ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ
ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

1η Ομάδα Ασκήσεων

Συστήματα Αναμονής (Queuing Systems)

ΔΗΜΗΤΡΙΟΣ ΓΕΩΡΓΟΥΣΗΣ

03119005

georg.dimi00@gmail.com

Κατανομή Poisson

Μερικά στοιχεία για τη κατανομή Poisson που θα μας φανούν χρήσιμα:

Έστω η τυχαία μεταβλητή X που ακολουθεί το κατανομή Poisson με παράμετρο λ ($X \sim Poisson(\lambda)$), τότε η συνάρτηση μάζας πιθανότητας είναι:

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

Και ισχύει λ = E(X) = Var(X), όπου E(X) και Var(X) η μέση τιμή και η διακύμανση της μεταβλητής X αντίστοιχα.

Επίσης, θα μας φανεί χρήσιμη η ιδιότητα:

Αν X1, X2, X3, …, Xn ανεξάρτητες κατανομές Poisson με παραμέτρους λ1, λ2, λ3, …, λn τότε το αποτέλεσμα είναι κατανομή Poisson με λ = λ1 + λ2 + λ3 + … + λn. Η ιδιότητα αυτή είναι συμπέρασμα του γνωστού αποτελέσματος ότι αν κάνουμε συνέλιξη των συναρτήσεων μάζας πιθανότητας δύο κατανομών Poisson με παραμέτρους λ1, λ2 παίρνουμε κατανομή Poisson με λ = λ1 + λ2.

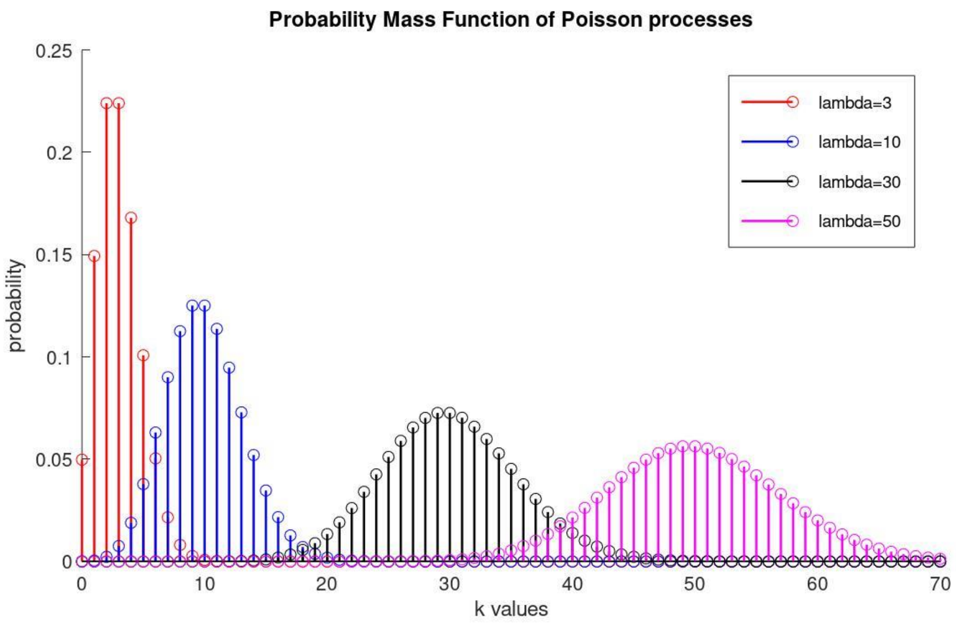
Τέλος, η συνάρτηση μάζας πιθανότητας της διωνυμικής κατανομής με παραμέτρους n, p (Binomial(n, p)) είναι:

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, \ldots, n$$

Τρέχουμε τον έτοιμο κώδικα.

A)

```
clc;
clear all;
close all;

% TASK: In a common diagram, design the Probability Mass Function of
Poisson processes
% with lambda parameters 3, 10, 30, 50. In the horizontal axes,
choose k parameters
% between 0 and 70.

pkg load statistics

k = 0:1:70;
lambda = [3, 10, 30, 50];

for i = 1 : columns(lambda)
  poisson(i, :) = poisspdf(k, lambda(i));
endfor

colors = "rbkmg";
%Added green as 5th color so that we can print all 5 distributions in
the last question
figure(1);
hold on;
for i = 1 : columns(lambda)
  stem(k, poisson(i, :), colors(i), "linewidth", 1.2);
endfor
hold off;

title("Probability Mass Function of Poisson processes");
xlabel("k values");
ylabel("probability");
legend("lambda=3", "lambda=10", "lambda=30", "lambda=50");
```



Επειδή το λ είναι η μέση τιμή βλέπουμε ότι κάθε κατανομή είναι «συγκεντρωμένη» γύρω από το αντίστοιχο λ στον οριζόντιο άξονα. Επίσης, το λ είναι η διακύμανση και βλέπουμε ότι, καθώς το λ αυξάνει, η αντίστοιχη κατανομή είναι πιο απλωμένη γύρω από τη μέση της τιμή.

