CS215 Random Variables

Harsh Shah

August 2021

1 Random Variables

A R.V. defined on probability space (Ω, β, P), is a function on domain Ω to R. It also induces a probability function P_X associated with range of X.

1.1 Discrete R.V.

Discrete R.V.: Cardinality of range is finite or countably infinite.Eg, Experiment: Coin toss, Map 'heads' to 1 and 'tails' to -1.

1.2 Continuous R.V.

CDF is continuous. Cardinality of range is uncountably infinite.
Eg:Experiment: Measurement of temperature. Take X to identity function, then range of X is R^+, and hence uncountably infinite.

1.3 Event probabilities via R.V.

$$P_X(a < X < b) := P(a < X < b) = P(s \in \Omega | a < X(s) < b)$$

1.4 Cumulative Distributive Function(CDF)

Definition:

$$CDF_X(x) := f_X(x) = P_X(X \leq x)$$

Properties:

1. $lim_{x \to \infty} f_X(x) = 1$
2. $lim_{x \to -\infty} f_X(x) = 0$
3. $P(a < X \leq b) = f_X(b) - f_X(a)$
4. $P(c) = f_X(c) - f_X(c^-)$

Absolute Continuity:
Refer this for definition:Absolute Continuity(Stronger than continuity and uniform continuity)
We assume absolute continuity of CDF to avoid cases like cantor function which are not integral of their derivative.
Support:Support of a r.v. X is the set of all points having $P_X(.) > 0$

2 Distributions

Generally distributions refer to PMF/PDF.

2.1 Bernoulli distribution

$$P_X(x = 1; \alpha) = \alpha$$
$$P_X(x = 0; \alpha) = 1 - \alpha$$

The r.v. X can model the failure/success of any event.

2.2 Binomial distribution

Can be considered repeated Bernoulli trials.
The r.v.X can model the number of successes in n trials.

$$P_X(x = k; p, n) =^n C_k p^k (1 - p)^{n-k}$$

2.3 Geometric distribution

The distribution of r.v. X modelling the number of Bernoulli trials until the first success.

$$P_X(x = k; p) = (1 - p)^{k-1}$$

Also,

$$\mathrm{CDF}_X(x = k) = P(X \leq k) = 1 - P(X > k) = (1 - p)^k$$

Memoryless Property of Geometric Distribution:
$P_X(x > k + m | x > k) = P(x > m)$
(Can be proved by definition of conditional probability and expression for CDF of geometric distribution).

2.4 Poisson Distribution

$P(k, \tau)$ is the probability of k arrivals in time interval of τ. Number of arrivals in disjoint time intervals is independent.

Small interval probability: For small interval δ

$$P(k,\delta) = \begin{cases} 1-\lambda\delta & if \quad k=0 \\ \lambda\delta & if \quad k=1 \\ 0 & if \quad k>1 \end{cases}$$

λ is called the "arrival rate".

2.4.1 PDF for Poisson process

Our aim is to get the probability of k arrivals in a given time interval τ. Lets divide the interval τ into n equal parts, such that each of them is of length:
$\delta = \frac{\tau}{n}$
Therefore, probability of k occurrences of successes is:

$$P(k, n \quad trials) = ^n C_k (\lambda\delta)^k (1-\lambda\delta)^{n-k}$$

where, $n \to \infty$, $\delta \to 0$, $\delta n = \tau$

In the above limit it can be proved that PDF is:

$$P(k,\tau) = \frac{(\lambda\tau)^k e^{-\lambda\tau}}{k!}$$

Note: If τ not mentioned then assume unit interval

3 Sum of R.V.

Let X and Y be two independent R.V. and define another R.V. $Z = X + Y$.
In general, $P_Z(z) \neq P_X(x) + P_Y(y)$
Consider the **special case** that X and Y have poisson distribution with arrival rate λ, μ respectively.

$$P_Z(z=k) = \sum_{i=0}^{k} P_X(x=i) P_Y(k-i)$$

Using the expression of distribution of poisson variables it is easy to prove $P_Z(z)$ is poisson distribution with arrival rate $\lambda + \mu$.
This operation,$h(a) = \sum_{i=0}^{a} f(i) g(a-i)$ is called convolution of functions f and g, also written as f*g. **For continuous R.V.**,

$$f_Z(z) = P_Z(Z \leq z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} P_X(x) P_Y(y) dy dx$$

