CS 70 Discrete Mathematics and Probability Theory
Fall 2026 Hug, Sabin

DIS 10

Random Variables and Expectation Intro

Note 19
Note 20

Random Variable: A random variable X is a function $X : \Omega \to \mathbb{R}$ that assigns a real number $X(\omega)$ to each outcome ω. The function itself is not random; what is random is which outcome ω occurs. The event "$X = a$" is shorthand for $\{\omega \in \Omega : X(\omega) = a\}$.

Distribution: The distribution of X is the collection $\{(a, \mathbb{P}[X = a]) : a \in \mathscr{A}\}$, where $\mathscr{A}$ is the set of values X can take. In other words, it lists every value X can take, along with the probability of each value. The events $X = a$ partition Ω, so

$$\sum_{a \in \mathscr{A}} \mathbb{P}[X = a] = 1.$$

Common Discrete Distributions: Here, "success" means an event with probability p, and all trials are independent.

Distribution	**Models**	$\mathbb{P}[X = k]$	$\mathbb{E}[X]$
$X \sim \text{Bernoulli}(p)$	one trial; 1 if success	p if $k = 1$, $1 - p$ if $k = 0$	p
$X \sim \text{Binomial}(n, p)$	# successes in n trials	$\binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, \ldots, n$	np
$X \sim \text{Uniform}\{1, \ldots, n\}$	each value equally likely	$\frac{1}{n}$, $k = 1, \ldots, n$	$\frac{n+1}{2}$
$X \sim \text{Geometric}(p)$	# trials until first success (inclusive)	$(1-p)^{k-1} p$, $k = 1, 2, \ldots$	$\frac{1}{p}$
$X \sim \text{Poisson}(\lambda)$	# of rare events, average rate λ	$\frac{\lambda^k}{k!} e^{-\lambda}$, $k = 0, 1, \ldots$	λ

For $X \sim \text{Geometric}(p)$, $\mathbb{P}[X > k] = (1-p)^k$ (the first k trials all fail). This gives the *memoryless property*: $\mathbb{P}[X > n + m \mid X > m] = \mathbb{P}[X > n]$. If $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\mu)$ are independent, then $X + Y \sim \text{Poisson}(\lambda + \mu)$.

Expectation: A weighted average of the values of X, weighted by their probabilities:

$$\mathbb{E}[X] = \sum_{a \in \mathscr{A}} a \cdot \mathbb{P}[X = a] = \sum_{\omega \in \Omega} X(\omega) \cdot \mathbb{P}[\omega].$$

Note that $\mathbb{E}[X]$ need not be a value that X can actually take (e.g., a die roll has $\mathbb{E}[X] = \frac{7}{2}$).

Tail Sum Formula: If X takes values in $\{0, 1, 2, \ldots\}$, then

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} \mathbb{P}[X \geq i].$$

Linearity of Expectation: For *any* random variables X, Y on the same probability space and constants a, b,

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

This holds *whether or not* X and Y are independent. However, it does *not* say that $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$ or $\mathbb{E}[\frac{1}{X}] = \frac{1}{\mathbb{E}[X]}$; these fail in general.

Indicator Variables: The indicator I_A of an event A is 1 if A occurs and 0 otherwise, so $I_A \sim \text{Bernoulli}(\mathbb{P}[A])$ and $\mathbb{E}[I_A] = \mathbb{P}[A]$. To find the expectation of a random variable that *counts* something, write it as a sum of indicators $X = I_1 + \cdots + I_n$, where I_i is the indicator of event A_i, and apply linearity:

$$\mathbb{E}[X] = \mathbb{E}[I_1] + \cdots + \mathbb{E}[I_n] = \mathbb{P}[A_1] + \cdots + \mathbb{P}[A_n].$$