Β) Για να υπολογίσουμε τη μέση τιμή και τη διακύμανση χρησιμοποιούμε τους τύπους:

$$E(X) = \sum_{k\geq 0} kP_X(k)\,, Var(X) = \sum_{k\geq 0} k^2 P_X(k) - E^2(X)$$

```
index = find(lambda == 30);
chosen = poisson(index, :);
mean_value = 0;
for i=0:(columns(poisson(index, :)) - 1)
  mean_value = mean_value + i .* poisson(index,i+1);
endfor

display("mean value of Poisson with lambda 30 is");
display(mean_value);

second_moment = 0;
for i = 0 : (columns(poisson(index, :)) - 1)
  second_moment = second_moment + i .* i .* poisson(index, i + 1);
endfor

variance = second_moment - mean_value .^ 2;
display("Variance of Poisson with lambda 30 is");
display(variance);
```

```
mean value of Poisson with lambda 30 is
mean_value = 30.000
Variance of Poisson with lambda 30 is
variance = 30.000
```

Παίρνουμε αυτή την έξοδο.

Γ) Έστω X, Y δύο τυχαίες μεταβλητές και Z = X + Y μία άλλη τυχαία μεταβλητή.

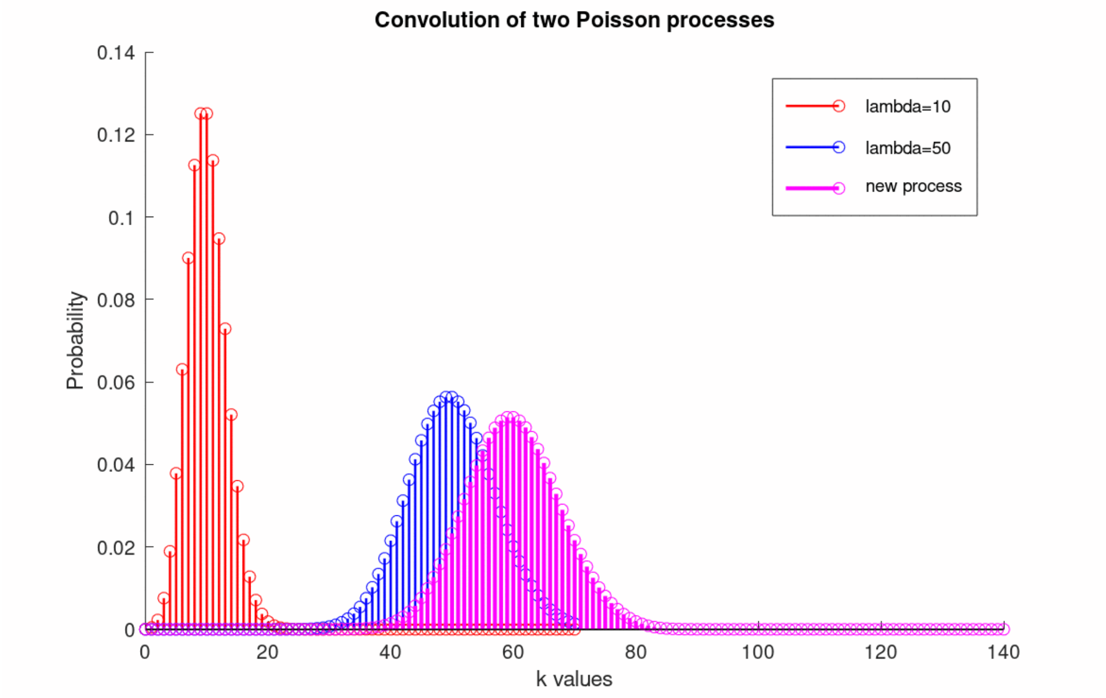
Έχω $P_Z(k) = P(X+Y=k) = \sum_{i=0}^{k} P(X=i, Y=k-i)$

Αν οι X, Y είναι ανεξάρτητες συνεχίζουμε ως εξής:

$$P_Z(k) = \sum_{i=0}^{k} P(X=i)P(Y=k-i) = \sum_{i=0}^{k} P_X(i)P_Y(k-i)$$

Το οποίο είναι η συνέλιξη των κατανομών των X, Y.

```
first = find(lambda == 10);
second = find(lambda == 50);
poisson_first = poisson(first, :);
poisson_second = poisson(second, :);

composed = conv(poisson_first, poisson_second);
new_k = 0 : 1 : (2 * 70);

figure(2);
hold on;
stem(k, poisson_first(:), colors(1), "linewidth", 1.2);
stem(k, poisson_second(:), colors(2), "linewidth", 1.2);
stem(new_k, composed, "mo", "linewidth", 2);
hold off;
title("Convolution of two Poisson processes");
xlabel("k values");
ylabel("Probability");
legend("lambda=10", "lambda=50", "new process");
```

