Συστήματα Αναμονής, 1η εργαστηριακή άσκηση

Νικόλαος Παγώνας, el18175

Απρίλιος 2021

Κατανομή Poisson

A.

Με τη βοήθεια του πακέτου **statistics** της Octave και της συνάρτησης **stem**, απεικονίζουμε τη συνάρτηση μάζας πιθανότητας της κατανομής Poisson, για $\lambda = \{3, 10, 50\}$.

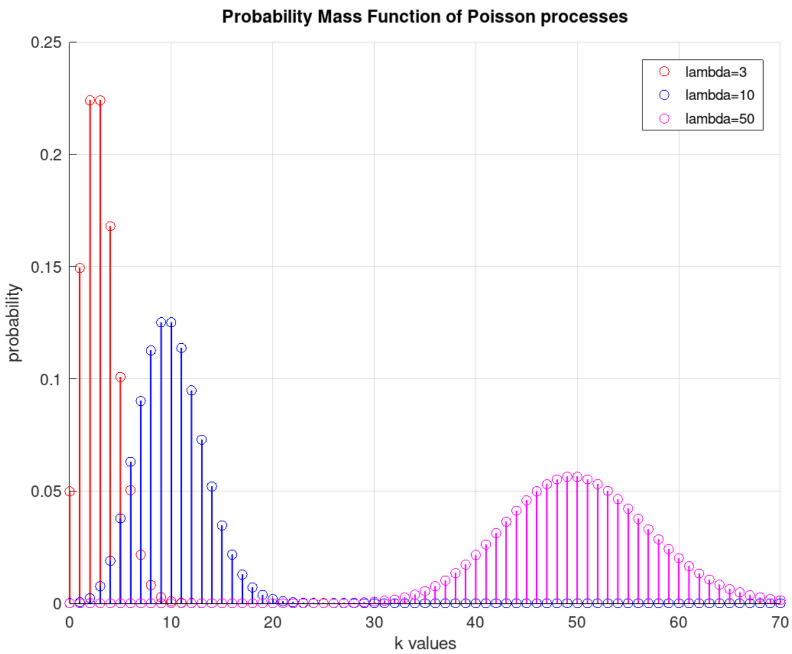

Figure 1: Probability Mass Function of Poisson processes

Παρατηρούμε ότι όσο μεγαλώνει η τιμή της παραμέτρου λ:

1. η κατανομή “απλώνεται”, δηλαδή οι τιμές των θέσεων αριστερά και δεξιά από την μέση τιμή της κατανομής αυξάνονται. Αυτό είναι λογικό, αφού η διακύμανσή της, που είναι ίση με λ, αυξάνεται.

2. η θέση όπου η κατανομή είναι μέγιστη, πηγαίνει προς τα δεξιά. Αυτό συμβαίνει γιατί η μέση τιμή της (επίσης ίση με λ) αυξάνεται.

3. η μέγιστη τιμή της κατανομής μειώνεται. Αυτό συμβαίνει γιατί πρέπει να ικανοποιείται η ιδιότητα:

$$\sum_x P(X = x) = 1,$$

   η οποία ισχύει για κάθε διακριτή τυχαία μεταβλητή. Έτσι, αφού οι τιμές αριστερά και δεξιά της μέσης τιμής αυξάνονται, το μέγιστο οφείλει να μειωθεί και η κατανομή “χαμηλώνει” συνολικά.

Β.

Με βάση τους ορισμούς της μέσης τιμής και της διακύμανσης:

$$E[X] = \sum_x xP(X = x)$$

και

$$Var[X] = E[X^2] - E^2[X]$$

Έτσι, για $\lambda = 30$ έχουμε:

```
mean value of Poisson with lambda 30 is 30
Variance of Poisson with lambda 30 is 30
```

Οπότε επαληθεύουμε και μέσω της προσομοίωσης αυτό που ξέρουμε ότι ισχύει από τη θεωρία για την κατανομή Poisson:

$$E[X] = Var[X] = \lambda = 30$$

Γ.

Επιλέγουμε τις κατανομές Poisson με παραμέτρους $\lambda = 10$ και $\lambda = 50$. Γνωρίζουμε από τη Θεωρία Πιθανοτήτων ότι η κατανομή της τυχαίας μεταβλητής $Z = X + Y$ προκύπτει από την συνέλιξη των κατανομών των X και Y (εδώ κάνουμε την υπόθεση ότι οι X και Y είναι ανεξάρτητες).