Joint Distribution: The joint distribution of X and Y is the collection of probabilities $\mathbb{P}[X = a, Y = b]$ for all $a \in \mathscr{A}$, $b \in \mathscr{B}$. The *marginal distributions* are found by summing out the other variable:

$$\mathbb{P}[X = a] = \sum_{b \in \mathscr{B}} \mathbb{P}[X = a, Y = b] \qquad \mathbb{P}[Y = b] = \sum_{a \in \mathscr{A}} \mathbb{P}[X = a, Y = b].$$

We can also condition one random variable on another:

$$\mathbb{P}[X = a \mid Y = b] = \frac{\mathbb{P}[X = a, Y = b]}{\mathbb{P}[Y = b]}.$$

Independence of Random Variables: X and Y are independent if the events $X = a$ and $Y = b$ are independent for *all* a, b:

$$\mathbb{P}[X = a, Y = b] = \mathbb{P}[X = a]\,\mathbb{P}[Y = b] \qquad \text{for all } a, b.$$

To show X and Y are *not* independent, it is enough to find a single pair (a, b) where this fails. Mutual independence of more than two random variables is defined similarly. For example, indicators of independent coin tosses are mutually independent and identically distributed (i.i.d.).

1 Head Count

Note 19

Consider a coin with $\mathbb{P}[\text{Heads}] = 2/5$. Suppose you flip the coin 20 times, and define X to be the number of heads.

(a) What is $\mathbb{P}[X = k]$, for some $0 \le k \le 20$? Express your answer in terms of k. (Do not just copy down a formula—re-derive it yourself!)

(b) What is the name of the distribution of X, and what are its parameters?

(c) What is $\mathbb{P}[12 \le X \le 14]$?

(d) What is $\mathbb{P}[X \ge 1]$? *Hint: You should be able to do this without a summation.*

(e) Now consider a second coin also with $\mathbb{P}[\text{Heads}] = 2/5$. Suppose you flip this second coin 30 times, and define Y to be the number of heads. What is the distribution of the *total* number of heads among these two coins, i.e. what is the distribution of $X + Y$?

Solution:

(a) There are a total of $\binom{20}{k}$ ways to select which k flips are heads. The probability that the selected k flips are heads is $(\frac{2}{5})^k$, and the probability that the rest are tails is $(\frac{3}{5})^{20-k}$. Putting this together, we have

$$\mathbb{P}[X = k] = \binom{20}{k} \left(\frac{2}{5}\right)^k \left(\frac{3}{5}\right)^{20-k}.$$

(b) Since we have 20 independent trials, with each trial having a probability $2/5$ of success, we can write $X \sim \text{Binomial}(20, \frac{2}{5})$.

(c) The only way to write out this expression is as a sum of 3 different probabilities:

$$\begin{aligned}\mathbb{P}[12 \le X \le 14] &= \mathbb{P}[X=12] + \mathbb{P}[X=13] + \mathbb{P}[X=14] \\ &= \binom{20}{12}\left(\frac{2}{5}\right)^{12}\left(\frac{3}{5}\right)^{8} + \binom{20}{13}\left(\frac{2}{5}\right)^{13}\left(\frac{3}{5}\right)^{7} + \binom{20}{14}\left(\frac{2}{5}\right)^{14}\left(\frac{3}{5}\right)^{6}.\end{aligned}$$

(d) Note that the probability that there is at least one head is the complement to the probability that there are zero heads. This means that

$$\mathbb{P}[X \ge 1] = 1 - \mathbb{P}[X=0] = 1 - \left(\frac{3}{5}\right)^{20}.$$

(e) Since these two coins have the exact same behavior (i.e. the same bias), we can treat this as an experiment where we flip just one coin 50 times. Here, this means that $X+Y \sim \text{Binomial}(50, \frac{2}{5})$.

Note that more generally, for two independent binomial random variables $X \sim \text{Binomial}(n_1, p)$ and $Y \sim \text{Binomial}(n_2, p)$ with identical success probabilities, we have $X+Y \sim \text{Binomial}(n_1+n_2, p)$, since we're essentially just increasing the number of trials we're considering.

2 Head Count II

Note 19

Consider a coin with $\mathbb{P}[\text{Heads}] = 3/4$. Suppose you flip the coin until you see heads for the first time, and define X to be the number of times you flipped the coin.

