Combinatorics: The Fine Art of Counting

Week 7 Lecture Notes – Discrete Probability Continued

Note – Binomial coefficients are written horizontally. The symbol ~ is used to mean approximately equal.

The Bernoulli Process

All of the probability measures we have looked at so far have used the uniform probability measure on a finite sample space. Today we will look at an example of a random process that may involve a non-uniform probability measure as well as a potentially infinite sample space.

A **Bernoulli process** consists of a sequence of repeated experiments, called trials. The canonical example is repeated coin tosses, but there are many other random processes that can be treated as a Bernoulli process. The three key requirements of any Bernoulli process are:

1. Each trial has exactly two possible outcomes. These are typically designated **success** or **failure**.

2. The trials are all identical – the probability of success is always **p**, where $0 < p < 1$, and the probability of failure is **q = 1 - p**.

3. The trials are all mutually independent. This means that the probability of any finite intersection of events relating to the different individual trials is simply the product of the probabilities of the events.

Some examples of Bernoulli trials include:

- a series of games – e.g. a best-of-7 world-series, a football season, a gambling session.

- a 1-dimensional random walk – e.g. stock market price changes

- a particular outcome – e.g. rolling a six on a die, shooting a hole-in-one, a baseball player getting a hit

- a medical trial conducted on a group of people

- a repeatable science experiment

- a consumer's response to an advertisement or survey question

As long as each trial can be categorized as a success or failure, the Bernoulli process may be applicable. It is critical, however, that the trials satisfy the requirements above. Note that when modeling a real world event that is not strictly random (e.g, a human response) the question of independence is always a tricky one, but in many cases it is a reasonable assumption.

A Bernoulli process may be limited to a particular number of trials (e.g. 7 games, or 10 coin-tosses), or it may go on indefinitely, in which case we may regard it as an infinite process. A finite Bernoulli sample space consists of all binary sequences of some particular length n (1 denotes success, 0 failure). In the infinite case, the sample space consists of all infinite binary sequences.

The probability measure on this sample space is simply defined in terms of the probability of the elementary events H_i which represent the set of sample points where the ith trial was successful.

$$P(H_i) = p, \text{ and } P(H_i^c) = 1\text{-}p = q.$$

The probability of a finite intersection of any of these elementary events (with distinct i's) is simply the product of the probabilities.

Unless p = ½, this probability measure is **not uniform**. To take a concrete example, if we consider a sequence of 3 tosses of a coin which comes up heads 2/3 of the time, the probability of the event containing the sequence HHT is $(2/3)^2*1/3 = 8/27$, while the probability of the event containing the sequence TTH is $(1/3)^3*2/3 = 2/27$ and is thus 4 times less likely.

For an infinite Bernoulli process, it worth noting that the individual sample points are not defined as events – the finite intersection of any of the elementary events as defined above is still going to contain many infinite sequences (uncountably many in fact). If an event were defined which contained only one sample point (or even a countably infinite number of sample points), it would necessarily have probability 0. Note that this does not mean that such an event is cannot exist or that it is the empty set, it simply means that the probability measure assigns the number 0 to it – this is the 1-d analog of the fact that a line segment has zero area – it's still a perfectly good line segment, it just has no area and doesn't "cover" a measurable portion of the plane. Many subsets of an infinite sample space may not cover enough of the space to have non-zero probability, and we are not really interested in considering such subsets as events.

If the notion of an infinite process seems unsettling, simply regard it as having an indefinite length. There are many games that could in theory go on indefinitely (e.g. baseball, or the dice game "craps"), but in practice will always end in a reasonable short number of steps. It is convenient in such cases to not have to set an arbitrary limit to the number of trials, but rather to regard the process as going on forever – this ensures that our sample space can handle all possible outcomes.

For a finite Bernoulli process, we are often interested in the probability of getting a specific number of successful trials. This probability is easy to compute. Let S_k be the event that exactly k successes occur in a Bernoulli process with n trials and success probability p. The probability of the event S_k is given by:

$$P(S_k) = (n\ k)*p^k*q^{n-k} \qquad \text{k successes in n trials}$$

In general, the probability of any event which is determined simply by the number of successes can be computed using binomial coefficients and appropriate powers of p and q.

Problem #1: What is the probability of getting exactly 5 heads in 10 tosses of a fair coin? How about 50 heads in 100 tosses?

Solution #1: $(10\ 5)*(1/2)^5*(1/2)^5 = 252/1024 =$ **63/256 < 1/4**. $(100\ 50)*(1/2)^{50}*(1/2)^{50} \sim$ **0.08**

Problem #2: If teams A and B play in a best of 7 series, what are the probabilities that the series ends in 4, 5, 6, or 7 games if team A wins each game with independent probability p? Given an explicit answer for the case where the teams are evenly matched.

