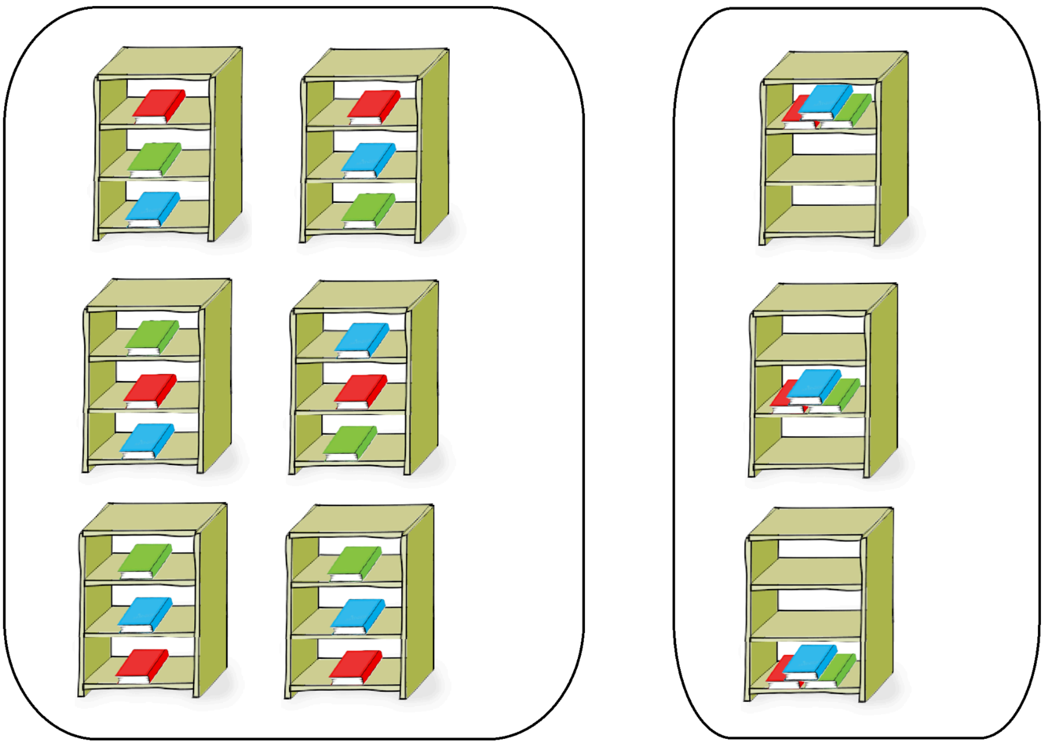
Ας επιστρέψουμε στο πρόβλημα με τα βιβλία. Αλλά αυτή την φορά μας ενδιαφέρει να υπολογίσουμε με πόσους τρόπους μπορούμε να χωρίσουμε 3 βιβλία σε 2 ομάδες (όπου επιτρέπονται κενές ομάδες). Η απάντηση σε αυτήν την ερώτηση δεν είναι 2^5 όπως ίσως λανθασμένα σκέφτηκες 🙅



Το μόνο που μας ενδιαφέρει είναι να είναι χωρισμένα σε 2 ομάδες, σε αντίθεση με πριν, το αν τα βιβλία είναι στο πάνω ή στο κάτω ράφι δεν έχει σημασία. Τώρα, λόγο περίπτωσης, στο πάνω παράδειγμα μπορούμε να διαιρέσουμε με 2, αλλά γενικότερα, δεν υπάρχει κλειστός τύπος για τέτοια προβλήματα, μόνο αναδρομικός.

Και θα βιαστεί κάποιος να πει, έστω πως έχουμε m ομάδες, και n στοιχεία, ο τύπος m^n / m!, δηλαδή το να διαιρούμε με τον αριθμό των δυνατών μεταθέσεων των ομάδων, δεν δουλεύει? Μην βγάζεις γρήγορα συμπεράσματα. Όπως είπα, στο πάνω παράδειγμα απλά έτυχε λόγο της περίπτωσης να ισχύει αυτό, γενικότερα όμως, επειδή μπορούν να μείνουν ομάδες χωρίς κανένα στοιχείο, δεν ισχύει ο τύπος, διαιρεί και για μεταθέσεις μεταξύ των κενών ομάδων, περιπτώσεις που δεν απαριθμίζει το m^n

Παράδειγμα: έστω έχουμε 3 ράφια αυτή την φορά. Οι περίπτωση όπου κάθε βιβλίο είναι μόνο του θα μετρηθεί 3! = 6 φορές, αλλά η περίπτωση όπου όλα τα βιβλία είναι μαζί θα μετρηθεί μόνο 3 φορές.



Πως θα μπορούσαμε λοιπόν να κάνουμε τον παραπάνω υπολογισμό? Για να είμαι ειλικρινής, **το παραπάνω δεν πρόκειται να το χρειαστείς για την εξέταση του συγκεκριμένου μαθήματος**, αλλά επειδή δεν θα μπορούσε μια ενότητα απαρίθμησης να είναι πλήρες χωρίς αυτό πρέπει να το αναφέρω.

Οι αριθμοί Στέρλινγκ δεύτερου είδους {n, k} αναφέρονται στον αριθμό των τρόπων με τους οποίους μπορούμε να χωρίσουμε n στοιχεία σε k ομάδες, χωρίς όμως να μπορεί να υπάρξει κενή ομάδα. Έχοντας ένα σύνολο A, μας βοηθούν να υπολογίσουμε με πόσους τρόπους μπορούμε να χωρίσουμε τα στοιχεία του συνόλου σε k ξένα μεταξύ τους υποσύνολα. Και ένα υποσύνολο πρέπει εξωρισμού να αποτελείται από τουλάχιστον 1 στοιχείο.

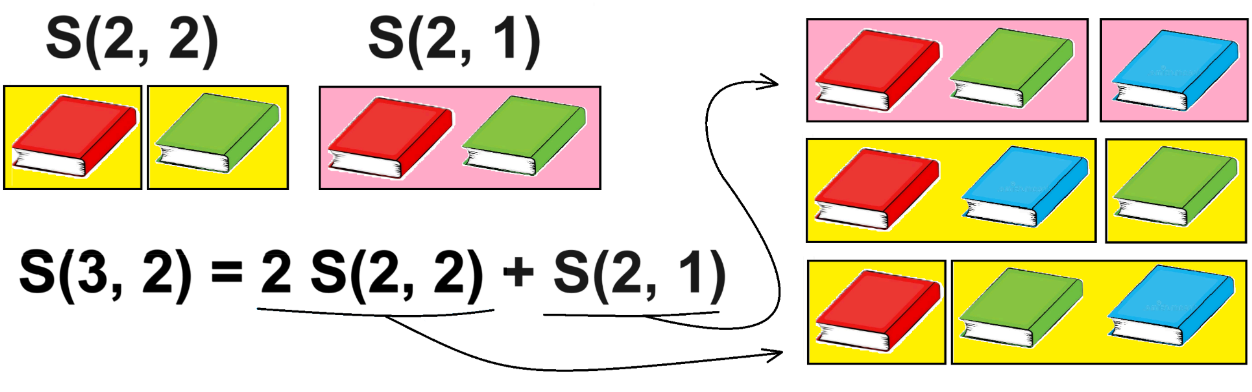
Οι αριθμοί υπολογίζονται από την σχέση **S(n, k) = k S(n-1, k) + S(n-1, k-1)** ⭐

Όπου **S(0, 0) = 1** (το κενό σύνολο μπορεί να χωριστεί σε 0 ομάδες με έναν μόνο τρόπο, δηλαδή χωρίς καμία ενέργεια). **S(n, 0) = 0 για n > 0** (δεν μπορούμε να χωρίσουμε ένα μη κενό σύνολο σε 0 ομάδες). **S(n, 1) = 1 για κάθε n > 0** (ένα σύνολο n στοιχείων μπορεί να χωριστεί σε 1 ομάδα με μόνο έναν τρόπο). Τέλος, **S(n, k) = 0 για k > n** (δεν μπορούμε να χωρίσουμε ένα σύνολο σε περισσότερες ομάδες από τα στοιχεία του).

$\left\{ {n \atop k} \right\}$	$k = 0$	1	2	3	4	5	6	7	8	b_n
$n = 0$	1									1
1	0	1								1
2	0	1	1							2
3	0	1	3	1						5
4	0	1	7	6	1					15
5	0	1	15	25	10	1				52
6	0	1	31	90	65	15	1			203
7	0	1	63	301	350	140	21	1		877
8	0	1	127	966	1701	1050	266	28	1	4140

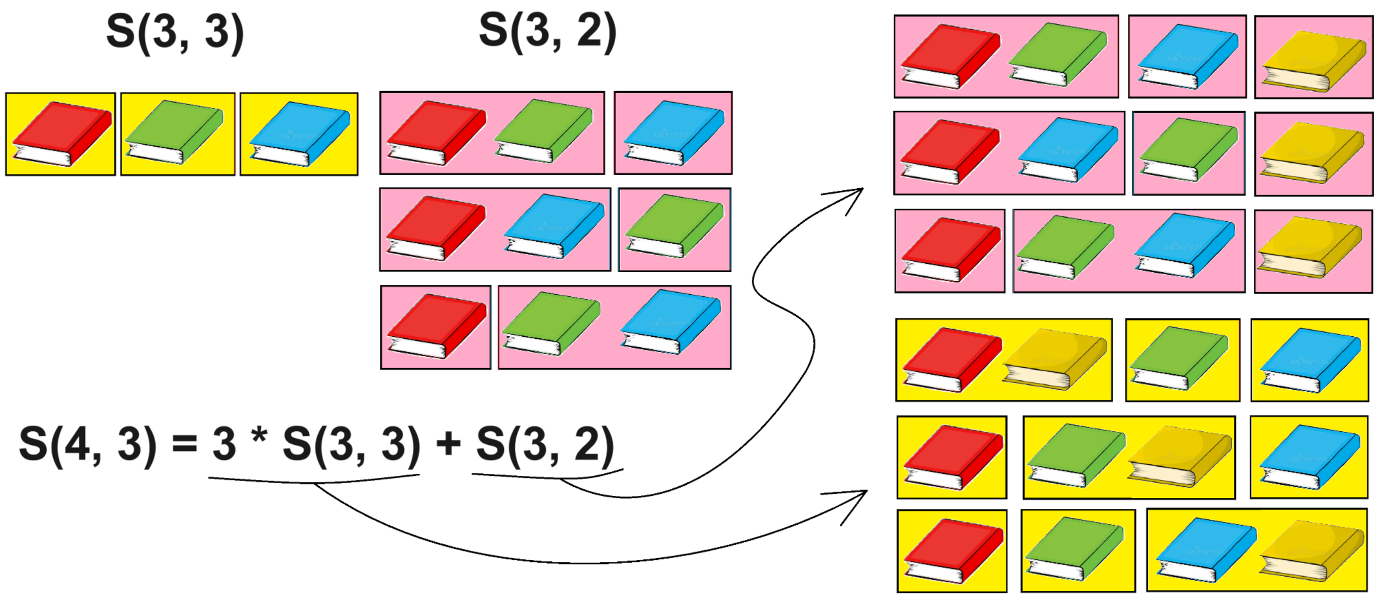
Η ακολουθία **Bn** της εικόνας είναι ίση με τον αριθμό των τρόπων να χωρίσεις n στοιχεία σε 1 ομάδα + τον αριθμό των τρόπων να χωρίσεις n στοιχεία σε 2 ομάδες + τον αριθμό των τρόπων να χωρίσεις n στοιχεία σε 3 ομάδες, κτλπ.

Παράδειγμα: Για να κατανοήσουμε τον λόγο που λειτουργεί η αναδρομή ας δούμε πως μπορούμε να βρούμε το S(3, 2) γνωρίζοντας τα S(2, 2) και S(2, 1)



Η εξήγηση του τύπου είναι η εξής: Πρώτα, αγνοούμε ένα από τα στοιχεία, όπως το μπλε βιβλίο στο παράδειγμα, και υπολογίζουμε τους τρόπους να χωρίσουμε τα υπόλοιπα n-1 στοιχεία σε k ομάδες, δίνοντας S(n-1, k) τρόπους. Στη συνέχεια, ενσωματώνουμε το στοιχείο που είχαμε αγνοήσει. Για κάθε τρόπο διαχωρισμού των n-1 στοιχείων σε k ομάδες, το νέο στοιχείο μπορεί να προστεθεί σε οποιαδήποτε από τις k ομάδες, δίνοντας k * S(n-1, k) πιθανούς τρόπους. Επιπλέον, πρέπει να μετρήσουμε και την περίπτωση όπου το νέο στοιχείο είναι μόνο του σε μία ομάδα, αφήνοντας τα υπόλοιπα n-1 στοιχεία να χωριστούν σε k-1 ομάδες, για αυτό προσθέτουμε και τον όρο S(n-1, k-1)

Παράδειγμα: ας δούμε πως υπολογίζεται το S(4, 3) γνωρίζοντας πως S(3, 2) = 3 (το υπολογίσαμε πριν), και S(3, 3) = 1 (κάθε στοιχείο δική του ομάδα).



Επιστροφή στο αρχικό πρόβλημα: Με πόσους τρόπους μπορείς να χωρίσεις 3 βιβλία σε 3 ομάδες, αν μπορούν να υπάρχουν κενές ομάδες?

Τα S(n, k) που μάθαμε μετράνε τους τρόπους να χωρίσουμε στοιχεία σε ομάδες όπου κάθε ομάδα έχει τουλάχιστον 1 στοιχείο. Μπορούμε όμως να κάνουμε το εξής. Μετράμε τους τρόπους όπου χωρίζουμε τα βιβλία σε μία μόνο ομάδα, τους τρόπους όπου τα χωρίζουμε σε 2 ομάδες, τους τρόπους όπου τα χωρίζουμε σε 3 ομάδες, και αθροίζουμε. Αυτό θα ισούται με S(3, 1) + S(3, 2) + S(3, 3) = 1+3+1 = 5 άρα υπάρχουν 5 τρόποι να χωρίσεις τα βιβλία!

Κλείνοντας αυτή την ενότητα, μένει ακόμα ένα είδος απαρίθμησης που μπορεί να συναντήσεις. Επιλογή με επανάληψη, χωρίς να μας ενδιαφέρει η σειρά.

Παράδειγμα: Έχεις κάποιο μπάτζετ στην διάθεσή σου με το οποίο ξέρεις ότι μπορείς να αγοράσεις μέχρι 30 βιβλία. Εσένα σου αρέσουν 3 κατηγορίες βιβλίων, επιστημονικά, περιπέτειας και μυστηρίου. Και σκέφτεσαι, πόσα βιβλία κάθε κατηγορίας να πάρεις? Θα μπορούσες για παράδειγμα, να πάρεις 30 βιβλία μυστηρίου και κανένα άλλης κατηγορίας. Η να πάρεις 20 επιστημονικά, 5 περιπέτειας και 5 μυστηρίου. Η οτιδήποτε άλλο συνδυασμό θες. Μας ενδιαφέρει να μετρήσουμε πόσοι δυνατοί συνδυασμοί υπάρχουν.

Ίσως κάποιος σκεφτεί ότι, έχεις 3 επιλογές κατηγορίας για το πρώτο βιβλίο, 3 επιλογές για το δεύτερο, 3 για το τρίτο, κτλπ. Οπότε, 3^30 δυνατοί συνδυασμοί. Αλλά ο παραπάνω συλλογισμός είναι λάθος 🙅

Θα ίσχυε αυτό αν σε ενδιέφερε η σειρά με την οποία θα αγοράσεις τα βιβλία. Με τον πάνω τύπο, το να πάρεις πρώτα 1 επιστημονικό βιβλίο και μετά 29 μυστηρίου το μετράει ως διαφορετικό συνδυασμό από το αν πάρεις πρώτα 29 μυστηρίου και μετά 1 επιστημονικό. Εμάς όμως δεν μας ενδιαφέρει αυτό, μας ενδιαφέρει μόνο πόσα μυστηρίου και πόσα επιστημονικά θα πάρεις.

Προηγουμένως με τα 3 ράφια, είχαμε το κόκκινο, το πράσινο και το μπλε βιβλίο, αλλά τώρα τα στοιχεία μας δεν μπορεις να τα ξεχωρίσεις μεταξύ τους. Τα βλέπεις απλά ως ποσότητες, όχι ως ξεχωριστές ατομικές οντότητες. Η κατηγορία μυστηρίου έχει 29 βιβλία και η κατηγορία επιστήμης έχει 1 βιβλίο.

