LECTURE: EXPONENTIAL DISTRIBUTION

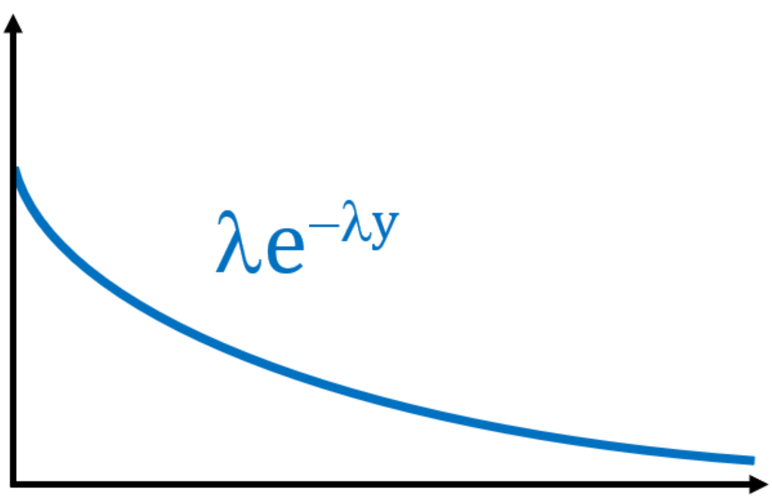
1. Exponential Distribution (continued)

Definition:

Y has an **exponential distribution** with parameter $\lambda > 0$ if

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

We write $Y \sim \text{Exp}(\lambda)$



It models the length of time between independent events with constant average rate, like phone calls at a call center.

Fact:

If $Y \sim$ Exp (λ) then $E(Y) = \frac{1}{\lambda}$ and $\text{Var}(Y) = \frac{1}{\lambda^2}$

Example 1:

While procrastinating studying for this course, you decide to create a Tiktok channel for your bunny Oreo. Suppose your bunny's channel gets an average of 5 unique visits per hour. What is the probability that your bunny will go more than 30 minutes without a visitor?

Let Y be the wait time between visits.

Here we have a Poisson-like situation and we are interested in the time between visits so let $Y \sim \text{Exp}(5)$.

Then Y has density $f(y) = 5e^{-5y}$

Then since we are working in hours, we want

$$P(Y \geq 1/2) = \int_{1/2}^{\infty} 5e^{-5t} dt = \left[-e^{-5t}\right]_{t=\frac{1}{2}}^{t=\infty} = -e^{-\infty} - \left(-e^{-5/2}\right) = e^{-5/2} \approx 0.082$$

Exponential cdf

Suppose $Y \sim$ Exp (λ) Then its cdf $F(y) = P(Y \leq y)$ is

$$F(y) = \begin{cases} 1 - e^{-\lambda y} & y > 0 \\ 0 & y \leq 0 \end{cases}$$

Why? For $y \leq 0$, $F(y) = 0$ since $f(y) = 0$

For $y > 0$, we have:

$$F(y) = P(Y \le y) = \int_{-\infty}^{y} f(t)dt = \int_{0}^{y} \lambda e^{-\lambda t} dt = \left[-e^{-\lambda t}\right]_0^y = 1 - e^{-\lambda y}$$

Example 2:

Find the median of the exponential distribution

By definition, all we need to do is solve $F(m) = P(Y \le m) = \frac{1}{2}$

$$\begin{aligned} F(m) &= \frac{1}{2} \\ 1 - e^{-\lambda m} &= \frac{1}{2} \\ e^{-\lambda m} &= \frac{1}{2} \\ -\lambda m &= \ln\left(\frac{1}{2}\right) = -\ln 2 \\ m &= \frac{\ln 2}{\lambda} \end{aligned}$$

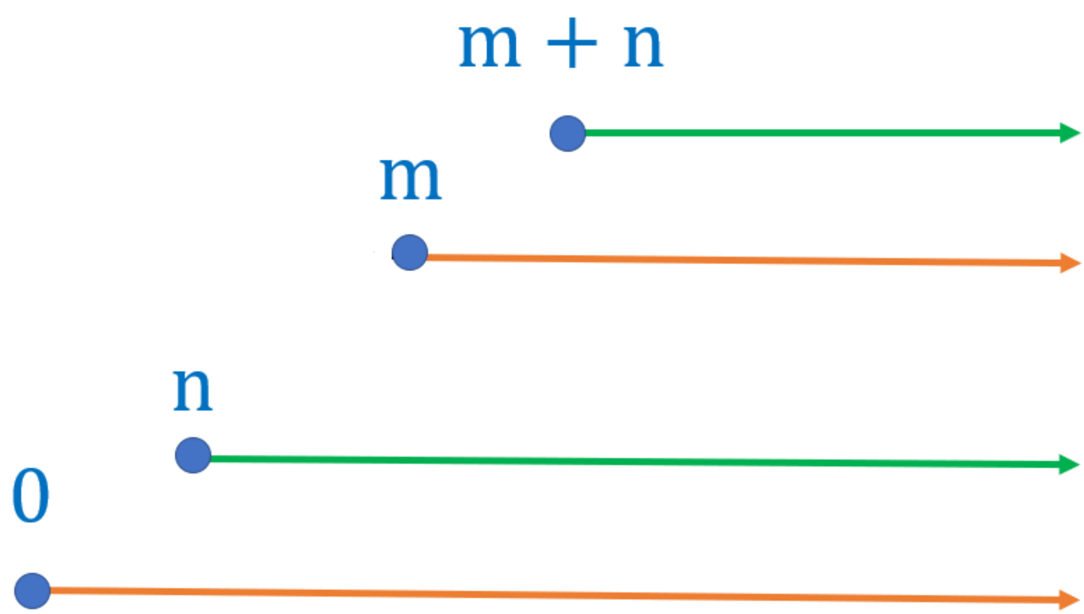
Hence median of the exponential distribution is a little smaller than the mean $\frac{1}{\lambda}$ (by a factor of $\ln 2 \approx 0.693$).

