<u>Διαμερίσεις</u>

Συναντάμε δύο ειδών διαμερίσεις και μάλιστα οι Μπερτσεκάς-Τσιτσικλής παρουσιάζουν μόνο την πρώτη έννοια ενώ οι Hoel-Port-Stone παρουσιάζουν μόνο τη δεύτερη. Οι δύο έννοιες αφορούν διαφορετικά προβλήματα και οδηγούν στη διατύπωση και λύση διαφορετικού τύπου πιθανοθεωρητικών μοντέλων. Η πρώτη αφορά διωνυμικά και πολυωνυμικά προβλήματα ενώ η δεύτερη υπεργεωμετρικά.

<u>*Α. Διαμέριση Συνόλου σε Υποσύνολα με Καθορισμένο Πλήθος Στοιχείων*</u>
(Μπερτσεκάς-Τσιτσικλής)

Έστω σύνολο από *n* διακεκριμένα[1] μέλη. Επιθυμούμε να το διαμερίσουμε σε δύο διακεκριμένα υποσύνολα, το ένα με *k* μέλη και το άλλο με *n-k*. Με πόσους τρόπους μπορούμε να κάνουμε μια τέτοια διαμέριση αν δε μας ενδιαφέρει η σειρά (διάταξη) των μελών ενός συνόλου; Παράδειγμα: Να διαμεριστεί το $\{1,2,\ldots,10\}$ σε δύο υποσύνολα, το πρώτο με 7 μέλη και το δεύτερο με 3 μέλη. Μια τέτοια διαμέριση είναι, ας πούμε, η $\{2,4,5,6,7,8,10\}$, $\{1,3,9\}$. Γνωρίζουμε από την τάξη ότι ο αριθμός των τρόπων με τους οποίους μπορούμε να κάνουμε αυτή τη διαμέριση είναι ο αριθμός των συνδυασμών των *n* ανά *k*, δηλαδή ο *διωνυμικός συντελεστής* $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$. Έτσι στο παράδειγμα ο ζητούμενος χωρισμός μπορεί να πραγματοποιηθεί με $\binom{10}{7}$ τρόπους.

Την ιδέα αυτή γενικεύουμε τώρα ως εξής: Έστω σύνολο με *n* διακεκριμένα μέλη. Επιθυμούμε να το διαμερίσουμε σε *r* το πλήθος διακεκριμένα υποσύνολα, το πρώτο με k_1 μέλη, το δεύτερο με k_2 μέλη,…, το τελευταίο με k_r μέλη, όπου βέβαια θα πρέπει $k_1 + k_2 + \cdots k_r = n$. Με πόσους τρόπους μπορούμε να κάνουμε μια τέτοια διαμέριση αν δε μας ενδιαφέρει η σειρά (διάταξη) των μελών των συνόλων; Ο συλλογισμός της γενίκευσης είναι εύκολος: Από τα *n* μέλη τσιμπάμε k_1. Αυτό μπορεί να γίνει με $\binom{n}{k_1}$ τρόπους. Από τα απομένοντα $n-k_1$ μέλη τσιμπάμε k_2. Αυτό μπορεί να γίνει με $\binom{n-k_1}{k_2}$ τρόπους. Συνεχίζουμε έτσι έως να φτάσουμε στον τελευταίο συντελεστή $\binom{n-k_1-\cdots-k_{r-1}}{k_r}$ (=1). Προφανώς, λόγω του πολλαπλασιαστικού κανόνα, ο συνολικός αριθμός τρόπων θα είναι το γινόμενο

[1] Όταν λέμε ότι μια συλλογή αποτελείται από διακεκριμένα μέλη εννοούμε ότι τα μέλη αυτά τα ξεχωρίζουμε (π.χ. έχουν ονόματα, είναι αριθμημένα, κ.λπ.)

$$\binom{n}{k_1}\binom{n-k_1}{k_2}\cdots\binom{n-k_1-\cdots-k_{r-1}}{k_r}=\frac{n!}{k_1!(n-k_1)!}\frac{(n-k_1)!}{k_2!(n-k_1-k_2)!}\cdots\frac{(n-k_1-k_2-\cdots-k_{r-1})!}{k_r!(n-k_1-k_2-\cdots k_r)!}$$

Οι όροι του αριθμητή από το δεύτερο και μετά απλοποιούνται με το 2^o, 4^o, κ.ο.κ. όρο του παρονομαστή. Ο τελευταίος όρος του παρονομαστή είναι 0!, δηλαδή 1, και καταλήγουμε έτσι στην απάντηση $\frac{n!}{k_1!k_2!\cdots k_r!}$.

Ο παραπάνω συντελεστής γενικεύει το διωνυμικό συντελεστή και ονομάζεται *πολυωνυμικός συντελεστής*.

