1 Επανάληψη Βασικών Μαθηματικών Εννοιών

Στο κεφάλαιο αυτό θα κάνουμε μια σύντομη επανάληψη της άλγεβρας, της θεωρίας αριθμών και των πιθανοτήτων. Επιπλέον υπενθυμίσεις αν είναι απαραίτητες θα δοθούν σε επόμενα κεφάλαια. [1]

1.1 Άλγεβρα και Θεωρία Αριθμών

Ομάδες

Ορισμός 1.1.1. Μια ***Ομάδα (Group)*** $(G, *)$ είναι ένα σύνολο G μαζί με μια διμελή πράξη $*$ που ικανοποιεί τα παρακάτω

- Το G είναι κλειστό ως προς την $*$: για κάθε $g, h \in G, g * h \in G$;
- Η πράξη $*$ είναι προσεταιριστική: για κάθε $g, h, \ell \in G,\ g * (h * \ell) = (g * h) * \ell \in G$;
- G περιέχει ένα μοναδιαίο στοιχείο e τέτοιο ώστε $g * e = e * g = g$ για κάθε $g \in G$;
- G είναι κλειστό ως προς την αντιστροφή: για κάθε $g \in G$, υπάρχει $g^{-1} \in G$ τέτοιο ώστε $g * g^{-1} = g^{-1} * g = e$.

Τυπικά, μια ομάδα συμβολίζεται ως ένα διατεταγμένο ζεύγος $(G, *)$. Θα συμβολίζουμε την ομάδα απλά γράφοντας το σύνολο G όταν η πράξη είναι το $*$.

Ορισμός 1.1.2. Μια ομάδα G ονομάζεται ***Αβελιανή (Abelian)*** αν για κάθε $g, h \in G$, $g * h = h * g$.

Θεώρημα 1.1.1. *Αν το G είναι μια Αβελιανή ομάδα ως προς τη πράξη $*$, τότε το G περιέχει ακριβώς ένα μοναδιαίο στοιχείο και κάθε στοιχείο της G έχει μοναδικό αντίστροφο.*

Ορισμός 1.1.3. Σε μια πεπερασμένη ομάδα G, η ***τάξη (order)*** της G είναι το πλήθος των στοιχείων που ανήκουν σε αυτή, και συμβολίζεται ως $\#G$ ή $|G|$.

Ορισμός 1.1.4. Για κάθε ομάδα G και στοιχείο $g \in G$, ορίζουμε την ***τάξη του (order of)*** g ως τον μικρότερο θετικό ακέραιο i τέτοιον ώστε $g^i = e$, ή ισοδύναμα, $\underbrace{g * g * \cdots * g}_{i \text{ times}} = e$. Θα συμβολίζουμε την τάξη του g ως $\text{ord}(g)$.

Θεώρημα 1.1.2 (Lagrange). *Σε μια πεπερασμένη ομάδα, η τάξη κάθε στοιχείου διαιρεί το πλήθος των στοιχείων της ομάδας.*

Ορισμός 1.1.5. Αν υπάρχει κάποιο στοιχείο $g \in G$ τέτοιο ώστε $\text{ord}(g) = \#G$, τότε το g είναι γεννήτορας της G και θα χαρακτηρίζουμε το G ως ***κυκλική ομάδα (cyclic group)***. Επιπλέον, θα γράφουμε $G = \langle g \rangle$.

Παράδειγμα. Θεωρείστε το $\mathbb{Z}_5^* = \mathbb{Z}_5 - \{0\}$. Αυτή η ομάδα είναι κυκλική ως προς τον πολλαπλασιασμό modulo 5. Αναζητούμε ένα $g \in \mathbb{Z}_5^*$ τέτοιο ώστε $\text{ord}(g) = \#\mathbb{Z}_5^* = 4$ και $\langle g \rangle = \mathbb{Z}_5^*$. Προφανώς ισχύει ότι $\langle 1 \rangle \neq \mathbb{Z}_5^*$, οπότε ας δοκιμάσουμε το στοιχείο 2:

$$2^0 \equiv 1 \bmod 5,\ 2^1 \equiv 2 \bmod 5,\ 2^2 \equiv 4 \bmod 5,\ 2^3 \equiv 3 \bmod 5, \text{ and } 2^4 \equiv 1 \bmod 5.$$

Αφού $\langle 2 \rangle = \{1, 2, 3, 4\}$ και το 2 έχει τάξη 4, το 2 είναι γεννήτορας του $\mathbb{Z}_5^*$.

Είναι δυνατόν περισσότερα από ένα στοιχεία να είναι γεννήτορες της ομάδας. Ας δοκιμάσουμε λοιπόν το στοιχείο 3. Από το θεώρημα Lagrange, ισχύει ότι $\text{ord}(3) \mid 4$. Από τους προηγούμενος υπολογισμούς μας ισχύει ότι $2^3 \equiv 3 \bmod 5$, συνεπώς $\text{ord}(2^3) = \text{ord}(3)$. Τότε ισχύει ότι $3^{\text{ord}(3)} \equiv 2^{3\text{ord}(3)} \equiv 1 \bmod 5$ και $3\text{ord}(3) \equiv \text{ord}(2) \bmod 4$. Αφού τα 3 και 4 είναι πρώτοι μεταξύ τους, ισχύει ότι $\text{ord}(3) = 4$. Συμπεραίνουμε λοιπόν ότι το 3 είναι ένας άλλος γεννήτορας του $\mathbb{Z}_5^*$.

Χρησιμοποιώντας την ίδια επιχειρηματολογία, μπορούμε να δείξουμε πως το 4 δεν είναι γεννήτορας. Από τα παραπάνω ισχύει ότι $2^2 \equiv 4 \bmod 5$. Συνεπώς $\text{ord}(2^2) = \text{ord}(4)$. Γνωρίζουμε ότι το 2 (ο εκθέτης στο 2^2) διαιρεί το 4, συνεπώς ο $\gcd(2, 4)$ διαιρεί το 4. Επιπλέον, έχουμε ότι $\text{ord}(4) = 4/\gcd(2, 4) = 2$. Αυτό συνεπάγεται ότι $\#\langle 4 \rangle = 2$, συνεπώς το 4 δεν είναι γεννήτορας: $\langle 4 \rangle = \{1, 4\}$.

[1] Για επιπλέον επανάληψη μαθηματικών εννοιών, βλ. [1].

Δακτύλιοι και Σώματα

Ορισμός 1.1.6. Ένας ***αντιμεταθετικός δακτύλιος (commutative ring)*** R είναι ένα σύνολο μαζί με δύο διμελείς πράξεις $+$ και $*$ έτσι ώστε

- Η $(R, +)$ είναι μια Αβελιανή ομάδα;
- Η πράξη $*$ είναι προσεταιριστική: $(r * s) * t = r * (s * t)$ για κάθε $r, s, t \in R$;
- Η επιμεριστική ιδιότητα ισχύει στο R: $r * (s + t) = r * s + r * t$ και $(r + s) * t = r * t + s * t$ για κάθε $r, s, t \in R$;
- Η πράξη $*$ είναι αντιμεταθετική: $r * s = s * r$ για κάθε $r, s \in R$; και
- Το R περιέχει ένα μοναδιαίο στοιχείο αν υπάρχει στοιχείο $1 \in R$ τέτοιο ώστε $1 * r = r * 1 = r$ για κάθε $r \in R$.

