MATH 340: PROBABILITY REVIEW.

Basic notions. A *sample space* S is a finite set, considered as a set of possible outcomes of an experiment.

An *event* is a subset of a sample space.

A function $p : S \to \mathbb{R}$ is a *probability distribution on* S if

- $p(x) \geq 0$ for every $x \in S$, and
- $\sum_{x \in S} p(x) = 1$.

The *probability* of an event A is defined as

$$p(A) = \sum_{x \in A} p(x).$$

The probability distribution is *uniform* if $p(x) = \frac{1}{|S|}$ for every $x \in S$, i.e. all outcomes of the experiment are equally likely.

Lemma 1. *For any pair of events* A, B *we have*

$$p(A \cup B) = p(A) + p(B) - p(A \cap B).$$

Lemma 2. *For any collection of events* $A_1, A_2, \ldots, A_n$ *we have*

$$p(A_1 \cup A_2 \cup \ldots \cup A_n) \leq p(A_1) + p(A_2) + \ldots + p(A_n).$$

Conditional probability and Bayes theorem. Let B be an event with $p(B) > 0$. The *conditional probability* of A given B is

$$p(A|B) = \frac{p(A \cap B)}{p(B)}.$$

Theorem 3 (Bayes)**.**

$$p(B|A) = \frac{p(A|B)p(B)}{p(A)}.$$

Let $B_1, B_2, \ldots, B_n$ be pairwise disjoint events such that $B_1 \cup B_2 \ldots \cup B_n = S$ then

$$p(A) = p(A|B_1)p(B_1) + p(A|B_2)p(B_2) + \ldots + p(A|B_n)p(B_n).$$

Independence. Events A and B are *independent* if

$$p(A \cap B) = p(A)p(B).$$

More generally, events $A_1, \ldots, A_k$ are *independent* if

$$p(\cap_{i \in I} A_i) = \prod_{i \in I} p(A_i)$$

for any $I \subseteq \{1, 2, \ldots, k\}$.

Random variables. A *random variable* X assigns a value $X(s)$ to each outcome $s \in S$.

The *probability density function* gives the probability that X takes a particular value:

$$p(X = v) = \sum_{s \in S, X(s)=v} p(s).$$

Expectation. The *expectation* of a random variable X is

$$\mathrm{E}[X] = \sum_{s \in S} X(s)p(s) = \sum_{v} p(X = v)v.$$

Theorem 4 (Linearity of expectation). *Given two random variables X and Y and two constans a and b we have*

$$\mathrm{E}[aX + bY] = a\,\mathrm{E}[X] + b\,\mathrm{E}[Y].$$

Two random variables are *independent* if

$$p(X = x \text{ and} Y = y) = p(X = x)p(Y = y)$$

for all values x and y.

Theorem 5. *If X and Y are independent random variables then*

$$\mathrm{E}[XY] = \mathrm{E}[X]\,\mathrm{E}[Y].$$

Theorem 6 (Markov). *If $a > 0$ and X is a random variable taking only non-negative values then*

$$p(X \geq a) \leq \frac{\mathrm{E}[X]}{a}$$

Bernoulli random variables. A *Bernoulli random variable* X or has two values 1 (success) with probability p and 0 failure with probability $1 - p$.

The birthday problemm, balls, bins. If we distribute k balls among n bins uniformly and independently at random, that no bin contains more than one ball is

$$\frac{n-1}{n} \cdot \frac{n-2}{n} \cdot \ldots \cdot \frac{n-k+1}{n} \leq e^{\frac{(k-1)k}{2n}}.$$

If we distribute k balls among n bins uniformly and independently at random and X_i is the random variable equal to the number of balls in the bin i then

$$p(X_i = r) = \binom{k}{r} p^r (1-p)^{k-r},$$

where $p = 1/n$ is the probability that we put any given ball into bin number i.

Random walks. A *random walk* in a graph is a walk in which at each vertex we arbitrarily choose an incident edge and move along it.

Theorem 7. *If G is a connected, non-bipartite graph then in the long run the probability that a given time the random walk is at a vertex v approaches*

$$x_v^* = \frac{\deg(v)}{2|E(G)|}.$$

Chernoff bounds.

Theorem 8 (Chernoff)**.** *Let* $X_1, X_2, \ldots, X_k$ *be independent Bernoulli random variables such that* $p(X_i = 1) = p_i$*. Let* $X = X_1 + X_2 + \ldots + X_k$ *and let* $\mu = \mathrm{E}[X] (= \sum_{i=1}^{k} p_i)$*. Then for* $\delta \geq 0$

$$p(X > (1+\delta)\mu) \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}} \right)^{\mu}.$$

Moreover, if $0 < \delta < 1$*, then*

$$p(X > (1+\delta)\mu) \leq e^{-\frac{1}{3}\mu\delta^2}.$$

Let X be the random variable counting the number of heads after n coin tosses using a fair coin. Then

$$p(X > \frac{n}{2} + 2\sqrt{n \ln n}) \leq \frac{1}{n^{8/3}}.$$

Ramsey theorem. The *Ramsey number* $R(k,l)$ denotes the minimum integer N such that in any coloring of edges of a complete graph K_N in two colors red and blue one can find K_k with all edges colored red or K_l with all edges colored blue.

Theorem 9 (Ramsey)**.** *The Ramsey number* $R(k,l)$ *exists for all* k, l*. For* $k, l \geq 2$

$$R(k,l) \leq R(k-1,l) + R(k,l-1).$$

Theorem 10 (Erdős)**.** *For* $k \geq 2$

$$R(k,k) \geq 2^{k/2}.$$