Programming in Python

Computational models and simulations
lecture 9

Department of Cybernetics and Artificial Intelligence
Technical University of Košice
Ing. Ján Magyar, PhD.

Random walk

It’s 3 a.m., party’s over, and students are slowly returning to their rooms. But since the party was a huge success, each of them has greatly limited cognitive and motoric skills. One of your friends had such a great time that his movement cannot be described other than random. Since you care about them, you want to find them before leaving. The last time you saw them five minutes ago.

- Where should you start looking for your friend?
- Will they be further from the starting point as time passes?

Analytical solution

- How far will they be after one step?
- How far will they be after two steps?
- How far will they be after three steps?
- …
- How far will they be after *n* steps?

Solving with a computational model

- we represent the map as a 2D space
- we don't expect any obstacles
- each step has an equal length
- your friend takes one step per second

Simulation structure

- inner loop executing one test run
- outer loop running multiple experiments (what is a suitable number?)
- we calculate statistical characteristics – we want to get as many views as possible, to get the global picture
- we visualize results in a suitable way

Random walk

- often-used model for different phenomena
- Brownian motion – motion of particles, molecules, etc.
- modelling biological processes, evolution, animal motion
- modelling brain signals
- betting
- approximating the size of the Internet
- modelling the stock exchange (nobody knows if it is suitable)

Biased random walk

- with random walk, each element is selected with the same probability
- biased random walk is a stochastic model where selection is not uniform, we define a different distribution
- sometimes selection is time-dependent or dependent on the previous selection

Simulation types

- stochastic vs. deterministic
- static vs. dynamic
- discrete vs. continuous

Monte Carlo simulation

- statistical inference method
- a random sample will usually have the same characteristics as the entire population
- the sample must be big enough
- not each sample of a certain size corresponds to the population

Sample size

- with simulations, the key problem is determining the necessary number of iterations and sample size
- with representative samples we never reach a 100% accuracy
- we can eliminate the chance of a non-representative sample only if we consider the entire population
- distribution within the sample should be approximately the same as in the population

Value distribution – rolling a die

What will be the distribution after rolling a die 1000 times?

What will be the probability of rolling 1?
What will be the probability of rolling 2?

Uniform distribution

- each value (or group of values) is represented in the same ratio in the population
- each value will be the result of a test run with the same probability
- we can define the distribution using an interval (smallest and largest value)
- not typical in reality, only in models

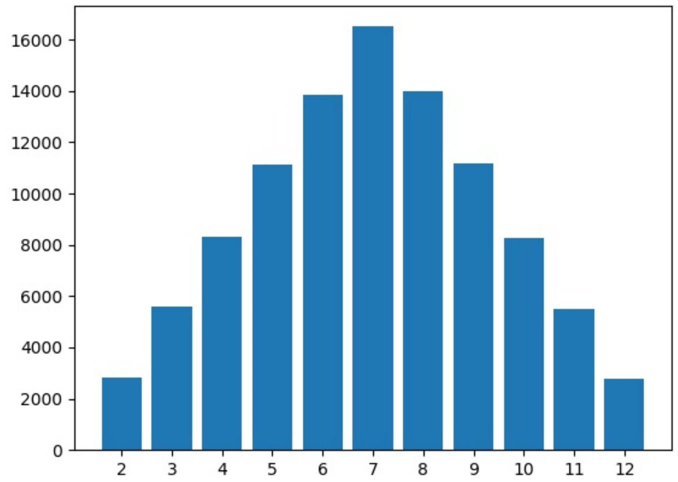
Value distribution – rolling two dice

What will be the distribution of the sum after rolling two dice 1000 times?

What will be the probability of the sum being 2?
What will be the probability of the sum being 7?

