Ανισότητες

Πιθανότητες και Στατιστική

Τμήμα Πληροφορικής, Πανεπιστήμιο Πειραιώς

2023-2024

Ανισότητα του Markov

Σε κάποια προβλήματα ενδέχεται να μην έχουμε πολλές πληροφορίες για μια ΤΜ μεταβλητή εκτός από τη μέση τιμή της, ή και τη διακύμανσή της. Σε αυτές τις περιπτώσεις, μπορούμε να βγάλουμε κάποια συμπεράσματα για την ΤΜ X, χρησιμοποιώντας ανισότητες όπως του Markov ή του Chebyshev, οι οποίες παρουσιάζονται στις επόμενες ενότητες.

Πρόταση (Ανισότητα του Markov)

Αν μια (διακριτή ή συνεχής) ΤΜ X λαμβάνει μόνο μη αρνητικές τιμές και έχει πεπερασμένη μέση τιμή $\mu = E(X)$, τότε

$$P(X \geq c) \leq \frac{\mu}{c}, \qquad \text{για κάθε πραγματική σταθερά } c > 0.$$

Ανισότητα του Markov

Απόδειξη.

Θα αποδείξουμε την ανισότητα στην περίπτωση όπου η X είναι συνεχής. (Αν η X είναι διακριτή η απόδειξη είναι ανάλογη.) Έστω $f(x)$ η PDF της X. Προφανώς, είναι $x < 0 \Rightarrow f(x) = 0$, οπότε

$$\begin{aligned}\mu &= \int_0^{+\infty} xf(x)dx = \int_0^c xf(x)dx + \int_c^{+\infty} xf(x)dx \\ &\geq \int_c^{+\infty} xf(x)dx \geq \int_c^{+\infty} cf(x)dx = c\int_c^{+\infty} f(x)dx = cP(X \geq c). \square\end{aligned}$$

Διαισθητικά, αν μια ΤΜ έχει μικρή μέση τιμή, τότε η πιθανότητα να λάβει μεγάλες τιμές είναι μικρή.

Η σημασία της ανισότητας του Markov είναι ότι αν και το μόνο που γνωρίζουμε για μια ΤΜ είναι η μέση τιμή της, εν τούτοις μπορούμε να βγάλουμε κάποιο χρήσιμο συμπέρασμα.

Ανισότητα του Markov

Παράδειγμα

Μια μη αρνητική ΤΜ X έχει μέση τιμή $\mu = 5$. Να βρεθεί ένα φράγμα για την πιθανότητα $P(X \geq 20)$.

Λύση

Από την ανισότητα του Markov έχουμε ότι

$$P(X \geq 20) \leq \frac{\mu}{20} = \frac{5}{20} = \frac{1}{4} = 0.25$$

δηλαδή το πολύ 25% των τιμών που λαμβάνει η X είναι μεγαλύτερες ή ίσες του 20, ενώ το 75% των τιμών της ανήκει στο διάστημα 0 έως 20.

Ανισότητα του Markov

Παράδειγμα

Ένας βιολόγος ισχυρίζεται ότι το μέσο βάρος του αφρικανικού χελιδονιού είναι 100 γραμμάρια, ενώ το 60% των αφρικανικών χελιδονιών έχει βάρος μεγαλύτερο ή ίσο από 200 γραμμάρια. Είναι δυνατόν να ευσταθεί ο ισχυρισμός αυτός;

Λύση

Έστω X η ΤΜ βάρος (σε γραμμάρια) του αφρικανικού χελιδονιού. Αν η ΤΜ X έχει μέση τιμή $\mu = 100$, τότε από την ανισότητα του Markov έχουμε ότι

$$P(X \geq 200) \leq \frac{\mu}{200} = \frac{100}{200} = 0.5,$$

το οποίο δεν συμφωνεί με τον ισχυρισμό του βιολόγου ότι η πιθανότητα αυτή είναι 60%.
Μάλιστα, βάσει της παραπάνω ανισότητας, θα πρέπει να είναι $\mu \geq 120$.

