Σημειώσεις Στοχαστικές Διεργασίες - 8ο Μάθημα

Τρίτη 21 Νοεμβρίου 2023

13. Κατανομή Poisson

Μία τ.μ. X ε N, ακολουθεί την κατανομή Poisson (X ~ Poisson(λ) ή απλά X ~ P(λ)), αν για κάποιο λ > 0, η συνάρτηση πιθανότητας δίνεται από τον τύπο:

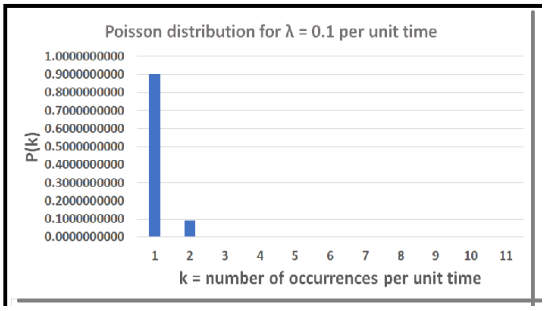
$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k}!,\ k = 0, 1, 2, \ldots$$

Η κατανομή Poisson εφαρμόζεται σε φαινόμενα με διακριτή συχνότητα που επιπλέον έχουν σταθερή πιθανότητα να συμβούν σε δοσμένο χρόνο ή χώρο. Η παράμετρος λ συμβολίζει την αναμενόμενη τιμή της X σε αυτόν το χρόνο ή χώρο, ενώ επίσης λ είναι και η διασπορά της (E(X) = Var(X) = λ). Παραδείγματα τ.μ. που ακολουθούν την Poisson κατανομή:

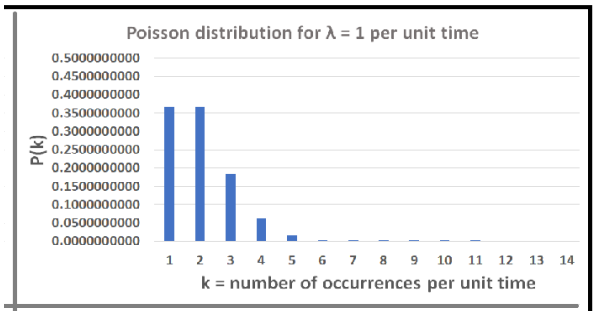
- οι τηλεφωνικές κλήσεις που φτάνουν σε ένα σύστημα (X: πλήθος κλήσεων, λ: μέσο πλήθος κλήσεων).
- τα φωτόνια που φτάνουν σε ένα τηλεσκόμπιο (X: πλήθος φωτονίων, λ: μέσο πλήθος φωτονίων που φτάνουν).
- ο αριθμός των μεταλλάξεων σε ένα κλώνο του DNA ανά μονάδα μήκους (X: πλήθος μεταλλάξεων, λ: μέσο πλήθος μεταλλάξεων που παρατηρούνται).
- Ο αριθμός των γκολ σε αγώνες ποδοσφαίρου (X: πλήθος γκολ, λ: μέσο πλήθος γκολ σε ανάλογους αγώνες).

Παράδειγμα 1

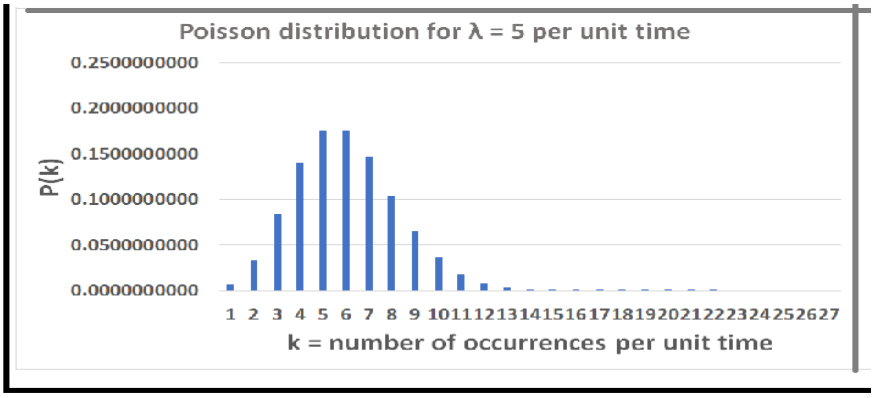
λ = 0,1: $P(X = \kappa) = e^{-0,1} \cdot 0,1^{\kappa} / \kappa!$, κ = 0, 1,…



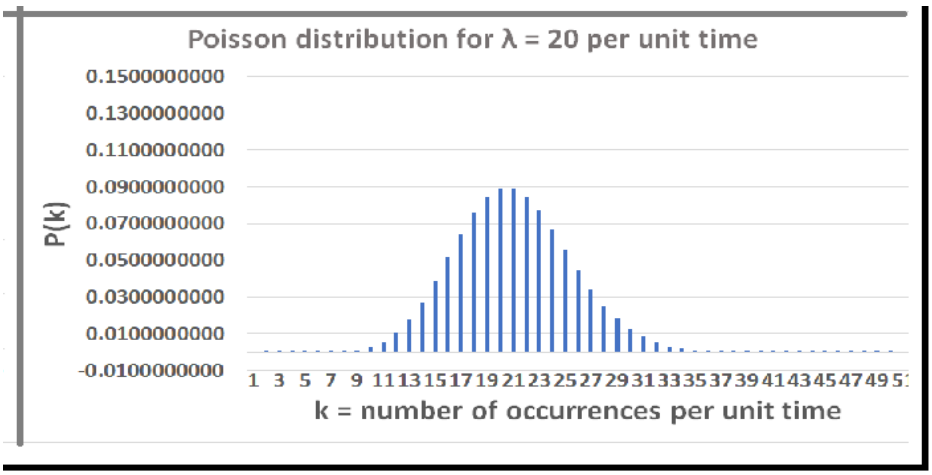
λ = 1: $P(X = \kappa) = e^{-1} / \kappa!$, κ = 0, 1,…



$\lambda = 5$: $P(X = \kappa) = e^{-5} \cdot 5^{\kappa} / \kappa!$, $\kappa = 0, 1, \ldots$



$\lambda = 20$: $P(X = \kappa) = e^{-20} \cdot 20^{\kappa} / \kappa!$, $\kappa = 0, 1, \ldots$



Παράδειγμα 2

Ο αριθμός των σωματιδίων που εκπέμπει μια ραδιενεργός πηγή ανά δευτερόλεπτο ακολουθεί την κατανομή Poisson με μέσο $\lambda = 5$. Να υπολογιστούν οι πιθανότητες:

$P(X = 0)$, $P(X = 1)$, $P(X = 2)$ και $P(X \geq 3)$.

Λύση

Είναι $X \sim \text{Poisson}(5)$ άρα $P(X = \kappa) = e^{-5} \cdot 5^{\kappa} / \kappa!$. Υπολογίζουμε:

$P(X = 0) = e^{-5} \cdot 5^{0} / 0! = e^{-5} = 0{,}0067 = 0{,}67\%$.

$P(X = 1) = e^{-5} \cdot 5^{1} / 1! = 5 \cdot e^{-5} = 0{,}0337 = 3{,}37\%$.

$P(X = 2) = e^{-5} \cdot 5^{2} / 2! = 25 \cdot e^{-5} / 2 = 0{,}0842 = 8{,}42\%$.

$P(X \geq 3) = 1 - P(X < 3) = 1 - [P(X = 0) + P(X = 1) + P(X = 2)] = 0{,}8753 = 87{,}5\%$.

Παράδειγμα 3

Έστω ότι η τ.μ. Χ ακολουθεί την κατανομή Poisson με παράμετρο λ. Αν ισχύει $P(X \leq 1) = 4 \cdot P(X = 2)$, να προσδιοριστεί η παράμετρος λ και να υπολογιστούν οι πιθανότητες $P(X < 3)$, $P(X \geq 1 \mid X < 3)$.

Λύση

Είναι $X \sim \text{Poisson}(\lambda)$ άρα $P(X = \kappa) = e^{-\lambda} \cdot \lambda^{\kappa} / \kappa!$. Τώρα,

$P(X \leq 1) = 4 \cdot P(X = 2) \Leftrightarrow P(X = 0) + P(X = 1) = 4 \cdot P(X = 2) \Leftrightarrow$

$\Leftrightarrow e^{-\lambda} + e^{-\lambda} \cdot \lambda = 4\, e^{-\lambda} \cdot \lambda^2 / 2$

$\Leftrightarrow 2\lambda^2 - \lambda - 1 = 0$

$\Leftrightarrow \lambda = -1/2$ ή $\lambda = 1$.

Καθώς, το λ αποτελεί ρυθμό εμφάνισης, πρέπει να είναι θετικός αριθμός. Άρα $\lambda = 1$.

$P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2)$

$= e^{-1} \cdot 1^0 / 0! + e^{-1} \cdot 1^1 / 1! + e^{-1} \cdot 1^2 / 2!$

$= e^{-1} \cdot 5 / 2$

$= 0{,}9196 \approx 92\%$

$P(X \geq 1 \mid X < 3) = P(X \geq 1, X < 3) / P(X < 3)$

$= P(X = 1$ ή $X = 2) / P(X < 3)$

$= [P(X = 1) + P(X = 2)] / P(X < 3)$

$= 3/5 = 0{,}6 = 60\%$.

13.1 Διεργασίες καταμέτρησης

Σε ορισμένα προβλήματα, μετράμε τις εμφανίσεις ορισμένων τύπων γεγονότων. Σε αυτές τις περιπτώσεις, έχουμε να κάνουμε με μια διεργασία καταμέτρησης.

Για παράδειγμα, η διεργασία {N(t), t ≥ 0} που δείχνει τον αριθμό των πελατών που φτάνουν σε ένα σούπερ μάρκετ μέχρι την ώρα t ξεκινώντας από την ώρα 0. Για τέτοιου είδους διαδικασίες, συνήθως υποθέτουμε N(0) = 0, άρα καθώς ο χρόνος περνάει και φτάνουν πελάτες, το N(t) παίρνει θετικές ακέραιες τιμές.

Ορισμός διεργασίας καταμέτρησης

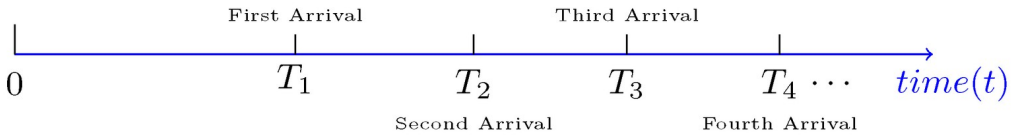
Μια στοχαστική διεργασία {N(t), t∈[0,∞)} λέγεται διεργασία καταμέτρησης εάν N(t) είναι ο αριθμός των γεγονότων που συνέβησαν από το χρόνο 0 έως και συμπεριλαμβανομένου του χρόνου t, δηλαδή

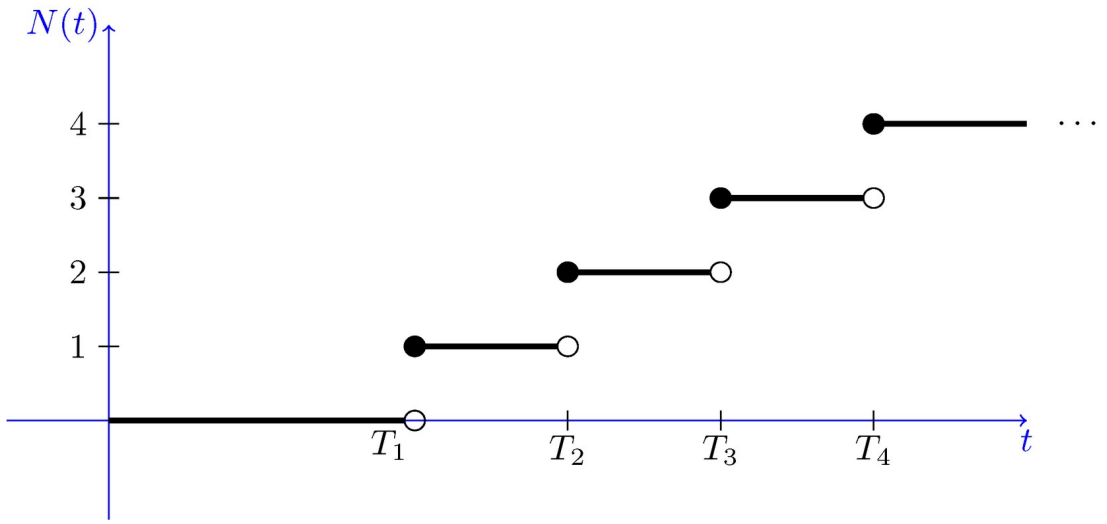
N(t) = {ο αριθμός των γεγονότων που συνέβησαν στο (0, t]}

Για μια διαδικασία μέτρησης, υποθέτουμε:

- N(0) = 0.
- N(t)∈{0, 1, 2, …}, για όλα τα t∈[0,∞)

Για 0 ≤ s < t η διαφορά N(t) − N(s) δείχνει τον αριθμό των γεγονότων που συμβαίνουν στο διάστημα (s, t].





Πηγή: https://www.probabilitycourse.com/chapter11/11_1_1_counting_processes.php

13.2 Διεργασίες Poisson

Η κατανομή Poisson δίνει τη δυνατότητα υπολογισμού συγκεκριμένων πιθανοτήτων για το πλήθος εμφανίσεων κάθε διακριτού φαινομένου που εξελίσσεται με τυχαίο τρόπο.

Αν ο ρυθμός άφιξης είναι σταθερός στην διάρκεια του χρόνου, το πλήθος αφίξεων σε οποιοδήποτε χρονικό διάστημα μήκους t, θα ακολουθεί την κατανομή Poisson με παράμετρο $\lambda \cdot t$. Με τον τρόπο αυτό, το ίδιο το φαινόμενο είναι δυνατό να μοντελοποιηθεί στα πλαίσια μίας διαδικασίας καταμέτρησης συνεχούς χρόνου, $\{N(t), t \geq 0\}$,

όπου $N(t) = \{$ο αριθμός των αφίξεων στο χρονικό διάστημα $(0, t]\}$ και $N(0) = 0$.

Ορισμός διεργασίας Poisson

Ως **διεργασία Poisson** με μέση ένταση (mean rate or intensity) λ, **ορίζεται** μία στοχαστική διεργασία καταμέτρησης $\{N(t), t \geq 0\}$ (άρα $N(0) = 0$), για την οποία επιπλέον:

(i) Η $N(t)$, $t \geq 0$, έχει ανεξάρτητες προσαυξήσεις, δηλαδή: για κάθε $0 \leq t_1 < t_2 < \ldots < t_n$, οι προσαυξήσεις $N(t_2) - N(t_1)$, $N(t_3) - N(t_2)$, …, $N(t_n) - N(t_{n-1})$ είναι ανεξάρτητες τ.μ..

(ii) Αν $X = \{$πλήθος γεγονότων σε χρονικό διάστημα μήκους $\tau\}$, τότε $X \sim \text{Poisson}(\lambda \cdot \tau)$.

(Είναι αξιοσημείωτο ότι εξ' ορισμού το πλήθος των γεγονότων εξαρτάται μόνο από το πλάτος του χρονικού διαστήματος και όχι από τη θέση του διαστήματος στην ημιευθεία του χρόνου)

Παρατηρήσεις

1. Η προϋπόθεση (ii) είναι ισοδύναμη με τη συνθήκη

$$N(t) - N(s) \sim \text{Poisson}(\lambda(t - s)) \sim N(t - s)$$

από την οποία προκύπτει ότι οι προσαυξήσεις είναι στάσιμες (ισόνομες). Ειδικότερα, συνάγουμε ότι:

$$E(N(t) - N(s)) = \text{Var}(N(t) - N(s)) = \lambda(t - s).$$

2. Η προϋπόθεση του ορισμού "(i) Η $N(t)$, $t \geq 0$, έχει ανεξάρτητες προσαυξήσεις" είναι ουσιαστική. Είναι δυνατόν μία διεργασία καταμέτρησης $\{N(t), t \geq 0\}$ να ικανοποιεί την

"(ii) Αν $X = \{$πλήθος γεγονότων σε χρονικό διάστημα μήκους $\tau\}$, τότε $X \sim \text{Poisson}(\lambda \cdot \tau)$.",

χωρίς ωστόσο να έχει ανεξάρτητες προσαυξήσεις.

