CS 107, Probability, Spring 2020
Lecture 23

Michael Poghosyan
mpoghosyan@aua.am

AUA

20 March 2020

Content

- Some important Discrete rv.s
  - Discrete Uniform Distribution;
  - Bernoulli Distribution;
  - Binomial Distribution;
  - Geometric Distribution

Reminder on rvs

Let us summarize what we have learned about r.v.s:

- R.v. X over some Experiment is a function taking real values for any outcome of our experiment. In words, we are calculating some quantity for an outcome;
- Usually we are not interested in outcomes, but rather in that quantity calculated, we are interested in the values of a r.v. X. So we are forgetting the Experiment and Sample Space behind our r.v. and talk about possible values of X and probabilities of that values.

Reminder on rvs

- We can describe any r.v. X completely (i.e., we can calculate Probabilities of the form $\mathbb{P}(X \in A)$), if we know its CDF;
- For a Discrete r.v., another complete characteristic of a r.v. is its PMF;
- For a Continuous r.v. another complete characteristic of a r.v. is its PDF;

Note: Besides these complete characteristics (the CDF and PD(M)F), there are others, that we are not covering in our course: two important ones are Characteristic Functions and Moment Generating Functions (the latter is not defined for any r.v.). Some properties of r.v. are easier to get/prove using these characteristics.

Reminder on Discrete rvs

- R.v X is Discrete, if it is taking only finite or countably infinite many values.
- Discrete r.v. X can be described through its PMF:

Values of X	x_1	x_2	x_3	...
$\mathbb{P}(X = x)$	p_1	p_2	p_3	...

Another motivation for a r.v.

In the first part of our course, we were dealing with different probability problems, mostly of combinatorial and geometric ones. Say, the following is a standard example of that type of problems:

Example 23.1: We are rolling two fair dice, independently.

a. What is the probability of having 2 sixes ?
b. What is the probability of having exactly one six?
c. What is the probability that we will have no sixes at all?

Now, if we will denote by *NoS* the number of sixes, then *NoS* will be a r.v., with possible values $0, 1, 2$. Now, constructing the PMF of *NoS* is the same as to solve the above 3 problems:

Values of *NoS*	0	1	2
$\mathbb{P}(NoS = x)$			

Example 23.1

Solution 1: Let us calculate

- $\mathbb{P}(NoS = 2)$
- $\mathbb{P}(NoS = 1)$
- $\mathbb{P}(NoS = 0)$

Example 23.1

Solution 2: We can visualize the Sample Space in the form of a table:

	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Example 23.1

Now we can construct the PMF of our r.v. *NoS*:

Values of *NoS*	0	1	2
$\mathbb{P}(NoS = x)$			

The idea of this example was that the Distribution of a this r.v. replaces the solution of the above three problems **a.-c.** So if you can find/identify the Distribution of a r.v., you can easily solve problems like the ones in **a.-c.**, as particular cases.

Now, let us forget the Experiments and Sample Spaces behind, and talk about the values of a r.v. and the probabilities of that values.

Reading/interpreting PMFs

Example 23.2: Assume X is a discrete r.v. given by its PMF:

Values of X	-1	1	4	6
$\mathbb{P}(X = x)$	0.2	0.1	0.6	0.1

a. Plot the PMF by a hand;
b. Plot the PMF in **R**;
c. Give the frequentist interpretation for the PMF;
d. Do a computer (**R**) experiment by generating 1000 values for X and calculating which percent of these values are equal to 4

Some important discrete univariate rvs

Notations

- From this lecture on, we will describe a r.v. X by giving its PMF (for the Discrete case) or PDF (for the Continuous case), and we will read/say that *"The Distribution of* X *is given by .... "* or *"*X *is .... Distributed "*;
- We will give names to standard Distributions, and write

$$X \sim DistributionName$$

  meaning that X is a r.v. with the Distribution described by $DistributionName$. Usually, our Distributions will have parameters, and we will write that parameters after the $DistributionName$, like $Bernoulli(p)$.
- So we will say, for example, X is a Bernoulli r.v. (with the parameter p), or the distribution of X is $Bernoulli(p)$, or X is a $Bernoulli(p)$ distributed

Notations

Example 23.3: Assume X is the discrete r.v. from the Example 23.1, given by its PMF:

Values of X	-1	1	4	6
$\mathbb{P}(X = x)$	0.2	0.1	0.6	0.1

Then (not a common notation) we can write

$$X \sim \begin{pmatrix} -1 & 1 & 4 & 6 \\ 0.2 & 0.1 & 0.6 & 0.1 \end{pmatrix}$$

Notations

Example 23.4: Assume

$$X, Y \sim \begin{pmatrix} \text{-10} & 10 \\ 0.5 & 0.5 \end{pmatrix}$$

a. Give the Distribution of X and Y. Are X, Y Discrete?
b. What is the Experiment behind X? Y?