Με απλά λόγια, ένας αντιμεταθετικός δακτύλιος είναι μια Αβελιανή ομάδα για την πράξη $+$ και αβελιανή ομάδα χωρίς αντίστροφους για την πράξη $*$. Δεν ισχύει ότι κάθε αντιμεταθετικός δακτύλιος περιέχει το 1. Συνεπώς η τελευταία ιδιότητα δεν είναι απόλυτη.

Παράδειγμα. Το $\mathbb{Z}$ είναι ένας δακτύλιος ως προς την συνήθη πρόσθεση και τον πολλαπλασιασμό.

Παράδειγμα. Το $\mathbb{Z}_n$ είναι ένας δακτύλιος ως προς την πρόσθεση και τον πολλαπλασιασμό modulo n.

Ορισμός 1.1.7. Ένα ***σώμα (field)*** F είναι ένα σύνολο μαζί με δύο διμελείς πράξεις $+$ και $*$ τέτοιες ώστε

- Η $(F, +)$ είναι μια Αβελιανή ομάδα με μοναδιαίο στοιχείο 0
- Η $(F - \{0\}, *)$ είναι μια Αβελιανή ομάδα με μοναδιαίο στοιχείο 1 και επιπλέον ισχύει η επιμεριστική ιδιότητα

Παράδειγμα. Τα $\mathbb{Q}, \mathbb{R}$, και $\mathbb{C}$ είναι όλα σώματα ως προς την πρόσθεση και τον πολλαπλασιασμό.

Παράδειγμα. Για κάθε πρώτο p, το $\mathbb{Z}_p$ είναι ένα σώμα ως προς την πρόσθεση και πολλαπλασιασμό υπόλοιπο p.

Ορισμός 1.1.8. Έστω p ένας πρώτος. Τότε το $\mathbb{Z}_p$ είναι ένα ***πεπερασμένο σώμα (finite field)***, και το συμβολίζουμε ως $\mathbb{F}_p$.

Κινέζικο Θεώρημα Υπολοίπων

Ορισμός 1.1.9. Συμβολίζουμε ***σχέσεις ισοδυναμίας (congruence relationships)*** πάνω στους ακεραίους με $a \equiv b \bmod n$ αν και μόνο αν $n \mid (a - b)$.

Θεώρημα 1.1.3 (Κινέζικο Θεώρημα Υπολοίπων). *Έστω $m_1, \ldots, m_k$ ανά μεταξύ τους πρώτοι θετικοί ακέραιοι και έστω $c_1, \ldots, c_k$ τυχαίοι ακέραιοι. Τότε υπάρχει ακέραιος x τέτοιος ώστε*

$$x \equiv c_i \bmod m_i$$

για κάθε $i = 1, \ldots, k$. Επιπλέον, κάθε ακέραιος x' είναι επιπλέον λύση σε αυτές τις ισοδυναμίες αν και μόνο αν $x \equiv x' \bmod M$ όπου $M = \prod m_i$ for $i = 1, \ldots, k$.

Απόδειξη. Έστω $M = \prod m_i$ για $i = 1, \ldots, k$. Ορίζουμε ως $m_i' = M/m_i$. Όλα τα m_is είναι πρώτοι μεταξύ τους, συνεπώς $\gcd(m_i, m_i') = 1$ για κάθε i. Έστω πως $u_i = (m_i')^{-1} \bmod m_i$ και $w_i = m_i' u_i$. Εκ κατασκευής, ισχύει ότι $w_i \equiv 1 \bmod m_i$ και $w_i \equiv 0 \bmod m_j$ όπου $i \neq j$. Από αυτό προκύπτει ότι $w_i \equiv \delta_{ij} \bmod m_j$ όπου

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j. \end{cases}$$

Θέτοντας $x = \sum w_i c_i$ για κάθε $i = 1, \ldots, k$, παρατηρούμε ότι

$$x \equiv \sum_{j=1}^{k} \delta_{ij} c_i \equiv c_j \bmod m_j$$

άρα ισχύει το ζητούμενο. ■

Παρατήρηση. Το Κινέζικο Θεώρημα Υπολοίπων συνεπάγεται τον ισομορφισμό ομάδων

$$\mathbb{Z}_n^* \cong \mathbb{Z}_{p_1^{e_1}}^* \times \ldots \times \mathbb{Z}_{p_m^{e_m}}^*,$$

ο οποίος δίνεται από $a \bmod n \mapsto (a \bmod p_1^{e_1}, \ldots, a \bmod p_m^{e_m})$, όπου $n = p_1^{e_1} \cdots p_m^{e_m}$ για ακεραίους e_i και διαφορετικούς πρώτους p_i.

Παράδειγμα. Ιστορικά, οι Κινέζοι χρησιμοποιούσαν αυτό το θεώρημα για να υπολογίσουν το πλήθος των στρατιωτών. Μετά από μια μάχη, οι στρατιώτες στοιχίζονταν σε σειρές (για παράδειγμα) των τριών, μετά των πέντε και μετά των εφτά. Υπολογίζοντας τους περισσευούμενος σε κάθε στοίχιση, οι στρατηγοί μπορούσαν να υπολογίσουν γρήγορα τον αριθμό των ανδρών και έτσι να υπολογίσουν και τις απώλειες.

Φανταστείτε πως υπάρχουν λιγότεροι από 100 στρατιώτες. Όταν στοιχίσουμε 3 στρατιώτες ανά σειρά, ένας περισσεύει. Όταν στοιχίσουμε 5 σε κάθε σειρά, 2 στρατιώτες περισσεύουν και όταν στοιχίσουμε 7 στη σειρά, 6 περισσεύουν. Θέλουμε να υπολογίσουμε τον ακριβή αριθμό στρατιωτών.

Έστω x ο συνολικός τους αριθμός. Τότε

$$x \equiv 1 \bmod 3$$
$$x \equiv 2 \bmod 5$$
$$x \equiv 6 \bmod 7.$$

Υπολογίζουμε το $M = 3 \cdot 5 \cdot 7 = 105$, και τα $m_1' = 35, m_2' = 21, m_3' = 15$. Στη συνέχεια υπολογίζουμε τους αντιστρόφους

$$u_1 = 35^{-1} \equiv 2 \bmod 3$$
$$u_2 = 21^{-1} \equiv 1 \bmod 5$$
$$u_3 = 15^{-1} \equiv 1 \bmod 7$$

Άρα έχουμε πως $w_1 = 70, w_2 = 21, w_3 = 15$ και $x = w_1 c_1 + w_2 c_2 + w_3 c_3 = 70(1) + 21(2) + 15(6) \equiv 97 \bmod 105$. Συνεπώς υπάρχουν 97 στρατιώτες.

1.2 Διακριτή πιθανότητα

Ορισμός 1.2.1. Μια ***διακριτή πιθανοτική κατανομή (discrete probability distribution)*** **D** σε ένα σύνολο $[\mathbf{D}]$ ορίζεται ως

- $\text{Prob}_{\mathbf{D}}[u] \in [0, 1]$ για κάθε $u \in [\mathbf{D}]$
- $\sum_{u \in \mathbf{D}} \text{Prob}_{\mathbf{D}}[u] = 1$.