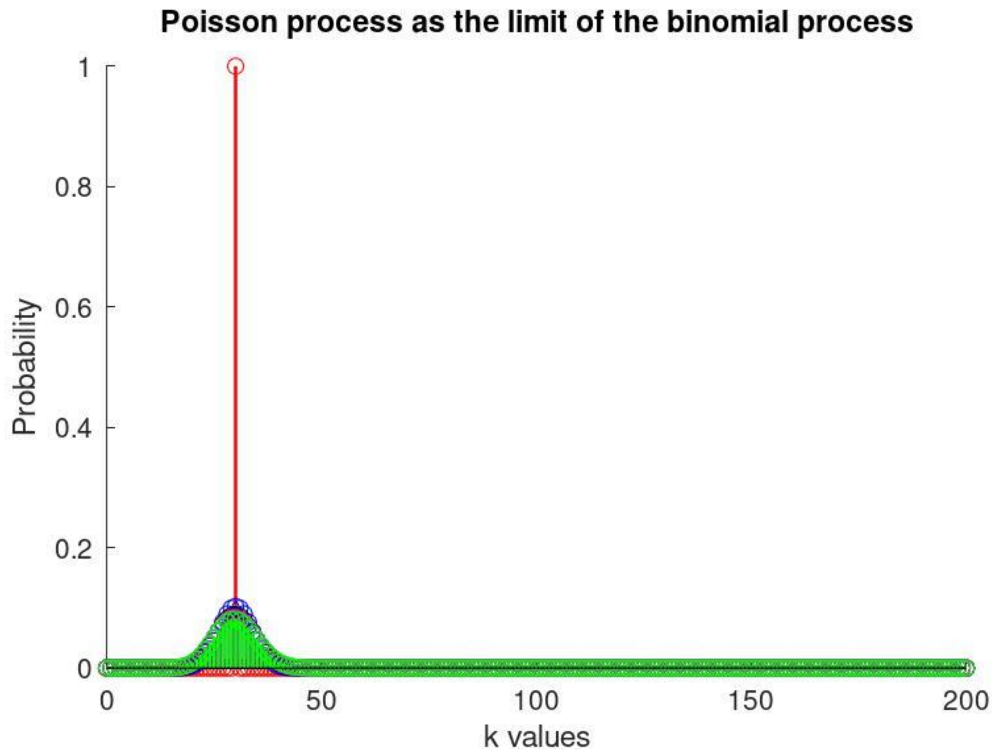


Κάναμε συνέλιξη Poisson(10), Poisson(50) άρα με βάση αυτά που είπαμε στην αρχή πήραμε Poisson(10+50), δηλαδή, Poisson(60). Επίσης, οι τιμές στον οριζόντιο άξονα φτάνουν μέχρι 140 γιατί οι 2 αρχικές έφταναν μέχρι 70. Για να μας δώσει η υπέρθεση 2 κατανομών τη συνέλιξη των κατανομών αυτών και, κατ' επέκταση το παραπάνω αποτέλεσμα, είδαμε στην αρχή του (Γ) ότι πρέπει οι 2 αυτές κατανομές να είναι ανεξάρτητες, αλλά και σε συνδυασμό με το σχόλιο στην αρχή της άσκησης.

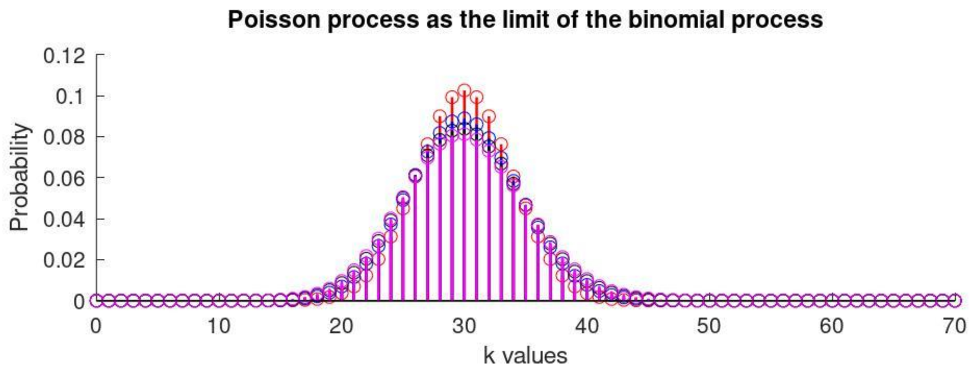
Δ)

Η Poisson(λ) είναι το όριο της Binomial(n, λ/n) καθώς το n $\rightarrow\infty$ και έχουμε περιγράψει στην αρχή την κατανομή Binomial(n, p). Θέτουμε, δηλαδή, p = 1/n.

Έγινε μια μικρή τροποποίηση στον δοσμένο κώδικα, γιατί δεν τύπωνε την Binomial(150, 1/5). Δείχνουμε το αποτέλεσμα παρακάτω:



Θα δείξουμε και τις κατανομές χωρίς αυτήν με n = 30 (p = 1) και για k μέχρι 70 (όπως στο πρώτο ερώτημα) για να είναι πιο εμφανές αυτό που θέλουμε να δείξουμε:



Είναι εμφανές ότι καθώς το n μεγαλώνει το σχήμα πάει να μοιάσει στο μαύρο διάγραμμα της 1ης εικόνας (Poisson(30)).

