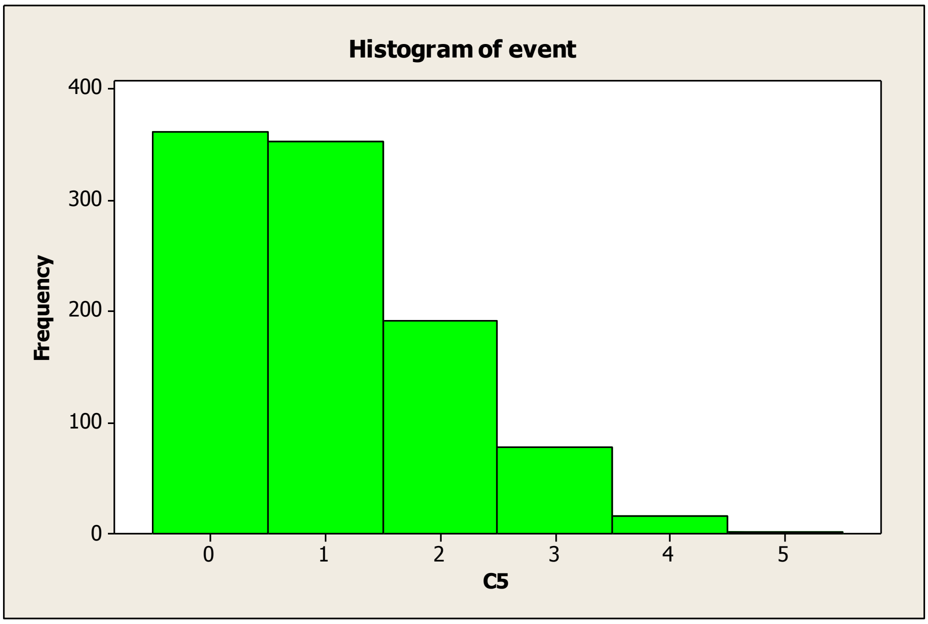
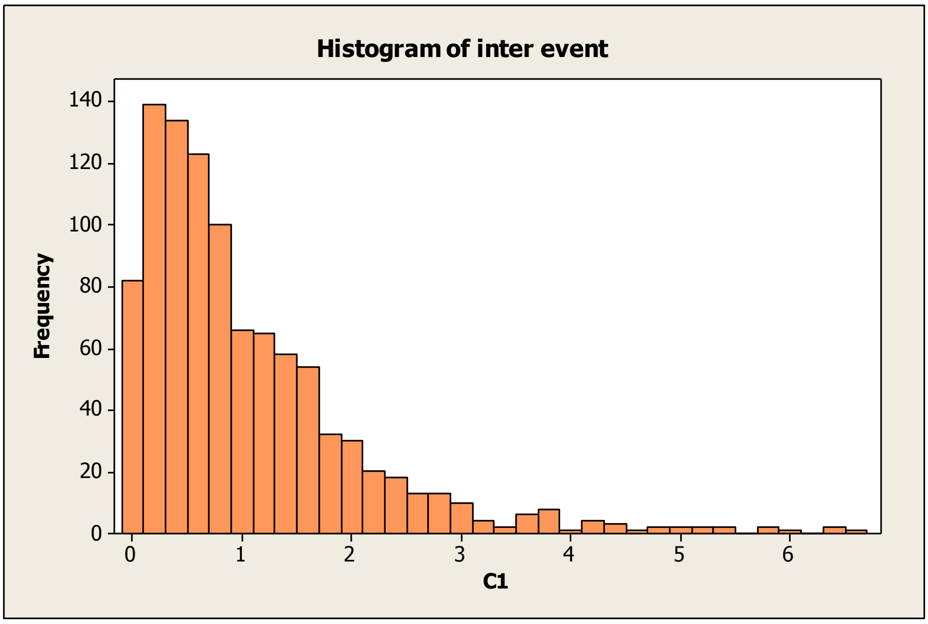
Exponential Distribution

Time or (length)	Inter arrival(event)	Number of Event
1.3923	1.39230	0
1.7123	0.31996	3
1.9331	0.22082	1
2.7428	0.80968	0
4.2009	1.45818	1
5.4410	1.24008	1
6.2672	0.82621	1
7.2707	1.00344	1
9.4782	2.20758	0
9.5625	0.08431	2
10.4470	0.88446	2
10.7162	0.26922	0
12.2204	1.50417	1
13.1075	0.88710	2
13.7518	0.64434	2
14.2307	0.47889	1
14.7596	0.52888	2
15.2832	0.52360	0
16.1347	0.85150	0
16.3408	0.20612	0
21.6632	5.32243	0
23.3079	1.64462	1
23.5427	0.23484	0
25.0680	1.52535	2
25.4869	0.41889	0
		2

Descriptive Statistics: event; inter event

Variable	N	Mean	StDev
event	26	1.000	0.938
inter event	25	1.019	1.050





Descriptive Statistics: C1; C5

Variable	N	Mean	StDev
event	1000	1.0370	1.0103
inter event	1000	1.0405	1.0155

The discussion of the Poisson distribution defined a random variable to be the **number of flaws** along a length of copper wire. The **distance between flaws is another** random variable that is often of interest.

- **Let the random variable X denote the length from any starting point on the wire until a flaw is detected.**
- **Let the random variable N denote the number of flaws in x millimeters of wire. If the mean number of flaws is λ per millimeter, N has a Poisson distribution with mean λx.**

We assume that the wire is longer than the value of x. Now

$$P(X > x) = P(N = 0) = \frac{e^{-\lambda x}(\lambda x)^0}{0!} = e^{-\lambda x}$$

Therefore

$$F(x) = P(X \leq x) = 1 - e^{-\lambda x}, \quad x \geq 0$$

is the cumulative distribution function of X. By differentiating F(x), the probability density function of X is calculated to be

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

The derivation of the distribution of X depends only on the assumption that the flaws in the wire follow a **Poisson process**. Also the starting point for measuring X doesn't matter because the probability of the number of flaws in an interval of a **Poisson process depends only on the length of the interval, not on the location**. For any Poisson process, the following general result applies.

Exponential Distribution:

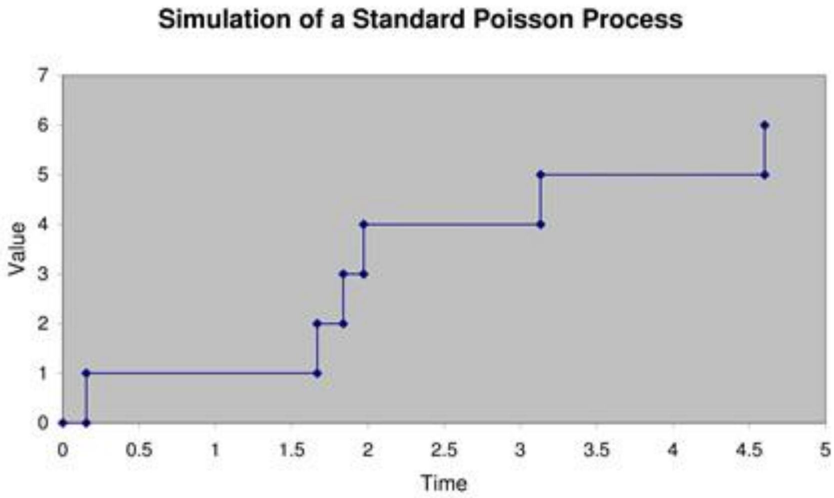
The random variable X that equals the distance between successive events of a Poisson process with mean $\lambda>0$ is an exponential random variable with parameter λ. The probability density function of X is