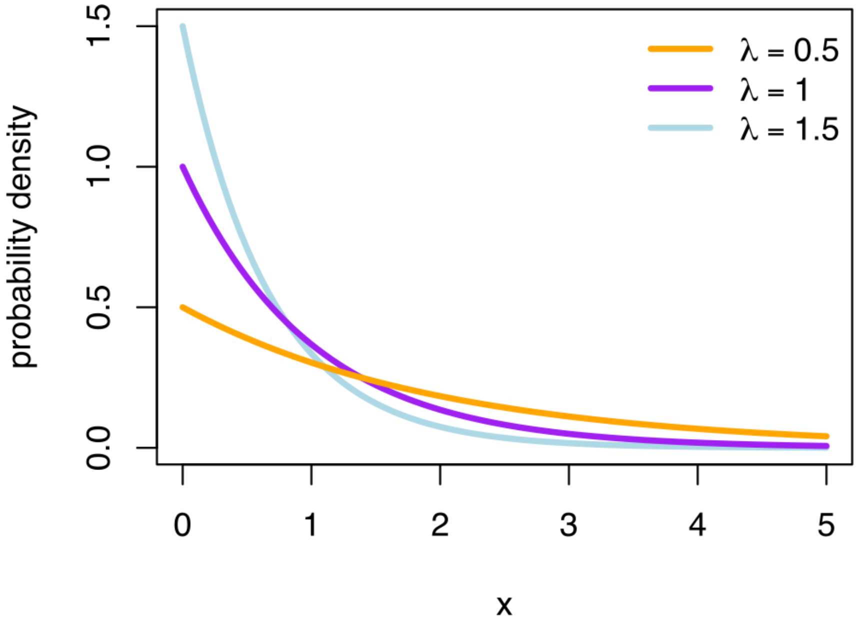
Normal distribution

- Gaussian distribution
- stable distribution – we get approximately the same result with multiple iterations
- the average value is the most likely (median at the same time)
- we can describe the distribution using the average and standard deviation
- most frequent distribution in nature



Exponential distribution

- asymmetric distribution
- the smallest value is the most likely
- described using a constant **λ**
- describes the time that passes between two instances of the same phenomenon, which happens with a constant average frequency (traffic, web page clicks)



Generating using distributions

- `random.uniform(a, b)`
  - random float from the interval `[a, b]` or `[b, a]`
- `random.gauss(mu, sigma)`
  - random number from a normal distribution
  - `mu` – most likely value
  - `sigma` – standard deviation
- `random.expovariate(lambd)`
  - random number from an exponential distribution
  - numbers from the interval `[0, ∞)` if `lambd > 0`, `(-∞, 0]` if `lambd < 0`

Craps

Craps is a popular dice game in casinos, where the player rolls two dice at the same time and wins according to the following rules:

- if the sum is 7 or 11, the player wins
- if the sum is 2, 3 or 12, the player loses
- in other cases the sum becomes the target number, and the player keeps rolling until he or she rolls
  - 7, in which case he or she loses
  - the target number, her or she wins

What are the player’s chances of winning? Does the casino have a higher chance?

