# Stat 230: Probability

## Lecture 30

Jeremy VanderDoes

University of Waterloo

Monday, July 18th

# Review

**Example**

We have $N$ letters to $N$ different people, and $N$ envelopes addressed to those $N$ people. One letter is put in each envelope at random. Find the mean and variance of the number of letters placed in the right envelope.

# Review

Last time we talked about:

(1) Indicator random variables

(2) Sums of normal random variables

For today:

(1) Sums of normal random variables

# Review

- Today: Quiz 6

# Sum of Random Variables

**Theorem**

*Suppose that $X_1, ..., X_n$ are independent random variables, and that $X_i \sim N(\mu_i, \sigma_i^2)$. Then*

$$\sum_{i=1}^{n} a_i X_i \sim N\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

We don't quite have enough probability machinery to prove this (Moment generating functions).

# Sum of Random Variables

## Remark

*Suppose $X_1, ..., X_n$ are independent, and $X_i \sim N(\mu, \sigma^2)$, then*

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu, \frac{\sigma^2}{n})$$

# Sum of Random Variables

## Example

The length of springs produced by the ACME company is assumed to be normally distributed with mean $10cm$ and standard deviation $0.5cm$. Since the ends of the spring do not have the same properties as the middle, a machine cuts off a section at each end of the spring. Suppose the length of the cut-off section is normally distributed with mean $1cm$ and standard deviation $0.2cm$, and is independent of the original length of the spring and the amount cut from the opposite end.

(1) Compute the distribution of the length of the remaining middle section of the spring.

(2) Compute the probability that the remaining middle section of the spring is longer than $7.5cm$.

# Sum of Random Variables

## Example

Suppose that the height of adult males in Canada is normally distributed with a mean of 70 inches and variance of $4^2$ inches, and let $X_1, ..., X_{10}$ denote the heights of a random sample of adult males. Suppose $\bar{X}_{10}$ denotes the sample mean of these heights.

(1) Compute the probability that $X_3$ exceeds 75.

(2) Compute the probability that $\bar{X}_{10}$ exceeds 75

# Sum of Random Variables

**Reminder:** If $X_1, ..., X_n$ are independent and $X_i \sim N(\mu, \sigma^2)$, then

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu, \frac{\sigma^2}{n})$$

What if $X_1, ..., X_n$ are not normally distributed? What is the distribution of the mean?

# Central Limit Theorem

Consider the following simulation experiment:

(1) Think of a number $n \geq 1$, and then generate random variables $X_1, ..., X_n$ with a common distribution (e.g. normal, exponential, etc.)

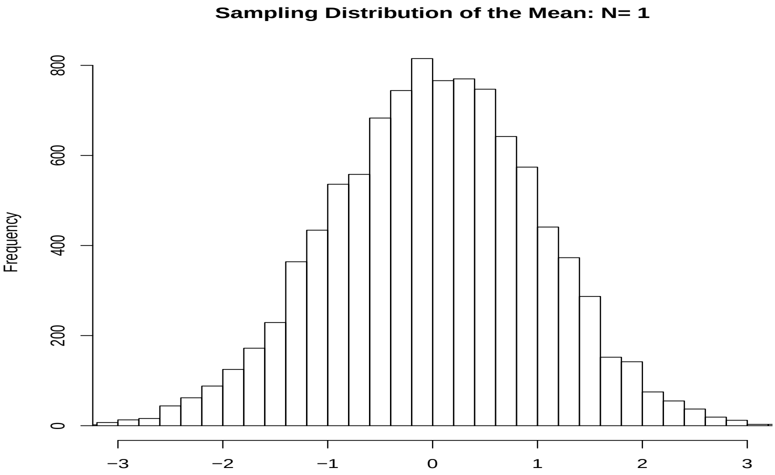
(2) Produce the sample mean $\bar{X}$ from the sample.

(3) Repeat this process 10,000 times to produce $\bar{X}_1, ..., \bar{X}_{10,000}$

(4) Plot a histogram of the $\bar{X}'s$.