Παράδειγμα

Έστω $\{N(t), t \geq 0\}$ μία διεργασία Poisson με παράμετρο λ: $N(t+s) - N(s) \sim \text{Poisson}(\lambda t)$, $s \geq 0$. Ορίζουμε

$$Z(t) = N(s + t) - N(s) \text{ για κάθε } t, s \geq 0.$$

Η διεργασία $\{Z(t), t \geq 0\}$ είναι διαδικασία καταμέτρησης, $Z(t) \sim \text{Poisson}(\lambda t)$, αλλά οι προσαυξήσεις της είναι εξαρτημένες καθώς για $m > s \geq 0$, $t_2 > t_1 \geq 0$ είναι

$$Z(t_2) - Z(t_1) = N(t_2+s) - N(t_1+s) \text{ και } Z(t_1) - Z(0) = N(t_1+m) - N(m),$$

τα διαστήματα $(t_1, t_2]$ και $(0, t_1]$ δεν επικαλύπτονται αλλά οι διαφορές $N(t_2+s) - N(t_1+s)$ και $N(t_1+m) - N(m)$ είναι εξαρτημένες καθώς $m > s$ και τα διαστήματα $[t_1+ s, t_2+ s]$ και $[m, t_1+ m]$ επικαλύπτονται.

Πηγή: https://math.stackexchange.com/questions/132892/counting-process-which-is-not-a-poisson-process

Πλήθος γεγονότων σε διάστημα

Αν Χ = {πλήθος αφίξεων σε χρονικό διάστημα μήκους τ}, τότε **Χ ~ Poisson(λ·τ).** Συμπεραίνουμε, ότι σε μια διεργασία Poisson, **το πλήθος των γεγονότων σε οποιοδήποτε διάστημα εξαρτάται μόνο από το μήκος του διαστήματος** και όχι από την ακριβή θέση του διαστήματος στην ημιευθεία του χρόνου. Πράγματι, για κάθε $t_2 > t_1 \geq 0$ και $r > 0$, είναι

$$N(t_2 + r) - N(t_1+ r) \sim \text{Poisson}(\lambda\cdot[t_2+ r - (t_1+ r)]) = \text{Poisson}(\lambda\cdot(t_2 - t_1))$$

και

$$N(t_2) - N(t_1) \sim \text{Poisson}(\lambda\cdot(t_2 - t_1)).$$

Συμπεραίνουμε ότι $N(t_2 + r) - N(t_1+ r) \sim N(t_2) - N(t_1)$, επομένως, η διαδικασία Poisson **είναι διεργασία με στάσιμες προσαυξήσεις**. Ιδιαίτερα, προκύπτει ότι:

$$E(N(t_2) - N(t_1)) = Var(N(t_2) - N(t_1)) = \lambda(t_2 - t_1), \text{ για κάθε } t_1 < t_2.$$

Από τον ορισμό της μία διεργασία Poisson ικανοποιεί την συνθήκη

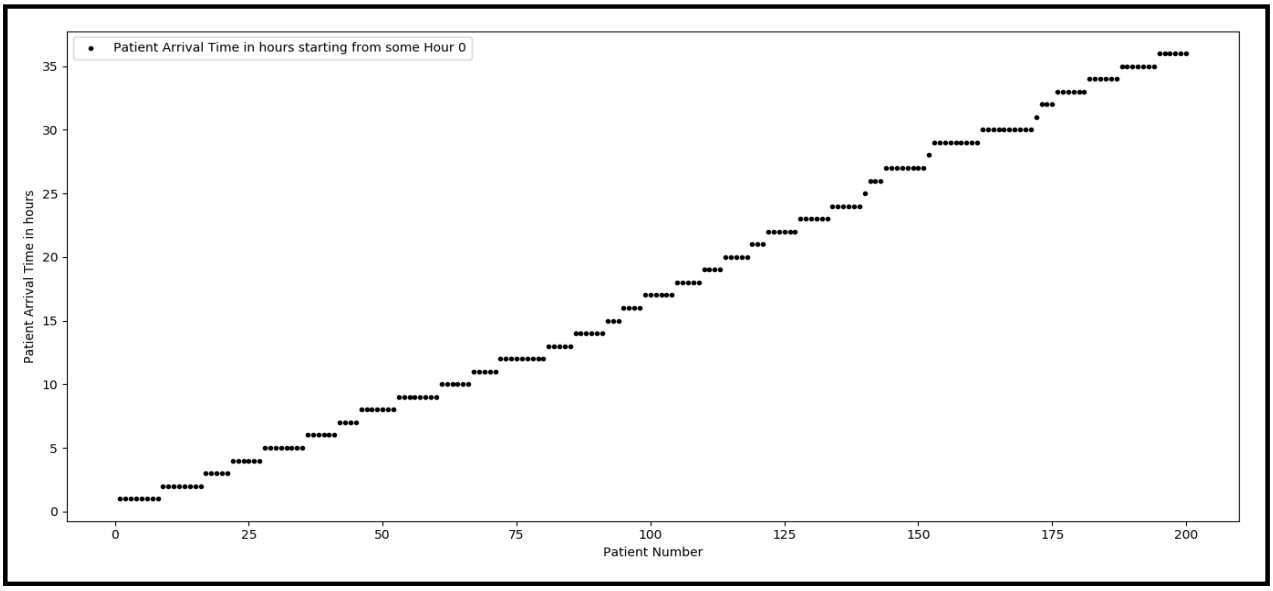
$$E(N(t)) = Var(N(t)) = \lambda t.$$

Παρατηρούμε ότι η αναμενόμενη τιμή της N(t) εξαρτάται από το t, ειδικότερα αυτή δεν είναι σταθερή. Προκύπτει ότι:

Η διεργασία Poisson δεν είναι ούτε ισχυρά στάσιμη (κοινή κατανομή για κάθε συνδυασμό των μεταβλητών), ούτε ασθενώς στάσιμη (κοινό EX, VarX για όλες τις μεταβλητές) στοχαστική διεργασία.

Γραφικό παράδειγμα

Αφίξεις ασθενών σε ένα ιατρείο ως μία διεργασία Poisson $\{N(t), t \geq 0\}$ με $\lambda = 1$ αφίξεις ανά ώρα.



Πηγή: https://towardsdatascience.com/the-poisson-process-everything-you-need-to-know-322aa0ab9e9a

Παράδειγμα: Ραδιενεργή αποσύνθεση

Όλες οι σύγχρονες θεωρίες της ραδιενεργής αποσύνθεσης θεωρούν ότι σε μία ομάδα πυρήνων ενός δεδομένου στοιχείου όλοι οι πυρήνες είναι ίδιοι, ανεξάρτητοι και έχουν την ίδια πιθανότητα αποσύνθεσης σε μοναδιαίο χρόνο. Μπορεί να αποδειχθεί ότι οι εκπομπές μίας ραδιενεργής πηγής αποτελούν μία διεργασία Poisson.

Παράδειγμα: Θόρυβος βολής σε ηλεκτρονικές λυχνίες

Η ευαισθησία που επιτυγχάνεται με τους ηλεκτρονικούς ενισχυτές περιορίζεται από τις στιγμιαίες μεταβολές ρεύματος που συμβαίνουν σε τέτοια όργανα και οι οποίες καλούνται θόρυβος βολής, ο οποίος οφείλεται στις τυχαίες εκπομπές ηλεκτρονίων από την θερμαινόμενη κάθοδο. Ας υποθέσουμε ότι η διαφορά δυναμικού μεταξύ καθόδου και ανόδου είναι τόσο μεγάλη ώστε όλα τα εκπεμπόμενα ηλεκτρόνια από την κάθοδο να έχουν πολύ μεγάλες ταχύτητες, με αποτέλεσμα να μην υπάρχει συγκέντρωση ηλεκτρονίων μεταξύ καθόδου και ανόδου. Ο αριθμός των ηλεκτρονίων που εκπέμπονται από την κάθοδο σε ένα χρονικό διάστημα $(0, t]$ αποτελεί μία διεργασία Poisson με παράμετρο λt, όπου λ είναι η μέση ένταση εκπομπής ηλεκτρονίων από την κάθοδο.

Παράδειγμα: Διακοπές μηχανών

Θεωρούμε ένα όργανο (όπως μία λυχνία κενού ή έναν απαριθμητή Geiger) το οποίο χρησιμοποιείται μέχρι να σταματήσει να λειτουργεί και μετά επισκευάζεται ή αντικαθίσταται από ένα όργανο του ίδιου τύπου. Η διάρκεια ζωής του οργάνου θεωρείται ότι είναι μία τυχαία μεταβλητή T. Οι διάρκειες ζωής T_1 , T_2 ,...,T_n, n οργάνων που τέθηκαν σε

λειτουργία θεωρούνται ανεξάρτητες τυχαίες μεταβλητές με την ίδια κατανομή όπως η τυχαία μεταβλητή T. Για $t > 0$, έστω $N(t)$ ο αριθμός των οργάνων που απέτυχαν στο χρονικό διάστημα $(0, t]$. Αν η διάρκεια ζωής κάθε οργάνου είναι μία εκθετική κατανομή, τότε μπορεί να αποδειχθεί ότι $\{N(t), t \geq 0\}$ είναι μία διεργασία Poisson.

Παρατήρηση

Η ιδιότητα των (στοχαστικά) ανεξάρτητων προσαυξήσεων απλοποιεί την ανάλυση μιας διαδικασίας καταμέτρησης.

Ειδικότερα, στα πλαίσια της διεργασίας Poisson, αν $0 \leq t_1 < t_2$, τότε

$P(N(t_1) = i, N(t_2) = j)$ (πιθανότητα να συμβούν i αφίξεις στο $[0, t_1)$ και j στο $[0, t_2)$)

$= P(N(t_1) = i, N(t_2) - N(t_1) = j - i)$ (πιθανότητα να συμβούν i αφίξεις στο $[0, t_1)$ και $j - i$ στο $[t_1, t_2)$)

$= P(N(t_1) = i, N(t_2 - t_1) = j - i)$ (ανεξάρτητες προσαυξήσεις: $N(t_2) - N(t_1) \sim N(t_2 - t_1)$)

$= P(N(t_1) = i) \cdot P(N(t_2 - t_1) = j - i)$. (ανεξαρτησία πλήθους αφίξεων από τη θέση των διαστημάτων)

$= e^{-\lambda t1} (\lambda t_1)^i/i! \cdot e^{-\lambda(t2 - t1)} [\lambda(t_2 - t_1)]^{j - i}/(j - i)!$

$= \lambda^j e^{-\lambda t2} \cdot t_1^i \cdot (t_2 - t_1)^{j - i}/[i!(j - i)!]$.

Για παράδειγμα, ας υποθέσουμε ότι θέλουμε να βρούμε την πιθανότητα να έχουμε 2 αφίξεις στο διάστημα (1,2] και 3 αφίξεις στο διάστημα (3, 5].

Επειδή τα διαστήματα (1, 2] και (3, 5] είναι ξένα, μπορούμε να γράψουμε

P(2 αφίξεις σε (1,2] και 3 αφίξεις σε (3,5])

= P(2 αφίξεις σε (1,2]) · P(3 αφίξεις σε (3,5])

Παράδειγμα

Ο αριθμός των πελατών που φτάνουν σε ένα μάρκετ μπορεί να θεωρηθεί μια διαδικασία Poisson με ρυθμό $\lambda = 10$ πελάτες ανά ώρα.

(α) Να βρεθεί η πιθανότητα να υπάρχουν 2 πελάτες μεταξύ 10:00 και 10:20.

(β) Να βρεθεί η πιθανότητα να υπάρχουν 3 πελάτες μεταξύ 10:00 και 10:20 και 7 πελάτες μεταξύ 10:20 και 11:00.

Λύση

(α) Ο ρυθμός με τον οποίο γίνεται η άφιξη των πελατών είναι 10 πελάτες / ώρα = 1/6 πελάτες / λεπτό. Καθώς η διαδικασία αφίξεων είναι μία διεργασία Poisson, από την προυπόθεση (ii) του ορισμού, για την τ.μ. X = {πλήθος αφίξεων μεταξύ 10:00 και 10:20}, είναι $X \sim \text{Poisson}(1/6 \cdot 20) = \text{Poisson}(10/3)$.

(α) P({2 πελάτες μεταξύ 10:00 και 10:20}) = P(X = 2)

$$= e^{-10/3} \cdot (10/3)^2 / 2!$$

$$\approx 0{,}198 = 19{,}8\%$$

(β) Έστω

X = {πλήθος αφίξεων μεταξύ 10:00 και 10:20} και

Y = {πλήθος αφίξεων μεταξύ 10:20 και 11:00}.

Είναι X ~ Poisson(1/6·20) και όπως στο (α) καταλήγουμε ότι Y ~ Poisson(20/3). Επίσης, αν N(t) = {πλήθος πελατών που φτάνουν στο μάρκετ μέχρι τη χρονική στιγμή t}, είναι N(t) ~ Poisson(1/6 · t) (t: χρόνος σε λεπτά), X = N(10:20) – N(10:00), Y = N(11:00) – N(10:20), και οι X, Y δηλώνουν τις προσαυξήσεις μεταξύ των διαστημάτων (10:00, 10:20] και (10:20, 11:00], άρα είναι ανεξάρτητες μεταξύ τους (ιδιότητα (i) του ορισμού). Τώρα, υπολογίζουμε:

P({3 πελάτες μεταξύ 10:00 και 10:20 και 7 πελάτες μεταξύ 10:20 και 11:00})

$$= P(X = 3, Y = 7)$$

$$= P(X = 3) \cdot P(Y = 7) \qquad \text{(ανεξάρτητες προσαυξήσεις)}$$

$$= e^{-10/3} \cdot (10/3)^3 / 3! \cdot e^{-20/3} \cdot (20/3)^7 / 7!$$

$$\approx 0{,}2202 \cdot 0{,}1477$$

$$= 0{,}0325 = 3{,}25\%.$$

Παράδειγμα

Έστω {N(t), t∈[0,∞)} μια διεργασία Poisson με ρυθμό λ = 0,5 αφίξεις / μονάδα χρόνου.

(α) Να βρεθεί η πιθανότητα να μην συμβεί καμία άφιξη στο διάστημα (3, 5].

(β) Να βρεθεί η πιθανότητα να υπάρχει ακριβώς μία άφιξη σε καθένα από τα ακόλουθα διαστήματα: (0, 1], (1, 2], (2, 3] και (3, 4].

Λύση

(α) Έστω Y το πλήθος των αφίξεων στο διάστημα (3, 5]. Καθώς, το μήκος του διαστήματος είναι δ = 2, θα είναι Y ~ Poisson(λ·δ) = Poisson(0,5·2) = Poisson(1).

Υπολογίζουμε, $P(Y = 0) = e^{-1}\, 1^0/0!$

$$= 1/e$$

$$= 0{,}368$$

$$= 36{,}8\%.$$

(β) Έστω Y_1, Y_2, Y_3, Y_4, το πλήθος των αφίξεων στα διαστήματα (0, 1], (1, 2], (2, 3] και (3, 4]. Κάθε ένα από τα διαστήματα έχει μήκος δ = 1, άρα $Y_i \sim \text{Poisson}(0{,}5)$, i = 1, 2, 3, 4. Επιπλέον, τα διαστήματα δεν είναι επικαλυπτόμενα, άρα οι Y_1, Y_2, Y_3, Y_4, είναι ανεξάρτητες μεταξύ τους.

Υπολογίζουμε: $P(Y_1 = 1, Y_2 = 1, Y_3 = 1, Y_4 = 1)$

$$= P(Y_1 = 1) \cdot P(Y_2 = 1) \cdot P(Y_3 = 1) \cdot P(Y_4 = 1)$$

$$= (0{,}5 \cdot e^{-0{,}5})^4 = 0{,}0085 = 0{,}85\%.$$

Παράδειγμα

Έστω $\{N(t), t \in [0,\infty)\}$ μια διεργασία Poisson με ρυθμό λ = 3 αφίξεις / μονάδα χρόνου. Να βρεθεί η πιθανότητα να υπάρχουν δύο αφίξεις στο (0, 2] και τρεις αφίξεις στο (1, 4].

