

Μάθημα 25 – 30.05.2023

ΓΡΑΦΕΙΣ

Γιάννης Βασιλόπουλος

Σωτηρία Χριστοδουλοπούλου

Πιθανότητες 1, Οριακά Θεωρήματα

30 Μαΐου 2023

Πώς συμπεριφέρται ένα άθροισμα πολλών (ανεξάρτητων) μεταβλητών

- Τιμές μετοχών στο χρηματιστήριο
- Κλίμα/Μετεωρολογία
- Κατανομή ταχυτήτων σε ένα αέριο
Έστω μια ωκολουσθία X_1, X_2, \dots, X_n από τυχαίες μεταβλητές ανεξάρτητες και ισόνομες με μέση τιμή μ και διασπορά σ^2 .

Πώς συμπεριφέρεται η $S_n = \sum_{i=1}^n X_i$ ως συνάρτηση του n

Έχουμε δει ότι: $E[S_n] = n\mu$ και $Var(S_n) = n\sigma^2$ (από ανεξαρτησία μεταβλητών).

- Κανονικοποιημένο άθροισμα: $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$ Η Z_n «κινείται» γύρω από τη μέση τιμή της $E[Z_n] = 0$ με διασπορά $Var(Z_n) = 1$.

Έχουμε ήδη λίγη πληροφορία στατιστικής φύσεων για τη Z_n , αλλά δεν αρκεί!
Αυτό που θα θέλαμε είναι η κατανομή $P(Z_n \geq z)$ ή την $f_n(z) \equiv f_{Z_n}(z)$.

- Μας ενδιαφέρουν οι ουρές $P(|Z_n| \geq z)$, δηλαδή η πιθανότητα να έχουμε απόκλιση πάνω από z από την κανονικοποιημένη μέση τιμή.
- Μας ενδιαφέρει η συμπεριφορά αυτών των ουρών για $n \rightarrow \infty$.

Πώς εξάγουμε πληροφορία για τις ουρές μιας μεταβλητής από τα στατιστικά της

- Ανισότητα Markov:

$$\text{Έστω τυχαία μεταβλητή } X > 0. \text{ Τότε: } P(X \geq a) \leq \frac{E[X]}{a} \quad \forall a > 0.$$

- Απόδειξη (για τη συνεχή περίπτωση):

$$E[X] = \int_0^\infty xf(x)dx = \int_0^a xf(x)dx + \int_a^\infty xf(x)dx \geq \int_a^\infty xf(x)dx \geq a \int_a^\infty af(x)dx = aP(X \geq a) \Rightarrow P(X \geq a) \leq \frac{E[X]}{a}$$

- Από Markov: $P(|Z_n| \geq z) \leq \frac{E[Z_n]}{z} \left(\leq \frac{\sqrt{E[Z_n^2]}}{z} = \frac{1}{z} \right).$
Το φράγμα $O(\frac{1}{z})$ δεν είναι αρκετά ακριβές.

- Ανισότητα Chebychev: Αν η τυχαία μεταβλητή X έχει $E[X] < \infty$ και

$$E[x^2] < \infty, \text{ τότε, } \forall a > 0, \text{ έχουμε: } P(|X - \mu| \geq a) \leq \frac{\sigma^2}{a^2}, \text{ όπου } \mu = E[X] \text{ και } \sigma^2 = Var(X).$$

- Απόδειξη:

Έστω $U = (X - \mu)^2$, οπότε:

$$P(|X - \mu| \geq a) = P((X - \mu)^2 \geq a^2) = P(U \geq a^2) \leq \frac{E[U]}{a^2} = \frac{E[(X - \mu)^2]}{a^2} = \frac{Var(X)}{a^2} = \frac{\sigma^2}{a^2}$$

- Από ανισότητα Chebychev στην Z_n , προκύπτει ότι $P(|Z_n| \geq z) \leq \frac{1}{z^2}$.

