22/02/2019

Στατιστική Συμπερασματολογία.

Διάλεξη 1η

Βιβλιογραφία: Παπαϊωάννου.

Αλλαγή Μεταβλητών:

▸ Μέθοδος Μετασχηματισμού: Έστω συνεχής τ.μ. X με γνωστή σ.π.π. f_X και τιμές στο $I \subseteq \mathbb{R}$. Θεωρούμε τον μετασχηματισμό $Y=g(X)$ με $g: I \to g(I)$,

$g(I)=\{y: y=g(x), x\in I\}$

Αν: i) Η g είναι "1-1"

ii) Η $\frac{dg^{-1}(y)}{dy} \neq 0$ και συνεχής.

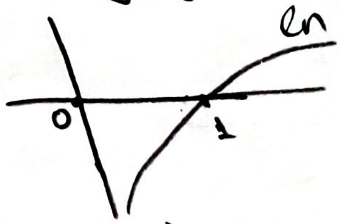
Τότε η τ.μ. $Y=g(X)$ είναι συνεχής και η σ.π.π της Y είναι:

$$f_Y(y)=f_X(g^{-1}(y))\left|\frac{d}{dy}g^{-1}(y)\right| \quad , y\in g(I)$$

π.χ.: Έστω τ.μ X με κατανομή Pareto με σ.π.π. $f_X(x)=\theta x^{-\theta-1}, x>1, \theta>0$.
Να βρεθεί η κατανομή της τ.μ. $Y=\ln X$.

Λύση

Θεωρώ τον μετασχηματισμό $y=g(x)=\ln x$, $x\in I=(1,\infty)$
$y\in g(I)=(0,\infty)$, οπότε



θα ελέγξω εάν ισχύουν οι προϋποθέσεις του παραπάνω θεωρήματος.

i) Προφανώς $\left\langle\begin{array}{l}\text{Αν } x_1,x_2 \text{ με } g(x_1)=g(x_2) \Rightarrow x_1=x_2 \\ \text{Η } g \text{ γνησίως μονότονη, δηλαδή } \frac{d}{dx}g(x) \left\langle \begin{array}{l} >0 \\ \text{ή} \\ <0 \end{array}\right. \end{array}\right.$

Άρα η g είναι "1-1" $\Rightarrow \exists$ ο g^{-1} και είναι η λύση ως προς x της $y=g(x)=\ln x$, δηλαδή $x=g^{-1}(y)=e^y$

ii) $\frac{d}{dy}g^{-1}(y)=\frac{d}{dy}e^y=e^y\neq 0$ και συνεχής $\forall y>0$

Άρα: ικανοποιούνται τα (i) και (ii)

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| = f_X(e^y) \cdot |e^y| = \theta (e^y)^{-\theta-1} e^y$$
$$= \theta e^{-\theta y} \quad , y>0$$

$Y = \ln X \sim Exp(\theta)$

Π.χ: Αν $X \sim Gamma(k,\lambda)$ τότε η τ.μ $Y = \frac{2}{\lambda} \cdot X \sim X^2_{2k}$

(*) Υπενθ: $X^2_v = X_1^2 + \dots + X_v^2$, X_i ανεξ $\sim N(0,1)$

Απόδ

Αφού $X \sim Gamma(k,\lambda)$ τότε η σ.π.π. της τ.μ X είναι:

$$f_X(x) = \frac{1}{\lambda^k \Gamma(k)} x^{k-1} e^{-x/\lambda} \quad , x>0 \, , \, k,\lambda>0$$

Θεωρώ τον μετασχηματισμό $y = g(x) = \frac{2}{\lambda} x$, $x>0$.

Προφανώς $y>0$

(i) Προφανώς η g είναι 1-1 άρα ∃ ο αντίστροφος μετασχ. g^{-1} τ.ω $x = g^{-1}(y) = \frac{\lambda}{2} y$

(ii) $\frac{d}{dy} g^{-1}(y) = \frac{\lambda}{2} \neq 0$ και συνεχής.