Λύση

Τα διαστήματα (0, 2] και (1, 4] επικαλύπτονται άρα το πλήθος αφίξεων που θα συμβούν στα δύο χρονικά διαστήματα δεν είναι ανεξάρτητα μεταξύ τους. Γράφουμε τα διαστήματα αυτά ως ένωση ξένων μεταξύ τους διαστημάτων.

$$(0, 2] \cup (1, 4] = (0, 1] \cup (1, 2] \cup (2, 4],$$

Αν X: αφίξεις στο (0, 1], Y: αφίξεις στο (1, 2] και Z: αφίξεις στο (2, 4], τότε παρατηρούμε ότι για να συμβούν δύο αφίξεις στο (0, 2] και τρεις αφίξεις στο (1, 4] πρέπει να συμβεί ένας από τους εξής καταμερισμούς στα διαστήματα (0, 1], (1, 2], (2, 4]:

$$(X, Y, Z) = (1, 1, 2), \text{ ή } (X, Y, Z) = (2, 0, 3), \text{ ή } (X, Y, Z) = (0, 2, 1).$$

Αν X, Y, Z είναι το πλήθος αφίξεων στα (0, 1], (1, 2], (2, 4] αντίστοιχα, τότε

$X \sim \text{Poisson}(\lambda \cdot 1)$, $Y \sim \text{Poisson}(\lambda \cdot 1)$, $Z \sim \text{Poisson}(\lambda \cdot 2)$ ή

$X \sim \text{Poisson}(3)$, $Y \sim \text{Poisson}(3)$, $Z \sim \text{Poisson}(6)$.

Υπολογίζουμε:

P({δύο αφίξεις στο (0, 2] και τρεις αφίξεις στο (1, 4]})

$$= P((X, Y, Z) = (1, 1, 2) \text{ ή } (X, Y, Z) = (2, 0, 3) \text{ ή } (X, Y, Z) = (0, 2, 1))$$

$$= P((X, Y, Z) = (1, 1, 2)) + P((X, Y, Z) = (2, 0, 3)) + P((X, Y, Z) = (0, 2, 1))$$

$$= P(X=1) \cdot P(Y=1) \cdot P(Z=2) + P(X = 2) \cdot P(Y = 0) \cdot P(Z = 3) + P(X = 0) \cdot P(Y = 2) \cdot P(Z = 1)$$

$$= e^{-\lambda} \cdot e^{-\lambda} \cdot e^{-2\lambda} (2\lambda)^2/2! + e^{-\lambda} \lambda^2/2! \cdot e^{-\lambda} \cdot e^{-2\lambda} (2\lambda)^3/3! + e^{-\lambda} \cdot e^{-\lambda} \lambda^2/2! \cdot e^{-2\lambda} (2\lambda)^1/1!$$

$$= e^{-3} \cdot e^{-3} \cdot e^{-6} \cdot 6^2/2 + e^{-3} \cdot 3^2/2 \cdot e^{-3} \cdot e^{-6} \cdot 6^3/6 + e^{-3} \cdot e^{-3} \cdot 3^2/2 \cdot e^{-6} \cdot 6$$

$$= e^{-12} \cdot (18 + 9 \cdot 18 + 27) = 0{,}00127... \approx 0{,}13\%$$

Παρατήρηση

Το ίδιο πρόβλημα θα μπορούσε να λυθεί με εφαρμογή του νόμου ολικής πιθανότητας, διαμερίζοντας τις πιθανές περιπτώσεις ως προς τις τιμές της τ.μ. Y ως εξής:

P({δύο αφίξεις στο (0, 2] και τρεις αφίξεις στο (1, 4]})

$= P(X + Y = 2, Y + Z = 3)$

$= \Sigma_{\kappa = 0, 1, \ldots} P(X + Y = 2, Y + Z = 3, Y = \kappa)$

$= \Sigma_{\kappa = 0, 1, \ldots} P(X + Y = 2, Y + Z = 3 \mid Y = \kappa) \cdot P(Y = \kappa)$

$= P(X + Y = 2, Y + Z = 3 \mid Y = 0) \cdot P(Y = 0)$

$+ P(X+Y = 2, Y + Z = 3 \mid Y = 1) \cdot P(Y = 1) + P(X + Y = 2, Y + Z = 3 \mid Y = 2) \cdot P(Y = 2)$

$= P(X = 2, Z = 3) \cdot P(Y = 0) + P(X = 1, Z = 2) \cdot P(Y = 1) + P(X = 0, Z = 1) \cdot P(Y = 2)$

$= P(X=1)\cdot P(Y=1)\cdot P(Z = 2) + P(X = 2)\cdot P(Y = 0)\cdot P(Z = 3) + P(X = 0)\cdot P(Y = 2)\cdot P(Z = 1)$

$= e^{-\lambda}\cdot e^{-\lambda}\cdot e^{-2\lambda}\,(2\lambda)^2/2! + e^{-\lambda}\,\lambda^2/2!\,\cdot e^{-\lambda}\,\cdot e^{-2\lambda}\,(2\lambda)^3/3! + e^{-\lambda}\cdot e^{-\lambda}\,\lambda^2/2!\cdot e^{-2\lambda}\,(2\lambda)^1/1!$

$= e^{-3}\cdot e^{-3}\cdot e^{-6}\cdot 6^2/2 + e^{-3}\cdot 3^2/2\,\cdot e^{-3}\,\cdot e^{-6}\,\cdot 6^3/6 + e^{-3}\cdot e^{-3}\cdot 3^2/2\cdot e^{-6}\,\cdot 6$

$= e^{-12}\,\cdot (18 + 9\cdot 18 + 27)$

$= 0{,}00127... \approx 0{,}13\%$

Παράδειγμα

Έστω $\{N(t), t \in [0,\infty)\}$ μια διεργασία Poisson με ρυθμό λ. Να βρεθεί η συνάρτηση συνδιακύμανσης

$$\mathrm{Cov}(N(t_1), N(t_2)) = E(N(t_1) - \lambda t_1)\cdot(N(t_2) - \lambda t_2),$$

συναρτήσει των t_1, t_2 για κάθε $t_1, t_2 \geq 0$.

Λύση

Θα αξιοποιήσουμε τις εξής τρεις ιδιότητες της συνδιακύμανσης:

(α) X, Y: ανεξάρτητες $\rightarrow$ Cov(X, Y) = 0.

(β) Cov(X, X) = Var(X).

(γ) Cov(X, Y + Z) = Cov(X, Y) + Cov(X, Z).

Επιπλέον, θα χρειαστούμε τις εξής ιδιότητες της διεργασίας Poisson:

(α) $N(t_2) - N(t_1) \sim N(t_2 - t_1)$, $t_2 > t_1 \geq 0$.

(β) Οι τ.μ. $N(t_1)$, $N(t_2 - t_1)$, είναι στοχαστικά ανεξάρτητες για κάθε $t_2 > t_1 \geq 0$ (ανεξάρτητες προσαυξήσεις μεταξύ των διαστημάτων $(0, t_1]$ και $(t_1, t_2]$).

(γ) $N(t_1) \sim \text{Poisson}(\lambda \cdot t_1)$ και $N(t_2) \sim \text{Poisson}(\lambda \cdot t_2)$ (άρα $E(N(t_i)) = Var(N(t_i)) = \lambda \cdot t_i$, $i = 1, 2$)

Για $t_2 > t_1 \geq 0$, είναι: $Cov(N(t_1), N(t_2)) = Cov(N(t_1), N(t_2) - N(t_1) + N(t_1))$

$$= Cov(N(t_1), N(t_2 - t_1) + N(t_1))$$

$$= Cov(N(t_1), N(t_2 - t_1)) + Cov(N(t_1), N(t_1))$$

$$= 0 + Var(N(t_1))$$

$$= \lambda t_1.$$

Ανάλογα, αποδεικνύεται ότι για $t_1 > t_2 \geq 0$, είναι $C_N(t_1, t_2) = \lambda t_2$.

Συμπεραίνουμε ότι: $\mathbf{Cov(N(t_1), N(t_2)) = \lambda \cdot min(t_1, t_2)}$, για κάθε $t_1, t_2 \geq 0$**.**

13.3 Χρόνος μέχρι το 1ο γεγονός

Έστω $\{N(t), t \geq 0\}$, διεργασία Poisson με παράμετρο λ. Έστω

X_1: ο χρόνος που περνάει μέχρι να συμβεί το 1ο γεγονός.

Αναζητούμε την κατανομή της τ.μ. X, δηλαδή, την τιμή της πιθανότητας $P(X_1 \leq t)$, για $t \geq 0$. Παρατηρούμε ότι για κάθε $t > 0$,

$\{X_1 > t\}$ = {δεν συμβαίνει κανένα γεγονός στο χρονικό διάστημα (0, t]} = {N(t) = 0},

άρα, $P(X_1 > t) = P(N(t) = 0)$

$$= e^{-\lambda t} \lambda^0/0!$$

$$= e^{-\lambda t} \text{ ή}$$

$P(X_1 \leq t) = 1 - e^{-\lambda t}$, $t \geq 0$ (και $P(X_1 \leq t) = 0$, $t < 0$),

Δηλαδή η τ.μ. X_1 έχει την αθροιστική συνάρτηση κατανομής της εκθετικής κατανομής με παράμετρο λ. Συμπεραίνουμε ότι $X_1 \sim \text{Exp}(\lambda)$. Άμεσες συνέπειες είναι οι εξής:

Αναμενόμενος χρόνος μέχρι το 1ο γεγονός: $E(X_1) = 1/\lambda$

Διασπορά χρόνου μέχρι το 1ο γεγονός: $Var(X_1) = 1/\lambda^2$.

Παράδειγμα

Έστω $\{N(t), t \in [0,\infty)\}$ μια διεργασία Poisson με ρυθμό λ = 5 αφίξεις / min.

(α) Να βρεθεί το χρονικό διάστημα που αναμένεται να περάσει έως την 1η άφιξη.

(β) Να βρεθεί η πιθανότητα η 1η άφιξη να καθυστερήσει περισσότερο από μισό λεπτό.

Λύση

(α) $E(X_1) = 1/\lambda = 1/5$ min = 0,2 min = 12 sec.

(β) Γνωρίζουμε ότι $X_1 \sim Exp(\lambda) = Exp(5)$. Είναι:

$P(X_1 > 0{,}5) = 1 - P(X_1 \le 0{,}5)$

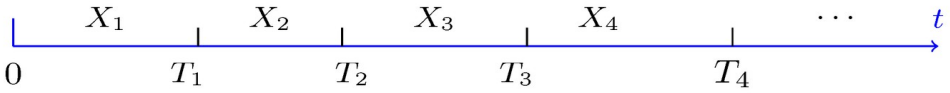
$= 1 - (1 - e^{-\lambda t})$

$= 1 - (1 - e^{-5\cdot 0{,}5})$

$= e^{-2{,}5} = 0{,}082 = 8{,}2\%$

13.3.1 Χρόνος μεταξύ δύο διαδοχικών γεγονότων

Το επόμενο βήμα είναι να βρούμε την κατανομή του χρόνου X_i, i = 2, 3, …, που μεσολαβεί μεταξύ οποιωνδήποτε δύο συμβάντων.



Έστω X_n το χρονικό διάστημα που μεσολαβεί μεταξύ του n-1 και του n γεγονότος. Αναζητούμε την κατανομή πιθανότητας της X_n. Είναι:

$P(X_n > t) = P(\{$κανένα γεγονός στο (s, s+t]$\})$ (s η στιγμή που συνέβη το προηγούμενο γεγονός)

$= P(N(s + t) - N(s) = 0)$

$= P(N(t) = 0)$ (ανεξάρτητες προσαυξήσεις)

$= e^{-\lambda t}$.

Όμοια με την περίπτωση της X_1 συμπεραίνουμε, ότι $P(X_n > t) = e^{-\lambda t}$ ή

$$P(X_n \le t) = 1 - e^{-\lambda t},\ t \ge 0 \text{ (και } P(X_n \le t) = 0,\ t < 0).$$

Παρατηρούμε ότι η τ.μ. X_n έχει την αθροιστική συνάρτηση κατανομής της εκθετικής κατανομής με παράμετρο λ, δηλαδή $X_n \sim Exp(\lambda)$. Ειδικότερα, $E(X_n) = 1/\lambda$ και $Var(X_n) = 1/\lambda^2$.

Παράδειγμα

Έστω $\{N(t), t \in [0,\infty)\}$ μια διεργασία Poisson με ρυθμό $\lambda = 6$ αφίξεις / min. Να βρεθεί το χρονικό διάστημα που αναμένεται να περάσει μεταξύ της 3ης και της 4ης άφιξης.

Λύση

Αν X_4 το χρονικό διάστημα μεταξύ της 3ης και της 4ης άφιξης, τότε $X_4 \sim \text{Poisson}(8)$ και

$E(X_4) = 1/6 \text{ min} = 10 \text{ sec}$.

Σύνοψη

Αν X το χρονικό διάστημα που μεσολαβεί μεταξύ δύο (οποιωνδήποτε) γεγονότων, τότε

$$X \sim \text{Exp}(\lambda),$$

όπου λ το αναμενόμενο πλήθος αφίξεων στη μονάδα του χρόνου. Η αναμενόμενη τιμή της χρονικής διάρκειας μεταξύ δύο αφίξεων είναι

$$E(X) = 1/\lambda$$

ενώ η διακύμανση αυτής είναι

$$\text{Var}(X) = 1/\lambda^2.$$

13.4 Χρόνος για την άφιξη, δεδομένου πως δεν αυτή δεν έχει συμβεί για χρόνο t_0

Σε μία διεργασία Poisson, **ο αριθμός των αφίξεων** που συμβαίνει σε μη επικαλυπτόμενα διαστήματα είναι **ανεξάρτητος ο ένας με τον άλλο** (ιδιότητα ανεξάρτητων προσαυξήσεων που αποδίδεται στη διεργασία Poisson από τον ορισμό της). Αυτή η ανεξαρτησία αντανακλάται και **στους χρόνους μεταξύ των αφίξεων**, καθώς αποδεικνύεται ότι:

Θεώρημα: Έστω X η τυχαία μεταβλητή του χρόνου μεταξύ δύο αφίξεων σε μία διεργασία Poisson. Για κάθε t_0, $x \geq 0$, είναι

$$P(X > x + t_0 \mid X > t_0) = P(X > x).$$

Απόδειξη

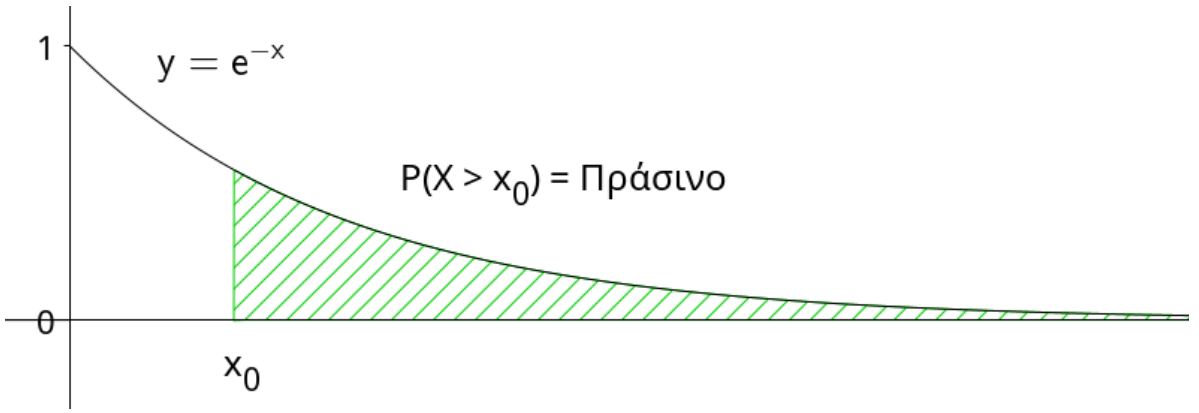
$P(X > x + t_0 \mid X > t_0) =$

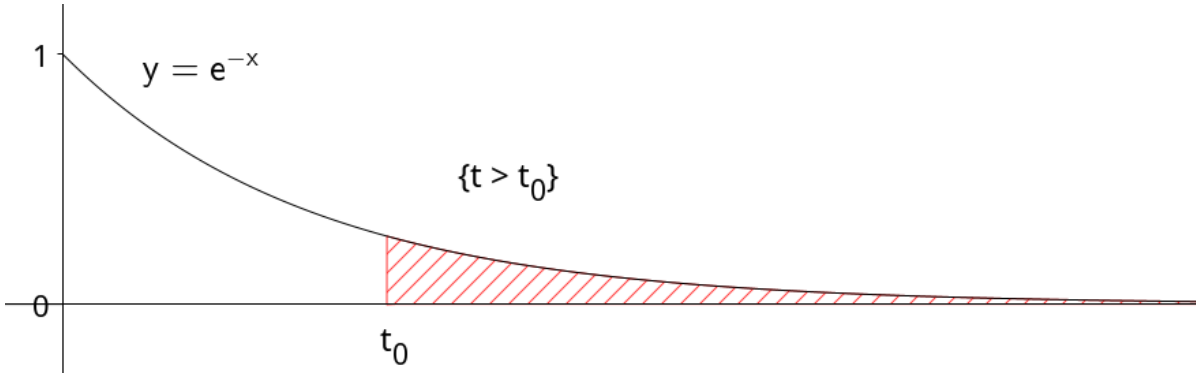
$= P(X > x + t_0, X > t_0) / P(X > t_0)$

