

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ**

**ΗΥ – 217
ΘΕΩΡΙΑ ΠΙΘΑΝΟΤΗΤΩΝ**

**ΣΗΜΕΙΩΣΕΙΣ – ΕΝΟΤΗΤΑ ΤΕΤΑΡΤΗ:
ΕΙΔΙΚΑ ΘΕΜΑΤΑ ΤΥΧΑΙΩΝ ΜΕΤΑΒΛΗΤΩΝ**

ΠΑΝΑΓΙΩΤΗΣ ΤΣΑΚΑΛΙΔΗΣ - 2002

ΕΝΟΤΗΤΑ 4 : ΕΙΔΙΚΑ ΚΕΦΑΛΑΙΑ ΘΕΩΡΙΑΣ ΠΙΘΑΝΟΤΗΤΩΝ (12 L)

Σε αυτή την ενότητα, θα αναπτύξουμε ορισμένες πιο προχωρημένες έννοιες της θεωρίας πιθανοτήτων καθώς και εφαρμογές, όπως μετασχηματισμούς και συνεπίξεις, που θα είναι χρήσιμα σε εφαρμογές όπως :

- (α) Η μελέτη των αθροισμάτων πολλών ανεξαρτήτων τ.φ. των οποίων ο αριθμός μπορεί να είναι επίσης τ.φ.
- (β) Η εκτίμηση ή πρόβλεψη μιας άγνωστης τ.φ. με βάση μετρήσεις τιμών άλλων τ.φ.

4.1 ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΙ

Μετασχηματισμοί των πιθανοτικών νόμων (PMF ή PDF) τ.φ. αποτελούν εναλλακτικούς τρόπους περιγραφής των τ.φ. και είναι χρήσιμοι σε ορισμένες μαθηματικές πράξεις με τ.φ.

Ορισμός : ΡΟΠΟΓΕΝΝΗΤΡΙΑ ΣΥΝΑΡΤΗΣΗ της τ.φ. X :
(Moment generating function)

$$M_X(s) \triangleq E[e^{sX}]$$

• Για διακριτές τ.φ. ο παραπάνω ορισμός δίνει :

$$M(s) = \sum_x e^{sx} \cdot p_X(x)$$

(Πιθανογεννήτρια)

• Για συνεχείς τ.φ. :

$$M(s) = \int_{-\infty}^{+\infty} e^{sx} \cdot f_X(x) dx$$

(ροπογεννήτρια)

* Όσοι από εσάς παρακολουθείτε τα "Εφαρμοσμένα Μαθηματικά για Μηχανικούς" (HY-215), παρατηρείτε την αναλογία των παραπάνω ορισμών με τους μετασχηματισμούς Z και Laplace, για τη διακριτή και συνεχή περίπτωση, αντίστοιχα.

- Με λόγια, οι παραπάνω ορισμοί λένε ότι ο μετασχηματισμός $M(s)$ είναι ένας γραμμικός συνδυασμός ενδεσμών της μορφής e^{sx} με βάση τις αντίστοιχες τιμές της $P_X(x)$ ή $f_X(x)$. (122)
- Παρατηρούμε ότι $M(0) = M(s)/_{s=0} = \sum_x e^{0 \cdot x} \cdot P_X(x) = \sum_x P_X(x) = 1$ (ομοίως για συνεχείς τ.τ.).

Παράδειγμα: Έστω $X \sim \text{Poisson}$ τ.τ. με παράμετρο λ :

$$P_X(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x=0, 1, 2, \dots$$

Η ποσογενήτρια συνάρτηση είναι:

$$M(s) = \sum_{x=0}^{+\infty} e^{sx} \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{+\infty} \frac{(e^s \lambda)^x}{x!} = e^{-\lambda} \cdot e^{e^s \lambda}$$

$$\therefore \boxed{M(s) = e^{\lambda(e^s - 1)}}$$

Παράδειγμα: Έστω $X \sim \text{ενδεσμη}$ τ.τ. με παράμετρο λ .

$$f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

$$M(s) = \lambda \int_0^{+\infty} e^{sx} \cdot e^{-\lambda x} dx = \lambda \left. \frac{e^{(s-\lambda)x}}{s-\lambda} \right|_0^{+\infty} = \frac{\lambda}{\lambda-s}, \quad \text{αν } s < \lambda.$$

$$\therefore \boxed{M(s) = \frac{\lambda}{\lambda-s} \quad \text{για } s < \lambda}$$

Παρατηρούμε ότι η $M(s)$ είναι μια συνάρτηση της μεταβλητής s που ορίζεται για ενεργές τις τιμές της s για τις οποίες ορίζεται η $E[e^{sX}]$, δηλ. $E[e^{sX}] < \infty$

Παράδειγμα: Ποσογενήτρια Γραμμικής Συνάρτησης τ.τ.

$$\underline{Y = aX + b} \leadsto M_Y(s) = E[e^{s(aX+b)}] = e^{sb} E[e^{saX}]$$

$$\therefore \boxed{M_Y(s) = e^{sb} M_X(as)}$$

Παράδειγμα: Ροπογενήτρια της Κανονικής τ.μ.

(123)

Έστω $X \sim N(\mu, \sigma^2)$. $M_X(s) = ?$

Ας θεωρήσουμε πρώτα την τυπική κανονική τ.μ., Y ,

$Y \sim N(0, 1)$ όπου $Y = \frac{X - \mu}{\sigma}$.

Για την Y : $f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$.

$$\begin{aligned} M_Y(s) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \cdot e^{sy} dy \\ &= e^{s^2/2} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2} + sy - \frac{s^2}{2}} dy \\ &= e^{s^2/2} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-(y-s)^2/2} dy \end{aligned}$$

$$\therefore \boxed{M_Y(s) = e^{s^2/2}}$$

Τώρα, $X = \sigma Y + \mu$ και κάνοντας χρήση του προηγούμενου παραδείγματος έχουμε:

$$M_X(s) = e^{s\mu} \cdot M_Y(\sigma s) = e^{\frac{\sigma^2 s^2}{2} + \mu s} \quad \square.$$

4.2.1 Υπολογισμός των Ρομών μιας τ.μ. από την $M(s)$

Παραγωγίζουμε ως προς s την $M(s)$:

$$\begin{aligned} \frac{d}{ds} M(s) &= \frac{d}{ds} \int_{-\infty}^{+\infty} e^{sx} \cdot f_X(x) dx = \int_{-\infty}^{+\infty} \frac{d}{ds} e^{sx} \cdot f_X(x) dx \\ &= \int_{-\infty}^{+\infty} x e^{sx} \cdot f_X(x) dx. \end{aligned}$$

$$\text{Για } s=0, \quad \left. \frac{d}{ds} M(s) \right|_{s=0} = \int_{-\infty}^{+\infty} x f_X(x) dx = E[X]$$

(124)

$$\therefore \boxed{E[X] = \left. \frac{d}{ds} M(s) \right|_{s=0}}$$

Παραγωγίζοντας η φορές ως προς s , προκύπτει:

$$\boxed{E[X^n] = \left. \frac{d^n}{ds^n} M(s) \right|_{s=0}}$$

Παράδειγμα: Έστω X διακριτή τ.φ. με $P_X(x)$:

$$P_X(x) = \begin{cases} 1/2, & x=2 \\ 1/6, & x=3 \\ 1/3, & x=5 \end{cases} \quad E[X] = ?$$

$$M(s) = \sum_x e^{sx} P_X(x) = \frac{1}{2} e^{2s} + \frac{1}{6} e^{3s} + \frac{1}{3} e^{5s}$$

$$E[X] = \left. \frac{dM(s)}{ds} \right|_{s=0} = \frac{1}{2} 2 e^{2s} + \frac{1}{6} 3 e^{3s} + \frac{1}{3} 5 e^{5s} \Big|_{s=0} = \frac{19}{6}$$

$$E[X^2] = \left. \frac{d^2}{ds^2} M(s) \right|_{s=0} = \frac{1}{2} 4 e^{2s} + \frac{1}{6} 9 e^{3s} + \frac{1}{3} 25 e^{5s} \Big|_{s=0} = \frac{71}{6}$$

Παράδειγμα: Ενδεστική τ.φ.

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \rightarrow M(s) = \frac{\lambda}{\lambda - s}, \quad s < \lambda$$

$$\frac{d}{ds} M(s) = \frac{\lambda}{(\lambda - s)^2}, \quad \frac{d^2}{ds^2} M(s) = \frac{2\lambda}{(\lambda - s)^3}$$

$$\text{Για } s=0 \quad E[s] = \frac{1}{\lambda}, \quad E[X^2] = \frac{2}{\lambda^2} \quad \square$$

4.1.2 Αντιστοιχία Ποσογεννητριών

Η Ποσογεννήτρια συνάρτηση $M_X(s)$ ορίζει πλήρως τον πιθανοτυπικό νόμο της τ.φ. X . Δηλαδή, αν $M_X(s) = M_Y(s) \forall s$, τότε οι τ.φ. X και Y έχουν την ίδια συνάρτηση (πυκνότητας) πιθανότητας.