Craps

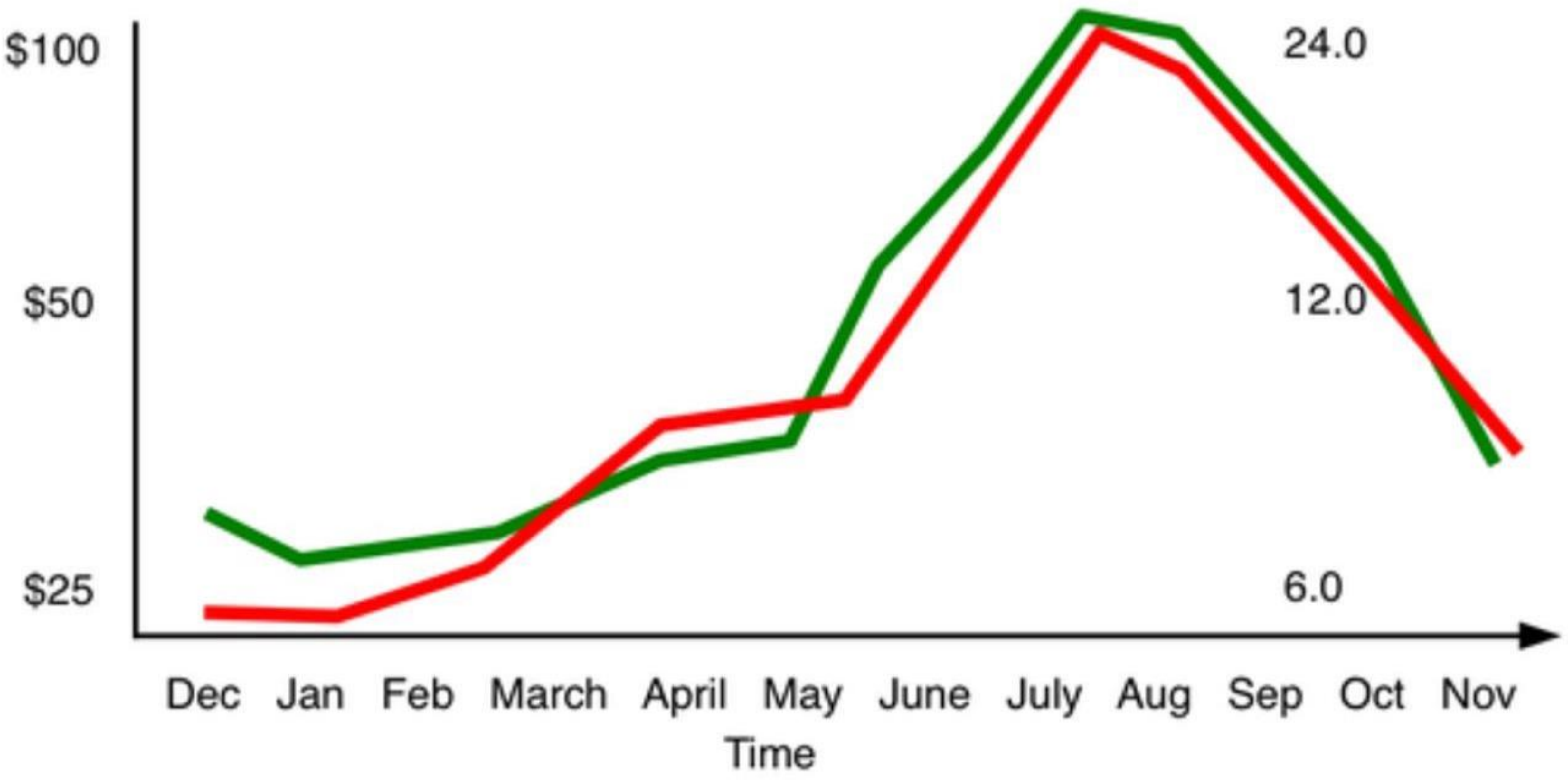
- How do we increase the chances of the player winning?
- How do we increase the chances of the casino winning?

