B. Tech Computer Science & Engineering
(School of Technology)
SEMESTER –V

MATH2361: PROBABITY AND STATISTCS

•••

UNIT-II
Probability Distributions

Dr. Mallikarjuna Reddy Doodipala
Department of Mathematics
GITAM University
Hyderabad Campus

Learning Objectives

At the end of the module students able to learn:

- Understand the density and distribution functions for discrete and continuous variables
- Know about probability distributions binomial, Poisson and Normal
- Applications

Learning Outcomes

After completion of this unit, the student will be able to

- apply Binomial and Poisson distributions to compute probabilities, theoretical frequencies (L3).
- explain the properties of normal distribution and its applications (L3).

Probability Distributions

- Probability mass and density functions,
- Probability distributions

1. Binomial,
2. Poisson,
3. Normal distributions and their properties(Only Mean Variance)

- Applications on Binomial, Poisson and Normal Distributions

Prerequisites

Before you start reading this unit, you should:

- Have some knowledge on definite integrals
- Know about probability and calculating it for simple problems
- Series & Exponential functions
- R.Vs Expectations and Variance

Probability Distributions

- Probability Distribution: Table, Graph, or Formula that describes values a random variable can take on, and its corresponding probability (discrete RV) or density (continuous RV)
- Discrete Probability Distribution: Assigns probabilities (masses) to the individual outcomes
- Continuous Probability Distribution: Assigns density at individual points, probability of ranges can be obtained by integrating density function
- Discrete Probabilities denoted by: $p(x) = P(X=x)$
- Continuous Densities denoted by: $f(x)$
- Cumulative Distribution Function: $F(x) = P(X \leq x)$

Expected Values of Discrete RV's

- Mean (Expected Value) – Long-Run average value an RV (or function of RV) will take on
- Variance – Average squared deviation between a realization of an RV (or function of RV) and its mean
- Standard Deviation(SD) – Positive Square Root of Variance (in same units as the data)
- Notation: Mean: $E(Y) = \mu$
- Variance: $V(Y) = \sigma^2$
- Standard Deviation: σ

Probability Distributions

-Bernoulli Distribution

Bernoulli distribution

- An experiment consists of one trial. It can result in one of 2 outcomes: Success or Failure (or a characteristic being Present or Absent).
- Probability of Success is p $(0<p<1)$
- X= 1 if Success (Characteristic Present), 0 if not

$$p(x) = \begin{cases} p & x = 1 \\ 1-p & x = 0 \end{cases}$$

$$E(Y) = \sum_{x=0}^{1} xp(x) = 0(1-p) + 1p = p$$

$$E(X^2) = 0^2(1-p) + 1^2 p = p$$

$$\Rightarrow V(X) = E(X^2) - [E(X)]^2 = p - p^2 = p(1-p) \text{ and S.D is}$$

$$\Rightarrow \sigma = \sqrt{p(1-p)}$$

Binomial Experiment

- Experiment consists of a series of n identical trials
- Each trial can end in one of 2 outcomes: Success or Failure
- Trials are independent (outcome of one has no bearing on outcomes of others)
- Probability of Success, p, is constant for all trials
- Random Variable X, is the number of Successes in the n trials is said to follow Binomial Distribution with parameters n and p
- X can take on the values $X = 0,1,\ldots,n$
- Notation: $X \sim Bin(n,p)$

Binomial Distribution

- A trial with only two possible outcomes is used so frequently as a building block of a random experiment that it is called a **Bernoulli trial**.

- It is usually assumed that the trials that constitute the random experiment are **independent.** This implies that the outcome from one trial has no effect on the outcome to be obtained from any other trial.

- Furthermore, it is often reasonable to assume that the probability of a success on each trial is constant.

Binomial Distribution

Consider the following random experiments and random variables. Do they meet the following criteria:

1. Does the experiment consist of **Bernoulli trials?**
2. Are the trials that constitute the random experiment are **independent?**
3. **Is probability of a success on each trial is constant?**

1. Flip a coin 10 times. Let X = the number of heads obtained..

How can you deal If a random experiment consists ?

- Finite number of trials say n
- All the trials are independent
- Each trail has only two possible outcomes
- The probability of success(p) in each trial, remains constant

Binomial Distribution

A random experiment consisting of n trials in which

1. All the trials are **independent**
2. Only two possible out comes **success and failure**
3. probability of a success event on each trial is constant

Let us define a r.v X as no.of successes of n trials then the the probability of X is given by

$$p(x) = \binom{n}{x} p^x q^{n-x} \quad (x = 0, 1, \ldots, n)$$

Mean and Variance of Binomial

Note that Binomial Probabilities Tables (available in many textbooks) can be used for certain values of p, n, and x. Also note that

$$E(X) = \mu = np$$

and

$$V(X) = \sigma^2 = npq$$

for the binomial distribution.

Binomial Distribution

Example: What is the probability of getting exactly three heads in seven flips of a fair coin?

$$f(3) = \binom{7}{3} 0.50^3 (1 - 0.50)^{(7-3)}$$

$$= \frac{7!}{3!(7-3)!}(0.125)(0.0625)$$

$$= 0.2734$$

Binomial Distribution

Example: What is the probability of getting at least one head in seven flips of a fair coin?

The sample space is $S = \{0, 1, 2, 3, 4, 5, 6, 7\}$

so $P(x \geq 1) = f(1) + f(2) + f(3) + f(4) + f(5) + f(6) + f(7)$

Is there an easier way?