Παράδειγμα. Η τ.μ. X έχει ροιογενήτση :

(125)

$$M(s) = \frac{1}{4} e^{-s} + \frac{1}{2} + \frac{1}{8} e^{4s} + \frac{1}{8} e^{5s}$$

Ποιά η PMF της X ?

- Συγκρίνουμε την $M(s)$ με τη γενική έκφραση:

$$M(s) = \sum_x e^{sx} \cdot P_X(x)$$

Οπότε έχουμε ότι

$$P_X(x) = \begin{cases} 1/4 & x = -1 \\ 1/2 & x = 0 \\ 1/8 & x = 4 \\ 1/8 & x = 5 \end{cases}$$

Παράδειγμα. Η ροιιογενήτση της Γεωμετρικής τ.μ.

Έστω X με $M(s) = \frac{pe^s}{1 - (1-p)e^s}$, όπου $0 < p < 1$.

Ποιά η PMF της X ?

- Έχουμε ότι $\frac{1}{1-a} = 1 + a + a^2 + \dots$ για $|a| < 1$.

Θέτουμε $a = (1-p)e^s$ με s κοντά στο 0 ώστε $a < 1$.

$$\begin{aligned} \text{Έχουμε: } M(s) &= pe^s (1 + (1-p)e^s + (1-p)^2 e^{2s} + \dots) \\ &= pe^s + p(1-p)e^{2s} + p(1-p)^2 e^{3s} + \dots \\ &= \sum_{k=1}^{\infty} p(1-p)^{k-1} e^{ks} \\ &= \sum_x P_X(x) e^{xs} \end{aligned}$$

$$\Rightarrow P_X(k) = p(1-p)^{k-1}, \quad k=1, 2, \dots$$

$\hookrightarrow X$: γεωμετρική τ.μ. με παράμετρο p .

Παράδειγμα: Ροπογεννήτρια Μιγμάτων δύο κατανομών (126)

Μία τράπεζα έχει τρία ταβέρια, δύο γρήγορα και ένα αργό. Ο χρόνος εξυπηρέτησης είναι εκθετικά κατανομημένος με $\lambda=6$ για τα γρήγορα ταβέρια και $\lambda=4$ για το αργό. Η Γιώτα μπαίνει στην τράπεζα και διαλέγει τυχαία ένα ταβέριο με πιθανότητα $1/3$. Ποια η PDF και η ροπογεννήτρια των χρόνων εξυπηρέτησης της Γιώτας?

Έχουμε $f_X(x) = \frac{2}{3} 6 e^{-6x} + \frac{1}{3} 4 e^{-4x}, \quad x \geq 0$

Επίσης:
$$\begin{aligned} M(s) &= \int_0^{+\infty} e^{sx} \left(\frac{2}{3} 6 e^{-6x} + \frac{1}{3} 4 e^{-4x} \right) dx \\ &= \frac{2}{3} \int_0^{+\infty} 6 e^{(s-6)x} dx + \frac{1}{3} \int_0^{+\infty} 4 e^{(s-4)x} dx \\ &= \frac{2}{3} \frac{6}{6-s} + \frac{1}{3} \frac{4}{4-s}, \quad s < 4 \end{aligned}$$

Στη γενική περίπτωση αν

$$f_Y(y) = p_1 \cdot f_{X_1}(y) + \dots + p_n \cdot f_{X_n}(y),$$

τότε

$$M_Y(s) = p_1 M_{X_1}(s) + \dots + p_n M_{X_n}(s).$$

4.4.3 Αθροίσματα Ανεξαρτήτων τ.φ.

Η Ροπογεννήτρια του αθροίσματος ανεξαρτήτων τ.φ. ισούται με το γινόμενο των ροπογεννητριών των τ.φ.

Πράγματι, έστω X και Y ανεξάρτητες τ.φ. και $Z = X + Y$.

Τότε $M_W(s) = E[e^{sW}] = E[e^{s(X+Y)}] = E[e^{sX} e^{sY}]$

Αφού οι X και Y είναι ανεξάρτητες, ομοίως είναι οι e^{sX} ή e^{sY} .

Επομένως, $E[e^{sX} \cdot e^{sY}] = E[e^{sX}] \cdot E[e^{sY}] = M_X(s) \cdot M_Y(s)$

Γενικεύοντας, αν $X_1, \dots, X_n$ είναι μια συλλογή ανεξαρτήτων τ.φ. και

$$W = X_1 + \dots + X_n,$$

τότε

$$M_W(s) = M_{X_1}(s) \dots M_{X_n}(s).$$

(127)

Παράδειγμα: Η Ρονογεωμετρία της Διωνυμικής

Έστω $X_1, \dots, X_n$ ανεξάρτητες τ.φ. Bernoulli με κοινή παράμετρο p . Τότε

$$M_{X_i}(s) = (1-p)e^{0 \cdot s} + pe^{1 \cdot s} = 1-p+pe^s, \quad \forall i$$

Η τ.φ. $Y = X_1 + \dots + X_n$ είναι διωνυμική με παράμετρο n και p . Η ρονογεωμετρία της Y δίνεται από

$$M_Y(s) = (1-p+pe^s)^n.$$

Παράδειγμα: Το άθροισμα ανεξαρτήτων τ.φ. Poisson είναι Poisson

Έστω X και Y , ανεξάρτητες Poisson τ.φ. με μέσες τιμές λ και μ , αντίστοιχα. Έστω $W = X + Y$. Τότε

$$M_X(s) = e^{\lambda(e^s-1)}, \quad M_Y(s) = e^{\mu(e^s-1)}$$

$$\text{και } M_W(s) = M_X(s) \cdot M_Y(s) = e^{(\lambda+\mu)(e^s-1)}$$

$\Rightarrow W$ είναι Poisson με παράμετρο $\lambda+\mu$.

Παράδειγμα: Το Άθροισμα Ανεξαρτήτων Κανονικών τ.φ. είναι Κανονική

$X \sim N(\mu_X, \sigma_X^2)$. $Y \sim N(\mu_Y, \sigma_Y^2)$. Έστω $W = X + Y$. Τότε:

$$\left. \begin{aligned} M_X(s) &= e^{\frac{\sigma_X^2 s^2}{2} + \mu_X s} \\ M_Y(s) &= e^{\frac{\sigma_Y^2 s^2}{2} + \mu_Y s} \end{aligned} \right\} M_W(s) = e^{\frac{\sigma_X^2 + \sigma_Y^2}{2} s^2 + (\mu_X + \mu_Y) s}$$

$$\Rightarrow W \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2).$$

4.1.4 Πολλαπλασιασμός Από κοινού Κατανομών

(128)

Έστω οι τ.μ. $X_1, \dots, X_n$ σχετικές με κάποιο πείραμα.

Τότε

$$M_{X_1, \dots, X_n}(s_1, \dots, s_n) \triangleq E[e^{s_1 X_1 + \dots + s_n X_n}]$$

4.2 ΑΘΡΟΙΣΜΑ ΑΝΕΞΑΡΤΗΤΩΝ Τ.Μ. : ΣΥΝΕΛΙΞΕΙΣ

Έστω X, Y ανεξάρτητες τ.μ. Όπως είδαμε στην 4.1.3, η κατανομή του αθροίσματος $W = X + Y$ μπορεί να υπολογιστεί μέσω της αντικατάστασης της πολλαπλασιαστικής της W :

$$M_W(s) = M_X(s) \cdot M_Y(s).$$

Αλλά μπορεί να υπολογιστεί και απ'ευθείας μέσω της παράστασης της συνελίξης.

4.2.1 Διακριτές Τ.Μ.

Έστω X, Y διακριτές ανεξάρτητες τ.μ. που παίρνουν τιμές στο σύνολο των ακεραίων. Τότε $W = X + Y$,

$$\begin{aligned} P_W(w) &= P(X + Y = w) = \sum_{(x,y): x+y=w} P(X=x \text{ και } Y=y) \\ &= \sum_x P(X=x \text{ και } Y=w-x) \end{aligned}$$

$$P_W(w) = \sum_x P_X(x) \cdot P_Y(w-x)$$

Συνελίξη των $P_X(k) \triangleright P_Y(k)$:

$$P_X(k) * P_Y(k)$$

4.2.2 Συνεχείς Τ.Μ.

Έστω X, Y συνεχείς ανεξάρτητες τ.μ. $W = X + Y$.

