Συστήματα Αναμονής Άσκηση 1^{η}

Αυγερινός Πέτρος 03115074

Contents

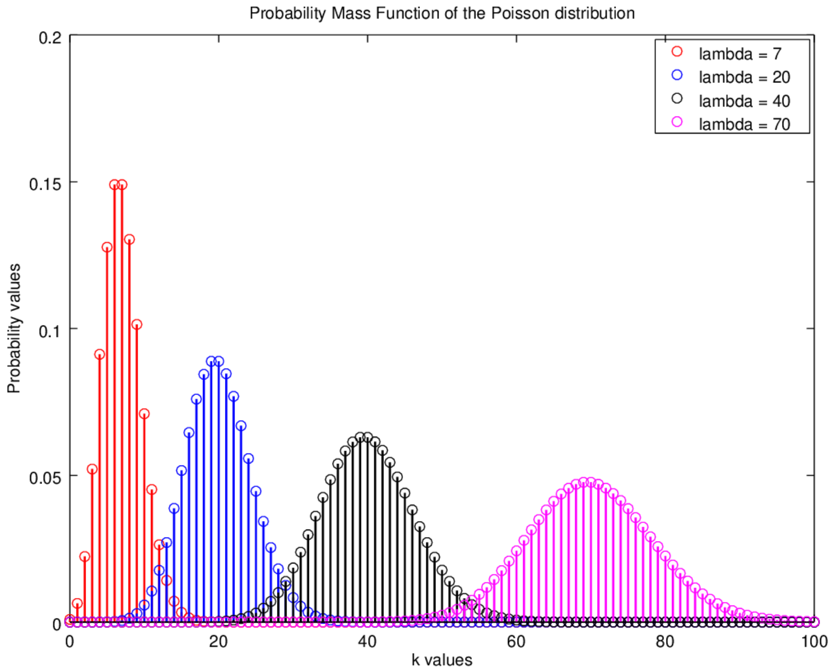
1 Κατανομη Poisson

1.1 Ερώτημα Α

Η διακριτή τυχαία μεταβλητή ακολουθεί κατανομή Poisson με παράμετρο $\lambda > 0$. Η συνάρτηση πυκνότητας πιθανότητας της μεταβλητής θα έχει ως εξής:

$$P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \tag{1}$$

Το κοινό διάγραμμα παρουσίασης της κατανομής Poisson για διαφορετικές τιμές του λ φαίνεται στο παρακάτω διάγραμμα:



Ο κώδικας που χρησιμοποιήθηκε για την παραγωγή του διαγράμματος είναι παρακάτω και παραδόθηκε από τους διδάσκοντες:

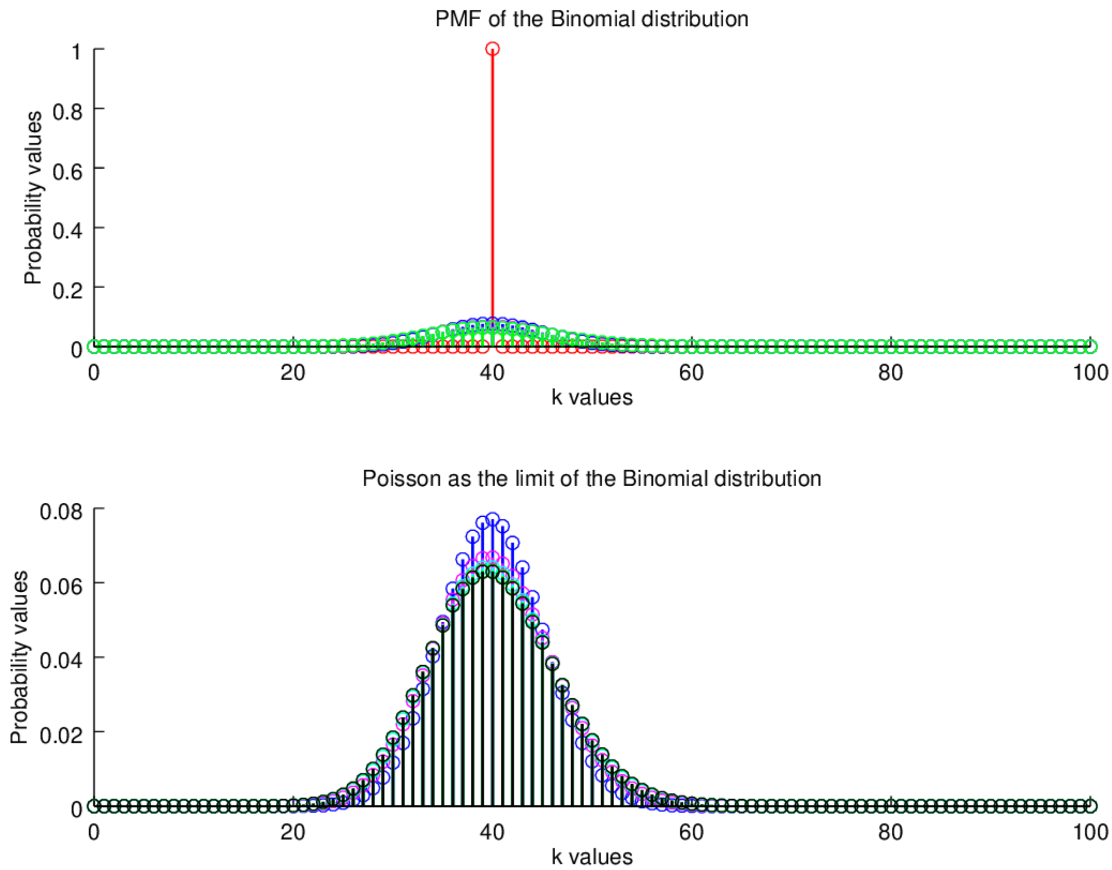
```
# TASK_A: In a common diagram, design the Probability Mass Function (PMF)
# of the Poisson distribution with lambda parameters 7, 20, 40, 70.
lambda = [7,20,40,70];

# For horizontal axis, choose k between 0 and 100, with step 1.
k = 0:1:100;

# Define the Poisson PMF values
for i=1:columns(lambda)
  poisson(i,:) = poisspdf(k,lambda(i));
endfor

# The colors for each plot-line are respectively:
colors = "rbkm";
figure(1);
hold on;

# Plotting the Poisson with parameter lambda, i.e. Po(lambda)
for i=1:columns(lambda)
  stem(k,poisson(i,:),colors(i),"linewidth",1.2);
endfor
hold off;
title("Probability Mass Function of the Poisson distribution");
xlabel("k values");
ylabel("Probability values");
legend("lambda = 7","lambda = 20","lambda = 40","lambda = 70");
```

Από το διάγραμμα είναι σαφές πως με την αύξηση της παραμέτρου λ μειώνεται το ύψος της καμπύλης και μετατοπίζεται προς τα δεξιά. Γνωρίζουμε από την θεωρία πως η μέση τιμή της κατανομής Poisson θα είναι ίση με την παράμετρο λ και η διακύμανση θα είναι ίση με την παράμετρο λ επίσης. Η παράμετρος λ αντιπροσωπεύει τον μέσο ρυθμό εμφάνισης των γεγονότων στο χρόνο. Όσο αυξάνεται η τιμή της παραμέτρου λ τόσο μεγαλύτερη είναι η πιθανότητα εμφάνισης γεγονότων σε ένα διάστημα χρόνου.

1.2 Ερώτημα Β

Μπορούμε να προσεγγίσουμε την κατανομή Poisson με την διωνυμική κατανομή ως το όριο της διωνυμικής κατανομής όταν ο αριθμός των δοκιμών της διωνυμικής κατανομής τείνει προς το άπειρο και η πιθανότητα επιτυχίας της διωνυμικής κατανομής τείνει προς το μηδέν όπου η παράμετρος λ της κατανομής Poisson είναι ίση με το γινόμενο των δύο παραμέτρων της διωνυμικής κατανομής.

Το διάγραμμα φαίνεται παρακάτω:



Ο κώδικας που χρησιμοποιήθηκε για την παραγωγή του διαγράμματος είναι ο εξής:

```
# TASK_B: Show that Poisson is the limit of the binomial distribution.
# Define the Poisson parameter lambda
lambda_constant = 40;

# Define the Binomial parameters n (number of trials) and p (probability for success)
# accordingly, for the approximation:
n = [40,120,360,1080,40000];
p = lambda_constant./n;

# Define the Binomial PMF values
for i=1:columns(n)
  binomial(i,:) = binopdf(k,n(i),p(i));
endfor

colors = "ygmbr";
figure(2);
subplot(2,1,1);
hold on;

# Plotting the Bino(n,p) for all the values of n
for i=1:columns(n)
  stem(k,binomial(i,:),colors(i),"linewidth",1.2);
endfor
title("PMF of the Binomial distribution");
xlabel("k values");
ylabel("Probability values");
legend();
hold off;
subplot(2,1,2);
hold on;

# In order to zoom in and notice the approximation
# better, we will work in the following way:

# Obtain the position of the Po(40) from the 1st Fig.
index = find(lambda == 40);
# Plot the Bino(n,p) without the first (n=40) value
for i=2:columns(n)
  stem(k,binomial(i,:),colors(i),"linewidth",1.2);
endfor
# Include the Po(40) from the 1st Fig. with the same color
stem(k,poisson(index,:),"k","linewidth",1.2);
hold off;
title("Poisson as the limit of the Binomial distribution");
xlabel("k values");
ylabel("Probability values");
legend();
```

