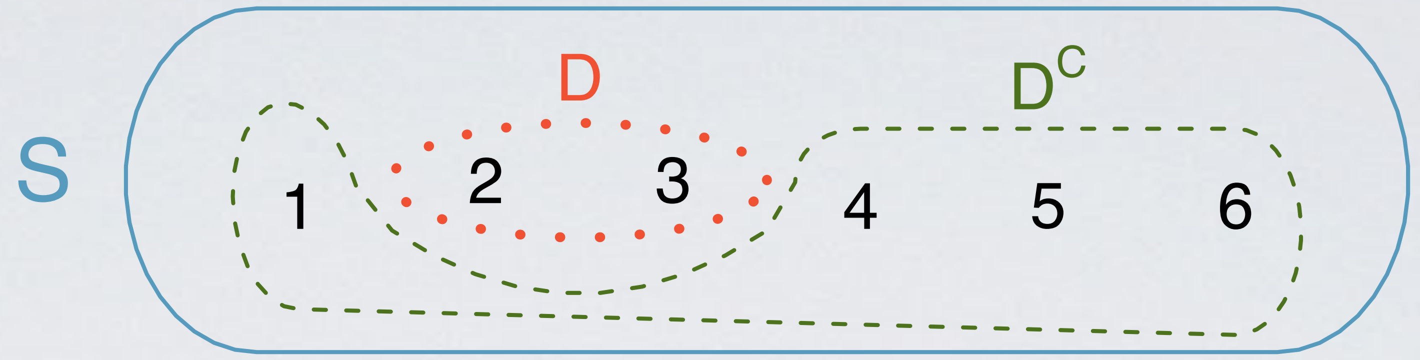
