Advanced Algorithms – COMS31900
2014/2015

Mathematical preliminaries
(probability)

Benjamin Sach
(slides by Markus Jalsenius)

Probability

- Sample space \( S \) is the set of outcomes of an experiment.
- For \( x \in S \), the probability of \( x \), written \( \Pr(x) \), is a real number between 0 and 1, such that \( \sum_{x \in S} \Pr(x) = 1 \).

**Example**

The outcome of a die roll: \( S = \{1, 2, 3, 4, 5, 6\} \).
\[
\Pr(1) = \Pr(2) = \Pr(3) = \Pr(4) = \Pr(5) = \Pr(6) = \frac{1}{6}.
\]

**Example**

Flip a coin: \( S = \{H, T\} \).
\[
\Pr(H) = \Pr(T) = \frac{1}{2}.
\]

**Example**

Amount of money you can win when playing some lottery:
\( S = \{£0, £10, £100, £1000, £10000, £100000\} \).
\[
\Pr(£0) = 0.9, \quad \Pr(£10) = 0.08, \ldots, \quad \Pr(£100000) = 0.0001.
\]

Event

- An event is a subset \( V \) of the sample space \( S \).
- The probability of event \( V \) happening, denoted \( \Pr(V) \), is
\[
\Pr(V) = \sum_{x \in V} \Pr(x).
\]

**Example**

Flip a coin 3 times.
- There are 8 outcomes, each has probability \( \frac{1}{2} \).
- Event \( V \) is “the first and last coin flips are the same”.
\[
\Pr(V) = \Pr(HTH) + \Pr(HHT) + \Pr(HTT) + \Pr(TTT) = 4 \times \frac{1}{8} = \frac{1}{2}.
\]

Random variable

- A random variable (r.v) \( Y \) over sample space \( S \) is a function \( S \to \mathbb{R} \)
  (a mapping from each outcome to a real number).
- The probability of \( Y \) taking value \( y \) is
\[
\Pr(Y = y) = \sum_{x \in S} \Pr(x) \quad \text{such that} \quad Y(x) = y.
\]

**Example**

Two coin flips.

<table>
<thead>
<tr>
<th>( S )</th>
<th>( Y )</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH</td>
<td>2</td>
</tr>
<tr>
<td>HT</td>
<td>1</td>
</tr>
<tr>
<td>TH</td>
<td>5</td>
</tr>
<tr>
<td>TT</td>
<td>2</td>
</tr>
</tbody>
</table>

The expected value (the mean) of a r.v. \( Y \), denoted \( \mathbb{E}(Y) \), is
\[
\mathbb{E}(Y) = \sum_{x \in S} Y(x) \cdot \Pr(x).
\]

**Example**

Flip a coin until first tail shows up:
\[
\Pr(\text{first tail}) = \sum_{n=1}^{\infty} \left( \frac{1}{2} \right)^n = 1.
\]

Linearity of expectation

**Theorem**

Let \( Y_1, Y_2, \ldots, Y_k \) be \( k \) random variables. Then
\[
\mathbb{E}\left( \sum_{i=1}^{k} Y_i \right) = \sum_{i=1}^{k} \mathbb{E}(Y_i)
\]

**Observation**

Linearity of expectation always holds, regardless of whether the random variables \( Y_i \) are dependent or not.

**Example**

What is the expected sum of two dice rolls?
- Sum over all possible outcomes:
\[
\mathbb{E}(\text{sum}) = 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + \ldots + 12 \cdot \frac{1}{6} = 7
\]
- Couple the dice so that their outcomes are always the same. Individually, they are still normal dice.
\[
\mathbb{E}(\text{sum}) = 2 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} + 8 \cdot \frac{1}{6} + 10 \cdot \frac{1}{6} + 12 \cdot \frac{1}{6} = 7
\]
**Indicator random variables**

- An indicator random variable is a r.v. that can only be 0 or 1.
- We usually use the letter \( I \) to emphasise that the r.v. is an indicator r.v.
- Property: \( E(I) = 0 \cdot \Pr(I = 0) + 1 \cdot \Pr(I = 1) = \Pr(I = 1) \).
- Often an indicator r.v. \( I \) is associated with an event such that if the event happens then \( I = 1 \), otherwise \( I = 0 \).