Ο κώδικας είναι:

```
% TASK: show that Poisson process is the limit of the binomial
distribution.
k = 0 : 1 : 200;
% Define the desired Poisson Process
lambda = 30;
i = 1 : 1 : 5;
n = lambda .* i;
p = lambda ./ n;

figure(3);
title("Poisson process as the limit of the binomial process");
xlabel("k values");
ylabel("Probability");
hold on;
for i = 1 : 5
%changed for range to include 5 so that we print the binomial(150,
1/5) too
  binomial = binopdf(k, n(i), p(i));
  stem(k, binomial, colors(i), 'linewidth', 1.2);
endfor
hold off;
```

Εκθετική κατανομή

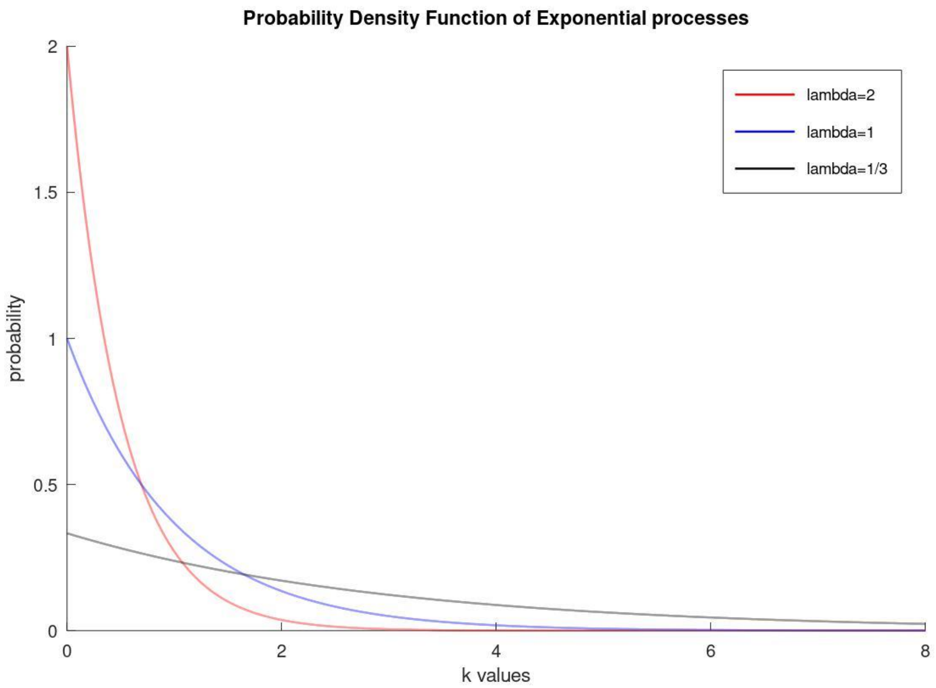
Έστω Χ μια τυχαία μεταβλητή που ακολουθεί την εκθετική κατανομή τότε για τη συνάρτηση πυκνότητας πιθανότητας της Χ έχουμε:

$$X \sim Exp(\lambda), f_X(t) = \lambda e^{-\lambda t}, t \geq 0$$

Έχουμε ότι E(X) = 1/λ και Var(X) = $1/\lambda^2$.

A)

```
clc;
clear all;
close all;

% TASK: In a common diagram, design the Probability Density Function
of
%Exponential processes.
% with mean values 0.5, 1, 3. In the horizontal axis the range is
from 0 to 8.

pkg load statistics

colors = "rbkmg";

step = 0.0001;

k = 0:step:8;
%changed the step to 0.0001 because it was too slow otherwise
mean_exp = [0.5 , 1, 3];

for i = 1 : columns(mean_exp)
  expon(i, :) = exppdf(k, mean_exp(i));
endfor

colors = "rbkmg";

figure(1);
hold on;
for i = 1 : columns(mean_exp)
  plot(k, expon(i, :), colors(i), "linewidth", 1.2);
endfor
hold off;

title("Probability Density Function of Exponential processes");
xlabel("k values");
ylabel("probability");
legend("lambda=2", "lambda=1", "lambda=1/3");
```