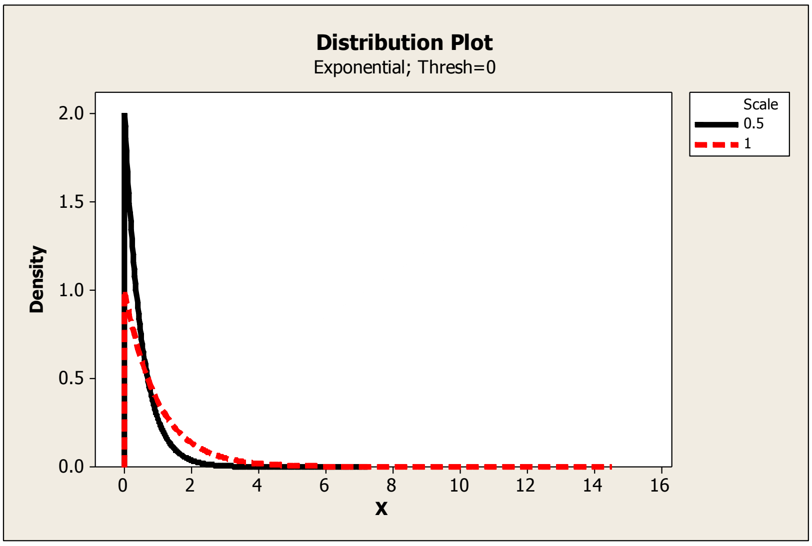
$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

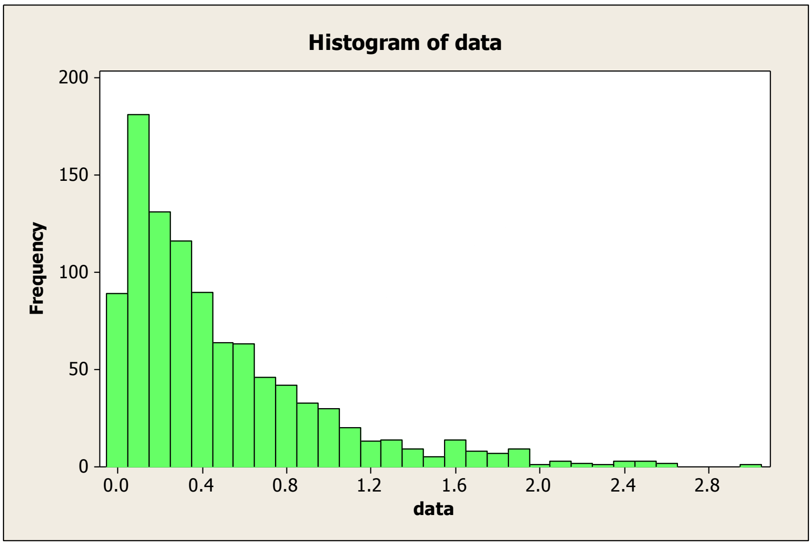
The exponential distribution obtains its name from the exponential function in the probability density function.

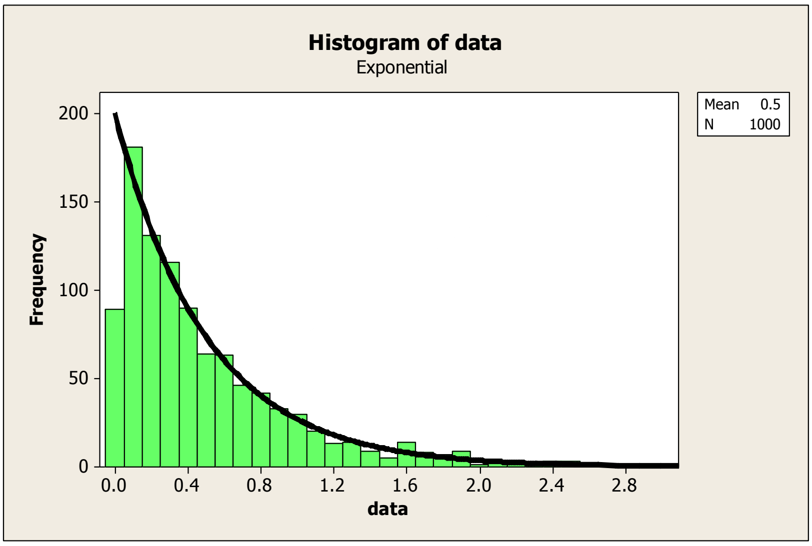
What is a Poisson Process?

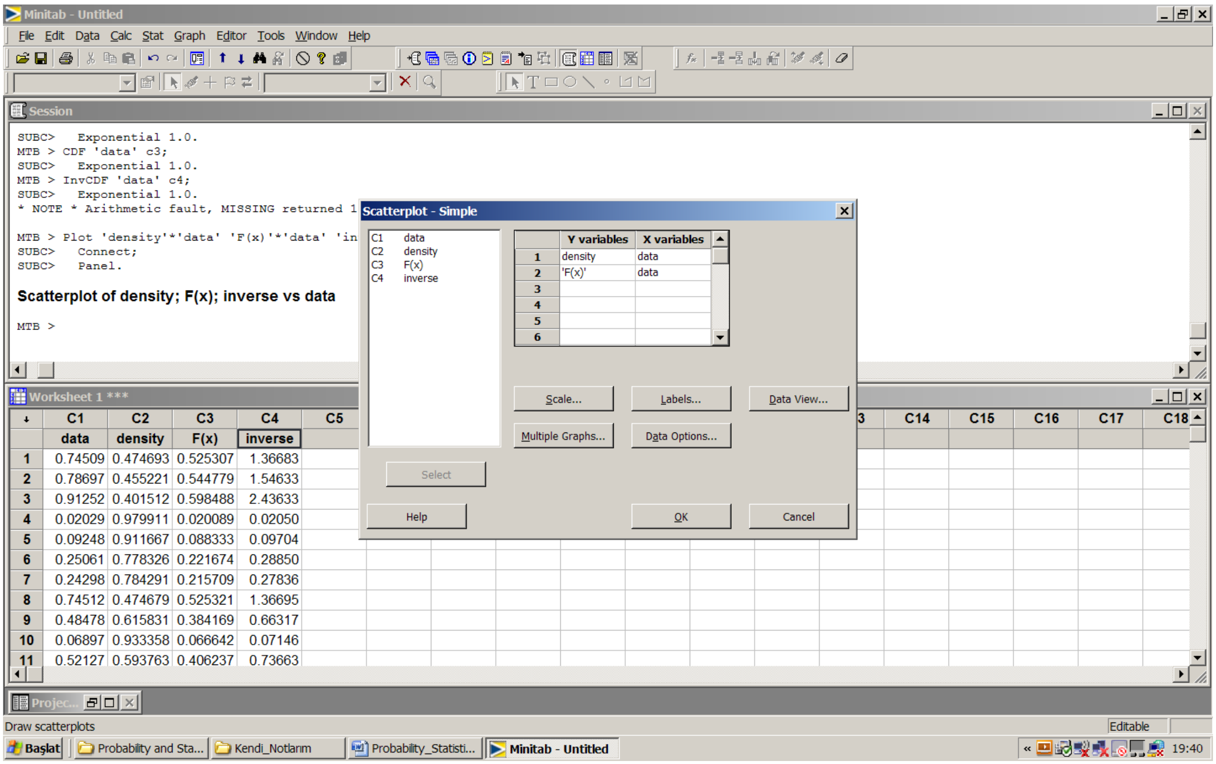
A Poisson process is a pure jump process: a process that changes instantaneously from one value to another at random times. The following is a simulation of a standard Poisson process (where the jump sizes are restricted to 1).

