Η Πιθανοθεωρητική Μέθοδος

Θερινό Σχολείο 2010

Διάλεξη 3η

Μιχάλης Κολουντζάκης

Πανεπιστήμιο Κρήτης

Ηράκλειο, Ιούλιος 2010

Πώς ελέγχουμε παραπάνω από μία τυχαία ποσότητα;

Αν X τ.μ. με $\mathbb{E}X \geq \alpha \Rightarrow$ υπάρχει τιμή της X που είναι $\geq \alpha$.

Αν $\mathbb{E}X \geq \alpha$ και $\mathbb{E}Y \geq \beta$
$\not\Rightarrow$ υπάρχει έκβαση του πειράματος όπου $X \geq \alpha$, $Y \geq \beta$.

Αν ξέρουμε όμως $\mathbb{P}[|X - \mathbb{E}X| > \epsilon_1] < \frac{1}{2}$ και $\mathbb{P}[|Y - \mathbb{E}Y| > \epsilon_2] < \frac{1}{2}$
$\Rightarrow$ με θετική πιθανότητα

$$\mathbb{E}X - \epsilon_1 \leq X \leq \mathbb{E}X + \epsilon_1 \quad \underline{\text{ΚΑΙ}} \quad \mathbb{E}Y - \epsilon_2 \leq Y \leq \mathbb{E}Y + \epsilon_2.$$

Χρειαζόμαστε λοιπόν
(α) Πληροφορία για τις μέσες τιμές των μεταβλητών μας, και
(β) Άνω φράγματα για τις <u>πιθανότητες απόκλισης</u> των τ.μ. από τις μέσες τιμές

Μερικές γενικές ανισότητες απόκλισης

Ανισότητα Markov

Αν $X \geq 0$ και $\mu = \mathbb{E}X$ τότε $\mathbb{P}\left[X \geq \lambda\mu\right] \leq \frac{1}{\lambda}$.

Ανισότητα Chebyshev

Αν $X \geq 0$, $\mu = \mathbb{E}X$ και $\sigma^2 = \mathbb{E}|X-\mu|^2$ τότε

$$\mathbb{P}\left[|X-\mu| \geq \lambda\sigma\right] \leq \frac{1}{\lambda^2}.$$

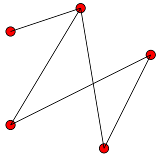
Ήδη χρησιμοποιήσαμε την ανισότητα Chebyshev για την ανάλυση της μεθόδου Monte Carlo για υπολογισμό εμβαδών (Διάλεξη 1η).

Πολύ εύχρηστες γιατί δεν έχουν ιδιαίτερες απαιτήσεις για την X.

Ασθενείς στο αποτέλεσμά τους όμως.

Παράδειγμα: Τυχαία γραφήματα Erdős–Rényi

Τυχαίο γράφημα $G(n,p)$:
(α) n κορυφές και
(β) η ακμή ανάμεσα στις i,j υπάρχει με πιθ. p, ανεξάρτητα από τις άλλες ακμές.



Πολύ σημαντικό μοντέλο διαφόρων καταστάσεων. Ορίστηκε από τους Erdős και Rényi (1959).

Για ποιες τιμές του p περιέχει το $G(n,p)$ "σχεδόν σίγουρα" ένα πλήρες υπογράφημα με 4 κορυφές (K_4);

Αλλαγή φάσης για $p \sim n^{-2/3}$:

Για $n \to \infty$:
(α) $\frac{p}{n^{-2/3}} \to \infty \Rightarrow$ η πιθανότητα να υπάρχει K_4 τείνει στο 1
(β) $\frac{p}{n^{-2/3}} \to 0 \Rightarrow$ η πιθανότητα να υπάρχει K_4 τείνει στο 0.

K_4 σε τυχαία γραφήματα Erdős–Rényi (συνέχεια)

Αραιά γραφήματα ($pn^{2/3} \to 0$)

Για $S \subseteq \{1, 2, \ldots, n\}$ μεγέθους 4:

$A_S = \{$υπάρχουν όλες οι ακμές ανάμεσα στις κορυφές του $S\}$

$X_S = \mathbf{1}(A_S)$

$\mathbb{P}[A_S] = \mathbb{P}[X_S = 1] = p^6$ (4 κορυφές ορίζουν 6 ακμές)

$X = \sum_S X_S =$ πόσα αντίγραφα του K_4 υπάρχουν στο G

$$\mathbb{E}X = \sum_S \mathbb{E}X_S = \sum_S \mathbb{P}[A_S] = \binom{n}{4} p^6 = \frac{n(n-1)(n-2)(n-3)}{4!} p^6 \sim \frac{1}{24} n^4 p^6$$

Αν $p = cn^{-2/3}$ τότε $\mathbb{E}X \sim \frac{c}{24}$

Ανισότητα Markov $\Rightarrow$ $\mathbb{P}[X \geq 1] = \mathbb{P}\left[X \geq \frac{1}{\mathbb{E}X}\mathbb{E}X\right] \leq \mathbb{E}X \sim \frac{c}{24}$

Όταν $c \to 0$ έπεται $\mathbb{P}[X_S \geq 1] \to 0$.