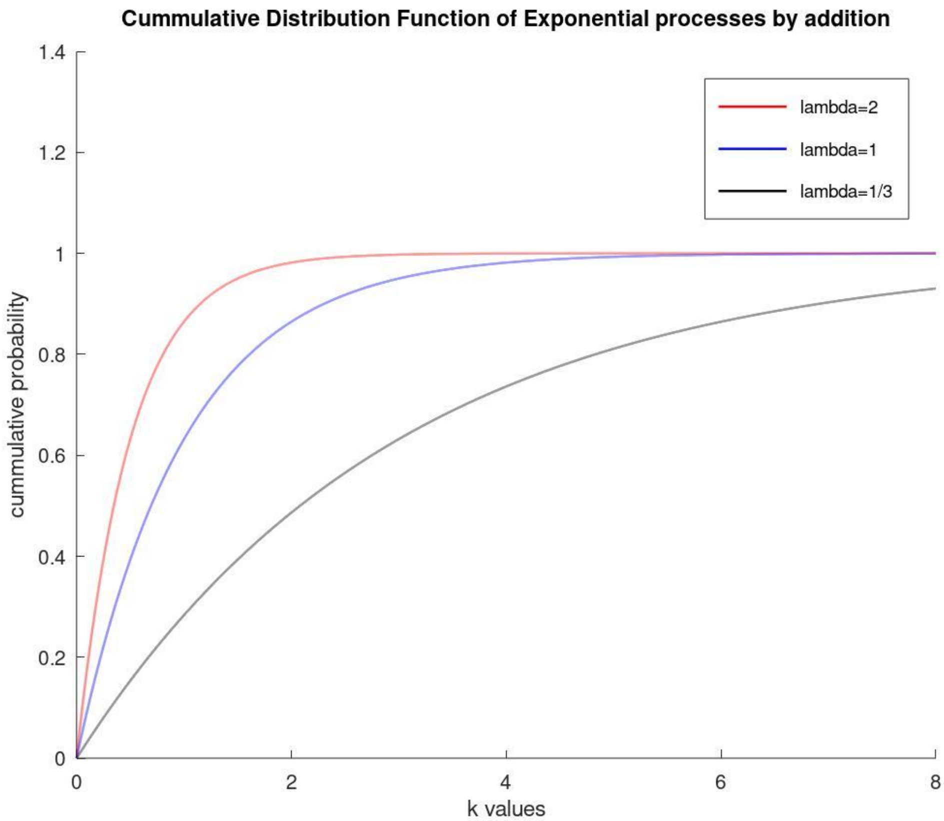
Παρατηρούμε ότι για μεγαλύτερα λ η συνάρτηση πυκνότητας πιθανότητας γίνεται πιο απότομη, καθώς και η μέση τιμή μικραίνει (αφού είναι το αντίστροφο του λ). Επίσης, τα γραφήματα ξεκινούν για k = 0 από το λ.

B) Έχουμε παραγάγει την αθροιστική συνάρτηση κατανομής με 2 τρόπους. Ο πρώτος είναι χρησιμοποιώντας τη σχέση

$$F_X(x+dx) = F_X(x) + f_X(x)dx$$

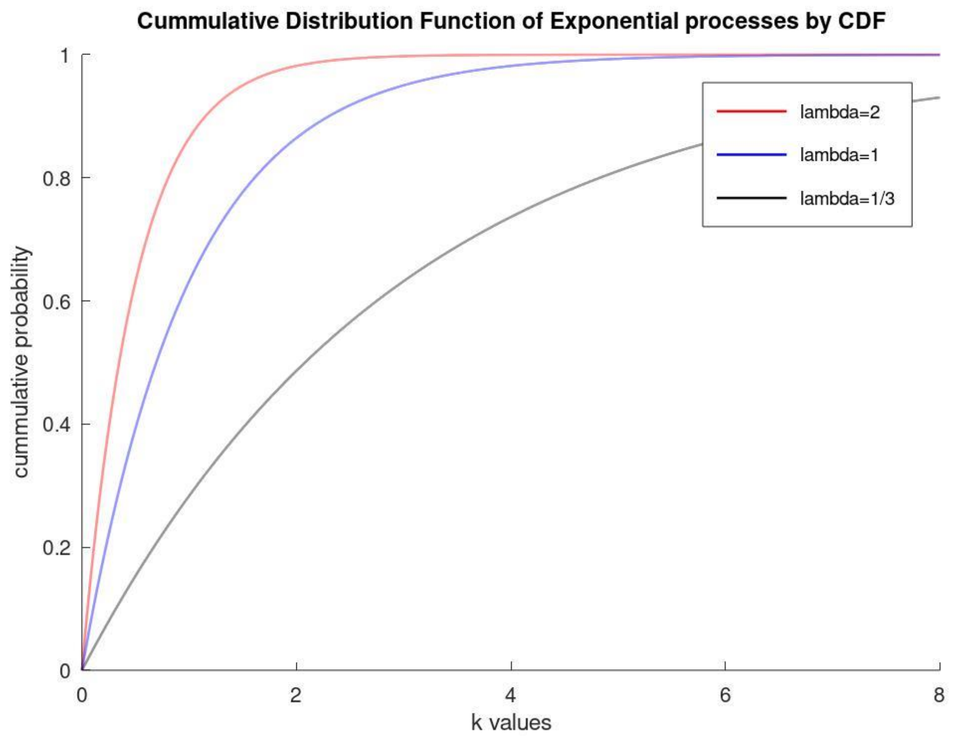
Όπου ως dx λαμβάνουμε το step που χρησιμοποιούμε στην οριζόντια κλίμακα.

```
for i = 1 : columns(mean_exp)
  cum_expon(i, 1) = expon(i, 1);
  for j = 2 : columns(expon(1, :))
    cum_expon(i, j) = cum_expon(i, j-1) + expon(i, j);
  endfor
endfor

%multiply with step, because dx = step (kind of...)
cum_expon = cum_expon*step;

figure(2);
hold on;
for i = 1 : columns(mean_exp)
  plot(k, cum_expon(i, :), colors(i), "linewidth", 1.2);
endfor
hold off;

title("Cummulative Distribution Function of Exponential processes by
addition");
xlabel("k values");
ylabel("cummulative probability");
legend("lambda=2", "lambda=1", "lambda=1/3");
```