Ανισότητα του Markov

Παράδειγμα

Έστω ότι ο μέσος βαθμός σε ένα μάθημα είναι 7 (με άριστα το 10). Να βρεθεί ένα κάτω φράγμα για το ποσοστό των φοιτητών που έχουν βαθμολογηθεί με βαθμό μικρότερο του 8.

Λύση

Αν η ΤΜ X αντιστοιχεί στον βαθμό ενός τυχαία επιλεγμένου φοιτητή, τότε

$$P(X < 8) = 1 - P(X \geq 8) \geq 1 - \frac{\mu}{8} = 1 - \frac{7}{8} = \frac{1}{8} = 0.125,$$

δηλαδή τουλάχιστον το $1/8$ των φοιτητών πήραν βαθμό κάτω του 8.

Ανισότητα του Markov

Παράδειγμα

Έστω $\sigma = \sigma(1)\sigma(2)\cdots\sigma(n)$ μια μετάθεση του $[n]$. Αν $\sigma(i) = i$ τότε λέμε ότι η σ έχει σταθερό σημείο το i. Να δειχθεί ότι με πιθανότητα τουλάχιστον $1 - \frac{1}{k}$ η σ έχει λιγότερα από k σταθερά σημεία.

Λύση

Έστω X ο αριθμός των σταθερών σημείων της μετάθεσης σ και έστω $X_i = \begin{cases} 1, & \sigma(i) = i \\ 0, & \sigma(i) \neq i \end{cases}$, για κάθε $i \in [n]$. Προφανώς, $X = X_1 + X_2 + \cdots + X_n$ και $P(X_i = 1) = \frac{1}{n}$, άρα $E(X_i) = \frac{1}{n}$. Επομένως,

$$E(X) = E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots E(X_n) = \frac{n}{n} = 1.$$

Επομένως, $P(X < k) = 1 - P(X \geq k) \geq 1 - \frac{E(X)}{k} = 1 - \frac{1}{k}$.

Ανισότητα του Markov

Παράδειγμα

Ο επόμενος αλγόριθμος εντοπίζει το μέγιστο στοιχείο μιας λίστας $a = (a[0], \ldots, a[n-1])$, με n διαφορετικά στοιχεία. Να υπολογισθούν:

i) Το αναμενόμενο πλήθος αναθέσεων (πόσες φορές αλλάζει η τιμή της μεταβλητής m).

ii) Ένα άνω φράγμα για την πιθανότητα να γίνουν περισσότερες από 15 αναθέσεις.

```python
def maxInList(a):
    m=a[0]
    for i in range(len(a)):
      if m < a[i]: m = a[i]
    return m
```

Λύση

i) Έστω X_i η δείκτρια ΤΜ του ενδεχομένου να γίνει ανάθεση στη θέση $i-1 \in [n]$. Επειδή η τιμή $a[i-1]$ είναι η μέγιστη των $a[0], a[1], \ldots, a[i-1]$ με πιθανότητα $1/i$, έπεται ότι $X_i \sim \mathrm{Bernoulli}(1/i)$, οπότε το πλήθος αναθέσεων είναι $X = X_1 + X_2 + \cdots + X_n$, με

$$E(X) = E(X_1 + X_2 + \cdots + X_n) = \sum_{i=1}^{n} \frac{1}{i},$$

όπου το τελευταίο άθροισμα είναι ο n-οστός αρμονικός αριθμός H_n.
ii) Ως γνωστό, ισχύει η σχέση $\ln(1+n) \leq H_n \leq 1 + \ln n$, οπότε, εφαρμόζοντας την ανισότητα Markov για $c = 16$, έχουμε ότι

$$P(X \geq 16) \leq \frac{E(X)}{16} \leq \frac{1 + \ln n}{16}.$$

Για παράδειγμα, για $n = 100$, βρίσκουμε ότι $P(X \geq 16) \leq 0.35$, ενώ για $n = 1000$, βρίσκουμε ότι $P(X \geq 16) \leq 0.494$.

Ανισότητα του Markov

Πρόταση

Έστω ΤΜ X και έστω $c \in \mathbb{R}$.

i) Αν $E(X) \leq c$, τότε $P(X \leq c) > 0$.

ii) Επιπλέον, αν $S_x \subseteq \mathbb{N}$ και $E(X) < 1$, τότε $P(X = 0) > 0$.