Επομένως η σ.π.π. της $Y = \frac{2}{\lambda} \cdot X$ θα είναι η εξής:

$$f_Y(y) = f_X\left(\frac{\lambda}{2} \cdot y\right) \left|\frac{\lambda}{2}\right| = \frac{1}{\lambda^k \Gamma(k)} \left(\frac{\lambda}{2} \cdot y\right)^{k-1} e^{-\frac{\frac{\lambda}{2} y}{\lambda}} \cdot \left|\frac{\lambda}{2}\right|$$

$$= \frac{1}{\lambda^k \Gamma(k)} \left(\frac{\lambda}{2} y\right)^{k-1} e^{-y/2} \left|\frac{\lambda}{2}\right| =$$

$$= \frac{1}{2^{2k/2} \Gamma\left(\frac{2k}{2}\right)} y^{\frac{2k}{2}-1} e^{-y/2} \quad , y>0 \quad (1)$$

Αλλά ισχύει ότι $f_{X^2_v}(x) = \frac{1}{2^{v/2} \Gamma\left(\frac{v}{2}\right)} x^{v/2-1} e^{-x/2}$, $x>0$ (2)

από την σύγκριση των (1) κ' (2)

$Y \sim X^2_{2k}$ και έπεται το ζητούμενο

Σημείωση: Αν $\lambda=2$ τότε $G(x,2)\equiv X^2_{2x}$

~o~o~o~o

Μέθοδος α.σ.κ:

Παράδειγμα: Έστω συνεχής τ.μ. X με (γνησίως) αύξουσα α.σ.κ. $F_X(x), x\in\mathbb{R}$. Αν $Y=F_X(X)$ τότε $Y\sim U(0,1)$.

Απόδ.

Είναι $F_Y(y) \overset{\text{ορ}}{=} P(Y\le y)=P(F_X(X)\le y)$

$= P(X\le F_X^{-1}(y))$

$\overset{\text{ορ}}{=} F_X(F_X^{-1}(y))=y \ , 0\le y\le 1$

Επομένως $F_Y(y)=y, \ 0\le y\le 1$

Αν $W\sim U(0,1)$ τότε

$$F_W(w)=\begin{cases}0 & , w<0\\ \frac{w-0}{1-0}=w & , 0\le w\le 1\\ 1 & , w>1.\end{cases}$$

Έτσι $Y=F_X(X)\sim U(0,1)$

~o~o~o~o~o~o~o~o

Κατανομή μεγίστου και ελαχίστου:

Έστω τ.μ. $X_1, X_2, \ldots, X_n$ ανεξάρτητες και ισόνομες, δηλαδή όλες οι $X_i, i=1,\ldots,n$ έχουν την ίδια κατανομή f_X ή F_X.

Έστω $X_{(n)}=\max\{X_1,\ldots,X_n\}$ και

$X_{(1)}=\min\{X_1,\ldots,X_n\}$.

Τότε $f_{X_{(n)}}(x)=n[F_X(x)]^{n-1} f_X(x), \ x\in\mathbb{R}$

$f_{X_{(1)}}(x)=n[1-F_X(x)]^{n-1} f_X(x), \ x\in\mathbb{R}$

(4)

Απόδ.

Κατανομή μεγίστου: $F_{X_{(n)}}(x) \overset{\text{ορ}}{=} P(X_{(n)} \le x)$

$= P(\max\{X_1, \dots, X_n\} \le x) = P(X_i \le x,\ \forall i = 1, \dots, n)$

$= P(X_1 \le x, X_2 \le x, \dots, X_n \le x)$

$\overset{\text{ανεξ.}}{=} \prod_{i=1}^{n} P(X_i \le x)$

Δηλαδή $F_{X_{(n)}}(x) = \prod_{i=1}^{n} F_{X_i}(x) = \prod_{i=1}^{n} F_X(x) = [F_X(x)]^n$.