- Αφού $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$, θα έχουμε:

$$P(|Z_n| \geq z) = P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \geq z\right) = P(|S_n - n\mu| \geq z\sqrt{n}\sigma) \leq \frac{1}{z^2}.$$

$$\text{Θέτοντας } a = z\sqrt{n}\sigma, \text{ παίρνουμε: } P(|S_n - n\mu| \geq a) \leq \frac{n\sigma^2}{a^2}.$$

Μελέτη του δειγματικού μέσου όρου

$$\bar{X}_n = \frac{S_n}{n} = \frac{X_1 + \dots + X_n}{n}$$

Εκτίμηση του μέσου όρου μ μιας τυχαίας μεταβλητής X από μια δειγματοληψία μεγέθους n :

$$P(|\bar{X}_n - \mu| \geq a) = P\left(\left|\frac{S_n}{n} - \mu\right| \geq a\right) = P(|S_n - n\mu| \geq an) \leq \frac{n\sigma^2}{n^2a^2} = \frac{\sigma^2}{na^2}$$

Δείξαμε, λοιπόν, ότι: $P(|S_n - n\mu| \geq an) \frac{\sigma^2}{na^2}$

Ασθενής Νόμος Μεγάλων Αριθμών (*Bernoulli, Khinchin*)

Έστω X_1, \dots, X_n ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με $E[X^2] < \infty$.

Τότε: $\forall \epsilon > 0, P(|\bar{X} - \mu| \geq \epsilon) \rightarrow 0, n \rightarrow \infty$.

Λέμε ότι η \bar{X}_n συγκλίνει στη μ κατά πιθανότητα και γράφουμε ότι $\bar{X}_n \xrightarrow{P} \mu, n \rightarrow \infty$.
(Καθώς $n \rightarrow \infty$, η τυχαιότητα «εξαφανίζεται».)

Ισχυρός Νόμος Μεγάλων Αριθμών (*Borel, Kolmogorov*)

Έστω X_1, \dots, X_n ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με μέση τιμή μ και διασπορά σ^2 .

Τότε: $P(\bar{X}_n \rightarrow \mu, \text{ καθώς } n \rightarrow \infty) = 1$.

Λέμε ότι η \bar{X}_n συγκλίνει στη μ σχεδόν βεβαίως.

Σημειώσεις:

1. Ισχυρός N.M.A. \Rightarrow Ασθενή N.M.A.
(Σχεδόν βέβαιη σύγκλιση \Rightarrow Σύγκλιση κατά πιθανότητα)
2. Ωστόσο, υπάρχουν περιπτώσεις (χαλαρότερες υποθέσεις) στις οποίες ισχύει ο Ασθενής N.M.A., αλλά όχι ο Ισχυρός N.M.A..

Πώς συμπεριφέρεται, τελικά, η $P(|Z_n| \geq z)$ για μεγάλο n

Δύο εφαρμογές:

1. Τυχαίοι περίπατοι
2. Πλήθος των πρώτων διαιρετών
Διαλέγουμε έναν τυχαίο αριθμό X ομοιόμορφα στην τύχη από το $\{1, \dots, 10^{12}\}$.
Έστω D ο αριθμός των πρώτων διαιρετών του X .
π.χ.: Αν $X = 30$, τότε $D_X = 3$.

Και στις δύο περιπτώσεις, η εμπειρική κατανομή είναι κανονική! (καθώς $n \rightarrow \infty$)
Θέλουμε να εκτιμήσουμε την $P(|Z_n| \geq z)$.

