Discrete Random Variables

Chapter 3: Discrete Random Variables

Anduo Wang
Temple University
Email: anduo.wang@gmail.com
https://cis-linux1.temple.edu/∼tug29203/25fall-2033/index.html

The Poisson process

- In many random phenomena we encounter, it is not just one or two random variables that play a role but a whole collection — **a random process**.
- The Poisson process: a random process that models the occurrence of random points in time or space.
  - describes in a certain sense **the most random** way to distribute points in time or space
  - the notions of **homogeneity** and **independence**.

Random points

- Typical examples of the occurrence of random time points
  - Arrival times of email messages at a server
  - The times at which asteroids hit the earth
  - Arrival times of radioactive particles at a Geiger counter
  - Times at which your computer crashes, the times at which electronic components fail
  - Arrival times of people at a pump in an oasis.

Random points

- Typical examples of the occurrence of random time points
  - Arrival times of email messages at a server
  - The times at which asteroids hit the earth
  - Arrival times of radioactive particles at a Geiger counter
  - Times at which your computer crashes, the times at which electronic components fail
  - Arrival times of people at a pump in an oasis.
- Examples of the occurrence of random points in space
  - The locations of asteroid impacts with earth (2-dimensional)
  - The locations of imperfections in a material (3-dimensional)
  - The locations of trees in a forest (2-dimensional).

The Poisson process

- The Poisson process model often applies in situations where there is a very large population, and each member of the population has a very small probability to produce a point of the process.
  - A property of the Poisson process: points may lie arbitrarily close together.

The Poisson process

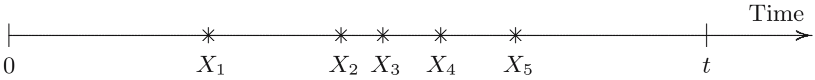
- The Poisson process model often applies in situations where there is a very large population, and each member of the population has a very small probability to produce a point of the process.
  - A property of the Poisson process: points may lie arbitrarily close together.
- The tree locations, **not** well modeled by the Poisson process.
- In a huge collection of atoms, just a few will emit a radioactive particle, well modeled

Random arrivals

Example

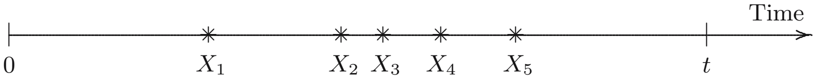
Calls arriving at a telephone exchange—the exchange is connected to a large number of people who make phone calls now and then.

Telephone calls arrive at random times $X_1, X_2, \cdots$ *at the telephone exchange during a time interval* $[0, t]$.



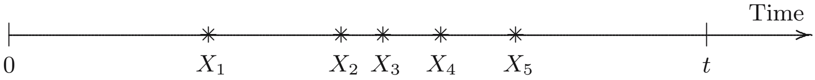
- The two basic assumptions on these random arrivals:
  - (**Homogeneity**) The rate λ at which arrivals occur is constant over time: in a subinterval of length δ_t the expectation of the number of telephone calls is $\lambda\delta_t$.
  - (**Independence**) The numbers of arrivals in disjoint time intervals are independent random variables.

An approximation



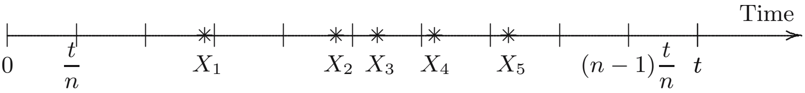
- Consider the total number of calls in an interval $[0, t]$, denoted by $N([0, t])$, abbreviating to N_t
- Homogeneity implies $E[N_t] = \lambda t$

An approximation



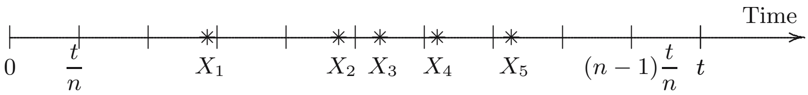
- Consider the total number of calls in an interval $[0, t]$, denoted by $N([0, t])$, abbreviating to N_t
- Homogeneity implies $E[N_t] = \lambda t$
- **Derive the probability distribution of N_t**