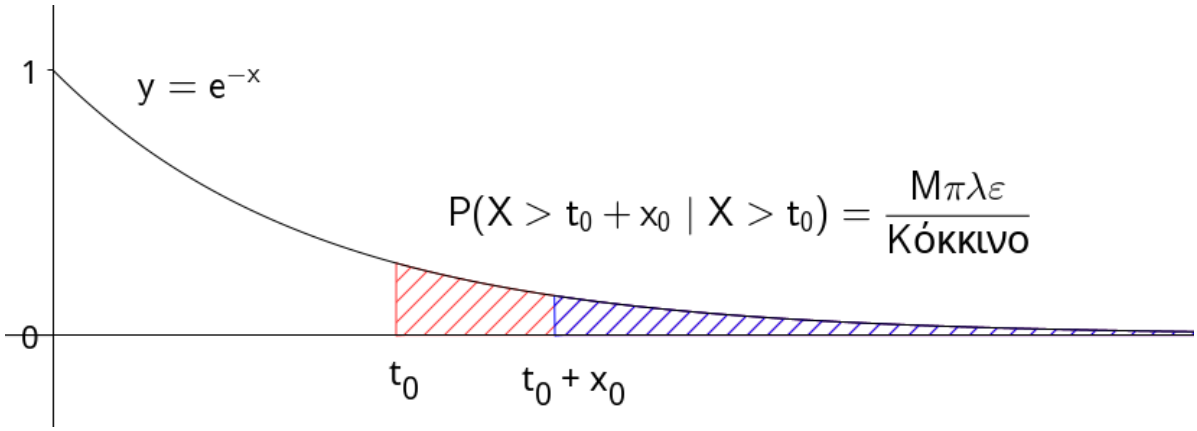
$= P(X > x + t_0) / P(X > t_0)$

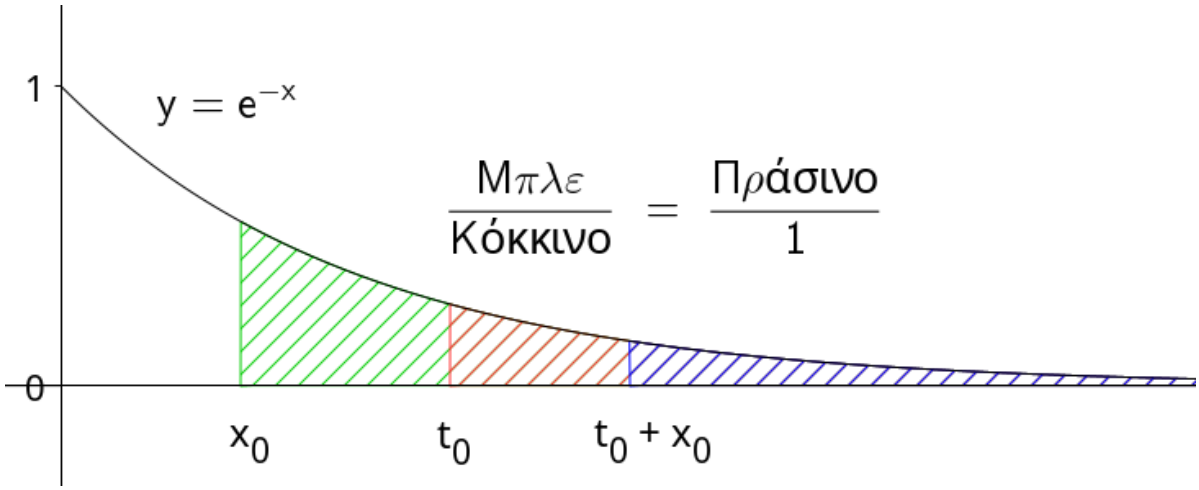
$= e^{-\lambda(x + t0)} / e^{-\lambda t0} = e^{-\lambda x}$

$= P(X > x).$


1
$y = e^{-x}$
$P(X > x_0)$ = Πράσινο
0
x_0


1
$y = e^{-x}$
$\{t > t_0\}$
0
t_0


1
$y = e^{-x}$
$P(X > t_0 + x_0 \mid X > t_0) = \frac{\text{Μπλε}}{\text{Κόκκινο}}$
0
t_0
$t_0 + x_0$


1
$y = e^{-x}$
$\frac{\text{Μπλε}}{\text{Κόκκινο}} = \frac{\text{Πράσινο}}{1}$
0
x_0
t_0
$t_0 + x_0$

Παρατήρηση

Το τελευταίο θεώρημα φαίνεται μη ρεαλιστικό, γιατί στην καθημερινότητα, αν ένα γεγονός έχει ήδη καθυστερήσει κάποιο χρονικό διάστημα τότε περιμένουμε να συμβεί πιο σύντομα στη συνέχεια, δηλαδή περιμένουμε η κατανομή του χρόνου πραγματοποίησης να έχει αλλάξει. Αυτή η αντίληψη υπάρχει διότι στις περισσότερες διαδικασίες αναμονής στην πραγματική ζωή, το θεωρητικό χρονικό διάστημα που αυτό θα συμβεί δεν είναι άπειρο όπως στην περίπτωση της εκθετικής κατανομής. Ωστόσο, το θεώρημα αφορά και ισχύει για την περίπτωση όπου μπορούμε να δεχθούμε ότι είναι δυνατόν η αναμονή να έχει απεριόριστα μεγάλες τιμές.

Ένα αξιοσημείωτο συμπέρασμα είναι πως η συνάρτηση πυκνότητας πιθανότητας της εκθετικής κατανομής $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$, είναι η μόνη συνάρτηση με “αυτοόμοια” γραφική παράσταση, υπό την έννοια ότι ο λόγος του μέρους $[0, x]$ προς το όλο 1, είναι όσο ο λόγος του $[t_0, x + t_0]$ προς το $[t_0, +\infty]$. Σε όρους πιθανότητας και δεσμευμένης πιθανότητας, η τελευταία ισότητα των λόγων, μεταφράζεται ακριβώς στο συμπέρασμα του θεωρήματος (Διαγράμματα προηγούμενης σελίδας).

Σημείωση: Μία απόδειξη για τη μοναδικότητα είναι διαθέσιμη εδώ:
https://math.stackexchange.com/questions/1664347/uniqueness-of-memoryless-property

Παράδειγμα

Έστω $N(t)$ μια διεργασία Poisson με μέσο ρυθμό $\lambda = 2$ συμβάντα / λεπτό και έστω $X_1, X_2, \ldots,$, οι αντίστοιχοι χρόνοι μεταξύ των γεγονότων.

(α) Να βρείτε την πιθανότητα η πρώτη άφιξη να καθυστερήσει περισσότερο από μισό λεπτό.

(β) Αν γνωρίζουμε ότι δεν συνέβη κάποιο γεγονός έως το 1ο λεπτό, να βρεθεί η πιθανότητα το 1ο γεγονός να συμβεί μετά το 3ο λεπτό.

(γ) Αν γνωρίζουμε ότι το 3ο γεγονός έγινε το 2ο λεπτό, να βρεθεί η πιθανότητα, το 4ο γεγονός να συμβεί μετά το 4ο λεπτό.

Ξεκινούμε να παρακολουθούμε τη διαδικασία το 10ο λεπτό. Έστω T η χρονική στιγμή που συμβαίνει το πρώτο γεγονός, μετά το 10ο λεπτό.

(δ) Να βρεθεί η αναμενόμενη τιμή και η διακύμανση της τ.μ. T.

Λύση

(α) Είναι $X_1 \sim Exp(\lambda)$ και $\lambda = 2$, άρα $X_1 \sim Exp(2)$. Υπολογίζουμε

$$P(X_1 > 0{,}5) = e^{-2\cdot 0{,}5}$$

$$= e^{-1} = 0{,}368$$

$$= 36{,}8\%.$$

(β) $P(X_1 > 3 \mid X_1 > 1) = P(X_1 > 3 - 1)$ (Θεώρημα 13.4)

$= P(X_1 > 2)$

$= e^{-2\cdot 2} = e^{-4} = 0{,}0183 = 1{,}83\%$

(γ) Αν X_3 ο χρόνος που μεσολαβεί μεταξύ 2η και 3η μέτρηση τότε $X_3 \sim Exp(2)$. Αν X_4 ο χρόνος που μεσολαβεί μεταξύ 3η και 4η μέτρηση τότε $X_4 \sim Exp(2)$.

{το 4ο γεγονός συμβαίνει μετά το 4ο λεπτό} = $\{X_1 + X_2 + X_3 + X_4 > 4\}$

{το 3ο γεγονός έγινε το 2ο λεπτό} = $\{X_1 + X_2 + X_3 = 2\}$

P({το 4ο γεγονός συμβαίνει μετά το 4ο λεπτό} | {το 3ο γεγονός έγινε το 2ο λεπτό}) =

$= P(X_1 + X_2 + X_3 + X_4 > 4 \mid X_1 + X_2 + X_3 = 2)$

$= P(X_4 > 2)$

$= e^{-2\cdot 2} = e^{-4} = 0{,}0183 = 1{,}83\%$

(δ) Μετά, το 10ο λεπτό παρακολουθούμε μία διεργασία Poisson με παράμετρο λ = 2. Συνεπώς, αν

T = {χρονικό διάστημα μέχρι το 1ο γεγονός (μετά το 10ο λεπτό)}

είναι τότε $T = 10 + X$, όπου $X \sim Exp(2)$.

Άρα, $E(T) = E(10 + X) = 10 + E(X) = 10 + 1/\lambda = 10 + ½ = 10{,}5$ και

$Var(T) = Var(10 + X) = Var(X) = 1/\lambda^2 = 1/2^2 = 1/4$.

13.5 Ομοιομορφία χρόνου δεδομένης της άφιξης

Θεώρημα

Έστω $\{N(t), t \in [0,\infty)\}$ μια διεργασία Poisson με ρυθμό λ. **Αν γνωρίζουμε πως έχει συμβεί μία άφιξη μέχρι τη χρονική στιγμή t (N(t) = 1)**, τότε, η χρονική στιγμή x που αυτή έχει συμβεί, $0 \le x \le t$, έχει ομοιόμορφη κατανομή στο (0, t]. Δηλαδή, αν X η χρονική στιγμή που έγινε η 1η άφιξη, τότε:

$$P(X \le x \mid N(t) = 1) = x/t,\ 0 \le x \le t.$$

Απόδειξη

Γνωρίζουμε ότι $N(t) \sim \text{Poisson}(\lambda \cdot t)$, άρα $P(N(t) = 1) = e^{-\lambda t} (\lambda t)^1/1! = \lambda \cdot t \cdot e^{-\lambda t}$. Είναι:

$$\begin{aligned} P(X \le x, N(t) = 1) &= P(\{1 \text{ άφιξη στο } (0, x] \text{ και } 0 \text{ αφίξεις στο } (x, t]\}) \\ &= P(\{1 \text{ άφιξη στο } (0, x]) \cdot P(\{0 \text{ αφίξεις στο } (x, t]\}) \\ &= P(\{1 \text{ άφιξη στο } (0, x]) \cdot P(\{0 \text{ αφίξεις στο } (0, t - x]\}) \\ &= e^{-\lambda x} (\lambda x)^1/1! \cdot e^{-\lambda(t - x)} (\lambda(t - x))^0/0! = \lambda \cdot x \cdot e^{-\lambda t}. \end{aligned}$$

Τώρα, $P(X \le x \mid N(t) = 1) = P(X \le x, N(t) = 1) / P(N(t) = 1) = \lambda \cdot x \cdot e^{-\lambda t} / \lambda \cdot t \cdot e^{-\lambda t} = x/t$.

Παρατηρήσεις

1. Το αποτέλεσμα αυτό δεν έρχεται σε αντίθεση με το $X_1 \sim \text{Exp}(\lambda)$ το οποίο εκφράζει το χρόνο που **αναμένεται να συμβεί** η 1η άφιξη. Γνωρίζοντας όμως ότι η πρώτη άφιξη συμβαίνει έως τη χρονική στιγμή t, η πιθανότητα να έχει συμβεί μέσα στο (0, t] κατανέμεται ομοιόμορφα σε αυτό.

2. Το παραπάνω αποτέλεσμα γενικεύεται για n αφίξεις. Δηλαδή, γνωρίζοντας ότι έχουν συμβεί n αφίξεις στο διάστημα (0, t] (N(t) = n), οι n χρόνοι άφιξης κατανέμονται όπως προβλέπεται από την κοινή συνάρτηση κατανομής n ανεξάρτητων ομοιόμορφων U(0,t) τυχαίων μεταβλητών, δηλαδή η κοινή συνάρτηση πυκνότητας πιθανότητας είναι η

$$f_{T_1, T_2, \ldots, T_n \mid N(t) = n}(t_1, t_1, \ldots, t_n) = \frac{n!}{t^n}.$$

Παράδειγμα

Αν γνωρίζουμε ότι την πρώτη ώρα λειτουργίας ενός εστιατορίου έχει έρθει ένας πελάτης, να βρεθεί η πιθανότητα αυτός

(α) να έχει έρθει στα πρώτα 10 λεπτά.

(β) να έχει έρθει μεταξύ του 25 και του 40 λεπτού της ώρας.

Λύση

Έστω N(t) = {πλήθος πελατών στο διάστημα (0, t]} και t: ο χρόνος σε λεπτά.

Γνωρίζουμε ότι ο πελάτης **έχει αφιχθεί** μέσα στα πρώτα 60 λεπτά, δηλαδή ότι N(60) = 1.

Αν X η τ.μ. που αναπαριστά τη στιγμή της άφιξης μέσα στο διάστημα (0, 60], τότε η τ.μ. X έχει κατανομή U(0, 60), άρα:

(α) $P(X \leq 10) = 10/60 = 1/6 = 16,7\%$.

(β) $P(25 \leq X \leq 40) = P(X \leq 40) - P(X \leq 25) = (40 – 25)/60 = 15/60 = 0,25 = 25\%$

Παράδειγμα

Να βρεθεί η πιθανότητα, δύο πελάτες που έχουν αφιχθεί σε ένα κατάστημα τα πρώτα δύο λεπτά από τη στιγμή που άνοιξε, να έχουν αφιχθεί στο πρώτο από τα δύο λεπτά. Οι αφίξεις των πελατών θεωρούνται ανεξάρτητες μεταξύ τους.

Λύση

Γνωρίζουμε ότι οι δύο πελάτες A και B **έχουν αφιχθεί** στα πρώτα δύο λεπτά. Κάθε πελάτης μπορεί να έχει μπει είτε στο 1ο είτε στο 2ο λεπτό και ο χρόνος προσέλευσης του κάθε ενός ακολουθεί την ομοιόμορφη κατανομή, άρα έχουν αφιχθεί κατά τη διάρκεια του 1ου λεπτού με πιθανότητα 50% στο κάθε ένα. Καθώς, οι αφίξεις των πελατών είναι ανεξάρτητες μεταξύ τους, βρίσκουμε ότι η πιθανότητα και οι δύο πελάτες να έχουν μπει στο κατάστημα κατά τη διάρκεια του πρώτου λεπτού είναι $0,5 \cdot 0,5 = 0,25 = 25\%$.

Σημείωση

Εναλλακτικά και ισοδύναμα μπορούμε να δούμε ότι από τους συνολικά τέσσερις πιθανούς συνδυασμούς προσέλευσης στα δύο λεπτά ((AB, 0), (A, B), (B, A), (0, AB)) ένας είναι ο επιθυμητός κατά τον οποίο και οι δύο μπαίνουν στο κατάστημα κατά τη διάρκεια του 1ου λεπτού.

