STA286 Lecture 15

Neil Montgomery

Last edited: 2017-02-28 09:09

“binomial process”

Let’s put a time scale on a Bernoulli process. One trial will now happen in every unit of time Δ (called a “frame”), and successes happen now with a “rate” λ per unit of time.

“binomial process”

Let’s put a time scale on a Bernoulli process. One trial will now happen in every unit of time Δ (called a “frame”), and successes happen now with a “rate” λ per unit of time.

The result of this re-jigging is a sequence $X_1, X_2, X_3, \ldots$ of random variables that are independent with Bernoulli($\lambda\Delta$) distribution.

“binomial process”

Let’s put a time scale on a Bernoulli process. One trial will now happen in every unit of time Δ (called a “frame”), and successes happen now with a “rate” λ per unit of time.

The result of this re-jigging is a sequence $X_1, X_2, X_3, \ldots$ of random variables that are independent with Bernoulli($\lambda\Delta$) distribution.

Consider the cumulative sums $X(t)$ of this Bernoulli process:

$$
\begin{aligned}
X(1) &= X_1 \\
X(2) &= X_1 + X_2 \\
\vdots \quad & \quad \vdots \\
X(n) &= \sum_{i=1}^{n} X_i \\
\vdots \quad & \quad \vdots
\end{aligned}
$$

“binomial process”

$X(t)$ is non-negative, integer valued, non-decreasing sequence of random variables. Such a sequence is called a *counting process*.

“binomial process”

$X(t)$ is non-negative, integer valued, non-decreasing sequence of random variables. Such a sequence is called a *counting process*.

This particular counting process is called a *binomial process*. It counts the number of 1s in the first n frames—equivalent to the number of 1s from time $t = 0$ to time $t = n\Delta$.

(Assume the Bernoulli trial occurs at the start of the frame.)

“binomial process”

$X(t)$ is non-negative, integer valued, non-decreasing sequence of random variables. Such a sequence is called a *counting process*.

This particular counting process is called a *binomial process*. It counts the number of 1s in the first n frames—equivalent to the number of 1s from time $t = 0$ to time $t = n\Delta$.

(Assume the Bernoulli trial occurs at the start of the frame.)

$X(t)$ has these properties:

- $X(0) = 0$

“binomial process”

$X(t)$ is non-negative, integer valued, non-decreasing sequence of random variables. Such a sequence is called a *counting process*.

This particular counting process is called a *binomial process*. It counts the number of 1s in the first n frames—equivalent to the number of 1s from time $t = 0$ to time $t = n\Delta$.

(Assume the Bernoulli trial occurs at the start of the frame.)

$X(t)$ has these properties:

- $X(0) = 0$
- $X(t) \sim \text{Binomial}\left(n = \frac{t}{\Delta}, p = \lambda\Delta\right)$.

“binomial process”

$X(t)$ is non-negative, integer valued, non-decreasing sequence of random variables. Such a sequence is called a *counting process*.

This particular counting process is called a *binomial process*. It counts the number of 1s in the first n frames—equivalent to the number of 1s from time $t = 0$ to time $t = n\Delta$.

(Assume the Bernoulli trial occurs at the start of the frame.)

$X(t)$ has these properties:

- $X(0) = 0$
- $X(t) \sim \text{Binomial}\left(n = \frac{t}{\Delta}, p = \lambda\Delta\right)$.
- $X(t)$ increments by 0 or 1 per frame.

“binomial process”

$X(t)$ is non-negative, integer valued, non-decreasing sequence of random variables. Such a sequence is called a *counting process*.

This particular counting process is called a *binomial process*. It counts the number of 1s in the first n frames—equivalent to the number of 1s from time $t = 0$ to time $t = n\Delta$.

(Assume the Bernoulli trial occurs at the start of the frame.)

$X(t)$ has these properties:

- $X(0) = 0$
- $X(t) \sim \text{Binomial}\left(n = \frac{t}{\Delta}, p = \lambda\Delta\right)$.
- $X(t)$ increments by 0 or 1 per frame.
- Number of frames between increments is Geometric(p).

“binomial process”

$X(t)$ is non-negative, integer valued, non-decreasing sequence of random variables. Such a sequence is called a *counting process*.

This particular counting process is called a *binomial process*. It counts the number of 1s in the first n frames—equivalent to the number of 1s from time $t = 0$ to time $t = n\Delta$.

(Assume the Bernoulli trial occurs at the start of the frame.)