Probability Distributions in **R**

The following names are used in **R** for distributions:

Distribution	R name		Distribution	R name
Beta	beta		Lognormal	lnorm
Binomial	binom		Negative Binomial	nbinom
Cauchy	cauchy		Normal	norm
Chisquare	chisq		Poisson	pois
Exponential	exp		Student t	t
F	f		Uniform	unif
Gamma	gamma		Tukey	tukey
Geometric	geom		Weibull	weib
Hypergeometric	hyper		Wilcoxon	wilcox
Logistic	logis			

Probability Distributions in R

Usage:

Name	Description
dname( )	PD(M)F of distribution
pname( )	CDF
qname( )	quantile function
rname( )	random numbers generated from distribution

Example:

```
rnorm(10)
dunif(1)
pexp(3, rate = 2)-pexp(-3, rate = 2)
qchisq(0.3, df = 1)
```

Probability Distributions in **R**

Note that the distributions above have some parameters (say, min and max for Uniform, mean and sd for Normal etc.), and some parameters have default values you can keep (say, for Normal Distribution one has mean=0 and sd=1 by default).

Discrete Uniform Distribution

Discrete Uniform Distribution

Distribution Name: Discrete Uniform, *DiscreteUnif*;
Parameters: $x_1, x_2, ..., x_n$ $(x_k \in \mathbb{R})$.

Discrete Uniform Distribution

We say that a r.v. X has a Discrete Uniform Distribution with (over) the values $x_1, x_2, ..., x_n$ $(x_i \neq x_j,\ i \neq j)$, and we will write $X \sim DiscreteUnif(x_1, ..., x_n)$, if

Values of X	x_1	x_2	...	x_n
$\mathbb{P}(X = x)$	$\frac{1}{n}$	$\frac{1}{n}$	...	$\frac{1}{n}$

that is, $\mathbb{P}(X = x_k) = \dfrac{1}{n}$, for any $k = 1, 2, ..., n$.

Discrete Uniform Distribution

So

$$X \sim DiscreteUnif(x_1, x_2, ..., x_n)$$

will denote that

$$X \sim \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ \dfrac{1}{n} & \dfrac{1}{n} & \dots & \dfrac{1}{n} \end{pmatrix}$$

Discrete Uniform Distribution: Examples

Usage: Discrete Uniform Distribution models equiprobable outcomes case.
Some examples:

- We are tossing a fair coin, and $X = 0$ if Heads appears, and $X = 1$ in the other case. Then

  $$X \sim \mathit{DiscreteUnif}(0, 1),$$

  meaning that

  $$X \sim \begin{pmatrix} 0 & 1 \\ 0.5 & 0.5 \end{pmatrix}$$

Discrete Uniform Distribution: Examples

- We are rolling a fair die, X is the top face number. Then

$$X \sim \mathit{DiscreteUnif}(1, 2, 3, 4, 5, 6),$$

i.e.

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$$

- Can you give some more?

Discrete Uniform Distribution

Example 23.5: Assume

$$X \sim \mathit{DiscreteUnif}(\{-1, 0, 4, 5\}).$$

a. Calculate $\mathbb{P}(X = 4)$;
b. Calculate $\mathbb{P}(X \geq 0)$.

Discrete Uniform Distribution: R Examples

R Code

```
#Discrete Uniform on data
data <- c(-1,2,4)
sample(data, size = 10, replace = T)

#Generating a sample of size 100
s <- sample(data, size = 100, replace = T)
#Calculating the number of -1 in the sample s
length(s[s == -1])
```

Bernoulli Distribution

Bernoulli Distribution

Distribution Name: Bernoulli;
Parameters: p, $(p \in [0, 1])$;

Bernoulli Distribution

We say that a r.v. X has a Bernoulli Distribution with the parameter (probability) $p \in [0, 1]$, and we will write $X \sim Bernoulli(p)$, if it has the following PMF:

Values of X	0	1
$\mathbb{P}(X = x)$	$1 - p$	p

that is, $\mathbb{P}(X = 1) = p$, and $\mathbb{P}(X = 0) = 1 - p$.