13.6 Άθροισμα ανεξάρτητων εκθετικών τυχαίων μεταβλητών

Ο νόμος ολικής πιθανότητας μπορεί να εκφραστεί και στην περίπτωση συνεχών μεταβλητών. Ειδικότερα, αξιοποιώντας τον ορισμό $f_{X|Y}(x|y) = f_{X,Y}(x,y) / f_Y(y)$, γράφουμε:

Νόμος Ολικής Πιθανότητας για Συνεχείς Μεταβλητές

Αν X, Y, συνεχείς τυχαίες μεταβλητές, τότε

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \int_{-\infty}^{\infty} f_{X|Y}(x|y) f_Y(y)\,dy$$

Πρόταση

Αν X, Y ανεξάρτητες τυχαίες μεταβλητές, X, Y ~ Exp(λ) και Z = X + Y, τότε

$$f_Z(z) = \lambda^2 z e^{-\lambda z},\ z > 0.$$

Απόδειξη

$$f_Z(z) = \int_{-\infty}^{\infty} f_{Z|X}(z|x) f_X(x)\,dy = \int_{-\infty}^{\infty} f_Y(z-x) f_X(x)\,dy$$

$$= \int_0^z \lambda e^{-\lambda(z-x)} \lambda e^{-\lambda x}\,dx = \lambda^2 e^{-\lambda z} \int_0^z dx = \lambda^2 z e^{-\lambda z} \quad \text{ή } Z \sim \text{Erlang}(2, \lambda).$$

Γενικότερα, αποδεικνύεται ότι:

Θεώρημα (κατανομή Erlang)

Αν $X = X_1 + X_2 + \ldots + X_\kappa$ όπου $X_i \sim \text{Exp}(\lambda)$, i = 1, 2,..., κ, τότε

$$P(X \le x) = 1 - \sum_{n=0}^{\kappa-1} \frac{1}{n!} e^{-\lambda x} (\lambda x)^n,\ x \ge 0 \quad \text{και} \quad f(x) = \frac{\lambda^n \cdot x^{n-1} \cdot e^{-\lambda x}}{(n-1)!},\ x > 0.$$

Η παραπάνω κατανομή ονομάζεται κατανομή **Erlang(κ, λ)**.

Σημειώσεις

1. Περισσότερες πληροφορίες για τη δεσμευμένη συνάρτηση μάζας πιθανότητας μπορούν να βρεθούν εδώ: **https://math.stackexchange.com/questions/2035418/can-we-prove-the-law-of-total-probability-for-continuous-distributions**

2. Μία απόδειξη για το τελευταίο θεώρημα είναι διαθέσιμη εδώ: **https://math.stackexchange.com/questions/250733/how-is-the-erlang-pdf-derived**

3. Η σχέση της εκθετικής κατανομής με την κατανομή Erlang είναι ανάλογη με τη σχέση που έχει η γεωμετρική κατανομή (πιθανότητα πρώτης επιτυχίας), με την αρνητική διωνυμική κατανομή (πιθανότητα πλήθους δοκιμασιών μέχρι τις n επιτυχίες)

4. Η κατανομή Erlang αποτελεί ειδική περίπτωση της κατανομής Γάμμα, καθώς

Γάμμα(α, β) ≡ Erlang(α, β), για α, β ϵ N

13.7 Χρόνος μέχρι το n – οστό γεγονός

Έστω T_1, T_2, …T_n, …, ο χρόνος μέχρι την υλοποίηση του $1^{ου}$, $2^{ου}$, …, n – οστού, … γεγονότος. Αν X_1: ο χρόνος που περνάει μέχρι να συμβεί το $1^{ο}$ γεγονός, και X_n: ο χρόνος που περνάει μεταξύ του n – 1 και του n γεγονότος, τότε:

$T_1 = X_1$,

$T_2 = X_1 + X_2$

$T_3 = X_1 + X_2 + X_3$

….

$T_n = X_1 + X_2 + \ldots + X_n$

Οι τ.μ. T_1, T_2, …T_n, …, δεν είναι ανεξάρτητες καθώς $T_1 \le T_2 \le \ldots \le T_n \le \ldots$. Ωστόσο, κάθε

$$T_n = X_1 + X_2 + \ldots + X_n,\ n = 1, 2, \ldots,$$

είναι μία συνεχής τ.μ. που προκύπτει ως άθροισμα ανεξάρτητων τυχαίων μεταβλητών που ακολουθούν την Exp(λ). Άρα, T_n ~ Erlang(n, λ) (κατανομή Erlang), με συνάρτηση πυκνότητας πιθανότητας

$$f_n(t) = \frac{\lambda^n \cdot t^{n-1} \cdot e^{-\lambda t}}{(n-1)!},\ t > 0.$$

και συνάρτηση κατανομής

$$F_n(t) = P(T_n \le t) = 1 - \sum_{k=0}^{n-1} \frac{1}{k!} e^{-\lambda t} (\lambda t)^k,\ x \ge 0$$

Επιπλέον,

- $E(T_n) = E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n) = n/\lambda$.
- $Var(T_n) = Var(X_1 + X_2 + \ldots + X_n)$

  $= Var(X_1) + Var(X_2) + \ldots + Var(X_n)$

  $= n/\lambda^2$ (X_i, i = 1, 2, … ανεξάρτητες).

Σύνοψη

Αν X είναι ο χρόνος μεταξύ της εμφάνισης γεγονότων σε μία Poisson διεργασία, τότε X ~ Exp(λ) (εκθετική κατανομή με παράμετρο λ). Έστω T_n η χρονική στιγμή κατά την οποία το νιοστό γεγονός συμβαίνει σε μία Poisson διεργασία. Η αναμενόμενη τιμή της T_n είναι $E(T_n) = n/\lambda$ και η διακύμανση αυτής $Var(T_n) = n/\lambda^2$.

Παράδειγμα

Ερωτήσεις καταφθάνουν σε ένα μηχανισμό καταγραφής μηνυμάτων σύμφωνα με μία Poisson διεργασία με συχνότητα 15 ερωτήσεις το λεπτό.

(α) Να βρεθεί η πιθανότητα ότι σε μία περίοδο 1 λεπτού, 3 ερωτήσεις θα φθάσουν κατά τα πρώτα 10 δευτερόλεπτα και 2 κατά τα τελευταία 15 δευτερόλεπτα.

(β) Να υπολογιστεί η πιθανότητα, ο χρόνος έως τη δεύτερη ερώτηση να είναι μεγαλύτερος από τα 10 sec.

(γ) Να υπολογισθεί η μέση τιμή και η τυπική απόκλιση του χρόνου μέχρι την άφιξη της $10^{ης}$ ερώτησης.

Λύση

Είναι λ = 15 / 60 = 0,25 ερωτήσεις / sec, άρα $P(N(t) = k) = e^{-0,25t} (0,25t)^k/k!$

(α) Τα χρονικά διαστήματα (0, 10] και (45, 60] δεν επικαλύπτονται άρα το πλήθος των ερωτήσεων στο ένα διάστημα είναι ανεξάρτητο με το πλήθος ερωτήσεων του άλλου. Η ζητούμενη πιθανότητα είναι η

$$P(N(10) = 3, N(60) - N(45) = 2) = P(N(10) = 3) \cdot P(N(15) = 2)$$

$$= e^{-2,5}\, 2,5^3/3! \cdot e^{-3,75}\, 3,75^2/2!$$

$$= 0,2138 \cdot 0,1654$$

$$\approx 0,035 = 3,5\%.$$

(β) Αν X ο χρόνος μέχρι τη 2η ερώτηση τότε $X = X_1 + X_2$, όπου X_i ~ Exp(0,25), i = 1, 2, άρα τότε X ~ Erlang(2, 0.25) και

$$P(X > 10) = \sum_{n=0}^{1} \frac{1}{n!} e^{-0,25 \cdot 10} (0,25 \cdot 10)^n = 3,5 \cdot e^{-2,5} = 0,287.$$

Σημείωση: Εναλλακτικά, P(X > 10) = P(N(10) = 0) + P(N(10) = 1) = 0,287.

(γ) Είναι $E(T_{10}) = 10 \cdot E(X) = 10 / \lambda = 40$ sec και

$Var(T_{10}) = 10 \cdot Var(X) = 10 / \lambda^2 = 160\ sec^2$, άρα $StDev(T_{10}) = 160^{0,5} = 12,6$ sec.

13.8 Άθροισμα ανεξάρτητων Poisson μεταβλητών

Πρόταση

Το άθροισμα δύο **ανεξάρτητων** τυχαίων μεταβλητών X ~ Poisson(λ) και Y ~ Poisson(μ) είναι ομοίως κατανομή Poisson με παράμετρο λ + μ, δηλαδή, αν Z ~ X + Y, και X, Y ανεξάρτητες τυχαίες μεταβλητές τότε Z ~ Poisson(λ + μ).

Απόδειξη

Αρκεί να δείξουμε ότι $P(Z = z) = e^{-\lambda+\mu}\cdot(\lambda + \mu)^z/z!$, για κάθε z ϵ N. Πράγματι, είναι

$P(Z = z) = \Sigma_{j = 0, 1, \ldots} P(Z = z, P(X = j)$ (Νόμος Ολικής Πιθανότητας)

$= \Sigma_{j = 0, 1, \ldots, z} P(Y = z - j, X = j)$ (αν j > z, τότε P(Z = z | X = j) = 0)

$= \Sigma_{j = 0, 1, \ldots, z} P(Y = z - j) \cdot P(X = j)$ (οι X, Y, είναι ανεξάρτητες τ.μ.)

$= \Sigma_{j = 0, 1, \ldots, z} e^{-\mu}\cdot\mu^{z - j}/(z - j)! \cdot e^{-\lambda}\cdot\lambda^j/j!$

$= \Sigma_{j = 0, 1, \ldots, z} \mu^{z - j} / [j!\cdot(z - j)!] \cdot e^{-(\lambda+\mu)}.\lambda^j$

$= e^{-(\lambda+\mu)}.\Sigma_{j = 0, 1, \ldots, z}$ (z ανά j) $\cdot \lambda^j \cdot \mu^{z - j} / z!$

$= e^{-(\lambda+\mu)}.(\lambda + \mu)^z / z!$

Παρατηρήσεις

1. Η υπόθεση της (στοχαστικής) ανεξαρτησίας των X, Y είναι ουσιαστική. Αν οι X, Y δεν είναι ανεξάρτητες τότε Cov(X, Y) > 0 και κατά συνέπεια θα είναι

$$Var(X + Y) = Var(X) + 2Cov(X, Y) + Var(Y) > Var(X) + Var(Y) = E(X + Y),$$

δηλαδή, η τ.μ. X + Y δεν μπορεί να ακολουθεί την κατανομή Poisson στην οποία $\mu = \sigma^2$.

2. Η παραπάνω πρόταση δεν αποκλείει το ενδεχόμενο το άθροισμα δύο **εξαρτημένων** Poisson μεταβλητών να είναι Poisson μεταβλητή. Ένα παράδειγμα αναφέρεται στη δημοσίευση [Jacod, J. (1975). Two Dependent Poisson Processes Whose Sum Is Still a Poisson Process. Journal of Applied Probability, 12(1), 170–172. https://doi.org/10.2307/3212423]

13.9 Συγχώνευση διεργασιών Poisson

Πολλές φορές είναι απαραίτητη η αναγνώριση δύο ανεξάρτητων διεργασιών καταμέτρησης ως μία κοινή διεργασία καταμέτρησης, όπως για παράδειγμα όταν σε ένα αεροδρόμιο οι επιβάτες από δύο διαφορετικά εκδοτήρια εισιτηρίων, περνάνε από τον ίδιο έλεγχο αποσκευών.

Αντίστροφα, είναι δυνατό να απαιτηθεί ο διαχωρισμός μίας διεργασίας καταμέτρησης σε δύο μέρη, όταν για παράδειγμα υπάρχουν δύο ταμεία σε ένα εμπορικό κατάστημα για να εξυπηρετήσουν μία ουρά στην οποία οι αφίξεις γίνονται σύμφωνα με μία κατανομή Poisson.

Θα αποδείξουμε ότι εάν τα γεγονότα που περιγράφονται από μια διεργασία Poisson χωριστούν σε δύο νέες διεργασίες με τυχαία κατανομή των συμβάντων στις δύο διεργασίες, τότε οι νέες διεργασίες είναι διεργασίες Poisson και είναι ανεξάρτητες μεταξύ τους.

Η πιο χρήσιμη συνέπεια αυτού του γεγονότος είναι ότι (αντίστροφα σκεπτόμενοι) οποιεσδήποτε δύο ανεξάρτητες διεργασίες Poisson με ρυθμούς λ_1 και λ_2 μπορούν να θεωρηθούν ότι προέρχονται από μια ενιαία διεργασία με ρυθμό $\lambda_1 + \lambda_2$.

Ορισμός

Δύο διακριτές διεργασίες συνεχούς χρόνου $\{N_1(t), t \geq 0\}$ και $\{N_2(t), t \geq 0\}$ ονομάζονται ανεξάρτητες εάν για όλους τους θετικούς ακέραιους αριθμούς k και όλα τα $0 < t_1 < t_2 < \ldots < t_k$, οι τυχαίες μεταβλητές $N_1(t_1),\ldots,N_1(t_k)$ είναι ανεξάρτητες από τις $N_2(t_1),\ldots,N_2(t_k)$.

Αποδεικνύεται το παρακάτω θεώρημα:

Θεώρημα

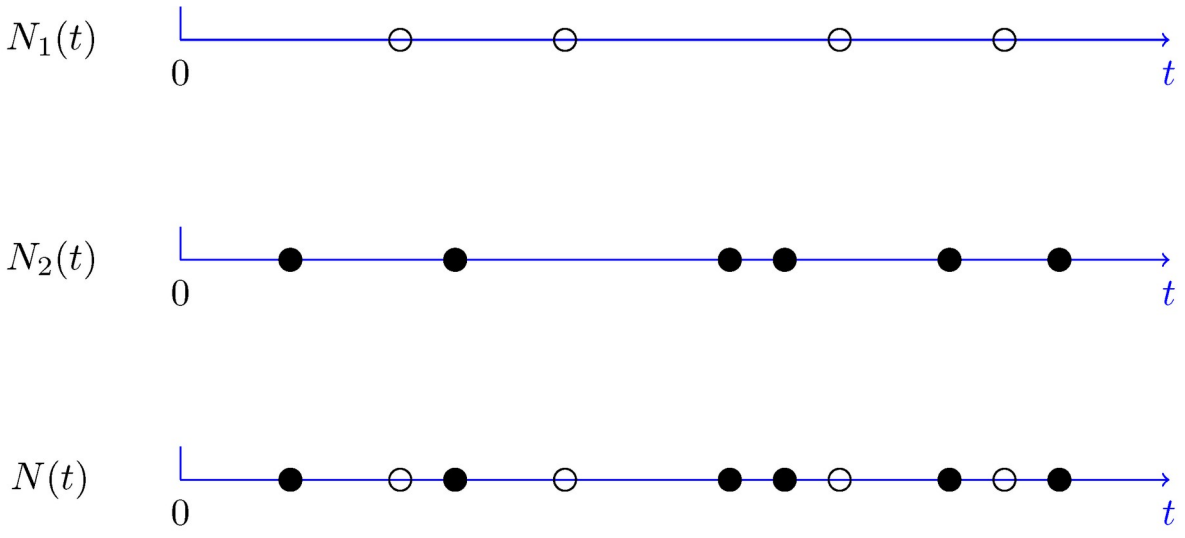
Οι διεργασίες καταμέτρησης $\{N_1(t), t \geq 0\}$ και $\{N_2(t), t \geq 0\}$ είναι ανεξάρτητες αν και μόνο αν οι τυχαίες μεταβλητές που εκφράζουν τα διαστήματα μεταξύ των γεγονότων για τη $N_1(t)$ είναι στοχαστικά ανεξάρτητες από τις αντίστοιχες μεταβλητές που εκφράζουν τα διαστήματα μεταξύ αφίξεων για τη $N_2(t)$.

Η απόδειξη παραλείπεται.

Έστω τώρα, $N_1(t)$ και $N_2(t)$ δύο *ανεξάρτητες* διεργασίες Poisson με ρυθμούς λ_1 και λ_2 αντίστοιχα. Ορίζουμε

$$N(t) = N_1(t) + N_2(t).$$

Η τυχαία διεργασία N(t) είναι το άθροισμα των γεγονότων στο διάστημα (0, t], που συμβαίνουν στην $N_1(t)$ και $N_2(t)$.



Πρόταση

Η N(t) = N_1(t) + N_2(t) είναι μια διαδικασία Poisson με ρυθμό λ = $\lambda_1 + \lambda_2$.

Απόδειξη

Πράγματι,

(α) N(0) = N_1(0) + N_2(0) = 0 + 0 = 0.

(β) Καθώς η N_1(t) και η N_2(t) είναι ανεξάρτητες και έχουν ανεξάρτητες προσαυξήσεις, συμπεραίνουμε ότι η N(t) έχει επίσης ανεξάρτητες προσαυξήσεις.

