CS 107, Probability, Spring 2020
Lecture 34

Michael Poghosyan
mpoghosyan@aua.am

AUA

29 April 2020

Content

- Partial Numerical Characteristics: Expectation of a r. v.

Partial Numerical Characteristics of a r.v.

Partial Characteristics of R.V.s: Motivation

So far we have considered Complete Characteristics of r.v.s: we were studying r.v.s given by their CDFs or PMFs/PDFs. Unfortunately, in real life, these complete characteristics are not given. Say, if X is measuring the

- the height of a (random) person;
- the insurance claim size for today for a company;
- the GPA of a (random) student at AUA;
- the AMZN stock closing price today;
- the no. of GG Taxi calls or some FB page visitors;
- the running time of Quicksort algo for a (random) input,

then nobody will give us the CDF/PDF/PMF of X.

So we want to give, at least, some partial information/characteristics for these r.v.s.

Partial Characteristics of R.V.s: Motivation

Two most important characteristics of a r.v. are

- the Expected Value or Expectation;
- the Variance and the Standard Deviation.

Other important characteristics are

- Skewness;
- Kurtosis;
- Higher order moments;
- Higher order central moments.

In this lecture, we will talk about the Expectation of a r.v. We will assume we know the distribution of a r.v., and we will define and study the properties of a the Expected value.

Expectation of a r.v.

Expectation of a R.V.

First, let's consider our previous example: if X is a r.v. showing

- the height of a (random) person;
- the insurance claim size for today;
- the GPA of a (random) student at AUA;
- the AMZN stock closing price today;
- the no. of GG Taxi calls or FB page visitors;
- the running time of Quicksort algo for a (random) input;

then one of the most important information is the Average Value, the Mean or the Expectation of X: say, the average height of a person, the average daily claim size, average running time of QuickSort algo on an array of size n etc.

Expectation of a R.V.

And if a tourist is interested in visiting Armenia in September, then he will not ask to give the temperature distribution in Armenia in September, but he will be interested in the average temperature.

So the first of our Partial Characteristics will be the Expectation of a r.v.

Note: We will denote the Expectation of a r.v. X by

$$\mathbb{E}(X) \qquad \text{or} \qquad \mu_X,$$

and read it as the Expectation of X, or the Expected Value of X, or the Mean Value of X, and, sometimes, the Average Value of X. Another interpretation/name will be a typical value of X.

We will give the definition of the Expected Value separately for the Discrete and Continuous r.v.s.

Expectation of a Discrete r.v.

Expectation of a Discrete r.v.

Expected Value of a Discrete r.v.

Assume X is a Discrete r.v. with the PMF

$$X \sim \begin{pmatrix} x_1 & x_2 & \dots \\ p_1 & p_2 & \dots \end{pmatrix}$$

Then the **Expected Value** or the **Expectation** of X is the following number:

$$\mathbb{E}(X) = \sum_i x_i \cdot p_i = \sum_i x_i \cdot \mathbb{P}(X = x_i),$$

if $\sum_i |x_i| \cdot p_i < +\infty$ (this is important if the sum is infinite).

Expectation: Example

Example 34.1: Assume

$$X \sim \begin{pmatrix} -1 & 0 & 3 \\ 0.3 & 0.1 & 0.6 \end{pmatrix}$$

Calculate $\mathbb{E}(X)$.

Expectation of a R.V.

Note: The Expectation of a Discrete r.v. X is the weighted average of its values, where the weights are the corresponding probabilities.

In other words, if we think about our distribution as a Probability Mass distribution, then the Expected Value is the Center of Mass.

Expectation of a R.V.

Example: For example, if $X \sim DiscreteUnif(\{a, b\})$, then

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

Here, the weights/probabilities are $\frac{1}{2}$ and $\frac{1}{2}$ for a and b, and we are obtaining that the Expectation is just the ordinary average of a and b.

Expectation of a R.V.

Example: If $Y \sim Bernoulli(0.99)$, then $\mathbb{E}(Y)$ is close to 1:

$$\mathbb{E}(Y) = 0 \cdot 0.01 + 1 \cdot 0.99 = 0.99.$$

Here, the chances that we will have 1 are much higher than we will have 0. And the Expectation is shifted towards 1: the weight of 0 is small, 0.01, and the weight of 1 is large, 0.99.