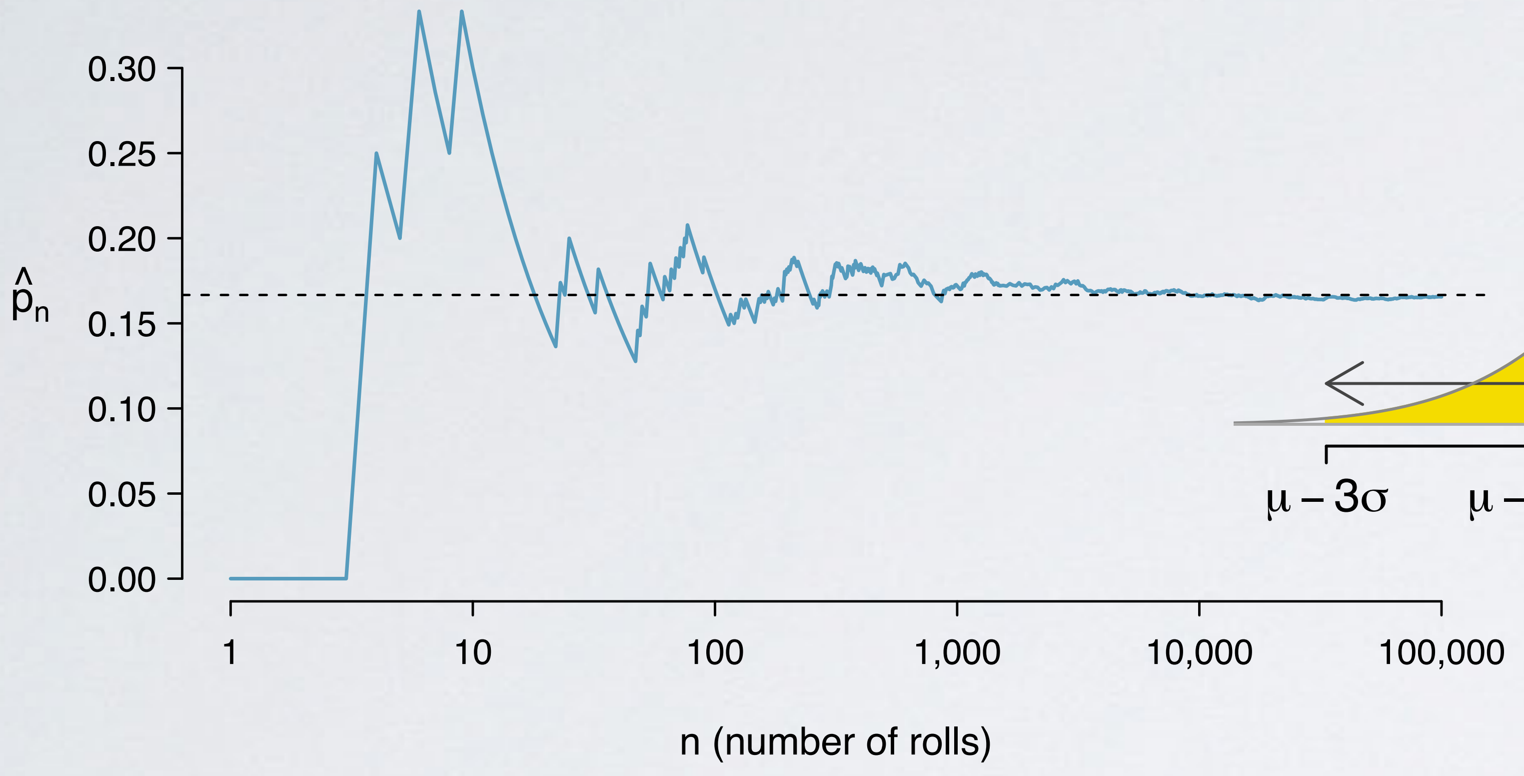
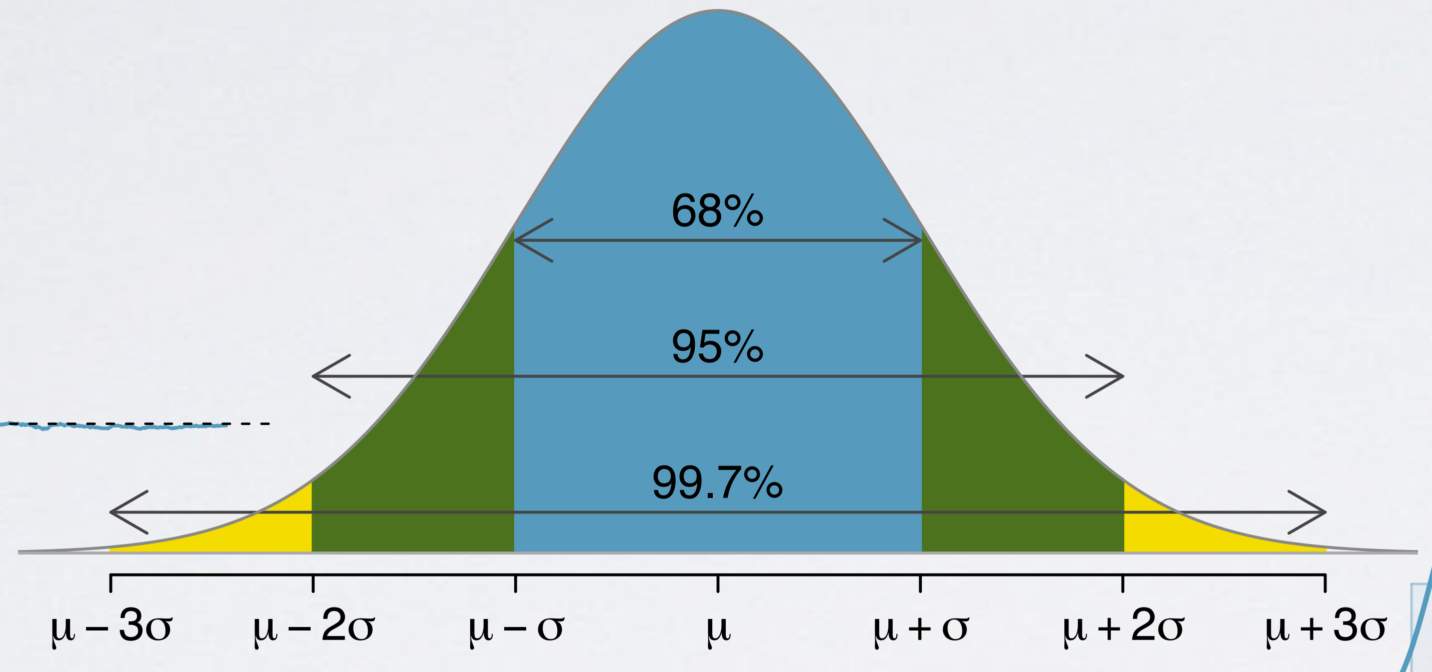