Απεικονίζουμε τις κατανομές των $X \sim Pois(10), Y \sim Pois(50)$ και $Z = X + Y$:

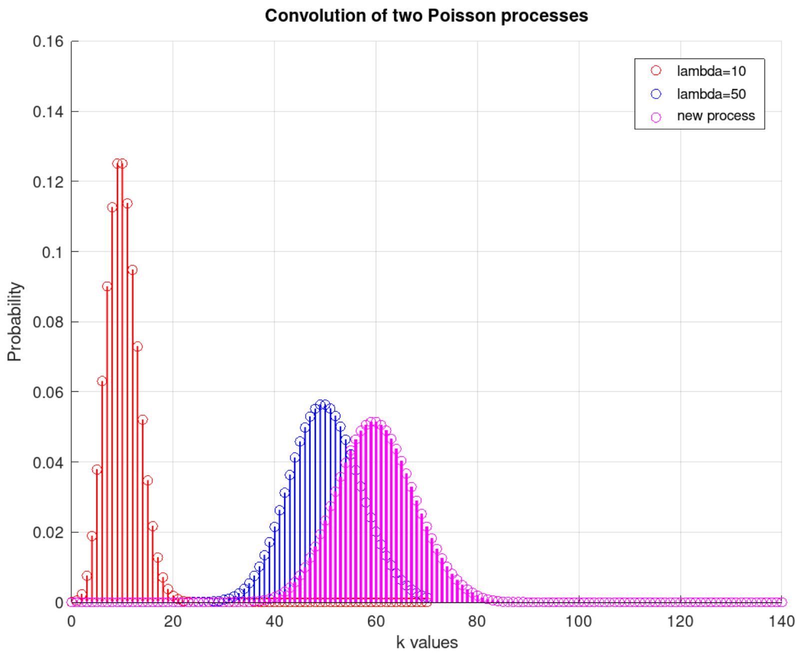

Figure 2: Convolution of two Poisson processes

Παρατηρούμε ότι προκύπτει κατανομή **Poisson** με μέση τιμή 60. Αυτό είναι λογικό, αφού γνωρίζουμε πως το άθροισμα δύο τυχαίων μεταβλητών Poisson με παραμέτρους λ_1, λ_2 ακολουθεί κατανομή Poisson με παράμετρο $\lambda_1 + \lambda_2$. Απαραίτητη προϋπόθεση για να ισχύει αυτό είναι οι τυχαίες μεταβλητές να είναι **ανεξάρτητες**.

Δ.

Αποδεικνύουμε τον παραπάνω ισχυρισμό:

Έστω ένα διάστημα χρόνου t.

- Διαιρούμε το διάστημα t σε n υποδιαστήματα, $t = n\Delta t$
- Πραγματοποιούμε n ανεξάρτητες δοκιμές Bernoulli, μία σε κάθε υποδιάστημα, με πιθανότητα επιτυχίας $p = \lambda\Delta t$
- Έτσι, η πιθανότητα k επιτυχιών σε n ανεξάρτητες δοκιμές δίνεται από την

Διωνυμική Κατανομή:

$$\begin{aligned} P[N(t)=k] &= \binom{n}{k} p^k (1-p)^{n-k},\ k=0,1,...,n \\ &= \binom{n}{k} (\lambda \Delta t)^k (1-\lambda \Delta t)^{n-k} \\ &= \binom{n}{k} \left(\frac{\lambda t}{n}\right)^k \left(1-\frac{\lambda t}{n}\right)^{n-k} \end{aligned}$$

- Και για $\Delta t \to 0, n \to \infty$, έχουμε

$$\frac{n!}{(n-k)!} \to n^k,\ \left(1-\frac{\lambda t}{n}\right)^{n-k} \to e^{-\lambda t},$$

οπότε

$$P[N(t)=k] = \frac{n!}{k!(n-k)!} \left(\frac{\lambda t}{n}\right)^k \left(1-\frac{\lambda t}{n}\right)^{n-k} \to \frac{(\lambda t)^k}{k!} e^{-\lambda t}. \quad \blacksquare$$

Με αυτόν τον τρόπο ($\lambda = np$, όπου n μεγάλο και p μικρό), μπορούμε να κατασκευάσουμε μία κατανομή Poisson παραμέτρου $\lambda = 30\ \frac{\text{σημεία}}{\text{sec}}$. Επιλέγουμε $n = 30, 60, 90, 120, 150$.

