CS 107, Probability, Spring 2020
Lecture 28

Michael Poghosyan
mpoghosyan@aua.am

AUA

10 April 2020

Content

- Transformations of r.v.s (functions of r.v.s)

Important Distributions

So far we have studied:

- The general theory of r.v.s, using their CDF or PD(M)F;
- Some important Discrete Distributions:
  - Discrete Uniform;
  - Bernoulli;
  - Binomial;
  - Geometric;
  - Poisson;
  - Negative Binomial;
  - Hypergeometric;
- Some important Continuous Distributions:
  - Uniform;
  - Exponential;
  - Normal

Important Distributions

But please note that this is a small list out of all important, named r.v.s. For example, if you are navigating through some Distribution page at Wikipedia, below, at the end of the page, you will find a list of different distributions.

Also here:

`http://www.math.wm.edu/~leemis/chart/UDR/UDR.html`

you will find a list of important univariate (1D) Distributions, along with the relationships between that distributions.

Let me just name some interesting distributions:

Some Continuous Distributions

- **Beta Distribution:** $X \sim Beta(\alpha, \beta)$, $(\alpha, \beta > 0)$, if the PDF of X is

$$f(x) = \begin{cases} C \cdot x^{\alpha-1} \cdot (1-x)^{\beta-1}, & x \in [0,1] \\ 0, & x \notin [0,1]. \end{cases}$$

- **Gamma Distribution:** $X \sim Gamma(\alpha, \beta)$, $(\alpha, \beta > 0)$, if the PDF of X is

$$f(x) = \begin{cases} C \cdot x^{\alpha-1} \cdot e^{-\beta x}, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

- **Student's t Distribution:** Wait for the Stat course!
- χ^2 **Distribution:** Wait for the Stat course!
- F **Distribution:** Wait for the Stat course!

Benford Distribution

Do you think all digits are equal? I mean, in the sense of rights? $\smile$

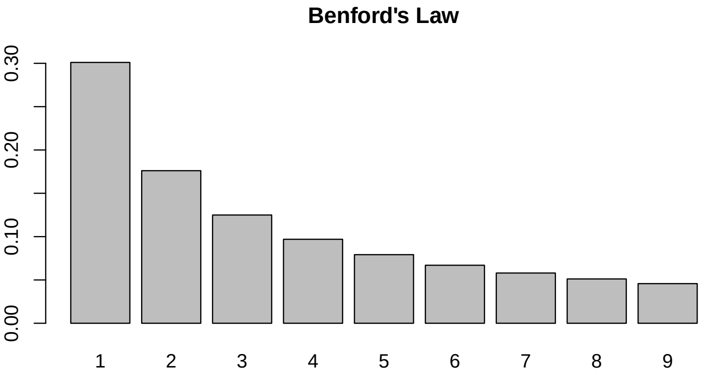
Benford's law

In many numerical datasets, the chance that the first digit of the chosen number will be $k \in \{1, 2, ..., 9\}$ is

$$\mathbb{P}(X = k) = \log_{10}\left(1 + \frac{1}{k}\right).$$

Here X is a Benford r.v. showing the first (non-zero) digit of a randomly chosen number from that dataset.
Say, the chance that our dataset number will start by 1 is 30.1%, by 2 is 17.6% etc.

Benford Distribution



Benford Distribution

Few resources for the Benford's law:

- https://www.youtube.com/watch?v=5HZdJlbxlOk
- https://www.youtube.com/watch?v=vIsDjbhbADY
- https://testingbenfordslaw.com/twitter-users-by-followers-count

Zipf Distribution

Zipf's Distribution

Assume $s \geq 0$ and $N \in \mathbb{N}$ are fixed. The r.v. X with a Zipf Distribution is given by its PMF

$$\mathbb{P}(X = k) = C \cdot \frac{1}{k^s}, \qquad k = 1, 2, ..., N,$$

where C is the normalizing constant:

$$C = \left(\sum_{k=1}^{N} \frac{1}{k^s} \right)^{-1}.$$

Zipf Distribution

Why is Zipf's law important?

