Discrete Structures

Discrete Probability

Finite Probability Space

- 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 possible outcomes.
- An event is an subset of the sample space.
- A (finite) probability space is a pair (Ω, P) where Ω is a finite set (the sample space) and P is an additive measure on subsets of Ω with $P(\Omega) = 1$. Any subset of Ω is called an event and each element of Ω is called an elementary event. For event A, $P[A]$ is called the probability of A. and is equal to $\sum_{a \in A} P[a]$.

Finite Probability Space

- The probability measure (probability distribution) is determined by its value on elementary events: in other words, by specifying a function $P\colon \Omega \to [0,1]$ with $\sum_{\omega \in \Omega} P(\omega) = 1$. Then, the probability measure on an event A is given by $\mathrm{P}[\mathrm{A}] = \sum_{a \in A} P[a]$

- The basic example of a probability measure is the uniform distribution on Ω, where $P[A] = \frac{|A|}{|\Omega|}$ for any $A \subseteq \Omega$. Such a distribution represents the situation where any outcome of an experiment (such as rolling a dice) is equally likely.

Examples

Example: Consider the experiment of tossing a fair coin. We have $\Omega = \{H, T\}$ and $P[H] = P[T] = 1/2$.

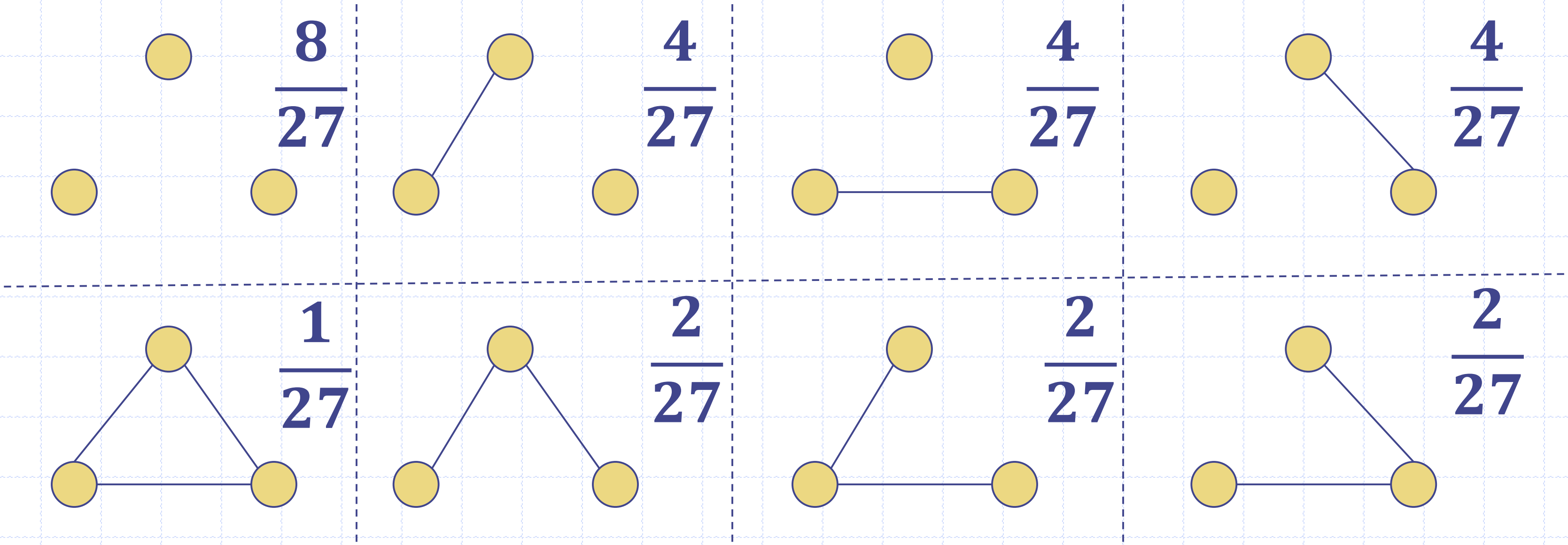
Example: Consider the experiment of tossing a fair coin twice. We have $\Omega = \{(H,H),(H,T),(T,H),(T,T)\}$ and $P[(H,H)] = P[(T,T)] = P[(H,T)] = P[(T,H)] = 1/4$.

Example: Consider the experiment of rolling a fair dice. We have $\Omega = \{1,2,3,4,5,6\}$ and $P[1] = \cdots = P[6] = 1/6$.

