CS204: Discrete Mathematics

Ch 9. Discrete Probability (3)

Sungwon Kang

Acknowledgement

- Michael P. Frank, University of Florida, COT 3100, Applications of Discrete Structures, Lecture Slides
- [Rosen 19] Kenneth H. Rosen, for Discrete Mathematics & Its Applications (8th Edition), Lecture slides
- [Hunter 11] David J. Hunter, Essentials of Discrete Mathematics, 2nd Edition, Jones & Bartlett Publishers, 2011, Lecture Slides
- Milos Hauskrecht, Pittsburgh University, Lecture Slides

Ch 7. Discrete Probability

7.1 An Introduction to Discrete Probability

7.2 Probability Theory

7.3 Bayes' Theorem

7.4 Expected Value and Variance ☚

Table of Contents

5. Integer random variables, Bernoulli trial

Random Variable

- **Random process**: an experiment in which
  we know what outcomes could happen (sample space),
  but we don't known which particular outcome will happen

Definition A **random variable** is a function f: S $\rightarrow$R
from the **sample space of an experiment**
to the set of real numbers.

- A ***random variable*** is a numerical result of a random experiment or process.

  => A random variable assigns a number to each possible outcome.

The probability distribution of a random variable X on the sample space S is a set of pairs (r, P(X=r)) for all r in S
where r is a number representing an outcome
and P(X=r) is the probability that X takes the value r.

Example 4.42 Rolling two standard six-sided dice is a random experiment.
The sum of the values on the two dice is a random variable X.
Recall the following example.

Example: If you roll two standard six-sided dice, what is the probability you roll 10 or less?

Solution

Compute the probability of rolling more than 10: There are two ways to roll 11 and one way to roll 12, so the probability of rolling more than 10 is $(2+1)/36 = 1/12$. *Thus the probability of rolling 10 or less is* $1 - 1/12 = 11/12$.

$P(X = 11) = 2/36$
$P(X = 12) = 1/36$
$P(X \le 10) = P(2 \le X \le 12) - \{P(X = 11) + P(X = 12)\}$

Random Variable

Example

Let S be the outcomes of a two-dice roll.

Let random variable X denote the sum of outcomes.

(1,1) $\rightarrow$ X = 2

(1,2) and (2,1) $\rightarrow$ X = 3

(1,3), (3,1) and (2,2) $\rightarrow$ X = 4

…

Probability Distribution of X:

2 $\rightarrow$ P(X=2) = 1/36,

3 $\rightarrow$ P(X=3) = 2/36,

4 $\rightarrow$ P(X=4) = 3/36,

…

12 $\rightarrow$ P(X=12) = 1/36

Bernoulli trial

- Suppose that an experiment can have only two possible outcomes.

  Example. When a coin is flipped, the possible outcomes are heads and tails.

- Each performance of an experiment with two possible outcomes is called a *Bernoulli trial*.
  - Possible outcome of a Bernoulli trial is called a **success** or a **failure.**
  - If p is the probability of a success and q is the probability of a failure,

$$p + q = 1$$

- Many problems can be solved by determining *the probability of k successes* when an experiment consists of n mutually independent Bernoulli trials.

Suppose we flip a coin flip repeatedly.

P(heads) = 0.6 and P(tails) = 0.4.

Each coin flip is independent of the previous flip.

- What is the probability of seeing HHHHH?

  $P(HHHHH) = 0.6^5$

- What is the probability of seeing TTHHT?

  $P(TTHHT) = 0.4^2 \times 0.6^2 \times 0.4 = 0.6^2 \times 0.4^3$

- What is the probability of seeing two heads and three tails?

  The number of two heads and three tails combinations = C(5,2)

  $P(\text{two-heads-three-tails}) = C(5,2) \times 0.6^2 \times 0.4^3$

A variant of a repeated coin flip problem.

Sample space: The number of occurrences of heads in 5 coin flips.

For example,

TTTTT yields outcome 0

HTTTT or TTHTT yields 1

HTHHT yields 3 …

What is the probability of an outcome i when i ranges from 0 to 5?

P(outcome = 0) = $C(5,0) \times 0.6^0 \times 0.4^5$

P(outcome = 1) = $C(5,1) \times 0.6^1 \times 0.4^4$

P(outcome = 2) = $C(5,2) \times 0.6^2 \times 0.4^3$

P(outcome = 3) = $C(5,3) \times 0.6^3 \times 0.4^2$

P(outcome = 4) = $C(5,4) \times 0.6^4 \times 0.4^1$

P(outcome = 5) = $C(5,5) \times 0.6^5 \times 0.4^0$

What is

$$\sum_{0 \le i \le 5} P(\text{outcome} = i) \; ?$$

$$(0.6 + 0.4)^5 = ?$$

Recall

The Binomial Theorem

Theorem

Let j and k be nonnegative integers such that $j + k = n$. The coefficient of the $a^j b^k$ term in the expansion of $(a+b)^n$ is $C(n, j)$.

Corollary

The Binomial Theorem.

$$\begin{aligned}(a+b)^n &= \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \cdots \\ &\quad + \binom{n}{j}a^{n-j}b^j + \cdots + \binom{n}{n}b^n\end{aligned}$$

Theorem
In *n* independent Bernoulli trials,
with probability of success p
and probability of failure $q = 1 - p$,

the probability of exactly *k* successes is

$$C(n, k) \times p^k \times q^{n-k}$$

Exercise. A coin is biased so that the probability of heads is 2/3.
What is the probability that exactly four heads come up
when the coin is flipped seven times, assuming that the flips are independent?

