Probability in Computing

LECTURE 7: CONTINUOUS DISTRIBUTIONS AND POISSON PROCESS

Agenda

- Continuous random variables.
  - Uniform distribution
  - Exponential distribution
- Poisson process
- Queuing theory: just a glance

Continuous Random Variables

- Consider a roulette wheel which has circumference 1. We spin the wheel, and when it stops, the outcome is the clockwise distance X from the "0" mark to the arrow.
- Sample space Ω consists of all real numbers in [0, 1).
- Assume that any point on the circumference is equally likely to face the arrow when the wheel stops. What's the probability of a given outcome x?
- Note: In an infinite sample space there maybe possible events that have probability = 0.
- Recall that the distribution function $F(x) = Pr(X \le x)$. and $f(x) = F'(x)$ then $f(x)$ is called the density function of $F(x)$.

Continuous Random Variables

- f(x)dx = probability of the infinitesimal interval [x, x + dx).
- Pr(a ≤X<b) = $\int_a^b f(x)dx$
- E[X] = $\int_{-\infty}^{\infty} xf(x)dx$
- E[g(X)] = $\int_{-\infty}^{\infty} g(x)f(x)dx$

Joint Distributions

- Def: The joint distribution function of X and Y is $F(x,y) = \Pr(X \le x, Y \le y)$. = $\int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)\,du\,dv$ where f is the joint density function.
- $f(x, y) = \dfrac{\partial^2}{\partial x \partial y} F(x, y)$
- Marginal distribution functions $F_X(x)=\Pr(X \le x)$ and $F_Y(y)=\Pr(Y \le y)$.
- Example: $F(x,y) = 1- e^{-ax} - e^{-by} + e^{-(ax+by)}$, $x, y >= 0$.
  - $F_X(x)=F(x,\infty) = 1-e^{-ax}$.
  - $F_Y(y)=1-e^{-by}$.
  - Since $F_X(x)F_Y(y) = F(x, y) \rightarrow$ X and Y are independent.

Conditional Probability

- What is Pr(X≤3|Y=4)? – Both numerator and denominator = 0.

- Rewriting Pr(X≤3|Y=4)? = $\lim_{\delta \to 0} \Pr(X \leq 3 \mid 4 \leq Y \leq 4+\delta)$

- Pr(X≤x|Y=y) = $\int_{u=-\infty}^{x} \frac{f(u,y)}{f_Y(y)} du$

Uniform Distribution

- Used to model random variables that tend to occur "evenly" over a range of values
- Probability of any interval of values proportional to its width
- Used to generate (simulate) random variables from virtually any distribution
- Used as "non-informative prior" in many Bayesian analyses

$$f(y) = \begin{cases} \dfrac{1}{b-a} & a \leq y \leq b \\ 0 & \text{elsewhere} \end{cases}$$

$$F(y) = \begin{cases} 0 & y < a \\ \dfrac{y-a}{b-a} & a \leq y \leq b \\ 1 & y > b \end{cases}$$

Uniform Distribution - expectation

$$E(Y)=\int_a^b y\left(\frac{1}{b-a}\right)dy=\left(\frac{1}{b-a}\right)\frac{y^2}{2}\bigg|_a^b=\frac{b^2-a^2}{2(b-a)}=\frac{(b-a)(b+a)}{2(b-a)}=\frac{b+a}{2}$$

$$E\left(Y^2\right)=\int_a^b y^2\left(\frac{1}{b-a}\right)dy=\left(\frac{1}{b-a}\right)\frac{y^3}{3}\bigg|_a^b=\frac{b^3-a^3}{3(b-a)}=\frac{(b-a)(a^2+b^2+ab)}{3(b-a)}=$$

$$=\frac{(a^2+b^2+ab)}{3}$$

$$\Rightarrow V(Y)=E\left(Y^2\right)-[E(Y)]^2=\frac{(a^2+b^2+ab)}{3}-\left[\frac{b+a}{2}\right]^2=$$

$$=\frac{4(a^2+b^2+ab)-3(b^2+a^2+2ab)}{12}=\frac{a^2+b^2-2ab}{12}=\frac{(b-a)^2}{12}$$

$$\Rightarrow \sigma=\sqrt{\frac{(b-a)^2}{12}}=\frac{b-a}{\sqrt{12}}\approx 0.2887(b-a)$$

Additional Properties

- Lemma 1: Let X be a uniform random variable on [a, b]. Then, for $c \le d$, $Pr(X \le c | X \le d) = (c-a)/(d-a)$.
- That is, conditioned on the fact that $X \le d$, X is uniform on [a, d].

