6/11/2019

• Ροπογεννήτρια τ.δ. $X=(X_1,X_2,\dots X_n)$

$$M_{X_1,X_2,\dots X_n}(t_1,\dots,t_n) = E(e^{t^T X}) = E(e^{t_1X_1+\dots+t_nX_n})$$

υπο την προϋπόθεση ότι η αναμενόμενη τιμή υπάρχει

• Χαρακτηριστική Συνάρτηση

$$\Phi_{X_1,\dots X_n}(t_1,\dots t_n) = E(e^{it^TX})$$

$t^T=(t_1,\dots,t_n)$ $\quad X=\begin{pmatrix} X_1\\ X_2\\ \vdots\\ X_n\end{pmatrix}$

Έστω $X_{n\times 1}$, $Y_{k\times 1}$ τ. διανύσματα, $A_{m\times n}$ και $B_{\ell\times k}$ έτσι ώστε
$U=(AX)$, $V=(BY)$

$E(U)=E(AX)=AE(X)$ $\qquad Var(U)=Var(\underbrace{AX}_{m\times 1}) = \underset{m\times n}{A}\ \underset{n\times n}{Var(X)}\ \underset{n\times m}{A^T}$

$E(V)=E(BY)=BE(Y)$

Όταν U είναι $m\times 1$ το $Var(U)$ τι διάσταση έχει?

Απαντ.: $m\times m$

$$Cov(A\underset{}{\overset{n\times 1}{X}}, B\overset{k\times 1}{Y}) = \underset{m\times n}{A}\ \underset{n\times k}{Cov(X,Y)}\ \underset{k\times l}{B^T}$$

διότι $Cov(X,Y) = \begin{bmatrix} Cov(X_1,Y_1) & Cov(X_1,Y_2) & \cdots \\ Cov(X_2,Y_1) & \cdots & \\ \vdots & \vdots & \\ \vdots & \vdots & \end{bmatrix} \rightarrow n\times k$

• Πολυωνυμική κατανομή: (γενίκευση της διωνυμικής)

Έστω ένα τυχαίο πείραμα σε κάθε επανάληψη του οποίου μπορεί να εμφανιστεί ένα μόνο από τα k το πλήθος διαφορετικά ενδεχόμενα $E_1, E_2, \ldots, E_k$ (ουσιαστικά αποτελούν διαμέριση του $S.X$) Επιπλέον, υποθέτουμε ότι οι επαναλήψεις του τ.πειράματος είναι ανεξάρτητες, και οι πιθανότητες $P(E_i) = P_i$ παραμένουν αμετάβλητες σε κάθε επανάληψη του τ.π. με $P_1 + P_2 + \ldots + P_k = 1$.

• Ορισμός

Έστω $X_1, X_2, \ldots, X_k$ οι τ.μ. που παριστάνουν τον αριθμό των εμφανίσεων των ενδεχομένων $E_1, E_2, \ldots, E_k$ αντίστοιχα σε n-ανεξάρτητες επαναλήψεις του πολυωνυμικού τ.π.. Το τ.δ.: $X = (X_1, X_2, \ldots, X_k)$ λέμε ότι ακολουθεί πολυωνυμική κατανομή με παραμέτρους $n, P_1, P_2, \ldots, P_k$ (*) και συνάρτηση πιθανότητας που δίνεται από τη σχέση $P_X(x) = P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k)$ όπου $x_1 + x_2 + x_3 + \ldots + x_k = n$.

• Συστηματικό / Απόδειξη

□ □ □ ... □ : n-δοκιμές

Θέλω στις n-δοκιμές να έχω x_1 επιτυχίες E_1. Με πόσους τρόπους μπορούν να γίνουν αυτές οι επιτυχίες?
$\binom{n}{x_1}$ και καθένας έχει πιθανότητα $p_1^{x_1}$

Θέλω στις υπόλοιπες $n-x_1$ επαναλήψεις να έχω x_2 φορές εμφάνισης E_2. Πόσοι τρόποι: $\binom{n-x_1}{x_2}$ και πιθανότητα $p_2^{x_2}$.