Example: Consider the experiment of rolling a fair dice first and then tossing a fair coin. We have $\Omega = \{(1,H), \ldots, (6,H), (1,T), \ldots, (6,T)\}$ and $P[(i,H)] = P[(i,T)\,] = 1/12$ for any i

The Probability Space $G(n,p)$

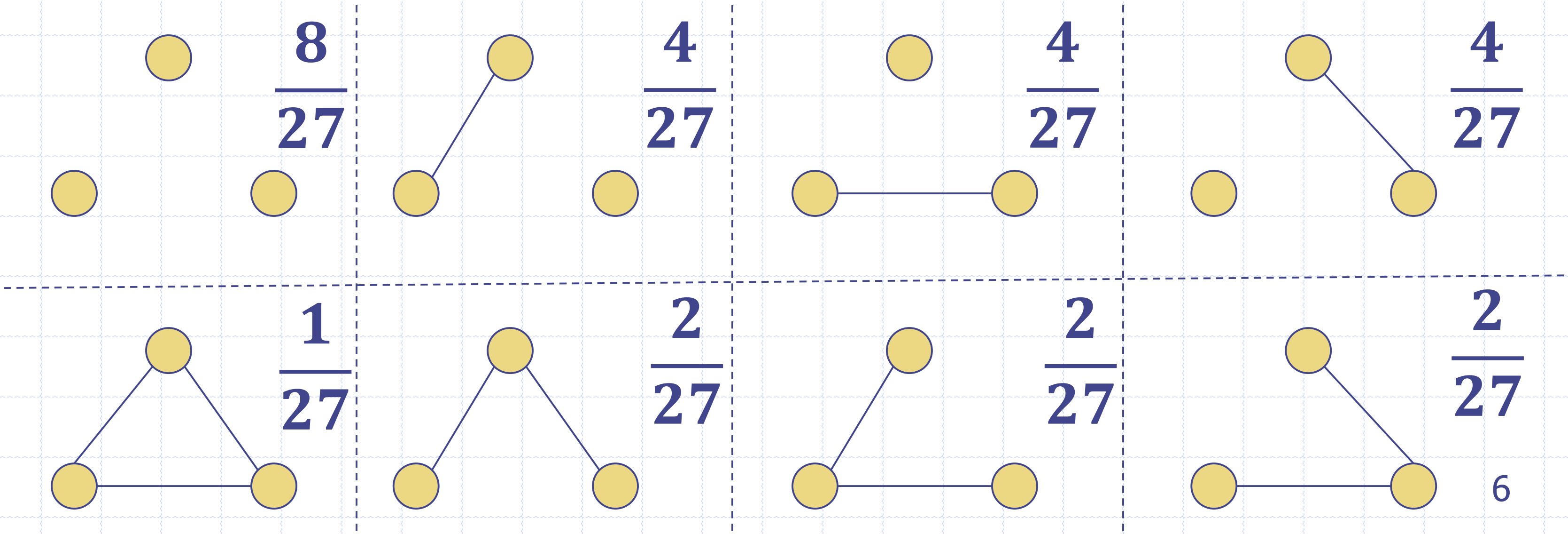
P.S. $G(n,p)$: Consider the experiment of constructing a graph with n vertices where each edge appears in the graph with probability p. The sample space Ω of $G(n,p)$ is the set of all graphs on a fix set of n vertices. The probability of graph with m edges is $p^m(1-p)^{\binom{n}{2}-m}$. For $n=3$ and $p=1/3$ elements of Ω with their probability are illustrated below.



The Probability Space $G(n,p)$

How to do the experiment (construct a random graph)

- Method 1: For each edge toss a coin where $P[H] = p, P[T] = 1 - p$
- Method 2: Produce all elements of the sample space with their probability like below. Then u.a.r. select an integer number in $[1,27]$. If it is between 1 and 8, select top-left one, if it is between 9 to 12, select the one next to top-left one, and so on.



Independent Sets

Given a graph G. Consider a probability space whose sample space Ω is the set of independent sets of G (an independent set is a subset of vertices that there is no edge between any two vertices of the subset) and whose probability measure is an uniform distribution.

How to do the experiment (select a random independent set)

- One simple way is to produce all independent set and number them from 1 to m where m is the number of independent set of G. Then select an integer k in $[1, m]$ u.a.r. and report the independent set whose number is k. This method is not efficient (takes too much time as m can be exponential). Any faster method?

Elementary and Useful Lemmas

Lemma: For any collection of events $A_1, \ldots, A_n$:

$$P[\cup A_i] \leq \sum_{i=1}^{n} P[A_i]$$

Proof:

Let $B_i = A_i \backslash A_1 \cup \cdots \cup A_{i-1}$.