$P(x \geq 1) = 1 - P(\overline{x \geq 1}) = 1 - P(x < 1) = 1 - P(x = 0) = 1 - f(0)$

and

$$f(0) = \binom{7}{0} 0.50^0 (1 - 0.50)^{(7-0)}$$

$$= \frac{7!}{0!(7-0)!}(1.00)(0.0078125)$$

$$= 0.0078125$$

so $P(x \geq 1) = 1 - 0.0078125 = 0.9921875$

Binomial Distribution

Example: What is the probability of getting no more than one head in seven flips of a fair coin?

$$P(x \le 1) = P(x = 0) + P(x = 1) = 0.0078125 + P(x = 1)$$

and

$$f(1) = \binom{7}{1} 0.50^1 (1 - 0.50)^{(7-1)}$$

$$= \frac{7!}{1!(7-1)!}(0.50)(0.015625)$$

$$= 0.0546875$$

so $P(x \le 1) = 0.0078125 + 0.0546875 = 0.0625$

Binomial Distribution

Example. Suppose X has a binomial distribution with $n = 10, p = .4$. Find

(i) $P(X \leq 4) = .633$

(ii) $P(X < 6) = P(X \leq 5) = .834$

(iii) $P(X > 4) = 1 - P(X \leq 4) = 1 - .633 = .367$

(iv) $P(X = 5) = P(X \leq 5) - P(X \leq 4) = .834 - .633 = .201$

Exercise: Answer the same question with $p = 0.7$

Binomial Distribution: Application

no. of Heads	0	1	2	3	4	5
Frequency	2	6	14	9	5	4

Fit the binomial distribution for the data observed data

SOL : To fit the binomial distribution to the given data. By binomial formula

$$p(x) = \binom{n}{x} p^x q^{n-x} \quad (x = 0, 1, \ldots, n)$$

n = no. of trials=5

X = no. of heads

Mean = np

P = mean/n **then q =** $\frac{1}{N}\sum fx$ **mean**=

=101/40=2.525

$\frac{1}{N}\sum\frac{1}{N}\sum fx$

P=2.525/5=0.505,q=1-0.505=0.495

Binomial Distribution: Application

X	F	f*x	p(x)	EXP F=N*p (x)
0	2	0	0.03	1.13
1	6	6	0.15	5.88
2	14	28	0.31	12.24
3	9	27	0.32	12.74
4	5	20	0.17	6.63
5	4	20	0.03	1.38
TOTAL	N=40	101	1.00	40.00

We conclude that binomial method is best fit to this observed Coin tossing experiment data.

Binomial Distribution – Expected Value(Mean)

$$p(x) = \frac{n!}{x!\,(n-x)!} p^x q^{n-x} \quad x = 0,1,\dots,n \quad q = 1-p$$

$$E(X) = \sum_{x=0}^{n} x p(x) = \sum_{x=0}^{n} x \left[\frac{n!}{x!\,(n-x)!} p^x q^{n-x} \right]$$

$$\Rightarrow E(X) = \sum_{x=0}^{n} \left[\frac{xn!}{x(x-1)!\,(n-x)!} p^x q^{n-x} \right] = p \sum_{x=1}^{n} \left[\frac{n!}{(x-1)!\,(n-x)!} p^{x-1} q^{n-x} \right]$$

$$np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!\,\left((n-x\quad\right)!} p^{x-1} q^{(n-x)} =$$

$$= np(p+q)^{n-1} = np(p+(1-p))^{n-1} = np(1) = np$$

Binomial Distribution – Variance and S.D.

$$p(x) = \frac{n!}{x!\,(n-x)!} p^x q^{n-x} \quad x = 0,1,\ldots,n \quad q = 1-p$$

Note: ⇄ $E(X^2)$ is difficult (impossible?) to get, but $E(X(X-1)) = E(X^2) - E(X)$ is not:

$$E(X(X-1)) = \sum_{x=0}^{n} x(x-1)\left[\frac{n!}{x!\,(n-x)!} p^x q^{n-x}\right] = \sum_{x=2}^{n} x(x-1)\left[\frac{n!}{x!\,(n-x)!} p^x q^{n-x}\right]$$

(Sum and $= 0$ when $x = 0,1$)

$$\Rightarrow E(X(X-1)) = \sum_{x=2}^{n} \frac{n!}{(x-2)!\,(n-x)!} p^x q^{n-x}$$

$$\Rightarrow E(X(X-1)) = p^2 \sum_{x=2}^{n} \frac{n(n-1)(n-2)!}{(x-2)!\,(n-x))!} p^{x-2} q^{n-x} = n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!\,((n-x))!} p^{x-2} q^{(n-x)}$$

$=$

$$= n(n-1)p^2(p+q)^{n-2} = n(n-1)p^2(p+(1-p))^{n-2} = n(n-1)p^2$$

$$\Rightarrow E(X^2) = E(X(X-1)) + E(X) = n(n-1)p^2 + np$$

$$\Rightarrow np[(n-1)p+1] = n^2p^2 - np^2 + np = n^2p^2 + np(1-p)$$

$$\Rightarrow V(X) = E(X^2) - [E(X)]^2 = n^2p^2 + np(1-p) - (np)^2 = np(1-p)$$

$$\Rightarrow \sigma = \sqrt{np(1-p)}$$

Poisson Distribution

- As a limit to binomial when n is large and p is small.
- A theorem by Simeon Denis Poisson(1781-1840).
- Parameter λ= np= mean
- As n is large and p is small, the binomial probability can be approximated by the Poisson probability function
- $P(X=x)= e^{-\lambda} \lambda^x / x!$, x= no. of counts: 0, 1, 2 , …

where e =2.71828 and

$\lambda =$ Parameter (mean)

Poisson Distribution

- The Poisson distribution is used to model the number of events occurring within a given time interval.
- The formula for the Poisson probability function is

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

- λ is the shape parameter which indicates the average(mean) number of events(counts) in the given time interval.

Poisson Distribution

- Distribution often used to model the number of incidences of some characteristic in time or space:
  - Arrivals of customers in a queue
  - Numbers of flaws in a roll of fabric
  - death of infants,
  - the number of customers arriving,

Poisson Distribution

- Distribution obtained as follows:
  - Break down the "area" into many small "pieces" (n pieces)
  - Each "piece" can have only 0 or 1 occurrences (p=P(1))
  - Let $\lambda=np \equiv$ Average number of occurrences over "area"
  - $Y \equiv$ # occurrences in "area" is sum of 0^s & 1^s over "pieces"
  - $Y \sim$ Bin(n,p) with $p = \lambda/n$
  - Take limit of Binomial Distribution as $n \to \infty$ with $p = \lambda/n$