Άρα $F_{X_{(n)}}(x) = [F_X(x)]^n$, $x \in \mathbb{R}$. Άρα $f_{X_{(n)}}(x) = \frac{d}{dx} F_{X_{(n)}}(x)$

$= \frac{d}{dx}[F_X(x)]^n = n[F_X(x)]^{n-1} \frac{d}{dx} F_X(x) = n[F_X(x)]^{n-1} f_X(x)$

Κατανομή ελαχίστου: $F_{X_{(1)}}(x) \overset{\text{ορ.}}{=} P(X_{(1)} \le x)$

$= P(\min\{X_1, \dots, X_n\} \le x) = 1 - P(\min\{X_1, \dots, X_n\} > x)$

$= 1 - P(X_i > x,\ \forall i = 1, \dots, n) = 1 - P(X_1 > x, X_2 > x, \dots, X_n > x)$

$\overset{\text{ανεξ. των } X_i}{\Longrightarrow} F_{X_{(1)}}(x) = 1 - \prod_{i=1}^{n} P(X_i > x) = 1 - \prod_{i=1}^{n} [1 - P(X_i \le x)]$

$= 1 - \prod_{i=1}^{n} [1 - F_{X_i}(x)]$

$\overset{X_i \text{ ισόνομες},\ F_{X_i} = F_X}{=} 1 - \prod_{i=1}^{n} [1 - F_X(x)]$

$\Rightarrow F_{X_{(1)}}(x) = 1 - [1 - F_X(x)]^n$, $x \in \mathbb{R}$

$f_{X_{(1)}}(x) = \frac{d}{dx} F_{X_{(1)}}(x) = \frac{d}{dx}\{1 - [1 - F_X(x)]^n\}$

$= -n[1 - F_X(x)]^{n-1} \left(\frac{d}{dx}(-F_X(x))\right)$

$= n[1 - F_X(x)]^{n-1} f_X(x)$, $x \in \mathbb{R}$.

Κατανομή αθροίσματος ανεξαρτήτων τ.μ.:

Έστω ανεξάρτητες (και ισόνομες) τ.μ. $X_1, \dots, X_n$.

Πρόβλημα: Η εύρεση κατανομής $T = \sum_{i=1}^{n} X_i$. Για την εύρεση κατανομής T αξιοποιείται αποκλειστικά η ροπογεννήτρια $m_T(t) \overset{\text{ορ.}}{=} E(e^{tT}) = E(e^{t \cdot \sum_{i=1}^{n} X_i})$

$$= E(e^{tX_1 + tX_2 + \dots + tX_n}) = E(e^{tX_1} e^{tX_2} \dots e^{tX_n})$$

$$\underset{X_i \text{ ανεξ.}}{=} E(e^{tX_1}) E(e^{tX_2}) \dots E(e^{tX_n})$$

$$\underset{\text{ορισμός ροπογ.}}{=} m_{X_1}(t) \dots m_{X_n}(t) \Rightarrow m_T(t) = \prod_{i=1}^{n} m_{X_i}(t) =$$

$$\underset{\text{αν } X_i \text{ ισόνομες}}{=} [m_X(t)]^n \text{ με } m_X \text{ την}$$

κοινή ροπογεννήτρια των $X_1, \dots, X_n$.

Ανάλογα και γενικότερα αν $T = \sum_{i=1}^{n} a_i X_i$. Τότε

$$\boxed{m_T(t) = \prod_{i=1}^{n} m_{X_i}(a_i t)}$$

~o~o~o

Παράδειγμα: Έστω ανεξ. τ.μ. $X_1, \dots, X_n$ με κατανομή Poisson(θ). Να βρεθεί η κατανομή $T = \sum_{i=1}^{n} X_i$.

Λύση

Εφόσον έχω κατανομή Poisson και την ίδια παράμετρο άρα εφόσον έχω ακριβώς την ίδια συναρτησιακή έκφραση είναι ισόνομες.

