# Lecture 13: Gaussian distribution

Statistics 251

Yi Sun and Zhongjian Wang

Department of Statistics
The University of Chicago

# Where are we?

Tossing coins

Gaussian random variables

Special case of central limit theorem

# Tossing coins

How many heads will we get from a million fair coin tosses?

- To first order, the expectation is $500,000$.
- How close will the answer be to the expectation?

# Tossing coins

What is the probability of getting $k$ heads from $n$ coin tosses?

$$\mathbb{P}(X = k) = \frac{\binom{n}{k}}{2^n}.$$

Let's see a rough plot...

# Where are we?

Tossing coins

Gaussian random variables

Special case of central limit theorem

# Standard Gaussian random variables

We define a **standard Gaussian** random variable to have density

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

This is also called a **normal** random variable.

Notice that

$$\begin{aligned}\left[\int_{-\infty}^{\infty} f_x(x)dx\right]^2 &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2+y^2)}dxdy \\ &= \frac{1}{2\pi}\int_{0}^{\infty}\int_{0}^{2\pi} e^{-r^2/2} r\, d\theta dr \\ &= [-e^{-r^2/2}]_0^{\infty} = 1.\end{aligned}$$

# Standard Gaussian random variables

We define a **standard Gaussian** random variable to have density

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

By symmetry we compute

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 0.$$

We may also compute

$$\mathrm{Var}(X) = \mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 1.$$

# Gaussian random variables

Let $X$ be standard Gaussian. Consider $Y = \sigma X + \mu$. It has density

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

and mean and variance

$$\mathbb{E}[Y] = \mu \qquad \text{and} \qquad \mathrm{Var}(Y) = \sigma^2.$$

# Cumulative distribution function

Let $X$ be standard Gaussian. The CDF is

$$F_X(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}x^2} dx.$$

It **does not have** explicit representation. Instead

$$\Phi(x) = F_X(x)$$

is called the **error function**. Some values are:

$$\Phi(-1) \approx 0.159 \qquad \Phi(-2) \approx 0.023 \qquad \Phi(-3) \approx 0.0013.$$

This means $\approx 95$ percent of the time a Gaussian is within 1 standard deviation of the mean.

# Where are we?

Tossing coins

Gaussian random variables

Special case of central limit theorem

# DeMoivre-Laplace Limit Theorem

Let $S_n$ be the number of heads in $n$ tosses of a coin which is heads with probability $p$. We have

$$\mathrm{Var}(S_n) = np(1-p) \implies \text{std. dev.} = \sqrt{np(1-p)}.$$

The number of standard deviations away from the mean is

$$\frac{S_n - np}{\sqrt{np(1-p)}}.$$

## Theorem (DeMoivre-Laplace)

*We have that*

$$\lim_{n\to\infty} \mathbb{P}\left(a \le \frac{S_n - np}{\sqrt{np(1-p)}} \le b\right) = \Phi(b) - \Phi(a),$$

*where* $\Phi(b) - \Phi(a) = \mathbb{P}(a \le X \le b)$ *for a standard Gaussian* $X$.

This is why we used normalizing factor in the movie.

# Examples

Approximate the probability that we get more than $501,000$ heads in a million fair coin tosses.

- The standard deviation is $\sqrt{np(1-p)} = 500$ and $1000 = 500 \cdot 2$.
- The answer is approximately

$$1 - \Phi(2) = \Phi(-2) \approx 0.023.$$