Math 425
Introduction to Probability
Lecture 12

Kenneth Harris
kaharri@umich.edu

Department of Mathematics
University of Michigan

February 4, 2009

Examples of infinite sample spaces

☞ We have been considering finite sample spaces, but some experiments are best modeled by infinite sample spaces.

- Experiment A: Toss a fair coin until the first head appears, then stop.
- Experiment C: Toss a fair coin until a run of ten heads in a row appears, then stop.
- Experiment B: Throw a pair of dice until either a 5 or 7 appears, then stop.

☞ No finite sample space is adequate in these cases, since each experiment has (potentially) infinitely many outcomes.

Example of coin tosses

Example A. Toss a fair coin until the first head appears, then stop.
The sample space for this experiment is

$$S = \{H,\ TH,\ ,TTH\ ,TTTH\ ,\dots, T^\infty = TTT\cdots\}$$

- For each k, a sequence of k tails followed by a single head: T^kH.
- an infinite sequence consisting only tails: T^∞.

☞ We can compute the probability of "most" of these outcomes:

$$\begin{aligned} \mathbf{P}(T^kH) &= 2^{-k-1} \\ \mathbf{P}(T^\infty) &= ??? \end{aligned}$$

Axiom 3 – Strong form

☞ We recall the strong form of the Addition Rule (Axiom 3):

Axiom (3 – Addition rule (strong form))

For any sequence of mutually exclusive events $E_1, E_2, \dots$ (so, $E_i \cap E_j = \emptyset$ whenever $i \neq j$),

$$\boldsymbol{P}\Big(\bigcup_{k=1}^{\infty} E_k\Big) = \sum_{k=1}^{\infty} \boldsymbol{P}(E_k).$$

Example of Coin Tosses

$$S = \{H,\ TH,\ ,TTH\ ,TTTH\ ,\dots, T^\infty = TTT\cdots\}$$

☞ Let H be the event that a heads is tossed. Then

$$H^c = \{T^\infty\} \qquad \mathbf{P}\left(T^\infty\right) = \mathbf{P}\left(H^c\right) = 1 - \mathbf{P}\left(H\right)$$

☞ Let $H_k = \{T^k H\}$ $(k \geq 0)$: the event that heads follows k tails.
So, the events $H_0, H_1, H_2, \dots,$ are mutually exclusive and

$$H = \bigcup_{k=0}^{\infty} H_k$$

Example – continued

☞ Then,

$$\mathbf{P}\left(H_k\right) = \mathbf{P}\left(T^k H\right) = 2^{-k-1}.$$

Use the strong form of the Addition Rule:

$$\begin{aligned} \mathbf{P}\left(\bigcup_{k=0}^{\infty} H_k\right) &= \sum_{k=0}^{\infty} \mathbf{P}\left(H_k\right) \\ &= \sum_{k=0}^{\infty} 2^{-k-1} = 1. \end{aligned}$$

We compute the infinite series shortly.

☞ So,

$$\mathbf{P}\left(H\right) = \mathbf{P}\left(\bigcup_{k=0}^{\infty} H_k\right) = 1$$

and thus,

$$\mathbf{P}\left(H^c\right) = \mathbf{P}\left(T^\infty\right) = 0.$$

Infinite series

Definition

Let $a_1, a_2, \dots$ be an infinite sequence of real numbers.
If the partial sums

$$s_n = \sum_{k=1}^{n} a_k.$$

have a finite limit:

$$\lim_{n\to\infty} s_n = s,$$

then we say the infinite series $\sum_{k=1}^{\infty} a_k$ converges with sum s.
Otherwise, it diverges.

If the series $\sum_{k=1}^{\infty} |a_k|$ converges as well,
then $\sum_{k=1}^{\infty} a_k$ converges absolutely.

Geometric series

Theorem

The geometric series converges absolutely:

$$\sum_{k=0}^{\infty} ax^k = \frac{a}{1-x} \qquad \text{for any real numbers } a, \text{ and } |x| < 1.$$

Example. From the previous section:

$$\begin{aligned} \sum_{k=0}^{\infty} 2^{-k-1} &= \sum_{k=0}^{\infty} \frac{1}{2} \cdot \left(\frac{1}{2}\right)^k \\ &= \frac{1}{2} \cdot \frac{1}{1-\frac{1}{2}} = 1. \end{aligned}$$

$a = \frac{1}{2}$ and $x = \frac{1}{2}$ from the Theorem.