1.3 Ερώτημα Γ

Είναι γνωστό πως για μια τυχαία μεταβλητή η οποία ακολουθεί κατανομή Poisson η μέση τιμή και η διακύμανση της θα είναι ίση με την παράμετρο λ της κατανομής. Θα το αποδείξουμε παρακάτω.

Η μέση τιμή έχει ως εξής:

$$\begin{aligned} \mu = E[X] = \sum_{x=0}^{\infty} xP(x) = \\ \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^x}{x!} = \\ e^{-\lambda}\sum_{x=0}^{\infty} x\frac{\lambda^x}{x!} = \\ e^{-\lambda}\lambda e^{\lambda} = \\ \lambda \end{aligned} \tag{2}$$

Γνωρίζουμε αρχικά ότι ισχύει:

$$\sigma^2 = E[X^2] - E[X]^2 \tag{3}$$

και ότι η ροπή δεύτερης τάξης είναι ίση με:

$$E[X^2] = E[X(X-1)] + E[X] \tag{4}$$

Αν υπολογίσουμε την ροπή δεύτερης τάξης της κατανομής Poisson θα έχουμε:

$$\begin{aligned} E[X(X-1)] = &\sum_{x=0}^{\infty} k(k-1)P(k) = \\ &\sum_{x=0}^{\infty} k(k-1)\frac{e^{-\lambda}\lambda^k}{k!} = \\ &e^{-\lambda}\sum_{x=0}^{\infty} k(k-1)\frac{\lambda^k}{(k)!} = \\ &\lambda^2 e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^{k-2}}{(k-2)!} = \\ &\lambda^2 e^{-\lambda} e^{\lambda} = \lambda^2 \end{aligned} \tag{5}$$

Άρα προκύπτει:

$$E[X^2] = E[X(X-1)] + E[X] = \lambda^2 + \lambda \tag{6}$$

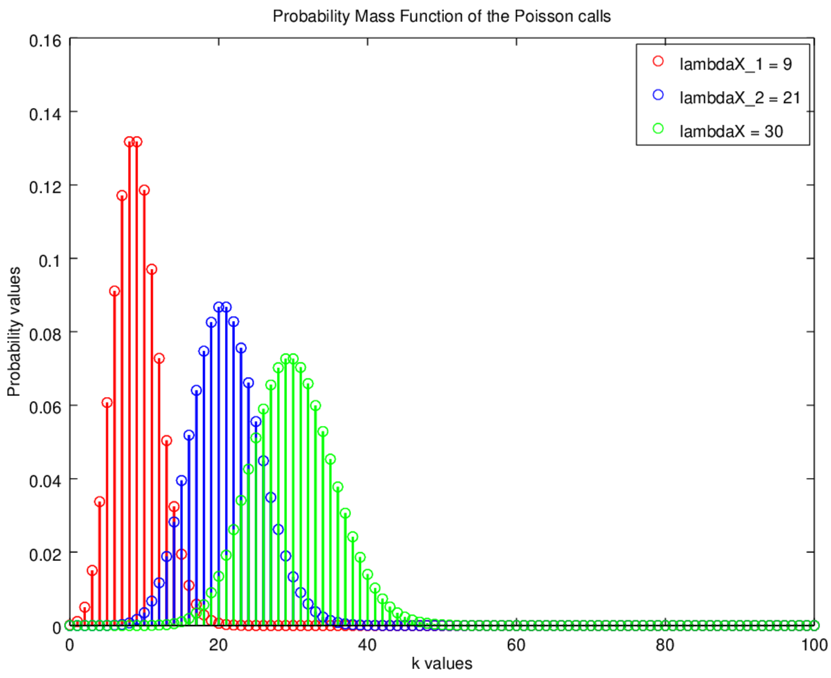
Και επομένως:

$$\sigma^2 = E[X^2] - E[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda \tag{7}$$

Άρα η διακύμανση της κατανομής Poisson είναι ίση με την παράμετρο λ της κατανομής και για το παρόν παράδειγμα η διακύμανση της κατανομής Poisson είναι ίση με 40.

