Discrete Math 37110 - Class 9 (2016-10-25)

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

9.1 Refresher of Last Class: Finite probability spaces, events, independence, random variables, expected value, indicator variables

Finite probability space: (Ω [a finite, nonempty set], P [a function called the probability distribution])
Regarding P, $P : \Omega \to \mathbb{R}$ such that $(\forall a \in \Omega)(P(a) \geq 0)$ and $\sum_{a\in\Omega} P(a) = 1$

An event A is a subset of Ω. $P(A) = \sum_{a\in A} P(a)$ (If P is uniform, $P(A) = \frac{|A|}{|\Omega|}$) $P(\overline{A}) = 1 - P(A)$.

A trivial event is an event where $P(A) = 0$ or $P(A) = 1$. Some examples include (but are not limited to) $\emptyset$ and Ω.

DO 9.1. These ($\emptyset$ and Ω) are the only trivial events if and only if $(\forall a \in \Omega)(P(a) > 0)$

Definition 9.2 (Correlation of events)**.** A, B are independent if $P(A \cap B) = P(A) \cdot P(B)$
A, B are positively correlated if $P(A \cap B) > P(A) \cdot P(B)$
A, B are negatively correlated if $P(A \cap B) < P(A) \cdot P(B)$

Definition 9.3 (Independence of events)**.** $A_1, \dots, A_k$ are independent if $(\forall I \subseteq [k])(P(\bigcap_{i\in I} A_i) = \Pi_{i\in I} P(A_i))$ (If $|I| \leq 1$, this condition automatically holds; the remaining $2^k - k - 1$ conditions are necessary)

DO 9.4. If $A_1, \dots, A_k$ are independent, $A_1, \dots, A_{k-1}, \overline{A_k}$ are independent.

DO 9.5. A, B, C independent implies $A, B \cup C$ independent

DO 9.6 (Boolean combinations of disjoint sets of independent events are independent)**.** If $A_1, \dots, A_k$ are independent, we have a partition $[k] = B_1 \dot\cup B_2 \dot\cup \cdots \dot\cup B_\ell$ and we set $C_i = f_i(A_j : j \in B_i)$, then $C_1, \dots, C_i$ are independent (where the f_i are boolean functions)

Definition 9.7. Random variable: $X : \Omega \to \mathbb{R}$
Expected value: $E(X) = \sum_{a\in\Omega} P(a)X(a)$

Theorem 9.8 (Alternate definition of expected value)**.**

$$E(X) = \sum_{y\in\mathbb{R}} yP(X = y)$$

DO 9.9. Prove the Theorem. Note that for finite Ω this is a finite sum.

Example 9.10. Expected number of heads in a sequence of n unbiased coin flips. (We conjecture that it is $\frac{n}{2}$.)

Let X denote the number of heads in the sequence.

$$\begin{aligned} E(X) &= \sum_{y=0}^{n} yP(X=y) \\ &= \sum_{y=0}^{n} y\binom{n}{y}\frac{1}{2^n} \\ &= \frac{1}{2^n}\sum_{y=1}^{n} y\binom{n}{y} \\ &= \frac{n}{2^n}\sum_{y=1}^{n}\binom{n-1}{y-1} \\ &= \frac{n}{2^n}2^{n-1} \\ &= \frac{n}{2} \end{aligned}$$

Definition 9.11. $T : \Omega \to \mathbb{R}$ is an **indicator variable** if $\text{range}(T) \subseteq \{0,1\}$.

Definition 9.12 (Indicator variable assoicated with an event)**.** For an event $A \subseteq \Omega$ define the indicator variable $\theta_A : \Omega \to \{0,1\}$ by setting $\theta_A(a) = \begin{cases} 1 & a \in A \\ 0 & a \notin A \end{cases}$

DO 9.13. The correspondence $A \mapsto \theta_A$ is a bijection between events and indicator variables.

DO 9.14 (Expectation of indicator variable)**.** $E(\theta_A) = P(A)$.

Hint. Notice that the event "$\theta_A = 1$" is identical with the event A. Now use Theorem 9.8.

Definition 9.15 (Linear Combination of Functions)**.** Given $f_i : \Omega \to \mathbb{R}$ and $c_i \in \mathbb{R}$, a *linear combination* is $\sum_{i=1}^{k} c_i f_i$

Definition 9.16 (Convex combination)**.** A *convex combination*, also known as a *weighted average*, is $\sum c_i f_i$ where $(c_1, \dots, c_k)$ is a probability distribution ($c_i \geq 0$, $\sum c_i = 1$).

DO 9.17. If $b_1, \dots, b_k \in \mathbb{R}$ and $(c_1, \dots, c_k)$ is a probability distribution then $\min b_i \leq \sum c_i b_i \leq \max b_i$.

