18.600: Lecture 15

Poisson processes

Scott Sheffield

MIT

Outline

Outline

Properties from last time...

- A **Poisson random variable** X with parameter λ satisfies $P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$.

Properties from last time...

- A **Poisson random variable** X with parameter λ satisfies $P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$.
- The probabilities are approximately those of a binomial with parameters $(n, \lambda/n)$ when n is very large.

Properties from last time...

- A **Poisson random variable** X with parameter λ satisfies $P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$.
- The probabilities are approximately those of a binomial with parameters $(n, \lambda/n)$ when n is very large.
- Indeed,

$$\binom{n}{k} p^k (1-p)^{n-k} = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} p^k (1-p)^{n-k} \approx$$

$$\frac{\lambda^k}{k!}(1-p)^{n-k} \approx \frac{\lambda^k}{k!}e^{-\lambda}.$$

Properties from last time...

- A **Poisson random variable** X with parameter λ satisfies $P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$.
- The probabilities are approximately those of a binomial with parameters $(n, \lambda/n)$ when n is very large.
- Indeed,

$$\binom{n}{k} p^k (1-p)^{n-k} = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} p^k (1-p)^{n-k} \approx$$

$$\frac{\lambda^k}{k!}(1-p)^{n-k} \approx \frac{\lambda^k}{k!}e^{-\lambda}.$$

- General idea: if you have a large number of unlikely events that are (mostly) independent of each other, and the expected number that occur is λ, then the total number that occur should be (approximately) a Poisson random variable with parameter λ.

Properties from last time...

- Many phenomena (number of phone calls or customers arriving in a given period, number of radioactive emissions in a given time period, number of major hurricanes in a given time period, etc.) can be modeled this way.

Properties from last time...

- Many phenomena (number of phone calls or customers arriving in a given period, number of radioactive emissions in a given time period, number of major hurricanes in a given time period, etc.) can be modeled this way.
- A **Poisson random variable** X with parameter λ has expectation λ and variance λ.

Properties from last time...

- Many phenomena (number of phone calls or customers arriving in a given period, number of radioactive emissions in a given time period, number of major hurricanes in a given time period, etc.) can be modeled this way.
- A **Poisson random variable** X with parameter λ has expectation λ and variance λ.
- Special case: if $\lambda = 1$, then $P\{X = k\} = \frac{1}{k!e}$.

Properties from last time...

- Many phenomena (number of phone calls or customers arriving in a given period, number of radioactive emissions in a given time period, number of major hurricanes in a given time period, etc.) can be modeled this way.
- A **Poisson random variable** X with parameter λ has expectation λ and variance λ.
- Special case: if $\lambda = 1$, then $P\{X = k\} = \frac{1}{k!e}$.
- Note how quickly this goes to zero, as a function of k.

Properties from last time...

- Many phenomena (number of phone calls or customers arriving in a given period, number of radioactive emissions in a given time period, number of major hurricanes in a given time period, etc.) can be modeled this way.
- A **Poisson random variable** X with parameter λ has expectation λ and variance λ.
- Special case: if $\lambda = 1$, then $P\{X = k\} = \frac{1}{k!e}$.
- Note how quickly this goes to zero, as a function of k.
- Example: number of royal flushes in a million five-card poker hands is approximately Poisson with parameter $10^6/649739 \approx 1.54$.

Properties from last time...

- Many phenomena (number of phone calls or customers arriving in a given period, number of radioactive emissions in a given time period, number of major hurricanes in a given time period, etc.) can be modeled this way.
- A **Poisson random variable** X with parameter λ has expectation λ and variance λ.
- Special case: if $\lambda = 1$, then $P\{X = k\} = \frac{1}{k!e}$.
- Note how quickly this goes to zero, as a function of k.
- Example: number of royal flushes in a million five-card poker hands is approximately Poisson with parameter $10^6/649739 \approx 1.54$.
- Example: if a country expects 2 plane crashes in a year, then the total number might be approximately Poisson with parameter $\lambda = 2$.

A cautionary tail

- Example: Joe works for a bank and notices that his town sees an average of one mortgage foreclosure per month.

A cautionary tail

- Example: Joe works for a bank and notices that his town sees an average of one mortgage foreclosure per month.
- Moreover, looking over five years of data, it seems that the number of foreclosures per month follows a rate 1 Poisson distribution.