Θυμόμαστε ότι αν $W \sim P(\theta)$ τότε $m_W(t) = e^{\theta(e^t - 1)}$, $t \in \mathbb{R}$

Με βάση το προηγούμενο και επειδή οι X_i, $i = 1, \dots, n$ ανεξ. και ισόνομες $P(\theta)$

$$m_T(t) = [m_X(t)]^n = [e^{\theta(e^t - 1)}]^n \Rightarrow m_T(t) = e^{n\theta(e^t - 1)}, \ t \in \mathbb{R}$$

Το T έχει ροπογεννήτρια την μορφή της ροπογ. Poisson με παράμ. $n\theta$. Από θεώρημα μοναδικότ. ροπογ.

$T \sim P(n\theta)$

Παράδειγμα 2:

Έστω ανεξ. $X_1,\dots,X_n$ με κατανομή $N(\mu,\sigma^2)$. Να βρεθεί η

κατανομή του $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ $\left(\bar{X}\sim N(\mu,\frac{\sigma^2}{n})\right)$

Λύση

Αν $W\sim N(\mu,\sigma^2)$ τότε $m_W(t)=e^{t\cdot\mu+\frac{1}{2}t^2\sigma^2}$, $t\in\mathbb{R}$

$$m_{\bar{X}}(t)\overset{op}{=}E(e^{t\bar{X}})=E\left(e^{t\cdot\frac{1}{n}\sum_{i=1}^{n}X_i}\right)=E\left(e^{t/n\cdot X_1+\frac{t}{n}\cdot X_2+\dots+\frac{t}{n}X_n}\right)$$

$$=E\left(e^{t/n\cdot X_1}\,e^{\frac{t}{n}\cdot X_2}\cdots e^{\frac{t}{n}\cdot X_n}\right)\xlongequal{X_i\text{ ανεξ.}}\prod_{i=1}^{n}E\left(e^{\frac{t}{n}\cdot X_i}\right)$$

$$\overset{\text{ορ. ροπογ.}}{=}\prod_{i=1}^{n}m_{X_i}\left(\frac{t}{n}\right)=\prod_{i=1}^{n}e^{\frac{t}{n}\cdot\mu+\frac{1}{2}\left(\frac{t}{n}\right)^2\sigma^2}$$

$$=\prod_{i=1}^{n}e^{\frac{t}{n}\mu+\frac{1}{2}\frac{t^2}{n^2}\cdot\sigma^2}$$

$$=e^{n\left[\frac{t}{n}\mu+\frac{1}{2}\frac{t^2}{n^2}\sigma^2\right]}=\underbrace{e^{t\mu+\frac{1}{2}t^2\frac{\sigma^2}{n}}}_{\text{ροπογεννήτρια της } N(\mu,\frac{\sigma^2}{n})}$$

Άρα από Θ.Μ.Ρ. $\bar{X}\sim N(\mu,\frac{\sigma^2}{n})$

HW

Παράδειγμα: Έστω X_i ανεξάρτητες και ισόνομες με $X_i\sim Εκθ(\frac{1}{\theta})$, $\forall i=1,\dots,n$. Να δειχθεί ότι $T=\sum_{i=1}^{n}X_i\sim G(n,\theta)$

Ανάλογα αν X_i ανεξ. και $X_i\sim Εκθ(\theta)$ τότε

$$T=\sum_{i=1}^{n}X_i\sim G\left(n,\frac{1}{\theta}\right)$$

18.