Απόδειξη.

i) Αν η X είναι διακριτή και $P(X \leq c) = 0$, τότε είναι

$$E(X) = \sum_{k \geq \lfloor c \rfloor + 1} kP(X = k) > \sum_{k \geq \lfloor c \rfloor + 1} cP(X = k) = c,$$

το οποίο είναι άτοπο.

Απόδειξη (συνέχεια).

Αν η X είναι συνεχής και $P(X \leq c) = F(c) = 0$, τότε υπάρχει $n \in \mathbb{N}^*$ ώστε $F(c + 1/n) < 1$, οπότε

$$\begin{aligned}
E(X) &= \int_c^{+\infty} xf(x)dx = \int_c^{c+1/n} xf(x)dx + \int_{c+1/n}^{+\infty} xf(x)dx \\
&\geq c \int_c^{c+1/n} f(x)dx + (c + 1/n) \int_{c+1/n}^{+\infty} f(x)dx \\
&= cF(c + 1/n) + (c + 1/n)(1 - F(c + 1/n)) \\
&= c + 1/n(1 - F(c + 1/n)) > c,
\end{aligned}$$

το οποίο είναι άτοπο.

ii) Εφαρμόζοντας, την ανισότητα Markov, έχουμε ότι

$$P(X = 0) = 1 - P(X \geq 1) \geq 1 - E(X) > 0.$$

□

Ανισότητα του Markov

Παράδειγμα

Αν $v_1, v_2, \ldots, v_n$ είναι μοναδιαία διανύσματα του $\mathbb{R}^n$, να δειχθεί ότι υπάρχουν συντελεστές $x_1, x_2, \ldots, x_n \in \{-1, +1\}$, ώστε $|x_1 v_1 + \cdots + x_n v_n| \leq \sqrt{n}$.

Λύση

Θεωρούμε την ΤΜ $Y = |X_1 v_1 + \cdots + X_n v_n|$, όπου οι ανεξάρτητες ΤΜ X_i παίρνουν τιμές ομοιόμορφα στο $\{-1, 1\}$. Τότε, είναι

$$E(X_i X_j) = \begin{cases} 0, & i \neq j, \\ 1, & i = j, \end{cases} \qquad \text{και}$$

$$E(Y^2) = E\left(\sum_{i,j} X_i X_j \langle v_i, v_j \rangle\right) = \sum_{i,j} \langle v_i, v_j \rangle E(X_i X_j) = \sum_i |v_i|^2 = n.$$

Επομένως, εφαρμόζοντας την προηγούμενη πρόταση για $c = n$, προκύπτει ότι $P(Y \leq \sqrt{n}) = P(Y^2 \leq n) > 0$.

Ανισότητα Cauchy-Schwarz

Πρόταση (Ανισότητα Cauchy-Schwarz)

Για οποιεσδήποτε ΤΜ X, Y ισχύει ότι

$$(E(XY))^2 \leq E(X^2)E(Y^2)$$

Απόδειξη.

Για $a \in \mathbb{R}$, η ΤΜ $(X + aY)^2$ παίρνει μη αρνητικές τιμές, άρα και η μέση τιμή της είναι μη αρνητική. Επομένως,

$$0 \leq E((X+aY)^2) = E(X^2+2aXY+a^2Y^2) = a^2E(Y^2)+2aE(XY)+E(X^2).$$

Αφού η ανισότητα ισχύει για κάθε a, έπεται ότι η διακρίνουσα $(2E(XY))^2 - 4E(Y^2)E(X^2)$ είναι μικρότερη ή ίση του 0, οπότε προκύπτει το ζητούμενο. $\square$

Ανισότητα Chernoff

Πρόταση (Ανισότητα του Chernoff)

Έστω ΤΜ X και έστω $c \in \mathbb{R}$. Τότε, για κάθε $t > 0$ ισχύει ότι

$$P(X \geq c) \leq \frac{E(e^{tX})}{e^{tc}}.$$

Απόδειξη.