We have $\cup A_i = \cup B_i$ and B_is are disjoint. Therefore

$$P[\cup A_i] = P[\cup B_i] = \sum_{i=1}^{n} P[B_i] \leq \sum_{i=1}^{n} P[A_i]$$

Lemma: For any event A, $P[\bar{A}] = 1 - P[A]$

Lemma: For any two events A and B, $P[A \cup B] = P[A] + P[B] - P[A \cap B]$

Independent Events

Definition: Two events A and B are called independent iff $P[A \cap B] = P[A]P[B]$

Problem: Consider the probability space of tossing a coin twice s.t. $P[H] = P[T] = 1/2$. Let A be the event that at the first toss H comes up and let B be the event that at the second toss T comes up. Are A and B independent?

Solution:

$\Omega = \{(H,H),(H,T),(T,H),(T,T)\}, A = \{(H,T),(H,H)\}, B = \{(H,T),(T,T)\}, A \cap B = \{(H,T)\}, \mathrm{P[A]} = \mathrm{P[B]} = \frac{2}{4} = \frac{1}{2}, P[A \cap B] = \frac{1}{4}$. Then $P[A \cap B] = P[A]P[B]$.

It was clear A and B are independent. No need all these calculations.

Independent Events

Problem: Consider the probability space of tossing a coin twice s.t. $P[H] = P[T] = 1/2$. Let A be the event that at the first toss H comes up and let B be the event that both tosses are the same. Are A and B independent?

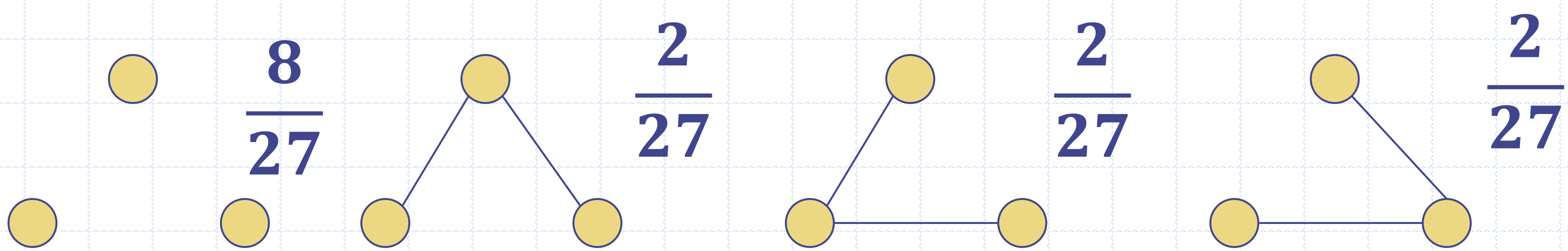
Solution: It is not clear whether A and B are independent. We have to do some calculations.

$\Omega = \{(H,H),(H,T),(T,H),(T,T)\}, A = \{(H,T),(H,H)\}, B = \{(H,H),(T,T)\}, A \cap B = \{(H,H)\}, \mathrm{P[A]} = \mathrm{P[B]} = \frac{2}{4} = \frac{1}{2}, P[A \cap B] = \frac{1}{4}$. Then $P[A \cap B] = P[A]P[B]$.

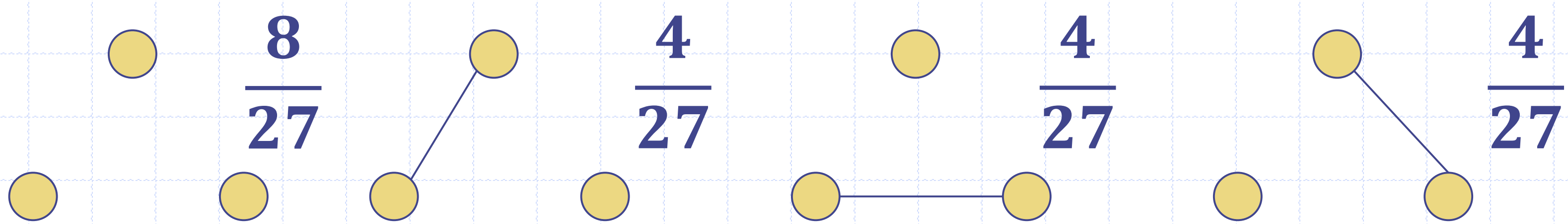
Independent Events

Problem: Consider the probability space $G(3,1/3)$. Let A be the event that the number of edges is even and Let B be the event that the maximum degree is at most 1. Are A and B independent?

Event A



Event B



, $\mathrm{P[A]} = \frac{14}{27}$, $\mathrm{P[B]} = \frac{20}{27}$, $P[A \cap B] = \frac{8}{27}$. Then $P[A \cap B] \neq P[A]P[B]$.