A cautionary tail

- Example: Joe works for a bank and notices that his town sees an average of one mortgage foreclosure per month.
- Moreover, looking over five years of data, it seems that the number of foreclosures per month follows a rate 1 Poisson distribution.
- That is, roughly a $1/e$ fraction of months has 0 foreclosures, a $1/e$ fraction has 1, a $1/(2e)$ fraction has 2, a $1/(6e)$ fraction has 3, and a $1/(24e)$ fraction has 4.

A cautionary tail

- Example: Joe works for a bank and notices that his town sees an average of one mortgage foreclosure per month.
- Moreover, looking over five years of data, it seems that the number of foreclosures per month follows a rate 1 Poisson distribution.
- That is, roughly a $1/e$ fraction of months has 0 foreclosures, a $1/e$ fraction has 1, a $1/(2e)$ fraction has 2, a $1/(6e)$ fraction has 3, and a $1/(24e)$ fraction has 4.
- Joe concludes that the probability of seeing 10 foreclosures during a given month is only $1/(10!e)$. Probability to see 10 or more (an extreme *tail event* that would destroy the bank) is $\sum_{k=10}^{\infty} 1/(k!e)$, less than one in million.

A cautionary tail

- Example: Joe works for a bank and notices that his town sees an average of one mortgage foreclosure per month.
- Moreover, looking over five years of data, it seems that the number of foreclosures per month follows a rate 1 Poisson distribution.
- That is, roughly a $1/e$ fraction of months has 0 foreclosures, a $1/e$ fraction has 1, a $1/(2e)$ fraction has 2, a $1/(6e)$ fraction has 3, and a $1/(24e)$ fraction has 4.
- Joe concludes that the probability of seeing 10 foreclosures during a given month is only $1/(10!e)$. Probability to see 10 or more (an extreme *tail event* that would destroy the bank) is $\sum_{k=10}^{\infty} 1/(k!e)$, less than one in million.
- Investors are impressed. Joe receives large bonus.

A cautionary tail

- Example: Joe works for a bank and notices that his town sees an average of one mortgage foreclosure per month.
- Moreover, looking over five years of data, it seems that the number of foreclosures per month follows a rate 1 Poisson distribution.
- That is, roughly a $1/e$ fraction of months has 0 foreclosures, a $1/e$ fraction has 1, a $1/(2e)$ fraction has 2, a $1/(6e)$ fraction has 3, and a $1/(24e)$ fraction has 4.
- Joe concludes that the probability of seeing 10 foreclosures during a given month is only $1/(10!e)$. Probability to see 10 or more (an extreme *tail event* that would destroy the bank) is $\sum_{k=10}^{\infty} 1/(k!e)$, less than one in million.
- Investors are impressed. Joe receives large bonus.
- But probably shouldn't....

Outline

Outline

How should we define the *Poisson process*?

- Whatever his faults, Joe was a good record keeper. He kept track of the precise *times* at which the foreclosures occurred over the whole five years (not just the total numbers of foreclosures). We could try this for other problems as well.

How should we define the *Poisson process*?

- Whatever his faults, Joe was a good record keeper. He kept track of the precise *times* at which the foreclosures occurred over the whole five years (not just the total numbers of foreclosures). We could try this for other problems as well.
- Let's encode this information with a function. We'd like a random function $N(t)$ that describe the number of events that occur during the first t units of time. (This could be a model for the number of plane crashes in first t years, or the number of royal flushes in first $10^6 t$ poker hands.)

How should we define the *Poisson process*?

- Whatever his faults, Joe was a good record keeper. He kept track of the precise *times* at which the foreclosures occurred over the whole five years (not just the total numbers of foreclosures). We could try this for other problems as well.
- Let's encode this information with a function. We'd like a random function $N(t)$ that describe the number of events that occur during the first t units of time. (This could be a model for the number of plane crashes in first t years, or the number of royal flushes in first $10^6 t$ poker hands.)
- So $N(t)$ is a **random non-decreasing integer-valued function** of t with $N(0) = 0$.

How should we define the *Poisson process*?

- Whatever his faults, Joe was a good record keeper. He kept track of the precise *times* at which the foreclosures occurred over the whole five years (not just the total numbers of foreclosures). We could try this for other problems as well.
- Let's encode this information with a function. We'd like a random function $N(t)$ that describe the number of events that occur during the first t units of time. (This could be a model for the number of plane crashes in first t years, or the number of royal flushes in first $10^6 t$ poker hands.)
- So $N(t)$ is a **random non-decreasing integer-valued function** of t with $N(0) = 0$.
- For each t, $N(t)$ is a random variable, and the $N(t)$ are functions on the same sample space.

