CMPSCI 240: Reasoning about Uncertainty

Lecture 8: Random Variables

Andrew McGregor

University of Massachusetts

Last Compiled: February 21, 2017

Outline

Experiments, Sample Spaces, Events

- Experiment: a process that results in exactly one of several possible outcomes, e.g., rolling a dice
- Sample space: the set of all possible outcomes of an experiment, e.g., $\Omega = \{1, 2, 3, 4, 5, 6\}$
- Event: a subset of Ω, e.g., $A =$ "odd number" $= \{1, 3, 5\}$
- Atomic event: event consisting of a single outcome, e.g., $\{1\}$
- Probability law: A function $P(\cdot)$ that maps event to a number between 0 and 1 that satisfies the probability axioms:
  1. Nonnegativity: $P(A) \geq 0$ for every $A \subseteq \Omega$
  2. Normalization: $P(\Omega) = 1$
  3. Additivity: $P(A \cup B) = P(A) + P(B)$ if A and B are disjoint
- The probability of any event $A = \{o_1, \ldots, o_N\}$ can be obtained from the probabilities of the $|\Omega|$ atomic events using the result that

$$P(A) = P(\{o_1\}) + \ldots + P(\{o_N\}) \; .$$

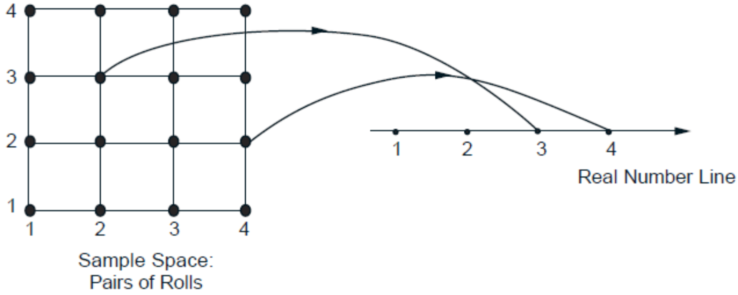
Outline

Random Variable

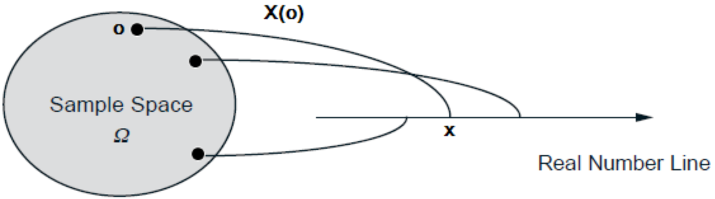
- A *random variable* is a function that maps from the sample space to the real numbers,
$$X : \Omega \to \mathbb{R}$$

Example: Maximum of Dice Rolls

- Example: Consider rolling two fair four sided dice.
- The outcomes $r \in \Omega$ are pairs $r = (r_1, r_2)$ where r_1, r_2 both take values on the set $\{1, ..., 4\}$
- We could consider the function $X(r_1, r_2) = \max(r_1, r_2)$



More Generally. . .



- For every outcome $o \in \Omega$, a random variable defines a single real number
$$X(o) \in \mathbb{R} \; .$$

- Note that it is possible for there to be multiple outcomes $o_1, o_2, \ldots$ such that $X(o_1) = X(o_2) = \ldots$.

Random Variables Give An Easy Way to Specify Events

- If we have a function $X : \Omega \to \mathbb{R}$, we can use it to construct a different event for each value of $k \in \mathbb{R}$:

$$\{X = k\} = \{o | o \in \Omega \text{ and } X(o) = k\}$$

- In the dice example, the event $\{X = k\}$ is the set of outcomes $o \in \Omega$ that are mapped to the the same value k by the function X. E.g.,

$$\{X = 2\} = \{(1, 2), (2, 1), (2, 2)\}$$

$$\{X = 3\} = \{(1, 3), (2, 3), (3, 3), (3, 1), (3, 2)\}$$

$$\{X = 1\} = \{(1, 1)\}$$

Outline

Random Variables and Probability

- We can compute the probability of an event $\{X = k\}$ for $k \in \mathbb{R}$ by decomposing it into atomic events and using the probability rule:

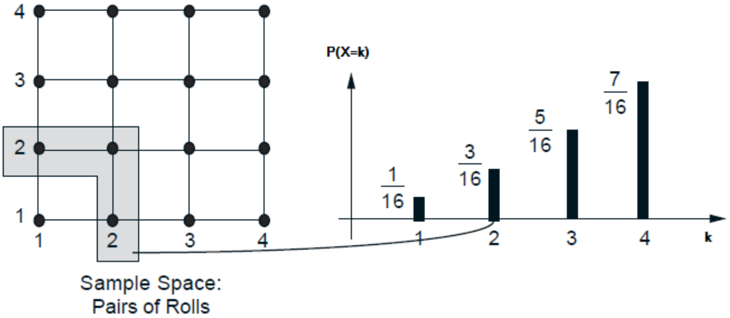
$$P(X = k) = P(\{o | o \in \Omega \text{ and } X(o) = k\})$$

- For example, in the event $\{X = 2\}$ in the case of rolling two dice and $X(r_1, r_2) = \max(r_1, r_2)$

$$\begin{aligned} P(X = 2) &= P(\{(1,2),(2,1),(2,2)\}) \\ &= P((1,2)) + P((2,1)) + P((2,2)) = 3/16 \end{aligned}$$

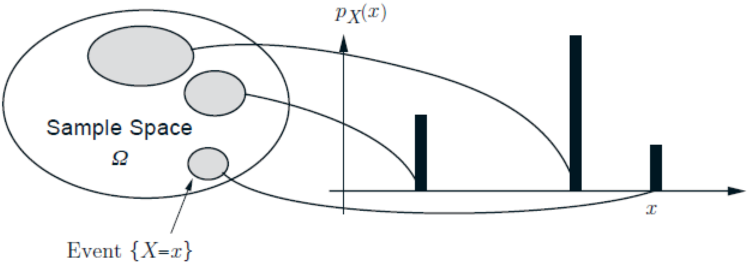
Example: Maximum of Dice Rolls

- We can work out the probability for each value of k from 1 to 4:



In general. . .

- The probability associated with the event $\{X = k\}$ for each element $k \in \mathbb{R}$ of a discrete random variable X is referred to as the **probability mass function** or **PMF** of the random variable.
- The probability mass function is denoted by $P(X = k)$ or $p_X(k)$



Outline

Binomial Random Variable

- Suppose we toss independent n coins where each coin has probability p of being heads
- The set of outcomes is:

$$\Omega = \{(TTT \dots TT), (TTT \dots TH), \dots, (HHH \dots HH)\}$$

- Define a random variable X where for each $o \in \Omega$,

$$X(o) = \text{ “the number of heads in outcome } o\text{”}$$

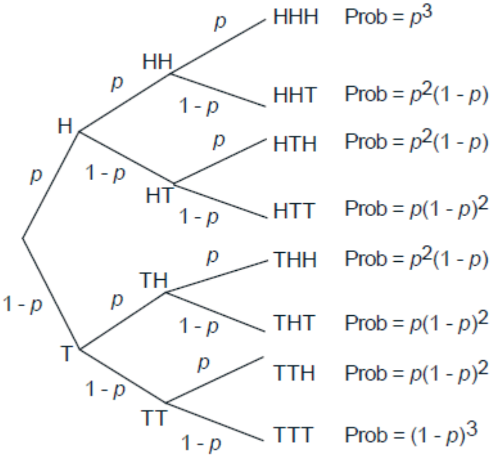
- We'll show that $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.

Binomial Random Variables: Examples

- The number of heads in N coin tosses.
- The number of severs that fail in a cluster of N servers.
- The number of games a soccer team wins in a season of N games.
- The number of multi-choice questions you get correct if you guess each of N questions.