$X(t)$ has these properties:

- $X(0) = 0$
- $X(t) \sim \text{Binomial}\left(n = \frac{t}{\Delta}, p = \lambda\Delta\right)$.
- $X(t)$ increments by 0 or 1 per frame.
- Number of frames between increments is Geometric(p).
- Given times $s_1 < t_1 < s_2 < t_2$, the differences $X(t_1) - X(s_1)$ and $X(t_2) - X(s_2)$ are *independent* random variables.

“binomial process”

$X(t)$ is non-negative, integer valued, non-decreasing sequence of random variables. Such a sequence is called a *counting process*.

This particular counting process is called a *binomial process*. It counts the number of 1s in the first n frames—equivalent to the number of 1s from time $t = 0$ to time $t = n\Delta$.

(Assume the Bernoulli trial occurs at the start of the frame.)

$X(t)$ has these properties:

- $X(0) = 0$
- $X(t) \sim \text{Binomial}\left(n = \frac{t}{\Delta}, p = \lambda\Delta\right)$.
- $X(t)$ increments by 0 or 1 per frame.
- Number of frames between increments is Geometric(p).
- Given times $s_1 < t_1 < s_2 < t_2$, the differences $X(t_1) - X(s_1)$ and $X(t_2) - X(s_2)$ are *independent* random variables.
- $E(X(t)) = np = \frac{t}{\Delta}\lambda\Delta = \lambda t$

pass to the limit

Fix a time t. Keep $E(X(t)) = np = \lambda t$ constant, so that $p = \lambda t/n$.

The goal is to find the limit of $P(X(t) = k)$ as $n \to \infty$ for any $k \geqslant 0$.

For any fixed n, $X(t) \sim \text{Binomial}\left(n, \frac{\lambda t}{n}\right)$

pass to the limit

$$\lim_{n\to\infty} P(X(t)=k) = \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda t}{n}\right)^{k} \left(1-\frac{\lambda t}{n}\right)^{n-k}$$

pass to the limit

$$\lim_{n\to\infty} P(X(t)=k) = \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda t}{n}\right)^k \left(1-\frac{\lambda t}{n}\right)^{n-k}$$

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

pass to the limit

$$
\begin{aligned}
\lim_{n\to\infty} P(X(t)=k) &= \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda t}{n}\right)^k \left(1-\frac{\lambda t}{n}\right)^{n-k} \\
&= \lim_{n\to\infty} \frac{n!}{k!(n-k)!} \left(\frac{\lambda t}{n}\right)^k \left(1-\frac{\lambda t}{n}\right)^{n-k} \\
&= \frac{(\lambda t)^k}{k!} \lim_{n\to\infty} \frac{n!}{(n-k)!} \frac{1}{n^k} \left(1-\frac{\lambda t}{n}\right)^n \left(1-\frac{\lambda t}{n}\right)^{-k}
\end{aligned}
$$

pass to the limit

$$
\begin{aligned}
\lim_{n\to\infty} P(X(t)=k) &= \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda t}{n}\right)^k \left(1-\frac{\lambda t}{n}\right)^{n-k} \\
&= \lim_{n\to\infty} \frac{n!}{k!(n-k)!} \left(\frac{\lambda t}{n}\right)^k \left(1-\frac{\lambda t}{n}\right)^{n-k} \\
&= \frac{(\lambda t)^k}{k!} \lim_{n\to\infty} \frac{n!}{(n-k)!} \frac{1}{n^k} \left(1-\frac{\lambda t}{n}\right)^n \left(1-\frac{\lambda t}{n}\right)^{-k} \\
&= \frac{(\lambda t)^k}{k!} \lim_{n\to\infty} \underbrace{\frac{n\cdot(n-1)\cdot(n-2)\cdots(n-k+1)}{n\cdot n\cdot n\cdots n}}_{k\text{terms}} \left(1-\frac{\lambda t}{n}\right)^n \left(1-\frac{\lambda t}{n}\right)^{-k}
\end{aligned}
$$

pass to the limit

$$
\begin{aligned}
\lim_{n\to\infty} P(X(t)=k) &= \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda t}{n}\right)^k \left(1-\frac{\lambda t}{n}\right)^{n-k} \\
&= \lim_{n\to\infty} \frac{n!}{k!(n-k)!} \left(\frac{\lambda t}{n}\right)^k \left(1-\frac{\lambda t}{n}\right)^{n-k} \\
&= \frac{(\lambda t)^k}{k!} \lim_{n\to\infty} \frac{n!}{(n-k)!} \frac{1}{n^k} \left(1-\frac{\lambda t}{n}\right)^{n} \left(1-\frac{\lambda t}{n}\right)^{-k} \\
&= \frac{(\lambda t)^k}{k!} \lim_{n\to\infty} \underbrace{\frac{n\cdot(n-1)\cdot(n-2)\cdots(n-k+1)}{n\cdot n\cdot n\cdots n}}_{k\text{terms}} \left(1-\frac{\lambda t}{n}\right)^{n} \left(1-\frac{\lambda t}{n}\right)^{-k} \\
&= \frac{(\lambda t)^k}{k!} e^{-\lambda t}
\end{aligned}
$$

Poisson process/distribution

Nothing to do with fish. Named after some French guy.

