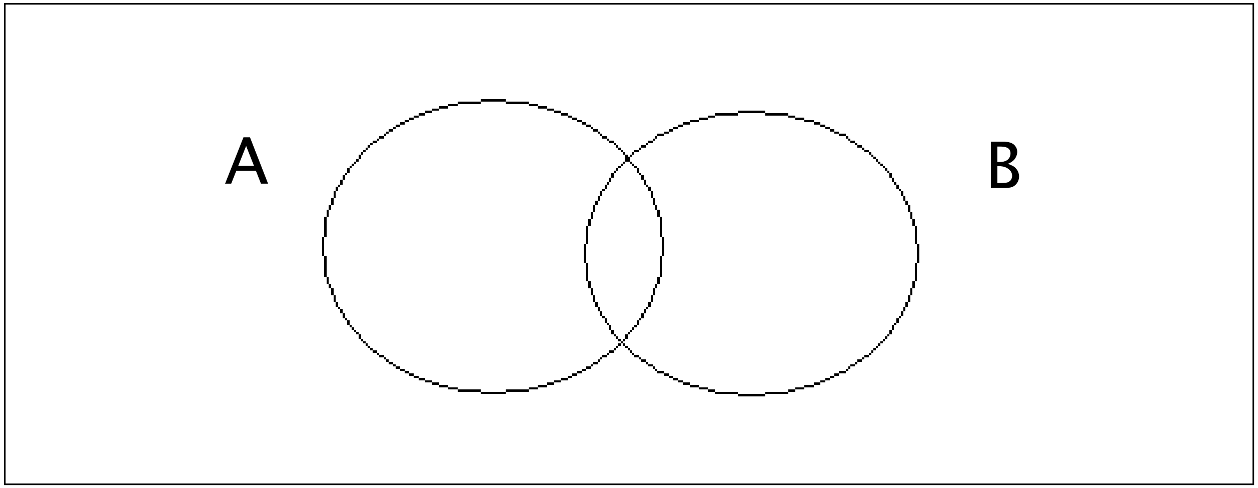
Review

Probability

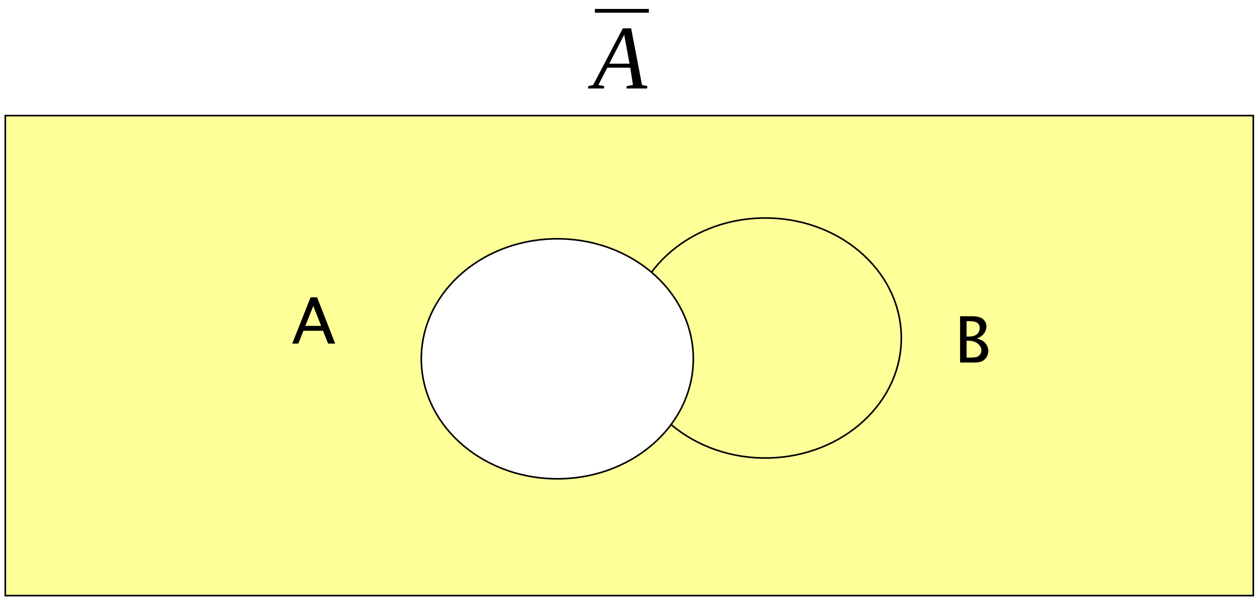
Events

- Definition: any collection of outcomes of an experiment.
- Events consisting of single outcomes in the sample space are called elementary or simple events.
- Events consisting of more than one outcome are called compound events.