ραδειγμα:

στω $X_1, \dots, X_n$ ανεξάρτητες με $X_i \sim Bernoulli(p)$

ΝΔΟ: $T = \sum_{i=1}^{n} X_i \sim B(n, p)$

$Bernoulli(p) \equiv B(1, p)$

ΕΠΑΝΑΛΗΨΗ

(Ι) ΚΑΤΑΝΟΜΕΣ

Κατανομή	Συνάρτηση Πιθανότητας / Συνάρτηση Πυκνότητας Πιθανότητας	E(X)	Var(X)	Ροπογεννήτρια Συνάρτηση $m_X(t) = E(e^{tX}),\ t \in R$
Διωνυμική $B(n,p)$	$\binom{n}{x} p^x (1-p)^{n-x}, x = 0,\dots,n$	np	$np(1-p)$	$(pe^t + 1 - p)^n, \quad t \in R$
Γεωμετρική Geo(p)	$p(1-p)^{x-1}, x = 1,2,\dots$	$1/p$	$(1-p)/p^2$	$\frac{pe^t}{1-qe^t}, \quad t < -\log q$
Poisson (λ)	$\frac{e^{-\lambda}\lambda^x}{x!}, x = 0,1,\dots$	λ	λ	$\exp\{\lambda(e^t - 1)\}, \quad t \in R$
Ομοιόμορφη $U(a,\beta)$	$\frac{1}{\beta - a}, x \in (a,\beta)$	$(a+\beta)/2$	$(\beta - a)^2/12$	$\frac{e^{\beta t} - e^{\alpha t}}{t(\beta - \alpha)}, \quad t \in R$
Εκθετική (λ)	$\lambda e^{-\lambda x}, x > 0$	$1/\lambda$	$1/\lambda^2$	$\frac{\lambda}{\lambda - t}, \quad t < \lambda$
Γάμμα $G(\alpha,\beta)$	$\frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, x > 0$	$\alpha\beta$	$\alpha\beta^2$	$(1-\beta t)^{-a}, \quad t < \frac{1}{\beta}$
Χι-τετράγωνο χ^2_ν	$\frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{(\nu/2)-1} e^{-x/2}, x > 0$	ν	2ν	$(1-2t)^{-\nu/2}, \quad t < \frac{1}{2}$
Βήτα $Beta(\alpha,\beta)$	$\frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, 0 < x < 1$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	----
Κανονική $N(\mu,\sigma^2)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \ x \in R$	μ	σ^2	$\exp\left\{\mu t + \frac{1}{2}\sigma^2 t^2\right\}, \quad t \in R$

<u>Παρατήρηση</u>: (α) $\chi^2_\nu \equiv G\left(\frac{\nu}{2}, 2\right)$, (β) $G\left(1, \frac{1}{\lambda}\right) \equiv E\kappa\theta(\lambda)$, (γ) $G(k,\lambda) \equiv \frac{\lambda}{2}\chi^2_{2k}$

Μέση Τιμή

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

Διακύμανση

$$\sigma^2 = Var(X) = E(X^2) - (EX)^2$$

Ροπογεννήτρια

$$m_X(t) = E(e^{tX}) = \begin{cases} \sum\limits_x e^{tx} p_X(x), & X \text{ διακριτή} \\ \int\limits_{-\infty}^{+\infty} e^{tx} f_X(x), & X \text{ συνεχής} \end{cases}$$

Συνάρτηση Βήτα

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx\,,\ \alpha > 0, \beta > 0$$

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

Συνάρτηση Γάμμα

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1}e^{-x}dx,\ \alpha > 0$$

$$\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1),\ \alpha > 1$$

$$\Gamma(n) = (n-1)!,\ n \in N$$

Θεώρημα Μονοσημάντου Ροπογεννητριών: $X \overset{d}{=} Y \ \Leftrightarrow\ m_X(t) = m_Y(t),\ |t| < c, c \in R$

Κατανομή Μεγίστου και Ελαχίστου: Έστω τυχαίο δείγμα $X_1, \dots, X_n$ από πληθυσμό με συνάρτηση πυκνότητας f και αντίστοιχη συνάρτησης κατανομής F. Έστω $X_{(n)} = max\{X_1, \dots, X_n\}$ και $X_{(1)} = min\{X_1, \dots, X_n\}$. Τότε:

$$f_{X_{(n)}}(x) = n[F(x)]^{n-1}f(x) \text{ και } f_{X_{(1)}}(x) = n[1-F(x)]^{n-1}f(x).$$