

## Lecture 6: Sept 1

### Last time

- Calculus of Probabilities (1.2)

### Today

- no class next Monday (Labor day)
- Conditional Probability (1.3)
- Independence (1.3)

**Theorem** If  $\Pr$  is a probability function, then

1.  $\Pr(A) = \sum_{i=1}^{\infty} \Pr(A \cap C_i)$  for any partition  $C_1, C_2, \dots$ ;
2.  $\Pr(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \Pr(A_i)$  for any sets  $A_1, A_2, \dots$

where (1) is also referred to as “Total probability” and (2) is Boole’s inequality.

*proof:*

## Conditional Probability

All of the probabilities that we have dealt with thus far have been unconditional probabilities. A sample space was defined and all probabilities were calculated with respect to that sample space. In many instances, however, we are in a position to update the sample space based on new information. In such cases we want to be able to update probability calculations or to calculate *conditional probabilities*.

**Definition** If  $A$  and  $B$  are events in  $S$ , and  $\Pr(B) > 0$ , then the *conditional probability* of  $A$  given  $B$ , written  $\Pr(A|B)$ , is

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}.$$

Note that  $B$  becomes the sample space now:  $\Pr(B|B) = 1$ .

**Example** Four cards are dealt from the top of a well-shuffled deck. What is the probability that they are the four aces? (there are in total 52 cards)

*solution:*

**Theorem** (Bayes' Rule) Let  $A_1, A_2, \dots$  be a partition of the sample space, and let  $B$  be any set. Then, for each  $i = 1, 2, \dots$ ,

$$\Pr(A_i|B) = \frac{\Pr(B|A_i) \Pr(A_i)}{\sum_{j=1}^{\infty} \Pr(B|A_j) \Pr(A_j)}.$$

*proof:*

## Independence

**Definition** Two events,  $A$  and  $B$ , are *statistically independent* if

$$\Pr(A \cap B) = \Pr(A) \Pr(B)$$

Note that independence could have been defined using Bayes' rule by  $\Pr(A|B) = \Pr(A)$  or  $\Pr(B|A) = \Pr(B)$  as long as  $\Pr(A) > 0$  or  $\Pr(B) > 0$ . More notation, often statisticians omit  $\cap$  when writing intersection in a probability function which means  $\Pr(AB) = \Pr(A \cap B)$ . Sometime, statisticians use comma  $(,)$  to replace  $\cap$  inside a probability function too,  $\Pr(A, B) = \Pr(A \cap B)$ .

**Theorem** If  $A$  and  $B$  are independent events, then the following pairs are also independent.

1.  $A$  and  $B^c$ ,
2.  $A^c$  and  $B$ ,
3.  $A^c$  and  $B^c$ .

**Theorem** If  $A$  and  $B$  are independent events, then the following pairs are also independent.

1.  $A$  and  $B^c$ ,
2.  $A^c$  and  $B$ ,
3.  $A^c$  and  $B^c$ .

*proof:*

**Example** Let the sample space  $S$  consist of the  $3!$  permutations of the letters  $a$ ,  $b$ , and  $c$  along with the three triples of each letter. Thus,

$$S = \left\{ \begin{array}{ccc} aaa & bbb & ccc \\ abc & bca & cba \\ acb & bac & cab \end{array} \right\}.$$

Furthermore, let each element of  $S$  have probability  $\frac{1}{9}$ . Define

$$A_i = \{i^{th} \text{ place in the triple is occupied by } a\}.$$

What are the values for  $\Pr(A_i), i = 1, 2, 3$ ? Are they pairwise independent?  
*solution*

**Definition\*** A collection of events  $A_1, \dots, A_n$  are *mutually independent* if for any subcollection  $A_{i_1}, \dots, A_{i_k}$ , we have

$$\Pr(\cap_{j=1}^k A_{i_j}) = \prod_{j=1}^k \Pr(A_{i_j}).$$

## Random Variables

In many experiments, it is easier to deal with a summary variable than with the original probability structure.

**Example** consider an opinion poll, we might decide to ask 50 people whether they agree or disagree with a certain issue. If we record a “1” for agree and “0” for disagree, the sample space for this experiment has  $2^{50}$  elements (all length 50 strings consist of 1s and 0s). However, if we are only interested in the number of people who agree, we may define a variable  $X =$  number of 1s recorded out of 50. Then, the sample space for  $X$  is the set of integers  $\{0, 1, 2, \dots, 50\}$ .

**Definition** A *random variable* (r.v.) is a function from a sample space  $S$  into the real numbers.

**Example** In some experiments random variables are implicitly used

### Examples of random variables

Experiment	Random variable
Toss two dice	$X =$ sum of numbers
Toss a coin 25 times	$X =$ number of heads in 25 tosses
Apply different amounts of fertilizer to corn plants	$X =$ yield / acre

In defining a random variable, we have also defined a new sample space (the range of the random variable).

**Induced probability function** Suppose we have a sample space  $S = \{s_1, s_2, \dots, s_n\}$  with a probability function  $\Pr$  defined on the original sample space. We define a random variable  $X$  with range  $\mathcal{X} = \{x_1, \dots, x_m\}$ . We can define a probability function  $\Pr_X$  on  $\mathcal{X}$  in the following way. We will observe  $X = x_i$  if and only if the outcome of the random experiment is an  $s_j \in S$  such that  $X(s_j) = x_i$ . Therefore,

$$\Pr_X(X = x_i) = \Pr(\{s_j \in S : X(s_j) = x_i\}),$$

defines an *induced* probability function on  $\mathcal{X}$ , defined in terms of the original function  $\Pr$ .

We will write  $\Pr(X = x_i)$  rather than  $\Pr_X(X = x_i)$  for simplicity. Note on notation: random variables will always be denoted with uppercase letters and the realized values of the variable (or its range) will be denoted by the corresponding lowercase letters.

**Example** Consider the experiment of tossing a fair coin three times. Define the random variable  $X$  to be the number of heads obtained in the three tosses. A complete enumeration of the value of  $X$  for each point in the sample space is

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

What is the range of  $X$ ? What is the induced probability function  $\Pr_X$ ?

*solution:*

So far, we have seen finite  $S$  and finite  $\mathcal{X}$ , and the definition of  $\Pr_X$  is straightforward. If  $\mathcal{X}$  is uncountable, we define the induced probability function,  $\Pr_X$  for any set  $A \subset \mathcal{X}$ ,

$$\Pr_X(X \in A) = \Pr(\{s \in S : X(s) \in A\}).$$

This defines a legitimate probability function for which the Kolmogorov Axioms can be verified.