Exercise. A coin is biased so that the probability of heads is 2/3. What is the probability that exactly four heads come up when the coin is flipped seven times, assuming that the flips are independent?

$$C(7,4)\,(2/3)^4(1/3)^3 = \frac{35 \cdot 16}{3^7} = \frac{560}{2187}$$

Exercise. A coin is biased so that the probability of heads is 2/3.
What is the probability that exactly four heads come up
when the coin is flipped seven times, assuming that the flips are independent?

Solution

When a coin is flipped 7 times, #possible outcomes = $2^7 = 128$

#ways 4 out of 7 flips are heads = C(7, 4).

Because the 7 flips are independent,
the probability of each such outcome (four heads and three tails) = $(2/3)^4(1/3)^3$.

Consequently, the probability that exactly four heads appear is

$$C(7,4)\,(2/3)^4(1/3)^3 = \frac{35 \cdot 16}{3^7} = \frac{560}{2187}$$

6. Expected Value, Linearity of Expectation, Variance

Expected value

Definition

Let $x_1, x_2, \ldots, x_n$ be all of the possible values of a random variable X. Then X's *expected value* $E(X)$ is the sum

$$\sum_{i=1}^{n} x_i \cdot P(X = x_i).$$

That is,
$E(X) = x_1 \cdot P(X = x_1) + x_2 \cdot P(X = x_2) + \cdots + x_n \cdot P(X = x_n).$

What is $\sum_{i=1}^{n}$ P(X = xi) ?

Example 1. Outcomes of rolling a die: 1, 2, 3, 4, 5, 6

E(X) = ?

$$
\begin{aligned}
E(X) &= 1*P(X=1) + 2*P(X=2) + 3*P(X=3) + 4* P(X=4) + 5*P(X=5) + 6*P(X=6) \\
&= 1*1/6 + 2*1/6+3*1/6 + 4*1/6 + 5*1/6 +6* 1/6 \\
&= 7/2
\end{aligned}
$$

Example 1. Outcomes of rolling a die: 1, 2, 3, 4, 5, 6

E(X) = ?

$$
\begin{aligned}
E(X) &= 1*P(X=1) + 2*P(X=2) + 3*P(X=3) + 4*P(X=4) + 5*P(X=5) + 6*P(X=6) \\
&= 1*1/6 + 2*1/6 + 3*1/6 + 4*1/6 + 5*1/6 + 6*1/6 \\
&= 7/2
\end{aligned}
$$

Example 2.

Flip a fair coin 3 times.

The outcome X of the trial is the number of heads.

What is the expected value of the trial?

Possible results = {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}

X= 3, 2, 2, 2, 1, 1, 1, 0

E(X) = ?

$$E(X) = 1/8 \times (3 \times 1 + 2 \times 3 + 1 \times 3 + 0 \times 1) = 12/8 = 3/2$$

Average

- The term ‘average’ expresses that something is statistically the norm.

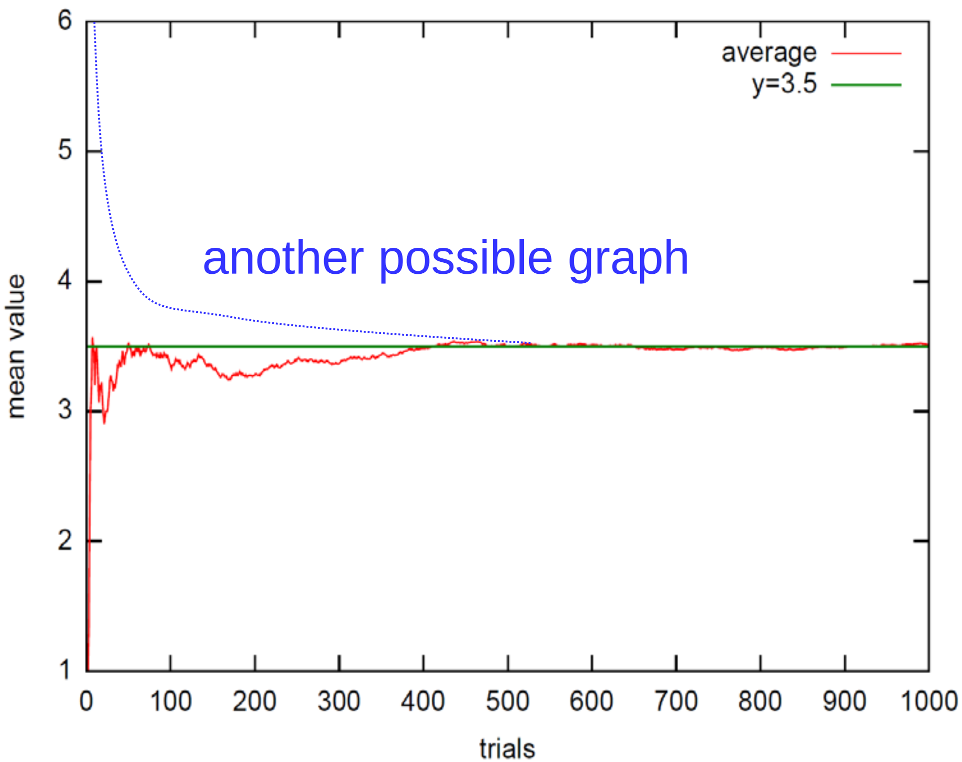
- Various measures of ‘average’:
  - The *(arithmetic) mean* of a finite set of real numbers is the sum of th e real numbers in the set divided by the number of real numbers in t he set.
    - Given two positive real numbers x and y,
      - their *arithmetic mean* is (x+y)/2.
      - their *geometric mean* is $\sqrt{(xy)}$.

