Part IV: Probability Theory

L13: Introduction to probability

L14: Conditional probability and Bayes' Theorem

L15: Expectation and variance

L13: Introduction to Probability

- Reading: Rosen 7.1, 7.2

The Hatcheck Problem Revisited

- Total number of permutations: $n!$
- Number of derangements:

$$D_n = n!\left[1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots (-1)^n \frac{1}{n!}\right]$$

- Probability of that a random permutation is a derangement:

$$p = \frac{D_n}{n!} = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots (-1)^n \frac{1}{n!}$$

Definitions

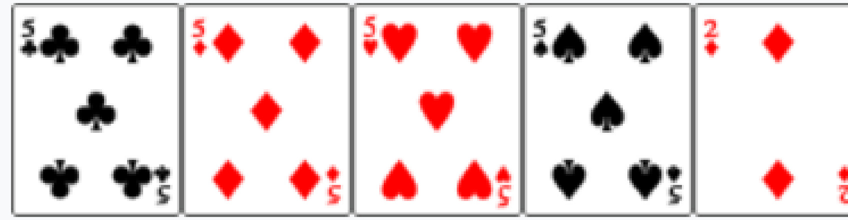
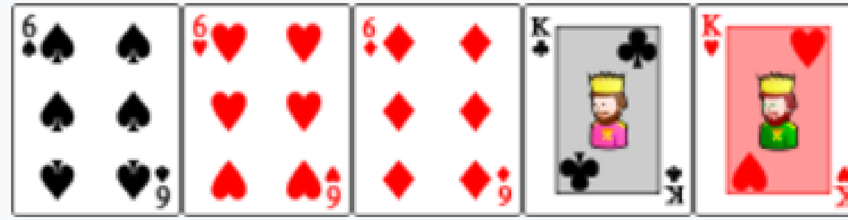
- An experiment is a procedure that yields one of a given set of possible outcomes.
- The sample space of the experiment is the set of all possible outcomes.
- An event is a subset of the sample space.
- If S is a finite sample space of equally likely outcomes, and E is an event, i.e., a subset of S, then the probability of E is

$$p(E) = \frac{|E|}{|S|}$$

Probability = Counting

- Example 1
  A bag contains 4 blue balls and 5 red balls. What is the probability that a ball chosen from the bag is blue?
- Solution
  4/9
- Example 2
  What is the probability that when two dice are rolled, the sum of the numbers on the two dice is 7?
- Solution
  By the product rule there are $6^2 = 36$ possible outcomes. Six of these sum to 7. Hence, the probability of obtaining a 7 is $6/36 = 1/6$.

Poker

Four of a kind	
Full house	

- Probability of four of a kind
  Total number of hands of five cards: $C(52,5)$
  Number of ways to get four of a kind: 13×48
  Probability $= \dfrac{13 \times 48}{C(52,5)} \approx 0.00024$
- Probability of full house
  Total number of hands of five cards: $C(52,5)$
  Number of ways to get full house: $P(13,2)C(4,3)C(4,2)$
  Probability $= \dfrac{P(13,2)C(4,3)C(4,2)}{C(52,5)} \approx 0.0014$

Mark Six (六合彩)

- Player chooses 6 numbers from 49
- 6 numbers are drawn + 1 extra

Prize	Criteria	Probability
1st Division	All 6 drawn numbers	$\frac{1}{\binom{49}{6}} = \frac{1}{13,983,816}$
2nd Division	5 out of 6 drawn numbers, plus the extra number	$\frac{\binom{6}{5}}{\binom{49}{6}} = \frac{1}{2,330,636}$
3rd Division	5 out of 6 drawn numbers	$\frac{\binom{6}{5}\binom{42}{1}}{\binom{49}{6}} \approx \frac{1}{55,491.33}$

Sampling

- What is the probability that the numbers 11, 4, 17, 39, and 23 are drawn in that order from a bin containing 50 balls labeled with the numbers 1, 2, …, 50 if
  (a) the ball selected is not returned to the bin before the next ball is selected (sampling without replacement), and
  (b) the ball selected is returned to the bin before the next ball is selected (sampling with replacement)?
- Solution
  Number of ways the event happens: 1
  Total number of ways to draw numbers:
  (a) $P(50, 5)$, probability is $1/P(50, 5)$
  (b) 50^5, probability is $1/50^5$