Independent Events

Definition: Three events A, B and C are called independent iff $P[A \cap B] = P[A]P[B]$, $P[A \cap C] = P[A]P[C]$, $P[B \cap C] = P[B]P[C]$, $P[A \cap B \cap C] = P[A]P[B]P[C]$. It is extendable to more events.

Problem: Give an example of three event A, B and C s.t. they are mutually independent but not all independent.

Solution: Consider the probability space of tossing a coin twice s.t. $P[H] = P[T] = 1/2$. Let A be the event that at the first toss H comes up and let B be the event that at the second toss T comes up and let C be the event that both tosses are the same. These three events are mutually independent but

$$0 = P[A \cap B \cap C] \neq P[A]P[B]P[C] = 1/8$$

Conditional Probability

Definition: For events A and B with $P[B] > 0$, we define the conditional probability of A, given that B occurs, as $P[A|B] = \frac{P[A \cap B]}{P[B]}$

Remark: Since B has been occurred, then we can pretend our sample space is B. Then to compute $P[A|B]$ we must just look at $A \cap B$ ($A - B$ does not happen (i.e. $P[A - B] = 0$) when we know B has been occurred). Therefore, $P[A|B]$ is the percentage of probability $A \cap B$ in probability B which is $\frac{P[A \cap B]}{P[B]}$.

Lemma: A and B are independent iff $P[A|B] = \mathrm{P[A]}$

Lemma: For any two events A and B, we have

$$P[A] = P[A|B]P[B] + P[A|\bar{B}]P[\bar{B}]$$

$$P[B|A] = \frac{P(A|B)P(B)}{P[A|B]P[B] + P[A|\bar{B}]P[\bar{B}]}$$

Bernoulli Trial

- Suppose that an experiment can have only two possible outcomes. Each performance of an experiment with two possible outcomes called a Bernoulli trial.
- A possible outcome of a Bernoulli trial is called a success or failure.
- If p is the probability of a success, then the probability of a failure is $1 - p$.
- The probability of k successes when an experiment consists of n mutually independent Bernoulli trials is $\binom{n}{k} p^k (1-p)^{n-k}$ (that the sample space is n-tuples $(t_1, \dots, t_n)$ where $t_i = S$ or $t_i = F$)

The Probabilistic Method

- Assume we would like to prove the existence of a object with specified properties (a good object)
- Sometimes, an explicit construction of such a good object does not seem feasible.
- Or maybe we do not even need a specific example of good object and we just want to prove something good exists
- Then we can consider a random object from a suitable probability space and calculate the probability that it satisfies our conditions.
- If we prove that the probability is strictly positive, then we prove that a good object must exist.
- If all objects were bad, the probability would be zero.

Ramsey Number

Definition: $R(k,l)$ =min{n: any graph on n vertices contains a clique of size k or an independent set of size l}. For instance $R(3,3) = 6, R(2,l) = l$

Problem: For any $\mathrm{k} \geq 3, R(k,k) \geq 2^{\frac{k}{2}-1}$

Solution:

- Let us consider a random graph $G(n, 1/2)$ on n vertices where every pair of vertices forms and edge with probability 1/2.
- For any fix set of k vertices, the probability that they form a clique is $2^{-\binom{k}{2}}$. The same goes for the occurrence of an independent set.
- There are $\binom{n}{k}$ k-tuples of vertices where a clique or an independent set might appear.

Ramsey Number

- We use the fact that the probability of a union of events is at most the sum of their respective probabilities, and we get

$P[G(n,1/2)$ contains a clique or an independent set of size $k] \leq 2\binom{n}{k}2^{-\binom{k}{2}}$

To clarify the above, for a subset V of vertices of size k, let's define two events $C_V$and I_V to be graphs that V is clique and independent set in those graphs.

We know $P[\cup C_V] \leq \sum P[C_V]$ and $P[\cup I_V] \leq \sum P[I_V]$

Where sum and union are over all subsets V of size k.

Ramsey Number

- It remains to choose n so that $2\binom{n}{k}2^{-\binom{k}{2}} < 1$.
- If we set $n = 2^{\frac{k}{2}-1}$ and use the inequality $\binom{n}{k} \le n^k$, the above holds.
- Therefore, there are graphs on $2^{\frac{k}{2}-1}$ vertices that contain neither a clique of size k or independent set of size k. This implies $R(k,k) \ge 2^{\frac{k}{2}-1}$

Random Variable

- A Random Variable X is a function from the sample space to real numbers (i.e $X: \Omega \to R$). A random variable assign a real number to each possible outcome. procedure that yields one of a given set of possible outcomes.
- Note that a random variable is a function. It is not a variable, and it is not random.
- The distribution of a random variable X:

$$P[X = r] = \sum_{\omega : X(\omega) = r} P[\omega]$$

- Consider event $A_r = \{\omega : X(\omega) = r\}$. So $P[X = r] = P[A_r]$
- Since the input of X is a random phenomena, the output of X is a random number in the codomain of X with the above distribution.