Bernoulli Distribution

So

$$X \sim Bernoulli(p)$$

means that

$$X \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$$

Note: The PMF of $Bernoulli(p)$ can be written in the form:

$$\mathbb{P}(X = x) = p^x \cdot (1-p)^{1-x}, \quad x \in \{0, 1\},$$

if[1] $p \in (0, 1)$.

Note: Usually, we decode 1 as **success**, and 0 as **failure**. So p is the **probability of success**.

[1] To avoid numbers like 0^0.

Bernoulli Distribution: Examples

Usage: Bernoulli Distribution models the binary outcomes case, the Success/Failure type results: Yes/No, Boy/Girl, Pass/Fail, Healthy/Ill, Smoker/Non-Smoker, Flan/Fstan, ...

Some examples:

- We are tossing a fair coin, and $X = 0$ if Heads appears, and $X = 1$ in the other case. Then $X \sim \textit{Bernoulli}\left(\frac{1}{2}\right)$;

- We are rolling a fair die. $X = 0$, if the shown number is > 4, and $X = 1$ otherwise. Then $X \sim \textit{Bernoulli}(\quad)$;

Bernoulli Distribution: Examples

- We are choosing a Prob student at random in AUA. We denote $X = 1$ in case the student will pass the course, and $X = 0$ otherwise. Say, we can model X as a Bernoulli r.v.: $X \sim \mathit{Bernoulli}(0.76)$;
- We are interested if a randomly chosen Prob student will get the A grade from Probability: we can take $X = 1$ to indicate that he/she will get A, and $X = 0$ otherwise. Then $X \sim \mathit{Bernoulli}(p)$, and p is the probability of having A from Probability $\ddot\smile$
- We are interested if the patient has a specific disease or not: we can take, for example, $X = 1$, if he/she has that disease, and $X = 0$ otherwise. p will be the probability of having that disease, and $X \sim \mathit{Bernoulli}(p)$;

Bernoulli Distribution: Examples

- We are interested if the nearest toilet room is busy at this moment $\ddot{\smile}$: if it is, we can denote that by $X = 1$, and $X = 0$ otherwise. Then $X \sim \mathit{Bernoulli}(p)$, but I do not have any ideas about p;
- An insurance company is interested if the claim size for today will exceed 10M AMD: we can take $X = 1$, if it will exceed and $X = 0$ if not. Again $X \sim \mathit{Bernoulli}(p)$ for some p.
- We are interested if the max stock price for FB next month will be higher than the max stock price today: we can take $X = 1$, if that max price next month will be larger, and $X = 0$ otherwise.

Bernoulli Distribution: Examples

- We are interested if some specific Football team will win the next game: $X = 1$ means will win, and $X = 0$ otherwise. Then $X \sim \mathit{Bernoulli}(p)$;
- We are interested if some specific Football team will win the next game with the difference in the scores more than 3: $X = 1$ if that will be the case, and $X = 0$ otherwise. Then $X \sim \mathit{Bernoulli}(p)$;
- And you can make infinitely many examples of Bernoulli distributed r.v.s

Interpretation of Bernoulli r.v.

Another interpretation of the Bernoulli r.v. X is that it is an Indicator r.v.:

- We have an experiment, and an event A in that experiment;
- the probability of A is $p = \mathbb{P}(A)$;
- If A happens, then $X = 1$, otherwise $X = 0$.

Bernoulli Distribution

Example 23.6: Assume

$$X \sim \mathit{Bernoulli}(0.2).$$

a. Calculate $\mathbb{P}(X \in [0.1, 0.7])$;
b. Calculate $\mathbb{P}(X \in [0.7, \sqrt{2}])$.

Bernoulli Distribution: R Examples

R Code

```
#Bernoulli Distribution
p <- 0.3
#In fact, we use the Binomial Distrib with size=1
x <- rbinom(10, size = 1, prob = p)
x
```

Binomial Distribution

Binomial Distribution

Distribution Name: *Binom*;
Parameters: n, p $(n \in \mathbb{N},\ p \in [0, 1])$

Binomial Distribution

We say that a r.v. X has a Binomial Distribution with the parameters n and p, $p \in [0, 1]$, and we will write $X \sim Binom(n, p)$, if it has the following PMF:

Values of X	0	1	2	...	n
$\mathbb{P}(X = x)$	$\binom{n}{0} p^0 (1-p)^{n-0}$	$\binom{n}{1} p^1 (1-p)^{n-1}$	$\binom{n}{2} p^2 (1-p)^{n-2}$	...	$\binom{n}{n} p^n (1-p)^{n-n}$

that is,

$$\mathbb{P}(X = x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}, \qquad x = 0, 1, 2, ..., n$$

Exercise: Check that the sum of probabilities is 1.