  - The *median* of a finite set of real numbers is the middle element in t he list when these integers are listed in the order of increasing size.

Average(= Arithmetic Mean) vs. Expected Value

"As the number of trials increases,
the average of the results obtained from them
gets close to the expected value." **- The Law of Large Numbers**

average dice value against number of rolls



Investment problem

You have 100 dollars and can invest into a stock.
The returns are volatile and
you may get either \$120 with probability of 0.4, or
\$90 with probability of 0.6.

What is the expected value of your investment?

$$E(X) = 0.4 \times 120 + 0.6 \times 90$$
$$= 48 + 54$$
$$= 102$$

Is it OK to invest?

Application: playing the lottery (More complicated problem)

In the "Both Ways" version of Australia's "Cash 3" lottery, you pick a three digit number from 000 to 999, and a randomly chosen winning three-digit number is announced every evening at 5:55pm. If you pick a number with three distinct digits, you win \$580 if your number matches the winning number exactly, and you win \$80 if the digits of your number match the digits of the winning number, but in a different order. What is the expected value of the amount of money you win?

Solution

Let X be the amount of money you win. The possible values of X are 0, 580, and 80. Out of 1000 possible winning numbers, only one matches your number exactly, and five $(3! - 1)$ have the same digits, but in a different order. Therefore your expected winnings are

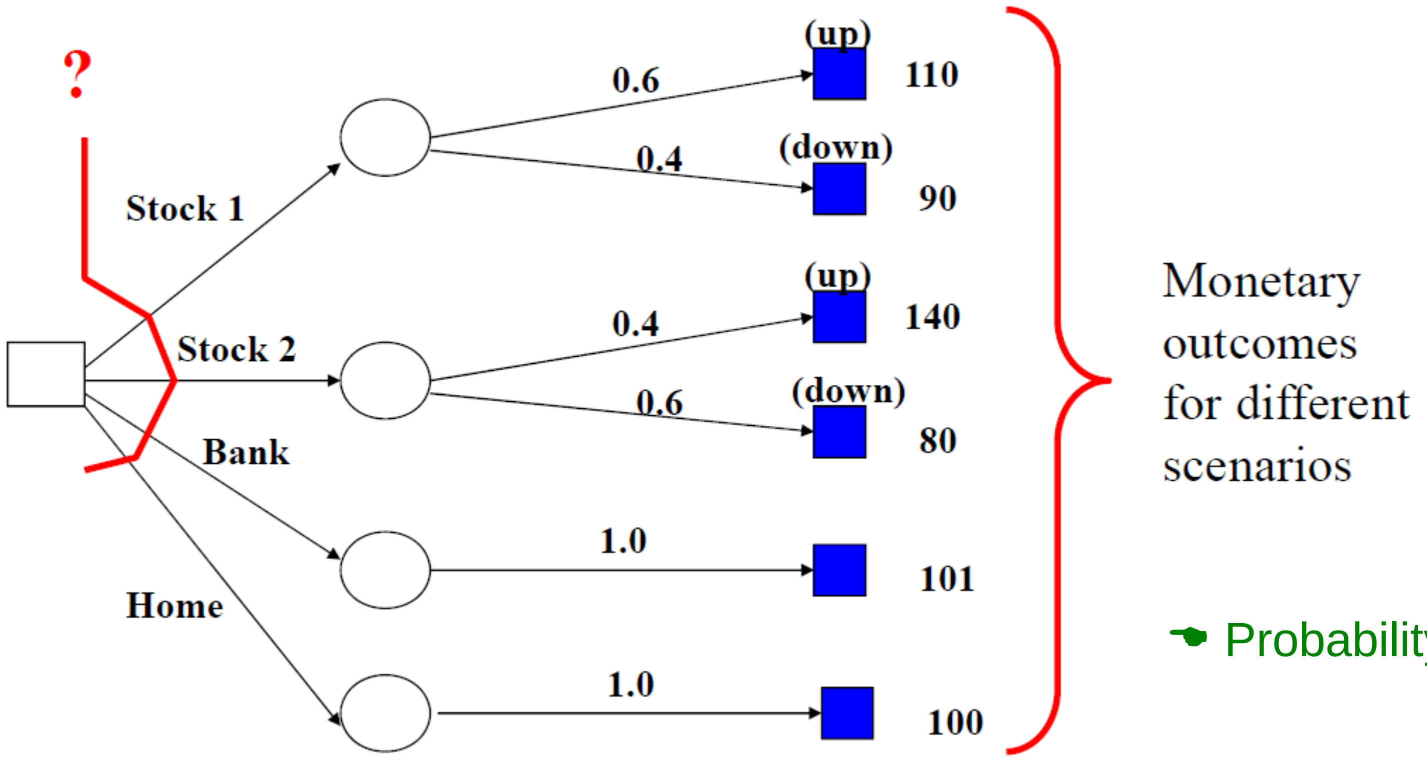
$$
\begin{aligned}
E(X) &= 0 \cdot P(X = 0) + 580 \cdot P(X = 580) + 80 \cdot P(X = 80) \\
&= 0 \cdot \frac{994}{1000} + 580 \cdot \frac{1}{1000} + 80 \cdot \frac{5}{1000} \\
&= 0.98
\end{aligned}
$$

How much would you pay to play?

Decision making

We will invest $100 for 6 months.

We need to make a choice whether to invest in Stock 1 or Stock 2, put money into bank or keep them at home.