Evaluating results

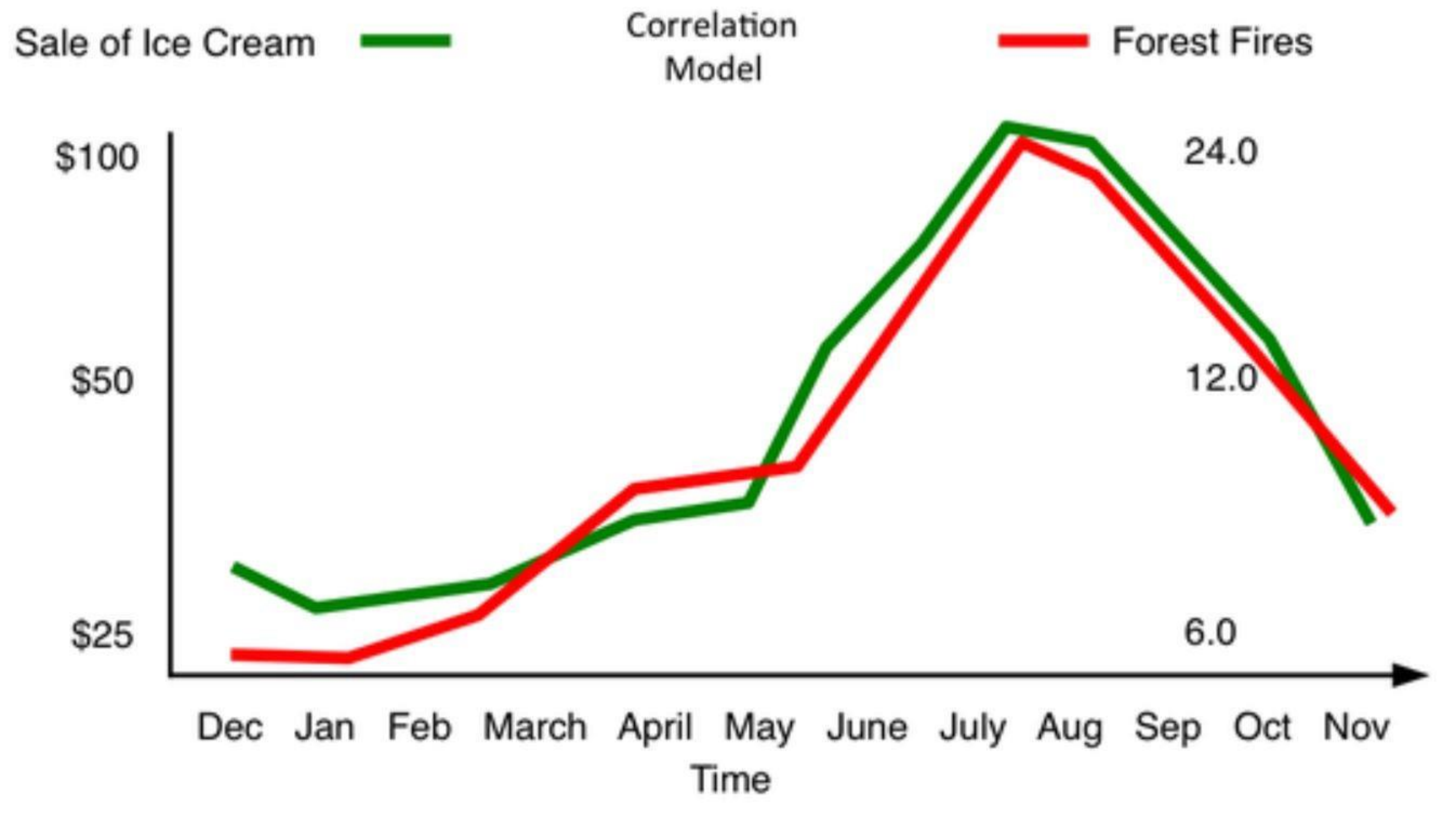
- How do we know that simulation results correspond to reality?
- we must test the implications on real examples (e.g. approximating π)
- the output of simulations are
  - data
  - models
  - implications or effects