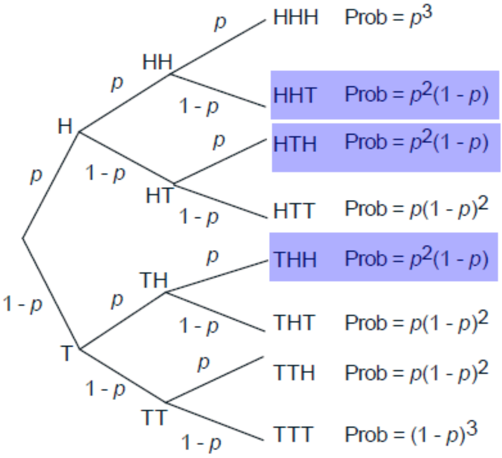
Example: Independent Coin Flips

- If we flip the coin 3 times, what is the probability that the **number** of heads in the outcome is 2?
- It's helpful to use a conditional probability tree.



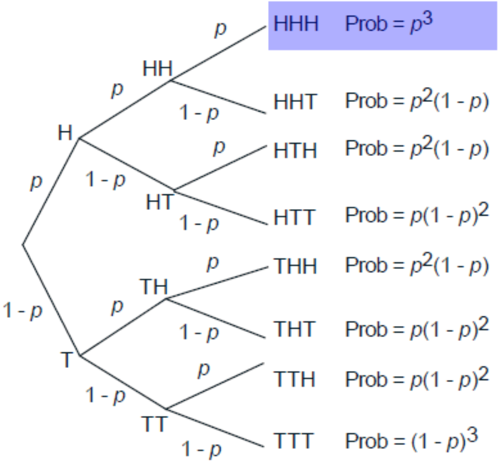
Example: Independent Coin Flips

- Probability of two heads? $3 \cdot p^2(1 - p)$



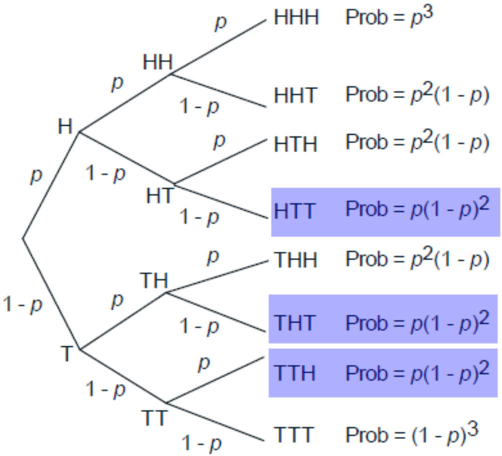
Example: Independent Coin Flips

- Probability of three heads? p^3



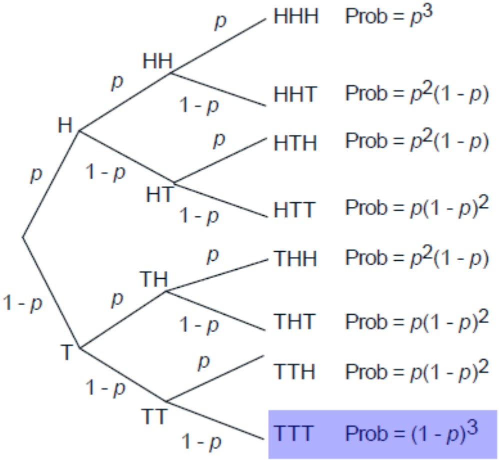
Example: Independent Coin Flips

- Probability of one head? $3 \cdot p(1-p)^2$



Example: Independent Coin Flips

- Probability of no heads? $(1-p)^3$



Generalizing to n Flips and k Heads

- If we toss n coins, what's the probability of seeing k heads?
- Any single sequence of length n with k heads has probability

$$p^k(1-p)^{n-k} \ .$$

- But how many different sequences of length n contain k heads?

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

where $\binom{n}{0} = 1$. This follows because any subset of k of the coin tosses could result in the k heads that are observed.