DO 9.18. In particular, from above, $\min X \leq E(X) \leq \max X$.

DO 9.19 (Linearity of Expectation)**.** If $c_1 \dots c_k \in \mathbb{R}$ and $X_1, \dots, X_k$ are random variables (over the same probability space) then $E(\sum c_i X_i) = \sum c_i E(X_i)$.

Example 9.20. Alternate proof of the expected number of heads:

$X = Y_1 + \cdots + Y_n$ where Y_i indicates that the i-th flip is heads. Then $E(X) = \sum_{i=1}^{n} E(Y_i) = \sum_{i=1}^{n} P(Y_i = 1) = \sum_{i=1}^{n}(1/2) = n/2$.

9.2 Independence of random variables

Definition 9.21 (Independence of a pair of random variables)**.** The random variables X, Y are **independent** if
$(\forall x, y \in \mathbb{R})(P(X = x \wedge Y = y) = P(X = x)P(Y = y))$

Definition 9.22 (Independence of random variables)**.** The random variables $X_1, \dots, X_k$ are **independent** if
$(\forall x_1, \dots, x_k \in \mathbb{R})(P(X_1 = x_1 \wedge \dots \wedge X_k = x_k) = \prod P(X_i = x_i)$

DO 9.23. Show that the events $A_1, \cdots A_k$ are independent if and only if their indicator variables $\theta_{A_1}, \dots, \theta_{A_k}$ are independent.

DO 9.24 (Multiplicativity of expectation 1)**.** If X, Y are independent random variables then $E(XY) = E(X)E(Y)$.

DO 9.25 (Multiplicativity of expectation 2)**.** If X, Y are independent random variables then $E(\prod X_i) = \prod E(X_i)$.

9.3 Variance, covariance

Definition 9.26. The *variance* of X is $\mathrm{Var}(X) = E((X - E(X))^2)$

DO 9.27. Prove: $\mathrm{Var}(X) \geq 0$. Moreover, $\mathrm{Var}(X) = 0$ if and only if X is almost constant, i.e., $(\exists c)(P(X = c) = 1)$.

Theorem 9.28. $\mathrm{Var}(X) = E(X^2) - (E(X))^2$

Proof.

$$\begin{aligned} E((X - m)^2) &= E(X^2 - 2mX + m^2) \\ &= E(X^2) - 2mE(X) + m^2 \\ &= E(X^2) - m^2 \end{aligned}$$

□

Corollary 9.29 (Cauchy-Schwarz)**.** $E(X^2) \geq (E(X))^2$

Note the spelling of "Schwarz."

DO 9.30. Relate the above corollary to more familiar forms of Cauchy-Schwarz.

DO 9.31 (Variance of indicator variable)**.** If $P(A) = p$ then $\mathrm{Var}(\theta_A) = p(1 - p)$.

Definition 9.32. The **covariance** of the random variables X, Y is
$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$

DO 9.33. $\mathrm{Var}(X) = \mathrm{Cov}(X, X)$.

DO 9.34 (Variance of sum)**.** If $X = Y_1 + \dots + Y_k$, then

$$\mathrm{Var}(X) = \sum_{i=1}^{k} \sum_{j=1}^{k} \mathrm{Cov}(Y_i, Y_j).$$

9.4 Random graphs

Definition 9.35 (Erdős–Rényi model of random graphs)**.** Erdős and Rényi defined the probability distribution $\mathbb{G}_{n,p}$ over the set of $2^{\binom{n}{2}}$ graphs on a given set of n vertices is by deciding adjacency of each pair of vertices by independent Bernoulli trials with probability p of success. (We flip a biased coin for each pair of vertices.) (This requires $\binom{n}{2}$ Bernoulli trials.)

Notation 9.36. The notation $G \sim \mathbb{G}_{n,p}$ means that G is a graph on a given set of n vertices, selected at random according to the distribution $\mathbb{G}_{n,p}$.

Let m_G denote the number of edges of the graph G.

DO 9.37. For $G \sim \mathbb{G}_{n,p}$ we have $E(m_G) = p\binom{m}{2}$ and $E(\text{number of triangles in } G) = p^3\binom{n}{3}$.

HW 9.38 (Due Tuesday, 2016-11-01)**.** Let $G \sim \mathbb{G}_{n,1/2}$ and let X denote the number of triangles in G.

(a) Find $\text{Var}(X)$ as a closed-form expression.

(b) Find an asymptotic formula for $\text{Var}(X)$ in the form $\text{Var}(X) \sim cn^d$. Determine the constants c and d. **(6+4 points)**