

Wk4_Notes: Continuous Random Variables

D. ODay

10/24/2021

Reading Notes

Poisson Distribution

Story

If we have many , with very small probabilities, chances at success then we can use the Poisson Distribution to model the number of occurrences of the event. An example would be lottery tickets.

Distribution Label

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

λ is the number of occurrences of the rare events/successes. Also labeled as the rate of occurrence. Example is if we say we expect 5 lottery winners out of all the tickets out there. $\lambda = 5$.

Expectation and Variance

$$\lambda$$

PMF

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

A binomial distribution converges to a poisson distribution when $n \rightarrow \infty$ and $p \rightarrow 0$.

Discrete vs. Continuous

PDF

Like a PMF for continuous random variables. In these, the probability that the continuous random variable takes on a specific value is 0; the intuition behind this is precision. Continuous values can have billions of decimal points. These PDFs return densities, not probabilities

In discrete values, we could use a summation operator to sum up all the different values. This wouldn't work for a PDF since continuous values don't fit into a summation. Officially, we say that a random variable has a PDF $f(x)$ and a CDF $F(x)$ if:

$$F(b) - F(a) = P(a \leq X \leq b) = \int_a^b f(x) dx$$

Properties of PDF

- They are positive
- $\int_{-\infty}^{\infty} f(x) = 1$

To find the probability the random variable took on values greater than a but less than b we would need to use the CDF, not the PDF. To do that, we would need to integrate the CDF

LoTUS

Stands for the the ‘Law of the Unthinking Statistician’.

The $E(x)$ of a continuous random variable is:

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

Can also find the expectation of any function with:

$$E(X) = \int_{-\infty}^{\infty} g(x) \cdot f(x) dx$$

This also holds in the discrete case:

$$E(X) = \int_{-\infty}^{\infty} g(x) \cdot P(X = x)$$

Using LoTUS we can find $E(X)^2$ just by plugging it in for $g(x)$:

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)$$

or for PDF:

$$E(X) = \int_{-\infty}^{\infty} x^2 \cdot P(X = x)$$

Example using the Bernoulli distribution ($X \sim \text{Bern}(.5)$)

$$E(X^2) = \sum_{x=0}^1 (x^2)(0.5^2)(1 - 0.5)^{1-x}$$

Overall, using these tools can be useful when needing to find the variance

Uniform Distribution

Story

Generates a completely random number in some segment. By random, we mean if we divide the spectrum into segments, we randomly choose one of the segments. for example, if we divide the segment into 4 equal pieces, they each have a .25 probability that the random variable will crystalize in it. Probability is proportional to length.

Distribution Label

$$X \sim U[a, b]$$

PDF We know the PDF must be constant. This is because the probability is proportional to length; the PDF cannot change with x . One area of the distribution that is the same length of another, cannot have a higher probability.

$$c = \frac{1}{b - a}$$

##CDF We integrate our PDF, from a to t , to get the following CDF

$$P(X \leq t) = \frac{t - a}{b - a}$$

Expectation

To find our expectation, we can use LOTUS to integrate x times the PDF to find our expectation. Remember, This is the same as what we do in the Discrete context. Multiply and sum all the X values by the probability within the density.

$$E(X) = \frac{b + a}{2}$$

Essentially, this is just the average of the two endpoints.

Variance

Since we know the variance is $E(X^2) - E(X)^2$. We can find $E(X^2)$ by adding x^2 to the pdf and integrating. Eventually, this comes out to:

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

Universality

You can generate any random variable you would like using just the CDF and the standard Uniform distribution. Basically, if you want to create a distribution X that has a CDF F then you just let $X = F^{-1}(U)$. Thus, X will have the CDF F .

In other words, if you start with any CDF you want (provided it's increasing and continuous), then plug in U to the inverse CDF, you have just created a random variable that follows the original CDF.

Example

Say we have the following CDF:

$$F(x) = 1 - e^{-x}$$

We want to simulate a random variable that follows this CDF, but your computer doesn't know how to generate random values with this structure. How would we do it?