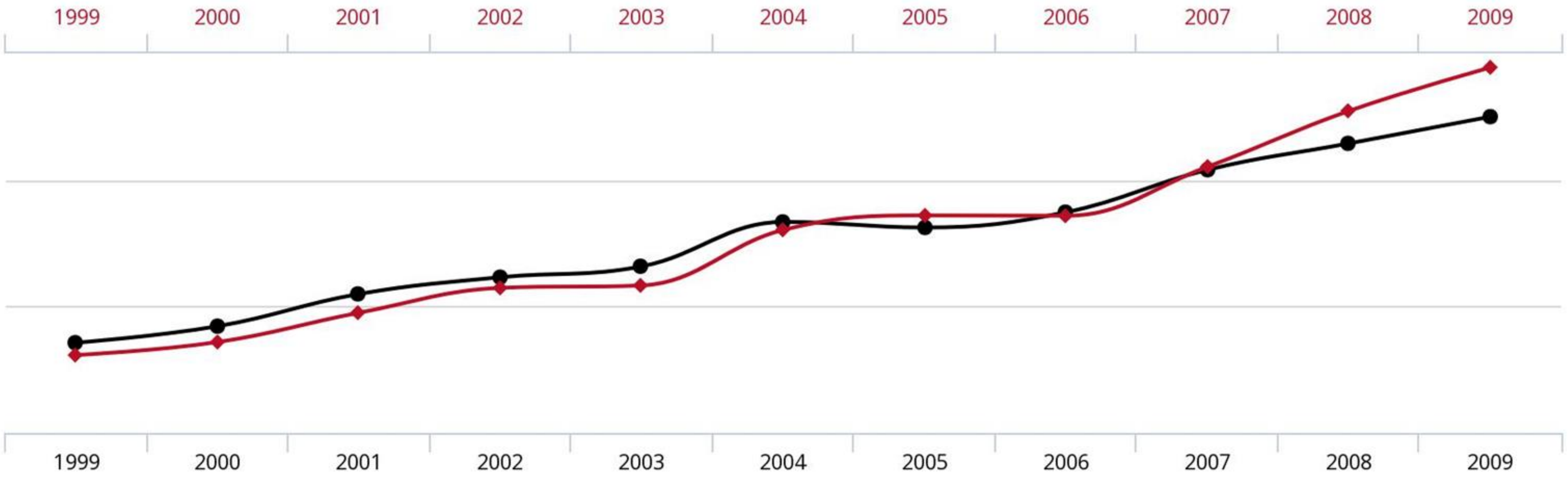
Evaluating results

1. we must analyse results not only their characteristics!
   e.g.: average points on a test
2. do we have a representative sample? AKA data enhancement
   our goal is to test, not to prove our hypothesis
3. Texas sharpshooter fallacy
   don’t ignore parts of data
4. rushing to conclusions
   e.g.: more accidents closer to home
5. correlation $\neq$ causation



$100
$50
$25
24.0
12.0
6.0
Dec
Jan
Feb
March
April
May
June
July
Aug
Sep
Oct
Nov
Time

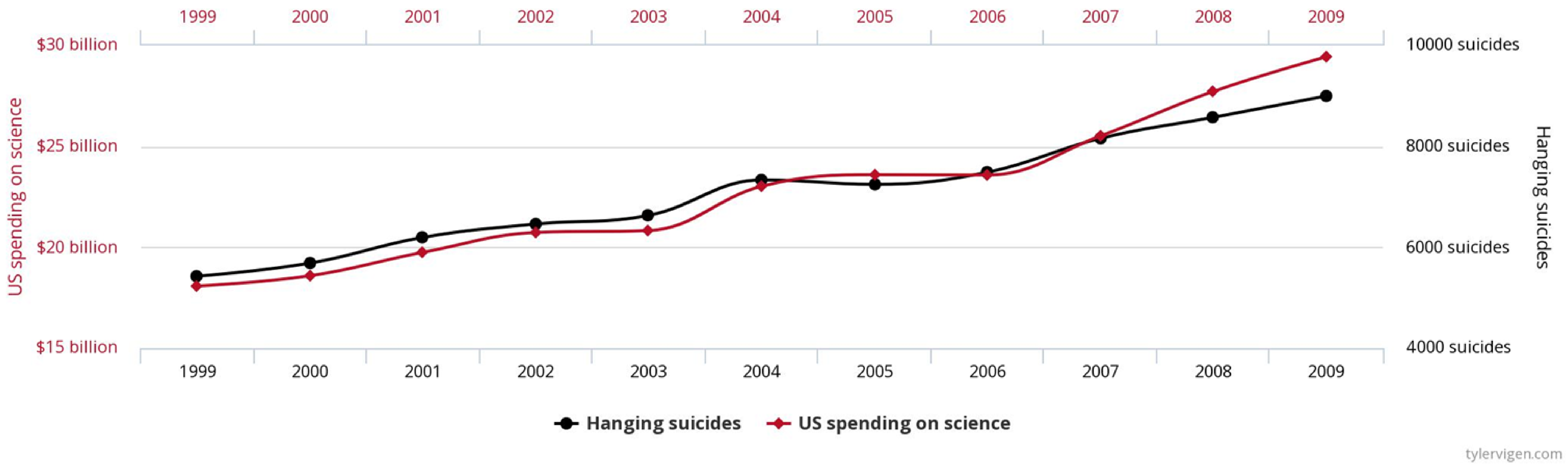

Sale of Ice Cream
Correlation Model
Forest Fires
$100
$50
$25
24.0
12.0
6.0
Dec
Jan
Feb
March
April
May
June
July
Aug
Sep
Oct
Nov
Time


1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009

US spending on science, space, and technology

correlates with

Suicides by hanging, strangulation and suffocation



Conclusion

- simulation structures
- simulation types
- random walk, biased random walk
- Monte Carlo simulations
- statistic distributions
- evaluating results
- fallacies when evaluating results