(γ) Αν I = (t, t + τ], τ > 0, τότε το πλήθος γεγονότων στο I που σχετίζονται με τις N_1(t) και N_2(t) είναι τυχαίες μεταβλητές που ακολουθούν τις κατανομές Poisson(λ_1τ) και Poisson(λ_2τ), συνεπώς, ο αριθμός των αφίξεων στο I που σχετίζεται με τη N(t) είναι Poisson((λ_1+λ_2)τ) (άθροισμα δύο ανεξάρτητων τυχαίων μεταβλητών Poisson).

Γενικότερα, ισχύει ότι:

Έστω N_1(t), N_2(t), … , N_m(t), m ανεξάρτητες διαδικασίες Poisson με ρυθμούς λ_1, λ_2, …, λ_m. Αν,

$$N(t) = N_1(t) + N_2(t) + \ldots + N_m(t),\ t \geq 0.$$

Τότε, η N(t) είναι μια διεργασία Poisson με ρυθμό $\lambda_1 + \lambda_2 + \ldots + \lambda_m$.

Παράδειγμα

Έστω $N_1(t)$ και $N_2(t)$ δύο ανεξάρτητες διεργασίες Poisson με ρυθμούς $\lambda_1 = 1$ και $\lambda_2 = 2$ συμβάντα / χρονική μονάδα και $N(t) = N_1(t) + N_2(t)$ η συγχωνευμένη διεργασία.

(α) Να βρείτε την πιθανότητα να συμβούν στη συγχωνευμένη διεργασία, 5 γεγονότα στο διάστημα (0, 2] από τα οποία τα 2 γεγονότα να συμβούν στο διάστημα (0, 1].

(β) Γνωρίζοντας ότι έχουν συμβεί 2 γεγονότα στη συγχωνευμένη διεργασία στο χρονικό διάστημα (0, 1] να βρεθεί η πιθανότητα 1 από αυτά να έχει έρθει από την $N_1(t)$.

Λύση

Είναι $N(t) \sim \text{Poisson}(\lambda_1 + \lambda_2) = \text{Poisson}(\lambda)$, όπου $\lambda = 3$.

(α) P(N(1) = 2, N(2) = 5) = P(2 γεγονότα στο [0, 1) και άλλα 3 γεγονότα στο [1, 2)) =

= P(2 γεγονότα στο [0, 1)) · P(3 γεγονότα στο [1, 2)) =

= P(2 γεγονότα στο [0, 1)) · P(3 γεγονότα στο [0, 1)) =

$= e^{-\lambda} \cdot \lambda^2/2! \cdot e^{-\lambda} \cdot \lambda^3/3! \approx 0{,}05 = 5\%$.

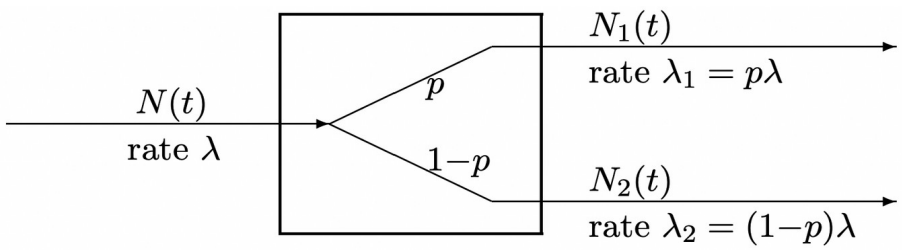
(β) $P(N_1(1) = 1 \mid N(1) = 2)$

$= P(N_1(1) = 1, N(1) = 2) / P(N(1) = 2)$

$= P(N_1(1) = 1) \cdot P(N_2(1) = 1) / P(N(1) = 2) =$

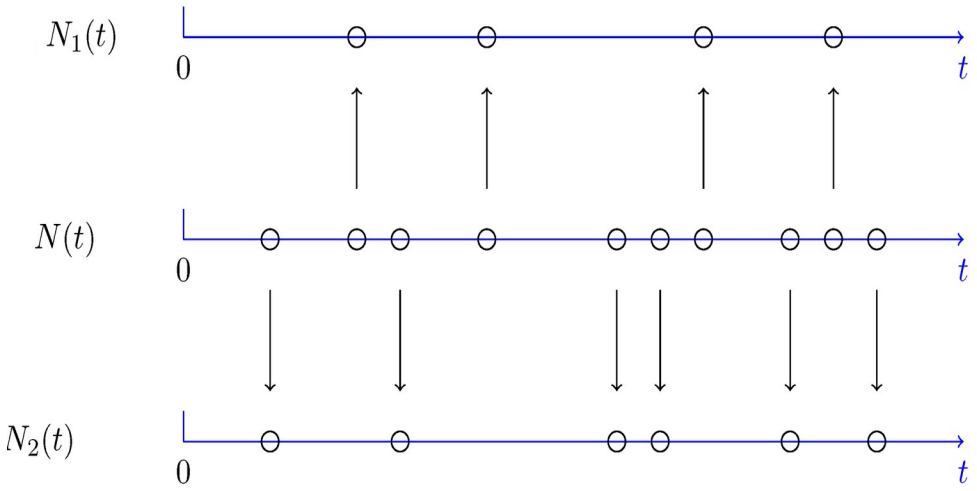
$= e^{-1} \cdot 2e^{-2} / e^{-3} \cdot 3^2/2! = 4/9 = 0{,}444\ldots \approx 44{,}4\%$

13.10 Διαίρεση διεργασιών Poisson



Θεώρημα (διαίρεσης διεργασίας Poisson)

Έστω $\{N(t), t \geq 0\}$ διαδικασία Poisson με ρυθμό λ της οποίας κάθε ένα γεγονός, μοιράζεται σε δύο άλλες διεργασίες $N_1(t)$ και $N_2(t)$ με πιθανότητα p και 1 – p αντίστοιχα. Τότε, οι διεργασίες $\{N_1(t), t \geq 0\}$ και $\{N_2(t), t \geq 0\}$ είναι δύο ανεξάρτητες μεταξύ τους διεργασίες Poisson με παραμέτρους λp και λ(1 – p) αντίστοιχα.



Παράδειγμα (αντί απόδειξης)

Ο αριθμός των πελατών N(t) που επισκέπτονται ένα εστιατόριο στο διάστημα (0, t], είναι διεργασία Poisson με παράμετρο μ. Η πιθανότητα ένας πελάτης να καταναλώσει και ποτό μαζί με το φαγητό του είναι p, ανεξάρτητα από άλλους πελάτες και ανεξάρτητα από την τρέχουσα τιμή του N(t). Έστω X(t) ο αριθμός των πελατών που αγοράζουν ποτό στο χρονικό διάστημα (0, t] και Y(t) ο αριθμός των πελατών που δεν αγοράζουν ποτό στο ίδιο διάστημα (δηλαδή, X(t) + Y(t) = N(t)).

(α) Να δείξετε ότι οι X(t), Y(t) είναι διεργασίες Poisson με ρυθμό pμ και (1 – p)μ αντίστοιχα.

(β) Να δείξετε ότι οι τ.μ. X(t) και Y(t) είναι ανεξάρτητες.

Λύση

(α) Φανερά, X(0) = Y(0) = 0. Επιπλέον, οι προσαυξήσεις των X(t), Y(t) αποτελούν ένα μέρος των προσαυξήσεων της N(t)) και η N(t) είναι διεργασία Poisson άρα τόσο η X(t) όσο και η Y(t) θα έχουν ομοίως ανεξάρτητες προσαυξήσεις.

Απομένει να δείξουμε ότι για δοσμένο χρονικό διάστημα Δ, είναι X(t) ~ Poisson(μpΔ) και Y(t) ~ Poisson(μ(1-p)Δ). Φανερά, τόσο η τ.μ. X(t) όσο και η Y(t) εξαρτώνται από την τρέχουσα τιμή που θα έχει αποκτήσει η τ.μ. N(t).

Γνωρίζουμε ότι σε δοσμένο διάστημα Δ, είναι N(t) ~ Poisson(μΔ). Αν N(t) = n τότε κάθε ένας από τους n πελάτες είτε παίρνει ποτό με πιθανότητα p είτε δεν παίρνει με πιθανότητα 1 – p και οι X(t), Y(t) μετράνε το πλήθος των αντίστοιχων συμβάντων.

Συνεπώς, είναι:

$$(X(t) \mid N(t) = n) \sim B(n, p) \text{ και } (Y(t) \mid N(t) = n) \sim B(n, 1 - p),$$

δηλαδή $P(X(t) = \kappa \mid N(t) = n) = (n \text{ ανά } \kappa)\, p^{\kappa} \cdot (1 - p)^{n - \kappa}$ και αντίστοιχα,

$$P(Y(t) = \kappa \mid N(t) = n) = (n \text{ ανά } \kappa)\, (1 - p)^{\kappa} \cdot p^{n - \kappa}, \ \kappa = 0, 1, \ldots, n.$$

Για κάθε κ = 0, 1, 2…, είναι

$P_X(\kappa) = P(X(t) = \kappa) = \Sigma_{n = 0, 1, \ldots} P(X(t) = \kappa \mid N(t) = n) \cdot P(N(t) = n)$ (νόμος ολικής πιθανότητας)

$= \Sigma_{n = \kappa, \kappa+1, \ldots} P(X(t) = \kappa \mid N(t) = n) \cdot P(N(t) = n)$ (αν $n < \kappa$, τότε $P(X = \kappa) = 0$)

$= \Sigma_{n = \kappa, \kappa+1, \ldots} (n \text{ ανά } \kappa) \cdot p^{\kappa} \cdot (1- p)^{n - \kappa} \cdot e^{-\mu\Delta} \cdot (\mu\Delta)^n / n!$

$= \Sigma_{n = \kappa, \kappa+1, \ldots} p^{\kappa} \cdot (1- p)^{n - \kappa} \cdot e^{-\mu\Delta} \cdot (\mu\Delta)^n / [\kappa!(n - \kappa)!]$

$= e^{-\mu\Delta} \cdot (\mu\Delta p)^{\kappa} \cdot 1/\kappa! \cdot \Sigma_{n = \kappa, \kappa+1, \ldots} [\mu\Delta(1- p)]^{n - \kappa} / (n - \kappa)!$

$= e^{-\mu\Delta} \cdot (\mu\Delta p)^{\kappa} \cdot 1/\kappa! \cdot \Sigma_{n = 0, 1, \ldots} [\mu\Delta(1- p)]^n / n!$

$= e^{-\mu\Delta} \cdot (\mu\Delta p)^{\kappa} \cdot 1/\kappa! \cdot e^{\mu\Delta(1-p)}$

$= e^{-\mu\Delta p} \cdot (\mu\Delta p)^{\kappa} / \kappa!$

Άρα, $X(t) \sim \text{Poisson}(\mu p\Delta)$. Συμπεραίνουμε ότι η X(t) είναι διεργασία Poisson με ρυθμό μp. Ανάλογα, αποδεικνύεται ότι $Y(t) \sim \text{Poisson}(\mu(1-p)\Delta)$

(β) Αρκεί να δείξουμε ότι $P(X(t) = i, Y(t) = j) = P(X(t) = i) \cdot P(Y(t) = j)$.

Πράγματι, για i, j ϵ N, είναι

$P(X(t) = i, Y(t) = j) =$

$= \Sigma_{n = 0, 1, \ldots} P(X(t) = i, Y(t) = j \mid N(t) = n) \cdot P(N(t) = n)$ (νόμος ολικής πιθανότητας)

$= P(X(t) = i, Y(t) = j \mid N(t) = i + j) \cdot P(N(t) = i + j)$ (είναι $P(X = i. Y = j \mid N = n) = 0$ αν $n \neq i + j$)

$= P(X(t) = i \mid N(t) = i + j) \cdot P(N(t) = i + j)$ (είναι $P(Y = j) = 1$, αν $X = i$ και $N = i + j$)

$= [(i + j) \text{ ανά } i] \cdot p^i \cdot (1 - p)^j \cdot e^{-\mu} \cdot \mu^{i+j} / (i + j)!$ (είναι $(X(t) \mid N(t) = i + j) \sim B(i + j, p)$ και $N(t) \sim \text{Poisson}(\mu t)$)

$= e^{-\mu} \cdot (\mu p)^i \cdot [\mu(1 - p)]^j \cdot / [i! \cdot j!]$

$= e^{-\mu p} \cdot (\mu p)^i / i! \cdot e^{-\mu(1 - p)} \cdot [\mu(1 - p)]^j \cdot / j!$

$= P(X(t) = i) \cdot P(Y(t) = j)$

Δείξαμε ότι $P(X(t) = i, Y(t) = j) = P(X(t) = i) \cdot P(Y(t) = j)$, άρα οι τυχαίες μεταβλητές X(t), Y(t) είναι στοχαστικά ανεξάρτητες.

Σημείωση

Το τελευταίο αποτέλεσμα φαίνεται παράδοξο γιατί οι X(t), Y(t) συνδέονται πάντα με τη σχέση X(t) + Y(t) = N(t). Ωστόσο, αυτή η σχέση δεν προσδιορίζει την τιμή καμίας από τις δύο καθώς η N(t) εξακολουθεί και παίρνει οποιαδήποτε τιμή στους φυσικούς αριθμούς. Ειδικότερα, η ανεξαρτησία τους αντανακλά το γεγονός πως η κατανομή των αφίξεων σε αυτές γίνεται με τυχαίο τρόπο.

Παράδειγμα

Έστω $\{N_1(t), t \geq 0\}$ και $\{N_2(t), t \geq 0\}$ δύο ανεξάρτητες διαδικασίες Poisson με ρυθμούς $\lambda_1 = 1$ και $\lambda_2 = 2$, αντίστοιχα. Βρείτε την πιθανότητα η $2^η$ άφιξη στο $N_1(t)$ να συμβεί πριν από την $3^η$ άφιξη στο $N_2(t)$.

Λύση

Θεωρούμε τη συγχωνευμένη διεργασία Poisson $\{N(t), t \geq 0\}$ με ρυθμό $\lambda = \lambda_1 + \lambda_2 = 1 + 2 = 3$. Οι διεργασίες $\{N_1(t), t \geq 0\}$ και $\{N_2(t), t \geq 0\}$ μπορούν να θεωρηθούν τα μέρη στα οποία διαιρείται η $\{N(t), t \geq 0\}$ αν κάθε γεγονός της αποδοθεί στη $\{N_1(t), t \geq 0\}$ με πιθανότητα $p = 1/3$ και στη $\{N_2(t), t \geq 0\}$ με πιθανότητα $1 – p = 2/3$.

Στην περίπτωση της παραπάνω θεώρησης, το θεώρημα διαίρεσης μίας διεργασίας Poisson μας διασφαλίζει ότι στην περίπτωση αυτή θα είναι $N_1(t) \sim \text{Poisson}(\lambda p) = \text{Poisson}(1)$ και $N_2(t) \sim \text{Poisson}(\lambda(1-p)) = \text{Poisson}(2)$ όπως προβλέπεται στην υπόθεση της άσκησης. Τώρα, είναι:

P({η δεύτερη άφιξη στο $N_1(t)$ να συμβεί πριν από την τρίτη άφιξη στο $N_2(t)$})

= P({ από τις πρώτες 4 αφίξεις στη N(t), οι 2 τουλάχιστον να διανεμηθούν στην $N_1(t)$})

= P({ από τις πρώτες 4 αφίξεις στη N(t), οι 2 ή οι 3 ή οι 4 να διανεμηθούν στην $N_1(t)$})

$= P_0$.

Αν, κάθε μία από τις πρώτες 4 αφίξεις θεωρηθεί ως δοκιμασία Bernoulli όπου “επιτυχία” είναι η διανομή της στη $N_1(t)$ με πιθανότητα $p = 1/3$, τότε αρκεί να βρούμε την πιθανότητα να έχουμε τουλάχιστον 2 “επιτυχίες” σε 4 ανεξάρτητες δοκιμές. Υπολογίζουμε:

$$P_0 = \Sigma_{\kappa = 2, 3, 4} (4 \text{ ανά } \kappa)\, p^{\kappa} (1 – p)^{n – \kappa}$$

$$= 6 (1/3)^2 (2/3)^2 + 4 (1/3)^3 (2/3) + (1/3)^4$$

$$= 33/81 = 0{,}407 = 40{,}7\%.$$

Εναλλακτική Λύση

Αν X = {ο χρόνος έως τη 2η άφιξη στο $N_1(t)$}, Y = {ο χρόνος έως τη 3η άφιξη στο $N_2(t)$}, τότε

X ~ Erlang(2, 1), Y ~ Erlang(3, 2) και $f_X(x) = x \cdot e^{-x}$, $f_Y(x) = 4 \cdot x^2 \cdot e^{-2x}$, x > 0.