(a) What is $\mathbb{P}[X=k]$, for some $k \ge 1$? Express your answer in terms of k. (Do not just copy down a formula—re-derive it yourself!)

(b) What is the name of the distribution of X, and what are its parameters?

(c) What is $\mathbb{P}[X > k]$, for some $k \ge 0$? (You should not have any summations.)

(d) What is $\mathbb{P}[X < k]$, for some $k \ge 1$? (You should not have any summations.)

(e) What is $\mathbb{P}[X > k \mid X > m]$, for some $k \ge m \ge 0$? Show that your answer is equal to $\mathbb{P}[X > k-m]$. Why do we call this the memoryless property?

(f) Suppose $Y \sim \text{Geometric}(p)$ and $Z \sim \text{Geometric}(q)$ are independent. Find the distribution of $\min(Y,Z)$ and justify your answer.

Hint: consider flipping two coins (with $\mathbb{P}[\text{Heads}] = p$ and $\mathbb{P}[\text{Heads}] = q$ respectively) simultaneously.

Solution:

(a) If we flipped k times, then we had $k-1$ tails and 1 head, in that order, giving us

$$\mathbb{P}[X=k] = \frac{3}{4}\left(1-\frac{3}{4}\right)^{k-1} = \frac{3}{4}\left(\frac{1}{4}\right)^{k-1}.$$

(b) $X \sim \text{Geometric}(\frac{3}{4})$

(c) If we had to flip *more than* k times before seeing our first heads, then our first k flips must have been tails, giving us

$$\mathbb{P}[X > k] = \left(1-\frac{3}{4}\right)^{k} = \left(\frac{1}{4}\right)^{k}.$$

You can alternatively write as the sum $\sum_{i=k+1}^{\infty} \mathbb{P}[X=i] = \sum_{i=k+1}^{\infty} \frac{3}{4} * (\frac{1}{4})^{i-1} = \frac{3}{4} * (\frac{1}{4})^{k} * \frac{1}{1-1/4} = (\frac{1}{4})^{k}$ using the formula for an infinite geometric sum.

(d) Notice $\mathbb{P}[X < k] = 1 - \mathbb{P}[X \geq k] = 1 - \mathbb{P}[X > k-1]$ since X can only take on integer values. Along similar lines to the previous part, we then have

$$\mathbb{P}[X < k] = 1 - \mathbb{P}[X > k-1] = 1 - \left(1 - \frac{3}{4}\right)^{k-1} = 1 - \left(\frac{1}{4}\right)^{k-1}.$$

(e) By part (c), we have

$$\mathbb{P}[X > k \mid X > m] = \frac{\mathbb{P}[X > k \cap X > m]}{\mathbb{P}[X > m]} = \frac{\mathbb{P}[X > k]}{\mathbb{P}[X > m]} = \left(\frac{1}{4}\right)^{k-m}.$$

However, note that this is exactly $\mathbb{P}[X > k-m]$. The reason this makes sense is that if we want to compute the probability that the first heads occurs after k flips, and we know that the first heads occurs after m flips, then the first m flips are tails. Thus, by the independence of the coin flips, the first m flips don't matter, and so we only need to compute the probability that the first heads occurs after $k-m$ flips. This is called the **memorylessness property** of the geometric distribution because having flipped the coin m times without seeing heads doesn't affect the distribution of future flips.

(f) Let Y be the number of coins we flip until we see a heads from flipping a coin with bias p, and let Z similarly be the number of coins we flip until we see a heads from flipping a coin with bias q.

Imagine we flip the bias p coin and the bias q coin at the same time. The minimum of the two random variables represents how many simultaneous flips occur before at least one head is seen.

The probability of not seeing a head at all on any given simultaneous flip is $(1-p)(1-q)$; this corresponds to a failure. This means that the probability that there will be a success on any particular trial is $1-(1-p)(1-q) = p+q-pq$. Therefore, $\min(Y,Z) \sim \text{Geometric}(p+q-pq)$.