Θεωρούμε την ΤΜ $Y = e^{tX} \geq 0$. Δεδομένου ότι $e^{tc} \geq 0$, εφαρμόζοντας την ανισότητα Markov, έχουμε ότι

$$P(X \geq c) = P(tX \geq tc) = P(e^{tX} \geq e^{tc}) \leq \frac{E(e^{tX})}{e^{tc}}.$$

□

Η παραπάνω ανισότητα δίνει συχνά καλύτερο άνω φράγμα από την ανισότητα Markov, διότι εφαρμόζεται για οποιοδήποτε $t > 0$ και το φράγμα που δίνει μειώνεται εκθετικά συναρτήσει του t.

Παράδειγμα

Έστω $X \sim E(\lambda)$, $\lambda > 0$ και $c > 0$. Να ευρεθεί ένα άνω φράγμα για την πιθανότητα $P(X \leq c)$ χρησιμοποιώντας την ανισότητα Markov και την ανισότητα Chernoff. Σε ποιες περιπτώσεις η δεύτερη δίνει μικρότερο άνω φράγμα;

Λύση

$$E(e^{tX}) = \int_0^{+\infty} e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^{+\infty} e^{(t-\lambda)x} dx = \lambda \left[\frac{e^{(t-\lambda)x}}{t-\lambda} \right]_0^{+\infty} dx$$

$$= \begin{cases} \dfrac{\lambda}{\lambda - t}, & 0 < t < \lambda, \\ +\infty, & t \geq \lambda. \end{cases}$$

Ανισότητα Chernoff

Λύση (συνέχεια)

Επομένως, σύμφωνα με την ανισότητα Chernoff, είναι

$$P(X \geq c) \leq \frac{E(e^{tX})}{e^{tc}} = \frac{\lambda e^{-tc}}{\lambda - t}$$

Θεωρώντας την συνάρτηση $g(t) = \dfrac{\lambda e^{-tc}}{\lambda - t}/(0, \lambda)$, βρίσκουμε ότι παρουσιάζει ολικό ελάχιστο όταν $t = \lambda - 1/c$. Επομένως,

$$P(X \geq c) \leq \frac{\lambda e^{(1/c-\lambda)c}}{1/c} = c\lambda e^{1-\lambda c}$$

Από την άλλη, εφαρμόζοντας την ανισότητα Markov, έχουμε ότι

$$P(X \geq c) \leq \frac{E(X)}{c} = \frac{1}{\lambda c}$$

Λύση (συνέχεια)

Για να εξετάσουμε πότε είναι $c\lambda e^{1-\lambda c} \leq \frac{1}{\lambda c}$, θεωρούμε τη συνάρτηση $h(x) = 1/x - xe^{1-x}/(0, +\infty)$ και βρίσκουμε ότι $h(x) \geq 0$ ανν $x \leq 1$ ή $x \geq 3.513$.
Όταν λοιπόν είναι $c\lambda \leq 1$ ή $c\lambda \geq 3.513$, το πρώτο φράγμα είναι μικρότερο.
Για παράδειγμα, αν $\lambda = 1$, $c = 10$, τότε $\frac{1}{c\lambda} = 0.1$ και $c\lambda e^{1-\lambda c} = 0.0012341$.

Ανισότητα του Chebyshev

Πρόταση (Ανισότητα του Chebyshev)

Αν μια (διακριτή ή συνεχής) ΤΜ X έχει πεπερασμένη μέση τιμή μ και διακύμανση σ^2, τότε

$$P(|X-\mu| \geq c) \leq \frac{\sigma^2}{c^2}, \qquad \text{για κάθε πραγματική σταθερά } c > 0.$$

ή, ισοδύναμα αν $c = k\sigma$,

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}, \qquad \text{για κάθε πραγματική σταθερά } k > 0.$$

Απόδειξη.