Poisson Distribution: Mean & Variance

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

$$E(X) = \sum_{x=0}^{\infty} x\left[\frac{e^{-\lambda}\lambda^x}{x!}\right] = \sum_{x=1}^{\infty} x\left[\frac{e^{-\lambda}\lambda^x}{x!}\right] = \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^x}{(x-1)!} = \lambda e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!}$$

$$E(X(X-1)) = \sum_{x=0}^{\infty} x(x-1)\left[\frac{e^{-\lambda}\lambda^x}{x!}\right] = \sum_{x=2}^{\infty} x(x-1)\left[\frac{e^{-\lambda}\lambda^x}{x!}\right] = \sum_{x=2}^{\infty}\frac{e^{-\lambda}\lambda^x}{(x-2)!} =$$

$$= \lambda^2 e^{-\lambda}\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!} = \lambda^2 e^{-\lambda}e^{\lambda} = \lambda^2$$

$$\Rightarrow E(X^2) = E(X(X-1)) + E(X) = \lambda^2 + \lambda$$
$$\Rightarrow V(X) = E(X^2) - [E(X)]^2 = \lambda^2 + \lambda - [\lambda]^2 = \lambda$$
$$\Rightarrow \sigma = \sqrt{\lambda}$$

Example problem

- Arrivals at a bus-stop follow a Poisson distribution with an average of 4.5 every quarter of an hour. calculate the probability of

1.No arrivals

2.One arrival

3. Two arrivals

3. fewer than 3 arrivals

Example problem

The probabilities of 0 up to 2 arrivals can be calculated directly from the formula

$$p(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$ with λ =4.5

X=0, 1, 2, …….

$$p(0) = \frac{e^{-4.5}4.5^{0}}{0!}$$

So p(0) = 0.01111

Example problem

Similarly, $p(1)=0.04999$

and $p(2)=0.11248$

So the probability of fewer than 3 arrivals is

$$
\begin{aligned}
p(x<3) &= p(x=0,1,2) \\
&= p(x=0)+p(x=1)+p(x=2) \\
&= 0.01111+0.04999+0.11248 \\
&= 0.17358 \\
&= 17.4\%
\end{aligned}
$$

Problem 2: Fitting of Poisson Distribution

The following is the data distribution of clinics in a city with respect to no. of persons who reported to have of +ve symptoms of AIDS a particular day

Effected persons	No. of clinics
0	109
1	65
2	22
3	3
4	1
total	200

Problem 2: Fitting of Poisson Distribution

sol: To fit the Poisson distribution to the data we have the formulae

- $$p(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$ X=0, 1, 2, …….

x =0,1,2,3,……=+ ve cases

N = total no. of clinics

Mean = λ = $\frac{1}{N}\sum fx = 122/200 = 0.61$

Mean = 0.61

Calculations : Fitting of Poisson Distribution

X	F	f*x	p(x)	EXP F=N*p (x)
0	109	0	0.54	108.67=109
1	65	65	0.33	66.29=66
2	22	44	0.10	20.22=20
3	3	9	0.02	4.11=4
4	1	4	0.00	0.63=1
TOTAL	N=200	122	1.00	199.92=200

We conclude that Poisson method is best fit to this observed clinical data.

Problems for Practice (Binomial-Poisson distribution)

- A multiple choice questionnaire has 12 questions with 5 options A-E. a student is completely un prepared and wrote the test. If test follows binomial probability law find (1) Exactly two answers are correct 2) At least two answers are correct 3) More than 7 answers are correct.

- With the usual notation find p for binomial random variable X if n = 6 and $9P(X = 4) = P(X = 2)$

- The mean and variance of a Binomial variate X with parameters n and p are 16 and 8. Find P (X = 1).
- A typist makes on average 2 mistakes per page. What is the probability of a particular page having no errors on it?
- Components are packed in boxes of 20. The probability of a component being defective is 0.1. What is the probability of a box containing 2 defective components?

Continuous distributions

Normal Distribution

Normal distribution

A continuous r.v X is said to have follows normal distribution if its pdf is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-\mu)^2/2\sigma^2}$$

where μ = mean

σ = standard deviation

π = 3.14159

e = 2.71828

Standard Normal Probability Density Function

A continuous r.v X is said to have follows standard normal distribution if its pdf is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-\mu)^2/2\sigma^2}$$

where μ = mean=0

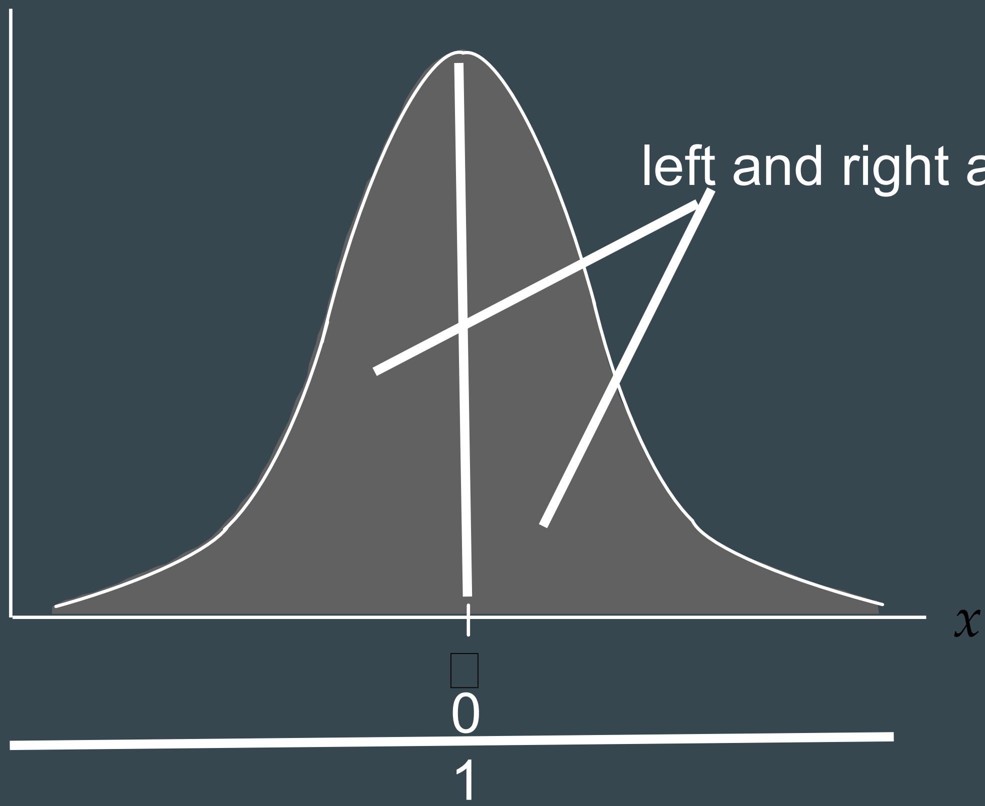
σ = standard deviation=1

π = 3.14159

e = 2.71828

The Normal Probability Distribution

- Graph of the Normal Probability Density Function $f(x)$ The total probability is one under the curve



The Normal Curve

- The shape of the normal curve is often illustrated as a bell-shaped curve.
- The highest point on the normal curve is at the mean of the distribution.
- The normal curve is symmetric.
- The standard deviation determines the width of the curve.