Για την CDF της W :

(129)

$$F_W(w) = P(W \leq w) = P(X + Y \leq w)$$

$$= \int_{x=-\infty}^{+\infty} \int_{y=-\infty}^{w-x} f_X(x) \cdot f_Y(y) dy dx$$

$$= \int_{x=-\infty}^{+\infty} f_X(x) \cdot \left[\int_{y=-\infty}^{w-x} f_Y(y) dy \right] \cdot dx$$

$$= \int_{x=-\infty}^{+\infty} f_X(x) \cdot F_Y(w-x) dx$$

H PDF τῆς W:

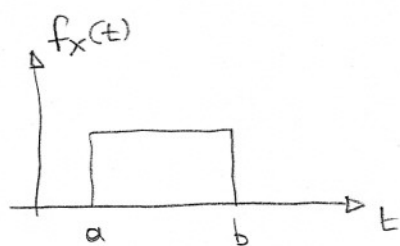
$$f_W(w) = \frac{dF_W(w)}{dw} = \int_{x=-\infty}^{+\infty} f_X(x) \cdot \frac{dF_Y(w-x)}{dw} dx$$

$$\therefore f_W(w) = \int_{x=-\infty}^{+\infty} f_X(x) \cdot f_Y(w-x) dx$$

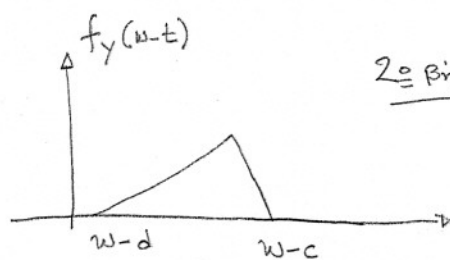
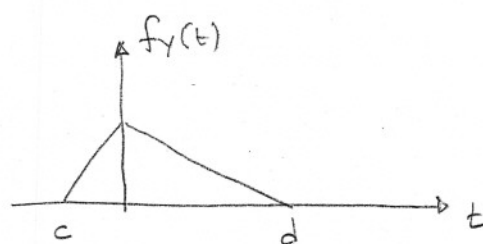
Συνέλιξη:

$$f_X(t) * f_Y(t)$$

Γραφικός Υπολογισμός Συνελίξεων

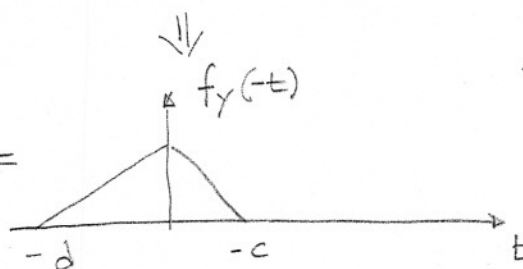


*

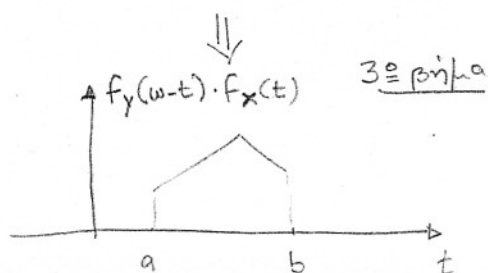


2^ο βήμα

⇐



1^ο βήμα



3^ο βήμα

4.4

ΑΘΡΟΙΣΜΑ ΕΝΟΣ ΤΥΧΑΙΟΥ ΑΡΙΘΜΟΥ ΑΝΕΞΑΡΤΗΤΩΝ Τ.Μ.

(135)

Θεωρούμε το άθροισμα:

$$Y = X_1 + \dots + X_N$$

όπου N : τυχαία μεταβλητή που παίρνει μη αρνητικές ακέραιες τιμές και $X_1, X_2, \dots, X_N$ τ.μ. με την ίδια κατανομή.

Υποθέτουμε ότι οι $N, X_1, \dots, X_N$ είναι ανεξάρτητες τ.μ., που σημαίνει ότι οποιοδήποτε ερωτησάκι αυτών είναι μια συλλογή ανεξάρτητων τ.μ.

Έστω μ και σ^2 η κοινή μέση τιμή και διασπορά των X_i .
Θέλουμε να υπολογίσουμε τις: $E[Y]$, $\text{var}(Y)$, και $M_Y(s)$.

Η μέθοδος που ακολουθούμε είναι να δουλέψουμε υπό συνθήκη του γεγονότος $N=n$, οπότε έχουμε να κάνουμε με το άθροισμα ενός καθορισμένου αριθμού ανεξάρτητων τ.μ.

$$\begin{aligned} E[Y/N=n] &= E[X_1 + \dots + X_N / N=n] \\ &= E[X_1 + \dots + X_n / N=n] \\ &= E[X_1 + \dots + X_n] = n \cdot \mu \end{aligned}$$

$$\therefore E[Y/N] = N \cdot \mu$$

Με χρήση του νόμου των διαδοχικών μέσων τιμών

$$E[Y] = E[E[Y/N]] = E[N \cdot \mu] = \mu E[N].$$

$$\therefore \boxed{E[Y] = \mu E[N]}$$

$$\begin{aligned} \text{Ομοίως, } \text{var}(Y/N=n) &= \text{var}(X_1 + \dots + X_N / N=n) \\ &= \text{var}(X_1 + \dots + X_n) = n \sigma^2 \end{aligned}$$

$$\therefore \text{var}(Y/N) = N \cdot \sigma^2$$

Με χρήση των νόμων των διαδοχικών διασπορών

$$\begin{aligned}\text{var}(Y) &= E[\text{var}(Y/N)] + \text{var}(E[Y/N]) \\ &= E[N] \cdot \sigma^2 + \text{var}(N \cdot \mu)\end{aligned}\quad (136)$$

$$\therefore \boxed{\text{var}(Y) = E[N] \cdot \sigma^2 + \mu^2 \text{var}(N)}$$

$$\begin{aligned}\text{Τέλος, } E[e^{sY/N=n}] &= E[e^{sX_1} \dots e^{sX_n}] \\ &= E[e^{sX_1}] \dots E[e^{sX_n}] = (M_X(s))^n\end{aligned}$$

$$\text{Επομένως } E[e^{sY}] = E[E[e^{sY/N}]] = E[(M_X(s))^N]$$

$$\therefore \boxed{M_Y(s) = \sum_{n=0}^{+\infty} (M_X(s))^n \cdot P_N(n)}$$

Παράδειγμα: Το άθροισμα ενός γεωμετρικά κατανοημένου αθροφού
Ενδετικών τ.τ.

Η Μαρία ψάχνει τα βιβλιοπωλεία ψάχνοντας τα "Εκατό χρόνια Μονασίας" του Marquez. Ένα τυχαίο βιβλιοπωλείο έχει το βιβλίο αυτό με πιθανότητα p . Η Μαρία μένει σε κάθε βιβλιοπωλείο, όπου επισκέπτεται χρόνο X_i μέχρις ότου είτε να βρει το βιβλίο είτε να αποφασίσει ότι δεν υπάρχει και να φύγει. Οι χρόνοι X_i είναι ανεξάρτητα κατανοημένα με παράμετρο λ , ανεξάρτητα μεταξύ τους. Η Μαρία συνεχίζει μέχρι να βρει το βιβλίο. Έστω Y ο συνολικός χρόνος που ψάχνει έως ότου το βρει. Ποια είναι τα ζητούμενα: $E[Y]$, $\text{var}(Y)$, $f_Y(y)$?

- Αν N είναι ο αριθμός των βιβλιοπωλείων που επισκέπτεται η Μαρία μέχρι να βρει το βιβλίο, προφανώς αυτό ακολουθεί γεωμετρική κατανομή ($N-1$ αποτυχίες $\cdot (1-p)$ και N -στή επιτυχία, p):

$$P_N(n) = (1-p)^{n-1} \cdot p, \quad n=1, 2, \dots \quad E[N] = \frac{1}{p}, \quad \text{var}[N] = \frac{1-p}{p^2}.$$

Αρα $Y = X_1 + \dots + X_N$ έχουμε ότι: (137)

$$E[Y] = E[N] \cdot E[X] = \frac{1}{p} \cdot \frac{1}{\lambda} = \frac{1}{\lambda p}.$$

$$\begin{aligned} \text{var}(Y) &= E[N] \cdot \text{var}(X) + (E[X])^2 \cdot \text{var}(N) \\ &= \frac{1}{p} \cdot \frac{1}{\lambda^2} + \frac{1}{\lambda^2} \cdot \frac{1-p}{p^2} = \frac{1}{\lambda^2 p^2}. \end{aligned}$$

Για τη φορογενή τιμή της Y έχουμε πρώτα ότι:

$$M_X(s) = \frac{\lambda}{\lambda - s}, \quad M_N(s) = \frac{p \cdot e^s}{1 - (1-p)e^s}.$$

$$\text{Επομένως} \quad M_Y(s) = \frac{p \cdot M_X(s)}{1 - (1-p)M_X(s)} = \frac{p \cdot \frac{\lambda}{\lambda - s}}{1 - (1-p) \frac{\lambda}{\lambda - s}}$$

$$\therefore M_Y(s) = \frac{p\lambda}{p\lambda - s}$$

Που σημαίνει ότι η Y είναι εκθετική τ.μ με παράμετρο $p\lambda$.

$$f_Y(y) = p\lambda \cdot e^{-p\lambda y}, \quad y \geq 0$$

Το αποτέλεσμα είναι ενδιαφέρον ιδίως αν αναλογιστούμε ότι τα άθροισμα ενός σταθερού αριθμού εκθετικών τ.μ. ΔΕΝ είναι εκθετικά κατανοημένο.