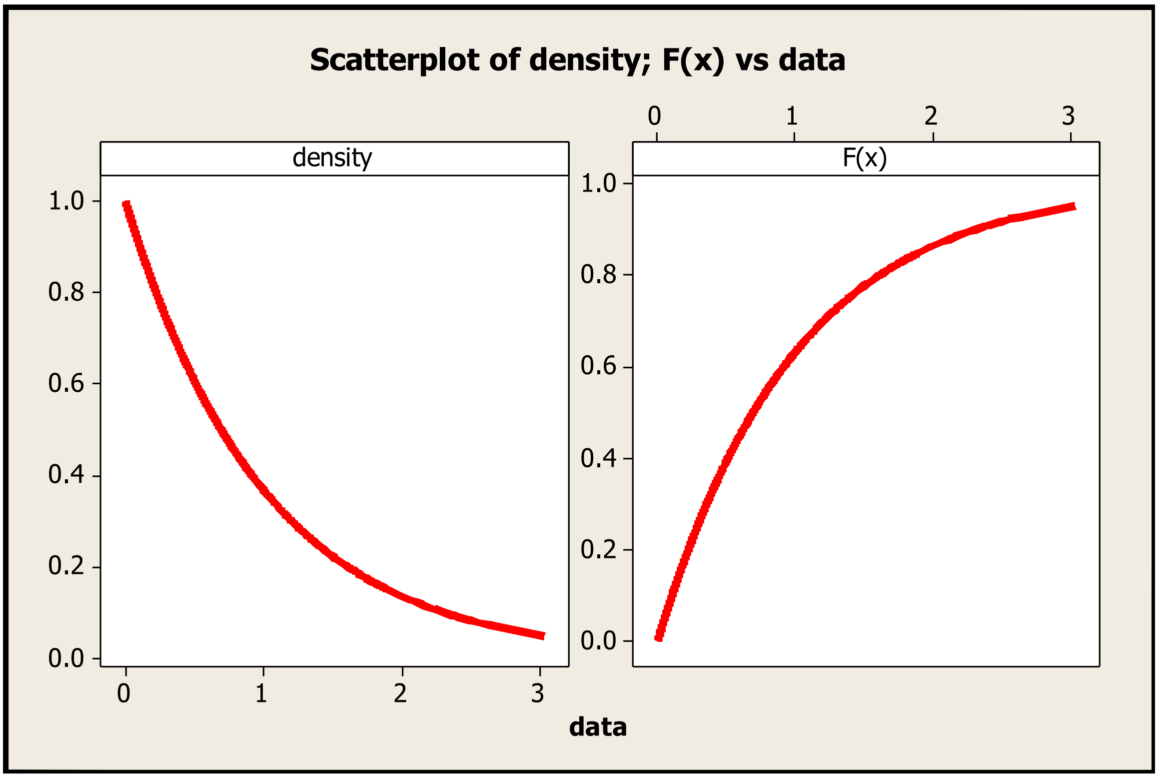



Distribution Plot
Exponential; Thresh=0
Scale
0.5
1
Density
2.0
1.5
1.0
0.5
0.0
0
2
4
6
8
10
12
14
16
X


Histogram of data
Frequency
200
150
100
50
0
0.0
0.4
0.8
1.2
1.6
2.0
2.4
2.8
data


Histogram of data
Exponential
Mean 0.5
N 1000
Frequency
200
150
100
50
0
0.0
0.4
0.8
1.2
1.6
2.0
2.4
2.8
data


Minitab - Untitled
Session
MTB > CDF 'data' c3;
SUBC> Exponential 1.0.
MTB > InvCDF 'data' c4;
SUBC> Exponential 1.0.
Scatterplot of density; F(x); inverse vs data
Scatterplot - Simple
Y variables
X variables
density
data
'F(x)'
data
Worksheet 1 ***
data
density
F(x)
inverse
Draw scatterplots
Minitab - Untitled


Scatterplot of density; F(x) vs data
density
F(x)
data

Mean and Variance:

If the random variable X has an exponential distribution with parameter λ,

$$E(X) = \frac{1}{\lambda} \quad \textit{and} \quad V(X) = \frac{1}{\lambda^2}$$

```
MTB > random 1000 c1;
SUBC> exponential 0.5.
```

???

Descriptive Statistics: data

```
Variable      N  N*    Mean  SE Mean   StDev  Minimum
Q1   Median      Q3
data       1000   0  0.4917   0.0153  0.4854   0.0003
0.1393  0.3345  0.6834

Variable  Maximum
data       3.0316
```

Negative Exponential Distribution

Alternative parameterization of the probability density function of exponential distribution is

$$f(x) = \frac{1}{\beta} e^{-x/\beta}, \quad x \geq 0$$

β>0 is the scale parameter of the distribution and is the reciprocal of the rate parameter λ.

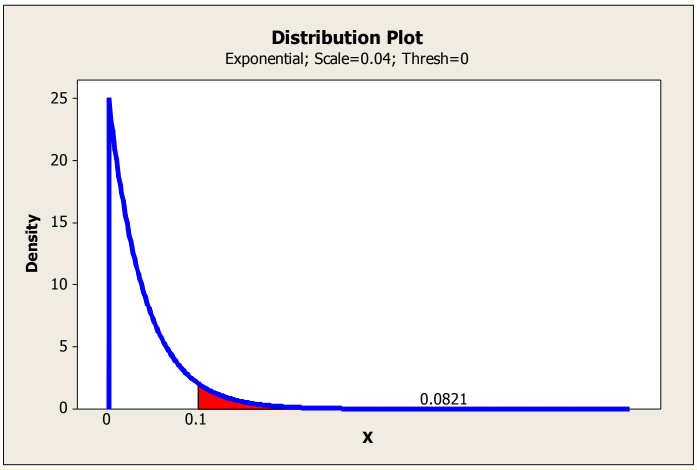
$$E(X) = \beta \quad V(X) = \beta^2$$

Example :Computer Usage

In a large corporate computer network, user log-ons to the system can be modeled as a Poisson process with a mean of 25 log-ons per hour.

i. What is the probability that there are no log-ons in an interval of 6 minutes?

Let X denote the time in hours from the start of the interval until the first log-on. Then, X has an exponential distribution with λ=25 log-ons per hour. We are interested in the probability that X exceeds 6 minutes. Because λ is given in log-ons per hour, we express all time units in hours. That is minutes=0.1 hour.



$$P(X > 0.1) = \int_{0.1}^{\infty} 25e^{-25x}dx = e^{-25(0.1)} = 0.082$$

Also, the cumulative distribution function can be used to obtain the same result as follows:

$$F(x) = P(X \leq x) = 1 - e^{-\lambda x}, \quad x \geq 0$$

$$P(X > 0.1) = 1 - F(0.1) = e^{-25(0.1)} = 0.082$$

With Poisson distribution

X has an exponential distribution with λ=25 log-ons per hour.