- https://www.youtube.com/watch?v=fCn8zs912OE
- https://en.wikipedia.org/wiki/Zipf's_law

Pareto Distribution

Pareto Distribution

Assume $a > 0$ and $\alpha > 0$ are fixed. The r.v. X has a Pareto Distribution with parameters a and α, if its PDF is given by

$$f(x) = \begin{cases} \dfrac{C}{x^{\alpha+1}}, & x \geq a \\ 0, & x < a \end{cases}$$

where C is the normalizing constant.

Pareto Distribution

Why is Pareto Distribution important? Have you heard about the famous Pareto 80-20 principle?

- https://www.youtube.com/watch?v=fCn8zs912OE
- https://www.youtube.com/watch?v=TcEWRykSgwE
- https://www.youtube.com/watch?v=U4GMUamUjT8
- https://en.wikipedia.org/wiki/Pareto_distribution
- https://en.wikipedia.org/wiki/Pareto_principle

Transformations of Random Variables: Making new R.V.s from the old ones

Functions of Random Variables

Given a r.v. X, one can form new r.v.s by applying functions on X. In fact, we already used transforms several times, say, when doing the Normal r.v. Standardization: for a Normal r.v. $X \sim \mathcal{N}(\mu, \sigma^2)$, the Standardization is:

$$Y = \frac{X - \mu}{\sigma}.$$

Here Y is a new r.v. obtained from X by applying some function (and Y is Standard Normal!).
Also, we talked about $X \sim \mathit{Unif}([0, 1])$, and noted that

$$Y = a + (b - a) \cdot X$$

is a new r.v. which has a Uniform Distribution on $[a, b]$: $Y \sim \mathit{Unif}([a, b])$.

Functions of Random Variables: Examples

Example: Let X be the price of the 1 FB share at the end of the next week, so X is a r.v. If my portfolio consists of 50 FB shares, then my portfolio price at the end of the next week will be $50 \cdot X$, which is a r.v. too.

Example: Now, again, let X (in USD) be the closing price of the 1 FB share in 5 days, and let my portfolio consist of 50 FB shares, and also of a cash in a deposit account, which is earning at the rate 1% annually (no capitalization of interests, simple interest). Assume I made the cash investment into that account with the size $\$1000$, 55 days ago. Let Y be the price of my portfolio at the end of that 5 days. Then

$$Y = 50 \cdot X + 1000 \cdot \left(1 + 0.01 \cdot \frac{60}{365}\right) = 50 \cdot X + 1001.64,$$

and Y is a new r.v. obtained from X.

Functions of Random Variables

Now, if X is a r.v., and $g : \mathbb{R} \to \mathbb{R}$ is some function[1] (it is enough to have g defined on the Range of X), then the r.v.

$$Y = g(X)$$

is called the transformation of X.

Example: For example,

$$Y = X^3, \quad L = 2X - 1, \quad Z = \ln(X), \quad U = \frac{X}{1+X}, \quad W = \sin(X)$$

are all (if defined, of course) new r.v.s obtained from the r.v. X.

[1] (18+) Measurable!

Example: Or, if X is a r.v., then we can take the function

$$g(t) = \begin{cases} t^2, & t < 3 \\ 1 - \text{arctg}(t), & t \geq 3, \end{cases}$$

and obtain the r.v.

$$Y = \begin{cases} X^2, & X < 3 \\ 1 - \text{arctg}(X), & X \geq 3, \end{cases}$$

Well, of course, we need to be sure that our new r.v.s are well-defined.
Example: If $X \sim Unif([-2, 2])$, then $Y = \sqrt{X}$ is not well-defined. Or, if $X \sim Pois(4)$, then $Y = \ln(X)$ is not well-defined.

Describing Transformations of X

Assume X is a r.v., and $Y = g(X)$ is the transformation of X. The general problem here is

Describe the Distribution of Y, having the Distribution of X.

For a warm-up, try to guess the answers of the following questions:

Example: Assume X is Discrete Uniform. What can be said about $Y = X^2$?

a. Will Y be discrete?

b. If yes, will Y be Discrete Uniform?

