The Normal Distribution

Gregory M. Shinault

Goals for this Lecture

1. Define the standard normal/Gaussian distribution, and the general normal/Gaussian distribution. This is the most important distribution in classical probability.
2. Learn the nice properties that the standard normal distribution possesses.
3. Learn how to compute probabilities for a general normal distribution from a table of standard normal values.

This material corresponds to section 3.5 of the textbook.

Introduction

The bell curve is known more formerly as the normal distribution or the Gaussian distribution. Today we present some facts about it, with no information about where it comes from.

Standard Normal Distribution

Definition

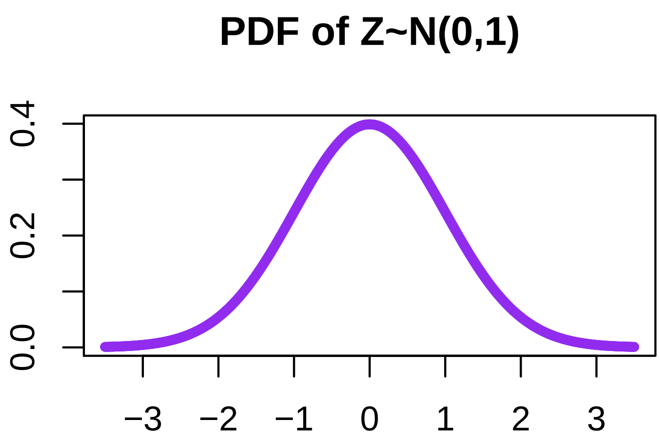
A continuous RV Z is called *standard normal* if it has the PDF

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

for all real x. We denote this by $Z \sim N(0,1)$.

The CDF a standard normal RV is denoted by $\Phi(x) = \mathbb{P}(Z \leq x)$.

PDF Plot



Example

Verify the the standard normal density is a PDF.

Example

Find the mean and variance of a standard normal RV.

Algebra of Standard Normal CDF

Fact:

1. $\Phi(-x) = 1 - \Phi(x)$.
2. $\Phi(x) - \Phi(-x) = 2\Phi(x) - 1$.

Exercise: Prove these facts.

General Normal Distribution

Definition

The standard normal distribution has very nice properties and its shape is useful, but most observables in application will not have mean 0 and variance 1.

A random variable X is Normal(μ, σ^2) if it is defined by

$$X = \mu + \sigma Z.$$

Alternate Definition: A random variable X is Normal(μ, σ^2) if it has the PDF
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$$
for all x.

Example

Prove that $X \sim$ Normal(μ, σ^2) has the PDF given in the alternate definition.

Example

The height of *the average american male* is 70 inches.
The standard deviation of American males' height is 3 inches.
We choose an adult American male at random.
Assuming heights are normally distributed (this is approximately true), what is the probability a randomly chosen man is between 68 and 73 inches tall?

Solution using R

```r
pnorm((73-70)/3) - pnorm((68-70)/3)
```

```
## [1] 0.5888522
```

```r
pnorm(73, mean=70, sd=3) - pnorm(68, mean=70, sd=3)
```

```
## [1] 0.5888522
```

Normally Distributed Data

The normal distribution is important because of there are countless types of data that are normally distributed, or can at least be accurately modelled by the normal distribution.

Student exam grades, heights, lumosity of stars, the modulus squared of the wave function for a quantum harmonic oscillator at its lowest energy level, some growth rates in finance, error analysis, etc.

The Wrap Up

Summary

1. The standard normal distribution is denoted by $N(0,1)$, usually as a RV Z. Its PDF is $f_Z(z) = e^{-z^2/2}/\sqrt{2\pi}$,

$$\mathbb{E}Z = 0, \quad \mathrm{Var}(Z) = 1.$$

2. The more general normal distribution with mean μ and variance σ^2 can be obtained by $X = \mu + \sigma Z$.
3. This is the most important continuous distribution in classical statistics.

Next Step

Now that we have introduced the Gaussian distribution, we are going to see where it actually comes from.