K_4 σε τυχαία γραφήματα Erdős–Rényi (συνέχεια)

Πυκνά γραφήματα ($pn^{2/3} \to \infty$)

$$\begin{aligned} \operatorname{Var} X &= \operatorname{Var} \sum_S X_S \\ &= \sum_S \operatorname{Var} X_S + \sum_{S \neq T} \operatorname{Cov}(X_S, X_T) \end{aligned}$$

όπου $\operatorname{Cov}(Z, W) = \mathbb{E}\left[(Z - \mathbb{E}Z) \cdot (W - \mathbb{E}W)\right] = \mathbb{E}(ZW) - \mathbb{E}Z \cdot \mathbb{E}W$.

$\operatorname{Var} X_S = \mathbb{E}X_S^2 - (\mathbb{E}X_S)^2 \leq \mathbb{E}X_S^2 = \mathbb{E}X_S = p^6$

$\sum_S \operatorname{Var} X_S \leq \frac{1}{24} n^4 p^6$ αφού υπάρχουν $\binom{n}{4} \leq \frac{1}{24} n^4$ σύνολα S

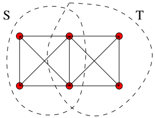
K_4 σε τυχαία γραφήματα Erdős–Rényi (συνέχεια)

Πυκνά γραφήματα ($pn^{2/3} \to \infty$)

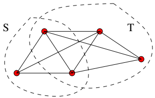
$S \cap T = \emptyset \Rightarrow X_S, X_T$ ανεξάρτητες $\Rightarrow \mathrm{Cov}\,(X_S, X_T) = 0$.

S και T πρέπει να έχουν κοινές ακμές $\Rightarrow$ ισχύει $|S \cap T| = 2$ ή 3.

Υπάρχουν $O(n^6)$ ζεύγη S, T με 2 κοινές κορυφές και $O(n^5)$ ζεύγη με 3 κοινές κορυφές.



1η περίπτωση: $\mathrm{Cov}\,(X_S, X_T) \le \mathbb{E}(X_S X_T) = p^{11}$



2η περίπτωση: $\mathrm{Cov}\,(X_S, X_T) \le \mathbb{E}(X_S X_T) = p^{9}$

K_4 σε τυχαία γραφήματα Erdős–Rényi (τέλος)

Ανισότητα Chebyshev:
$\mathbb{P}[X=0] \le \mathbb{P}\left[|X-\mathbb{E}X| \le \frac{\mathbb{E}X}{\sigma(X)}\sigma(X)\right] \le \frac{\mathrm{Var}\,X}{(\mathbb{E}X)^2}$
$\mathrm{Var}\,X = \sum_S \mathrm{Var}\,X_S + \sum_{S\neq T}\mathrm{Cov}\,(X_S, X_T) = O(n^4p^6 + n^6p^{11} + n^5p^9)$

$(\mathbb{E}X)^2 \sim (n^4p^6)^2 = n^8p^{12}$
Αν $p = \phi(n)n^{-2/3}$ όπου $\phi(n)\to\infty$ τότε

$$(\mathbb{E}X)^2 \sim n^{8-12(2/3)}\phi(n)^{12} = \phi(n)^{12}$$

και

$$\begin{aligned}
\mathrm{Var}\,X &= O(n^4p^6 + n^6p^{11} + n^5p^9)\\
&= O(n^{4-6(2/3)}\phi(n)^6 + n^{6-11(2/3)}\phi(n)^{11} + n^{5-9(2/3)}\phi(n)^9)\\
&= O(\phi(n)^6 + n^{-4/3}\phi(n)^{11} + n^{-1}\phi(n)^9)\\
&= o(\phi(n)^{12}).
\end{aligned}$$

άρα $\mathbb{P}[X=0]\to 0$.