The mean of the logon events in 6 minutes:25/10=2.5

Probability of no event in 6 minutes

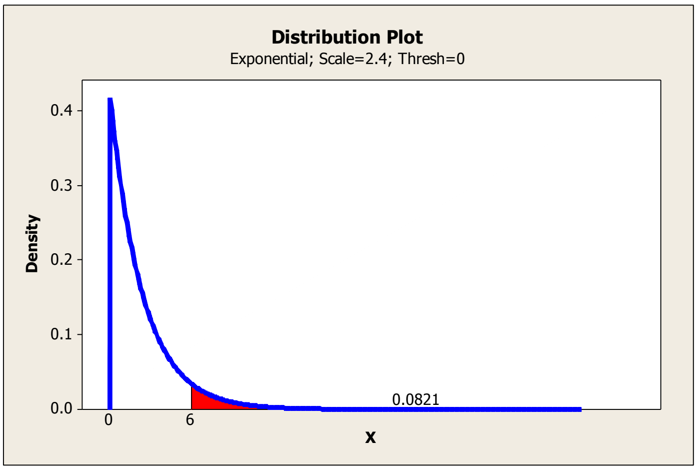
$$P(N = 0) = \frac{e^{-\lambda}\lambda^{x}}{x!} = \frac{e^{-2.5}(2.5)^{0}}{0!} = e^{-2.5} = 0.082$$

!!!!!

X discrete variable in Poisson distribution but

continuous variable in Exponential distribution.

An identical answer is obtained by expressing the mean number of log-ons as 25/60=0.417 <u>log-ons per minute</u> and computing the probability that the next log-on exceeds 6 minutes.



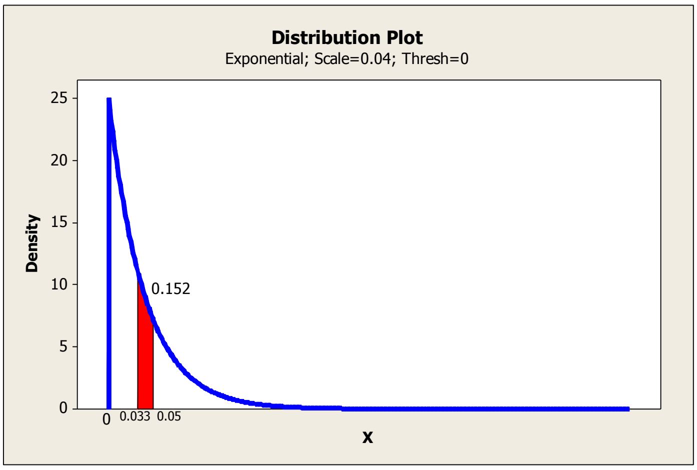
$$P(X>6)=\int_{6}^{\infty}0.417e^{-0.417x}dx=e^{-0.417(6)}=0.082$$

Or using cumulative distribution function

$$P(X>6)=1-F(6)=e^{-0.417(6)}=0.082$$

ii. What is the probability that the time until the next log-on is between 2 and 3 minutes?

Upon converting all units to hours,



$$P(0.033 < X > 0.05) = \int_{0.033}^{0.05} 25e^{-25x}dx = e^{-25x}\Big|_{0.033}^{0.05} = 0.152$$

Or computing as minutes

$$P(2 < X > 3) = \int_{2}^{3} 0.417e^{-0.417x}dx = e^{-0.417x}\Big|_{2}^{3} = 0.152$$

<u>It is important</u> to use consistent units in the calculation of probabilities, means and variances involving exponential random variables.

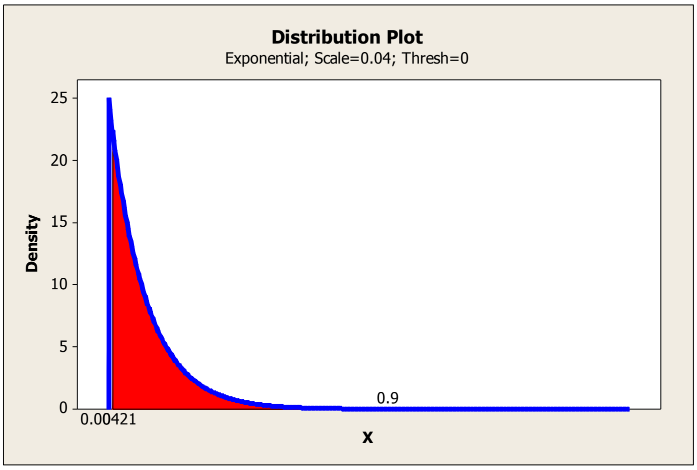
iii. Determine the interval of time such that the probability that no log-on occurs in the interval is 0.90.

$$P(X > x) = 0.90$$

$$P(X > x) = e^{-25x} = 0.90$$

$$-25x = \ln(0.90) = -0.1054$$

Therefore $x = 0.00421 \quad \text{hour} = 0.25 \quad \text{minute}$



Furthermore, the mean time until the next log-on is

$$\mu = 1/25 = 0.04 \, \text{hour} = 2.4 \quad \text{minutes}$$

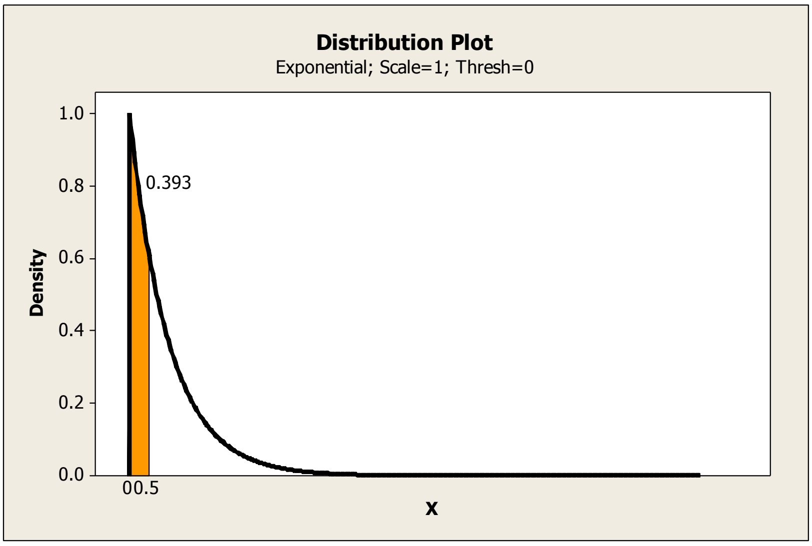
The standard deviation of the time until the next log-on is

$$\sigma = 1/25 = 0.04 \, \text{hour} = 2.4 \quad \text{minutes}$$

Lack of Memory Property

Let X denote the time between detections of particle with a Geiger counter and assume that X has an exponential distribution with E(X) =1 minute. The probability that we detect a particle within 30 seconds of starting the counter is
E(X)=1/λ
λ=1
30seconds=0.5 minute therefore

$$P(X < 0.5) = F(0.5) = 1 - e^{-0.5(1)} = 0.39345$$



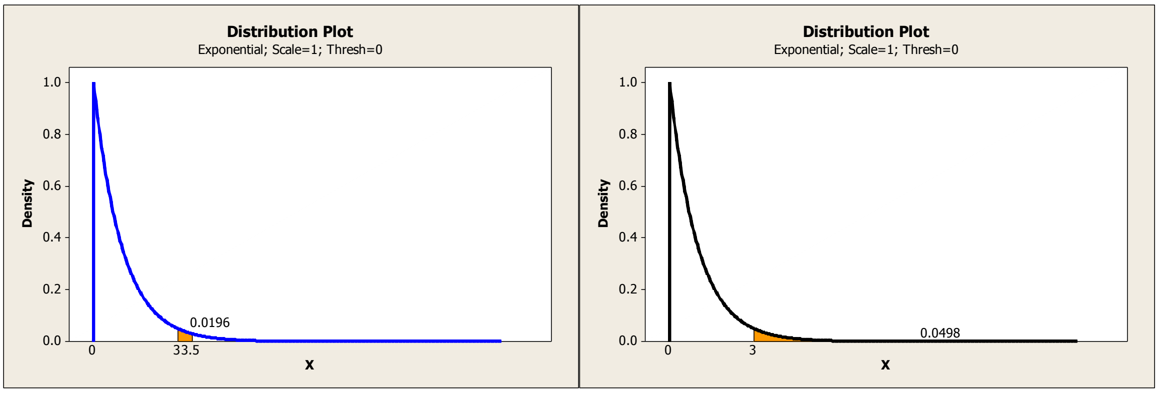
Now suppose we turn on the Geiger counter and wait 3 minutes without detecting a particle. What is the probability that a particle is detected in the next 30 seconds?

