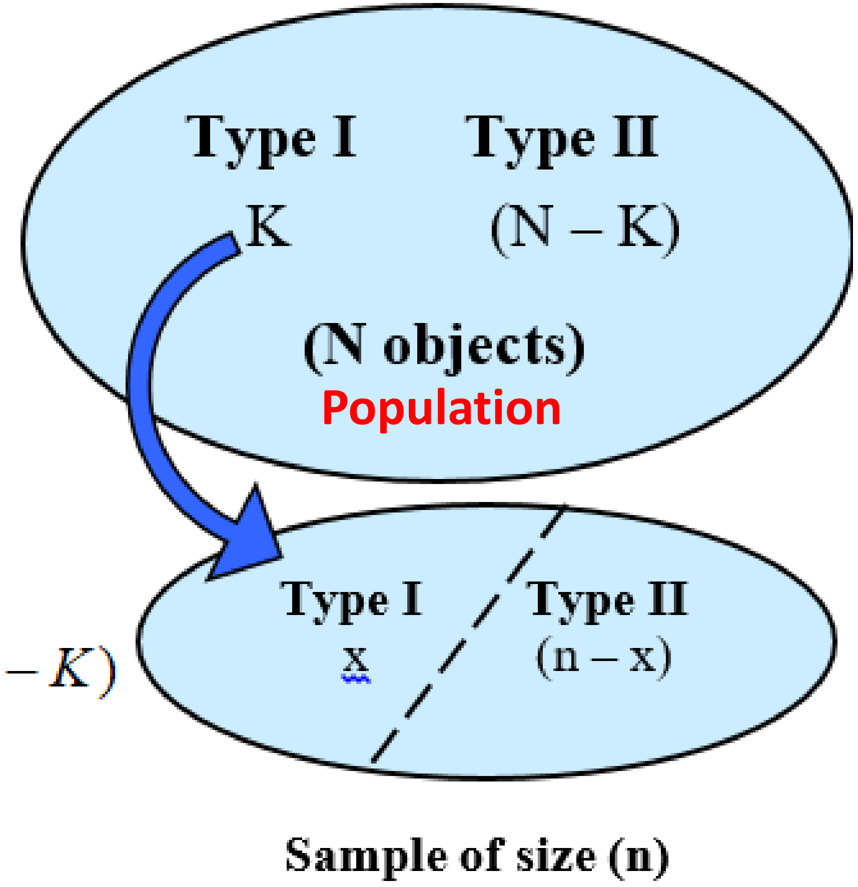
Hypergeometric Distribution

Consider the sampling **without replacement** of a lot of (N) items, (K) of which are of one type and (N – K) of a second type. The probability of obtaining (x) items in a selection of (n) items without replacement obeys the hyper-geometric distribution:

$$P(X=x)=\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}},\ x = 0, 1, \ldots, \min(n, K)$$



$\binom{K}{x}$: Number of ways x items are chosen out of K.

$\binom{N-K}{n-x}$: Number of ways (n-x) items are chosen out of $(N-K)$

$\binom{N}{n}$: Number of ways the n items are chosen out of N.

$\min(n,\ K)$: x cannot be larger than k, the number of items of type I and also cannot be larger than n, the sample itself.

Activate

Theorem: Mean and Variance of the Hypergeometric Distribution

(The theorem is stated without proof)

The mean and the variance of the hyper-geometric random variable are given by:

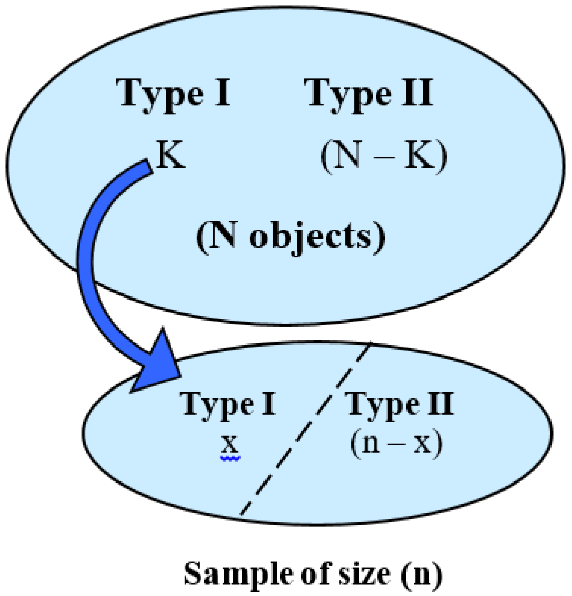
$$\mu_X = E(X) = n\frac{K}{N} = n\,p \;\left(p = \frac{K}{N}\right.$$ is the ratio of items of type (I) to the total population)

$$\sigma_X^2 = Var(X) = n\left(\frac{K}{N}\right)\left(1 - \frac{K}{N}\right)\left(\frac{N-n}{N-1}\right) = n\,p\,(1-p)\left(\frac{N-n}{N-1}\right)$$

Remark: These expressions exhibit some similarities with the mean and variance of the binomial distribution, which were derived earlier and given by

Binomial Distribution:

$$\mu_X = E(X) = n\,p \qquad \sigma_X^2 = Var(X) = n\,p\,(1-p)$$



In one of the examples considered later, we will demonstrate the differences between the binomial and the hypergeometric distributions.

Hypergeometric Distribution

EXAMPLE: A committee of seven members is to be formed at random from a class with 25 students of whom 15 are girls. Find the probability that:

a- No girls are among the committee
b- All committee members are girls
c- The majority of the members are girls

$$P(X=x)=\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}},\quad x = 0, 1, \ldots, \min(n, K)$$

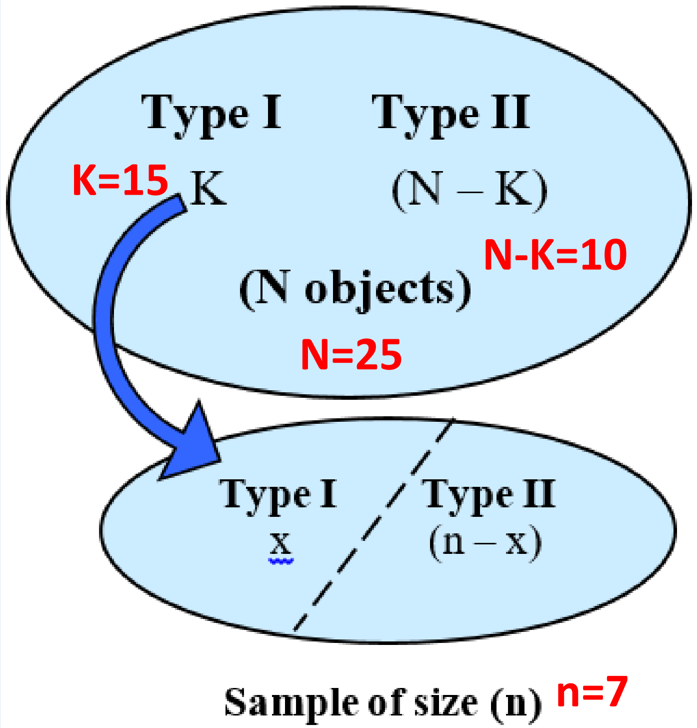
N=25
K=15
n=7

SOLUTION: Let X be the r.v. represents the number of girls in the committee.

a- $$\text{P(No girls among the committee)} = P(X=0)=\frac{\binom{15}{0}\binom{10}{7}}{\binom{25}{7}}$$

b- $$\text{P(All committee members are girls)} = P(X=7)=\frac{\binom{15}{7}\binom{10}{0}}{\binom{25}{7}}$$

c- $$\text{P(Majority members are girls)} = \sum_{x=4}^{7} P(X=x)=\sum_{x=4}^{7}\frac{\binom{15}{x}\binom{10}{7-x}}{\binom{25}{7}}$$



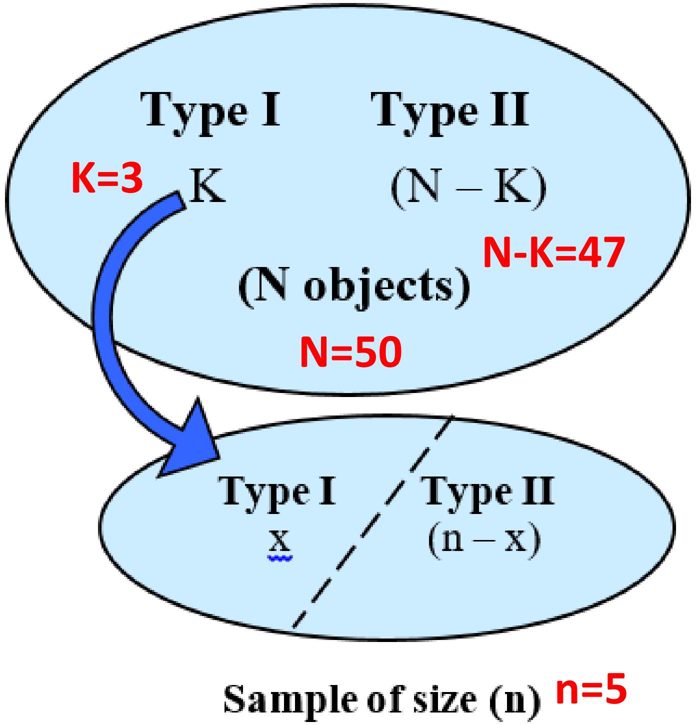
EXAMPLE: Fifty small electric motors are to be shipped. However, before such a shipment is accepted, an inspector chooses 5 of the motors randomly and inspects them. If none of these tested motors is defective, the lot is accepted. If one or more are found to be defective, the entire shipment is inspected. Suppose that there are, in fact, three defective motors in the lot. What is the probability that the entire shipment is inspected?

SOLUTION: Let X be the number of defective motors found in the sample.

- X assumes the values (0, 1, 2, 3) according to the hypergeometric distribution given as

$$P(X=x)=\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}},\ x = 0, 1, \ldots, \min(n, K)$$

$$P(X=x)=\frac{\binom{3}{x}\binom{47}{5-x}}{\binom{50}{5}};\ x = 0, 1, 2, 3;$$



$$\text{P(Shipment is accepted)} = P(X=0)=\frac{\binom{3}{0}\binom{47}{5}}{\binom{50}{5}}=0.72$$

$$\text{P(Shipment is inspected)} = P(X\geq 1) = 1-P(X=0)=0.28$$

Theorem: Binomial Approximation to the Hyperbolic Distribution

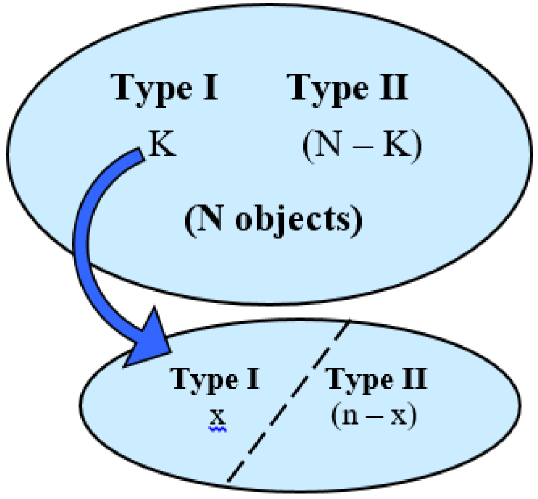
For large (N), one can use the approximation

$$P(X=x) \cong \binom{n}{x} P^x (1-P)^{n-x}; \quad x=0,1,\ldots n \quad ; P=\frac{k}{N}$$

$$P(X=x)=\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}, \quad x = 0, 1, \ldots, \min(n, K)$$

This approximation gives very good results if the sample size to the total population, $\frac{n}{N} \leq 0.1$.

- **Binomial Distribution**: When we considered the binomial distribution, (sampling with replacement), we assumed that the probability of a success remains constant over the n repeated trials.
- When the sampling is without replacement, the probability of a success changes as more items are drawn.
- When the ratio n/N < 0.1, the probability p = K/N of type I does not change much from trial to trial. Hence, p can be considered a constant, even though the sampling is without replacement.



Sample of size (n)

Without replacement, K/N changes from trial to trial.
With replacement, K/N remains constant from trial to trial.

Hyper-geometric Versus Binomial Distributions

EXAMPLE: A box contains 20 good (G) items and 5 bad (B) items. Consider the following

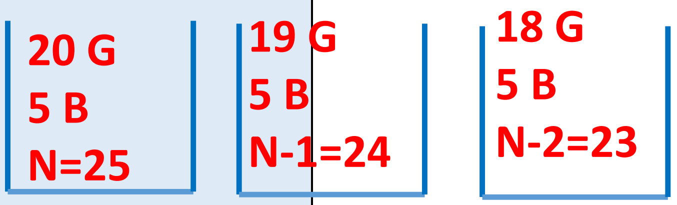
a. Three items are drawn **without replacement**, find the probability that the sequence of objects obtained is (GGB) in the given order.

b. Three items are drawn **with replacement**, find the probability that the sequence of objects obtained is (GGB) in the given order.

c. Three items are drawn **without replacement**, find the probability that exactly one bad item is obtained.

d. Three items are drawn **with replacement**, find the probability that exactly one bad item is obtained.

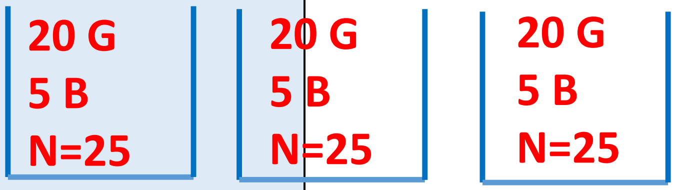
Part a: $P(GGB)=\left(\frac{20}{25}\right)\left(\frac{20-1}{25-1}\right)\left(\frac{5}{25-2}\right)=\left(\frac{20}{25}\right)\left(\frac{19}{24}\right)\left(\frac{5}{23}\right)$



Sampling without replacement

Part b: $P(GGB)=\left(\frac{20}{25}\right)\left(\frac{20}{25}\right)\left(\frac{5}{25}\right)$

Part c: $P(GGB)+P(GBG)+P(BGG)=3\left(\frac{20}{25}\right)\left(\frac{20-1}{25-1}\right)\left(\frac{5}{25-2}\right)=3\left(\frac{20}{25}\right)\left(\frac{19}{24}\right)\left(\frac{5}{23}\right)$



Sampling with replacement

Part c: $P(X=1)=\dfrac{\binom{5}{1}\binom{20}{2}}{\binom{25}{3}}=3\left(\frac{20}{25}\right)\left(\frac{19}{24}\right)\left(\frac{5}{23}\right)$; Hypergeometric distribution **0.413**

Part d; $P(X=1)=\binom{3}{1}(p)^1(1-p)^2=3\left(\frac{5}{25}\right)\left(\frac{20}{25}\right)\left(\frac{20}{25}\right)$; binomial distribution (n=3, p=5/25) **0.384**

EXAMPLE: Binomial versus hypergeometric distribution

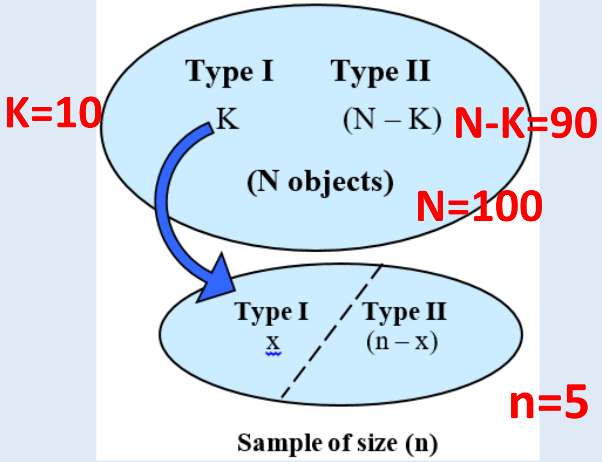
One hundred small electric motors are to be shipped. However, before such a shipment is accepted, an inspector chooses 5 of the motors randomly and inspects them. If none of these tested motors is defective, the lot is accepted. If one or more are found to be defective, the entire shipment is inspected. Suppose that 10% of the shipment are defective, find the probability that the shipment is accepted

a. Using the exact hypergeometric distribution

b. Using the binomial approximation, if valid.