Alternative 1: We can also solve this algebraically. The probability that $\min(Y,Z) = k$ for some positive integer k is the probability that the first $k-1$ coin flips for both Y and Z were tails, and we get heads on the kth toss (this can come from either Y or Z). Specifically, this occurs with probability

$$((1-p)(1-q))^{k-1} \cdot (p+q-pq)$$

We recognize this as the formula for a geometric random variable with parameter $p+q-pq$.

Alternative 2: An alternative, slightly cleaner approach is to work with the *tail probabilities* of the geometric distribution, rather than with the usual point probabilities as above. Let $W = \min(Y,Z)$. We can work with $\mathbb{P}[W \geq k]$ rather than with $\mathbb{P}[W = k]$; clearly the values $\mathbb{P}[W \geq k]$ specify the values $\mathbb{P}[W = k]$ since $\mathbb{P}[W = k] = \mathbb{P}[W \geq k] - \mathbb{P}[W \geq (k+1)]$, so it suffices to calculate them instead. We then get the following argument:

$$\begin{aligned}
\mathbb{P}[W \geq k] &= \mathbb{P}[\min(Y,Z) \geq k] \\
&= \mathbb{P}[(Y \geq k) \cap (Z \geq k)] \\
&= \mathbb{P}[Y \geq k] \cdot \mathbb{P}[Z \geq k] && \text{since } Y,Z \text{ are independent} \\
&= (1-p)^{k-1}(1-q)^{k-1} && \text{since } Y,Z \text{ are geometric} \\
&= ((1-p)(1-q))^{k-1} \\
&= (1-p-q+pq)^{k-1}.
\end{aligned}$$

This is the tail probability of a geometric distribution with parameter $p+q-pq$, thus we can conclude that $W \sim \text{Geometric}(p+q-pq)$, which is the same result as before!

3 Shuttles and Taxis at Airport

Note 20

In front of terminal 3 at San Francisco Airport is a pickup area where shuttles and taxis arrive according to a Poisson distribution. The shuttles arrive at a rate $\lambda_1 = 1/20$ (i.e. 1 shuttle per 20 minutes) and the taxis arrive at a rate $\lambda_2 = 1/10$ (i.e. 1 taxi per 10 minutes) starting at 00:00. The shuttles and the taxis arrive independently.

(a) Write in terms of a Poisson distribution, the distribution of the number of taxis that arrive between 00:00 and 00:01:

(b) What is the distribution of the following:

(i) The number of taxis that arrive between times 00:00 and 00:20?

(ii) The number of shuttles that arrive between times 00:00 and 00:20?

(iii) The total number of vehicles (shuttles and taxis) that arrive between times 00:00 and 00:20?

(c) What is the probability that exactly 1 shuttle and 3 taxis arrive between times 00:00 and 00:20?

(d) Given that exactly 1 pickup vehicle arrived between times 00:00 and 00:20, what is the conditional probability that this vehicle was a taxi?

(e) Suppose you reach the pickup area at 00:20. You learn that you missed 3 taxis and 1 shuttle in those 20 minutes. What is the probability that you need to wait for more than 10 mins until either a shuttle or a taxi arrives?

Solution:

(a) Let $T([0,1])$ denote the number of taxis that arrive between times 00:00 and 00:01. This interval has length 1 minute, so the number of taxis $T([0,1])$ arriving in this interval is distributed according to $\text{Poisson}(\lambda_2) = \text{Poisson}(\frac{1}{10})$, i.e.

$$\mathbb{P}[T([0,1]) = t] = \frac{\left(\frac{1}{10}\right)^t e^{-\frac{1}{10}}}{t!}, \text{ for } t = 0,1,2,\ldots.$$

(b) (i) Let $T([0,20])$ denote the number of taxis that arrive between times 00:00 and 00:20. This interval has length 20 minutes, so the number of taxis $T([0,20])$ arriving in this interval is distributed according to $\text{Poisson}(\lambda_2 \cdot 20) = \text{Poisson}(2)$, i.e.