?

Because we have already been waiting for 3 minutes, we feel that we are "due". That is, the probability of a detection in the next 30 seconds should be greater than 0.3935. However, for an exponential distribution, this is not true.

From the definition of conditional probability,

$$P(X<3.5|X>3)=\frac{P(3.0<X<3.5)}{P(X>3)}$$



$$P(3<X<3.5)=F(3.5)-F(3)$$
$$=[1-e^{-3.5}]-[1-e^{-3}]=0.0196$$

And $P(X>3)=1-F(3)=e^{-3}=0.0498$

Therefore,

$$P(X<3.5|X>3)=\frac{0.0196}{0.0498}=0.39345$$

After waiting for 3 minutes without detection, the probability of detection in the next 30 seconds is the same as the probability of detection in the 30 seconds immediately after starting the counter.
The fact that we have waited 3 minutes without detection does not change the probability of detection in the next 30 seconds.

Now suppose we turn on the Geiger counter and wait 10 minutes <u>without detecting</u> a particle. What is the probability that a particle is detected in the next 30 seconds?

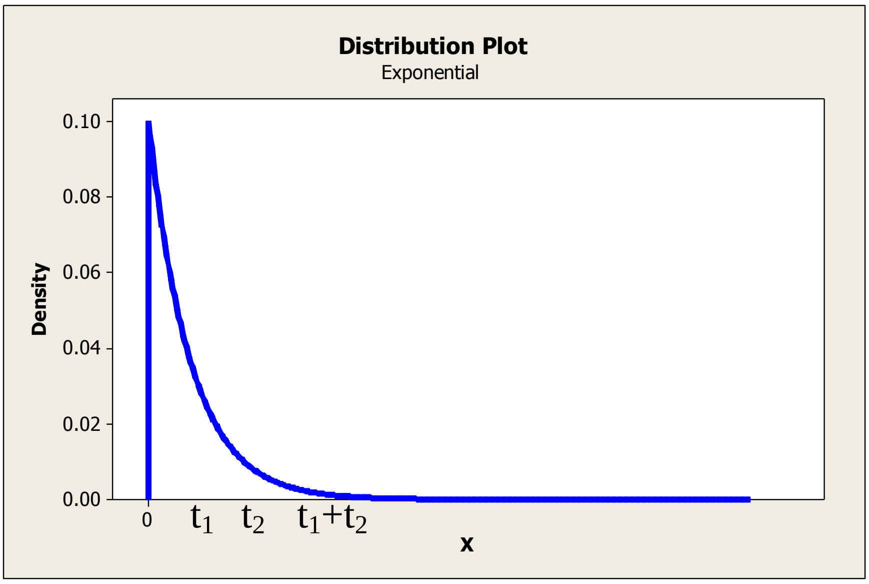
$$P(X < 10.5 | X > 10) = \frac{P(10.0 < X < 10.5)}{P(X > 10)}$$

$$P(10 < X < 10.5) = F(10.5) - F(10)$$
$$= [1 - e^{-10.5}] - [1 - e^{-10}] = 0.000017863$$

$$P(X > 10) = 1 - F(10) = e^{-10} = 0.00000454$$

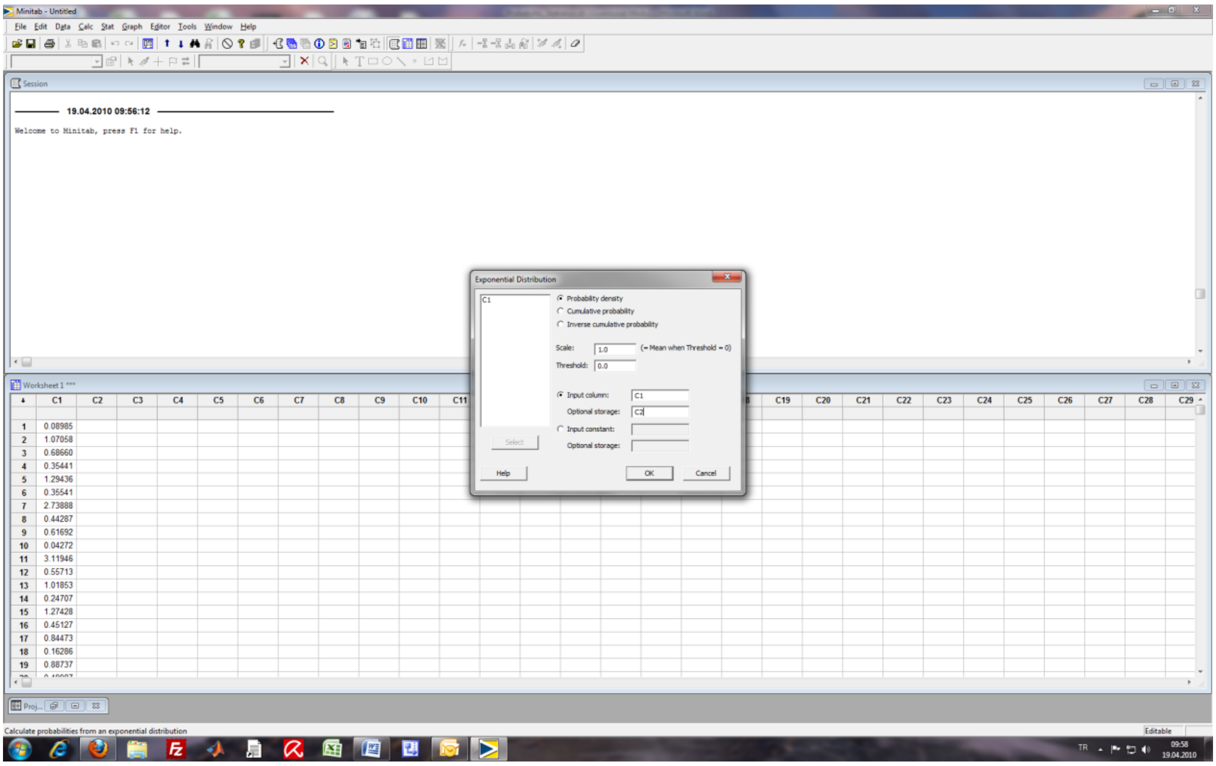
$$P(X < 10.5 | X > 10) = \frac{0.000017863}{0.0000454} = 0.39345$$