Εφαρμόζοντας την ανισότητα του Markov για την ΤΜ $Y = (X-\mu)^2 \geq 0$, προκύπτει ότι

$$P(|X-\mu| \geq c) = P((X-\mu)^2 \geq c^2) \leq \frac{E((X-\mu)^2)}{c^2} = \frac{\sigma^2}{c^2}. \qquad \square$$

Παράδειγμα

Έστω ότι η ΤΜ X έχει πεπερασμένες μέση τιμή μ και διακύμανση σ^2. Να βρεθεί ένα κάτω φράγμα για την πιθανότητα $P(\mu - 2\sigma < X < \mu + 2\sigma)$.

Λύση

Εφαρμόζοντας την ανισότητα του Chebyshev για $c = 2\sigma > 0$, έχουμε ότι

$$\begin{aligned} P(\mu - 2\sigma < X < \mu + 2\sigma) &= P(-2\sigma < X - \mu < 2\sigma) = P(|X - \mu| < 2\sigma) \\ &= 1 - P(|X - \mu| \geq 2\sigma) \geq 1 - \frac{\sigma^2}{(2\sigma)^2} = \frac{3}{4}, \end{aligned}$$

δηλαδή τουλάχιστον το 75% των τιμών που λαμβάνει η ΤΜ X είναι στο διάστημα $[\mu - 2\sigma, \mu + 2\sigma]$, δηλαδή απέχει το πολύ 2σ (2 φορές την τυπική απόκλιση) από τη μέση τιμή μ.

Παράδειγμα

Έστω ότι η ΤΜ X έχει μέση τιμή $\mu = 0$ και διακύμανση $\sigma^2 = 6^2$. Να βρεθεί μια τιμή $\theta > 0$, ώστε με πιθανότητα τουλάχιστον $a = 0.99$ οι τιμές της X να βρίσκονται στο διάστημα $(-\theta, \theta)$.

Λύση

Θέλουμε $P(|X| < \theta) \geq a$, δηλαδή $P(|X| \geq \theta) \leq 1 - a$. Από την ανισότητα του Chebyshev, έχουμε ότι

$$P(|X| \geq \theta) = P(|X - 0| \geq \theta) \leq \frac{\sigma^2}{\theta^2}.$$

Αρκεί λοιπόν $\dfrac{\sigma^2}{\theta^2} \leq 1 - a \Leftrightarrow \dfrac{\sigma^2}{\theta^2} \leq \sqrt{1-a} \Leftrightarrow \theta \geq \dfrac{\sigma}{\sqrt{1-a}}$.

Άρα, για $\theta = \dfrac{\sigma}{\sqrt{1-a}} = \dfrac{6}{0.1} = 60$, με πιθανότητα τουλάχιστον $a = 99\%$, η X λαμβάνει τιμές στο διάστημα $(-60, 60)$.

Ανισότητα του Chebyshev

Πόσο ακριβές είναι το φράγμα της ανισότητας του Chebyshev

Έστω ότι η συνεχής ΤΜ X λαμβάνει τιμές στο διάστημα $[-k, k]$, $k > 0$, και ακολουθεί την ομοιόμορφη κατανομή. Να βρεθεί η πιθανότητα $P(|X - \mu| \geq c)$, $c > 0$ και να συγκριθεί με το φράγμα που δίνει η ανισότητα του Chebyshev.

Λύση

Αφού η X ακολουθεί την ομοιόμορφη κατανομή στο διάστημα $[-k, k]$ θα είναι $\mu = \dfrac{-k+k}{2} = 0$ και $\sigma^2 = \dfrac{(k-(-k))^2}{12} = \dfrac{(2k)^2}{12} = \dfrac{k^2}{3}$.
Οι PDF και CDF της X είναι αντίστοιχα οι

$$f(x) = \begin{cases} \dfrac{1}{2k}, & x \in [-k, k], \\ 0, & x \notin [-k, k], \end{cases} \qquad F(x) = \begin{cases} 0, & x < -k, \\ \dfrac{x+k}{2k}, & x \in [-k, k], \\ 1, & x > k. \end{cases}$$

Λύση (συνέχεια)

