Poisson Processes

Poisson and Gamma

Radu Trîmbițaș

UBB

1st Semester 2010-2011

The Poisson Process I

- The *Poisson process* having rate $\lambda > 0$ is a collection $\{N(t) : t \geq 0\}$ of random variables, where $N(t)$ is the number of events that occur in the time interval $[0, t]$, which fulfill the following conditions:
  (a) $N(0) = 0$
  (b) The number of events occuring in disjoint time intervals are independent.
  (c) The distribution of the number of events that occur in a given interval depends only on the length of the interval and not on its location.
  (d) $\lim_{h\to 0} \frac{P(N(h)=1)}{h} = \lambda$.
  (e) $\lim_{h\to 0} \frac{P(N(h)\geq 2)}{h} = 0$.
- Condition (b), the *independent increment* assumption, states that $N(t))$ is independent of $N(t+s) - N(t)$.
- Condition (c), the *stationary increment* assumption, states that the probability distribution of $N(t+s) - N(t)$ is the same for all values of t.

The Poisson Process II

- Conditions (d) and (e) states that in a small interval of length h, the probability of one event occuring is approximately λh, whereas the probabilty of two or more is approximately 0.

Theorem

Number of events occuring in an interval of length t is a Poisson r.v. with mean λt.

Proof.

We break the interval $[0, t]$ into n nonoverlapping subintervals of length t/n. Consider the number of these intervals that contain an event. Conditions (b) and (c) imply each interval contains an event with the same probability $\lambda t/n \Longrightarrow N(t)$ is a binomial r.v with parameters n and $p \approx \lambda t/n$. When $n \to \infty$, binomial converges to Poisson with mean λt. (e) implies $P(2 \textit{ or more events}) \to 0$, so $N(t)$ is a Poisson r.v with mean λt. □

The Poisson Process II

X_1 -the time of first event; ... X_n the elapsed time time between the $(n-1)$st and the nth event

$\{X_n : n = 1, 2, \dots\}$ the *sequence of interarrival times*

Theorem

The interarrival times X_1, X_2, ... are i.i.d. exponential variables with parameters λ.

Proof.

$P(X_1 > t) = P(N(t) = 0) = e^{-\lambda t}$ ($(X_1 > t)$ occurs if no events of the Poisson process occur in $[0, t]$) X_1 is $Exp(\lambda)$.

$$\begin{aligned} P(X_2 > t | X_1 = s) &= P(0 \text{ events in}(s, s+t] | X_1 = s)) \\ &= P(0 \text{ events in}(s, s+t]) = e^{-\lambda t} \end{aligned}$$

analogously for $X_3, X_4, \dots, X_n \Longrightarrow X_n$ is $Exp(\lambda)$. $\square$

Applications

- A Poisson process, named after the French mathematician Siméon-Denis Poisson (1781–1840), is a stochastic process in which events occur continuously and independently of one another (the word event used here is not an instance of the concept of event frequently used in probability theory). Examples that are well-modeled as Poisson processes include the radioactive decay of atoms, telephone calls arriving at a switchboard, page view requests to a website, and rainfall.

- X has a gamma distribution with parameters $a > 0$ and λ if its pdf is

$$f(x|a,\lambda) = \begin{cases} \dfrac{\lambda^a x^{a-1} e^{-\lambda x}}{\Gamma(a)}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

  where $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$ is the Euler's gamma function
- alternative parametrization $\lambda = \frac{1}{\theta}$(e.g. as in MATLAB)

$$f(x|a,\lambda) = \begin{cases} \dfrac{x^{a-1} e^{-\frac{x}{\theta}}}{\theta^a \Gamma(a)}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Gamma distribution II

- mean, variance

$$E(X) = \frac{a}{\lambda} = a\theta$$
$$V(X) = \frac{a}{\lambda^2} = a\theta^2$$

- particular cases:
  - $a = 1$ exponential,
  - $a = \nu/2$, $\lambda = 1/2$ chi-square with ν degres of freedom,
  - $a \in \mathbb{N}^*$ Erlang
- Summation: If X_i has a $\Gamma(a_i, \lambda)$ distribution for $i = 1, 2, \ldots, n$, i.r.v., then

$$\sum_{i=1}^{n} X_i \sim \Gamma\left(\sum_{i=1}^{n} a_i, \lambda\right)$$

The gamma distribution exhibits infinite divisibility.

Gamma distribution III

- In particular, if X_i is $Exp(\lambda)$ the sum is Erlang of n and λ
- Erlang distribution could arise in the following context: consider k servers in series to complete the service of a customer; an additional customer cannot enter the first station until the customer in process has negotiated all the stations; each station has an exponential service time with parameter λ.

