Continuous Zoo

CSE 312 Winter 26
Lecture 15

Let's start with the pmf

For discrete random variables, we defined the pmf: $p_Y(k) = \mathbb{P}(Y = k)$.

We can't have a pmf quite like we did for discrete random variables. Let X be a random real number between 0 and 1.

$\mathbb{P}(X = .1) = \frac{1}{\infty}$??

Let's try to maintain as many rules as we can…

Discrete	Continuous
$p_Y(k) \geq 0$	$f_X(k) \geq 0$
$\sum_{\omega} p_Y(\omega) = 1$	$\int_{-\infty}^{\infty} f_X(k)\ \mathrm{d}k = 1$

Use f_X instead of p_X to remember it's different .

The probability density function

For Continuous random variables, the analogous object is the "probability density function" we write $f_X(k)$ instead of $p_X(k)$

Idea: Make it "work right" for **events** since single outcomes don't make sense.

$\mathbb{P}(a \leq X \leq b) = c$

$$\int_a^b f_X(z)\,\mathrm{dz} = c$$

integrating is analogous to sum.

Let's derive an example PDF together!

For a uniform random real number in [0,1]

CDFs

What's a CDF?

The Cumulative Distribution Function $F_X(k) = \mathbb{P}(X \leq k)$ analogous to the CDF for discrete variables.

$$F_X(k) = \mathbb{P}(X \leq k) = \int_{-\infty}^{k} f_X(z)\,\mathrm{d}z$$

So how do I get from CDF to PDF? Taking the derivative!

$$\frac{\mathrm{d}}{\mathrm{d}k} F_X(k) = \frac{\mathrm{d}}{dk}\left(\int_{-\infty}^{k} f_X(z)\,\mathrm{d}z\right) = f_X(k)$$

Comparing Discrete and Continuous

	Discrete Random Variables	Continuous Random Variables
Probability 0	Equivalent to impossible	All impossible events have probability 0, but not conversely.
Relative Chances	PMF: $p_X(k) = \mathbb{P}(X = k)$	PDF $f_X(k)$ gives chances relative to $f_X(k')$
Events	Sum over PMF to get probability	Integrate PDF to get probability
Convert from CDF to PMF	Sum up PMF to get CDF. Look for “breakpoints” in CDF to get PMF.	Integrate PDF to get CDF. Differentiate CDF to get PDF.
$\mathbb{E}[X]$	$\sum_{\omega} X(\omega) \cdot p_X(\omega)$	$\int_{-\infty}^{\infty} z \cdot f_X(z)\,\mathrm{dz}$
$\mathbb{E}[g(X)]$	$\sum_{\omega} g(X(\omega)) \cdot p_X(\omega)$	$\int_{-\infty}^{\infty} g(z) \cdot f_X(z)\,\mathrm{dz}$
$\mathrm{Var}(X)$	$\mathbb{E}[X^2] - (\mathbb{E}[X])^2$	$\mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \int_{-\infty}^{\infty} (z - \mathbb{E}[X])^2 f_X(z)\,\mathrm{dz}$

What about expectation?

For a random variable X, we define:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} z \cdot f_X(z)\, \mathrm{d}z$$

Just replace summing over the pmf with integrating the pdf.

It still represents the average value of X.

Expectation of a function

$(z^2+3) \cdot f_X(z)$

For any function g and any continuous random variable, X:

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(X(z)) \cdot f_X(z)\, \mathrm{dz}$$

Again, analogous to the discrete case; just replace summation with integration and pmf with the pdf.

We're going to treat this as a definition.

Technically, this is really a theorem; since $f()$ is the pdf of X and it only gives relative likelihoods for X, we need a proof to guarantee it "works" for $g(X)$.

Sometimes called "Law of the Unconscious Statistician."

Linearity of Expectation

Still true!

$$\mathbb{E}[aX + bY + c] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c$$

For all X, Y; even if they're continuous.