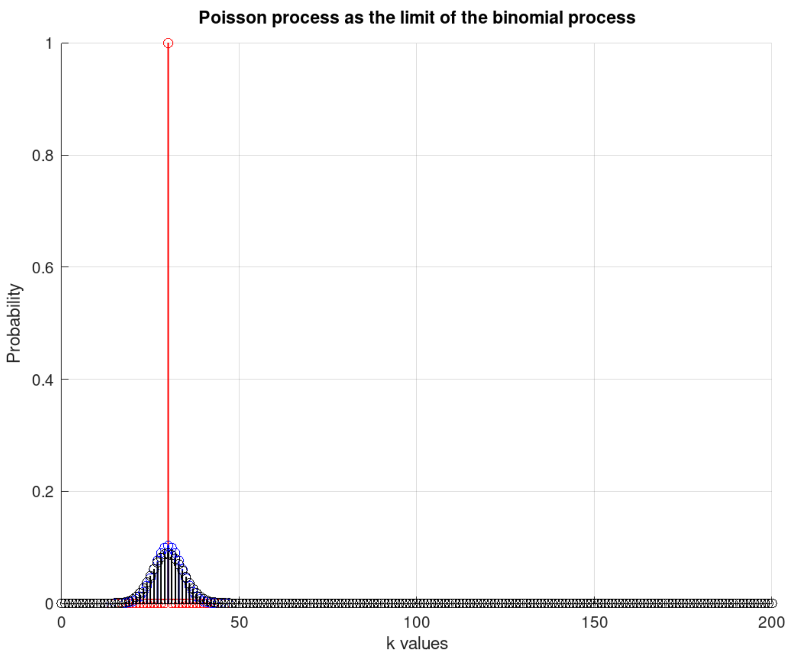

Figure 3: Poisson process as the limit of the binomial process

Παρατηρούμε ότι όσο το n αυξάνει, τόσο η διωνυμική κατανομή προσεγγίζει καλύτερα την Poisson.

Εκθετική Κατανομή

A.

Απεικονίζουμε στο Octave την συνάρτηση πυκνότητας πιθανότητας της εκθετικής κατανομής, με μέσο όρο $\frac{1}{\lambda} = \{0.5, 1, 3\}$. Χρησιμοποιώντας πολύ μικρό βήμα (k = 0:0.00001:8) μπορούμε να προσεγγίσουμε με μεγάλη ακρίβεια την συνεχή εκθετική κατανομή:

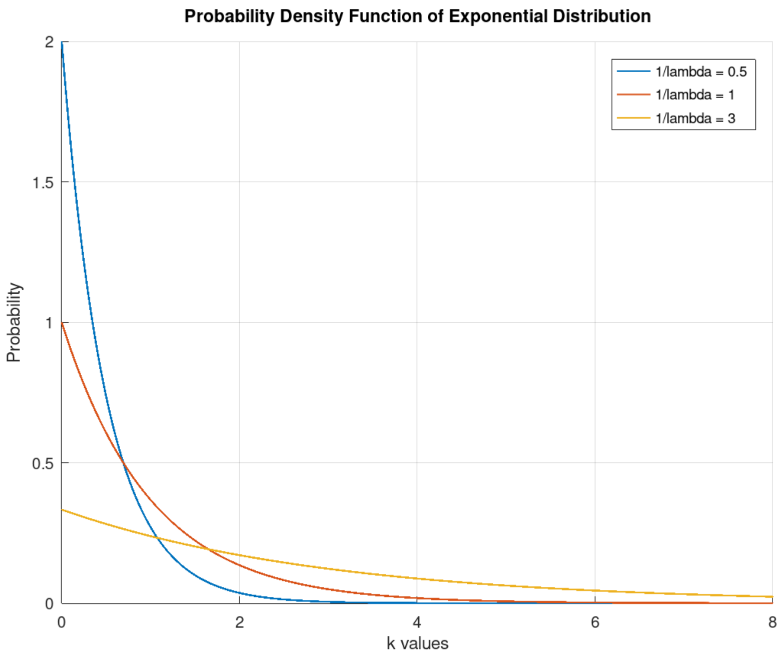

Figure 4: Probability Density Function of Exponential Distribution

B.