$$\mathbb{P}[T([0,20]) = t] = \frac{2^t e^{-2}}{t!}, \text{ for } t = 0,1,2,\ldots.$$

(ii) Let $S([0,20])$ denote the number of shuttles that arrive between times 00:00 and 00:20. This interval has length 20 minutes, so the number of shuttles $S([0,20])$ arriving in this interval is distributed according to $\text{Poisson}(\lambda_1 \cdot 20) = \text{Poisson}(1)$, i.e.

$$\mathbb{P}[S([0,20]) = s] = \frac{1^s e^{-1}}{s!}, \text{ for } s = 0,1,2,\ldots.$$

(iii) Let $N([0,20]) = S([0,20]) + T([0,20])$ denote the total number of pickup vehicles (taxis and shuttles) arriving between times 00:00 and 00:20. Since the sum of independent Poisson random variables is Poisson distributed with parameter given by the sum of the individual parameters, we have $N([0,20]) \sim \text{Poisson}(3)$, i.e.

$$\mathbb{P}[N([0,20]) = n] = \frac{3^n e^{-3}}{n!}, \text{ for } n = 0,1,2,\ldots.$$

(c) We have

$$\mathbb{P}[T([0,20]) = 3] = \frac{2^3 e^{-2}}{3!} \text{ and } \mathbb{P}[S([0,20]) = 1] = \frac{1^1 e^{-1}}{1!}.$$

Since the taxis and the shuttles arrive independently, the probability that exactly 3 taxis and 1 shuttle arrive in this interval is given by the product of their individual probabilities, i.e.

$$\frac{2^3 e^{-2}}{3!} \frac{1^1 e^{-1}}{1!} = \frac{4}{3} e^{-3} \approx 0.0664.$$

(d) Let A be the event that exactly 1 taxi arrives between times 00:00 and 00:20. Let B be the event that exactly 1 vehicle arrives between times 00:00 and 00:20. We have

$$\mathbb{P}[B] = \frac{3^1 e^{-3}}{1!}.$$

Event $A \cap B$ is the event that exactly 1 taxi and 0 shuttles arrive between times 00:00 and 00:20. Hence

$$\mathbb{P}[A \cap B] = \frac{2^1 e^{-2}}{1!} \frac{1^0 e^{-1}}{0!}.$$

Thus, we get

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = 2/3.$$

(e) The event that you need to wait for more than 10 minutes starting 00:20 is equivalent to the event that no vehicle arrives between times 00:20 and 00:30. Let $N([20,30])$ denote the number of vehicles that arrive between times 00:20 and 00:30. This interval has length 10 minutes, so $N([20,30]) \sim \text{Poisson}((\lambda_1 + \lambda_2) \cdot 10) = \text{Poisson}(3/2)$. Since Poisson arrivals in disjoint intervals are independent, we have

$$\mathbb{P}[N([20,30]) = 0 | T([0,20]) = 3, S([0,20]) = 1] = \mathbb{P}[N([20,30]) = 0] = \frac{1.5^0 e^{-1.5}}{0!} = e^{-1.5} \approx 0.2231.$$

4 Linearity

Note 20

Solve each of the following problems using linearity of expectation. Explain your methods clearly.

(a) In an arcade, you play game A 10 times and game B 20 times. Each time you play game A, you win with probability $1/3$ (independently of the other times), and if you win you get 3 tickets (redeemable for prizes), and if you lose you get 0 tickets. Game B is similar, but you win with probability $1/5$, and if you win you get 4 tickets. What is the expected total number of tickets you receive?

(b) A monkey types at a 26-letter keyboard with one key corresponding to each of the lower-case English letters. Each keystroke is chosen independently and uniformly at random from the 26 possibilities. If the monkey types 1 million letters, what is the expected number of times the sequence "book" appears? (*Hint*: Consider where the sequence "book" can appear in the string.)