Interpretation: If, approximately, in 100 numbers generated from Y, 99 are 1, then the average value of that 100 numbers will be approximately 0.99. Or, maybe more convincing, if we will generate 1M numbers from this distribution, then approximately 99% of them will be 1s, and the average of all that 1M numbers will be approximately 0.99.

Expectation of a R.V.

Note: The Expected Value of X is not necessary a value of X. Say, the Expected Value of points X we will get when rolling a fair die is 3.5, but we can't roll 3.5 $\ddot\smile$. Then what the Expected Value in this case shows?

We will see later that it shows the long range average of the values of X: if we will keep rolling our die, fixing the outcomes and calculating the average of all outcomes received so far, then that averages will approach 3.5 when we will roll our die again and again and again.

Expectation: Example

Example 34.2: Assume we are playing an American Roulette. The board has 37 numbers, from 0 to 36, and an additional slot 00. The slots 0 and 00 are in green, half of the other numbers are in black, and the other half are in red. I am betting on red. If the ball will stop at a red pocket, I will get $1\$$, otherwise I will loose $1\$$. Calculate my expected winning.

Expectation: Example

Example 34.3: Assume

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

Find $\mathbb{E}(X)$.

Expectation: Example

Example 34.4: Assume

$$X \sim Bernoulli(p).$$

Find $\mathbb{E}(X)$.

Expectation: Bernoulli Distribution

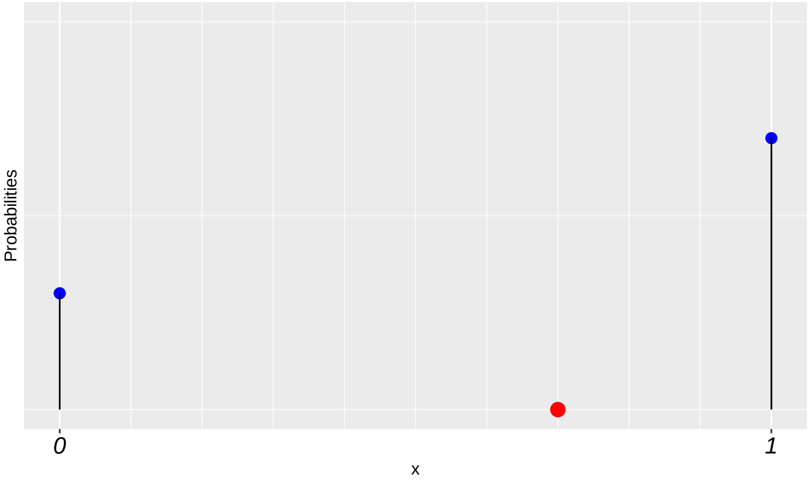

Figure: PMF of $Bernoulli(0.7)$, red dot is the Expected Value

Expectation: Example

Example 34.5: Assume

$$X \sim Binom(n, p).$$

Find $\mathbb{E}(X)$.

Interpretation: If in an Experiment, the chances that Success will happen are p, and if we will repeat that Experiment independently n times, then we expect to have $n \cdot p$ Successes! Say, we have a box containing 80% black and 20% white balls, and we are picking at random 100 balls, with replacements. Our Success is to have a white ball. Then we expect to see 20 white balls, which is exactly $n \cdot p = 100 \cdot 0.2$.

Expectation: Binomial Distribution

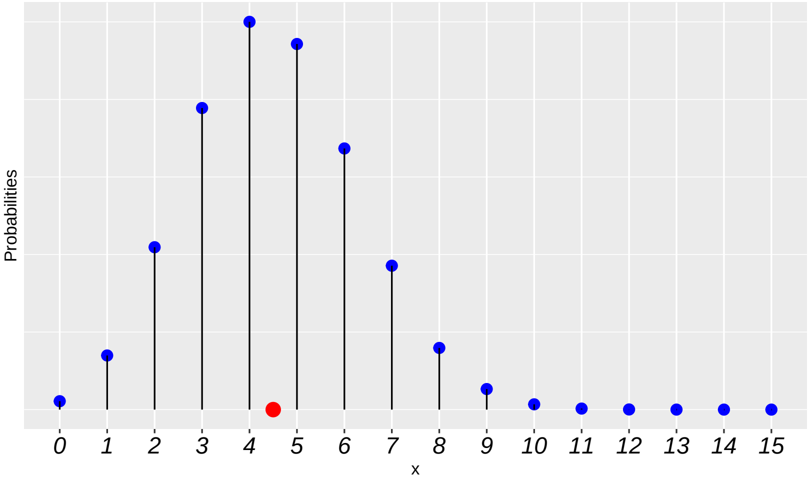

Figure: PMF of $Binom(15, 0.3)$, red dot is the Expected Value

Expectation: Example

Example 34.6: Assume

$$X \sim Geom(p).$$

Calculate $\mathbb{E}(X)$.

