0368-3248-01-Algorithms in Data Mining **Fall 2011**

Lecture 1: Preliminaries

Lecturer: Edo Liberty

Warning: *This note may contain typos and other inaccuracies which are usually discussed during class. Please do not cite this note as a reliable source. If you find mistakes, please inform me.*

---

1 Preliminaries

A variable X is a random variable if it assumes different values according to probability distribution. For example, X can denote the outcome of a three sided die throw and X taking the values $x = 1, 2, 3$ with equal probabilities. The expectation of X is the sum over the possible values times the probability of the events.

$$E[X] = \sum_{x=1}^{3} x \Pr(X = x) = 1\frac{1}{3} + 2\frac{1}{3} + 3\frac{1}{3} = 2 \tag{1}$$

Another example is a continuous variable Y taking the values $[0, 1]$ uniformly. Meaning that the probability of Y being in the interval $[t, t + dt]$ is exactly dt. And so the expectation of Y is:

$$E[Y] = \int_{t=0}^{1} dt \cdot t = \frac{1}{2}t^2|_0^1 = 1/2 \tag{2}$$

1.1 Dependence and Independence

A variable X is said to be *dependent* on Y if given the value of Y changes the probability distribution of X. For example. Assume the variable X takes the value 1 if Y takes a value of less than 1/3 and the values 2 or 3 with equal probably otherwise (1/2 each).

Clearly, the probability of X assuming each of its values is still 1/3. however, if we know that Y is 0.7234 the probability of X assuming the value 1 is zero.

This is denoted by $E(X|Y)$ (read: expectation of X given Y).

$$E(X|Y) = \sum_{x=1}^{3} x \Pr(X = x|Y \leq 1/3) = 1 \cdot 1 \tag{3}$$

$$E(X|Y) = \sum_{x=1}^{3} x \Pr(X = x|Y > 1/3) = 1 \cdot 0 + 2\frac{1}{2} + 3\frac{1}{2} = 2.5 \tag{4}$$

Note that $E(X|Y)$ is a function of y!! $E(X|Y) = 1$ for $y \in [0, 1/3]$ and $E(X|Y) = 2.5$ for $y \in (1/3, 1]$.

Definition 1.1. *Two variables are said to be Independent if:*

$$\forall y, \quad E[X|Y = y] = E[X].$$

They are dependent otherwise.

Fact 1.1. *For any two random variables (even if they are dependent):*

$$E_Y E_X[X|Y] = E[X] \tag{5}$$

Checking this for our example:

$$E_Y E_X[X|Y] = \int_{y=0}^{1} E_X[X|Y=y] = \int_{y=0}^{1/3} 1dy + \int_{y=1/3}^{1} 2.5dy = \frac{1}{3}\cdot 1 + \frac{2}{3}\cdot 2.5 = 2 \tag{6}$$

Fact 1.2 (Linearity of expectation 1)**.** *For any random variable and any constant α:*

$$E[\alpha X] = \alpha E[X]$$

Fact 1.3 (Linearity of expectation 2)**.** *For any random variable and any constant α: $E[\alpha X] = \alpha E[X]$ For any two random variables (even if they are dependent):*

$$E_{X,Y}[X+Y] = E[X] + E[Y] \tag{7}$$

Given the previous fact we can convince ourselves that this is true. $E_{X,Y}[x+y] = E_Y E_X[X+Y] = E_Y[X|Y] + E_Y[Y] = E_X[X] + E_Y[Y]$.

Fact 1.4 (Markov's inequality)**.** *For any positive random variable X:*

$$\Pr(X > t) \leq \frac{E[X]}{t} \tag{8}$$

Definition 1.2. *The variance of a random variable x is:*

$$\sigma^2[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2 \tag{9}$$

Fact 1.5 (Chebyshev's inequality)**.** *For any random variable X*

$$\Pr[|X - E[X]| > t] \leq \frac{\sigma^2(X)}{t^2} \tag{10}$$