Complement of Event

- Theorem
  Let E be an event in a finite sample space S. The probability of the event $\bar{E}$, the complementary event of E, is given by $p(\bar{E}) = 1 - p(E)$.
- Proof
  Using the fact that $|\bar{E}| = |S| - |E|$,

$$p(\overline{E}) = \frac{|S| - |E|}{|S|} = 1 - \frac{|E|}{|S|} = 1 - p(E).$$

Complement of Event

- **Example**:
  A sequence of 10 bits is chosen randomly. What is the probability that at least one of these bits is 0?
- **Solution**:
  E: the event that at least one of the 10 bits is 0.
  $\bar{E}$: the event that all of the bits are 1s.
  The size of the sample space S is 2^{10}. Hence,

$$p(E) = 1 - p(\overline{E}) = 1 - \frac{|\overline{E}|}{|S|} = 1 - \frac{1}{2^{10}} = 1 - \frac{1}{1024} = \frac{1023}{1024}.$$

Union of Events

- **Theorem**
  Let E_1 and E_2 be events in the sample space S. Then

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$$

- **Proof**:
  Given the inclusion-exclusion formula $|A \cup B| = |A| + |B| - |A \cap B|$, it follows that

$$p(E_1 \cup E_2) = \frac{|E_1 \cup E_2|}{|S|} = \frac{|E_1| + |E_2| - |E_1 \cap E_2|}{|S|}$$

$$= \frac{|E_1|}{|S|} + \frac{|E_2|}{|S|} - \frac{|E_1 \cap E_2|}{|S|}$$

$$= p(E_1) + p(E_2) - p(E_1 \cap E_2).$$

Union of Events

- **Example**:
  What is the probability that a positive integer selected at random from the set of positive integers not exceeding 100 is divisible by either 2 or 5?
- **Solution**:
  $E_{1:}$ the event that the integer is divisible by 2;
  $E_{2:}$ the event that it is divisible 5;
  $E_1 \cup E_{2:}$ The event that the integer is divisible by 2 or 5;
  $E_1 \cap E_{2:}$ The event that it is divisible by 2 and 5.

  It follows that:

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$$
$$= 50/100 + 20/100 - 10/100 = 3/5.$$

Inclusion-Exclusion Principle for Probability

- **Theorem**

$$p\Big(\bigcup_{i=1}^{n} E_i\Big) = \sum_{1\le i\le n} p(E_i) - \sum_{1\le i<j\le n} p(E_i \cap E_j) + \sum_{1\le i<j<k\le n} p(E_i \cap E_j \cap E_k)$$
$$- \cdots + (-1)^{n+1} p(E_1 \cap E_2 \cap \cdots \cap E_n).$$

Probability Distribution

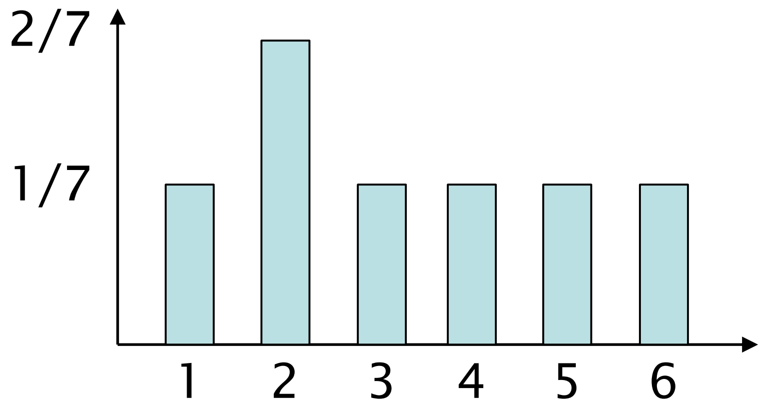
- **Definition**
  Let S be the sample space of an experiment with a finite number of outcomes. A probability distribution on S is characterized by a probability mass function (pmf) $p: S \to \mathbf{R}$ such that
  (a) $0 \leq p(s) \leq 1$ for each $s \in S$,
  (b) $\sum_{s \in S} p(S) = 1$,
  where $p(s)$ is the probability of an outcome s.

Uniform Distribution

- **Definition**
  Suppose that S is a set with n elements. The uniform distribution assigns the probability $1/n$ to each element of S.
- **Example**:
  For a fair dice with 6 sides, we have $p(x) = \frac{1}{6}$ for all x.
  What is the probability that an odd number appears when we roll this dice?