1.4 Ερώτημα Δ

α) Είναι σαφές πως η κατανομή Poisson με $\lambda = 30$ είναι η υπέρθεση των δύο άλλων κατανομών Poisson με $\lambda_1 = 9$ και $\lambda_2 = 21$ όπου $\lambda = \lambda_1 + \lambda_2$. Η Poisson αυτή μπορεί να διασπαστεί σε δύο διαφορετικές κατανομές Poisson με διαφορετικές παραμέτρους λ_1 και λ_2 όπου $\lambda = \lambda_1 + \lambda_2$ και $p_{ext} = 0.3$ και $q_{int} = 0.7$ αντίστοιχα αφού $\lambda_1 = \lambda p_{ext}$ και $\lambda_2 = \lambda q_{int}$. Το διάγραμμα φαίνεται παρακάτω:



β) Η πιθανότητα αυτή είναι ίση με 0.09 και μπορεί να υπολογιστεί με την χρήση του τύπου της διωνυμικής κατανομής όπως φαίνεται παρακάτω:

$$P(X = k) = \binom{n}{k} p^k q^{n-k} = \binom{2}{2} 0.3^2 0.7^{2-2} = 0.09 \tag{8}$$

γ) Η πιθανότητα αυτή είναι ίση με 0.3087 και μπορεί να υπολογιστεί με την χρήση του τύπου της διωνυμικής κατανομής όπως φαίνεται παρακάτω:

$$P(X = k) = \binom{n}{k} p^k q^{n-k} = \binom{5}{2} 0.3^2 0.7^{5-2} = 0.3087 \tag{9}$$

Ο κώδικας που χρησιμοποιήθηκε φαίνεται παρακάτω

```
# TASK_D: Poisson decomposition/superposition.

# a) Plot in a common diagram the 3 distributions

# Define the lambda parameter
lambdaX = 30;

# Define the Poisson PMF values regarding the number of calls
PMF_X = poisspdf(k,lambdaX);

# Define the probabilities regarding the calls (external, internal)
p_ext = 0.3;
q_int = 1-p_ext;

# Thus, each type of call-event will follow a Poisson distribution

# Define the lambda parameters:
lambdaX_1 = lambdaX*p_ext;
lambdaX_2 = lambdaX*q_int;

# Define the PMF values for each type of call
PMF_X1 = poisspdf(k,lambdaX_1);
PMF_X2 = poisspdf(k,lambdaX_2);

colors = "rbg";
figure(3);
hold on;

# Plotting the Poisson calls
stem(k,PMF_X1,colors(1),"linewidth",1.2);
stem(k,PMF_X2,colors(2),"linewidth",1.2);
stem(k,PMF_X,colors(3),"linewidth",1.2);
hold off;
title("Probability Mass Function of the Poisson calls");
xlabel("k values");
ylabel("Probability values");
# legend for each case
legend1 = sprintf('lambdaX_1 = %d', lambdaX_1);
legend2 = sprintf('lambdaX_2 = %d', lambdaX_2);
legend3 = sprintf('lambdaX = %d', lambdaX);
legend(legend1,legend2,legend3)

# Calculation of probabilities

# using the number of combinations a.k.a. "binomial coefficient"
# (or directly using the binopdf from the statistics pkg)

# the probability (given n and k) can be found in either of the following ways:

# b)
# Define the number of n trials (number of calls made)
number_of_calls_b = 2;
# Define the k "successes", that, in our case, is
# the number of external calls with probability p_ext
k_ext_b = 2;
Prob_b = (nchoosek(number_of_calls_b,k_ext_b))*(p_ext^k_ext_b)*(q_int^(number_of_calls_b-k_ext_b));
# or directly with the use of binopdf (requires the statistics pkg)
Prob_b_bino = binopdf(k_ext_b,number_of_calls_b,p_ext);
display("The probability that both of the next two calls are external is");
display(Prob_b);

# c)
number_of_calls_c = 5;
k_ext_c = 2;
Prob_c = (nchoosek(number_of_calls_c,k_ext_c))*(p_ext^k_ext_c)*(q_int^(number_of_calls_c-k_ext_c));
# or, with the use of binopdf,
Prob_c_bino = binopdf(k_ext_c,number_of_calls_c,p_ext);
display("The probability that out of the 5 calls, two calls exactly are external, is");
display(Prob_c);
```