□.

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ**

**ΗΥ – 217
ΘΕΩΡΙΑ ΠΙΘΑΝΟΤΗΤΩΝ**

ΣΗΜΕΙΩΣΕΙΣ – ΕΝΟΤΗΤΑ ΠΕΜΠΤΗ:

ΟΡΙΑΚΑ ΘΕΩΡΗΜΑΤΑ

ΠΑΝΑΓΙΩΤΗΣ ΤΣΑΚΑΛΙΔΗΣ - 2002

ΕΝΟΤΗΤΑ 5: ΟΡΙΑΚΑ ΘΕΩΡΗΜΑΤΑ

(154)

Θεωρούμε μια ακολουθία $X_1, X_2, \dots$ ανεξάρτητων όμοια κατανοημένων (independent identically distributed, iid) τ.μ. με μέση τιμή μ και διασπορά σ^2 . Έστω:

$$S_n = X_1 + \dots + X_n$$

Το άθροισμα των πρώτων n από αυτές. Τα οριακά θεωρήματα ασχολούνται με τις ιδιότητες της τ.μ. S_n και συναφών τ.μ. καθώς το n γίνεται μεγάλο.

Προφανώς, λόγω ανεξαρτησίας, $\text{var}(S_n) = \sum_{i=1}^n \text{var}(X_i) = n\sigma^2$ κάτι που υποδηλώνει ότι η κατανομή (PDF) της S_n "απλώνει" καθώς το n αυξάνεται, και δεν έχει κάποιο όριο.

Αν θεωρήσουμε το δειγματικό μέσο όρο (sample mean)

$$M_n = \frac{X_1 + \dots + X_n}{n} = \frac{S_n}{n}$$

Θα έχουμε ότι: $E[M_n] = \mu$ και $\text{var}(M_n) = \frac{\sigma^2}{n}$.

Προκύπτει δηλαδή ότι η διασπορά του M_n τείνει στο 0 καθώς $n \rightarrow \infty$ που δηλώνει ότι η κατανομή του M_n έχει τον κύριο όγκο της πολύ κοντά γύρω από το μ .

Το φαινόμενο αυτό είναι το αντικείμενο ειδικών κανόνων μεγάλων αριθμών, που δείχνουν ότι ο δειγματικός μέσος όρος M_n (μια τ.μ.) τείνει στην σταθερή μέση τιμή μ (έναν αριθμό), με εξαιρετικό τρόπο.

Οι κανόνες αυτοί μας παρέχουν τη μαθηματική βάση για την ερμηνεία της στατιστικής μέσης τιμής $E[X] = \mu$ ως το μέσο όρο (average) ενός μεγάλου αριθμού ανεξάρτητων δειγμάτων επιλεγμένων από την κατανομή της τ.μ. X .

Σε αυτή των ενόχτων, θα δώμε επίσης μια ενδιαμέση μεταξύ των S_n και M_n τ.μ.: Πρώτα αφαιρούμε $n\mu$ από των S_n , ώστε να προκύψει η τ.μ. $S_n - n\mu$ με μηδενική μέση τιμή, και κατόπιν διαιρούμε με $\sigma\sqrt{n}$, ώστε η διασπορά να γίνει ίση με τη μονάδα:

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}, \quad E[Z_n] = 0, \quad \text{var}(Z_n) = 1.$$

Εφόσον η μέση τιμή και η διασπορά της τ.μ. Z_n παραμένουν αμετάβλητες καθώς $n \rightarrow \infty$, η κατανομή της ούτε "απλώνει" ούτε "μαζεύει" σε κάποιο σημείο. Το λεγόμενο Θεώρημα Κεντρικού Ορίου ασχολείται με την ασυμπτωτική μορφή της κατανομής της Z_n , και δείχνει ότι τείνει οριακά στην κανονική (Gaussian) κατανομή.

Τα οριακά θεωρήματα είναι χρήσιμα για δύο λόγους:

- (α) Εννοιολογικά, μας προσφέρουν μια ερμηνεία των μέσων τιμών και των πιθανοτήτων βάσει μακρών ακολουθιών ομοίων και ανεξαρτητών πειραμάτων.
- (β) Επιτρέπουν την προσεγγιστική ανάλυση των ιδιοτήτων τ.μ. όπως οι S_n , M_n , Z_n . Ακριβής ανάλυση τέτοιων τ.μ. απαιτεί συναρτήσεις πυκνότητας πιθανότητας που είναι δύσκολο να υπολογιστούν όταν $n \rightarrow \infty$.

Στην παράγραφο αυτή παρουσιάζουμε ορισμένες σημαντικές ανισότητες οι οποίες χρησιμοποιούν τη μέση τιμή και τη διασπορά μιας τ.μ. X ώστε να φράζουν την πιθανότητα γεγονότων.

5.1.1 Ανισότητα Markov

Η ανισότητα Markov δηλώνει ότι αν μια μη-αρνητική τ.μ. X έχει μικρή μέση τιμή, τότε η πιθανότητα να πάρει μια μεγάλη τιμή είναι επίσης μικρή:

$$P(X \geq a) \leq \frac{E[X]}{a}, \quad \forall a > 0$$

(Η τ.μ. X παίρνει μη-αρνητικές τιμές)

Απόδειξη: Έστω a θετική σταθερά. Ορίζουμε την τ.μ. Y_a :

$$Y_a = \begin{cases} 0 & \text{αν } X < a \\ a & \text{αν } X \geq a \end{cases}$$

Από τον ορισμό ησυχόντα ότι πάντα

$$Y_a \leq X$$

$$\Rightarrow E[Y_a] \leq E[X]$$

$$\text{Αλλά, } E[Y_a] = a P(Y_a = a) = a P(X \geq a).$$

$$\text{Συνεπώς } a P(X \geq a) \leq E[X] \Rightarrow P(X \geq a) \leq \frac{E[X]}{a} \quad \text{ο.ε.δ.} //$$

5.1.2 Ανισότητα Chebyshev

Η ανισότητα Chebyshev δηλώνει ότι αν η τ.μ. X έχει μικρή διασπορά, τότε η πιθανότητα να πάρει τιμή μακριά από τη μέση τιμή της είναι επίσης μικρή:

$$P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2} \quad \forall c > 0$$

Απόδειξη: Θεωρούμε τη μη-αρνητική τιμ. $(X-\mu)^2$ (157)

και εφαρμόζουμε την ανισότητα Markov με $a=c^2$:

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

$$\therefore P(|X-\mu| \geq c) \leq \frac{\sigma^2}{c^2} \quad \text{Q.E.D.} //$$

Μια εναλλακτική έκφραση της ανισότητας Chebyshev προκύπτει αν θέσουμε $c=k\sigma$ ($k>0$). Τότε:

$$P(|X-\mu| \geq k\sigma) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$$

που λέει ότι η πιθανότητα να πάρει η τιμ. X τιμή περισσότερο από k τυπικές αποκλίσεις μακριά από τη μέση τιμή της είναι το πολύ $1/k^2$.

5.2 Ο ΑΣΘΕΝΗΣ ΝΟΜΟΣ ΤΩΝ ΜΕΓΑΛΩΝ ΑΡΙΘΜΩΝ (Weak Law of Large Numbers).

Ο ασθενής νόμος των μεγάλων αριθμών λέει ότι ο μέσος όρος ενός μεγάλου αριθμού ανεξαρτήτων και όμοια κατανεμημένων (iid) τ.τ. είναι πολύ κοντά στην στατιστική μέση τιμή, με μεγάλη πιθανότητα.

Θεωρούμε $X_1, X_2, \dots$ ακολουθία iid τ.τ. με μέση τιμή μ και διασπορά σ^2 . Ορίζουμε τον μέσο όρο:

$$M_n = \frac{X_1 + \dots + X_n}{n}$$

$$\text{Έχουμε } E[M_n] = \frac{E[X_1] + \dots + E[X_n]}{n} = \frac{n\mu}{n} = \mu \quad \text{και}$$

$$\text{var}(M_n) = \frac{\text{var}(X_1 + \dots + X_n)}{n^2} = \frac{\text{var}(X_1) + \dots + \text{var}(X_n)}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

όπου χρησιμοποίησαμε το γεγονός ότι οι τ.τ. X_i είναι ανεξάρτητες.

