ΑΡΙΣΤΟΤΕΛΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΟΝΙΚΗΣ

ΑΝΟΙΚΤΑ
ΑΚΑΔΗΜΑΙΚΑ
ΜΑΘΗΜΑΤΑ

Θεωρία Πιθανοτήτων & Στατιστική

Ενότητα 6η: Θεωρητικές Κατανομές Πιθανότητας για Συνεχή Τυχαία Μεταβλητή.

Γεώργιος Ζιούτας
Τμήμα Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών Α.Π.Θ.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Άδειες Χρήσης

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Θεωρητικές Κατανομές Πιθανότητας για Συνεχή Τυχαία Μεταβλητή.

Περιεχόμενα ενότητας

1. Ομοιόμορφη Κατανομή.
2. Εκθετική Κατανομή.
3. Κανονική Κατανομή (Normal ή Gauss).
   i. Τυπική Κανονική Κατανομή
   ii. Κανονική προσέγγιση στην Διωνυμική και Poisson Κατανομή.

$11^η$ Διάλεξη

Συνεχής Τυχαία Μεταβλητή X

Χρήσιμες Κατανομές

1. Ομοιόμορφη

2. Εκθετική

3. Κανονική

Ομοιόμορφη (Κατανομή)

Κατανομές για Συνεχή Τυχαία Μεταβλητή

$$f(X) = c = \frac{1}{\beta - \alpha}$$

$$E(X) = \frac{\beta + \alpha}{2}$$

$$VAR(X) = \frac{(\beta - \alpha)^2}{12}$$

Αθροιστική της *f(x)*:

$$F(X) = \frac{x - a}{\beta - \alpha}$$

Γέννηση τυχαίων αριθμών από το 0 μέχρι το 1 με ομοιόμορφη κατανομή:

$$X = K[0,1] \Rightarrow x_1, x_2, x_3, \ldots, x_n$$

Γέννηση τυχαίων αριθμών με εκθετική κατανομή:

$$f(y) = \lambda \cdot e^{-\lambda y}$$

Αθροιστική της f(y):

$$F(y) = 1 - e^{-\lambda y} \Rightarrow x_i = 1 - e^{\lambda y_i} \Rightarrow y_i = -\frac{1}{\lambda} \cdot \ln(1 - x_i) \Rightarrow y = F^{-1}(x)$$

$$x_1, x_2, x_3, \ldots, x_n \Rightarrow y_1, y_2, y_3 \ldots, y_n$$

Εκθετική Κατανομή

Έστω ότι έχουμε μία Poisson κατανομή $X_t = \{0,1,2,3,\ldots\}$ με συνάρτηση μάζας πιθανότητας

$$P(X_t = x) = e^{-\lambda \cdot t}\frac{(\lambda \cdot t)^x}{x!}$$

Το χρονικό διάστημα ανάμεσα σε δύο διαδοχικά γεγονότα Α είναι μία τυχαία μεταβλητή T:

$$T = \{\chi\rho\acute{o}\nu o\varsigma\_\mu\varepsilon\tau\alpha\xi\acute{\upsilon}\_\delta\acute{\upsilon}o\_\delta\iota\alpha\delta o\chi\iota\kappa\acute{\omega}\nu\_\gamma\varepsilon\gamma o\nu\acute{o}\tau\omega\nu\_\tau\eta\varsigma\_Poisson\}$$

$$\left.\begin{array}{l} f_T(t) = ? \\ \\ F_T(t) = ? \end{array}\right\} \begin{array}{l} \Rightarrow \quad f_T(t) = \dfrac{dF_T(t)}{dt} = \lambda \cdot e^{-\lambda t} \\ \\ \Rightarrow F_T(t) = P(T \le t) = 1 - P(T > t) = 1 - P(X_t = 0) = 1 - e^{-\lambda t} \end{array}$$

$$E(T) = \frac{1}{\lambda} \qquad VAR(T) = \frac{1}{\lambda^2}$$

Έτσι, αν συμβολίσουμε τον χρόνο με $X: \; f(x) = e^{-\lambda x} \; F(x) = 1 - e^{-\lambda x} \;\; P(X > x) = e^{-\lambda x}$

Ιδιότητα Εκθετικής

Έλλειψη Μνήμης

Για παράδειγμα αν X παριστάνει τον χρόνο καλής λειτουργίας μιας ηλεκτρικής μηχανής, εκθετικής κατανομής, η πιθανότητα να λειτουργήσει στο μέλλον t χρόνο δεδομένου μέχρι τώρα παρήλθε χρόνος λειτουργίας s, Δεν επηρεάζεται από την παλαιότητα της μηχανής (*αλλά από τα τυχαία γεγονότα που συμβαίνουν στο χρόνο*),

$$P(X > t + s / X > s) = e^{-\lambda t} = P(X > t)$$

Παράδειγμα

Έστω μία σεισμογενής περιοχή όπου στα 125 χρόνια σημειώθηκαν 16 σεισμοί. Αν οι σεισμοί ακολουθούν την Poisson διαδικασία να υπολογιστεί η πιθανότητα να μη γίνει σεισμός στα επόμενα 2 χρόνια.