Interpretation: Recall that X is showing the number of trials until having the first Success. Now, if the probability of having a Success in one trial is p, then we expect to have $\frac{1}{p}$ trials until having the first Success.

Say, if the Probability of Success in one Trial is $\frac{1}{3}$, then we expect to have one success in 3 trials.

Example: What is the average number of rolls to get 3 on a fair die?

Expectation: Geometric Distribution

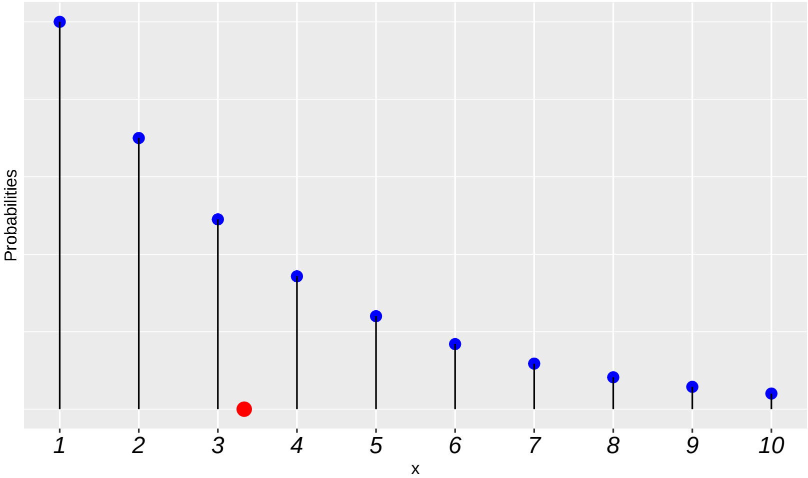

Figure: PMF of $Geom(0.3)$, red dot is the Expected Value

Expectation: Example

Example 34.7: Assume

$$X \sim Pois(\lambda).$$

Calculate $\mathbb{E}(X)$.

Interpretation: Recall that X was showing the number of some (rare) events in a unit of time. Now, this result is showing that the average, expected value of X is λ, and this explains why we were choosing the parameter value λ when modeling with the Poisson Distribution.

Expectation: Poisson Distribution

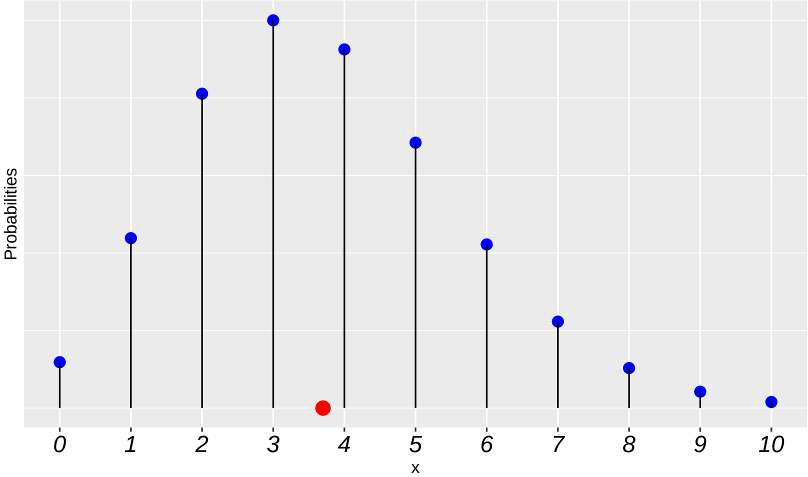

Figure: PMF of $Pois(3.7)$, red dot is the Expected Value

Expectation: Example

Example 34.8: A freshly-formed family plans to have children until the first boy will be born. Assume that the child-births are independent, and the probability of a boy child to be born is 0.5, calculate the expected number of children they will have. Interpret the result.