Ο δεύτερος τρόπος υπολογισμού έγινε με χρήση της expcdf που δίνει έτοιμη το octave.

```
for i = 1 : columns(mean_exp)
  cum_expon_2(i, :) = expcdf(k, mean_exp(i));
endfor

colors = "rbkmg";

figure(3);
hold on;
for i = 1 : columns(mean_exp)
  plot(k, cum_expon_2(i, :), colors(i), "linewidth", 1.2);
endfor
hold off;

title("Cummulative Distribution Function of Exponential processes by
CDF");
xlabel("k values");
ylabel("cummulative probability");
legend("lambda=2", "lambda=1", "lambda=1/3");
```



Θεωρητικά, αναμένουμε η αθροιστική συνάρτηση κατανομής της X να είναι:

$$\alpha\nu\ X \sim Exp(\lambda), F_X(t) = 1 - e^{-\lambda t}, t \geq 0$$

Επαληθεύεται από τα γραφήματα. Παρατηρούμε ότι για μεγάλα λ (άρα μικρές μέσες τιμές) η γραφική παράσταση της αθροιστικής κατανομής πιθανότητας φτάνει γρηγορότερα στη μονάδα, αφού, όπως είδαμε και από την κατανομή πυκνότητας πιθανότητας, τα ενδεχόμενα για μικρές τιμές του άξονα k είναι πιο πιθανά (η μέση τιμή είναι μικρότερη και μεγαλύτερο μέρος της γραφικής παράστασης είναι συγκεντρωμένο κοντά της).

Γ) Αφού εμείς χρησιμοποιήσαμε για step το 0.0001 θα έχουμε (3000, 5000, 2000) αντί για (30000, 50000, 20000). Γενικά, τα σημεία που θέλουμε είναι (0.3, 0.5, 0.2)/step. Χρησιμοποιούμε τη τελευταία έκφραση για να μπορούμε να βρούμε τη πιθανότητα αυτού του ενδεχομένου ανεξάρτητα από την επιλογή step στον κώδικα.

Για τα 3 αυτά σημεία θα βρούμε τις πιθανότητες η τυχαία μεταβλητή να έχει τιμή μεγαλύτερη από αυτά μέσω της ιδιότητας:

$$P(X \geq t) = 1 - F_X(t)$$

Και $P(X > 50000|X > 20000) = \frac{P(X>50000)}{P(X>20000)}$

Για αυτό το ερώτημα, μιας και δεν έχει γραφικές παραστάσεις χρησιμοποιήθηκε κλίμακα k = 0:0.00001:8, ωστόσο, επειδή στον κώδικά μας τα πάντα είναι παραμετρικά με το step αυτό δεν επηρεάζει τα υπόλοιπα.

```
Pr_30000 = 0.8869
Pr_50000_given_20000 = 0.8869
```

Παίρνουμε αυτή την έξοδο. Παρατηρούμε ότι οι 2 τιμές είναι ίδιες. Το αποτέλεσμα αυτό είναι αναμενόμενο, αφού η κατανομή που εξετάζουμε είναι εκθετική και άρα έχει έλλειψη μνήμης. Δηλαδή, το παραπάνω είναι αποτέλεσμα της εξής ιδιότητας της εκθετικής κατανομής:

$$P(X > t + s|X > s) = P(X > t)$$

```
mean_expC = 2.5;
cum_exponC = expcdf(k, mean_expC);

%rounding needed because matrices accept only integers as indexes
points = round([0.3, 0.5, 0.2]/step);

Pr_30000=1-cum_exponC(points(1))
%Pr(X>30000) in original scale

Pr_50000=1-cum_exponC(points(2));
Pr_20000=1-cum_exponC(points(3));

Pr_50000_given_20000=Pr_50000/Pr_20000
%Pr(X>50000|X>20000) in original scale
```

Διαδικασία Καταμέτρησης Poisson

Α) Έχουμε μια διαδικασία καταμέτρησης Poisson N(t). Θεωρούμε ρυθμό λ = 5 γεγονότα/sec.

