GB13604 - Maths for Computer Science

Lecture 9 – Probability, Part II

Claus Aranha
caranha@cs.tsukuba.ac.jp

College of Information Science

2018-12-12

Last updated December 12, 2018

Introduction

- Independence and Causality
- Random Variables
- Expectation

Independence and Causality

Independent Events

- If it is raining, the probability that people will take the bus from tsukuba station is higher than if it is sunny.

$$\Pr(\text{Take the bus} \mid \text{raining}) \neq \Pr(\text{Take the bus} \mid \text{sunny})$$

- If your father and mother have type A blood, the probability that you have type A blood is higher.

$$\Pr(A_{me} | A_P \& A_M) \geq \Pr(A_{me})$$

- If you flip two separate coins, the probability of the second coin being heads is not influenced by the result of the first coin.

$$\Pr(H_2 | H_1) = \Pr(H_2 | \bar{H}_1) = \Pr(H_2)$$

Independent Events

Two events, A and B are independent if the probability of A does not depend on the occurrence of B.

Mathematical Definitions:

- **Definition 1:** Events A and B are independent **iff**:

$$\Pr(A) = \Pr(A|B) \tag{1}$$

- **Definition 2:** Events A and B are independent **iff**:

$$\Pr(A) \cdot \Pr(B) = \Pr(A \cap B) \tag{2}$$

Proof of Equivalence of Definitions 1 and 2

Definition 1:

$$\Pr(A) = \Pr(A|B) \textbf{ iff}$$

Conditional Prob Definition:

$$\Pr(A) = \frac{\Pr(A \cap B)}{\Pr(B)} \textbf{ iff}$$

Algebra:

$$\Pr(A) \cdot \Pr(B) = \Pr(A \cap B).$$

Independence: Characteristics

Simmetry:

$$\Pr(A) \cdot \Pr(B) = \Pr(A \cap B).$$

If A is independent from B, then B is independent from A.

Zero probability event:

If $\Pr(B) = 0$ then:

- **Definition 1:** Does not work
- **Definition 2:** $\Pr(A) \cdot 0 = \Pr(A \cap \emptyset) = 0.$

So, B is independent from all events!

Independence: Complement

A independent of B means probability of A does not change if B happens or not. **Therefore** A should also be independent of **Not B.**

Lemma: A independent of B **iff** A independent of NOT B.

Proof: Proof can be derived using algebra from:

$$\Pr(A - B) = \Pr(A) - \Pr(A \cap B)$$

Mutual Independence: Idea

Experiment: Two coins are thrown

- **Event 1**: Coin 1 is heads: $\Pr(H_1) = (0,0),(0,1),(1,0),(1,1) = 0.5$
- **Event 2**: Coin 2 is heads: $\Pr(H_2) = (0,0),(0,1),(1,0),(1,1) = 0.5$
- **Event 3**: # of heads is odd:
  $\Pr(O) = (0,0),(0,1),(1,0),(1,1) = 0.5$

Mutual Independence: Idea

Experiment: Two coins are thrown

- **Event 1**: Coin 1 is heads: $\Pr(H_1) = (0,0),(0,1),(1,0),(1,1) = 0.5$
- **Event 2**: Coin 2 is heads: $\Pr(H_2) = (0,0),(0,1),(1,0),(1,1) = 0.5$
- **Event 3**: # of heads is odd:
  $\Pr(O) = (0,0),(0,1),(1,0),(1,1) = 0.5$

Events H_1, H_2, H_1, O, H_2, O are, 2-by-2, independent:

- $\Pr(O|H_1) = (1,0),(1,1) = 0.5$
- $\Pr(O|H_2) = (0,1),(1,1) = 0.5$
- $\Pr(H_1|H_2) = (0,1),(1,1) = 0.5$

These 3 events are 2-way independent.

Mutually Independent Events – Definition

Events $A_1, A_2, A_3, \ldots, A_n$ are Mutually Independent when:

The probability that A_i occurs is not changed by the occurence of the other events.