2. Memoriless Property

Recall: Memoriless Property for Geometric Distribution

If $Y \sim$ Geom (p) then for all m and n

$$P(Y > m + n | Y > m) = P(Y > n)$$



Just like the geometric distribution is memoriless, so is the exponential distribution. If we think of this in terms of the lifetime of a light bulb, this means that the probability of the bulb burning out in the next 30 minutes is the same whether we just turned the light bulb on or whether it has been on for 5 days.

Memoriless Property:

Suppose $Y \sim$ Exp (λ) Then for all a and b we have

$$P(Y > a + b | Y > a) = P(Y > b)$$

Why?

STEP 1: First use the exponential cdf $F(y)$ to see that

$$P(Y > y) = 1 - P(Y \leq y) = 1 - F(y) = 1 - (1 - e^{-\lambda y}) = e^{-\lambda y}$$

STEP 2: Using this and the definition of conditional probability:

$$
\begin{aligned}
P(Y > a+b | Y > a) &= \frac{P((Y > a+b) \cap (Y > a))}{P(Y > a)} = \frac{P(Y > a+b)}{P(Y > a)} \\
&= \frac{e^{-\lambda(a+b)}}{e^{-\lambda a}} = e^{-\lambda b} = P(Y > b) \checkmark
\end{aligned}
$$

3. Markov's Inequality

Video: Markov and Chebyshev's Inequalities

In the previous examples, we knew the distribution of Y and were able to calculate $P(Y \geq a)$ explicitly.

Question: What if we don't know *anything* about Y except for its mean $E(Y)$? Can we still say something about $P(Y \geq a)$?

Answer: **YES**!! And this is the essence of Markov's inequality:

Markov's Inequality:

Let $Y \geq 0$ be a non-negative random variable. If $a > 0$, then

$$
P(Y \geq a) \leq \frac{E(Y)}{a}
$$

In other words, Markov's inequality estimates the probability that a non-negative random variable exceeds a certain threshold

Example 3:

Suppose we randomly select an article from a journal where the mean length is known to be 1000 words. Find an upper bound on the probability that an article exceeds 1400 words in length.

Let Y be the length of an article in this journal.

Then by Markov's Inequality

$$P(Y \geq 1400) \leq \frac{E(Y)}{1400} = \frac{1000}{1400} \approx 0.71$$

So there's an at most 71% chance the article exceeds 1400 words

Proof of Markov: Define I_a by

$$I_a = \begin{cases} 1 & \text{if } Y \geq a \\ 0 & \text{if } Y < a \end{cases}$$

Claim: $aI_a \leq Y$

This is because if $Y < a$, then $aI_a = 0 \leq Y$ (since Y is non-negative). And if $Y \geq a$, then $aI_a = a \leq Y$ as well ✓

Using the Claim, we get

$$\begin{aligned} aI_a &\leq Y \\ E(aI_a) &\leq E(Y) \\ a\left[1 \cdot P(Y \geq a) + 0 \cdot P(Y < a)\right] &\leq E(Y) \\ aP(Y \geq a) &\leq E(Y) \\ P(Y \geq a) &\leq E(Y) \end{aligned}$$

4. Chebyshev's Inequality

The second inequality we will look at is Chebyshev's Inequality. The advantage is that it gives a better estimate, but the disadvantage is that you need to know both $E(Y)$ and $\text{Var}(Y)$.

So more information gives you a better result, which makes sense!

Chebyshev's Inequality:

If Y is any random variable and $a > 0$, then

$$P(|Y - E(Y)| \geq a) \leq \frac{\text{Var}(Y)}{a^2}$$

In other words, Chebyshev gives the probability that a random variable deviates from its mean by more than a certain amount

Note: Notice the absolute value sign inside the probability for Chebyshev's Inequality. This means that Chebyshev's Inequality gives a bound *in either direction*. Contrast this to Markov's Inequality, which is a bound on the probability of *exceeding* a certain value.