The Normal Curve

- The total area under the curve the same as any other probability distribution is 1.

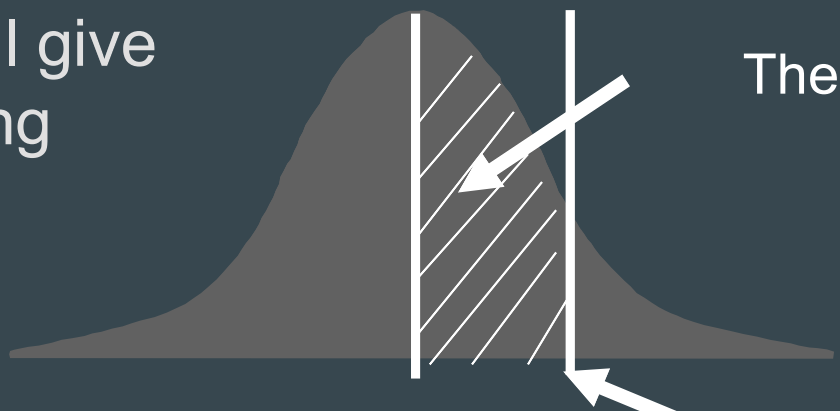
- The probability of the normal random variable assuming a specific value the same as any other continuous probability distribution is 0.

- Probabilities for the normal random variable are given by areas under the curve.

The Normal Curve- Probability

Given any positive value for *z*, the table will give us the following

probability



The table will give this probability

Given **positive** z

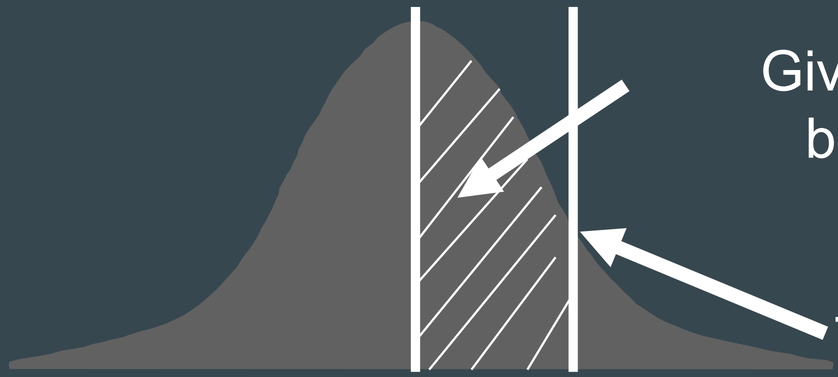
The probability that we find using the table is the probability of having a standard normal variable

between 0 and the given positive z.

The Normal Tables- Probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389

The Normal Curve- Probability

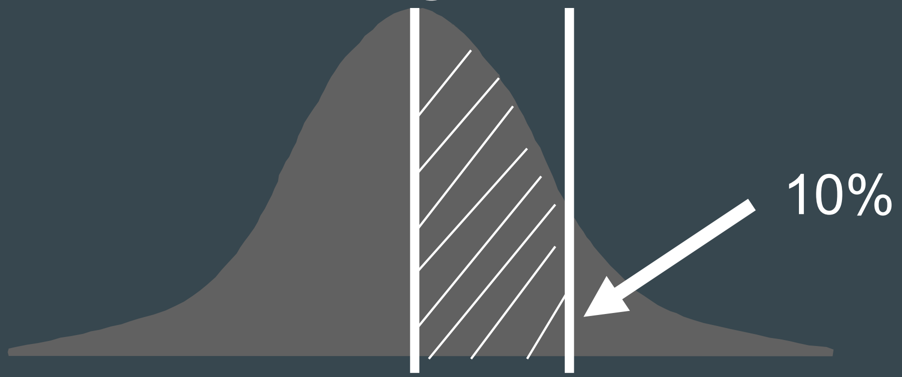


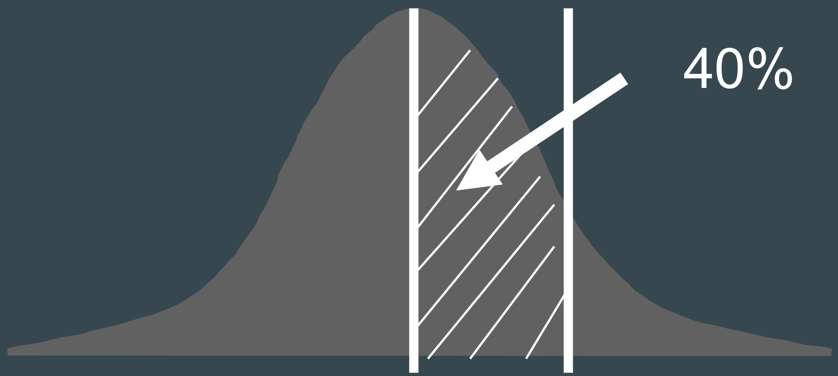
Given any probability between 0 and 0.5,,
the table will give us the following positive *z*
value

