Chapter 4
Probability theory

Probability theory

The difficulties in understanding inferential statistics branch are due to the probability or chance of drawing wrong conclusions

The probability of drawing wrong conclusions attributed to:

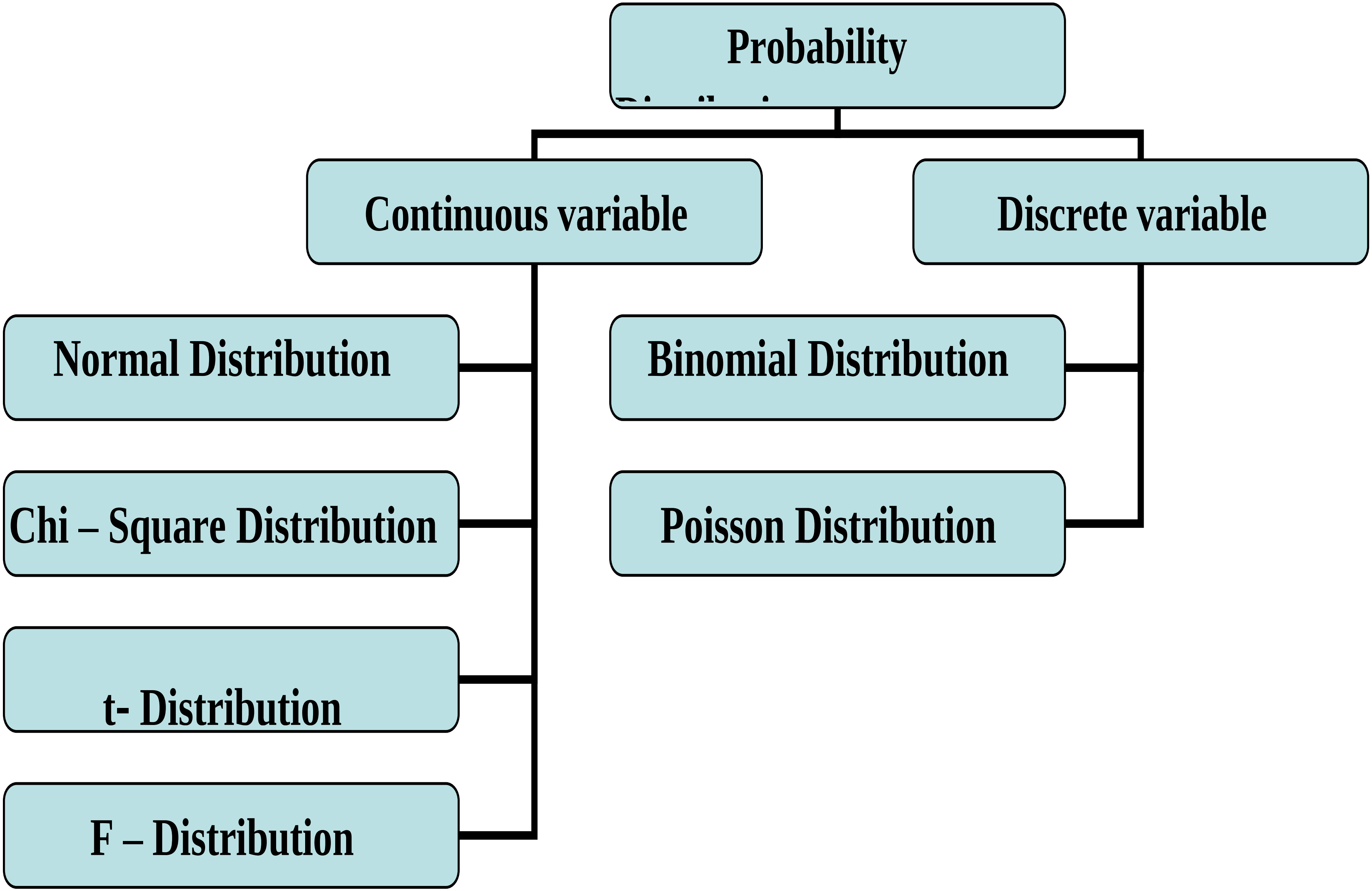
① Dealing with samples to draw conclusions about populations.
② Drawing wrong conclusions due to unobservable(uncontrolled) factors that affect on your results or observations

- The probability of occurrence of an event is called the probability laws.