Venn Diagrams

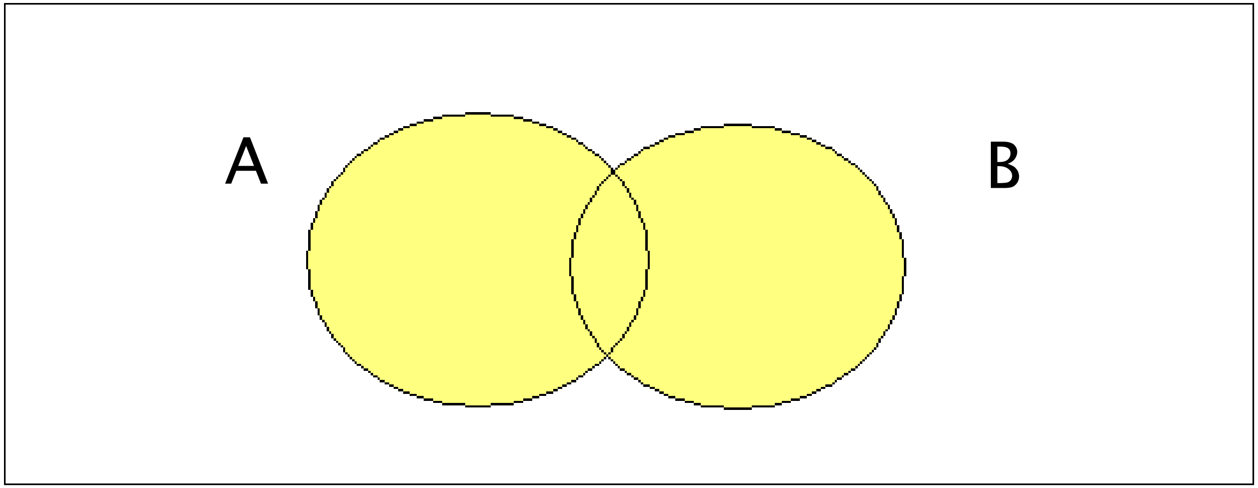


Venn Diagrams



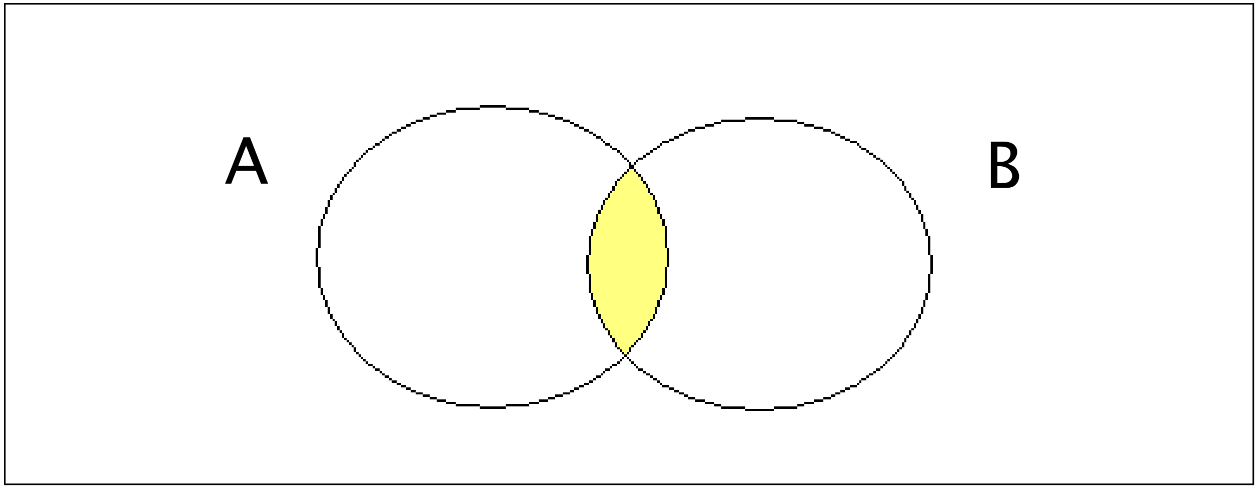
Venn Diagrams

$$A \cup B$$



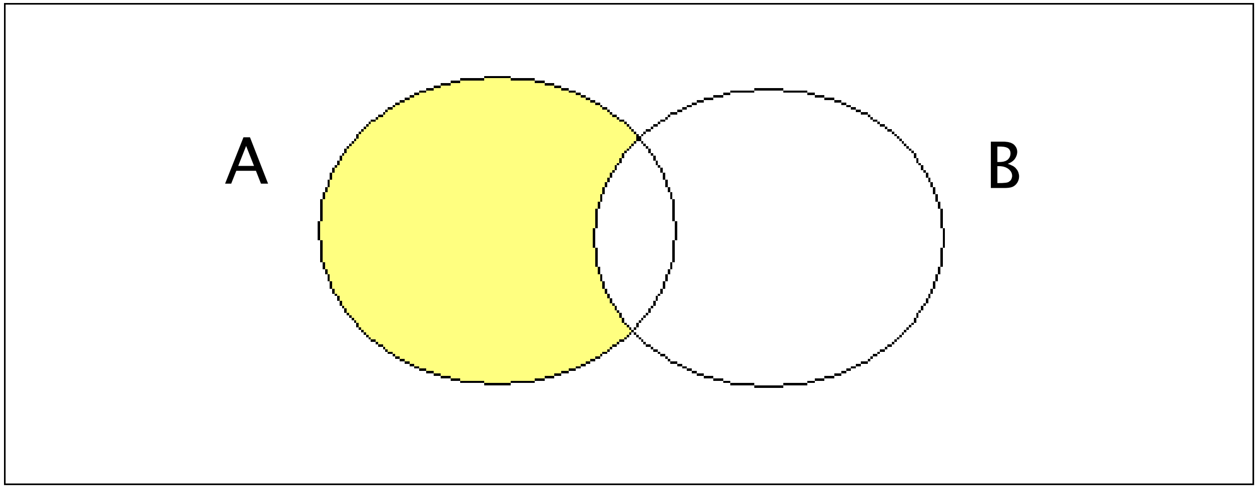
Venn Diagrams

$$A \cap B$$



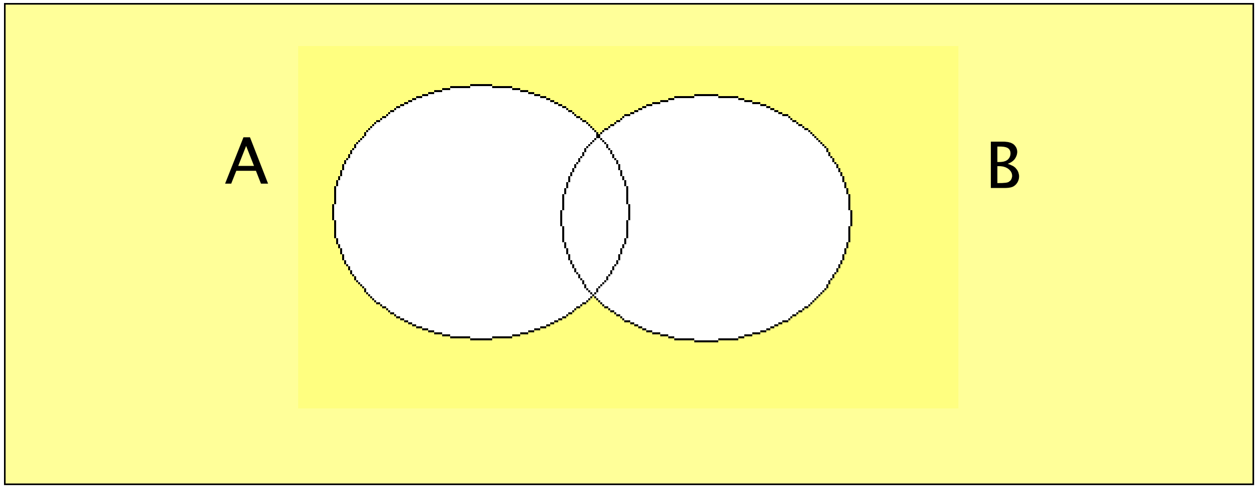
Venn Diagrams

$$A - B = A \cap \overline{B}$$



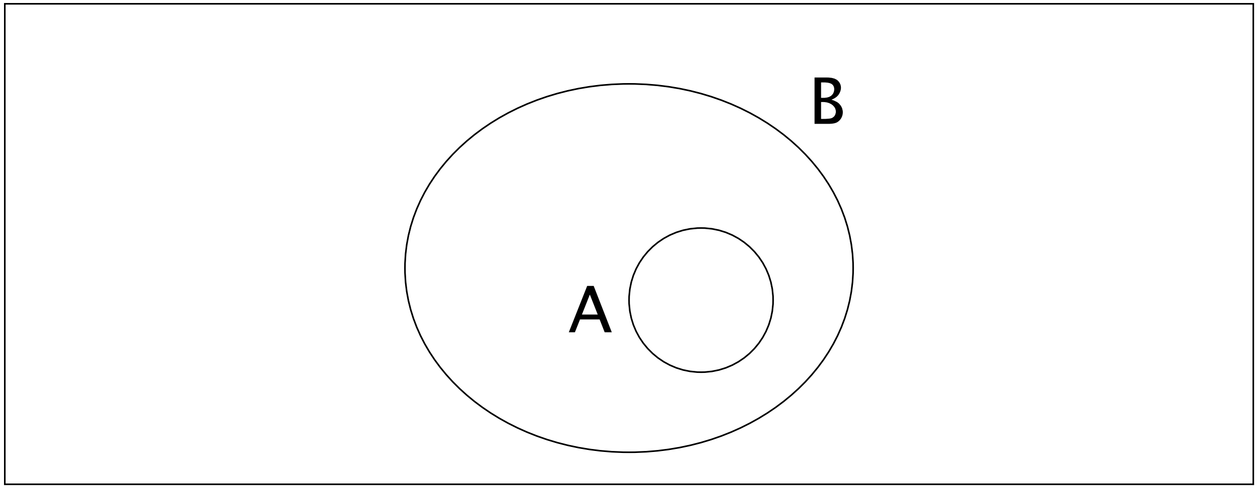
Venn Diagrams

$$\overline{A} \cap \overline{B}$$



Venn Diagrams

$$A \subseteq B$$



Relationships among events

- If A and B are two events in the sample space S, then
  - $A \cup B$ (A union B) = 'either A or B occurs or both occur'
  - $A \cap B$ (A intersection B) = 'both A and B occur'
  - $A \subseteq B$ (A is a subset of B) = 'if A occurs, so does B'
  - A' or $\overline{A}$ = 'event A does not occur'
  - ϕ (the empty set) = an impossible event
  - S (the sample space) = an event that is certain to occur

Probability in discrete space

Probability Axioms:

$$P(A) \geq 0$$

$$P(\Omega) = 1$$

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$$

Probability in discrete space

Lemma:

$$P(A) = 1 - P(\overline{A})$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Probability in discrete space

Example 1:

Jerry and Susan have a joint bank account.

Jerry goes to the bank 20% of the days.

Susan goes there 30% of the days.