Example: Assume $X \sim Unif[-2, 2]$.

a. What can be said about $Y = 3X + 1$? Will it be Continuous? Will it be Uniform in some interval?

b. What about $Z = X^2$?

Describing Transformations of X through CDFs

Describing Transformations of X through CDFs

Assume X is a r.v. (either discrete or continuous, or other) with the CDF $F_X(x)$, and $Y = g(X)$ is the transformation of X.

Problem:

Find the CDF of Y in terms of the CDF of X, $F_X(x)$.

Solution: The general idea: let $F_Y(x)$ be the CDF of Y. Then, by definition,

$$F_Y(x) = \mathbb{P}(Y \leq x) = \mathbb{P}(g(X) \leq x).$$

At this point we need to solve $g(X) \leq x$ in terms of X. Then we can express F_Y in terms of F_X.

Describing Transformations of X through CDFs

Solving $g(X) \leq x$ in terms of X can be not so simple (and sometimes, impossible) - that depends on g. Let us consider two particular cases now, when we can solve that easily.

Case 1: g is **strictly increasing**. Then g is invertible, the inverse g^{-1} is defined, and, for any $x \in \mathbb{R}$,

$$\{g(X) \leq x\} = \{X \leq g^{-1}(x)\}, \quad \text{so}$$

$$F_Y(x) = \mathbb{P}(g(X) \leq x) = \mathbb{P}(X \leq g^{-1}(x)) = F_X(g^{-1}(x)).$$

Case 2: g is **strictly decreasing**. Then it is again invertible, and, for any $x \in \mathbb{R}$,

$$F_Y(x) = \mathbb{P}(Y \leq x) = \mathbb{P}(g(X) \leq x) = \mathbb{P}(X \geq g^{-1}(x)) =$$

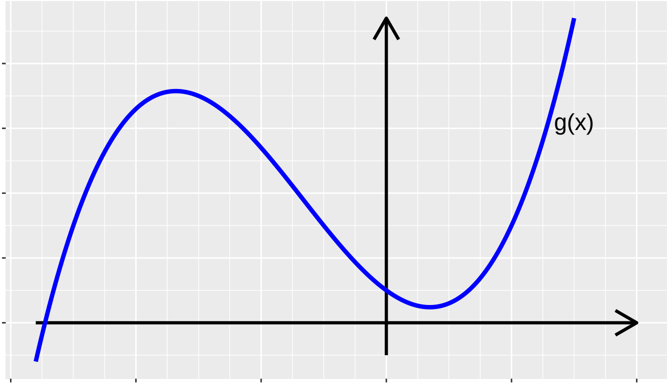
$$= 1 - \mathbb{P}(X < g^{-1}(x)) = 1 - F_X\big(g^{-1}(x) - \big).$$

Describing Transformations of X through CDFs

In the general case, as stated above, one needs to solve

$$g(X) \leq x$$

in terms of X, to express F_Y in terms of F_X.

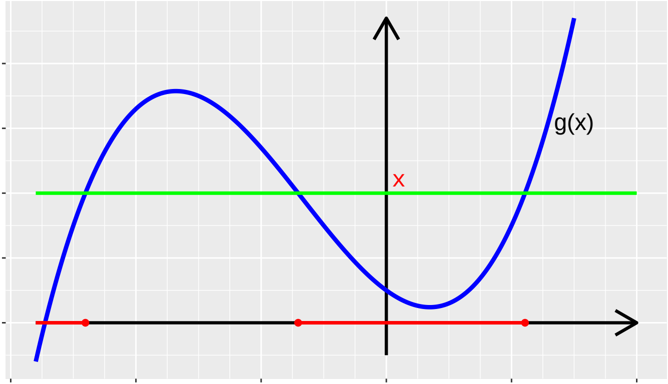


Describing Transformations of X through CDFs

In the general case, as stated above, one needs to solve

$$g(X) \leq x$$

in terms of X, to express F_Y in terms of F_X.



Transformations: Example

Example 28.1: Assume $X \sim Unif[-1, 3]$, and $Y = 3 - X$. Find the CDF of Y, and describe the distribution of Y.

Solution:

Transformations: Example

Example 28.2: Assume $X \sim Exp(\lambda)$, and $Y = X^2$. Find the CDF of Y.