Partial derivativ w.r.t. z will yield,

$$P_Z(z) = \int_{-\infty}^{\infty} P_X(x) P_Y(z-x) dx$$

4 Poisson thinning

Consider two R.V. X and Y, $X \sim Poisson(\lambda)$ and
$P(Y|X=j) = P_{Binomial}(Y;p,j)$
Now,

$$P(Y=k) = \sum_{j=k}^{\infty} P(X=j, Y=k)$$

Writing the joint distribution as product, then using the expression for the respective distributions, it can be proved after much simplification that,

$$P(Y=k) = P(Y=k; p\lambda)$$

that is, poisson distribution with rate $p\lambda$

5 Exponential distribution

This distribution can model the time difference between two consecutive successes (or arrival) for a Poisson process. How?

$$f_{expo}(x) = P_{expo}(X \leq x) = 1 - P_{Poisson}(0 \text{ occurrences}, \text{x time})$$

$$f_{expo}(x) = 1 - e^{-\lambda x}$$

$$P_{expo}(x) = \lambda e^{-\lambda x}$$

Exponential distribution also satisfies **memoryless property** that is,

$$P(X > x+t | X > t) = P(X > x)$$

(Proof similar to that of geometric distribution) The only continuous r.v. satisfying memoryless property is the exponential one
Proof: Suppose there exists a continuous r.v.(X) with memoryless property. Therefore,

$$\begin{aligned}
&P_X(X > x+t | X > t) = P_X(X > x) \\
&\Rightarrow \frac{P_X(X > x+t)}{P_X(X > t)} = P_X(X > x) \\
&\Rightarrow P_X(X > x+dx) = P_X(X > x)P_X(X > dx) \\
&\Rightarrow f_X(x+dx) - f_X(x) = f_X(dx)(1 - f_X(x))
\end{aligned}$$

Thereafter, using $f_X(x=0) = 0$ we get that

$$f_X(x) = 1 - e^{-f'(0)x}$$

(QED)

6 Gaussian Distribution

PDF for a Gaussian distribution is given by:

$$P_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-\mu)^2}{2\sigma^2}}$$

μ is called location parameter
σ is called scale parameter

6.1 Central Limit Theorem

Consider continuous r.v.s X_i such that each of them has same distribution. Then the distribution of the mean r.v. defined as

$$\hat{X_n} = \frac{X_1 + X_2 \dots X_n}{n}$$

converges to that of normal r.v. . That is,

$$\text{lim}_{n\to\infty} P_{\hat{X_n}} = P_{normal}(x)$$

OR, the distribution of r.v. Y defined as $Y = \frac{\hat{X}-\mu}{(\sigma/\sqrt{n})}$ would tend to normal distribution with $\mu = 0$ and $\sigma^2 = 1$ as $n \to \infty$

6.2 Gaussian as limiting case of Bernoulli

To prove Gaussian distribution is a limiting case of Bernoulli distribution, we will use **Stirling's Approximation** for factorials given by,

$$n! = n^n e^{-n} \sqrt{2\pi n} [1 + O\left(\frac{1}{n}\right)]$$

Now, consider a Bernoulli distribution ($P(x)$) with n trials, probability of success in a single trial p and failure q ($p + q = 1$). Now,

$$P(x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

Using Stirling's Approximation for the factorials, we get

$$P(x) = \left(\frac{np}{x}\right)^x \left(\frac{nq}{n-x}\right)^{n-x} \sqrt{\frac{n}{2\pi x(n-x)}} [1 + O\left(\frac{1}{n}\right)]$$

Assume, $\delta = x - np$ which also implies $nq - \delta = n - x$
It can be proved using Taylor Series for $ln(1 + x)$ and much simplification that,

$$ln\left[\left(\frac{np}{x}\right)^x \left(\frac{nq}{n-x}\right)^{n-x}\right] = \frac{-\delta^2}{2npq} + O\left(\frac{\delta^3}{n^3}\right)$$

As $n \to \infty$ applying similar approximations on the square root part, we finally get:

$$P(x) = e^{\frac{-(x-np)^2}{2npq}} \sqrt{\frac{1}{2\pi npq}}$$

Now, lets try converting this P(X) into a distribution of continuous random variable Z which is related to X by,

$$Z = \Delta z(2 \cdot X - n)$$

A unit change in X results in $2 \cdot \Delta z$ change in Z, which implies that the distribution of Z (by probability mass conservation) is given as:

$$P_Z(z) \cdot (2 \cdot \Delta z) = P_X(x)$$

Now, suppose p=q=0.5, and define $D = (\Delta z)^2/(2\Delta t)$ where $n\Delta t = t$. Then,

$$P_Z(z) = \frac{1}{\sqrt{4\pi Dt}} e^{\frac{-z^2}{4Dt}}$$

The above is a solution to diffusion equation in one variable.