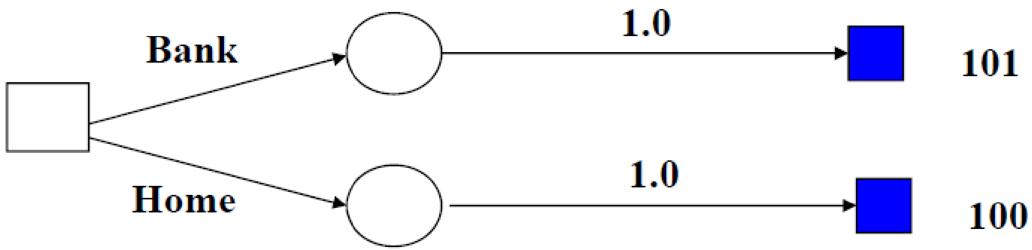


☚ Probability tree

Deterministic outcome (1/2)

Assume the following simplified problem with the Bank and Home choices only.

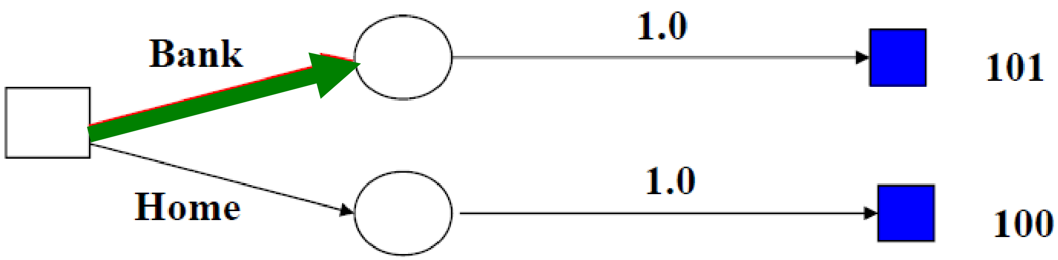
The result is guaranteed – the outcome is deterministic.



What is the rational choice assuming our goal is to make money?

Deterministic outcome (2/2)

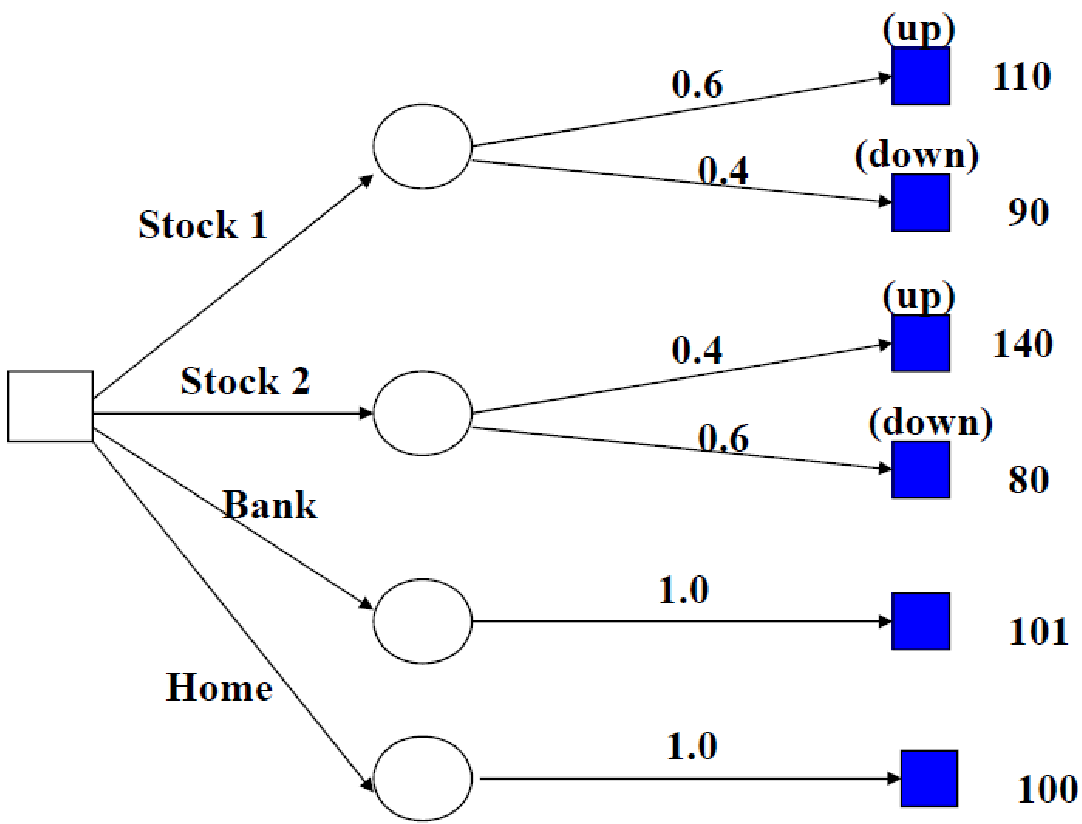
Assume the simplified problem with the Bank and Home choices only.
The result is guaranteed – the outcome is deterministic



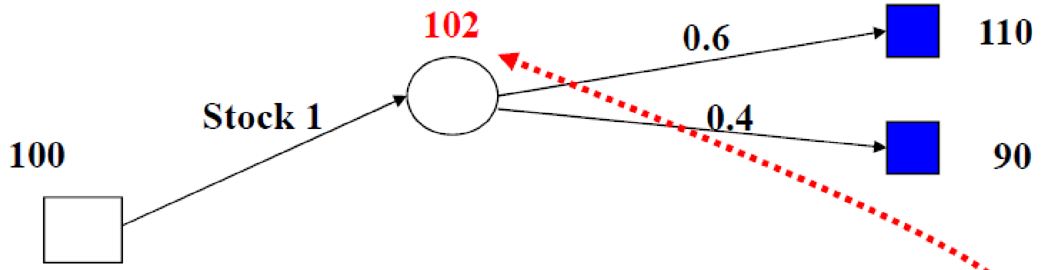
What is the rational choice assuming our goal is to make money?