Won't show you the proof – for just $\mathbb{E}[aX + b]$, it's

$\mathbb{E}[aX + b] = \int_{-\infty}^{\infty} [aX(k) + b] f_X(k)\, \mathrm{d}k$

$= \int_{-\infty}^{\infty} aX(k) f_X(k) dk + \int_{-\infty}^{\infty} b f_X(k) dk$

$= a \int_{-\infty}^{\infty} X(k) f_X(k) dk + b \int_{-\infty}^{\infty} f_X(k) dk$

$= a\mathbb{E}[X] + b$

Variance

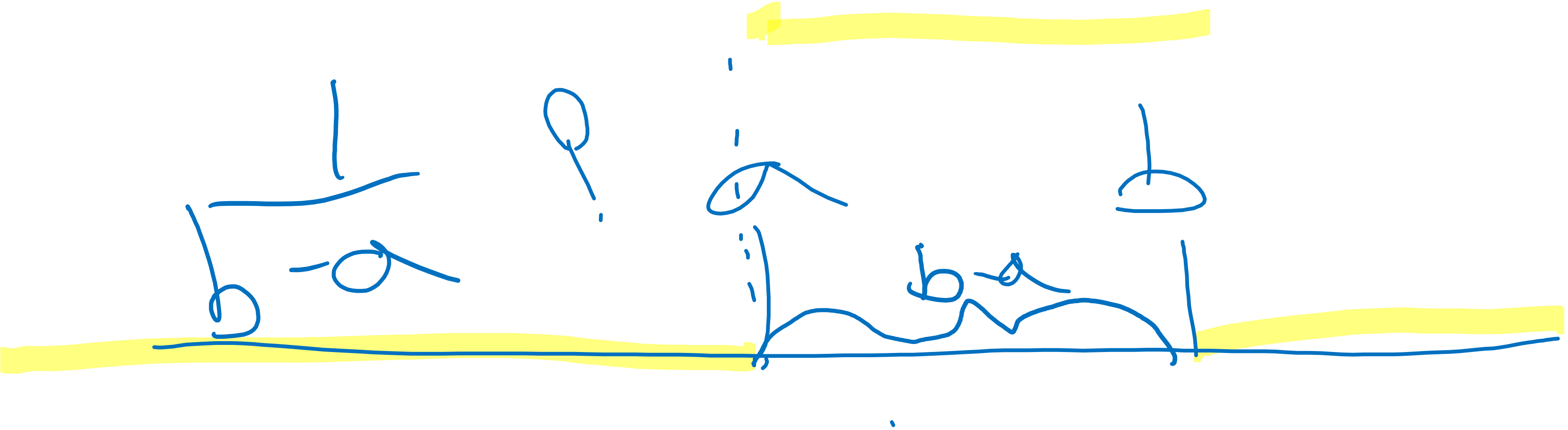
No surprises here

$$\mathbf{Var}(\boldsymbol{X}) = \mathbb{E}[\boldsymbol{X}^2] - (\mathbb{E}[\boldsymbol{X}])^2 = \int_{-\infty}^{\infty} \boldsymbol{f}_{\boldsymbol{X}}(\boldsymbol{k})(\boldsymbol{X}(\boldsymbol{k}) - \mathbb{E}[\boldsymbol{X}])^2 \, \mathbf{d}\boldsymbol{k}$$

Let's calculate an expectation

Let X be a uniform random number between a and b.

$\mathbb{E}[X] = \int_{-\infty}^{\infty} z \cdot f_X(z)\, \mathrm{d}z$



Let's calculate an expectation

Let X be a uniform random number between a and b.

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} z \cdot f_X(z)\,\mathrm{dz}$$

$$= \int_{-\infty}^{a} z \cdot 0\,\mathrm{dz} + \int_{a}^{b} z \cdot \frac{1}{b-a}\,\mathrm{dz} + \int_{b}^{\infty} z \cdot 0\,\mathrm{dz}$$

$$= 0 + \int_{a}^{b} \frac{z}{b-a}\,\mathrm{dz} + 0$$

$$= \frac{z^2}{2(b-a)}\Bigg|_{z=a}^{b} = \frac{b^2}{2(b-a)} - \frac{a^2}{2(b-a)} = \frac{b^2-a^2}{2(b-a)} = \frac{(b+a)(b-a)}{2(b-a)} = \frac{a+b}{2}$$

What about $\mathbb{E}[g(X)]$

Let $X \sim \mathrm{Unif}(a, b)$, what about $\mathbb{E}[X^2]$?

$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} z^2 f_X(z) \mathrm{dz}$$

$$= \int_{-\infty}^{a} z^2 \cdot 0 \ \mathrm{dz} + \int_{a}^{b} z^2 \cdot \frac{1}{b-a} \mathrm{dz} + \int_{b}^{\infty} z^2 \cdot 0 \ \mathrm{dz}$$

$$= 0 + \int_{a}^{b} z^2 \cdot \frac{1}{b-a} \mathrm{dz} + 0$$

$$= \frac{1}{b-a} \cdot \frac{z^3}{3} \bigg|_{z=a}^{b} = \frac{1}{b-a} \left(\frac{b^3}{3} - \frac{a^3}{3} \right) = \frac{1}{3(b-a)} \cdot (b-a)(a^2 + ab + b^2)$$