Proof

Proof . The nth partial sum is

$$s_n = \sum_{k=0}^{n} ax^k.$$

Multiply by x (shifting powers by 1):

$$x \cdot s_n = \sum_{k=1}^{n+1} ax^k.$$

Subtract the previous two results (most terms cancel)

$$s_n - x \cdot s_n = a - ax^{k+1}.$$

Solve for s_n,

$$s_n = \frac{a(1 - x^{n+1})}{1 - x}$$

Proof

☞ By definition

$$\sum_{k=0}^{\infty} ax^k = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \frac{a(1 - x^{n+1})}{1 - x}.$$

Since $|x| < 1$,

$$\lim_{k\to\infty} x^k = 0,$$

so

$$\sum_{k=0}^{\infty} ax^k = \lim_{n\to\infty} \frac{a(1 - x^{n+1})}{1 - x} = \frac{a}{1 - x}.$$

The monkey and the bard

Example

A monkey is sitting at a typewriter and randomly hits keys in a never-ending sequence. What is the probability the monkey eventually types the complete works of Shakespeare?

☞ A real monkey will eventually tire of pecking at a typewriter and go searching for a banana.

☞ However, this problem is really about a probability model, not a real monkey.

Bernoulli Trials

Definition

By a sequence of Bernoulli trials, we mean a sequence of trials (repetitions of an experiment) satisfying the following

1. Only two possible mutually exclusive outcomes on each trial.
   One arbitrarily called success and the other failure.
2. The probability of success on each trial is the same for each trial.
3. The trials are independent.

Examples of Bernoulli Trials

Examples. The following are examples of Bernoulli trials:

- Flip a coin (heads, tails),
- Each computer chip in a production line tested (chip passes test, fails test),
- Rolling a pair of dice for “snake-eyes" (double ones, any other value),
- A patient is prescribed a drug treatment (cured, not cured).
- A monkey types the complete works of Shakespeare (success, failure).

Sample space

☞ A typical experiment involving a sequence of Bernoulli trials is as follows.

Example. Suppose you are willing to wait indefinitely for the first success in a sequence of Bernoulli trials. What is the sample space of such an experiment?

☞ Let s be success, f be failure. Then the sample space S consists of

$$(s)$$
$$(f, f, \ldots, f, s) \qquad \text{only success is last trial}$$
$$(f, f, f \ldots) \qquad \text{infinite sequence of failures .}$$

Waiting for success

Problem. Suppose the probability of success is p and failure is $1 - p$ for a sequence of Bernoulli trials.
What is the probability of each outcome?

Solution. Let E_k be the event that the first success is on the $k + 1$st trial. The finite outcomes, success on the $k + 1$st trial, are

$$E_k = \{(f, f, \ldots, f, s)\} \qquad k \text{ failures } f\text{, first success } s$$

Since the trials are independent

$$\mathbf{P}(E_k) = p(1-p)^k.$$

Waiting for success

☞ Let E be the event that there is eventually some success. So,

$$E = \bigcup_{k=0}^{\infty} E_k.$$

The event of no success is $E^c = \{(f, f, f, \ldots)\}$.

☞ Since the events $E_0, E_1, \ldots$ are mutually exclusive,

$$\begin{aligned} \mathbf{P}(E) &= \mathbf{P}\left(\bigcup_{k=0}^{\infty} E_k\right) = \sum_{k=0}^{\infty} \mathbf{P}(E_k) \\ &= \sum_{k=0}^{\infty} p(1-p)^k \\ &= \frac{p}{1-(1-p)} = 1 \end{aligned}$$

So,

$$\mathbf{P}(E^c) = 1 - \mathbf{P}(E) = 0.$$

Back to the Monkey

☞ The complete works of Shakespeare consists of L symbols (spaces, letters, punctuation, etc.). A monkey beating on the keys at random has exceeding small, but nonzero probability p of typing a sequence of L symbols which is exactly the complete works of Shakespeare.

☞ Let each trial consist of L symbols banged-out by the Monkey. Success in a trial occurs when the monkey has typed the complete works of Shakespeare in the trial. The probability of eventually succeeding is one, since the only other outcome, the infinite sequence of failures, $(f, f, f, \dots)$, has probability zero.

