Discrete Math 37110 - Class 8 (2016-10-20)

Instructor: László Babai
Notes taken by Jacob Burroughs
Revised by instructor

Midterm is Tuesday, November 8 in class.

Grading:

HW	40%
Midterm	20%
Final	35%
Participation	5%

8.1 Finite Probability Spaces

Study the instructor's online "Discrete Mathematics" lecture notes (LN) for further definitions and exercises.

Example 8.1. Poker hand: draw 5 cards from a *standard deck* of 52 cards

Ω is the "sample space," which is the set of possible outcomes of the experiment. In this case, $|\Omega| = \binom{52}{5}$.

This is a uniform distribution, where $P(\text{a particular hand}) = \frac{1}{\binom{52}{5}}$

Example 8.2. Rolling n dice, represented as a sequence of numbers.

$|\Omega| = 6^n$

If the dice are fair, this is a uniform distribution, otherwise, it might not be.

Definition 8.3 (Finite Probability Space)**.** A *finite probability space* (Ω, P) consists of:

The "sample space" Ω, a non-empty finite set

A **probability distribution** P over Ω, i.e., a function $P : \Omega \to \mathbb{R}$ such that $(\forall a \in \Omega)(P(a) \geq 0)$, and $\sum_{a\in\Omega} P(a) = 1$.

Definition 8.4 (Event)**.** An *event* A is a subset of Ω.

We define $P(A) := \sum_{a\in A} P(a)$

We say an event is trivial if $P(A) = 0$ or $P(A) = 1$

An *elementary event* is an element of the sample space, $a \in \Omega$, viewed as the set $\{a\}$.

DO 8.5. Show that $0 \leq P(A) \leq 1$.

DO 8.6 (Union bound)**.** Show that $P(\bigcup_{i=1}^{k} A_i) \leq \sum_{i=1}^{k} P(A_i)$.

DO 8.7. Show that $P(A \cup B) + P(A \cap B) = P(A) + P(B)$.

Definition 8.8. We say a probability distribution is **uniform** if $(\forall a \in \Omega)(P(a) = \frac{1}{|\Omega|})$.

Then $P(A) = \frac{|A|}{|\Omega|}$ (naive probability).

Example 8.9. A **Bernoulli trial** ("flipping a biased coin") is an experiment with two outcomes: "success" ("Heads") with probability p and "failure" ("Tails") with probability $1-p$.

Consider a sequence of n independent identical Bernoulli trials. The size of the sample space is $|\Omega| = 2^n$ (binary strings of length n). An elementary event is a particular outcome (a binary string, such as $HHTHTTTH$). The probability of an elementary event is $p^k(1-p)^{n-k}$ where k is the number of successes in the sequence.

Consider the event that there are exactly k successes.
$P(\text{exactly } k \text{ sucesses}) = \binom{n}{k}p^k(1-p)^{n-k}$.

Note that by the binomial theorem, $\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k} = 1$.

Definition 8.10. A sequence $a_0, a_1, \ldots, a_n$ is *unimodal* if $(\exists \ell)(a_0 \le \cdots \le a_\ell \ge \cdots \ge a_n)$.

HW 8.11. (a) Prove that the sequence $a_k = \binom{n}{k}p^k(1-p)^{n-k}$ $(k = 0, \ldots, n)$ is unimodal.

(b) What is the most likely number of successes in a sequence of n independent identical Bernoulli trials? Prove that it is one of $\lfloor np \rfloor$ or $\lceil np \rceil$.

(5+4 points)

Definition 8.12. A sequence of $a_0, a_1, \ldots, a_n$ of positive numbers is **log-concave** if $(\forall i)(a_{i-1}a_{i+1} \le a_i^2)$.

XC 8.13. Prove: If $\{a_i\}$ is log-concave, then it is unimodal. **(4 points)**

PFF 8.14. Show that $a_k = \binom{n}{k}p^k(1-p)^{n-k}$ is log-concave.

(PFF: Problems for fun – optional, do not hand in.)

CH 8.15. Given a sequence $b_1, \cdots, b_n$ of positive numbers, let $f(t) = \prod_{i=1}^{n}(t + b_i) = a_0 + a_1 t + \cdots + a_n t^n$. Prove that the sequence $\{a_i\}$ is log-concave.

Definition 8.16 (Conditional Probability)**.** Suppose $P(A) > 0$. Given $B \subseteq \Omega$, define $P(B \mid A) = \dfrac{P(A \cap B)}{P(A)}$, which is the "probability of B given A." Here A is the *condition.*

DO 8.17. If $\Omega = \Omega_1 \dot\cup \cdots \dot\cup \Omega_k$ is a partition of the sample space then $\sum_{i=1}^{k} P(\Omega_i) = 1$.