Το σύνολο $[\mathbf{D}]$ ονομάζεται ως ***στήριγμα (support)*** του **D**.

Παράδειγμα (Επιτυγχάνοντας μέσω Επανάληψης). Θεωρείστε ένα πείραμα όπου η πιθανότητα επιτυχίας είναι p. Υποθέστε πως το πείραμα επαναλαμβάνεται n φορές. Θέλουμε να φράξουμε την πιθανότητα του ενδεχομένου "Όλες οι n δοκιμές αποτυγχάνουν". Αφού κάθε δοκιμή είναι ανεξάρτητη από την επόμενη, η πιθανότητα n δοκιμές να αποτυγχάνουν είναι $(1 - p)^n$. Θυμίζουμε πως $1 - x \leq e^{-x}$ για κάθε x. Θέτοντας

$x = p$ και υψώνοντας και τις δύο πλευρές στην n-οστή δύναμη, παίρνουμε άνω φράγμα $(1-p)^n \leq e^{-pn}$. Then

$$\begin{aligned}\text{Prob[At least 1 success]} &= 1 - \text{Prob[All fail]} \\ &\geq 1 - e^{-pn}.\end{aligned}$$

Αν το p είναι σταθερό, η πιθανότητα πως κάθε δοκιμή αποτυγχάνει μειώνεται εκθετικά στον αριθμό των επαναλήψεων n.

Παράδειγμα (Σφαιρίδια και Κουτιά). Θεωρείστε ένα πείραμα με n κουτιά και k σφαιρίδια, κάθε ένα με διαφορετικό χρώμα. Ρίχνουμε κάθε σφαιρίδιο τυχαία σε κάποιο κουτί. Ορίζουμε ως σύγκρουση να είναι το γεγονός πως 2 σφαιρίδια διαφορετικού χρώματος πέφτουν στο ίδιο κουτί. Θέλουμε να υπολογίσουμε την πιθανότητα σύγκρουσης. Σε μια τέτοια περίπτωση είναι ευκολότερο να υπολογίσουμε την πιθανότητα του συμπληρωματικού γεγονότος:

$$\text{Prob}_{\mathbf{D}}[\text{No collision}] = \frac{n(n-1)\cdots(n-k+1)}{n^k} = \prod_{j=0}^{k-1}\left(\frac{n-j}{n}\right).$$

Χρησιμοποιώντας πάλι το γεγονός πως $1 - x \leq e^{-x}$ για κάθε x, έχουμε ότι

$$\prod_{j=0}^{k-1}\left(\frac{n-j}{n}\right) = \prod_{j=1}^{k-1}\left(1-\frac{j}{n}\right) \leq \prod_{j=1}^{k-1} e^{-j/n} = e^{-k(k-1)/2n}.$$

Αφού ισχύει πως Prob[Collision] $= 1 -$ Prob[No collision],

$$\text{Prob}_{\mathbf{D}}[\text{Collision}] \geq 1 - e^{-k(k-1)/2n}.$$

Παράδειγμα (Το παράδοξο των γενεθλίων). Το παράδοξο των γενεθλίων είναι ένα κλασικό πρόβλημα που χρησιμοποιούμε την προηγούμενη μέθοδο. Θέλουμε να μάθουμε πόσοι άνθρωποι πρέπει να είναι παρόντες σε ένα δωμάτιο έτσι ώστε να υπάρχει τουλάχιστον 50% πιθανότητα δύο άνθρωποι έχουν γενέθλια την ίδια μέρα (η σύγκρουση σε αυτή τη περίπτωση). Έστω $n = 365$ και υποθέστε πως τα γενέθλια των ανθρώπων είναι ομοιόμορφα κατανεμημένα σε όλες τις ημέρες του χρόνου. Αν θέλουμε $\text{Prob}_{\mathbf{D}}[\text{Collision}] \geq 1/2$, τότε

$$\begin{aligned}1 - e^{-k(k-1)/2n} &\geq \frac{1}{2} \\ e^{-k(k-1)/2n} &\leq \frac{1}{2} \\ e^{k(k-1)/2n} &\geq 2 \\ \frac{k(k-1)}{2n} &\geq \ln 2 \\ \frac{k^2}{2n} &\geq \ln 2 \\ k &\geq \sqrt{2n\ln 2}\end{aligned}$$

Συνεπώς αν υπάρχουν περισσότερα από 23 άτομα σε ένα δωμάτιο, τότε υπάρχει μεγαλύτερη από 50% πιθανότητα δύο άτομα να έχουν γενέθλια την ίδια μέρα. Το αποτέλεσμα αυτό είναι ενάντια στη διαίσθησή μας και για το λόγο αυτό ονομάστηκε παράδοξο.

Παράδειγμα (Διωνυμική κατανομή (Binomial Distribution)). Μία ***διωνυμική δοκιμή (binomial trial)*** είναι ένα πείραμα με μόνο δύο δυνατά αποτελέσματα: επιτυχία και αποτυχία. Έστω $[\mathbf{D}] = \{0, 1, \ldots, n\}$ και έστω ότι η πιθανότητα της επιτυχίας είναι p. Τότε η διωνυμική κατανομή είναι η πιθανότητα u επιτυχιών σε μια ακολουθία από n ανεξάρτητες δοκιμές:

$$\text{Prob}_{\mathbf{D}}[u] = \binom{n}{u} p^u (1-p)^{n-u}.$$

Ορισμός 1.2.2. Ένα υποσύνολο $\mathbf{A} \subseteq [\mathbf{D}]$ θα συμβολίζει ένα ***γεγονός (event)***. Η πιθανότητα του γεγονότος $\mathbf{A}$ είναι

$$\text{Prob}_{\mathbf{D}}[\mathbf{A}] = \sum_{u \in \mathbf{A}} \text{Prob}_{\mathbf{D}}[u].$$

Είναι επίσης δυνατό να αποδείξουμε διάφορες προτάσεις σχετικά με τις θεωρητικές πράξεις πάνω σε γεγονότα, όπως η ένωση και η τομή. Για παράδειγμα, αν $\mathbf{A}, \mathbf{B} \subseteq [\mathbf{D}]$, τότε έχουμε

$$\text{Prob}_{\mathbf{D}}[\mathbf{A} \cup \mathbf{B}] = \text{Prob}_{\mathbf{D}}[\mathbf{A}] + \text{Prob}_{\mathbf{D}}[\mathbf{B}] - \text{Prob}_{\mathbf{D}}[\mathbf{A} \cap \mathbf{B}].$$

Αυτό ονομάζεται ***η αρχή εγκλεισμού/αποκλεισμού (inclusion-exclusion principal)***.