Distribution of N_t



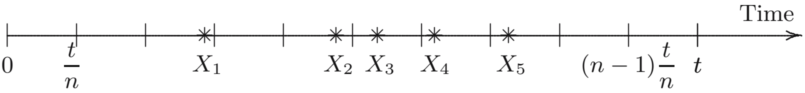
- Divide the interval $[0, t]$ into n intervals of length t/n
- Every interval $I_{j,n} = ((j-1)t/n, jt/n]$ will contain either 0 or 1 arrival, **when n is large enough**
- Let R_j be the number of arrivals in the time interval $I_{j,n}$
  - Since R_j is 0 or 1, R_j has a $Ber(p_j)$ distribution for some p_j

Distribution of N_t



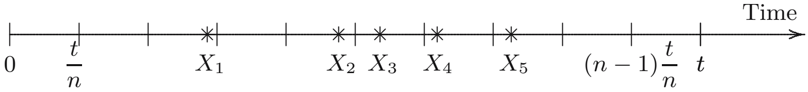
- Divide the interval $[0, t]$ into n intervals of length t/n
- Every interval $I_{j,n} = ((j-1)t/n, jt/n]$ will contain either 0 or 1 arrival, **when n is large enough**
- Let R_j be the number of arrivals in the time interval $I_{j,n}$
  - Since R_j is 0 or 1, R_j has a $Ber(p_j)$ distribution for some p_j
  - $E[R_j] = 0 \cdot (1 - p_j) + 1 \cdot p_j = p_j$

Distribution of N_t



- Divide the interval $[0, t]$ into n intervals of length t/n
- Every interval $I_{j,n} = ((j-1)t/n, jt/n]$ will contain either 0 or 1 arrival, **when n is large enough**
- Let R_j be the number of arrivals in the time interval $I_{j,n}$
  - Since R_j is 0 or 1, R_j has a $Ber(p_j)$ distribution for some p_j
  - $E[R_j] = 0 \cdot (1-p_j) + 1 \cdot p_j = p_j$
  - (by homogeneity) $p_j = \lambda \cdot \text{length} \quad \text{of} I_{j,n} = \lambda t/n$
- $N_t = R_1 + R_2 + \cdots + R_n$ has a $Bin(n, p)$ distribution, with $p = \lambda t/n$

Distribution of N_t



- $N_t = R_1 + R_2 + \cdots + R_n$ has a $Bin(n,p)$ distribution, with $p = \lambda t/n$
- $P(N_t = k) = \binom{n}{k}(\frac{\lambda t}{n})^k(1-\frac{\lambda t}{n})^{n-k}$ for $k = 0, \cdots, n$

$$\lim_{n\to\infty}\binom{n}{k}\frac{1}{n^k} = \lim_{n\to\infty}\frac{n}{n}\cdot\frac{n-1}{n}\cdots\frac{n-k+1}{n}\cdot\frac{1}{k!} = \frac{1}{k!}$$

$$\lim_{n\to\infty}(1-\frac{\lambda t}{n})^n = e^{-\lambda t}$$

$$\lim_{n\to\infty}(1-\frac{\lambda t}{n})^{-k} = 1$$

- $\lim_{n\to\infty} P(N_t = k) = \lim_{n\to\infty}\binom{n}{k}\frac{1}{n^k}\cdot(\lambda t)^k\cdot(1-\frac{\lambda t}{n})^n\cdot(1-\frac{\lambda t}{n})^{-k} = \frac{(\lambda t)^k}{k!}e^{-\lambda t}$

Poisson distribution

- $\lim_{n\to\infty} P(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$ is indeed a probability distribution on the numbers $0, 1, 2, \cdots$
  - Since $e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} = e^{-\lambda t} e^{\lambda t} = 1$

Poisson distribution

- $\lim_{n\to\infty} P(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$ is indeed a probability distribution on the numbers $0, 1, 2, \cdots$
  - Since $e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} = e^{-\lambda t} e^{\lambda t} = 1$

Definition

A discrete random variable X has a Poisson distribution with parameter μ, where $\mu > 0$ if its probability mass function p is given by

$$p(k) = P(X = k) = \frac{\mu^k}{k!} e^{-\mu} \quad \text{for} \quad k = 0, 1, 2, \cdots$$

Poisson Distribution

- The Poisson distribution can be viewed as an approximation of the binomial distribution.
- This is useful as the Poisson PMF is much easier to compute than the binomial.
- Thus we have

  Theorem: Let $X \sim Binomial(n, p = \frac{\mu}{n})$, where $\mu > 0$ is fixed. Then for any $k \in \{0, 1, 2, ...\}$, we have

$$\lim_{n\to\infty} p(k) = P(X = k) = \frac{e^{-\mu}\mu^k}{k!}.$$

Poisson Distribution

- Very widely used probability distribution.
- Used in counting the occurrences of certain events in an interval of time or space.
- Suppose we are counting the number of customers who visit a certain store from 1pm to 2pm.
- Based on data from previous days, we know that on average $\mu = 15$ customers visit the store.
- We can model the random variable X showing the number of customers as Poisson random variable with parameter $\mu = 15$.