A Poisson process $N(t)$ is a counting process which counts the number of "events" that happen inside the time interval $[0, t]$, subject to the following:

- $N(0) = 0$

Poisson process/distribution

Nothing to do with fish. Named after some French guy.

A Poisson process $N(t)$ is a counting process which counts the number of "events" that happen inside the time interval $[0, t]$, subject to the following:

- $N(0) = 0$
- Given times $s_1 < t_1 < s_2 < t_2$, the differences $N(t_1) - N(s_1)$ and $N(t_2) - N(s_2)$ are *independent* random variables.

Poisson process/distribution

Nothing to do with fish. Named after some French guy.

A Poisson process $N(t)$ is a counting process which counts the number of "events" that happen inside the time interval $[0, t]$, subject to the following:

- $N(0) = 0$
- Given times $s_1 < t_1 < s_2 < t_2$, the differences $N(t_1) - N(s_1)$ and $N(t_2) - N(s_2)$ are *independent* random variables.
- $P(N(t) - N(s) = k) = \frac{(\lambda(t-s))^k}{k!} e^{-\lambda(t-s)}$ for $0 \leqslant s < t$

Poisson process/distribution

Nothing to do with fish. Named after some French guy.

A Poisson process $N(t)$ is a counting process which counts the number of "events" that happen inside the time interval $[0, t]$, subject to the following:

- $N(0) = 0$
- Given times $s_1 < t_1 < s_2 < t_2$, the differences $N(t_1) - N(s_1)$ and $N(t_2) - N(s_2)$ are *independent* random variables.
- $P(N(t) - N(s) = k) = \frac{(\lambda(t-s))^k}{k!} e^{-\lambda(t-s)}$ for $0 \leqslant s < t$

Poisson process/distribution

Nothing to do with fish. Named after some French guy.

A Poisson process $N(t)$ is a counting process which counts the number of "events" that happen inside the time interval $[0, t]$, subject to the following:

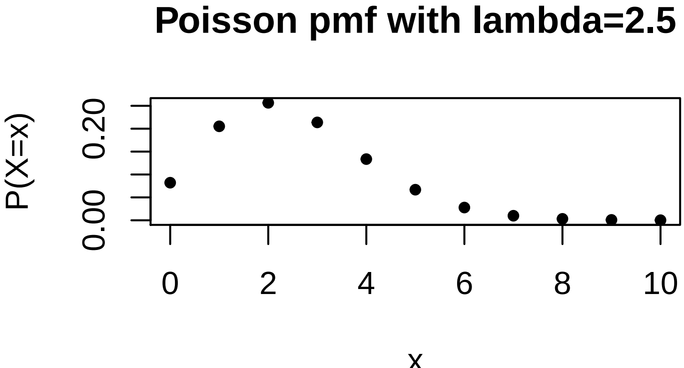
- $N(0) = 0$
- Given times $s_1 < t_1 < s_2 < t_2$, the differences $N(t_1) - N(s_1)$ and $N(t_2) - N(s_2)$ are *independent* random variables.
- $P(N(t) - N(s) = k) = \frac{(\lambda(t-s))^k}{k!} e^{-\lambda(t-s)}$ for $0 \leqslant s < t$

The Poisson process is a common model for events that happen "completely randomly" in time (to be further discussed)

Poisson distribution

There is also the closely related Poisson distribution, in which the t is not explicitly used. We say $X \sim \text{Poisson}(\lambda)$ if it has pmf:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{for } k \in \{0, 1, 2, \ldots\}$$



Poisson examples

Customers enter a store at a rate of 1 per minute. Find the probabilities that:

1. More than one will enter in the first minute.

Poisson examples

Customers enter a store at a rate of 1 per minute. Find the probabilities that:

1. More than one will enter in the first minute.
2. More than two will enter in the first two minutes.

Poisson examples

Customers enter a store at a rate of 1 per minute. Find the probabilities that:

1. More than one will enter in the first minute.
2. More than two will enter in the first two minutes.
3. More than one will enter in each of the first two minutes.