Εφαρμόζουμε την ανισότητα Chebyshev και παίρνουμε ότι: (158)

$$P(|M_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}, \quad \forall \varepsilon > 0.$$

Παρατηρούμε ότι για σταθερό ε , η ποσότητα $\frac{\sigma^2}{n\varepsilon^2} \xrightarrow{n \rightarrow \infty} 0$.

Συνεπώς, προκύπτει ότι:

□ Weak Law of Large Numbers (WLLN)

Έστω $X_1, X_2, \dots$ iid τ.μ. με μέση τιμή $E[X_i] = \mu$.

Για κάθε $\varepsilon > 0$, ισχύει ότι:

$$P(|M_n - \mu| \geq \varepsilon) = P\left(\left|\frac{X_1 + \dots + X_n}{n} - \mu\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow \infty} 0$$

Ο WLLN λέει ότι για μεγάλα n , ο κύριος όγκος της κατανομής του M_n βρίσκεται κοντά στο μ . Δηλαδή, αν θεωρήσουμε το διάστημα $[\mu - \varepsilon, \mu + \varepsilon]$ γύρω από το μ , με μεγάλη πιθανότητα το M_n θα βρίσκεται μέσα στο διάστημα. Καθώς $n \rightarrow \infty$, η πιθανότητα αυτή τείνει στο 1. Φυσικά αν το ε είναι πολύ μικρό, πρέπει να περιμένουμε περισσότερα (δηλαδή, χρειαζόμαστε μεγαλύτερα n) πριν μπορέσουμε να πούμε ότι είναι πολύ πιθανό το $M_n \in [\mu - \varepsilon, \mu + \varepsilon]$.

Παράδειγμα: Δημοσκοπήσεις

Έστω p το ποσοστό των ψηφοφόρων οι οποίοι υποστηρίζουν κάποιον πολιτικό. Ρωτάμε n "τυχαία επιλεγμένους" ψηφοφόρους και μετράμε το κλάσμα M_n αυτών που υποστηρίζουν τον πολιτικό. Θεωρούμε την ποσότητα M_n ως τον εκτιμητή του p , και θέλουμε να μελετήσουμε τις ιδιότητές του.

Εφαρμόζουμε το "τυχαία επιλεγμένοι ψηφοφόροι" να σημαίνει (159) ότι η άποψη επιλέγεται ανεξάρτητα από τον ημιδοσικό. Επομένως, η απάντηση καθενός μπορεί να θεωρηθεί ως ανεξάρτητη Βερνούλλι. Χι με π.δ. επιτυχίας p και διασπορά $\sigma^2 = p(1-p)$.

Η Ανισότητα Chebyshev δίνει ότι:

$$P(|M_n - p| \geq \varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2}.$$

Βέβαια, η ακριβής τιμή της p είναι άγνωστη (για αυτό γίνεται η δημοσκοπήση!!). Από την άρα, ισχύει ότι $p(1-p) \leq 1/4$ + $0 \leq p \leq 1$. Επομένως

$$P(|M_n - p| \geq \varepsilon) \leq \frac{1}{4n\varepsilon^2}.$$

Έτσι, αν $\varepsilon = 0.1$ και $n = 100$, έχουμε ότι

$$P(|M_{100} - p| \geq 0.1) \leq \frac{1}{4 \times 100 \times 0.1^2} = 0.25.$$

Με άλλα λόγια: Με ένα δείγμα 100 ατόμων, η πιθανότητα ότι ο εκτιμητής μας M_n είναι λάθος περισσότερο από 10% είναι μικρότερη ή ίση των 25%.

Έστω τώρα ότι έχουμε σωστής ηρωδία γραφές ως προς την ακρίβεια της δημοσκοπήσης (οι εκποχές ημεδιάζουν!!). Θα θέλαμε να επιτύχουμε υψηλή εμπιστοσύνη (high confidence), δηλαδή πιθανότητα τουλάχιστον 95% ότι η εκτίμησή μας θα είναι πολύ ακριβής (εφάρκτα το ποσό 0.01 του p). Πόσο μεγάλο δείγμα πρέπει να χρησιμοποιήσουμε;

$$P(|M_n - p| \geq 0.01) \leq \frac{1}{4n(0.01)^2}$$

Πρέπει να επιλέξουμε το n αρκετά μεγάλο ώστε

$$\frac{1}{4n(0.01)^2} \leq 1 - 0.95 = 0.05 \Rightarrow \underline{\underline{n \geq 50,000}}$$

5.3

ΣΥΓΚΛΙΣΗ ΩΣ ΠΡΟΣ ΤΗΝ ΠΙΘΑΝΟΤΗΤΑ (Convergence in Probability).

(160)

Μπορούμε να ερμηνεύσουμε την έκφραση $P(|M_n - \mu| \geq \varepsilon) \xrightarrow{n \rightarrow \infty} 0$

του WLLN ως: "Το M_n συγκλίνει στο μ ."

Βέβαια, αφού η $M_1, M_2, \dots$ είναι μια ακολουθία τ.μ. και όχι απλών αριθμών, θα πρέπει να ορίσουμε σύγκλιση ακολουθιών τ.μ. Ας θυμόμαστε πρώτα τον ορισμό της σύγκλισης μιας ακολουθίας αριθμών:

Έστω $a_1, a_2, \dots$ μια ακολουθία πραγματικών αριθμών, και έστω a ένας άλλος πραγματικός αριθμός. Λέμε ότι η ακολουθία a_n συγκλίνει στο a ($\lim_{n \rightarrow \infty} a_n = a$) αν για κάθε $\varepsilon > 0$ υπάρχει κάποιο n_0 ώστε

$$|a_n - a| \leq \varepsilon \quad \forall n \geq n_0.$$

Ανάλογα ορίζουμε τη σύγκλιση ακολουθιών τ.μ.:

Έστω $Y_1, Y_2, \dots$ μια ακολουθία τ.μ. (όχι απαραίτητα ανεξάρτητων), και έστω a ένας πραγματικός αριθμός. Λέμε ότι η ακολουθία Y_n συγκλίνει στο a ως προς την πιθανότητα, αν για κάθε $\varepsilon > 0$, έχουμε ότι:

$$\lim_{n \rightarrow \infty} P(|Y_n - a| \geq \varepsilon) = 0.$$

Η παραπάνω σχέση δηλώνει ότι "εχεδόν όλη" η κατανομή (PMF ή PDF) της Y_n είναι συγκεντρωμένη στο διάστημα $[a - \varepsilon, a + \varepsilon]$ για μεγάλες τιμές του n .

Ο παραπάνω ορισμός μπορεί επίσης να ερμηνευτεί ως
 ακολουθώς: Για κάθε $\varepsilon > 0$ και για κάθε $\delta > 0$, υπάρχει
 κάποιο n_0 ώστε: (161)

$$\underline{P(|Y_n - a| \geq \varepsilon) \leq \delta, \quad \forall n \geq n_0}$$

Το ε ονομάζεται "επίπεδο ακρίβειας" (accuracy level) και
 το δ ονομάζεται "επίπεδο εμπιστοσύνης" (confidence level).

Οπότε, η σχέση δηλώνει ότι: για οποιαδήποτε επίπεδα
 ακρίβειας και εμπιστοσύνης, η Y_n θα είναι ίση με a
 (μέσα σε αυτά τα επίπεδα) δεδομένου ότι το n είναι αρκετά
 μεγάλο.

Παράδειγμα: Έστω μια ακολουθία ανεξάρτητων τ.τ. X_n που είναι
 ομοιόμορφα κατανεμημένες στο διάστημα $[0, 1]$. Έστω

$$Y_n = \min \{X_1, \dots, X_n\}.$$

Προφανώς, η ακολουθία τιμών της Y_n δεν μπορεί να αυξάνει
 καθώς το n αυξάνει, και πιθανώς θα μειώνεται (όταν
 προκύπτει μια τιμή της X_n που είναι μικρότερη από τις προηγού-
 μενες τιμές.) Επομένως περιμέναμε η Y_n να συγκλίνει
 στο 0. Πράγματι, $\forall \varepsilon > 0$ (προφανώς $\varepsilon < 1$)

$$\begin{aligned} P(|Y_n - 0| \geq \varepsilon) &= P(X_1 \geq \varepsilon, \dots, X_n \geq \varepsilon) \\ &= P(X_1 \geq \varepsilon) \dots P(X_n \geq \varepsilon) \\ &= (1 - \varepsilon) \dots (1 - \varepsilon) = (1 - \varepsilon)^n. \end{aligned}$$

$$\Rightarrow \lim_{n \rightarrow \infty} P(|Y_n - 0| \geq \varepsilon) = \lim_{n \rightarrow \infty} (1 - \varepsilon)^n = 0.$$

ΟΕΔ //

Παράδειγμα: Έστω Y εκθετική τ.μ. με $\lambda = 1$. Σχηματίζουμε (162)
την ακολουθία $Y_n = Y/n$ όπου $n = 1, 2, 3, \dots$ (Προφανώς, οι Y_n δεν είναι ανεξάρτητες τ.μ.) Θέλουμε να ελέγξουμε αν η ακολουθία Y_n συγκλίνει στο Φ .

$$- \forall \varepsilon > 0, \quad P(|Y_n - 0| \geq \varepsilon) = P(Y_n \geq \varepsilon) = P(Y \geq n\varepsilon) = e^{-n\varepsilon}.$$

$$\text{Επομένως, } \lim_{n \rightarrow \infty} P(|Y_n - 0| \geq \varepsilon) = \lim_{n \rightarrow \infty} e^{-n\varepsilon} = 0. \quad \text{ΟΕΔ} //$$

Παράδειγμα. Το ότι η ακολουθία Y_n συγκλίνει στον αριθμό a ,
δεν σημαίνει ότι η ακολουθία $E[Y_n]$ συγκλίνει στο a .