Exact Solution: The problem parameters are: population Size: N = 100; Sample Size n = 5

Number of defective items K = (0.1)(100) =10



$$\text{P(Shipment is accepted)} = P(X=0) = \frac{\binom{10}{0}\binom{90}{5}}{\binom{100}{5}} = 0.58375$$

Approximate Solution: The ratio n/N = 5/100 = 0.05. This justifies the use of the binomial distribution. Here, p is assumed constant and the binomial parameters are n = 5, p = 0.1

$$\text{P(Shipment is accepted)} = P(X=0) = \binom{5}{0} p^0 (1-p)^5 = (0.9)^5 = 0.59049$$

Poisson Distribution

Definition: A discrete random variable (X) is said to have a *Poisson distribution* if it has the following probability mass function:

$$P(X = x) = e^{-b}\frac{b^x}{x!};\quad x = 0, 1, 2, \ldots;\ b > 0$$

The sample space is discrete and countably infinite

We can verify that this is, indeed, a valid probability mass function by summing over all values of x.

$$\sum_{x=0}^{\infty} P(X = x) = \sum_{x=0}^{\infty} e^{-b}\frac{b^x}{x!} = e^{-b}\left[1 + \frac{b}{1!} + \frac{b^2}{2!} + \frac{b^3}{3!} + \ldots\right]$$

The summation on the right side is easily recognized as the power series expansion of e^b.

Therefore, $\sum_{x=0}^{\infty} e^{-b}\frac{b^x}{x!} = e^{-b}e^{b} = 1$.

Mean Value of the Poisson Distribution

Theorem: If X is a Poisson random variable with parameter b, then its mean value is given as

$$\mu_X = E(X) = b$$

Proof: To find the mean value of X, we follow the same procedure we used to find value of the binomial and the geometric distributions.

$$E(X) = \sum_{x=0}^{\infty} x e^{-b} \frac{b^x}{x!} = \sum_{x=1}^{\infty} x e^{-b} \frac{b^x}{x!}$$

$$= \sum_{x=1}^{\infty} x e^{-b} \frac{b^x}{x(x-1)!} = \sum_{x=1}^{\infty} e^{-b} \frac{b^x}{(x-1)!}$$

Let u = x-1 (or x = u + 1), and change the index of the summation from x to u. The result is

$$E(X) = \sum_{u=0}^{\infty} e^{-b} \frac{b^{u+1}}{u!} = b \underbrace{\sum_{u=0}^{\infty} e^{-b} \frac{b^u}{u!}}_{=1} = b$$

As was shown earlier, the summation on the right side equals 1.

Variance of the Poisson Distribution

Theorem: If X is a Poisson random variable with parameter b, then its variance is

$$\sigma_X^2 = Var(X) = b$$

Proof: First, we find $E(X(X-1))$

$$E(X(X-1)) = \sum_{x=0}^{\infty} x(x-1)e^{-b}\frac{b^x}{x!} = \sum_{x=2}^{\infty} x(x-1)e^{-b}\frac{b^x}{x(x-1)(x-2)!}$$

$$E(X(X-1)) = \sum_{x=2}^{\infty} e^{-b}\frac{b^x}{(x-2)!}$$

Let u = x-2 in the above summation, or x = u + 2, then

$$E(X(X-1)) = \sum_{u=0}^{\infty} e^{-b}\frac{b^{u+2}}{u!} = b^2\sum_{u=0}^{\infty} \mathrm{e}^{-b}\frac{b^u}{u!} = b^2$$

But, $E(X(X-1)) = E(X^2) - E(X) \Rightarrow E(X^2) = E(X(X-1)) + E(X)$

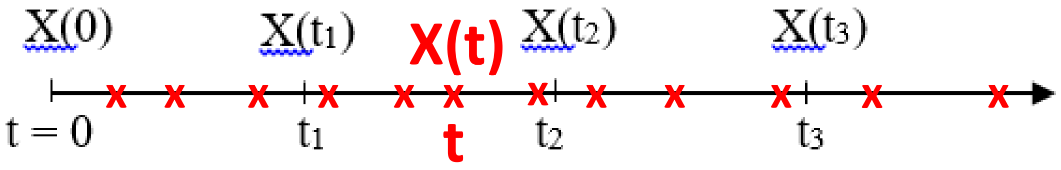
The variance of X can therefore be obtained as

$$\sigma_X^2 = E(X^2) - \mu_X^2 = E(X(X-1)) + E(X) - \mu_X^2 \Rightarrow \sigma_X^2 = b^2 + b - b^2 = b$$

Poisson Process

Consider a counting process in which events occur at a rate of (λ) occurrences per unit time. Let X(t) be the number of occurrences recorded in the interval (0, t). We define the *Poisson process* by the following assumptions:

1- $X(0) = 0$, i.e., the counting begins at t = 0 with the counter set to 0.
2- For non-overlapping time intervals $(0, t_1)$, (t_2, t_3), the number of occurrences $\{X(t_1) - X(0)$ and $\{X(t_3) - X(t_2)$ are independent.
3- The probability distribution of the number of occurrences in any time interval depends only on the length of that interval.
4- The probability of an occurrence in a small time interval (Δt) is approximately ($\lambda \Delta t$).



The idea is to divide one of the intervals of length T into N subintervals, each with length $\Delta t = \frac{T}{N}$. The probability of a success in each subinterval (based on 4) is $p = \lambda\left(\frac{T}{N}\right)$.

The probability of x successes in the N subintervals obeys the binomial distribution

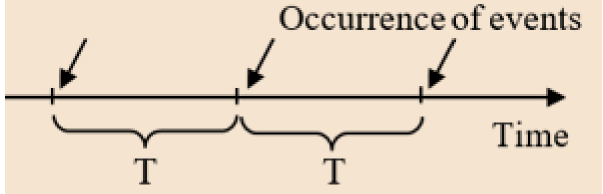
$$P(X=x) = \binom{N}{x} p^x (1-p)^{N-x}; \quad x = 0,1,\ldots N;\ p = \lambda\left(\frac{T}{N}\right)$$

It can be shown that in the limit as $N \to \infty$ so that $\Delta t \to 0$, we get

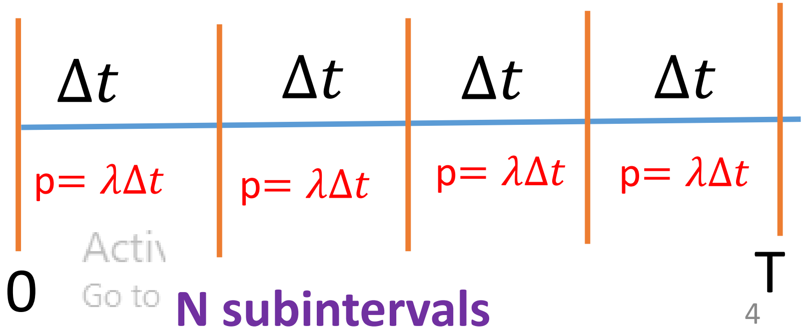
$$\lim_{N\to\infty} P(X=x) = \binom{N}{x}\left(\frac{\lambda T}{N}\right)^x \left(1-\frac{\lambda T}{N}\right)^{N-x} = e^{-\lambda T}\frac{(\lambda T)^x}{x!}; x = 0,1,2\ldots$$

In the next lecture, we will see that the inter-arrival time between occurrences follows the exponential distribution

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$



An example will be presented at the end of the lecture that compares the Poisson distribution with its binomial approximation



EXAMPLE: Messages arrive to a computer server according to a Poisson distribution with a mean rate of 10 messages/hour

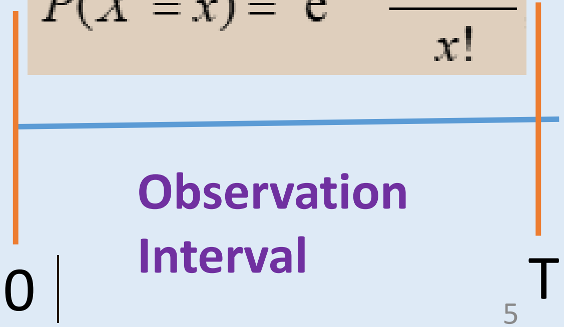
c- What is the probability that exactly three messages will arrive in one hour?

d- What is the probability that at most two messages will arrive in 30 minutes?

e- Find the expected number of messaging arriving in observation time of 90 minutes.

SOLUTION:

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



a- $\lambda = 10$ messages/hour ➔ T = 1 hour

$$P(X=x)=e^{-\lambda T}\frac{(\lambda T)^x}{x!};x=0,1,2,\ldots$$

With λ=10 and T = 1, we get

$$P(X=x)=e^{-10\times 1}\frac{(10\times 1)^x}{x!}=e^{-10}\frac{(10)^x}{x!}\Rightarrow P(X=3)=e^{-10}\frac{(10)^3}{3!}$$

b- With λ=10 and T = 1/2, $P(X=x)=e^{-10\times 1/2}\frac{(10\times 1/2)^x}{x!}=e^{-5}\frac{(5)^x}{x!};x=0,1,2,\ldots$

$$\text{P(at most two)}=P(X\le 2)=e^{-5}\left[\frac{(5)^0}{0!}+\frac{(5)^1}{1!}+\frac{(5)^2}{2!}\right]=e^{-5}\left[1+\frac{(5)}{1}+\frac{25}{2}\right]=\left(\frac{37}{2}\right)e^{-5}$$

c. With λ=10 and T = 3/2, $E(X)=\lambda T=(10)(3/2)=15$ messages

EXAMPLE: The number of cracks in a section of a highway that are significant enough to require repair is assumed to follow a Poisson distribution with an average of two cracks per mile.

a- What is the probability that there are no cracks in a section of 5 miles?

b- What is the probability that there is at least two crack in a section of 3 miles?

SOLUTION: With λ=2 and T = 5, we get

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

a-

$$P(X=x) = e^{-2\times 5}\frac{(2\times 5)^x}{x!} = e^{-10}\frac{(10)^x}{x!}; x = 0,1,2,\ldots$$

$$\text{P(no cracks)} = P(X=0) = e^{-10}$$

With λ=2 and T = 3, we get

b-

$$P(X=x) = e^{-2\times 3}\frac{(2\times 3)^x}{x!} = e^{-6}\frac{(6)^x}{x!}; x = 0,1,2,\ldots$$

$$\text{P(at least two cracks)} = P(X\geq 2) = 1-\left[P(X=0)+P(X=1)\right]$$

$$=1-\left[e^{-6}\frac{(6)^0}{0!}+e^{-6}\frac{(6)^1}{1!}\right] = 1- 8e^{-6}$$

1 – 7exp(-6)

EXAMPLE: The number of telephone calls that arrives at a certain office is modeled by a Poisson random variable. Assume that on the average there are five calls per hour.

a. What is the average (mean) time between phone calls?
b. What is the probability that at least 30 minutes will pass without receiving any phone call?
c. What is the probability that there are exactly three calls in an observation interval of two consecutive hours?
d. What is the probability that there is exactly one call in the first hour and exactly two calls in the second hour of a two-hour observation interval?

SOLUTION: $P(X = x) = e^{-\lambda T}\dfrac{(\lambda T)^x}{x!}\,; x = 0, 1, 2, \ldots\,; \lambda = 5$

a. Here, λ = 5 calls/hour ➔ mean time between calls = 60minutes/5calls = 12 minutes

b. Here, T = ½. X=0. Therefore, $\Rightarrow P(X = 0) = e^{-5\times 1/2}\dfrac{(5 \times 1/2)^0}{0!} = e^{-5/2}$

c. Here, T = 2, X=3 Therefore, $\Rightarrow P(X = 3) = e^{-5\times 2}\dfrac{(5 \times 2)^3}{3!} = \dfrac{e^{-10}(10)^3}{3!}$

d. Here, $(T_1 = 1,\ X_1 = 1)$, $(T_2 = 1,\ X_2 = 2)$

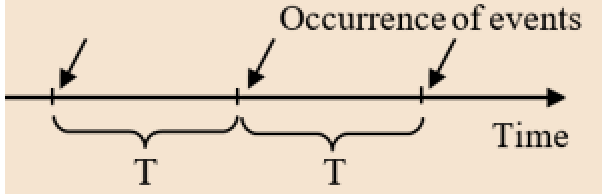
$$\Rightarrow P(X_1 = 1 \cap X_2 = 2) = (X_1 = 1)\ (X_2 = 2) \text{ due to independence}$$

$$\Rightarrow P(X_1 = 1 \cap X_2 = 2) = \left[e^{-5\times 1}\frac{(5 \times 1)^1}{1!}\right]\left[e^{-5\times 1}\frac{(5 \times 1)^2}{2!}\right] = \left[5e^{-5}\right]\left[\frac{25}{2}e^{-5}\right]$$

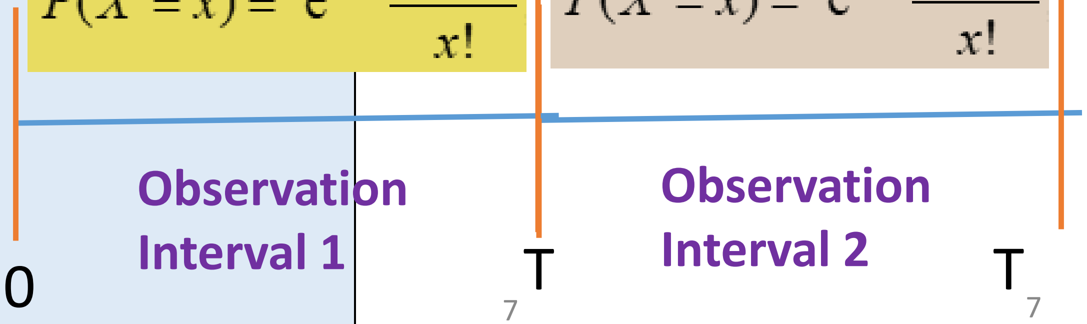
In the next lecture, we will see that the inter-arrival time between occurrences follows the exponential distribution

