Πιθανότητες 17/12

• **Ροπές**

1) Απλές ροπές ή ροπές περί το μηδέν

ΟΡΙΣΜΟΣ

Έστω μια τ.μ X. Η απλή ροπή k-τάξης συμβολίζεται με μ_k και ορίζεται

$$\mu_k = E(X^k) = \begin{cases} \sum_x x^k p_X(x) & , X \text{ διακριτή} \\ \int_{-\infty}^{+\infty} x^k f_X(x)\,dx & , X \text{ συνεχής} \end{cases}$$

Παρατήρηση

Αν $k=1$ τότε $\mu_1 \overset{\text{ορ}}{=} E(X) = \mu$

2) Κεντρικές ροπές ή ροπές περί τη μέση τιμή

ΟΡΙΣΜΟΣ

Έστω τ.μ. X με μέση τιμή $\mu = E(X)$. Η κεντρική τιμή k-τάξης συμβολίζεται με V_k και ορίζεται

$$V_k = \begin{cases} \sum_x (x-\mu)^k p_X(x) & , X \text{ διακρ} \\ \int_{-\infty}^{+\infty} (x-\mu)^k f_X(x)\,dx & , X \text{ συνεχ} \end{cases}$$

Παρατήρηση

Αν $k=2$ τότε $V_2 = E(X-\mu)^2 = Var(X)$

ΠΡΟΤΑΣΗ

$$V_k = \sum_{r=0}^{k} \binom{k}{r} (-1)^r \mu_1^r \mu_{k-r} \quad , k=1,2,\dots$$

Απόδειξη

$$V_k \overset{\text{ορ}}{=} E(X-\mu)^k = E\left[(-\mu+X)^k\right] \xrightarrow[\text{Νεύτωνα}]{\text{διώνυμο}} E\left\{\sum_{r=0}^{k} \binom{k}{r} (-\mu)^r (X)^{k-r}\right\}$$

$$= \sum_{r=0}^{k} \binom{k}{r} (-1)^r \mu^r E(X)^{k-r} = \sum_{r=0}^{k} \binom{k}{r} (-1)^r \mu_1^r \mu_{k-r}$$

ΠΡΟΤΑΣΗ

Έστω X συνεχής τ.μ. με $E(X)=\mu$. Αν η f είναι συμμετρική γύρω από το μ $\left(f_X(x+\mu) = f_X(\mu-x)\right)$ τότε οι κεντρικές ροπές περιττού τάξου είναι ίσες με 0 δηλ $V_k = 0$ $\quad k = 2n+1 \quad n=0,1,2,\dots$

Απόδειξη

Βασίζεται σε αλλαγές μεταβλητών σε ολοκληρώματα.

Συντελεστής Λοξότητας - Συντελεστής κύρτωσης

ΟΡΙΣΜΟΣ (Λοξότητας)

Ο συντελεστής λοξότητας συμβολίζεται με β και ορίζεται $\beta = \dfrac{V_3}{\sigma^3}$

$\sigma = +\sqrt{Var(X)} = +\sqrt{V_2}$

Παρατηρηση

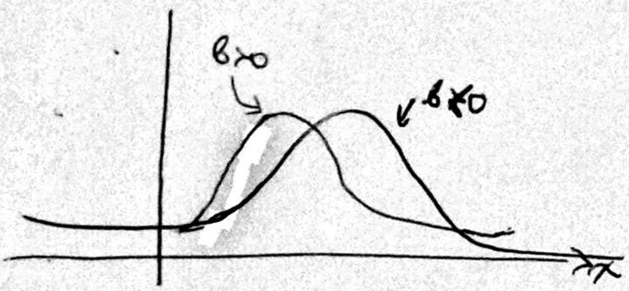
α) Αν η κατανομή είναι συνεχής και συμμετρική τότε $\beta=0$
Το αντίστροφο δεν ισχύει πάντα.