Solution:

Discrete r.v. Transforms through PMF

Describing Transformations of X through PMFs

Assume X is a Discrete r.v., and $Y = g(X)$ is the transformation of X. Then Y will be a Discrete r.v.

Problem:

Express the PMF of Y in terms of the PMF of X.

Solution: Easy, very easy: say, the PMF of X is:

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

Then the PMF of $Y = g(X)$ will be:

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

Describing Transformations of X through PMFs

Important Note: Above, if two values $g(x_i)$ and $g(x_j)$ coincide for $i \neq j$, then we write $g(x_i)$ only once, and add the corresponding Probabilities: $p_i + p_j$.
Formally,

$$\mathbb{P}(Y = y) = \mathbb{P}(g(X) = y) = \sum_{g(x_k)=y} p_k.$$

Example 28.3: Assume

$$X \sim \begin{pmatrix} 1 & 2 & 3 \\ 0.2 & 0.1 & 0.7 \end{pmatrix}.$$

a. Find the distribution of $Y = X - 3$;
b. Find the distribution of $Y = \ln(X)$.

Solution:

Discrete r.v. Transforms: Example

Example 28.4: Assume X is a Discrete r.v. given by its PMF

$$X \sim \begin{pmatrix} -2 & 0 & 1 & 2 \\ 0.1 & 0.3 & 0.2 & 0.4 \end{pmatrix}.$$

Find the PMF of $Y = X^2 + 2$.

Solution:

Discrete r.v. Transforms: Example

Example 28.5: Assume $X \sim Bernoulli(p)$.

a. Find the distribution of $Y = 1 - X$;
b. Find the distribution of $Z = X^2$;
c. Will the r.v. $V = -X$ be again Bernoulli r.v.?

Discrete r.v. Transforms: Example

Example 28.6: Assume $X \sim Binom(3, 0.7)$.

a. Find the distribution of $Y = X - 1$. Will it be again Binomial r.v.?
b. Find the distribution of $Z = 2 \cdot X$. Will it be again Binomial r.v.?
c. Find the distribution of $T = 3 - X$. Will it be again Binomial r.v.? Interpret and explain the result.

Discrete r.v. Transforms: Example

Example 28.7: Assume $X \sim Binom(n, 0.5)$, and let $Y = n - X$. Show that $Y \sim Binom(n, 0.5)$.

Note: This result means that X, $Y \sim Binom(n, 0.5)$, so X and Y have the same distribution (but $X \not\equiv Y$). In this case we say that X and Y are **Identically Distributed**, X and Y are ID.

Discrete r.v. Transforms: Example

Example 28.8: Assume $X \sim Geom(p)$.

a. Will $Y = 2 \cdot X$ be Geometric r.v.?
b. Will $T = X^2 - 1$ be Geometric r.v.?
c. Find the distribution of $Z = X - 1$, and give an interpretation;
d. If $U \sim NBinom(n, p)$, find the distribution of $V = U - n$, and give an interpretation.

Discrete r.v. Transforms: Example

Example 28.9: Assume $X \sim Pois(\lambda)$. Describe the distribution of $Y = 3 - 2X$. Is it again Poisson r.v.?

Continuous r.v. transforms through PDF

Describing Transformations of X through PDFs

We have already considered how to find the CDF of the transformed r.v., having the CDF of the original r.v. Now, assume X is a Continuous r.v. with the PDF $f_X(x)$, and $Y = g(X)$. Then, we can have the following cases:

- Y is Continuous;
- Y is Discrete;
- other

Example: Assume $X \sim Unif[0,4]$, and $g(t) = t^3$, so $Y = g(X) = X^3$. Then Y is continuous, and we can find its PDF.

Example: Assume $X \sim Unif[0,4]$, and $g(t) = [t]$, so $Y = g(X) = [X]$. Then Y is discrete, and we can find its PMF.

Describing Transformations of X through PDFs

First, we consider the case when Y is continuous. Here, our problem is:

Problem:

Express the PDF of Y, $f_Y(x)$, in terms of the PDF of X, $f_X(x)$.