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & t \ge 0 \\ 0 & t < 0 \end{cases}$$

$E(T) = 1/\lambda$



$P(X = x) = e^{-\lambda T}\dfrac{(\lambda T)^x}{x!}$ $P(X = x) = e^{-\lambda T}\dfrac{(\lambda T)^x}{x!}$



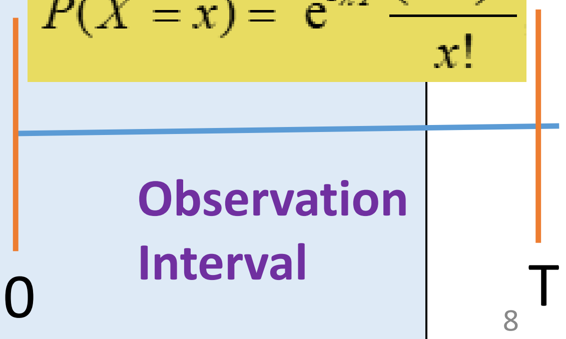
EXAMPLE: Binomial Distribution and the Poisson Distribution

A web server receives an average number of 2 queries per minute. We will compute the probability of no queries, one query, and two queries in one minute using

a. The Poisson distribution

b. The binomial approximation

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



SOLUTION:

a. Poisson Distribution: With λ=2 and T = 1, we get

$$P(X=x) = e^{-2\times 1}\frac{(2\times 1)^x}{x!} = e^{-2}\frac{(2)^x}{x!}; x = 0, 1, 2,$$

$$P(X=0) = e^{-2} = 0.13533; \quad P(X=1) = 2e^{-2} = 0.27067; \quad P(X=2) = 2e^{-2} = 0.27067$$

b. Binomial Distribution

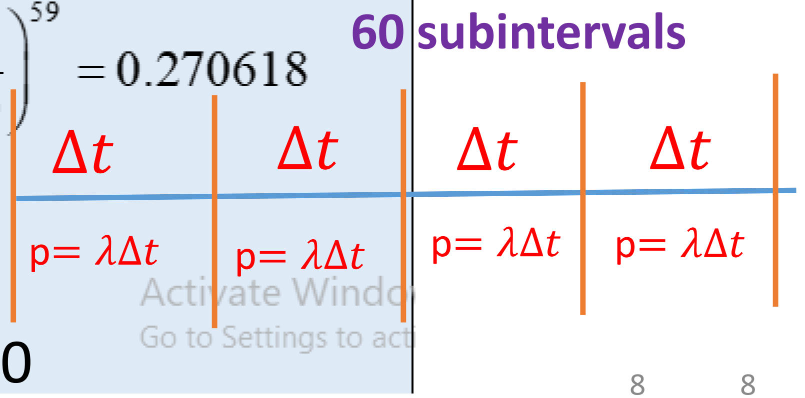
Let us divide the one minute interval into 60 seconds and use the approximation considered earlier

$$P(X=x) = \binom{N}{x}\left(\frac{\lambda T}{N}\right)^x\left(1-\frac{\lambda T}{N}\right)^{n-x} = \binom{60}{x}\left(\frac{2}{60}\right)^x\left(1-\frac{2}{60}\right)^{n-x}; x = 0, 1, 2, \ldots 60$$

$$P(X=x) = \binom{N}{x}p^x(1-p)^{N-x};$$

$$P(X=0) = \binom{60}{0}\left(\frac{2}{60}\right)^0\left(1-\frac{2}{60}\right)^{60} = 0.130799; \quad P(X=1) = \binom{60}{1}\left(\frac{2}{60}\right)^1\left(1-\frac{2}{60}\right)^{59} = 0.270618$$

$$P(X=2) = \binom{60}{2}\left(\frac{2}{60}\right)^2\left(1-\frac{2}{60}\right)^{58} = 0.27528$$

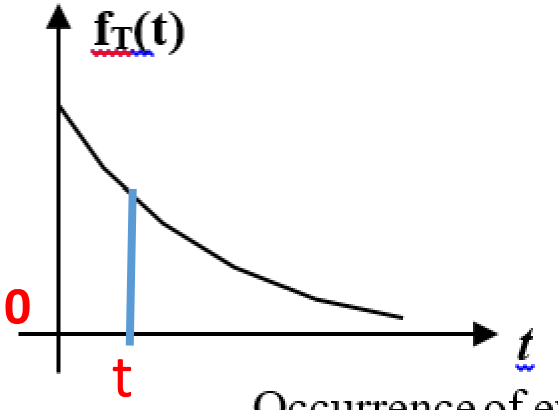


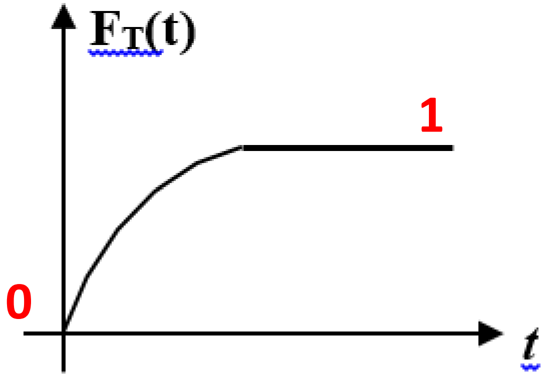
Remark: making N larger and larger yields a more accurate approximation.

Exponential Distribution

Definition: A random variable T is said to have an exponential distribution with a parameter λ ($\lambda > 0$) if T has a continuous distribution for which the pdf $f_T(t)$ is:

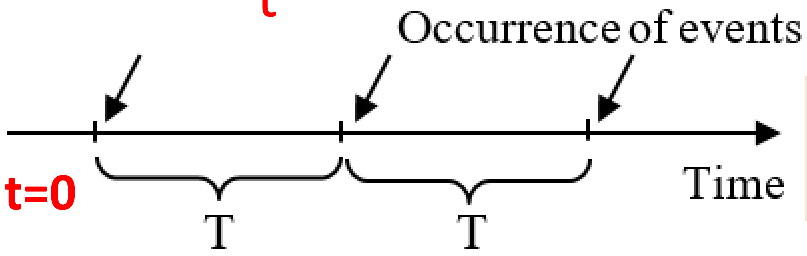
$$f_T(t) = \begin{cases} \lambda\, e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$





The cumulative distribution function is

$$F_T(t) = P(T \leq t) = \int_{-\infty}^{t} f_T(t)\,dt = \begin{cases} 1 - e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$



$$P(X = x) = e^{-\lambda T_1} \frac{(\lambda T_1)^x}{x!}$$

Applications of the Exponential Distribution

- The exponential distribution is used to represent the distribution of the time that elapses between occurrences of a Poisson process.
- It also represents the waiting time until the first event occurs in a Poisson process.
- It has been used to represent the period for which a machine or an electronic component will operate without breaking down.
- The period required to take care of a customer at some service facility (the service time).
- The period between the arrivals of two successive customers at a facility (bank, supermarket,).

Exponential Distribution

Theorem: If the random variable (T) has an exponential distribution with parameter (λ), then:

$$\mu_T = E(T) = \int_0^{\infty} t\lambda e^{-\lambda t} dt = \frac{1}{\lambda}$$

$$E(T^2) = \int_0^{\infty} t^2 \lambda e^{-\lambda t} dt = \frac{2}{\lambda^2}$$

$$\sigma_T^2 = E(T^2) - E^2(T) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

$$f_T(t) = \begin{cases} \lambda\, e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

Proof: These results can be easily obtained using integration by parts.

$$E\{g(X)\} = \int_{-\infty}^{\infty} g(x)\, f_X(x) dx \; ; \text{ X is continuous}$$

Exponential Distribution

EXAMPLE: Find the mode and median of the exponential distribution

$$f_T(t) = \begin{cases} \lambda\, e^{-\lambda t} & t \ge 0 \\ 0 & t < 0 \end{cases}$$

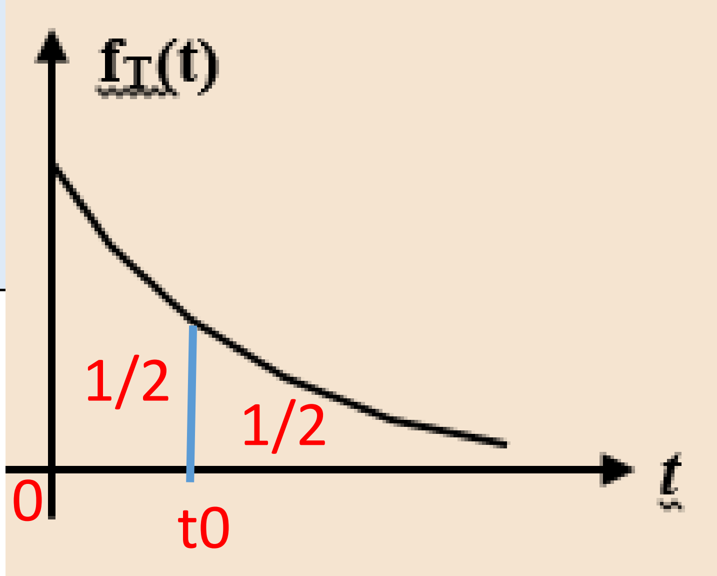
SOLUTION:

a. The mode is $t_{\text{mode}} = 0$, since this is the point which has the highest likelihood of occurrence.

Note that here no differentiation is needed to find the point at which the pdf is maximum.

b. The median is some value t_0 such that $P(\mathrm{T} \le t_0) = P(\mathrm{T} > t_0) = 1/2$

That is, $$\int_0^{t_0} \lambda\, e^{-\lambda t}\, dt = \int_{t_0}^{\infty} \lambda\, e^{-\lambda t}\, dt = 1/2 \Rightarrow e^{-\lambda t_0} = 1/2 \quad \Rightarrow t_0 = \frac{\ln(2)}{\lambda}$$



EXAMPLE: Suppose that the lifetime T of a power transmission tower, measured in years, is described by an exponential distribution with mean equals to 25 years

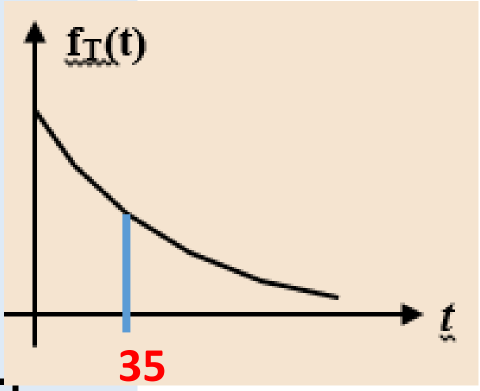
$$f_T(t) = \begin{cases} \frac{1}{25}\, e^{-t/25} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

If three towers, operated independently, were erected at the same time, what is the probability that at least two towers will still stand after 35 years.

SOLUTION: First, we find the probability that the lifetime of one tower is greater than 35 years

$$P(\mathrm{T} \geq 35) = \int_{35}^{\infty} \frac{1}{25}\, e^{\frac{-t}{25}}\, dt = e^{-35/25} = e^{-1.4}$$

$$P(X = x) = \binom{n}{x}\, p^x (1-p)^{n-x}$$

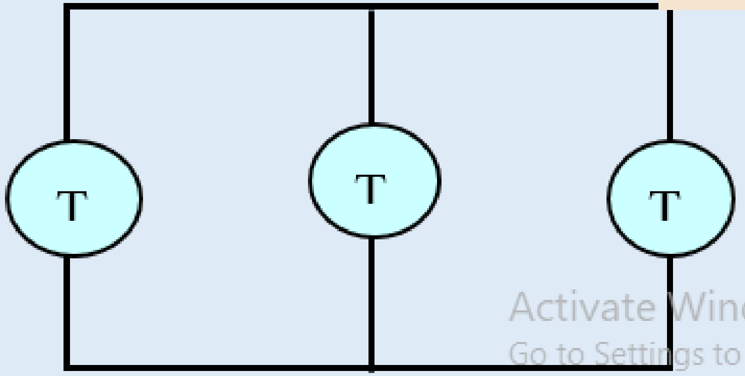


Next, let X be the random variable representing the number of operating towers after 35 years.

X follows the binomial distribution with parameters n=3 and $p = e^{-1.4}$

$$P(X \geq 2) = \binom{3}{2} (p)^2 (1-p)^1 + \binom{3}{3} (p)^3 (1-p)^0$$

$$= 3\,(p)^2 (1-p) + (p)^3$$



EXAMPLE: The number of telephone calls that arrive at a certain office is modeled as a Poisson process with an average arrival rate of $\lambda = 1/12$ calls minute (5 calls/hour)

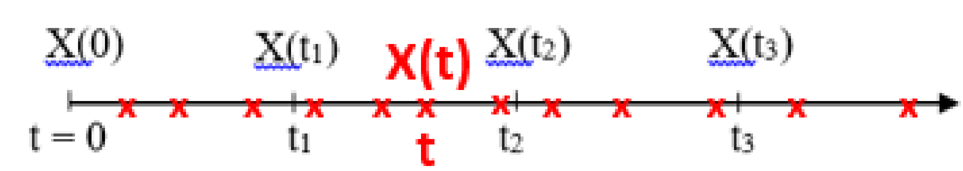
a. What is the average (mean) time between phone calls?

b. What is the probability that the waiting time between successive occurrences is between 5 and 10 minutes?

c. What is the probability that at least 30 minutes will pass without receiving any phone call?