Mutually Independent Events – Example

Experiment: You throw the coin n times
Event: H_i the i-th throw is heads

What happens in the 5th flip is independent from the 1st flip, or the 7th flip, etc.

$$\Pr(H_5) = \Pr(H_5 | H_1 \cap H_3 \cap H_7 \cap \ldots)$$

Mutiually Independent Events – Math Def

Events $A_1, A_2, \ldots A_n$ are mutually independent when:

Definition 1:

$$Pr(A_i) = Pr(A_i | A_1 \cap A_2 \cap \ldots \cap A_j \cap \ldots \cap A_n); j \neq i \tag{3}$$

Definition 2:

$$Pr(A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_n) = Pr(A_1) \cdot Pr(A_2) \cdot Pr(A_3) \cdot \ldots \cdot Pr(A_n) \tag{4}$$

K-wise independence

Experiment: Two coins are thrown

- **Event 1**: Coin 1 is heads: $\Pr(H_1) = (0,0),(0,1),(1,0),(1,1) = 0.5$
- **Event 2**: Coin 2 is heads: $\Pr(H_2) = (0,0),(0,1),(1,0),(1,1) = 0.5$
- **Event 3**: # of heads is odd:
  $\Pr(O) = (0,0),(0,1),(1,0),(1,1) = 0.5$

- Each pair $(H_1, H_2), (H_1, O), (H_2, O)$ is mutually independent.
- However, (H_1, H_2, O) are not mutually independent:
  $\Pr(O|H_1 \cap H_2) = 0$

K-wise independence

Experiment: n coins are thrown

- **Event i**: Coin i is heads: $\Pr(H_i) = 0.5$
- **Event** $i+1$: # of heads is odd: $\Pr(O) = 0.5$

Any subset of i events is mutually independent.
However, the set of all $i+1$ events is **not**.

K-wise independence

A set of m events is k-wise independent if:

$A_1, A_2, \ldots, A_m$

Any subset of k events is mutually independent.

Random Variables

Playing tRPGs

In the game 3:16, you win a space battle if you roll a d10 dice, and the result is below the value FA, which you decide before the game.

- If you lose one battle you lose your armor
- If you lose another battle you are wounded
- If you lose a third battle you die

Events and Numbers

Until now, we approached probability from a Set and Events point of view:

- **Event 1:** Result of the dice: $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$
- **Event 2:** Value of **FA**: $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$
- **Event 3:** You win a combat with FA = 6: $\{1, 6\}, \{2, 6\}, \{3, 6\}, \ldots$
- **Event 4:** 3 victories in 5 battles with FA = 6: $\{E_3, E_3, E_3, \bar{E}_3, \bar{E}_3\}, \{E_3, E_3, \bar{E}_3, E_3, \bar{E}_3\}, \ldots$

Events and Numbers

However, many times we want to see probability in terms of Numbers:

- What is the probability that I can survive 7 battles if my FA is 6?
- What is the minimum FA I need to have more than .8 chance to survive 5 battles?
- How many battles in a row I expect to win if my FA is 6?

Intuitively, a Random Variable is a number that comes from a Random Process

- # of hours to the next System Crash
- # faulty pixels in the monitor
- # heads in a coin flip

Random Variable: Formal Definition

A random variable is a total function from a sample space to a numeric domain.

$$R : S \to \mathbb{R} \text{ or } R : S \to \mathbb{N}, \text{ etc...} \tag{5}$$

So a random *variable* is really a function...

(... maybe that is a bad name)

Random Variable: Formal Example

Experiment: We throw three *fair* coins independently.

- Variable C ::= Number of heads (**C**ount)
- Variable M ::= 1 if all coins **M**atch, 0 if they don't **M**atch

- $M = 1$ when $\{H, H, H\}$ or $\{T, T, T\}$,
  $\Pr(M = 1) = 2/8$
- $C = 1$ when $\{H, T, T\}, \{T, H, T\}, \{T, T, H\}$,
  $\Pr(C = 1) = 3/8$
- $\Pr(C \cdot M > 0) = \Pr(C > 0 \cap M > 0) = 1/8$

“Indicator” random variables

An **indicator random variable** takes a value of 0 or 1 for a subset of the sample space.

