Specific Probability Density Functions

Normal or Gaussian Random Variable and Distribution

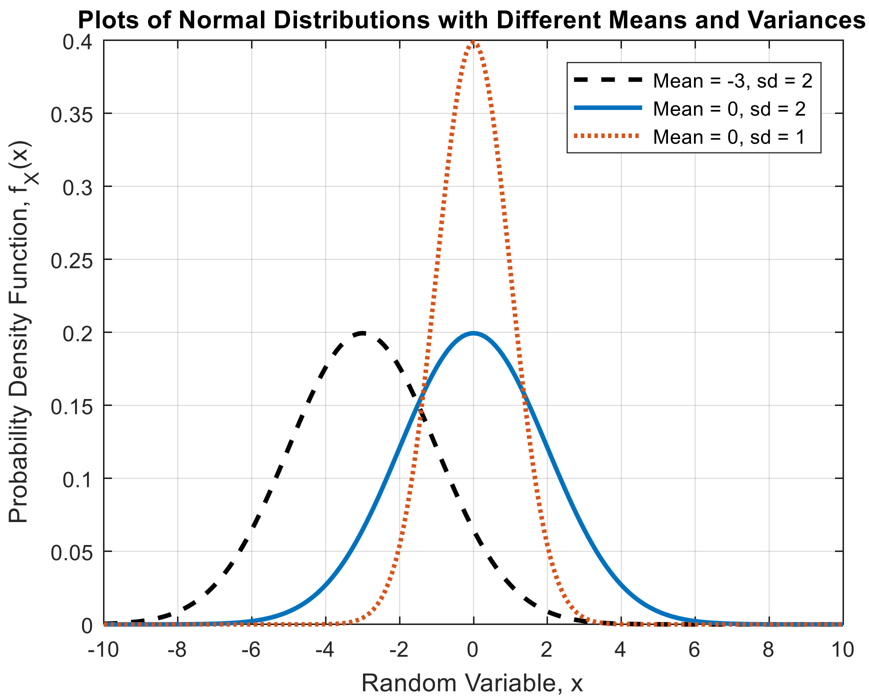
The normal distribution is the most widely used continuous probability distribution, mainly because it is tractable analytically, it follows the familiar bell shape which is consistent with a lot of population models, and the Central Limit Theorem says that, with a large enough sample, the normal distribution can be used to approximate a large variety of other distributions (e.g., Normal approximation to the Binomial). There are other continuous distributions such as chi-square, t and F distributions which are all by-products of the normal distribution.

$\mathbf{X} \sim \mathbf{N}(\mu, \sigma^2)$

$$\mathbf{pdf}: \; f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} \quad ; \; -\infty < x < \infty, \qquad -\infty < \mu < \infty, \; \sigma > 0$$

Mean : $E(X) = \mu$

Variance : $Var(X) = \sigma^2$




```
figure(1)
x = -10:0.1:10; % Set the range of values for Normal random variable%
mean1 = -3; std1 = 2; %Set the mean and standard deviation of 1st Normal distribution %
mean2 = 0; std2 = 2; %Set the mean and standard deviation of 2nd Normal distribution %
mean3 = 0; std3 = 1; %Set the mean and standard deviation of 3rd Normal distribution %
y1 = normpdf(x, mean1, std1); % Normal with mean1 and std1 %
y2 = normpdf(x, mean2, std2); % Normal with mean2 and std2 %
y3 = normpdf(x, mean3, std3); % Normal with mean3 and std3 %
plot(x, y1, '--k', x, y2, '-', x, y3, ':','LineWidth',2)
xlabel('Random Variable, x')
ylabel('Probability Density Function, f_X(x)')
title(['Plots of Normal Distributions with Different Means and Variances'])
legend([' Mean = ' num2str(mean1) , ', sd = ' num2str(std1)] , [' Mean = '
num2str(mean2), ', sd = '  num2str(std2)]  , [' Mean = ' num2str(mean3), ', sd = '
num2str(std3)])
grid on; % Adds grid on the plot %
```

If X has a $N(\mu, \sigma^2)$, then $Z = \dfrac{X-\mu}{\sigma}$ has a $N(0, 1)$

The disribution of Z is called the **standard normal distribution.**

This is a special normal distribution.