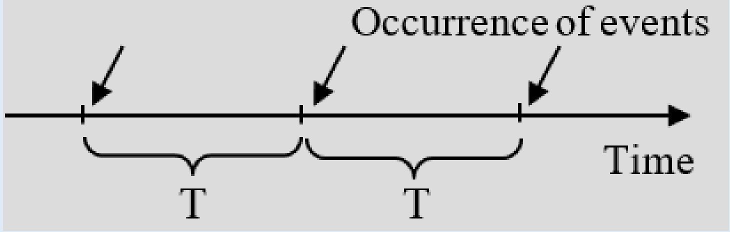
Solution:

The waiting time between arrivals follows the exponential distribution: $f_T(t) = \begin{cases} \lambda\, e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$



a. The mean value of the waiting time $E(T) = \dfrac{1}{1/12} = 12$ minutes

b. $P(5 \leq T \leq 10) = \displaystyle\int_5^{10} \frac{1}{12}\, e^{\frac{-t}{12}}\, dt = e^{-5/12} - e^{-10/12}$



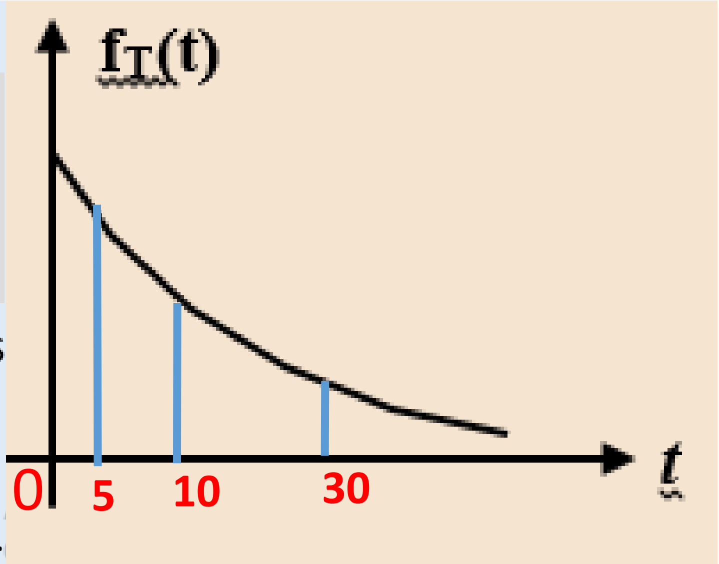
c. $P(T \geq 30) = \displaystyle\int_{30}^{\infty} \frac{1}{12}\, e^{\frac{-t}{12}}\, dt = e^{-30/12} = e^{-2.5}$

We have also solved this part in the previous lecture using the Poisson distribution as

$$P(X = 0) = e^{-\lambda T_1} \frac{(\lambda T_1)^x}{x!} = e^{-\frac{1}{12}30} \frac{(\frac{1}{12} 30)^0}{0!} = e^{-2.5}$$

$$P(X = x) = e^{-\lambda T_1} \frac{(\lambda T_1)^x}{x!}$$

0 **Observation Interval** T1

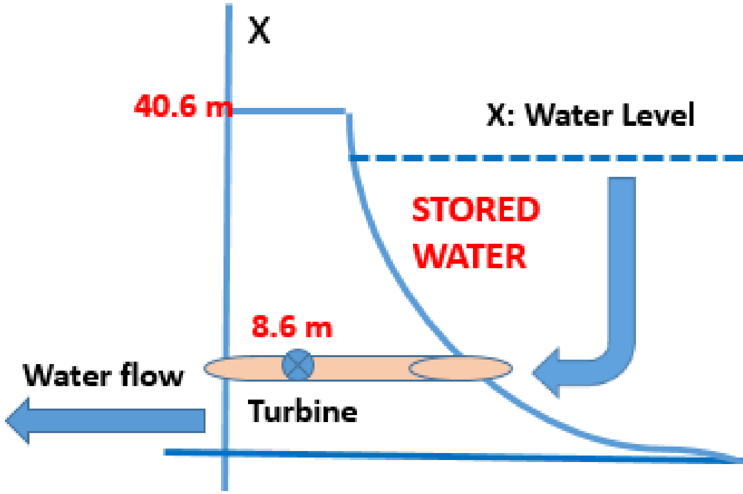


EXAMPLE: Suppose that the depth of water, measured in meters, behind a dam is described by an exponential random variable with pdf:

$$f_X(x) = \begin{cases} \frac{1}{13.5} \; e^{\frac{-x}{13.5}} & x > 0 \\ 0 & o.w \end{cases}$$

There is an emergency overflow at the top of the dam, which prevents the depth from exceeding 40.6 m. There is a pipe placed 32.0 m below the overflow that feeds water to a hydroelectric generator (turbine).

a- What is the probability that water is wasted though emergency overflow?
b- What is the probability that water will be too low to produce power?
c- Given that water is not wasted in overflow, what is the probability that the generator will have water to derive it?



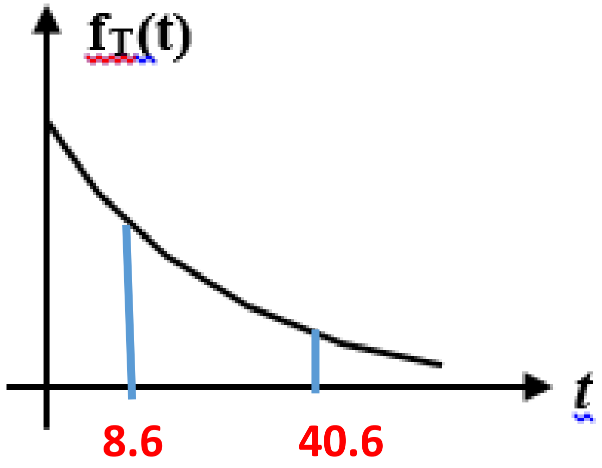
SOLUTION:

a- P(water wasted through emergency) = $P(X \geq 40.6\ m) = \int_{40.6}^{\infty} \frac{1}{13.5} \; e^{\frac{-x}{13.5}} \, dx \;=\; e^{-3}$

b- P(water too low to produce power) = $P(X \leq 8.6m) = \int_{0}^{8.6} \frac{1}{13.5} \; e^{\frac{-x}{13.5}} \, dx \;=\; \left(1 - e^{-0.637}\right) = 0.47$

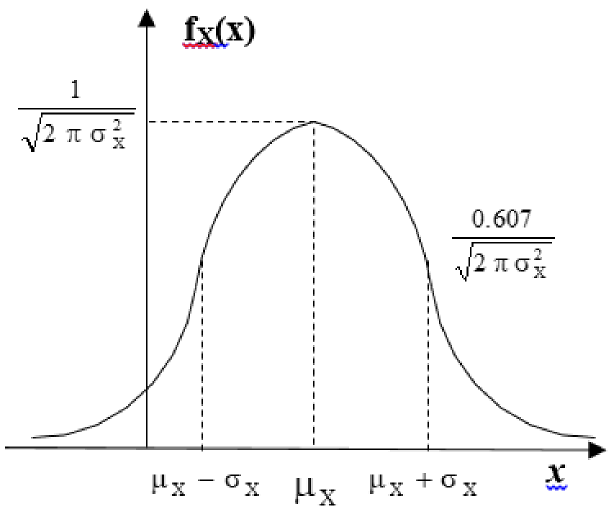
c- P(generator has water to derive it / water is not wasted)

$$P(X > 8.6 \,|\, X < 40.6) = \frac{P(X > 8.6 \cap X < 40.6)}{P(X < 40.6)} = \frac{P(8.6 < X < 40.6)}{P(X < 40.6)} = \frac{\int_{8.6}^{40.6} \frac{1}{13.5} \; e^{\frac{-x}{13.5}} \, dx}{\int_{0}^{40.6} \frac{1}{13.5} \; e^{\frac{-x}{13.5}} \, dx} = 0.504$$



Gaussian (Normal) Distribution

Definition: A random variable X is said to have a Gaussian (normal) distribution with a mean value μ_X and variance σ_X^2 if X has a continuous distribution for which the pdf $f_X(x)$ is:

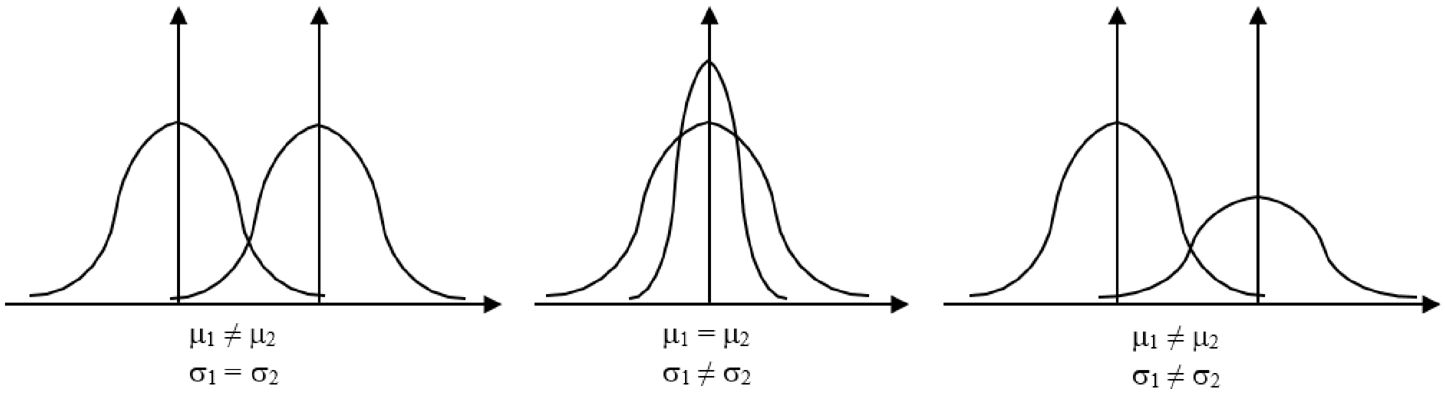


$$f_X(x) = \frac{1}{\sqrt{2\ \pi\ \sigma_X^2}} e^{\frac{-(x\ -\ \mu_X)^2}{2\ \sigma_X^2}} \ ; \quad -\infty < x < \infty$$

Note that:

$$E(x) = \mu_X = \int_{-\infty}^{\infty} x f_X(x) dx \qquad Var(x) = \sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx$$

$$P(x_1 \le X \le x_2) = \int_{x_1}^{x_2} \frac{1}{\sqrt{2\ \pi\ \sigma_X^2}} e^{\frac{-(x\ -\ \mu_X)^2}{2\ \sigma_X^2}} dx$$; cannot be evaluated in closed form, but can only be evaluated numerically. In this lecture, we will learn how to use standard tables to evaluated probabilities arising from various Gaussian distributions with different means and variances.



Standard Gaussian (Normal) Distribution

Definition: A normal random variable with mean zero and variance one is called a standard normal random variable. A standard normal random variable is denoted by Z and is obtained from any normal random variable X with mean μ_X and variance σ_X^2 through the linear transformation

$$Z = \frac{X - \mu_X}{\sigma_X}$$

Z can be written as $Z = \frac{X}{\sigma_X} - \frac{\mu_X}{\sigma_X}$ which is of the form $Z = aX + b$

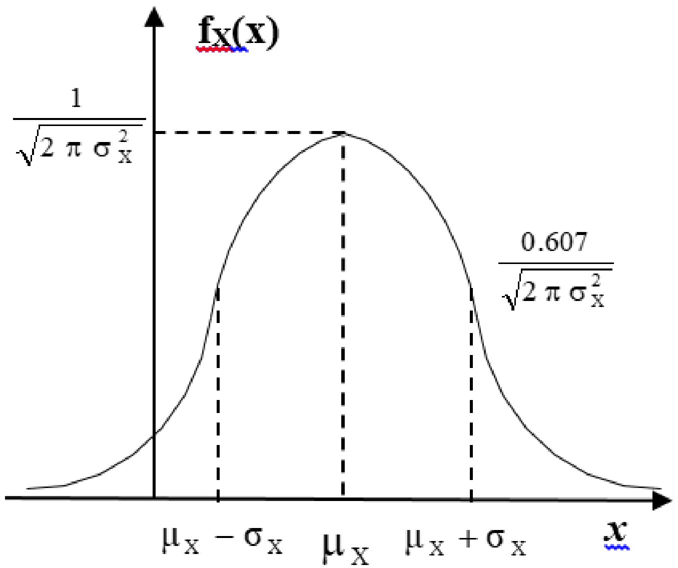
$$Mean(Z) = \mu_Z = E\{Z\} = aE\{X\} + b = \frac{1}{\sigma_X}\mu_X - \frac{1}{\sigma_X}\mu_X = 0$$

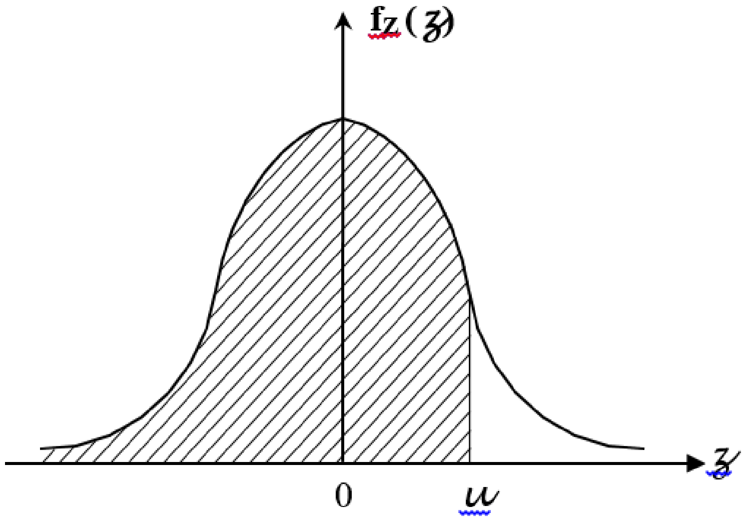
$$Var(Z) = \sigma_Z^2 = a^2\sigma_X^2 = \frac{1}{\sigma_X^2}\sigma_X^2 = 1$$

$$\phi(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{\frac{-(x-\mu_X)^2}{2\sigma_X^2}}; \quad -\infty < x < \infty$$

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}; -\infty < z < \infty$$





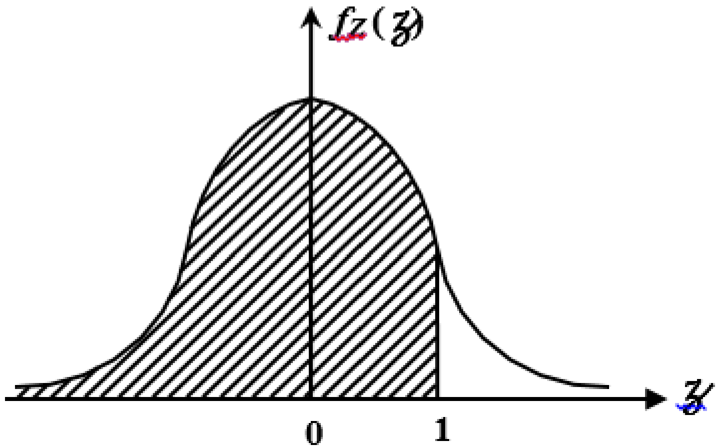
Definition: The Cumulative Distribution Function for the Standard Gaussian Distribution

The function $\phi(u) = P(Z \le u)$ is used to denote the cumulative distribution function of a standard normal random variable:

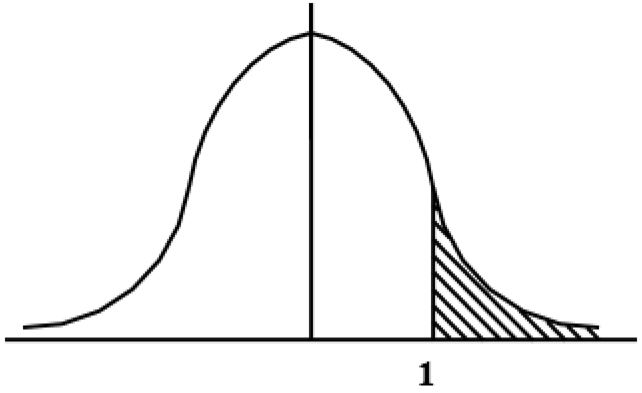
$$\phi(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{z^2}{2}}\, dz$$

