273 - Lecture 7

Sariel Har-Peled

February 17, 2004

1 Previous Lecture

1. Discrete probability.
2. Probability space: (S, F, P).
3. Conditional probability

2 Independent Events

Definition 2.1 (Independent Events) Two events E and F are *independent* iff $\mathbf{Pr}[E \cap F] = \mathbf{Pr}[E] \cdot \mathbf{Pr}[F]$.

If E and F are independent, then

$$\mathbf{Pr}\left[E \,\middle|\, F\right] \mathbf{Pr}[F] = \mathbf{Pr}[E]\,\mathbf{Pr}[F] \quad \text{and} \quad \mathbf{Pr}\left[F \,\middle|\, E\right] \mathbf{Pr}[E] = \mathbf{Pr}[F]\,\mathbf{Pr}[E]\,.$$

2.0.1 Bayes Rule

$\mathbf{Pr}[E \cap F] = \mathbf{Pr}[F]\,\mathbf{Pr}\left[E \,\middle|\, F\right] = \mathbf{Pr}[E]\,\mathbf{Pr}\left[F \,\middle|\, E\right]$.

From this we get *Bayes rule*:

$$\mathbf{Pr}\left[F \,\middle|\, E\right] = \frac{\mathbf{Pr}\left[E \,\middle|\, F\right] \mathbf{Pr}[F]}{\mathbf{Pr}[E]}.$$

Bayes Rule
Let $A_1, A_2, \ldots, A_n$ be a partition of S. Then

$$\mathbf{Pr}\left[A_k \,\middle|\, E\right] = \frac{\mathbf{Pr}\left[E \,\middle|\, A_k\right] \mathbf{Pr}[A_k]}{\sum_{i=1}^{n} \mathbf{Pr}\left[E \,\middle|\, A_i\right] \mathbf{Pr}[A_i]}.$$

Why should we care? Imagine, that a patient comes to a doctor. The doctor does some checks, which gives it the information E. Now, the doctor knows what is the probability to get E (E - might for example the result of a blood pressure check, and the sugar level in the blood of the patient) if the patient have the disease A_i. But the doctor whats is the other direction, given the "evidence" E, what is the probability that the patient have the disease A_i. Namely, we want to compute $\mathbf{Pr}\left[A_i \,\middle|\, E\right]$ (i.e., the posterior probabilities) while given the probabilities $\mathbf{Pr}\left[E \,\middle|\, A_i\right]$ and $\mathbf{Pr}[A_i]$, for $i = 1, \ldots, n$.

3 Binomial Distribution

Definition 3.1 (Bernoulli Trials) A sequence of independent experiments, each with 2 possible outcomes: success or failure (i.e., head or tail for a coin). The probability for success is p, and the probability for failure is $1 - p = q$.

Example 3.2 If there are n coin flips (of a fair coin), what is the probability of having k heads?

Lets look on one possible outcome $HHHTHT \ldots TTTHT$ (where there are k heads). Clearly, the probability for getting this sequence is $1/2^n$. How many different sequence do we have with k heads? Clearly, this is $\binom{n}{k}$. Thus, the probability of having k heads in n coin flips is $\binom{n}{k}2^{-n}$.

In general, if the probability for success is p, the probability of having k successes in n experiments is $\binom{n}{k}p^k q^{n-k}$, using the same argumentation as above.

Definition 3.3 (Random Variable) For a probability space $(S, \mathcal{F}, P)$, a random variable is a function $f : S \to \mathbb{R}$ which assigns a value for every element (i.e., atomic event) in S. We usually would refer to the function $f(\cdot)$ using capital letter, such as X.[1]

Example 3.4 Toss a coin 5 times, and let X the random variable be the number of heads in those coin tosses. Thus:

$$HHHTT \to 3, \qquad HTTHT \to 2 \qquad TTTTT \to 5.$$

It is important to realize, that usually when we speak about an experiment, we usually know the value of some random variable of the probability space. Very rarely we have the "real" atomic event that happened, and usually we do not care about the real event, because we care about how the random variable behaves.

Definition 3.5 (Expected Value) For a random variable, the *expectation* of X, is its average value. Formally:

$$\mathbf{E}[X] = \sum_{s \in S} X(s)\,\mathbf{Pr}[s] = \sum_{s \in S} X(s)\,\mathbf{Pr}[\{s\}]\,.$$

[1]Since a random variable is neither random nor variable, one has to wonder what the connection between the name and what it means.

Example 3.6 For the probability space

$$S = \left\{ \begin{array}{ll} A, & (* \ \mathbf{Pr}[\{A\}] = 1/4, X = 3 \ *) \\ B, & (* \ \mathbf{Pr}[\{B\}] = 1/8, X = 1 \ *) \\ C, & (* \ \mathbf{Pr}[\{C\}] = 1/8, X = 4 \ *) \\ D, & (* \ \mathbf{Pr}[\{D\}] = 1/2, X = 2 \ *) \end{array} \right\}$$

As such,

$$\begin{aligned} \mathbf{E}[X] &= \mathbf{Pr}[\{A\}]\, X(A) + \mathbf{Pr}[\{B\}]\, X(B) + \mathbf{Pr}[\{C\}]\, X(C) + \mathbf{Pr}[\{D\}]\, X(D) \\ &= \frac{1}{4} \cdot 3 + \frac{1}{8} \cdot 1 + \frac{1}{8} \cdot 4 + \frac{1}{2} \cdot 2 = \frac{19}{8} = 2\frac{3}{8}. \end{aligned}$$

Note that the expected value is neither integer, nor a likely value. It is the average value.

The concept of expectation is related to the center of mass, average value of a function:

$$\int_x x f(x) \mathrm{dx} \approx \sum_i x_i f(x_i) \Delta x_i.$$

Example 3.7 Roll a die. What is the expected value?

$$\frac{1}{6} \cdot 1 + \cdots + \frac{1}{6} 6 = \frac{1}{6} = \frac{1}{6} \frac{7 \cdot 6}{2} = \frac{21}{6} = 3.5.$$

This is the average value I expect to get, if I throw a die enough times.