Examples

Problem: Suppose that a fair coin is flipped three times. Let $X(\omega)$ be the random variable that equal to the number of heads that appear when ω is the outcome. Then $X(\omega)$ takes on the following values.

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

$$P[X = 0] = \frac{1}{8}, P[X = 1] = \frac{3}{8}, P[X = 2] = \frac{3}{8}, P[X = 3] = \frac{1}{8}$$

Examples

Problem: Let X be the sum of the numbers that appear when a pair of dice is rolled. What are the values of this random variable for the 36 possible outcome (i, j)?

Solution:

$$
\begin{aligned}
&X((1,1)) = 2,\\
&X((1,2)) = X((2,1)) = 3,\\
&X((1,3)) = X((2,2)) = X((3,1)) = 4,\\
&X((1,4)) = X((2,3)) = X((3,2)) = X((4,1)) = 5,\\
&X((1,5)) = X((2,4)) = X((3,3)) = X((4,2)) = X((5,1)) = 6,\\
&X((1,6)) = X((2,5)) = X((3,4)) = X((4,3)) = X((5,2)) = X((6,1)) = 7,\\
&X((2,6)) = X((3,5)) = X((4,4)) = X((5,3)) = X((6,2)) = 8,\\
&X((3,6)) = X((4,5)) = X((5,4)) = X((6,3)) = 9,\\
&X((4,6)) = X((5,5)) = X((6,4)) = 10,\\
&X((5,6)) = X((6,5)) = 11,\\
&X((6,6)) = 12.
\end{aligned}
$$

The Expected Value

The Expected Value of a random variable X on the sample space Ω is equal to $E(X) = \sum_{\omega \in \Omega} P[\omega] X(\omega)$

Problem: what is the expected edges of a random graph in $G(3,1/3)$.

Solution:

Let $X(\omega)$ be the number of edges ofω where ω is a graph in the sample space of $G(3,1/3)$. We should compute $\mathrm{E}(X)$ which is $\sum_{\omega \in \Omega} P[\omega] X(\omega) = \frac{8}{27} \times 0 + \frac{4}{27} \times 1 + \frac{4}{27} \times 1 + \frac{4}{27} \times 1 + \frac{2}{27} \times 2 + \frac{2}{27} \times 2 + \frac{2}{27} \times 2 + \frac{1}{27} \times 3 = \frac{8}{27} \times 0 + \left(\frac{4}{27} + \frac{4}{27} + \frac{4}{27}\right) \times 1 + \left(\frac{2}{27} + \frac{2}{27} + \frac{2}{27}\right) \times 2 + \frac{1}{27} \times 3 = \sum_{r \in R} P[X = r] r$

$$E(X) = \sum_{\omega \in \Omega} P[\omega] X(\omega) = \sum_{r \in R} P[X = r] r$$

The Expected Value

Problem: Show that the expected number of successes when n independent Bernoulli trials are performed, where p is the probability of success on each trial, is np

Solution: Let X be the random variable equal to the number of successes in n trials. We know $\mathrm{P}[\mathrm{X} = \mathrm{k}] = \binom{n}{k} p^k (1-p)^{n-k}$. So

$$E(X) = \sum_{k=0}^{n} kP[X = k] = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{n} n \binom{n-1}{k-1} p^k (1-p)^{n-k}$$

$$= np \sum_{k=0}^{n} n \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} = np$$

Linearity of Expectations

Theorem: Let X_i $(i = 1, \ldots, n)$ be random variables on the sample space Ω, and let a and b be two real numbers, then

(i) $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$

(ii) $E(aX + b) = aE(X) + b$

Proof:

(i) $E(X_1 + X_2) = \sum_{\omega\in\Omega} P[\omega](X_1(\omega) + X_2(\omega)) =$
$\sum_{\omega\in\Omega} P[\omega]X_1(\omega) + \sum_{\omega\in\Omega} P[\omega]X_2(\omega) = E(X_1) + E(X_2)$

(ii) $E(aX + b) = \sum_{\omega\in\Omega} P[\omega](aX(\omega) + b) =$
$a\sum_{\omega\in\Omega} P[\omega]X(\omega) + b\sum_{\omega\in\Omega} P[\omega] = aE(X) + b$

Linearity of Expectations

Problem: Compute the expected value of the sum of the numbers that appear when a pair of fair dice is rolled.