Craps

Example

Craps is played with a pair of dice. The player (or “shooter") rolls once:

- If 7 or 11 show, she wins,
- if 2, 3 or 12 shows she loses,
- and if any other number shows (the “gambler's point"), she must keep rolling the dice until she gets a 7 before her point appears (she loses), or her point appears before the 7 (she wins).

What is the probability of winning at craps?

Sample space

☞ The sample space S consists of the following sequences:

$$\begin{array}{ll} (7), (11), (2), (3), (12) & \\ (d, a_2, a_3, \dots, a_n, d) & \text{where } d \neq 7 \text{ and } a_i \neq d, 7 \\ (d, a_2, a_3, \dots, a_n, 7) & \text{where } d \neq 7 \text{ and } a_i \neq d, 7 \\ (d, a_2, a_3, \dots) & \text{where } d \neq 7 \text{ and } a_i \neq d, 7 \end{array}$$

☞ Let E_d be the event that the first throw is d and the gambler eventually wins. Let $E_{d,n}$ be the event that the first roll is d and the gambler wins on the nth throw. (So, $d = 4, 5, 6, 8, 9, 10$).

Computing the odds

☞ Let $d = 4$. The probability that the gambler rolls 4 and wins on the nth throw, the outcome $(4, a_2, a_3, \dots, a_{n-1}, 4)$:

$$\mathbf{P}(E_{4,n}) = (\frac{3}{36})^2 \cdot (\frac{27}{36})^{n-2}$$

☞ The probability that the gambler wins some time (when she threw a 4 on the first toss):

$$\mathbf{P}(E_4) = \mathbf{P}(\bigcup_{n=2}^{\infty} E_{4,n}) = (\frac{3}{36})^2 \cdot \sum_{k=0}^{\infty} (\frac{27}{36})^k.$$

This is a geometric series,

$$\begin{aligned} \mathbf{P}(E_4) &= (\frac{3}{36})^2 \cdot \sum_{k=0}^{\infty} (\frac{27}{36})^k \\ &= (\frac{3}{36})^2 \cdot \frac{1}{1 - \frac{2}{3}} = \frac{1}{38}. \end{aligned}$$

Computing the odds

☞ The probability of rolling a 4 or 10 are $\frac{3}{36}$, so

$$\mathbf{P}(E_4) = \mathbf{P}(E_{10}) = (\frac{3}{36})^2 \cdot \sum_{k=0}^{\infty} (\frac{27}{36})^k = \frac{1}{38}$$

☞ The probability of rolling a 5 or 9 are $\frac{4}{36}$, so

$$\mathbf{P}(E_5) = \mathbf{P}(E_9) = (\frac{4}{36})^2 \cdot \sum_{k=0}^{\infty} (\frac{26}{36})^k = \frac{2}{45}$$

☞ The probability of rolling a 6 or 8 are $\frac{5}{36}$, so

$$\mathbf{P}(E_6) = \mathbf{P}(E_8) = (\frac{5}{36})^2 \cdot \sum_{k=0}^{\infty} (\frac{25}{36})^k = \frac{25}{396}$$

Computing the odds

☞ The probability of winning at black jack is

$$\frac{8}{36} + 2 \cdot \frac{1}{38} + 2 \cdot \frac{2}{45} + 2 \cdot \frac{25}{396} \approx 0.4783$$

The probability of throwing a 7 or 11 on the first throw is $\frac{8}{36}$.

Example

☞ The ideas of independence and conditioning are remarkably effective at working together to provide neat solutions to a wide range of problems.

Example

A coin shows a head with probability p, or a tail with probability $1 - p$. It is flipped repeatedly until the first head appears.

- What is the probability that a heads appears on an even number of tosses?

Example – continued

☞ Let E be the event of heads eventually appearing on an even toss. By the partition rule, conditioning on the outcome of the first toss (H or T):

$$\begin{aligned}\mathbf{P}(E) &= \mathbf{P}(E \mid H) \cdot \mathbf{P}(H) + \mathbf{P}(E \mid T) \cdot \mathbf{P}(T) \\ &= 0 \cdot p + \mathbf{P}(E \mid T) \cdot (1 - p),\end{aligned}$$

since 1 is odd, $\mathbf{P}(E \mid H) = 0$.

☞ What about $\mathbf{P}(E \mid T)$? The key is that we now need an odd number of tosses to succeed.

