CHAPTER 3
DISCRETE PROBABILITY DISTRIBUTION

L 1 – Binomial and Poisson

DISCRETE PROBABILITY DISTRIBUTION

Learning Objectives:

:

At the end of the lecture, you will be able to :

- select an appropriate discrete probability distribution
  * binomial distribution or
  * poisson distribution

to calculate probabilities in specific application

- calculate the probability, means and variance for each of the discrete distributions presented

BINOMIAL DISTRIBUTION Bin(n,p)

<u>Bernoulli Trials</u>: experiment with two possible outcomes, either 'Success' or 'failure'.

Probability of success is given as p and probability of failure is 1- p

Requirements of a binomial experiment:

* n Bernoulli trials

* trials are independent

* that each trial have a constant probability p of success.

Example binomial experiment: tossing the same coin successively and independently n times

BINOMIAL DISTRIBUTION Bin(n,p)

A binomial random variable X associated with a binomial experiment consisting of n trials is defined as:

X = the number of 'success' among n trials

The probability mass function of X is

$$\boxed{f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \; x = 0, 1, \cdots, n}$$

Mean and Variance: $$\boxed{E(X) = np \;, V(X) = np(1-p)}$$

A random variable that has a binomial distribution with parameters n and p, is denoted by $X \sim \text{Bin}\,(n,p)$

Example 1:

In the experiment of tossing a fair coin for 10 times, X is a random variable, the number of head.

(a) What is the pmf of X ?
(b) Find the probability the head will appear exactly 5 times
(c) What is the probability no head?
(d) Find the mean and the variance of X.

Solution:

Example 1:

In the experiment of tossing a fair coin for 10 times, X is a random variable, the number of head.

(a) What is the pmf of X ?
(b) Find the probability the head will appear exactly 5 times
(c) What is the probability no head?
(d) Find the mean and the variance of X.

Solution:

(a) $P(X = x) = \binom{n}{x} p^x q^{n-x} \qquad x = 0, 1, 2, \ldots\ldots n$

(b) $P(X = 5) = \binom{10}{5} (0.5)^5 (0.5)^5 = 0.246$

(b) $P(X = 0) = \binom{10}{0} (0.5)^0 (0.5)^{10} = 0.0009756$

(d) $E(X) = np = 5, \quad Var(X) = npq = 2.5$

Example 2:

$Let\ X \sim Bin(8,\ 0.45)$

Find

$(a) P(X = 2)$

$(b) P(X \geq 1)$

$(c) Mean\ X\ and\ \mathrm{var}(X)$

Solution :

$$Let\ X \sim Bin(8,0.45)$$

$$\Rightarrow P(X = x) = \binom{8}{x}(0.45)^x(0.55)^{8-x}\ , x = 0, 1, 2, \ldots 8$$

$$(a) P(X = 2) = \binom{8}{2}(0.45)^2(0.55)^6 = 0.157$$

$$(b) P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0)$$

$$= 1 - \binom{8}{0}(0.45)^0(0.55)^8 = 1 - 0.00837 = 0.992$$

$$(c) Mean\ X = np = 3.6\ and\ \mathrm{var}(X) = npq = 1.98$$

Poisson Probability Distribution

Conditions to apply the Poisson Probability distribution ar

1. x is a discrete random variable
2. The occurrences are random
3. The occurrences are independent

Useful to model the number of times that a certain event occurs per unit of time, distance, or volume. Examples of application of Poisson probability distribution

i) The number of telephone calls received by an office during a given day

ii) The number of defects in a five-foot-long iron rod.

Poisson Probability Distribution, $X \sim P(\lambda)$

The probability of x occurrences in an interval is

$$f(x) = P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \quad x = 0,1,2,3,\ldots$$

where λ is the mean number of occurrences in that interval.
(per unit time or per unit area)

Mean and Variance:

$$E(X) = V(X) = \lambda$$

Poisson Process

- Carry out experiment to estimate λ that represents the mean number of events that occur in one unit time/ space

* The number of events *X* that occur in *t* units of time is counted and λ is estimated.