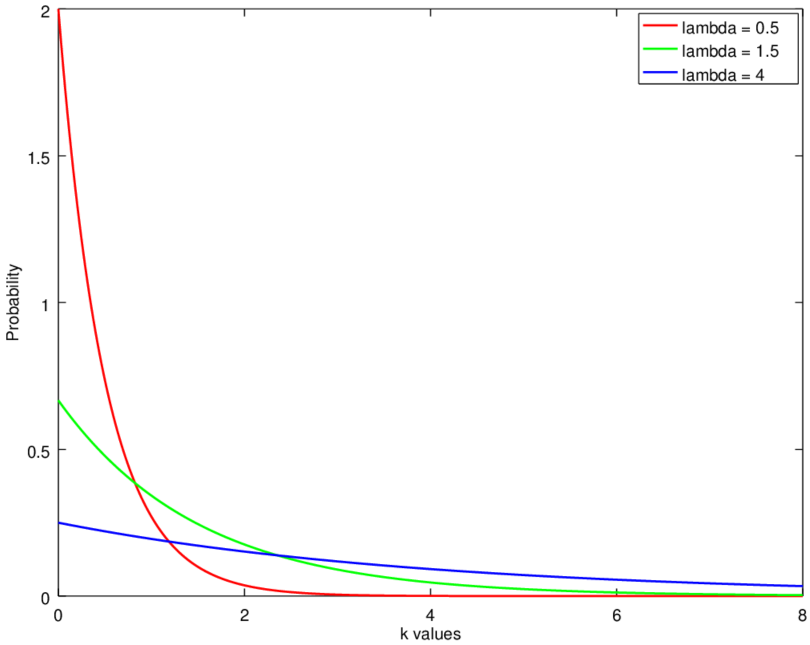
2 Εκθετικη κατανομη

2.1 Ερώτημα Α

Η εκθετική κατανομή ορίζεται ως εξής:

$$P(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \tag{10}$$

Το διάγραμμα παρουσίασης της εκθετικής κατανομής για τις διαφορετικές τιμές του λ φαίνεται στο παρακάτω διάγραμμα:



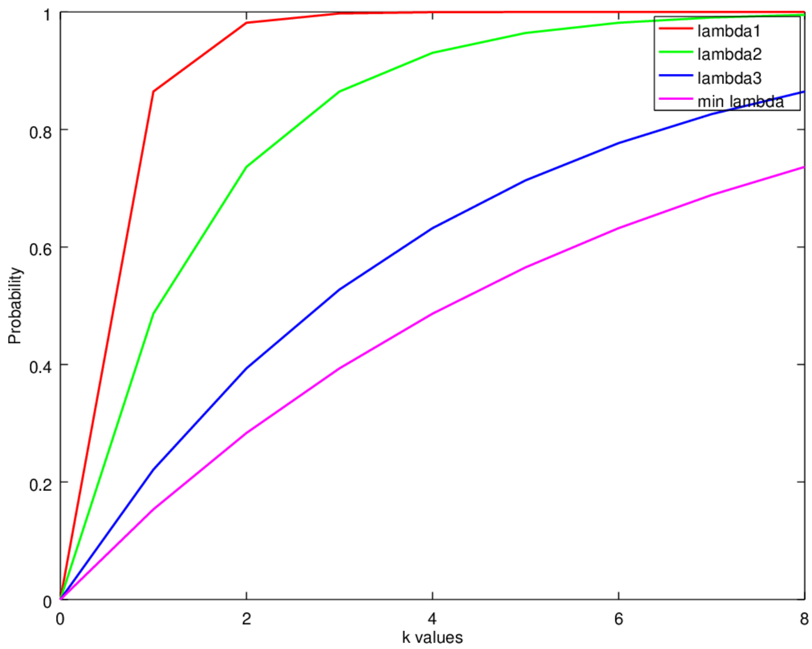
Ο κώδικας που χρησιμοποιήθηκε για την παραγωγή του διαγράμματος είναι ο εξής:

```
lambda = [0.5, 1.5, 4];

x = 0:0.001:8;

pdf(1,:) = exppdf(x,lambda(1));
pdf(2,:) = exppdf(x,lambda(2));
pdf(3,:) = exppdf(x,lambda(3));

figure(1);

hold on;

plot(x,pdf(1,:), 'r', 'LineWidth', 1.2);
plot(x,pdf(2,:), 'g', 'LineWidth', 1.2);
plot(x,pdf(3,:), 'b', 'LineWidth', 1.2);

hold off;

xlabel('k values');
ylabel('Probability');
legend('lambda = 0.5', 'lambda = 1.5', 'lambda = 4');
```

2.2 Ερώτημα Β

Γνωρίζουμε ότι $F(x) = P(X \leq x) = 1 - e^{-\frac{x}{\lambda}}$. Επίσης γνωρίζουμε ότι η κατανομή της τυχαίας μεταβλητής $X = min(X_1, X_2, X_3)$ είναι η εκθετική κατανομή με παράμετρο $\lambda = \lambda_1 + \lambda_2 + \lambda_3$.