Θέλω στις υπόλοιπες $n-x_1-x_2$ επαναλήψεις να έχω x_3 φορές εμφάνισης E_3. Πόσοι τρόποι? $\binom{n-x_1-x_2}{x_3}$ και πιθανότητα $p_3^{x_3}$

⋮ Επαγωγικά έχω ότι:

$$P(X_1=x_1, X_2=x_2, \dots, X_k=x_k) =$$

$$= \binom{n}{x_1}\binom{n-x_1}{x_2}\binom{n-x_1-x_2}{x_3}\cdots\binom{n-x_1-x_2-\dots-x_{k-1}}{x_k}\cdot p_1^{x_1}p_2^{x_2}\cdots p_k^{x_k} =$$

$$= \frac{n!}{x_1!\cancel{(n-x_1)!}}\cdot\frac{\cancel{(n-x_1)!}}{x_2!\cancel{(n-x_1-x_2)!}}\cdots\frac{\cancel{(n-x_1-x_2-\dots-x_{k-1})!}}{x_k!\underbrace{(n-x_1-x_2-\dots-x_k)!}_{=0!=1}}\cdot p_1^{x_1}p_2^{x_2}\cdots p_k^{x_k} =$$

$$= \frac{n!}{x_1!x_2!\cdots x_k!}p_1^{x_1}\cdot p_2^{x_2}\cdots p_k^{x_k} \quad \text{με} \quad \begin{array}{l} x_1+x_2+\dots+x_k=n \\ p_1+p_2+\dots+p_k=1 \end{array}$$

Άρα με τους περιορισμούς: $\begin{cases} x_1+x_2+\dots+x_k=n \\ p_1+p_2+\dots+p_k=1 \end{cases}$ έχω:

$$P_X(x) = \frac{n!}{x_1!x_2!\cdots x_k!}p_1^{x_1}p_2^{x_2}\cdots p_k^{x_k} \quad (1)$$

(*) Εναλλακτικά δίνεται η σχέση:

(2) $$P_X(x)=\frac{n!\,p_1^{x_1}p_2^{x_2}\cdots p_{k-1}^{x_{k-1}}(1-p_1-p_2-\dots-p_{k-1})^{n-x_1-x_2-\dots-x_{k-1}}}{x_1!\,x_2!\cdots x_{k-1}!\,(n-x_1-x_2-\dots-x_{k-1})!}$$

[Η διαφορά είναι ότι εδώ δε χρειάζεται να βάλω τους περιορισμούς]

$$X\sim M(n,p_1,p_2,\dots,p_k) \longrightarrow P_X(x) \overset{(1)}{=} \frac{n!}{x_1!\,x_2!\cdots x_k!}\,p_1^{x_1}p_2^{x_2}\cdots p_k^{x_k}$$

με $p_1+p_2+\dots+p_k=1$ και $x_1+x_2+\dots+x_k=n$.

ή

$$X\sim M(n,p_1,p_2,\dots,p_{k-1}) \longrightarrow P_X(x) \overset{(2)}{=} \frac{n!\,p_1^{x_1}p_2^{x_2}\cdots p_{k-1}^{x_{k-1}}(1-p_1-\dots-p_{k-1})^{n-x_1-\dots-x_{k-1}}}{x_1!\,x_2!\cdots x_{k-1}!\,(n-x_1-\dots-x_{k-1})!}$$

• **ΠΡΟΤΑΣΗ**

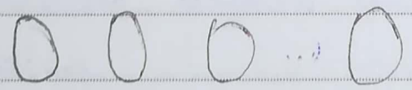
Έστω $X\sim M(n,p_1,p_2,\dots,p_k)$ $\quad p_1+p_2+\dots+p_k=1$

Τότε:

(1) Η περιθώρια κατανομή του X_i, $i=1,\dots,k$ είναι διωνυμική με παραμέτρους n, p_i

1ος Τρόπος: X_i τ.μ. παριστάνει αριθμό εμφανίσεων E_i στις n-δοκιμές του τ. πειράματος: $P(X_i=x_i) \underset{B(n,p_i)}{=} \binom{n}{x_i}p_i^{x_i}(1-p_i)^{n-x_i}$

2ος Τρόπος



n-δοκιμές

Θέλω x_i φορές εμφάνιση του E_i. Πόσοι τρόποι?