- Variable M ::= 1 if all coins **M**atch, 0 if they don't **M**atch
- Variable O ::= 1 if number of heads is **O**dd.

Other integer random variables partition the sample space:

- $C = 0 \rightarrow \{T, T, T\}$
- $C = 1 \rightarrow \{H, T, T\}, \{T, H, T\}, \{T, T, H\}$
- $C = 2 \rightarrow \{H, T, H\}, \{H, H, T\}, \{T, H, H\}$
- $C = 3 \rightarrow \{H, H, H\}$

Independence and Random Variables

So, a Random Variable is a function that assigns a value to a set of outcomes.

Two or more Random Variables $R_1, R_2, \ldots, R_n$ are mutually independent **iff** $(R_1 = a_1), (R_2 = a_2), \ldots, (R_n = a_n)$ are all mutually indepedent events.

Alternatively:

$$\Pr(R_1 = a_1 \cap R_2 = a_2 \cap R_3 = a_3 \cap \ldots) = \Pr(R_1 = a_1) \cdot \Pr(R_2 = a_2) \ldots$$

For all values of a_1

Independence and Random Variables

Are R.V.s C and M independent?

- Variable M ::= 1 if all coins **M**atch, 0 if they don't **M**atch
- Variable C ::= number of heads (**C**ount).

Answer: NO!

- $\Pr(M = 1) \cdot \Pr(C = 1) > 0$
- $\Pr(M = 1 \cap C = 1) = 0$

Independence and Random Variables

Are R.V.s O and M independent?

- Variable M ::= 1 if all coins **M**atch, 0 if they don't **M**atch
- Variable O ::= 1 if # heads is **O**dd, 0 if it is even.

Answer: YES!

- $\Pr(M = 1) \cdot \Pr(O = 1) = \frac{2}{8} \cdot \frac{1}{2} = \frac{1}{8}$
- $\Pr(M = 1 \cap C = 1) = \frac{1}{8}$

- Don't forget to test for the other combinations:
  $(M = 0, C = 1), (M = 1, C = 0), \ldots$

- To prove NOT independence, you just need one counterexample.

Uniform Random Variables

A Uniform Random Variable is one where the all values are equally likely.

- D_n ::= result of a fair dice with n sides.

$$\Pr(D_6 = 1) = \Pr(D_6 = 2) = \ldots = \Pr(D_6 = 6) = \frac{1}{6}$$

- S_4 ::= lottery number with 4 digits.

$$\Pr(S_4 = 0000) = \Pr(S_4 = 1234) = \ldots = \Pr(S_4 = 9998) = 0.0001$$

Binomial Random Variable

Experiment: We throw a biased coin n times, each time with p chance to turn heads.

Event: $B_{n,p}$::= # of heads in n mutually independent flips.

For $n = 5, p = 2/3$, $\Pr(HHTTH)$:

Binomial Random Variable

Experiment: We throw a biased coin n times, each time with p chance to turn heads.

Event: $B_{n,p}$::= # of heads in n mutually independent flips.

For $n = 5, p = 2/3$, Pr($HHTTH$):

- = Pr(H)·Pr(H)·Pr(T)·Pr(T)·Pr(H)

Binomial Random Variable

Experiment: We throw a biased coin n times, each time with p chance to turn heads.

Event: $B_{n,p}$::= # of heads in n mutually independent flips.

For $n = 5, p = 2/3$, $\Pr(HHTTH)$:

- = $\Pr(H) \cdot \Pr(H) \cdot \Pr(T) \cdot \Pr(T) \cdot \Pr(H)$
- $= \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3}$

Binomial Random Variable

Experiment: We throw a biased coin n times, each time with p chance to turn heads.

Event: $B_{n,p}$::= # of heads in n mutually independent flips.