* If the number of events are independent and events cannot occur simultaneously, then *X* follows a Poisson distribution. A process that produces such events is a **Poisson process**

Let λ denote the mean number of events that occur in one unit of time.Let N_T denote the number of events that are observed to occur in T units of time or space, then
$N_T \sim P\ (\lambda T)$,

Example 1

Anne's answering machine receives about 6 calls between 8am to 10am. What is the probability Anne will receives At least 1 call in the next 20 minutes?

Solution:

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

Given for 120 min. 6 calls. It means that for 20 minutes she will receives 1 call.
So the $\lambda = 1$.

$$\text{Hence, } P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{e^{-1}1^{0}}{0!}$$
$$= 1 - 0.368 = 0.632$$

Example 2

Suppose X has a Poisson distribution a mean of 5. Determine the following.

(a) $P(X = 0)$;

(b) $P(X < 3)$;

(c) $P(X \geq 4)$

(d) What is the mean and the standard deviation of X?

Example 2

Suppose X has a Poisson distribution a mean of 5.

Determine the following.

(a) P(X = 0); (b) P(X < 3); (c) $P(X \geq 4)$

(d) What is the mean and the standard deviation of X?

Solution:

(a) $P(X=0)=\dfrac{5^0 e^{-5}}{0!}=e^{-5}=0.00673$

(b) $P(X<3)=P(X=0)+P(X=1)+P(X=2)$

$$=\frac{5^0 e^{-5}}{0!}+\frac{5^1 e^{-5}}{1!}+\frac{5^2 e^{-5}}{2!}=0.125$$

(c) $P(X \geq 4)=1-P(X \leq 3)=0.735$

(d) $E(X)=Var(X)=5$

Exercise

At the Mc Donald drive-thru window of food establishment, it was found that during slower periods of the day, vehicles visited at the rate of 12 per hour. Determine the probability that

(a) no vehicles visiting the drive-thru within a ten-minute interval during one of these slow periods;

(b) only 2 vehicles visiting the drive-thru within a ten-minute interval during one of these slow periods; and

(c) at least three vehicles visiting the drive-thru within a ten-minute interval during one of these slow periods.

CHAPTER 3
CONTINUOUS PROBABILITY DISTRIBUTION

L1- Exponential distribution

L2- Normal and Nonstandard Normal Distributions

Exponential and Normal Distribution

Learning Objectives:

At the end of the lecture, you will be able to :

- select an appropriate continuous probability distribution to calculate probabilities in specific application

- calculate the probability, means and variance for each of the continuous distributions presented

CONTINUOUS PROBABILITY DISTRIBUTION

The probability functions are described by formulas that depend on some parameter values:

The expectations and variances of the distributions are also specified in terms of these parameters.

Common continuous distributions: uniform distribution, exponential distribution, gamma distribution, Weibull distribution, beta distribution and the normal distribution.