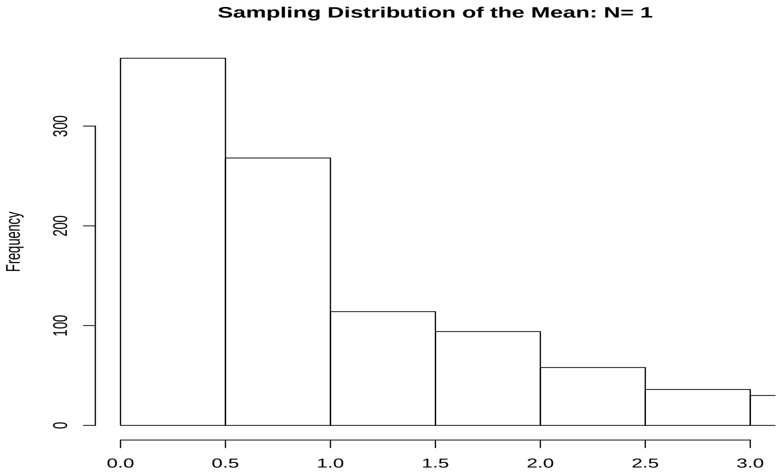
When the $X_i \sim N(\mu, \sigma^2)$ are normal, for example, then $\bar{X} \sim N(\mu, \sigma^2/n)$, and so the histograms from the $\bar{X}'s$ should look normal in that case.

# Central Limit Theorem



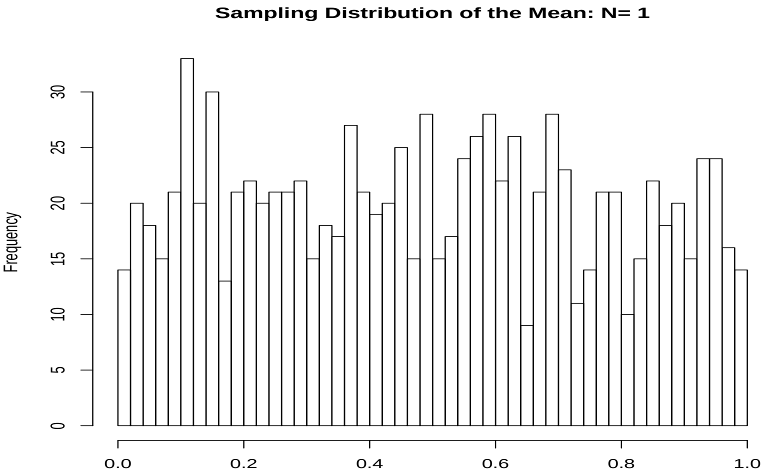

Normal Distribution

# Central Limit Theorem



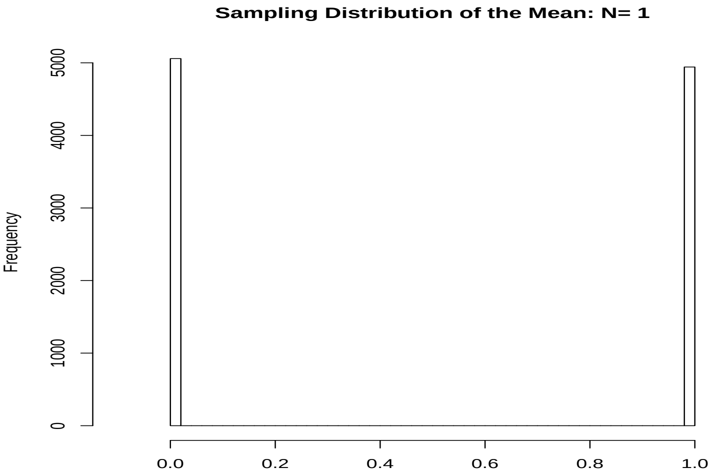

Exponential Distribution

# Central Limit Theorem




Uniform Distribution

# Central Limit Theorem




Bernoulli Distribution

# Central Limit Theorem

**Theorem (The Central Limit Theorem)**

*Suppose that $X_1, ..., X_n$ are independent random variables, each with a common cumulative distribution function $F$ with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2 < \infty$.*

*Then for all $x \in \mathbb{R}$*

$$P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq x\right) \to \Phi(x)$$

*as $n \to \infty$.*

In other words, if $n$ is large

$$\bar{X} \overset{approx}{\sim} N(\mu, \frac{\sigma^2}{n}) \text{ and } \sum_{i=1}^{n} X_i \overset{approx}{\sim} N(n\mu, n\sigma^2)$$

# Central Limit Theorem

## Example

Billy rolls a six sided die 1000 times, and records the results. If the die is a fair die, estimate the probability that the sum of the die rolls is less than 3400.