Exercise: What if they will stop after the second boy, not the first one?

Expectation: Example

Example 34.9: Assume I found a mobile phone, and it has a 4 digit password. I want to unlock the phone (to find out who is the owner and return the phone, of course $\ddot\smile$). Now, I will try some 4-digit numbers at random. Because I am not fixing that numbers, I can have repetitions (so my trials are independent). What is the expected number of trials to unlock the phone?

Exercise: What if I am doing it wisely, writing down and not repeating already entered number?

Expectation: Example

Example 34.10: Casino is suggesting the following game: you are tossing a fair coin, until having the first Heads. If the first Heads appears on the n-th toss, then you get $2^n\,\$$.

a. Calculate your Expected winning;

b. How much casino should ask for this game?

c. Will you pay 1024$\$$ to play this game?

Note: This is very famous St. Petersburg Paradox. You can read about this in almost every Probability textbook, and in many other places, say, `https://plato.stanford.edu/entries/paradox-stpetersburg/`.

Exercise: Write an **R** code to simulate this game.

Expectation of a Continuous r.v.

Expectation of a Continuous r.v.

Expected Value of a Continuous R.V.

Assume X is a Continuous r.v. with the PDF $f(x)$. Then the **Expected Value** or the **Expectation** of X is the following number:

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

assuming that $\displaystyle\int_{-\infty}^{\infty} |x| \cdot f(x)\, dx < +\infty$.

Expectation: Example

Example 34.11: Assume X is a Continuous r.v. with the following PDF:

$$f(x) = \begin{cases} K \cdot (x+2), & x \in [1,2] \\ 0, & otherwise \end{cases}$$

a. Find K;
b. Calculate $\mathbb{E}(X)$.

Expectation: Example

Example 34.12: Assume $X \sim Unif[a, b]$. Find $\mathbb{E}(X)$.

Interpretation: If we will generate a lot of numbers from $[a, b]$, uniformly, and calculate their average, that will be approximately

$$\frac{a+b}{2}.$$

Expectation: Uniform Distribution

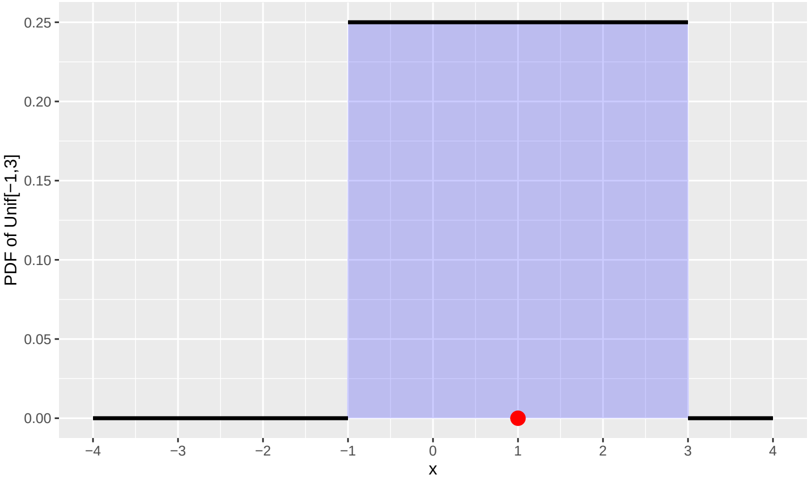

Figure: PDF of $Unif([-1,3])$, red dot is the Expected Value

Expectation: Example

Example 34.13: Assume $X \sim Exp(\lambda)$.

a. Find $\mathbb{E}(X)$;
b. Use **R** simulations to show the obtained result.

