CS 107, Probability, Spring 2020
Lecture 26

Michael Poghosyan
mpoghosyan@aua.am

AUA

06 April 2020

Content

- Some important Continuous rv.s
  - Uniform Distribution;
  - Exponential Distribution;

Important Continuous Distributions

Reminder for Continuous Distributions

Recall that if X is a Continuous r.v., then we describe it by its CDF $F(x)$ or its PDF $f(x)$, and we have the following relationship between F, f and X:

- $F(x) = \mathbb{P}(X \leq x) = \displaystyle\int_{-\infty}^{x} f(t)\,dt$, for all $x \in \mathbb{R}$;
- $f(x) = F'(x)$, for (almost) all $x \in \mathbb{R}$.

And, having F or f, we can calculate probabilities concerning X, say:

$$\mathbb{P}(X \in [a, b]) = F(b) - F(a) = \int_{a}^{b} f(x)\,dx.$$

Uniform Distribution

Uniform Distribution

Distribution Name: $Unif$;
Parameters: a, b $(a < b)$

Uniform Distribution in $[a, b]$ or (a, b)

We say that the r.v. X has a uniform distribution in $[a, b]$ (or in (a, b)), and we will write $X \sim Unif([a, b])$ if its PDF is given by

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a, b]; \\ 0, & \text{otherwise} \end{cases}$$

Exercise: Check that this function is a PDF of some r.v.
The Graph of the PDF: OTB.

The CDF of the Uniform Distribution

Fact: The CDF of $X \sim Unif[a, b]$ is given by:

$$F(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ 1, & x > b \end{cases}$$

Graph: OTB

Idea of Uniform Distribution

Usage: Uniform Distribution Models situations, where our r.v. X can take any value from $[a, b]$, and we want to have "all values equiprobable". In fact, the probability of each value is 0, so we describe equiprobability in the following way: if we will take two subintervals in $[a, b]$ of the **same length**, then the probability that X is in each subinterval is the same. That is, if $A, B \subset [a, b]$ are two intervals (or just subsets) of the same length, $length(A) = length(B)$, then

$$\mathbb{P}(X \in A) = \mathbb{P}(X \in B).$$

And this is because the Density of the Probability is the same everywhere in $[a, b]$, $f \equiv const$ in $[a, b]$!

Idea of Uniform Distribution

For example, if $X \sim Unif([a, b])$, then, for example (assuming $[a, b]$ contains all intervals below)

- $\mathbb{P}\left(X \in \left[a, \frac{a+b}{2}\right]\right) = \mathbb{P}\left(X \in \left[\frac{a+b}{2}, b\right]\right) = \frac{1}{2}$;
- $\mathbb{P}\left(a < X \leq a + \frac{1}{100}\right) = \mathbb{P}\left(a + \frac{3}{100} \leq X \leq a + \frac{4}{100}\right) = \frac{1/100}{b-a}$;
- $\mathbb{P}(a < X < a+h) = \mathbb{P}(b-2h \leq X < b-h) = \mathbb{P}(t < X < t+h) = \frac{h}{b-a}$;
- And, in general, $\mathbb{P}(X \in I) = \frac{length(I)}{b-a}$, for any interval $I \subset [a, b]$.

Uniform Distribution: Remark

Remark: Recall the experiments of picking a random number from some interval - we were talking, actually, about the Uniform Distribution. So if we are picking from $[a, b]$, and if we will denote the randomly chosen number by X, then $X \sim Unif[a, b]$ (if, of course, we are choosing *uniformly* $\ddot\smile$).

Remark: The distribution $Unif([0, 1])$ is usually called the Standard Uniform Distribution, and the r.v. $X \sim Unif([0, 1])$ is called a Standard Uniform r.v..

Uniform Distribution: Example

Example 26.1: Assume $X \sim Unif[1, 4]$.

a. Give some interpretation for X;
b. Write down the PDF of X;
c. Write down the CDF of X;
d. Calculate $\mathbb{P}(X = 2)$;
e. Calculate $\mathbb{P}(X \leq 2)$;
f. Calculate $\mathbb{P}(2 \leq X \leq 3.5)$;
g. Calculate $\mathbb{P}(X > 1.8)$.