Given the probability find z find

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389

***What is the z* value** where probability of a standard normal variable to be greater than z is .1





The Standard Normal Variate- Z

- A random variable that has a normal distribution with a mean of zero and a standard deviation of one is said to have a **standard normal probability distribution.**
- The letter *z* is commonly used to designate this normal random variable.
- The following expression convert any Normal Distribution into the Standard Normal Distribution

$$z = \frac{x - \mu}{\sigma}$$

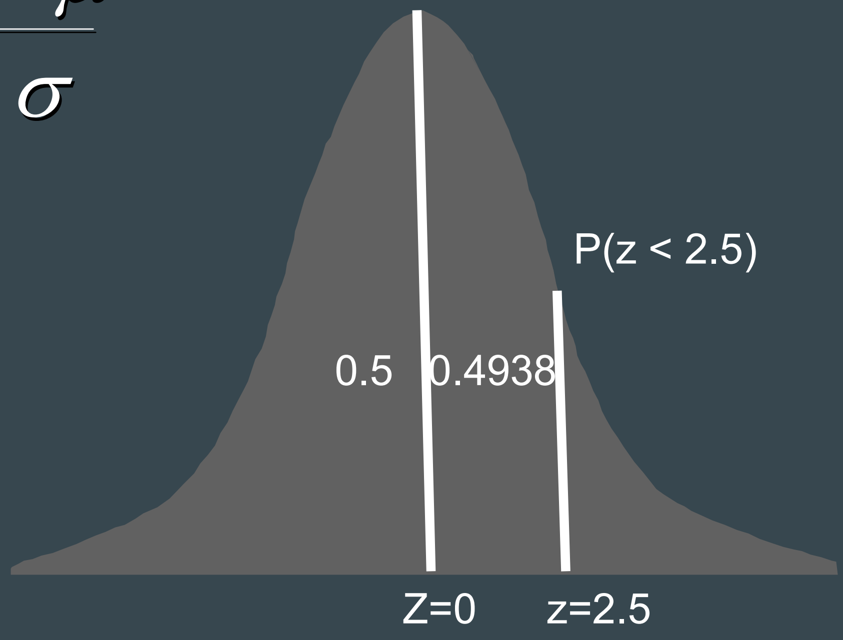
Problems with Solutions

Problem-1 X is a normally distributed variable with mean μ = 30 and standard deviation σ = 4. Find a) P(x < 40) b) P(x > 21)
c) P(30 < x < 35)
Sol: the area under the standard normal curve is unity.

$$z = \frac{x - \mu}{\sigma}$$

By area property of SNV

a) For x = 40, μ = 30 ,σ = 4 then z-value
z = (40 - 30) / 4 = 2.5
Hence P(x < 40) = P(z < 2.5)
= [area to the left of 2.5]
= 0.5+p(0<z<2.5)=0.5+0.4938
=0.9938
From the Table p(0<z<2.5)=0.4938



Problems with Solutions

Cont'd

b) For $x = 21$, $z = (21 - 30) / 4 = -2.25$

Hence $P(x > 21) = P(z > -2.25)$ = [total area] - [area to the left of -2.25]

$= 1 - 0.0122 = 0.9878$

c) For $x = 30$, $z = (30 - 30) / 4 = 0$ and for $x = 35$, $z = (35 - 30) / 4 = 1.25$

Hence $P(30 < x < 35) = P(0 < z < 1.25)$ = [area to the left of $z = 1.25$] - [area to the left of 0]

$= 0.8944 - 0.5 = 0.3944$

Problems with Solutions

A radar unit is used to measure speeds of cars on a motorway. The speeds are normally distributed with a mean of 90 km/hr and a standard deviation of 10 km/hr. What is the probability that a car picked at random is travelling at more than 100 km/hr?

Solution:

Let x be the random variable that represents the speed of cars.

x has $\mu = 90$ and $\sigma = 10$.

We have to find the probability that x is higher than 100 or $P(x > 100)$

For $x = 100$, $z = (100 - 90) / 10 = 1$

$P(x > 90) = P(z > 1) =$

[total area] - [area to the left of z = 1]

$= 1 - 0.8413 = 0.1587$

The probability that a car selected at a random has a speed greater than 100 km/hr is equal to 0.1587

Problems with Solutions

Problem-3

For a certain type of computers, the length of time between charges of the battery is normally distributed with a mean of 50 hours and a standard deviation of 15 hours. one of these computers and wants to know the probability that the length of time will be between 50 and 70 hours.

Solution:

Let x be the random variable that represents the length of time. It has a mean of 50 and a standard deviation of 15.

We have to find the probability that x is between 50 and 70 or P(50< x < 70)

Problems with Solutions

For $x = 50$, $z = (50 - 50) / 15 = 0$

For $x = 70$, $z = (70 - 50) / 15$
$= 1.33$ (rounded to 2 decimal places)

$P(50< x < 70) = P(0< z < 1.33)$
= [area to the right of z from 0 to 1.33] = 0.4082

The probability that computer has a length of time between 50 and 70 hours is equal to 0.4082 =40.82%.

Problems with Solutions

Find the following probabilties at mean=15,sd=6.
1.p(15<x<20) 2.p(x>=20).

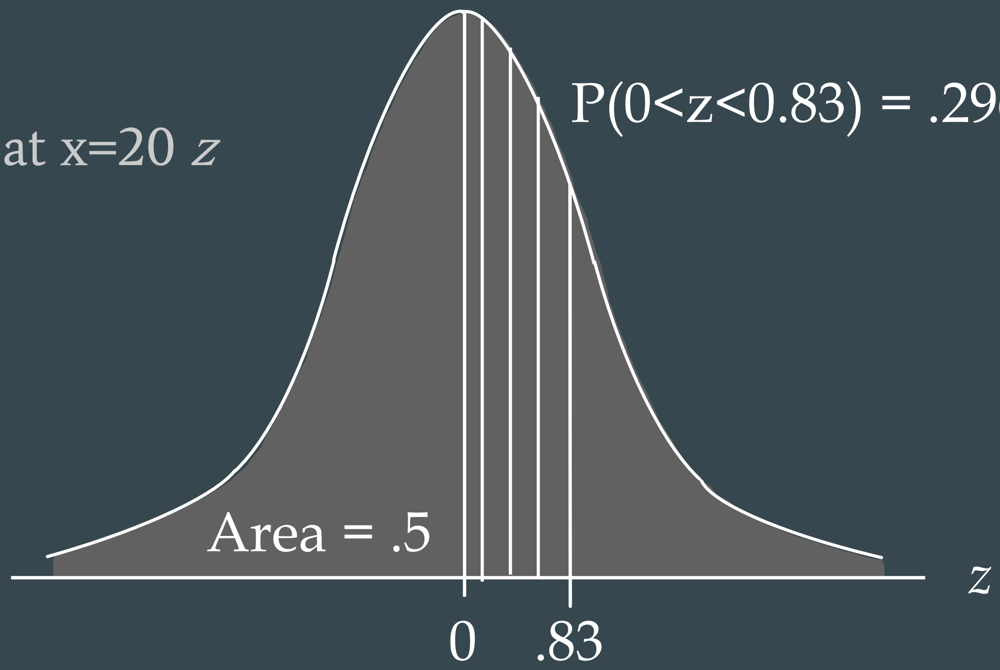
Solution: 1.p(x<20)