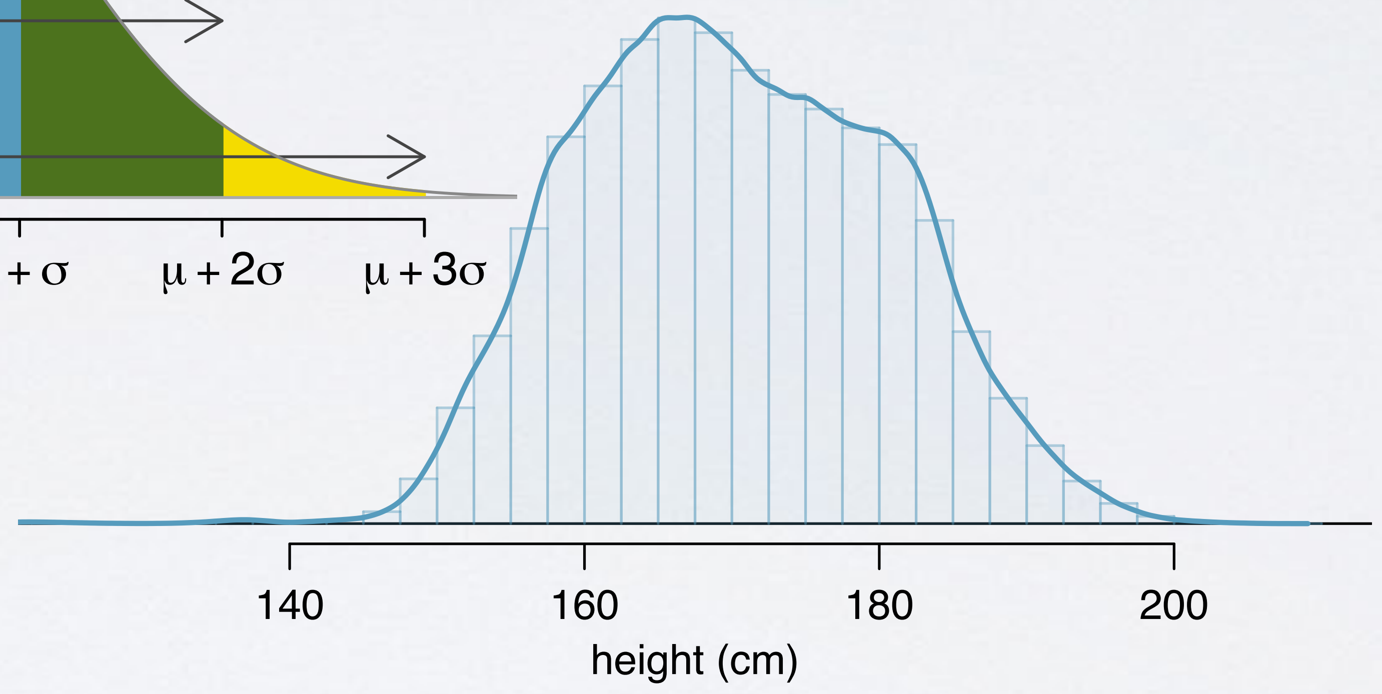