Each value of Z describes the number of standard deviations each value of X is away from mean of X.

$\mathbf{Z} \sim \mathbf{N}(0\,,\ 1)$

pdf : $f(z) = \dfrac{1}{\sqrt{2\pi}}\, e^{-\frac{z^2}{2}} \qquad ; \qquad -\infty < z < \infty$

Mean : $E(Z) = 0$

Variance : $Var(Z) = 1$

Use MATLAB to compute probability from Normal (Gaussian) Density:

Let random variable X has a normal distribution with mean, $\mu = 5$ and standard deviation, $\sigma = 2$.

(a) Find $P(X < 2.85)$ using **normcdf(2.85, 5, 2)** and get **0.1412**

(b) Find $P(X > 2.85)$ using **normcdf(2.85, 5, 2, 'upper')** and get **0.8588**

(c) Find $P(1 < X < 2.85)$ using **p = normcdf([1, 2.85], 5, 2)** and then using **p(2)-p(1)** to get **0.1184**

Let random variable Z has a normal distribution with mean, $\mu = 0$ and standard deviation, $\sigma = 1$. Find $P(Z < 2.85)$ using **normcdf(2.85)** and get **0.9978**

Use MATLAB to get the value of variable for a given probability from Normal Density:

Let random variable X has a normal distribution with mean, $\mu = 5$ and standard deviation, $\sigma = 2$, then

a) to find k so that $P(X < k) = 0.548$, use **norminv(0.548, 5, 2)** and get **5.2412**

b) to find s so that $P(X > s) = 0.7251$, use **norminv(1- 0.7251, 5, 2)** and get **3.8039**

Let random variable Z has a normal distribution with mean, $\mu = 0$ and standard deviation, $\sigma = 1$. Find d such that $P(Z < d) = 0.548$ using **norminv(0.548)** and get **0.1206**

Ex 4: A soft-drink machine is regulated so that it discharges an average of 200 milliliters per cup. The amount of drink has a normal distribution with a standard deviation equal to 15 milliliters.

(a) What is the probability that a cup contains more than 230 milliliters?

(b) What is the probability that a cup contains between 191 and 209 milliliters?

(c) What is the median amount filled by this machine?

(d) Below what value do we get the smallest 25% of the drinks?

(e) What is the minimum amount per cup for largest 5% of the drinks?

(f) Find the 75th percentile of the amount filled by this machine.

(g) If 230 milliliter cups are used for the next 100 drinks, then

(i) what is the expected number of cups that will overflow?

(ii) what is the probability that 20 out of 100 such cups will overflow?

(h) If 230 milliliter cups are used, what is the expected number of cups to be used until the 1st cup overflows?

Ex 5: A company manufactures electoral resistors with mean of 3 ohms and standard deviation of 0.1 ohms. Assume that the distribution of resistance is a normal distribution.
Find a value of d such that 95% of all manufactured resistors have resistance in the range 3 ± d ohms.

Useful Facts about the Normal Distribution:

- If $Y_1, Y_2, \dots, , Y_k$ are **independent** random variables so that $Y_i \sim N(\mu_i\,,\ \sigma_i^2)$ for $i = 1, 2, \dots, k$, then the sum of those random variables

$$X = Y_1 + Y_2 + \cdots + Y_k \sim N\left(\sum_{i=1}^{k} \mu_i \quad , \quad \sum_{i=1}^{k} \sigma_i^2\right)$$

- The normal distribution can be used to **approximate Binomial distribution when n is large**.

Exponential Random Variable and Distribution

This distribution occurs in applications such as reliability theory and queuing theory. Reasons for its use include its memoryless property (and resulting analytical tractability) and its relation to the (discrete) Poisson distribution. Thus the following random variables will often be modeled as exponential:

1. Time between two successive job arrivals to a file server (often called inter arrival time).
2. Service time at a server in a queuing network; the server could be a resource such as a CPU, an I/O device, or a communication channel.
3. Time to failure (lifetime) of a component.
4. Time required repairing a component that has malfunctioned.