at x=0,z=0 and at x=20 z

$= (x - \text{mu}\ )/\ \sigma$

$= (20 - 15)/6$

$= .83$



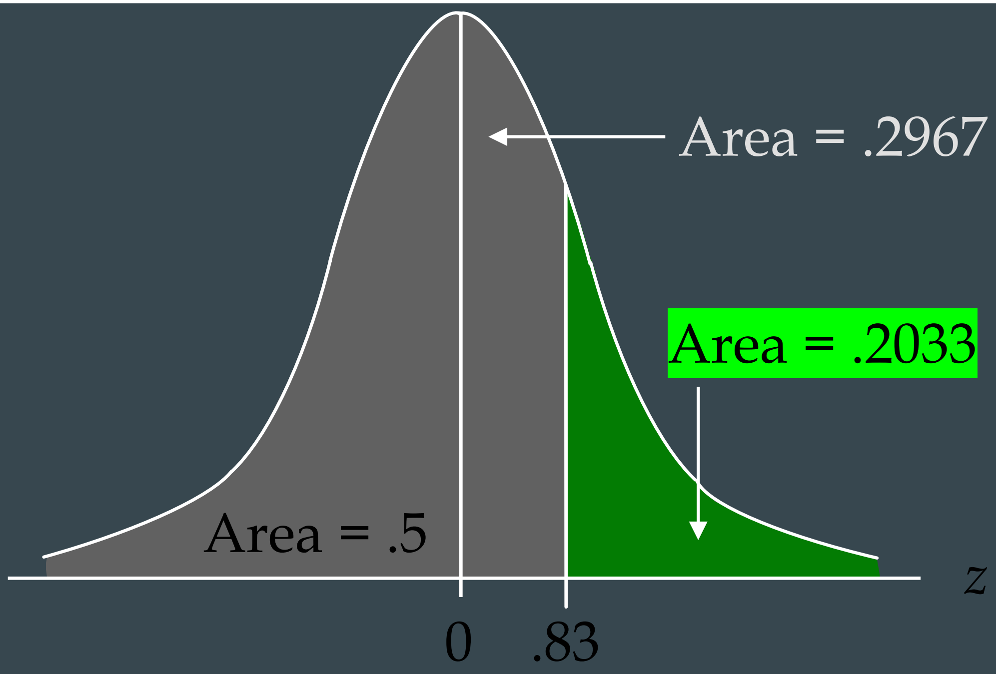
The Standard Normal table shows an area of .2967 for the region between the z = 0 line and the z = .83 line above.

Sample Normal Tables

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389

Problems with Solutions

2. P(x>=20)
=p(z>=0.83)
=0.5-p(<0.83)
=0.5-0.2967
=0.2033



The shaded tail area is 0.5 - .2967 = .2033.

The probability of x>=20 is
0.2033=20.33%.

Problems with Solutions

If probability is given how to find Z-Value

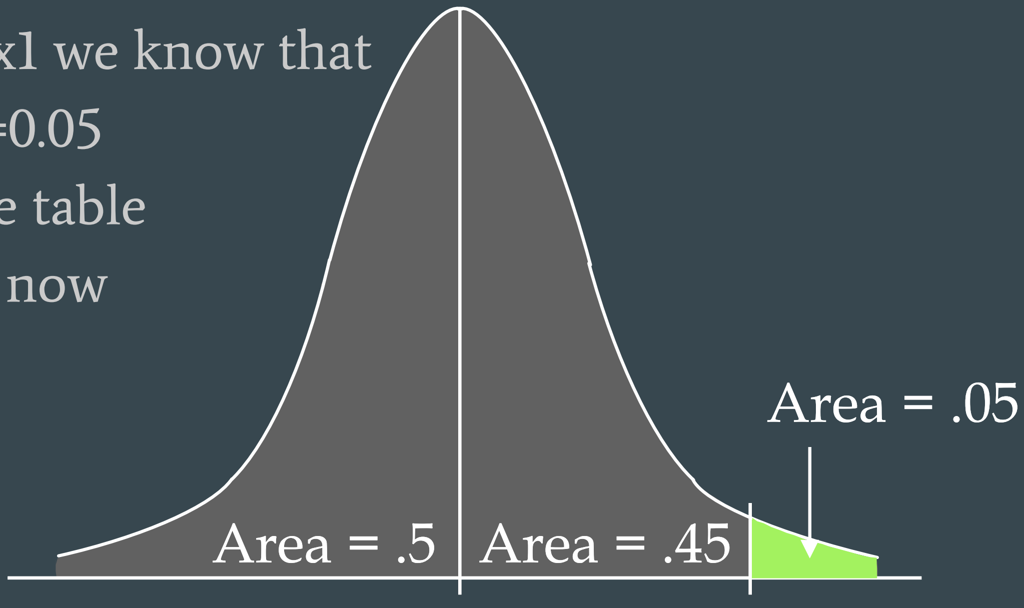
Find p(X>X1)=0.05 Find X1

To find x1 we know that

P(z>z1)=0.05

From the table

z1=1.645 now



Let z_1 represent the z value cutting the tail area of .05.

Problems with Solutions

Example: Pep Zone

- Using the Standard Normal Probability Table

  We now look-up the .4500 area in the Standard Normal Probability table to find the corresponding $z_{.05}$ value. $z_{.05} = 1.645$ is a reasonable estimate.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.										
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
.										

Problems with Solutions

Continued

The corresponding value of x is given by

$$x1- \mu \,/\, \sigma = z1$$

$$x1 = \mu + \sigma z_1$$

$= 15 + 1.645(6)$ where mean=15, sd=6

$= 24.87$

Therefore x1=24.87

Problems with Solutions: Aptitude Test

A firm has assumed that the distribution of the aptitude test of people applying for a job in this firm is normal. The following sample is available.