(c) Sam repeatedly flips a coin with probability of Heads = 0.2 until he gets Heads. Then he rolls a fair 6-sided die with sides $\{1,1,1,1,2,3\}$, and flips the coin that many more times. What is the expected number of coin flips Sam makes in total?

Solution:

(a) Let T be the random variable for the total number of tickets you receive. Then, we can express T as

$$\begin{aligned} T &= 3 \cdot (\text{number of wins in game A}) + 4 \cdot (\text{number of wins in game B}) \\ &= 3 \cdot (A_1 + A_2 + \cdots + A_{10}) + 4 \cdot (B_1 + B_2 + \cdots + B_{20}). \end{aligned}$$

where A_i is the indicator variable that is 1 if you win on the i-th play of game A, and B_i is the indicator variable that is 1 if you win on the i-th play of game B. The expected value of A_i and B_i are

$$\begin{aligned} \mathbb{E}[A_i] &= 1 \cdot \frac{1}{3} + 0 \cdot \frac{2}{3} = \frac{1}{3}, \\ \mathbb{E}[B_i] &= 1 \cdot \frac{1}{5} + 0 \cdot \frac{4}{5} = \frac{1}{5}. \end{aligned}$$

Then by linearity of expectation,

$$\begin{aligned} \mathbb{E}[T] &= 3\,\mathbb{E}[A_1] + \cdots + 3\,\mathbb{E}[A_{10}] + 4\,\mathbb{E}[B_1] + \cdots + 4\,\mathbb{E}[B_{20}] \\ &= 10\left(3 \cdot \frac{1}{3}\right) + 20\left(4 \cdot \frac{1}{5}\right) \\ &= 26. \end{aligned}$$

Note that $10\left(3 \cdot \frac{1}{3}\right)$ and $20\left(4 \cdot \frac{1}{5}\right)$ matches the expression directly gotten using the expected value of a binomial random variable.

(b) There are $1,000,000 - 4 + 1 = 999,997$ places where "book" can appear, each with a (non-independent) probability of $1/26^4$ of happening. If A is the random variable that tells how many times "book" appears, and A_i is the indicator variable that is 1 if "book" appears starting at the ith letter, then

$$\begin{aligned} \mathbb{E}[A] &= \mathbb{E}[A_1 + \cdots + A_{999,997}] \\ &= \mathbb{E}[A_1] + \cdots + \mathbb{E}[A_{999,997}] \\ &= \frac{999,997}{26^4} \approx 2.19. \end{aligned}$$

(c) Define C as the number of coin flips Sam makes until (and including) his first Head, and define D as the number of coin flips he makes after rolling the die. We want to find $\mathbb{E}[C + D]$.

First, since $C \sim \text{Geometric}(0.2)$, $\mathbb{E}[C] = \frac{1}{0.2} = 5$.

Second, D has the following distribution:

$$\begin{aligned} \mathbb{P}[D = 1] &= \frac{4}{6} \\ \mathbb{P}[D = 2] &= \frac{1}{6} \\ \mathbb{P}[D = 3] &= \frac{1}{6} \end{aligned}$$

So

$$\begin{aligned} \mathbb{E}[D] &= 1 \cdot \mathbb{P}[D = 1] + 2 \cdot \mathbb{P}[D = 2] + 3 \cdot \mathbb{P}[D = 3] \\ &= 1 \cdot \frac{4}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} \\ &= 1.5 \end{aligned}$$

Using Linearity of Expectation,

$$\begin{aligned}\mathbb{E}[C+D] &= \mathbb{E}[C]+\mathbb{E}[D]\\ &= 5+1.5\\ &= \boxed{6.5}\end{aligned}$$

5 Throwing Frisbees

Note 19
Note 20

Andy and $n-1$ other students play a frisbee-passing game, where $n \geq 3$. Andy makes the first throw. On each throw, the current holder chooses uniformly at random from the other $n-1$ people, using a fresh random choice. Nobody can throw to themselves. The game stops the first time the frisbee returns to Andy.

