Lecture 8: Sept 5

Last time

- Conditional Probability (1.3)

Today

- HW2 posted
- Independence
- Random variables

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

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

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.

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,

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

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 $A_1, \ldots, A_n$ are *mutually independent* if for any subcollection $A_{i_1}, \ldots, A_{i_k}$, we have

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

Random Variables

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

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 $S = \{s_1, s_2, \ldots, s_n\}$ with a probability function Pr defined on the original sample space. We define a random variable X with range $\mathcal{X} = \{x_1, \ldots, x_m\}$. We can define a probability function $\Pr_X$ on $\mathcal{X}$ in the following way. We will observe $X = x_i$ if an 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 leeters 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 anyset $A \subset \mathcal{X}$,

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

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

Distribution Functions

Distribution Functions are used to describe the behavior of a r.v.

Cumulative distribution function

Definition The *cumulative distribution function* or *cdf* of a random variable X, denoted by $F_X(x)$, is defined by

$$F_X(x) = \Pr_X(X \leqslant x), \text{ for all } x.$$

Definition The *survival function* of a random variable X, is defined by

$$S_X(x) = 1 - F_X(x) = \Pr_X(X > x).$$

Example Consider the experiment of tossing three fair coins, and let X = number of heads observed. The cdf of X is

$$F_X(x) = \begin{cases} 0 & \text{if } -\infty < x < 0 \\ \frac{1}{8} & \text{if } 0 \leqslant x < 1 \\ \frac{1}{2} & \text{if } 1 \leqslant x < 2 \\ \frac{7}{8} & \text{if } 2 \leqslant x < 3 \\ 1 & \text{if } 3 \leqslant x < \infty \end{cases}$$