For $n = 5, p = 2/3$, Pr($HHTTH$):

- = $\Pr(H) \cdot \Pr(H) \cdot \Pr(T) \cdot \Pr(T) \cdot \Pr(H)$
- $= \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3}$
- $= \frac{2}{3}^{3} \cdot \frac{1}{3}^{2}$

Binomial Random Variable

Experiment: We throw a biased coin n times, each time with p chance to turn heads.

Event: $B_{n,p}$::= # of heads in n mutually independent flips.

For $n = 5, p = 2/3$:

- Pr(ONE seq with i H and (n-i) T):
  $p^i(1-p)^{n-i}$

- Pr(ANY seq with i H and (n-i) T):
  $\binom{n}{i}p^i(1-p)^{n-i}$

Probability Density Function

The Probability Density Function (PDF) of a random variable R defines the probability that R will have a given value i

$$\mathrm{PDF}_R(i) ::= \Pr(R = i)$$

Example 1: for the binomial random variable:

$$\mathrm{PDF}_{B_{n,p}}(i) = \binom{n}{i} p^i (1-p)^{n-i}$$

Example 2: for the uniform random variable:

$$\mathrm{PDF}_U(v) = \text{constant}$$

(for v in range of U)

Cumulative Distribution

The Cumulative Distribution Function is the function that measures the probability that random variable R will have a value smaller or equal than i

$$\mathrm{CDF}_R(i) ::= \Pr(R \le i)$$

Important points

- PDF and CDF capture similar information.
- PDF and CDF do not involve the actual sample space.
- Many different **experiment** have very similar PDFs

PDF and CDF example: 2d6

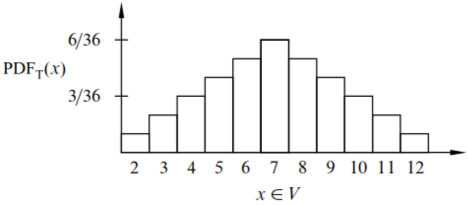


Figure 18.1 The probability density function for the sum of two 6-sided dice.

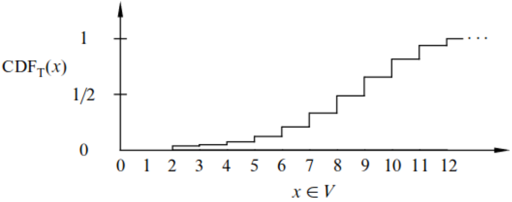


Figure 18.2 The cumulative distribution function for the sum of two 6-sided dice.

Expectation

Game: Carnival Dice

Consider the following game: Choose a number of 1 to 6 and roll 3 dice. Each dice that rolls the same number you win $1. If no dice roll that number, you lose $1.

Example: You choose number 5:

Game: Carnival Dice

Consider the following game: Choose a number of 1 to 6 and roll 3 dice. Each dice that rolls the same number you win \$1. If no dice roll that number, you lose \$1.

Example: You choose number 5:

- 2, 3, 4: Lose \$1

Game: Carnival Dice

Consider the following game: Choose a number of 1 to 6 and roll 3 dice. Each dice that rolls the same number you win $1. If no dice roll that number, you lose $1.

Example: You choose number 5:

- 2, 3, 4: Lose $1
- **5**, 4, 6: Win $1

Game: Carnival Dice

Consider the following game: Choose a number of 1 to 6 and roll 3 dice. Each dice that rolls the same number you win \$1. If no dice roll that number, you lose \$1.

Example: You choose number 5:

- 2, 3, 4: Lose \$1
- **5**, 4, 6: Win \$1
- **5**, **5**, 1: Win \$2

Game: Carnival Dice

Consider the following game: Choose a number of 1 to 6 and roll 3 dice. Each dice that rolls the same number you win $1. If no dice roll that number, you lose $1.