Answer: Put money into the bank.
The choice is strictly better in terms of the outcome.

Expected value for the outcome of the Stock 1 option (1/2)

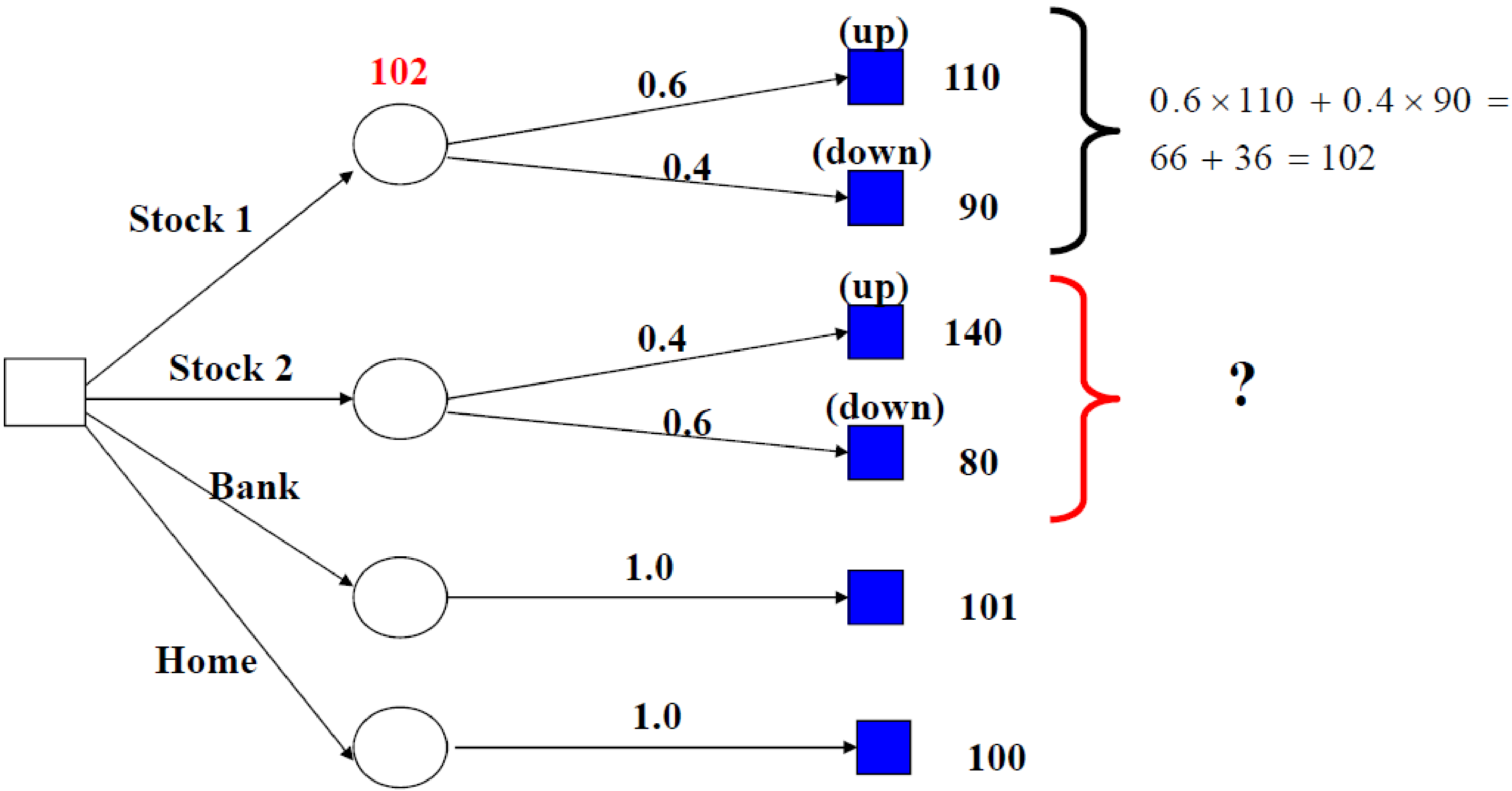


Expected value for the outcome of the Stock 1 option (2/2)