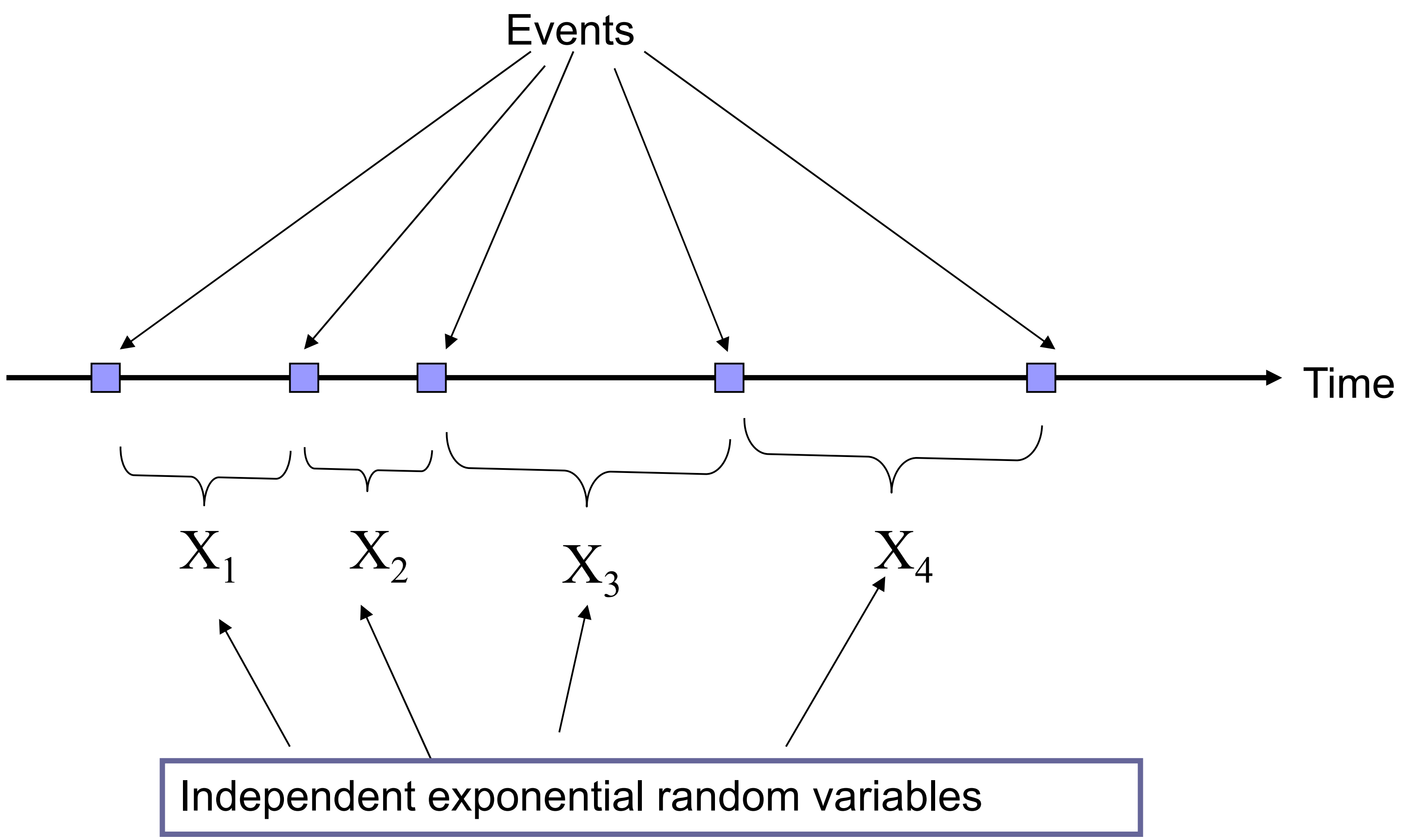
EXPONENTIAL DISTRIBUTION

The exponential distribution has a state space $x \geq 0$

The exponential distribution is often used to model failure or waiting times and inter arrival times.

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 .