unit 2
probability and distributions



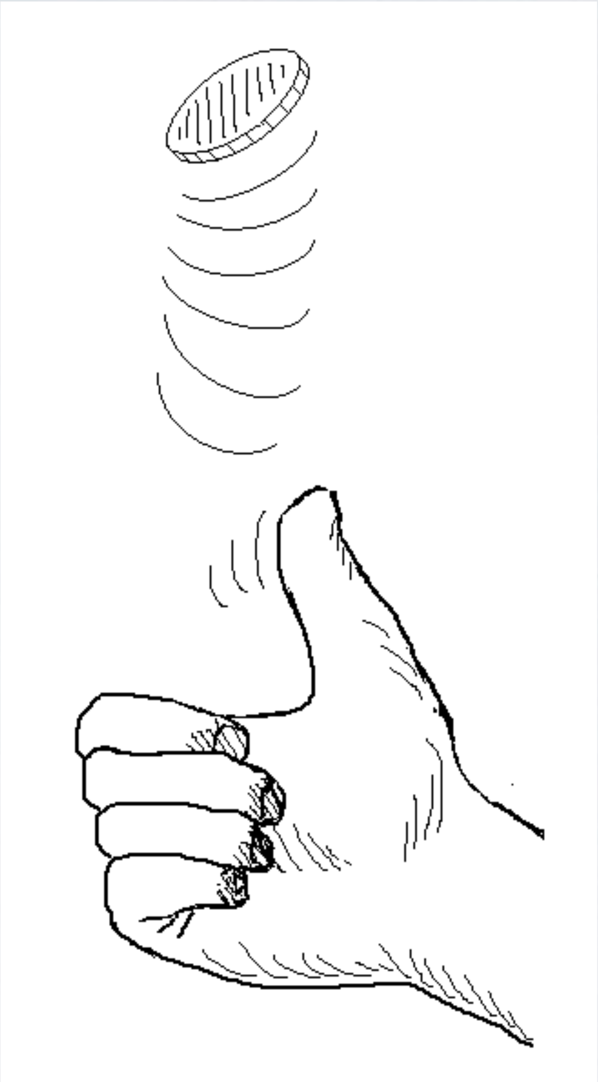




Dr. Mine Çetinkaya-Rundel
Duke University

random process

In a random process we know what outcomes could happen, but we don't know which particular outcome will happen.