Επομένως,

$$
\begin{aligned}
P(|X-\mu| \geq c) &= P(|X| \geq c) = 1 - P(|X| \leq c) = 1 - P(-c \leq X \leq c) \\
&= 1 - (F(c) - F(-c)) = \begin{cases} 1 - \dfrac{c}{k}, & c \in [0, k] \\ 0, & c > k. \end{cases}
\end{aligned}
$$

Η ανισότητα του Chebyshev δίνει $P(|X-\mu| \geq c) \leq \dfrac{\sigma^2}{c^2} = \dfrac{k^2}{3c^2}$.
Αν για παράδειγμα $k = 10$ και $c = 7$, τότε έχουμε ότι
$P(|X-\mu| \geq 7) = 1 - \dfrac{7}{10} = \dfrac{3}{10} = 0.3$, ενώ η ανισότητα Chebyshev δίνει
$P(|X-\mu| \geq 7) \leq \dfrac{10^2}{3 \cdot 7^2} = 0.68$, δηλαδή, ενώ η πραγματική πιθανότητα είναι 0.3, το φράγμα 0.68 του δίνει η ανισότητα Chebyshev είναι υπερεκτίμηση της πραγματικής τιμής.

Ο νόμος των μεγάλων αριθμών

Πρόταση

Αν $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες τυχαίες μεταβλητές με μέση τιμή μ και διακύμανση σ^2 και $\epsilon > 0$ και $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, τότε

$$P\left(|\overline{X} - \mu| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2}.$$

Απόδειξη.

Από την γραμμικότητα της μέσης τιμής, ισχύει ότι

$$E(\overline{X}) = \frac{1}{n}\left(E(X_1) + E(X_2) + \cdots + E(X_n)\right) = \frac{1}{n} n\mu = \mu.$$

Επιπλέον, επειδή οι X_i είναι ανεξάρτητες, ισχύει ότι

$$V(\overline{X}) = \frac{1}{n^2}\left(V(X_1) + V(X_2) + \cdots + V(X_n)\right) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

Επομένως, εφαρμόζοντας την ανισότητα του Chebyshev, προκύπτει ότι

$$P\left(\left|\overline{X} - \mu\right| \geq \epsilon\right) = P\left(|X - E(X)| \geq \epsilon\right) \leq \frac{V(X)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}. \qquad \square$$

Πρόταση (Ασθενής νόμος των μεγάλων αριθμών)

Έστω $X_1, X_2, \ldots, X_n$ ανεξάρτητες ΤΜ που ακολουθούν την ίδια κατανομή με μέση τιμή μ. Τότε, για κάθε $\epsilon > 0$ έχουμε ότι

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + X_2 + \cdots X_n}{n} - \mu\right| < \epsilon\right) = 1.$$

Με άλλα λόγια, αν καταγράψουμε τις τιμές X_1, X_2, X_3, …, X_n, …, μιας ΤΜ X, τότε ο εμπειρικός μέσος των τιμών αυτών με πιθανότητα που τείνει στο 1 είναι οσοδήποτε κοντά στην μέση τιμή μ της X.
Μια εφαρμογή του ασθενούς νόμου των μεγάλων αριθμών είναι ότι μπορούμε να εκτιμήσουμε την μέση τιμή μιας ΤΜ X με βάση τον εμπειρικό μέσο όρο των τιμών της.

Ο νόμος των μεγάλων αριθμών

Πρόταση (Ισχυρός νόμος των μεγάλων αριθμών)

Έστω $X_1, X_2, \dots, X_n$ ανεξάρτητες ΤΜ που ακολουθούν την ίδια κατανομή με μέση τιμή μ. Τότε,

$$P\left(\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu\right) = 1.$$

Με άλλα λόγια, αν καταγράψουμε τις τιμές X_1, X_2, X_3, ..., X_n, ... μιας ΤΜ X, τότε ο εμπειρικός μέσος των τιμών αυτών τείνει στην μέση τιμή με πιθανότητα 1.