Binomial Distribution: Interpretation

Usage: Binomial Distribution gives the model for the repeated trials.

If $X \sim Binom(n, p)$, then

- we are doing n independent repetitions of some Simple Experiment,
- in each Simple Experiment some Event can happen (Success) with the Probability p;
- X shows the number of Successes we will have during that n trials

Binomial Distribution: Examples

- We are tossing a fair coin 12 times, and we are interested in how many heads we will get during that tosses. If we will denote by X the number of heads we will have, then $X \sim \mathit{Binom}(12, 0.5)$.
- Assume we have 5 red and 10 green balls in the box. We are drawing a ball at random, fixing its color, returning it to the box, and then doing that again 100 times. We are interested in the number of green balls we will pick. If we will denote by X the number of green balls taken out of the box during this 100 trials, then $X \sim \mathit{Binom}(100, 2/3)$, since the probability to get a green ball in each trial is $p = \frac{10}{15} = \frac{2}{3}$ (we assume drawing each ball has the same probability for all balls).
- Your examples?

Binomial Distribution

Example 23.7: Assume

$$X \sim Binom(20, 0.4) \qquad \text{and} \qquad Y \sim Binom(20, 0.4).$$

a. Calculate $\mathbb{P}(X > 2)$;
b. Calculate $\mathbb{P}(X \text{ is Even})$;
c. Calculate $\mathbb{P}(Y \text{ is Even})$.

Example 23.8: 5% of population in Armenia is using VK regularly. We are choosing 50 persons at random (with replacements!). What is the probability that more than 20 of them will be a regular VK users?

Binomial Distribution: R Examples

R Code

```
#Binomial Distribution
p <- 0.45
#Size is our n
x <- rbinom(20, size = 3, prob = p)
x

#Plotting the PMF
size = 21
x <- 0:size
PMF <- dbinom(x, size = size, prob = p)
plot(x,PMF, pch = 19)
```

Binomial Distribution

Example 23.9: Plot in **R** and compare the PMFs of two Binomial r.v.s with the same no. of trials, but different values of p.

Example 23.10: One of the Jazz Clubs has 40 seats, and asks for making reservations for that seats. Historically, 25% of persons who make reservations are not showing up. The manager of our Club is a smart guy, with some Probability knowledge, and is accepting a little bit more reservations than 40, i.e., the number of places - not to have many empty seats during a performance. But then there is a chance that everybody will show up, and some persons who reserved seats, will not have a place to sit on. Our manager wants to limit this risk to 5%. What is the maximal number of reservations that he can accept?

Solution:

Geometric Distribution

Geometric Distribution

Distribution Name: *Geom*;
Parameters: p ($p \in (0, 1]$)

Geometric Distribution

We say that a r.v. X has a Geometric Distribution with the parameter p, $p \in (0, 1]$, and we will write $X \sim Geom(p)$, if its PMF is given by

$$\mathbb{P}(X = x) = p \cdot (1 - p)^{x-1}, \qquad x \in \mathbb{N}$$

that is,

Values of X	1	2	3	$\dots$	k	$\dots$
$\mathbb{P}(X = x)$	p	$p(1-p)$	$p(1-p)^2$	$\dots$	$p(1-p)^{k-1}$	$\dots$

- **Exercise:** Prove that the probabilities sum up to 1.
- **Exercise:** Calculate the CDF of $X \sim Geom(p)$.

Geometric Distribution: Interpretation

Usage: Geometric Distribution gives a model for counting the number of repeated trials until the first Success.
If $X \sim Geom(p)$, then

- we are doing independent repetitions of some Simple Experiment;
- in each Simple Experiment some Event can happen (Success) with the Probability p;
- X shows the number of trials we will have until the first Success will be shown.

In other words, Geometric r.v is the Discrete **waiting time** (for the Success) r. v.