Solution #2: We will assume the outcomes of the games are independent and treat this as a Bernoulli process with 7 trials and p = 1/2. The probability of the series ending after 4 games is simply the probability of getting either 4 heads or 4 tails in a row which is $(4\ 4) * (p^4 + q^4) =$ **2/16** when p=q. For a five game series, we are interested in sequences where exactly 3 of the first 4 trials are heads (or tails) and the 5th trial is a head (or a tail). This is $(4\ 3) * (p^3q*p + q^3p*q) =$ **4/16** when p=q. (Note the difference in the two terms when p≠q). The computations for six and seven games series are similar. For six games we get $(5\ 3) * (p^3q^2*p + q^3p^2*q) =$ **5/16** when p=q and for seven games we get $(6\ 3) * (p^3q^3*p + q^3p^3*q) =$ **5/16** when p=q.

Problem #3: If teams A and B play in a best of 3 series and if team A wins each game with independent probability p > ½, what is the probability that team A wins the series in terms of p? What about a best of 5 series? Compute actual values for p = 0.75.

Solution #3: This problem can be solved by partitioning based A winning in either 2 or 3 games, but a simpler approach (especially for longer series) is to simply consider all sequences of three tosses of a coin with bias p that have more heads than tails (if the series really only lasts 2 games, we can imagine them playing a third consolation game, it won't change the answer). This

2

approach gives $(3\ 2)*p^2*(1-p) + (3\ 3)*p^3 = \mathbf{3p^2\text{-}2p^3}$. Note that this value is equal to ½ for p=1/2 but is always greater than p when p > ½. Applying the same approach for a best of 5 series gives: $(5\ 3)*p^3*(1-p)^2 + (5\ 4)*p^4*(1-p) + (5\ 5)*p^5 = \mathbf{10p^3\text{-}15p^4\text{+}6p^5}$. Plugging in p = 0.75 gives probability of ~**0.84** for team A to win a best of 3 series and **~0.90** for a best of 7 series.

Infinite Bernoulli Processes

We now want to take a look at some problems involving infinite Bernoulli processes. Consider a game where two players take turns flipping a fair coin. If a player gets heads on their turn the win and the game is over, otherwise they give the coin to the other player who then takes his or her turn. What is the probablility that the player who goes first wins?

First note that while this game will end on the first turn half the time, it could take an arbitrarily long sequence of turns before someone gets a head. Note also that the player who goes first clearly has an advantage - half the time the second player never even gets to flip the coin, and when they do there is no guarantee they will win before the first player has another flip. To model this game we will use the infinite Bernoulli sample space with p = ½. Let's assume the players are named Alice and Bob (A and B) and that Alice goes first. Let A_k be the event that Alice wins on her k^{th} turn:

$P(A_1)$ = ½   A gets heads on the first flip

$P(A_2) = [(½)*(½)]*(½)$ A gets tails, B gets tails, then A gets heads

$P(A_2) = [(½)*(½)]^2*(½)$ A and B both get tails twice, then A gets heads

$P(A_k) = [(½)*(½)]^{k-1}*(½)$ A and B get tails k times, then A gets heads

These events are all disjoint, so the probability that A wins is just:

$P(A) = P(A_1) + P(A_2) + P(A_3) + P(A_4) + \ldots$

(Note that this infinite sum is bounded (it is a probability so it is at most 1) and it converges – we won't worry about the details behind this or any other issues of convergence necessary to justify the steps below - they are all straight-forward).

$P(A) = ½ + (½)*(¼) + (½)*(¼)^2 + (½)*(¼)^3 + \ldots$

$P(A) = ½ * (1 + ¼ + (¼)^2 + (¼)^3 + \ldots)$

The infinite sum in brackets is known as a geometric series and there is a simple way to solve such infinite sums which we will review briefly here for the benefit of those who may not have seen it before.

Recall the following facts from algebra for factoring the difference of two squares or two cubes and it's generalization to higher powers:

$1^2 – x^2 = (1 - x) * (1 + x)$ $\Rightarrow$ $1 + x$ $= (1^2 – x^2) / (1 - x)$

$1^3 – x^3 = (1 - x) * (1 + x + x^2)$ $\Rightarrow$ $1 + x + x^2$ $= (1^3 – x^3) / (1 - x)$

$1^4 – x^4 = (1 - x) * (1 + x + x^2 + x^3)$ $\Rightarrow$ $1 + x + x^2 + x^3$ $= (1^3 – x^3) / (1 - x)$

