Random Variables and Expectation

Anjali Bhavan

20-12-2020

1 Random variables

Definition 1 (Random variable). *A function that assigns a real value $X(w)$ to every outcome w in the sample space.*

Two types: discrete and continuous.

1. Distribution function for discrete RV:

$$F(x) = \sum_{i:x_i \leq x} p(x_i)$$

   where $p(x_i)$ is the probability mass function (basically probability value).

2. Distribution function for continuous RV:

$$F(x) = \int_{-\infty}^{x} f(x)\,dx$$

   where $f(x)$ is the probability density function.

2 Jointly distributed random variables

For two RVs X and Y associated with the same random experiment,

1. when X and Y are discrete:
   Probability mass function:

$$p(x_i, y_i) = P(X = x_i, Y = y_i)$$

   and

$$\sum_x \sum_y p(x_i, y_j) = 1$$

   Marginal probability mass functions are given by:

$$p(x_i) = P(X = x_i) = \sum_j P(X = x_i, Y = y_j) = \sum_j p_{ij}$$

$$p(y_j) = P(Y = y_j) = \sum_i P(X = x_i, Y = y_j) = \sum_i p_{ij}$$

2. when X and Y are continuous:
   The joint probability density function is given by $f(x,y)$, and

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$$

   The marginal p.d.fs are given by:

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

   Important: if $C = \{(x,y) : a \le x \le b, c \le y \le d\}$, then

$$P((X,Y) \in C) = \int_a^b \int_c^d f(x,y)\,dx\,dy$$

3 Expectation

Definition 2 (Expectation)**.** *Expectation of a random variable X is the weighted average of all the possible values X can take. For a discrete RV, it is given by:*

$$E(X) = \sum_i p_i x_i$$

And for a continuous RV,

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,dx$$

Note: Addition property of expectation is valid for any two random variables X and Y, but multiplication property is valid only if X and Y are independent. Additive property states that $E(X+Y) = E(X)+E(Y)$. Multiplication property states that $E(XY) = E(X)E(Y)$ iff X and Y are independent. **Note:** Variance is given by $E(X^2) - E(X)^2$.

3.1 Covariance

Definition 3 (Covariance)**.** *Covariance of two random variables X and Y is a measure of how the two are related to each other. It is given by:*

$$\begin{aligned} Cov(X,Y) &= E(X-\bar{X})(Y-\bar{Y}) \\ &= E(XY - X\bar{Y} - Y\bar{X} + \bar{X}\bar{Y}) \\ &= E(XY) - \bar{Y}E(X) - \bar{X}E(Y) + \bar{X}\bar{Y} \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

Positive covariance means X and Y are increasing or decreasing together, negative means the opposite. When X and Y are independent, covariance is 0. But converse is not always true.

Some important formulae:

- $Var(aX+b) = a^2 Var(X)$
- $Var(aX+bY) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X,Y)$

4 Moment generating function

The moment generating function of a RV X is given by:

$$M(X) = E[e^{tX}] = \begin{cases} \sum_x e^{tx} p(x) & X is discrete \\ \int_{-\infty}^{\infty} e^{tx} f(x)\,dx & X is continuous \end{cases}$$

Moments are usually measures for a function's graph shape. Particular moments have particular values: for instance, the zeroth moment of a random variable is 1, the first the mean, the second the variance, the third the skewness and fourth the kurtosis. MGF can describe a distribution uniquely if it exists. If X and Y are two RVs, their sum's MGF is given by the product of the individual MGFs.
Important: The MGF of a RV about any arbitrary point can be derived from the MGF about origin by:

$$M_a(t) = E[e^{t(x-a)}] = e^{-at}E[e^{tx}] = e^{-at}M_o(t)$$

5 Chebyshev's Inequality

This basically shows that no more than a certain fraction of values can lie more than a few standard deviations away from the mean. Specifically, no more than $1/k^2$ of the distribution's values can be more than k standard deviations away from the mean. Mathematically, let X be a random variable with mean μ and variance σ^2. Then,

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

This is used to prove the weak law of large numbers, which states:

Theorem 1. *The probability that the average of a sequence of i.i.d. RVs differs from its mean by more than ϵ tends to 0 as the number of terms tends to infinity, provided the RVs have finite variance.*