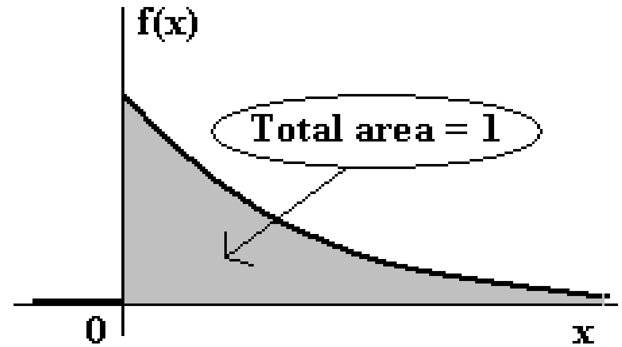
EXPONENTIAL DISTRIBUTION



EXPONENTIAL DISTRIBUTION

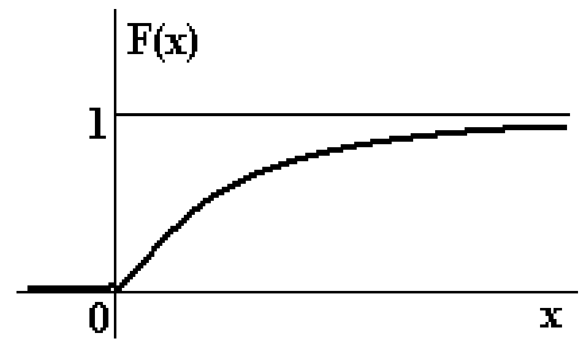
- Density function of exponential distribution

$$f(x)=\begin{cases}\lambda e^{-\lambda x} & , \quad x\geq 0\\ 0 & , \quad otherwise\end{cases}$$



- Cumulative Distribution Function

$$F(x)=\begin{cases}0, & x<0,\\ 1-e^{-\lambda x} & x\geq 0\end{cases}$$



- Mean and Variance

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

Example 1:

Let *X* is the number of individuals failing in a large group and has a exponential distribution. If we assume that the mean of *X* is 10, what is the probability that more than 20 will fail at the same time?

Solution:

$$f(x) = \begin{cases} (0.1)e^{-0.1x} & , x \geq 0 \\ 0 & , \text{otherwise} \end{cases}$$

$$\text{(a)} \;\; P(x > 20) = \int_{20}^{\infty} (0.1)e^{-0.1x} dx = -e^{-0.1x}\Big|_{20}^{\infty}$$

$$= e^{-2} = 0.135$$

Example 2:

Let X be an exponential random variable with $\lambda = 0.05$ Calculate the following probabilities:

a. $P(X < 50)$

b. $P(x > 60)$

c. $P(50 < x < 60)$

d. What is the mean and the variance of X.

$$f(x) = \begin{cases} (0.05)e^{-0.05x} & , x \geq 0 \\ 0 & , \text{otherwise} \end{cases}$$

$$\text{(a)} \;\; P(X \langle 50) = \int_0^{50} (0.05)e^{-0.05x} dx = -e^{-0.05x} \Big|_0^{50}$$

$$= 1 - e^{-2.5} = 1 - 0.082 = 0.918$$

$$(b)\ P(x>60)=\int_{60}^{\infty}(0.05)e^{-0.05x}dx=-e^{-0.05x}\Big|_{60}^{\infty}$$

$$=e^{-3}=0.05$$

$$(c)\ P(50<x<60)=\int_{50}^{60}(0.05)e^{-0.05x}dx=-e^{-0.05x}\Big|_{50}^{60}$$

$$=e^{-2.5}-e^{-3}=0.082-0.05=0.032$$

$$(d)\ E(x)=\frac{1}{0.05}=20,\quad V(x)=\frac{1}{(0.05)^2}=400$$

Exercise:

The lifetime of a certain electronic component is known to be exponentially distributed with a mean lifetime of 100 hours. What is the probability that

(i)the lifetime of the component is more than 100hours?

(ii)the lifetime of the component is between 50 to 100hours?

(iii)a component will fail before 50hours?

NORMAL DISTRIBUTION

- The most important of all continuous probability distribution

- Used extensively as the basis for many statistical inference methods.

- The probability density function is a bell-shaped curve that is symmetric about μ.

- The notation $X \sim N(\mu, \sigma^2)$ denotes the random variable X has a normal distribution with mean μ and variance σ^2.