Ένας τρόπος να προσεγγίσεις το πρόβλημα είναι ο εξής: Φαντάσου 32 κενές θέσεις. Επιλέγοντας 2 από αυτές, στις θέσεις αριστερά της πρώτης επιλεγμένης θα τοποθετούμε επιστημονικά βιβλία, στο διάστημα μεταξύ των δύο επιλεγμένων βιβλία περιπέτειας, και στις θέσεις δεξιά της δεύτερης επιλεγμένης βιβλία μυστηρίου. Στις επιλεγμένες βάζω **|** και στις άλλες βάζω *

8 επιστημονικά **20 περιπέτειας** **2 μυστηρίου**

* * * * * * * * | * | * *

Και τα 30 επιστημονικά

* | |

Και τα 30 περιπέτειας

| * |

Και τα 30 μυστηρίου

| | *

Με τον πάνω τρόπο μπορεί να αναπαρασταθεί οποιοσδήποτε δυνατός συνδυασμούς για το πόσα βιβλία θα έχει η κάθε κατηγορίας. Επομένως, το πρόβλημα ανάγεται στον υπολογισμό των τρόπων με των οποίων μπορείς να διαλέξεις 2 θέσεις από ένα σύνολο 32 θέσεων, που ισούται με C(32, 2)

Γενικότερα, ο αριθμός των τρόπων να επιλέξουμε m στοιχεία από ένα σύνολο n στοιχείων, όταν επιτρέπεται η επανάληψη επιλογής κάποιου στοιχείου, και δεν μας ενδιαφέρει η σειρά επιλογής, υπολογίζεται ως **C(m+n-1, n-1)** ⭐
— Από το σύνολο 3 κατηγοριών κάναμε 30 επιλογές χωρίς να έχει σημασία η σειρά

Ανακεφαλαίωση: Παραθέτω μια λίστα όλων των τύπων που έχουμε δει:

1. Τρόποι για να κάνεις n βήματα με k1, …, kn επιλογές το καθένα: **n1 * … nk**

2. Τρόποι για να διατάξεις n στοιχεία: **n!**

3. Τρόποι για να επιλέξεις m από n στοιχεία όταν σε ενδιαφέρει η σειρά επιλογής: **P(n, m) =** $\frac{n!}{(n-m)!}$

4. Τρόποι για να επιλέξεις m από n στοιχεία όταν δεν σε ενδιαφέρει η σειρά επιλογής: **C(n, m) =** $\frac{n!}{(n-m)!\,m!}$

5. Τρόποι για να βάλεις n στοιχεία σε m κουτιά τα οποία χωράνε k1, k2, …, km στοιχεία το καθένα: **C(n, k1, k2, …, km) =** $\frac{n!}{k1!*k2!*\ldots*km!}$

6. Τρόποι για να επιλέξεις m στοιχεία από n στοιχεία όταν επιτρέπεται η επαναληπτική επιλογή ίδιου στοιχείου και σε ενδιαφέρει η σειρά επιλογής των στοιχείων: **n^m**

7. Τρόποι για να χωρίσεις n στοιχεία σε k μη κενές ομάδες: **S(n, k) = k S(n-1, k) + S(n-1, k-1)**

8. Τρόποι για να επιλέξεις m στοιχεία από n στοιχεία όταν επιτρέπεται η επαναληπτική επιλογή ίδιου στοιχείου και δεν σε ενδιαφέρει η σειρά επιλογής των στοιχείων: **C(m + n - 1, n - 1)**

Η απαρίθμηση είναι μια διαδικασία που απαιτεί φαντασία. Συχνά, σε διάφορα προβλήματα, μπορεί να χρειαστεί να υπολογίσεις κάτι που είναι ισοδύναμο με την επιλογή k από n στοιχεία. Αλλά αυτό δεν είναι πάντα άμεσα προφανές. Απαιτεί λίγη δημιουργικότητα για να κάνεις τη σύνδεση στο μυαλό σου. Στην εξέταση όμως μπαίνουν πάντα παρόμοιες ασκήσεις, οπότε μην αγχώνεσαι.

Οι παραπάνω είναι ασκήσεις που είδαμε σε διαλέξεις του μαθήματος ⚔️

Άσκηση1: Από μία τράπουλα με 52 φυλλα, τραβώ τυχαία 10 φύλλα. Υπάρχουν συνολικά 4 άσοι στην τράπουλα. Ποια είναι η πιθανότητα να μην τραβήξω κανέναν άσο;

Υπάρχουν C(52, 10) τρόποι να επιλέξεις 10 φύλλα από την τράπουλα. Για να βρεις πόσοι από αυτούς τους τρόπους δεν περιλαμβάνουν κανέναν άσο, σκέψου το εξής: αφαιρώντας τους 4 άσους από την τράπουλα, απομένουν 48 φύλλα. Ο αριθμός των τρόπων να επιλέξεις 10 φύλλα από τα 48 αυτά φύλλα είναι C(48, 10). Αυτός ο υπολογισμός καλύπτει όλους τους τρόπους με τους οποίους μπορείς να αποκτήσεις 10 φύλλα χωρίς να έχεις επιλέξει κανέναν άσο. Επομένως η πιθανότητα είναι C(48, 10) / C(52, 10) = 0.413

Άσκηση2: Από μία τράπουλα με 52 φυλλα, τραβώ τυχαία 10 φύλλα. Υπάρχουν συνολικά 4 άσοι στην τράπουλα. Ποια είναι η πιθανότητα να τραβήξω το πολύ 3 άσους;

Επειδή υπάρχουν 4 άσοι, οι μόνοι τρόποι με τους οποίους τραβάς περισσότερους από 3 είναι αν τραβήξεις και τους 4. Μπορούμε να υπολογίσουμε την πιθανότητα να τραβήξεις 4 άσους, και να αφαιρέσουμε από 1 για να βρούμε πόση πιθανότητα αντιστοιχεί στα υπόλοιπα σενάρια όπου τραβάς λιγότερους από 4 άσους. Για να υπολογίσεις με πόσους τρόπους μπορεις να τραβήξεις 4 άσους, σκέψου το εξής: έχοντας τραβήξει 4 άσους, στην τράπουλα μένουν 48 φύλλα, ο αριθμός των τρόπων να επιλέξεις 6 ακόμα φύλλα από αυτά είναι C(48, 6). Αυτός ο υπολογισμός καλύπτει όλους τους τρόπους με τους οποίους μπορείς να αποκτήσεις 10 φύλλα με 4 άσους. Οπότες C(48, 6) / C(52, 10) = 0.0007757 θα είναι η πιθανότητα να τραβήξεις 4 άσους, και 1 - 0.0007757 = 0.9992243 να τραβήξεις το πολύ 3 άσους.

Άσκηση3: Σε ένα παιχνίδι πόκερ, ένας παίκτης λαμβάνει 5 φύλλα από μια στάνταρ τράπουλα που αποτελείται από 52 φύλλα. Η τράπουλα περιλαμβάνει 4 διαφορετικά σχέδια ή "χρώματα": κούπες, σπαθιά, καρό και μπαστούνια. Κάθε χρώμα έχει 13 φύλλα, τα οποία περιλαμβάνουν τους αριθμούς από 2 έως 10, τον Βαλέ, τη Δάμα, τον Ρήγα και τον Άσσο. Ποια είναι η πιθανότητα ο παίκτης να έχει 4 φύλλα με την ίδια τιμή (δηλαδή, τέσσερα Βαλέδες, τέσσερις Δάμες, κ.λπ.) στο χέρι του;

Υπάρχουν C(52, 5) τρόποι να τραβήξεις 5 φύλλα από την τράπουλα. Για κάθε τετράδα μιας τιμής, υπάρχουν 52 - 4 = 48 πιθανές επιλογές για το ποιο θα είναι το πέμπτο φύλλο. Δηλαδή, 48 τρόποι να λάβεις 4 άσους, 48 τρόποι να λάβεις 4 Δάμες, κτλπ. Επειδή λαμβάνουμε μόνο 5 φύλλα, τα παραπάνω σύνολα είναι ξένα μεταξύ τους. Η τρόποι να λάβεις 4 άσους δεν περιέχουν μέσα τρόπους όπου λαμβάνεις 4 Δάμες. Αν λάμβανες πάνω από 5 φύλλα τότε θα ήθελε πιο προσεκτική ανάλυση. Στην περίπτωση μας, εφόσον τα σύνολα είναι ξένα, μπορούμε να τα προσθεσουμε, οπότε υπάρχουν 13 * 48 τρόποι ο παίκτης να λάβει 4 φύλλα ίδιας τιμής, τραβόντας 5 φύλλα από την τράπουλα. Η πιθανότητα μπορεί να υπολογιστεί ως 13 * 48 / C(52, 5) = 0.0002401

Άσκηση4: Σε ένα παιχνίδι πόκερ, όπου η τράπουλα περιλαμβάνει 52 φύλλα, τραβάς 5 φύλλα. Εάν από τα φύλλα αυτά, τρία έχουν την ίδια τιμή και τα άλλα δύο έχουν μια διαφορετική τιμή αλλά μεταξύ τους ίδια, τότε αυτό ονομάζεται Full House. Ποια είναι η πιθανότητα να έχεις τραβήξει Full House;

Βήμα1: 13 επιλογές για το ποια τιμή θα κάνει την τριάδα
Βήμα2: C(4, 3) επιλογές για το ποια φύλλα της τιμής θα κάνουν την τριάδα
Βήμα3: 12 επιλογές για το ποια τιμή θα κάνει την δυάδα
Βήμα4: C(4, 2) επιλογές για το ποια φύλλα της τιμής θα κάνουν την δυάδα
Δυνατοί συνδυασμοί: 13 * C(4, 3) * 12 * C(4, 2) = 3744

Συνολικά υπάρχουν C(52, 5) τρόποι να τραβήξεις 5 φύλλα και επομένως η πιθανότητα υπολογίζεται ως 3744 / C(52, 5) = 0.00144

Άσκηση5: Έξι φίλοι σχεδιάζουν να συναντηθούν στο ξενοδοχείο "Ακρόπολης" στην Αθήνα. Παρ' όλα αυτά, υπάρχουν τέσσερα διαφορετικά ξενοδοχεία με το ίδιο όνομα στην πόλη. Κάθε φίλος επιλέγει τυχαία ένα από τα τέσσερα ξενοδοχεία για να πάει. Ποια είναι η πιθανότητα να συναντηθούν σε ομάδες των 2 στα ξενοδοχεία;

Βημα1: C(4, 3) επιλογες για το σε ποια ξενοδοχεία θα συναντηθούν
Βήμα2: C(6, 2, 2, 2) τρόποι να τους βάλεις στα 3 ξενοδοχεία σε ομάδες των 2
Δυνατοι συνδυασμοί: C(4, 3) * C(6, 2, 2, 2)

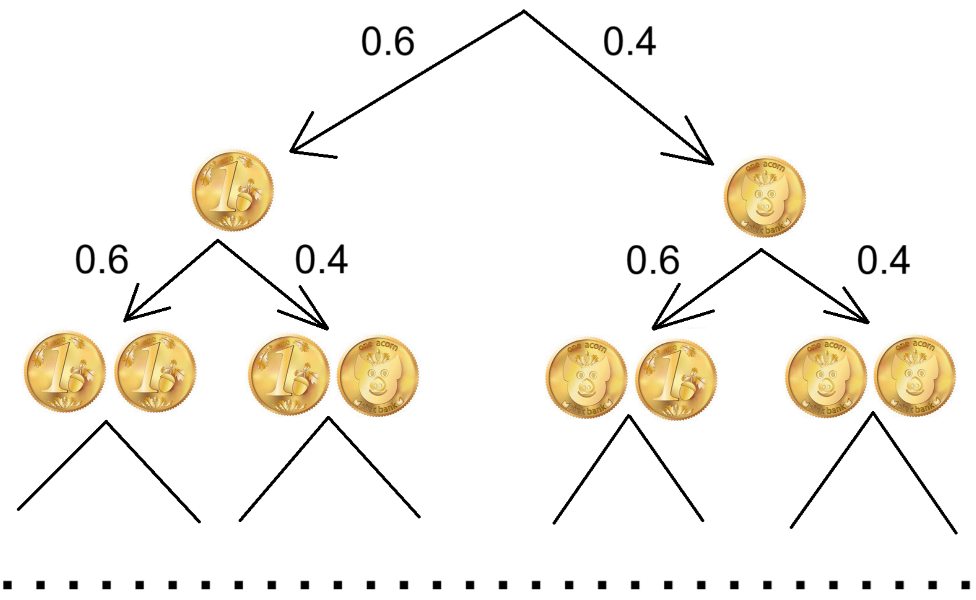
Συνολικά υπάρχουν 4^6 δυνατοι τρόποι να πάνε στα ξενοδοχεία αφού ο κάθε ένας φίλος έχει 4 επιλογες για το σε ποιο ξενοδοχειο θα πάει. Επομένως η πιθανότητα υπολογίζεται ως C(4, 3) * C(6, 2, 2, 2) / 4^6

Άσκηση6: Έξι φίλοι σχεδιάζουν να συναντηθούν στο ξενοδοχείο "Ακρόπολης" στην Αθήνα. Παρ' όλα αυτά, υπάρχουν τέσσερα διαφορετικά ξενοδοχεία με το ίδιο όνομα στην πόλη. Κάθε φίλος επιλέγει τυχαία ένα από τα τέσσερα ξενοδοχεία για να πάει. Ποια είναι η πιθανότητα να συναντηθούν 2 φίλοι σε ένα ξενοδοχείο ο καθένας ενώ οι υπόλοιποι 4 να πάνε 2 και 2 στα υπόλοιπα;

Βήμα1: C(6, 2) επιλογες για το ποιοι θα συναντηθούν σε ένα ο καθένας
Βήμα2: C(4, 2) για το ποια θα είναι τα δύο αυτά ξενοδοχεία
Βήμα3: 2 πιθανές διατάξεις με τις οποίες μπορούν να πάνε σε αυτά
Βήμα4: C(4, 2, 2) τρόποι να πάνε οι 4 στα άλλα ξενοδοχεία σε ομάδες των 2
Δυνατοι συνδυασμοί: C(6, 2) * C(4, 2) * 2 * C(4, 2, 2)

Συνολικά υπάρχουν 4^6 δυνατοι τρόποι να πάνε στα ξενοδοχεία αφού ο κάθε ένας φίλος έχει 4 επιλογες για το σε ποιο ξενοδοχειο θα πάει. Επομένως η πιθανότητα υπολογίζεται ως C(6, 2) * C(4, 2) * 2 * C(4, 2, 2) / 4^6

Άσκηση 7: Έστω ότι ρίχνουμε ένα κέρμα 10 φορές, όπου η πιθανότητα να έρθει κορόνα είναι 0,4 και η πιθανότητα να έρθει γράμματα είναι 0,6. Πόση είναι η πιθανότητα να έρθει κορόνα ακριβώς 6 φορές;



Η πιθανότητα ενός συγκεκριμένου φύλλου στο δέντρο είναι το γινόμενο των πιθανοτήτων κατά μήκος της διαδρομής που οδηγεί σε αυτό το φύλλο. Για παράδειγμα, η πιθανότητα να έρθει κορώνα και στις 10 ρίψεις είναι 0.4^10. Για ευκολία μπορούμε να αναπαριστανουμε τα πιθανά αποτελέσματα με συμβολοσειρές, όπου το 10 κορώνες θα αντιστοιχούσε στην “KKKKKKKKKK”.