Applications

1. Plot gamma pdfs and cdfs for $a = 1, 2, 4$ and $\lambda = 1$.
2. Four-week summer rainfall totals in a section of the midwest United States have approximately a gamma distribution with $a = 1.6$ and $\theta = 2.0$. Find the probability to have an amount of rainfall between 3 and 57. What is the median of the rainfall?
3. Annual incomes for heads of household in a section of a city have approximately a gamma distribution with $a = 1000$ and $\theta = 20$. Find the mean and the variance of these incomes. Would you expect to find many incomes in excess of \$40,000 in this section of the city?
4. The gamma distribution is also used to model errors in multi-level Poisson regression models, because the combination of the Poisson distribution and a gamma distribution is a negative binomial distribution.

- $N(t)$ the number of events that occur by time t, $\{N(t), t \geq 0\}$ is a *nonhomogeneous (nonstationary) Poisson* process with *intensity function* $\lambda(t), t \geq 0$ if
  (a) $N(0) = 0$
  (b) The number of events occuring in disjoint time intervals are independent.
  (c) $\lim_{h\to 0} \frac{P(N(h)=1)}{h} = \lambda.$
  (d) $\lim_{h\to 0} \frac{P(N(h)\geq 2)}{h} = \lambda.$
- the mean value function is defined by

$$m(t) = \int_0^t \lambda(s)ds, \qquad t \geq 0.$$

- **Proposition**. $N(t+s) - N(t)$ is a Poisson r.v. with mean $m(t+s) - m(t)$.

The Nonhomogeneous (Nonstationary) Poisson Process II

- $\lambda(t)$ - intensity at time t; it indicates how likely it is that an event will occur around the time t. If $\lambda(t) \equiv \lambda$ (constant) we obtain an usual Poisson process
- **Proposition**. Suppose that the events are occuring according to a Poisson proces with rate λ, and suppose that, independently of anything that came before, an event that occurs at time t is counted with probbility $p(t)$. Then the process of counted events constitutes an nonhomogeneous Poisson process with intensity function $\lambda(t) = \lambda p(t)$.
- **Proof**. Conditions (a), (b), (c) holds for all (not just the counted) events. Check (c)

$$
\begin{aligned}
&P(1 \; event \; in \; [t, t+h]) \\
&= P(1 \; event \; and \; counted) + P(2 \text{ or more events, 1 counted}) \\
&\approx \lambda h p(t).
\end{aligned}
$$

Conditional Expectation I

- X, Y d.r.v., $E[X|Y=y]$, the *conditional expectation* of X, given that $Y=y$, is defined by

$$E[X|Y=y] = \sum_x xP(X=x|Y=y) = \frac{\sum_x xP(X=x, Y=y)}{P(Y=y)}$$

- X, Y c.r.v., $E[X|Y=y]$ is defined by

$$E[X|Y=y] = \frac{\int_{\mathbb{R}} xf(x,y)dx}{\int_{\mathbb{R}} f(x,y)dx}$$

- like an ordinary expectation, but the distribution is conditional
- **Proposition**.

$$E[E[X|Y]] = E(X) \tag{1}$$

- If Y discrete (1) states

$$E(X) = \sum_{y} E\left[X|Y = y\right] P\left(Y = y\right)$$

- If Y continuous with density g, (1) states

$$E(X) = \int_{\mathbb{R}} E\left[X|Y = y\right] g(y) dy$$

Conditional Expectation III

- Proof for discrete case

$$
\begin{aligned}
\sum_y E\left[X|Y=y\right] P\left(Y=y\right) &= \sum_y \sum_x xP\left(X=x|Y=y\right) P(Y=y) \\
&= \sum_y \sum_x xP\left(X=x, Y=y\right) \\
&= \sum_y x \sum_x P\left(X=x, Y=y\right) \\
&= \sum_y xP\left(X=x\right) \\
&= E(X)
\end{aligned}
$$

Conditional Variance

- Conditional variance of X, given the value of Y

$$V(X|Y) = E\left[(X - E[X|Y])^2 \,|Y\right]$$

- We have

$$V(X|Y) = E\left[X^2|Y\right] - (E[X|Y])^2$$

- Take expectation

$$\begin{aligned} E(V(X|Y)) &= E\left(E\left[X^2|Y\right]\right) - E\left((E[X|Y])^2\right) && (2) \\ &= E[X^2] - E\left((E[X|Y])^2\right) && (3) \end{aligned}$$

- Using (1)

$$V(E[X|Y]) = E\left((E[X|Y])^2\right) - E(X)^2 \qquad (4)$$

- (3)+(4) yields

$$V(X) = E[V(X|Y)] + V(E[X|Y])$$