MIDTERM TWO

MODULE 7	
Probability Space	A **probability space** $(\Omega, \Pr)$ consists of • a finite non-empty set Ω of **outcomes** and • a **probability distribution** function $\Pr : \Omega \to [0,1]$ that associates with each outcome $w \in \Omega$ its **probability** $\Pr[w]$ which is a real number between 0 and 1 (inclusive), such that $$\sum_{w \in \Omega} \Pr[w] = 1$$
Event	Let $(\Omega, \Pr)$ be a probability space. An **event** in this space is a subset $E \subseteq \Omega$. We extend the probability function from outcomes to events as follows $$\Pr[E] = \sum_{w \in E} \Pr[w]$$ Note that • $\Pr[E] \in [0,1]$ • $\Pr[\emptyset] = 0$ • $\Pr[\{w\}] = \Pr[w]$
Uniform	A probability space $(\Omega, \Pr)$ is called **uniform** if all the outcomes have the **same** probability. Denote $n = \|\Omega\|$. Since the probabilities are equal and sum up to 1: . $\Pr[w] = 1/n$ for each outcome $w \in \Omega$. **Proposition.** In a uniform probability space $\Pr[E] = m/n$ where $m = \|E\|$ and $n = \|\Omega\|$. **Proof.** $$\Pr[E] = \sum_{w \in E} \Pr[w] = \sum_{w \in E} \frac{1}{n} = m \cdot \frac{1}{n} = \frac{m}{n}$$

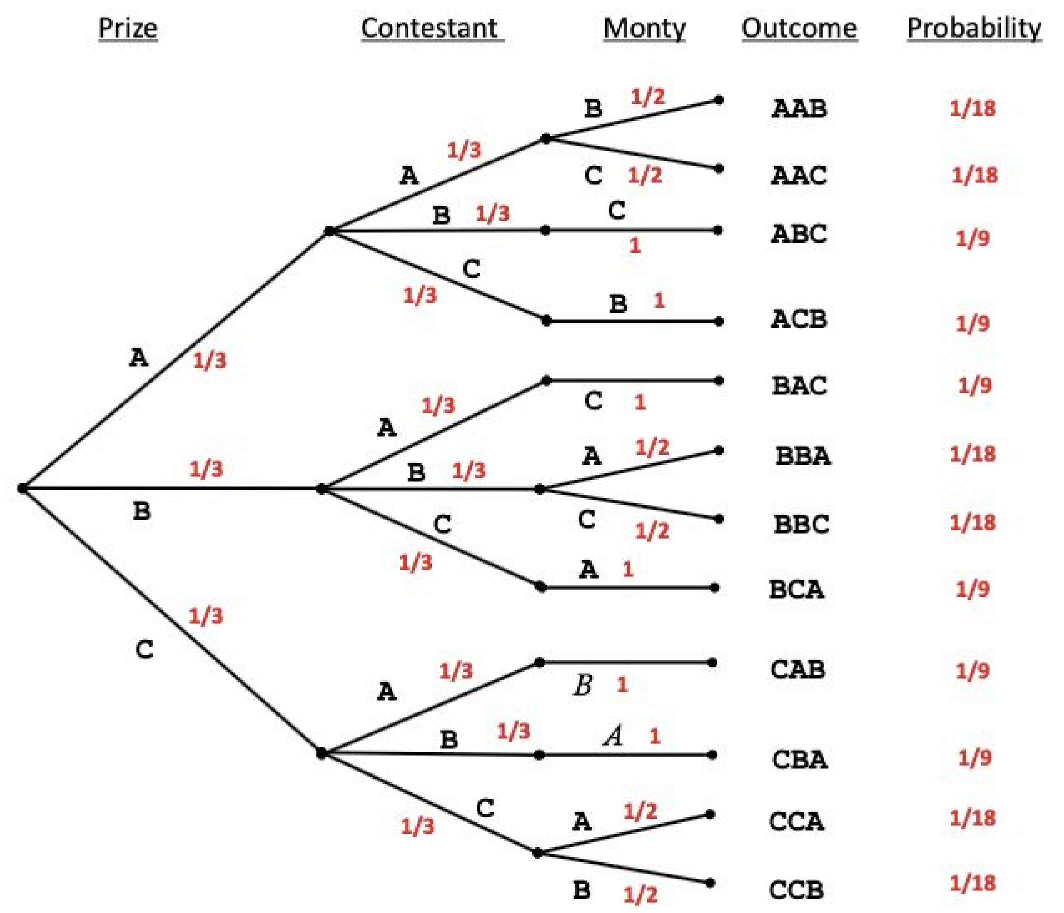
Random Permutation	Distinct objects $a_1, \ldots, a_n$. A **random permutation** of $a_1, \ldots, a_n$ is an element of the **uniform** probability space whose outcomes are all the permutations . Each outcome has probability $1/n!$.
Properties	Consider an arbitrary probability space (Ω, Pr) and arbitrary events E, A, B in this space. **Property P0.** $\mathrm{Pr}[E] \geq 0$ Since it's the sum of non-negative numbers. **Property P1.** $\mathrm{Pr}[\Omega] = 1$ Since it adds up the probabilities of all the outcomes in the space. **Property P2.** If A, B are disjoint then $\mathrm{Pr}[A \cup B] = \mathrm{Pr}[A] + \mathrm{Pr}[B]$ This is called the **addition rule** and is analogous to the addition rule in counting applied to set cardinality: $A \cap B = \emptyset \Rightarrow \|A \cup B\| = \|A\| + \|B\|$. **Property P3.** If $A \subseteq B$ then $\mathrm{Pr}[A] \leq \mathrm{Pr}[B]$ This is called **monotonicity** and it has an analogous property of set cardinality: $A \subseteq B \Rightarrow \|A\| \leq \|B\|$. If $E \subseteq \Omega$ is an event then the **complement** of E is the event $\overline{E} = \Omega \setminus E$. **Property P4.** $\mathrm{Pr}[\overline{E}] = 1 - \mathrm{Pr}[E]$ **Property P5.** $\mathrm{Pr}[\emptyset] = 0$ By the definition of event probability, this is a sum with **no terms**! A common convention is that such a sum is 0. However, see the next activity. **Property P6.** $\mathrm{Pr}[A \cup B] = \mathrm{Pr}[A] + \mathrm{Pr}[B] - \mathrm{Pr}[A \cap B]$ This is called (of course!) **inclusion-exclusion** for two events and is analogous to PIE for two sets. **Property P7.** $\mathrm{Pr}[A \cup B] \leq \mathrm{Pr}[A] + \mathrm{Pr}[B]$ This is called the **union bound**, and it plays a major role in the analysis of **probabilistic algorithms**. It's quite clear that it follows immediately from **P6**.