Υπολογίζουμε $P(X < Y) = \int_0^\infty P(X < y) f_Y(y)\, dy = \int_0^\infty \left(\int_0^y f_X(x)\, dx \right) f_Y(y)\, dy$

Είναι $\int_0^y f_X(x)\, dx = \int_0^y x \cdot e^{-x}\, dx = (-y - 1)e^{-y} + 1$

Τελικά: $P(X < Y) = \int_0^\infty [(-y - 1)e^{-y} + 1]\, 4y^2 e^{-2y}\, dy = \frac{11}{27} = 0{,}407.$

Σημειώσεις

1. Για την πιθανότητα της μορφής P(X < Y) περισσότερες πληροφορίες μπορούν να βρεθούν εδώ: ***https://math.stackexchange.com/questions/261073/finding-probability-pxy***

2. Κωδικός Octave για την επαλήθευση των υπολογισμών:

```
syms x y; f = x*e^(-x); p = int (int(f, 0, y)*4*y^2*e^(-2*y), 0, inf)
```

Ασκήσεις στις διεργασίες Poisson

1. Γνωρίζοντας ότι $e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$ να αποδείξετε ότι $e^{-\lambda} = \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n$.

Υπόδειξη: Υπολογίστε το όριο του λογαρίθμου της παράστασης.

2. Οι ακόλουθες δύο πιθανότητες προκύπτουν από τη διωνυμική κατανομή και τη κατανομή Poisson, αντίστοιχα. Ποια από τις παρακάτω είναι την πιθανότητα, αν ρίξουμε αν ζάρι 6 φόρες, να έρθει ο αριθμός 3 δύο φορές;

(α) $\binom{6}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^4 = 0{,}2009$ (β) $e^{-3{,}5}\frac{3{,}5^2}{2!} = 0{,}1850$

3. Ο αριθμός των σωματιδίων που εκπέμπει μια ραδιενεργός πηγή ανά δευτερόλεπτο ακολουθεί την κατανομή Poisson με παράμετρο λ = 5. Να υπολογιστούν οι πιθανότητες: P(X = 0), P(X = 1), P(X = 2) και P(X ≥ 3).

Υπόδειξη: Αν X ~ Poisson(λ) τότε $P(X = \kappa) = e^{-\lambda} \cdot \lambda^{\kappa} / \kappa!$.

4. Έστω ότι η τ.μ. X ακολουθεί την κατανομή Poisson με παράμετρο λ.

Αν ισχύει P(X ≤ 1) = 4·P(X = 2), να προσδιοριστεί η παράμετρος λ και να υπολογιστούν οι πιθανότητες P(X < 3), P(X ≥ 1 | X < 3).

Υπόδειξη: P(A | B) = P(AB) / P(B).

5. Τα τελευταία 100 χρόνια, έχουν σημειωθεί 93 σεισμοί μεγέθους 6,0 και άνω της κλίμακας Ρίχτερ. Ποια είναι η πιθανότητα να έχουμε 3 σεισμούς την ίδια χρονιά που έχουν όλοι μέγεθος 6,0 ή περισσότερο;

6. Ένα τηλεφωνικό κέντρο δέχεται 2 κλήσεις κάθε 3 λεπτά. Να υπολογιστεί η πιθανότητα να φτάσουν 5 ή περισσότερες κλήσεις σε διάστημα 9 λεπτών.

7. Να αποδειχθεί ότι το άθροισμα δύο **ανεξάρτητων** τυχαίων μεταβλητών $X_1 \sim \text{Poisson}(\lambda_1)$ και $X_2 \sim \text{Poisson}(\lambda_2)$ είναι ομοίως κατανομή Poisson με παράμετρο $\lambda_1 + \lambda_2$, δηλαδή, αν $X \sim X_1 + X_2$, τότε $X \sim \text{Poisson}(\lambda_1 + \lambda_2)$.

Υπόδειξη: Αρκεί να δειχθεί ότι $P(X = z) = e^{-\lambda}\cdot\lambda^z/z!$, για κάθε z ϵ N και $\lambda = \lambda_1 + \lambda_2$.

8. Ο αριθμός των πελατών που φτάνουν σε ένα μάρκετ μπορεί να θεωρηθεί μια διαδικασία Poisson με ρυθμό λ = 10 πελάτες ανά ώρα.

(α) Να βρεθεί η πιθανότητα να υπάρχουν 2 πελάτες μεταξύ 10:00 και 10:20.

(β) Να βρεθεί η πιθανότητα να υπάρχουν 3 πελάτες μεταξύ 10:00 και 10:20 και 7 πελάτες μεταξύ 10:20 και 11.

9. Έστω $\{N(t), t \in [0,\infty)\}$ μια διεργασία Poisson με ρυθμό λ = 0,5 αφίξεις / μονάδα χρόνου.

(α) Να βρεθεί η πιθανότητα να μην συμβεί καμία άφιξη στο διάστημα (3, 5].

(β) Να βρεθεί η πιθανότητα να υπάρχει ακριβώς μία άφιξη σε καθένα από τα ακόλουθα διαστήματα: (0, 1], (1, 2], (2, 3] και (3, 4].

10. Έστω $\{N(t), t \in [0,\infty)\}$ μια διεργασία Poisson με ρυθμό λ = 3 αφίξεις / μονάδα χρόνου. Να βρεθεί η πιθανότητα να υπάρχουν δύο αφίξεις στο (0, 2] και τρεις αφίξεις στο (1, 4].

11. Έστω $\{N(t), t \in [0,\infty)\}$ μια διεργασία Poisson με ρυθμό λ. Να βρεθεί η συνάρτηση συνδιακύμανσης $C_N(t_1, t_2) = Cov(N(t_1), N(t_2)) = E(N(t_1) – \lambda t_1, N(t_2) – \lambda t_2)$, συναρτήσει των t_1, t_2 για κάθε $t_1, t_2 \geq 0$.

12. Έστω $\{N(t), t \in [0,\infty)\}$ μια διεργασία Poisson με ρυθμό λ = 5 αφίξεις / min.

(α) Να βρεθεί το χρονικό διάστημα που αναμένεται να περάσει έως την 1η άφιξη.

(β) Να βρεθεί η πιθανότητα η 1η άφιξη να καθυστερήσει περισσότερο από μισό λεπτό.

13. Έστω $\{N(t), t \in [0,\infty)\}$ μια διεργασία Poisson με ρυθμό λ. Να δειχθεί ότι, αν γνωρίζουμε πως έχει συμβεί μία άφιξη μέχρι τη χρονική στιγμή t (N(t) = 1), τότε ο χρόνος x που αυτή έχει συμβεί, $0 \leq x \leq t$, έχει ομοιόμορφη κατανομή στο (0, t].

Υπόδειξη
Αν X_1 η χρονική στιγμή που έγινε η $1^η$ άφιξη, τότε, αρκεί να δειχθεί ότι $P(X_1 \leq x \mid N(t) = 1) = x/t$, $0 \leq x \leq t$.

14. Αν γνωρίζουμε ότι την πρώτη ώρα λειτουργίας ενός εστιατορίου έχει έρθει ένας πελάτης, να βρεθεί η πιθανότητα αυτός

(α) να έχει έρθει στα πρώτα 10 λεπτά.

(β) να έχει έρθει μεταξύ του 25 και του 40 λεπτού της ώρας.

15. Να βρεθεί η πιθανότητα, δύο πελάτες που έχουν αφιχθεί σε ένα κατάστημα τα πρώτα δύο λεπτά από τη στιγμή που άνοιξε, να έχουν αφιχθεί στο πρώτο από τα δύο λεπτά.

16. Έστω X ο χρόνος που μεσολαβεί μεταξύ δύο αφίξεων σε μία διεργασία Poisson με παράμετρο λ. Να αποδείξετε ότι τότε για κάθε t_0, $x \geq 0$, $P(X > x + t_0 \mid X > t_0) = P(X > x)$

Υπόδειξη: Είναι $X \sim Exp(\lambda)$

17. Έστω $T_1, T_2, \ldots T_n, \ldots$, ο χρόνος που απαιτείται μέχρι την υλοποίηση του 1ου, 2ου, …, n – οστού, … γεγονότος. Να δείξετε ότι (α) $E(T_n) = n/\lambda$. (β) $Var(T_n) = n/\lambda^2$.

18. Ερωτήσεις καταφθάνουν σε ένα μηχανισμό καταγραφής μηνυμάτων σύμφωνα με μία Poisson διεργασία με ένταση 15 ερωτήσεων το λεπτό.

(α) Να βρεθεί η πιθανότητα ότι σε μία περίοδο 1 λεπτού, 3 ερωτήσεις να φθάσουν κατά τα πρώτα 10 δευτερόλεπτα και 2 κατά τα τελευταία 15 δευτερόλεπτα.

(β) Να υπολογισθεί η μέση τιμή και η διασπορά του χρόνου μέχρι την άφιξη της $10^{ης}$ ερώτησης.

19. Έστω $N_1(t)$ και $N_2(t)$ δύο ανεξάρτητες διεργασίες Poisson με ρυθμούς $\lambda_1 = 1$ και $\lambda_2 = 2$ συμβάντα / χρονική μονάδα και $N(t) = N_1(t) + N_2(t)$ η συγχωνευμένη διεργασία.

(α) Να βρείτε την πιθανότητα να συμβούν στη συγχωνευμένη διεργασία 2 γεγονότα στο χρονικό διάστημα (0, 1] και συνολικά 5 γεγονότα στο χρονικό διάστημα (0, 2].

(β) Γνωρίζοντας ότι έχουν συμβεί 2 γεγονότα στη συγχωνευμένη διεργασία στο χρονικό διάστημα (0, 1] να βρεθεί η πιθανότητα 1 από αυτά να έχει έρθει από την $N_1(t)$.

Παράρτημα 1: Γέννεση της κατανομής Poisson ως όριο δοκιμασιών Bernoulli

Διαχωρίζουμε τη μονάδα του χρόνου σε n μέρη. Σε κάθε ένα από αυτά είτε συμβαίνει είτε δεν συμβαίνει μία άφιξη και έστω p η πιθανότητα άφιξης σε κάθε ένα από αυτά τα μέρη. Αν X_n η τ.μ. που μετράει το πλήθος αφίξεων στη μονάδα του χρόνου, τότε $X_n \sim B(n, p)$ και $E(X_n) = np$.

Αν λ το (γνωστό) πλήθος αφίξεων στη μονάδα του χρόνου, τότε ορίζουμε $p = \lambda / n$ και θεωρούμε την τ.μ. $X_n \sim B(n, \lambda/n)$ για $n \geq \lambda$. Η τ.μ. X_n, εξακολουθεί να έχει $E(X_n) = \lambda$ και τώρα μετράει τις αφίξεις στην μονάδα του χρόνου, όταν αυτή η χρονική μονάδα διαχωρίζεται σε n μέρη και σε κάθε ένα από αυτά υπάρχει πιθανότητα άφιξης λ / n.

Για $p = \lambda / n$ και για $k \leq n$, γράφουμε:

$$P(X_n = k) = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k} = \frac{\lambda^k}{k!} \frac{n!}{(n-k)!n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}.$$

Όμως

- $\lim\limits_{n \to \infty} \frac{n!}{(n-k)!n^k} = \lim\limits_{n \to \infty} \frac{n(n-1)(n-2) \cdot \ldots \cdot (n-k+1)}{n^k} = 1.$
- $\lim\limits_{n \to \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}.$ (*)
- $\lim\limits_{n \to \infty} \left(1 - \frac{\lambda}{n}\right)^{-k} = 1.$

Από τα παραπάνω συμπεραίνουμε ότι

$$\lim_{n \to \infty} P(X_n = k) = \lim_{n \to \infty} \binom{n}{k} p^k (1-p)^{n-k} = e^{-\lambda} \frac{\lambda^k}{k!} = P(Y = k), \quad Y \sim \text{Poisson}(\lambda).$$

(*) Απόδειξη του $\lim\limits_{n \to \infty} \left(1 + \frac{\lambda}{n}\right)^n = e^{\lambda}, \lambda \in \mathbb{R}$.

Θέτουμε $y_n = \left(1 + \frac{\lambda}{n}\right)^n$, άρα $\ln y_n = n \ln\left(1 + \frac{\lambda}{n}\right) = \dfrac{\ln\left(1 + \frac{\lambda}{n}\right)}{\frac{1}{n}}$

Είναι $\lim\limits_{n \to \infty} \ln y_n = \lim\limits_{n \to \infty} \dfrac{\ln\left(1 + \frac{\lambda}{n}\right)}{\frac{1}{n}} = \lim\limits_{n \to \infty} \dfrac{\frac{1}{1 + \frac{\lambda}{n}}\left(-\frac{\lambda}{n^2}\right)}{-\frac{1}{n^2}} = \lambda.$ (κανόνας De L'Hospital)

Άρα $\lim\limits_{n \to \infty} \left(1 + \frac{\lambda}{n}\right)^n = \lim\limits_{n \to \infty} y_n = \lim\limits_{n \to \infty} e^{\ln y_n} = e^{\lambda}.$

Παράρτημα 2: Ανεξάρτητες και Στάσιμες Προσαυξήσεις

Ορισμός (Ανεξάρτητες Προσαυξήσεις)

Έστω $\{X(t), t \geq 0\}$ μια στοχαστική διεργασία συνεχούς χρόνου. Λέμε ότι η $X(t)$ έχει **ανεξάρτητες προσαυξήσεις** εάν, για κάθε $0 \leq t_1 < t_2 < t_3 <\ldots < t_n$, οι τυχαίες μεταβλητές

$$X(t_2) - X(t_1), X(t_3) - X(t_2),\ldots ,X(t_n) - X(t_{n-1})$$

είναι ανεξάρτητες. Έστω τώρα $\{N(t), t \geq 0\}$ μια διεργασία καταμέτρησης και

$$N(t) = \{\text{πλήθος αφίξεων που συμβαίνει στο χρονικό διάστημα } (0, t]\}$$

Η διαφορά $N(t_i) - N(t_{i-1})$ εκφράζει τον αριθμό των αφίξεων στο διάστημα $(t_{i-1}, t_i]$. Έτσι, μια διαδικασία καταμέτρησης **έχει ανεξάρτητες προσαυξήσεις εάν το πλήθος αφίξεων σε ξένα μεταξύ τους χρονικά διαστήματα $(t_1, t_2], (t_2, t_3],\ldots,,(t_{n-1}, t_n]$ είναι ανεξάρτητες.**

Ορισμός (Στάσιμες Προσαυξήσεις)

Έστω $\{X(t), t \geq 0\}$ μια στοχαστική διεργασία συνεχούς χρόνου. Λέμε ότι η $X(t)$ έχει **στάσιμες προσαυξήσεις** εάν, για κάθε $0 \leq t_1 < t_2$, και όλα τα $r > 0$, οι τυχαίες μεταβλητές

$$X(t_2 + r) - X(t_1 + r), X(t_2) - X(t_1)$$

είναι ισόνομες τυχαίες μεταβλητές (δηλ. έχουν τις ίδιες κατανομές πιθανότητας).

Έστω τώρα $\{N(t), t \geq 0\}$ μια διεργασία καταμέτρησης. Η ποσότητα $N(t_2) - N(t_1)$ εκφράζει τον αριθμό των αφίξεων στο διάστημα $(t_1, t_2]$.

Καθώς $N(0) = 0$, **μια διαδικασία καταμέτρησης έχει στάσιμες προσαυξήσεις εάν** για όλα τα $0 \leq t_1 < t_2$, το $N(t_2) - N(t_1)$ έχει την ίδια κατανομή με το $N(t_2 - t_1) - N(0)$ δηλαδή με το $N(t_2 - t_1)$:

$$\mathbf{N(t_2) - N(t_1) \sim N(t_2 - t_1)}$$

Αυτό σημαίνει ότι **η κατανομή του πλήθους των αφίξεων** σε οποιοδήποτε διάστημα **εξαρτάται μόνο από το μήκος του διαστήματος**, και όχι από την ακριβή θέση του διαστήματος στην πραγματική γραμμή.

Παρατήρηση: Μία διεργασία καταμέτρησης $\{N(t), t \geq 0\}$ μπορεί να έχει στάσιμες αλλά όχι ανεξάρτητες προσαυξήσεις.

Παράδειγμα

Ας υποθέσουμε ότι τα γεγονότα συμβαίνουν στις χρονικές στιγμές X + n για n ≥ 0 και έστω t ≥ 0. Αν X ~ U(0, 1) και θεωρήσουμε αφίξεις που συμβαίνουν στις χρονικές στιγμές X + i, i = 0, 1, 2, …, τότε η διεργασία καταμέτρησης {N(t), t ≥ 0} που μετρά το πλήθος των αφίξεων στο διάστημα (0, t]:

- Δεν έχει ανεξάρτητες προσαυξήσεις, καθώς για κάθε s ϵ (0, 1) οι τυχαίες μεταβλητές U = N(t+s) − N(t) και V = N(t+1) − N(t+s) δεν είναι στοχαστικά ανεξάρτητες από τη στιγμή που είναι Bernoulli τυχαίες μεταβλητές με U + V = 1.