Non-Uniform Distribution

- **Example**:
  Suppose that a dice is biased so that
  (1) $p(2) = \frac{2}{7}$,
  (1) $p(x) = \frac{1}{7}$ for all other x.
  What is the probability that an odd number appears when we roll this dice?



Histogram of the pmf

Probability of an Event

- **Definition**:
  The probability of the event E is the sum of the probabilities of the outcomes in E.

$$p(E) = \sum_{s \in S} p(s)$$

- Uniform distributions: probability = count
- General distributions: probability = sum

Complement and Union of Events

- Complement: $p(\overline{E}) = 1 - p(E)$
- Union: $p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$
- Inclusion-exclusion principle for probability

$$p\Big(\bigcup_{i=1}^{n} E_i\Big) = \sum_{1 \le i \le n} p(E_i) - \sum_{1 \le i < j \le n} p(E_i \cap E_j) + \sum_{1 \le i < j < k \le n} p(E_i \cap E_j \cap E_k)$$
$$- \cdots + (-1)^{n+1} p(E_1 \cap E_2 \cap \cdots \cap E_n).$$

- Disjoint union:
  If $E_1, E_2, \ldots$ is a sequence of pairwise disjoint events in a sample space S, then

$$p\left(\bigcup_i E_i\right) = \sum_i p(E_i)$$

Independence

- **Example**
  Roll two dice. What's the probability that they both turn up heads?
- **Definition**
  Two events E and F are independent if and only if $p(E \cap F) = p(E)p(F)$.
- How to check independence?
  - Two unrelated events are independent
  - Use definition

Example

- **Example**
  In a randomly generated bit string of length 4:
  E: it begins with a 1
  F: it contains an even number of 1s.
  Are E and F independent?
- **Solution**
  - $p(E) = 1/2$.
  - $p(F) = 1/2$.
  - $p(E \cap F) = |\{1111, 1100, 1010, 1001\}|/16 = 1/4$.
  - So E and F are independent.

Example

- **Example**:
  Suppose a family have 3 children, each of which has equal probability to be a boy or a girl. Are these two events independent?
  - E: The family has children of both sexes
  - F: The family has at most one boy
- **Solution**:
  - $p(E) = \frac{|\{BBG, BGB, BGG, GBB, GBG, GGB\}|}{8} = \frac{3}{4}$
  - $p(F) = \frac{|\{BGG, GBG, GGB, GGG\}|}{8} = \frac{1}{2}$
  - $p(E \cap F) = \frac{|\{BGG, GBG, GGB\}|}{8} = \frac{3}{8}$
  - So they are independent

Pairwise and Mutual Independence

- **Definition**:
  The events $E_1, E_2, \ldots, E_n$ are pairwise independent if and only if $p(E_i \cap E_j) = p(E_i)\, p(E_j)$ for all pairs i and j with $1 \leq i < j \leq n$.
- **Definition**:
  The events are mutually independent if
  $$p(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_m}) = p(E_{i_1})p(E_{i_2}) \cdots p(E_{i_m})$$
  whenever $i_j, j = 1,2, \ldots, m$, are integers with
  $$1 \leq i_1 < i_2 < \cdots < i_m \leq n \text{ and } m \geq 2.$$
- Note
  - Mutual independence ➔ pairwise independence,
  - but the reverse is not necessarily true.

Random Variables

- **Definition**
  A random variable is a function from the sample space of an experiment to the set of real numbers, i.e., a random variable assigns a real number to each possible outcome.
- Note: This function is not random!
- **Example**
  Suppose that a coin is flipped three times. Let $X(t)$ be the random variable that equals the number of heads that appear when t is the outcome. Then $X(t)$ is the function where:

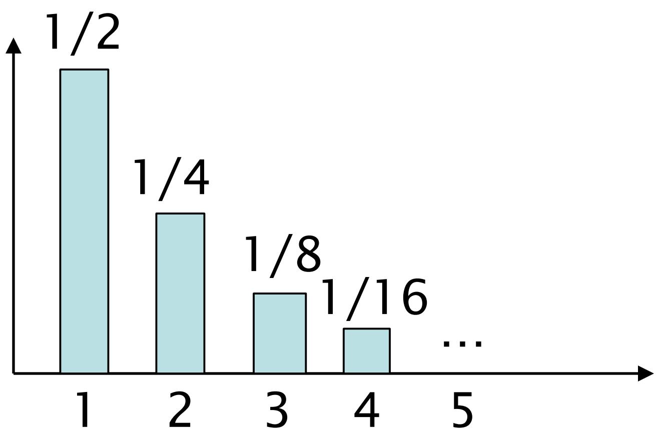
$$X(HHH) = 3, X(TTT) = 0,$$
$$X(HHT) = X(HTH) = X(THH) = 2,$$
$$X(TTH) = X(THT) = X(HTT) = 1.$$