$$\lambda = \frac{16}{128} \quad \text{σεισμοί/χρόνο}$$

Αν X η τυχαία μεταβλητή που παριστάνει το διάστημα μεταξύ δύο σεισμών, τότε αυτή ακολουθεί εκθετική κατανομή και έχουμε ότι:

$$P(X > 2) = e^{-\lambda 2} \cong 0,78$$

Η περίοδος επαναφοράς του σεισμού:

$$E(X) = \frac{1}{\lambda} = \frac{125}{16} = 7,8 \text{ χρόνια}$$

Κανονική Κατανομή (Κατανομή Gauss)

$$X \sim N(\mu, \sigma^2)$$

$$f(X) = \frac{1}{\sqrt{2\pi \cdot \sigma}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx = \ldots = \mu$$

$$VAR(X) = \ldots = \sigma_x{}^2 = E(X^2) - \mu^2 = \sigma^2$$

Εκτίμηση Πιθανότητας

Η δυσκολία της εκτίμησης πιθανότητας στην κανονική κατανομή έγκειται στο ότι δε μπορούμε να ολοκληρώσουμε τη συνάρτηση πυκνότητας πιθανότητας.

$$P(X \le x) = \int_{-\infty}^{x} f(u)du = ;$$

Δε γνωρίζω το <u>αναλυτικό ολοκλήρωμα</u>.

Έτσι οδηγούμαστε στην <u>τυποποίηση της X</u>: $\longrightarrow X^* = \dfrac{X-\mu}{\sigma}$ $\left\} \begin{aligned} E(X^*) &= 0 \\ VAR(X^*) &= 1 \end{aligned} \right.$

$$P(X \le x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) \xrightarrow{\text{Τυποποίηση z}} P\left(Z \le \frac{x-\mu}{\sigma}\right) = P(Z \le z) = F_Z(z) = \Phi(z)$$

Η Z ακολουθεί τυπική κανονική κατανομή: $Z \sim N(0,1^2)$

και από τους στατιστικούς πίνακες που περιέχεται σε κάθε βιβλίο Στατιστικής βρίσκουμε την ζητούμενη πιθανότητα.

Παράδειγμα

Μία βιομηχανία παράγει καρφιά το μήκος των καρφιών X είναι τυχαία μεταβλητή κανονικής κατανομής

$$X \sim N(\mu = 4, \sigma^2 = 1)$$

Τα καρφιά κάτω των 3 πόντων απορρίπτονται. Ποια η πιθανότητα το μήκος; Του καρφιού που θα επιλέξει ένας πελάτης θα είναι μεγαλύτερο των 4,5 πόντους.

$$P(X > 4,5 / X > 3) = \frac{P(X > 4,5)}{P(X > 3)} = \frac{1 - P(X \leq 4,5)}{1 - P(X \leq 3)} = \frac{1 - \Phi\left(\frac{4,5-4}{1}\right)}{1 - \Phi\left(\frac{3-4}{1}\right)} = \frac{1 - \Phi(0,5)}{1 - \Phi(-1)} = \frac{1 - 0,69}{1 - 0,16} = \frac{0,31}{0,84}$$

Παράδειγμα

Η διάρκεια ζωής X ενός μηχανήματος είναι τυχαία μεταβλητή και ακολουθεί κανονική κατανομή:

$$X \sim N(\mu = 400 ώρες, \sigma^2)$$

Απαιτείται άνω του 90% των μηχανημάτων να έχει διάρκεια ζωής πάνω από 350 ώρες. Ποια πρέπει να είναι η διακύμανσή ;

$$P(X \leq 350) = 0,10$$

$$\Phi\left(\frac{350-400}{\sigma}\right) = 0,10$$

Από πίνακες: $\frac{350-400}{\sigma} = -1,28 \Leftrightarrow \sigma = \frac{50}{1,28}$

συνέχεια....

