Exercise Walkthrough: Transformation of a Uniform Random Variable

Justin Lanfermann

25. June 2025

Abstract

This document provides a detailed, step-by-step walkthrough for an exercise on the transformation of random variables. We will derive the cumulative distribution function (CDF) and probability density function (PDF) for a squared uniform random variable, and then use these to calculate a specific probability. The explanation is based on the definitions and theorems from the "Discrete Probability Theory" script by Niki Kilbertus (Summersemester 2025) and includes in-depth explanations of key concepts.

1 Problem Statement

We uniformly choose a real number from the interval $[0, 1]$. We then square this number. Let the result be represented by a real-valued random variable X.

(i) What is the cumulative distribution function of X?

(ii) What is the probability density function of X?

(iii) What is the probability that $X > \frac{1}{4}$? Try to find an answer in your head intuitively first.

2 Overview and Strategy

This exercise involves a **transformation of a random variable**. We start with a random variable we know well, in this case, a uniform distribution, and apply a function to it (squaring) to create a new random variable. Our goal is to find the distribution of this new variable.

The most robust way to solve this is the **CDF method**:

1. Define the initial random variable (let's call it U) and the new random variable $X = g(U)$.
2. Use the definition of the CDF for X: $F_X(x) = P(X \leq x)$.
3. Substitute X with its definition in terms of U: $F_X(x) = P(g(U) \leq x)$.
4. Solve the inequality for U and express the probability in terms of the CDF of U, which we know.
5. Once we have the CDF of X, we can find its PDF by differentiation.
6. Finally, we use the derived CDF or PDF to compute the required probability.

Let's apply this strategy.

3 Step-by-Step Solution

3.1 Part (i): Cumulative Distribution Function (CDF) of X

Step 1: Formalize the setup. Let U be the random variable representing the number chosen uniformly from $[0, 1]$. This means U follows a Uniform distribution, denoted $U \sim \text{Unif}(0, 1)$. The random variable X is the square of this number, so we have the transformation:

$$X = U^2$$

From *Example 1.56 (i)* of the script, we know the CDF of U is $F_U(u) = u$ for $u \in [0, 1]$, and its PDF is $f_U(u) = 1$ for $u \in [0, 1]$ (and 0 otherwise).

Step 2: Apply the definition of the CDF. We want to find the CDF of X, which we denote by $F_X(x)$. According to *Definition 1.21 (cdf)*, this is:

$$F_X(x) = P(X \leq x)$$

Substituting $X = U^2$, we get:

$$F_X(x) = P(U^2 \leq x)$$

Step 3: Solve the inequality for U. To evaluate this probability, we need to solve the inequality $U^2 \leq x$ for U. This gives us $-\sqrt{x} \leq U \leq \sqrt{x}$. So,

$$F_X(x) = P(-\sqrt{x} \leq U \leq \sqrt{x})$$

However, we must consider the domain (or support) of U. We know that U can only take values in $[0, 1]$. Therefore, the condition $-\sqrt{x} \leq U \leq \sqrt{x}$ combined with $0 \leq U \leq 1$ simplifies to:

$$F_X(x) = P(0 \leq U \leq \sqrt{x})$$

Step 4: Analyze the cases for x. The value of this probability depends on the value of x. We must consider all possible real values for x.

- **Case 1: $x < 0$.** Since U is a real number, $X = U^2$ cannot be negative. Therefore, the event $X \leq x$ is impossible.

$$F_X(x) = P(U^2 \leq x) = 0 \quad \text{for } x < 0.$$

- **Case 2: $0 \leq x \leq 1$.** In this range, $\sqrt{x}$ is between 0 and 1. The probability $P(0 \leq U \leq \sqrt{x})$ can be found using the CDF of U:

$$P(0 \leq U \leq \sqrt{x}) = F_U(\sqrt{x}) - F_U(0) = \sqrt{x} - 0 = \sqrt{x}.$$

  So, $F_X(x) = \sqrt{x}$ for $0 \leq x \leq 1$.