- Γνωρίζουμε ότι $E[Z_n] = 0$ και $E[Z_n^2] = 1$.
- Ιδανικά, όταν θέλαμε να γνωρίζουμε όλες τις ροπές $E[Z_n^r]$ της Z_n .
(ή όσο περισσότερες γίνεται)
- Εναλλακτικά, όταν θέλουμε να υπολογίζουμε τη ροπογεννήτρια συνάρτηση $M_n(t) := E[\exp(tZ_n)]$.
- Εξ ορισμού, έχουμε:

$$M_n(t) = E[\exp(tZ_n)] = E\left[\exp\left(t\frac{S_n - n\mu}{\sigma\sqrt{n}}\right)\right] = E\left[\exp\left(t\frac{X_1 + \dots + X_n}{\sigma\sqrt{n}}\right)\right] \exp\left(-\frac{\mu\sqrt{n}t}{\sigma}\right).$$
- Για απλότητα, όταν συνεχίσουμε με την περίπτωση $\mu = 0$ και $\sigma = 1$.

$$\begin{aligned} M_n(T) &= E\left[\exp\left(t\frac{X_1 + \dots + X_n}{\sigma\sqrt{n}}\right)\right] = E\left[\exp\left(t\frac{X_1}{\sqrt{n}}\right) \cdot \dots \cdot \exp\left(t\frac{X_n}{\sqrt{n}}\right)\right] = \\ &= E\left[\exp\left(t\frac{X_1}{\sqrt{n}}\right)\right] \cdot \dots \cdot E\left[\exp\left(t\frac{X_n}{\sqrt{n}}\right)\right] = E\left[\exp\left(t\frac{X}{\sqrt{n}}\right)\right]^n = \left[M_X\left(\frac{t}{\sqrt{n}}\right)\right]^n \end{aligned}$$

 για X μια τυχαία μεταβλητή ισόνομη με X_1, \dots, X_n .

Εξ ορισμού:

$$\begin{aligned} M_X(u) &= E[\exp(uX)] = E[1] + uE[X] + \frac{u^2}{2}E[X^2] + O(u^3) = \\ &= 1 + \frac{u^2}{2} + O(u^3), E[X] = 0, E[X^2] = 1 \Rightarrow \\ M_n(t) &= \left[1 + \frac{t^2}{2n} + O\left(\frac{t^3}{n^{3/2}}\right)\right]^n \rightarrow e^{\frac{t^2}{2}} \end{aligned}$$

- Δείξαμε ότι οι ροπογεννήτριες συναρτήσεις $M_n(t)$ συγκλίνουν
 $\forall t \in \mathbb{R}, M_n(t) \rightarrow e^{\frac{t^2}{2}}$ καθώς $n \rightarrow \infty$,
 δηλαδή στη ροπογεννήτρια συνάρτηση της $f(x) = \frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}$.
- Λήμμα *Levy*: Αν η ροπογεννήτρια συνάρτηση $M_n(t)$ της Z_n συγκλίνει στη ροπογεννήτρια συνάρτηση $M(t)$ της Z , τότε:
 $F_{Z_n}(z) \rightarrow F_Z(z)$ καθώς $n \rightarrow \infty$ σε κάθε σημείο συνέχειας της F_Z .

Κεντρικό Οριακό Θεώρημα

Έστω ότι X_1, \dots, X_n ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με μέσο μ και διασπορά σ^2 . Τότε, η κατανομή του κανονικοποιημένου ανθροίσματος $Z_n = \frac{X_1 + \dots + X_n}{\sigma\sqrt{n}}^*$ τείνει καθώς $n \rightarrow \infty$ στην τυπική κανονική κατανομή $f_Z(z) \rightarrow \frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}$ καθώς $n \rightarrow \infty, \forall z \in \mathbb{R}$. ** -ημ*

30/5/23

Ορικά Θεωρήματα [κεφ. 8 από Ross]

Πώς συμπληρώνεται η αίσθοση πολλών (aveγαρίνιων) ωχαιών μεταβλητών;

- π.χ.
- Επιπρόσυνη → τύπος μετωπών σε κριτικότητα
 - τάχιση / Μετεωρολογία
 - κατανομής σαχατήτων σε ένα αεροίο

» Εάντομο μία αριθμούδια X_1, X_2, \dots, X_n από ωχαιες

μεταβλητές που είναι

- i) ανεγαρίνιτες } δεν πας είναι αναράτητα ποντών
- ii) ισορρόπες } αλλι τα υποστηρίζουν ποντών