Η γραφική παράσταση των συναρτήσεων σε ίδιο διάγραμμα φαίνονται παρακάτω:



Ο κώδικας που χρησιμοποιήθηκε για την παραγωγή του διαγράμματος είναι ο εξής:

```
lambdas = [0.5, 1.5, 4];

min_lambda = 0.5 + 1.5 + 4;

x = 0:00001:8;

for i=1:columns(lambdas)
        cdf(i,:) = expcdf(x,lambdas(i));
endfor

cdf_min = expcdf(x, min_lambda);

colors = 'rgbm';

figure(1);

hold on;

for i=1:columns(lambdas)
        plot(x, cdf(i,:), colors(i), 'LineWidth', 1.2);
endfor

plot(x, cdf_min, colors(4), 'LineWidth', 1.2);

hold off;

xlabel('k-values');
ylabel('Probability');
legend('lambda1', 'lambda2', 'lambda3', 'min-lambda');
```

Αυτό που παρατηρούμε σχετικά με τη μέση τιμή των εκθετικών κατανομών είναι ότι όσο μεγαλύτερη είναι η τιμή της παραμέτρου λ τόσο μικρότερη είναι η μέση τιμή της κατανομής.

2.3 Ερώτημα Γ

Από τον ορισμό της δεσμευμένης πιθανότητας έχουμε:

$$
\begin{aligned}
P(X \geq x+t | X \geq t) = & \frac{P(X \geq x+t \cap X \geq t)}{P(X \geq t)} = \\
& \frac{P(X \geq x+t)}{P(X \geq t)} = \\
& \frac{1 - F(x+t)}{1 - F(t)} = \\
& \frac{1 - (1 - e^{-\frac{x+t}{\lambda}})}{1 - (1 - e^{-\frac{t}{\lambda}})} = \\
& \frac{e^{-\frac{x+t}{\lambda}}}{e^{-\frac{t}{\lambda}}} = e^{-\frac{x}{\lambda}} = 1 - F(x) = P(X \geq x)
\end{aligned} \tag{11}
$$

Επομένως εύκολα προκύπτει ότι:

$$
P(X > 45000 | X > 25000) = P(X > 20000 + 25000 | X > 25000) = P(X > 20000) \tag{12}
$$

Επομένως οι δύο πιθανότητες θα πρέπει να είναι ίσες όπως και επιβεβαιώνεται από τον κώδικα που παρατίθεται παρακάτω:

```
x = 0:0.001:8;

exp = expcdf(x,2.5);

p_1 = 1 - exp(20000);
display(p_1);

p_2 = (1 - exp(45000))./(1 - exp(25000));
display(p_2);
```

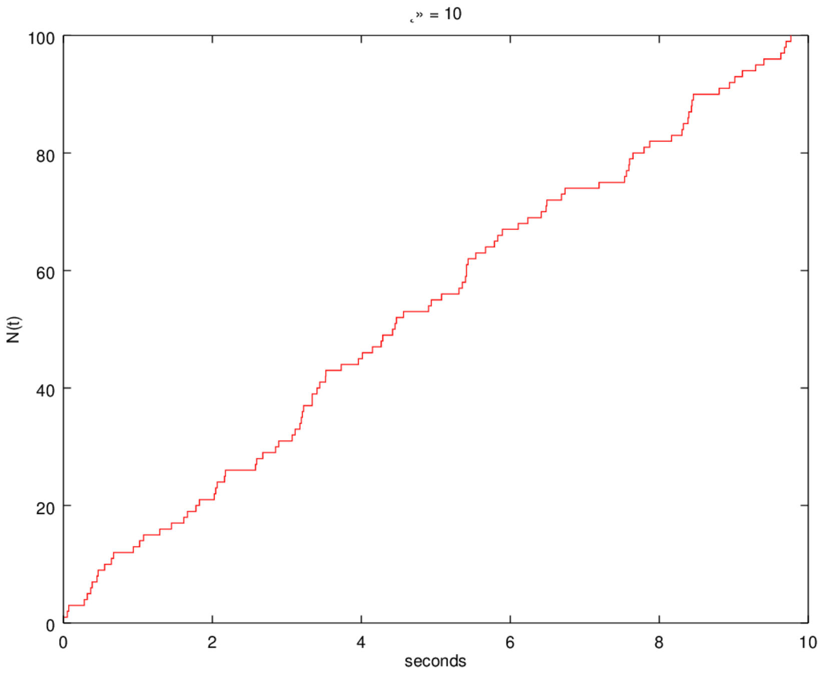
Οι πιθανότητες που υπολογίστηκαν είναι ίσες και ίσες με 0.92312.