P(E) = number of occurrence of an event E / all the possible outcomes

P(E) = the probability of the occurrence of an event E “$0 \leq P(E) \leq 1$”

The Probability distributions (Chapter 5)



The Binomial (Bernoulli) Distribution

- The binomial distribution is used when a researcher is interested in the occurrence of an event, not in its magnitude.

- For example:-
- The animal is pregnant or not
- The animal male or female
- The animal shows the clinical signs of infectious disease or it not (not the severity)

The Poisson distribution

- This is often known as the distribution of rare events
- DISCRETE events occur in a CONTINUOUS interval of time or space (this interval is finite)

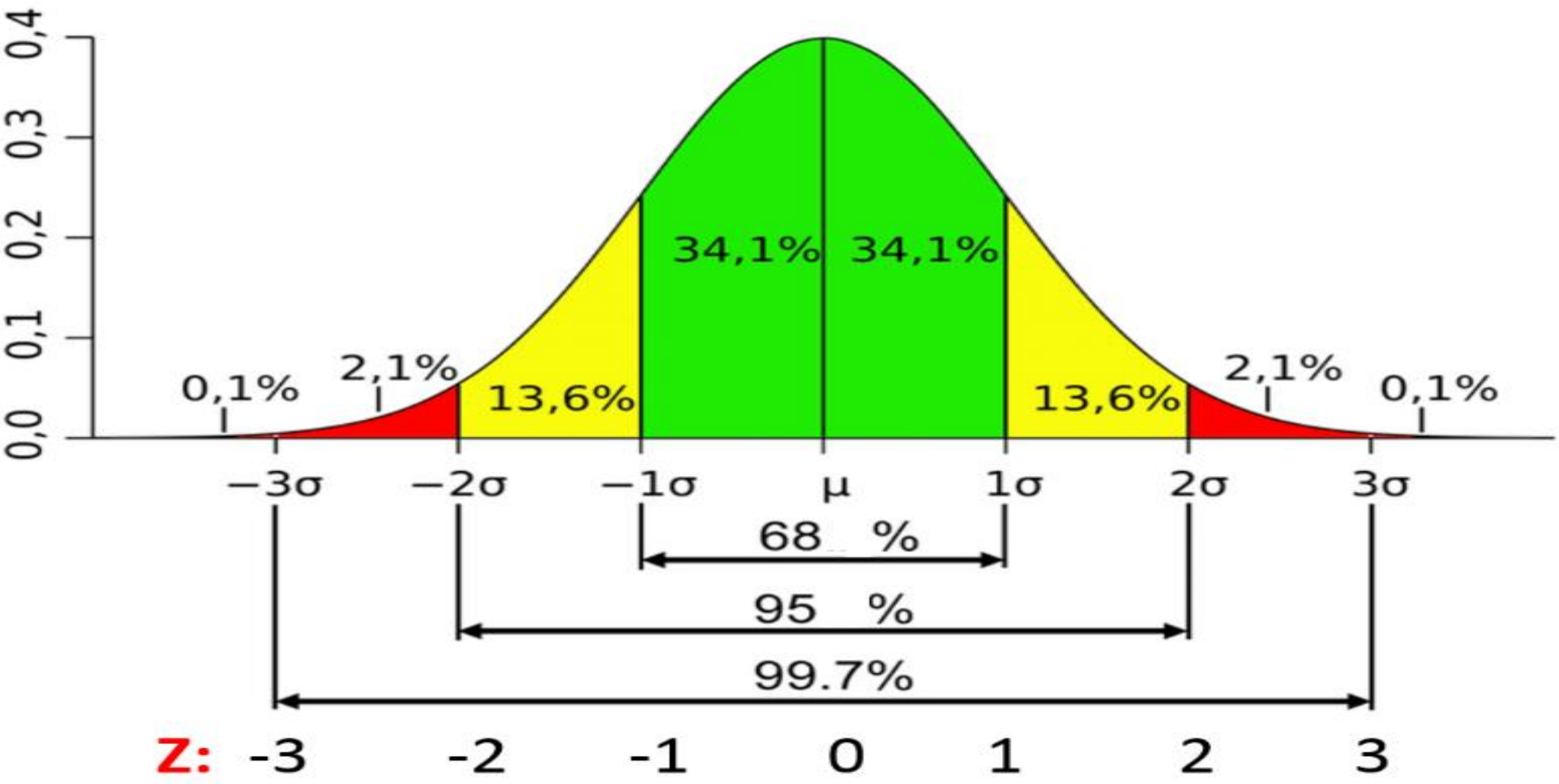
Examples:

- The number of meteorites >1 meter diameter that strike earth in a year
- The number of patients arriving in an emergency room between 10 and 11 pm

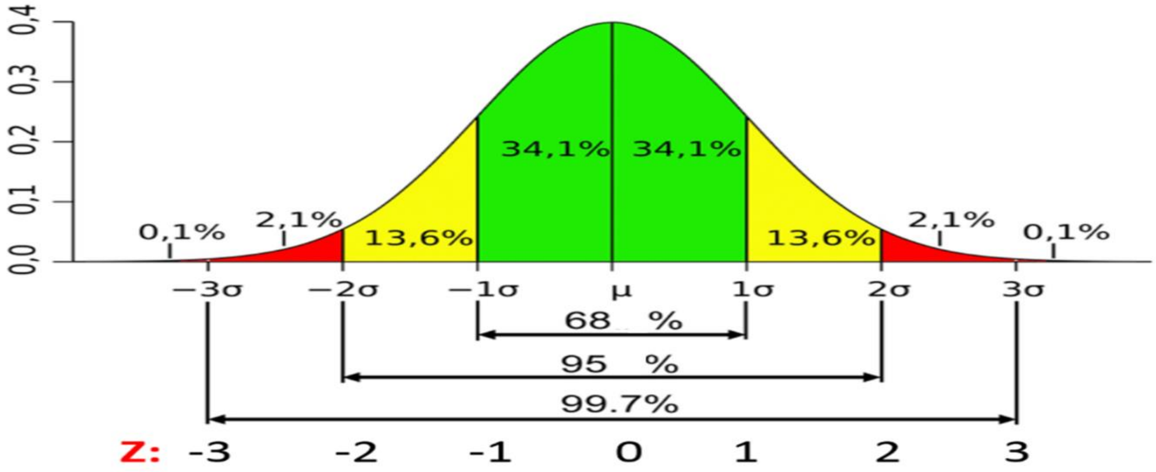
The Normal Distribution (Gaussian, 1809)

The Normal Distribution

- The normal distribution deals with continuous random variables
- Examples: blood pressure, body weight, milk production, body height, concentration of various blood constituents



Properties of the normal distribution



① It is symmetrical around the mean "bell-shaped".
② Its mean, median and mode are equal.
③ Continuous and more denser at the center (mean).
④ Never touches the x-axis.
⑤ Total area under curve (total probabilities) = 1 or 100% (half of the curve = 0.5 or 50 %).
⑥ The limits (μ –σ) and (μ+σ) contain 68% of the distribution.
⑦ The limits (μ –2σ) and (μ+2σ) contain 95% of the distribution.
⑧ The limits (μ –3σ) and (μ+3σ) contain 99.7% of the distribution.
⑨ The normal distribution is completely described or shaped by mean (μ) and the standard deviation (σ).