MODULE 8

Three Events	**Proposition.** For any events A, B, C in the same probability space $$\Pr[A \cup B \cup C] = \Pr[A] + \Pr[B] + \Pr[C] - \Pr[A \cap B] - \Pr[B \cap C] - \Pr[C \cap A] + \Pr[A \cap B \cap C]$$ The proof is in the segment entitled "Inclusion-exclusion for three events". **Answer (continued).** We will apply this to $\Pr[D_1 \cup D_2 \cup D_3]$.
Independent	Let $(\Omega, \Pr)$ be a probability space. Two events $A, B \subseteq \Omega$ are **independent**, write $A \perp B$, when $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$. Note that $A \perp B$ iff $B \perp A$ (independence is **symmetric**).
Pairwise Independent	Events $A_1, \ldots, A_n$ are called **pairwise independent** when for any $1 \leq i < j \leq n$ we have $A_i \perp A_j$.
Mutually Independent	Events $A_1, \ldots, A_n$ are called **mutually independent** when for any $\{i_1, \ldots, i_k\} \subseteq [1..n]$ we have $$\Pr[A_{i_1} \cap \cdots \cap A_{i_k}] = \Pr[A_{i_1}] \cdots \Pr[A_{i_k}]$$ Mutual independence implies pairwise independence but the converse is not true, as we saw in the first proposition on the previous slide.