Uniform Distribution: Example

Example 26.2: Which of the following r.v.s can be modeled by the Uniform Distribution:

a. Assume that at the Baghramyan Metro Station trains arrive in 10 min intervals, and I am visiting that station at some random time instant. Let X be my waiting time (in minutes) for the train;

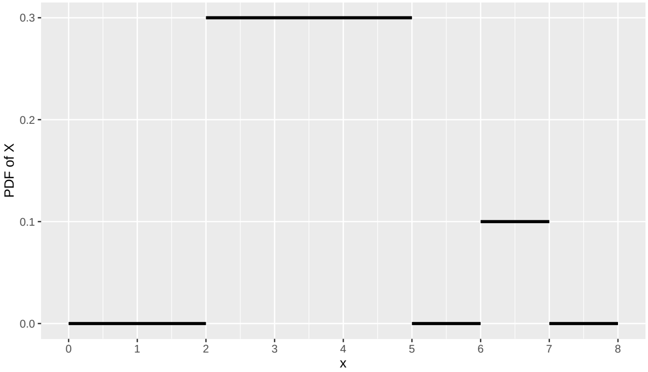
b. $Y =$ the age of a random person at AUA;

c. $Z =$ the number of AUA cafeteria visitors from 12PM till 1PM in a day;

d. Say, at the end of the day, supermarket manager is calculating the weight of the leftover rice from a 10 Kg rice bag. And let W be that weight.

Uniform Distribution: Example

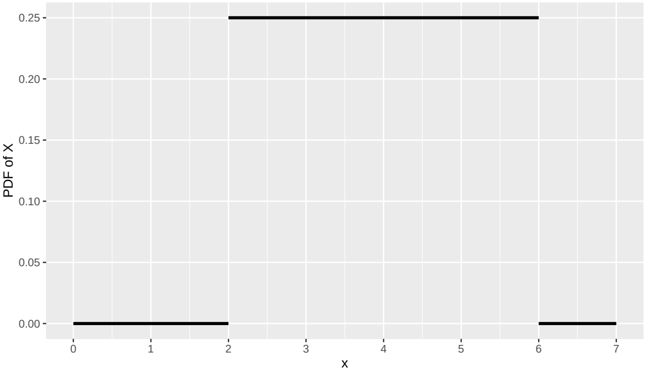
Example 26.3: Assume the PDF of the r.v. X is given by the following graph:



Is X continuous? Is X Uniform? If it is, write the exact distribution (i.e., give the parameters).

Uniform Distribution: Example

Example 26.4: Assume the PDF of the r.v. X is given by the following graph:



Is X continuous? Is X Uniform? If it is, write the exact distribution (i.e., give the parameters).

Uniform Distribution: Example

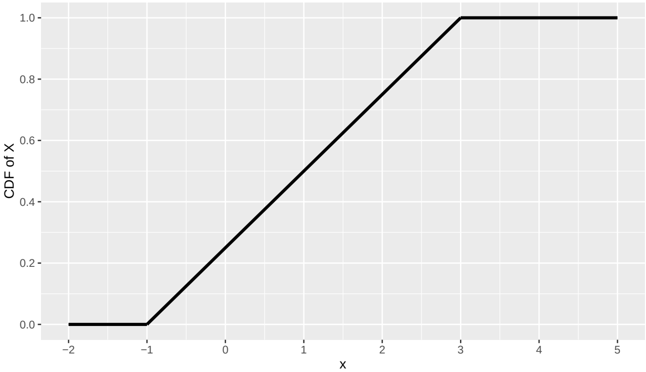
Example 26.5: I have generated, using **R**, a random number from some distribution, and obtained the value

$$2.31030241.$$

is it possible that I have generated from the $Unif([-1, 2])$?

Uniform Distribution: Example

Example 26.6: Assume the CDF of the r.v. X is given by the following graph:



What is the distribution of X?

Uniform Distribution: Example

Example 26.7: Usually, for my quizzes in our AUA classroom, the duration is something between 10min and 12min. Assuming the duration D is uniformly distributed, find the probability that we next quiz will last more than 11.5min.

Uniform Distribution: Example

Example 26.8: Assume $X \sim Unif([-2, 3])$. For any $x \in \mathbb{R}$, calculate

$$\mathbb{P}(X \leq x | X > 0),$$

in other words, calculate the distribution of the r.v.

$$Y = (X | X > 0).$$

Uniform Distribution: Example

Example 26.9:

a. Let $X \sim Unif([0,1])$. Find the distribution of the r.v. $Y = 1 - X$.

b. Assume $X, Y \sim Unif([0,1])$. Is it true that $X = Y$?