This function is tabulated for $u \ge 0$

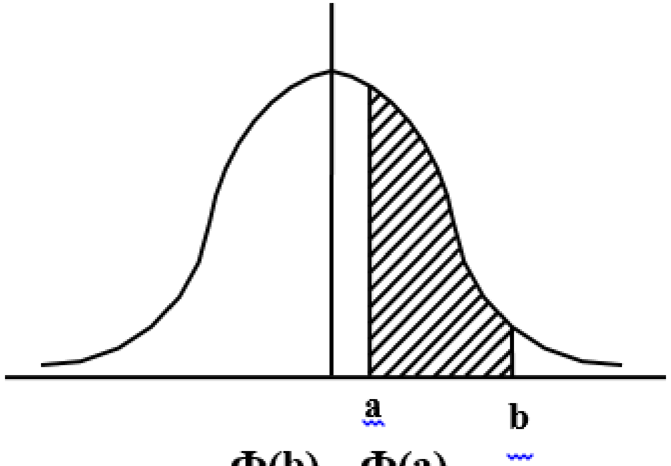
For $u < 0$; $\phi(u) = 1 - \phi(-u)$



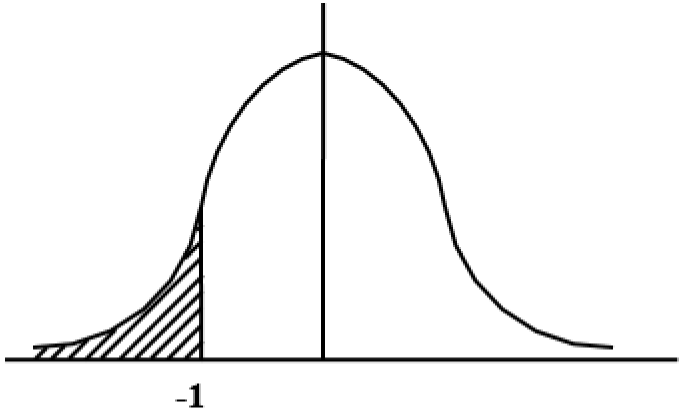
Φ(1)



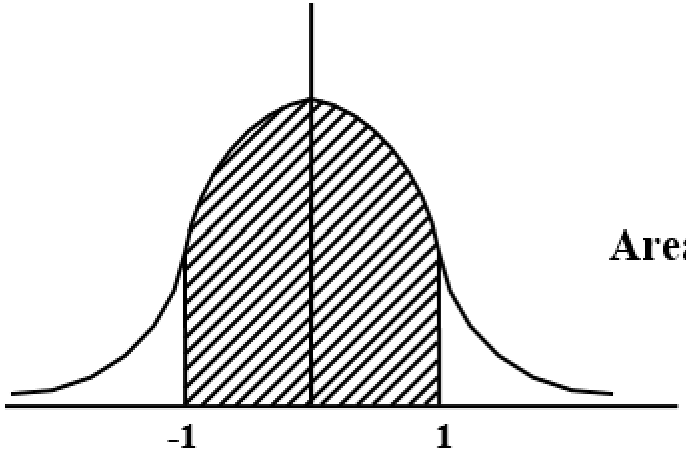
Area = 1 - Φ(1)



Φ(b) - Φ(a)



Φ(-1) = 1 - Φ(1)



Area = Φ(1) - Φ(-1)

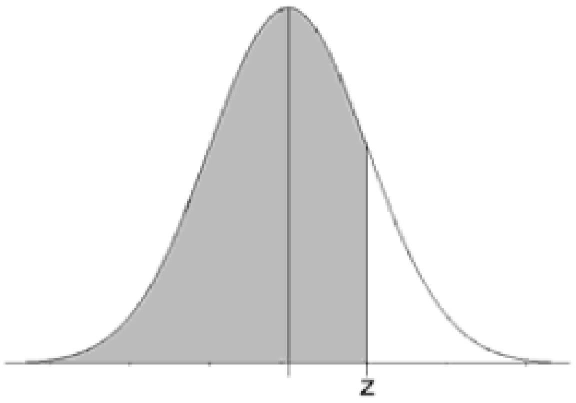
= Φ(1) – [1 – Φ(1)]

= 2Φ(1) – 1

Gaussian Distribution

Standard Normal Cumulative Probability Table

$$\Phi(z) = P(Z \leq z)$$



Cumulative probabilities for POSITIVE z-values are shown in the following table:

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621

Cumulative Distribution Function:

$$P(X \le x_0) = F_X(x_0) = \int_{-\infty}^{x_0} \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{\frac{-(x-\mu_X)^2}{2\sigma_X^2}} dx$$

$$\text{Let } Z = \left(\frac{X - \mu_X}{\sigma_X}\right) \Rightarrow dz = \frac{dx}{\sigma_X} \Rightarrow dx = \sigma_X du$$

$$F_X(x_0) = \int_{-\infty}^{\frac{x_0-\mu_X}{\sigma_X}} \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{\frac{-z^2}{2}} \sigma_X \, dz = \int_{-\infty}^{\frac{x_0-\mu_X}{\sigma_X}} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz$$

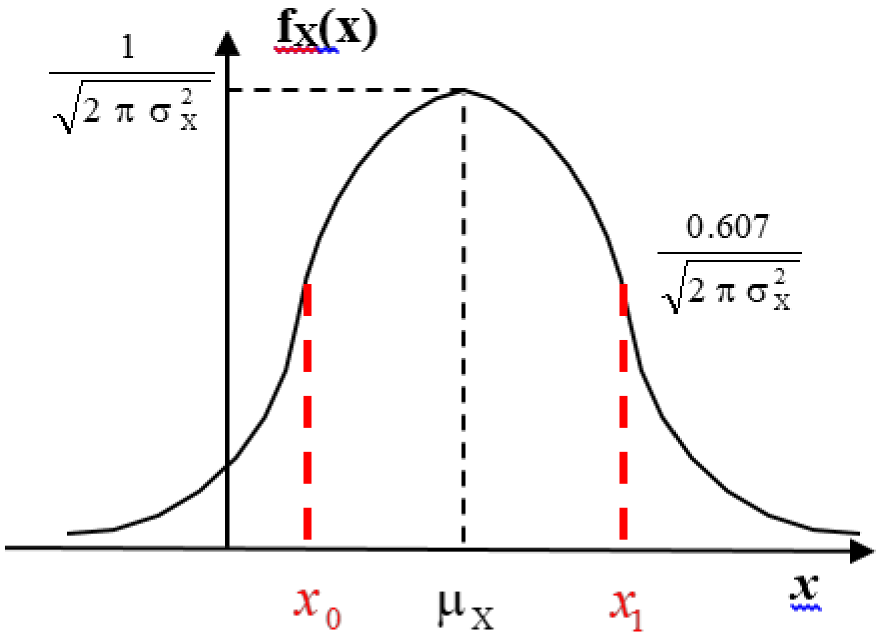
$$\Phi(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz$$

$$P(X \le x_0) = \Phi\left(\frac{x_0 - \mu_X}{\sigma_X}\right)$$

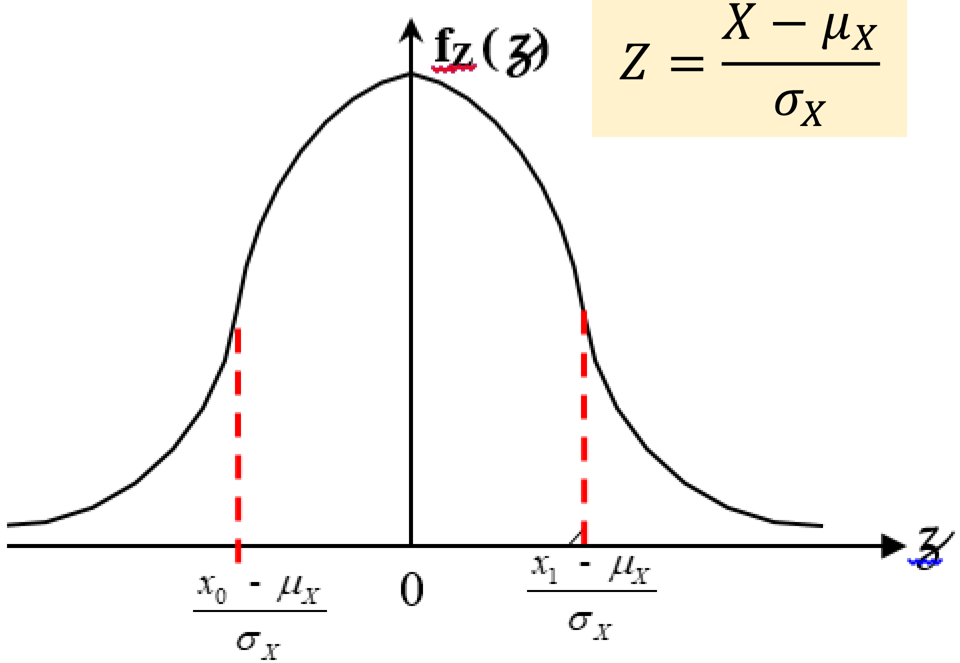
Therefore, we conclude that:

1- $P(X \le x_0) = \Phi\left(\frac{x_0 - \mu_X}{\sigma_X}\right)$

2- $P(x_0 \le X \le x_1) = \Phi\left(\frac{x_1 - \mu_X}{\sigma_X}\right) - \Phi\left(\frac{x_0 - \mu_X}{\sigma_X}\right)$



$$Z = \frac{X - \mu_X}{\sigma_X}$$

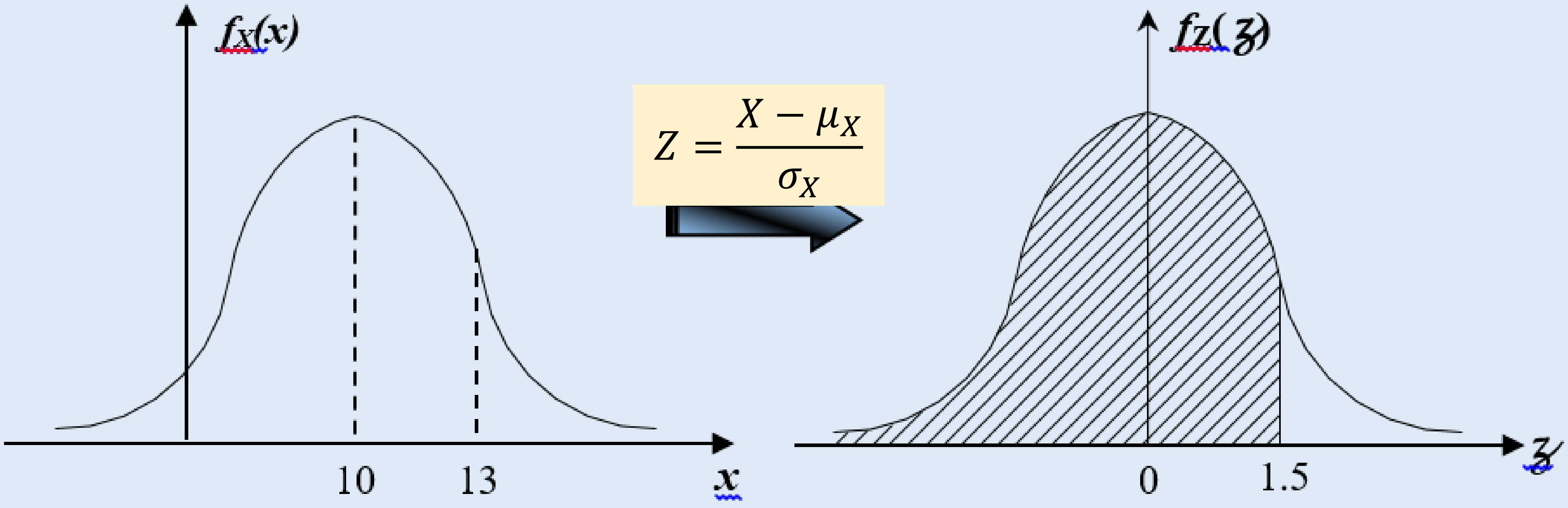


EXAMPLE: Suppose the current measurements in a strip of wire are assumed to follow a normal distribution with a mean of 10 mA and variance 4 $(mA)^2$. What is the probability that a measurement will exceed 13 mA?

SOLUTION: $X \sim N(\mu_X,\ \sigma_X^2) \Rightarrow N(10,\ 4)$

$$P(X \geq x_0) = 1 - P(X < x_0) = 1 - \Phi\left(\frac{x_0 - \mu_X}{\sigma_X}\right) \Rightarrow P(X > 13) = 1 - \Phi\left(\frac{13 - 10}{\sqrt{4}}\right)$$

$$P(X > 13) = 1 - \Phi\ (1.5) = 1\ -\ 0.93319\ =\ 0.06681$$



Gaussian Distribution

EXAMPLE

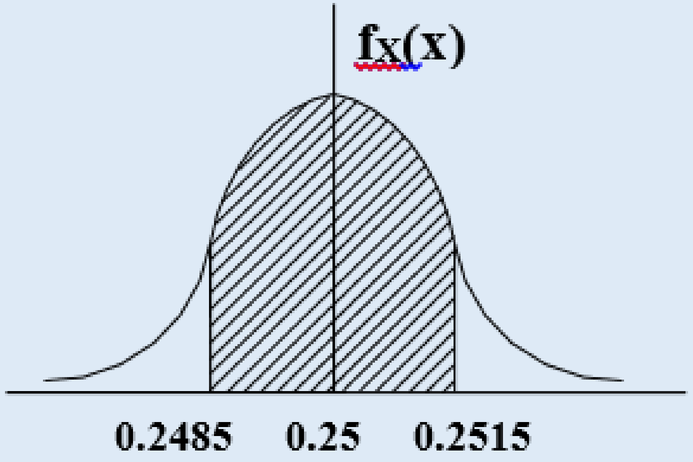
The diameter of a shaft in an optical storage drive is normally distributed with mean 0.25 inch and standard deviation of 0.0005 inch. The specifications on the shaft are 0.25 ± 0.0015 inch. What proportion of shafts conforms to specifications?

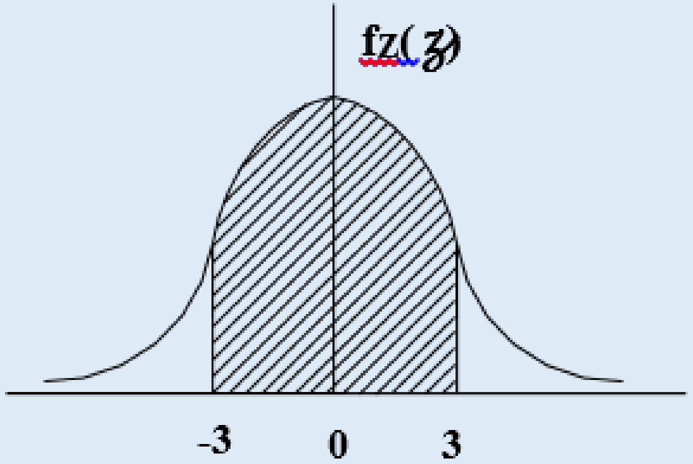
SOLUTION:

$$P(0.2485 < X < 0.2515) = \Phi\left(\frac{x_1 - \mu_X}{\sigma_X}\right) - \Phi\left(\frac{x_0 - \mu_X}{\sigma_X}\right)$$

$$= \Phi\left(\frac{0.2515 - 0.25}{0.0005}\right) - \Phi\left(\frac{0.2485 - 0.25}{0.0005}\right) = \Phi(3) - \Phi(-3)$$

$$= \Phi(3) - (1 - \Phi(3)) = 2\Phi(3) - 1 = 2(0.99865) - 1 = 0.9973$$





$$Z = \frac{X - \mu_X}{\sigma_X}$$

EXAMPLE: The tensile strength of paper is modeled by a normal distribution with a mean of 35 lb/in^2 and standard deviation of 2 lb/in^2.

a. If the specifications require the tensile strength to exceed 33 lb/in^2, what is the probability that a given sample will pass the specification test?

b. If 10 samples undergo the specification test, what is the probability that at least 9 will pass the test?

c. If 20 samples undergo the test, what is the expected number of samples that pass the test?