They are generally used when the scenario being modeled results in **nonnegative** observations with a **right-skewed** distribution.

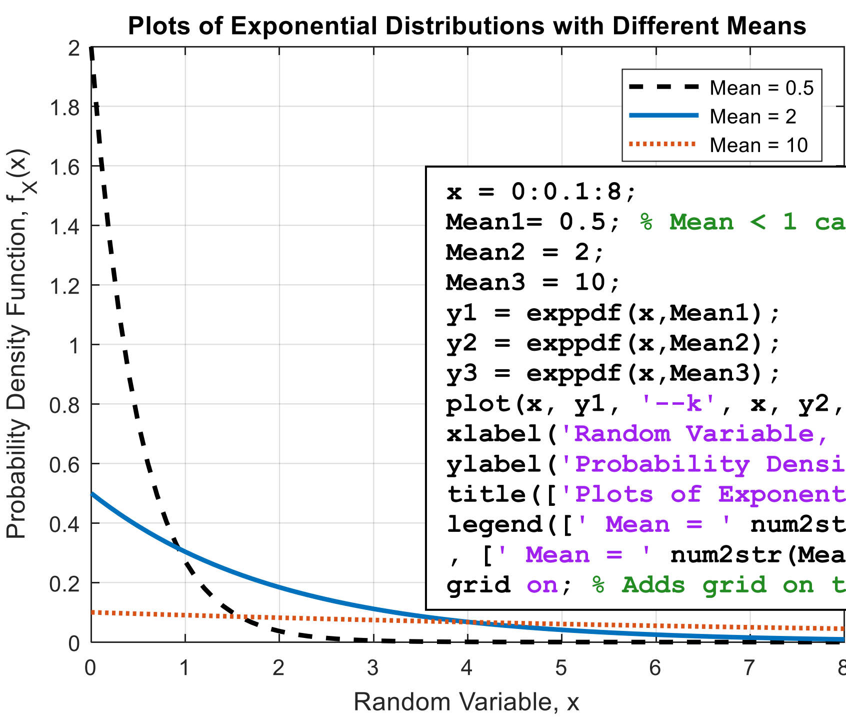
$\lambda =$ Rate of occurrences

$\mathbf{X \sim Exp}(\lambda)$

pdf : $f(x) = \lambda\, e^{-\lambda x}$; $x > 0,\ \lambda > 0$

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

Variance : $Var(X) = 1/\lambda^2$




```
x = 0:0.1:8;
Mean1= 0.5; % Mean < 1 case
Mean2 = 2;
Mean3 = 10;
y1 = exppdf(x,Mean1);
y2 = exppdf(x,Mean2);
y3 = exppdf(x,Mean3);
plot(x, y1, '--k', x, y2, '-', x, y3, ':','LineWidth',2)
xlabel('Random Variable, x')
ylabel('Probability Density Function, f_X(x)')
title(['Plots of Exponential Distributions with Different Means'])
legend([' Mean = ' num2str(Mean1) ] , [' Mean = ' num2str(Mean2)]
, [' Mean = ' num2str(Mean3)])
grid on; % Adds grid on the plot %
```

❖ If Y is the number of independent occurrences in a **unit** time interval, then $\boldsymbol{Y}$ **has a Poisson distribution with mean $\boldsymbol{\lambda}$**.
Let $\boldsymbol{X}$ **be the time between two consecutive occurrences**.
Then $\boldsymbol{X}$ **has an Exponential distribution with $\boldsymbol{\lambda}$.**

Ex 7: Consider Web server with an average rate of requests $\lambda = 0.1$ jobs per second. Assuming that the number of arrivals per unit time is Poisson distributed, (the elapse time without requests) arrival time, X, is an exponential random variable with $\lambda = 0.1$ jobs per second.

(a) What is the mean request arrival time?

(b) Compute the probability that at most 15 seconds elapses without requests.

Ex 8: Let T be the random variable which measures the time to failure of a certain electronic component.
T has an exponential distribution with mean of 5 years.
If 5 of these components are randomly selected, then what is the probability that more than 3 are still functioning at the end of 8 years?