1.3 Δεσμευμένη Πιθανότητα

Ορισμός 1.3.1. Έστω τα γεγονότα $\mathbf{A}$ και $\mathbf{B}$. Η πιθανότητα να συμβεί το γεγονός $\mathbf{A}$, δεδομένου ότι το γεγονός $\mathbf{B}$ έχει ήδη συμβεί ονομάζεται ***δεσμευμένη πιθανότητα (conditional probability)***. Η πιθανότητα αυτή δίνεται από το

$$\text{Prob}_{\mathbf{D}}[\mathbf{A} \mid \mathbf{B}] = \frac{\text{Prob}_{\mathbf{D}}[\mathbf{A} \cap \mathbf{B}]}{\text{Prob}_{\mathbf{D}}[\mathbf{B}]}.$$

Το επόμενο θεώρημα είναι χρήσιμο για τον υπολογισμό δεσμευμένων πιθανοτήτων.

Θεώρημα 1.3.1 (Bayes). *Για δύο γεγονότα* $\mathbf{A}$ *και* $\mathbf{B}$*, ισχύει ότι*

$$\text{Prob}_{\mathbf{D}}[\mathbf{B} \mid \mathbf{A}] = \frac{\text{Prob}_{\mathbf{D}}[\mathbf{A} \mid \mathbf{B}] \cdot \text{Prob}_{\mathbf{D}}[\mathbf{B}]}{\text{Prob}_{\mathbf{D}}[\mathbf{A}]}.$$

Επιπλέον, αν τα $\mathbf{D}_1, \ldots, \mathbf{D}_n$ *είναι ξένα γεγονότα που διαμερίζουν το* $[\mathbf{D}]$*, δηλαδή* $[\mathbf{D}] = \bigcup \mathbf{D}_i$ *για* $1 \leq i \leq n$*, τότε για κάθε γεγονότα* $\mathbf{A}$ *και* $\mathbf{B}$*, ισχύει πως*

$$\text{Prob}_{\mathbf{D}}[\mathbf{B} \mid \mathbf{A}] = \frac{\text{Prob}_{\mathbf{D}}[\mathbf{A} \mid \mathbf{B}] \cdot \text{Prob}_{\mathbf{D}}[\mathbf{B}]}{\sum_{i=1}^{n} \text{Prob}_{\mathbf{D}}[\mathbf{A} \mid \mathbf{D}_i] \cdot \text{Prob}_{\mathbf{D}}[\mathbf{D}_i]}.$$

θα χρησιμοποιήσουμε το $\overline{\mathbf{B}}$ για να συμβολίσουμε το συμπλήρωμα : $\overline{\mathbf{B}} = [\mathbf{D}] \setminus \mathbf{B}$. Χρησιμοποιώντας το θεώρημα του Bayes έχουμε

$$\text{Prob}_{\mathbf{D}}[\mathbf{B} \mid \mathbf{A}] = \frac{\text{Prob}_{\mathbf{D}}[\mathbf{A} \mid \mathbf{B}] \cdot \text{Prob}_{\mathbf{D}}[\mathbf{B}]}{\text{Prob}_{\mathbf{D}}[\mathbf{A} \mid \mathbf{B}] \cdot \text{Prob}_{\mathbf{D}}[\mathbf{B}] + \text{Prob}_{\mathbf{D}}[\mathbf{A} \mid \overline{\mathbf{B}}] \cdot \text{Prob}_{\mathbf{D}}[\overline{\mathbf{B}}]}.$$

Παράδειγμα. Τώρα θα δούμε μια εφαρμογή του θεωρήματος του Bayes. Έστω $\mathbf{D}$ η πιθανοτική κατανομή σε ένα δοσμένο πληθυσμό και έστω το γεγονός $\mathbf{S}$ να αντιστοιχεί στο πληθυσμό που έχει κάποια συγκεκριμένη ασθένεια. Υποθέστε πως υπάρχει ένα ιατρικό τεστ για την διάγνωση της ασθένειας και έστω $\mathbf{T}$ το γεγονός πως για κάποιο άτομο του πληθυσμού το τεστ είναι θετικό.

Η παρουσία της ασθένειας στον πληθυσμό είναι $\text{Prob}_{\mathbf{D}}[\mathbf{S}] = 1\%$, η πιθανότητα ενός επιτυχούς τεστ είναι $\text{Prob}_{\mathbf{D}}[\mathbf{T} \mid \mathbf{S}] = 99\%$, και η πιθανότητα ενός λανθασμένου τεστ είναι $\text{Prob}_{\mathbf{D}}[\mathbf{T} \mid \overline{\mathbf{S}}] = 5\%$. Θέλουμε να βρούμε την πιθανότητα πως κάποιος είναι ασθενής δεδομένου ενός θετικού τεστ. Ένα σύνηθες λάθος είναι να θεωρήσουμε πως η πιθανότητα είναι 99%- δηλαδή η πιθανότητα επιτυχίας του τεστ. Αυτό είναι λάθος αφού δεν λαμβάνουμε υπ' όψιν πως ήδη γνωρίζουμε πως το τεστ είναι θετικό. Χρησιμοποιώντας το θεώρημα του Bayes, μπορούμε να υπολογίσουμε αυτή την πιθανότητα και έτσι υπολογίζουμε

$$\begin{aligned}
\text{Prob}_{\mathbf{D}}[\mathbf{S} \mid \mathbf{T}] &= \frac{\text{Prob}_{\mathbf{D}}[\mathbf{T} \mid \mathbf{S}] \cdot \text{Prob}_{\mathbf{D}}[\mathbf{S}]}{\text{Prob}_{\mathbf{D}}[\mathbf{T} \mid \mathbf{S}] \cdot \text{Prob}_{\mathbf{D}}[\mathbf{S}] + \text{Prob}_{\mathbf{D}}[\mathbf{T} \mid \overline{\mathbf{S}}] \cdot \text{Prob}_{\mathbf{D}}[\overline{\mathbf{S}}]} \\
&= \frac{(0.99)(0.01)}{(0.99)(0.01) + (0.05)(1 - 0.01)} \\
&= \frac{1}{6}.
\end{aligned}$$

Αυτό μπορεί να φαίνεται παράλογο, αλλά επειδή η αρρώστια είναι τόσο σπάνια, ένα θετικό τεστ είναι πιο πιθανό να προκύψει επειδή το τεστ δεν είναι ακριβές παρά από την πραγματική ύπαρξη της ασθένειας.

1.4 Τυχαίες μεταβλητές

Ορισμός 1.4.1. Για μια πιθανοτική κατανομή **D**, ορίζουμε μια ***τυχαία μεταβλητή (random variable)*** X να είναι μια συνάρτηση
$X : [\mathbf{D}] \longrightarrow R$. Για κάθε $x \in R$, θα χρησιμοποιούμε τον συμβολισμό

$$\text{Prob}[X = x] = \sum_{X(u)=x} \text{Prob}_{\mathbf{D}}[u].$$

Θα λέμε ότι μια τυχαία μεταβλητή X κατανέμεται ανάλογα με το **D** αν το $X : [\mathbf{D}] \longrightarrow [\mathbf{D}]$ είναι η ταυτοτική συνάρτηση. Θα το συμβολίζουμε με

$$\underset{X \leftarrow \mathbf{D}}{\text{Prob}}[X = x] = \sum_{X(u)=x} \text{Prob}_{\mathbf{D}}[u].$$