3 Άσκηση 3

3.1 Ερώτημα Α

Η κατανομή που ακολουθείτε από τους χρόνους ανάμεσα στην εμφάνιση δύο διαδοχικών γεγονότων είναι μία εκθετική κατανομή με μέση τιμή $\frac{1}{\lambda}$.

Το διάγραμμα φαίνεται παρακάτω:



Το κώδικας που χρησιμοποιήθηκε για την παραγωγή του διαγράμματος είναι ο εξής:

```
x = exprnd(0.1, 1,100);
y = ones(100,1);

 for i=1:99
  x(i+1) = x(i+1) + x(i);
  y(i+1) = y(i+1) + y(i);
endfor


figure(1);

stairs(x, y, color='r');
xlabel('seconds');
ylabel('N(t)');
title('  -=- 10');
```

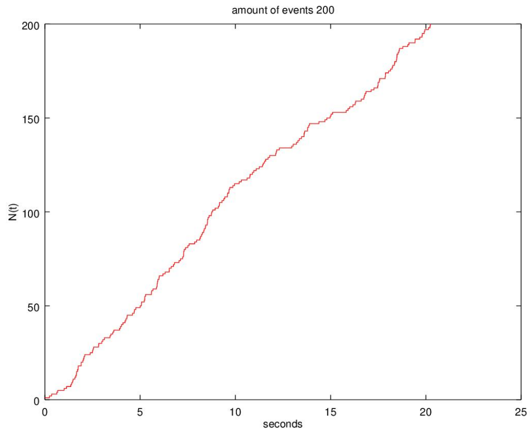
3.2 Ερώτημα Β

Σε ένα τέτοιο χρονικό παράθυρο ακολουθείτε κατανομή Poisson με μέσο αριθμό γεγονότων $\lambda \cdot \Delta T$. Η παράμετρος λ είναι προσεγγίσιμη από το πηλίκο του αριθμού των γεγονότων με το χρονικό παράθυρο ΔT.

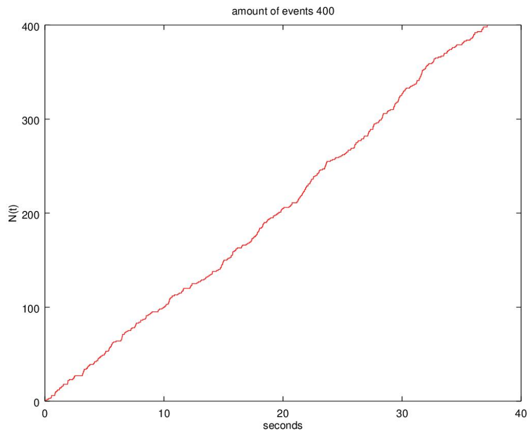
Ο μέσος ρυθμός για το προηγούμενο ερώτημα είναι ίσος με $100/exprnd(100) = 11.813$.

Για διαφορετικές τιμές αριθμού τυχαίων γεγονότων έχουμε:

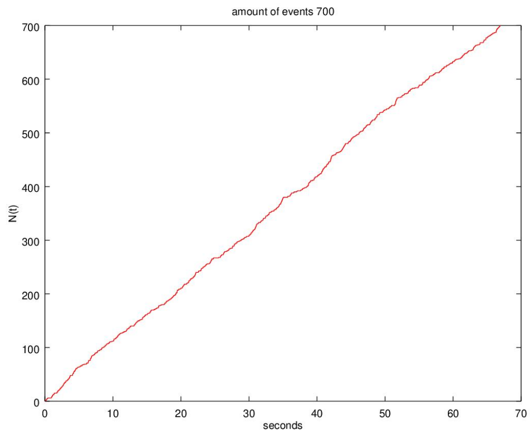
1. Για $N = 200$ έχουμε $\lambda = 9.6447$