Ex 9: A certain type of electronic component has an exponential distribution with a mean life of 500 hours.
Let X be the life (or the time to failure) of such component.
If such a component has been in operation for more than 300 hours, find the conditional probability that it will last at least for another 600 hours.

❖ This is called the **memoryless property** of the **exponential distribution**.

Lognormal Random Variable and Distribution

The random variable X has a lognormal distribution if the random variable Y = ln(X) has a normal (Gaussian) distribution with mean μ and variance σ^2.

$\mathbf{X \sim logNormal}(\mu\,, \sigma^2)$

$$\textbf{pdf} : \ f(x) = \frac{1}{x\sigma\sqrt{2\pi}}\, e^{-\,(\,ln(x)-\mu)^2/(2\sigma^2)} \quad ; \ x > 0, \ \ \mu > 0, \ \ \sigma > 0$$

$\textbf{Mean} : E(X) = e^{\mu+\frac{\sigma^2}{2}}$

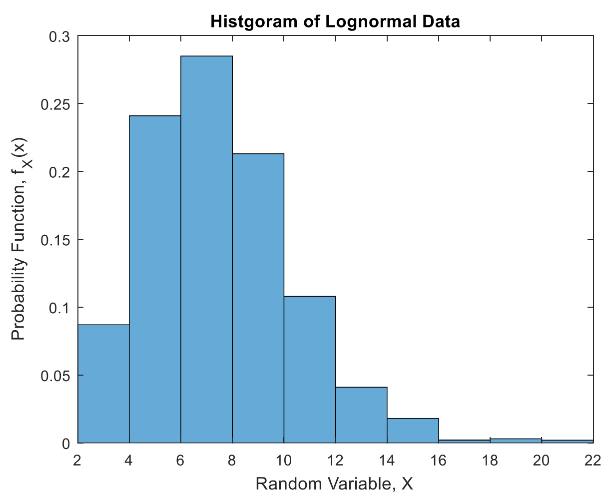
$\textbf{Variance} : Var(X) = \left(e^{\sigma^2} - 1\right)e^{(2\mu+\sigma^2)}$

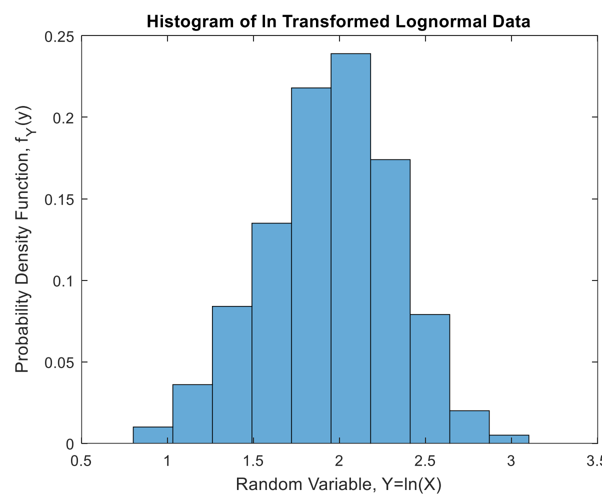
Ex 10: The life, in thousands of miles, of a certain type of electronic control for locomotives has an approximately lognormal distribution with μ = 5.149 and σ = 0.737.

(a) Compute the average life of a randomly selected such an electronic control.

(b) Find the 5th percentile of the life of such an electronic control.

(c) Compute the probability that a randomly selected locomotive still works after 100 thousand miles.






```
%% Log transformation of skewed data to Normal
Clear  ; clc
n= 1000;
A= 7.5; B= 1;
y1=gamrnd(A,B,n,1); % Randomly generate data from a skewed
distribution called gamma(shape=A, scale=B);
figure(1);
histogram(y1, 10, 'Normalization', 'probability')
xlabel('Random Variable, X')
ylabel('Probability Function, f_X(x)')
title(['Histgoram of Lognormal Data'])
%Natural log transformation;
x=log(y1);
figure(2);
histogram(x, 10, 'Normalization', 'probability')
xlabel('Random Variable, Y=ln(X)')
ylabel('Probability Density Function, f_Y(y)')
title(['Histogram of ln Transformed Lognormal Data'])
```