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

Let's assemble the variance

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

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

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

$$= \frac{a^2-2ab+b^2}{12}$$

$$= \frac{(a-b)^2}{12}$$

Continuous Uniform Distribution

$X \sim \text{Unif}(a, b)$ (uniform real number between a and b)

PDF: $f_X(k) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq k \leq b \\ 0 & \text{otherwise} \end{cases}$

CDF: $F_X(k) = \begin{cases} 0 & \text{if } k < a \\ \frac{k-a}{b-a} & \text{if } a \leq k \leq b \\ 1 & \text{if } k \geq b \end{cases}$

$\mathbb{E}[X] = \frac{a+b}{2}$

$\text{Var}(X) = \frac{(b-a)^2}{12}$

Comparing Discrete and Continuous

	Discrete Random Variables	Continuous Random Variables
Probability 0	Equivalent to impossible	All impossible events have probability 0, but not conversely.
Relative Chances	PMF: $p_X(k) = \mathbb{P}(X = k)$	PDF $f_X(k)$ gives chances relative to $f_X(k')$
Events	Sum over PMF to get probability	Integrate PDF to get probability
Convert from CDF to PMF	Sum up PMF to get CDF. Look for “breakpoints” in CDF to get PMF.	Integrate PDF to get CDF. Differentiate CDF to get PDF.
$\mathbb{E}[X]$	$\sum_{\omega} X(\omega) \cdot p_X(\omega)$	$\int_{-\infty}^{\infty} z \cdot f_X(z)\, \mathrm{d}z$
$\mathbb{E}[g(X)]$	$\sum_{\omega} g(X(\omega)) \cdot p_X(\omega)$	$\int_{-\infty}^{\infty} g(z) \cdot f_X(z)\, \mathrm{d}z$
$\mathrm{Var}(X)$	$\mathbb{E}[X^2] - (\mathbb{E}[X])^2$	$\mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \int_{-\infty}^{\infty} (z - \mathbb{E}[X])^2 f_X(z)\, \mathrm{d}z$

Continuous Zoo

$X \sim \text{Unif}(a, b)$	$X \sim \text{Exp}(\lambda)$	$X \sim \mathcal{N}(\mu, \sigma^2)$
$f_X(k) = \frac{1}{b-a}$	$f_X(k) = \lambda e^{-\lambda k}$ for $k \geq 0$	$f_X(k) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$
$\mathbb{E}[X] = \frac{a+b}{2}$	$\mathbb{E}[X] = \frac{1}{\lambda}$	$\mathbb{E}[X] = \mu$
$\text{Var}(X) = \frac{(b-a)^2}{12}$	$\text{Var}(X) = \frac{1}{\lambda^2}$	$\text{Var}(X) = \sigma^2$

It's a smaller zoo, but it's just as much fun!

Exponential Random Variable

Like a geometric random variable, but continuous time. How long do we wait until an event happens? (instead of "how many flips until a heads")

Where waiting doesn't make the event happen any sooner. (memoryless)

Geometric: $\mathbb{P}(X = k + 1 \,|X \geq 1) = \mathbb{P}(X = k)$

When the first flip is tails, the coin doesn't remember it came up tails, you've made no progress.

For an exponential random variable:

$\mathbb{P}(X \geq k + 1 \,|X \geq 1) = \mathbb{P}(Y \geq k)$

Are these memoryless?

You arrive to a bus stop at a (uniformly) random time, to a bus that arrives every 10 minutes. How long until the bus arrives? How long conditioned on you've already waited 8 minutes?

You put everyone in class into a random order. You'll iterate through that list. What is the probability of being next? Probability of being next conditioned on not selected yet AND half the class has gone?

You flip a coin (independently) until you see a heads. How many flips do you need? How many additional flips after seeing 4 tails?

Are these memoryless?

You arrive to a bus stop at a (uniformly) random time, to a bus that arrives every 10 minutes. How long until the bus arrives? How long conditioned on you've already waited 8 minutes?

Not memoryless! (bus must arrive in 10 minutes total, must be soon!)

You put everyone in class into a random order. You'll iterate through that list. What is the probability of being next? Probability of being next conditioned on not selected yet AND half the class has gone?

Not memoryless ($1/n$ of being first $1/(n/2)$ after half class gone)

You flip a coin (independently) until you see a heads

Memoryless!

A continuous memoryless RV?

Poisson random variables come from a memoryless-type process.