- Έχει στάσιμες προσαυξήσεις καθώς για κάθε t ≥ 0 και s ϵ (0,1), η N(t+s) − N(t) είναι Bernoulli(s) (άρα για κάθε επιλογή t_1, t_2 είναι ανεξάρτητες μεταξύ τους), ενώ για κάθε s ≥ 0, N(t+s) − N(t) είναι της μορφής [s] + Y, όπου Y ~ Bernoulli(s − [s]) (όπου [s] το ακέραιο μέρος του t). Άρα, οι προσαυξήσεις είναι ισόνομες, δηλαδή η διεργασία έχει στάσιμες προσαυξήσεις.

Παρατήρηση: Μία διεργασία καταμέτρησης {N(t), t ≥ 0} μπορεί να έχει ανεξάρτητες αλλά όχι στάσιμες προσαυξήσεις.

Παράδειγμα

Έστω {N(t), t ≥ 0} μία διεργασία που *έχει* στάσιμες προσαυξήσεις. Ορίζουμε τη διεργασία {f(t), t ≥ 0} ως

$$f(t) = [t]\cdot N(1) + N(t),$$

όπου [t] το ακέραιο μέρος του t. Η {f(t), t ≥ 0} δεν έχει ανεξάρτητες προσαυξήσεις καθώς

$$f(n+1) - f(n) = f(n) - f(n-1)\ (= 2N(1)),\ n \in N,\ n \geq 1,$$

αλλά έχει στάσιμες προσαυξήσεις καθώς

$f(t_2 + r) – f(t_1 + r) = [t_2 + r]\cdot N(1) + N(t_2 + r) – [t_1 + r]\cdot N(1) – N(t_1 + r) = [t_2 – t_1]\cdot N(1) + N(t_2 – t_1)$: εξαρτάται μόνο από τη διαφορά $t_2 – t_1$ και όχι από το r > 0.

ΣημείωσηΑν $x = n + c_0$, $y = m + c_1$, τότε [x – y] = [x] – [y], αν $0 \leq c_0 – c_1 < 1$ και [x – y] = [x] – [y] + 1, αν $-1 < c_0 – c_1 < 0$.Στην περίπτωση του παραδείγματος $c_0 = c_1 = c$.

Παράρτημα 3: (Ισχυρά) Στάσιμες στοχαστικές διεργασίες

Λέμε ότι μία στοχαστική ανέλιξη είναι ισχυρά στάσιμη όταν για κάθε n και για κάθε t_1, t_2, …, t_n ϵ T, και για κάθε τ > 0, τ.ω. $t_1 + \tau$, $t_2 + \tau$, …,. $t_n + \tau$ ϵ T, τα διανύσματα

$$[X(t_1), X(t_2), \ldots, X(t_n)] \text{ και } [X(t_1 + \tau), X(t_2 + \tau), \ldots, X(t_n + \tau)]$$

έχουν την **ίδια κοινή κατανομή πιθανότητας**, δηλαδή

$$\mathbf{F_{t1, t2, \ldots, tn} = F_{t1 + \tau, t2 + \tau, \ldots, tn + \tau}}$$

Άμεση συνέπεια: για κάθε t, τ, τα X(t) και X(t + τ) είναι ισόνομα (όχι απαραίτητα ανεξάρτητα). Από την ισότητα $F_{t1, t2, \ldots, tn} = F_{t1 + \tau, t2 + \tau, \ldots, tn + \tau}$, θέτοντας n = 1, έχουμε $F_{t1} = F_{t1 + \tau}$, τ > 0, δηλαδή η κατανομή πιθανότητας των $\{X_t\}$, t Є T, δεν εξαρτάται από το δείκτη t, ή

$$F_t(x) = F(x), \ t \in T.$$

Από την παραπάνω ισότητα προκύπτει ότι

$$m(t) = E(X(t)) = \int_{-\infty}^{+\infty} t dF(t) = m = \text{σταθερά}$$

και

$$Var(X(t)) = E(X(t)^2) - E(X(t))^2 = \int_{-\infty}^{+\infty} t^2 dF(t) - m^2 = \text{σταθερά}$$

δηλαδή τόσο **η τάση (trend)** όσο και **η διακύμανση (variance)** είναι σταθερές για κάθε $\{X_t\}$, t Є T.

Από την ισότητα $F_{t1, t2, \ldots, tn} = F_{t1 + \tau, t2 + \tau, \ldots, tn + \tau}$, θέτοντας n = 2, $t_1 = 0$, $t_2 = t - s$, και τ = s, όπου s < t, παίρνουμε

$$F_{0, t - s} = F_{s, t},$$

δηλαδή η κοινή κατανομή $F_{s, t}$ των X_t, X_s, εξαρτάται αποκλειστικά από τη διαφορά δ = t – s, και όχι από τις συγκεκριμένες τιμές των t, s. Καθώς η E[X(t)X(s)] προσδιορίζεται πλήρως από την $F_{s, t}$, συνάγουμε ότι και **η συνάρτηση της συνδιακύμανσης**

$$C(s, t) = Cov(X(t), X(s)) = E[X(t)X(s)] - m(t)m(s) = E[X(t)X(s)] - m^2,$$

εξαρτάται αποκλειστικά από τη διαφορά δ = t – s. Ιδιαίτερα, βρίσκουμε ότι

$$C(s, t) = C(s, s + \delta) = C(0, \delta)$$

Καθώς

$$\rho(t,s) = \frac{Cov(X(t),X(s))}{\sqrt{Var(X(t))}\cdot\sqrt{Var(X(t))}} = \frac{C(t,s)}{\sqrt{Var(X(0))}\cdot\sqrt{Var(X(0))}} = \frac{C(0,\delta)}{VarX(0)},$$

συνάγεται ότι και **ο συντελεστής συσχέτισης** των X_t, X_s, εξαρτάται αποκλειστικά από τη διαφορά δ = t – s, και όχι από τις συγκεκριμένες τιμές των t, s.

Συνοψίζοντας

Σε κάθε ισχυρά στάσιμη στοχαστική διεργασία

(α) Η αναμενόμενη τιμή E(X(t)) είναι σταθερή στο χρόνο.

(β) Η διακύμανση Var(X(t)) είναι σταθερή στο χρόνο.

(γ) Η συνδιακύμανση Cov(X(t), X(s)), εξαρτάται αποκλειστικά από τη διαφορά |s – t|.

(δ) Ο συντελεστής συσχέτισης ρ(X(t), X(s)), εξαρτάται αποκλειστικά από τη διαφορά |s – t|.

Παράρτημα 4: (Ασθενώς) Στάσιμες στοχαστικές διεργασίες

Μία στοχαστική διεργασία λέμε ότι είναι **διεργασία δεύτερης τάξης** αν για κάθε t ≥ 0, ισχύει

$$E(X^2(t)) < +\infty.$$

Για τις διεργασίες δεύτερης τάξης, είναι βέβαιο πως θα υπάρχουν και θα ορίζονται για κάθε t ≥ 0, ως πραγματικοί αριθμοί οι

$E(X(t)) \le [E(X^2(t))]^{0,5} < +\infty$, και $(var(X(t)) = E(X^2(t)) - E(X(t))^2 \ge 0 \leftrightarrow E(X(t)) \le [E(X^2(t))]^{0,5})$

$Var(X(t)) = E(X^2(t)) - E(X(t))^2 < +\infty$.

Ορισμός

Μία στοχαστική διεργασία λέμε ότι είναι **ασθενώς στάσιμη**, όταν

- $E(X^2(t)) < +\infty$.
- $m(t) = E(X(t)) = m \in R$
- Η συνδιακύμανση των $X(t_1)$ και $X(t_2)$ εξαρτάται μόνο από την απόσταση των t_1, t_2 και όχι από την θέση τους στον πραγματικό άξονα, δηλαδή

$$Cov(X(t_1), X(t_2)) = Cov(X(t_1+ h), X(t_2+ h)),\ t_1, t_2, h \ge 0.$$

Σημείωση: Παρατηρούμε ότι $\mathrm{Var}X(t) = \mathrm{Cov}(X(t), X(t))$ = σταθερό για κάθε $t \geq 0$.

Δηλαδή, μια στοχαστική διαδικασία είναι ασθενώς στάσιμη αν

(α) έχει την ίδια μέση τιμή, μ, σε όλα τα χρονικά σημεία.

(β) έχει την ίδια διακύμανση σε όλα τα χρονικά σημεία.

(γ) η συνδιακύμανση μεταξύ των τιμών σε οποιαδήποτε δύο χρονικά σημεία, $t_1 + h$, $t_2 + h$, εξαρτώνται μόνο από το $t_2 - t_1$, τη διαφορά μεταξύ των δύο στιγμών, και όχι από τη θέση των σημείων στον άξονα του χρόνου.

Παρατήρηση: Ασθενώς στάσιμη ≠ Ισχυρά στάσιμη

Μία ασθενώς στάσιμη διεργασία δεν είναι απαραίτητα ισχυρά στάσιμη.

Παράδειγμα

Έστω ότι $X \sim N(0, 1)$ κατανομή. Θεωρούμε την τυχαία μεταβλητή $Z = X^2$.

Τότε $Z \sim \chi^2_1$ (χι τετράγωνο κατανομή με 1 βαθμό ελευθερίας), συνεπώς για τις ροπές της θα ισχύει:

$$E(Z^n) = 1 \cdot (1 + 2) \cdot \ldots \cdot (1 + 2n - 2) = 2^n\, \Gamma(n + ½) / \Gamma(½)$$

Συμπεραίνουμε ότι $\mathbf{E(Z) = 1}$ και $\mathbf{E(Z^2)} = 1 \cdot (1 + 2) = 3$.

Έστω τώρα ότι $\{U(t), t \in N\}$ μία οικογένεια *ανεξάρτητων* τυχαίων μεταβλητών που στο σύνολό τους ακολουθούν την $N(0, 1)$ κατανομή. Θεωρούμε τη διεργασία $\{X(t), t \in N\}$, με $X(t) = U(t)$, αν t άρτιος

και $X(t) = \frac{1}{\sqrt{2}}\left(U^2(t) - 1\right)$, αν t περιττός.

Παρατηρούμε ότι

- $E(X^2(t)) < +\infty$,
- $E(X(t)) = 0$,
- $\mathrm{Var}(X(t)) = 1$ και
- $\mathrm{Cov}(X(t_1), X(t_2)) = 0$,

δηλαδή η $\{X(t), t \in N\}$ είναι ασθενώς στάσιμη. Ωστόσο δεν είναι ισχυρά στάσιμη, γιατί δύο διαδοχικές μεταβλητές της δεν έχουν την ίδια κατανομή.

Πιο συγκεκριμένα, παρατηρούμε ότι

- Για άρτιο t: $P(X(t) \le 0) = P(U(t) \le 0) = 0{,}5$ ενώ
- Για περιττό t: $P(X(t) \le 0) = P(U^2(t) - 1 \le 0) = P(|U(t)| \le 1) = P(-1 \le U(t) \le 1) = 0{,}683 \ne 0{,}5$

Παρατήρηση: Ισχυρά στάσιμη ≠ Ασθενώς στάσιμη

Μία ισχυρά στάσιμη διεργασία, δεν είναι πάντα ασθενώς στάσιμη καθώς είναι πιθανό να μην ικανοποιεί την απαραίτητη συνθήκη δεύτερης τάξης $E(X^2(t)) < +\infty$.

Παράδειγμα

Θεωρούμε τη διεργασία {X(t), t 0ε0 N0}, όπου X(t) ~ Cauchy(0, 1). Η {X(t), t 0ε0 N0} είναι ισχυρά στάσιμη καθώς όλες οι μεταβλητές που την αποτελούν είναι ισόνομες και ανεξάρτητες μεταξύ τους, αλλά δεν είναι ασθενώς στάσιμη καθώς δεν είναι διεργασία δεύτερης τάξης. Πράγματι, η συνάρτηση πυκνότητας πιθανότητας μίας Cauchy(0, 1) είναι η

$$f(x) = 1/[\pi(1 + x^2)]$$

και είναι εύκολο να δειχθεί ότι το ολοκλήρωμα στο $(-\infty, +\infty)$ της $x^2 f(x)$ δεν είναι πεπερασμένο, δηλαδή $E(X^2(t)) = +\infty$.

Σημείωση

Η έννοια της “κεντρικής τάσης” είναι γενικότερη της έννοιας της αναμενόμενης 0τιμής μίας κατανομής. Μία διεργασία {X(t), t 0ε N0}, που είναι ισχυρά στάσιμη αλλά 0έχει $E(X^2(t)) = +\infty$ εξακολουθεί να έχει σταθερή “κεντρική τάση” αν και αυτή δεν μπορεί να προσδιοριστεί από τα συνηθισμένα μέτρα κεντρικής τάσης όπως η αναμενόμενη τιμή.

Στην περίπτωση αυτή η κεντρική τάση *μπορεί να εκτιμηθεί από άλλα μέτρα* που στοχεύουν σε αυτήν όπως **η διάμεσος της κατανομής** ή **η επικρατούσα τιμή** της.

Παράδειγμα

Αν X(t) ~ Cauchy(0, 1) με $f(x) = 1/[\pi(1 + x^2)]$ τότε η αναμενόμενη τιμή δεν ορίζεται καλώς, καθώς

$$E(X(t)) = \int_{-\infty}^{+\infty} x f(x)\, dx = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{x}{1+x^2}\, dx = \frac{1}{\pi}\int_{-\infty}^{\alpha} \frac{x}{1+x^2}\, dx + \frac{1}{\pi}\int_{\alpha}^{+\infty} \frac{x}{1+x^2}\, dx = -\infty + \infty.$$

(Υπενθυμίζεται πως για να ορίζεται καλώς το γενικευμένο ολοκλήρωμα $\int_{-\infty}^{+\infty} \frac{x}{1+x^2}\, dx$ πρέπει και αρκεί τουλάχιστον ένα από τα δύο ολοκληρώματα $\int_{-\infty}^{\alpha} \frac{x}{1+x^2}\, dx$ και $\int_{\alpha}^{+\infty} \frac{x}{1+x^2}\, dx$ να υπάρχουν ως

πεπερασμένοι πραγματικοί αριθμοί. Ιδιαίτερα, το γενικευμένο ολοκλήρωμα δεν ορίζεται ως

$$\int_{-\infty}^{+\infty} \frac{x}{1+x^2}\,dx = \lim_{\alpha \to +\infty} \int_{-\alpha}^{\alpha} \frac{x}{1+x^2}\,dx \quad)$$

Ωστόσο

- Διάμεσος δ = 0 και
- Επικρατούσα τιμή M = 0.

Σημείωση:

Αν X είναι συνεχής τ.μ. τότε η διάμεσος της X είναι ο αριθμός δ για τον οποίο $F_X(\delta) = 0{,}5$, ενώ η επικρατούσα τιμή M_X, είναι ο αριθμός στον οποίο η συνάρτηση πυκνότητας πιθανότητας f_X αποκτά τη μέγιστη τιμή της.

Παρατήρηση: Ισχυρά στάσιμη και $E(X^2(t)) < +\infty$ → Ασθενώς στάσιμη

Μία ισχυρά στάσιμη διεργασία για την οποία ισχύει $E(X^2(t)) < +\infty$, θα είναι και ασθενώς στάσιμη.

Απόδειξη

Πράγματι, αν η διεργασία $\{X(t), t \in N\}$, είναι ισχυρά στάσιμη τότε οι τυχαίες μεταβλητές

$$X(0), X(1), X(2), \ldots$$

Θα έχουν ίδια κατανομή, άρα $E(X(t)) = m$ = σταθερό και $E(X^2(t)) = w$ = σταθερό.

Επιπλέον, κάθε ζεύγος τυχαίων μεταβλητών $(X(t_1), X(t_2))$ και $(X(t_1+ h), X(t_2+ h))$ θα έχει ομοίως ίδια κοινή κατανομή. Καθώς, $E(X^2(t_1)), E(X^2(t_2)) < +\infty$ θα είναι

$$Cov(X(t_1), X(t_2)) = Cov(X(t_1+ h), X(t_2+ h)) < +\infty,$$

καθώς $Cov(X(t_1), X(t_2)) = E(X^2(t_1)) - [E(X(t_1))]^2$.