Poisson Distribution

- Example: The number of emails that I get in a weekday can be modeled by a Poisson distribution with an average of 0.2 emails per minute.
    - What is the probability that I get no emails in an interval of length 5 minutes?
    - What is the probability that I get more than 3 emails in an interval of length 10 minutes?

Poisson Distribution

- The PMF of a *Poisson* random variable with $\mu = 1$

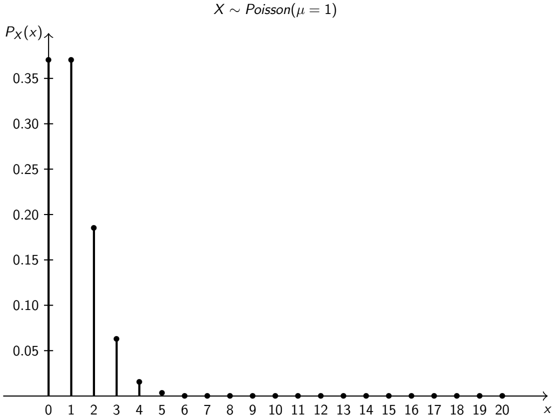


Figure: PMF of a *Poisson*(1) random variable

Poisson Distribution

- The PMF of a *Poisson* random variable with $\mu = 10$

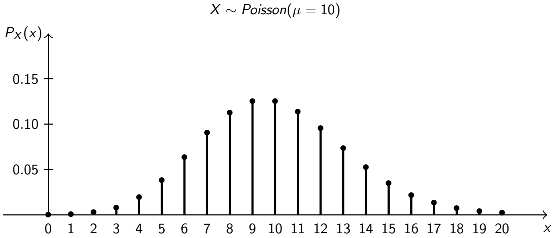


Figure: PMF of a *Poisson*(10) random variable

Poisson distribution: more examples

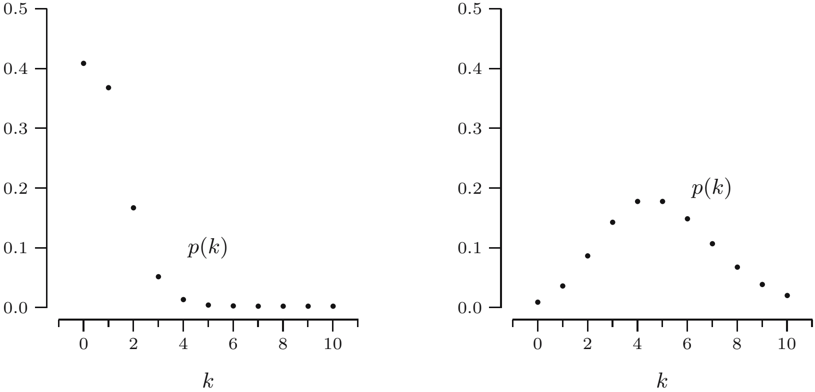


Figure: The probability mass functions of the Pois(0.9) and the Pois(5) distributions.

Expectation & Variance

- Expectation: the limiting Poisson distribution will have expectation λt
  - since $E[Nt] = \lambda t$ for all n
- Variance: $\lim_{n \to \infty} Var(N_t) = \lim_{n \to \infty} n \cdot \frac{\lambda t}{n} \cdot (1 - \frac{\lambda t}{n}) = \lambda t$
  - since N_t has a $Bin(n, \lambda t)$ distribution

Expectation & Variance

- Expectation: the limiting Poisson distribution will have expectation λt
  - since $E[Nt] = \lambda t$ for all n
- Variance: $\lim_{n\to\infty} Var(N_t) = \lim_{n\to\infty} n \cdot \frac{\lambda t}{n} \cdot (1 - \frac{\lambda t}{n}) = \lambda t$
  - since N_t has a $Bin(n, \lambda t)$ distribution

The expectation and variance of a Poisson distribution.

Let X have a Poisson distribution with parameter μ; then

$$E[X] = \mu \text{ and } Var(X) = \mu.$$