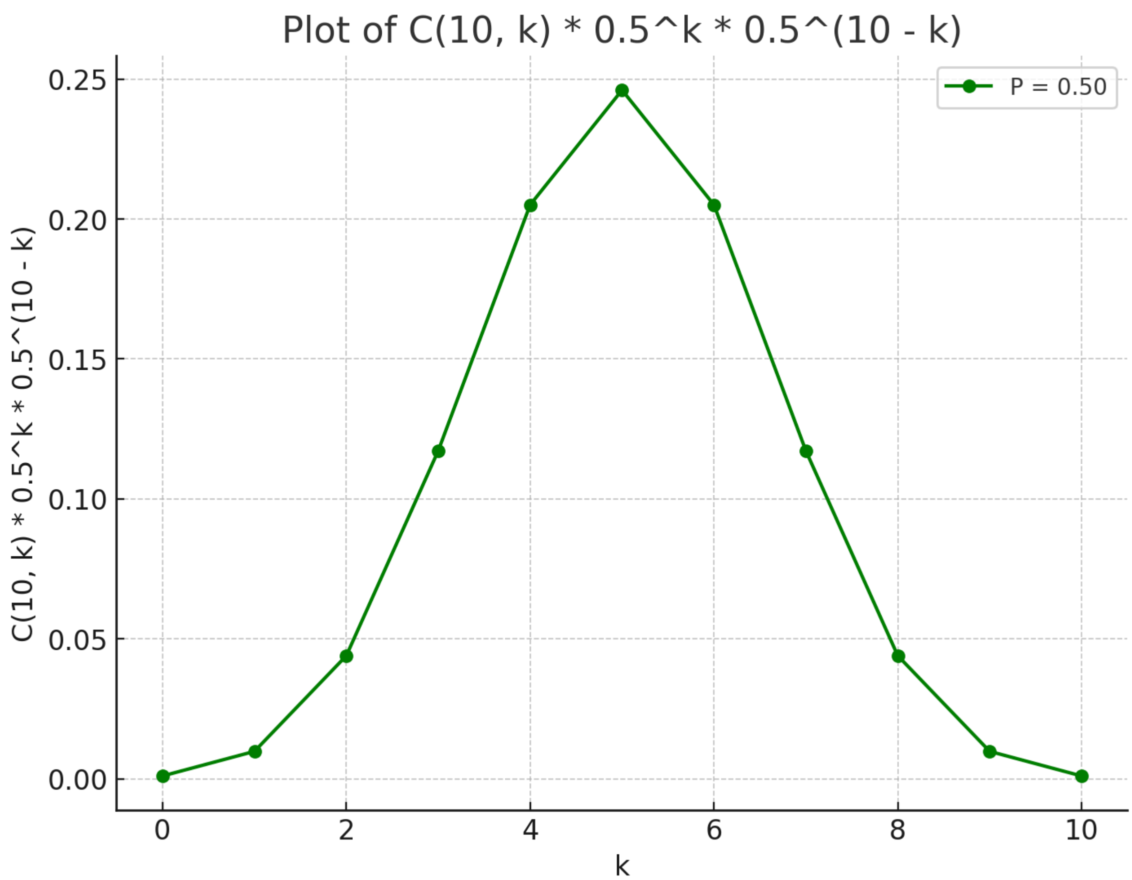
Η πιθανότητα να έρθει κορώνα 6 φορές είναι το άθροισμα της πιθανότητας όλων των περιπτώσεων όπου εμφανίζονται ακριβώς 6 κορώνες. Για παράδειγμα, μια τέτοια περίπτωση είναι η 'ΚΚΚΚΚΚΓΓΓΓ', ενώ μια άλλη είναι η 'ΚΓΚΚΚΚΚΓΓΓ'. Παρατηρούμε ότι, καθώς η σειρά στον πολλαπλασιασμό των πιθανοτήτων δεν παίζει ρόλο, το γινόμενο παραμένει το ίδιο για κάθε περίπτωση και ισούται με 0.4^6 * 0.6^4.

Πόσες τέτοιες συμβολοσειρές υπάρχουν; Αυτό υπολογίζεται βρίσκοντας με πόσους τρόπους μπορούμε να επιλέξουμε 6 από τις 10 θέσεις για τις κορώνες το οποία θα ισούται με C(10, 6). Κάθε τέτοια περίπτωση έχει πιθανότητα 0.4^6 * 0.6^4. Συνεπώς, η συνολική πιθανότητα είναι C(10, 6) * 0.4^6 * 0.6^4."

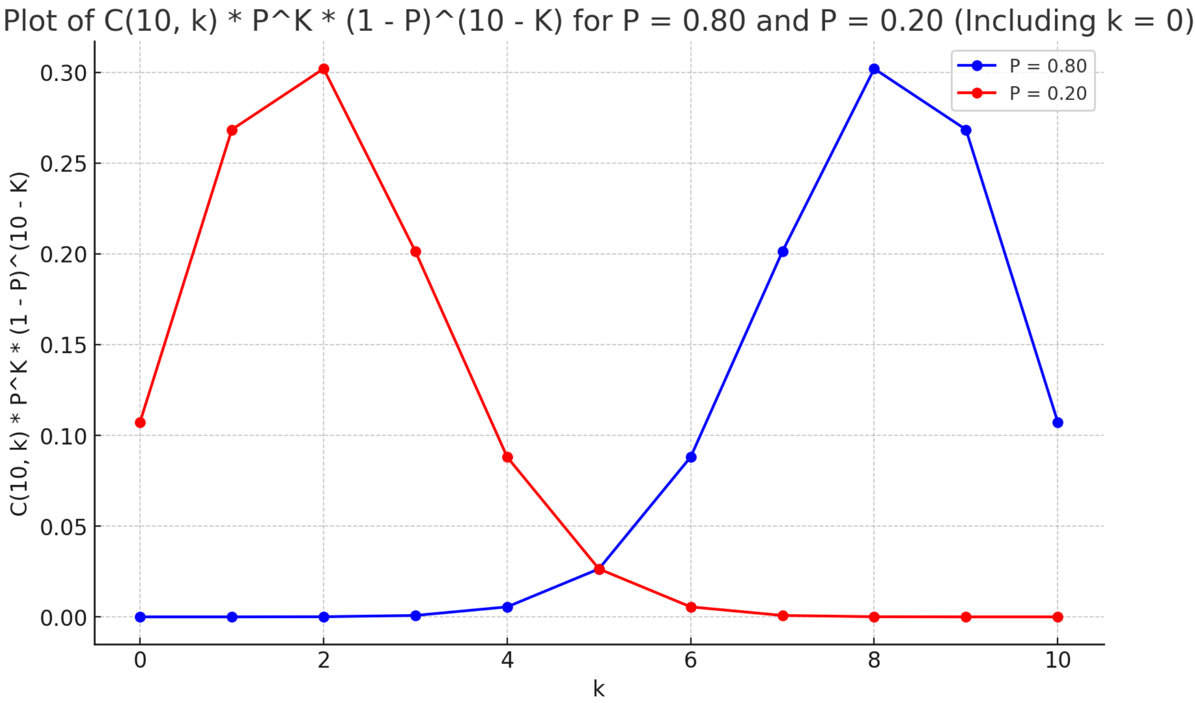
Γενικότερα: Αν η πιθανότητα να έρθει κορώνα είναι P, τότε η πιθανότητα να έρθει κορώνα K φορές σε N ρίψεις είναι C(N, K) * P^K * (1 - P)^(N - K) ⭐

Όταν P = 0.5, παρατηρούμε ότι οι πιθανότητες συγκεντρώνονται γύρω από την τιμή K = N/2. Αυτό σημαίνει ότι οι μεγαλύτερες πιθανότητες εμφανίζονται όταν ο αριθμός των κορώνων K είναι κοντά στο μισό του συνολικού αριθμού των ρίψεων. Δηλαδή, η πιθανότητα να έρθουν 5 κορώνες στις 10 ρίψεις θα έχει μεγαλύτερη πιθανότητα από το να έρθουν 4 η 6 κορώνες.

.



Αλλάζοντας το P, μπορούμε να μετακινήσουμε το σημείο συγκέντρωσης των πιθανοτήτων αριστερά ή δεξιά. Αλλά πάντα σχηματίζεται μια καμπύλη πιθανοτήτων. Ας αναλύσουμε για P = 0.2 και P = 0.8 για παράδειγμα:



Αυτή η καμπύλη αντικατοπτρίζει το Κεντρικό Οριακό Θεώρημα (Central Limit Theorem), το οποίο λέει ότι οι πιθανότητες κατανέμονται κανονικά γύρω από τη μέση τιμή τους, ειδικά όταν ο αριθμός των πειραμάτων είναι μεγάλος.

Η μέση τιμή μπορεί να υπολογιστεί με τον τύπο N * P, το τι σημαίνει αυτό θα το αναλύσουμε περισσότερο αργότερα. Στην ουσία, αυτό που λέει είναι πως, κάνοντας N δοκιμές, με πιθανότητα επιτυχίας P σε κάθε δοκιμή, όπου επιτυχία θεωρούμε την κορώνα, κατα μέσο όρο θα έχουμε N * P επιτυχίες. Μπορει να έχουμε λιγότερες η περισσότερες, αλλά επαναλαμβάνοντας το πείραμα των 10 ρίψεων ξανά και ξανά ο μέσος όρος επιτυχιών θα πλησιάζει το N * P.

Τα παραπάνω δεν είναι τόσο σημαντικά για την εξέταση, απλά ήταν ενδιαφέρον σαν μια εισαγωγή στην μέση τιμή που θα μάθουμε αργότερα. Δεν θα έχει θεωρία όμως η εξέταση, μόνο ασκήσεις.

Κάποια ενδιαφέρον βίντεο που προτείνω:
But what is the Central Limit Theorem? (Εμβαθύνει πιο πολύ από ότι χρειάζεται)
Cute puppy doing cute things (Πολύ σημαντικό, να το δεις)

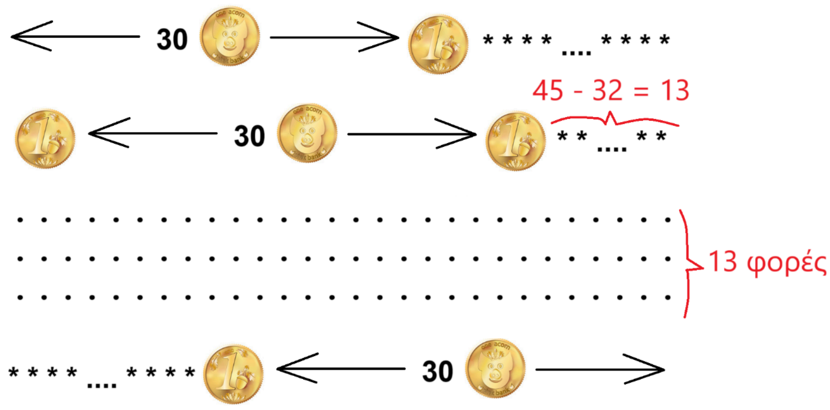
Άσκηση 8: Έστω ότι ρίχνουμε ένα κέρμα 45 φορές, όπου η πιθανότητα να έρθει κορόνα είναι 0,6 και η πιθανότητα να έρθει γράμματα είναι 0,4. Πόση είναι η πιθανότητα να έρθει σερί από κορώνες ακριβώς 30 φορές;

Πιθανότητα: 0.6^30 * 0.4

Πιθανότητα: 0.6^30 * 0.4^2

Πιθανότητα: 0.6^30 * 0.4^2 * 13

Πιθανότητα: 0.6^30 * 0.4



Υπολογίζουμε πιθανότητες περιπτώσεων, και τις αθροίζουμε. Η διαδικασία υπολογισμού των πιθανοτήτων κάθε περίπτωσης βασίζεται στον κανόνα του γινομένου για ανεξάρτητες πιθανότητες, καθώς κάθε ρίψη είναι ανεξάρτητη.

Στην ακολουθία ρίψεων, κάποιες θέσεις είναι αδιάφορες (δηλώνονται με "*"), που σημαίνει ότι δεν μας ενδιαφέρει αν είναι κορώνα ή γράμματα. Αυτές οι θέσεις έχουν πιθανότητα 1, για αυτό και αγνοούνται στον υπολογισμό.

Για παράδειγμα, στην ακολουθία [*Γ←30Κ→Γ**,,,**], η πρώτη ρίψη έχει δηλωθεί με "*", που σημαίνει ότι θέλουμε να έρθει είτε γράμματα είτε κορώνα, κάτι που έχει πιθανότητα 1, αφού το κέρμα θα έρθει αναγκαστικά ένα από τα δύο. Στη συνέχεια, θέλουμε η επόμενη ρίψη να είναι γράμματα, με πιθανότητα 0.4. Αμέσως μετά, ακολουθούν 30 κορώνες, με πιθανότητα 0.6^30, και επειδή η ακολουθία πρέπει να είναι ακριβώς 30 κορώνες, η επόμενη ρίψη πρέπει να είναι γράμματα, με πιθανότητα 0.4. Συνολικά, η πιθανότητα αυτής της συγκεκριμένης ακολουθίας υπολογίζεται ως 0.6^30 * 0.4^2

Για να υπολογίσουμε πιθανότητα να έρθουν 30 κορώνες σερί, θα αθροίσουμε πιθανότητες όλων τον περιπτώσεων, των οποίων έχω καταγράψει στην φωτογραφία πάνω. Τελικά βρίσκουμε μια αρκετά μικρή πιθανότητα που θα ισούται με 2 * 0.6^30 * 0.4 + 14 * 0.6^30 * 0.4^2

Αν αντί για σερί από 30 είχαμε σερί από 10, η άσκηση θα ήταν αρκετά πιο δύσκολη, γιατί θα μπορούσαμε να έχουμε πάνω από μία συνεχόμενη ακολουθία από 10 κορώνες.

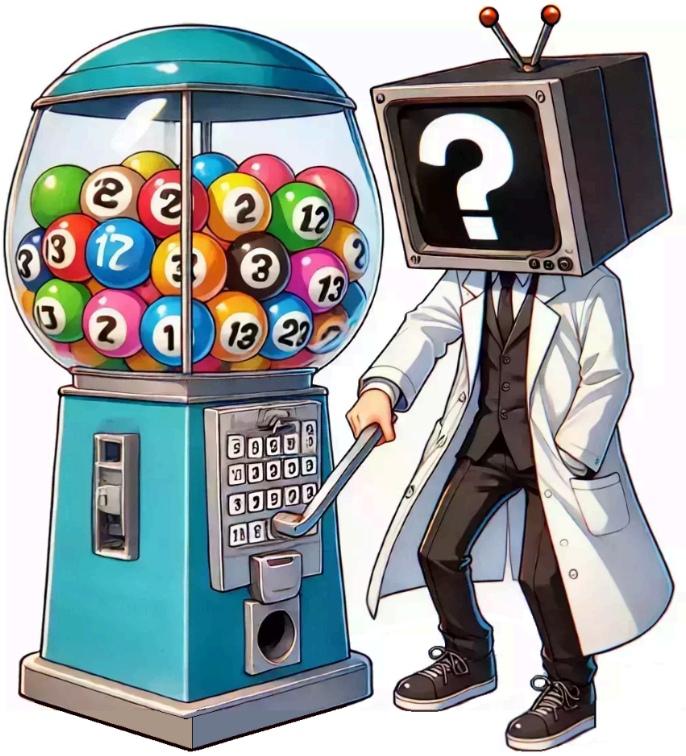
Επίσης, η άσκηση θα γινόταν πιο περίπλοκη αν, αντί να ζητάμε ακριβώς 30, ζητούσαμε τουλάχιστον 30 κορώνες, καθώς αυτό θα επέτρεπε πολλαπλές διαφορετικές ακολουθίες που πληρούν την προϋπόθεση, και θα έπρεπε να υπολογίσουμε την πιθανότητα για καθεμία από αυτές.

Ευτυχώς, κάτι τέτοιο δεν πρόκειται να το ζητήσει σε εξέταση 😌

Ενότητα 4: Τυχαίες Μεταβλητές και άλλα

Μια **τυχαία μεταβλητή** είναι μια μαθηματική έννοια που περιγράφει την τιμή που μπορεί να προκύψει από ένα τυχαίο πείραμα. Στην ουσία, η τυχαία μεταβλητή συνδέει κάθε πιθανό αποτέλεσμα ενός τυχαίου γεγονότος με έναν αριθμό, και μέσω αυτής μπορούμε να αναλύσουμε και να κατανοήσουμε τα πιθανά αποτελέσματα του πειράματος.

Ένα παράδειγμα είναι μια **μηχανή με τσίχλες**, όπου τραβάς έναν μοχλό και βγαίνει μια τσίχλα τυχαίας γεύσης. Στην περίπτωση αυτή, το "πείραμα" είναι το τράβηγμα του μοχλού, και η "τυχαία μεταβλητή" θα μπορούσε να είναι η γεύση της τσίχλας που βγαίνει. Κάθε φορά που τραβάς τον μοχλό, το αποτέλεσμα είναι απρόβλεπτο και μπορεί να είναι, για παράδειγμα, φράουλα, λεμόνι ή μέντα. Η τυχαία μεταβλητή αναλαμβάνει να "καταγράψει" αυτή την τυχαία γεύση, συνδέοντάς την με μια τιμή (π.χ. 1 για φράουλα, 2 για λεμόνι, 3 για μέντα).



Το όνομα "τυχαία μεταβλητή" προκύπτει ακριβώς από το γεγονός ότι **η τιμή της μεταβάλλεται τυχαία** κάθε φορά που εκτελείται το πείραμα, δηλαδή κάθε φορά που τραβάς τον μοχλό, η γεύση αλλάζει με τρόπο που δεν μπορούμε να προβλέψουμε με βεβαιότητα εκ των προτέρων.

Στο παράδειγμα με τη μηχανή με τις τσίχλες, το "πείραμα" είναι το τράβηγμα του μοχλού. Ο χώρος των πιθανών αποτελεσμάτων **Ω** μπορεί να περιλαμβάνει όλες τις πιθανές γεύσεις, π.χ. **Ω = { φράουλα, λεμόνι, μέντα }**. Η τυχαία μεταβλητή **X** αντιστοιχίζει κάθε γεύση σε έναν αριθμό, π.χ:

X(φράουλα) = 1, X(λεμόνι) = 2, X(μέντα) = 3

Για τις **διακριτές τυχαίες μεταβλητές**, όπως η τυχαία μεταβλητή **X** στο παράδειγμα της μηχανής με τις τσίχλες, το αποτέλεσμα που παίρνουμε είναι **μετρήσιμο και συγκεκριμένο**. Δηλαδή, μπορεί να πάρει μόνο κάποιες συγκεκριμένες τιμές, όπως οι γεύσεις (φράουλα, λεμόνι, μέντα).