Example – continued

☞ Let O be the event that the first head is on odd throw, and N be that no head is tossed.

$$\mathbf{P}(E^c) = \mathbf{P}(O) + \mathbf{P}(N) = \mathbf{P}(O) + 0 = \mathbf{P}(O).$$

But,

$$\mathbf{P}(O) = \mathbf{P}(E \mid T),$$

so

$$\mathbf{P}(E \mid T) = 1 - \mathbf{P}(E).$$

☞ Thus,

$$\mathbf{P}(E) = \mathbf{P}(E \mid T) \cdot (1-p) = (1 - \mathbf{P}(E)) \cdot (1-p)$$

Example – continued

$$\mathbf{P}(E) = (1 - \mathbf{P}(E)) \cdot (1-p)$$

☞ Let $q = 1 - p$, and solve for $\mathbf{P}(E)$:

$$\mathbf{P}(E) = (1 - \mathbf{P}(E)) \cdot q$$
$$\mathbf{P}(E) = \frac{q}{1+q} = \frac{1-p}{1+(1-p)} = \frac{1-p}{2-p}.$$

Example – continued

$$\mathbf{P}(E) = \frac{1-p}{2-p}.$$

Consider a fair coin ($p = \frac{1}{2}$).

- The probability of tossing heads on an even throw is

$$\mathbf{P}(E) = \frac{1}{3}$$

- The probability of tossing heads on an odd throw is

$$\mathbf{P}(O) = 1 - \frac{1}{3} = \frac{2}{3}$$

Does this seem right?

Example – Method 2

Method 2. Let E_n ($n \geq 0$) be the event that heads first appears on the $2n$th toss. Let $q = 1 - p$.

$$\mathbf{P}(E_n) = pq^{2n-1} = pq(q^2)^n$$

Since the events $E_1, E_2, \ldots$ are mutually exclusive,

$$\begin{aligned}\mathbf{P}(E) &= \sum_{n=0}^{\infty} \mathbf{P}(E_n) \\ &= \sum_{n=0}^{\infty} pq(q^2)^n \\ &= \frac{pq}{1-q^2} = \frac{p(1-p)}{1-(1-p)^2} \\ &= \frac{1-p}{2-p}\end{aligned}$$

Example: Huygen's problem

Example

Two players, A and B, take turns at throwing dice; each needs some score to win. If one player does not throw the required score, the play continues with the next person throwing. At each of their attempts A wins with probability α and B wins with probability β.

(a) What is the probability that A wins if A throws first?

(b) What is the probability that A wins if A throws second?

Solution to Huygen's problem

☞ Let

- p_1 be the probability A wins when A has the first throw, and
- p_2 be the probability A wins when B has the first throw.

By conditioning on the outcome of the first throw, when A is first,

$$p_1 = \alpha + (1-\alpha)p_2.$$

When B is first, conditioning on the first throw gives

$$p_2 = (1-\beta)p_1.$$

(If B throws their score, A loses.)

☞ Solving this pair gives

$$p_1 = \frac{\alpha}{\alpha+\beta-\alpha\beta}$$

$$p_2 = \frac{(1-\beta)\alpha}{\alpha+\beta-\alpha\beta}$$

Example

☞ Suppose A and B are tossing a fair coin, and each needs a head. A throws first. (Here, $\alpha = \beta = \frac{1}{2}$).

- The probability that A wins is

$$p_1 = \frac{\frac{1}{2}}{\frac{1}{2}+\frac{1}{2}-\frac{1}{4}} = \frac{2}{3}$$

The problem is equivalent to tossing a heads on an odd throw.

- The probability that B wins is

$$p_2 = \frac{1}{2}\cdot p_1 = \frac{1}{3}$$

The problem is equivalent to tossing a heads on an even throw.

Example

☞ Suppose A and B are throwing a pair of dice. A needs a 5 and B a 7. (Here, $\alpha = \frac{1}{12}$ and $\beta = \frac{1}{6}$).

- The probability that A wins if A throws first is

$$p_1 = \frac{\frac{1}{9}}{\frac{1}{6}+\frac{1}{9}-\frac{1}{54}} = \frac{3}{7} \approx 0.43$$

- The probability that A wins if A throws second is

$$p_2 = \frac{5}{6}\cdot p_1 = \frac{15}{42} \approx 0.36$$