Together they are at the bank 8% of the days.

What is the probability that in a particular day at least one of them is visiting the bank?

42%

Events – class assignments

An insurance company offers four different deductible level- none(N), low(L), medium(M), and high (H). for its homeowner's policyholders, and three different for its automobile policyholders. Given the following random sample of policyholders.

- What is the probability that the individual has a medium auto deductible and a high homeowner's deductible?
- What is the probability that the individual has a medium auto deductible ?
- What is the probability that the individual has a high homeowner's deductible ?
- What is the probability that the individual is in the same category for both auto and homeowner's deductibles?
- What is the probability that the individual is in two different categories?

Home

Auto	N	L	M	H
L	40	60	50	30
M	70	100	200	100
H	20	30	150	150

Conditional Probability

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

Events – class assignments

Another insurance company offers four different deductible level- none(N), low(L), medium(M), and high (H). for its homeowner's policyholders, and three different for its automobile policyholders. Given the following random sample of policyholders.

Home

Auto	**N**	**L**	**M**	**H**
L	20	40	80	60
M	50	100	200	150
H	30	60	120	90

Independent Events

A and B are independent (no additional information) if:

$$P(A \mid B) = P(A) \qquad \text{or} \qquad P(A \cap B) = P(A)P(B)$$

Example

Home

Auto	N	L	M	H	
L	4%	6%	5%	3%	18%
M	7%	10%	20%	10%	47%
H	2%	3%	15%	15%	35%
	13%	19%	40%	28%	100%

- What is P(Auto=H/Home=H)=?

15% **/** 28% = 53.57%

Example

You roll 2 dice

P(First Die =1) =?
P(Total = 7) = ?
P(First Die=1/Total =7)= ?
P(First Die=1/Total=8)=?
Does Total=7 give any additional information about the first die?

Example

		Die 2					
		1	2	3	4	5	6
	1	2	3	4	5	6	7
	2	3	4	5	6	7	8
Die 1	3	4	5	6	7	8	9
	4	5	6	7	8	9	10
	5	6	7	8	9	10	11
	6	7	8	9	10	11	12

Law of Total Probability

Let $B_1, B_2, \ldots, B_n$ be disjoint sets

$$P(B_1 \cup B_2 \cup \ldots \cup B_n) = \Omega$$

Then, for each event A

$$P(A) = P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + \ldots + P(A \mid B_n)P(B_n)$$

Bayes' Formula

$$P(B_1|A) = \frac{P(B_1 \cap A)}{P(A)} = \frac{P(B_1)P(A|B_1)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)}$$

Bayes' Formula

The probability of an event, before obtaining information, is called "prior probability".

The probability of an event, after obtaining information, is called "posterior probability".

Using Bayes' Formula we can calculate posterior probabilities using signals.

Medical Testing

A new and very accurate test has been developed for the detection of a disease (e.g. Cancer). The test is 99.9 percent accurate with error rates of 0.1 % for both types of errors. In other words:

- Out of 1,000 sick patients, the test misses only 1 patient, and
- Results in only 1 false positive for every 1,000 healthy individuals.

The prevalence rate of the disease in the general populations is about 1,000 per million. Given the positive result of the test, what is the probability that the individual is in fact sick?

The second test of the patient is also positive. Now, what is probability that the individual is in fact sick?

Medical Testing

	Sick	Not-sick	Total
Positive	999	999	1,998
Not Positive	1	998,001	998,002
	1,000	999,000	1,000,000

P(sick/positive)=999/1998=50%

Medical Testing

The second test of the patient is also positive. Now, what is probability that the individual is in fact sick?

	Sick	No-sick	Total
Positive	998	1	999
Not Positive	1	998	999
	999	999	1,998

P(sick/1st & 2nd positive)=998/999=99.9%

Game Show

You are the finalist in a game show and you have the option of choosing one of the three briefcases. One of the briefcases contains a $1 million prize, and the other two are empty. To make the show more exciting, after you choose a briefcase, the host opens one of the remaining two briefcases, shows you that it is empty, and gives you the option of either keeping your original briefcase, or switching to the remaining briefcase. Should you switch?

