Stats 130
Day 6 Notes

Elijah Hantman

Examples

A gene has two alleles A and a. The gene exhibits itself through a trait, such as hair color or blood type with two versions. A is dominant and a is recessive. Individuals with AA and Aa show the same versino of the trait, while individuals with aa show the other version. Assume that the genotypes AA, Aa, and aa occur with probability $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{1}{4}$ respectively. Let E be the event that an individual has the dominant trait. There are six possible parental genotypes.

	(AA, AA)	(AA, Aa)	(AA, aa)	(Aa, Aa)	(Aa, aa)	(aa, aa)
Event	B_1	B_2	B_3	B_4	B_5	B_6
Prob	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{16}$
$Pr(A\|B_i)$	1	1	1	$\frac{3}{4}$	$\frac{1}{2}$	0

Partitions

The notion of a partition is useful for conditional probability. The idea is to split the sample space into disjoint events.

Example Continued

$$Pr(E) = \sum_{i=1}^{5} Pr(B_i)Pr(E|B_i) = \frac{3}{4} \tag{1}$$

This is the denominator used in Bayes' theorem. We can now calculate the probabilities of different genotypes.

$$Pr(B_1|E) = \frac{1 \times 16}{\frac{3}{4}} \tag{2}$$

$$Pr(B_4|E) = \frac{3/4 \times 1/4}{\frac{3}{4}} \tag{3}$$

$$Pr(B_6|E) = 0 \tag{4}$$

Bayes' Theorem

Given a partition $B_1, ...B_n$ of S, such that $Pr(B_j) > 0, \forall j$, and given an event $A, Pr(A) > 0$ then

$$Pr(B_i|A) = \frac{Pr(A|B_i)Pr(A)}{Pr(B_i)}$$

Random Variable

You toss a coin three times to participate in a game where you make \$1 for each H, and you pay \$1 for each T. Let X be the total payoff. There are eight possible outcomes to this game.

Outcome	Value
HHH	+3
HHT	+1
HTH	+1
THH	+1
THT	-1
TTH	-1
HTT	-1
TTT	-3

The payoff is a random variable. Something to note is that there are only 4 distinct values for 8 outcomes. The mapping from outcome to value is not symmetric, it is surjective but not injective.

$$X : \{\omega_1, ...\omega_8\} \to \{-3, -1, 1, 3\}$$

Random variables are functions.

Example 2 Revisited

In Example 2 we considered the problem of power and water demand of a new building. The sample space is given by all points that correspond to water demand between 4 and 200 (1000G/day) and power demand between 1 and 150 (MM-Kw/h).

We considered the set E, that corresponds to high water and power demand.

Take a building at random and define

$$Z = \begin{cases} 1 & if \quad (x, y) \in E \\ 0 & if \quad (x, y) \notin E \end{cases}$$

Then $Z : S \to \{0, 1\}$ is a function that depends on the outcome of a random experiment.

Often Z is denoted $\mathbf{1}_E$ and is called the indicator.

Definition: A random variable is a function on S which returns a real value for each outcome of an experiemtn with random outcomes. Thus $X : S \to \mathbb{R}$

Random variables are characterized by their probabilistic behavior. For each possible value of X we want to obtain the associated events and calculate their probabilities.

Notation:

$$Pr(X \in A) = Pr(s \in S : X(s) \in A), A \subseteq \mathbb{R}$$

The probability that X is in A is shorthand for arbitrary s in the Sample space what is the probability that the mapping X takes s to a subset of the reals A.

A case of interest is when $A = \{x\}$, then $Pr(X = x) = Pr(s \in S : X(s) = x)$

Note the difference between X and x. X denotes the random variable, which is a function of the outcomes, and x is a value, a real number.

Example

In the three coin toss example, what is the probability of making money?

$$Pr(X > 0) = Pr(\omega_i : X(\omega_i) > 0 = Pr(\omega_1, \omega_2, \omega_3, \omega_4) = \frac{4}{8} = \frac{1}{2}$$

Example

Suppose you toss a coin 10 times. Let X be the number of H. What is the probability of x heads?

$$Pr(X = x) = \binom{10}{x}\left(\frac{1}{2}\right)^{10}$$

This is valid for $x = \{0, 1, ...10\}$ All other probabilities of x are equal to 0. $Pr(X > 3)1 - Pr(X \leq 3)$
A Bernolli equation maps from E to 1, and everthing else to 0.