- A normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$ is known as the standard normal distribution. $Z \sim N(0, 1)$

NORMAL DISTRIBUTION

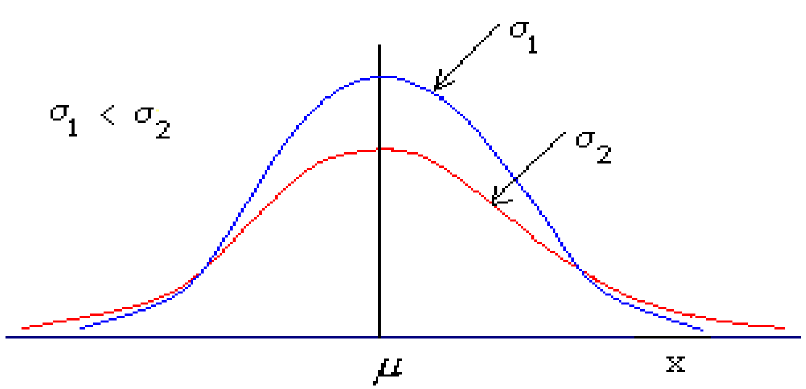
- **Probability Density function of $X \sim \mathcal{N}(\mu, \sigma^2)$**

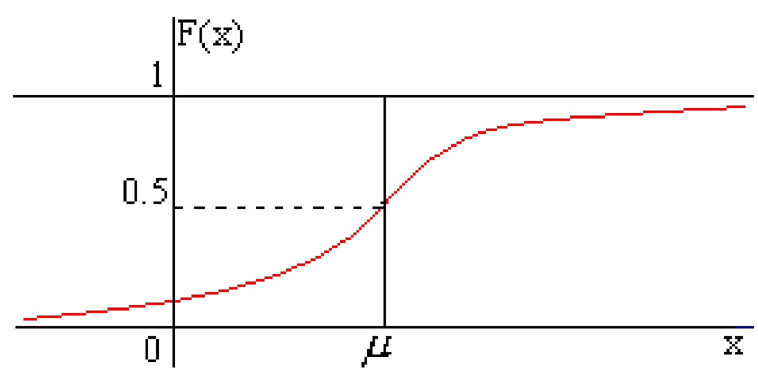
$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}[(x-\mu)/\sigma]^2} \quad , \quad -\infty < x < +\infty \quad ,$$

$$\text{where } \mu \text{ is a real number and } \sigma > 0.$$

$$E(X) = \mu \quad \text{and} \quad V(X) = \sigma^2$$

- **Graph**





In order to compute $P(a \leq X \leq b)$ when X is a normal rv with parameters μ and σ, we must determine

$$\int_a^b \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}[(x-\mu)/\sigma]^2},$$

However, it is not easy to evaluate this expression, so numerical techniques have been used to evaluate the integral when $\mu = 0$ and $\sigma = 1$, for certain values of *a* and *b* and results are tabulated. This table is often used to compute probabilities for any other values of μ and σ under consideration.

standard normal distribution : The normal distribution with parameter values $\mu = 0$ and $\sigma = 1$.
standard normal random variable (Z) : A random variable having this distribution.