Geometric Distribution: Note

Recall that the definition of the Geometric Distribution was:

Values of X	1	2	3	...	k	...
$\mathbb{P}(X = x)$	p	$p(1-p)$	$p(1-p)^2$	...	$p(1-p)^{k-1}$	...

X is counting **the number of trial we will have until the first Success**.

In some books, people define the Geometric Distribution in the other way, starting from the value 0:

Values of X	0	1	2	...	k	...
$\mathbb{P}(X = x)$	p	$p(1-p)$	$p(1-p)^2$	...	$p(1-p)^{k}$	...

In this case X shows **the number of Failures before the first Success**.

Geometric Distribution: Examples

- Assume we are rolling a die several times until first time we will get 5 shown. Then p is the probability to get 5 in one roll, and we know that for a fair die $p = \frac{1}{6}$. Let X be the number of roll that will show 5 first time. Then $X \sim Geom(\frac{1}{6})$.
- Assume a couple wants to calculate the number of children they will have until the first boy. Assume also we know the probability of having a boy, $p = 0.49$, and suppose that having a boy or girl is independent of the previous children. Let X be the number of children they will have until the first boy (say, $X = 1$ means that the first child was a boy, $X = 3$ means the first 2 children were girls and 3rd one was a boy etc.). Then $X \sim Geom(0.49)$.

Geometric Distribution: Examples

- Say, we are asking randomly AUA students until we will find a GPA 4.0 student. Then the number of students we will ask until meeting a GPA 4.0 student is $X \sim Geom(p)$, where p is the probability of AUA student to be a 4.0 GPA student (the ratio of 4.0 GPA students to all AUA students). BTW, this can give you a method to estimate the proportion (i.e., the probability p) of AUA 4.0 GPA students. Just go out and randomly ask students until getting the first 4.0 GPA-shnik. Then use some statistics to estimate p $\ddot\smile$
- Your turn!

Geometric Distribution

Example 23.11: Assume

$$X \sim Geom(0.8)$$

a. Calculate $\mathbb{P}(X > 2)$;
b. Calculate $\mathbb{P}(X \in (4, 8))$.

Geometric Distribution

Example 23.12: Two persons, A and B, are playing the following game, many times: every time B is keeping in a mind an integer from 1 to 10 (including), and A tries to guess that number (we assume independence in games). What is the probability that A will give the first correct guess in less than 20 plays/trials?

Geometric Distribution: R Examples

R Code

```
#Geomteric Distribution
p = 0.7
smpl <- rgeom(20, prob = p)
smpl

#Plotting the PMF
n = 10
x <- 0:n
PMF <- dgeom(x, prob = p)
plot(x,PMF, pch = 19)
#Now, the CDF
t <- seq(-2,5, 0.1)
CDF <- pgeom(t, p)
plot(t, CDF, type = "s")
```

Example 23.13: Plot and compare the PMFs of two Geometric r.v.s with different Probabilities, say, compare the PMFs of

$$X \sim Geom(0.4) \qquad \text{and} \qquad Y \sim Geom(0.8).$$

Geometric Distribution

An important property of a Geometric Distribution is the *Memoryless* (forgetfulness) property:

Memoryless Property of Geometric Distribution

If $X \sim Geom(p)$, then

- For any $n \in \mathbb{N}$,
$$\mathbb{P}(X > n) = (1-p)^n;$$
- for any $n, m \in \mathbb{N}$,
$$\mathbb{P}(X = m + n | X > n) = \mathbb{P}(X = m);$$
- for any $n, m \in \mathbb{N}$,
$$\mathbb{P}(X > m + n | X > n) = \mathbb{P}(X > m).$$

Example: Assume that the probability that a customer will buy Suriki Lavash is 0.05. Then the following probabilities are equal:

- The 5-th customer will buy Suriki Lavash;
- The 15-th customer will buy Suriki Lavash, given that the first 10 customers have not bought Suriki Lavash.

Remark: An important remark is that the **only** Discrete r.v. with the Memoryless Property is the Geometric r.v. . Mathematically, if X is a Discrete r.v. with the Range $\{1, 2, 3, 4, 5, ...\}$, and satisfying

$$\mathbb{P}(X > m + n | X > n) = \mathbb{P}(X > m), \qquad \forall n, m \in \mathbb{N},$$

then there exists $p \in (0, 1]$ such that $X \sim Geom(p)$.
Remark: Later we will learn that the only Memoryless Continuous r.v. is the Exponential one.

Poisson Distribution