Παραδείγματα

(1) Σε n ρίψεις ενός νομίσματος πόσες διαφορετικές ακολουθίες μπορώ να πάρω με k κεφάλια; Το πρόβλημα είναι ισοδύναμο με τη διαμέριση του συνόλου $\{1,2,\ldots,n\}$, όπου το σημείο $i, i=1,\ldots,n$, εκπροσωπεί την i ρίψη, σε δύο υποσύνολα, το ένα με k μέλη (οι ρίψεις με αποτέλεσμα Κ) και το άλλο με n-k μέλη (οι ρίψεις με αποτέλεσμα Γ). Επομένως η απάντηση είναι $\binom{n}{k}$.

Αν οι ρίψεις είναι ανεξάρτητες με πιθανότητα p σε μία ρίψη να έρθει Κ, τότε κάθε τέτοια ακολουθία θα έχει πιθανότητα να εμφανιστεί ίση με $p^k(1-p)^{n-k}$. Θεωρώντας ως δειγματικό χώρο όλες τις ακολουθίες (λέξεις) μήκους n με γράμματα τα Κ και Γ, τώρα θέλω να υπολογίσω την πιθανότητα σε μία ακολουθία τέτοιων ρίψεων να φέρω k κεφάλια. Σύμφωνα με τη συζήτησή μας, αυτό το ενδεχόμενο, ας το ονομάσουμε A, απαρτίζεται από $\binom{n}{k}$ στοιχεία (ακολουθίες μήκους n), το κάθε ένα με πιθανότητα $p^k(1-p)^{n-k}$. Άρα, η απάντηση θα είναι $P(A)=\binom{n}{k}p^k(1-p)^{n-k}$.

(2) Σε n ρίψεις ενός ζαριού πόσες διαφορετικές ακολουθίες μπορώ να πάρω με k_1 άσσους και k_2 δυάρια; Πάλι έχω να διαμερίσω το $\{1,2,\ldots,n\}$ σε υποσύνολα, τρία τώρα. Το πρώτο υποσύνολο θα είναι οι ρίψεις με αποτέλεσμα 1, το δεύτερο οι ρίψεις με αποτέλεσμα 2 και το τρίτο οι ρίψεις με αποτέλεσμα διάφορο των 1 και 2. Άρα, ο ζητούμενος αριθμός είναι $\frac{n!}{k_1!k_2!(n-k_1-k_2)!}$.

Αν οι ρίψεις είναι ανεξάρτητες και το ζάρι αμερόληπτο, ας υπολογίσουμε την πιθανότητα σε 100 ρίψεις να φέρω 10 άσσους και 10 δυάρια. Τώρα, η πιθανότητα μιας λέξης που περιέχει 10 άσσους, 10 δυάρια και 80 στοιχεία από το σύνολο $\{2,3,4,5,6\}$ θα είναι

$\left(\frac{1}{6}\right)^{10}\left(\frac{1}{6}\right)^{10}\left(\frac{4}{6}\right)^{80}$ (προσέξτε ότι ο δειγματικός χώρος παύει να είναι ομοιόμορφος αφού η πιθανότητα ρίψεως από την τρίτη κατηγορία είναι 4/6). Με συλλογισμούς ανάλογους του Παραδείγματος (1), συμπεραίνω ότι η ζητούμενη πιθανότητα είναι ίση με $\frac{100!}{10!10!80!}\left(\frac{1}{6}\right)^{10}\left(\frac{1}{6}\right)^{10}\left(\frac{4}{6}\right)^{80}$.

(3) (Αναγραμματισμοί-παράδειγμα από το βιβλίο των Μπερτσεκά-Τσιτσικλή) Πόσες διαφορετικές λέξεις (ακολουθίες γραμμάτων) μπορώ να φτιάξω με τα γράμματα της λέξης TATTOO. Έχουμε 6 μέλη στο σύνολο (1^o γράμμα της λέξης, 2^o γράμμα, κ.λπ.) και θέλουμε να το διαμερίσουμε σε 3 υποσύνολα, το πρώτο με 3 στοιχεία (γράμμα T), το δεύτερο με 2 στοιχεία (γράμμα O) και το τρίτο με 1 στοιχείο (γράμμα A). Σύμφωνα με τη συζήτηση, ο αριθμός αυτός θα είναι ίσος με $\frac{6!}{3!2!1!} = 60$.

Β. Δειγματοληψία Χωρίς Επανάθεση από Έναν Πληθυσμό Χωρισμένο σε Κατηγορίες (Hoel-Port-Stone)