Ορισμός 1.4.2. Για κάθε πιθανοτική κατανομή **D** με τυχαία μεταβλητή X, η ***μέση τιμή της (expectation)*** είναι

$$\text{E}[X] = \sum_{x \in R} x\text{Prob}[X = x].$$

Ορισμός 1.4.3. Η ***τυπική απόκλιση (variance)*** μιας διακριτής τυχαίας μεταβλητής X μετράει το εύρος της, ή το πόσο αποκλίνει η κατανομή από τη μέση τιμή. Ορίζεται ως

$$\text{Var}[X] = \text{E}[X^2] - \text{E}[X]^2.$$

1.5 Ουρές Πιθανοτικών κατανομών

Όταν αναλύουμε τυχαίες διαδικασίες, συχνά θέλουμε να εκτιμήσουμε τα φράγματα στις ***ουρές (tails)*** μιας πιθανοτικής κατανομής. Ο όρος "ουρά" αναφέρεται στα άκρα της γραφικής αναπαράστασης μιας πιθανοτικής κατανομής, όπου η κατανομή αποκλίνει από την μέση τιμή. Τα επόμενα θεωρήματα θα φανούν χρήσιμα

Θεώρημα 1.5.1 (Ανισότητα του Markov (Markov's Inequality)). *Έστω X μια τυχαία μεταβλητή που παίρνει πραγματικές μη αρνητικές τιμές. Τότε για κάθε $t > 0$,*

$$\text{Prob}[X \geq t] \leq \frac{\text{E}[X]}{t}.$$

Θεώρημα 1.5.2 (Ανισότητα του Chebyshev (Chebyshev's Inequality)). *Έστω X μια τυχαία μεταβλητή. Για κάθε $t > 0$ έχουμε ότι*

$$\text{Prob}[|X - \text{E}(X)| \geq t] \leq \frac{\text{Var}[X]}{t^2}.$$

Θεώρημα 1.5.3 (Το φράγμα του Chernoff (Chernoff's Bound)). *Έστω $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές που παίρνουν τιμές στο $\{0, 1\}$ με πιθανότητα $\text{Prob}[X_i = 1] = p_i$. Τότε ισχύει ότι*

$$\text{Prob}\left[\sum_{i=1}^{n} X_i \leq (1-\delta)\mu\right] \leq e^{-\mu\delta^2/2} \qquad \textit{and} \qquad \text{Prob}\left[\sum_{i=1}^{n} X_i \geq (1+\delta)\mu\right] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$$

όπου $\mu = \sum p_i$ και $\delta \in (0, 1]$.

Εδώ το μ είναι η μέση τιμή και τα $(1-\delta)\mu$ και $(1+\delta)\mu$ είναι οι ουρές

Παράδειγμα (Μαντεύοντας από την Πλειοψηφία). Υποθέτουμε πως υπάρχει ένα μαντείο που απαντά ερωτήσεις με "Ναι" ή "Όχι", και απαντά σωστά με πιθανότητα $1/2+\alpha$. Έστω ότι ρωτάμε το μαντείο n ερωτήσεις και έστω X_i μια τυχαία μεταβλητή όπου

$$X_i = \begin{cases} 1, & \text{oracle answers the } i\text{th query correctly} \\ 0, & \text{otherwise.} \end{cases}$$

Αν ορίσουμε ως αποτυχία το ενδεχόμενο να λάβουμε λιγότερες σωστές απαντήσεις από λανθασμένες, τότε η πιθανότητα αποτυχίας είναι

$$\text{Prob[Failure]} = \text{Prob}\left[\#\text{ of correct answers } \leq \frac{n}{2}\right] = \text{Prob}\left[\sum_{i=1}^{n} X_i \leq \frac{n}{2}\right].$$

Εφαρμόζουμε το φράγμα του Chernoff θέτοντας $n/2 = (1-\delta)\mu$ και έχουμε ότι

$$\text{Prob}\left[\sum_{i=1}^{n} X_i \leq (1-\delta)\mu\right] \leq e^{-\mu\delta^2/2}. \tag{1}$$

Σημειώνουμε ότι $\mu = (1/2+\alpha)n$. Έτσι μπορούμε να λύσουμε ως προς δ.

$$\begin{aligned}
\frac{n}{2} &= (1-\delta)\mu \\
\frac{n}{2} &= (1-\delta)\left(\frac{1}{2}+\alpha\right)n \\
\delta &= \frac{\alpha}{1/2+\alpha}
\end{aligned}$$

Για να υπολογίσουμε την πιθανότητα μίας αποτυχίας, αντικαθιστούμε την τιμή δ στο (1).

$$\text{Prob[Failure]} \leq e^{-\alpha^2 n/(1+2\alpha)}.$$

Αυτό συνεπάγεται πως αν το μαντείο μεροληπτεί α, μπορούμε να την εξαλείψουμε τη μεροληψία αυτή μετά από n επαναλήψεις. Εξαιτίας αυτού η πιθανότητα αποτυχίας μειώνεται εκθετικά ανάλογα με το βαθμό μεροληψίας και το πλήθος των δοκιμών. Αν θέλουμε η πιθανότητα αποτυχίας να πέσει κάτω από κάποιο ε, μπορούμε να βρούμε το αντίστοιχο κάτω φράγμα για το n.

$$\begin{aligned}
e^{-\alpha^2 n/(1+2\alpha)} &< \varepsilon \\
\frac{-\alpha^2 n}{(1+2\alpha)} &< \ln(\varepsilon) \\
n &> \alpha^{-2}(1+2\alpha)\ln\left(\frac{1}{\varepsilon}\right)
\end{aligned}$$

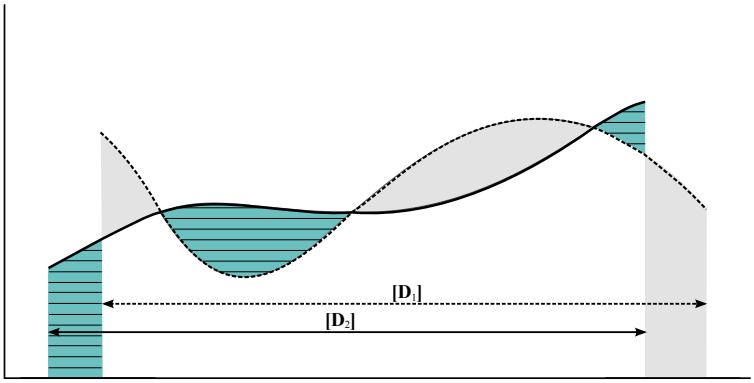
Θεωρώντας το n αρκετά μεγάλο, εγγυόμαστε ότι η πιθανότητα αποτυχίας είναι ικανοποιητικά χαμηλή.