Outline

Outline

Poisson random variables

What should a Poisson point process be?

Poisson point process axioms

Consequences of axioms

Poisson process axioms

- Let's back up and give a precise and minimal list of properties we want the random function $N(t)$ to satisfy.

Poisson process axioms

- Let's back up and give a precise and minimal list of properties we want the random function $N(t)$ to satisfy.
- 1. $N(0) = 0$.

Poisson process axioms

- Let's back up and give a precise and minimal list of properties we want the random function $N(t)$ to satisfy.
- 1. $N(0) = 0$.
- 2. **Independence:** Number of events (jumps of N) in disjoint time intervals are independent.

Poisson process axioms

- Let's back up and give a precise and minimal list of properties we want the random function $N(t)$ to satisfy.
- 1. $N(0) = 0$.
- 2. **Independence:** Number of events (jumps of N) in disjoint time intervals are independent.
- 3. **Homogeneity:** Prob. distribution of # events in interval depends only on length. (Deduce: $E[N(h)] = \lambda h$ for some λ.)

Poisson process axioms

- Let's back up and give a precise and minimal list of properties we want the random function $N(t)$ to satisfy.
- 1. $N(0) = 0$.
- 2. **Independence:** Number of events (jumps of N) in disjoint time intervals are independent.
- 3. **Homogeneity:** Prob. distribution of # events in interval depends only on length. (Deduce: $E[N(h)] = \lambda h$ for some λ.)
- 4. **Non-concurrence:** $P\{N(h) \geq 2\} << P\{N(h) = 1\}$ when h is small. Precisely:

Poisson process axioms

- Let's back up and give a precise and minimal list of properties we want the random function $N(t)$ to satisfy.
- 1. $N(0) = 0$.
- 2. **Independence:** Number of events (jumps of N) in disjoint time intervals are independent.
- 3. **Homogeneity:** Prob. distribution of # events in interval depends only on length. (Deduce: $E[N(h)] = \lambda h$ for some λ.)
- 4. **Non-concurrence:** $P\{N(h) \geq 2\} << P\{N(h) = 1\}$ when h is small. Precisely:
  - $P\{N(h) = 1\} = \lambda h + o(h)$. (Here $f(h) = o(h)$ means $\lim_{h \to 0} f(h)/h = 0$.)

Poisson process axioms

- Let's back up and give a precise and minimal list of properties we want the random function $N(t)$ to satisfy.
- 1. $N(0) = 0$.
- 2. **Independence:** Number of events (jumps of N) in disjoint time intervals are independent.
- 3. **Homogeneity:** Prob. distribution of # events in interval depends only on length. (Deduce: $E[N(h)] = \lambda h$ for some λ.)
- 4. **Non-concurrence:** $P\{N(h) \geq 2\} << P\{N(h) = 1\}$ when h is small. Precisely:
  - $P\{N(h) = 1\} = \lambda h + o(h)$. (Here $f(h) = o(h)$ means $\lim_{h \to 0} f(h)/h = 0$.)
  - $P\{N(h) \geq 2\} = o(h)$.

Poisson process axioms

- Let's back up and give a precise and minimal list of properties we want the random function $N(t)$ to satisfy.
- 1. $N(0) = 0$.
- 2. **Independence:** Number of events (jumps of N) in disjoint time intervals are independent.
- 3. **Homogeneity:** Prob. distribution of # events in interval depends only on length. (Deduce: $E[N(h)] = \lambda h$ for some λ.)
- 4. **Non-concurrence:** $P\{N(h) \geq 2\} << P\{N(h) = 1\}$ when h is small. Precisely:
  - $P\{N(h) = 1\} = \lambda h + o(h)$. (Here $f(h) = o(h)$ means $\lim_{h \to 0} f(h)/h = 0$.)
  - $P\{N(h) \geq 2\} = o(h)$.
- A random function $N(t)$ with these properties is a **Poisson process with rate λ**.

Outline

Outline

Consequences of axioms: time till first event

- Can we work out the probability of no events before time t?