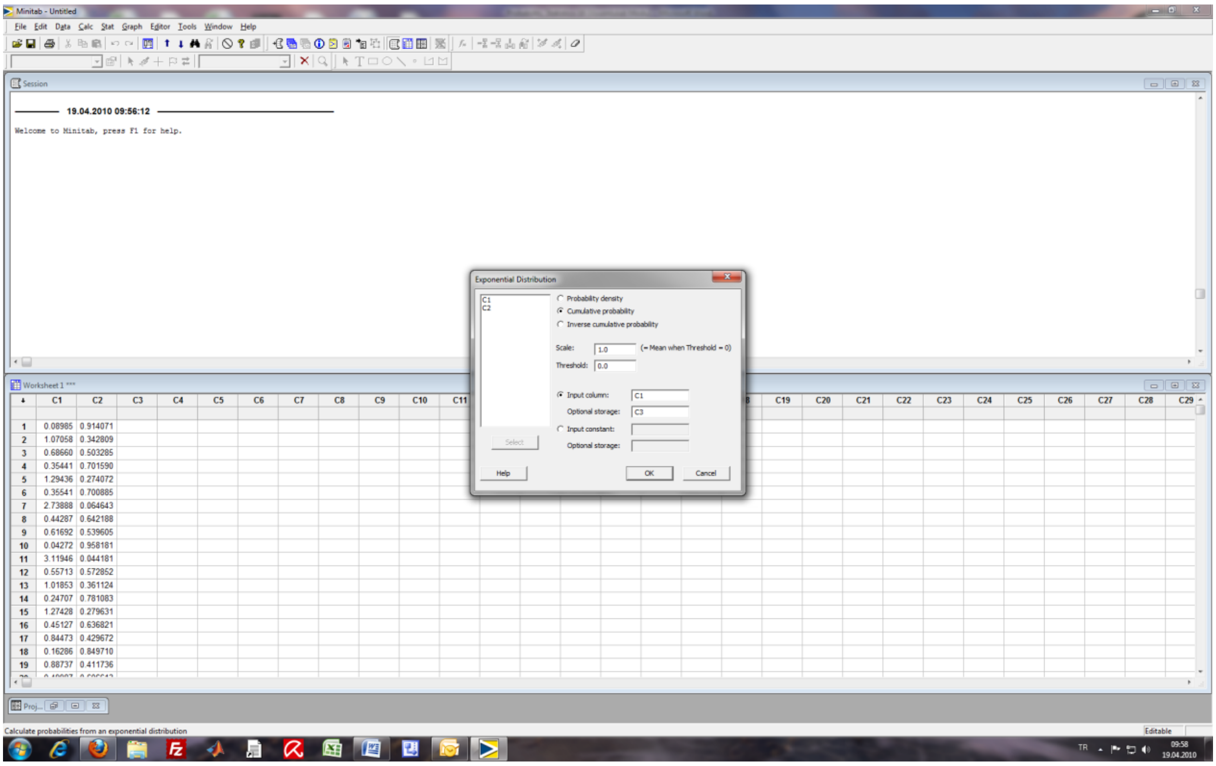
```
MTB > cdf 10 k1;
SUBC> expo 1.
MTB > cdf 10.5 k2;
SUBC> expo 1.
MTB > print k1 k2
```
Data Display
```
K1     0.999955
K2     0.999972
MTB > let k3=k2-k1
MTB > print k3
MTB > print k4
```
Data Display
```
K4     0.000045400
```
Data Display
```
K3     0.000017863
MTB > let k4=1-k1
MTB > let k5=k3/k4
MTB > print k5
```
Data Display
```
K5     0.393469
```

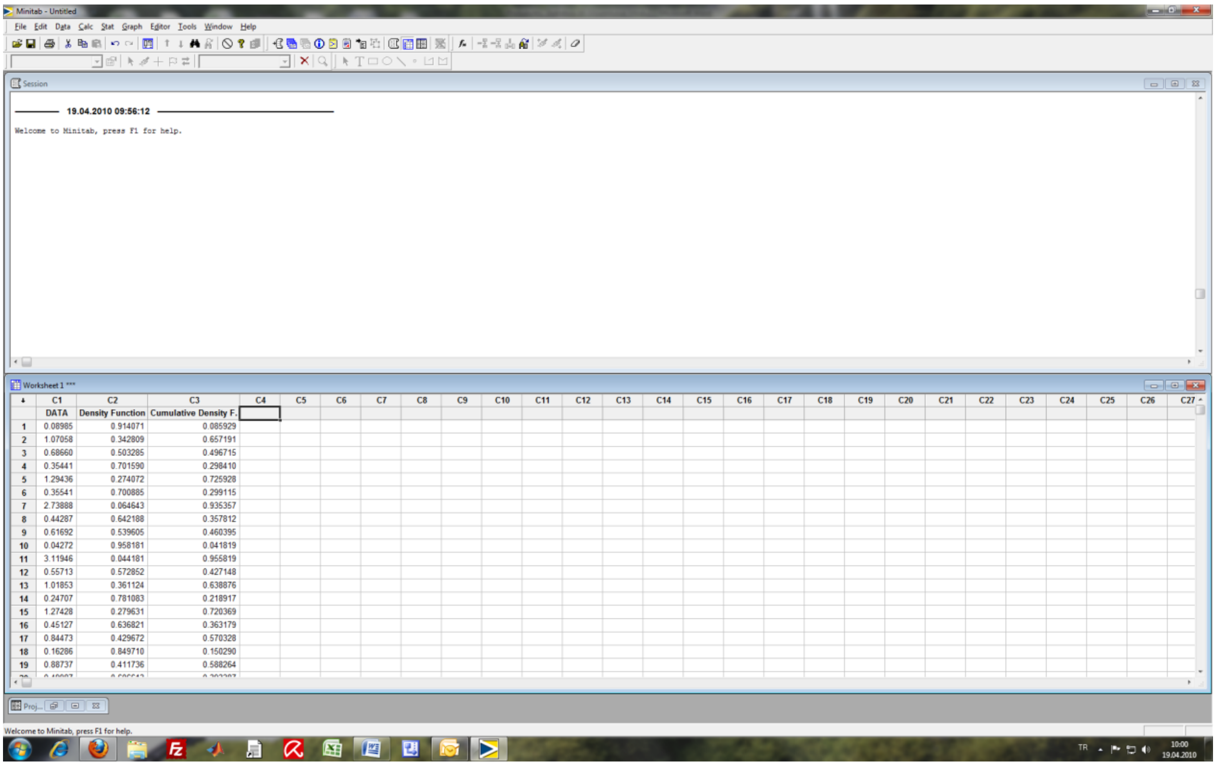


For an exponential random variable X,

$$P(X < t_1 + t_2 | X > t_2) = F(X < t_1)$$


Minitab - Untitled
File Edit Data Calc Stat Graph Editor Tools Window Help
Session
19.04.2010 09:56:12
Welcome to Minitab, press F1 for help.
Worksheet 1 ***
Exponential Distribution
C1
Probability density
Cumulative probability
Inverse cumulative probability
Scale: 1.0 (= Mean when Threshold = 0)
Threshold: 0.0
Input column: C1
Optional storage: C2
Input constant:
Optional storage:
Select
Help
OK
Cancel
C1
0.08985
1.07058
0.68660
0.35441
1.29436
0.35541
2.73888
0.44287
0.61692
0.04272
3.11946
0.55713
1.01853
0.24707
1.27428
0.45127
0.84473
0.16286
0.88737
Calculate probabilities from an exponential distribution
Editable