(a) Let T be the total number of throws, including Andy's first throw and the final throw back to Andy. Identify the distribution of $T-1$ and find $\mathbb{E}[T]$.

(b) A student concludes from part (a) that everyone gets a turn on average. Find the expected number of students *other than Andy* who never receive the frisbee during the game. Also find the expected number of those students who do receive it.

Explain why the expected number of throws alone does not determine how many different people participate.

Solution:

(a) The first throw cannot end the game. After that, whenever the frisbee is held by another student, the next throw returns to Andy with probability $1/(n-1)$, regardless of the earlier throws. Thus

$$T-1 \sim \text{Geometric}\left(\frac{1}{n-1}\right), \qquad \boxed{\mathbb{E}[T] = 1+(n-1) = n.}$$

Here a geometric variable counts trials up to and including the first success.

(b) Fix one of the other students, v. Let I_v be 1 if v never receives the frisbee, and 0 otherwise.

Andy avoids v on the first throw with probability $(n-2)/(n-1)$. After such a throw, someone other than Andy or v holds the frisbee. Until either Andy or v receives it, these two people play symmetric roles: every throw has the same chance of reaching either one. Eventually one of them is reached with probability 1, and each is first with probability $1/2$.

Therefore v is never reached precisely when the first throw avoids v and Andy is reached before v afterwards. Hence

$$\mathbb{P}[I_v = 1] = \frac{n-2}{n-1}\cdot\frac{1}{2} = \frac{n-2}{2(n-1)}.$$

Let U count the other students who never receive the frisbee. Since $U = \sum_{v \neq \text{Andy}} I_v$, linearity gives

$$\boxed{\mathbb{E}[U] = (n-1)\frac{n-2}{2(n-1)} = \frac{n-2}{2}.}$$

The indicators need not be independent. The number of other students who receive the frisbee is $n-1-U$, so its expected value is

$$\boxed{n-1-\frac{n-2}{2} = \frac{n}{2}.}$$

A throw can go to someone who has already received the frisbee. Counting throws counts repeat visits, so an expected n throws does not mean all n people participate.

6 Family Planning

Note 20

Mr. and Mrs. Johnson decide to continue having children until they either have their first girl or until they have three children. Assume that each child is equally likely to be a boy or a girl, independent of all other children, and that there are no multiple births. Let G denote the number of girls that the Johnsons have. Let C be the total number of children they have.

(a) Determine the sample space, along with the probability of each sample point.

(b) Compute the joint distribution of G and C. Fill in the table below.

	$C=1$	$C=2$	$C=3$
$G=0$			
$G=1$			

(c) Use the joint distribution to compute the marginal distributions of G and C and confirm that the values are as you'd expect. Fill in the tables below.

$\mathbb{P}[G=0]$	
$\mathbb{P}[G=1]$	

$\mathbb{P}[C=1]$	$\mathbb{P}[C=2]$	$\mathbb{P}[C=3]$

(d) Are G and C independent?

(e) What is the expected number of girls the Johnsons will have? What is the expected number of children that the Johnsons will have?

Solution:

(a) The sample space is the set of all possible sequences of children that the Johnsons can have: $\Omega = \{g, bg, bbg, bbb\}$. The probabilities of these sample points are:

$$\mathbb{P}[g] = \frac{1}{2}$$
$$\mathbb{P}[bg] = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$
$$\mathbb{P}[bbg] = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$
$$\mathbb{P}[bbb] = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

(b)

	$C=1$	$C=2$	$C=3$
$G=0$	0	0	$\mathbb{P}[bbb]=1/8$
$G=1$	$\mathbb{P}[g]=1/2$	$\mathbb{P}[bg]=1/4$	$\mathbb{P}[bbg]=1/8$

(c) Marginal distribution for G:

$$\mathbb{P}[G=0] = 0+0+\frac{1}{8} = \frac{1}{8}$$
$$\mathbb{P}[G=1] = \frac{1}{2}+\frac{1}{4}+\frac{1}{8} = \frac{7}{8}$$