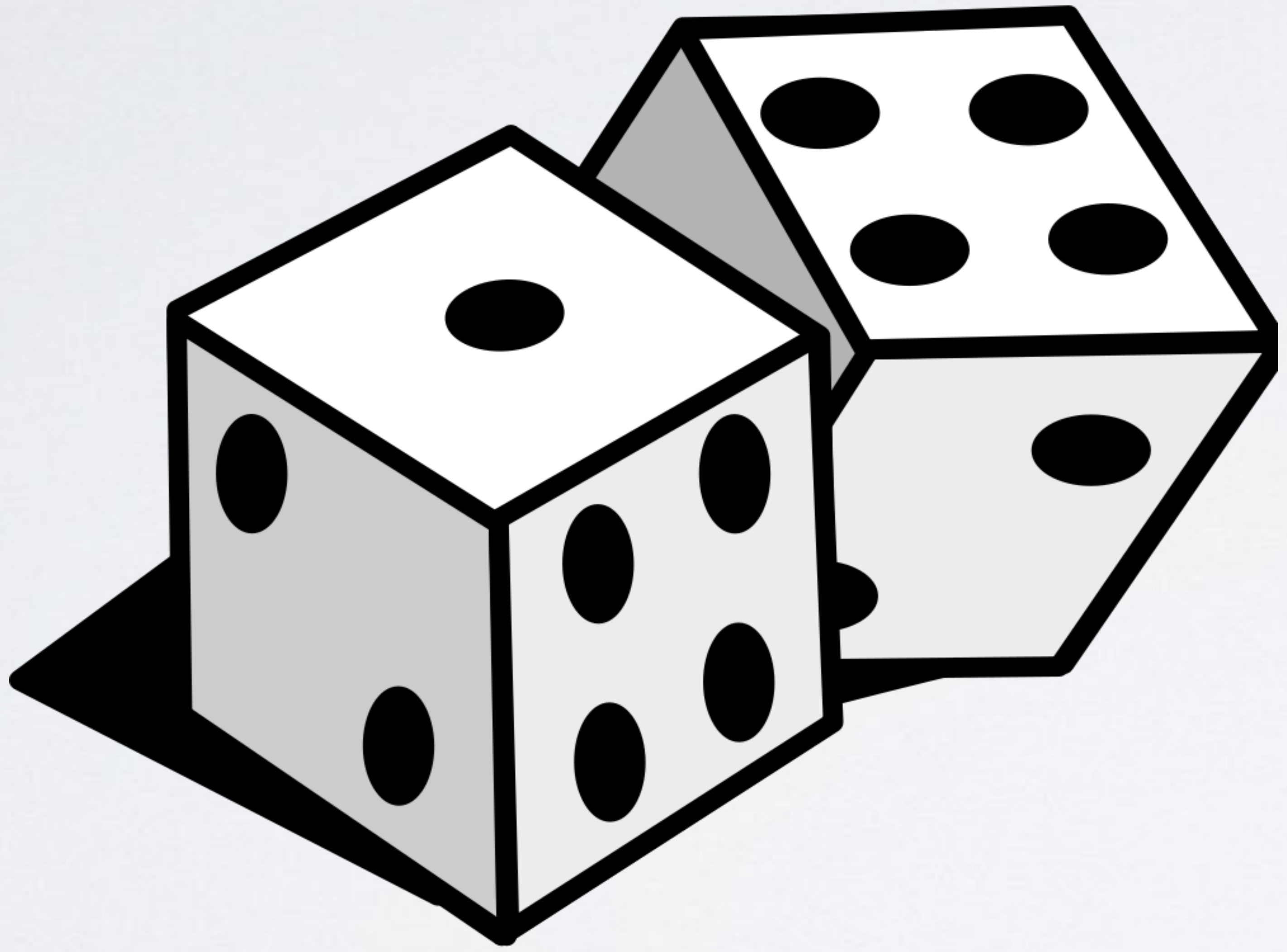


Image sources:

NASDAQ: By bfishadow on Flickr (http://flickr.com/photos/61368956@N00/3100369536)

iPod: http://tango-project.org/releases/tango-icon-theme-0.7.1.tar.gz

probability

P(A) = Probability of event A

There are several possible interpretations of probability but they (almost) completely agree on the mathematical rules probability must follow:

$$0 \leq P(A) \leq 1$$

frequentist interpretation

The probability of an outcome is the proportion of times the outcome would occur if we observed the random process an infinite number of times.

bayesian interpretation

A Bayesian interprets probability as a subjective degree of belief.