Αν η κανονική κατανομή μας είναι: $X \sim N(\mu = 400, \sigma^2 = 5^2)$

$$\left. P(X \le x) = 0,90 \right\} \quad x = ;$$

$$\left. \begin{array}{l} \Phi\left(\dfrac{x-400}{5}\right) = 0,90 \\ \\ z = 1,28 \end{array} \right\} \dfrac{x-400}{5} = 1,28 \Rightarrow x = 400 + 1,28 \cdot 5 \Rightarrow x = \ldots$$

Προσέγγιση Διωνυμικής με Κανονική

Ένα ανταλλακτικό έχει πιθανότητα p=0,01 να είναι ελαττωματικό. Αν έχουμε n=1000 ανταλλακτικά ποια η πιθανότητα ο αριθμός των ανταλλακτικών να είναι μικρότερος των 20.

$$X = \{1, 2, 3, \ldots, 1000\}$$

$$P(X \le x) = \sum_{x=0}^{20} \binom{n}{x} p^x (1-p)^x$$

<u>Προσέγγιση Διωνυμικής με Poisson</u>

$$E(X) = np = 10$$

$$\sigma_x{}^2 = np(1-p) = 9,9$$

$$\Rightarrow$$

$$P(X_t = x) = e^{-\lambda t} \frac{(\lambda t)^x}{x!} = e^{-10} \frac{10^x}{x!}$$

$$P(X \le 20) = \sum_{x=0}^{20} e^{-10} \frac{10^x}{x!}$$

$$E(X_t) = \lambda t = 10$$

Προσέγγιση Διωνυμικής με Κανονική

$$\left.\begin{array}{l} X \sim N(\mu = np = 10, \sigma^2 = np(1-p) = 9,9) \\ P(X \le 20) = \Phi\left(\dfrac{20-10}{3,1}\right) = \Phi(3,2) \end{array}\right\}$$

➡ Από στατιστικό πίνακα για z=3,2

πιθανότητα:~0,99

Άσκηση

Ο αριθμός κακοτεχνιών σε ένα δρόμο ακολουθεί Poisson κατανομή με 0,10 κακοτεχνίες ανά χιλιόμετρο. Ποια η πιθανότητα να μη βρεθούν κακοτεχνίες στα επόμενα 10 χιλιόμετρα;

$$X = \{απόσταση\_μέχρι\_να\_βρεθεί\_κακοτεχνία\}$$

$f(x) = \lambda e^{-\lambda x}$, η X ακολουθεί εκθετική κατανομή

$$E(X) = 10km$$

$$\lambda = \frac{1}{E(X)} = \frac{1}{10} κακοτεχνία / km$$

$$P(X > 10) = e^{-\frac{1}{10}(10)} = e^{-1}$$

Η πιθανότητα να μη παρουσιάσει κακοτεχνία στα επόμενα 10 km, όταν έχουμε Ήδη διανύσει κάποια απόσταση χωρίς κακοτεχνίες, είναι η ίδια γιατί η εκθετική έχει έλλειψη μνήμης.

συνέχεια…

Ποια η πιθανότητα ότι υπάρχουν 4 κακοτεχνίες στα επόμενα 40km.

$$Y_{t=40} = \{0, 1, 2, \ldots\}$$

Ακολουθεί Poisson κατανομή (εδώ δε μπορούμε να χρησιμοποιήσουμε εκθετική κατανομή):

$$P(Y_{40} = 4) = e^{-\lambda t} \frac{(\lambda t)^4}{4!}$$

Άσκηση

Η ημερήσια ζήτηση X σε Ηλεκτρικό ρεύμα ακολουθεί εκθετική κατανομή με:

$$\left.\begin{array}{l} f(x) = ce^{-cx} \\ \\ E(X) = \dfrac{1}{c} \quad \text{ή} \quad c = \dfrac{1}{E(X)} \\ \\ P(X > 4) = e^{-c\cdot 4} = 0,30 \end{array}\right\} \Rightarrow c = -\frac{1}{4}\ln(30)$$

α) Ποια η πιθανότητα διακοπής ρεύματος σε μία συγκεκριμένη μέρα:

$$P(X > 8) = e^{-8c} = p$$

συνέχεια…

β) Ποια είναι η περίοδος επιστροφής (σε μέρες) της διακοπής μεταξύ δύο διαδοχικών black-out:

$$Y = \{1, 2, \ldots\}$$

$$E(Y) = \frac{1}{p}$$

γ) Αν το black-out μπορεί να συμβεί και από μία άλλη αιτία με πιθανότητα 0,05 , να απαντηθούν τα ίδια ερωτήματα…

Προσθέτουμε την πιθανότητα 0,05 και προχωράμε:

$$P(X > 8) = e^{-8c} = p + 0,05$$

ΑΡΙΣΤΟΤΕΛΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΟΝΙΚΗΣ

ΑΝΟΙΚΤΑ
ΑΚΑΔΗΜΑΙΚΑ
ΜΑΘΗΜΑΤΑ

Τέλος Ενότητας

Επεξεργασία: Καρανάσιος Αναστάσιος-Νικόλαος

Θεσσαλονίκη, Εαρινό Εξάμηνο 2013-2014