$\binom{n}{x_i}$ με πιθανότητα $p_i^{x_i}$.

Θέλω στις υπόλοιπες $n-x_i$ δοκιμές να μην εμφανιστεί E_i.

Πόσοι τρόποι; $\underbrace{\binom{n-x_i}{n-x_i}}_{=1}$ με πιθανότητα $(1-p_i)^{n-x_i}$

ένας τρόπος δηλ.

Αν τα πολλαπλασιάσω προκύπτει πάλι ο ίδιος τύπος

3ος Τρόπος

$$P(X_i = x_i) = \sum_{\substack{x_1 \dots x_{i-1} \\ x_{i+1}, \dots x_k}} \frac{n! \cdot p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}}{x_1! \, x_2! \cdots x_k!}$$

Η/Ω Υπόδειξη: να φτιάξω μέσα στο άθροισμα μία γνωστή κατανομή έτσι ώστε $\sum \dots = 1$

(2) $X \sim M(n, p_1, p_2, \dots p_k)$: $p_1 + p_2 + \dots + p_k = 1$

Τότε η περιθώρια κατανομή των:

$(X_i, X_j) \sim M(n, p_i, p_j, 1-p_i-p_j)$ με $p_i + p_j + (1-p_i-p_j) = 1$

$i = 1, \dots, k$

$j = 1, \dots, k$

$i \neq j$

1ος Τρόπος

$$\underbrace{\sum_{\substack{x_1, \dots, x_{i-1} \\ x_{i+1} \dots x_{j-1}, x_{j+1} \dots x_k}}}_{} \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \cdot p_2^{x_2} \cdots p_k^{x_k}$$

$\Leftrightarrow (x_l, \ l \neq i, \ l = j, \ l = 1, \dots k)$

2ος Τρόπος

O O … O n- δοκιμές

Θέλω x_i φορές E_i. Πόσοι τρόποι; $\binom{n}{x_i}$ με πιθανότητα $p_i^{x_i}$

Θέλω x_j φορές E_j: $\binom{n-x_i}{x_j}$ με πιθανότητες $p_j^{x_j}$

Στις υπόλοιπες: $\binom{n-x_i-x_j}{x_3}$ με πιθανότητα $(1-p_i-p_j)^{x_3}$ με

$x_i+x_j+x_3=n$

Άρα: $\binom{n}{x_i}\binom{n-x_i}{x_j}\binom{n-x_i-x_j}{x_3} p_i^{x_i} p_j^{x_j} (1-p_i-p_j)^{n-x_i-x_j} =$

$$= \frac{n!}{x_i!\cancel{(n-x_i)!}} \cdot \frac{\cancel{(n-x_i)!}}{x_j!\cancel{(n-x_i-x_j)!}} \cdot \frac{\cancel{(n-x_i-x_j)!}}{x_3!\underbrace{\cancel{(n-x_i-x_j-x_3)!}}_{1}} p_i^{x_i} p_j^{x_j} (1-p_i-p_j)^{n-x_i-x_j}$$

$$= \frac{n!}{x_i!\,x_j!\,x_3!} p_i^{x_i} p_j^{x_j} (1-p_i-p_j)^{n-x_i-x_j}$$

- Πολυδιάστατη κανονική κατανομή

Η μονοδιάστατη κανονική κατανομή όπως είδαμε τόσο στην Εισαγωγή στις Πιθανότητες, όσο και στη Στατιστική, είναι η πιο ευρέως χρησιμοποιούμενη μονοδιάστατη κατανομή, τόσο γιατί χρησιμοποιείται για τη μοντελοποίηση πλήθους πραγματικών τυχαίων φαινομένων σε διάφορους επιστημονικούς κλάδους, όσο και γιατί αποτελεί προσέγγιση του αθροίσματος ανεξάρτητων τυχαίων μεταβλητών (βλ. Κεντρικό Οριακό Θεώρημα)

(*) ΠΑΡΑΤΗΡΗΣΗ: Παρ' όλο που τα διανύσματα τα γράφω ως γραμμές κανονικά είναι ΣΤΗΛΕΣ !!!