Uniform Distribution

Now, if we have a r.v. $X \sim Unif[a, b]$, and assume $[\alpha, \beta] \subset [a, b]$. Then

$$\mathbb{P}(\alpha \le X \le \beta) = \int_\alpha^\beta \frac{1}{b-a} \cdot dx = \frac{\beta - \alpha}{b - a} =$$

$$= \frac{length([\alpha, \beta])}{length([a, b])}.$$

Interpretation: We can interpret X as a randomly (and uniformly!) chosen point from $[a, b]$, and the probability of having that point in $[\alpha, \beta]$ is the length of $[\alpha, \beta]$ divided by the length of $[a, b]$. This is exactly our Geometric Probabilities model (1D case). So Uniform Distribution is the model behind Geometric Probabilities.

Question: Can we choose a point at random, from $[a, b]$, but not uniformly?

Uniform Distribution: Example

Example 26.10:

a. Generate, in **R**, 10 random numbers from $Unif([0,5])$;
b. Now, generate in **R** 1000 random numbers from $Unif([0,5])$, and plot the histogram of that dataset;
c. You are asking me to give 9 uniformly distributed random numbers from $[0,5]$, and I am giving you

$$0.5,\ 1,\ 1.5,\ 2,\ 2.5,\ 3,\ 3.5,\ 4,\ 4.5.$$

Is this a reasonable answer?

Uniform Distribution: Example

Example 26.11: Assume $X \sim Unif([4, 10])$.

a. Let $\alpha = 0.4$. Find the value of $q \in \mathbb{R}$ such that

$$\mathbb{P}(X \leq q) = \alpha;$$

b. Solve the previous problem for any $\alpha \in (0, 1)$.

Remark: The number q, satisfying

$$F(q) = \mathbb{P}(X \leq q) = \alpha,$$

(F is the CDF of X) is called the α-quantile of the distribution of X.

Uniform Distribution: Example

Example 26.12:

a. Assume $X \sim Unif([0,1])$. Let $Y = a + (b-a)\cdot X$. Prove that $Y \sim Unif([a,b])$;

b. Assume $X \sim Unif([a.b])$. Prove that $Y = \dfrac{X-a}{b-a} \sim Unif([0,1])$.

Uniform Distribution: Remarks

Remark: We cannot define a Uniform Distribution on an infinite interval. Say, we cannot define $Unif([0,+\infty))$. So a statement like:

we are choosing a random real number

is not correct, one needs to give what is meant by this - one needs to give the distribution (say, PDF) of the chosen number.

Remark: In fact, we can define a Uniform Distribution over any (measurable!) subset $A \subset \mathbb{R}$, with finite length (measure) $length(A)$. We define $X \sim Unif(A)$, if the PDF of X is given by

$$f(x) = \begin{cases} constant, & x \in A \\ 0, & x \notin A. \end{cases}$$

Uniform Distribution: Remarks

Important Remark: The Uniform Distribution is very important for computer generation of random numbers. It is the basis to generate (pseudo-) random numbers from different Distributions, both Discrete or Continuous, since almost every algorithm of (pseudo-)random number generation is starting first by generating $Unif([0,1])$ numbers.

Uniform Distribution: Example

Example 26.13: Let $X \sim \mathit{Unif}([0,1])$. Now, let y_1, y_2, y_3 be some real numbers and p_1, p_2, p_3 be non-negative numbers with

$$p_1 + p_2 + p_3 = 1.$$

We define

$$Y = \begin{cases} y_1, & \text{if } X \in [0, p_1] \\ y_2, & \text{if } X \in (p_1, p_1 + p_2] \\ y_3, & \text{if } X \in (p_1 + p_2, p_1 + p_2 + p_3] = [p_1 + p_2, 1]. \end{cases}$$

a. Is Y discrete or continuous (or singular) ?
b. If Y is discrete, find its PMF. If it is continuous, find its PDF. Otherwise, do nothing $\ddot{\smile}$;
c. Write an **R** code to generate a random number from

$$Y \sim \begin{pmatrix} -1 & 3 \\ 0.2 & 0.8 \end{pmatrix}$$

Exponential Distribution

Exponential Distribution

Distribution Name: Exp;
Parameters: λ $(\lambda > 0)$

Exponential Distribution

We say that the r.v. X has an Exponential Distribution with the parameter $\lambda > 0$ (rate), and we write $X \sim Exp(\lambda)$, if its PDF is given by

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

Exercise: Check that this function is a PDF of some r.v.
Exercise: Prove that the CDF of $X \sim Exp(\lambda)$ is given by:

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0 \end{cases}$$

Exponential Distribution Examples

Usage: We use the Exponential Distribution to model the waiting time for some event to happen, the time until the next event will happen, or the time until the first time the event will happen.