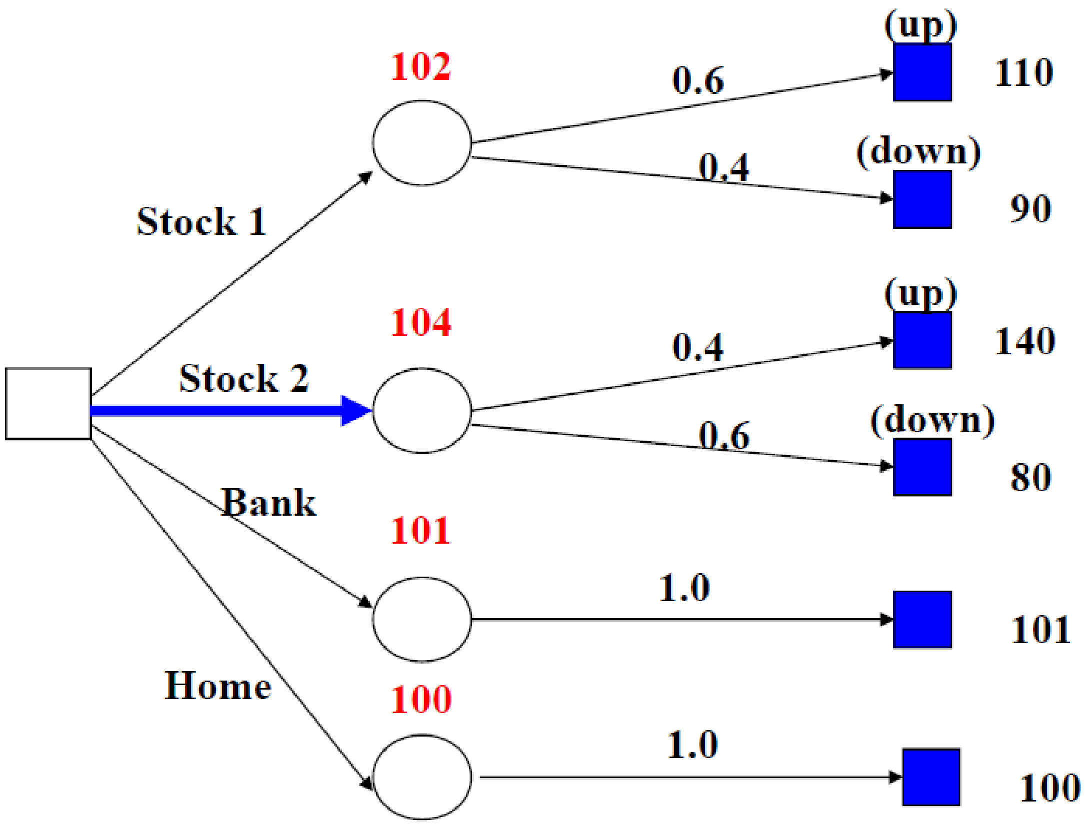
$$0.6 \times 110 + 0.4 \times 90 = 66 + 36 = \boxed{102}$$

Expected value for the outcome of the Stock 2 option



Selection based on expected values

The optimal action is the option that maximizes the expected outcome:



$0.4 \times 140 + 0.6 \times 80$
$= 56 + 48$
$= 104$

Linearity of Expectation

Theorem If Xi, i=1,2,3, ..., n, with a positive integer n, are random variables on S, and a and b are real numbers, then

(1) $E(X1 + X2 + ... + Xn) = E(X1) + E(X2) + ... + E(Xn)$

(2) $E(aX + b) = aE(X) + b$

Example

Roll a pair of dice. What is the expected value of the sum of outcomes?

Approach 1:

Outcomes: (1,1),(1,2), ...,(1,6),(2,1),(2,2), ...,(2,6), ...,(6,1),(6,2), ...,(6,6)

X= 2, 3, ..., 7, 3, 4, ..., 8, ..., 7, 8, ..., 12

Expected value: $\sum_{i=1}^{n} xi \times P(X = xi) =$

$1/36 \times ((2+3+...+7)+(3+4+ ...+8)+ ...+(7+8+...+12)) = 7$

Approach 2 (Using the Linearity Theorem):

$E(x1+x2) = E(x1) + E(x2)$

$E(x1) = E(x2) = 7/2$

$E(x1+x2) = 7$

Variance

- Variance measures how far a data set is spread out.

- The *variance* of a random variable *X* is :

$$\mathbf{V}[X] = \sum_{s \in S} (X(s) - \mathbf{E}[X])^2 P(s)$$

What is the following?

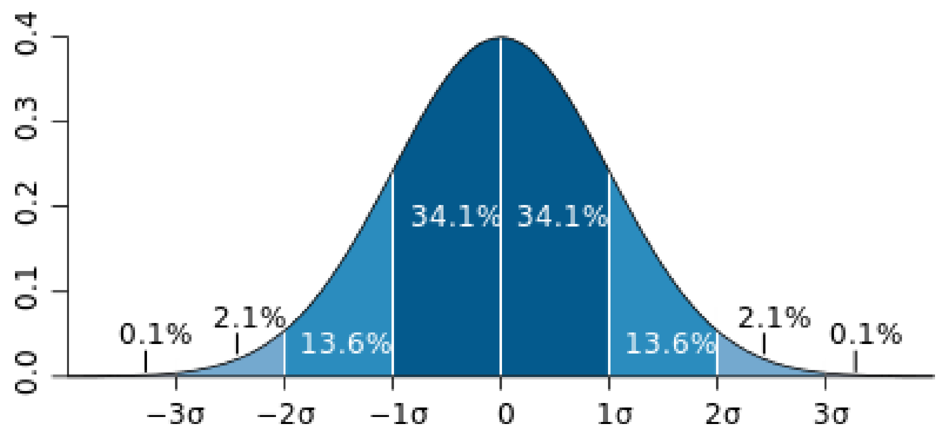
$$\sum_{s \in S} (X(s) - \mathbf{E}[X]\,\mathbf{)}^{-} P(s)$$