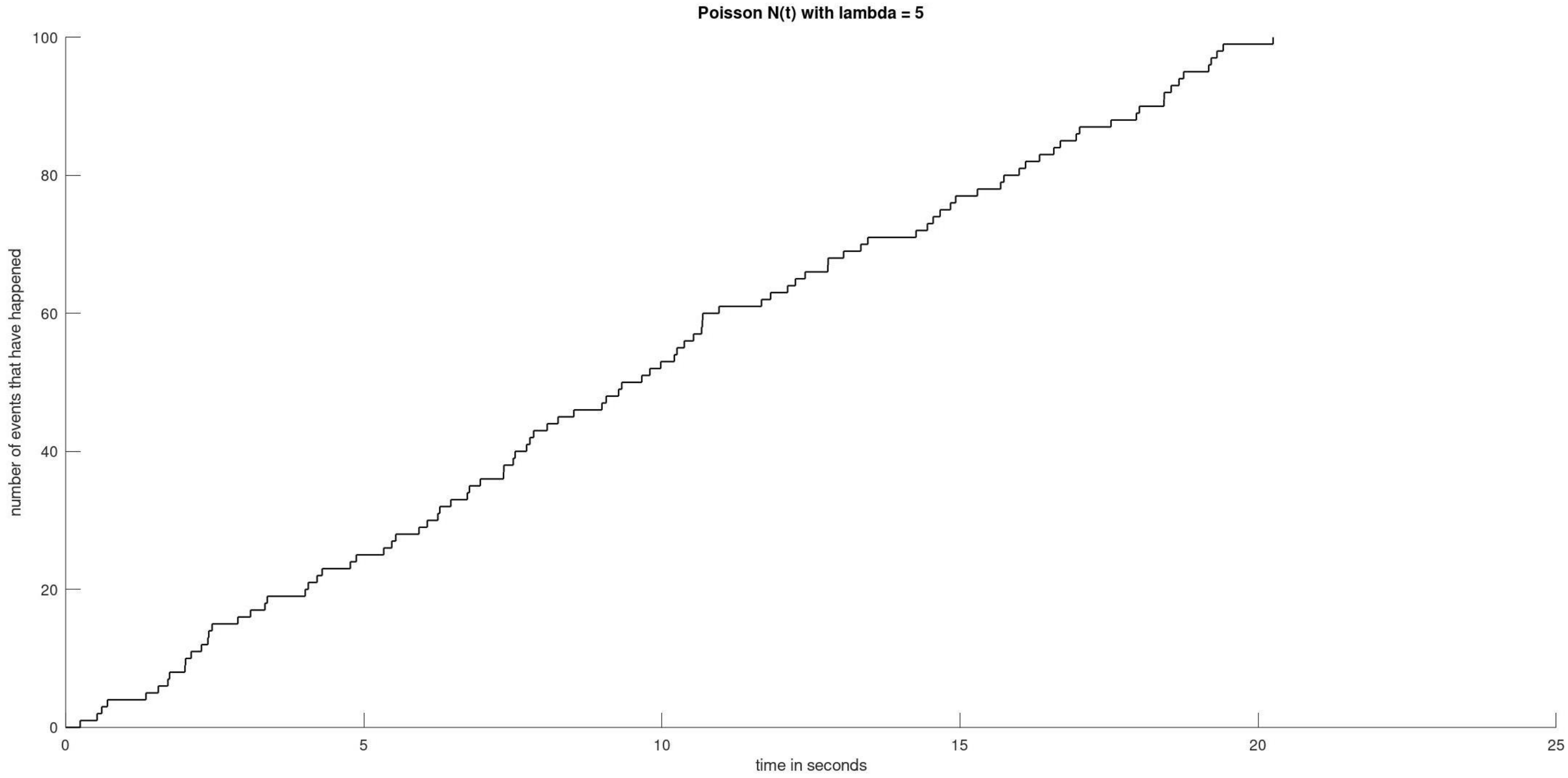
Έστω T η τυχαία μεταβλητή που παριστάνει το χρονικό διάστημα μεταξύ 2 γεγονότων. Τότε γνωρίζουμε ότι:

$$N(t) \sim Poisson(\lambda t) \;\; και \;\; T \sim Exponential(\lambda)$$

Στον κώδικα έχουμε βάλει τα διανύσματα που μπαίνουν σε άξονες να έχουν ((πλήθος γεγονότων που συνέβησαν) + 1) στοιχεία. Αυτό διότι θεωρούμε ότι τη στιγμή 0 συμβαίνει ένα γεγονός, άρα αυτό δεν είναι τυχαίο, και μετά θα συμβούν όλα τα υπόλοιπα τυχαία γεγονότα.

