The Distribution of a Function of a RV

Gregory M. Shinault

The Problem

X is a RV, with known PDF $f_X(x)$ or PMF $p_X(k)$. The function $g : \mathbb{R} \to \mathbb{R}$ is given. Find the PDF/PMF of $Y = g(X)$.

This is the only type of problem we solve in this section. The case where X is a discrete RV is usually easy, so we will primarily focus on the case where X is a continuous RV.

Goals for this Lecture

1. Learn how to find the PMF of $g(X)$ if X is discrete.
2. Learn how to find the PDF of $g(X)$ if X is continuous.

This material corresponds to section 5.2 of the textbook.

Discrete Case

Example

Example 5.14. Suppose the range of X is $\{-1, 0, 1, 2\}$ with $P\{X = k\} = \frac{1}{4}$ for each $k \in \{-1, 0, 1, 2\}$. Let $Y = X^2$. Find the PMF of Y.

Example

Example 5.15. Suppose that X takes values in the set $\{-1, 0, 1, 2\}$ with

$$P\{X = -1\} = \frac{1}{10} \quad P\{X = 0\} = \frac{2}{10}$$
$$P\{X = 1\} = \frac{3}{10} \quad P\{X = 2\} = \frac{4}{10}.$$

Let $Y = 2X^3$. Find the PMF of Y.

Examples for Continuous RVs

General Strategy

1. Find the CDF: $F_Y(y) = \mathbb{P}(g(X) \leq y)$.
2. Differentiate the CDF: $f_Y(y) = F_Y'(y)$.

Easiest Example

Suppose $X \sim U([0,4])$. Set $Y = X^2$. Find $f_Y(y)$.

Easy Example

Suppose $X \sim U([-4,4])$. Set $Y = X^2$. Find $f_Y(y)$.

Moderate Example

Suppose $X \sim U([-3,4])$. Set $Y = X^2$. Find $f_Y(y)$.

Theory and Special Examples for Continuous RVs

Simplest General Case

Fact: If $g(u)$ has a differentiable inverse then

$$f_{g(X)}(y) = \frac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|}.$$

Most General Case

Fact: If $g(u)$ is differentiable and $g'(u) = 0$ at only finitely many points then

$$f_{g(X)}(y) = \sum_{x:g(x)=y} \frac{f_X(x)}{|g'(x)|}.$$

Special Example

Suppose X is $N(\mu, \sigma^2)$. Prove that the distribution of $Y = aX + b$ is $N(a\mu + b, a^2\sigma^2)$.

This is a general fact worth remembering, especially for the duration of this course.

Log-Normal Distribution

Suppose X is $N(\mu, \sigma^2)$. Find the PDF of $Y = e^X$.

Definition: The distribution of Y is called the *log-normal* distribution.

Application of Log-Normal

1. Modeling in mathematical finance
2. Distribution of income in the USA
3. Number of alcoholic drinks consumed by an individual per week in every culture

The Wrap Up

Summary

1. Use the CDF method or the formula to find the PDF of $g(X)$.
2. The CDF method is probably better practice because it requires more thought about probabilities, rather than analysis. (Personal preference)