- The *standard deviation* of *X is:*

$$\sigma(X) := \mathbf{V}[X]^{1/2}.$$

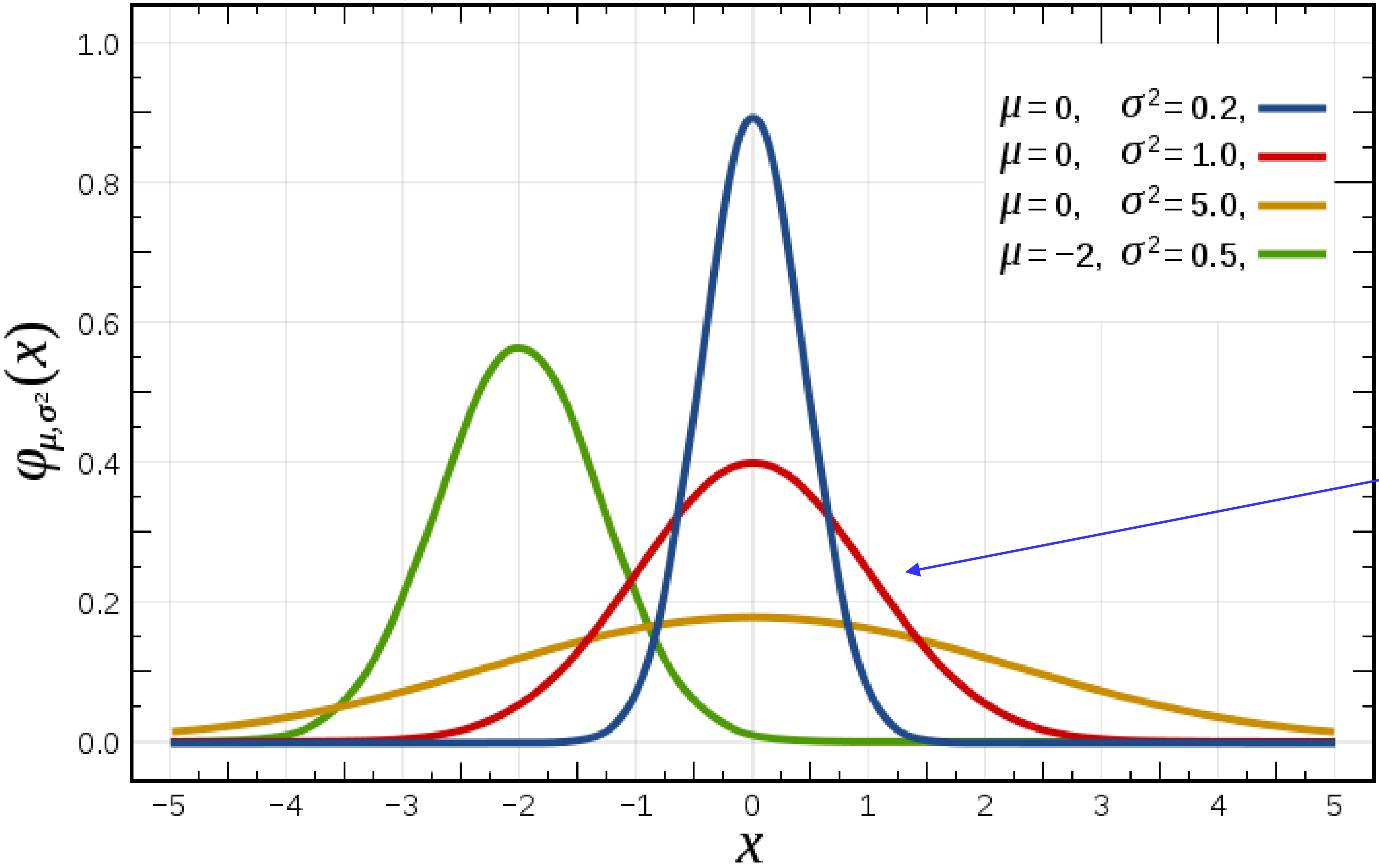
- A normal distribution, sometimes called the *bell curve*, is a distribution that occurs naturally in many situations.

Examples SAT scores, GRE scores, CS204 exam scores.



Normal distribution

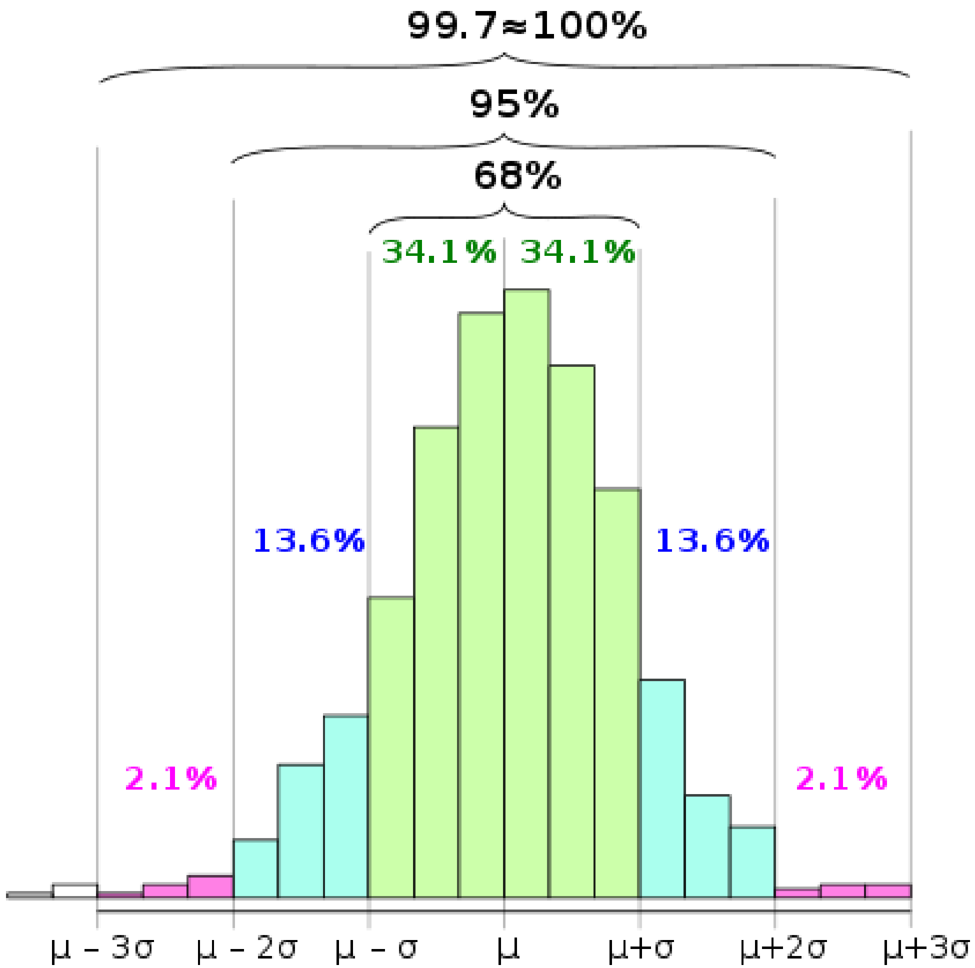
Normal distributions with various variances