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \ldots + P(A|B_n)P(B_n)$$

A= Choosing the Prize B= Choosing Empty

Policy 1: Always Switch

P(wining)=P(prize)* P(wining/prize)+P(empty)*P(wining/empty).

P(wining)= 1/3 * 0 + 2/3 * 1= 2/3

Policy 2: Do Not Switch

P(wining)=1/3*1 + 2/3 * 0 = 1/3

Expectation of function variables

$$E[f(x)] = \sum_x f(x)p(x)$$

Expectation of a random variables

$$\mu = E[X] = \sum_x xp(x)$$

E[a]=a; E[aX]=aE[X]

Variance of a random variables

$$\mathrm{Var}(X) = E[(X-\mu)^2]$$

$$\sigma^2 = \sum_x (x-\mu)^2 p(x)$$

Var(a)=0; Var(aX)=a^2Var(X)

Example

You roll a die.

a. What are the possible outcomes?

b. What are the probabilities for each outcome?

c. What's the expected value?

d. What is the variance? The standard deviation?

Example

	X	P	X*P	(X-Mue)^2	((X-Mue)^2)*p
	1	1/6	0.167	6.25	1.04
	2	1/6	0.333	2.25	0.38
	3	1/6	0.500	0.25	0.04
	4	1/6	0.667	0.25	0.04
	5	1/6	0.833	2.25	0.38
	6	1/6	1	6.25	1.04
Overall	**21**		**3.5**		**2.92**

Std= sqrt(2.92)=1.70

Example

You roll 2 dice.

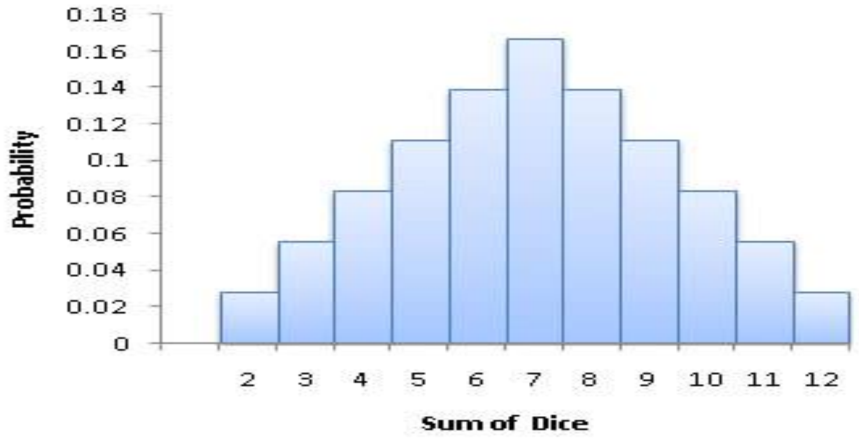
a. What are the possible outcomes? Show the histogram

b. What are the probabilities for each outcome?

c. What’s the expected value?

d. What is the variance? The standard deviation?