Η ίδια την μέση μ και σιασηρά σ^2

» Πώς συμπληρώνεται η $S_n = \sum_{i=1}^n X_i$ ως συνάρτηση των;

Έχουμε δει ότι $E[S_n] = n\mu$ $\text{Var}(S_n) = n\sigma^2$

(δεν χρειάζομαι ανεγαρίνια) \quad (χρειάζομαι ανεγαρίνια)

Κανονικοποίηση αριθμητών:

→ Κανονικοποιηθέν αίσθοση

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

Η Z_n "κυριεύει" πάνω από την μέση την $E[Z_n] = 0$

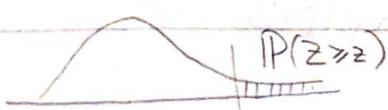
Η σιασηρά $\text{Var}(Z_n) = 1$

Έχουμε την αριθμητή πιθανότητα για την Z_n να είναι μεταξύ των z_1 και z_2

Αυτό που θα δείξουμε είναι η κανονική

$$\underbrace{\Pr(Z_n \geq z)}_{\text{αριθμητή}} \in \text{την } F_n(z) \equiv F_{Z_n}(z)$$

αριθμητή



»» Μας ενδιαφέρουν οι ουπες $P(|Z_n| \geq z)$ στησι
η οιδαντίτικα να εχουμε ανάλογη θέση ανo z
ανo την καν/νη μέση την
Μας ενδιαφέρει η συμπεριφορά αυτών των ουπών
για n → ∞

»» Έως έργουμε οληροφορία για τις ουπες πιας μεταβλητής
ανo τα γεωγεύματα της?

Anigōniza Markov: Είνω r.p. $X \geq 0$. Τότε

$$P(X > a) \leq \frac{E[X]}{a} \quad \text{Τότε}$$

Είναι ανισότητα συρπίζεις
ανοσήγην (πιο ζην γενεχρις νερπίζων)

$$E[X] = \int_0^\infty x f(x) dx = \int_0^a x f(x) dx + \int_a^\infty x f(x) dx \geq$$

$\xrightarrow{\text{οχι } -\infty}$

$\xrightarrow{\text{αφού } X > 0}$

≥ 0 απo ρηγήw

$$\geq \int_a^\infty x f(x) dx \geq \int_a^\infty a f(x) dx = a \int_a^\infty f(x) dx = a P(X > a)$$

$$\Rightarrow P(X > a) \leq \frac{E[X]}{a}$$

»» Στην $P(|Z_n| \geq z)$: Ανo Markov: $P(|Z_n| \geq z) \leq \frac{E[|Z_n|]}{z}$

$$\left(\leq \frac{\sqrt{E[Z_n^2]}}{z} = \frac{1}{z} \right) \Rightarrow \text{Το υραγμα } O(1/z)$$

Ανo Jensen

δεν είναι αρκετά
ακριβής

Anigortza Chebyshov: Avn t.u. X exi

$$\cdot \mathbb{E}[X] < \infty$$

$$\cdot \mathbb{E}[X^2] < \infty$$

Tote $\forall a > 0$ exoupe:

$$\text{ono } \mu = \mathbb{E}[X], \sigma^2 = \text{Var}(X)$$

$$\boxed{\mathbb{P}(|X-\mu| > a) \leq \frac{\sigma^2}{a^2}}$$

anoðsyn: Egw $U = (X-\mu)^2$ onore

$$\mathbb{P}(|X-\mu| > a) = \mathbb{P}((X-\mu)^2 > a^2) = \mathbb{P}(U > a^2)$$

Martov

$$\mathbb{P}(U > a^2) \stackrel{Martov}{\leq} \frac{\mathbb{E}[U]}{a^2} = \mathbb{E}[(X-\mu)^2] = \frac{\text{Var}(X)}{a^2} = \frac{\sigma^2}{a^2}$$

