Heuristic Optimization

Lecture 2

Algorithm Engineering Group
Hasso Plattner Institute, University of Potsdam

21 April 2015

Basic Definitions

- $\mathbb{N} = \{0, 1, 2, \ldots\}$.
- (Ω, P): discrete probability space;
  - Ω is a countable set;
  - $P : \Omega \to [0, 1]$ a function such that $\sum_{\omega \in \Omega} P(\omega) = 1$.
- $\omega \in \Omega$: elementary events.
- For $\omega \in \Omega$: $P(\omega)$ is the probability of ω.
- $A \subseteq \Omega$: event.
- For $A \subseteq \Omega$: $P(A) = \sum_{a \in A} P(a)$.

Example (rolling a die; generalizes to uniform distribution):

- $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- For all $\omega \in \Omega$, $P(\omega) = 1/6$.

Another example (flipping a coin until we see tails):

- $\Omega = \mathbb{N}$.
- For all $n \in \mathbb{N}$, $P(n) = 2^{-n-1}$.

Note that $\sum_{n \in \mathbb{N}} 2^{-n-1} = 1$ (geometric sum).

Some Laws

For all events $A, B \subseteq \Omega$:

1. Recall: $P(A) = \sum_{a \in A} P(a)$.
2. If $A \subseteq B$, then $P(A) \leq P(B)$.
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
4. $P(\Omega \setminus A) = 1 - P(A)$.
5. For all sequences $(A_i)_i$ of events, $P(\bigcup_i A_i) \leq \sum_i P(A_i)$.

The last item is called the union bound.

Random Variables

- Mapping $X : \Omega \to \mathbb{R}$: random variable.

Example (rolling two dice):

- $\Omega = \{1, 2, 3, 4, 5, 6\}^2$, P uniform.
- For $a, b \in \{1, 2, 3, 4, 5, 6\}$, $X(a, b) = a + b$.
- "$X = 12$" and similar things are now events ($= \{\omega \in \Omega \mid X(\omega) = 12\}$).

Exercise: What are the following probabilities?

$$P(X = 12) \quad P(X = 11) \quad P(X = 0) \quad P(X \geq 7)$$

Another example (rolling two dice):

- Same (Ω, P).
- For all $a, b \in \{1, 2, 3, 4, 5, 6\}$, $X(a, b) = a$ and $Y(a, b)$.
- "$X = Y$" and similar things are now events ($= \{\omega \in \Omega \mid X(\omega) = Y(\omega)\}$).

Exercise: What are the following probabilities?

$$P(X = 3) \quad P(X = Y) \quad P(X = Y + 2) \quad P(X \geq Y)$$

Properties of Random Variables

- X, Y identically distributed ($X \sim Y$): for all $r \in \mathbb{R}$, $P(X = r) = P(Y = r)$.
- X, Y identical ($X = Y$): implies $P(X = Y) = 1$.
- X, Y independent: for all sets $A, B \subseteq \mathbb{R}$ we have

$$P(X \in A \text{ and } Y \in B) = P(X \in A) \cdot P(Y \in B).$$

- $(X_i)_i$ independent:for all sequences $(A_i)_i$ of subsets of real numbers

$$P(\bigwedge_i X_i \in A_i) = \prod_i P(X_i \in A_i).$$

- X, Y i.i.d.: identically distributed and independent.
- Extends to sequences of random variables.
- When X, Y are r.v., $X + Y$ is r.v.: for $\omega \in \Omega$, $(X + Y)(\omega) = X(\omega) + Y(\omega)$.
- When X is r.v. and $r \in \mathbb{R}$, rX is r.v.: $(rX)(\omega) = rX(\omega)$.

Expected Value

- Expected value of X: $E(X) = \sum_{\omega \in \Omega} P(\omega) \cdot X(\omega)$.
- Variance: $\mathrm{Var}(X) = E((X - E(X))^2)$

$$\begin{aligned} E(X) &= \sum_{\omega \in \Omega} P(\omega) \cdot X(\omega) \\ &= \sum_{r \in \mathbb{R}} \sum_{\omega : X(\omega) = r} P(\omega) \cdot r \\ &= \sum_{r \in \mathbb{R}} r \cdot \sum_{\omega : X(\omega) = r} P(\omega) \\ &= \sum_{r \in \mathbb{R}} r \cdot P(X = r). \end{aligned}$$

Exercise (linearity of E):

1. $E(X + Y) = E(X) + E(Y)$;
2. $E(rX) = rE(X)$.

Some Theorems

Products of independent random variables

Let X, Y be independent random variables. We have $E(XY) = E(X)E(Y)$.

Markov's Inequality

Let X be a random variable with $P(X < 0) = 0$. For all $a > 0$ we have

$$P(X \geq a) \leq \frac{E(X)}{a}.$$

Variance

Let X, Y be a random variables. We have

1. $\mathrm{Var}(X) = E(X^2) - E(X)^2$; and
2. $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.