Counting Sequences

Sequences	Subset of Positions where H Occurs
HHHH	{1,2,3,4}
HHHT	{1,2,3}
HHTH	{1,2,4}
HHTT	{1,2}
HTHH	{1,3,4}
HTHT	{1,3}
HTTH	{1,4}
HTTT	{1}
THHH	{2,3,4}
THHT	{2,3}
THTH	{2,4}
THTT	{2}
TTHH	{3,4}
TTHT	{3}
TTTH	{4}
TTTT	{}

Note that there are $\binom{4}{k}$ sequences where the number of H's is k.

The Binomial Law

- The probability of observing k heads in n independent trials where the probability of success is p in each trial is thus:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

The Binomial Law: $n = 3$

- Consider the special case where $n = 3$:

$$\binom{3}{3} = \frac{3!}{3!0!} = 1\ ,\ \binom{3}{2} = \frac{3!}{2!1!} = 3\ ,\ \binom{3}{1} = \frac{3!}{1!2!} = 3\ ,\ \binom{3}{0} = \frac{3!}{0!3!} = 1$$

- This gives us back the probabilities we worked out using the tree:

$$P_3(3) = \binom{3}{3} p^3 (1-p)^0 = p^3$$

$$P_3(2) = \binom{3}{2} p^2 (1-p) = 3p^2(1-p)$$

$$P_3(1) = \binom{3}{1} p(1-p)^2 = 3p(1-p)^2$$

$$P_3(0) = \binom{3}{0} p^0 (1-p)^3 = (1-p)^3$$

Outline

Geometric Random Variables

- Suppose we flip a biased coin repeatedly until it lands heads. Let X be the flips of the coin up to and including the first head.
- The PMF of a *geometric random* variable X is

$$P(X = k) = (1 - p)^{k-1} \cdot p \quad \text{for } k = 1, 2, 3, \ldots$$

- Used to model the number of repeated independent trials up to (and including) the first "successful" trial, e.g., the number of patients we test before the first one we find who has a given disease.

Bernoulli Random Variables

- Suppose we have an experiment with two outcomes H and T. H happens with probability $(1-p)$ and T with probability p.
- We define a random variable X such that $X(H)=0$ and $X(T)=1$.
- This is called a Bernoulli random variable X:

$$P(X=k)=\begin{cases} 1-p & \text{if } k=0 \\ p & \text{if } k=1 \end{cases}$$

Bernoulli Random Variables: Examples

- Whether a coin lands heads or tails.
- Whether a server is online or offline.
- Whether an email is spam or not.
- Whether a pixel in a black and white image is black or white.
- Whether a patient has a disease or not.

Discrete Uniform Random Variables

- A discrete uniform random variable X with range $[a, b]$ takes on any integer value between a and b inclusive
- The PMF of a discrete uniform random variable X is

$$P(X = k) = \frac{1}{b - a + 1} \text{ for } k = a, \ldots, b$$

- Used to model probabilistic situations where each of the values $a, \ldots, b$ are equally likely. E.g., the random variable that maps a six-sided dice roll to the number that comes up is a uniform random variable with $a = 1$, $b = 6$ and $P(X = k) = 1/6$ for $k = 1, \ldots, 6$.

Poisson Random Variables

- Processes that involve counting up many different independent events that occur within a given time interval can often be modeled as Poisson random variables.
- The PMF of a Poisson random variable X is

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad \text{for } k = 0, 1, 2, \ldots$$

- e.g., the number of typos in a book, number of cars that crash in a city on a given day, the number of phone calls arriving at a call center per minute etc.

Discrete Random Variables

- **Uniform:** For $k = a, \ldots, b$:

$$P(X = k) = \frac{1}{b - a + 1}$$

- **Bernoulli:** For $k = 0$ or 1:

$$P(X = k) = \begin{cases} 1 - p & \text{if } k = 0 \\ p & \text{if } k = 1 \end{cases}$$

- **Binomial:** For $k = 0, \ldots, N$

$$P(X = k) = \binom{N}{k} p^k (1 - p)^{N-k}$$

- **Geometric:** For $k = 1, 2, 3, \ldots$, $P(X = k) = (1 - p)^{k-1} \cdot p$
- **Poisson:** $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ for $k = 0, 1, 2, \ldots$