18.440: Lecture 20

Exponential random variables

Scott Sheffield

MIT

Outline

Exponential random variables

Minimum of independent exponentials

Memoryless property

Relationship to Poisson random variables

Outline

Exponential random variables

Minimum of independent exponentials

Memoryless property

Relationship to Poisson random variables

Exponential random variables

- Say X is an **exponential random variable of parameter** λ when its probability distribution function is

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}.$$

- For $a > 0$ have

$$F_X(a) = \int_0^a f(x)dx = \int_0^a \lambda e^{-\lambda x} dx = -e^{-\lambda x}\Big|_0^a = 1 - e^{-\lambda a}.$$

- Thus $P\{X < a\} = 1 - e^{-\lambda a}$ and $P\{X > a\} = e^{-\lambda a}$.
- Formula $P\{X > a\} = e^{-\lambda a}$ is very important in practice.

Moment formula

- Suppose X is exponential with parameter λ, so $f_X(x) = \lambda e^{-\lambda x}$ when $x \geq 0$.
- What is $E[X^n]$? (Say $n \geq 1$.)
- Write $E[X^n] = \int_0^\infty x^n \lambda e^{-\lambda x} dx$.
- Integration by parts gives $E[X^n] = -\int_0^\infty n x^{n-1} \lambda \frac{e^{-\lambda x}}{-\lambda} dx + x^n \lambda \frac{e^{-\lambda x}}{-\lambda} \Big|_0^\infty$.
- We get $E[X^n] = \frac{n}{\lambda} E[X^{n-1}]$.
- $E[X^0] = E[1] = 1$, $E[X] = 1/\lambda$, $E[X^2] = 2/\lambda^2$, $E[X^n] = n!/\lambda^n$.
- If $\lambda = 1$, the $E[X^n] = n!$. Could take this as definition of $n!$. It makes sense for $n = 0$ and for non-integer n.
- Variance: $\mathrm{Var}[X] = E[X^2] - (E[X])^2 = 1/\lambda^2$.

Outline

Exponential random variables

Minimum of independent exponentials

Memoryless property

Relationship to Poisson random variables

Outline

Exponential random variables

Minimum of independent exponentials

Memoryless property

Relationship to Poisson random variables

Minimum of independent exponentials is exponential

- **CLAIM:** If X_1 and X_2 are independent and exponential with parameters λ_1 and λ_2 then $X = \min\{X_1, X_2\}$ is exponential with parameter $\lambda = \lambda_1 + \lambda_2$.
- How could we prove this?
- Have various ways to describe random variable Y: via density function $f_Y(x)$, or cumulative distribution function $F_Y(a) = P\{Y \leq a\}$, or function $P\{Y > a\} = 1 - F_Y(a)$.
- Last one has simple form for exponential random variables. We have $P\{Y > a\} = e^{-\lambda a}$ for $a \in [0, \infty)$.
- Note: $X > a$ if and only if $X_1 > a$ *and* $X_2 > a$.
- X_1 and X_2 are independent, so $P\{X > a\} = P\{X_1 > a\}P\{X_2 > a\} = e^{-\lambda_1 a}e^{-\lambda_2 a} = e^{-\lambda a}$.
- If $X_1, \dots, X_n$ are independent exponential with $\lambda_1, \dots \lambda_n$, then $\min\{X_1, \dots X_n\}$ is exponential with $\lambda = \lambda_1 + \dots + \lambda_n$.

Outline

Exponential random variables

Minimum of independent exponentials

Memoryless property

Relationship to Poisson random variables

Outline

Exponential random variables

Minimum of independent exponentials

Memoryless property

Relationship to Poisson random variables

Memoryless property

- Suppose X is exponential with parameter λ.
- **Memoryless property:** If X represents the time until an event occurs, then *given* that we have seen no event up to time b, the conditional distribution of the remaining time till the event is the same as it originally was.
- To make this precise, we ask what is the probability distribution of $Y = X - b$ *conditioned on* $X > b$?
- We can characterize the conditional law of Y, given $X > b$, by computing $P(Y > a | X > b)$ for each a.
- That is, we compute $P(X - b > a | X > b) = P(X > b + a | X > b)$.
- By definition of conditional probability, this is just $P\{X > b + a\}/P\{X > b\} = e^{-\lambda(b+a)}/e^{-\lambda b} = e^{-\lambda a}$.
- Thus, conditional law of $X - b$ *given* that $X > b$ is same as the original law of X.