Example: You choose number 5:

- 2, 3, 4: Lose $1
- **5**, 4, 6: Win $1
- **5**, **5**, 1: Win $2
- **5**, **5**, **5**: Win $3

Game: Carnival Dice

Consider the following game: Choose a number of 1 to 6 and roll 3 dice. Each dice that rolls the same number you win \$1. If no dice roll that number, you lose \$1.

Example: You choose number 5:

- 2, 3, 4: Lose \$1
- **5**, 4, 6: Win \$1
- **5**, **5**, 1: Win \$2
- **5**, **5**, **5**: Win \$3

- Is this a fair game?

Game: Carnival Dice Analysis

- Pr(zero 5s) = $\frac{5}{6}^3 = \frac{125}{216}$ = Lose \$1
- Pr(one 5) = $\binom{3}{1}\frac{5}{6}^2\frac{1}{6} = \frac{75}{216}$ = Win \$1
- Pr(two 5) = $\binom{3}{2}\frac{5}{6}\frac{1}{6}^2 = \frac{15}{216}$ = Win \$2
- Pr(three 5) = $\frac{1}{6}^3 = \frac{1}{216}$ = Win \$3

Carnival Game: Law of Averages

Every 216 Games, you expect 125 games with 0 matches, 75 games with 1 match, 15 games with 2 matches, and only one game with 3 matches!

$$\frac{216 \times (-1) + 75 \times (1) + 15 \times (2) + 1 \times (3)}{216} = \frac{-17}{216} = -0.08$$

So on average, you expect to lose \$0.08 per game. Not a fair game!

Carnival Game: Impossible expectations

Note that, for the carnival game, you expect to lose 0.08 per game but **the result -0.08 does not exist!**

Another example: Rolling one 6-sided die:

Pr(1) = 1/6, Pr(2) = 1/6, Pr(3) = 1/6, . . .

$$\text{Expected value } = \frac{1 \cdot 1 + 1 \cdot 2 + 1 \cdot 3 + 1 \cdot 4 + 1 \cdot 5 + 1 \cdot 6}{6} = \frac{21}{6} = 3.5$$

You expect an average result of 3.5, but the 3.5 result does not exist!

Expectation: Definitions

The Expected Value of a variable R is the Average Value of R, weighted by the probabilities.

$$E[R] ::= \sum_{v \in \text{range}(R)} v \cdot \Pr[R = v] \tag{6}$$

So, $E[\$ \text{ at carnival game}] = -0.08$

Note 1: We can use sums in this definition because the probability space is countable.
Note 2: For continuous probability spaces we need to use integrals.

Total Expectation

Law of Total Expectation is useful for reasoning by cases:

- **Conditional Expectation:** $E[R|A] = \sum v \cdot \Pr[R = v|A]$
- **Total Expectation:** $E[R] = E[R|A] \cdot \Pr[A] + E[R|\bar{A}] \cdot \Pr[\bar{A}]$

In other words: The total expectation is The expectation when A happens plus The expectation when A does not happen.

General Total Expectation

More generally:

$$E[R] = E[R|A_1] \cdot \Pr[A_1] + E[R|A_2] \cdot \Pr[A_2] + \ldots + E[R|A_n] \cdot \Pr[A_n] \quad (7)$$

when $\Pr[A_1] + \Pr[A_2] + \ldots + \Pr[A_n] = 1$

Total Expectation: How many heads?

What is expected number of Heads when we flip a coin n times with bias p?

A coin with bias p:

- Will turn **Heads** with probability p
- Will turn **Tails** with probability $q = 1 - p$

$E[n]$ expectation of heads after n flips - study by cases:

- **Case 1:** First coin is Heads: $E[n|H_1] = 1 + E[n-1]$
- **Case 2:** First coin is Tails: $E[n|T_1] = E[n-1]$

Total Expectation: How many heads?

$E[n]$ expectation of heads after n flips - study by cases:

- **Case 1:** First coin is Heads: $E[n|H_1] = 1 + E[n-1]$
- **Case 2:** First coin is Tails: $E[n|T_1] = E[n-1]$

By total expectation:

Total Expectation: How many heads?

