Πιθανότητες 1

28/09/2023

ΓΡΑΦΕΙΣ:
Ειρήνη Θεοδώρου
Κωνσταντίνος Κράββαρης

Πρόβλημα γενεθλίων

Δεδομένης μιας τάξης με n άτομα:

1) Ποια είναι η πιθανότητα τουλάχιστον δύο εξ' αυτών να έχουν γενέθλια την ίδια μέρα;

2) Ποια είναι η ελάχιστη τιμή του n ώστε η πιθανότητα του ερωτήματος (1) να είναι μεγαλύτερη του $1/2$;

Σημείωση: Θεωρούμε ότι κάθε έτος έχει 365 μέρες και κάθε μέρα του έτους είναι ισοπίθανη.

Σκέψη: Γνωρίζουμε ότι η πιθανότητα είναι $\dfrac{\#\text{ ευνοϊκών αποτελεσμάτων}}{\#\text{ συνολικών αποτελεσμάτων}}$

Στο πρόβλημα αυτό όμως είναι δύσκολο να υπολογίσουμε τα ευνοϊκά αποτελέσματα. Είναι πιο εύκολο να βρούμε τα μη ευνοϊκά αποτελέσματα

Μη ευνοϊκά αποτελέσματα = {κανένα ζεύγος}.

365	364	...	365−n+1

Ο πρώτος από τα n άτομα έχει γενέθλια σε οποιαδήποτε μέρα από τις 365 ημέρες του χρόνου. $\rightsquigarrow$ 365

Ο δεύτερος έχει γενέθλια σε οποιαδήποτε μέρα του έτους εκτός της μέρας που έχει γενέθλια ο πρώτος. $\rightsquigarrow$ 364

⋮

Ο n-οστός έχει γενέθλια σε οποιαδήποτε μέρα του χρόνου εκτός των ημερών που έχουν γενέθλια οι $n-1$ προηγούμενοι. $\rightsquigarrow$ $365-n+1$

Άρα # Μη ευνοϊκών αποτελεσμάτων $= 365 \cdot 364 \cdot \ldots \cdot (365-n+1) = \dfrac{(365)!}{(365-n)!}$

Συνεπώς P(τουλάχιστον 2 έχουν γενέθλια την ίδια μέρα) $= 1 - \dfrac{365!}{(365-n)!} \cdot \dfrac{1}{(365)^n}$

Το $(365)^n$ προκύπτει επειδή για τα συνολικά αποτελέσματα κάθε άτομο έχει 365 πιθανές ημέρες γενεθλίων.

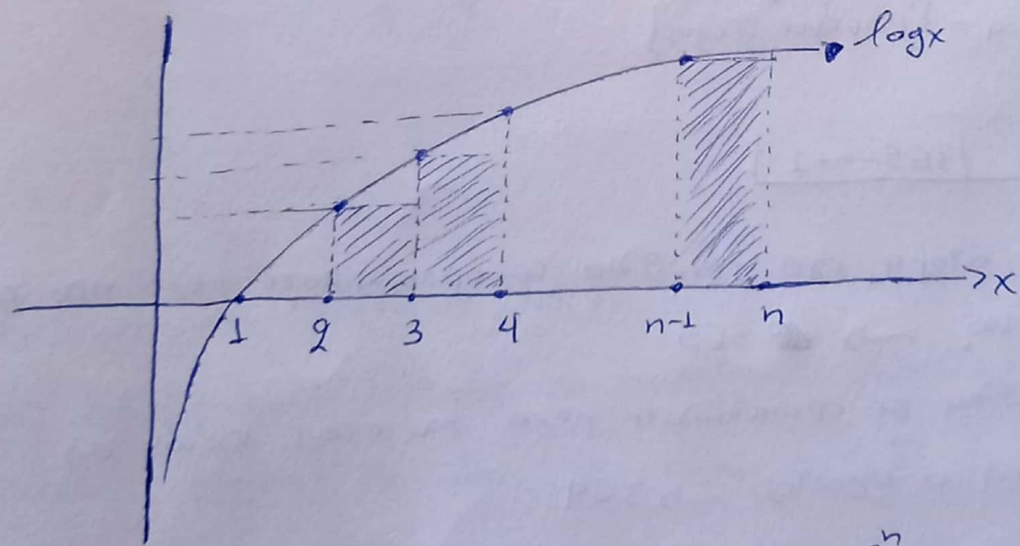
β) Για το ερώτημα αυτό θα χρειαστούμε έναν τρόπο ικανοποιητικής προσέγγισης παραστάσεων που περιέχουν το $n!$ Αυτό μας οδηγεί στον τύπο του Stirling.

Σκοπός: Να βρεθεί μια προσεγγιστική έκφραση για το $n!$

Για να έχουμε καλύτερο έλεγχο της προσέγγισης θα προσπαθήσουμε να προσεγγίσουμε το $\log n!$

Έχουμε $\log n! = \log \prod_{k=1}^{n} k = \sum_{k=1}^{n} \log k$

Θα δοκιμάσουμε να προσεγγίσουμε το άθροισμα με ένα ολοκλήρωμα.