Effect of mean (μ)

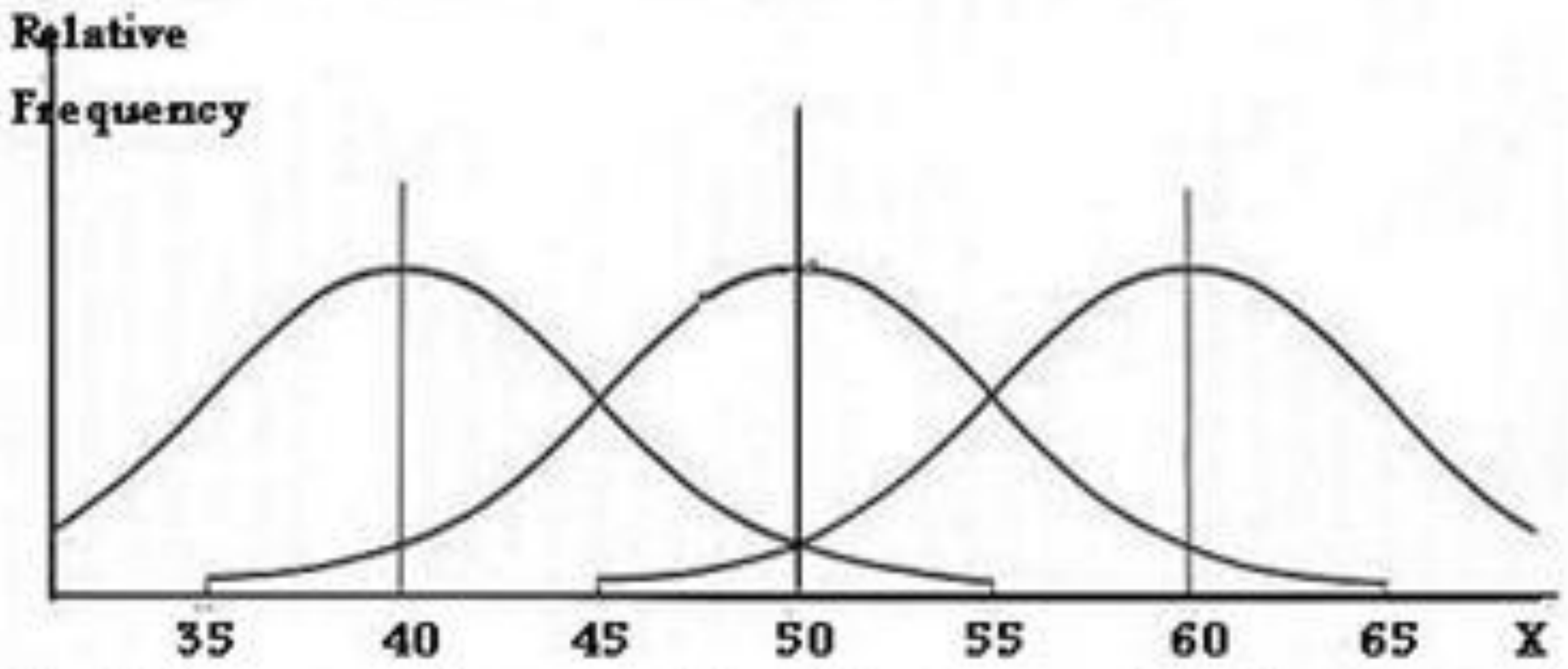


Fig 2: Three normal curves with different means but the same variance

- If σ unchanged, increasing the value of the mean, shifts the curve horizontally to the right and vice versa .

Effect of standard deviation (σ)