Solution:

One way is to list 36 outcomes and compute the value of X and its probability like below.

$p(X=2) = p(X=12) = 1/36,$

$p(X=3) = p(X=11) = 2/36 = 1/18,$

$p(X=4) = p(X=10) = 3/36 = 1/12,$

$p(X=5) = p(X=9) = 4/36 = 1/9,$

$p(X=6) = p(X=8) = 5/36,$

$p(X=7) = 6/36 = 1/6.$

$$E(X) = 2\cdot\frac{1}{36} + 3\cdot\frac{1}{18} + 4\cdot\frac{1}{12} + 5\cdot\frac{1}{9} + 6\cdot\frac{5}{36} + 7\cdot\frac{1}{6} + 8\cdot\frac{5}{36} + 9\cdot\frac{1}{9} + 10\cdot\frac{1}{12} + 11\cdot\frac{1}{18} + 12\cdot\frac{1}{36} = 7.$$

Other way is to define $X_1(i,j) = i, X_2(i,j) = j$. We have $E(X_1) = \mathrm{E}(X_2) = \frac{1+2+3+4+5+6}{6} = \frac{7}{2}$. So $E(X_1+X_2) = E(X_1) + \mathrm{E}(X_2) = 7$

Linearity of Expectations

Problem: What is the expected number of trials in the Bernoulli trials to get the fist success where the probability of success is p.

Solution:

Let X be number of trials in the Bernoulli trials to get the fist success. It is easy to see

$P[X = k] = (1-p)^{k-1}p$. So

$$E(X) = \sum_{k=1}^{\infty}(1-p)^{k-1}pk = p\sum_{k=1}^{\infty}(1-p)^{k-1}k = p \times \frac{1}{p^2} = \frac{1}{p}$$

Note that $\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k \rightarrow \frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} kx^{k-1}$

Indicator Random Variable

Definition: For an event A, we define the indicator variable I_A:

- $I_A(\omega) = 1$ if $\omega \in A$, and
- $I_A(\omega) = 0$ if $\omega \notin A$

Lemma: $E(I_A) = 1 \times P[A] + 0 \times (1 - \mathrm{P[A]}) = \mathrm{P[A]}$

Problem: Compute the expected number of successes when n independent Bernoulli trials are performed, the probability of success is p.

Solution: Let A_i be the event that the i-th trial yields to success and let X_i be the I.R.V. for event A_i. Then the random variable X, the number of successes, is equal to $X_1 + \cdots + X_n$. Therefore, we have

$$E(X) = E(X_1) + \cdots + E(X_n) = np$$

as we know $E(X_i) = \mathrm{P}[A_i] = p$

Random Permutation

- **Problem:** For given n produce a random permutation σ of 1, 2, …,n with uniform distribution.
- **Solution:** As usual, the simple way is to produce all $n!$ permutation and number them from 1 to $n!$ and then select a number k u.a.r. and report the permutation whose number is k. A more efficient way is to select $\sigma(1)$, the leftmost item of the permutation, u.a.r. from set $\{1,2,\ldots,n\}$, the select $\sigma(2)$ u.a.r from set $\{1,2,\ldots,n\} - \{\sigma(1)\}$, and so on. How to program this?
- **Problem:** Let σ be random permutation and let X be the number of i s.t. $\sigma(i) = i$. Show that $E(X) = 1$
- **Solution:** Let A_i be the event $\sigma(i) = i$ and let X_i be its I.R.V. So we have $\mathrm{X} = X_1 + \cdots + X_n$ and $E(X) = E(X_1) + \cdots + E(X_n) = 1/\mathrm{n} + \cdots + 1/\mathrm{n} = 1$ as $E(X_i) = \mathrm{P}[A_i] = 1/n$

Independent Random Variable

Definition: Two random variables X and Y on the sample space Ω are independent iff

$$\forall x, y\colon\ P[X = x \wedge Y = y] = P[X = x]P[Y = y]$$

In other words, let events $A_r = \{\omega \in \Omega : X(\omega) = r\}$ and $B_r = \{\omega \in \Omega : Y(\omega) = r\}$, then $\forall x, y\colon A_x, B_y$ are independent

Problem: Let $\{(i, j) : 1 \le i, j \le 6\}$ be the sample space of rolling a dice twice. Specify whether X and Y are independent or not

(i) $X(i,j) = i, Y(i,j) = j$

(ii) $X(i,j) = i, Y(i,j) = i + j$

Solution:

(i) $\mathrm{P}[X = i \wedge Y = j] = \frac{1}{36}, P[X = i] = P[Y = j] = \frac{6}{36} = \frac{1}{6}$

(ii) $\mathrm{P}[X = 1 \wedge Y = 12] = 0, P[X = 1] = \frac{1}{6}, P[Y = 12] = \frac{1}{36}$

Independent Random Variable

Theorem: Let X and Y be two independent random variables on Ω, then $E(XY) = E(X)E(Y)$.