Interpretation: Recall that, when modeling with Exponential Distribution, we were using, say $X \sim Exp(\lambda)$ to measure the waiting time for some event, and we were choosing the parameter λ as the inverse of the average waiting time. And this comes from

$$\lambda = \frac{1}{\mathbb{E}(X)} = \frac{1}{\text{(theoretical) average value of } X}.$$

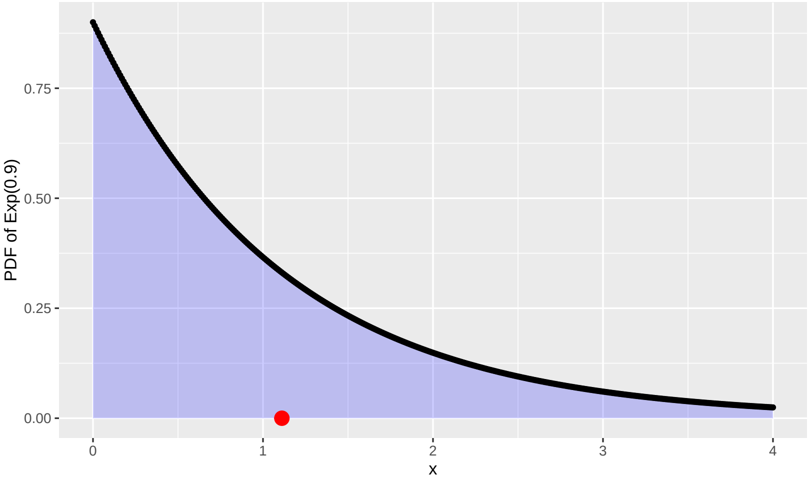

Figure: PDF of $Exp(0.9)$, red dot is the Expected Value

Expectation: Example

Example 34.14: Assume $X \sim \mathcal{N}(\mu, \sigma^2)$. Show that $\mathbb{E}(X) = \mu$.

Interpretation: Well, we were calling μ the Mean of X. This is because of the above result.

Expectation: Normal Distribution

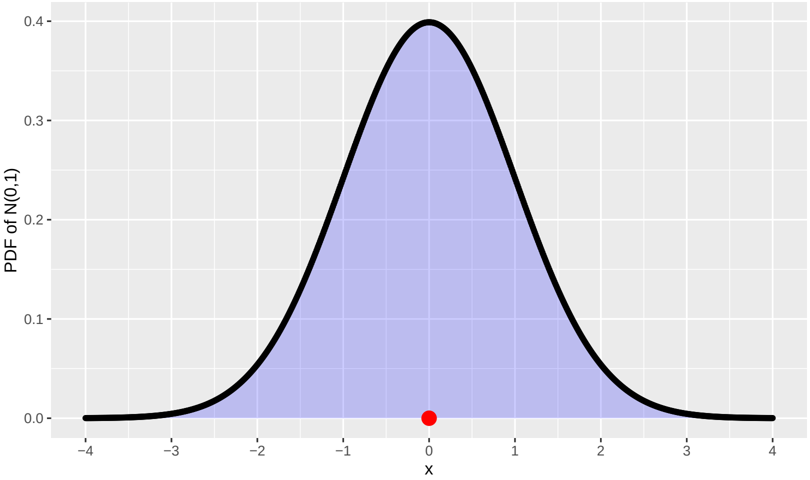

Figure: PDF of $\mathcal{N}(0,1)$, red dot is the Expected Value

Expectation: Example

Example 34.15: Assume X is a Continuous r.v. with the PDF $f(x)$, and assume f is symmetric around the point a, i.e., its graph is symmetric around the line $x = a$. Show that

$$\mathbb{E}(X) = a.$$

Expectation: Example

Example 34.16: Assume X is a Continuous r.v. with the PDF $f(x)$. Is it true that the point $\mu = \mathbb{E}(X)$ divides the area under the graph of f into two equal parts, i.e., is it true that

$$\mathbb{P}(X \leq \mu) = \mathbb{P}(X \geq \mu) = \frac{1}{2} \text{ ?}$$

Note: The point $m \in \mathbb{R}$ satisfying the above condition is called a Median of the Distribution of X, and is denoted by $Median(X)$. And, if f is symmetric around some point a, then

$$Median(X) = \mathbb{E}(X) = a.$$

Expectation: Note

Note: It is possible that the Expectation of the r.v. X is infinite, or is not defined. This is because we can deal with infinite sums or with improper integrals. The example in the St. Petersburg Paradox was giving infinite Expectation. And, for the continuous case, if we will consider, say, a r.v. X with the PDF

$$f(x) = \begin{cases} \dfrac{K}{x^2}, & x \geq 1 \\ 0, & \text{otherwise} \end{cases}$$

then $\mathbb{E}(X) = +\infty$. And, for the standard Cauchy r.v. X, with the PDF

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2},$$

the Expectation $\mathbb{E}(X)$ is not defined.