SOLUTION: First, we find the probability that one sample will pass the test

$$P(X>33) = 1-\Phi\left(\frac{x_0 - \mu_X}{\sigma_X}\right) = 1-\Phi\left(\frac{33-35}{2}\right) = 1-\Phi(-1) = 1-(1-\Phi(1)) = \Phi(1) = 0.8413$$

Next, let Y be the random variable representing the number of samples that pass the test out of the 10 tested samples.

$$P(Y=y) = \binom{n}{y} p^y (1-p)^{n-y}$$

Y follows the binomial distribution with parameters n = 10 and $p = 0.8413$

10

$$P(Y \geq 9) = \binom{10}{9}(p)^9(1-p)^1 + \binom{10}{10}(p)^{10}(1-p)^0 = 9(p)^9(1-p) + (p)^{10} = 0.4791$$

Finally, here m =20, p 0.8413 and $E(Y) = mp = 20(0.8413) = 16.626$

Gaussian Distribution

EXAMPLE

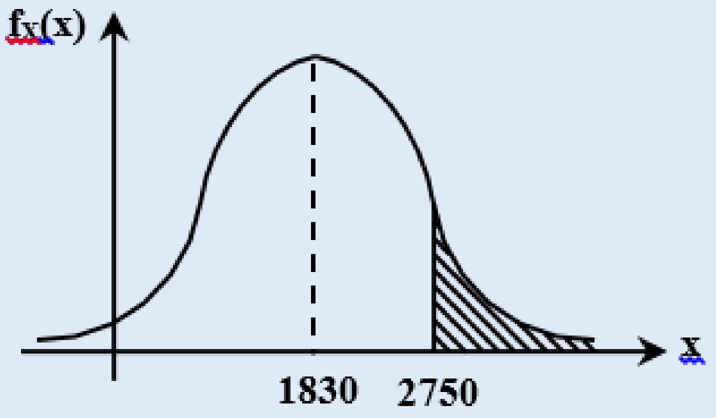
Assume that the height of clouds above the ground at some location is a Gaussian random variable (X) with mean 1830 m and standard deviation 460 m. find the probability that clouds will be higher than 2750 m.

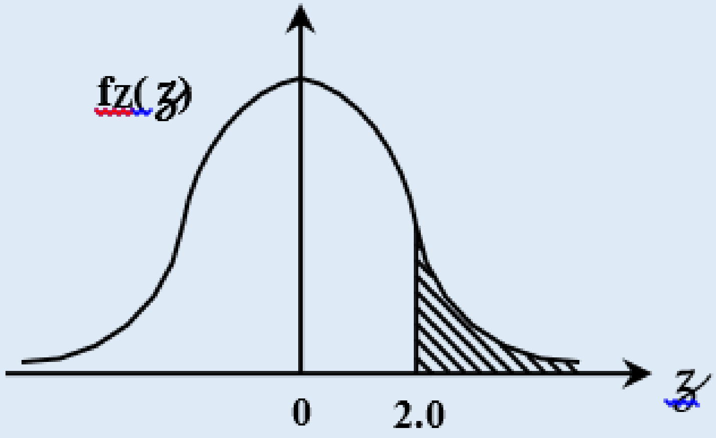
SOLUTION:

$$P(X > 2750) = 1 - \Phi\left(\frac{x_0 - \mu_X}{\sigma_X}\right)$$

$$1 - \Phi\left(\frac{2750 - 1830}{460}\right) = 1 - \Phi\ (2)$$

$$= 1 - 0.9772 = 0.0228$$





Normal Approximation of the Binomial and Poisson Distribution

De Moivre - Laplace Theorem: For large n, the binomial distribution with parameters n and p can be approximated by a Gaussian distribution with the same mean $n\,p$ and the same variance $n\,p(1-p)$.

$$\binom{n}{x} p^x (1-p)^{n-x} \simeq \frac{1}{\sqrt{2\pi np(1-p)}} e^{\frac{-(x-np)^2}{2np(1-p)}};$$ Here, $\simeq$ means asymptotically equal

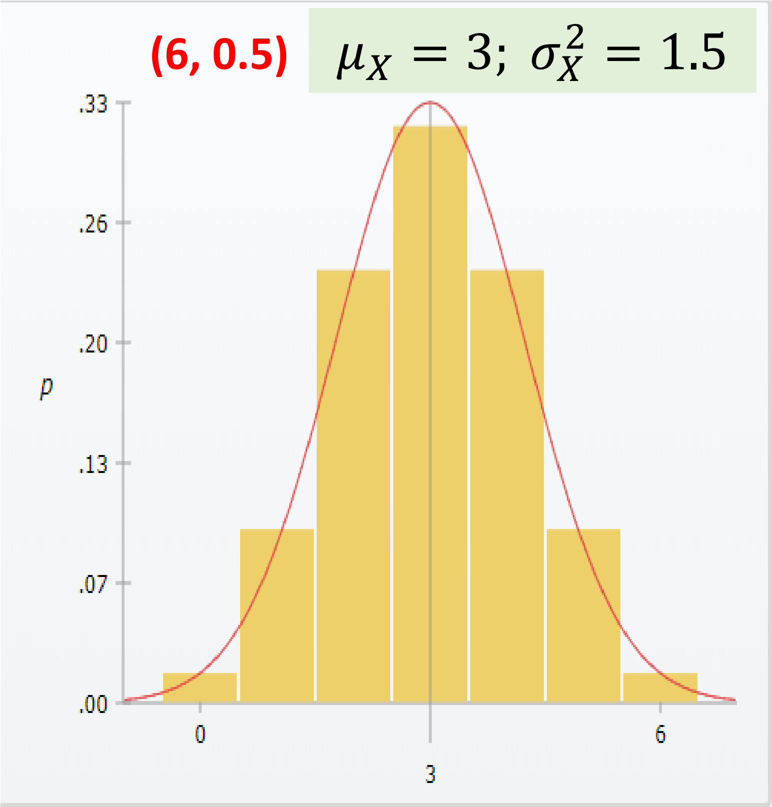
Remarks:

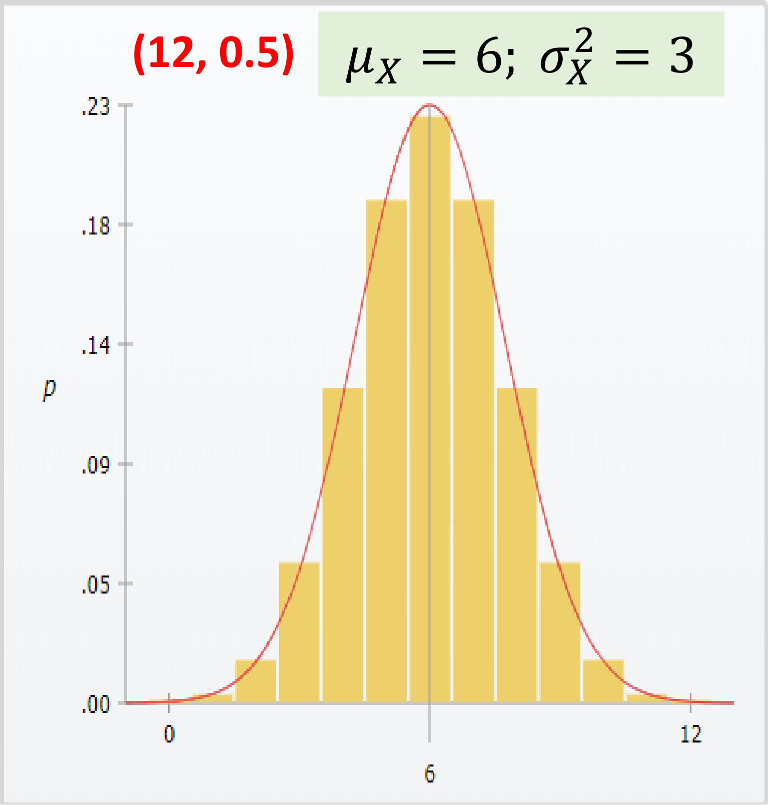
- The binomial distribution is discrete while the Gaussian distribution is continuous. Therefore, the above formula does not mean equality.
- The theorem gives good results when the mean $n\,p > 5$ and the variance $n\,p(1-p) > 5$.
- The theorem can be used to approximate binomial probabilities by their Gaussian probabilities.

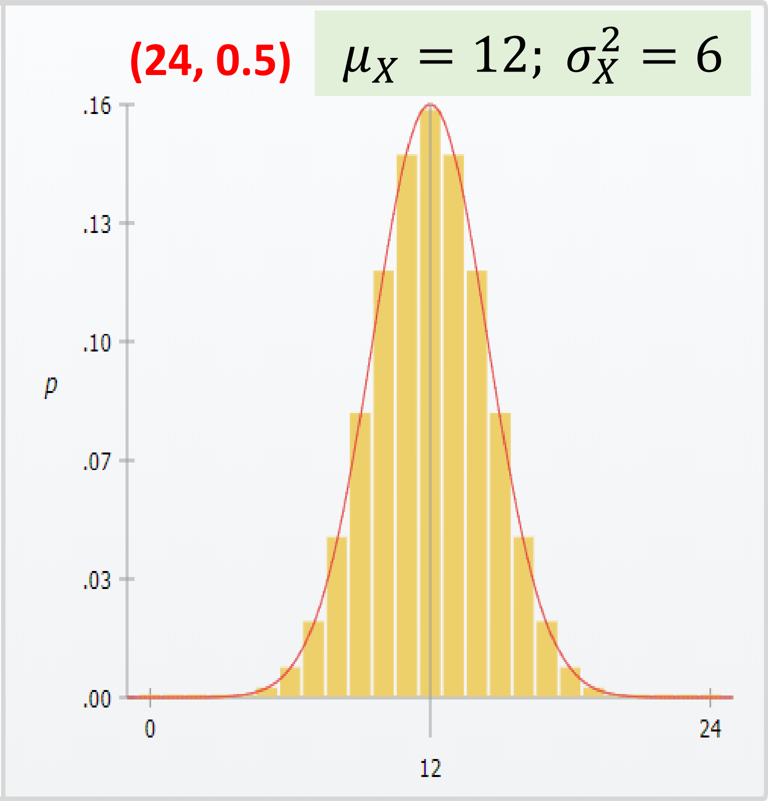
$$P(x_0 \le X \le x_1) = \sum_{x=x_0}^{x_1} \binom{n}{x} p^x (1-p)^{n-x} \sim \Phi\left(\frac{x_1 - \mu_X}{\sigma_X}\right) - \Phi\left(\frac{x_0 - \mu_X}{\sigma_X}\right)$$

Here, $\mu_X = np; \quad \sigma_X^2 = np(1-p)$.

Normal Approximation of the Binomial and Poisson Distribution







$$\mathrm{P(X = x)} = \binom{n}{x}\, p^x (1-p)^{n-x}$$

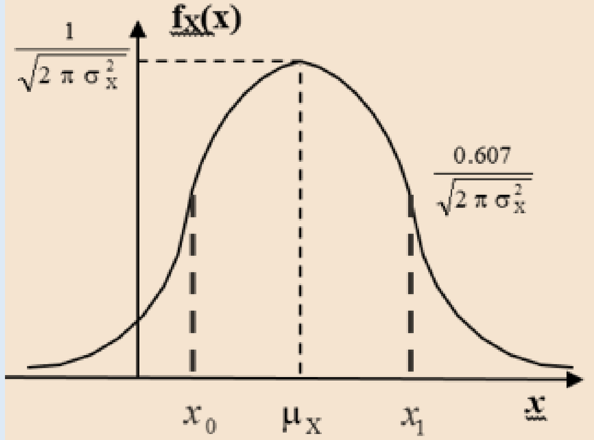
$$\mu_X = np; \quad \sigma_X^2 = np(1-p)$$

$$f_X(x) = \frac{1}{\sqrt{2\,\pi\,\sigma_X^2}}\, e^{\frac{-(x-\mu_X)^2}{2\,\sigma_X^2}}\ ; \quad -\infty < x < \infty$$

Source: https://digitalfirst.bfwpub.com/stats_applet/stats_applet_2_cltbinom.html

Normal Approximation of the Binomial and Poisson Distribution

EXAMPLE: Consider a binomial experiment with n = 50 and p = 0.2. If X is the number of successes, find the probability that $P(12 \le X \le 16)$.



SOLUTION:

Exact solution: $P(12 \le X \le 16) = \sum_{x=12}^{16} \binom{50}{x} (0.2)^x (1-0.2)^{50-x} = 0.2749$

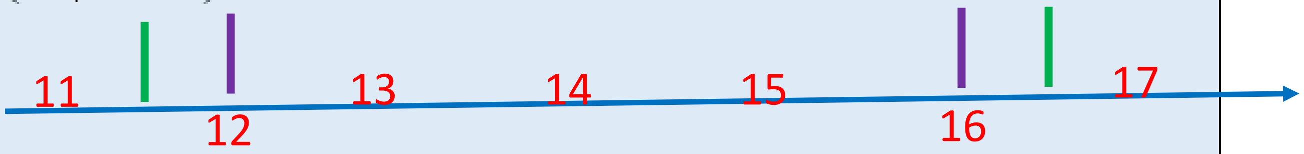
Applying the De Moivre - Laplace theorem, we can approximate the probability as:

$$\Phi(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz$$

Mean and Variance of X: $\mu_X = np = 50(0.2) = 10;$ $\sigma_X^2 = np(1-p) = 50(0.2)(0.8) = 8$, note that both are > 5.

$$P(12 \le X \le 16) = \Phi\left(\frac{x_1 - \mu_X}{\sqrt{\sigma_X^2}}\right) - \Phi\left(\frac{x_0 - \mu_X}{\sqrt{\sigma_X^2}}\right) = \Phi\left(\frac{16-10}{\sqrt{8}}\right) - \Phi\left(\frac{12-10}{\sqrt{8}}\right) = 0.2228$$

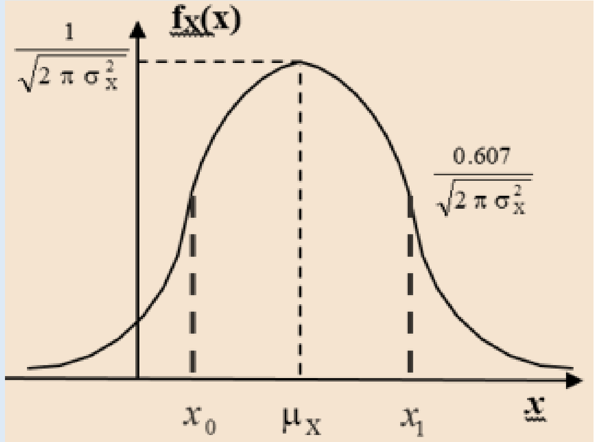
A more accurate result is:



$$P(11.5 \le X \le 16.5) = \Phi\left(\frac{x_1 - \mu_X}{\sqrt{\sigma_X^2}}\right) - \Phi\left(\frac{x_0 - \mu_X}{\sqrt{\sigma_X^2}}\right) = \Phi\left(\frac{16.5-10}{\sqrt{8}}\right) - \Phi\left(\frac{11.5-10}{\sqrt{8}}\right) = 0.2872$$

Normal Approximation of the Binomial and Poisson Distribution

EXAMPLE: Consider a binomial experiment with n = 1000 and p = 0.2. If X is the number of successes, find the probability that $P(X \leq 240)$.