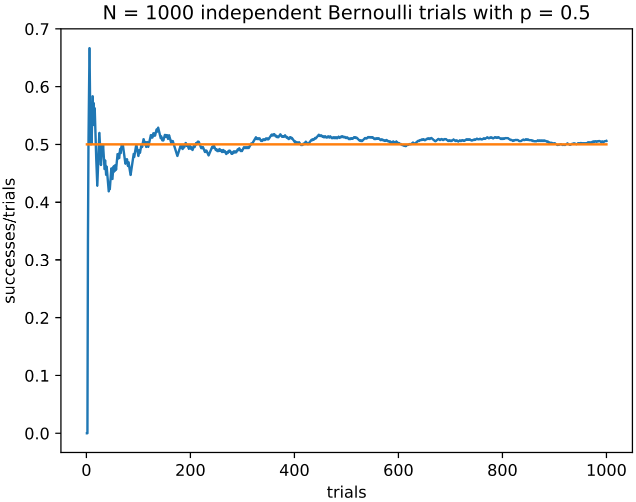
Ο νόμος των μεγάλων αριθμών

```
import scipy.stats as st
import matplotlib.pyplot as plt

p, N = 0.5, 1000
rv = st.bernoulli(p)
x = rv.rvs(N) # a random sample of size N
f = len(x)*[0]; s = 0
for i in range(len(x)): #calculate success frequencies
    if x[i]==1: s+=1;
    f[i] = s/(i+1)

plt.plot(range(1,len(x)+1),f)
plt.plot(range(1,len(x)+1),len(x)*[0.5])
plt.xlabel("trials")
plt.ylabel("successes/trials")
plt.title("N = %s independent Bernoulli trials with p = %s"
    %(N, p))
plt.show()
```

Ο νόμος των μεγάλων αριθμών



Σχήμα: Ο λόγος επιτυχίες δια επαναλήψεις συγκλίνει στο $p = 1/2$ όταν το N γίνεται μεγάλο.

Ο νόμος των μεγάλων αριθμών

Παράδειγμα

Να εκτιμηθεί ο αριθμός των ψαριών μιας λίμνης.

Λύση

Έστω n ο αριθμός των ψαριών της λίμνης. Ψαρεύουμε k ψάρια, τα μαρκάρουμε με κάποιο τρόπο και τα ρίχνουμε πάλι στη λίμνη.

Η πιθανότητα να ψαρέψουμε ένα μαρκαρισμένο ψάρι είναι $p = \dfrac{k}{n}$.

Ψαρεύουμε ξανά λ ψάρια. Μέσα σε αυτά θα υπάρχουν μ μαρκαρισμένα.

Από τον νόμο των μεγάλων αριθμών έχουμε ότι

$$\frac{\mu}{\lambda} \approx \frac{k}{n} \Leftrightarrow n \approx \frac{k\lambda}{\mu}$$

Για παράδειγμα, αν $k = 1000$, $\lambda = 200$, $\mu = 17$ τότε
$n \approx \dfrac{1000 \cdot 200}{17} = 11764.7$.

Άσκηση

Αν παίζεις τάβλι με έναν ισοδύναμο αντίπαλο, ποιο από τα ενδεχόμενα
A: Να κερδίσεις τρεις από τις τέσσερις παρτίδες,
B: Να κερδίσεις έξι από τις οκτώ παρτίδες,
έχει μεγαλύτερη πιθανότητα;
Παρατηρήστε ότι $\frac{3}{4} = \frac{6}{8}$, δηλαδή η αναλογία των παιχνιδιών είναι η ίδια. Πώς ερμηνεύετε το αποτέλεσμα που βρήκατε;

Λύση

Ο αριθμός X των παρτίδων που κερδίζουμε ακολουθεί την διωνυμική κατανομή:
Στην περίπτωση A, με παραμέτρους $N = 4$ και $p = 1/2$ και
στην περίπτωση B, με παραμέτρους $N = 8$ και $p = 1/2$.