- Say, the waiting time X until the next order at the GG taxi service during the time interval $9AM$ and $10AM$ can be modeled as an Exponential r.v., $X \sim Exp(\lambda)$.
- Say, the time between clicks on the webpage in some fixed time interval can be modeled as an Exponential r.v.;
- The time between goals scored in a World Cup football match; Time between meteors greater than 1 meter diameter striking earth; Time between successive failures of a machine; Time until the discovery of a new Bitcoin block, Time a storekeeper must wait before the arrival of their next customer; Time between late arrivals to Prob class ...

Exponential Distributions: Example

Example 26.14: Assume $X \sim Exp(2.6)$.

a. Give some interpretation for X;
b. Write down the PDF of X;
c. Write down the CDF of X;
d. $\mathbb{P}(X = 2.6)$;
e. $\mathbb{P}(X \leq 2)$;
f. $\mathbb{P}(X \geq -2)$;
g. $\mathbb{P}(2 \leq X \leq 3.5)$;
h. $\mathbb{P}(X > 3 | X > 1)$.

Exponential Distributions: Example

Example 26.15: Assume $X \sim Exp(\lambda)$. Prove that

$$\mathbb{P}(X > t) = e^{-\lambda \cdot t}, \qquad \forall t \geq 0.$$

Note: The function

$$S(t) = \mathbb{P}(X > t)$$

is usually called the **Survival or Reliability function**. If we will interpret X as a lifetime of something, then $X > t$ will mean that the lifetime is larger than t.

Exponential Distribution: Example

Example 26.16:

a. Generate, in **R**, 10 random numbers from $Exp(2)$;
b. Now, generate in **R** 1000 random numbers from $Exp(2)$, and plot the histogram of that dataset.

Exponential Distribution: Example

Example 26.17: Plot and compare the PDFs of two different Exponential Distributions, say, for

$$Exp(0.5) \qquad \text{and} \qquad Exp(2).$$

Exp and Pois Distributions

Note: There is an important relation between the Poisson and Exponential Distributions: assume we are considering the occurrence of some rare events in some unit of time (say, in 10min):

- Poisson r.v. is calculating the **number of events (occurrences)** in that time interval; and it is discrete;
- Exponential r.v. is calculating the **time between two successive events**, also the **time until the first event will happen**; and it is continuous

The Rate Parameter

Note: If we model by $Exp(\lambda)$ the waiting time for something, then the rate parameter λ is the inverse of the average waiting time. That is, if the average time between two phone calls is 2.5min, then we can model the waiting time X (in min) as

$$X \sim Exp\left(\frac{1}{2.5}\right).$$

This is a result from Statistics; also, we will consider the notion of the mean value of a r.v., and show that for $X \sim Exp(\lambda)$, the mean value of X is $\frac{1}{\lambda}$. Hence, if we interpret X as a waiting time, then we will obtain

$$\text{mean of } X = \text{average of waiting time} = \frac{1}{\lambda}.$$

Exponential Distribution: Example

Example 26.18: Assume that the average duration of a taxi order is 22 min. Let T be the duration of a taxi order, and assume T can be modeled by the Exponential distribution.

a. Write the distribution of T;
b. What is the probability that the duration of the order will be larger than 30 min?
c. What is the probability that the duration of the order will be between 20 and 30 min? Calculate in two ways, using both the PDF and CDF.

Note: Now, assume again that we are considering the occurrence of some rare events in some unit of time. Let Y be the Poisson r.v. showing the number of events in that time interval, $Y \sim Pois(\lambda)$, and let X be the Exponential r.v. measuring the waiting time for (time between) the events, in the same units, $X \sim Exp(\lambda_1)$. We want to give some relation between λ and λ_1. Now,

- The Poisson parameter λ is counting the average number of events in that time interval;

- The Exponential parameter λ_1 is the inverse of the average waiting time. Now, if, in average, λ events happen in the unit time interval, then the average waiting time will be $\frac{1}{\lambda}$ units, so $\lambda_1 = \lambda$.

Rate Parameters; Exp and Pois Distributions

Note: (cont'd) So if X and Y are, respectively, the waiting time and number of events (in the same experiment!) in a unit time, then

$$X \sim Exp(\lambda) \qquad \text{and} \qquad Y \sim Pois(\lambda),$$

where λ is the rate: the average number of events in a unit time, or the inverse of the average waiting time between that events.