SOLUTION:

Exact solution: $P(X \leq 240) = \sum_{x=0}^{240} \binom{1000}{x} (0.2)^x (1-0.2)^{1000-x} = 0.999141$

Applying the De Moivre - Laplace theorem, we can approximate the probability as:

Mean and Variance of X: $\mu_X = np = 1000(0.2) = 200; \quad \sigma_X^2 = np(1-p) = 1000(0.2)(0.8) = 160$, note that both are > 5.

$$\Phi(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz$$

$$P(X < 240) = \Phi\left(\frac{x_0 - \mu_X}{\sqrt{\sigma_X^2}}\right) = \Phi\left(\frac{240-200}{\sqrt{160}}\right) = \Phi(3.162) = 0.999216$$

Normal Approximation of the Binomial and Poisson Distribution

Theorem: If X is a Poisson distribution with mean $\mu_X = b$ and variance $\sigma_X^2 = b$, then when b is sufficiently large, the distribution can be approximated by a Gaussian distribution with the same mean $\mu_X = b$ and the same variance $\sigma_X^2 = b$.

$$e^{-b}\frac{b^x}{x!} \simeq \frac{1}{\sqrt{2\pi b}} e^{\frac{-(x-b)^2}{2b}};$$ Here, $\simeq$ means asymptotically equal

Remarks:

- The Poisson distribution is discrete while the Gaussian distribution is continuous. Therefore, the above formula does not mean equality.
- The theorem gives good results when the mean $b > 5$ and the variance $b > 5$.
- The theorem can be used to approximate Poisson probabilities by their Gaussian probabilities.

$$P(x_0 \le X \le x_1) = \sum_{x=x_0}^{x_1} e^{-b}\frac{b^x}{x\,!} \simeq \Phi\left(\frac{x_1 - b}{\sqrt{b}}\right) - \Phi\left(\frac{x_0 - b}{\sqrt{b}}\right)$$

Here, $\mu_X = b; \quad \sigma_X^2 = b$.

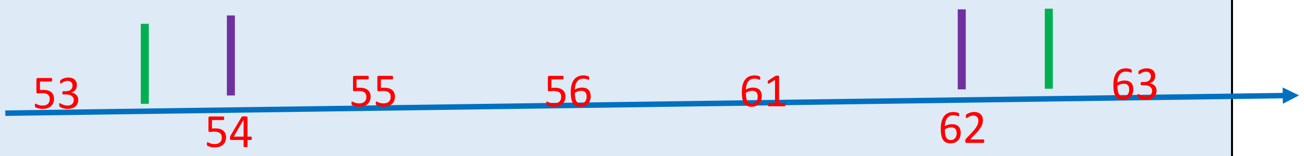
Normal Approximation of the Binomial and Poisson Distribution

EXAMPLE: Suppose cars arrive at a parking lot at a rate of 50 per hour according to a Poisson process. Compute the probability that the number of arriving cars in one hour will be between 54 and 62

Using the Poisson distribution

Using the normal approximation.

SOLUTION:



Exact solution:

$$P(X=x)=e^{-\lambda T}\frac{(\lambda T)^x}{x\,!}=e^{-50(1)T}\frac{(50)^x}{x\,!}\Rightarrow P(54\le X\le 62)=\sum_{54}^{62}e^{-50}\frac{(50)^x}{x\,!}=0.2616$$

Approximate Solution Using Normal Distribution:

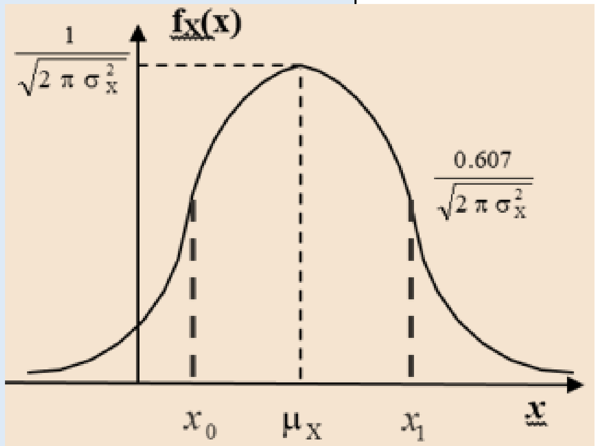
$$b=\lambda T\;\Rightarrow\;\mu_X=\lambda T=50;\;\sigma_X^2=\lambda T$$

$$P(x_0\le X\le x_1)\simeq\;\Phi\left(\frac{x_1-b}{\sqrt{b}}\right)-\Phi\left(\frac{x_0-b}{\sqrt{b}}\right)$$

$$P(54\le X\le 62)\simeq\;\Phi\left(\frac{62-50}{\sqrt{50}}\right)-\Phi\left(\frac{54-50}{\sqrt{50}}\right)=0.241$$

A more accurate result is obtained using

$$P(53.5\le X\le 62.5)\simeq\;\Phi\left(\frac{62.5-50}{\sqrt{50}}\right)-\Phi\left(\frac{53.5-50}{\sqrt{50}}\right)=0.2717$$



$$\Phi(u)=\int_{-\infty}^{u}\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}dz$$

Normal Approximation of the Binomial and Poisson Distribution

EXAMPLE: Assume the number of asbestos particles in one cm^3 of dust follow a Poisson distribution with a mean of 1000. If one cm^3 of dust is analyzed, what is the probability that less than 950 particles are found?

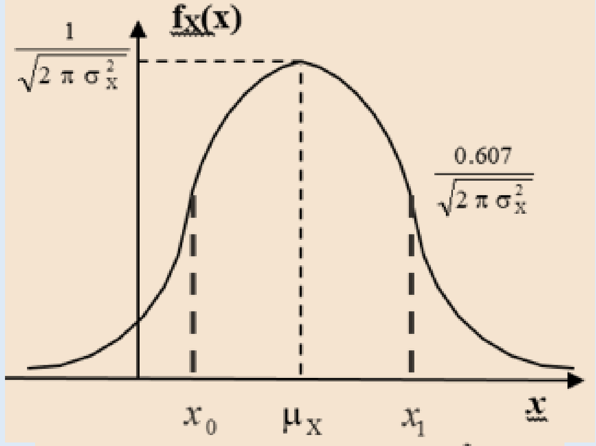
SOLUTION:

Exact solution: $P(X \le 950) = \sum_{x=0}^{950} e^{-1000} \frac{(1000)^x}{x\,!}$ = 0.0578

Approximate:

$$\mu_X = b = 1000; \; \sigma_X^2 = b = 1000$$

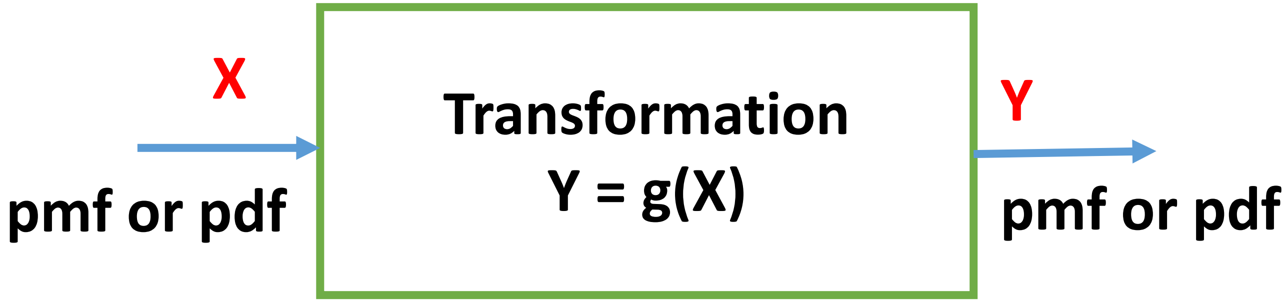
$$P(X \le x_1) \simeq \Phi\left(\frac{x_1 - b}{\sqrt{b}}\right) = \Phi\left(\frac{950 - 1000}{\sqrt{1000}}\right) = \Phi(-1.58) = 0.057$$



$$\Phi(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz$$

Transformation of Random Variables

- Let X be a random variable with a given pmf P(X=x), if discrete, or pdf $f_X(x)$, if continuous.
- Let Y = g(X) be a single-valued function of X, then Y is a random variable. The objective is to find its pmf/pdf



An Example on Discrete Random Variables: An intercom system master station provides power to three offices. The probability that any one office will be switched on and draw power at any time is 0.4. When on, an office draws 0.5 Watts

a. Find the pmf of the power drawn by the intercom system.

b. Find the mean value of the power delivered by the master station.

SOLUTION: Let X be the random variable representing the number of ON offices. X is binomial with parameters n=3 and p=0.4.

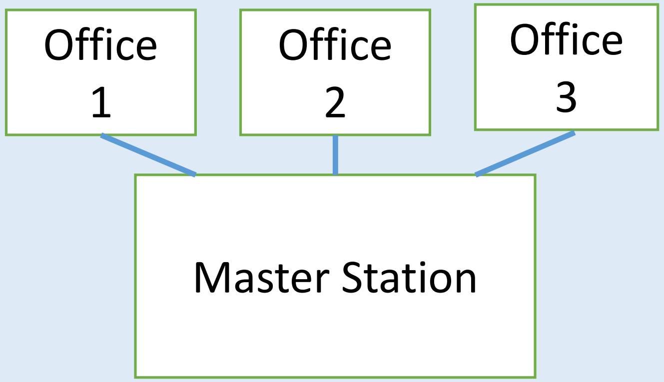
Let Y be the random variable representing the delivered power.

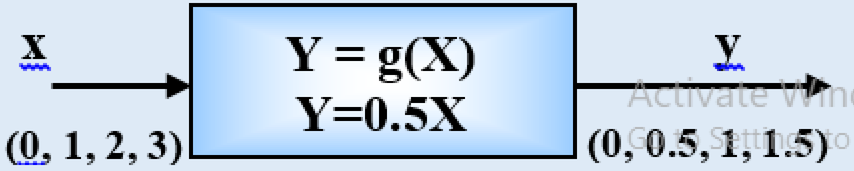
Y is related to X by **Y=0.5 X**. Note that: **P(Y = y) = P(X =x), where** $P(X=x)=\binom{n}{x} p^x (1-p)^{n-x}$

The table below shows the X and Y values and their probabilities.

$$E(Y)=E(0.5X)=0.5E(X)=0.5(3)(0.4)=0.6W$$

x	**y**	**P(X = x)**	**P(Y = y)**
0	0	$(1-p)^3$	$(1-p)^3$
1	0.5	$3\,p\,(1-p)^2$	$3\,p\,(1-p)^2$
2	1	$3\,p^2\,(1-p)$	$3\,p^2\,(1-p)$
3	1.5	p^3	p^3





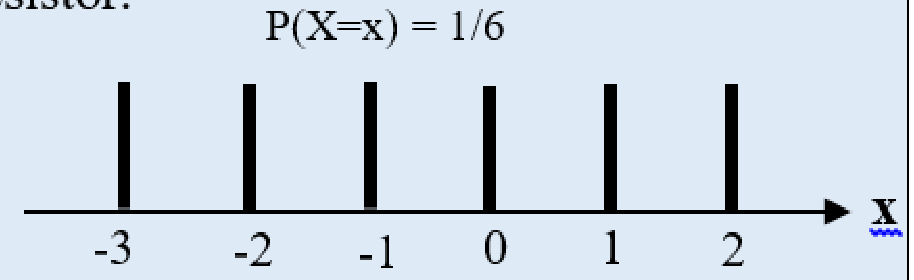
EXAMPLE: The DC current X that flows through a $1-\Omega$ resistor R is a discrete random variable with the following pmf: $P(X=x)=1/6$; X = -3 , -2 , -1 , 0 , 1 , 2.

a. Find the pmf of the power in the resistor defined as $Y=g(X)=RX^2=X^2$.

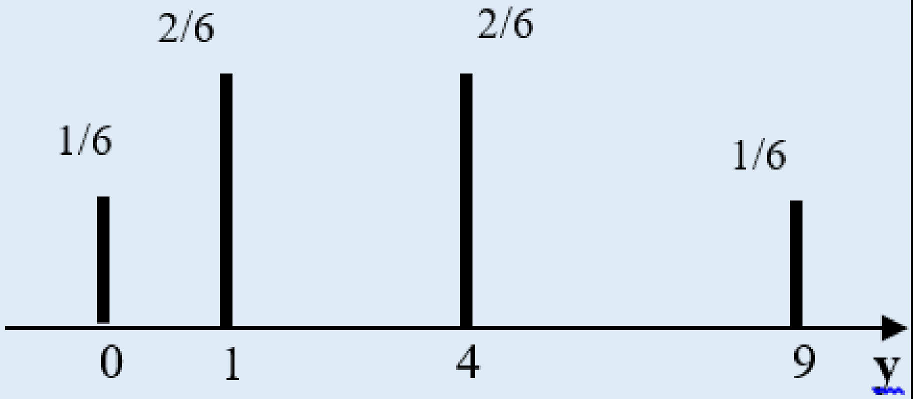
b. Find the average power dissipated in the resistor.