Properties of Expectation

Properties of the Expectation

Assume X, Y are r.v. defined on the same Experiment, with finite Expectations. Then

- $\mathbb{E}(X)$ is a (deterministic) number, is not random;
- $\mathbb{E}(C) = C$ for any constant C;
- $\mathbb{E}(\mathbb{E}(X)) = \mathbb{E}(X)$;
- If X and Y are ID, then $\mathbb{E}(X) = \mathbb{E}(Y)$;
- If $X \geq 0$, then $\mathbb{E}(X) \geq 0$;
- If $X \geq 0$ and $\mathbb{E}(X) = 0$, then $X = 0$ a.s., i.e., $\mathbb{P}(X = 0) = 1$;

Properties of the Expectation

- $\mathbb{E}(\alpha \cdot X + \beta \cdot Y) = \alpha \cdot \mathbb{E}(X) + \beta \cdot \mathbb{E}(Y)$;
- In general, for any r.v. X_k with finite Expectation and for any $\alpha_k \in \mathbb{R}$,

$$\mathbb{E}\left(\sum_{k=1}^{n} \alpha_k X_k\right) = \sum_{k=1}^{n} \alpha_k \mathbb{E}(X_k);$$

- If $X \leq Y$, then $\mathbb{E}(X) \leq \mathbb{E}(Y)$;
- $\big|\mathbb{E}(X)\big| \leq \mathbb{E}(|X|)$;

Properties of the Expectation

- If $\mathbb{1}_A$ is the **characteristic (indicator) function** of the Event A, i.e.,

$$\mathbb{1}_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A, \end{cases}$$

then $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$;

Expectation Properties: Example

Example 34.17: Calculate

$$\mathbb{E}(3X - 4Y + 7),$$

if $X \sim Bernoulli(0.2)$ and Y is continuous r.v. with the PDF

$$f_Y(x) = 2x, \ x \in [0,1] \qquad \text{and} \qquad f_Y(x) = 0, \ \text{otherwise.}$$

Expectation Properties: Example

Example 34.18: Calculate

$$\mathbb{E}\Big(X - \mathbb{E}(X)\Big)$$

and interpret the result.

Expectation Properties: Example

Example 34.19: Assume $X, Y, Z \sim Unif[0,1]$. What can be said about the Distribution of

$$S = X + Y + Z?$$

Solution: Well, in general, because we do not know the relationship between X, Y, Z, we cannot say too much. Say, if we will have that X, Y, Z are Independent, then, after some technical calculations, we can find the PDF of S.
But, one thing can be said about S without any assumption:

$$\mathbb{E}(S) = \mathbb{E}(X) + \mathbb{E}(Y) + \mathbb{E}(Z) = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} = \frac{3}{2},$$

so the values of S are around $\frac{3}{2}$, with the mean at that point!

Expectation Properties: Example

Example 34.20: Assume $X_1, X_2, ..., X_n$ are ID r.v.s with

$$\mathbb{E}(X_1) = \mu.$$

Find

a. $\mathbb{E}(X_1 + X_2 + ... + X_n)$;

b. $\mathbb{E}\left(\dfrac{X_1 + X_2 + ... + X_n}{n}\right)$.

Expectation: Example

Example 34.21: What is the expected number of fixed points in a random permutation of elements $\{1, 2, 3, ..., n\}$.

Expectation of the Transformed r.v.

Expectation of the transformed r.v.

Now, assume X is a r.v., $g : \mathbb{R} \to \mathbb{R}$ is a given function, and $Y = g(X)$. We want to calculate the Expected Value $\mathbb{E}(Y) = \mathbb{E}(g(X))$.

In general, we have 2 methods to calculate $\mathbb{E}(Y)$:

M1: First, we can find the Distribution of Y, say, its PDF/PMF, then calculate $\mathbb{E}(Y)$;

M2: We can calculate $\mathbb{E}(Y) = \mathbb{E}(g(X))$ straightforwardly from the Distribution of X, without doing the transform.

Let us know consider separately Discrete/Continuous cases.

Expectation of the transformed Discrete r.v.

Discrete Case: Assume X is Discrete, with

$$X \sim \begin{pmatrix} x_1 & x_2 & \dots \\ p_1 & p_2 & \dots \end{pmatrix}$$

and $Y = g(X)$.