Αντίστοιχα, με τον ίδιο τρόπο απεικονίζουμε την αθροιστική συνάρτηση κατανομής, πάλι για $\frac{1}{\lambda} = \{0.5, 1, 3\}$:

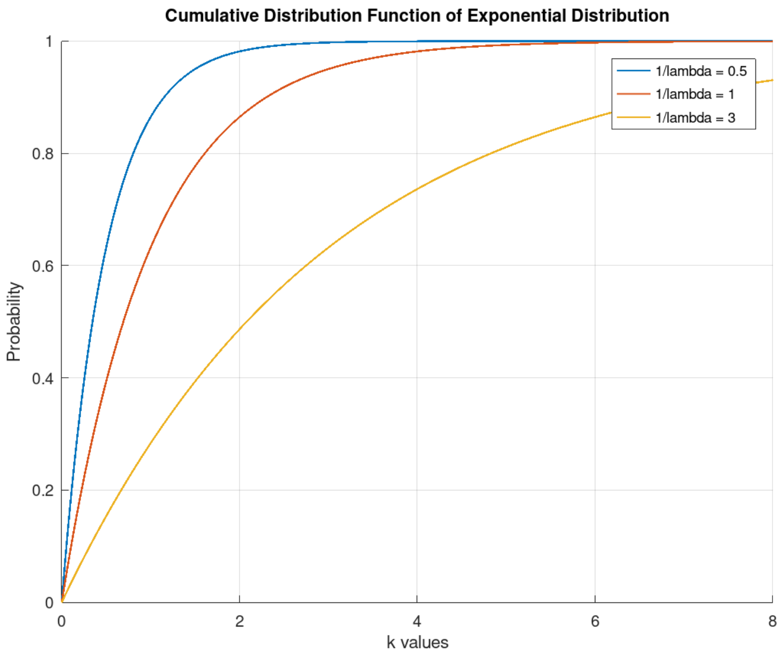

Figure 5: Cumulative Distribution Function of Exponential Distribution

Γ.

Χρησιμοποιώντας την συνάρτηση **exppdf()** και τον ορισμό της δεσμευμένης πιθανότητας, υπολογίζουμε τις ζητούμενες τιμές:

```
P(X > 50000 | X > 20000) with lambda 0.4 = 0.88692
P(X > 30000) with lambda 0.4 = 0.88692
```

Παρατηρούμε ότι οι δύο πιθανότητες είναι ίσες. Για να δούμε γιατί ισχύει αυτή η ισότητα, υπολογίζουμε λίγο πιο αναλυτικά τις πιθανότητες και έχουμε:

$$
\begin{aligned}
P(X > k[50000] | X > k[20000]) &= \frac{P(X > k[50000] \cap X > k[20000])}{P(X > k[20000])} \\
&= \frac{P(X > k[50000])}{P(X > k[20000])} \\
&= \frac{e^{\lambda \cdot k[50000]}}{e^{\lambda \cdot k[20000]}} \\
&= e^{\lambda \cdot k[30000]} \\
&= P(X > k[30000])
\end{aligned}
$$

Στην ουσία αποδείξαμε μία ειδική περίπτωση της ιδιότητας απώλειας μνήμης, χαρακτηριστικής ιδιότητας της εκθετικής κατανομής. Η ιδιότητα αυτή υπαγορεύει ότι:

$$P[X > t + s|X > s] = P[X > t]$$

Μία πιο πρακτική ερμηνεία αυτής της ιδιότητας είναι ότι, αν ο χρόνος αναμονής κάποιας διαδικασίας μοντελοποιηθεί ως εκθετική τυχαία μεταβλητή, τότε ο υπολειπόμενος χρόνος θα έχει συνεχώς την ίδια εκθετική κατανομή, χωρίς να έχει σημασία το πόσος χρόνος έχει περάσει ήδη.

Διαδικασία Καταμέτρησης Poisson

A.

Γνωρίζουμε ότι οι χρόνοι που μεσολαβούν ανάμεσα στην εμφάνιση δύο διαδοχικών γεγονότων της διαδικασίας Poisson με παράμετρο λ ακολουθούν εκθετική κατανομή με μέση τιμή $\frac{1}{\lambda}$.