Μια εκθετική ανισότητα απόκλισης

$X_1, \ldots, X_N$ ανεξάρτητες, $\mathbb{E}X_j = 0$ και για κάθε j: $\max X_j - \min X_j \le A$,

$$X = X_1 + \cdots + X_N.$$

Τότε

$$\mathbb{P}\left[|X| > a\right] \le 2e^{-2a^2/(A\cdot N)}.$$

Παρατηρείστε την εκθετική εξάρτηση από το a.

Οφείλεται στην εσωτερική δομή της X ως άθροισμα πολλών ανεξαρτήτων μεταβλητών.

Παράδειγμα: Monte Carlo για εμβαδά, 2η ανάλυση.

Έστω $\Omega \subseteq [0,1]^2$, $Y_1, \ldots, Y_N \in [0,1]^2$ ανεξάρτητα και ομοιόμορφα

$$X_j = \mathbf{1}\left(Y_j \in \Omega\right) - |\Omega| = \begin{cases} 1 - |\Omega| & \text{με πιθ. } p_j = |\Omega| \\ -|\Omega| & \text{με πιθ. } 1 - p_j = 1 - |\Omega| \end{cases}$$

$$X = X_1 + \cdots + X_N = (\text{αριθμ. σημείων στο } \Omega) - N|\Omega| = S - N|\Omega|$$

όπου $S = (\text{αριθμ. σημείων στο } \Omega)$. Με $a = \epsilon N^{1/2}$, $A = 1$

$$\begin{aligned} \mathbb{P}\left[\left|S - N|\Omega|\right| > \epsilon N^{1/2}\right] &\leq 2e^{-2a^2/(A\cdot N)} = 2e^{-2\epsilon^2} \Rightarrow \\ \mathbb{P}\left[\left|S/N - |\Omega|\right| > \epsilon N^{-1/2}\right] &\leq 2e^{-2\epsilon^2} \Rightarrow \\ \mathbb{P}\left[\left|S/N - |\Omega|\right| > \lambda\right] &\leq 2e^{-2\lambda^2 N} \text{ με } \lambda = \epsilon/N^{1/2} \end{aligned}$$

Με $\lambda = 10^{-3}$ και $N = \frac{\ln 200}{2\cdot 10^{-6}} = 2.64 \cdot 10^6$ πετυχαίνουμε το δεξί μέλος $\leq 10^{-2}$.

Σύγκριση: Με ανισότητα Chebyshev (1η διάλεξη) θέλαμε $N \geq 10^8$.

Παράδειγμα: Προσθαφαιρέσεις διανυσμάτων

Για $v = (v_1, \ldots, v_d) \in \mathbb{R}^d$ γράφουμε $\|v\|_\infty = \max\{|v_1|, \ldots, |v_d|\}$.

N διανύσματα στο $\mathbb{R}^d$: $v_1, \ldots, v_N$ με $\|v_j\|_\infty \le 1$. Τότε μπορούμε να επιλέξουμε τα πρόσημα $\epsilon_1 = \pm 1, \ldots, \epsilon_N = \pm 1$ ώστε το

$$u = \epsilon_1 v_1 + \cdots + \epsilon_N v_N$$

να έχει $\|u\|_\infty \le \sqrt{N \ln(2d)}$.

Επιλέγουμε και πάλι $\epsilon_j = \pm 1$ με ίση πιθανότητα και ανεξάρτητα.

Παράδειγμα: Προσθαφαιρέσεις διανυσμάτων (συνέχεια)

Γράφουμε $v_i = (v_{i,1}, v_{i,2}, \ldots, v_{i,d})$.

Η τ.μ. $u_j = \sum_{i=1}^{N} \epsilon_i v_{i,j}$, για $j = 1, 2, \ldots, d$.

Έχουμε $\max(\epsilon_i v_{i,j}) - \min(\epsilon_i v_{i,j}) \le 2$, άρα $A = 2$.

$B_j = \{|u_j| > t\}, j = 1, 2, \ldots, d$ (κακά ενδεχόμενα)

$\mathbb{P}[B_j] \le 2e^{-t^2/N}$ (από εκθετική ανισότητα απόκλισης)

$\Rightarrow \mathbb{P}\left[\bigcup_{j=1}^{d} B_j\right] \le \sum_{j=1}^{d} \mathbb{P}[B_j] \le 2de^{-t^2/N} =: p$

Αν $p < 1$ τότε με θετική πιθανότητα δεν ισχύει κανένα B_j

$\Rightarrow \|u\|_\infty \le t$.

Επιλέγοντας $t > \sqrt{N \ln(2d)}$ παίρνουμε $p < 1$