Simulation

(Benzerini Matematiksel Modelleme Kullanılarak Yazılımsal Oluşturulması)

Dr. Cahit Karakuş

What is Big Data Analytics and Machine Intelligence?

Big Data Analytics

Ölçek, çeşitlilik ve karmaşıklığı yönetmek, değer ve bilgi elde etmek için yeni teknikler ve algoritmalar gerektiren veri analizidir.

Veri Analizi: Görselleştirme ve istatistiksel olasılık metotları kullanılarak veri yığınından yorum yapmaya yönelik çıkarımlar elde etmek.

- Volume: scale of data
- Velocity: analysis of streaming data
- Variety: numerous forms of data
- Veracity: mitigating uncertainty of data
- Hacim: veri büyüklüğü
- Hız: Veri tansferi ya da veri işleme hızı
- Çeşitlilik: çok sayıda veri biçimi
- Doğruluk: veri içindeki belirsizliği (eksik, hata, kasdi, bozulma, kalinrasyon) azaltmak

Machine Learning

Karar vermeye yönelik tahminlerde bulunmak için hem verilerden hem de insan etkileşiminden öğrenerek performans artırımını sağlayan algoritmalar

Machine Intelligence

Makine Zekası, çevresini gözlemleyebilen ve gezgin olan ve kararlar alabilen özerk bir varlıktır.

Simülasyon

- Simüle etmek, fiziksel bir sistemin davranışını mantıksal ve matematiksel bir modelle yeniden oluşturmaktır.
- Pratik olarak bilgisayarlar, mantıksal bir modeli sayısal olarak simüle etmek için kullanılır. Bilgisayarlar ikili sayı sisteminde çalışmaktadır.
- Simülasyonlar, karmaşık sistemlerin performans değerlendirmesi ve tahmini davranışlarını gözetlemek için kullanılır:
  - akışkanlar dinamiği, kimya reaksiyonları (sürekli)
  - iletişim ağı modelleri: yönlendirme, tıkanıklıktan kaçınma, mobil… (ayrık)
  - Eğitim: Uçak, …
- Simülasyon, analitik yöntemlerden daha esnektir

What is a simulation?

- Simülasyon – genellikle bir bilgisayar kullanarak bir tesisin, bir cihazın, bir makinenin, bir organizasyonun, bir fabrikanın veya bir sistemin işleyişinin taklit edilmesidir.
  - Simüle edilen tesis aynı zamanda “sistem” olarak da adlandırılır.
  - Sistemin nasıl çalıştığı hakkında hem mantıksal hem de matematiksel varsayımlar/tahminler yapılır.
  - Bu varsayımlar sistemin mantıksal ya da matematiksel bir modelini oluşturur.
- İlgilenilen sorular hakkında kesin bilgi almak için matematiksel yöntemler de kullanabilir.
- Avantaj — Şüphesiz doğru şeye bakmak gerekir. Ancak bunu gerçek sistemle gerçekte denemeler yapmak çoğu zaman risklidir, imkansızdır. Sistem mevcut değil. Mevcut sisemlerde deneme yapmak risklidir, çünkü gerçek sistemler ile denenirse sonucu yıkıcı, pahalı veya tehlikeli olabilir.
- Modellerin birçok uygulaması vardır ve aşağıdaki gibi soruları yanıtlayabilir:
  - Bir kasırganın sonucu ne olur?