Θεωρούμε τ.μ. Y_n :

$$P(Y_n = y) = \begin{cases} 1 - \frac{1}{n} & , \quad y = 0 \\ 1/n & , \quad y = n^2 \\ 0 & , \quad \text{αλλιώς} \end{cases}$$

Έχουμε ότι $\forall \varepsilon > 0$,

$$\lim_{n \rightarrow \infty} P(|Y_n| \geq \varepsilon) = \lim_{n \rightarrow \infty} \frac{1}{n} = 0$$

$\Rightarrow$ Η ακολουθία των τ.μ. Y_n συγκλίνει στο 0 .

$$\text{Από την άλλη, } E[Y_n] = \frac{n^2}{n} = n \xrightarrow{n \rightarrow \infty} +\infty \quad \text{ΟΕΔ } 1.$$

5.4ΘΕΩΡΗΜΑ ΚΕΝΤΡΙΚΟΥ ΟΡΙΟΥ

(163)

(Central Limit Theorem, CLT)

Σύμφωνα με τον WLLN, η κατανομή του μέσου όρου $\bar{X}_n$ τείνει να επικεντρώνεται στη γειτονιά της στατιστικής μέσης τιμής $\mu = E[X_i]$. Πιο συγκεκριμένα, η διασπορά του $\bar{X}_n$ τείνει στο μηδέν. Από την άλλη, η διασπορά του αθροίσματος $S_n = X_1 + \dots + X_n = n\bar{X}_n$ τείνει στο άπειρο και η κατανομή του S_n δεν συγκρίνεται σε κάποια κατανομή.

Ένα ενδιαφέρον αθροισμα με πολύ ενδιαφέρουσες ιδιότητες προκύπτει αν θεωρήσουμε την απόκλιση $S_n - n\mu$ του S_n από τη μέση του τιμή $n\mu$, και διαιρέσουμε με συντελεστή ανάλογο του $1/\sqrt{n}$ ώστε η διασπορά που θα προκύψει να είναι σταθερή καθώς το n αυξάνει. Το Θεώρημα Κεντρικού Ορίου λέει ότι η κατανομή αυτής της νέας τ.μ. συγκρίνεται στην κανονική (Gaussian) κατανομή.

Θεωρούμε λοιπόν μια ακολουθία $X_1, X_2, \dots$ ανεξάρτητων και όμοια κατανομημένων (iid) τ.μ. με μέση τιμή μ και διασπορά σ^2 . Ορίζουμε την τ.μ.

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{X_1 + \dots + X_n - n\mu}{\sigma\sqrt{n}}$$

Εύκολα προκύπτει ότι:

$$E[Z_n] = \frac{E[X_1 + \dots + X_n] - n\mu}{\sigma\sqrt{n}} = 0,$$

και

$$\text{var}(Z_n) = \frac{\text{var}(X_1 + \dots + X_n)}{\sigma^2 n} = \frac{\text{var}(X_1) + \dots + \text{var}(X_n)}{\sigma^2 n} = \frac{n\sigma^2}{n\sigma^2} = 1.$$

Το Θεώρημα Κεντρικού Ορίου (Central Limit Theorem, CLT)

Έστω $X_1, X_2, \dots$ μια ακολουθία ανεξαρτήτων και όμοιων κατανομών (iid) τ.μ. με κοινή μέση τιμή μ και διασπορά σ^2 . Ορίζουμε την τ.μ. Z_n :

$$Z_n = \frac{X_1 + \dots + X_n - n\mu}{\sigma\sqrt{n}}$$

Η αθροιστική συνάρτηση πυκνότητας (CDF) της Z_n συγκρίνεται στην standard κανονική CDF:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-x^2/2} dx,$$

με την έννοια:

$$\lim_{n \rightarrow \infty} P(Z_n \leq z) = \Phi(z), \quad \forall z.$$

Το CLT είναι σχεδόν καθολικά εφαρμόσιμο: Υποθέτει μόνο την ανεξαρτησία των τ.μ. X_i και την ύπαρξη πεπερασμένης μέσης τιμής και διασποράς. Δεν βάζει κανένα περιορισμό στην κατανομή των τ.μ. X_i που μπορεί να είναι διακριτή, συνεχής, ή μικτή.

Εννοιολογικά δηλώνει ότι το άθροισμα ενός μεγάλου αριθμού ανεξαρτήτων τ.μ. τείνει να έχει κανονική κατανομή. Επομένως εφαρμόζεται σε πρακτικές καταστάσεις όπου ένα τυχαίο αποτέλεσμα είναι το άθροισμα ενός μεγάλου αριθμού μικρών αλλά ανεξαρτήτων τυχαίων παραγόντων. Ο δόρυπος σε πολλά φυσικά ή μηχανικά συστήματα έχει αυτή την ιδιότητα. Εμπειρικά έχει δείξει ότι όπως ο δόρυπος ακολουθεί Gaussian κατανομή.

Πρακτικά, το CLT μας απαλλάσσει από την ανάγκη υπολογισμού πολύπλοκων στατιστικών μοντέλων (PMF και PDF).

(15)

Υπολογισμοί για γεγονότα που περιέχουν την τ.μ. Z_n μπορούν να γίνουν πολύ απλά με βάση τους πίνακες της standard κανονικής CDF και απαιτούν μόνο τη γνώση μέσων τιμών και διασπορών.

5.4.1 Προσεγγίσεις Βασισμένες στο CLT.

Το CLT μας επιτρέπει τον υπολογισμό πιθανοτήτων σχετικών με την τ.μ. Z_n εάν η Z_n να ήταν κανονική τ.μ. Εφόσον, όπως έχουμε δει, η κανονική κατανομή διατηρείται με γραμμικούς μετασχηματισμούς, προκύπτει ότι μπορούμε, προσεγγιστικά πάντα, να θεωρούμε την τ.μ. S_n ($Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$) ως κανονική τ.μ. με μέση τιμή $n\mu$ και διασπορά $n\sigma^2$.

Κανονική Προσέγγιση Βασισμένη στο CLT

Έστω $S_n = X_1 + \dots + X_n$, όπου X_i iid τ.μ. με μέση τιμή μ και διασπορά σ^2 . Αν το πλήθος n είναι μεγάλο, η πιθανότητα $P(S_n \leq c)$ μπορεί να προσεγγιστεί θεωρώντας την τ.μ. S_n κανονική, σύμφωνα με την ακόλουθη διαδικασία:

1. Υπολογίζουμε τη μέση τιμή $n\mu$ και τη διασπορά $n\sigma^2$ της S_n .
2. Υπολογίζουμε την ανηγμένη τιμή $z = (c - n\mu) / \sigma\sqrt{n}$.
3. Χρησιμοποιούμε την προσέγγιση

$$P(S_n \leq c) \approx \Phi(z)$$

όπου $\Phi(z)$ είναι διαδεδομένη σε πίνακες.

Παράδειγμα. Φορτώνουμε ένα φορτίο με 100 δέματα (166)

των οποίων τα βάρη είναι iid τ.μ. ομοιόμορφα κατανοημένες μεταξύ 5 και 50 κιλών. Ποια η πιθανότητα ότι το συνολικό βάρος θα είναι μεγαλύτερο από 3.000 κιλά;

- = Πραφανώς για την ακριβή απάντηση, θα έπρεπε να υπολογίσουμε την CDF του συνολικού φορτίου $S_{100} = X_1 + \dots + X_{100}$ και κατόπιν την πιθανότητα $P(S_{100} > 3000)$. Μια πολύ καλή προσέγγιση όμως μπορεί εύκολα να προκύψει με εφαρμογή του CLT:

Η μέση τιμή, μ , και η διασπορά, σ^2 , κάθε δέματος προκύπτει από όσα έχουμε πει για ομοιόμορφες κατανομές:

$$\mu = \frac{5+50}{2} = 27.5, \quad \sigma^2 = \frac{(50-5)^2}{12} = 168.75.$$

Υπολογίζουμε τώρα την ανηγμένη τιμή.

$$Z = \frac{3000 - 100 \times 27.5}{\sqrt{168.75 \times 100}} = 1.92,$$

και χρησιμοποιούμε τους πίνακες για να υπολογίσουμε

$$P(S_{100} \leq 3000) \approx \Phi(1.92) = 0.9726.$$

$$\text{Τελικά, } P(S_{100} > 3000) = 1 - P(S_{100} \leq 3000) \approx 0.0274 //$$

Παράδειγμα: Μια μηχανή επεξεργάζεται ελαφύματα, ένα κάθε φορά.