Marginal distribution for C:

$$\mathbb{P}[C=1] = 0 + \frac{1}{2} = \frac{1}{2}$$
$$\mathbb{P}[C=2] = 0 + \frac{1}{4} = \frac{1}{4}$$
$$\mathbb{P}[C=3] = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$$

(d) No, G and C are not independent. If two random variables are independent, then

$$\mathbb{P}[X=x, Y=y] = \mathbb{P}[X=x]\,\mathbb{P}[Y=y].$$

To show this dependence, consider an entry in the joint distribution table, such as $\mathbb{P}[G=0, C=3] = 1/8$. This is not equal to $\mathbb{P}[G=0]\,\mathbb{P}[C=3] = (1/8)\cdot(1/4) = 1/32$, so the random variables are not independent.

(e) We can apply the definition of expectation directly for this problem, since we've computed the marginal distribution for both random variables.

$$\mathbb{E}[G] = 0\cdot\mathbb{P}[G=0] + 1\cdot\mathbb{P}[G=1] = 1\cdot\frac{7}{8} = \frac{7}{8}$$
$$\mathbb{E}[C] = 1\cdot\mathbb{P}[C=1] + 2\cdot\mathbb{P}[C=2] + 3\cdot\mathbb{P}[C=3] = 1\cdot\frac{1}{2} + 2\cdot\frac{1}{4} + 3\cdot\frac{1}{4} = \frac{7}{4}$$

7 Pairwise and Mutual Independence

Note 20

Let X and Y be independent Bernoulli random variables with parameter p, where $0 < p < 1$. Define

$$Z = \begin{cases} 1, & X \neq Y, \\ 0, & X = Y. \end{cases}$$

This operation is also called exclusive OR (XOR).

(a) Find the joint distribution of X and Z, and the marginal distribution of Z.

(b) For which values of p are X and Z independent? What about Y and Z? Justify your answers.

(c) Set $p = 1/2$. Every pair among X, Y, Z is independent. Are all three mutually independent? Prove your answer.

How can every pair be independent even though knowing any two variables determines the third?

Solution:

(a) The four possible pairs (X, Y) determine (X, Z), giving

	$Z=0$	$Z=1$	$\mathbb{P}[X=x]$
$X=0$	$(1-p)^2$	$p(1-p)$	$1-p$
$X=1$	p^2	$p(1-p)$	p
$\mathbb{P}[Z=z]$	$(1-p)^2+p^2$	$2p(1-p)$	1

Thus $\boxed{Z \sim \text{Bernoulli}(2p(1-p))}$.

(b) Given $X = 0$, we have $Z = Y$, whereas given $X = 1$, we have $Z = 1 - Y$. Since X and Y are independent,

$$\mathbb{P}[Z = 1 \mid X = 0] = p, \qquad \mathbb{P}[Z = 1 \mid X = 1] = 1 - p.$$

Both conditioning events have positive probability. Independence requires these conditional probabilities to be equal, so $p = 1 - p$, or $p = 1/2$.

Conversely, take $p = 1/2$. Each joint probability is $1/4$, while each marginal probability is $1/2$. Thus the joint table factors into the marginals, proving independence. Interchanging X and Y gives the same conclusion for Y and Z:

$$\boxed{X \text{ and } Z \text{ are independent iff } p = 1/2,}$$

$$\boxed{Y \text{ and } Z \text{ are independent iff } p = 1/2.}$$

(c) They are *not mutually independent*. If $X = Y = 1$, then $Z = 0$, so

$$\mathbb{P}[X = 1, Y = 1, Z = 1] = 0 \neq \frac{1}{8} = \mathbb{P}[X = 1]\,\mathbb{P}[Y = 1]\,\mathbb{P}[Z = 1].$$

Mutual independence would require this factorization for every triple of values.

There is no contradiction. Pairwise independence says that knowing *one* variable does not change the distribution of another. It does not say that knowing *two together* supplies no information. Here, knowing X, Y gives Z; knowing X, Z gives Y; and knowing Y, Z gives X.