**EXAMPLE**

- Roll a die \( n \) times. How many die rolls do we expect to show a value that is at least the value of the previous roll?
  - For \( j \in \{2, \ldots, n\} \), indicator r.v. \( I_j = 1 \) if the value of the \( j \)th roll is at least the value of the previous roll, otherwise \( I_j = 0 \).
  - \( \Pr(I_j = 1) = \frac{6}{j} \).
  - Use linearity of expectation:
    \[
    E\left(\sum_{j=2}^{n} I_j\right) = \sum_{j=2}^{n} E(I_j) = (n-1) \cdot \frac{7}{12}.
    \]

**Mean**

- Flip a coin \( n \) times.
- Let r.v. \( T \) be the number of tails.
  - \( E(T) = \frac{n}{2} \).
  - How likely is \( T \) to be within 5% of \( \frac{n}{2} \)?

<table>
<thead>
<tr>
<th>( n )</th>
<th>Chance</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>24%</td>
</tr>
<tr>
<td>100</td>
<td>36%</td>
</tr>
<tr>
<td>200</td>
<td>52%</td>
</tr>
<tr>
<td>1000</td>
<td>89%</td>
</tr>
<tr>
<td>100,000</td>
<td>( 100 - 2.6 \cdot 10^{-14}) %</td>
</tr>
</tbody>
</table>

**Markov’s inequality**

**THEOREM**

If \( X \) is a non-negative r.v. then for all \( a > 0 \),
\[
\Pr(X \geq a) \leq \frac{E(X)}{a}.
\]

**EXAMPLE**

- Suppose average (mean) speed on the motorway is 60 mph.
  - Then at most
    - \( \frac{1}{3} \) of all cars run at 120 mph or more,
    - \( \frac{1}{6} \) of all cars run at 90 mph or more, otherwise the mean must be higher than 60 mph.

**Markov’s inequality**

**EXAMPLE**

- \( n \) people go to a party, leaving their hats at the door.
  - Each person leaves with a random hat.
  - How many people leave with their own hat?

  For \( j \in \{1, \ldots, n\} \), let indicator r.v. \( I_j = 1 \) if the \( j \)th person gets their own hat, otherwise \( I_j = 0 \). By linearity of expectation,
\[
E\left(\sum_{j=1}^{n} I_j\right) = \sum_{j=1}^{n} E(I_j) = \sum_{j=1}^{n} \Pr(I_j = 1) = n \cdot \frac{1}{n} = 1.
\]

**Union bound**

**EXAMPLE**

- \( S = \{1, \ldots, 6\} \) is the set of outcomes of a die roll.
  - Three events:
    - \( V_1 = \{3, 4, 5\} \) (between 3 and 5),
    - \( V_2 = \{1, 2, 3\} \) (3 or less),
    - \( V_3 = \{2, 3, 5\} \) (a prime).
  - \( \Pr(V_1 \cup V_2 \cup V_3) \leq \Pr(V_1) + \Pr(V_2) + \Pr(V_3) = \frac{3}{2} \).
  - Double counting \( \Pr(2) \) and \( \Pr(5) \),
  - Triple counting \( \Pr(3) \).
Union bound

**Theorem**
Let $V_1, \ldots, V_k$ be $k$ events. Then
\[
\Pr\left(\bigcup_{i=1}^{k} V_i\right) \leq \sum_{i=1}^{k} \Pr(V_i).
\]

**Observe**
The bound is tight when the events are all disjoint.

**Proof**
- Indicator r.v. $I_j = 1$ if event $V_j$ happens, otherwise $I_j = 0$.
- Let $X = \sum_{j=1}^{k} I_j$ be the number of events happening.
- $\Pr\left(\bigcup_{j=1}^{k} V_j\right) = \Pr(X > 0) \leq \mathbb{E}(X) = \sum_{j=1}^{k} \mathbb{E}(I_j) = \sum_{j=1}^{k} \Pr(V_j)$. Using previous corollary Linearity of expectation

Binomial coefficients

The number of ways to choose $k$ elements from a set of $n$ elements is
\[
\binom{n}{k} = \frac{n!}{k!(n-k)!}.
\]

**Lemma**
\[
\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{n \cdot e}{k}\right)^k
\]