Proof:

$$E(XY) = \sum_{\omega \in \Omega} P[\omega] X(\omega) Y(\omega)$$

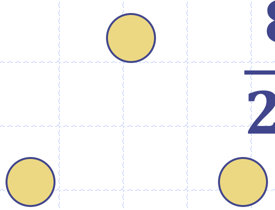
$$= \sum_{x \in X(\Omega), y \in Y(\Omega)} xyP[X = x \wedge Y = y]$$

$$= \sum_{x \in X(\Omega), y \in Y(\Omega)} xyP[X = x]P[Y = y]$$

$$= \sum_{x \in X(\Omega)} xP[X = x] \sum_{y \in Y(\Omega)} yP[Y = y] = E(X)E(Y)$$

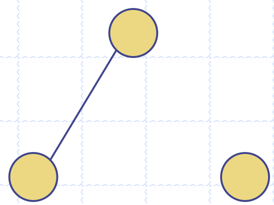
Independent Random Variable

Problem: Let X and Y be the minimum degree and the number of edges of graph ω from the sample space of $G(3,1/3)$. check whether $E(XY)=E(X)E(Y)$.



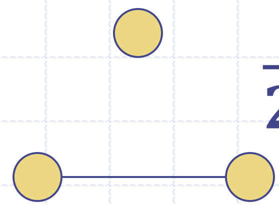
$\frac{8}{27}$

$X=0$, $Y=0$, $XY=0$



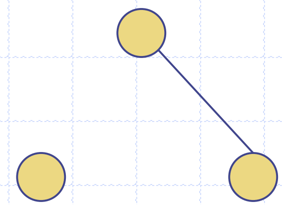
$\frac{4}{27}$

$X=0$, $Y=1$, $XY=0$



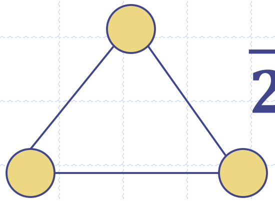
$\frac{4}{27}$

$X=0$, $Y=1$, $XY=0$



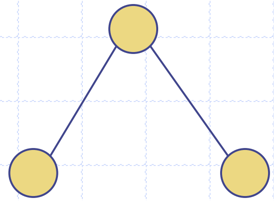
$\frac{4}{27}$

$X=0$, $Y=1$, $XY=0$



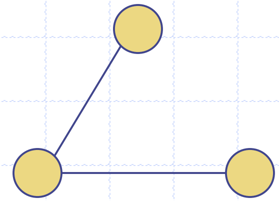
$\frac{1}{27}$

$X=2$, $Y=3$, $XY=6$



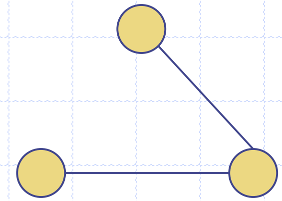
$\frac{2}{27}$

$X=1$, $Y=2$, $XY=2$



$\frac{2}{27}$

$X=1$, $Y=2$, $XY=2$



$\frac{2}{27}$

$X=1$, $Y=2$, $XY=2$

$E(X)=8/27, E(Y)=1, \mathrm{E(XY)}=18/27 \rightarrow E(XY)\neq E(X)E(Y)$

Variance

- The expected value of a random variable tells us its average value, but nothing about how widely its values are distributed. For instance define $X(H) = X(T) = 0$, and $Y(H) = 1000, Y(T) = -1000$ for tossing a fair coin. We have $E(X) = E(Y)$ but X never varies from 0 while Y always differ from 0 by 1000.
- Then we need a measure telling us how far the values of a random variable X from $E(X)$ are (what the deviation of X is). One good option is to measure this by $\sum_{\omega\in\Omega} P[\omega]|X(\omega) - E(X)| = E(|X - E(X)|)$. But unfortunately the absolute value annoys us while doing calculation. Due to this $Var(X)$ is define to be $\sum_{\omega\in\Omega} P[\omega](X(\omega) - E(X))^2 = E((X - E(X))^2)$ and deviation of X denoted by $\sigma(X)$ is defined to be $\sqrt{Var(X)}$.