71	66	61	65	54	93
60	86	70	70	73	73
55	63	56	62	76	54
82	79	76	68	53	58
85	80	56	61	61	64
65	62	90	69	76	79
77	54	64	74	65	65
61	56	63	80	56	71
79	84				

Example: Mean and Standard Deviation

We first need to estimate mean and standard deviation

$$\bar{x} = \frac{\sum x_i}{n} = \frac{3421}{50} = 68.42$$

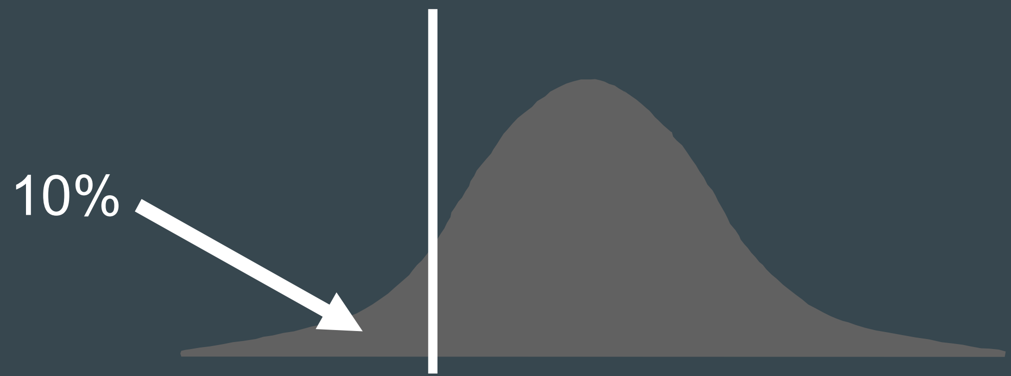
$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}} = \sqrt{\frac{5310.04}{49}} = 10.41$$

z Values

What test mark has the property of having 10% of test marks being less than or equal to it

To answer this question, we should first answer the following ….

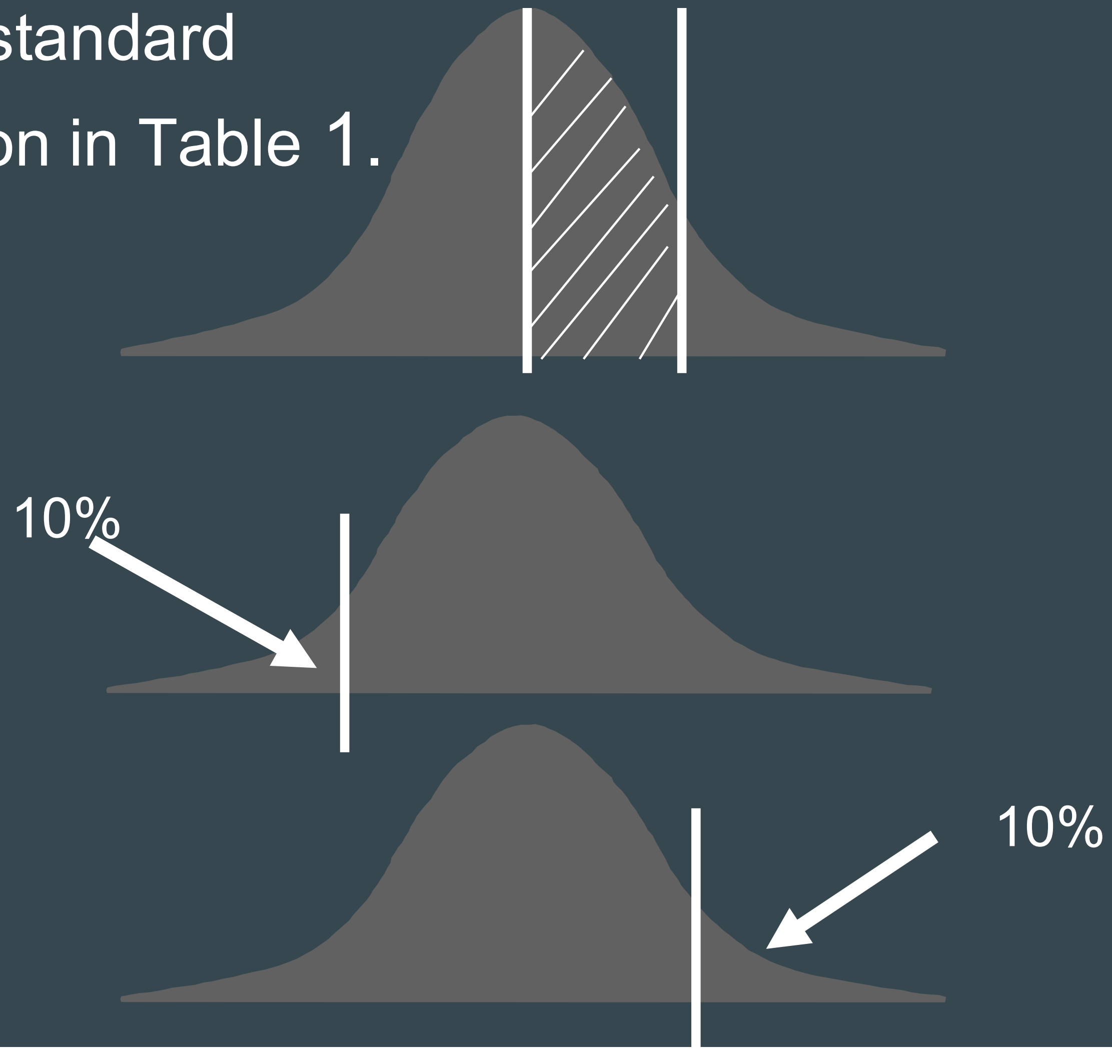
What is the standard normal value (z value), such that 10% of z values are less than or equal to it?