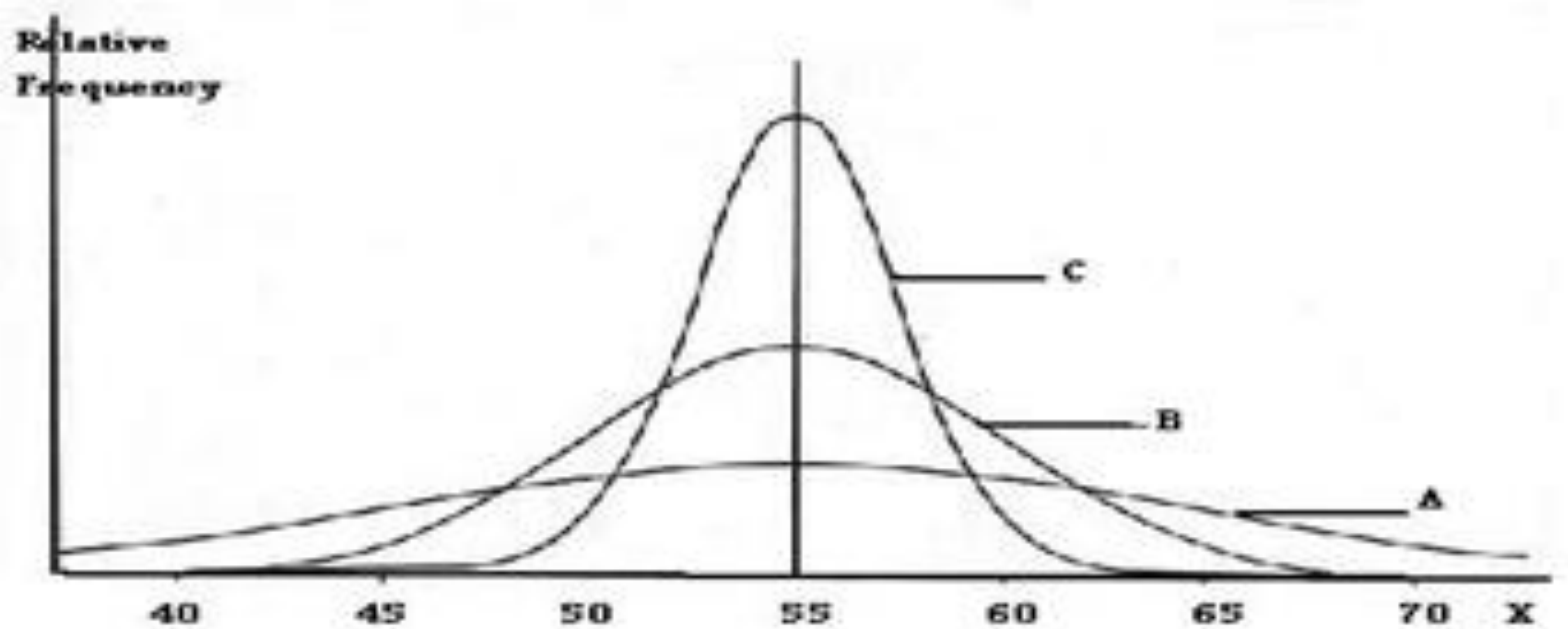


Fig. 3. Three normal curves with different variances but the same mean

- A decrease in the σ of the curve makes the curve more thinner, taller and peaked.
- Conversely, an increase in the standard deviation makes the curve more fatter, shorter and flatter.

The standard normal distribution "ND (0,1)"

- The standard normal distribution has $\mu = 0$ and $\sigma = 1$.
- We can change any normal distribution curve with any μ and any σ to standard normal distribution curve by using this equation

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

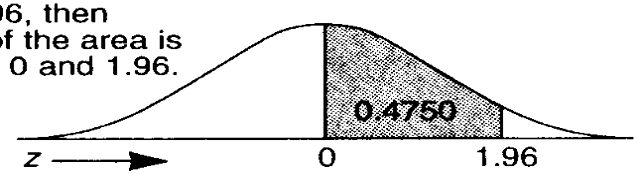
- X is the value or mean of sample
- µ is the population mean
- Σ is the SD of the population

The standard normal distribution "ND (0,1)"