β) • Αν $\beta \neq 0$ τότε η κατανομή δεν είναι συμμετρική

Απόδειξη

Αν $A \Rightarrow B$ τότε $\sim B \Rightarrow \sim A$

• Αν $\beta>0$ τότε η κατανομή είναι λοξή δεξιά



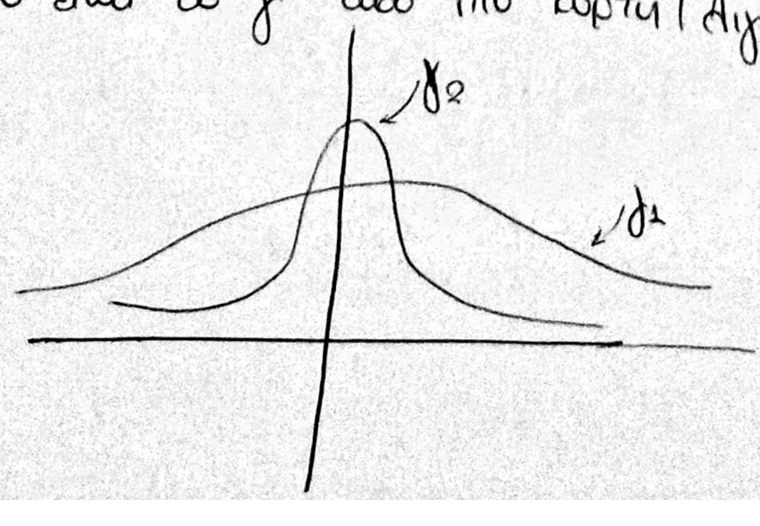
• Αν $\beta<0$ τότε η κατανομή είναι λοξή αριστερά

ΟΡΙΣΜΟΣ (Κυρτώσεις)

Ο συντελεστής κυρτώσεως συμβολίζεται με γ και ορίζεται $\gamma = \frac{V_4}{\sigma^4}$

$\sigma = +\sqrt{Var(x)} = +\sqrt{V_2}$

Όσο πιο μεγάλο είναι το γ τόσο πιο κυρτή (λιγότερο επίπεδη) είναι η κατανομή



$\gamma_1 < \gamma_2$

Ροπογεννήτρια

ΟΡΙΣΜΟΣ

Έστω τ.μ X. Η ροπογεννήτρια της X συμβολίζεται με $m_X(t)$ και ορίζεται $m_X(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx} P_X(x) & , X \text{ διακριτή} \\ \int_{-\infty}^{+\infty} e^{tx} f_X(x)\,dx & , X \text{ συνεχής.} \end{cases}$

Ιδιότητες της ροπογεννήτριας

1) Για $t=0$ $\quad m_X(0) = E(e^{0X}) = E(1) = 1$

2) $m_{aX+b}(t) \overset{\text{ορ}}{=} E(e^{t(aX+b)}) = E(e^{tb} e^{taX}) = e^{tb} E(e^{taX}) \overset{\text{ορ}}{=}$

$= e^{tb} m_X a$

Άρα $m_{aX+b}(t) = e^{tb} m_X(at)$

3) $m_X(t) = \int_{-\infty}^{+\infty} e^{tx} f_X(x)\,dx \quad , \quad t \in \mathbb{R}$

Αν περιοριστώ στα t σε μια περιοχή του 0 $\left(|t| < c \quad c > 0 \right)$

$$\frac{d}{dt} m_X(t) = \frac{d}{dt} \int_{-\infty}^{+\infty} e^{tx} f_X(x)\,dx = \int_{-\infty}^{+\infty} \left(\frac{d}{dt} e^{tx} \right) f_X(x)\,dx$$

Ομοίως $\quad \frac{d^2}{dt^2} m_X(t) = \int_{-\infty}^{+\infty} x^2 e^{tx} f_X(x)\,dx$

$\vdots$

$$\frac{d^k}{dt^k} m_X(t) = \int_{-\infty}^{+\infty} x^k e^{tx} f_X(x)\,dx$$

Άρα $\left. \frac{d^k}{dt^k} m_X(t) \right|_{t=0} = \int_{-\infty}^{+\infty} x^k f_X(x)\,dx = \mu_k \quad k = 1, 2, \ldots$

Αντίστροφα $\quad m_X(t) \overset{op}{=} E(e^{tX}) = E\left(\sum_{k=0}^{\infty} \frac{(tX)^k}{k!}\right) \Rightarrow$

$$\Rightarrow m_X(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} E(X^k) \Rightarrow m_X(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mu_k$$

4) <u>Θεώρημα μονοσημάντου των ροπογεννητριών</u>

Έστω τ.μ X και Y με αδκ F_X και F_Y. Αν οι ροπογεννήτριες $m_X(t)$ και $m_Y(t)$ υπάρχουν και $m_X(t) = m_Y(t)$, $|t| < c$, $c > 0$ τότε $F_X(x) = F_Y(x)$, $x \in \mathbb{R}$

Παράδειγμα.

$$m_X(t) = \frac{1}{6} e^{-t} + \frac{1}{2} e^{-2t} + \frac{1}{3} e^{t} \quad \text{όπου } X \text{ διακριτή}$$

Να υπολογιστούν : $E(X) = ?$, $Var(X) = ?$

Ποια είναι η $P_X(x)$?