M1: We first calculate the PMF of Y: if the possible values of Y are y_j, and the corresponding probabilities are $\mathbb{P}(Y = y_j)$, then

$$\mathbb{E}(Y) = \sum_j y_j \cdot \mathbb{P}(Y = y_j);$$

M2: We just calculate

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = \sum_i g(x_i) \cdot \mathbb{P}(X = x_i) = \sum_i g(x_i) \cdot p_i.$$

Expectation of the transformed Continuous r.v.

Continuous Case: Assume X is Continuous, with PDF $f_X(x)$, and $Y = g(X)$.

M1: If Y is Discrete, we can find its PMF, and calculate

$$\mathbb{E}(Y) = \sum_j y_j \cdot \mathbb{P}(Y = y_j);$$

If Y is Continuous, we can find its PDF, $f_Y(x)$, and calculate

$$\mathbb{E}(Y) = \int_{-\infty}^{\infty} x \cdot f_Y(x)\, dx.$$

M2: We can use the PDF of X to calculate $\mathbb{E}(Y)$, both if Y is Discrete or Continuous:

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) \cdot f_X(x)\, dx.$$

Expectation of the transformed r.v.

Note: This method is called the **law of the unconscious statistician (LOTUS)**[1].

Example: Assume X is continuous r.v with the PDF $f(x)$. Then

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx; \qquad \mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)\, dx;$$

$$\mathbb{E}(X^3 + 2X - 1) = \int_{-\infty}^{\infty} (x^3 + 2x - 1) \cdot f(x)\, dx;$$

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

[1] See, e.g. https://en.wikipedia.org/wiki/Law_of_the_unconscious_statistician

Example

Example 34.22: Assume $X \sim Binom(3, 0.2)$. Calculate

$$\mathbb{E}(\sin\left(\frac{\pi}{2} \cdot X\right)).$$

Expectation: Example

Example 34.23: The company wants to suggest a warranty service in the following form: if the product will go order within the first year, they will reimburse in the amount of $100\$$. If it fails within the second year, they will pay $50\$$. In all other cases they will not pay any money. For this service, how much they need to add to the price of the product, if the lifetime of the product is estimated to be an Exponential r.v., and the average lifetime for a product is 5 years?

Example

Example 34.24: Assume $X \sim Unif[0, 1]$. Calculate $\mathbb{E}(X^n)$, for any n.

Note: For any r.v. X, $\mathbb{E}(X^n)$ is called its n-th order Moment. The first order Moment is just $\mathbb{E}(X)$, the Expected Value of X. It is giving some information about X, but not too much. The second order Moment is $\mathbb{E}(X^2)$, and if we know also this, then we have more info about X. Knowledge of each additional Moment is giving more information about X.

Example

Example 34.25: Assume $X \sim \mathcal{N}(0, 1)$. Show that all odd order moments of X are equal to 0.

Expectation of the transformed r.v.

Important Note: Please note that, in general,

$$\mathbb{E}(X^2) \neq \left(\mathbb{E}(X)\right)^2,$$

and

$$\mathbb{E}(g(X)) \neq g\big(\mathbb{E}(X)\big).$$

Example: If $X \sim Unif[-1, 1]$, then $\mathbb{E}(X) = 0$, but $\mathbb{E}(X^2) > 0$.

Jensen's Inequality

Now, concerning the last question, I want to give the Jensen's Inequality for r.v.s.

Recall the Jensen's Inequality for convex functions: if $g : [a, b] \to \mathbb{R}$ is a convex function, $x_1, ..., x_n \in [a, b]$ and $\alpha_1, ..., \alpha_n \in [0, 1]$ with $\alpha_1 + \alpha_2 + ... + \alpha_n = 1$, then

$$g\left(\sum_{k=1}^{n} \alpha_k x_k\right) \leq \sum_{k=1}^{n} \alpha_k \cdot g(x_k).$$

Now, the generalization, and Probabilistic Interpretation is:

Jensen's Inequality

If X is a r.v. and $g : \mathbb{R} \to \mathbb{R}$ is a **convex function**, then

$$g(\mathbb{E}(X)) \leq \mathbb{E}(g(X)).$$

Expectation of a r.v. obtained from 2 Jointly Distributed r.v.s

Expectation for the r.v. obtained from 2 r.v.s

Assume X and Y are Jointly Distributed r.v.s, and $Z = g(X, Y)$ for some known function g. We want to calculate $\mathbb{E}(Z)$.