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

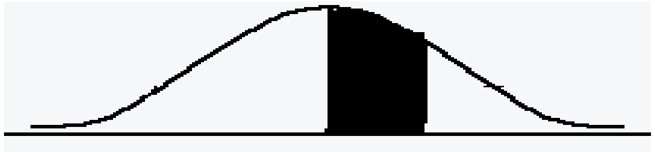
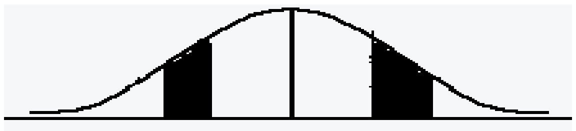
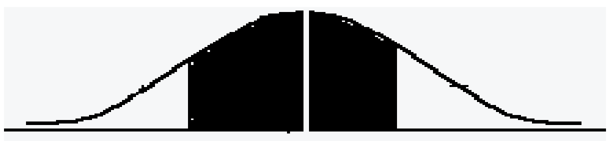
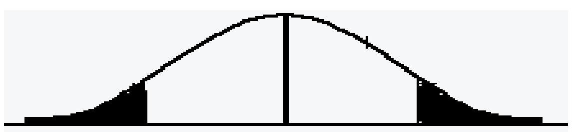
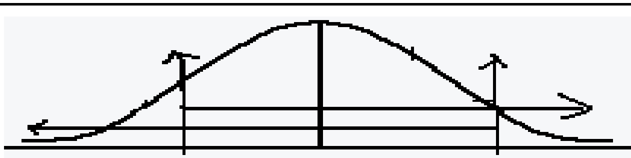
- We can measure the area (probability) under the curve by this equation
- The units on this scale are the values of σ below and above μ
- The cumulative normal frequency distribution (Z table) will be used to measure the area under the standard curve from **zero** to **Z**
- Or to find the probability between **zero** and a specified value of **Z**

Example:
If z = 1.96, then
0.4750 of the area is
between 0 and 1.96.



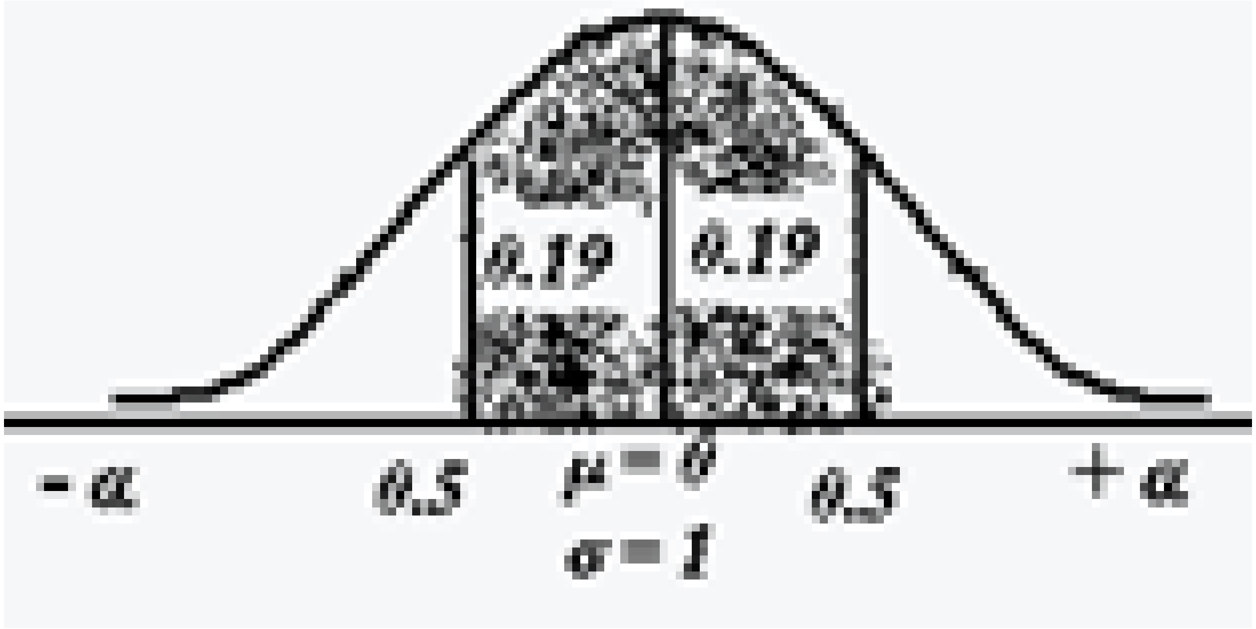
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