Number of earthquakes (people in bakery, days with snow) would be memoryless under assumption that events are independent of each other!

Same experiments, but now ask a different question:

Poisson: how many incidents occur in fixed interval?

Exponential: how long do I have to wait to see the next incident?

Exponential random variable

If you take a Poisson random variable and ask "what's the time until the next event" you get an exponential distribution!

Let's find the CDF for an exponential.

Let $Y \sim \text{Exp}(\lambda)$, be the time until the first event, when we see an average of λ events per time unit.

What's $\mathbb{P}(Y > t)$?

What Poisson are we waiting on, and what event for it tells you that $Y > t$?

Exponential random variable

If you take a Poisson random variable and ask "what's the time until the next event" you get an exponential distribution!

Let's find the CDF for an exponential.

Let $Y \sim \text{Exp}(\lambda)$, be the time until the first event, when we see an average of λ events per time unit. What's $\mathbb{P}(Y > t)$?

What Poisson are we waiting on? For $X \sim \text{Poi}(\lambda t)$ $\mathbb{P}(Y > t) = \mathbb{P}(X = 0)$

$$\mathbb{P}(X = 0) = \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = e^{-\lambda t}$$

$F_Y(t) = \mathbb{P}(Y \leq t) = 1 - e^{-\lambda t}$ (for $t \geq 0$, $F_Y(x) = 0$ for $x < 0$)

Where did the t come from?

Why did we switch from Exp(λ) to Poi(λt)?

Let's make our units "incidents/second", so $\lambda = 3$ says we average 3 incidents per second.

What if I want to know the probability of waiting at least 5 seconds? Well then on average how many incidents do we see in a 5 second period?

15 incidents! So the Poisson (how many incidents in fixed interval) now refers to a larger interval, so averages more events; specifically λt.

Find the density

We know the CDF, $F_Y(t) = \mathbb{P}(Y \leq t) = 1 - e^{-\lambda t}$

What's the density?

$f_Y(t) =$

Find the density

We know the CDF, $F_Y(t) = \mathbb{P}(Y \leq t) = 1 - e^{-\lambda t}$

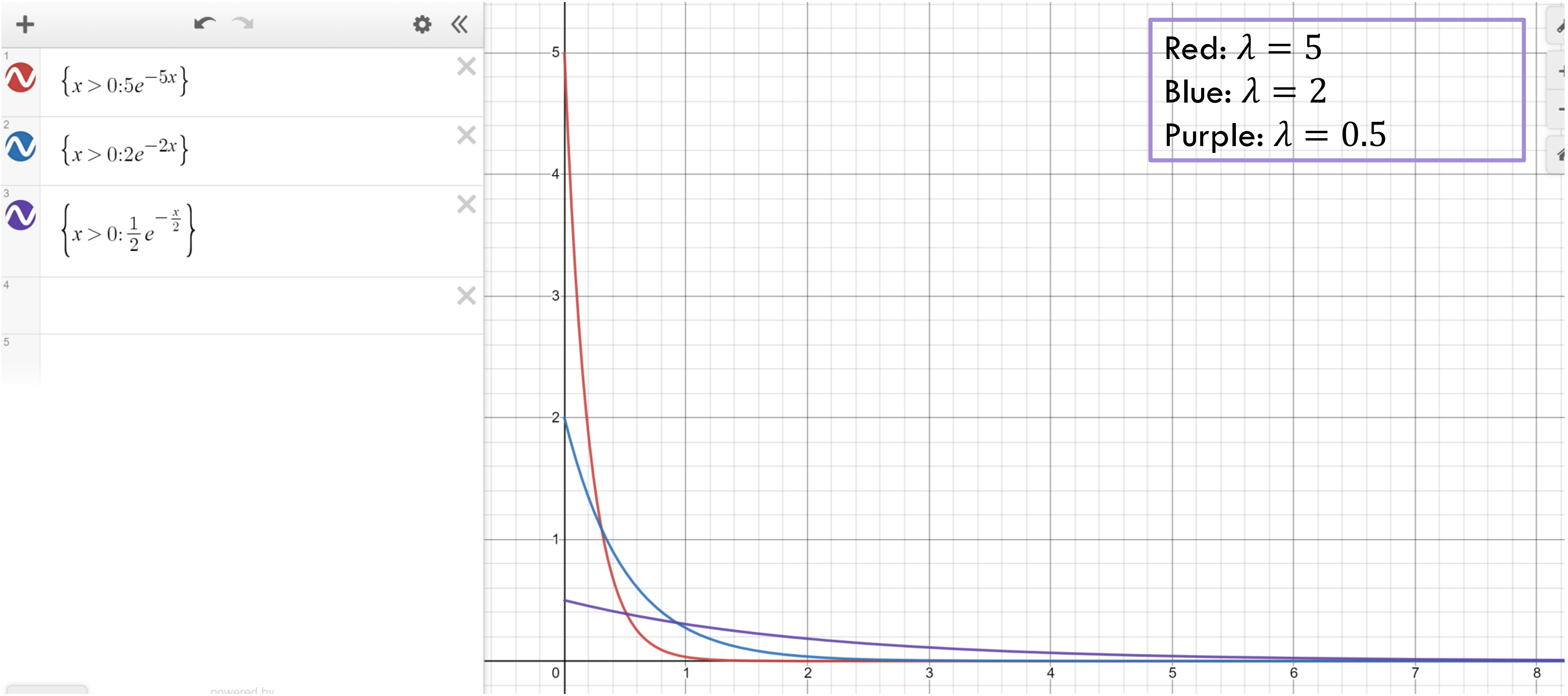
What's the density?