- Lemma 2: Let $X_1, X_2, \ldots, X_n$ be independent uniform random variables over [0, 1], Let $Y_1, Y_2, \ldots, Y_n$ be the same values as $X_1, X_2, \ldots, X_n$ in increasing sorted order. Then $E[Y_k] = k/(n+1)$.

Work in class

- Show lemma 2 for $n=2$, $k=1$
  - i.e. show that $E[\min\{X_1, X_2\}] = 1/3$
  - Hint: try a discretized version first
- Extend what you have done for arbitrary n
- Now for arbitrary k

Exponential Distribution

- Right-Skewed distribution with maximum at y=0
- Random variable can only take on positive values
- Used to model inter-arrival times/distances for a Poisson process

$$F(x) = \begin{cases} 1 - e^{-\theta x} & \textit{for } x \geq 0, \\ 0 & \textit{otherwise.} \end{cases}$$

$$f(x) = \theta e^{-\theta x} \quad \text{for } x \geq 0.$$

$$\mathbf{E}[X] = \int_0^\infty t\theta e^{-\theta t}\, dt = \frac{1}{\theta},$$

$$\mathbf{E}[X^2] = \int_0^\infty t^2\theta e^{-\theta t}\, dt = \frac{2}{\theta^2}.$$

$$\mathbf{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \frac{1}{\theta^2}.$$

Additional Properties

- Lemma 3: $Pr(X>s+t|X>t) = Pr(X>s)$
  - The exponential distribution is the only continuous memory-less distribution: time until the 1^{st} event in a memoryless continuous time stochastic process.
  - Similarly, geometric is the only discrete memoryless distribution: time until 1st success in a sequence of independent identical Bernoulli trials.
- Reliability: Amount of time a component has been in service has no effect on the amount of time until it fails
- Inter-event times: Amount of time since the last event contains no information about the amount of time until the next event
- Service times: Amount of remaining service time is independent of the amount of service time elapsed so far

Additional Properties

- The minimum of several exponential random variables also exhibits some interesting properties.

Lemma 8.5: *If $X_1, X_2, \ldots, X_n$ are independent exponentially distributed random variables with parameters $\theta_1, \theta_2, \ldots, \theta_n$, respectively, then $\min(X_1, X_2, \ldots, X_n)$ is exponentially distributed with parameter $\sum_{i=1}^{n} \theta_i$ and*

$$\Pr(\min(X_1, X_2, \ldots, X_n) = X_i) = \frac{\theta_i}{\sum_{i=1}^{n} \theta_i}.$$

Example: An airline ticket counter has n service agents, where the time that agent I takes per customer has an exponential distribution with parameter θ. You stand at the head of the line at time To, and all of the n agents are busy. What is the average time you wait for an agent?

- Because service time is exponentially distributed → the remaining time for each customer is still exponentially distributed.
- Apply Lemma 8.5, time until 1st agent is free is exponentially distributed with parameter $\sum\theta_i$. → expected time = $1 / \sum\theta_i$.
- The j^{th} agent will become free first with prob. $\theta_j / \sum\theta_i$.

Work in Class

- Understanding the concepts of service time and related
- Show the lemma for n=2

Counting Process

A stochastic process $\{N(t), t \geq 0\}$ is a *counting process* if $N(t)$ represents the total number of events that have occurred in $[0, t]$

Then $\{ N(t), t \geq 0 \}$ must satisfy:

a) $N(t) \geq 0$
b) $N(t)$ is an integer for all t
c) If $s < t$, then $N(s) \leq N(t)$ and
d) For $s < t$, $N(t) - N(s)$ is the number of events that occur in the interval $(s, t]$.

Stationary & Independent Increments

independent increments

A counting process has *independent increments* if for any $0 \leq s \leq t \leq u \leq v$,

$$N(t) - N(s) \text{ is independent of } N(v) - N(u)$$

i.e., the numbers of events that occur in non-overlapping intervals are independent r.v.s

stationary increments

A counting process has *stationary increments* if the distribution if, for any $s < t$, the distribution of

$$N(t) - N(s)$$

depends only on the length of the time interval, $t - s$.

Work in Class

- Create your own practical examples,
  - i.e. in practice can you see situations that can be simulated by counting processes?
  - and ones with each of the two properties?