Η πιθανότητα κάθε συγκεκριμένης τιμής περιγράφεται από τη **συνάρτηση μάζας πιθανότητας (pmf)**:

$$p_X(x) = \mathbb{P}(X = x)$$ ⭐

Στο παράδειγμά μας, οι πιθανότητες για κάθε γεύση είναι:

- $p_X(1) = 0.5$ (πιθανότητα να βγει φράουλα)
- $p_X(2) = 0.3$ (πιθανότητα να βγει λεμόνι)
- $p_X(3) = 0.2$ (πιθανότητα να βγει μέντα)

Αυτές οι πιθανότητες μάς δίνουν την πιθανότητα να πάρουμε μια συγκεκριμένη τιμή για την τυχαία μεταβλητή X. Στην περίπτωση των διακριτών μεταβλητών, έχει νόημα να μιλάμε για την πιθανότητα να συμβεί ακριβώς ένα συγκεκριμένο αποτέλεσμα, όπως η πιθανότητα $X=1$ (δηλαδή να βγει φράουλα).

Όταν περνάμε σε **συνεχείς τυχαίες μεταβλητές**, η κατάσταση αλλάζει. Σε αυτή την περίπτωση, η τυχαία μεταβλητή X μπορεί να πάρει οποιαδήποτε τιμή μέσα σε ένα διάστημα, και όχι απλά κάποιες συγκεκριμένες διακριτές τιμές. Ένα παράδειγμα συνεχούς τυχαίας μεταβλητής είναι η θερμοκρασία που καταγράφεται σε μια ημέρα, η οποία μπορεί να είναι οποιοσδήποτε αριθμός, όπως 23.15°C ή 23.151°C, κ.ο.κ.

Σε τέτοιες περιπτώσεις, η πιθανότητα να πάρει η τυχαία μεταβλητή ακριβώς μια συγκεκριμένη τιμή $X = x$ είναι **μηδενική**. Αυτό συμβαίνει επειδή σε ένα συνεχές σύνολο, υπάρχουν άπειρες δυνατές τιμές που η μεταβλητή μπορεί να πάρει, και επομένως η πιθανότητα να "πέσει" ακριβώς σε μία συγκεκριμένη τιμή είναι σχεδόν ανύπαρκτη.

Γι' αυτό, αντί να μιλάμε για πιθανότητες μεμονωμένων τιμών, χρησιμοποιούμε την έννοια της **συνάρτησης κατανομής (CDF)**.

Η συνάρτηση κατανομής εκφράζει την πιθανότητα να πάρει η τυχαία μεταβλητή τιμή **μικρότερη ή ίση** από μια συγκεκριμένη τιμή x:

$$\mathbf{F_X(x) = \mathbb{P}(X \leq x)}$$ ⭐

Αυτή η συνάρτηση μάς δίνει τη **συσσωρευμένη πιθανότητα** μέχρι την τιμή x. Για παράδειγμα, αν μετράμε την κατανομή της θερμοκρασίας, η $F_X(23.15)$ θα μας έλεγε την πιθανότητα να είναι η θερμοκρασία **μικρότερη ή ίση** από 23.15°C. Αυτή η συνάρτηση **είναι πάντα αύξουσα** ή τουλάχιστον μη φθίνουσα, διότι η πιθανότητα να πάρει η τυχαία μεταβλητή τιμή μικρότερη ή ίση από το x μπορεί μόνο να αυξάνεται ή να μένει σταθερή καθώς το x μεγαλώνει.

Στις συνεχείς τυχαίες μεταβλητές, η πιθανότητα ότι η μεταβλητή παίρνει τιμή μέσα σε ένα μικρό διάστημα εκφράζεται μέσω της **συνάρτησης πυκνότητας πιθανότητας (pdf)**, την οποία συμβολίζουμε με $\mathbf{f_X(x)}$. Η συνάρτηση αυτή δεν μας δίνει την πιθανότητα να πάρει η τυχαία μεταβλητή μια συγκεκριμένη τιμή, αλλά μας δείχνει **πόσο πυκνά** είναι τα αποτελέσματα γύρω από μια συγκεκριμένη τιμή.

Για παράδειγμα, η πιθανότητα ότι η θερμοκρασία βρίσκεται **μέσα σε ένα μικρό διάστημα** μεταξύ των τιμών 22.9°C και 23.1°C, μπορεί να υπολογιστεί ως το ολοκλήρωμα της συνάρτησης πυκνότητας πιθανότητας $\mathbf{f_X(x)}$ σε αυτό το διάστημα:

$$\mathbb{P}(22.9 \leq X \leq 23.1) = \int_{22.9}^{23.1} fx(x)dx$$

$$= Fx(23.1) - Fx(22.9)$$ ⭐

Αυτό μας δίνει την πιθανότητα ότι η τυχαία μεταβλητή X θα βρεθεί σε ένα συγκεκριμένο διάστημα, κάτι που είναι ουσιαστικό στις συνεχείς μεταβλητές, όπου δεν μπορούμε να μιλάμε για πιθανότητες ακριβούς τιμής.

Η συνάρτηση πυκνότητας πιθανότητας (pdf) $\mathbf{f_\chi(x)}$ ορίζεται ως η παράγωγος της συνάρτησης κατανομής (CDF) $\mathbf{F_\chi(x)}$. Δηλαδή: $f\chi(x) = \frac{d}{dx}F\chi(x)$ ⭐

Παράδειγμα1: Ρίχνουμε ένα νόμισμα, όπου η πιθανότητα να έρθει "κεφαλή" είναι ίση με p. Ποια θα ήταν η γραφική παράσταση της συνάρτησης κατανομής (CDF) για την τυχαία μεταβλητή που δηλώνει το αποτέλεσμα του κέρματος;

Η τυχαία μεταβλητή που δηλώνει το αποτέλεσμα του κέρματος είναι μια **διακριτή τυχαία μεταβλητή** και ακολουθεί την **διωνυμική κατανομή** για ένα μόνο πείραμα (δηλαδή μια **Bernoulli** κατανομή). Αυτή η τυχαία μεταβλητή μπορεί να πάρει δύο τιμές:

- 0 (αν φέρει "γράμματα"),
- 1 (αν φέρει "κεφαλή").

Η **συνάρτηση κατανομής (CDF)**, Fx(x), μάς δίνει την πιθανότητα ότι η τυχαία μεταβλητή X θα πάρει τιμή μικρότερη ή ίση με το x. Στην περίπτωση του κέρματος, η CDF θα έχει τα εξής χαρακτηριστικά:

- Για τιμές **x < 0**, η πιθανότητα είναι μηδενική, δηλαδή **Fx(x) = 0** γιατί δεν μπορεί να έχουμε αρνητική τιμή για το αποτέλεσμα του κέρματος.
- Για **0 ≤ x < 1**, η πιθανότητα **Fx(x)** είναι ίση με την πιθανότητα να έρθουν "γράμματα", δηλαδή **Fx(x) = 1 − p**.
- Για **x ≥ 1**, η πιθανότητα **Fx(x)** είναι ίση με 1, αφού καλύπτει όλες τις πιθανότητες, δηλαδή είτε έρθει "κεφαλή" είτε "γράμματα".

Παράδειγμα2: Ρίχνουμε ένα νόμισμα συνεχόμενα μέχρι να έρθει “Κεφαλή”, με πιθανότητα να έρθει "Κεφαλή" σε κάθε ρίψη να είναι ίση με p. Κάθε ρίψη είναι ανεξάρτητη από τις προηγούμενες, και όπως και στο προηγούμενο παράδειγμα, για κάθε i = 1, 2, …, m ορίζουμε τις τυχαίες μεταβλητές Xi, οι οποίες περιγράφουν το αποτέλεσμα της i-στης ρίψης.

$$X_i = \begin{cases} 1, & \text{αν φέραμε "Κ" στη ρίψη } i \\ 0, & \text{αν φέραμε "Γ" στη ρίψη } i \end{cases}$$

Τώρα, ορίζουμε μια νέα τυχαία μεταβλητή Z, η οποία περιγράφει **την πρώτη φορά που θα φέρουμε "K"**. Η τυχαία μεταβλητή Z λαμβάνει ακέραιες τιμές και υποδηλώνει τον αριθμό της ρίψης στην οποία εμφανίζεται το πρώτο "K". Συγκεκριμένα, η Z ορίζεται ως εξής:

$$Z = \begin{cases} 1, & \text{αν } X_1 = 1 \\ 2, & \text{αν } X_1 = 0 \text{ και } X_2 = 1 \\ 3, & \text{αν } X_1 = 0,\, X_2 = 0, \text{ και } X_3 = 1 \\ \vdots & \\ m, & \text{αν } X_1 = X_2 = \cdots = X_{m-1} = 0 \text{ και } X_m = 1 \end{cases}$$

Με αυτόν τον τρόπο, η μεταβλητή Z μάς δείχνει τον αριθμό των ρίψεων που χρειάστηκαν για να φέρουμε "Κεφαλή" πρώτη φορά.

Η **συνάρτηση μάζας πιθανότητας (pmf)** της Z, δηλαδή η πιθανότητα να πάρει η Z την τιμή m (η πρώτη φορά που φέρνουμε "K" να είναι στην m-οστή ρίψη), είναι:

$\mathbb{P}(Z = m) = (1 - p)^{m-1}\, p$ ⭐

Επειδή κάθε ρίψη είναι ανεξάρτητη, και θέλουμε να έρθουν (m - 1) γράμματα με πιθανότητα (1 - p) σε κάθε ρίψη και μετά μία κεφαλή με πιθανότητα p.

Η τυχαία μεταβλητή Z ακολουθεί τη **γεωμετρική κατανομή**, που περιγράφει πότε θα εμφανιστεί η πρώτη επιτυχία σε μια σειρά ανεξάρτητων δοκιμών. Τις βασικές κατανομές που βλέπουμε θα τις ορίσω ξανά αργότερα πιο αναλυτικά ώστε να είναι μαζεμένες κάπου καθώς είναι ιδιαίτερα χρήσιμες.

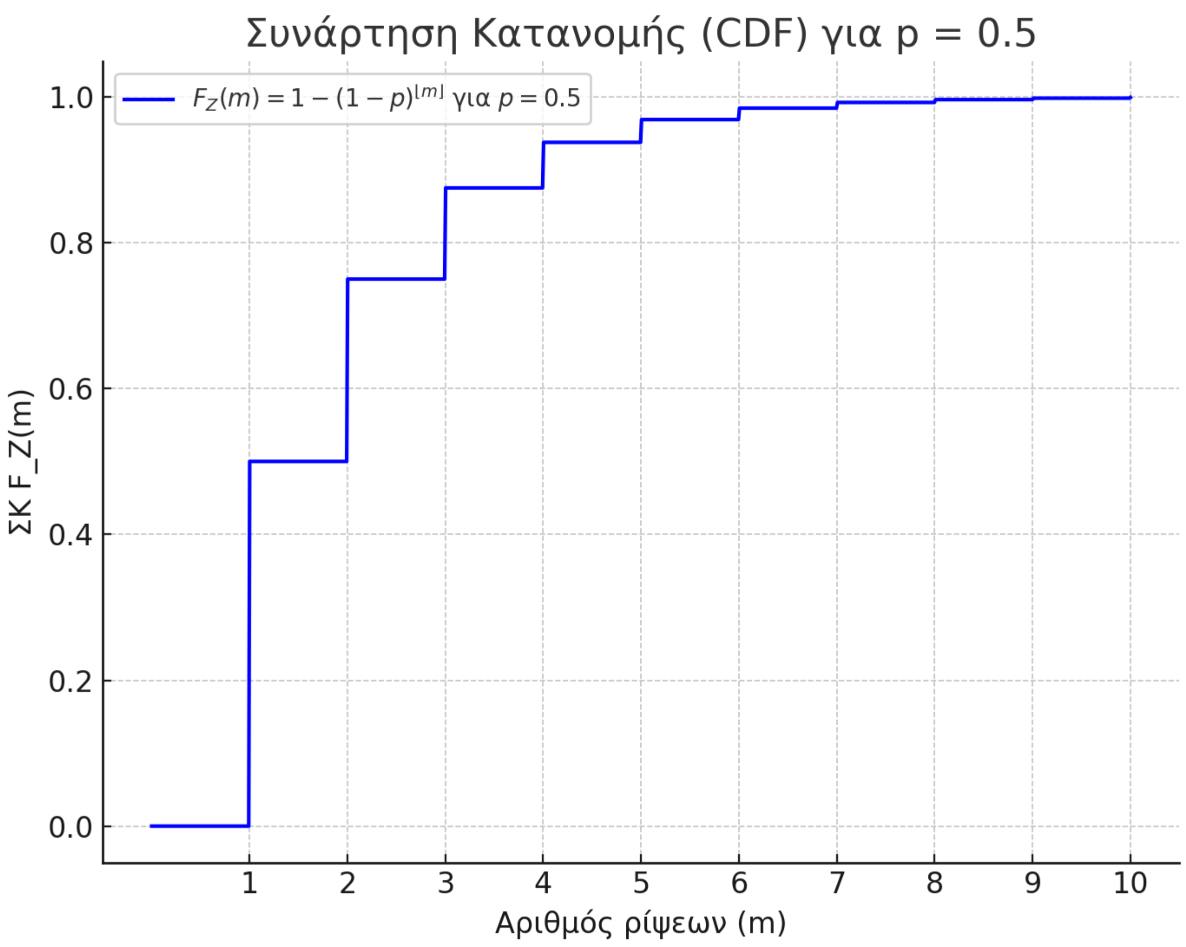
Η **συνάρτηση κατανομής πιθανότητας (CDF)** Fz(m) δίνει την πιθανότητα η τυχαία μεταβλητή Z να είναι μικρότερη ή ίση μιας συγκεκριμένης τιμής m, δηλαδή:

$$F_Z(m) = \mathbb{P}(Z \leq \lfloor m \rfloor)$$

Όπου ⌊m⌋ είναι το ακέραιο μέρος του m. Αυτό σημαίνει ότι θέλουμε να υπολογίσουμε την πιθανότητα να φέρουμε "K" στην 1η, 2η, ..., ή στην ⌊m⌋-οστή ρίψη. Επειδή οι ρίψεις είναι ανεξάρτητες, η πιθανότητα να φέρουμε "K" σε οποιαδήποτε από τις ρίψεις έως την ⌊m⌋-οστή είναι η συμπληρωματική πιθανότητα του να μην έχουμε φέρει "K" μέχρι και την ⌊m⌋-οστή ρίψη, δηλαδή:

$$F_Z(m) = \sum_{k=1}^{\lfloor m \rfloor} \mathbb{P}(Z = k) = 1 - (1-p)^{\lfloor m \rfloor}$$ ⭐

Όπου ο όρος (1-p)^⌊m⌋ εκφράζει την πιθανότητα να έρθουν γράμματα σε ⌊m⌋ ρίψεις. Τότες η συμπληρωματική πιθανότητα 1 - (1-p)^⌊m⌋ ισούται με την πιθανότητα να μην έρθουν γράμματα σε ⌊m⌋ ρίψεις, δηλαδή την πιθανότητα να φέρουμε κορώνα στην 1η, 2η, …, ή στην ⌊m⌋-οστή ρίψη.



Άσκηση1: Έστω πως το νόμισμα είναι δίκαιο, δηλαδή p = 0.5, και κάποιος μας προτείνει το εξής παιχνίδι:

- Θα του δώσουμε 2.50 ευρώ και θα ρίξουμε το νόμισμα επαναλαμβανόμενα μέχρι να φέρουμε κορώνα ("Κ") για πρώτη φορά.
- Η αμοιβή που θα πάρουμε εξαρτάται από τον αριθμό των ρίψεων που κάναμε για να φέρουμε "Κ":
    - Αν φέρουμε "Κ" στην 1η ρίψη, θα μας δώσει 1 ευρώ.
    - Αν φέρουμε "Γ" στην πρώτη ρίψη και "Κ" στη 2η ρίψη, θα μας δώσει 2 ευρώ.
    - Γενικά, αν φέρουμε m−1 φορές "Γ" και φέρουμε "Κ" στην m-οστή ρίψη, θα μας δώσει m ευρώ.