Memoryless property for geometric random variables

- Similar property holds for geometric random variables.
- If we plan to toss a coin until the first heads comes up, then we have a .5 chance to get a heads in one step, a .25 chance in two steps, etc.
- Given that the first 5 tosses are all tails, there is conditionally a .5 chance we get our first heads on the 6th toss, a .25 chance on the 7th toss, etc.
- Despite our having had five tails in a row, our expectation of the amount of time remaining until we see a heads is the same as it originally was.

Exchange overheard on Logan airport shuttle

- **Bob:** There's this really interesting problem in statistics I just learned about. If a coin comes up heads 10 times in a row, how likely is the next toss to be heads?
- **Alice:** Still fifty fifty.
- **Bob:** That's a common mistake, but you're wrong because the 10 heads in a row increase the conditional probability that there's something funny going on with the coin.
- **Alice:** You never said it might be a funny coin.
- **Bob:** That's the point. You should always suspect that there might be something funny with the coin.
- **Alice:** It's a math puzzle. You always assume a normal coin.
- **Bob:** No, that's your mistake. You should never assume that, because maybe somebody tampered with the coin.

Exchange overheard on a Logan airport shuttle

- **Alice:** Yeah, yeah, I get it. I can't win here.
- **Bob:** No, I don't think you get it yet. It's a subtle point in statistics. It's very important.
- Exchange continued for duration of shuttle ride (Alice increasingly irritated, Bob increasingly patronizing).
- Raises interesting question about memoryless property.
- Suppose the duration of a couple's relationship is exponential with λ^{-1} equal to two weeks.
- Given that it has lasted for 10 weeks so far, what is the conditional probability that it will last an additional week?
- How about an additional four weeks? Ten weeks?

Remark on Alice and Bob

- Alice assumes Bob means "independent tosses of a fair coin." Under this assumption, all 2^{11} outcomes of eleven-coin-toss sequence are equally likely. Bob considers HHHHHHHHHHH more likely than HHHHHHHHHHT, since former could result from a faulty coin.
- Alice sees Bob's point but considers it annoying and churlish to ask about coin toss sequence and criticize listener for assuming this means "independent tosses of fair coin".
- Without that assumption, Alice has no idea what context Bob has in mind. (An environment where two-headed novelty coins are common? Among coin-tossing cheaters with particular agendas?...)
- Alice: you need assumptions to convert stories into math.
- Bob: good to question assumptions.

Radioactive decay: maximum of independent exponentials

- Suppose you start at time zero with n radioactive particles. Suppose that each one (independently of the others) will decay at a random time, which is an exponential random variable with parameter λ.
- Let T be amount of time until no particles are left. What are $E[T]$ and $\mathrm{Var}[T]$?
- Let T_1 be the amount of time you wait until the first particle decays, T_2 the amount of *additional* time until the second particle decays, etc., so that $T = T_1 + T_2 + \ldots T_n$.
- Claim: T_1 is exponential with parameter $n\lambda$.
- Claim: T_2 is exponential with parameter $(n-1)\lambda$.
- And so forth. $E[T] = \sum_{i=1}^{n} E[T_i] = \lambda^{-1} \sum_{j=1}^{n} \frac{1}{j}$ and (by independence) $\mathrm{Var}[T] = \sum_{i=1}^{n} \mathrm{Var}[T_i] = \lambda^{-2} \sum_{j=1}^{n} \frac{1}{j^2}$.

Outline

Exponential random variables

Minimum of independent exponentials

Memoryless property

Relationship to Poisson random variables

Outline

Exponential random variables

Minimum of independent exponentials

Memoryless property

Relationship to Poisson random variables

Relationship to Poisson random variables

- Let $T_1, T_2, \ldots$ be independent exponential random variables with parameter λ.
- We can view them as waiting times between "events".
- How do you show that the number of events in the first t units of time is Poisson with parameter λt?
- We actually did this already in the lecture on Poisson point processes. You can break the interval $[0, t]$ into n equal pieces (for very large n), let X_k be number of events in kth piece, use memoryless property to argue that the X_k are independent.
- When n is large enough, it becomes unlikely that any interval has more than one event. Roughly speaking: each interval has one event with probability $\lambda t/n$, zero otherwise.
- Take $n \to \infty$ limit. Number of events is Poisson λt.

18.440 Probability and Random Variables
Spring 2014