<u>Λύση</u>

$$E(X) = \mu_1 = \frac{d}{dt} m_X(t)\Big|_{t=0} = \left[-\frac{1}{6} e^{-t} - e^{-2t} + \frac{1}{3} e^{t}\right]_{t=0} = -\frac{5}{6}$$

$$Var(X) = E(X^2) - (E(X))^2 = \mu_2 - \left(-\frac{5}{6}\right)^2 = \frac{d^2}{dt^2} m_X(t)\Big|_{t=0} - \frac{25}{36} = \frac{15}{6} - \frac{25}{36} = \frac{65}{36}$$

Αφού η X διακριτή έστω $x_1, x_2, x_3 \ldots$ οι τιμές της και $P_X(x_1), P_X(x_2), P_X(x_3), \ldots$ η σπ της.

$$m_X(t) \overset{op}{=} E(e^{tX}) = \sum_{i=1}^{\infty} e^{t x_i} P_X(x_i) =$$

$= e^{tx_1} p_X(x_1) + e^{tx_2} p_X(x_2) + e^{tx_3} p_X(x_3) + \cdots$

Συγκεκριμένα $\quad m_X(t) = \frac{1}{6} e^{-t} + \frac{1}{2} e^{-2t} + \frac{1}{3} e^{t}$

$$m_X(t) = e^{tx_1} p_X(x_1) + e^{tx_2} p_X(x_2) + e^{tx_3} p_X(x_3) + \cdots$$

Από θεώρημα μονοσημάντου αφού οι ροπογεννήτριες ταυτίζονται θα πρέπει $x_1 = -1$ και $p_X(x_1) = \frac{1}{6}$

$x_2 = -2$ και $p_X(x_2) = \frac{1}{2}$

$x_3 = 1$ και $p_X(x_3) = \frac{1}{3}$

$x_4 = 0$ και $p_X(x_4) = 0$ κ.ο.κ.

Αν η X είναι διακριτή με τιμές $-1, -2, 1$ και $p_X(x)$

$$p_X(x) = \begin{cases} \frac{1}{6} & , x = -1 \\ \frac{1}{2} & , x = -2 \\ \frac{1}{3} & , x = 1 \\ 0 & , \text{αλλού} \end{cases}$$

Είναι σ.π.π. αφού $\frac{1}{6} + \frac{1}{2} + \frac{1}{3} = 1$

Μέσες Τιμές - Διακυμάνσεις - Ροπογεννήτριες Ειδικών Κατανομών

Κατανομή	$E(X)$	$Var(X)$	$M_X(t)$
$B(n,p)$	np	npq, $q=1-p$	$(pe^t+1-p)^n$ $t\in\mathbb{R}$
$Hg(M,N,n)$	$\frac{nM}{M+N}$	$\frac{nMN}{(M+N)^2}\cdot\frac{M+N-n}{M+N-1}$	Δεν υπάρχει
$Geo(p)$	$\frac{1}{p}$	$\frac{q}{p^2}$, $q=1-p$	$\frac{pe^t}{1-qe^t}$, $t<-\log q$
$NB(k,p)$	$\frac{k}{p}$	$\frac{kq}{p^2}$, $q=1-p$	$\frac{p^k e^{kt}}{(1-qe^t)^k}$, $t<-\log q$
$U(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt}-e^{at}}{t(b-a)}$ $t\in\mathbb{R}$
$Εκθ(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda-t}$ $t<\lambda$
$G(a,b)$	ab	ab^2	$\frac{1}{(1-bt)^a}$ $t<\frac{1}{b}$
$Beta(a,b)$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)(a+b+1)}$	Δεν υπάρχει
$N(\mu,\sigma^2)$	μ	σ^2	$e^{\mu t+\frac{1}{2}\sigma^2t^2}$ $t\in\mathbb{R}$

Ανισότητες Markov - Chebyshev

ΠΡΟΤΑΣΗ

α) Ανισότητα Markov:

Έστω X μια αρνητική τ.μ. με μέση τιμή $E(X)$. Τότε $\forall$ σταθερά $a>0$

$$P(X\geq a)\leq\frac{E(X)}{a} \quad ή \quad P(X<a)\geq 1-\frac{E(X)}{a}$$

β) Ανισότητα Chebyshev:

Έστω τ.μ. X με μέση τιμή $\mu=E(X)$ και $\sigma^2=Var(X)$

Τότε $\forall$ σταθερά $a>0$

$$P(|X-\mu|\geq a)\leq\frac{Var(X)}{a^2} \quad ή \quad P(|X-\mu|<a)\geq 1-\frac{Var(X)}{a^2}$$

Απόδειξη

α Ορίζω την τ.μ $Y = \begin{cases} a & αν\ X \geq a \\ 0 & αν\ X < a \end{cases}$

Παρατηρώ ότι $Y \leq X \Rightarrow E(Y) \leq E(X)$

$\Rightarrow aP(Y=a) + 0\,P(Y=0) \leq E(X) \Rightarrow a\,P(X \geq a) \leq E(X) \Rightarrow$
$\Rightarrow P(X \geq a) \leq \frac{E(X)}{a}$

β. $P(|X-\mu| \geq a) = P((X-\mu)^2 \geq a^2) \overset{markov}{\leq} \frac{E(X-\mu)^2}{a^2} = \frac{Var(X)}{a^2}$

Πόρισμα

Για την τ.μ X ισχύουν

1) $P(|X| \geq a) \leq \frac{E(|X|)}{a}$, $a>0$

2) Αν $X \geq 0$ τότε $P(X^k \geq a) \leq \frac{E(X^k)}{a} = \frac{\mu_k}{a}$

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

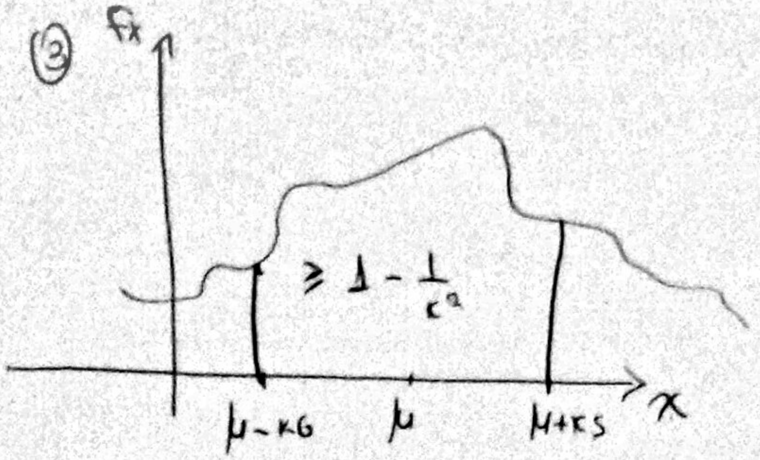
4) $P(X \geq a) \leq \inf_{t \geq 0} \{e^{-at} m_X(t)\}$

Απόδειξη 4)

$P(X \geq a) \overset{t \geq 0}{=} P(tX \geq ta) = P(e^{tX} \geq e^{ta}) \overset{markov}{\leq} \frac{E(e^{tX})}{e^{ta}} \Rightarrow$

$\Rightarrow P(X \leq a) \leq e^{-ta} m_X(t) \quad \forall t \geq 0$

3



Παράδειγμα

Έστω $X \sim B(16, \frac{1}{2})$ $\qquad P(4<X<12) = ?$

Να βρεθεί ένα κάτω φράγμα για την πιθανότητα αυτή.

<u>Λύση</u>

Αφού $X \sim B(16, \frac{1}{2})$ $\qquad p_X(x) = \binom{16}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{16-x}$ $\quad x=0,\dots,16$

$E(X) = np = 8$, $\quad Var(X) = npq = 4$, $\quad \sigma = \sqrt{4} = 2$

$$P(4<X<12) = \sum_{4<x<12} p_X(x) = \sum_{x=5}^{11} \binom{16}{x}\left(\frac{1}{2}\right)^{16} = 0,9234$$

$$P(4<X<12) = P(4-8<X-8<12-8) = P(-4<X-8<4) =$$

$$= P(|X-8|<4) = P(|X-\mu| < \overset{k}{2}\cdot\overset{\sigma}{2}) \geq 1 - \frac{1}{k^2} = 1 - \frac{1}{2^2} = 1 - \frac{1}{4} = \frac{3}{4} = 0,71$$

Παράδειγμα (Σπίτι)

Ο μέσος χρόνος σε λεπτά στον οποίο ένας οδηγός αγώνων Formula I διανύει μια πίστα αγώνων είναι 3 με τυπική απόκλιση $\frac{1}{2}$. Είναι ρεαλιστικό να πιστεύει ότι σε μια προσπάθεια του να διασχίσει την πίστα θα πετύχει χρόνο μεταξύ ~~[illegible]~~ 1 και 5 λεπτών?