Παίρνουμε τις κάτω προσεγγίσεις και ύστερα τις πάνω

Έχουμε $\log 1 \cdot 1 + \log 2 \cdot 1 + \ldots + \log(n-1) \cdot 1 < \int_1^n \log x \, dx < \log 1 \cdot 1 + \log 2 \cdot 1 + \ldots + \log n \cdot 1$

Συμπεραίνουμε ότι $\int_1^n \log x \, dx < \sum_{k=1}^{n} \log k < \int_1^{n+1} \log x \, dx \implies$

$n\log n - n + 1 < \sum_{k=1}^{n} \log k < (n+1)\log(n+1) - n$

Καθώς $n \to \infty$ έχουμε $\dfrac{(n+1)\log(n+1)-n}{n\log n - n+1} \to 1$. Συμπεραίνουμε λοιπόν ότι

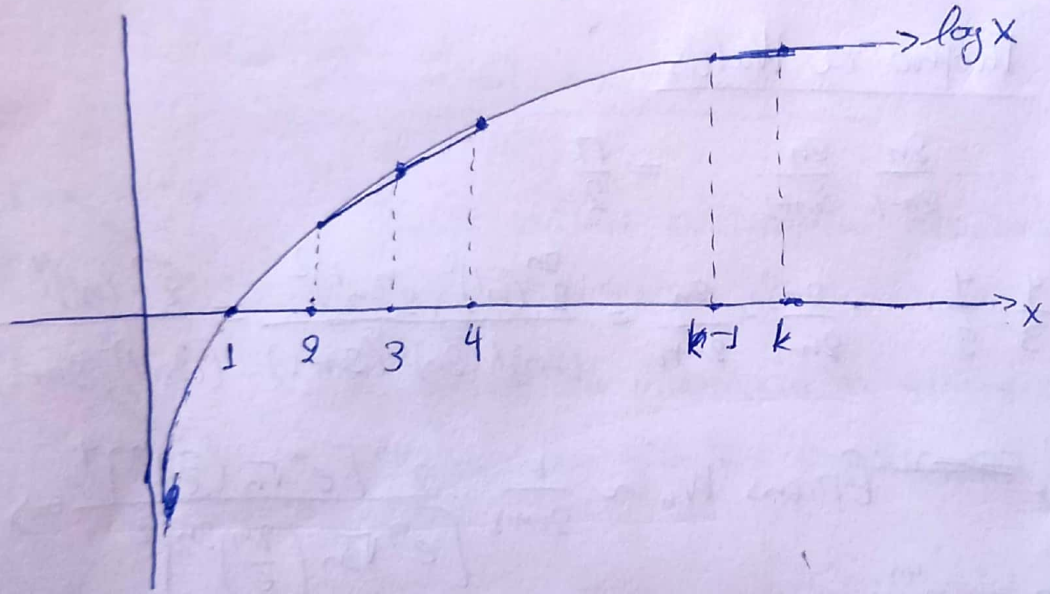
$\log n! \sim n\log n - n$

↳ ασυμπτωτική ισότητα

Σημείωση: $a_n \sim b_n \Longleftrightarrow \lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = 1$ για $(a_n), (b_n)$ ακολουθίες με $a_n > 0$, $b_n > 0$ $\forall n \in \mathbb{N}$.

Έχουμε λοιπόν $n! \sim e^{n\log n - n} = \dfrac{e^{n\log n}}{e^n} = \dfrac{n^n}{e^n} = \left(\dfrac{n}{e}\right)^n$.

Θα προσπαθήσουμε να βελτιώσουμε την προσέγγιση.



Αντί για τα ορθογώνια, θα δουλέψουμε τώρα με τα τραπέζια.
Θα προσεγγίσουμε δηλαδή το εμβαδό της καμπύλης $\log x$ με τραπέζια βάσεων $\log(k-1)$ και $\log k$ που έχουν εμβαδό $E_k = \dfrac{\log(k-1)+\log k}{2}$

Άρα $\int_1^n \log x\,dx \sim \sum_{k=1}^{n} \dfrac{\log(k-1)+\log k}{2} = \dfrac{\log 1+\log 2}{2} + \dfrac{\log 2+\log 3}{2} + \dots + \dfrac{\log(k-1)+\log k}{2}$

$= \log 2 + \log 3 + \dots + \log(n-1) + \log n - \dfrac{\log n}{2}$. Οπότε $\int_1^n \log x\,dx + \dfrac{\log n}{2} \sim \log n!$

Άρα $n\log n - n + 1 + \frac{\log n}{2} \sim \log n!$ Άρα "προκύπτει" $n! \sim e^{n\log n - n} \cdot e \cdot e^{\log n/2}$, δηλαδή,
$n! \sim \left(\frac{n}{e}\right)^n \cdot e\sqrt{n}$

Η προσέγγιση $n! \sim e\sqrt{n}\left(\frac{n}{e}\right)^n$ είναι σωστή μέχρις μίας πολλαπλασιαστικής σταθεράς. Αυτό συμβαίνει επειδή $\log n! \sim n\log n - n + 1 + \frac{\log n}{2}$ είναι σωστή μέχρις επιπέδου $O(1)$.