- **Discrete Case:** If the values of X are $x_1, x_2, \dots$ and the values of Y are $y_1, y_2, \dots$, then

$$\mathbb{E}(Z) = \mathbb{E}(g(X, Y)) = \sum_{i,j} g(x_i, y_j) \cdot \mathbb{P}(X = x_i, Y = y_j)$$

- **Continuous Case:** If $f(x, y)$ is the Joint PDF of X and Y, then

$$\mathbb{E}(Z) = \mathbb{E}(g(X, Y)) = \iint_{\mathbb{R}^2} g(x, y) \cdot f(x, y)\, dxdy.$$

Example

Example: Assume X and Y are Jointly Continuous r.v.s with the Joint PDF $f(x, y)$. Then

$$\mathbb{E}(X + Y) = \iint_{\mathbb{R}^2} (x + y) \cdot f(x, y)\, dxdy;$$

$$\mathbb{E}(X \cdot Y) = \iint_{\mathbb{R}^2} x \cdot y \cdot f(x, y)\, dxdy;$$

$$\mathbb{E}(\sqrt{X^2 + \cos(Y)}) = \iint_{\mathbb{R}^2} \sqrt{x^2 + \cos(y)} \cdot f(x, y)\, dxdy.$$

Note: Similar formulas, using sums, work also for the case when X and Y are discrete.

Example

Example: Assume X and Y are Jointly Continuous r.v.s with the Joint PDF $f(x, y)$. And assume we want to calculate the Expected Value of X. We can do that in 2 ways:

M1: First calculate the Marginal PDF of X, $f_X(x)$, and then use

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx.$$

M2: Just use Double Integration:

$$\mathbb{E}(X) = \iint_{\mathbb{R}^2} x \cdot f(x, y)\, dxdy.$$

Example

Example 34.26: Assume X and Y are Jointly Discrete r.v.s with the following Joint PMF:

$Y \backslash X$	0	1	2
-1	0.3	0.1	0.1
1	0.2	0.2	0.1

a. Calculate $\mathbb{E}(\sqrt{X^2 + Y^2})$;

b. Calculate $\mathbb{E}(X)$.

Example

Example 34.27: Assume X and Y are Jointly Continuous r.v.s with the following Joint PDF:

$$f(x, y) = \begin{cases} K \cdot xy, & x, y \in [0, 1], \ x + y \leq 1 \\ 0, & \text{otherwise.} \end{cases}$$

a. Find $\mathbb{E}\left(\frac{e^X}{1+X+Y}\right)$;
b. Find $\mathbb{E}(Y)$.

Example

Example 34.28: Assume we are picking 2 random points in $[0, 1]$, independently. What is the expected distance between that points?

Expectation and Independence

Assume X and Y are Independent r.v. Then

$$\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y).$$

Moreover, if g and h are any (nice!) functions, then

$$\mathbb{E}(g(X) \cdot h(Y)) = \mathbb{E}(g(X)) \cdot \mathbb{E}(h(Y)).$$

Note (18+): The inverse of the last statement is also true: if

$$\mathbb{E}(g(X) \cdot h(Y)) = \mathbb{E}(g(X)) \cdot \mathbb{E}(h(Y))$$

holds for any (nice) functions g and h, then X and Y are Independent.

Note: But please note that only $\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$, is not guaranteeing that X and Y are independent.

Example

Example 34.29: Assume $\mathbb{E}(X) = -2$, $\mathbb{E}(Y) = 4$.

a. Calculate $\mathbb{E}\Big(X(1-Y)\Big)$,

b. Calculate $\mathbb{E}\Big(X(1-Y)\Big)$, if $X \perp\!\!\!\perp Y$.

On the Definition of the Expectation

Note (18+): In advanced Probability textbooks the Expected value of the r.v. X is defined as the Lebesgue integral of X with respect to the Probability measure $\mathbb{P}$:

$$\mathbb{E}(X) = \int_\Omega X(\omega)\, d\,\mathbb{P}$$

Please note that the integration is over Ω here.
And, using the CDF $F(x)$, one is proving that

$$\mathbb{E}(X) = \int_{\mathbb{R}} x\, dF(x),$$

in terms of a Stieltjes integral.