Distribution of a Random Variable

- **Definition**
  The distribution of a random variable X on a sample space S is the set of pairs $(r, p(X = r))$ for all $r \in X(S)$, where $p(X = r)$ is the probability that X takes the value r, i.e., $p(X = r) = \sum_{s \in S, X(s) = r} p(s)$
- **Example**
  Continuing with the previous example, we have
  $p(X = 3) = 1/8$
  $p(X = 2) = 3/8$
  $p(X = 1) = 3/8$
  $p(X = 0) = 1/8$

Infinite Sample Space

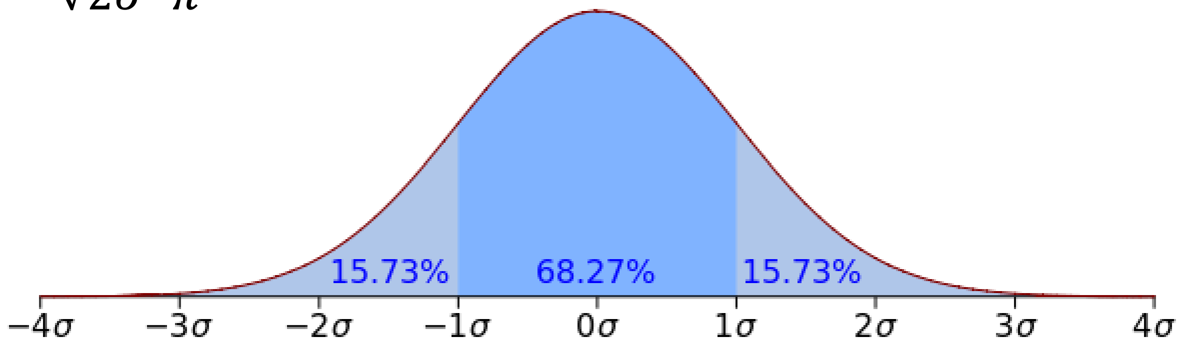
- Example
  Experiment: Flip a fair coin until it turns up heads.
  Let random variable X = the number of flips.



- Verify $\sum_{x=1}^{\infty} p(X = x) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 1$
- The probability of the event that we get 3 or more flips = $p(X = 3) + p(X = 4) + \cdots = \frac{1}{8} + \frac{1}{16} + \cdots = \frac{1}{4}$.

Continuous Probability Distribution

- Defined over a domain of real numbers
- Characterized by a probability density function (pdf)
- Example: The pdf of a normal (Gaussian) distribution is $f(x) = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.



- Remark: $f(x)$ does not represent a probability!
- The probability of an interval: $\Pr[a \leq X \leq b] = \int_a^b f(x)dx$

Bernoulli Trials

- Definition:
  Each performance of an experiment with two possible outcomes is called a Bernoulli trial. The two possible outcomes are generally referred to as success and failure.
- Example:
  A coin is biased so that the probability of heads is $2/3$. What is the probability that exactly four heads occur when the coin is flipped seven times?
- Solution:
  $C(7,4)\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^3$

Binomial Distribution

- **Theorem**
  Let X be the number of successes in n independent Bernoulli trials, with probability of success p and probability of failure $q = 1 - p$. Then
  $$p(X = k) = b(k: n, p) = C(n, k)p^k q^{n-k}$$
- **Definition**:
  X follows the binomial distribution with parameters n and p, and $b(k: n, p)$ is the pmf.
- Check (from the binomial theorem)
  $$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p + q)^n = 1$$

Binomial Distribution

- **Question**
  What is $\Pr(10 \le X \le 20)$?
- **Solution**
  $\sum_{k=10}^{20} \binom{n}{k} p^k q^{n-k}$,
  which, unfortunately, has no closed form.
- For large n, and p not too close to 0 or 1, the normal distribution approximates binomial distribution well.
  - A special case of the central limit theorem.