Έτσι, με την εντολή **exprnd()** ($\lambda = 5$) δημιουργούμε 100 τέτοιους χρόνους, δηλαδή 100 διαδοχικά τυχαία γεγονότα, και μέσω της συνάρτησης stairs() απεικονίζουμε μία διαδικασία καταμέτρησης Poisson:

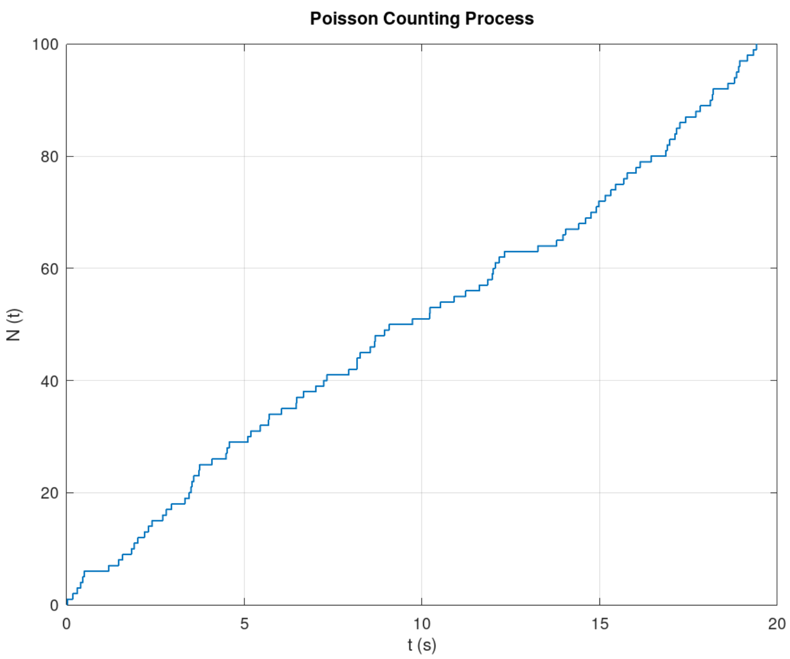

Figure 6: Poisson Counting Process

B.

Γνωρίζουμε ότι ο αριθμός γεγονότων σε ένα χρονικό παράθυρο $\Delta T = t_1 - t_2$ ακολουθεί κατανομή Poisson με παράμετρο $\lambda\Delta T$ (όπου λ η παράμετρος της διαδικασίας καταμέτρησης Poisson), δηλαδή ο μέσος αριθμός εμφανίσεων είναι ανάλογος του διαστήματος ΔT.

Βρίσκουμε τον μέσο αριθμό γεγονότων στην μονάδα του χρόνου $\left(\frac{\text{αριθμός γεγονότων}}{\text{συνολικός χρόνος}}\right)$, για $\{200, 300, 500, 1000, 10000, 100000, 1000000\}$ διαδοχικά γεγονότα. Προκύπτει:

```
Average for 100 events: 5.15191
Average for 200 events: 4.66471
Average for 300 events: 4.71886
Average for 500 events: 4.61717
Average for 1000 events: 4.98296
Average for 10000 events: 4.99382
Average for 100000 events: 5.01452
Average for 1000000 events: 5.00161
```

Παρατηρούμε ότι όσο ο αριθμός γεγονότων μεγαλώνει, τόσο ο μέσος αριθμός γεγονότων πλησιάζει την παράμετρο $\lambda = 5 \ \frac{\text{γεγονότα}}{\text{sec}}$, ως οφείλει, αφού πρόκειται για διαδικασία καταμέτρησης Poisson.