1.6 Στατιστική απόσταση

Ορισμός 1.6.1. Έστω X και Y τυχαίες μεταβλητές που κατανέμονται σύμφωνα με τα $\mathbf{D}_1$ και $\mathbf{D}_2$ αντίστοιχα και έστω $\mathbf{V} = X([\mathbf{D}_1]) \cup Y([\mathbf{D}_2])$. Ορίζουμε την ***στατιστική απόσταση (statistical distance)*** Δ ως

$$\Delta[X,Y] = \frac{1}{2}\sum_{u\in\mathbf{V}}\left|\Pr_{X\leftarrow\mathbf{D}_1}[X=u] - \Pr_{Y\leftarrow\mathbf{D}_2}[Y=u]\right|.$$

Η Εικόνα 1 δίνει μια γραφική αναπαράσταση της στατιστικής απόστασης δύο τυχαίων μεταβλητών και Y. Η καμπύλη ζωγραφισμένη ως συνεχόμενα σημεία αναπαριστά την κατανομή της X στο $\mathbf{D}_1$ και η μαύρη καμπύλη αντιστοιχεί το Y στο $\mathbf{D}_2$. Εξ ορισμού, το άθροισμα των πιθανοτήτων στο πεδίο ορισμού είναι ένα, συνεπώς η περιοχή κάτω από κάθε καμπύλη έχει εμβαδόν 1. Το μισό από το άθροισμα των σκιαγραφημένων περιοχών αναπαριστά την στατιστική απόσταση μεταξύ των X και Y. Επειδή η περιοχή με τις γραμμές είναι ίση με την γκρίζα περιοχή, διαιρώντας τη συνολική σκιαγραφημένη περιοχή με το 2 παίρνουμε μια από τις δύο σημαδεμένες περιοχές σαν την στατιστική απόσταση.
Άσκηση: Δείξε ότι για κάθε δύο πεδία ορισμού $[\mathbf{D}_1]$ και $[\mathbf{D}_2]$, η περιοχή με τις γραμμές είναι ίση με τη γκρίζα περιοχή και έτσι η στατιστική απόσταση είναι ίση με μια από τις δυο περιοχές.



Σχήμα 1: Δύο πιθανοτικές κατανομές με διαφορετικά πεδία ορισμού $[\mathbf{D}_1]$ και $[\mathbf{D}_2]$. Οι σκιαγραφημένες περιοχές διαχωρίζουν την στατιστική απόσταση μεταξύ των τυχαίων μεταβλητών.

Ορισμός 1.6.2. Έστω $\varepsilon > 0$, τότε δύο τυχαίες μεταβλητές X και Y λέμε πως είναι ***ε-κοντά (ε-close))*** αν $\Delta[X, Y] \leq \varepsilon$.

Παράδειγμα. Έστω $\mathbf{D}_1$ η ομοιόμορφη κατανομή στο $[0, A)$ όπου $2^n \leq A < 2^{n+1}$ και έστω $\mathbf{D}_2$ η ομοιόμορφη κατανομή στο $[0, 2^n)$. Θέλουμε να υπολογίσουμε τη στατιστική απόσταση των $\mathbf{D}_1$ και $\mathbf{D}_2$.

Αφού το $\mathbf{D}_1$ είναι ομοιόμορφα κατανεμημένο στο $[0, A)$, έχουμε ότι $\text{Prob}_{\mathbf{D}_1}[u] = 1/A$ για κάθε $u \in [0, A)$. Αντίστοιχα, επεκτείνουμε το $\mathbf{D}_2$ στο δειγματικό χώρο $[0, A)$ ορίζοντας

$$\text{Prob}_{\mathbf{D}_2}[u] = \begin{cases} 1/2^n, & u \in [0, 2^n) \\ 0, & u \in [2^n, A). \end{cases}$$

Υποθέστε πως τα X και Y είναι τυχαίες μεταβλητές που κατανέμονται σύμφωνα με τα $\mathbf{D}_1$ και $\mathbf{D}_2$ αντίστοιχα όπου $[\mathbf{D}_1] = [\mathbf{D}_2] = [0, A)$. Τότε

$$\begin{aligned} \Delta[X, Y] &= \frac{1}{2} \sum_{u \in [0,A)} \left| \Pr_{X \leftarrow \mathbf{D}_1}[X = u] - \Pr_{X \leftarrow \mathbf{D}_2}[X = u] \right| \\ &= \frac{1}{2} \left(\sum_{u \in [0,2^n)} \left| \frac{1}{A} - \frac{1}{2^n} \right| + \sum_{u \in [2^n, A)} \left| \frac{1}{A} - 0 \right| \right) \\ &= \frac{1}{2} \left(\sum_{u \in [0,2^n)} \left(\frac{1}{2^n} - \frac{1}{A} \right) + \sum_{u \in [2^n, A)} \frac{1}{A} \right) \\ &= \frac{1}{2} \left(\left(\frac{1}{2^n} - \frac{1}{A} \right) 2^n + \frac{1}{A}(A - 2^n) \right) \\ &= \frac{A - 2^n}{A}. \end{aligned}$$

Θέτοντας $d = A - 2^n$, έχουμε ότι $\Delta[X, Y] = d/(d + 2^n)$. Όταν το A είναι σχετικά κοντά στο 2^n, τότε $\Delta[X, Y]$ προσεγγίζει το 0. Για παράδειγμα, αν $d = 2^{n/2}$ έτσι ώστε $A = 2^{n/2} + 2^n$, η στατιστική απόσταση μειώνεται εκθετικά

$$\frac{d}{d + 2^n} = \frac{2^{n/2}}{2^{n/2} + 2^n} = \frac{1}{1 + 2^{n/2}} \approx 2^{-n/2}.$$

Ορισμός 1.6.3. Μία συνάρτηση f είναι ***αμελητέα (negligible)*** αν για κάθε $c \in \mathbb{R}$ υπάρχει $n_0 \in \mathbb{N}$ τέτοιο ώστε $f(n) \leq 1/n^c$ για κάθε $n \geq n_0$.

Ορισμός 1.6.4. Ένα ***(πιθανοτικό) σύνολο (probability ensemble)*** είναι μια συλλογή από κατανομές $\mathcal{D} = \{\mathbf{D}_n\}_{n\in\mathbb{N}}$.

Παίρνουμε τώρα μια συλλογή X στο σύνολο $\mathcal{D}$ και θα εννοούμε μια συλλογή από τυχαίες μεταβλητές στο $\mathbf{D}_n \in \mathcal{D}$. Ως κατάχρηση του συμβολισμού Θα αναφερόμαστε στη συλλογή X, ως τυχαία μεταβλητή.

Ορισμός 1.6.5. Έστω X και Y τυχαίες μεταβλητές στα σύνολα $\mathcal{D}$ και $\mathcal{D}'$. Θα λέμε ότι τα $\mathcal{D}$ και $\mathcal{D}'$ είναι ***στατιστικά αδιαχώριστα (statistically indistinguishable)***, αν $\Delta[X,Y]$ είναι μια αμελητέα συνάρτηση στο n.

Θα πρέπει να τονίσουμε πως αν ισχύει $\Delta[X,Y] \to 0$ για δύο σύνολα δεν συνεπάγεται ότι είναι αδιαχώριστα. Αν δύο τυχαίες μεταβλητές είναι στατιστικά αδιαχώριστες συνεπάγεται πως η στατιστική απόσταση, αν την δούμε ως συνάρτηση του n, θα πρέπει να είναι μικρότερη κάθε πολυωνυμικής συνάρτησης στο n για αρκετά μεγάλες τιμές του n.