Έστω σύνολο («ο πληθυσμός») από k μέλη τα οποία χωρίζονται σε r κατηγορίες. Η πρώτη κατηγορία απαρτίζεται από k_1 μέλη, η δεύτερη από k_2 μέλη, κ.ο.κ., με την τελευταία να απαρτίζεται από k_r μέλη, όπου βέβαια θα πρέπει $k_1 + k_2 + \cdots k_r = k$. Υποθέτοντας ότι η σειρά με την οποία εμφανίζονται τα μέλη στο δείγμα μας δεν ενδιαφέρει, το ερώτημα είναι με πόσους τρόπους μπορούμε να πάρουμε δείγμα χωρίς επανάθεση μεγέθους n το οποίο να περιέχει n_1 μέλη από την πρώτη κατηγορία, n_2 μέλη από τη δεύτερη κατηγορία, κ.ο.κ., μέχρι n_r μέλη από την τελευταία κατηγορία, όπου βέβαια θα πρέπει $0 \le n_i \le k_i, i = 1, \ldots, r$ και $n_1 + n_2 + \cdots n_r = n$; Η απάντηση προκύπτει ως εξής: Αφού δε μας ενδιαφέρει η διάταξη, ο αριθμός των τρόπων με τους οποίους μπορούμε να πάρουμε n_1 μέλη από την πρώτη κατηγορία είναι $\binom{k_1}{n_1}$. Όμοια, ο αριθμός των τρόπων με τους οποίους μπορούμε να πάρουμε n_2 μέλη από τη δεύτερη κατηγορία είναι $\binom{k_2}{n_2}$, κ.ο.κ. Σύμφωνα με τον πολλαπλασιαστικό κανόνα, ο συνολικός αριθμός τρόπων βρίσκεται συνδυάζοντας κάθε έναν από τους τρόπους για την πρώτη κατηγορία με κάθε έναν από τους τρόπους για τη δεύτερη κατηγορία, κ.ο.κ. Άρα, η απάντηση θα είναι $\binom{k_1}{n_1}\binom{k_2}{n_2}\cdots\binom{k_r}{n_r}$.

Παραδείγματα

(1) Ένα κουτί περιέχει r κόκκινες και b μαύρες μπάλες. Παίρνετε χωρίς επανάθεση τυχαίο δείγμα[2] μεγέθους n. Με πόσους τρόπους μπορείτε να πάρετε r_0 κόκκινες (και επομένως $n - r_0$ μαύρες) μπάλες στο δείγμα; Για να μην έχουμε τετριμμένη απάντηση πρέπει να υποθέσουμε ότι $\max(0, n-b) \le r_0 \le \min(r,n)$ (γιατί;). Τότε, σύμφωνα με όσα έχουμε πει, ο αριθμός των τρόπων με τους οποίους παίρνουμε r_0 κόκκινες μπάλες στο δείγμα θα είναι $\binom{r}{r_0}\binom{b}{n-r_0}$. Αν τώρα θέλουμε να υπολογίσουμε την αντίστοιχη πιθανότητα (να πάρουμε δηλαδή r_0 κόκκινες μπάλες στο δείγμα), θα πρέπει να διαιρέσουμε τον παραπάνω αριθμό «ευνοϊκών» περιπτώσεων με τον αριθμό των «δυνατών» περιπτώσεων, δηλαδή με το $\binom{r+b}{n}$.

(2) Ποια η πιθανότητα σε ένα χέρι πόκερ (5 φύλλα) να έχουμε δύο ζεύγη (δηλ. μια πεντάδα της μορφής (x,x,y,y,z), όπου τα x, y, z συμβολίζουν όχι ίσες τιμές—τιμές ονομάζουμε την «τάξη» του χαρτιού, π.χ. άσσος, ρήγας, δυάρι, κ.λπ., σε αντιδιαστολή με το «χρώμα»); Αφού η διάταξη με την οποία μοιράστηκαν τα χαρτιά δεν ενδιαφέρει βρισκόμαστε στα πλαίσια του μοντέλου μας. Η ζητούμενη πιθανότητα θα έχει ως παρονομαστή τον αριθμό $\binom{52}{5}$, αφού αυτός είναι ο αριθμός των δυνατών χεριών πόκερ που μπορεί να μοιραστούν. Τα φύλλα της τράπουλας μπορούν να πάρουν 13 τιμές (4 φύλλα την κάθε τιμή). Για κάθε δύο συγκεκριμένες τιμές, η τράπουλα διαμερίζεται σε τρεις κατηγορίες: τα φύλλα που παίρνουν την πρώτη τιμή, τα φύλλα που παίρνουν τη δεύτερη τιμή και τέλος, όλα τα υπόλοιπα. Επομένως για δύο συγκεκριμένες τιμές, ο αριθμός των χεριών πόκερ με ζεύγη (από αυτές τις τιμές) που μπορούμε να σχηματίσουμε είναι $\binom{4}{2}\binom{4}{2}\binom{44}{1}$. Αλλά, μπορούμε να επιλέξουμε τις 2 συγκεκριμένες τιμές κατά $\binom{13}{2}$ τρόπους και επομένως ο αριθμητής της ζητούμενης πιθανότητας (αριθμός ευνοϊκών τρόπων) θα είναι $\binom{13}{2}\binom{4}{2}\binom{4}{2}\binom{44}{1}$.

[2] Υπενθυμίζουμε ότι στην ορολογία οι λέξεις «τυχαίο δείγμα» σημαίνουν ότι κάθε μία από τις $(r+b)_n$ διατάξεις είναι εξίσου πιθανή.