»» Tm $\mathbb{P}(|Z_n| > z)$: Ano Chebyshov:

$$\boxed{\mathbb{P}(|Z_n| > z) \leq \frac{1}{z^2}}$$

Aþou $Z_n = \frac{S_n - n\mu}{\sqrt{n}\sigma}$ da exoupe

$$\mathbb{P}(|Z_n| > z) = \mathbb{P}\left(\frac{|S_n - n\mu|}{\sqrt{n}\sigma} > z\right) = \mathbb{P}(|S_n - n\mu| > z\sqrt{n}\sigma)$$

Ano Chebyshov: $\mathbb{P}(|S_n - n\mu| > z\sqrt{n}\sigma) \leq \frac{1}{z^2}$

$$\Rightarrow \boxed{\mathbb{P}(|S_n - n\mu| > a) \leq \frac{n\sigma^2}{a^2}}$$

»» Medein wu seifharakoi pegas ou $\bar{X}_n = \frac{S_n}{n} = \frac{X_1 + \dots + X_n}{n}$
 (exipnon tou pegas b. tias ar.t.u. X)
 ano mia seifharakotia pegasous n
 $\Rightarrow \mathbb{P}(|\bar{X}_n - \mu| > a) = \mathbb{P}\left(|\frac{S_n}{n} - \mu| > a\right) \leq$

$$= \boxed{\mathbb{P}(|S_n - n\mu| > an) \leq \frac{n\sigma^2}{n^2 a^2} = \frac{\sigma^2}{na^2}}$$

Definisioun ou:

$$P(|\bar{X}_n - \mu| > \alpha) \leq \frac{\sigma^2}{n\alpha^2}$$

Alephis vofos Megadwv Apidmuwv (Bernoulli Khinchin)

Egitw X_1, \dots, X_n akoloudia aveg. leovotuvv p.t. $\mu \in \mathbb{E}[X_i] < \infty$

Tote $\forall \epsilon > 0$ $P(|\bar{X}_n - \mu| > \epsilon) \rightarrow 0$ kaws $n \rightarrow \infty$

Nefte ou n \bar{X}_n gudkivei gnv μ kawta nisavortia
kai gudkivei ou $\bar{X}_n \xrightarrow{P} \mu$ kaws $n \rightarrow \infty$

Kaws $n \rightarrow \infty$ n twnxiotria "Egavavijetai"

Igxupos vofos Megadwv Apidmuwv (Borel Kolmogorov)

Egitw X_1, \dots, X_n akoloudia leovotuvv kai aveg. p.t.
t.p. μ metan twn μ kaws Siagnopai 6.

Tote $P(\bar{X}_n \rightarrow \mu \text{ kaws } n \rightarrow \infty) = 1$

Nefte ou n \bar{X}_n gudkivei gnv μ ~~ges~~ gesuv Bebara

Inferwes

1) Igxupos NMA \Rightarrow Alephis NMA

(gesuv Bebara gudkivei \Rightarrow gudkivei kawta nisavortia)

2) Ω_{67060} unoihxou nferintwes (xalupozepes unodewes)

Onou igxun o alephis NMA alli oxi o igxupos NMA

➡ Πώς διμηριφερεται τελικά η $P(|Z_n| \geq z)$ στα μεγάλα;

→ Δύο εφαρμογές:

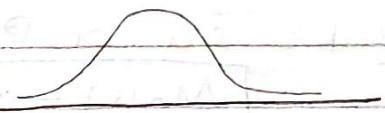
- ωχαιοι περιπτώσει

- ηλίθιος πρώτων σταρέτων

Διαδεχούμε εναν αριθμό αθοιούμερα Z_n στην οποία αντιτίθεται $\{Z_1, \dots, Z_N = 10^{12}\}$. Εγτών θέλουμε να αριθμήσουμε την πρώτη σταρέτη του X .