$E[n]$ expectation of heads after n flips - study by cases:

- **Case 1:** First coin is Heads: $E[n|H_1] = 1 + E[n-1]$
- **Case 2:** First coin is Tails: $E[n|T_1] = E[n-1]$

By total expectation:

- $E[n] = E[n|H_1] \cdot p + E[n|T_1] \cdot q$

Total Expectation: How many heads?

$E[n]$ expectation of heads after n flips - study by cases:

- **Case 1:** First coin is Heads: $E[n|H_1] = 1 + E[n-1]$
- **Case 2:** First coin is Tails: $E[n|T_1] = E[n-1]$

By total expectation:

- $E[n] = E[n|H_1] \cdot p + E[n|T_1] \cdot q$
- $E[n] = (1 + E[n-1]) \cdot p + (E[n-1]) \cdot q$

Total Expectation: How many heads?

$E[n]$ expectation of heads after n flips - study by cases:

- **Case 1:** First coin is Heads: $E[n|H_1] = 1 + E[n-1]$
- **Case 2:** First coin is Tails: $E[n|T_1] = E[n-1]$

By total expectation:

- $E[n] = E[n|H_1] \cdot p + E[n|T_1] \cdot q$
- $E[n] = (1 + E[n-1]) \cdot p + (E[n-1]) \cdot q$
- $E[n] = p + E[n-1](p+q) = p + E[n-1]$

Total Expectation: How many heads?

$E[n]$ expectation of heads after n flips - study by cases:

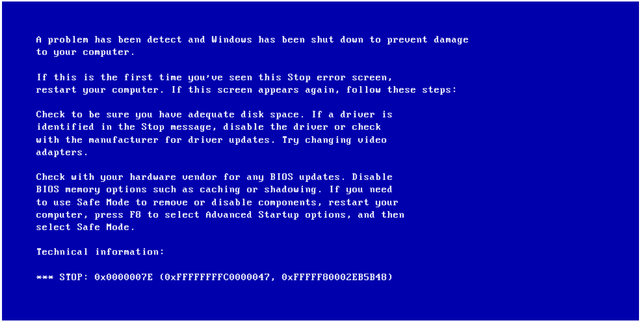
- **Case 1:** First coin is Heads: $E[n|H_1] = 1 + E[n-1]$
- **Case 2:** First coin is Tails: $E[n|T_1] = E[n-1]$

By total expectation:

- $E[n] = E[n|H_1] \cdot p + E[n|T_1] \cdot q$
- $E[n] = (1 + E[n-1]) \cdot p + (E[n-1]) \cdot q$
- $E[n] = p + E[n-1](p+q) = p + E[n-1]$
- $E[n] = p + E[n-1] = p + p + E[n-2] = 3p + E[n-3] = np + E[0] = np$

Mean Time To Failure

Suppose that a computer has a chance to fail of p every hour.



How many hours does it take for the computer to fail on average?

Mean Time To Failure

- Chance to fail in the first hour = p
- Chance to fail in the second hour = qp
- Chance to fail in the third hour = $q^2 p$
- $\dots$
- Chance to fail in the n-th hour = $q^{n-1} p$

Geometric Distribution / Geometric Sum

Mean Time To Failure

How many hours does it take for the computer to fail on average?

- $E[F] = \sum_{n>0} n \cdot \Pr[F = n]$
- $E[F] = \sum_{n>0} nq^{n-1}p = p\sum_{n>0} nq^{n-1}$
- $E[F] = p \cdot \frac{1}{(1-q)^2}$
- $E[F] = p \cdot \frac{1}{p^2} = \frac{1}{p}$

So, in average, the server will fail after $\frac{1}{p}$ hours.

Class Summary

- Random Variable is a total function that describes a numeric value for a set of probability events.
- Probability Density Function describes the probability of a Random Variable to have a certain value *a*
- Expectation is the weighted average of a PDF over its domain.

Next Class

- No Exercise Sheet for Next class!
- Sample Exam;
- Lecture Evaluation;

Please come!