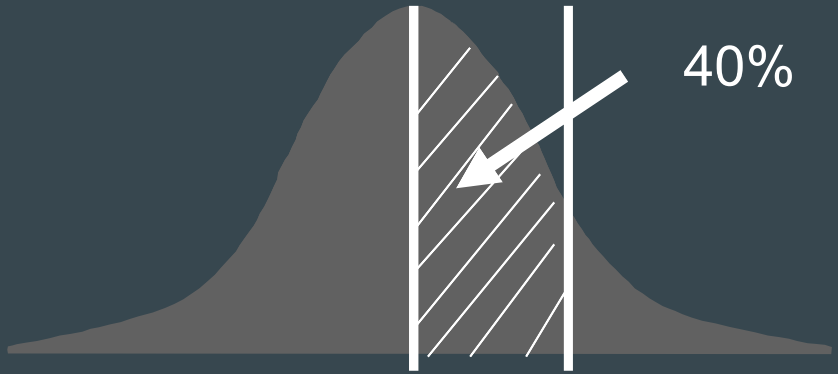
z Values

We need to use standard

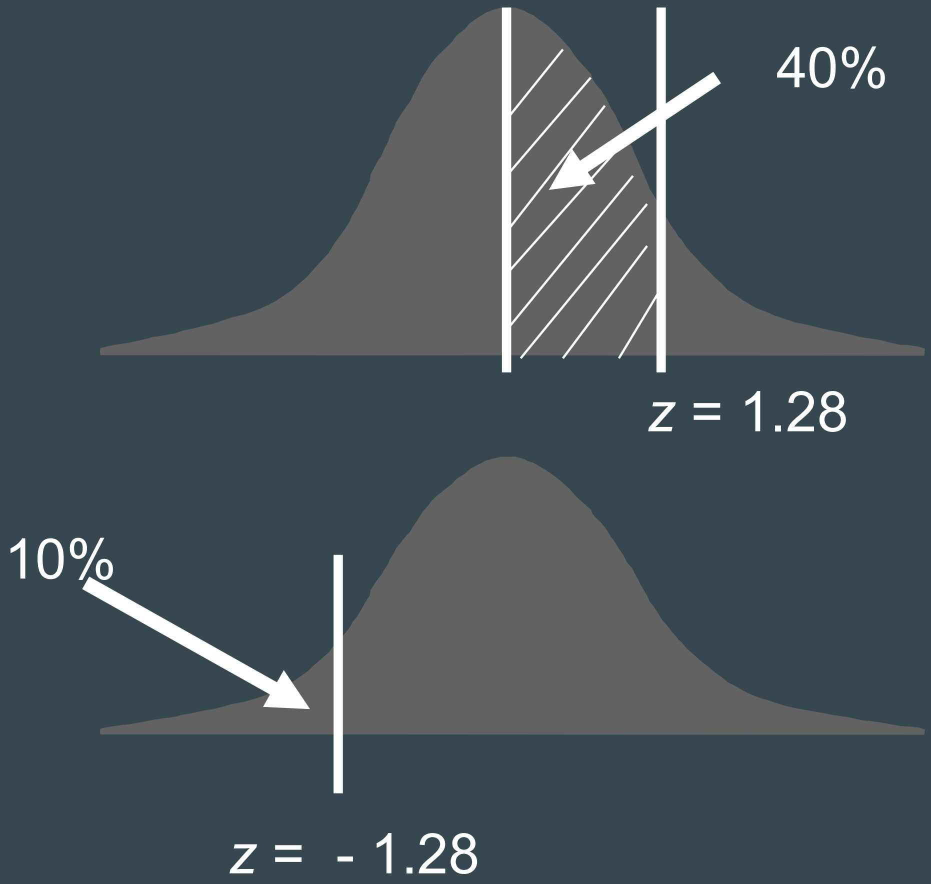
Normal distribution in Table 1.



z Values



z Values



z Values and *x* Values

The standard normal value (*z* value), such that 10% of z values are less than or equal to it is z = -1.28

To transform this standard normal value to a similar value in our example, we use the following relationship

$$\frac{x-\mu}{\sigma} = z$$

$$\frac{x-68.42}{10.41} = -1.28$$

$$x = 10.41(-1.28) + 68.42$$
$$= 55.1$$

The normal value of test marks such that 10% of random variables are less than it is 55.1.

z Values and x Values

Following the same procedure, we could find **z values** for cases where 10 %, 20%, 30%, 40%, …of random variables are less than these values. Following the same procedure, we could transform z values into x values.

$$\frac{x-\mu}{\sigma} = z$$

Lower 10%	-1.28	55.1
Lower 20%	-.84	59.68
Lower 30%	-.52	63.01
Lower 40%	-.25	65.82
Lower 50%	0	68.42
Lower 60%	.25	71.02

Properties

- Continuous Random Variable
- Mound or Bell-shaped curve
- The normal curve extends indefinitely in both directions, approaching, but never touching, the horizontal axis as it does so.
- Unimodal
- Mean = Median = Mode
- Symmetrical with respect to the mean. That is, 50% of the area (data) under the curve lies to the left ofthe mean and 50% of the area (data) under the curve lies to the right of the mean.

Properties

≈

- 68% of the area (data) under the curve is within one standard deviation of the mean
- 95% of the area (data) under the curve is within two standard deviations of the mean
- 99.7% of the area (data) under the curve is within three standard deviations of the mean
- The total area under the normal curve is equal to 1

Problems for practice (Normal distribution)

1.X is a normally distributed with mean $\mu = 30$ and SD $\sigma = 4$. Find a) $P(x < 40)$

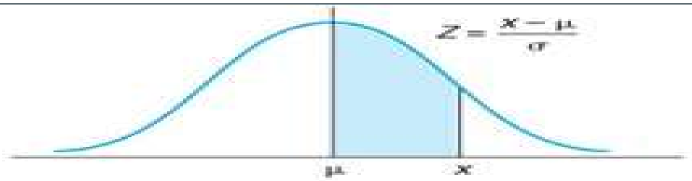
b) $P(x > 21)$

c) $P(30 < x < 35)$

2. In a normal distribution 10.03% of the items are under 25 kilogram and 89.97% of the items are under 70 kilogram. What are the mean and standard deviation of the distribution.

3. A large group of students took a test in Physics and the final grades have a mean of 70 and a standard deviation of 10. If we can approximate the distribution of these grades by a normal distribution, what percent of the students

a) scored higher than 80? b) Should pass the test (grades≥60)? c) Should fail the test (grades<60)?



Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Feed Back to mdoodipa@gmail.com