Συμβολισμός: Για ακολουθίες (a_n), (b_n), $a_n = O(b_n)$ όταν $\limsup\limits_{n\to\infty} \frac{a_n}{b_n} < \infty$

Stirling: Έστω ότι $n! \sim e^a \sqrt{n}\left(\frac{n}{e}\right)^n$ όπου a πρέπει να υπολογισθεί.
Η σκέψη του Stirling ήταν ότι αν μπορέσουμε να βρούμε μια 2η ασυμπτωτική έκφραση για το $n!$, ίσως μπορέσουμε να βγάλουμε μια εξίσωση για το a.

Γινόμενο του Wallis

$$\frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdots\cdots\frac{2n}{2n-1}\cdot\frac{2n}{2n+1}\cdots = \frac{\pi}{2}$$

Έστω $W_n = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdots\cdot\frac{2n}{2n-1}\cdot\frac{2n}{2n+1} = \frac{2^{2n}(n!)^2(2^n\cdot n!)^2}{(2n)!\,(2n)!\,(2n+1)} = \frac{2^{4n}\cdot(n!)^4}{((2n)!)^2(2n+1)}$

Έχουμε $\lim\limits_{n\to\infty} W_n = \frac{\pi}{2}$, ~~επειδή~~ Επίσης $W_n \sim \frac{1}{2n+1}\cdot\frac{2^{4n}\left[e^a\sqrt{n}\left(\frac{n}{e}\right)^n\right]^4}{\left[e^a\sqrt{2n}\left(\frac{2n}{e}\right)^{2n}\right]^2} \Rightarrow$

$$W_n \sim \frac{1}{2n+1}\cdot\frac{2^{4n}\cdot e^{4a}\cdot n^2\cdot\frac{n^{4n}}{e^{4n}}}{e^{2a}\cdot 2n\cdot 2^{4n}\cdot\frac{n^{4n}}{e^{4n}}} \Rightarrow W_n \sim \frac{e^{2a}}{2n+1}\cdot\frac{n}{2}$$

Οπότε $\lim\limits_{n\to\infty} W_n = \frac{\pi}{2} \Rightarrow \lim \frac{e^{2a}}{2n+1}\cdot\frac{n}{2} = \frac{\pi}{2} \Rightarrow \frac{e^{2a}}{4} = \frac{\pi}{2} \Rightarrow e^{2a} = 2\pi \Rightarrow$
$e^a = \sqrt{2\pi}$. Συνεπώς $\boxed{n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n}$ Τύπος του Stirling

Επιστρέφουμε στο πρόβλημα των γενεθλίων. Είχαμε υπολογίσει ότι

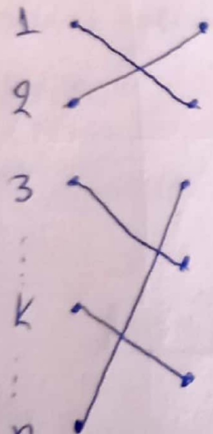
$$P(\text{καμιά κοινή μέρα γενεθλίων}) = \frac{365!}{(365-n)!\,(365)^n} = \frac{\sqrt{365}\cdot\left(\frac{365}{e}\right)^{365}}{(365)^n\cdot\sqrt{365-n}\cdot\left(\frac{365-n}{e}\right)^{365-n}}$$

Κάνοντας πράξεις προκύπτει ότι το ελάχιστο n για το οποίο η παραπάνω πιθανότητα είναι μικρότερη από $1/2$ είναι $n=23$. Άρα το ελάχιστο n τέτοιο ώστε $P(\text{τουλ. μία κοινή ημέρα γενεθλίων}) > \frac{1}{2}$ είναι $n=23$

Πρόβλημα αναδιατάξεων-ανακατατάξεων (derangement)

Έχουμε n αριθμημένες κάρτες $\{1,\dots,n\}$ τις οποίες τοποθετούμε σε μια τυχαία διάταξη. Αν a_n ο αριθμός της n-οστής κάρτας, να υπολογιστεί η πιθανότητα του ενδεχομένου $E=\{\exists n : a_n = n\}$.

Σχηματική αναπαράσταση



Κάθε πιθανό ενδεχόμενο αντιστοιχεί σε μια απεικόνιση $k \to a_k \in \{1,2,\dots,n\}$.

Αναζητούμε την πιθανότητα η απεικόνιση αυτή να έχει ένα τουλάχιστον σταθερό σημείο.

Συμβολισμός: Έστω $D_n =$ αριθμός απεικονίσεων χωρίς κανένα σταθερό σημείο.

Δύο σκέψεις: → Επαγωγικά
→ "Μέτρηση τουλάχιστον k σταθερών σημείων

1) Επαγωγικά: Στα n στοιχεία πόσες απεικονίσεις δεν έχουν σταθερό σημείο;

$$D_n = \underline{n-1} \begin{cases} D_{n-2} \\ D_{n-1} \end{cases}$$ Άρα $D_n = (n-1)\cdot(D_{n-1}+D_{n-2})$