Minitab - Untitled
File Edit Data Calc Stat Graph Editor Tools Window Help
Session
19.04.2010 09:56:12
Welcome to Minitab, press F1 for help.
Worksheet 1 ***
Exponential Distribution
C1
C2
Probability density
Cumulative probability
Inverse cumulative probability
Scale: 1.0 (= Mean when Threshold = 0)
Threshold: 0.0
Input column: C1
Optional storage: C3
Input constant:
Optional storage:
Select
Help
OK
Cancel
C1 C2
0.08985 0.914071
1.07058 0.342809
0.68660 0.503285
0.35441 0.701590
1.29436 0.274072
0.35541 0.700885
2.73888 0.064643
0.44287 0.642188
0.61692 0.539605
0.04272 0.958181
3.11946 0.044181
0.55713 0.572852
1.01853 0.361124
0.24707 0.781083
1.27428 0.279631
0.45127 0.636821
0.84473 0.429672
0.16286 0.849710
0.88737 0.411736
Calculate probabilities from an exponential distribution
Editable



```
MTB > print c1-c3
```

Data Display

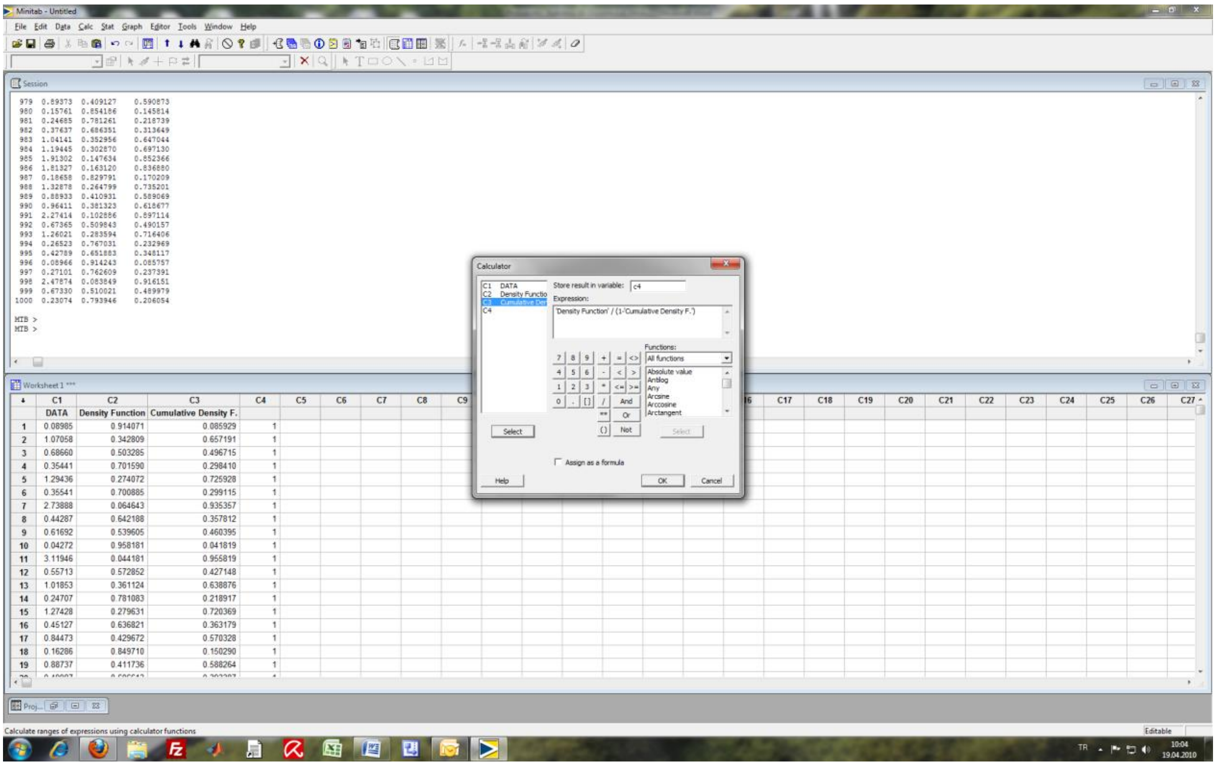
Row	DATA	Density Function	Cumulative Density F.
1	0.08985	0.914071	0.085929
2	1.07058	0.342809	0.657191
3	0.68660	0.503285	0.496715
4	0.35441	0.701590	0.298410
5	1.29436	0.274072	0.725928
6	0.35541	0.700885	0.299115
7	2.73888	0.064643	0.935357
8	0.44287	0.642188	0.357812
9	0.61692	0.539605	0.460395
10	0.04272	0.958181	0.041819
11	3.11946	0.044181	0.955819
12	0.55713	0.572852	0.427148
13	1.01853	0.361124	0.638876
14	0.24707	0.781083	0.218917
15	1.27428	0.279631	0.720369
16	0.45127	0.636821	0.363179
17	0.84473	0.429672	0.570328
18	0.16286	0.849710	0.150290
19	0.88737	0.411736	0.588264
20	0.49987	0.606613	0.393387

```
Let Hazard RAte = 'Density Function'
/ (1-'Cumulative Density F.')
```

```
MTB > print c1-c4
```