Poisson Process Definition 1

A counting process $\{N(t), t \geq 0\}$ is a *Poisson process with rate l*, $l > 0$, if

$N(0) = 0$

The process has independent increments

The number of events in any interval of length t follows a Poisson distribution with mean λt

$$\Pr\{ N(t+s) - N(s) = n \} = (\lambda t)^n e^{-\lambda t}/n!\ ,\ \ n = 0, 1, \ldots$$

Where λ is arrival rate and t is length of the interval

Notice, it has stationary increments

Poisson Process Definition 2

Definition 8.4: *A* Poisson process *with parameter (or rate)* λ *is a stochastic counting process* $\{N(t), t \geq 0\}$ *such that the following statements hold.*

1. $N(0) = 0$.
2. *The process has independent and stationary increments. That is, for any* $t, s > 0$, *the distribution of* $N(t+s) - N(s)$ *is identical to the distribution of* $N(t)$, *and for any two disjoint intervals* $[t_1, t_2]$ *and* $[t_3, t_4]$, *the distribution of* $N(t_2) - N(t_1)$ *is independent of the distribution of* $N(t_4) - N(t_3)$.
3. $\lim_{t \to 0} \Pr(N(t) = 1)/t = \lambda$. *That is, the probability of a single event in a short interval* t *tends to* λt.
4. $\lim_{t \to 0} \Pr(N(t) \geq 2)/t = 0$. *That is, the probability of more than one event in a short interval* t *tends to zero.*

Theorem 8.7: *Let* $\{N(t) \mid t \geq 0\}$ *be a Poisson process with parameter* λ. *For any* $t, s \geq 0$ *and any integer* $n \geq 0$,

$$P_n(t) = \Pr(N(t+s) - N(s) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.$$

Work in Class

- Can you see why Poison distribution come out here?
  - Hint: relate to discretized versions which could be simulated by using binomial distribution

Limit of the Binomial Distribution

We have shown that, when throwing m balls randomly into b bins, the probability p_r that a bin has r balls is approximately the Poisson distribution with mean m/b. In general, the Poisson distribution is the limit distribution of the binomial distribution with parameters n and p, when n is large and p is small. More precisely, we have the following limit result.

Theorem 5.5: *Let X_n be a binomial random variable with parameters n and p, where p is a function of n and $\lim_{n\to\infty} np = \lambda$ is a constant that is independent of n. Then, for any fixed k,*

$$\lim_{n\to\infty} \Pr(X_n = k) = \frac{e^{-\lambda}\lambda^k}{k!}.$$

This theorem directly applies to the balls-and-bins scenario. Consider the situation where there are m balls and b bins, where m is a function of b and $\lim_{n\to\infty} m/b = \lambda$. Let X_n be the number of balls in a specific bin. Then X_n is a binomial random variable with parameters m and $1/b$. Theorem 5.5 thus applies and says that

$$\lim_{n\to\infty} \Pr(X_n = r) = \frac{e^{-m/n}(m/n)^r}{r!},$$

Inter-Arrival and Waiting Times

The times between arrivals $T1$, $T2$, … are independent exponential random variables with mean $1/\lambda$:

$$P(T1>t) = P(N(t) = 0) = e^{-\lambda t}$$

The (total) waiting time until the n^{th} event has a gamma distribution

$$S_n = \sum_{i=1}^{n} T_i$$

An Example

Suppose that you arrive at a single teller bank to find five other customers in the bank. One being served and the other four waiting in line. You join the end of the line. If the service time are all exponential with rate 5 minutes.

What is the prob. that you will be served in 10 minutes ?

What is the prob. that you will be served in 20 minutes ?

What is the expected waiting time before you are served?

Queuing Theory

- Many applications:
  - In OS: Schedulers hold tasks in queue until required resources are available.
  - In parallel/distributed processing: threads can queue for a critical section that allows access to only one thread at a time.
  - In networks: packets are queued while waiting to be forwarded by a router.

- We are going to:
  - Analyze one of the most basic queue model.
  - It uses Poisson process to model how customers arrive
  - Exponentially distributed r.v. to model the time required for service.

Notations

- Typical performance characteristics of queuing models are:

L : Ave. number of customers in the system

L_Q : Ave. number of customers waiting in queue

W : Ave. time customer spends in the system

W_Q: Ave. time customer spends waiting in the queue

Queue notation

M/M/k/c

Arrival process
M = Markovian
GI = General

Departure process
(Service time distribution)
M = Markovian
G = General

Number of servers

Capacity of the queue
If nothing is specified,
we assume infinite capacity

M/M/1 queue

Special Birth - Death process, where arrival rate is λ and service rate is μ.

$$P_0 = 1 - \frac{\lambda}{\mu}, \qquad P_n = \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right), n \geq 1$$

$$L = \sum_{n=0}^{\infty} nP_n = \frac{\lambda}{\mu - \lambda}, \qquad W = \frac{L}{\lambda} = \frac{1}{\mu - \lambda}$$

$$W_Q = W - E[S] = \frac{\lambda}{\mu(\mu - \lambda)} \qquad L_Q = \lambda W_Q = \frac{\lambda^2}{\mu(\mu - \lambda)}$$