Consequences of axioms: time till first event

- Can we work out the probability of no events before time t?
- We assumed $P\{N(h) = 1\} = \lambda h + o(h)$ and $P\{N(h) \geq 2\} = o(h)$. Taken together, these imply that $P\{N(h) = 0\} = 1 - \lambda h + o(h)$.

Consequences of axioms: time till first event

- Can we work out the probability of no events before time t?
- We assumed $P\{N(h) = 1\} = \lambda h + o(h)$ and $P\{N(h) \geq 2\} = o(h)$. Taken together, these imply that $P\{N(h) = 0\} = 1 - \lambda h + o(h)$.
- Fix λ and t. Probability of no events in interval of length t/n is $(1 - \lambda t/n) + o(1/n)$.

Consequences of axioms: time till first event

- Can we work out the probability of no events before time t?
- We assumed $P\{N(h) = 1\} = \lambda h + o(h)$ and $P\{N(h) \geq 2\} = o(h)$. Taken together, these imply that $P\{N(h) = 0\} = 1 - \lambda h + o(h)$.
- Fix λ and t. Probability of no events in interval of length t/n is $(1 - \lambda t/n) + o(1/n)$.
- Probability of no events in first n such intervals is about $\big(1 - \lambda t/n + o(1/n)\big)^n \approx e^{-\lambda t}$.

Consequences of axioms: time till first event

- Can we work out the probability of no events before time t?
- We assumed $P\{N(h) = 1\} = \lambda h + o(h)$ and $P\{N(h) \geq 2\} = o(h)$. Taken together, these imply that $P\{N(h) = 0\} = 1 - \lambda h + o(h)$.
- Fix λ and t. Probability of no events in interval of length t/n is $(1 - \lambda t/n) + o(1/n)$.
- Probability of no events in first n such intervals is about $\big(1 - \lambda t/n + o(1/n)\big)^n \approx e^{-\lambda t}$.
- Taking limit as $n \to \infty$, can show that probability of no event in interval of length t is $e^{-\lambda t}$.

Consequences of axioms: time till first event

- Can we work out the probability of no events before time t?
- We assumed $P\{N(h) = 1\} = \lambda h + o(h)$ and $P\{N(h) \geq 2\} = o(h)$. Taken together, these imply that $P\{N(h) = 0\} = 1 - \lambda h + o(h)$.
- Fix λ and t. Probability of no events in interval of length t/n is $(1 - \lambda t/n) + o(1/n)$.
- Probability of no events in first n such intervals is about $\big(1 - \lambda t/n + o(1/n)\big)^n \approx e^{-\lambda t}$.
- Taking limit as $n \to \infty$, can show that probability of no event in interval of length t is $e^{-\lambda t}$.
- $P\{N(t) = 0\} = e^{-\lambda t}$.

Consequences of axioms: time till first event

- Can we work out the probability of no events before time t?
- We assumed $P\{N(h) = 1\} = \lambda h + o(h)$ and $P\{N(h) \geq 2\} = o(h)$. Taken together, these imply that $P\{N(h) = 0\} = 1 - \lambda h + o(h)$.
- Fix λ and t. Probability of no events in interval of length t/n is $(1 - \lambda t/n) + o(1/n)$.
- Probability of no events in first n such intervals is about $\big(1 - \lambda t/n + o(1/n)\big)^n \approx e^{-\lambda t}$.
- Taking limit as $n \to \infty$, can show that probability of no event in interval of length t is $e^{-\lambda t}$.
- $P\{N(t) = 0\} = e^{-\lambda t}$.
- Let T_1 be the time of the first event. Then $P\{T_1 \geq t\} = e^{-\lambda t}$. We say that T_1 is an **exponential random variable with rate** λ.

Consequences of axioms: time till second, third events

- Let T_2 be time between first and second event. Generally, T_k is time between $(k-1)$th and kth event.

Consequences of axioms: time till second, third events

- Let T_2 be time between first and second event. Generally, T_k is time between $(k-1)$th and kth event.
- Then the $T_1, T_2, \dots$ are independent of each other (informally this means that observing some of the random variables T_k gives you no information about the others). Each is an exponential random variable with rate λ.

Consequences of axioms: time till second, third events

- Let T_2 be time between first and second event. Generally, T_k is time between $(k-1)$th and kth event.
- Then the $T_1, T_2, \ldots$ are independent of each other (informally this means that observing some of the random variables T_k gives you no information about the others). Each is an exponential random variable with rate λ.
- This finally gives us a way to construct $N(t)$. It is determined by the sequence T_j of independent exponential random variables.