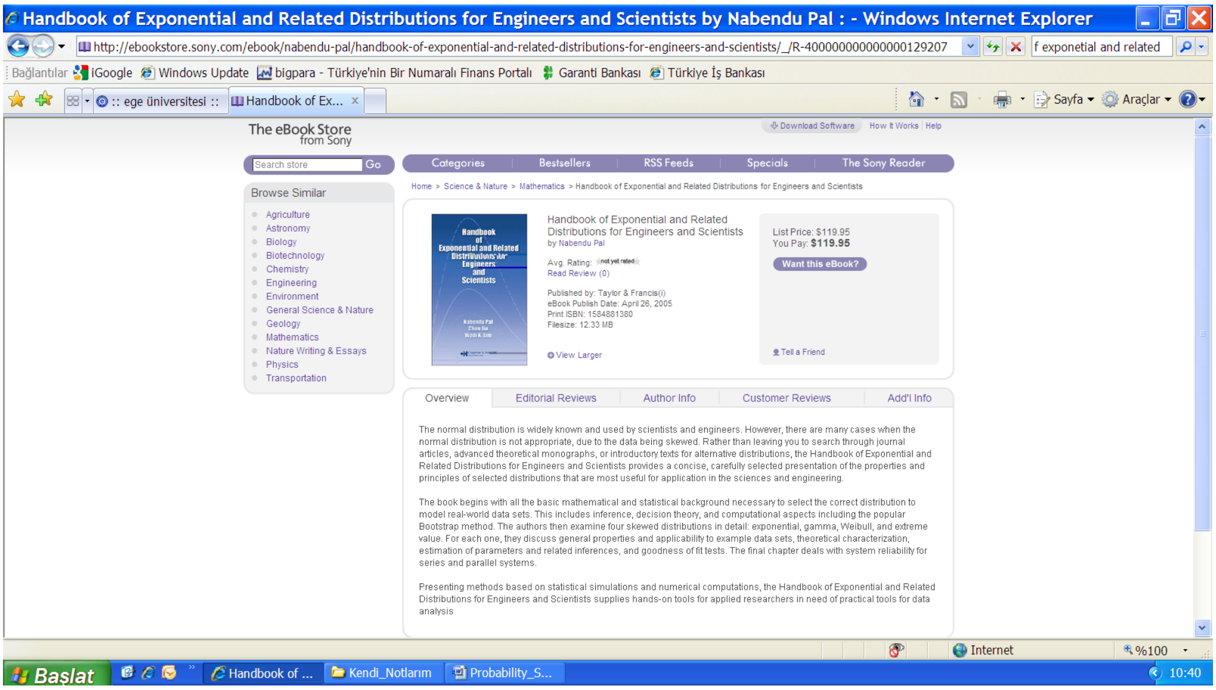
Data Display

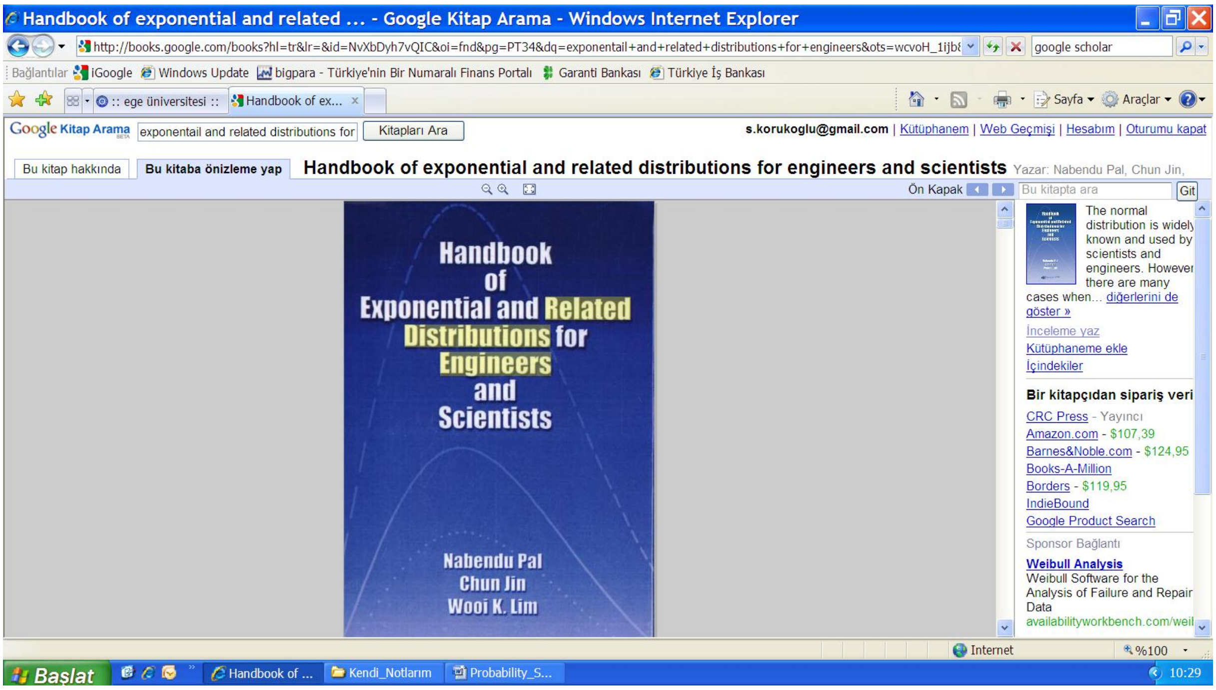
Row	DATA	Density Function	Cumulative Density F.	Hazard Rate
1	0.08985	0.914071	0.085929	1
2	1.07058	0.342809	0.657191	1
3	0.68660	0.503285	0.496715	1
4	0.35441	0.701590	0.298410	1
5	1.29436	0.274072	0.725928	1
6	0.35541	0.700885	0.299115	1
7	2.73888	0.064643	0.935357	1
8	0.44287	0.642188	0.357812	1
9	0.61692	0.539605	0.460395	1
10	0.04272	0.958181	0.041819	1
11	3.11946	0.044181	0.955819	1
12	0.55713	0.572852	0.427148	1
13	1.01853	0.361124	0.638876	1
14	0.24707	0.781083	0.218917	1
15	1.27428	0.279631	0.720369	1
16	0.45127	0.636821	0.363179	1
17	0.84473	0.429672	0.570328	1
18	0.16286	0.849710	0.150290	1
19	0.88737	0.411736	0.588264	1



Inference of Lack of Memory Property

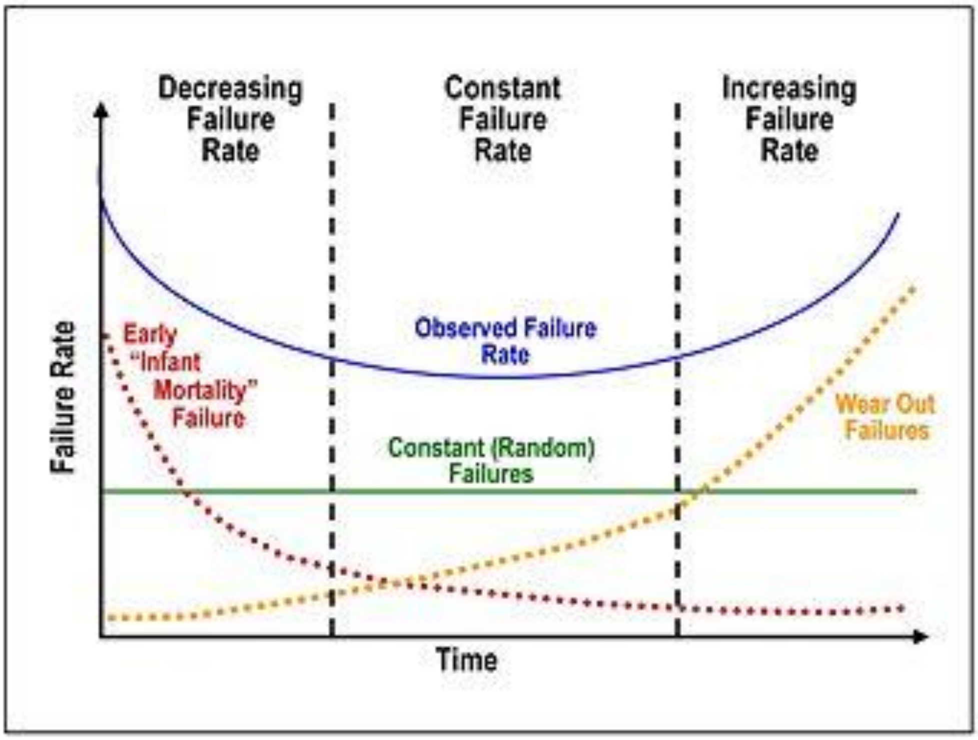
- **The lack of memory property is not that surprising when we consider the development of a Poisson process.**
- **In that development, we assumed that an interval could be partitioned into small intervals that were independent.**
- **These subintervals are similar to independent Bernoulli trials that comprise a binomial process; knowledge of previous results does not affect the probabilities of events in future subintervals.**
- **An exponential random variable is the continuous analog of a geometric random variable, and they share a similar lack of memory property.**





Reliability Studies and Exponential Distribution

- **The exponential distribution is often used in reliability studies as the model for the time until failure of a device.**
- **For example, the life time of a semiconductor chip might be modeled as an exponential random variable with a mean 50000 hours.**
- **The lack of memory property of the exponential distribution implies that the device does not wear out.**



- **That is, regardless of how long the device has been operating, the probability of a failure in the next 1000 hours is the same as the probability of a failure in the first 1000 hours of operation.**
- **The lifetime L of a device with failures caused by random shocks might be appropriately modeled as an exponential random variable.**
- **However the lifetime L of a device that suffers slow mechanical wear is better modeled a distribution such that**

$P(L < t + \Delta t / L > T)$ **increases with t.**