Η ερώτηση είναι:

1. Είναι δίκαιη η αρχική αμοιβή των 2.50 ευρώ που μας ζητά να πληρώσουμε για να συμμετάσχουμε στο παιχνίδι;
2. Αν όχι, ποια θα ήταν μια δίκαιη αμοιβή για να συμμετάσχουμε στο παιχνίδι;

Για να απαντήσουμε σε αυτές τις ερωτήσεις, χρειαζόμαστε να μάθουμε πρώτα την έννοια της αναμενόμενης τιμής:

Η **αναμενόμενη τιμή** μιας τυχαίας μεταβλητής είναι η **μέση τιμή** που θα αναμέναμε να πάρουμε αν εκτελούσαμε το πείραμα πολλές φορές. Στην ουσία, η αναμενόμενη τιμή μας δείχνει τον **μακροχρόνιο μέσο όρο** των αποτελεσμάτων.

Για μια διακριτή τυχαία μεταβλητή X, η αναμενόμενη τιμή **E[X]** υπολογίζεται με τον εξής τρόπο:

$$\mathbf{E[X]} = \sum_x x \cdot \mathbb{P}(X = x)$$ ⭐

Για παράδειγμα έστω μια τυχαία μεταβλητή με τιμές 1, 2 και 3, και πιθανότητες 0.3, 0.5 και 0.2 αντίστοιχα. Σε N πειράματα, αναμένουμε ότι η τιμή 1 θα εμφανιστεί N * 0.3 φορές, η τιμή 2 θα εμφανιστεί N * 0.5 φορές και η τιμή 3 N * 0.2 φορές. Αν συμβεί αυτό, ο μέσος όρος θα είναι ίσος με:

$$\frac{1\cdot 0.3\cdot N + 2\cdot 0.5\cdot N + 3\cdot 0.2\cdot N}{N} = 1 \cdot 0.3 + 2 \cdot 0.5 + 3 \cdot 0.2 = E[X]$$

Στην πράξη, τα αποτελέσματα μπορεί να διαφέρουν για μικρά N, αλλά όσο αυξάνεται ο αριθμός των πειραμάτων, ο μέσος όρος των παρατηρήσεων τείνει να πλησιάζει την αναμενόμενη τιμή E[X]. Αυτό το φαινόμενο περιγράφεται από το **Νόμο των Μεγάλων Αριθμών**.

Στην συνέχεια, θα δείξω πως για μια τυχαία μεταβλητή Z η οποία ακολουθεί την **γεωμετρική κατανομή**, με πιθανότητα επιτυχίας p, ισχύει ότι:

$$E[Z] = \frac{1}{p}$$ ⭐

Διαισθητικά και μόνο αν το σκεφτείς, αν το να έρθει μια επιτυχία έχει πιθανότητα 1/n τότε κατά μέσο όρο, θα έρχεται επιτυχία 1 φορά σε κάθε n δοκιμές, και οπότε περιμένεις ότι θα χρειαστείς περίπου n δοκιμές κατα μέσο όρο για να έρθει η επιτυχία.

Μπορούμε να το αποδείξουμε και αλγεβρικά όμως αυτό μέσα από τον τύπο της αναμενόμενης τιμής. Θυμίζω πως εφόσον η Z είναι **γεωμετρικής κατανομής** με πιθανότητα επιτυχίας p αυτό σημαίνει ότι:

$$\mathbb{P}(Z = m) = (1 - p)^{m-1} p$$

Επομένως, η αναμενόμενη τιμή E[Z] μπορεί να υπολογιστεί ως:

$$\mathbf{E[Z] =} \sum_{n=1}^{\infty} n \cdot (1 - p)^{n-1} p \mathbf{=} p \cdot \sum_{n=1}^{\infty} n \cdot (1 - p)^{n-1}$$

Η παραπάνω σειρά είναι γνωστή. Δεν θα αναλύσω το πως υπολογίζεται, καθώς αυτό βγαίνει εκτός των ορίων του συγκεκριμένου μαθήματος. Δεν θα χρειαστεί να λύσεις τέτοιο άθροισμα σε εξέταση. Αν σε ενδιαφέρει όμως μπορείς να το ψάξεις περισσότερο μόνος σου:

$$\mathbf{E[Z] =} p \cdot \frac{1}{(1-(1-p))^2} = p \cdot \frac{1}{p^2} = \frac{1}{p}$$

Επιστροφή στο παιχνίδι με το νόμισμα:

Έστω τώρα Z τυχαία μεταβλητή που εκφράζει την αμοιβή που παίρνουμε όταν παίζουμε έναν γύρο του παιχνιδιού με το νόμισμα. Η αμοιβή είναι ανάλογη των ρίψεων που χρειάστηκαν μέχρι να έρθει κεφαλή.

Παρατηρούμε πως η Z ακολουθεί την **γεωμετρική κατανομή** για πιθανότητα επιτυχίας 1/2. Έτσι υπολογίζουμε αναμενόμενη τιμή E[Z] = 2 που σημαίνει ότι η μέση τιμή που θα αναμέναμε να πάρουμε αν εκτελούσαμε το πείραμα πολλές φορές είναι 2.

Η αναμενόμενη αμοιβή των 2 ευρώ είναι **μικρότερη** από τα 2.50 ευρώ που πρέπει να πληρώσουμε για να παίξουμε. Αυτό σημαίνει ότι, κατά μέσο όρο, θα κερδίζουμε λιγότερα από όσα πληρώνουμε για να συμμετέχουμε στο παιχνίδι. Συμπερασματικά, **η αρχική αμοιβή των 2.50 ευρώ δεν είναι δίκαιη**.

Όσο για το δεύτερο ερώτημα, η δίκαιη τιμή για να συμμετάσχουμε στο παιχνίδι είναι **2 ευρώ**. Αν η συμμετοχή κοστίζει αυτό το ποσό, τότε το παιχνίδι θα είναι δίκαιο, καθώς κατά μέσο όρο θα κερδίζουμε ακριβώς όσα πληρώνουμε.

Συνεχίζοντας ας δούμε κάποιες Βασικές ιδιότητες της αναμενόμενης τιμής. Μία από τις πιο βασικές είναι η **γραμμικότητα**. Αυτή η ιδιότητα μάς λέει ότι για δύο τυχαίες μεταβλητές X και Y, ισχύει πως:

E(X + Y) = E(X) + E(Y) ⭐

Για να κατανοήσουμε γιατί η **γραμμικότητα** της αναμενόμενης τιμής ισχύει, ας θυμηθούμε μια βασική ιδιότητα των πιθανοτήτων, την οποία ήδη έχουμε δει στην **ενότητα 2** της παρουσίασης:

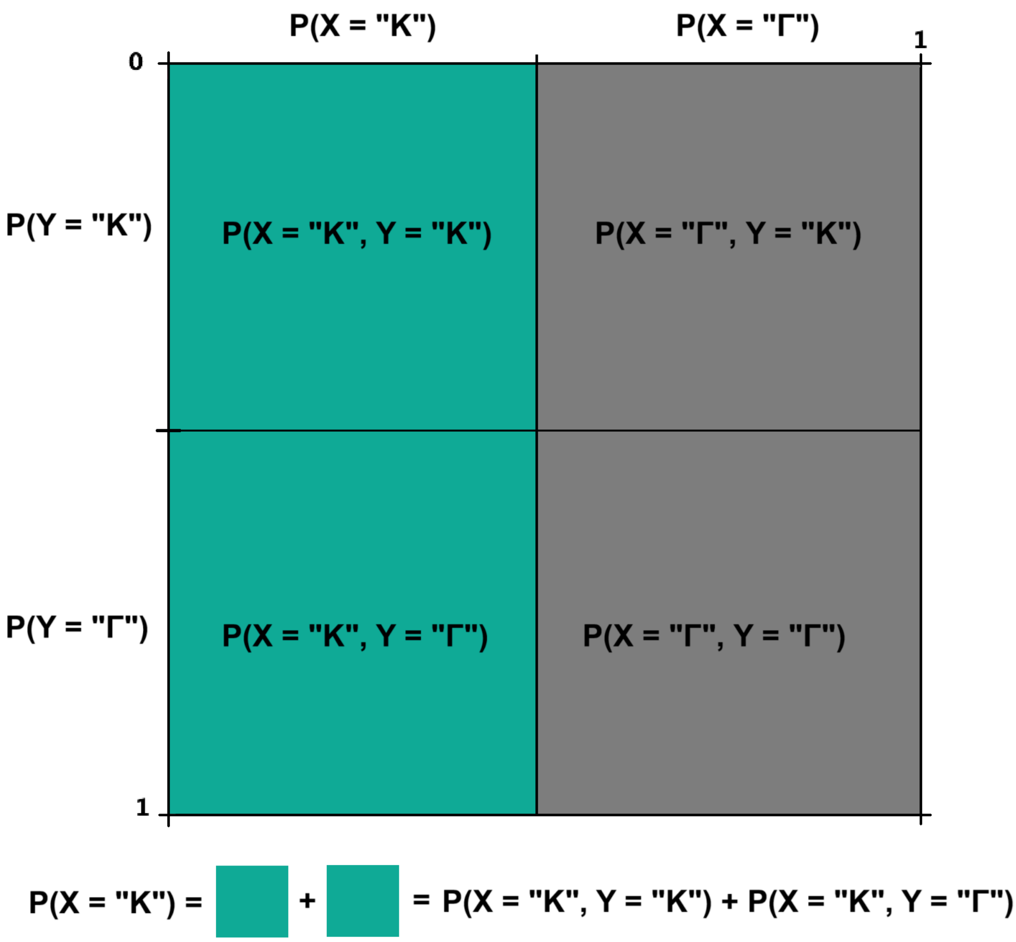
$$P(X = x) = \sum_{y} \mathbb{P}(X = x, Y = y)$$ ⭐

Μπορούμε να υπολογίσουμε την πιθανότητα ενός συγκεκριμένου αποτελέσματος X = x συνυπολογίζοντας όλες τις πιθανότητες όπου X = x ανεξάρτητα από την τιμή του Y. Όπως συζητήσαμε και στην **ενότητα 2**, αντιστοιχίζουμε κάθε πιθανό συνδυασμό των X και Y σε συγκεκριμένα τμήματα του επιπέδου. Έτσι, η πιθανότητα του X = x αντανακλά το μέγεθος της περιοχής του επιπέδου που καταλαμβάνουν τα γεγονότα όπου X = x.

Για παράδειγμα, αν X και Y είναι τα αποτελέσματα δύο ρίψεων ενός νομίσματος, η πιθανότητα να έρθει κορώνα για το X είναι ίση με το άθροισμα των πιθανοτήτων όπου:

1. Το X φέρνει κορώνα και το Y φέρνει κορώνα,
2. Το X φέρνει κορώνα και το Y φέρνει γράμματα.

Αν το νόμισμα είναι δίκαιο, τότε η πιθανότητα να φέρει κορώνα το X σε κάθε περίπτωση είναι 1/4 (από κανόνα του γινομένου ανεξάρτητων γεγονότων), και έτσι η συνολική πιθανότητα να φέρει κορώνα το X είναι 1/4 + 1/4 = 1/2:



Ακόμα και διαισθητικά, η ιδιότητα δεν λέει κάτι περίπλοκο. Σκέψου, για παράδειγμα, ένα νόμισμα. Η πιθανότητα να ρίξεις το νόμισμα και να έρθει κορώνα είναι ίση με το άθροισμα δύο περιπτώσεων:

1. Η πιθανότητα να έρθει κορώνα και να συμβεί κάτι άλλο, π.χ. ο Ξέργιας να γίνει στρίπερ, και
2. Η πιθανότητα να έρθει κορώνα και ο Ξέργιας να **μην** γίνει στρίπερ.

Εφόσον αποδεχόμαστε και τα δύο πιθανά σενάρια (ο Ξέργιας είτε γίνεται στρίπερ είτε όχι), το μόνο που μας ενδιαφέρει είναι αν το νόμισμα φέρει κορώνα. Και αυτή η πιθανότητα είναι απλώς 1/2, ανεξάρτητα από το τι κάνει ο Ξέργιας.

Ας δούμε τώρα πως χρησιμοποιώντας αυτή την ιδιότητα μπορούμε να αποδειξουμε την γραμμικότητα της αναμενόμενης τιμής. Για δύο τυχαίες μεταβλητές X και Y, μπορούμε να υπολογίσουμε την αναμενόμενη τιμή του αθροίσματος τους από την σχέση:

$$E(X + Y) = \sum_{x}\sum_{y}(x + y) \cdot \mathbb{P}(X = x, Y = y)$$

Αυτό προκύπτει από τον τύπο της αναμενόμενης τιμής. Σκέψου το άθροισμα X + Y ως μια νέα τυχαία μεταβλητή Z = X + Y όπου οι τιμές που μπορεί να πάρει είναι τα διάφορα αθροίσματα (x+y), με πιθανότητα P(X = x,Y = y).

Μπορεί το ίδιο άθροισμα x + y να προκύψει από διαφορετικά ζευγάρια τιμών x και y. Αυτό όμως λαμβάνεται υπόψη στον τύπο χάρη στην ιδιότητα του κοινού παράγοντα του πολλαπλασιασμού, όλα τα κοινά αθροίσματα x + y θα βγουν ως κοινός παράγοντας και θα πολλαπλασιαστούν με το άθροισμα των πιθανοτήτων όλων των ζευγαριών που οδηγούν στο κοινό άθροισμα.

Στην συνέχεια, κάνουμε απλοποιήσεις στην εξίσωση για να δείξουμε πως τελικά είναι ίση με E(X) + E(Y):

$$E(X + Y) = \sum_{x} x \sum_{y} \mathbb{P}(X = x, Y = y) + \sum_{y} y \sum_{x} \mathbb{P}(X = x, Y = y)$$

Εδώ είναι που μας χρειάζεται η ιδιότητα που είδαμε πριν:

$$E(X + Y) = \sum_{x} x \ \mathbb{P}(X = x) + \sum_{y} y \ \mathbb{P}(Y = y) = E(X) + E(Y)$$

Και έτσι αποδείξαμε την ιδιότητα της γραμμικότητας. **Δεν είναι απαραίτητο να γνωρίζεις αποδείξεις όμως**, καθώς στην εξέταση δεν θα ζητηθεί θεωρία.

Μία ακόμα σημαντική ιδιότητα της αναμενόμενης τιμής είναι η εξής:

E(c · X) = c · E(X) ⭐

Η ιδιότητα προκύπτει άμεσα από τον τύπο της αναμενόμενης τιμής:

E(c · X) = $\sum_{x} c \cdot x \cdot \mathbb{P}(X = x)$ = c · $\sum_{x} x \cdot \mathbb{P}(X = x)$ = c · E(x)

Συνδυάζοντας τις δύο προηγούμενες ιδιότητες μπορείς να δείξεις ότι:

E(a · X + b · Y) = a · E(X) + b · E(Y) ⭐

Ακόμα, ισχύει πως:

E(X + c) = a · E(X) + c ⭐

Η ιδιότητα προκύπτει άμεσα από τον τύπο της αναμενόμενης τιμής:

E(X + c) = $\sum_{x} (x + c) \cdot \mathbb{P}(X = x)$

= $\sum_{x} x \cdot \mathbb{P}(X = x)$ + $\sum_{x} c \cdot \mathbb{P}(X = x)$ = E[x] + c

Τέλος, για **ανεξάρτητες τυχαίες μεταβλητές** X και Y ισχύει ότι:

E(X · Y) = E(X) · E(Y) ⭐

Η ιδιότητα προκύπτει άμεσα από τον τύπο της αναμενόμενης τιμής:

E(X · Y) = $\sum_{x}\sum_{y}(x \cdot y) \cdot \mathbb{P}(X = x, Y = y)$

$$= \sum_{x}\sum_{y}(x \cdot y) \cdot \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y)$$

$$= \sum_{x} x \cdot \mathbb{P}(X = x) \sum_{y} y \cdot \mathbb{P}(Y = y) = E(x) \cdot E(y)$$

Για **εξαρτημένες τυχαίες μεταβλητές**, η αναμενόμενη τιμή του γινομένου E(X·Y) εξαρτάται από την **κοινή κατανομή** των X και Y.