1.7 Ένας Εναλλακτικός Ορισμός της Στατιστικής Απόστασης

Ορισμός 1.7.1. Ένα ***στατιστικό τεστ (statistical test)*** $\mathcal{A}$ για ένα σύνολο $\mathcal{D} = \{\mathbf{D}_n\}_{n\in\mathbb{N}}$ είναι ένας αλγόριθμος που παίρνει ως είσοδο στοιχεία από το $\mathbf{D}_n$ και επιστρέφει τιμές στο $\{0,1\}$ για κάθε $n \in \mathbb{N}$.

Θεώρημα 1.7.1. *Θεωρήστε το στατιστικό τεστ $\mathcal{A}$ ως συνάρτηση του n και έστω X και Y τυχαίες μεταβλητές στα σύνολα $\mathcal{D}_1$ και $\mathcal{D}_2$ αντίστοιχα. Ορίζουμε ως*

$$\Delta_{\mathcal{A}}[X,Y] = \left| \Pr_{X\leftarrow\mathcal{D}_1}[\mathcal{A}(X)=1] - \Pr_{Y\leftarrow\mathcal{D}_2}[\mathcal{A}(Y)=1] \right|$$

να είναι η ***στατιστική απόσταση (statistical distance)*** *αναφορικά με το τεστ $\mathcal{A}$. Τότε για κάθε $\mathcal{A}$, ισχύει ότι $\Delta[X,Y] \geq \Delta_{\mathcal{A}}[X,Y]$ και υπάρχει κάποιο $\mathcal{A}^*$ τέτοιο ώστε $\Delta[X,Y] = \Delta_{\mathcal{A}^*}[X,Y]$.*

Το πρώτο κομμάτι του θεωρήματος αποδεικνύεται ως εξής. Για κάθε $\mathcal{A}$ ισχύει

$$\Delta_{\mathcal{A}}[X,Y] = \left| \sum_{a\in A_n} \text{Prob}_{\mathbf{D}_1}[a] - \text{Prob}_{\mathbf{D}_2}[a] \right| \leq \sum_{a\in A_n} |\text{Prob}_{\mathbf{D}_1}[a] - \text{Prob}_{\mathbf{D}_2}[a]| =_{\text{df}} N_1$$

όπου $A_n = \{a \in \mathbf{D}_n : \mathcal{A}(a) = 1\}$.

Τώρα θεωρούμε το στατιστικό τεστ $\overline{\mathcal{A}}$ που ενεργεί ακριβώς όπως το $\mathcal{A}$ με τη διαφορά όμως ότι αντιστρέφει την απάντηση. Άμεσα βλέπουμε ότι $\Delta_{\overline{\mathcal{A}}}[X,Y] = \Delta_{\mathcal{A}}[X,Y]$ βασιζόμενοι στον ορισμό του $\Delta_{\mathcal{A}}[\cdot,\cdot]$. Χρησιμοποιώντας ίδια επιχειρηματολογία όπως παραπάνω έχουμε ότι

$$\Delta_{\mathcal{A}}[X,Y] = \Delta_{\overline{\mathcal{A}}}[X,Y] \leq \sum_{a\in\overline{A_n}} |\text{Prob}_{\mathbf{D}_1}[a] - \text{Prob}_{\mathbf{D}_2}[a]| =_{\text{df}} N_2$$

όπου το $\overline{A_n}$ είναι το συμπλήρωμα του A_n στο $\mathbf{D}_n$.

Τώρα παρατηρούμε ότι $N_1 + N_2 = \Delta[X,Y]$ και δεδομένου ότι $N_1, N_2 \in [0,1]$ ισχύει ότι ένα από αυτά είναι το πολύ $\frac{1}{2}\Delta[X,Y]$. Το αποτέλεσμα έπεται.

Για το δεύτερο τμήμα του θεωρήματος, ορίζουμε έναν διαχωριστή $\mathcal{A}^*$ ως εξής:

$$\mathcal{A}^*(a) = \begin{cases} 1, & \text{Prob}_{\mathbf{D}_1}[a] \geq \text{Prob}_{\mathbf{D}_2}[a] \\ 0, & \text{otherwise,} \end{cases}$$

εύκολα προκύπτει ότι $\Delta[X,Y] = \Delta_{\mathcal{A}^*}[X,Y]$. Πράγματι για $A_n^* = \{a \in \mathbf{D}_n : \mathcal{A}^*(a) = 1\} \subseteq \mathbf{D}_n$,

$$\Delta_{\mathcal{A}^*}[X,Y] = \left| \sum_{a\in A_n^*} \text{Prob}_{\mathbf{D}_1}[a] - \sum_{a\in A_n^*} \text{Prob}_{\mathbf{D}_2}[a] \right| = \sum_{a\in A_n^*} (\text{Prob}_{\mathbf{D}_1}[a] - \text{Prob}_{\mathbf{D}_2}[a])$$

από το οποίο το καταλήγουμε στο ζητούμενο.

Για μια γραφική ερμηνεία του $\Delta[X,Y] = \Delta_{\mathcal{A}^*}[X,Y]$ για αυτόν τον διαχωριστή, επιστρέφουμε στην εικόνα 1. Η περιοχή με τις γραμμές συμβολίζει το χώρο όπου $\text{Prob}_{\mathbf{D}_1}[u] \geq \text{Prob}_{\mathbf{D}_2}[u]$, ο οποίος έχουμε ήδη δει πως είναι ακριβώς η στατιστική απόσταση.

Παράδειγμα. Θεωρήστε τις δύο πιθανοτικές κατανομές $\mathbf{D}_1$ και $\mathbf{D}_2$ όπου

$b_1 b_0$	$\mathbf{D}_1$	$\mathbf{D}_2$
0 0	$0.25 - \varepsilon$	0.25
0 1	0.25	0.25
1 0	$0.25 + \varepsilon$	0.25
1 1	0.25	0.25

Έστω X και Y τυχαίες μεταβλητές που ακολουθούν τις κατανομές $\mathbf{D}_1$ και $\mathbf{D}_2$ αντίστοιχα. Η στατιστική τους απόσταση είναι

$$\Delta[X,Y] = \frac{1}{2}\left(|(0.25-\varepsilon)-0.25| + |0.25-0.25| + |(0.25+\varepsilon)-0.25| + |0.25-0.25|\right) = \varepsilon.$$

Θεωρούμε το σύνολο των στατιστικών τεστ $\mathcal{A}_1, \ldots, \mathcal{A}_5$ που διαχωρίζουν την προηγούμενες πιθανοτικές κατανομές. Υποθέστε πως δίνονται δύο bits b_0 και b_1. Το τεστ $\mathcal{A}_1$ επιστρέφει b_1. Εξαιτίας αυτού, είναι φανερό ότι $\Delta_{\mathcal{A}_1}[X,Y] = |(0.25+\varepsilon)-0.25| + |0.25-0.25| = \varepsilon$. Το τεστ $\mathcal{A}_2$ επιστρέφει b_0, έτσι λοιπόν $\Delta_{\mathcal{A}_2}[X,Y] = |0.25-0.25| + |0.25-0.25| = 0$. Αν το τεστ $\mathcal{A}_3$ επιστρέφει $b_0 + b_1 \bmod 2$, που επιπλέον συμβολίζεται με την πράξη "αποκλειστικό ή", $b_0 \oplus b_1$, τότε η στατιστική του απόσταση είναι $\Delta_{\mathcal{A}_3}[X,Y] = |0.25-0.25| + |(0.25+\varepsilon)-0.25| = \varepsilon$. Το τεστ $\mathcal{A}_4$ επιστρέφει $b_0 \vee b_1$ και έτσι ισχύει ότι $\Delta_{\mathcal{A}_4}[X,Y] = |0.25-0.25| + |(0.25+\varepsilon)-0.25| + |0.25-0.25| = \varepsilon$. Τέλος αν το τεστ $\mathcal{A}_5$ επιστρέφει $b_0 \wedge b_1$, τότε η στατιστική του απόσταση είναι $\Delta_{\mathcal{A}_5}[X,Y] = |0.25-0.25| = 0$.