Τα παραπάνω οδηγούν στη γενίκευση της και στην εισαγωγή της k-διάστατης κανονικής κατανομής.

Το τυχαίο διάνυσμα(*) $X=(X_1,X_2,\dots,X_k)$ λέμε ακολουθεί k-διάστατη κανονική κατανομή αν η σ.π.π. του δίνεται από τη σχέση.

$$f_X(x) = (2\pi)^{-\frac{k}{2}} |\Sigma|^{-1/2} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}$$ όπου:

$x \in \mathbb{R}^k$, $\mu \in \mathbb{R}^k$ Σ: Θετικά ορισμένος

$$\underset{k\times k}{\Sigma} = Var\, x \qquad \underset{k\times 1}{\mu} = E(x)$$

Ειδικότερα: για $k=1$

$$(2\pi)^{1/2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Συμβολισμός: $x \sim N_k(\mu, \Sigma)$

μ, Σ οι παράμετροι της κανονικής k-διάστατης κατανομής

• ΠΡΟΤΑΣΗ:

$x \sim N_k(\mu, \Sigma)$ αν-ν κάθε γραμμική συνάρτηση $\alpha^T x$ έχει μονοδιάστατη κανονική

ΠΑΡΑΔΕΙΓΜΑ Κατανόησης:

$x = (x_1, x_2, x_3)$. Με ρωτάει κάποιος: Ποια η κατανομή του x_1?

του (x_1, x_2)?

του $x_1 + 7x_2 + 8x_3$?

• Για το x_1

$$x_1 = (1,0,0)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Τότε σύμφωνα με την προηγούμενη πρόταση έχω $x_1 \sim N_1(E(x_1), Var(x_1))$

• Για το $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

• Πρόταση:

Αν $X \sim N_k(\mu, \Sigma)$ τότε:

$$\underset{m\times k}{A}\underset{k\times 1}{X} + \underset{m\times 1}{b} \sim N_m(A\mu + b, A\Sigma A^T)$$

$$E(AX+b) = E(AX) + E(b) = AE(X) + b$$

• Για το $X_1 + 7X_2 + 8X_3$

$$\underbrace{X_1 + 7X_2 + 8X_3}_{1\times 1} = \begin{bmatrix} 1 & 7 & 8 \end{bmatrix}_{1\times 3} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}_{3\times 1}$$

Άρα από πρόταση $X_1 + 7X_2 + 8X_3 \sim N_1$

$$E(X_1 + 7X_2 + 8X_3) = \begin{pmatrix} 1 & 7 & 8 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}$$

$$Var(X_1 + 7X_2 + 8X_3) = \cancel{Var(X_1) + 7^2 Var(X_2) + 8^2 Var(X_3)}$$

αυτό ισχύει μόνο όταν είναι ανεξάρτητα!

$$Var(X_1 + 7X_2 + 8X_3) = \underset{1\times 3}{\begin{pmatrix} 1 & 7 & 8 \end{pmatrix}} \underset{3\times 3}{\Sigma} \begin{pmatrix} 1 \\ 7 \\ 8 \end{pmatrix}_{3\times 1}$$

• Συμπέρασμα:

Αν $X_{k\times 1} \sim N_k(\mu, \Sigma)$

$\Rightarrow$ κάθε X_i ακολουθεί κανονική κατανομή

$$X_i = \underset{1\times k}{(0\ 0 \ldots 0\ \underset{i\text{ θέση}}{1}\ 0 \ldots 0)} \begin{pmatrix} X_1 \\ \vdots \\ X_k \end{pmatrix}_{k\times 1}$$

ΠΡΟΣΟΧΗ!!!