STANDARD NORMAL R.V $Z \sim N(0\,,1)$

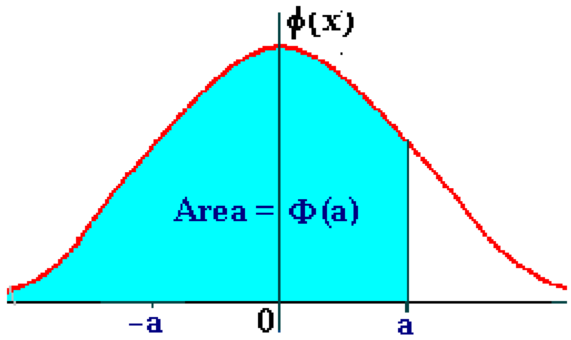
UNIVERSITI TEKNOLOGI PETRONAS

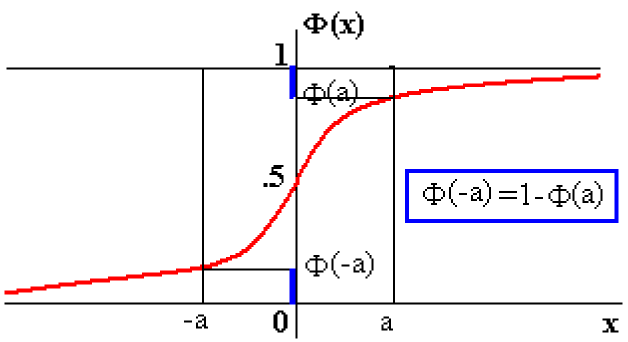
- Probability Density of Standard Normal RV

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

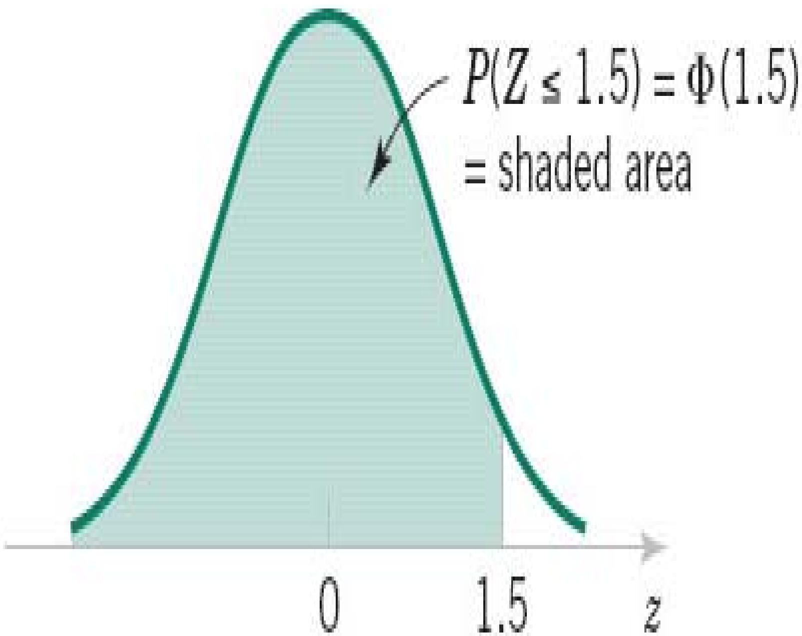
- Cumulative Distribution Function

$$P(Z \le z) = \Phi(z) = \int_{-\infty}^{z} \phi(t)\, dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt \;, \quad -\infty < z < \infty$$





The function $\Phi(z) = P(Z \leq z)$ is the area under the standard normal density curve to the left of z



z	0.00	0.01	0.02	0.03
0	0.50000	0.50399	0.50398	0.51197
⋮		⋮		
1.5	0.93319	0.93448	0.93574	0.93699

Example 1: Evaluate

$$a)\ \ P(Z \le 1.25)$$

$$= \Phi(1.25) = 0.8944$$

$$b)\ \ P(Z > 1.25)$$

$$= 1 - P(Z \le 1.25)$$

$$= 1 - \Phi(1.25)$$

$$= 0.1056$$

$$c)\ \ P(Z \le -1.25)$$

$$= \Phi(-1.25) = 0.1056$$

$$d)\ \ P(-0.38 \le Z \le 1.25)$$

$$= \Phi(1.25) - \Phi(-0.38)$$

$$= 0.8944 - 0.3520$$

$$= 0.5424$$

Example 1: Evaluate

$$a)\ \ P(Z \le 1.25) \quad = \Phi(1.25) = 0.8944$$

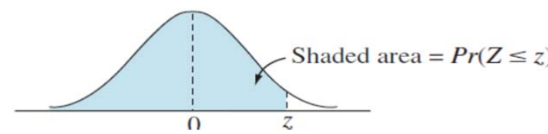


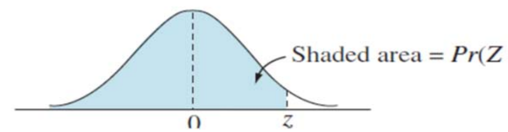
TABLE 1

Standard normal curve areas

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
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	$\boxed{0.8944}$	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319