How to measure the area under the curve from the Z table in case of given Z value

Situation	Instructions
 Between zero and any number	**Look up** the area in the table
 Between two positives, or Between two negatives	**Look up** both areas in the table and **Subtract** the smaller from the larger.
 Between a negative and a positive	**Look up** both areas in the table and add them together
 Less than a negative, or Greater than a positive	**Look up** the area in the table and **subtract from 0.5**
 Greater than a negative, or Less than a positive	**Look up** the area in the table and **add** to 0.5

1) Examples of given Z-value to find an area

1. Find the prob. that a random value Z lies in the interval - 0.5 and + 0.5??

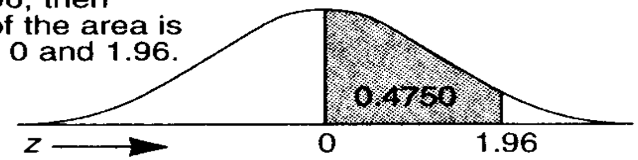


Procedure:

P (Z between Zero and - 0.5) = 0.19.

P (Z between 0.5 and -0.5) = 0.19 + 0.19 = 0.38

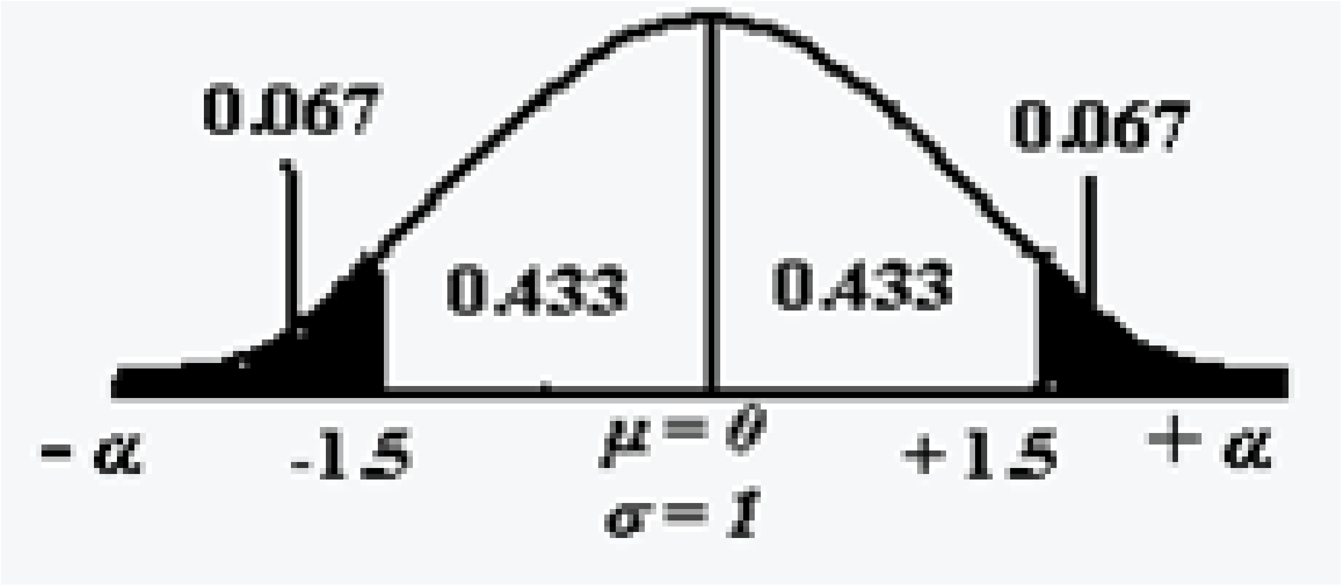
Example:
If z = 1.96, then
0.4750 of the area is
between 0 and 1.96.