- **Case 3: $x > 1$.** Since the maximum value of U is 1, the maximum value of $X = U^2$ is also 1. Therefore, for any $x > 1$, the event $X \leq x$ is certain to happen.

$$F_X(x) = P(U^2 \leq x) = 1 \quad \text{for } x > 1.$$

Step 5: Combine the pieces. Putting all the cases together, we get the complete CDF of X:

$$F_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ \sqrt{x} & \text{if } 0 \leq x \leq 1 \\ 1 & \text{if } x > 1 \end{cases}$$

This is the final answer for part (i).

3.2 Part (ii): Probability Density Function (PDF) of X

Step 1: Differentiate the CDF. According to *Lemma 1.44 (ii)*, the PDF $f_X(x)$ is the derivative of the CDF $F_X(x)$ with respect to x.

$$f_X(x) = \frac{d}{dx} F_X(x)$$

Step 2: Differentiate the piecewise function. We differentiate each part of the CDF we found in part (i):

- For $x < 0$ and $x > 1$, $F_X(x)$ is constant, so its derivative is 0.
- For $0 < x < 1$, we differentiate $F_X(x) = \sqrt{x} = x^{1/2}$:

$$\frac{d}{dx}(\sqrt{x}) = \frac{1}{2}x^{-1/2} = \frac{1}{2\sqrt{x}}.$$

  Note that the derivative is not defined at $x = 0$, but this is not a problem for a PDF, as the probability of any single point for a continuous variable is zero.

Step 3: Combine into the PDF. The resulting PDF is:

$$f_X(x) = \begin{cases} \frac{1}{2\sqrt{x}} & \text{if } 0 < x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

This is our answer for part (ii). As a quick sanity check, we can verify that this integrates to 1:

$$\int_{-\infty}^{\infty} f_X(x)dx = \int_0^1 \frac{1}{2\sqrt{x}}dx = \left[\sqrt{x}\right]_0^1 = \sqrt{1} - \sqrt{0} = 1.$$

The PDF is valid.

3.3 Part (iii): Calculating the Probability

We need to find $P(X > \frac{1}{4})$.

Intuitive Approach: The transformation $X = U^2$ is not linear. It "squishes" the numbers in $[0, 1]$. Let's think about the condition $X > 1/4$. This is the same as $U^2 > 1/4$. Since U is always non-negative, this is equivalent to $U > \sqrt{1/4}$, which means $U > 1/2$. The question now becomes: "What is the probability that a uniformly chosen number from $[0, 1]$ is greater than $1/2$?" Since the distribution is uniform, the probability is simply the length of the favorable interval, which is $(1/2, 1]$. The length is $1 - 1/2 = 1/2$. So, intuitively, the answer should be $\frac{1}{2}$.

Formal Calculation: We can use the CDF we derived in part (i), which is the most direct method.

$$P\left(X > \frac{1}{4}\right) = 1 - P\left(X \leq \frac{1}{4}\right)$$

By definition, $P(X \leq 1/4)$ is just the CDF evaluated at $x = 1/4$:

$$P\left(X > \frac{1}{4}\right) = 1 - F_X\left(\frac{1}{4}\right)$$

Since $0 \leq 1/4 \leq 1$, we use the formula $F_X(x) = \sqrt{x}$:

$$P\left(X > \frac{1}{4}\right) = 1 - \sqrt{\frac{1}{4}} = 1 - \frac{1}{2} = \frac{1}{2}.$$

This confirms our intuition. Alternatively, we could have integrated the PDF from part (ii):

$$P\left(X > \frac{1}{4}\right) = \int_{1/4}^{\infty} f_X(x)dx = \int_{1/4}^{1} \frac{1}{2\sqrt{x}}dx = \left[\sqrt{x}\right]_{1/4}^{1} = \sqrt{1} - \sqrt{\frac{1}{4}} = 1 - \frac{1}{2} = \frac{1}{2}.$$

Both formal methods give the same correct result.

4 Summary and Takeaways

In this exercise, we analyzed the random variable $X = U^2$ where $U \sim \text{Unif}(0, 1)$.