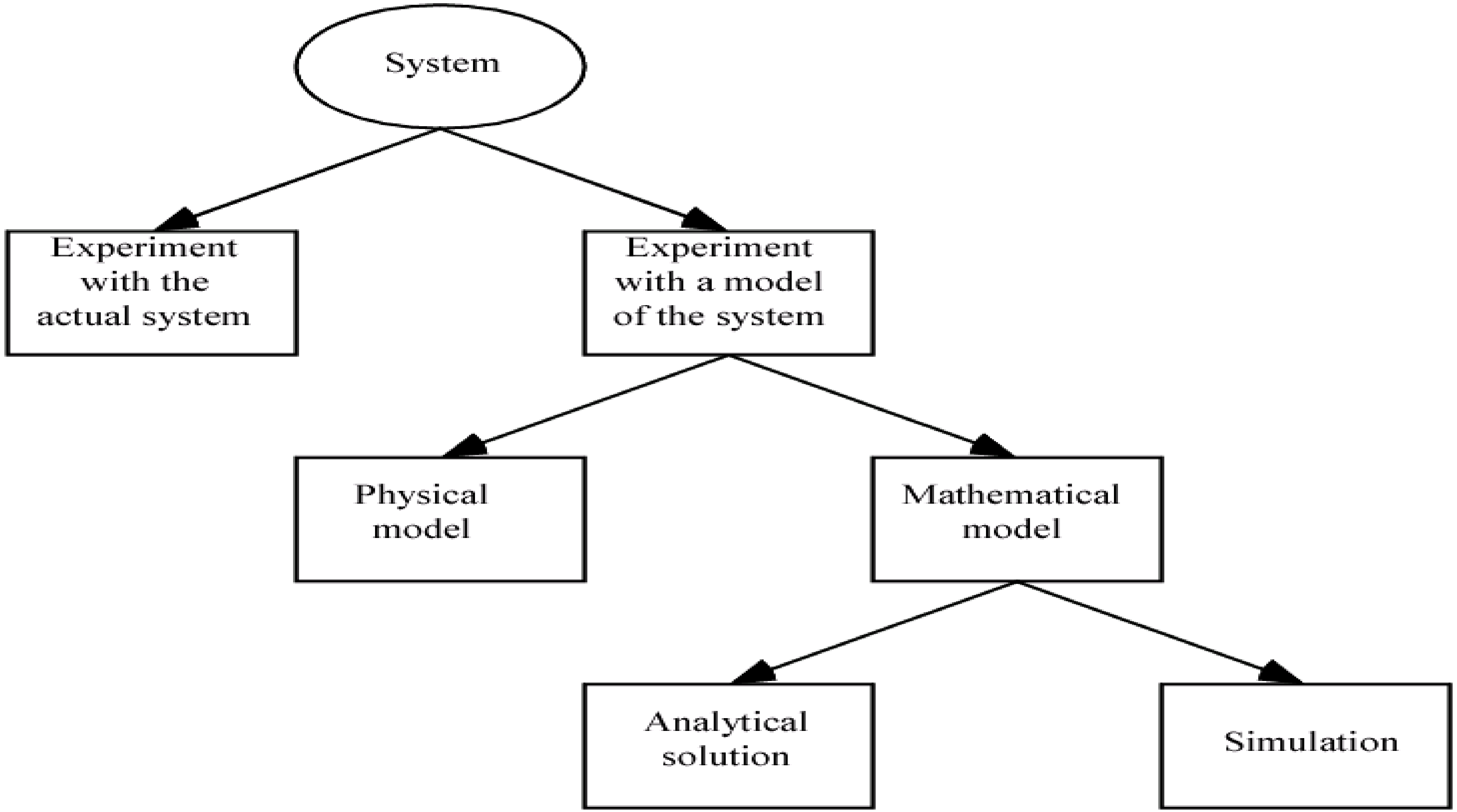
Simülasyon

- Eğitim ya da sistem performansını artırmak veya test etmek amacıyla, genellikle bilgisayar aracılığıyla gerçek sistemleri taklit etmek için yöntemler ve uygulamalar gerçekleştirilir.
- Artık “son çare” yaklaşımı olarak görülmemelidir. Birinci adım olmaldır.
- Günümüzde mühendisler, tasarımcılar ve yöneticiler için vazgeçilmez bir problem çözme metodolojisi olarak görülmektedir.
- Analitik çözüm mevcut olsa bile modelleri incelemek için kullanılabilir.
- Simülasyonun gerçek gücü, karmaşık mantıksal ya da matematiksel modelleri incelemektir.