$1^n – x^n = (1 - x) * (1 + x + x^2 + \ldots + x^{n-1})$ $\Rightarrow$ $1 + x + x^2 + \ldots + x^{n-1} = (1^n – x^n) / (1 - x)$

These equations hold for any x not equal to 1. Now when x happens to be a probability less than 1 (or any number with absolute value less than 1), as n gets large, x^n gets very small and rapidly becomes indistinguishable from 0. Pick any number between 0 and 1, punch it in to your calculator and then press the x^2 button over and over. It won't take long before your calculator display is nothing but 0's. Thus to compute the infinite sum $1 + x + x^2 + \ldots$, provided |x| < 1 we can assume that x^n is zero in the equation on the right to obtain a very simple solution:

$\mathbf{1 + x + x^2 + x^{3} + \ldots = 1 / (1 – x)}$ **provided that |x| < 1**

We can now use this handy equation to solve our problem by plugging in ¼ for x. which gives us $P(A) = ½ * (1 / (1- ¼)) = ½ * 4/3 = $ **2/3**. Thus Alice will win 2/3 of the time. The analysis above can easily be generalized to a non-fair coin – simply replace ½ by p and ¼ by $(1-p)^2$.

The last problem we will look at is one that is often considered a difficult problem, but can be solved quite easily using the approach above.

Problem #5: A fair coin is flipped repeatedly. What is the probability of getting 3 heads in a row before 5 tails in a row.

Non-Solution #5: At first glance it might seem that since the probability of getting 3 heads in a row is 1/8 and the probability of getting 5 tails in a row is 1/32, that getting 3 heads in a row first should occur 4 times as often. This is not quite right however – consider for example that there are 8 ways to get a run of 3 heads in the first 5 flips but only 1 way to get 5 tails. Trying to compute the exact probability of runs of various lengths quickly becomes complex, especially since we don't know exactly how many flips may be required to produce a run of 3 heads or 5 tails. Another approach is to try to count sequences of binary strings of a given length that end in a run of 3 heads or 5 tails but don't contain any other runs of either type. This can be done, but it requires finding and solving the right recurrence relation. Fortunately there is an easier way.

Solution #5: Let's look at the situation as a game between two players Alice and Bob, where Alice is trying to get a run of 3 heads and Bob is trying for a run of 5 tails. They use the first coin toss to decide who goes first, and then that player continues flipping until either they win, or the get the wrong toss in which case they hand the coin to the other person who takes their turn. For example if the first toss is heads, Alice then flips the coin hoping to get 2 more heads in a row, but if she flips a tail before then, Bob takes over and tries to get 4 more tails (the first tail occurred when Alice's turn ended). Note that after the initial flip, every turn looks the same – Alice tries to get 2 heads in a row, and Bob tries to get 4 tails in a row. Alice always has a $1/4^{th}$ chance of winning on her turn, and Bob always has a $1/16^{th}$ chance.

We can now compute the probabilities using an analysis similar to the one above, except there are two cases depending on the outcome of the first flip. As above, let Let A_k be the event that Alice wins on her k^{th} turn, and let H be the event that the first flip is heads and let T be the complementary event.

$P(A_1|H) = ¼$	A wins on her first turn
$P(A_2|H) = (3/4)*(15/16)*(1/4)$	A doesn't win, B doesn't win, A wins
$P(A_3|H) = [(3/4)*(15/16)]^2*(1/4)$	A and B don't win twice, A wins
$P(A_k|H) = [(3/4)*(15/16)]^k*(1/4)$	A and B don't win twice, A wins

As above, these events are disjoint and we have:

$P(A|H) = P(A_1|H) + P(A_2|H) + P(A_3|H) + …$

$P(A|H) = ¼ * (1 + 45/64 + (45/64)^2 + (45/64)^3 + …)$

$P(A|H) = ¼ * [1 / (1 – 45/64)] = ¼ * 64/19 = 16/19$

If Bob goes first, the situation is identical except that Bob needs to lose on his first turn:

$P(A_1|T) = (15/16)*¼$	B doesn't win then A wins on her first turn
$P(A_k|T) = (15/16)[(3/4)*(15/16)]^k*(1/4)$	B doesn't win, A and B don't win, then A wins

Every equation simply has the factor 15/16 in it, so we have $P(A|T) = 15/16 * P(A|H) = 15/19$.

Putting it all together using the Law of Alternatives we have:

$P(A) = P(H)P(A|H) + P(T)P(A|T) = (½) * (16/19) + (½) * (15/19) = $ **31/38**

Note that this is slightly higher than we would get if we used our initial erroneous assumption that Alice wins 4 times as often as Bob which would have given 4/5.

MIT OpenCourseWare
http://ocw.mit.edu

Combinatorics: The Fine Art of Counting

Summer 2007

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.