Expectation of Discrete Random Variables

Saravanan Vijayakumaran
sarva@ee.iitb.ac.in

Department of Electrical Engineering
Indian Institute of Technology Bombay

February 13, 2013

Expectation of Discrete Random Variables

Definition

The expectation of a discrete random variable X with probability mass function f is defined to be

$$E(X) = \sum_{x:f(x)>0} xf(x)$$

whenever this sum is absolutely convergent. The expectation is also called the mean value or the expected value of the random variable.

Example

- Bernoulli random variable

  $\Omega = \{0, 1\}$

$$f(x) = \begin{cases} p & \text{if } x = 1 \\ 1 - p & \text{if } x = 0 \end{cases}$$

  where $0 \leq p \leq 1$

$$E(X) = 1 \cdot p + 0 \cdot (1 - p) = p$$

What is absolute convergence?

- A discrete random variable can take a countable number of variables
- Its expectation may involve a countable sum $\sum_{i=1}^{\infty} a_i$
- The sum $\sum_{i=1}^{\infty} a_i$ is called an infinite series and $s_n = \sum_{i=1}^{n} a_i$ are called the partial sums of the series
- The series is said to converge if the sequence $\{s_n\}$ converges
- A series $\sum a_i$ is said to converge absolutely if the series $\sum |a_i|$ converges

Example

- A series which converges but does not converge absolutely

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \cdots$$

Why do we need absolute convergence?

- Consider a rearrangement of the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \cdots$$

where two positive terms are followed by one negative term

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \cdots$$

- Since

$$\frac{1}{4k-3} + \frac{1}{4k-1} - \frac{1}{2k} > 0$$

the rearrranged series sums to a value greater than $\frac{5}{6}$

Theorem

If $\sum a_i$ is a series which converges absolutely, then every rearrangement of $\sum a_i$ converges, and they all converge to the same sum

Considering only absolutely convergent sums makes the expectation well-defined.

$$E(X) = \sum_{x:f(x)>0} xf(x)$$

Expectations of Functions of Discrete RVs

- If X has pmf f and $g : \mathbb{R} \to \mathbb{R}$, then

$$E(g(X)) = \sum_x g(x)f(x)$$

whenever this sum is absolutely convergent.

Example

- Suppose X takes values $-2, -1, 1, 3$ with probabilities $\frac{1}{4}, \frac{1}{8}, \frac{1}{4}, \frac{3}{8}$ respectively.
- Consider $Y = X^2$. It takes values $1, 4, 9$ with probabilities $\frac{3}{8}, \frac{1}{4}, \frac{3}{8}$ respectively.

$$E(Y) = \sum_y yP(Y = y) = 1 \cdot \frac{3}{8} + 4 \cdot \frac{1}{4} + 9 \cdot \frac{3}{8} = \frac{19}{4}$$

Alternatively,

$$E(Y) = E(X^2) = \sum_x x^2 P(X = x) = 4 \cdot \frac{1}{4} + 1 \cdot \frac{1}{8} + 1 \cdot \frac{1}{4} + 9 \cdot \frac{3}{8} = \frac{19}{4}$$

Some Properties of Expectation

- If $a, b \in \mathbb{R}$, then $E(aX + bY) = aE(X) + bE(Y)$
- If X and Y are independent, $E(XY) = E(X)E(Y)$
- X and Y are said to be uncorrelated if $E(XY) = E(X)E(Y)$
- Independent random variables are uncorrelated but uncorrelated random variables need not be independent

Example

Y and Z are independent random variables such that Z is equally likely to be 1 or -1 and Y is equally likely to be 1 or 2.

Let $X = YZ$. Then X and Y are uncorrelated but not independent.

Moments

- If k is a positive integer, the kth moment m_k of X is defined to be

$$m_k = E(X^k)$$

- The kth central moment σ_k is

$$\sigma_k = E\left[(X - m_1)^k\right]$$

- The first moment is the same as the expectation $m_1 = E(X)$
- The second central moment $\sigma_2 = E[(X - m_1)^2]$ is called the variance
- The positive square root of the variance is called the standard deviation

$$\sigma = \sqrt{E\left[(X - m_1)^2\right]}$$

Some Properties of Variance

- $\text{var}(X) = E(X^2) - [E(X)]^2$
- For $a \in \mathbb{R}$, $\text{var}(aX) = a^2 \, \text{var}(X)$
- $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$ if X and Y are uncorrelated

Questions?