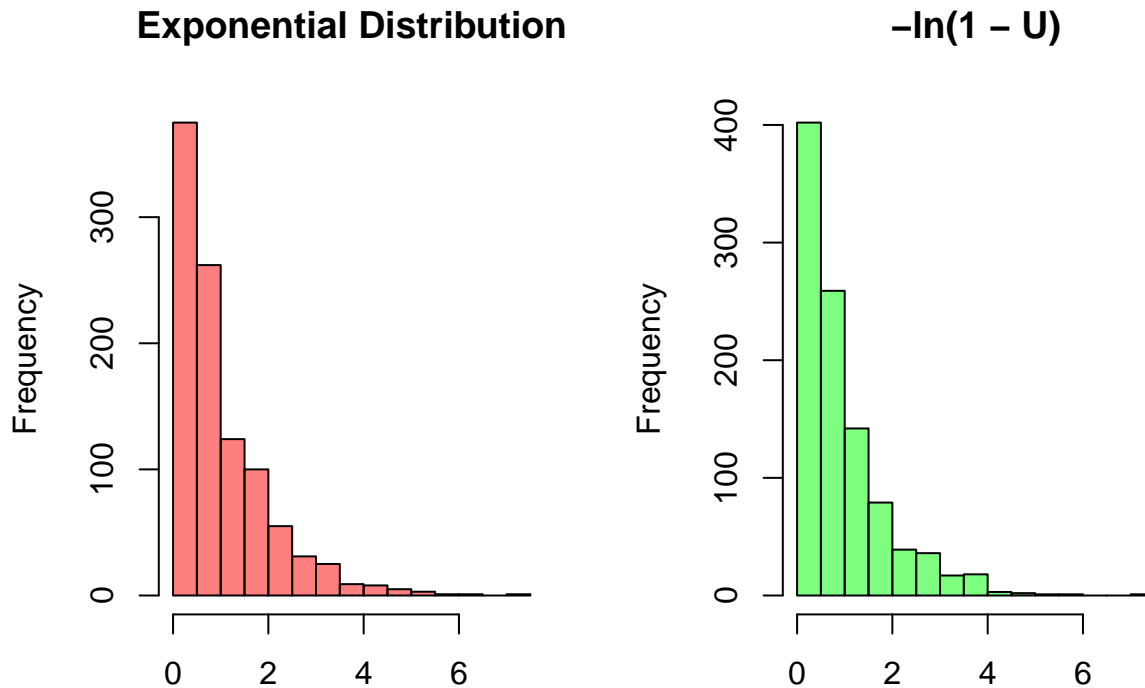
Knowing the Universality of the Uniform, all we have to do is set the CDF equal to U and solve for x :

$$1 - e^{-x} \rightarrow -\ln(1 - U) = x$$

```
#set grid
par(mfrow = c(1,2))

#Exponential
hist(rexp(1000), col = rgb(1, 0, 0, 1/2),
     main = "Exponential Distribution",
     xlab = "")

#transformed Uniform
hist(-log(1 - runif(1000)), col = rgb(0, 1, 0, 1/2),
     main = "-ln(1 - U)", xlab = "")
```



Basically, with this code we simulated U a bunch of times and plugged these values into $-\ln(1 - U)$. This followed the same distribution as the exponential distribution. Think of it this way, when we plug in X into the CDF, it has to result into a value between $[0,1]$ based on the axiom of probability. Therefore, by taking the inverse and using the uniform distribution, plug in a value between 0 and 1 gives us x .

Normal Distribution

Story

- **Central Limit Theorem** - adding up lots of identically and independently distributed random variables with it eventually leading up to a Normal distribution.

Distribution Label

$$X \sim N(\mu, \sigma^2)$$

The standard normal distribution is just a normal distribution centered at 0 with a variance of 1 ($Z \sim N(0, 1)$).