SOLUTION:

x	y	P(X = x)	P(Y = y)
-3	9	1/6	1/6
-2	4	1/6	1/6
-1	1	1/6	1/6
0	0	1/6	1/6
1	1	1/6	1/6
2	4	1/6	1/6



Probability Mass Function P(X)



Probability Mass Function P(Y)

The distribution of Y is:

P(Y = 0) = 1/6 P(Y = 1) = 2/6
P(Y = 4) = 2/6 P(Y = 9) = 1/6

Average power = $E(Y)=(0)\left(\frac{1}{6}\right)+(1)\left(\frac{2}{6}\right)+(4)\left(\frac{2}{6}\right)+(9)\left(\frac{1}{6}\right)=\left(\frac{19}{6}\right)W$

Transformation of Random Variables

Continuous Case:

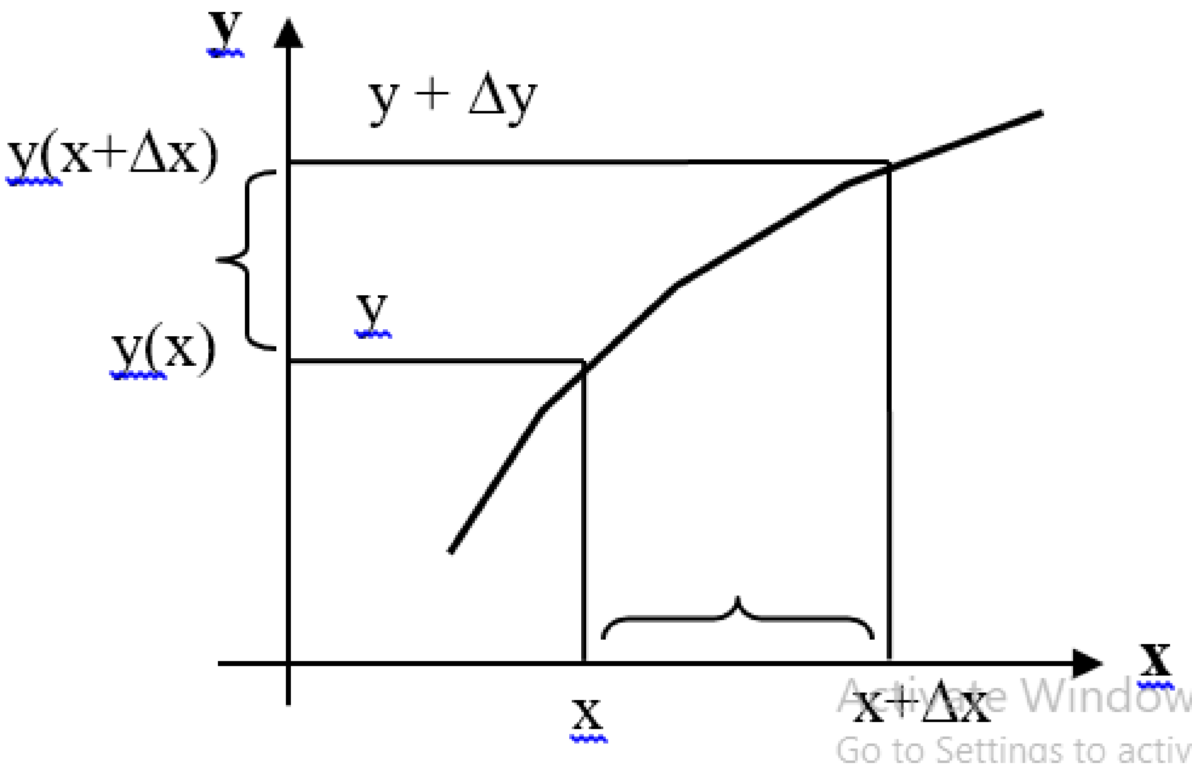
Let Y = g(X) be a monotonically increasing or decreasing function of (x).

$$P(x < X < x + \Delta x) = P\{y(x) < Y < y(x + \Delta x)\}$$

$$P(x < X < x + \Delta x) = P\{y < Y < y + \Delta y\}$$

$$f_Y(y)\Delta y = f_X(x)\Delta x$$

$$f_Y(y) = f_X(x)\left|\frac{\Delta x}{\Delta y}\right| = \frac{f_X(x)}{\left|\frac{\Delta y}{\Delta x}\right|} = \frac{f_X(x)}{\left|\frac{dy}{dx}\right|}; y_1 < y < y_2$$



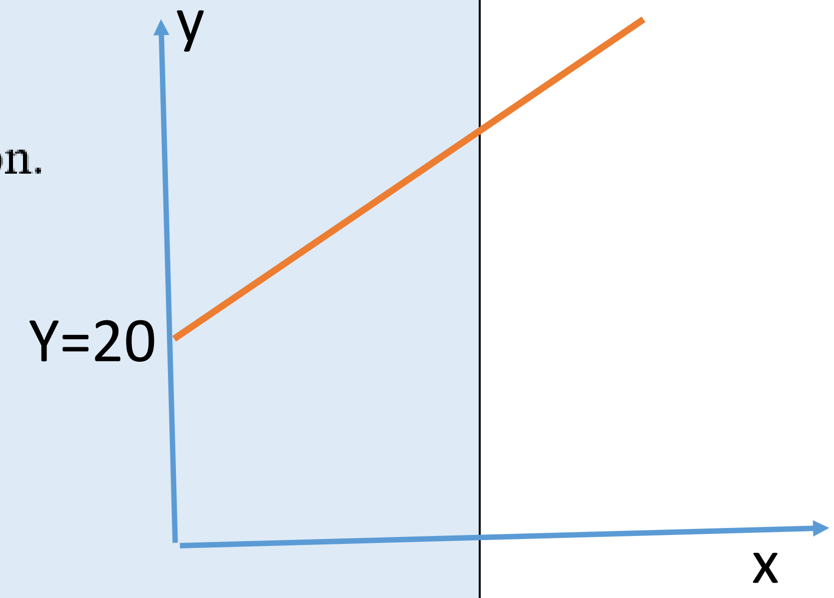
$$P(\mathrm{x} \le X \le \mathrm{x} + \Delta x) = \int_{x}^{\mathrm{x}+\Delta x} f_X(u)du = f_X(x)\Delta x$$

$$P(\mathrm{y} \le \mathrm{Y} \le \mathrm{y} + \Delta \mathrm{y}) = \int_{x}^{\mathrm{y}+\Delta y} f_Y(u)du = f_Y(y)\Delta y$$

EXAMPLE: The amount, in dollars, charged by a technician is related to the time X, in hours, needed to complete a task by the formula Y = 10X + 20. The time, X, needed to complete a task is a random variable which follows the exponential distribution $f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x < 0 \end{cases}$

a. Find $f_Y(y)$ and the region over which it is defined.

b. Find the mean value of Y.



SOLUTION: Y = 10X + 20; This is a monotonically increasing function.

$$f_Y(y) = \frac{f_X(x)}{|dy/dx|} \quad ; \quad \left|\frac{dy}{dx}\right| = 10|$$

$$f_Y(y) = \frac{f_X(x)}{10} \;,\; but\ x = \frac{y - 20}{10} \quad \Rightarrow \quad f_Y(y) = \frac{f_X\left(\frac{y-20}{10}\right)}{10}$$

$$f_Y(y) = \begin{cases} \frac{\lambda}{10} e^{-\lambda(\frac{y-20}{10})} & \frac{y-20}{10} > 0 \\ 0 & \frac{y-20}{10} < 0 \end{cases} \quad \rightarrow \quad f_Y(y) = \begin{cases} \frac{\lambda}{2} e^{-\lambda\left(\frac{y-20}{10}\right)} & y > 20 \\ 0 & y < 20 \end{cases}$$

Mean value of Y: $E(Y) = E(10X + 20) = 10E(X) + 20 = \frac{10}{\lambda} + 20$

EXAMPLE: The profit Y of a manufacturing plant is related to the demand X by the relationship $Y = aX + b$. Let X be a Gaussian r.v with mean μ_X variance σ_X^2. Find $f_Y(y)$.

SOLUTION: $Y = aX + b$ is a monotonic function.

$$f_Y(y) = \frac{f_X(x)}{|dy/dx|}; \qquad \left|\frac{dy}{dx}\right| = |a|; \qquad x = \frac{y-b}{a}$$

$$f_Y(y) = \frac{1}{a}\frac{1}{\sqrt{2\pi\sigma_X^2}} e^{\frac{-(x-\mu_X)^2}{2\sigma_X^2}} = \frac{1}{\sqrt{2\pi(a\sigma_X)^2}} e^{\frac{-(\frac{y-b}{a}-\mu_X)^2}{2\sigma_X^2}}$$

$$= \frac{1}{\sqrt{2\pi(a\sigma_X)^2}} e^{\frac{-(y-(b+a\mu_X))^2}{2(a\sigma_X)^2}} = \frac{1}{\sqrt{2\pi\sigma_Y^2}} e^{\frac{-(y-\mu_Y)^2}{2\sigma_Y^2}}$$

Therefore, $Y = aX + b$ **is Gaussian with mean** $\mu_Y = a\mu_x + b$ **and variance** $\sigma_Y^2 = a^2 \sigma_X^2$

Result: A linear transformation of a Gaussian random variable is also Gaussian

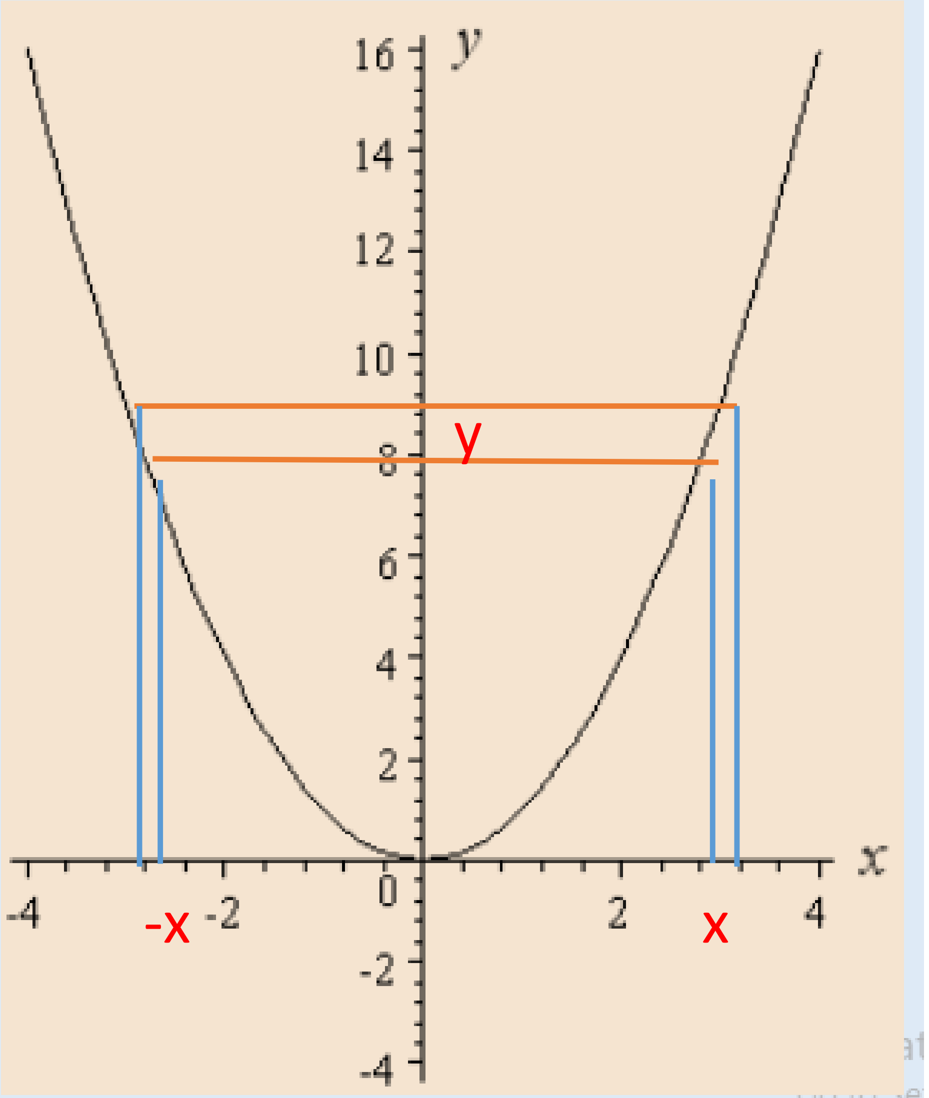
EXAMPLE: Let (X) be a Gaussian r.v with mean 0 variance 1. Define $Y = X^2$. Find $f_Y(y)$

SOLUTION: From the figure, we note that

$$P(y < Y < y + \Delta y) = 2P(x < X < x + \Delta x)$$

$$f_Y(y)\Delta y = 2 f_X(x)\Delta x$$

$$f_Y(y) = 2 f_X(x)\left|\frac{\Delta x}{\Delta y}\right| = \frac{2f_X(x)}{\left|\frac{\Delta y}{\Delta x}\right|} = \frac{2f_X(x)}{\left|\frac{dy}{dx}\right|}; y \geq 0$$



Here, $y = x^2$; $\left|\frac{dy}{dx}\right| = |2x|$

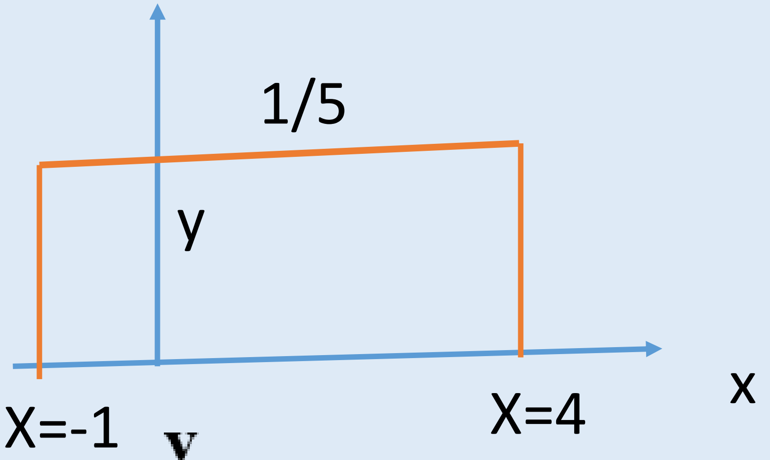
$$f_Y(y) = \frac{2}{|2x|}\frac{1}{\sqrt{2\ \pi}}e^{\frac{-(x)^2}{2}} \quad , x = \sqrt{y}$$

$$f_Y(y) = \frac{1}{\sqrt{y}}\frac{1}{\sqrt{2\pi}}e^{\frac{-y}{2}}$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi y}}e^{\frac{-y}{2}} \quad ; \quad y \geq 0$$

EXAMPLE: Let (X) be a uniform r.v in the interval (-1, 4). If Y = X^2. Find $f_Y(y)$

SOLUTION: $$f_X(x) = \begin{cases} \frac{1}{5} & -1 \le x \le 4 \\ 0 & otherwise \end{cases}$$



For $(-1 \le X \le 1)$, $f_Y(y) = 2\dfrac{f_X(x)}{|dy/dx|} = \dfrac{2 \times 1/5}{|2x|} = \dfrac{1}{5\sqrt{y}}$

For $(1 < X \le 4)$, $f_Y(y) = \dfrac{f_X(x)}{|dy/dx|} = \dfrac{1/5}{|2x|} = \dfrac{1}{10\sqrt{y}}$

$$f_Y(y) = \begin{cases} \dfrac{1}{5\sqrt{y}} & 0 < y \le 1 \\ \dfrac{1}{10\sqrt{y}} & 1 < y \le 16 \\ 0 & Otherwise \end{cases}$$