Λύση (συνέχεια)

Επομένως,

$$P(A) = P(X = 3|N = 4) = \binom{4}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^1 = 0.2,$$

$$P(B) = P(X = 6|N = 8) = \binom{8}{6}\left(\frac{1}{2}\right)^6\left(\frac{1}{2}\right)^2 = 0.11.$$

Παρόλο που τα δύο ενδεχόμενα φαίνονται ίδια: 3 στα 4 στο ένα και 6 στα 8 στο άλλο, παρατηρούμε ότι $P(A) > P(B)$. Η εξήγηση είναι ότι όσο περισσότερα παιχνίδια παίζουν οι παίχτες, τόσο περισσότερο η X θα είναι κοντά στην μέση τιμή της (νόμος των μεγάλων αριθμών), δηλαδή στο $E(X) = Np = N/2$ και αποκλίσεις της τάξης του $75\% - 25\%$ θα είναι αρκετά σπάνιες.

Λυμένες ασκήσεις

Άσκηση

Έστω (απλό μη κατευθυνόμενο) γράφημα $G = (V, E)$ με n κορυφές και m ακμές. Αν $\frac{n}{2} \leq m \leq \frac{n^2}{4}$, τότε το G περιέχει ένα ανεξάρτητο σύνολο με τουλάχιστον $\frac{n^2}{4m}$ κορυφές.

Λύση

Θέτουμε $V = \{v_1, v_2, \ldots, v_n\}$ και $E = \{e_1, e_2, \ldots, e_m\}$. Για κάθε $S \subseteq V$, συμβολίζουμε με $G(S) = (S, E(S))$ το υπογράφημα του G με σύνολο κορυφών το S και σύνολο ακμών το $E(S)$, όπου κάθε ακμή $e \in E$ ανήκει στο $E(S)$ ανν τα δύο άκρα της ανήκουν στο S. Αρκεί να δειχθεί ότι υπάρχει σύνολο $S \subseteq V$, ώστε το $G(S)$ να έχει τουλάχιστον $\frac{n^2}{4m}$ περισσότερες κορυφές από ακμές, αφού αν διαγράψουμε μια κορυφή από κάθε ακμή του, θα προκύψει ένα υπογράφημα με τουλάχιστον $\frac{n^2}{4m}$ κορυφές και χωρίς ακμές, δηλαδή το ζητούμενο ανεξάρτητο σύνολο.

Λύση (συνέχεια)

Επιλέγουμε τυχαία ένα σύνολο S ως εξής: Κάθε κορυφή $v \in V$ επιλέγεται να ανήκει στο S με πιθανότητα p (η τιμή της θα επιλεχθεί κατάλληλα στη συνέχεια). Το $G(S)$ που προκύπτει είναι τυχαίο από τον τρόπο κατασκευής του, οπότε θεωρούμε τις τυχαίες μεταβλητές

$$X_i \sim \mathrm{Bernoulli}(p), \text{ με } X_i = 1 \Leftrightarrow v_i \in S, \qquad i \in [n],$$

και

$$Y_j \sim \mathrm{Bernoulli}(p^2), \text{ με } Y_j = 1 \Leftrightarrow e_j \in E(S), \qquad j \in [m].$$

Το πλήθος κορυφών και το πλήθος ακμών του $G(S)$ είναι αντίστοιχα $X = X_1 + X_2 + \cdots + X_n$ και $Y = Y_1 + Y_2 + \cdots + Y_m$, οπότε

$$E(X) = E(X_1 + \cdots + X_n) = np \qquad \text{και} \qquad E(Y) = E(Y_1 + \cdots + Y_m) = mp^2.$$

Λύση (συνέχεια)

Θεωρώντας την ΤΜ $Z = X - Y$ (πλήθος κορυφών μείον πλήθος ακμών στο $G(S)$), έχουμε ότι

$$E(Z) = np - mp^2.$$

Η συνάρτηση

$$f(p) = np - mp^2, \qquad p \in (0, 1],$$

παίρνει μέγιστη τιμή όταν $p = \dfrac{n}{2m}$, δηλαδή

$$E(Z) \leq f(\frac{n}{2m}) = \frac{n^2}{2m} - \frac{n^2}{4m} = \frac{n^2}{4m} = c$$

οπότε $P(Z \leq c) > 0$ (βλ. Πρόταση 10), δηλαδή υπάρχει S ώστε το $G(S)$ να έχει τουλάχιστον c περισσότερες κορυφές από ότι ακμές, όταν επιλέξουμε $p = \dfrac{n}{2m}$.