Οι χρόνοι επεξεργασίας είναι iid τ.μ., ομοιόμορφα κατανοημένες στο διάστημα $[L, 5]$ sec. Θέχουμε να προσεγγίσουμε την πιθανότητα ότι ο αριθμός των ελαφτημάτων που η μηχανή επεξεργάζεται σε 320 secs είναι τουλάχιστον 100.

(167)

- Ας ονομάσουμε N_{320} αυτών τον αριθμό. Θέλουμε να υπολογίσουμε την $P(N_{320} \geq 100)$. Δεν υπάρχει κάποιος προφανής τρόπος να εκφράσουμε την τ.μ. N_{320} ως ένα άθροισμα iid τ.μ. οπότε πρέπει να βρούμε κάποιο ισοδύναμο του γεγονότος $\{N_{320} \geq 100\}$.

Έστω X_i ο χρόνος επεξεργασίας του i -ετού εξαρτήματος. Θέτουμε $S_{100} = X_1 + \dots + X_{100}$ το συνολικό χρόνο επεξεργασίας των 100 εξαρτημάτων. Το γεγονός $\{N_{320} \geq 100\}$ είναι το ίδιο με το γεγονός $\{S_{100} \leq 320\}$, που μπορεί να προσεγγιστεί με χρήση του CLT για την S_{100} .

Έχουμε ότι: $\mu = E[X_i] = \frac{1+5}{2} = 3$, $\sigma^2 = \text{var}(X_i) = \frac{(5-1)^2}{12} = \frac{4}{3}$.

Υπολογίζουμε την ανηγμένη τιμή:

$$z = \frac{320 - n\mu}{\sigma\sqrt{n}} = \frac{320 - 300}{\sqrt{100 \cdot 4/3}} = 1.73,$$

και υπολογίζουμε

$$P(S_{100} \leq 320) \approx \Phi(1.73) = 0.9582.$$

//

- Αν η διασπορά των X_i είναι άγνωστη, αλλά ένα άνω όριο της σ^2 υπάρχει, το CLT μπορεί να χρησιμοποιηθεί για τον υπολογισμό ορίων στις πιθανότητες που μας ενδιαφέρουν.

Παράδειγμα. Δημοσκοπήσεις (Δες παράγραφο 7.2)

Ρωτούμε n ψηφοφόρους και καταγράφουμε το κλάσμα M_n αυτών που υποστηρίζουν κάποιον πολιτικό. Αν p είναι το ποσοστό του συνολικού πληθυσμού που υποστηρίζει τον συγκεκριμένο πολιτικό, τότε

$$M_n = \frac{X_1 + \dots + X_n}{n},$$

όπου X_i είναι ανεξάρτητες Bernoulli τ.μ. με παράμετρο p .

Επομένως, το M_n έχει μέση τιμή p και διασπορά $\frac{p(1-p)}{n}$. (168)

Από το CLT, το άθροισμα $X_1 + \dots + X_n$ είναι προσεγγιστικά Gaussiano, και επομένως το M_n είναι επίσης προσεγγιστικά Gaussiano.

Ενδιαφερόμαστε για την πιθανότητα $P(|M_n - p| \geq \varepsilon)$ ότι το αράχτι δημοκρίτους είναι μεγαλύτερο από κάποια επιθυμητή ακρίβεια, ε . Λόγω συμμετρίας της κανονικής CDF, γύρω από τη μέση τιμή, έχουμε

$$P(|M_n - p| \geq \varepsilon) \approx 2 P(M_n - p \geq \varepsilon).$$

Η διασπορά $p(1-p)/n$ της τ.μ. $M_n - p$ εξαρτάται από το p και είναι επομένως άγνωστη. Παρατηρούμε ότι η πιθανότητα μιας μεγάλης απόκλισης από τη μέση τιμή αυξάνεται με τη διασπορά. Άρα μπορούμε να βρούμε ένα άνω όριο για την $P(M_n - p \geq \varepsilon)$ υποθέτοντας ότι η τ.μ. $M_n - p$ έχει τη μέγιστη δυνατή διασπορά $\left(\frac{p(1-p)}{n}\right)_{\max_p} = \frac{1}{4n}$.

Για να υπολογίσουμε αυτό το άνω όριο, βρίσκουμε την ανηγμένη τιμή:

$$z = \frac{\varepsilon - 0}{\sqrt{1/4n}} = 2\varepsilon\sqrt{n}.$$

Επομένως, $P(M_n - p \geq \varepsilon) \leq 1 - \Phi(z) = 1 - \Phi(2\varepsilon\sqrt{n})$.

Π.χ. για $n=100$ και $\varepsilon=0.1$, έχουμε ότι

$$\begin{aligned} P(|M_{100} - p| \geq 0.1) &\approx 2 P(M_{100} - p \geq 0.1) \\ &\leq 2 - 2\Phi(2 \times 0.1 \times \sqrt{100}) = 2 - 2\Phi(2) = \underline{0.054} \end{aligned}$$

Το όριο αυτό είναι πολύ μικρότερο, και πολύ πιο ακριβές, από ό,τι είχε υπολογιστεί στην παράγραφο 7.2 μέσω της ανισότητας Chebyshev (0.25).

Μελετάμε τώρα το αντίστροφο πρόβλημα: Πόσο μεγάλο δείγμα (n) πρέπει να χρησιμοποιηθεί ώστε το φράγμα εκτίμησης μέσω του μ_n να είναι το πολύ $0.01 \times \rho$ με πιθανότητα τουλάχιστον 0.95; (169)

Υποθέτοντας και πάλι την μεγαλύτερη δυνατή διασπορά για το μ_n , έχουμε ότι:

$$P(|\mu_n - \rho| \geq 0.01) \leq 2 - 2\Phi(2 \times 0.01 \times \sqrt{n}) \leq 1 - 0.95$$

$$\therefore \Phi(2 \times 0.01 \times \sqrt{n}) \geq 0.975$$

Από τους πίνακες έχουμε ότι $\Phi(1.96) = 0.975$, οπότε

$$2 \times 0.01 \times \sqrt{n} \geq 1.96 \Rightarrow \underline{n \geq 9604}$$

Αυτό μας λέει ότι χρειάζεται ένα πολύ μικρότερο δείγμα από ότι μας είχε ότι απαιτείται η κρισιμότητα Chebyshev ($n \geq 50000$). //

Ένα από τα κύρια ζητήματα σε ότι αφορά την πρακτική εφαρμογή του CLT, είναι πόσο μεγάλο πρέπει να είναι το n ώστε να αρχίσει να ισχύει η προσέγγιση της κανονικής κατανομής. Δεν υπάρχει κάποιος «πρό» γενικός κανόνας ως προς αυτό.

Ένας παράγοντας είναι το πόσο κοντά στην κανονική είναι η κατανομή των X_i και κυρίως αν αυτή είναι συμμετρική.

Π.χ. αν οι τ.μ. X_i είναι ομοιόμορφα κατανομημένες, τότε η τ.μ. S_n είναι ήδη πολύ κοντά στην κανονική. Αλλά αν οι X_i είναι εκθετικές τ.μ. μία πολύ μεγαλύτερη τιμή του n απαιτείται για να είναι η S_n περίπου κανονική.

Επιπλέον, η κανονική προσέγγιση για την πιθανότητα $P(S_n \leq c)$ είναι γενικά πιο ακριβής όσο πιο κοντά φτάνεται η τιμή c στη μέση τιμή της S_n .

5.4.2 Η De Moivre-Laplace Προσέγγιση της Διωνυμικής (470)

Έχουμε δει ότι η Διωνυμική τ.φ. S_n με παραμέτρους n και p εκφράζεται ως το άθροισμα n ανεξάρτητων τ.φ. Bernoulli $X_1, \dots, X_n$, με κοινή παράμετρο p :

$$S_n = X_1 + \dots + X_n.$$

Έχουμε δείξει ότι

$$\mu = E[X_i] = p, \quad \sigma = \sqrt{\text{Var}(X_i)} = \sqrt{p(1-p)}.$$

Θα χρησιμοποιήσουμε τώρα το CLT για να προσεγγίσουμε την πιθανότητα του γεγονότος $\{k \leq S_n \leq l\}$, όπου l και k είναι κάποιοι ακέραιοι αριθμοί.