DO 8.18 (Theorem of Complete Probability)**.** Let $\Omega = \Omega_1 \dot\cup \cdots \dot\cup \Omega_k$ be a partition of the sample space. Assume each block Ω_i has positive probability. Prove:

$$P(A) = \sum_{i=1}^{k} P(A \mid \Omega_i) \cdot P(\Omega_i).$$

8.2 Independence of events

Definition 8.19. Events A, B are *independent* if $P(A \cap B) = P(A)P(B)$.

DO 8.20. Show that the above definition holds almost if and only if $P(A \mid B) = P(A)$. (Exception: when $P(B) = 0$).

Definition 8.21. We define A, B, C to be independent if $P(A \cap B \cap C) = P(A)P(B)P(C)$ and the sets are pairwise independent.

For the following 3 exercises, use the **smallest possible sample space**. Explicitly define your probability space in each case and state the size of your sample space.

HW 8.22. Find a probability space with 3 non-trivial events A, B, C that satisfy $P(A \cap B \cap C) = P(A)P(B)P(C)$ but are not pairwise independent. **(4 points)**

XC 8.23. Find a probability space with 3 non-trivial events that satisfy the following conditions: $P(A \cap B \cap C) = P(A)P(B)P(C)$, and A, B are independent and A, C are independent, but B, C are not independent. **(4 points)**

HW 8.24. Find 3 non-trivial events which are pairwise independent but not independent. **(5 points)**

Definition 8.25. $A_1, \ldots A_k \subseteq \Omega$ are **independent** if $\forall I \subseteq [k]$ we have $P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i)$.

This gives 2^k conditions. Note that those where $|I| = 1$ and $I = \emptyset$ are self-evident. (This relies on the convention that $\bigcap_{i \in \emptyset} A_i = \Omega$.) So in effect we have $2^k - k - 1$ conditions.

XC 8.26. Construct k events that are $(k-1)$-wise independent but not fully independent. Make your sample space as small as possible. State the size of your sample space. Your events should have a very simple description. Elegance counts. **(5 points)**

DO 8.27. If $A_1, \ldots, A_k$ are independent, and A_{k+1} is a trivial event, then $A_1, \ldots, A_k, A_{k+1}$ are independent.

DO 8.28 (Probability of complement)**.** $P(\overline{A}) = 1 - P(A)$.

DO 8.29. If $A_1, \ldots, A_k$ are independent then $A_1, \ldots, A_{k-1}, \overline{A_k}$ are independent.

It follows by induction that we can negate any subset of a set of independent events and they remain independent.

DO 8.30. If A, B, C are independent then A and $B \cup C$ is also independent.

XC 8.31. Prove: If there exist k non-trivial independent events in the probability space (Ω, P) then $|\Omega| \geq 2^k$. **(5 points)**

CH 8.32. If exist k non-trivial pairwise independent events then $|\Omega| \geq k + 1$.

CH 8.33. The above bound is tight whenever $k = 2^t - 1$.

Definition 8.34. We say A, B are *positively* correlated if $P(A \cap B) > P(A)P(B)$.

They are *negatively* correlated if $P(A \cap B) < P(A)P(B)$.

8.3 Random variables, expected value

Definition 8.35 (Random variable)**.** A *random variable* is a function $X : \Omega \to \mathbb{R}$.

Definition 8.36 (Expected value)**.** The expected value of X is $E(X) = \sum_{a \in \Omega} P(a)X(a)$.

Note that this is a weighted average of the values taken by X.

DO 8.37. Prove: $\min X \leq E(X) \leq \max X$.

DO 8.38 (Linearity of expectation)**.** Let $X_1, \ldots, X_k$ be random variables on the probability space (Ω, P) and let $c_1, \ldots, c_k \in \mathbb{R}$. We consider the linear combination of the X_i with coefficients c_i. Prove:

$$E(c_1 X_1 + \cdots + c_k X_k) = \sum_{i=1}^{k} c_i E(X_i).$$

Definition 8.39 (Indicator variable)**.** An indicator variable is a random variable $\Omega \to \{0, 1\}$ (takes values $0, 1$ only). Let $A \subseteq \Omega$ be an event. The indicator of A is the indicator variable θ_A defined as $\theta_A(a) = \begin{cases} 1 \text{ if } a \in A \\ 0 \text{ if } a \in \Omega \setminus A \end{cases}$

DO 8.40. The correspondence $A \mapsto \theta_A$ is a bijection between events and indicator variables.

DO 8.41. Prove: $E(\theta_A) = P(A)$.