Properties	Consider an arbitrary probability space (Ω, Pr) and arbitrary events E, A, B in this space. **Property Ind (i).** If $\mathrm{Pr}[A] = 0$ then $A \perp B$ for any B. In particular, $\emptyset \perp E$ for any E. **Proof.** $A \cap B \subseteq A$ so by **P3** (monotonicity) $\mathrm{Pr}[A \cap B] \leq \mathrm{Pr}[A] = 0$. If $\mathrm{Pr}[A] = 0$ then $\mathrm{Pr}[A \cap B] = 0$. $A \perp B$ follows. **Property Ind (ii).** $\Omega \perp E$ for any E. The proof is in the segment entitled "Proofs of independence properties". Don't confuse "independent" with "disjoint"! In fact, disjoint events are typically **not** independent of each other! **Proposition.** Let A, B be disjoint events in (Ω, Pr). If $A \perp B$ then at least one of A, B has probability 0. **Proof.** If A, B are both disjoint and independent then by **Ind (iii)** and by **P2** (addition rule) we have: . $1 - (1 - \mathrm{Pr}[A])(1 - \mathrm{Pr}[B]) = \mathrm{Pr}[A \cup B] = \mathrm{Pr}[A] + \mathrm{Pr}[B]$. $\mathrm{Pr}[A] + \mathrm{Pr}[B] - \mathrm{Pr}[A] \cdot \mathrm{Pr}[B] = \mathrm{Pr}[A] + \mathrm{Pr}[B]$. $\mathrm{Pr}[A] \cdot \mathrm{Pr}[B] = 0$. **Corollary.** If $E \perp \overline{E}$ then $\mathrm{Pr}[E]$ is 0 or 1.
Unions of Mutually Independent Events	**Proposition. (generalizes Ind (iii))** Let $A_1, \ldots, A_n$ be mutually independent events in the same probability space. Then we have $\mathrm{Pr}[A_1 \cup \cdots \cup A_n] = 1 - (1 - \mathrm{Pr}[A_1]) \cdots (1 - \mathrm{Pr}[A_n])$ **Proof.** We could use induction. We present a more interesting method, for the case $n = 3$. We use **P4** and a De Morgan law for sets: $\overline{A \cup B \cup C} = \overline{A} \cap \overline{B} \cap \overline{C}$ together with a generalization of **Ind (iv)**: **Lemma.** A, B, C are mutually independent iff $\overline{A}, \overline{B}, \overline{C}$ are mutually independent.

Monty Hall



Conditional Probability

$$\Pr[E|U] \;=\; \frac{\sum_{w\in E\wedge w\in U}\Pr[w]}{\sum_{w\in U}\Pr[w]} \;=\; \frac{\Pr[E\cap U]}{\Pr[U]} \qquad (\text{provided} \quad \Pr[U]\neq 0)$$

When $\Pr[U] = 0$ the conditional probability $\Pr[E|U]$ is **undefined**.

Independent and Conditional

Proposition. For any two events A, B in the same probability space the following two statements are **equivalent**:

(i) $A \perp B$ (ii) $\Pr[B] = 0$ or ($\Pr[B] \neq 0$ and $\Pr[A|B] = \Pr[A]$)

Chain Rule

Proposition (The chain rule). For any events A, B, C in the same probability space we have

$$\Pr[A\cap B\cap C] \;=\; \Pr[A]\cdot\Pr[B|A]\cdot\Pr[C|A\cap B]$$

For any events $A_1, \ldots, A_n$ in the same probability space we have

$$\Pr[A_1\cap A_2\cap A_3\cdots\cap A_n] \;=\; \Pr[A_1]\cdot\Pr[A_2|A_1]\cdot\Pr[A_3|A_1\cap A_2]\cdots\Pr[A_n|A_1\cap\cdots\cap A_{n-1}]$$

MODULE 9

Random Variable	A **random variable** on $(\Omega, \Pr)$ is a function $X : \Omega \to \mathbb{R}$. Denote $\mathsf{Val}(X) = \{x \in \mathbb{R} \mid \exists w \in \Omega \;\; X(w) = x\}$. (The set of values **taken** by X.) Like Ω, $\mathsf{Val}(X)$ is also a finite set. Denote with x the real values that X takes and with $X = x$ the **event** $\{w \in \Omega \mid X(w) = x\}$. Its probability $\Pr[X = x]$ is of particular interest. The **distribution** of the random variable X is the function $f : \mathsf{Val}(X) \to [0, 1] \quad$ where $\quad f(x) = \Pr[X = x]$.
Uniform Random Variable	Let $v_1, \ldots, v_n$ be n distinct values in $\mathbb{R}$. Given $(\Omega, \Pr)$, an r.v. $U : \Omega \to \mathbb{R}$ is **uniform** with these values when $\mathsf{Val}(U) = \{v_1, \ldots, v_n\}$ and $\Pr[U = v_i] = 1/n$ for $i = 1, \ldots, n$. The corresponding distribution . $f : \{v_1, \ldots, v_n\} \to [0, 1] \qquad f(v_i) = 1/n \quad$ for $\quad i = 1, \ldots, n$ is also called **uniform**. The r.v. D associated with a fair die is uniform with $n = 6$ and $v_i = i$ for $i = 1, \ldots, 6$.