μ: mean

The standard normal distribution

68-95-99.7 rule



Percentages of the values that lie within a band around the mean in a normal distribution

Six Sigma in the manufacturing world

6 σ: Defect ratio of less than 0.002 out of 100Milliion

Theorem If X is a random variable on a sample space S, then

$$V(X) = E(X^2) - E(X)^2$$

Proof

$$\begin{aligned} V(X) &= \underline{\sum_{s \in S} (X(s) - E(X))^2 p(s)} && \text{- by definition of V(X)} \\ &= \sum_{s \in S} X(s)^2 p(s) - \underline{2E(X) \sum_{s \in S} X(s) p(s)} + E(X)^2 \sum_{s \in S} p(s) && \text{- by definition of E(X)} \\ &= E(X^2) - 2E(X)E(X) + E(X)^2 \\ &= E(X^2) - E(X)^2. \end{aligned}$$

We have used the fact that $\sum_{s \in S} p(s) = 1$ in the next-to-last step.

Example

Two tetrahedral dice are rolled.
(A tetrahedral die has four faces, which are numbered 1, 2, 3, 4.)
Let X(i,j) = i + j, where the first die shows i and the second die shows j.
Find E(X) and V(X).

Solution

The sample space S consists of 16 outcomes: $S = \{(i,j) \mid i,j = 1,2,3,4\}$. We have the following probabilities:

$$p(2) = 1/16,\ \ p(3) = 2/16,\ \ p(4) = 3/16,\ \ p(5) = 4/16,\ \ p(6) = 3/16,\ \ p(7) = 2/16,\ \ p(8) = 1/16.$$

Therefore,

$$E(X) = 2\cdot\frac{1}{16} + 3\cdot\frac{2}{16} + 4\cdot\frac{3}{16} + 5\cdot\frac{4}{16} + 6\cdot\frac{3}{16} + 7\cdot\frac{2}{16} + 8\cdot\frac{1}{16} = \frac{80}{16} = 5.$$

To find $V(X)$, we use the equality $V(X) = E(X^2) - E(X)^2$:

Example

Two tetrahedral dice are rolled.
(A tetrahedral die has four faces, which are numbered 1, 2, 3, 4.)
Let X(i,j) = i + j, where the first die shows i and the second die shows j.
Find E(X) and V(X).

Solution

The sample space S consists of 16 outcomes: $S = \{(i,j) \mid i,j = 1,2,3,4\}$. We have the following probabilities:

$$p(2) = 1/16,\ \ p(3) = 2/16,\ \ p(4) = 3/16,\ \ p(5) = 4/16,\ \ p(6) = 3/16,\ \ p(7) = 2/16,\ \ p(8) = 1/16.$$

Therefore,

$$E(X) = 2\cdot\frac{1}{16} + 3\cdot\frac{2}{16} + 4\cdot\frac{3}{16} + 5\cdot\frac{4}{16} + 6\cdot\frac{3}{16} + 7\cdot\frac{2}{16} + 8\cdot\frac{1}{16} = \frac{80}{16} = 5.$$

To find $V(X)$, we use the equality $V(X) = E(X^2) - E(X)^2$:

$\sum_{s\in S} (X - E(X))^2 \times P(s)$

$$= (2-5)^2\times\frac{1}{16} + (3-5)^2\times\frac{2}{16} + (4-5)^2\times\frac{3}{16} + (5-5)^2\times\frac{4}{16} + (6-5)^2\times\frac{3}{16} + (7-5)^2\times\frac{2}{16} + (8-5)^2\times\frac{1}{16}$$

$$= \frac{1\times 9}{16} + \frac{2\times 4}{16} + \frac{3\times 1}{16} + \frac{4\times 0}{16} + \frac{3\times 1}{16} + \frac{2\times 4}{16} + \frac{1\times 9}{16} = \frac{40}{16} = 2.5$$

Example

Two tetrahedral dice are rolled.
(A tetrahedral die is a die with four faces, which are numbered 1, 2, 3, 4.)
Let X(i,j) = i + j, where the first die shows i and the second die shows j.
Find E(X) and V(X).

Solution

:

E(X) = 5

To find $V(X)$, we use the equality $V(X) = E(X^2) - E(X)^2$:

$$
\begin{aligned}
V(X) &= E(X^2) - E(X)^2 \\
&= \left(4\cdot\frac{1}{16} + 9\cdot\frac{2}{16} + 16\cdot\frac{3}{16} + 25\cdot\frac{4}{16} + 36\cdot\frac{3}{16} + 49\cdot\frac{2}{16} + 64\cdot\frac{1}{16}\right) - 5^2 \\
&= \frac{1}{16}(440) - 5^2 = 27.5 - 25 = 2.5.
\end{aligned}
$$

Quiz 22-1

What is the smallest sigma value we can use when we want a less than 0.3% defect ratio?

1) 1 σ
2) 2 σ
3) 3 σ
4) 4 σ
5) 6 σ