PDF

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

CDF

$$f(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{z^2}{2}} dz$$

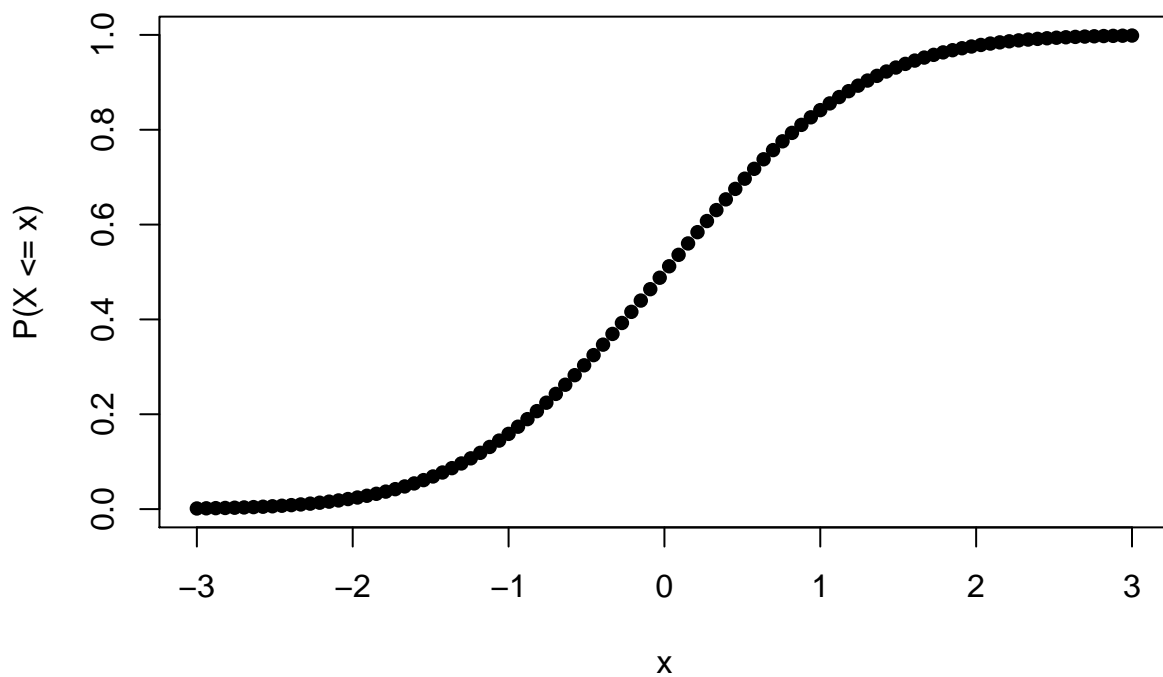
Can calculate in R doing the following:

```
pnorm(0)
```

```
## [1] 0.5
```

```
plot(seq(from = -3, to = 3, length.out = 100), pnorm(seq(from = -3, to = 3, length.out = 100)),  
     xlab = "x", ylab = "P(X <= x)", main = "CDF of X where X ~ N(0, 1)",  
     type = "p", pch = 16)
```

CDF of X where X ~ N(0, 1)



Transformations

We can make any Normal Distribution X with mean μ and variance σ^2 :

$$X = \mu + \sigma Z$$

Let c be a constant. As long as X and Y are independent we have the following:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$$

$$Var(cX) = c^2 Var(X)$$

The process of converting a Normal Distribution back to the Standard Normal is a process called **standardization**

PDF of a Normal Random Variable

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Exponential Distribution

Story

Basically, the distribution gives the time waited for the first success in a continuous interval of time, where λ is the rate parameter for the success. Basically, the continuous version of the Geometric function. ###
Distribution Label

$$X \sim Expo(\lambda)$$

##PDF

$$f(x) = \lambda e^{-\lambda x}$$

The support for the distribution is $x > 0$.

Expectation

$$\frac{1}{\lambda}$$

Variance

$$\frac{1}{\lambda^2}$$

Scaling an Exponential

$$(\lambda X) \sim Expo(1)$$

Lecture Notes

Cdf - We expect the following • $0 \leq F(x) \leq 1$

- limit of x to $-\infty$ is 0 and x to ∞ is 1
- $F'(x) = f(x)$ (Fundamental Theorem of Calculus)
- $F(x)$ is always increasing

Exponential Random Variable

The family of Exponential Distributions provides probability models that are widely used in engineering and science disciplines to describe time-to-event data. Examples are:

- Time until birth
- Time until a light bulb fails
- Waiting time in a queue
- Length of service time
- Time between customer arrivals

Also, we have the memoryless property of the exponential random variable:

$$P(X > s + t | X > s) = P(X > t)$$

Basically, this states no matter what information we know before hand. Our probability will be the same whether we have a 'given' attachment or not.

Normal Random Variable

What assumptions do we make with these problems?

If $X \sim N(\mu, \sigma^2)$, then $\frac{X - \mu}{\sigma} \sim N(0, 1)$

Formulas

Poisson Distribution

PMF

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

Expected Value and Variance

$$\lambda$$

Continuous Distribution

General Formula

$$F(b) - F(a) = P(a \leq X \leq b) = \int_a^b f(x) dx$$

Mean of Distribution (LoTUS)

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

Properties

- They are positive
- $\int_{-\infty}^{\infty} f(x) = 1$
- $F(x)$ is always increasing

Uniform Distribution

PDF

$$c = \frac{1}{b - a}$$

CDF

$$P(X \leq t) = \frac{t - a}{b - a}$$

Expectation

$$E(X) = \frac{b + a}{2}$$

Variance

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

Exponential Distribution

PDF

$$f(x) = \lambda^{-\lambda x}$$

Expectation

$$\frac{1}{\lambda}$$

Variance

$$\frac{1}{\lambda^2}$$