Bernoulli Random Variable	Recall **Bernoulli trials**. Similarly, we can define: Given $(\Omega, \Pr)$, an r.v. $X : \Omega \to \mathbb{R}$ with $\mathsf{Val}(X) = \{0, 1\}$ and $\Pr[X = 1] = p$ is called a **Bernoulli random variable** with parameter p. A Bernoulli r.v. X defines implicitly a Bernoulli trial where "success" is $X = 1$ and "failure" is $X = 0$. The corresponding distribution $f : \{0,1\} \to [0,1]$ where $f(1) = p$ and $f(0) = 1 - p$ is called a **Bernoulli distribution** with parameter p.
Mean	**Average**, or **mean**, value returned by a random variable. Such an average should take into account that some values may "weigh" more than others. The **weights** are given by the probability distribution! **Notation** $\mathsf{E}[X]$. Two candidates, but they give the same answer: **Proposition.** For a random variable X defined on $(\Omega, \Pr)$ we have $$\mathsf{E}[X] = \sum_{x \in \mathsf{Val}(X)} x \cdot \Pr[X = x] = \sum_{w \in \Omega} X(w) \cdot \Pr[w]$$ The first expression corresponds directly to a **weighted average** of the values taken by random variable. Recall that the weights $\Pr[X = x]$ sum up to 1. The second expression takes the average of values by outcomes they map from. Multiple outcomes may be mapped to the same value taken by the r.v.
Expectation of Bernoulli	$\mathsf{E}[X] = 1 \cdot \Pr[X = 1] + 0 \cdot \Pr[X = 0] = 1 \cdot p + 0 \cdot (1 - p) = p.$
Linearity of Expectation	$\mathsf{E}[c_1 X_1 + \cdots + c_n X_n] = c_1 \mathsf{E}[X_1] + \cdots + c_n \mathsf{E}[X_n]$

Indicator Random Variable	Let A be an event in a probability space $(\Omega, \Pr)$. The **indicator** random variable of the event A, notation I_A, is defined by $$I_A(w) = \begin{cases} 1 & \text{if } w \in A \\ 0 & \text{if } w \notin A \end{cases}$$ Note that I_A is a Bernoulli random variable with success probability $\Pr[I_A = 1] = \Pr[A]$. As we have shown before, its expectation is $E[I_A] = \Pr[I_A = 1] = \Pr[A]$.

MODULE 10	
Expectation	If X is an r.v. its **expectation**, $\mathrm{E}[X]$, also called its **mean**, is often denoted by $\mu = \mathrm{E}[X]$.
Variance	$\mathrm{Var}[X] \;=\; \mathrm{E}[(X-\mu)^2] \quad (\text{where } \mu = \mathrm{E}[X])$ $\mathrm{Var}[X] \;=\; \mathrm{E}[X^2] - \mu^2$
Standard Deviation	$\sigma[X] \;=\; \sqrt{\mathrm{Var}[X]}$
Variance of Bernoulli Random Variable	$\mathrm{Var}[X] \;=\; \mathrm{E}[X^2] - \mu^2 \;=\; p - p^2 \;=\; p(1-p)$.
Linearity of Variance	$\mathrm{Var}[c\,X] \;=\; c^2\,\mathrm{Var}[X]$ if $X \perp Y$ then $\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$ **Proposition.** If r.v.'s $X_1, \ldots, X_n$ are **pairwise independent** then $\mathrm{Var}[X_1 + \cdots + X_n] \;=\; \mathrm{Var}[X_1] + \cdots + \mathrm{Var}[X_n]$ $\mathrm{Var}[B] \;=\; p(1-p) + \cdots + p(1-p) \;=\; np(1-p)$.
Binomial Random Variable	An r.v. $B : \Omega \to \mathbb{R}$ is called **binomial** with parameters $n \in \mathbb{N}$ and $p \in [0,1]$ when $\mathrm{Val}(B) = [0..n]$ and $\forall k \in [0..n]$ $\Pr[B = k] = \binom{n}{k} p^k (1-p)^{n-k}$.
Expectation for Binomial + Bernoulli	$\mathrm{E}[B] \;=\; \mathrm{E}[J_1] \;+\; \cdots \;+\; \mathrm{E}[J_n]$. $\mathrm{E}[B] \;=\; p \;+\; \cdots \;+\; p \;=\; np$.