Με βάση τις παρατηρήσεις για N(t) και T. Φτιάχνουμε n (όπου n το πληθος γεγονότων που συμβαίνουν) τυχαία διαστήματα με βάση την Exponential(λ) και σε κάθε βήμα προσθέτουμε στον χρόνο ένα από αυτά καθώς κα 1 μονάδα στα τυχαία γεγονότα που έχουν συμβεί. Αυτό μας δίνει τη διαδικασία καταμέτρησης Poisson μέχρι να συμβούν n τυχαία γεγονότα.

```
clc;
clear all;
close all;

pkg load statistics

lambda = 5;

mean_exp = ones(1, 100)/lambda;
%the mean is the inverse of lambda
random_intervals = exprnd(mean_exp);
%create 100 random intervals from exponential distribution with mean
1/lambda

k(1) = 0;
steps(1) = 0;
%we suppose that the first event happens at time 0, thus making it
not random,
%however, the next 100 events are random
for i = 1 : columns(random_intervals)
  k(i+1) = k(i)+random_intervals(i);
  steps(i+1) = steps(i)+1;
endfor
%k is now a vector that contains all the exact moments in time that
events happen

figure(1);
hold on;

stairs(k, steps, 'k', "linewidth", 1);
hold off;

title("Poisson N(t) with lambda = 5");
xlabel("time in seconds");
ylabel("number of events that have happened");

%mean number of events per unit of time is lambda
mean(1) = 100/k(101);
```



Poisson N(t) with lambda = 5
number of events that have happened
time in seconds
0
20
40
60
80
100
0
5
10
15
20
25

B) Σε ένα χρονικό παράθυρο ΔT = t1 – t2 ο αριθμός των γεγονότων ακολουθεί την ίδια κατανομή με την N(t), διότι οι μεταβολές της διαδικασίας είναι στατικές (Η κατανομή του αριθμού γεγονότων σε οποιοδήποτε χρονικό διάστημα εξαρτάται μόνο από το μήκος του διαστήματος και όχι από την αρχή του).

Για να το εκφράσουμε καλύτερα έχουμε ότι ο αριθμός των γεγονότων στο διάστημα ΔT = t1 – t2 ακολουθεί την κατανομή Poisson(λ*ΔT), όπου λ το ίδιο με πριν.

Στο ερώτημα αυτό δεν μας ζητείται να επαναλάβουμε το σχεδιασμό των γραφικών παραστάσεων. Μόνον να υπολογίσουμε το μέσο αριθμό γεγονότων στη μονάδα του χρόνου, δηλαδή, το λ. Υπολογίζουμε τη μέση τιμή αυτή ως εξής:

$$\mu = \frac{(\pi\lambda\eta\theta o\varsigma\ \gamma\varepsilon\gamma o\nu\acute{o}\tau\omega\nu)}{(\chi\rho\acute{o}\nu o\varsigma\ \alpha\pi\acute{o}\ \tau\eta\nu\ \alpha\rho\chi\acute{\eta}\ \mu\acute{\varepsilon}\chi\rho\iota\ \kappa\alpha\iota\ \tau\eta\ \sigma\tau\iota\gamma\mu\acute{\eta}\ \pi o\upsilon\ \sigma\upsilon\mu\beta\alpha\acute{\iota}\nu\varepsilon\iota\ \tau o\ \tau\varepsilon\lambda\varepsilon\upsilon\tau\alpha\acute{\iota}o\ \gamma\varepsilon\gamma o\nu\acute{o}\varsigma)}.$$

```
n_events = [200, 300, 500, 1000, 10000];

for i = 1 : columns(n_events)

  clear mean_exp
  mean_exp= ones(1, n_events(i))/lambda;

  clear random_intervals
  random_intervals = exprnd(mean_exp);

  clear k
  k(1) = 0;

  clear steps
  steps(1) = 0;

%we clear all variables that we will reuse
  for j = 1 : columns(random_intervals)
    k(j+1) = k(j) + random_intervals(j);
    steps(j+1) = steps(j) + 1;
  endfor
  mean(i+1) = n_events(i)/k(n_events(i)+1);
endfor

mean.'
%print all the lambdas
```

Στην έξοδο του octave παίρνουμε τον διπλανό πίνακα μέσων τιμών. Ο οποίος, για να τον εμφανίσουμε πιο δομημένα στο word είναι ο παρακάτω:

```
ans =

   4.9378
   5.1709
   5.1080
   5.2522
   5.0938
   5.0484
```

Πλήθος γεγονότων	Μέσο πλήθος στη μονάδα του χρόνου
100	4.9378
200	5.1709
300	5.1080
500	5.2522
1000	5.0938
10000	5.0484

Όπως είπαμε και ανωτέρω ο μέσος ρυθμός γεγονότων ανά μονάδα χρόνου είναι το λ, που εδώ, σε όλες τις περιπτώσεις είναι 5 γεγονότα/sec. Βλέπουμε ότι οι παραπάνω μέσοι ρυθμοί πλησιάζουν την τιμή αυτή.

Τα αποτελέσματα της άσκησης με τίτλο «Διαδικασία Καταμέτρησης Poisson» πάρθηκαν όλα από το ίδιο τρέξιμο του προγράμματος.

Ο κώδικας που χρησιμοποιήθηκε υπάρχει στα αντίστοιχα σημείο. Συμπεριλαμβάνονται όμως και τα αρχεία στο αρχείο που θα υποβληθεί.