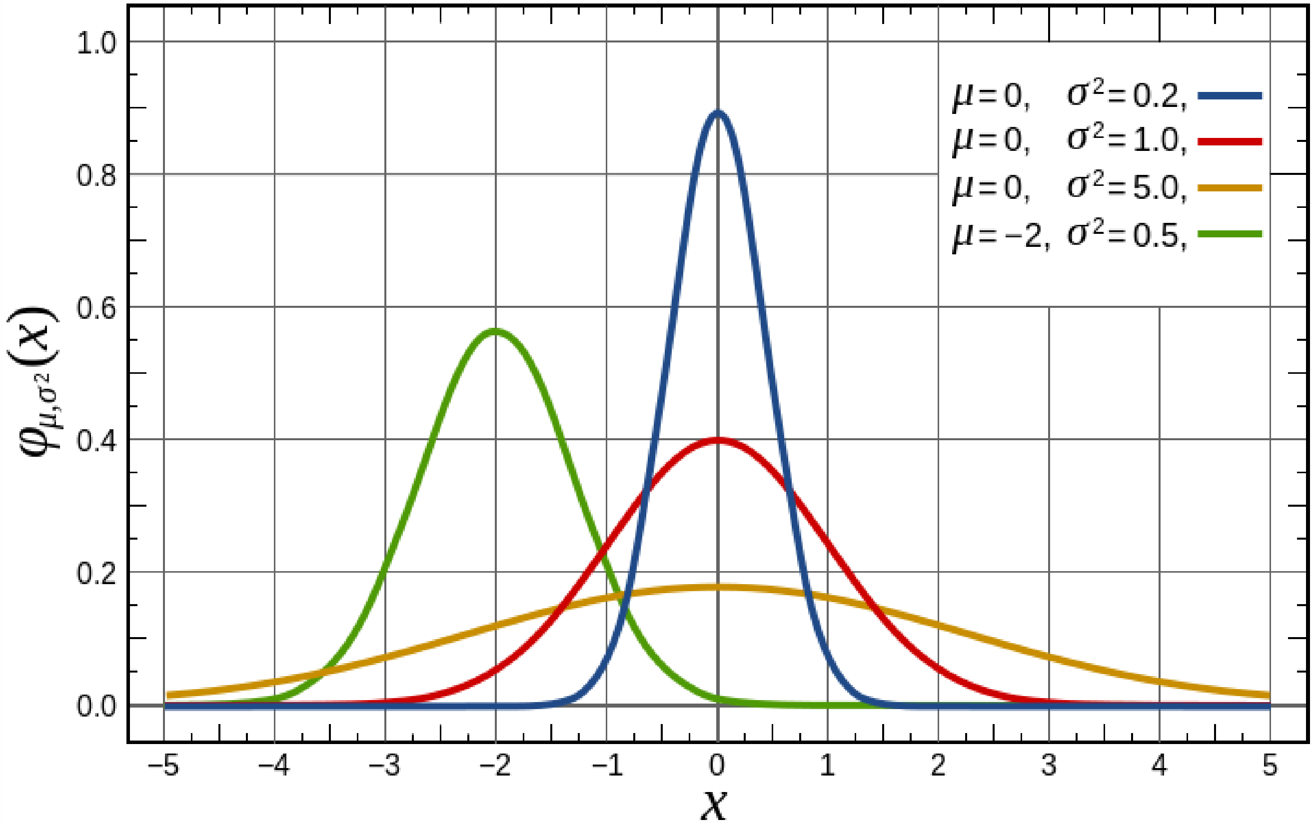
Example



X	P	X*P	(X-Mue)^2	((X-Mue)^2)*p
2	1/36	0.056	25.00	0.69
3	2/36	0.167	16.00	0.89
4	3/36	0.333	9.00	0.75
5	4/36	0.556	4.00	0.44
6	5/36	0.833	1.00	0.14
7	6/36	1.167	0.00	0.00
8	5/36	1.111	1.00	0.14
9	4/36	1.000	4.00	0.44
10	3/36	0.833	9.00	0.75
11	2/36	0.611	16.00	0.89
12	1/36	0.333	25.00	0.69
Overall		**7.000**		**5.83**