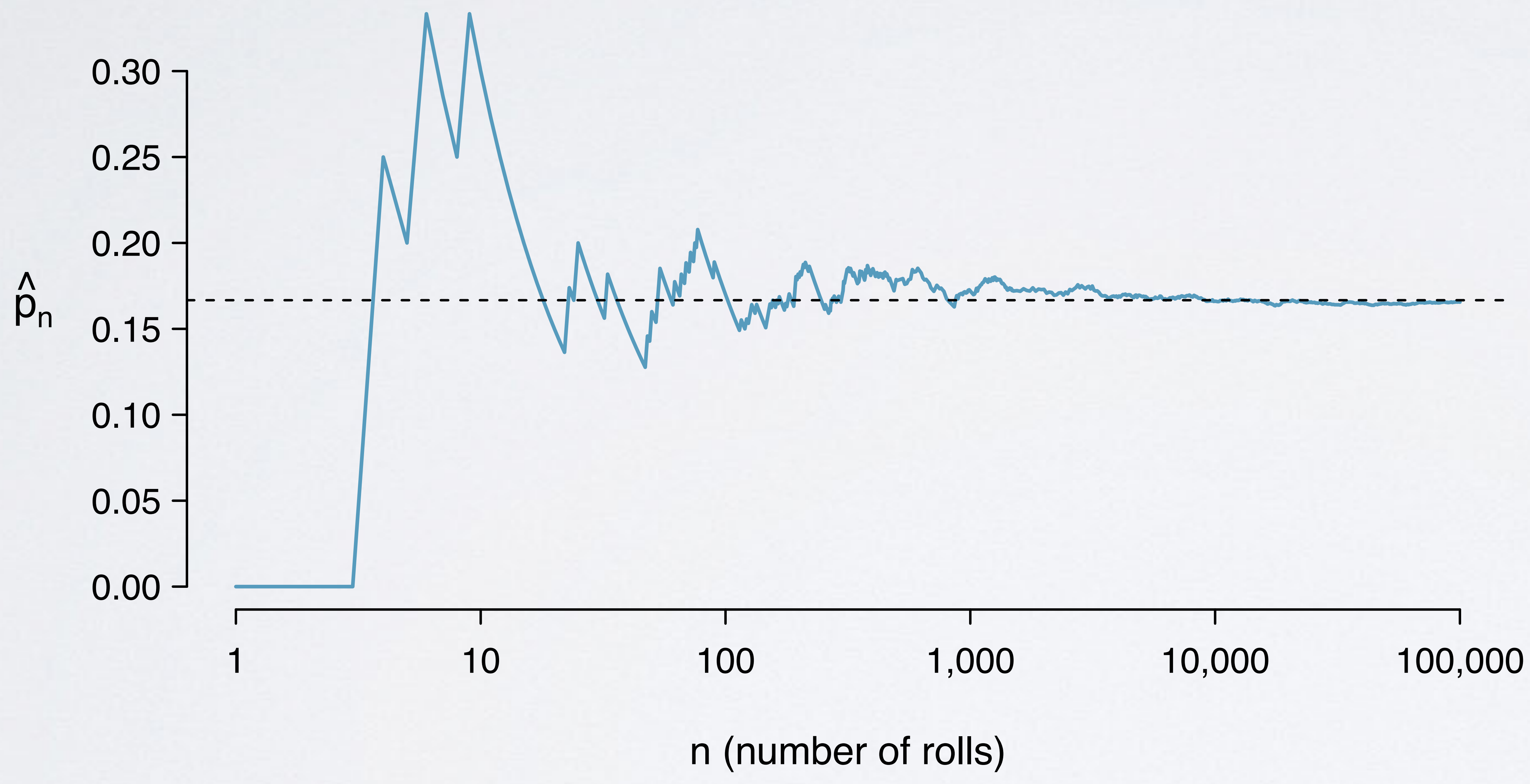
Largely popularized by revolutionary advance in computational technology and methods during the last twenty years.

Photo by dahlstroms on Flickr (http://www.flickr.com/photos/dahlstroms/527634847

law of large numbers

law of large numbers states that as more observations are collected, the proportion of occurrences with a particular outcome converges to the probability of that outcome.

examples



- exactly 3 heads in 10 coin flips
- exactly 3 heads in 100 coin flips
- exactly 3 heads in 1000 coin flips

Say you toss a coin 10 times, and it lands on Heads each time. What do you think the chance is that another head will come up on the next toss? 0.5, less than 0.5, or more than 0.5?

H H H H H H H H H H ?

The probability is still 50%:
P(H on the 11th toss)
= P(H on the 10th toss)
= 0.50

The coin is **not** due for a tail.

Common misunderstanding of law of large numbers: gambler's fallacy (law of averages)

Photo by thegrid-ch on Flickr (http://www.flickr.com/photos/thegrid-ch/5087085391) probability rules

conditional probability

probability distributions

binomial

normal