Variance

Theorem: If X is a random variable on the sample space Ω, we have $Var(X) = E(X^2) - E(X)^2$

Proof:

$$Var(X) = E\left(\left(X - E(X)\right)^2\right) = \sum_{\omega\in\Omega} P[\omega]\left(X(\omega) - E(X)\right)^2$$

$$= \sum_{\omega\in\Omega} P[\omega]X(\omega)^2 - 2E(X)\sum_{\omega\in\Omega} P[\omega]X(\omega) + E(X)^2\sum_{\omega\in\Omega} P[\omega]$$

$$= E(X^2) - 2E(X)^2 + E(X)^2 = E(X^2) - E(X)^2$$

Variance

Theorem: If X and Y are two independent random variable on the sample space Ω, we have $Var(X+Y) = Var(X) + Var(Y)$. This is hold for n pairwise independent random variables as well.

Proof:

$$Var(X + \mathrm{Y}) = \mathrm{E}\left((X+Y)^2\right) - E(X+Y)^2$$

$$= E(X^2) + 2\mathrm{E}(\mathrm{X}Y) + E(Y^2) - E(X)^2 - 2E(X)E(Y) - E(Y)^2$$

$$= E(X^2) - E(X)^2 + E(Y^2) - E(Y)^2 = \mathrm{Var(X)} + \mathrm{Var(Y)}$$

Variance

Problem: Consider n Bernoulli trails with success probability p. Let X be the number of successes. Compute $Var(X)$.

Solution:

As you see before $X = X_1 + \cdots + X_n$ where X_is are independent. So $\mathrm{Var}(X) = Var(X_1) + \cdots + Var(X_n)$.

We have

$$E(X_i) = p, Var(X_i) = p(1-p)^2 + (1-p)(0-p)^2$$
$$= p(1-p)$$

So $\mathrm{Var}(X) = \mathrm{np}(1-\mathrm{p})$

Recall that $E(X) = np$

Markov's Inequality

Theorem: If X is a non-negative random variable and t is real number, we have $P[X \geq t] \leq \frac{\mathrm{E(X)}}{t}$.

Proof:

$$E(X) = \sum_{0 \leq r} rP(X = r) = \sum_{0 \leq r < t} rP(X = r) + \sum_{t \leq r} rP(X = r)$$

$$\geq \sum_{t \leq r} rP(X = r) \geq \sum_{t \leq r} tP(X = r) = t \sum_{t \leq r} P(X = r) = \mathrm{tP[X}$$

$$\geq t] \rightarrow P[X \geq t] \leq \frac{\mathrm{E(X)}}{t}$$

Markov's inequality can be written of form

$$P[X \geq t\mathrm{E(X)}] \leq \frac{1}{t}$$

Chebyshev's Inequality

Theorem: If X is a random variable and t is positive real number, we have $P[|X - \mathrm{E}(\mathrm{X})| \geq t\sigma(X)] \leq \frac{1}{t^2}$.

Proof:

Let $Y = (X - E(X))^2$. We know $Y \geq 0$. So,

$$P\left[Y \geq \left(t\sigma(X)\right)^2\right] \leq \frac{E(Y)}{\left(t\sigma(X)\right)^2} = \frac{Var(X)}{t^2 Var(X)} = \frac{1}{t^2} \rightarrow$$

$$P[\left(X - E(X)\right)^2 \geq \left(t\sigma(X)\right)^2] \leq \frac{1}{t^2} \rightarrow$$

$$P[|X - \mathrm{E}(\mathrm{X})| \geq t\sigma(X)] \leq \frac{1}{t^2}$$

Variance

Problem: Consider n Bernoulli trails with success probability 1/2. Let X be the number of successes. Show $P[X \geq \frac{n}{2} + \sqrt{n}] \leq \frac{1}{4}$.

Solution:

We know $E(X) = \frac{n}{2}, Var(X) = \frac{n}{4}$. Then, if we set $t = 2$

$$P[(\left|X - \frac{n}{2}\right| \geq 2\frac{\sqrt{n}}{2}] \leq \frac{1}{2^2} \rightarrow$$

$$P\left[X \geq \frac{n}{2} + \sqrt{n}\right] \leq P[(\left|X - \frac{n}{2}\right| \geq 2\frac{\sqrt{n}}{2}] \leq \frac{1}{4}$$

If we set $t = \sqrt{n}$, we have $P[X \geq n] \leq \frac{1}{n}$. On the other hand $P[X \geq n] = (\frac{1}{2})^n$. So $(\frac{1}{2})^n \leq \frac{1}{n} \rightarrow n \leq 2^n$

Which is obvious inequality. Just consider this as a probabilistic method to show $n \leq 2^n$