E(X · Y) = $\sum_{x}\sum_{y}(x \cdot y) \cdot \mathbb{P}(X = x, Y = y)$

$$= \sum_{x}\sum_{y}(x \cdot y) \cdot \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y \mid X = x)$$

$$= \sum_{x}\sum_{y}(x \cdot y) \cdot \mathbb{P}(Y = y) \cdot \mathbb{P}(X = x \mid Y = y)$$

Δύο τυχαίες μεταβλητές X και Y είναι **εξαρτημένες** όταν η τιμή της μίας επηρεάζει ή παρέχει πληροφορίες για την τιμή της άλλης. Αυτό σημαίνει ότι:

$$\mathbb{P}(X = x, Y = y) \neq \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y)$$

Ισχύουν όλα όσα έχουμε μάθει στην **ενότητα 2,** το **P(X = x, Y = y)** μπορεί να υπολογιστεί ως το γινόμενο **P(x = x) * P(Y = y | X = x)**, και μπορεις να χρησιμοποιήσεις όλους τους τύπους, όπως τον **κανόνα το Bayes** που λέει ότι:

$$P(Y = y \mid X = x) = \frac{P(X=x, Y=y)}{P(X=x)} = \frac{P(X=x \mid Y=y)\, P(Y=y)}{P(X=x)}$$ ⭐

Οι τυχαίες μεταβλητές απλά αλλάζουν λίγο το συντακτικό αλλά ότι έχουμε δει πιο πριν από θεωρία παραμένει το ίδιο. Ας συνεχίσουμε με μια άσκηση 💨

Άσκηση2: Κάνουμε ένα πείραμα όπου κάποιος ρίχνει ένα ζάρι συνεχώς, με στόχο να φέρει μια ακολουθία από 2 και 4 που να καταλήγει σε 6. Κάθε φορά που φέρνει 3, 1 ή 5, η ακολουθία απορρίπτεται και ξεκινάει από την αρχή. Αυτό γίνεται μέχρι να φέρουμε κάποια ακολουθία που να τηρεί τις προϋποθέσεις. Όταν προκύψει μια τέτοια ακολουθία, τι αριθμό ρίψεων αναμένουμε να έχει?

Μια αποδεκτή ακολουθία θα μπορούσε να είναι η 2242424222424242446, ενώ μια άλλη μπορεί να είναι μια απλή ρίψη 6 με την πρώτη προσπάθεια. Υπάρχουν άπειρες δυνατές ακολουθίες που πληρούν τις προϋποθέσεις.

Το ζάρι έχει ίση πιθανότητα να φέρει οποιαδήποτε από τις έξι πλευρές του. Η διαδικασία της δημιουργίας ακολουθιών μοιάζει με μια μηχανή όπου:

1. Πατάς το κουμπί και δημιουργείται μια τυχαία ακολουθία.
2. Αν η ακολουθία πληροί τα κριτήρια (αποτελείται από 2 και 4 και τελειώνει με 6), την κρατάς.
3. Αν όχι, την πετάς και επαναλαμβάνεις.

Αυτό το ονομάζουμε **Rejection Sampling (Απορριπτική Δειγματοληψία)**.

Μια πρώτη ιδέα θα ήταν να μοντελοποιήσουμε το ζάρι ως ένα ζάρι με μόνο τρεις όψεις: 2, 4, και 6. Σκεπτόμενοι έτσι, θα υποθέταμε ότι οι πιθανότητες είναι:

- Πιθανότητα 2/3 να έρθει 2 η 4
- Πιθανότητα 1/3 να έρθει 6

(γεωμετρική κατανομή με πιθανότητα επιτυχίας 1/3)

Πιθανότητα να έρθει 6 στην m-οστή ρίψη: **(2/3)^(m-1)** · **(1/3)**

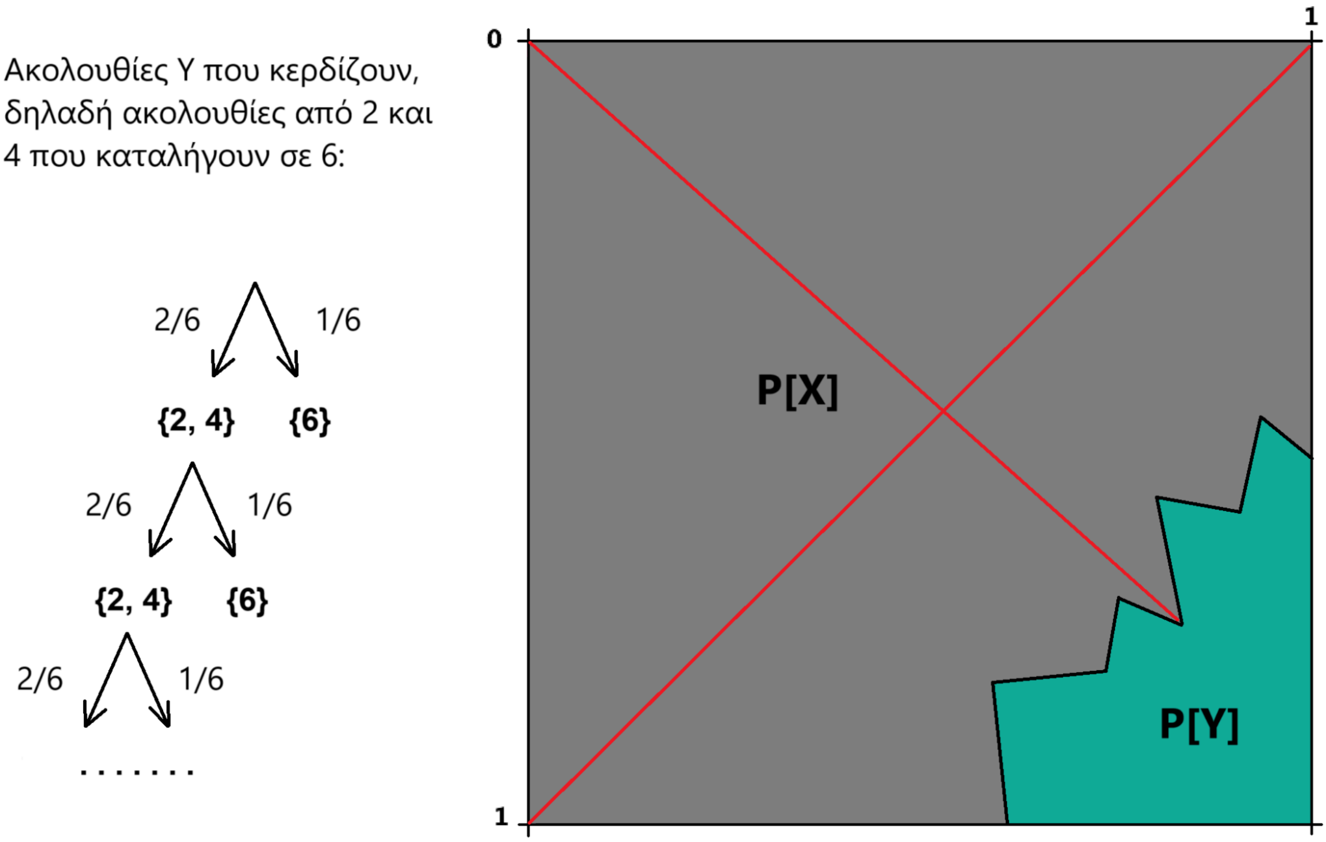
Με αυτή τη λογική, καταλήγουμε στο συμπέρασμα ότι κατά μέσο όρο θα χρειαζόμασταν 3 ρίψεις για να φέρουμε 6. Ωστόσο, αυτός ο υπολογισμός είναι λάθος. Το πρόβλημα έγκειται στο ότι δεν υπολογίσαμε σωστά τις πιθανότητες, αφού παραλείψαμε τον τρόπο με τον οποίο οι απορριφθείσες ακολουθίες επηρεάζουν τις τελικές πιθανότητες.

Ας εξετάσουμε ξανά τις πιθανότητες με μεγαλύτερη προσοχή. Ο δειγματικός χώρος Ω περιλαμβάνει όλες τις πιθανές ακολουθίες. Από αυτές, ορισμένες ανήκουν στο σύνολο αποδεκτών ακολουθιών Y, τις οποίες διατηρούμε, ενώ οι υπόλοιπες ανήκουν στο σύνολο μη αποδεκτών ακολουθιών X, που απορρίπτονται.

Η πιθανότητα να κερδίσεις, δηλαδή να προκύψει μια ακολουθία Y με 2 και 4 που καταλήγει σε 6, υπολογίζεται ως εξής:

- Η πιθανότητα να φέρεις 6 με την πρώτη ρίψη είναι 1/6.
- Η πιθανότητα να φέρεις πρώτα 2 ή 4 και μετά 6 είναι 2/6 * 1/6.
- Η πιθανότητα να φέρεις δύο φορές 2 ή 4 και μετά 6 είναι (2/6)^2 * 1/6.
- Για τρεις φορές 2 ή 4 και μετά 6, η πιθανότητα είναι (2/6)^3 * 1/6
- …
- Για n φορές 2 η 4 και μετά 6, η πιθανότητα είναι (2/6)^n * 1/6

Αθροίζεις τα παραπάνω για n από 0 μέχρι άπειρο.



$$P(Y) = \frac{1}{6} \cdot \left(1 + \frac{1}{3} + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^3 + \ ...\right) \quad \Leftrightarrow$$

$$P(Y) = \frac{1}{6} \cdot \sum_{n=0}^{\infty} \left(\frac{1}{3}\right)^n = \frac{1}{6} \cdot \frac{3}{2} = \frac{1}{4}$$

Η πιθανότητα να κερδίσεις, δηλαδή να φέρεις μια ακολουθία από 2 και 4 και στο τέλος 6, είναι 1/4. Αυτή η πιθανότητα σημαίνει ότι από όλες τις πιθανές ακολουθίες ρίψεων, μόνο το 25% ανήκουν στο σύνολο των νικητήριων ακολουθιών.

Όμως, επειδή ενδιαφερόμαστε μόνο για τις ακολουθίες που κερδίζουν, μπορούμε να απορρίψουμε τις υπόλοιπες ακολουθίες (που περιέχουν 1, 3 ή 5). Έτσι, εργαζόμαστε μόνο με το 25% των πιθανών ακολουθιών που έχουν πιθανότητα 1/4 στο αρχικό δειγματικό χώρο.

Αφού απορρίψουμε τις μη αποδεκτές ακολουθίες, μένουμε με τον χώρο των νικητήριων ακολουθιών. Αυτό σημαίνει ότι η συνολική μάζα πιθανότητας που έχουμε τώρα είναι ίση με 1 (ή 100% των αποδεκτών ακολουθιών).

Για να το επιτύχουμε αυτό, οι πιθανότητες των επιτρεπτών ακολουθιών πρέπει να αναθεωρηθούν έτσι ώστε να καταλάβουν όλο το νέο δειγματικό χώρο.

Η διαδικασία της αναθεώρησης των πιθανοτήτων γίνεται όπως μάθαμε στην **ενότητα 2**:

Κάθε πιθανότητα μιας νικητήριας ακολουθίας αυξάνεται πολλαπλασιαζόμενη με 4, αφού απορρίψαμε τις υπόλοιπες ακολουθίες (επειδή η πιθανότητα των νικητήριων ακολουθιών ήταν αρχικά 1/4).

Για παράδειγμα:

- Η πιθανότητα να φέρουμε 6 στην **πρώτη ρίψη** ήταν αρχικά 1/6. Μετά την αναθεώρηση, γίνεται:

$$P'(6 \text{ στην 1η ρίψη}) = \frac{1}{6} \cdot 4 = \frac{4}{6} = \frac{2}{3}$$

- Η πιθανότητα να φέρουμε 6 στη **δεύτερη ρίψη** (αφού πρώτα φέρουμε 2 ή 4) ήταν αρχικά (2/6) * (1/6). Μετά την αναθεώρηση, γίνεται:

$$P'(6 \text{ στην 2η ρίψη}) = \frac{2}{6} \cdot \frac{1}{6} \cdot 4 = \frac{1}{3} \cdot \frac{2}{3}$$

- Γενικότερα η πιθανότητα για την m-οστή ρίψη γίνεται:

$$P'(6 \text{ στην m-οστή ρίψη}) = \left(\frac{1}{3}\right)^{m-1} \frac{2}{3}$$

Η πάνω σχέση είναι ο τύπος μιας τυχαίας μεταβλητής **γεωμετρικής κατανομής** με πιθανότητα επιτυχίας (2/3). Επομένως, εύκολα συμπεράνουμε πως η αναμενόμενη τιμή για το σε ποια ρίψη θα έρθει 6 μπορεί να υπολογιστεί από τον γνωστό τύπο **E[m] = 1 / (2/3) = 3/2 = 1.5**

Το 1.5 είναι ο **αναμενόμενος αριθμός ρίψεων** που απαιτούνται για να πετύχεις 6 **υπό την προϋπόθεση ότι έχεις ήδη αφαιρέσει όλες τις απορριπτέες περιπτώσεις**. Δεν μετράει τις ρίψεις που οδηγούν σε αποτυχία, μόνο τις ρίψεις μέσα στον νέο χώρο των έγκυρων ακολουθιών.

Η συγκεκριμένη άσκηση ήταν λίγο περίπλοκη, δεν θα πέσει κάτι παρόμοιο σε εξέταση. Την κάναμε κυρίως για να δούμε πώς λειτουργούν οι δεσμευμένες πιθανότητες, οι οποίες μπορεί να είναι περίπλοκες και όχι πάντα προφανείς.

Άσκηση3: Έστω ότι 100 άτομα είναι προγραμματισμένο να πάνε σε έναν χορό, σχηματίζοντας 50 ζευγάρια. Στην καλύτερη περίπτωση, αν όλα τα άτομα εμφανιστούν, θα έχουμε 50 ζευγάρια παρόντα στον χορό. Ωστόσο, κάθε άτομο θα εμφανιστεί με πιθανότητα 90%. Κάποιο άτομο μπορεί να μην εμφανιστεί λόγω απρόβλεπτων καταστάσεων.

Ποια είναι η **αναμενόμενη τιμή** του πλήθους των ζευγαριών τα οποία θα εμφανιστούν και οι δύο στον χορό;

Ορίζουμε τυχαίες μεταβλητές:

- **X**: αντιπροσωπεύει τον συνολικό αριθμό των ζευγαριών που θα εμφανιστούν στον χορό και τα δύο μέλη τους. Δηλαδή, αν ένα ζευγάρι εμφανιστεί πλήρως, τότε αυτό προσμετράται στον X.
- **xi**: Για κάθε ζευγάρι i (όπου i=1,2,…,50), ορίζουμε την τυχαία μεταβλητή xi που δηλώνει αν το συγκεκριμένο ζευγάρι i εμφανίστηκε πλήρως στον χορό ή όχι:

$$x_i = \begin{cases} 1 & \text{αν και τα δύο μέλη του ζευγαριού } i \text{ εμφανιστούν} \\ 0 & \text{αν κάποιο ή και τα δύο μέλη του ζευγαριού } i \text{ δεν εμφανιστούν} \end{cases}$$

Το xi είναι μια δυαδική τυχαία μεταβλητή (Bernoulli), που παίρνει την τιμή 1 με κάποια πιθανότητα και την τιμή 0 με την υπόλοιπη πιθανότητα.

Κάθε άτομο έχει πιθανότητα 90% να εμφανιστεί στον χορό. Η πιθανότητα να εμφανιστούν και τα δύο μέλη ενός ζευγαριού είναι το γινόμενο των πιθανοτήτων εμφάνισης κάθε ατόμου, δηλαδή:

$$\mathbb{P}(xi = 1) = 0.9 \cdot 0.9 = 0.81$$

Άρα, η πιθανότητα να εμφανιστεί πλήρως ένα ζευγάρι στον χορό είναι 0.81, και η πιθανότητα να μην εμφανιστεί τουλάχιστον ένα από τα δύο μέλη είναι:

$$\mathbb{P}(xi = 0) = 1 - 0.81 = 0.19$$

Η τυχαία μεταβλητή X, που μετράει τον συνολικό αριθμό των ζευγαριών που εμφανίζονται και τα δύο μέλη τους, είναι το άθροισμα των xi για τα 50 ζευγάρια:

$$X = x_1 + x_2 + x_3 + \ldots + x_{50}$$

Η αναμενόμενη τιμή της τυχαίας μεταβλητής X είναι το άθροισμα των αναμενόμενων τιμών των επιμέρους μεταβλητών xi λόγω της ιδιότητας της γραμμικότητας

$$E[X] = E[x_1] + E[x_2] + E[x_3] + \ldots + E[x_{50}]$$

Για κάθε ζευγάρι i, η αναμενόμενη τιμή E[xi] είναι ίση με την πιθανότητα να εμφανιστεί το ζευγάρι, δηλαδή 0.81 καθώς E[χi] = 0 * 0.19 + 1 * 0.81 = 0.81 από ορισμό αναμενόμενης τιμής. Επομένως, υπολογίζουμε ότι:

$$E[X] = 50 * 0.81 = 40.5$$

Και αν το σκεφτούμε διαισθητικά, είναι λογικό. Εφόσον κάθε ζευγάρι έχει πιθανότητα 0.81 να εμφανιστεί στον χορό (δηλαδή να εμφανιστούν και τα δύο μέλη), αν έχουν προγραμματιστεί να έρθουν 50 ζευγάρια, τότε αναμένουμε ότι περίπου το 81% από αυτά τα ζευγάρια θα εμφανιστούν.