Well, the general scheme is the following: let $F_X(x)$ be the CDF of X and $F_Y(x)$ be the CDF of Y. First we express F_Y in terms of F_X, we already know how to do that. Then we use

$$f_Y(x) = \Big(F_Y(x)\Big)'$$

to find the PDF of Y.

Describing Transformations of X through PDFs

For example, assume $g(x)$ is **strictly increasing** and differentiable, and $Y = g(X)$. We know that

$$F_Y(x) = F_X(g^{-1}(x)).$$

Then,

$$f_Y(x) = \frac{d}{dx}F_Y(x) = \frac{d}{dx}F_X(g^{-1}(x)) =$$

$$= F'_X(g^{-1}(x)) \cdot (g^{-1}(x))' = f_X(g^{-1}(x)) \cdot (g^{-1}(x))'.$$

Describing Transformations of X through PDFs

Now assume $g(x)$ is **strictly decreasing** and differentiable, and $Y = g(X)$. In this case we have

$$F_Y(x) = 1 - F_X\big(g^{-1}(x) - \big) = 1 - F_X\big(g^{-1}(x)\big),$$

because F_X is continuous everywhere. Then,

$$f_Y(x) = \Big(F_Y(x)\Big)' = -F'_X(g^{-1}(x))\cdot(g^{-1}(x))' = -f_X(g^{-1}(x))\cdot(g^{-1}(x))'$$

Note: We can combine these two cases (when g is strictly increasing or decreasing) in the following way: If $Y = g(X)$, where g is either strictly increasing or decreasing differentiable function, then

$$f_Y(x) = f_X(g^{-1}(x)) \cdot \left|(g^{-1}(x))'\right|.$$

Describing Transformations of X through PDFs

Summarizing, if we have the PDF f_X of a r.v. X, and we want to find the PDF f_Y of $Y = g(X)$, we perform the following steps:

- We express the CDF of Y in terms of the CDF of X;
- We calculate the derivative of the CDF of Y, to find the PDF of Y.

Continuous r.v. Transforms: Example

Example 28.10: Assume $X \sim Unif([0,1])$. Find the PDF of

$$Y = X^2$$

a. Step by step, by doing all calculations;
b. By using the above PDF transform formula;
c. Check the result in **R**: generate 100000 random numbers from standard uniform distribution, calculate their squares, plot the histogram of relative frequencies of that squared dataset. Then add the theoretical PDF to that histogram.

Continuous r.v. Transforms: Example

Example 28.11: Assume $X \sim Exp(\lambda)$. Find the PDF of

$$Y = \sqrt{X}$$

a. Step by step, by doing all calculations;
b. By using the above PDF transform formula.

Example 28.12: Assume X is a r.v. with the PDF

$$f_X(x) = C \cdot x^5, \quad x \in [0,1], \qquad f(x) = 0, \text{ otherwise}$$

and assume $Y = X^4 + 2$. Find the PDF of Y.

Continuous r.v. Transforms: Example

Example 28.13: Assume $X \sim \mathit{Unif}([-1,2])$. Find the distribution of

a. $Y = g(X)$, if
$$g(t) = \begin{cases} 0, & t < 1 \\ 1, & t \geq 1 \end{cases}$$

b. $Z = [X] + 2$.

Continuous r.v. Transforms: Example

Example 28.14: Assume $X \sim Exp(0.3)$. Find the distribution of $Y = 2 \cdot [X]$.

Example 28.15: Assume $X \sim Exp(\lambda)$. Find the distribution of $Y = 3 \cdot X$.

Remark: In general, if $X \sim Exp(\lambda)$, then for any $\alpha > 0$, $\alpha \cdot X \sim Exp\left(\frac{\lambda}{\alpha}\right)$.

Interpretation: Say, $X \sim Exp(\lambda)$ is measuring the waiting time for some event, in days. Then the average waiting time will be $\frac{1}{\lambda}$ days. If, say, Y is the waiting time in hours, $Y = 24 \cdot X$, then the average waiting time will be

$$\frac{1}{\lambda} \text{ days} = \frac{24}{\lambda} \text{ hours}, \qquad \text{so } Y \sim Exp\left(\frac{\lambda}{24}\right), \quad \text{in hours.}$$

Question: If $X \sim Exp(\lambda)$, then what about $X - 3$, will it be Exponential?