Με βάση όσα είπαμε στην παράγραφο 7.4.1,

$$k \leq S_n \leq l \Leftrightarrow \frac{k-np}{\sqrt{np(1-p)}} \leq \frac{S_n-np}{\sqrt{np(1-p)}} \leq \frac{l-np}{\sqrt{np(1-p)}}.$$

Από το CLT, η τ.φ. $\frac{S_n-np}{\sqrt{np(1-p)}}$ ακολουθεί κατά προσέγγιση standard κανονική κατανομή, και επομένως

$$\begin{aligned} P(k \leq S_n \leq l) &= P\left(\frac{k-np}{\sqrt{np(1-p)}} \leq \frac{S_n-np}{\sqrt{np(1-p)}} \leq \frac{l-np}{\sqrt{np(1-p)}}\right) \\ &\approx \Phi\left(\frac{l-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k-np}{\sqrt{np(1-p)}}\right). \end{aligned}$$

που είναι το ολοκλήρωμα της standard κανονικής PDF από το k στο l . Στην πράξη, για λόγους προσέγγισης και της $P(S_n=k)$, ολοκληρώνουμε από το $k-\frac{1}{2}$ στο $l+\frac{1}{2}$ και έχουμε:

Αν η S_n είναι διωνυμική τ.φ. με παραμέτρους n, p και το n είναι μεγάλο και k, l είναι μη αρνητικοί ακέραιοι αριθμοί, τότε

$$P(k \leq S_n \leq l) \approx \Phi\left(\frac{l+\frac{1}{2}-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k-\frac{1}{2}-np}{\sqrt{np(1-p)}}\right).$$

Παράδειγμα. Έστω S_n διωνυμική τ.μ. με $n=36$ (171)
και $p=0.5$. Ποια η πιθανότητα του $\{S_n \leq 21\}$;

- Αντίστροφη Υπολογισμός:
$$P(S_n \leq 21) = \sum_{k=0}^{21} \binom{36}{k} (0.5)^k (1-0.5)^{36-k}$$
$$= \sum_{k=0}^{21} \binom{36}{k} (0.5)^{36} = 0.8785$$

• De Moivre-Laplace Προσέγγιση:

$$P(S_n \leq 21) \approx \Phi\left(\frac{21.5 - np}{\sqrt{np(1-p)}}\right) = \Phi\left(\frac{21.5 - 18}{3}\right) = \Phi(1.17)$$
$$= 0.879$$

Για το γεγονός $\{S_n = 19\}$, έχουμε:

$$P(S_n = 19) = \binom{36}{19} (0.5)^{36} = 0.1251$$

$$\approx \Phi\left(\frac{19.5 - 18}{3}\right) - \Phi\left(\frac{18.5 - 18}{3}\right) = 0.124$$

5.5 Ο ΙΣΧΥΡΟΣ ΝΟΜΟΣ ΤΩΝ ΜΕΓΑΛΩΝ ΑΡΙΘΜΩΝ
(Strong Law of Large Numbers, SLLN).

Ο SLLN είναι παραπλήσιος στον WLLN της παραγράφου 7.2 στο ότι έχει να κάνει με την σύγκλιση του μέσου όρου στην στατιστική μέση τιμή. Διαφέρει όμως στο ότι χρησιμοποιεί ισχυρό νόμο κανόνα σύγκλισης.

▮ Strong Law of Large Numbers (SLLN)

Έστω $X_1, X_2, \dots$ iid τ.μ. με κοινή μέση τιμή μ . Τότε η ακολουθία των μέσων όρων $M_n = (X_1 + \dots + X_n) / n$ συγκλίνει στο μ , με πιθανότητα 1, που σημαίνει ότι

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

Για να κατανοήσουμε το SLLN, πρέπει να θυμόμαστε την (172)
αρχική περιγραφή των πιθανοτικών μοντέλων βάση δειγματικών
χώρων: Το πείραμά μας δημιουργεί πειραματικές τιμές
για κάθε μία από τις z στην (άπειρη) ακολουθία $X_1, X_2, \dots$

Επομένως, ο δειγματοχώρος μας Ω είναι το σύνολο των
άπειρων ακολουθιών $\omega = (x_1, x_2, \dots)$ πραγματικών αριθμών:
κάθε μία από αυτές αποτελεί ένα από ενδεχόμενο του
πείραματος.

Ορίζουμε τώρα το υποσύνολο A του Ω το οποίο αποτε-
λείται από αυτές τις ακολουθίες $(x_1, x_2, \dots)$ των οποίων
ο μέσος όρος είναι μ , δηλαδή

$$(x_1, x_2, \dots) \in A \iff \lim_{n \rightarrow \infty} \frac{x_1 + \dots + x_n}{n} = \mu.$$

Ο SLLN λέει ότι όλη η πιθανότητα είναι συγκεντρωμένη
σε αυτό το υποσύνολο του Ω . Ισοδύναμα, η συλλογή των
ενδεχομένων που δεν ανήκουν στο A έχει πιθανότητα μηδέν.

Η διαφορά μεταξύ SLLN και WLLN είναι λεπτή: Ο WLLN
λέει ότι η πιθανότητα, $P(|M_n - \mu| \geq \varepsilon)$, απόκλισης του M_n
από το μ τείνει στο μηδέν καθώς $n \rightarrow \infty$. Αλλά, για
πεντησάρη n , αυτή η πιθανότητα μπορεί να είναι δευτική, και
υπάρχουν (πιθανόν άχρηστες) φορές που το M_n θα αποκλίνει
σημαντικά από το μ . Ο WLLN δε μας δίνει πληροφορία
πόσο συχνές μπορεί να είναι αυτές οι αποκλίσεις, κάτι που
κάνει ο SLLN: Σύμφωνα με το SLLN, με πιθανότητα 1,
το M_n συγκλίνει στο μ . Αυτό σημαίνει ότι $\forall \varepsilon > 0$, η
διαφορά $|M_n - \mu|$ θα είναι μεγαλύτερη από ε για πεπερασμένο
αριθμό φορές.

5.5.1 Σύγκριση με Πιθανότητα 1

(173)

(Convergence with Probability 1)

Έστω $Y_1, Y_2, \dots$ μια ακολουθία τ.κ. (όχι απαραίτητα ανεξάρτητων) σχετιζόμενες με το ίδιο πιθανοτικό μοντέλο. Έστω c ένας πραγματικός αριθμός. Λέμε ότι η ακολουθία Y_n συγκλίνει στο c με πιθανότητα 1 (ή σχεδόν βέβαια) αν

$$P\left(\lim_{n \rightarrow \infty} Y_n = c\right) = 1.$$

Με άλλα λόγια το πείραμά μας δημιουργεί όποιες ακολουθίες και όλη η πιθανότητα συγκεντρώνεται σε αυτές που συγκλίνουν στο c . Αυτό δεν σημαίνει ότι άλλες ακολουθίες είναι αδύνατες, απλά ότι είναι εξαιρετικά κινδυνώδεις, με την έννοια ότι η συνολική τους πιθανότητα είναι μηδέν.

Παράδειγμα: Έστω μια ακολουθία ανεξάρτητων τ.κ. X_n που είναι ομοιόμορφα κατανοημένες στο $[0, 1]$. Έστω $Y_n = \min\{X_1, \dots, X_n\}$. Θέλουμε να δείξουμε ότι η ακολουθία Y_n συγκλίνει στο 0 με πιθανότητα 1.

Προφανώς (όπως είδαμε και στο αντίστοιχο παράδειγμα στην παράγραφο 7.2) η ακολουθία Y_n είναι μη αύξουσα:

$$Y_{n+1} \leq Y_n \quad \forall n.$$

Αφού είναι και κάτω φραγμένη από το 0, θα έχει ένα όριο, έστω το Y . Έστω κάποιο $\varepsilon > 0$. Αν $Y \geq \varepsilon$, τότε $X_i \geq \varepsilon \quad \forall i$ το οποίο σημαίνει ότι

$$P(Y \geq \varepsilon) \leq P(X_1 \geq \varepsilon, \dots, X_n \geq \varepsilon) = (1 - \varepsilon)^n.$$

Αφού αυτό ισχύει $\forall n$, τότε $P(Y \geq \varepsilon) \leq \lim_{n \rightarrow \infty} (1 - \varepsilon)^n = 0 \quad \forall \varepsilon > 0$

$\Rightarrow P(Y > 0) = 0 \Rightarrow P(Y = 0) = 1$. Αφού Y είναι το όριο της Y_n

συμπεραίνουμε ότι η Y_n συγκλίνει στο 0 με πιθανότητα 1. //