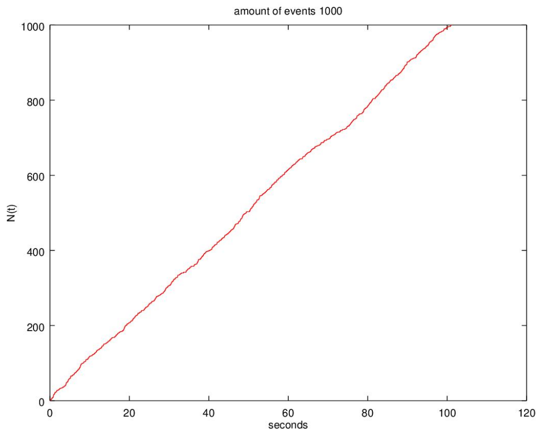
2. Για $N = 400$ έχουμε $\lambda = 9.6523$



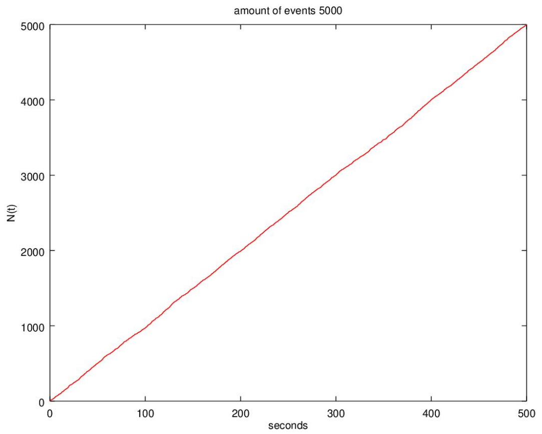
3. Για $N = 700$ έχουμε $\lambda = 9.9795$



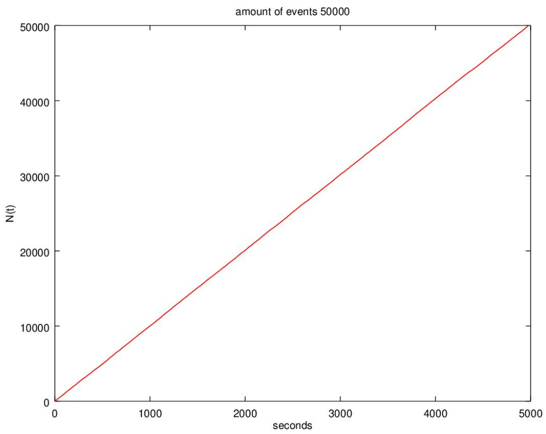
4. Για $N = 1000$ έχουμε $\lambda = 10.127$



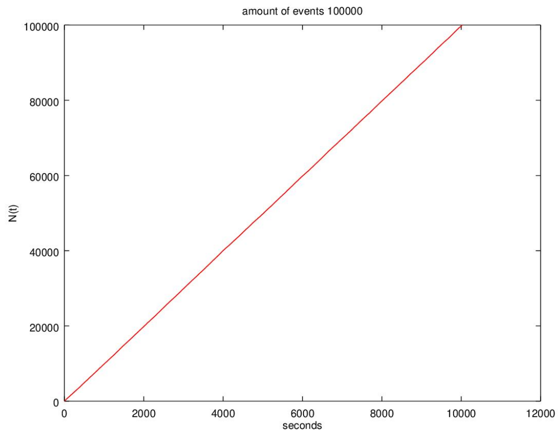
5. Για N = 5000 έχουμε λ = 10.133



6. Για N = 50000 έχουμε λ = 10.018



7. Για $N = 100000$ έχουμε $\lambda = 9.9654$



Ο κώδικας φαίνεται παρακάτω:

```
ev = [200, 400, 700, 1000, 5000, 50000, 100000];

for j=1:columns(ev)
        x = exprnd(0.1, 1,ev(j));
        y = ones(ev(j),1);
        k = ev(j) - 1;
        for i=1:k
                x(i+1) = x(i+1) + x(i);
                y(i+1) = y(i+1) + y(i);
        endfor


        figure(1);

        titlename = sprintf('amount of events %d', ev(j));

        stairs(x, y, color='r');
        xlabel('seconds');
        ylabel('N(t)');
        title(titlename);

        filename = sprintf('../images/events_%d.jpg', ev(j));
        print('-djpg',filename);

        solution = ev(j) / x(ev(j));
        display(ev(j));
        display(solution);


endfor
```

Από τα διαγράμματα φαίνεται ότι με την αύξηση των γεγονότων ο μέσος ρυθμός των γεγονότων ταλαντώνεται γύρω από την τιμή του $\lambda = 10$ πλησιάζοντας διαρκώς όλο και πιο κοντά. Η ευκρίνεια στα διαγράμματα επίσης αυξάνεται με την αύξηση των γεγονότων.