Example 1: Evaluate

b) $P(Z \le -1.25) \;= \Phi(-1.25) = 0.1056$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	$\boxed{0.1056}$	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170

Example 2:

After completing a study, the civil engineering department in Universiti Teknologi PETRONAS (UTP) concluded that the time UTP employees spend commuting to work each day is normally distributed with a mean equal to 15 minutes and a standard deviation equal to 5 minutes. One employee has indicated that he commutes 25 minutes per day. Find the probability that an employee would commute 25 or more minutes per day,

Solution:

The random variable ***X*** is the time employees spend commuting to work and

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

where $\mu = 15$ and $\sigma^2 = (5)^2$

Hence,

$$P(X \geq 25) = P\left(\frac{x-\mu}{\sigma} \geq \frac{25-15}{5}\right) = P(z \geq 2)$$

$$= 1 - P(z \leq 2) = 1 - \Phi(2) = 1 - 0.977 = 0.023$$

TABLE 1

Standard normal curve areas

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						
0.1	0.5398	0.5438	0.5478	0.5517						
0.2	0.5793	0.5832	0.5871	0.5910						
0.3	0.6179	0.6217	0.6255	0.6293						
0.4	0.6554	0.6591	0.6628	0.6664						
0.5	0.6915	0.6950	0.6985	0.7019						
0.6	0.7257	0.7291	0.7324	0.7357						
0.7	0.7580	0.7611	0.7642	0.7673						
0.8	0.7881	0.7910	0.7939	0.7967						
0.9	0.8159	0.8186	0.8212	0.8238						
1.0	0.8413	0.8438	0.8461	0.8485						
1.1	0.8643	0.8665	0.8686	0.8708						
1.2	0.8849	0.8869	0.8888	0.8907						
1.3	0.9032	0.9049	0.9066	0.9082						
1.4	0.9192	0.9207	0.9222	0.9236						
1.5	0.9332	0.9345	0.9357	0.9370						
1.6	0.9452	0.9463	0.9474	0.9484						
1.7	0.9554	0.9564	0.9573	0.9582						
1.8	0.9641	0.9649	0.9656	0.9664						
1.9	0.9713	0.9719	0.9726	0.9732						
2.0	$\boxed{0.9772}$	0.9778	0.9783	0.9788						
2.1	0.9821	0.9826	0.9830	0.9834						
2.2	0.9861	0.9864	0.9868	0.9871						
2.3	0.9893	0.9896	0.9898	0.9901						
2.4	0.9918	0.9920	0.9922	0.9925						

$$P(X \geq 25) = P\left(\frac{x-\mu}{\sigma} \geq \frac{25-15}{5}\right)$$

$$= P(z \geq 2)$$

$$= 1 - P(z \leq 2)$$

$$= 1 - \Phi(2)$$

$$= 1 - 0.977$$

$$= 0.023$$

Exercise:

The line width of a tool used for semiconductor manufacturing is assumed to be normally distributed with a mean of 0.55 μm and variance of 0.0025 μm.

(a) What is the probability that a line width is greater than 0.60 μm?
(b) What is the probability that a line width is between 0.45 μm and 0.65μm?
(c) The line width of 95% of samples is below what value?

CRITICAL VALUES OF THE STANDARD NORMAL Random Variable

- Critical Value z_α: for $Z \sim \mathcal{N}(0, 1)$
- denote the value on the z axis for which α of the area under the z curve lies to the right of z_α

$$shaded\ \ area = P(Z \geq z_\alpha) = \alpha$$