Συνεχίζοντας, θέλω να μάθουμε μία χρήσιμη ανισότητα η οποία θα μας χρησιμεύσει αργότερα. Η **Ανισότητα Markov** είναι ένα εργαλείο που μας επιτρέπει να **εκτιμήσουμε** την πιθανότητα μια **μη αρνητική** τυχαία μεταβλητή X να πάρει μεγάλες τιμές.

Συγκεκριμένα, λέει το εξής:

$$P(X \geq a) \leq \frac{E[X]}{a}$$ ⭐

Όπου:

- α είναι μια θετική σταθερά,
- E[X] είναι η αναμενόμενη τιμή της τυχαίας μεταβλητής X.

Έστω μεταβλητή X που αντιπροσωπεύει το ποσό χρημάτων που μπορείς να κερδίσεις. Η **αναμενόμενη τιμή** των χρημάτων που κερδίζεις είναι 10 ευρώ. Αν σου πουν ότι η πιθανότητα να κερδίσεις πάνω από 20 ευρώ (δηλαδή X ≥ 20) είναι μεγαλύτερη από 10 / 20 = 0.5, τότε κάτι δεν πάει καλά!

Αυτό γιατί αν η πιθανότητα να κερδίσεις πάνω από 20 ευρώ ήταν μεγαλύτερη από 50%, η μέση τιμή (αναμενόμενη τιμή) δεν θα μπορούσε να είναι 10 ευρώ. Θα έπρεπε να είναι πολύ μεγαλύτερη, επειδή οι τιμές που βρίσκονται πάνω από 20 ευρώ θα "τραβούσαν" τη μέση τιμή προς τα πάνω.

Αυτό ακριβώς εξηγεί η Ανισότητα Markov: αν η πιθανότητα να έχεις μεγάλες τιμές ήταν μεγαλύτερη από το E[X] / α, τότε η αναμενόμενη τιμή E[X] θα έπρεπε να είναι μεγαλύτερη από ό,τι είναι.

Μέχρι στιγμής έχουμε δει την έννοια της **αναμενόμενης τιμής**, που μας βοηθά να καταλάβουμε ποια τιμή αναμένουμε κατά μέσο όρο από μια τυχαία μεταβλητή. Ωστόσο, η αναμενόμενη τιμή από μόνη της δεν μας δίνει την πλήρη εικόνα. Πόσο μακριά μπορεί να βρίσκονται οι πραγματικές τιμές από την αναμενόμενη τιμή; Μας ενδιαφέρει μόνο ο μέσος όρος ή και η απόκλιση από αυτόν;

Σκεφτείτε ότι προσπαθείτε να προβλέψετε τα **μηνιαία έξοδα θέρμανσης** για το χειμώνα. Κάποιος σας λέει ότι, κατά μέσο όρο, τα έξοδά σας για θέρμανση θα είναι περίπου 150 ευρώ τον μήνα. Αυτή είναι η **αναμενόμενη τιμή**. Ακούγεται καλό, αλλά είναι αρκετό αυτό το νούμερο για να καταλάβετε πόσα πρέπει να προϋπολογίσετε;

Ας εξετάσουμε δύο σενάρια:

- Στο πρώτο σενάριο, τα έξοδα θέρμανσης σας είναι συνήθως 140, 150, ή 160 ευρώ τον μήνα. Δηλαδή, κυμαίνονται κοντά στα 150 ευρώ.
- Στο δεύτερο σενάριο, τα έξοδα θέρμανσης μπορεί να είναι 50 ευρώ σε έναν μήνα και 250 ευρώ στον επόμενο, παρόλο που ο μέσος όρος εξακολουθεί να είναι 150 ευρώ.

Και στα δύο σενάρια η αναμενόμενη τιμή των εξόδων είναι 150 ευρώ, αλλά η **διασπορά** των τιμών είναι εντελώς διαφορετική. Στο πρώτο σενάριο, τα έξοδα παραμένουν κοντά στην αναμενόμενη τιμή, οπότε δεν υπάρχει μεγάλη αβεβαιότητα. Στο δεύτερο σενάριο, οι αποκλίσεις από τα 150 ευρώ μπορεί να είναι πολύ μεγάλες, και αυτό μας δημιουργεί μεγαλύτερη αβεβαιότητα για το πόσο θα κοστίσει η θέρμανση κάθε μήνα.

Εδώ είναι που η **διασπορά** γίνεται σημαντική. Η διασπορά μετρά πόσο "διασκορπισμένες" είναι οι τιμές γύρω από την αναμενόμενη τιμή.

Στο πρώτο σενάριο, η διασπορά είναι μικρή, οπότε ξέρουμε ότι τα έξοδα θα είναι σταθερά γύρω από τα 150 ευρώ. Στο δεύτερο σενάριο, η διασπορά είναι μεγάλη, και αυτό σημαίνει ότι μπορεί να χρειαστούμε περισσότερα χρήματα σε ορισμένους μήνες για να καλύψουμε τα έξοδα θέρμανσης.

Η αναμενόμενη τιμή μας λέει ποια είναι η "μέση" τιμή, αλλά αν δεν γνωρίζουμε τη διασπορά, δεν μπορούμε να προετοιμαστούμε για πιθανές αποκλίσεις. Σε καταστάσεις όπου υπάρχει αβεβαιότητα ή αστάθεια (όπως τα μηνιαία έξοδα), η διασπορά μας βοηθά να προβλέψουμε πόσο μεγάλη μπορεί να είναι η διακύμανση από τον μέσο όρο. Στο παράδειγμα των εξόδων θέρμανσης, αυτό είναι σημαντικό γιατί μας επιτρέπει να υπολογίσουμε αν χρειαζόμαστε περισσότερα χρήματα σε κάποιο μήνα, πέρα από την αναμενόμενη τιμή των 150 ευρώ.

Η διασπορά μετράει πόσο αναμένεται να αποκλίνουν οι τιμές μιας τυχαίας μεταβλητής από την αναμενόμενη τιμή της. Ο τύπος για τη διασπορά είναι:

$$\textbf{Var(X)} = E[\,(X - E[X])^2\,] = \sum_x (x - E[X])^2\, \mathbb{P}(X = x)$$ ⭐

Το X − E[X] μας δείχνει την απόκλιση της κάθε πιθανής τιμής της X από τον μέσο όρο της. Αν η απόκλιση είναι μεγάλη, τότε η τιμή X είναι μακριά από τον μέσο όρο, ενώ αν η απόκλιση είναι μικρή, η τιμή είναι κοντά στον μέσο όρο.

Το τετράγωνο της απόκλισης χρησιμοποιείται για να **εξαλείφει τα αρνητικά πρόσημα**. Διασφαλίζει ότι οι αποκλίσεις είναι πάντα θετικές. Αυτό μας επιτρέπει να μετρήσουμε το μέγεθος της απόκλισης χωρίς να ανησυχούμε για την κατεύθυνση της (αν δηλαδή η απόκλιση είναι προς τα πάνω ή προς τα κάτω).

Επίσης, όταν υψώνουμε την απόκλιση στο τετράγωνο, **δίνουμε μεγαλύτερη έμφαση σε μεγάλες αποκλίσεις**. Δηλαδή, μεγάλες αποκλίσεις από την αναμενόμενη τιμή αυξάνουν σημαντικά τη διασπορά, ενώ μικρές αποκλίσεις έχουν μικρότερη επίδραση. Αυτό είναι ιδιαίτερα χρήσιμο όταν θέλουμε να μετρήσουμε περιπτώσεις όπου υπάρχουν σπάνια αλλά μεγάλα άλματα στις τιμές. .

Θεωρητικά, θα μπορούσαμε να χρησιμοποιήσουμε μεγαλύτερες δυνάμεις, όπως τον κύβο ή την τέταρτη δύναμη, για να μετρήσουμε την απόκλιση. Όμως, το τετράγωνο έχει συγκεκριμένες μαθηματικές ιδιότητες που το κάνουν πολύ βολικό.

Για παράδειγμα, αν χρησιμοποιήσουμε τρίτη δύναμη, θα ξαναεμφανιστούν αρνητικά πρόσημα, καθώς οι αρνητικές αποκλίσεις από τον μέσο όρο θα υψωθούν σε περιττή δύναμη και θα παραμείνουν αρνητικές, κάτι που θέλουμε να αποφύγουμε.

Η τέταρτη δύναμη και υψηλότερες δυνάμεις θα έδιναν **ακόμη μεγαλύτερη έμφαση στις μεγάλες αποκλίσεις**, αλλά αυτό θα μπορούσε να κάνει τη μέτρηση υπερβολικά ευαίσθητη σε σπάνιες, ακραίες τιμές. Το τετράγωνο θεωρείται ένας καλός συμβιβασμός: καταγράφει τη μεταβλητότητα χωρίς να δίνει υπερβολική έμφαση σε ακραίες τιμές.

Παράδειγμα: Ας υποθέσουμε ότι προσπαθείτε να υπολογίσετε τα μηνιαία έξοδα θέρμανσης για το χειμώνα. Γνωρίζετε ότι υπάρχουν τρία πιθανά σενάρια για τα μηνιαία έξοδά σας, με τις αντίστοιχες πιθανότητες:

- 100 ευρώ με πιθανότητα 0.2
- 150 ευρώ με πιθανότητα 0.5
- 250 ευρώ με πιθανότητα 0.3

Η **αναμενόμενη τιμή** E[X] εκφράζει την αναμενόμενη τιμή των εξόδων, και υπολογίζεται ως το σταθμισμένο άθροισμα των τιμών, πολλαπλασιασμένων με τις αντίστοιχες πιθανότητες:

E[X] = (100 * 0.2) + (150 * 0.5) + (250 * 0.3) = 20 + 75 + 75 = 170 ευρω

Αυτό σημαίνει ότι αναμένεται να πληρώνετε 170 ευρώ τον μήνα για θέρμανση.

Ωστόσο, η αναμενόμενη τιμή **δεν μας λέει πόσο μπορεί να διαφέρει ο πραγματικός λογαριασμός από μήνα σε μήνα**. Για να καταλάβουμε τη μεταβλητότητα αυτών των εξόδων, πρέπει να υπολογίσουμε τη **διασπορά**.

Var(X) = (100-170)^2 * 0.2 + (150-170)^2 * 0.5 + (250- 70)^2 * 0.3 = 3100 ευρω^2

Η διασπορά των μηνιαίων εξόδων θέρμανσης είναι **3100 ευρώ²**. Αυτό είναι ένα μέτρο σχετικό του πόσο αποκλίνουν τα έξοδα από την αναμενόμενη τιμή, αλλά το αποτέλεσμα είναι σε τετραγωνικές μονάδες (ευρώ²), κάτι που δεν είναι άμεσα κατανοητό.

Για να κατανοήσουμε καλύτερα τι σημαίνει αυτή η διασπορά, μπορούμε να υπολογίσουμε την **τυπική απόκλιση**. Η τυπική απόκλιση δίνει μια μέτρηση της "τυπικής" απόκλισης των τιμών γύρω από τον μέσο όρο. Η τυπική απόκλιση **δεν** αντιστοιχεί στην **πραγματική μέση τιμή των αποκλίσεων**.

$$\sigma(X) = \sqrt{Var(X)} = \sqrt{3100} = 55.68 \text{ ευρώ}$$ ⭐

Αν και **είναι διαφορετική έννοια από τη πραγματική μέση απόλυτη απόκλιση**, που θα μπορούσαμε να υπολογίσουμε ως $E[\,|X - E[X]|\,]$, η τυπική απόκλιση χρησιμοποιείται περισσότερο γιατί έχει συγκεκριμένες μαθηματικές ιδιότητες που διευκολύνουν τους υπολογισμούς και την ανάλυση.

Η τυπική απόκλιση μπορεί να αποκαλύψει διάφορα στοιχεία σχετικά με τη μεταβλητότητα της μεταβλητής μας. Για παράδειγμα, ένας τρόπος να παρέχει πληροφορίες για την **πραγματική μέση απόκλιση** είναι μέσω της **ανισότητας του Chebyshev**, η οποία αναφέρει πως:

$$P(|X - E[X]| > k \cdot \sigma) \leq \frac{1}{k^2}$$ ⭐

Από αυτή μπορούμε να συμπεράνουμε ότι η πιθανότητα να έχουμε απόκλιση μεγαλύτερη των 1.5 * 55.68 = 83.52 ευρώ είναι μικρότερη η ίση με 44.44%. Για k = 1 από την άλλη, η ανισότητα δεν προσφέρει καμία πληροφορία, αφού απλά λέει ότι η πιθανότητα να έχουμε απόκλιση μεγαλύτερων των 55.68 ευρώ είναι μικρότερη η ίση με 100%.

Για k μικρότερο η ίσο του 1 θα παρατηρήσεις ότι το αποτέλεσμα δεν είναι πολύ ενδιαφέρον. Δεν είναι απαραίτητα λάθος, απλά δεν προσφέρει καμία πληροφορία. Για παράδειγμα, για k = 0.5 η ανισότητα εκφράζει πως η πιθανότητα θα είναι μικρότερη η ίση από 4, το οποίο έτσι και αλλιώς είναι προφανές αφού σε πιθανότητες η μεγαλύτερη τιμή είναι η μονάδα.

Ας δούμε από που βγαίνει η ανισότητα:

Ξεκινάμε από την πιθανότητα που θέλουμε να εκτιμήσουμε:

$$P(|X - E[X]| \geq k \cdot \sigma) = P(|x - E[X]|^2 \geq k^2 \cdot \sigma^2)$$

Αυτή η εξίσωση προκύπτει από την ύψωση των δύο μελών στο τετράγωνο, ώστε να αφαιρέσουμε την απόλυτη τιμή.

Εφαρμόζουμε την **Ανισότητα Markov** στην τυχαία μεταβλητή **(X − E[X])^2**, η οποία είναι μη αρνητική. Σύμφωνα με την ανισότητα Markov:

$$P\left((X - E[X])^2 \geq k^2 \cdot \sigma^2\right) \leq \frac{\mathrm{E}\left[(\mathrm{X}-\mathrm{E}[\mathrm{X}])^2\right]}{k^2 \cdot \sigma^2}$$

Θυμόμαστε ότι **E[(X − E[X])^2]** είναι η διασπορά **V(X) = σ^2**.

Άρα, η εξίσωση γίνεται:

$$P\left((X - E[X])^2 \geq k^2 \cdot \sigma^2\right) \leq \frac{\sigma^2}{k^2 \cdot \sigma^2}$$

Απλοποιώντας τα σ^2, καταλήγουμε στην **ανισότητα του Chebyshev**:

$$P\left(|X - E[X]| \geq k \cdot \sigma\right) \leq \frac{1}{k^2}$$

Ας συνεχίσουμε βλέποντας κάποιες χρήσιμες ιδιότητες της διασποράς:
Ο υπολογισμός της διασποράς μπορει να γίνει πιο απλός με τον τύπο:

Var(X) = E(X^2) - E(X)^2 ⭐

Το παραπάνω προκύπτει από τις γνωστές ιδιότητες της αναμενόμενης τιμής:
Θέτω **E[X] = μ** για να είναι πιο φανερό ότι αναφέρεται σε μια σταθερά.

$$Var(X) = E[(X - \mu)^2]$$

$$= E[X^2 - 2 \cdot X \cdot \mu + \mu^2]$$

$$= E[X^2] - E[2 \cdot X \cdot \mu] + \mu^2$$

$$= E[X^2] - 2 \cdot \mu \cdot E[X] + \mu^2$$

$$= E[X^2] - 2 \cdot E[X]^2 + E[X]^2$$

$$= E[X^2] - E[X]^2$$

Άλλη μία χρήσιμη ιδιότητα είναι η:

Var(X + c) = Var(X) ⭐

Για να δείξουμε το παραπάνω, να θυμίσω πως E[x + c] = E[x] + c, και οπότε εύκολα με πράξεις δείχνεις την ισότητα:

$$Var(X + c) = E[(X + c - E[X + c])^2]$$

$$= E[(X + c - c - E[X])^2]$$

$$= E[(X - E[X])^2]$$

$$= Var(X)$$

Μια τελευταία (για τώρα) ιδιότητα είναι η:

Var(c * X) = c^2 Var(X) ⭐

Η οποία μπορεί να αποδειχθεί εύκολα ως εξής:

$$Var(c \cdot X) = E[(c \cdot X - E[c \cdot X])^2]$$

$$= E[(c \cdot X - c \cdot E[X])^2]$$

$$= E[c^2 \cdot (X - E[X])^2]$$

$$= c^2 \cdot E[(X - E[X])^2]$$

$$= c^2 \cdot Var(X)$$

Μένει ακόμα να μάθεις το πως μπορείς να υπολογίσεις το **Var(X + Y)** δύο τυχαίων μεταβλητών. Θα δείξω πως ο υπολογισμός εξαρτάται από το αν οι δύο είναι ανεξάρτητες η όχι. Πριν από αυτό όμως θα χρειαστεί να εισάγω πρώτα την εννοια της **συνδιακύμανση του X και Y.**

Η **συνδιακύμανση Cov(X,Y)** είναι ένα μέτρο του πώς οι τυχαίες μεταβλητές X και Y μεταβάλλονται από κοινού. Δηλαδή, μας δείχνει πώς οι τιμές της X σχετίζονται με τις τιμές της Y.