Παράρτημα Κώδικα

Κατανομή Poisson

```
pkg load statistics

clc;
clear all;
close all;

nfig=0;
# Poisson Distribution A
# TASK: In a common diagram, design the Probability Mass Function of Poisson
# processes with lambda parameters 3,10,50. In the horizontal axes, choose
# k parameters between 0 and 70.

k = 0:1:70;
lambda = [3,10,30,50];
index = find(lambda == 30);

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

colors = "rbkm";

nfig=nfig+1; figure(nfig);
hold on;
for i=1:columns(lambda)
```

```
  if (i == index)
    continue
  endif
  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=50");
grid on;

#name = "../images/Poisson_Distribution_A.png";
#saveas(nfig,name);



# Poisson Distribution B
# TASK: regarding the poisson process with parameter lambda 30, compute its mean
# value and variance

#fd = fopen("../images/Poisson_Distribution_B.txt","w");

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

fprintf("mean value of Poisson with lambda 30 is %d\n", 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;

fprintf("Variance of Poisson with lambda 30 is %d\n",variance);
# fclose(fd);
# Poisson Distribution C
# TASK: consider the convolution of the Poisson distribution with lambda 20 with
# the Poisson distribution with lambda 30.

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);

nfig=nfig+1; figure(nfig);
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");
grid on;

#name = "../images/Poisson_Distribution_C.png";
#saveas(nfig,name);

# Poisson Distribution D
# TASK: show that Poisson process is the limit of the binomial
   distribution.
k = 0:1:200;
# Define the desired Poisson Process
lambda = 30;
n = [30, 60, 90, 120, 150];
p = lambda./n;

nfig=nfig+1; figure(nfig);

hold on;
for i=1:3
  binomial = binopdf(k,n(i),p(i));
  stem(k,binomial,colors(i),'linewidth',1.2);
endfor
hold off;

title("Poisson process as the limit of the binomial process");
xlabel("k values");
ylabel("Probability");
grid on;

#name = "../images/Poisson_Distribution_D.png";
#saveas(nfig,name);
```

Εκθετική Κατανομή

```
pkg load statistics

clc;
clear all;
close all;

nfig=0;
# Exponential Distribution A

k = 0:0.00001:8;
means = [0.5,1,3];

for i = 1:3
  exp_pdf(i,:) = exppdf(k, means(i));
endfor

nfig=nfig+1; figure(nfig);
hold on;
for i = 1:3
  plot(k,exp_pdf(i,:),"linewidth",1.2);
endfor
hold off;
```

```
title("Probability Density Function of Exponential Distribution");
xlabel("k values");
ylabel("Probability");
legend("1/lambda = 0.5","1/lambda = 1","1/lambda = 3");
grid on;

#name = "../images/Exponential_Distribution_A.png";
#saveas(nfig,name);

# Exponential Distribution B

for i = 1:3
  exp_cdf(i,:) = expcdf(k, means(i));
endfor

nfig=nfig+1; figure(nfig);
hold on;
for i = 1:3
  plot(k,exp_cdf(i,:),"linewidth",1.2);
endfor
hold off;

title("Cumulative Distribution Function of Exponential Distribution");
xlabel("k values");
ylabel("Probability");
legend("1/lambda = 0.5","1/lambda = 1", "1/lambda = 3");
grid on;

#name = "../images/Exponential_Distribution_B.png";
#saveas(nfig,name);

# Exponential Distribution C

mean = 2.5; # mean=1/lambda

P_30000 = 1 - expcdf(k(30000), mean);
P_50000_20000 = (1 - expcdf(k(50000), mean))/(1 - expcdf(k(20000), mean));

#fd = fopen("../images/Exponential_Distribution_C.txt","w");
fprintf("P(X > 50000 | X > 20000) with lambda 0.4 = %.5d\n", P_50000_20000);
fprintf("P(X > 30000) with lambda 0.4 = %.5d\n", P_30000);
#fclose(fd);
```

Διαδικασία Καταμέτρησης Poisson

```
pkg load statistics

clc;
clear all;
close all;

nfig = 0;
# Poisson Counting A
sample_sizes = [100,200,300,500,1000,10000,100000,1000000];

#fd = fopen("../images/Poisson_Counting_B.txt","w");
```

```
for j = 1:length(sample_sizes)
  lambda = 5;
  mean = 1/lambda;
  sample_size = sample_sizes(j);

  rnd_events = exprnd(mean,1,sample_size);
  t = zeros(1,sample_size+1);

  for i = 1:sample_size
    t(i+1) = t(i) + rnd_events(i);
  endfor

  N = 0:1:sample_size;


  if sample_size == 100
    nfig=nfig+1; figure(nfig);
    stairs(t,N,"linewidth",1.2);
    title("Poisson Counting Process");
    xlabel("t (s)");
    ylabel("N (t)");
    grid on;

 # saveas(nfig,"../images/Poisson_Counting_A.png");
  endif

  # Poisson Counting B

  average = sample_size/t(end);

  fprintf("Average for %d events: %d\n", sample_size, average);

endfor

# fclose(fd);
```