Continuous r.v. Transforms: Example

Example 28.16: Assume $X \sim \mathcal{N}(0,1)$. Find the distribution of $Y = e^X$.

Remark: The Distribution of Y is called LogNormal distribution with parameters 0 and 1: $Y \sim LNorm(0,1)$. This name is because $\ln(Y)$ is Normal, $\ln(Y) \sim \mathcal{N}(0,1)$.

Continuous r.v. Transforms: Example

Example 28.17: Assume $X \sim Unif([0, 20])$. Find the shape of the distribution of

$$Y = \frac{X^2}{1+X} + \ln(X),$$

in **R**.

Note on the linear transform

Let us note some facts about **linear transforms** of r.v.s:

- If $X \sim Unif[a, b]$ and $\alpha > 0$, then

$$\alpha \cdot X + \beta \sim Unif([\alpha \cdot a + \beta, \alpha \cdot b + \beta]);$$

- If $X \sim Exp(\lambda)$ and $\alpha > 0$, then

$$\alpha \cdot X \sim Exp\left(\frac{\lambda}{\alpha}\right);$$

- If $X \sim \mathcal{N}(\mu, \sigma^2)$ and $\alpha \neq 0$, then

$$\alpha \cdot X + \beta \sim \mathcal{N}(\alpha \cdot \mu + \beta, \alpha^2 \cdot \sigma^2).$$

Example 28.18:

a. Assume X is a continuous r.v. with a strictly increasing CDF $F(x)$. We define

$$Y = F(X).$$

Show that $Y \sim Unif[0,1]$;

b. Assume $Y \sim Unif[0,1]$, and $F(x)$ is a continuous, strictly increasing function with $F(-\infty) = 0$ and $F(+\infty) = 1$. We define

$$X = F^{-1}(Y).$$

Show that the CDF of the r.v. X is F.

Note: Informally,

$$Y = F(X) \qquad \Leftrightarrow \qquad X = F^{-1}(Y)$$

with $Y \sim Unif([0,1])$ and $X \sim F$ (i.e., the CDF of X is F).

On previous Example

Example 28.19:

a. If $X \sim \mathcal{N}(0,1)$, and $\Phi(x)$ is its CDF, then what is the distribution of $Y = \Phi(X)$?

b. Find a r.v. for which the CDF is the Sigmoid function:

$$\sigma(x) = \frac{1}{1+e^{-x}}.$$

c. To check the last result, generate 10000 random numbers from that distribution, plot in **R** their Empirical CDF, and put the graph of the Sigmoid function over that.

On previous Example

Note: The results obtained in the above Problem are very important. This is one of the methods to generate random numbers from any distribution given by strictly increasing CDF $F(x)$, and is called the Inverse Transformation Method for RNG (Random Number Generation).

The method is as follows: say, we want to generate random numbers from the distribution given by strictly increasing CDF $F(x)$. Then we can do the following:

- calculate the inverse function $F^{-1}(x)$;
- generate standard uniform random numbers;
- calculate the value of F^{-1} at that generated numbers.

Note: The drawback of this method is that we need to be able to calculate the inverse function $F^{-1}(x)$ for any x.

On previous Example

Note: The above example was for any **strictly increasing** CDF $F(x)$. Nice thing is that the above method works also in the case when F^{-1} is well-defined on $(0, 1)$, i.e., when F has one of the following forms:

$$F(x) = \begin{cases} 0, & x < a \\ \text{strictly increasing}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

or

$$F(x) = \begin{cases} 0, & x < a \\ \text{strictly increasing}, & x \geq a \end{cases}$$

or

$$F(x) = \begin{cases} \text{strictly increasing}, & x \leq b \\ 1, & x > b \end{cases}$$

Example 28.20: Using only standard uniform random numbers,

a. generate random numbers from the $Exp(2)$ distribution in **R**;
b. plot the relative frequency histogram of that dataset;
c. plot over the histogram the PDF of $Exp(2)$ to compare.