Poisson examples

Customers enter a store at a rate of 1 per minute. Find the probabilities that:

1. More than one will enter in the first minute.
2. More than two will enter in the first two minutes.
3. More than one will enter in each of the first two minutes.

Poisson examples

Customers enter a store at a rate of 1 per minute. Find the probabilities that:

1. More than one will enter in the first minute.
2. More than two will enter in the first two minutes.
3. More than one will enter in each of the first two minutes.

Why or why not might a Poisson process model be suitable here?

other Poisson distribution properties

Suppose $X \sim \text{Poisson}(\lambda)$. Then (using s as the meaningless "dummy" variable):

other Poisson distribution properties

Suppose $X \sim \text{Poisson}(\lambda)$. Then (using s as the meaningless "dummy" variable):

$$M_X(s) = E(e^{sX}) = \sum_{k=0}^{\infty} e^{sk} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^s \lambda)^k}{k!}$$

other Poisson distribution properties

Suppose $X \sim \text{Poisson}(\lambda)$. Then (using s as the meaningless "dummy" variable):

$$M_X(s) = E(e^{sX}) = \sum_{k=0}^{\infty} e^{sk} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^s \lambda)^k}{k!} = e^{-\lambda} e^{e^s \lambda} = e^{\lambda(e^s - 1)}$$

other Poisson distribution properties

Suppose $X \sim \text{Poisson}(\lambda)$. Then (using s as the meaningless "dummy" variable):

$$M_X(s) = E(e^{sX}) = \sum_{k=0}^{\infty} e^{sk} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^s \lambda)^k}{k!} = e^{-\lambda} e^{e^s \lambda} = e^{\lambda(e^s - 1)}$$

Then it's easy to show $E(X) = \lambda$ and $\text{Var}(X) = \lambda$.

Poisson - when did the event happen?

In a Poisson process we might have observed that $N(t) = 1$. (We will condition on this fact throughout.)

Poisson - when did the event happen?

In a Poisson process we might have observed that $N(t) = 1$. (We will condition on this fact throughout.)

When inside the interval $[0, t]$ did that single event occur?

Poisson - when did the event happen?

In a Poisson process we might have observed that $N(t) = 1$. (We will condition on this fact throughout.)

When inside the interval $[0, t]$ did that single event occur?

The time when the event happened is a continuous random variable, which we can call X. What is it's distribution?

Poisson - when did the event happen?

In a Poisson process we might have observed that $N(t) = 1$. (We will condition on this fact throughout.)

When inside the interval $[0, t]$ did that single event occur?

The time when the event happened is a continuous random variable, which we can call X. What is it's distribution?

We can derive its cumulative disribution function from first principles. We know $P(X \leqslant x | N(t) = 1) = 0$ for $x \leqslant 0$ and $P(X \leqslant x | N(t) = 1) = 1$ for $x > t$.

cdf of "the time when that one thing happened" - cont'

Between 0 and t we have:

$$P(X \leqslant x | N(t) = 1) = 1 - P(X > x | N(t) = 1)$$

cdf of "the time when that one thing happened" - cont'

Between 0 and t we have:

$$
\begin{aligned}
P(X \leqslant x | N(t) = 1) &= 1 - P(X > x | N(t) = 1) \\
&= 1 - P(N(x) = 0 | N(t) = 1)
\end{aligned}
$$

cdf of "the time when that one thing happened" - cont'

Between 0 and t we have:

$$
\begin{aligned}
P(X \leqslant x | N(t) = 1) &= 1 - P(X > x | N(t) = 1) \\
&= 1 - P(N(x) = 0 | N(t) = 1) \\
&= 1 - \frac{P(N(x) = 0, N(t) = 1)}{P(N(t) = 1)}
\end{aligned}
$$

cdf of "the time when that one thing happened" - cont'

Between 0 and t we have:

$$
\begin{aligned}
P(X \leqslant x | N(t) = 1) &= 1 - P(X > x | N(t) = 1) \\
&= 1 - P(N(x) = 0 | N(t) = 1) \\
&= 1 - \frac{P(N(x) = 0, N(t) = 1)}{P(N(t) = 1)} \\
&= 1 - \frac{P(N(x) = 0, N(t) - N(x) = 1)}{P(N(t) = 1)}
\end{aligned}
$$

cdf of "the time when that one thing happened" - cont'