$$f_Y(t) = \frac{d}{dt}\left(1 - e^{-\lambda t}\right) = 0 - \frac{d}{dt}\left(e^{-\lambda t}\right) = \lambda e^{-\lambda t}.$$

For t ≥ 0 it's that expression

For $t < 0$ it's just 0.

Exponential PDF



Memorylessness

$\mathbb{P}(X \geq k + 1 | X \geq 1) = \frac{\mathbb{P}(X \geq k+1 \cap X \geq 1)}{\mathbb{P}(X \geq 1)} = \frac{\mathbb{P}(X \geq k+1)}{1-(1-e^{-\lambda \cdot 1})}$

$= \frac{e^{-\lambda(k+1)}}{e^{-\lambda}} = e^{-\lambda k}$

What about $\mathbb{P}(X \geq k)$ (without conditioning on the first step)?

$1 - (1 - e^{-\lambda k}) = e^{-\lambda k}$

It's the same!!!

More generally, for an exponential rv X, $\mathbb{P}(X \geq s + t | X \geq s) = \mathbb{P}(X \geq t)$

Side note

I hid a trick in that algebra,

$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X < 1) = 1 - \mathbb{P}(X \leq 1)$

The first step is the complementary law.

The second step is using that $\int_1^1 f_X(z)\mathrm{d}z = 0$

In general, for continuous random variables we can switch out $\leq$ and $<$ without anything changing.

We can't make those switches for discrete random variables.

Expectation of an exponential

Don't worry about the derivation (it's here if you're interested; you're not responsible for the derivation. Just the value.

Let $X \sim \mathrm{Exp}(\lambda)$

$\mathbb{E}[X] = \int_{-\infty}^{\infty} z \cdot f_X(z)\mathrm{d}z$

$= \int_0^{\infty} z \cdot \lambda e^{-\lambda z}\, dz$

Let $u = z;\ dv = \lambda e^{-\lambda z} dz\ (v = -e^{-\lambda z})$

Integrate by parts: $-ze^{-\lambda z} - \int -e^{-\lambda z}\, dz = -ze^{-\lambda z} - \frac{1}{\lambda} e^{-\lambda z}$

Definite Integral: $-ze^{-\lambda z} - \frac{1}{\lambda} e^{-\lambda z} |_{z=0}^{\infty} = (\lim_{z \to \infty} -ze^{-\lambda z} - \frac{1}{\lambda} e^{-\lambda z})\ - (0 - \frac{1}{\lambda})$

By L'Hopital's Rule $(\lim_{z \to \infty} -\frac{z}{e^{\lambda z}} - \frac{1}{\lambda e^{\lambda z}})\ - (0 - \frac{1}{\lambda}) = \left(\lim_{z \to \infty} - \frac{1}{\lambda e^{\lambda z}}\right) + \frac{1}{\lambda} = \frac{1}{\lambda}$

Variance of an exponential

If $X \sim \text{Exp}(\lambda)$ then $\text{Var}(X) = \frac{1}{\lambda^2}$

Similar calculus tricks will get you there.

Exponential

$X \sim \text{Exp}(\lambda)$

Parameter $\lambda \geq 0$ is the average number of events in a unit of time.

$$f_X(k) = \begin{cases} \lambda e^{-\lambda k} & \text{if } k \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$F_X(k) = \begin{cases} 1 - e^{-\lambda k} & \text{if } k \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{E}[X] = \frac{1}{\lambda}$$

$$\text{Var}(X) = \frac{1}{\lambda^2}$$