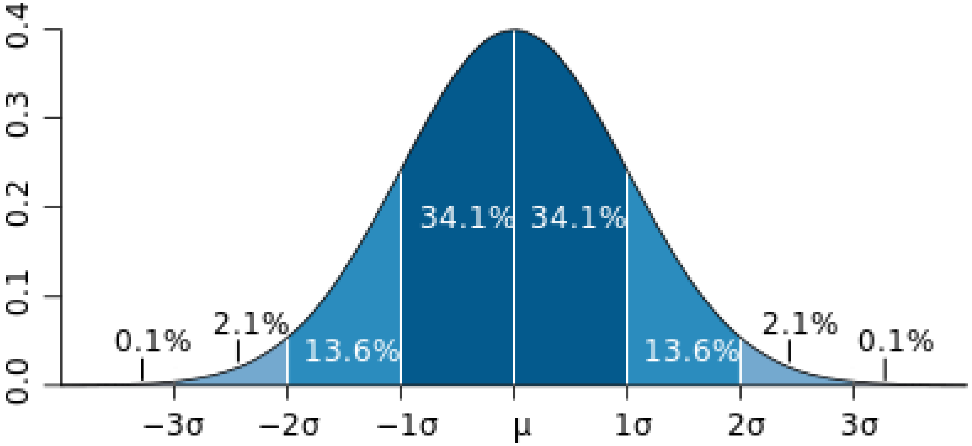
Mean	7
Var	5.83
Std	2.42

Normal Distribution

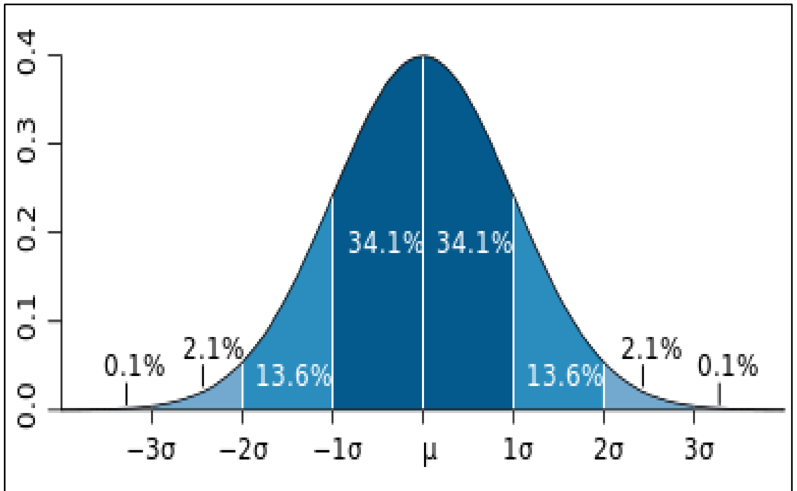


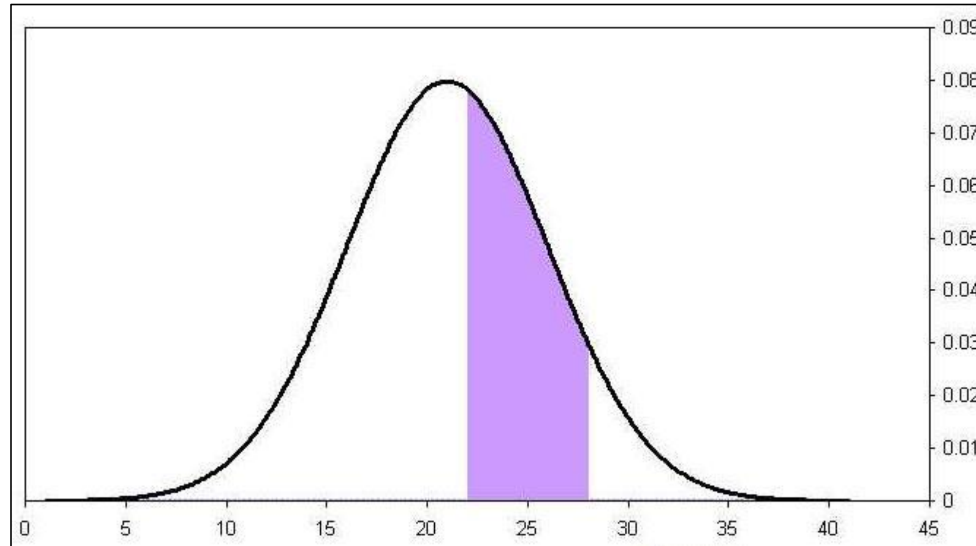
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Standard Normal Distribution



Z- Score





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

$\mu =$ Mean

$\sigma =$ Standard Deviation

Example

The amount of distilled water dispensed by a certain machine is normally distributed with mean value 64 oz and standard deviation 0.78 oz. What container size (c) will ensure that overflow occurs only one-half of one percent of the time?

C=66