Between 0 and t we have:

$$
\begin{aligned}
P(X \leqslant x | N(t) = 1) &= 1 - P(X > x | N(t) = 1) \\
&= 1 - P(N(x) = 0 | N(t) = 1) \\
&= 1 - \frac{P(N(x) = 0, N(t) = 1)}{P(N(t) = 1)} \\
&= 1 - \frac{P(N(x) = 0, N(t) - N(x) = 1)}{P(N(t) = 1)} \\
&= 1 - \frac{P(N(x) = 0)P(N(t) - N(x) = 1)}{P(N(t) = 1)}
\end{aligned}
$$

cdf of "the time when that one thing happened" - cont'

Between 0 and t we have:

$$
\begin{aligned}
P(X \leqslant x | N(t) = 1) &= 1 - P(X > x | N(t) = 1) \\
&= 1 - P(N(x) = 0 | N(t) = 1) \\
&= 1 - \frac{P(N(x) = 0, N(t) = 1)}{P(N(t) = 1)} \\
&= 1 - \frac{P(N(x) = 0, N(t) - N(x) = 1)}{P(N(t) = 1)} \\
&= 1 - \frac{P(N(x) = 0)P(N(t) - N(x) = 1)}{P(N(t) = 1)} \\
&= 1 - \frac{e^{-\lambda x}\lambda(t - x)e^{-\lambda(t-x)}}{\lambda t e^{-\lambda t}}
\end{aligned}
$$

cdf of “the time when that one thing happened” - cont’

Between 0 and t we have:

$$
\begin{aligned}
P(X \leqslant x | N(t) = 1) &= 1 - P(X > x | N(t) = 1) \\
&= 1 - P(N(x) = 0 | N(t) = 1) \\
&= 1 - \frac{P(N(x) = 0, N(t) = 1)}{P(N(t) = 1)} \\
&= 1 - \frac{P(N(x) = 0, N(t) - N(x) = 1)}{P(N(t) = 1)} \\
&= 1 - \frac{P(N(x) = 0)P(N(t) - N(x) = 1)}{P(N(t) = 1)} \\
&= 1 - \frac{e^{-\lambda x}\lambda(t - x)e^{-\lambda(t-x)}}{\lambda t e^{-\lambda t}} \\
&= 1 - \frac{t - x}{t} = \frac{x}{t}
\end{aligned}
$$

density of "the time when that one thing happened"

Putting it all together, the cdf of the time X when the event occurred given $N(t) = 1$ is:

$$F_X(x) = \begin{cases} 0 & : x \leqslant 0 \\ \frac{x}{t} & : 0 < x \leqslant t \\ 1 & : x > t \end{cases}$$

density of "the time when that one thing happened"

Putting it all together, the cdf of the time X when the event occurred given $N(t) = 1$ is:

$$F_X(x) = \begin{cases} 0 & : x \leqslant 0 \\ \frac{x}{t} & : 0 < x \leqslant t \\ 1 & : x > t \end{cases}$$

So the density is:

$$f_X(x) = \frac{d}{dx} F_X(x) = \begin{cases} \frac{1}{t} & : 0 < x \leqslant t \\ 0 & : \text{ otherwise} \end{cases}$$

density of "the time when that one thing happened"

Putting it all together, the cdf of the time X when the event occurred given $N(t) = 1$ is:

$$F_X(x) = \begin{cases} 0 & : x \leqslant 0 \\ \frac{x}{t} & : 0 < x \leqslant t \\ 1 & : x > t \end{cases}$$

So the density is:

$$f_X(x) = \frac{d}{dx} F_X(x) = \begin{cases} \frac{1}{t} & : 0 < x \leqslant t \\ 0 & : \text{ otherwise} \end{cases}$$

The density is flat between 0 and t. We call this a "uniform distribution" between 0 and t.

the uniform distributions

The random variable X, the result of "picking a real number at random between a and b", is modeled using a flat density:

$$f(x) = \begin{cases} \frac{1}{b-a} & : a < x < b \\ 0 & : \text{ otherwise} \end{cases}$$

the uniform distributions

The random variable X, the result of "picking a real number at random between a and b", is modeled using a flat density:

$$f(x) = \begin{cases} \frac{1}{b-a} & : a < x < b \\ 0 & : \text{ otherwise} \end{cases}$$

Easy to show $E(X) = (b-a)/2$ and $\mathrm{Var}(X) = (b-a)^2/12$.

the uniform distributions

The random variable X, the result of "picking a real number at random between a and b", is modeled using a flat density:

$$f(x) = \begin{cases} \frac{1}{b-a} & : a < x < b \\ 0 & : \text{ otherwise} \end{cases}$$

Easy to show $E(X) = (b-a)/2$ and $\mathrm{Var}(X) = (b-a)^2/12$.

mgf is easy to determine, but not really useful for anything.