Αν οι τιμές της μίας μεταβλητής αυξάνονται όταν αυξάνονται οι τιμές της άλλης, τότε η συνδιακύμανση θα είναι θετική. Αν οι τιμές της μίας μεταβλητής αυξάνονται όταν οι τιμές της άλλης μειώνονται, τότε η συνδιακύμανση θα είναι αρνητική.

Ο τύπος για τη συνδιακύμανση είναι:

Cov(X, Y) $= E[(X - E[X]) \cdot (Y - E[Y])]$ ⭐

$$= \sum_x \sum_y (x - E[X]) \cdot (Y - E[Y]) \cdot P(X = x, Y = y)$$

Η συνδιακύμανση, με απλά λόγια, μας δείχνει αν και πως οι αλλαγές στην τυχαία μεταβλητή X επηρεάζουν την τυχαία μεταβλητή Y. Αυτό συνδέεται άμεσα με το εάν οι δύο μεταβλητές είναι εξαρτημένες ή ανεξάρτητες.

Ειδικότερα, αν οι μεταβλητές είναι ανεξάρτητες, η συνδιακύμανση τους θα είναι ίση με μηδέν, αφού οι τιμές της μίας δεν θα επηρεάζουν τις τιμές της άλλης.

- Αν **Cov(X,Y) >** 0, τότε οι X και Y έχουν θετική γραμμική συσχέτιση. Αυτό σημαίνει ότι όταν η X αυξάνεται, η Y τείνει να αυξάνεται επίσης.
- Αν **Cov(X,Y) < 0**, τότε υπάρχει αρνητική γραμμική συσχέτιση. Δηλαδή, όταν η X αυξάνεται, η Y τείνει να μειώνεται.
- Αν **Cov(X,Y) = 0**, αυτό υποδηλώνει ότι οι αλλαγές στην X δεν σχετίζονται γραμμικά με τις αλλαγές στην Y. Δεν σημαίνει απαραίτητα ότι δεν υπάρχει καθόλου σχέση μεταξύ τους, αλλά ότι η σχέση τους δεν είναι γραμμική.

Αν οι X και Y είναι ανεξάρτητες, τότε η συνδιακύμανση Cov(X,Y) = 0. Ωστόσο, το αντίθετο δεν ισχύει πάντα. Δηλαδή, αν Cov(X,Y) = 0, δεν μπορούμε να συμπεράνουμε ότι οι X και Y είναι ανεξάρτητες.

Αποδειξη πως αν X, Y ανεξάρτητες τότες Cov(X, Y) = 0:

$$\text{Cov(X, Y)} = E[\,(X - \mu[X]) \cdot (Y - \mu[Y])\,]$$

$$= E[\,(X \cdot Y - X \cdot \mu[Y] - Y \cdot \mu[X] + \mu[X] \cdot \mu[Y]\,]$$

$$= E[X \cdot Y] - \mu[Y] \cdot E[X] - \mu[X] \cdot E[Y] + \mu[X] \cdot \mu[Y]$$

$$= E[X] \cdot E[Y] - E[Y] \cdot E[X] - E[X] \cdot E[Y] + E[X] \cdot E[Y]$$

$$= 2 \cdot E[X] \cdot E[Y] - 2 \cdot E[X] \cdot E[Y] = 0$$

Το ότι είναι ανεξάρτητες μας επέτρεψε να πούμε ότι E[X * Y] = E[X] * E[Y] και έτσι να καταλήξουμε σε ισότητα με το 0.

Δεν θα εξηγήσω περισσότερο την θεωρία πίσω από τον τύπο, καθώς αυτή έχει να κάνει με θεωρία στατιστικής και δεν χρειάζεται για το συγκεκριμένο μάθημα. Αφήνω όμως ένα ενδιαφέρον βίντεο πάνω στο θέμα να το δει όποιος ενδιαφέρεται: Covariance, Clearly Explained!!!

Καθώς το βλέπεις αυτό, κράτησε στο μυαλό σου όσα έχουμε εξηγήσει μέχρι τώρα. Κάνοντας Ν πειράματα για τις μεταβλητές (X,Y), θα πάρεις ζευγάρια τιμών με κάποια πιθανότητα P(X,Y). Για κάθε ζευγάρι, θα υπολογίσουμε την απόκλιση της X από την αναμενόμενη τιμή της E[X], και αντίστοιχα, την απόκλιση της Y από την E[Y].

Πολλαπλασιάζουμε αυτές τις αποκλίσεις για να δούμε τη σχέση τους. Αν οι αποκλίσεις έχουν το ίδιο πρόσημο (δηλαδή και οι δύο τιμές είναι είτε μεγαλύτερες είτε μικρότερες από τις αναμενόμενες τιμές τους), τότε το γινόμενο είναι θετικό. Αν έχουν διαφορετικά πρόσημα (δηλαδή όταν η μία τιμή είναι πάνω από την αναμενόμενη τιμή και η άλλη κάτω), το γινόμενο είναι αρνητικό.

Το πιο σημαντικό από το αποτέλεσμα είναι το αν η συνδιακύμανση είναι θετική, αρνητική ή μηδενική. Το **ακριβές νούμερο** της συνδιακύμανσης από μόνο του **δεν** μας δίνει μια εύκολη ερμηνεία για το πόσο η μία μεταβλητή επηρεάζει την άλλη, όπως ίσως θα περίμενες.

Έχοντας μάθει τι είναι η συνδιακύμανση τώρα μπορώ να ορίσω την ιδιότητα:

Var(X + Y) = Var(X) + Var(Y) + 2 Cov(X, Y) ⭐

Όταν οι μεταβλητές X και Y είναι ανεξάρτητες: Εφόσον Cov(X,Y) θα είναι 0, ο τύπος απλοποιείται σε: **Var(X + Y) = Var(X) + Var(Y)**

Απόδειξη του τύπου:

$$Var(X + Y) = E[\,(X + Y - E[X + Y]\,)^2\,]$$

$$= E[\,(\,(X - E[X]) + (Y - E[Y])\,)^2\,]$$

$$= E[\,(X - E[X])^2 + 2\,(X - E[X])\,(Y - E[Y]) + (Y - E[Y])^2\,]$$

$$= E[\,(X - E[X])^2\,] + 2\,E[\,(X - E[X])\,(Y - E[Y])\,] + E[\,(Y - E[Y])^2\,]$$

$$= Var(X) + 2\,Cov(X, Y) + Var(Y)$$

Μέχρι τώρα, έχουμε ασχοληθεί με **διακριτές τυχαίες μεταβλητές**, όπου η αναμενόμενη τιμή και η διασπορά υπολογίζονται μέσω **αθροισμάτων**.

Στις συνεχείς μεταβλητές, η βασική διαφορά είναι ότι δεν μπορούμε να μιλήσουμε για πιθανότητες μεμονωμένων τιμών (καθώς αυτές είναι μηδέν), αλλά για **πυκνότητες πιθανότητας**. Για αυτόν τον λόγο, οι βασικοί τύποι της αναμενόμενης τιμής και της διασποράς για συνεχείς μεταβλητές χρησιμοποιούν **ολοκληρώματα** αντί για αθροίσματα.

Για μια **συνεχή τυχαία μεταβλητή** Χ με **συνάρτηση πυκνότητας πιθανότητας fχ(x)**, η **αναμενόμενη τιμή** (ή μέση τιμή) ορίζεται ως:

$$E[X] = \int_{-\infty}^{\infty} x \cdot fx(x)\, dx$$ ⭐

Αντί για το άθροισμα που χρησιμοποιούμε στις διακριτές μεταβλητές, εδώ παίρνουμε το **ολοκλήρωμα** πάνω από όλες τις δυνατές τιμές του x, και ζυγίζουμε την τιμή x με την πυκνότητα πιθανότητας fx(x). Το ολοκλήρωμα αυτό μας δίνει τον μέσο όρο των τιμών που μπορεί να πάρει η μεταβλητή Χ.

Η **διασπορά** μιας συνεχούς τυχαίας μεταβλητής ορίζεται επίσης με ολοκλήρωμα. Ο τύπος είναι παρόμοιος με αυτόν που είδαμε για τις διακριτές μεταβλητές, αλλά τώρα ολοκληρώνουμε αντί να αθροίζουμε. Εναλλακτικά, μπορείς να χρησιμοποιήσεις και αμέσως την ιδιότητα **Var(x) = E[X^2] - E[X]^2** που είδαμε πριν, αν αυτό βολεύει περισσότερο (συνήθως είναι πιο γρήγορο).

Παρόλο που για τις **συνεχείς τυχαίες μεταβλητές** χρησιμοποιούμε ολοκληρώματα αντί για αθροίσματα, οι ιδιότητες που έχουμε δει για τις διακριτές μεταβλητές **εξακολουθούν να ισχύουν** και για τις συνεχείς. Αυτό συμβαίνει επειδή το άθροισμα και το ολοκλήρωμα είναι μαθηματικά εργαλεία που σχετίζονται στενά, και χρησιμοποιούνται αντίστοιχα για διακριτά και συνεχή σύνολα τιμών.

Παράδειγμα1: Έστω ότι η τυχαία μεταβλητή Χ έχει ομοιόμορφη κατανομή στο διάστημα [a,b]. Αυτό σημαίνει ότι η πιθανότητα η Χ να παίρνει οποιαδήποτε τιμή μέσα στο διάστημα είναι σταθερή. Ζητείται να υπολογιστεί η αναμενόμενη τιμή της Χ.

Αφού η κατανομή είναι ομοιόμορφη στο διάστημα [a,b], η **συνάρτηση πυκνότητας πιθανότητας** fx(x) πρέπει να είναι σταθερή εντός του διαστήματος [a,b].

Αυτή η συνάρτηση πυκνότητας ορίζεται έτσι ώστε το **ολοκλήρωμα** της να ισούται με 1 (δηλαδή η συνολική πιθανότητα να βρίσκεται η X εντός του διαστήματος να είναι 100%).

Επομένως, η συνάρτηση πυκνότητας πιθανότητας για την ομοιόμορφη κατανομή στο διάστημα [a,b] πρέπει να είναι ίση με:

$$fx(x) \ = \ \frac{1}{b-a} \ \text{για} \ x \ \in [a, \ b]$$

Καθώς μόνο τότε το ολοκλήρωμα θα ισούται με 1:

$$\int_a^b \frac{1}{b-a}\,dx \ = \left[\frac{x}{b-a}\right]_a^b \ = \frac{b-a}{b-a} = 1$$

Η αναμενόμενη τιμή E[X] μπορεί οπότες να υπολογιστεί ως:

$$E[X] \ = \int_a^b x \cdot \frac{1}{b-a}\,dx$$

$$= \frac{1}{b-a} \cdot \int_a^b x \ dx$$

$$= \frac{1}{b-a} \cdot \left[\frac{x^2}{2}\right]_a^b$$

$$= \frac{1}{b-a} \cdot \frac{b^2 - a^2}{2}$$

$$= \frac{b-a}{b-a} \cdot \frac{b+a}{2}$$

$$= \frac{a+b}{2}$$ ⭐

Διαισθητικά, αυτό το αποτέλεσμα είναι λογικό. Αφού η κάθε τιμή στο διάστημα έχει ίση πιθανότητα να επιλεγεί λόγω της ομοιόμορφης κατανομής, ο μέσος όρος των τιμών θα βρίσκεται στη μέση του διαστήματος, δηλαδή στο μέσο των άκρων a και b. Για αυτό, η αναμενόμενη τιμή **E[X]** ισούται με **(a + b) / 2**

Παράδειγμα2: Έστω ότι η τυχαία μεταβλητή X έχει ομοιόμορφη κατανομή στα διαστήματα [a,b] ∪ [c,d]. Ζητείται να υπολογιστεί η αναμενόμενη τιμή της X και της Y, όπου Y = 2X + 3

Για τον ίδιος λόγο που εξήγησα στο **παράδειγμα1**, η συνάρτηση πυκνότητας πιθανότητας fX(x) θα είναι:

$$fx(x) \;=\; \frac{1}{(b-a)+(d-c)} \;\; \text{για} \;\; x \;\in\; [a,\; b] \;\cup\; [c, d]$$

Για να βρούμε την αναμενόμενη τιμή της X, ξεκινάμε από τον τύπο του ολοκληρώματος για την αναμενόμενη τιμή:

$$E[X] \;= \int_a^b x \cdot \frac{1}{(b-a)+(d-c)}\, dx \;+ \int_c^d x \cdot \frac{1}{(b-a)+(d-c)}\, dx$$

$$= \frac{1}{(b-a)+(d-c)} \;\cdot \left(\int_a^b x \;\, dx \;+ \int_c^d x \;\, dx \right)$$

$$= \frac{1}{(b-a)+(d-c)} \;\cdot \left(\frac{b^2-a^2}{2} + \frac{c^2-d^2}{2} \right)$$

$$= \frac{a+b}{2} \cdot \frac{b-a}{(b-a)\;+(d-c)} \;+ \frac{c+d}{2} \cdot \frac{d-c}{(b-a)\;+(d-c)}$$

Ορίζουμε **L= (b−a) + (d−c)**, που εκφράζει το συνολικό μήκος των δύο διαστημάτων. Παράλληλα, **L1 = b − a** είναι το μήκος του πρώτου διαστήματος και **L2 = d − c** το μήκος του δεύτερου διαστήματος. Επιπλέον, η αναμενόμενη τιμή του X στο πρώτο διάστημα είναι **E[X | X∈[a,b]] = (a + b) / 2**, ενώ στο δεύτερο διάστημα, η αναμενόμενη τιμή είναι **E[X | X∈[c,d]] = (c + d) / 2**

Και έτσι έχουμε έναν γενικό τύπο που μπορείς να χρησιμοποιείς σε τέτοια προβλήματα για να κάνεις την ζωή σου πιο εύκολη: ⭐

$$E[X] = E[X \mid X \in [a,b]] \cdot P(X \in [a,b]) + E[X \mid X \in [c,d]] \cdot P(X \in [c,d])$$

Όπου **P(X ∈ [a,b]) = (L1 / L)** και **P(X ∈ [c,d]) = (L2 / L)** αφού η πιθανότητα κατανέμεται ομοιόμορφα σε όλο το μήκος L. Έτσι, όπως αναφέραμε και στην **ενότητα 2**, η πιθανότητα το X να πάρει τιμή εντός του πρώτου διαστήματος εξαρτάται από το ποσοστό του συνολικού χώρου που καλύπτει το διάστημα. Αν τα δύο διαστήματα είχαν το ίδιο μήκος, οι πιθανότητες θα ήταν 50-50..

Ο τύπος που έχουμε δώσει για την αναμενόμενη τιμή μπορεί να προσαρμοστεί εύκολα και σε προβλήματα με περισσότερα διαστήματα. Αν η X κατανέμεται ομοιόμορφα σε περισσότερα από δύο διαστήματα, ο συνολικός χώρος L θα είναι το άθροισμα των μηκών όλων των διαστημάτων. Για κάθε διάστημα, η πιθανότητα ότι X θα πάρει τιμή σε αυτό θα είναι το μήκος του διαστήματος διαιρούμενο με το συνολικό μήκος L. Η αναμενόμενη τιμή θα είναι το άθροισμα των μέσων των άκρων κάθε διαστήματος, σταθμισμένα με την πιθανότητα το X να πάρει τιμή στο διάστημα, όπως κάναμε και για τα δύο διαστήματα

Στη συνέχεια, πριν ολοκληρώσουμε αυτή την ενότητα, θα ήθελα να ορίσουμε μερικές βασικές κατανομές που θα σου φανούν χρήσιμες. Η κατανόησή τους είναι σημαντική, καθώς μπορούν να χρησιμοποιηθούν σε διάφορα προβλήματα που ενδέχεται να συναντήσεις στο μέλλον.

Στην αρχή της ενότητας, είχα ήδη κάνει μια σύντομη εισαγωγή στις κατανομές Bernoulli και Γεωμετρική. Ωστόσο, θα τις αναφέρω ξανά εδώ συνοπτικά, ώστε να τις έχουμε συγκεντρωμένες, μαζί με τις νέες κατανομές που θα περιγράψω.

1. Κατανομή Bernoulli

Η **Bernoulli** είναι η απλούστερη διακριτή κατανομή και περιγράφει ένα πείραμα με δύο πιθανά αποτελέσματα: "επιτυχία" με πιθανότητα p και "αποτυχία" με πιθανότητα 1 − p. Η τυχαία μεταβλητή X, που ακολουθεί την κατανομή Bernoulli, μπορεί να πάρει τις τιμές 0 και 1.