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

2) Examples of given Z-value to find an area

2. What is the probability to get Z < -1.5 or > +1.5?



Procedure:

P(Z >1.5) = 0.5 – P(Z between zero and 1.5) = 0.500 - 0.433 = 0.067

P(Z < -1.5) = 0.5 – P(Z between zero and -1.5) = 0.500 - 0.433 = 0.067

P (Z< -1.5 or Z > +1.5) = .067 + .067 = 0.134

How to solve simple normal distribution problems using the standard normal curve

There are two types of simple normal distribution problems

A) Given an X-value to find an area

- Draw a picture
- Calculate the Z score
- Use the table to look up the area

B) Given an area to find an X-value

- Draw a picture
- Use the table to look up the Z score
- Calculate the X-value

A) Examples of given X-value to find an area

- It is essential to transform given **X** value of the normal distribution to **Z** value of the standard normal distribution.

- Deviations from the mean expressed by σ units distribute normally with a μ equals zero and variance equals one.

$$Z = \frac{X - \mu}{\sigma}$$

<u>**Example (1):**</u>

The mean packed cell volume (PCV) of normal cat population approximates follow a normal distribution with a mean 0.37 ml/ml and a standard deviation of 0.066 ml/ml.

(a) What the percentage of cats has values above 0.40 ml/ml?

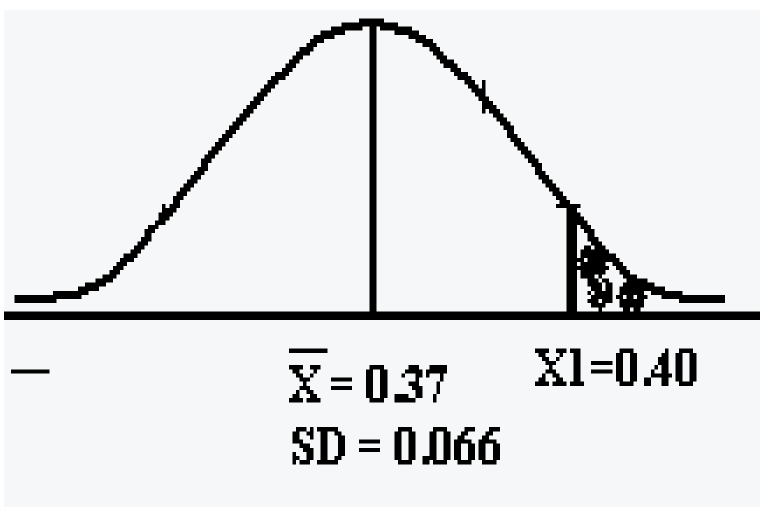
Solution

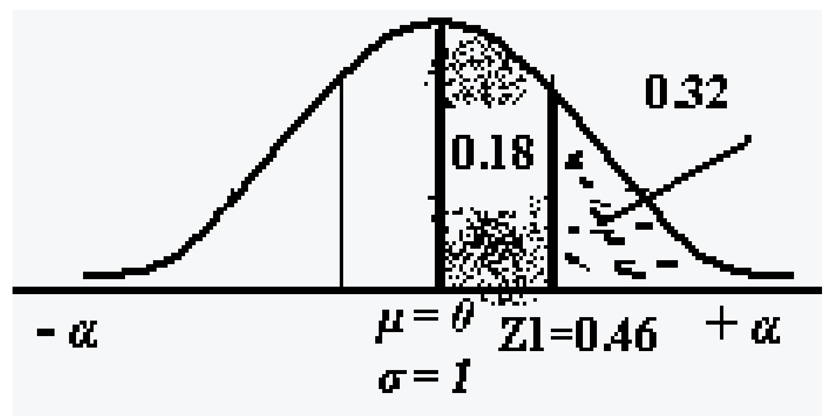
$$\text{(a)}\quad Z_1 = \frac{0.40 - 0.37}{0.066} = 0.46$$

- Reference to Table A gives the area between Z_1 and zero as **0.18** (approx.)

- Area > Z_1 = 0.5 - 0.18 = 0.32 i.e. 32 % of cats have values >0.40 ml/ml.

(a)





Thanks for your attention