Βασισμένοι στα παραπάνω, μπορούμε να θεωρήσουμε πως τα $\mathcal{A}_1, \mathcal{A}_3$, και $\mathcal{A}_4$ είναι "καλά" τεστ σχετικά με τα $\mathbf{D}_1$ και $\mathbf{D}_2$ επειδή αντίστοιχα οι στατιστικές τους αποστάσεις είναι ακριβώς $\Delta[X,Y]$. Αντίστοιχα τα τεστ $\mathcal{A}_2$ και $\mathcal{A}_5$ θεωρούνται "κακά" επειδή και τα δύο έχουν στατιστική απόσταση 0.

1.8 Πιθανοτικοί Αλγόριθμοι

Οι αλγόριθμοι θα μπορούν να χρησιμοποιούν επιπλέον την εντολή $x \xleftarrow{r} \{0,1\}$ για μια τυχαία μεταβλητή X ομοιόμορφα κατανεμημένη στο $\mathbf{D} = \{0,1\}$. Αυτοί οι αλγόριθμοι ονομάζονται ***πιθανοτικοί (probabilistic)*** και θα λέμε ότι "στρίβουν νομίσματα"

Για κάθε πιθανοτικό αλγόριθμο, το σύνολο των δυνατών αποτελεσμάτων διαμορφώνουν το πεδίο ορισμού της πιθανοτικής κατανομής. Συγκεκριμένα, αν $a \in \{0,1\}$ είναι ένα πιθανό αποτέλεσμα ενός πιθανοτικού αλγορίθμου $\mathcal{A}$ με είσοδο x, ορίζουμε το

$$\text{Prob}[\mathcal{A}(x) = a] = \frac{\#\{b \in \{0,1\}^n : \mathcal{A} \text{ flips } b \text{ and outputs } a\}}{2^n},$$

όπου το n συμβολίζει το πλήθος των στρέψεων νομισμάτων που πραγματοποίησε ο $\mathcal{A}$ δεδομένου του x. Ανάλογα των προδιαγραφών του αλγορίθμου, ο καθορισμός του n μπορεί να είναι περίπλοκος. Μπορούμε να θεωρήσουμε χωρίς βλάβη της γενικότητας όμως, πως ένας πιθανοτικός αλγόριθμος $\mathcal{A}$ κάνει τον ίδιο πλήθος στρέψεων νομισμάτων για κάθε είσοδο ιδίου μήκους. Αυτή η απαίτηση δεν επηρεάζει την υπολογιστική δύναμη του πιθανοτικού μας μοντέλου.

Παράδειγμα. Θεωρείστε τον ακόλουθο αλγόριθμο. Θα τον ονομάζουμε $\mathcal{A}_1$.

1: Είσοδος 1^n
2: Επέλεξε $x_0, \cdots, x_{n-1} \xleftarrow{r} \{0,1\}$
3: αν $\sum_{i=0}^{n-1} 2^i x_i \geq 2^{n-1}$
4: τότε επέστρεψε 1
5: αλλιώς επέστρεψε 0

Αφού το 1 είναι ένα πιθανό αποτέλεσμα,

$$\text{Prob}[\mathcal{A}_1(1^n) = 1] = \frac{\#\{b \in \{0,1\}^n : \mathcal{A} \text{ flips } b \text{ and outputs } 1\}}{2^n} = \frac{2^{n-1}}{2^n} = \frac{1}{2}.$$

Παράδειγμα. Θα ονομάσουμε τον επόμενο αλγόριθμο $\mathcal{A}_2$.

1: Είσοδος 1^n
2: Επανέλαβε n φορές
3: $x \xleftarrow{r} \{0,1\}$
4: αν $x = 1$, Επέστρεψε 1 και Τερμάτισε
5: Επέστρεψε Αποτυχία

Έχουμε τις εξής πιθανότητες

$$\text{Prob}[\mathcal{A}_2(1^n) = \text{Fail}] = \frac{1}{2^n}$$
$$\text{Prob}[\mathcal{A}_2(1^n) = 1] = 1 - \frac{1}{2^n}.$$

Έστω A ένας αριθμός με n bits. Θα ονομάζουμε το αριστερότερο bit στην δυαδική αναπαράσταση του A ως ***περισσότερο σημαντικό bit (most significant bit)***. Για να αποφύγουμε τετριμμένες περιπτώσεις, θα απαιτήσουμε το περισσότερο σημαντικό bit του να είναι 1.

Ακολουθούν τρεις πιθανοτικοί αλγόριθμοι που προσπαθούν να δειγματοληπτήσουν ομοιόμορφα στο $[0, A)$. Για να μετρήσουμε την ποιότητα του δειγματολήπτη, πρέπει να υπολογίσουμε την στατιστική απόσταση της εξόδου του αλγορίθμου απο την ομοιόμορφη κατανομή στο σύνολο $\{0, 1, 2, \ldots, A-1\}$.

Άσκηση: Θεωρήστε το ακόλουθο σύνολο δειγματοληπτών. Μελετήστε την πιθανοτική κατανομή καθενός για να δείτε ποιος έχει είναι πιο κοντά στην ομοιόμορφη κατανομή.

Δειγματολήπτης 1:

1: $n := \lceil \log_2 A \rceil$
2: Επέλεξε: $x_0, x_1, \ldots, x_{n-1} \xleftarrow{r} \{0,1\}$
3: $y := \sum_{i=0}^{n-1} 2^i x_i$
4: Επέστρεψε $y \bmod A$

Δειγματολήπτης 2:

1: Επέλεξε: $x_0, x_1, \ldots, x_{A-1} \xleftarrow{r} \{0,1\}$
2: $y := \sum_{i=0}^{A-1} x_i$
3: Επέστρεψε y

Δειγματολήπτης 3:

1: $n := \lceil \log_2 A \rceil$
2: Επανάλαβε
3: Επέλεξε: $x_0, x_1, \ldots, x_{n-1} \xleftarrow{r} \{0,1\}$
4: $y := \sum_{i=0}^{n-1} 2^i x_i$
5: Αν $y < A$ Επέστρεψε y και Τερμάτισε
6: Διαφορετικά Επανέλαβε

Αναφορές

[1] V. Shoup, *A Computational Introduction to Number Theory and Algebra*, Cambridge University Press, Cambridge, UK, 2005. http://www.shoup.net/ntb/.

Επιμέλεια σημειώσεων Αγγλικά: S. Pehlivanoglu, J. Todd, & H.S. Zhou. Ελληνικά : Μ. Πουντουράκης, Ν. Καρβέλας.