- **Key Technique:** The CDF method is a powerful and reliable way to find the distribution of a transformed random variable. It involves expressing the new CDF in terms of the old one.
- **CDF of $X = U^2$:** We found $F_X(x) = \sqrt{x}$ for $x \in [0, 1]$.
- **PDF of $X = U^2$:** By differentiating the CDF, we found $f_X(x) = \frac{1}{2\sqrt{x}}$ for $x \in (0, 1]$.
- **Probabilities:** We calculated $P(X > 1/4) = 1/2$. The intuitive approach was very effective here because the transformation and the initial distribution were simple. For more complex problems, the formal CDF method is indispensable.

Follow-up question for you: What if the transformation was $Y = \sqrt{U}$ instead? Can you try to find the CDF and PDF of Y?

5 In-depth Explanations

Here are more detailed explanations of the core concepts used in this walkthrough.

In-depth Concepts

[1] Probability Space $(\Omega, \mathcal{A}, P)$

- A probability space is the mathematical foundation for any probability problem. It consists of three components as per *Definition 1.18 (probability space)*:
  1. **Sample Space** Ω: The set of all possible outcomes of an experiment. In our exercise, $\Omega = [0, 1]$ for the initial choice of U.
  2. σ**-algebra** $\mathcal{A}$: A collection of subsets of Ω that we call "events". We assign probabilities to these events. For continuous spaces like $[0, 1]$, this is typically the *Borel* σ*-algebra* $\mathcal{B}$, which contains all intervals and any sets you can form from them using countable unions, intersections, and complements.
  3. **Probability Measure** P: A function that assigns a probability (a number in $[0, 1]$) to each event in $\mathcal{A}$. It must satisfy certain axioms, like $P(\Omega) = 1$.

In-depth Concepts

[2] Random Variable (RV)

- As per *Definition 1.45 (random variable)*, a random variable is not a variable in the traditional sense, but a **function** that maps outcomes from the sample space Ω to a set of values (usually real numbers).

- **Analogy:** Think of rolling a die. The sample space is $\Omega = \{\text{face with 1 dot}, \ldots, \text{face with 6 dots}\}$. The random variable X maps these abstract outcomes to numbers: $X(\text{face with } k \text{ dots}) = k$.

- In our exercise, U is a random variable mapping an abstract notion of "a random choice" to a number in $[0, 1]$. X is another random variable that further processes this number.

In-depth Concepts

[3] Uniform Distribution Unif(a, b)

- This distribution models the idea of "choosing a point completely at random" from an interval $[a, b]$. Every point is "equally likely", which for a continuous space means every sub-interval of the same length has the same probability.

- Its **PDF** is constant over the interval, as described in *Example 1.56 (i)*: $f(x) = \frac{1}{b-a}$ for $x \in [a, b]$. For $\text{Unif}(0, 1)$, this is just $f(x) = 1$ for $x \in [0, 1]$.

- Its **CDF** is a straight line, representing the accumulation of this constant density: $F(x) = \frac{x-a}{b-a}$ for $x \in [a, b]$. For $\text{Unif}(0, 1)$, this is $F(x) = x$.

In-depth Concepts

[4] Cumulative Distribution Function (CDF)

- The CDF of a random variable X, denoted $F_X(x)$, gives the probability that X will take a value less than or equal to x (*Definition 1.21*).

$$F_X(x) = P(X \leq x)$$

- It's "cumulative" because it adds up all the probability from $-\infty$ up to the point x.
- It has key properties (*Lemma 1.22*): it's non-decreasing, starts at 0 ($F_X(-\infty) = 0$), and ends at 1 ($F_X(\infty) = 1$).

In-depth Concepts

[5] Probability Density Function (PDF)

- For a continuous random variable, the PDF $f_X(x)$ describes the *relative likelihood* of the variable taking on a value near x.
- **Important:** $f_X(x)$ is NOT a probability. $P(X = x) = 0$ for any continuous RV. The PDF represents a *density*. To get a probability, you must integrate the PDF over an interval:
$$P(a \leq X \leq b) = \int_a^b f_X(x)dx$$
- As stated in *Lemma 1.44 (ii)*, the PDF is the derivative of the CDF. This makes sense: the density at a point is the rate at which the cumulative probability is changing at that point.