Note: And, in general, if some events happen with the waiting time between them $X \sim Exp(\lambda)$ in some units, then, for the number Y of events in the time interval of length T units, we will have

$$Y \sim Pois(\lambda \cdot T).$$

Note: Let me obtain the PDF of $Y \sim Exp(\lambda)$ r.v., using the above relation between the Poisson and Exponential r.v.s. Let X be the r.v. counting the number of events in the unit of time, $X \sim Pois(\lambda)$. And let Y be modeling the waiting time between the events, in the same time units. Let's find the CDF/PDF of Y, just by using this relation with the Poisson r.v.

Now, assume one of that Poisson events just happened. Then Y will show the time that will pass until the next event will happen. We fix some $t \geq 0$, and consider the probability of $Y > t$:

$$\mathbb{P}(Y > t) = \mathbb{P}(\text{the next event will take more than } t \text{ time}) =$$

$$= \mathbb{P}(\text{No event in } [0, t]).$$

Exp and Pois Distributions

Now, to calculate this Probability, let X_t be the r.v. showing the number of events in $[0, t]$. Since the average number of events was λ in 1 time unit, then the average number of events in t time units, say, in $[0, t]$, will be $\lambda \cdot t$. Hence, $X_t \sim Pois(\lambda \cdot t)$. This yields, for any $t \geq 0$,

$$\mathbb{P}(Y > t) = \mathbb{P}(\text{No event in } [0, t]) = \mathbb{P}(X_t = 0) = e^{-\lambda \cdot t}.$$

This shows that $\mathbb{P}(Y \leq t) = 1 - e^{-\lambda \cdot t}$, for $t \geq 0$. Clearly, also $\mathbb{P}(Y \leq t) = 0$, if $t < 0$ (since Y is nonnegative). This means that the CDF of Y is of the form

$$F_Y(t) = \begin{cases} 1 - e^{-\lambda \cdot t}, & t \geq 0, \\ 0, & t < 0 \end{cases}$$

and, by calculating $f(t) = F'(t)$, we will obtain the PDF of Y, which coincides with the formula given in the definition of $Exp(\lambda)$ distribution.

Exponential Distributions: Example

Example 26.19: Assume that phone calls at some taxi service are arriving at a rate 2 calls in 5 min (i.e., we have, in average, 2 calls in 5 min). Let X be the number of calls in 10 min, and let WT be the waiting time (in min) between two successive calls.

a. Give appropriate models for X and WT;
b. Calculate the probability that the no. of calls in 10 min will be less than 4;
c. Calculate the probability that we will wait less than 1 min for the first call;
d. Calculate the probability that we will wait more than 3 min for the next call.

Memoryless Property

The following property is a very important characteristic of the Exponential Distribution: this distribution is the only continuous distribution with the Memoryless Property:

Memoryless Property

a. If $X \sim Exp(\lambda)$, then

$$\mathbb{P}(X > t) = e^{-\lambda \cdot t}, \qquad \forall t > 0;$$

b. $X \sim Exp(\lambda)$ if and only if

$$\mathbb{P}(X > t + s | X > t) = \mathbb{P}(X > s), \qquad \forall t, s \in [0, +\infty).$$

Memoryless Property

Interpretation: If X is the waiting time for smthng, then $\mathbb{P}(X > s)$ is the Probability that the waiting time will be more than s. Now, $\mathbb{P}(X > t + s|X > t)$ is the Probability that the waiting time will be more than $t+s$, if you have waited already t units of time. That is, waiting for *another* s units is independent how long you have already waited!

Recall that the only Discrete Distribution sharing the Memoryless Property:

$$\mathbb{P}(X > m + n|X > m) = \mathbb{P}(X > n), \qquad \forall m, n \in \{0, 1, 2, ...\}$$

is the Geometric Distribution (which can be described as the waiting time for the Bernoulli process).

Memoryless Property

So, according to the Memoryless property, we can use the Exponential Distribution to model the lifetime of something. But that something needs to satisfy the following: it *does not age*, in the sense that the Probability of functioning yet another time is independent on its current age. So our something is not wearing out gradually, it is just, at some unpredictable time, stopping functioning (say, like memory chips).

Note: As we noted, the only two distributions with Memoryless propery are the Exponential (continuous) and Geometric (discrete) ones. And no wonder that it can be shown that the Exponential r.v. can be obtained as a limit of appropriately chosen Geometric r.v.s.

Normal Distribution