Consequences of axioms: time till second, third events

- Let T_2 be time between first and second event. Generally, T_k is time between $(k-1)$th and kth event.
- Then the $T_1, T_2, \ldots$ are independent of each other (informally this means that observing some of the random variables T_k gives you no information about the others). Each is an exponential random variable with rate λ.
- This finally gives us a way to construct $N(t)$. It is determined by the sequence T_j of independent exponential random variables.
- Axioms can be readily verified from this description.

Back to Poisson distribution

- Axioms should imply that $P\{N(t) = k\} = e^{-\lambda t}(\lambda t)^k / k!$.

Back to Poisson distribution

- Axioms should imply that $P\{N(t) = k\} = e^{-\lambda t}(\lambda t)^k/k!$.
- One way to prove this: divide time into n intervals of length t/n. In each, probability to see an event is $p = \lambda t/n + o(1/n)$.

Back to Poisson distribution

- Axioms should imply that $P\{N(t)=k\} = e^{-\lambda t}(\lambda t)^k/k!$.
- One way to prove this: divide time into n intervals of length t/n. In each, probability to see an event is $p = \lambda t/n + o(1/n)$.
- Use binomial theorem to describe probability to see event in exactly k intervals.

Back to Poisson distribution

- Axioms should imply that $P\{N(t) = k\} = e^{-\lambda t}(\lambda t)^k/k!$.
- One way to prove this: divide time into n intervals of length t/n. In each, probability to see an event is $p = \lambda t/n + o(1/n)$.
- Use binomial theorem to describe probability to see event in exactly k intervals.
- Binomial formula: $\binom{n}{k}p^k(1-p)^{n-k} = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!}p^k(1-p)^{n-k}$.

Back to Poisson distribution

- Axioms should imply that $P\{N(t) = k\} = e^{-\lambda t}(\lambda t)^k/k!$.
- One way to prove this: divide time into n intervals of length t/n. In each, probability to see an event is $p = \lambda t/n + o(1/n)$.
- Use binomial theorem to describe probability to see event in exactly k intervals.
- Binomial formula: $\binom{n}{k}p^k(1-p)^{n-k} = \frac{n(n-1)(n-2)\ldots(n-k+1)}{k!}p^k(1-p)^{n-k}$.
- This is approximately $\frac{(\lambda t)^k}{k!}(1-p)^{n-k} \approx \frac{(\lambda t)^k}{k!}e^{-\lambda t}$.

Back to Poisson distribution

- Axioms should imply that $P\{N(t) = k\} = e^{-\lambda t}(\lambda t)^k/k!$.
- One way to prove this: divide time into n intervals of length t/n. In each, probability to see an event is $p = \lambda t/n + o(1/n)$.
- Use binomial theorem to describe probability to see event in exactly k intervals.
- Binomial formula: $\binom{n}{k}p^k(1-p)^{n-k} = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!}p^k(1-p)^{n-k}$.
- This is approximately $\frac{(\lambda t)^k}{k!}(1-p)^{n-k} \approx \frac{(\lambda t)^k}{k!}e^{-\lambda t}$.
- Take n to infinity, and use fact that expected number of intervals with two or more points tends to zero (thus probability to see any intervals with two more points tends to zero).

Summary

- We constructed a random function $N(t)$ called a Poisson process of rate λ.

Summary

- We constructed a random function $N(t)$ called a Poisson process of rate λ.
- For each $t > s \geq 0$, the value $N(t) - N(s)$ describes the number of events occurring in the time interval (s, t) and is Poisson with mean $(t - s)\lambda$.

Summary

- We constructed a random function $N(t)$ called a Poisson process of rate λ.
- For each $t > s \geq 0$, the value $N(t) - N(s)$ describes the number of events occurring in the time interval (s, t) and is Poisson with mean $(t - s)\lambda$.
- The numbers of events occurring in disjoint intervals are independent random variables.

Summary

- We constructed a random function $N(t)$ called a Poisson process of rate λ.
- For each $t > s \geq 0$, the value $N(t) - N(s)$ describes the number of events occurring in the time interval (s, t) and is Poisson with mean $(t - s)\lambda$.
- The numbers of events occurring in disjoint intervals are independent random variables.
- Let T_k be time elapsed, since the previous event, until the kth event occurs. Then the T_k are independent random variables, each of which is exponential with parameter λ.