π.χ. αν $X_0 = 30$, ώστε $D_X = 3$

Kai ους δύο περιπτώσεις ή εμπειρική κατανομή είναι κανονική (καθώς $n \rightarrow \infty$)



➡ Θέλουμε να εξιμογούμε την $P(|Z_n| \geq z)$

- Γνωριζόμε $E[Z_n] = 0$

$$E[Z_n^2] = 1$$

- Ιστορικά θα έλθει να γνωριζόμετε ότις τις πονεί $E[Z_n]$ της Z_n [n ορο ηπειροτερες γίνεται]

- Ειδικάρικα θα έλθει να υπολογίζουμε τη $P[\sum_{n=1}^{\infty} Z_n(t) \geq z]$

Εγι' οριγκούς ξέρουμε: $M_n(t) = E[\exp(tZ_n)] = E[\exp(t \frac{S_n - n\mu}{\sqrt{n}})]$
 $M_n(t) = E[\exp(t \frac{X_1 + \dots + X_n}{\sqrt{n}})] \exp(-\frac{\mu \sqrt{n} t}{6})$

Για αποτίνω θα γνωριζόμετε ότι την περιπτώση $\mu = 0$ $\sigma = 1$

$\Rightarrow M_n(t) = E[\exp(t \frac{X_1 + \dots + X_n}{\sqrt{n}})] = E[\exp(t \frac{X_1}{\sqrt{n}}) \times \dots \times \exp(t \frac{X_n}{\sqrt{n}})]$

$\stackrel{\text{εγτών } X_i \text{ i.i.d.}}{=} E[\exp(t \frac{X_1}{\sqrt{n}})] \times \dots \times E[\exp(t \frac{X_n}{\sqrt{n}})] \stackrel{\text{εγτών } X_i \text{ i.i.d.}}{=} [E[\exp(t \frac{X}{\sqrt{n}})]]^n$

$$M_X(t/\sqrt{n})$$

$$\Rightarrow M_n(t) = [M_X(t/\sqrt{n})]^n \quad (\text{X})$$

$$\begin{aligned} \text{Εγι' ορισμα } M_X(u) &= \mathbb{E}[\exp(uX)] = \mathbb{E}[1 + uX + \frac{u^2}{2} X^2 + O(u^3)] \\ &= \mathbb{E}[1] + u \mathbb{E}[X] + \frac{u^2}{2} \mathbb{E}[X^2] + O(u^3) \\ &= 1 + \frac{u^2}{2} + O(u^3) \end{aligned}$$

$$(\text{X}) \Rightarrow M_n(t) = \left[1 + \frac{t^2}{2n} + O\left(\frac{t^3}{n^{3/2}}\right)\right]^n \rightarrow e^{t^2/2} \quad \lim_{n \rightarrow \infty} \left(1 + \frac{a}{n}\right)^n = e^a$$

Δείγαμε ου αι ΡΓΣ $M_n(t)$ γυρκλινου $\forall t \in \mathbb{R}$

$$M_n(t) \rightarrow e^{t^2/2} \quad \text{καθως } n \rightarrow \infty$$

$$\text{ΡΓΣ της } f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Λιτφα (Lévy):

Αν οι ΡΓΣ $M_n(t)$ της Z_n γυρκλινει στη ΡΓΣ

$M(t)$ της Z , τοτε $F_{Z_n}(z) \xrightarrow{n \rightarrow \infty} F_Z(z)$ σε κάθε σημειο
γυρκλινας της F_Z

Κεντρικο οριακο δεύτρυνα!

Εστω οι X_1, \dots, X_n ανεγαπιντες και ισονομει $\mathbb{E}X_i = \mu$ μερικα
και σιασηραι σ^2 . Τοτε η κανονικη των κανινου
ασποισματος $Z_n = \frac{X_1 + \dots + X_n - n\mu}{\sqrt{n}\sigma}$ τεινε καθως $n \rightarrow \infty$

ετην των κανονικην κανανοτη σημειων

$$f_{Z_n}(z) \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad \forall z \in \mathbb{R}$$