Το αντίστροφο δεν ισχύει! (Θα δούμε αντιπαράδειγμα σε επόμενο μάθημα)

Δηλαδή αν
$$\left.\begin{array}{l} X_1 \text{ μονοδ. καν.} \\ X_2 \text{ μονοδ. καν.} \\ \vdots \\ X_k \text{ μονοδ. καν.} \end{array}\right\} \not\Rightarrow (X_1, X_2, \ldots, X_k) \text{ k-διάστατη καν.}$$

• Ανεξαρτησία

• Ορισμός

Οι τ.μ. $X_1, \ldots, X_k$ λέγονται στοχαστικά ανεξάρτητες αν για κάθε συλλογή από υποσύνολα $B_1, B_2, \ldots, B_k$ του $\mathbb{R}$ ισχύει:

$$P(X_1 \in B_1, X_2 \in B_2, \ldots, X_k \in B_k) = \prod_{i=1}^{k} P(X_i \in B_i)$$

• 1ο Κριτήριο Ανεξαρτησίας

Οι τ.μ. $X_1, X_2, \ldots, X_k$ είναι ανεξάρτητες $\Leftrightarrow$
$$F_{X_1 \ldots X_k}(x_1, \ldots, x_k) = \prod_{i=1}^{k} F_{X_i}(x_i)$$

• 2ο Κριτήριο Ανεξαρτησίας

Οι $X_1, X_2, \ldots, X_k$ είναι ανεξάρτητες αν η σ.π.π ή σ.π της:

$$f_{X_1 \ldots X_k}(x_1 \ldots x_k) = \prod_{i=1}^{k} f_{X_i}(x_i)$$

• 3^o Κριτήριο Ανεξαρτησίας

Οι τ.μ $X_1, X_2, \dots, X_k$ είναι ανεξάρτητες $\Leftrightarrow$

$\forall (t_1, \dots, t_k) \in (-h_1, h_1) \times \dots \times (-h_k, h_k)$ με $h_1, h_2, \dots, h_k > 0$ ισχύει:

$$m_{X_1 \dots X_k}(t_1, \dots, t_k) = \prod_{i=1}^{k} m_{X_i}(t_i)$$

• Παρατηρηση / Προταση :

α) Αν οι τ.μ. $X_1, X_2, \dots, X_k$ είναι ανεξάρτητες, τότε οποιεσδήποτε ℓ - απο αυτές με $\ell \leq k$ είναι ανεξάρτητες.

β) Η ανα ζεύγη ανεξαρτησία δεν συνεπάγεται την "ολική" ανεξαρτησία.

• Πορισμα :

Έστω $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ → ισχύει και για περισσότερες διαστάσεις

Αν οι τ.μ. X_1, X_2 είναι ανεξάρτητες τότε:

$$F_{X_1|X_2}(x_1|x_2) = F_{X_1}(x_1)$$

$$f_{X_1|X_2}(x_1|x_2) = f_{X_1}(x_1)$$

• Θεωρημα :

Αν $X_1, X_2, \dots, X_k$ k - ανεξάρτητες τ.μ και $Y_1 = h_1(X_1)$, $Y_2 = h_2(X_2), \dots$
$\dots, Y_k = h_k(X_k)$ μετασχηματισμοί / συναρτήσεις των $X_1, X_2, \dots, X_k$

τότε: $Y_1, Y_2, \dots, Y_k$: ανεξάρτητες

ΘΕΩΡΗΜΑ: Μέση Τιμή Γινομένου Ανεξάρτητων τ.μ.

Έστω $X_1, \dots, X_k$ ανεξάρτητες τ.μ. και h_i: κ-συναρτήσεις αυτών, με $E(h_i(X_i)) < \infty$

Τότε: $E(h_1(X_1) h_2(X_2) \cdots h_k(X_k)) = \prod_{i=1}^{k} E(h_i(X_i))$

$$= \int \cdots \int h_1(x_1) \cdots h_k(x_k) f_{X_1, \dots, X_k}(x_1, \dots, x_k)\, dx_1 \cdots dx_k$$

$$= \int \cdots \int h_1(x_1) f_{X_1}(x_1)\, h_2(x_2) f_{X_2}(x_2) \cdots h_k(x_k) f_{X_k}(x_k)\, dx_1 \cdots dx_k$$