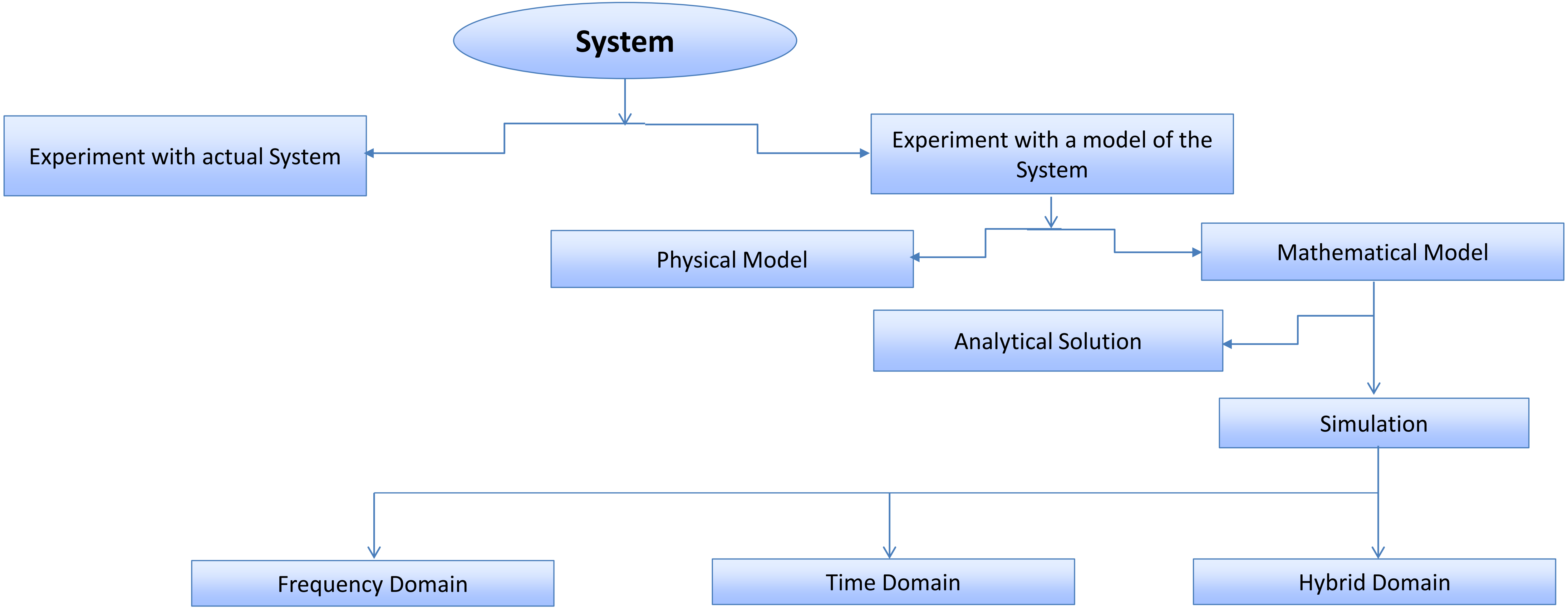
Preliminary (crude) definitions: real vs. simulated

- **Bir deneyim**, standart olmayan yöntemlerle kontrollü bir şekilde üretiliyorsa simüle edilir.
  - standart dışı = fiziksel çevre ile etkileşime giren bedenler/beyinler tarafından değil
  - Örneğin. şeytanlar, bilgisayar yazılımları
- **Bir yaşam (ya da bir davranış parçası),** simüle edilmiş deneyimlerden oluşuyorsa sanaldır.

How to study a system?



Ways to Study a System



When to use simulations?

- Simülasyonlar kullanılabilir:
  - Karmaşık sistemleri, yani analitik çözümlerin mümkün olmadığı sistemleri incelemek.
  - Var olmayan bir sistem için tasarım alternatiflerini karşılaştırmak.
  - Değişikliklerin mevcut bir sistem üzerindeki etkisini incelemek. Neden sistemi değiştirmiyorsunuz?
  - Analitik çözümleri güçlendirmek/doğrulamak için.
- Simülasyonlar kullanılmamalıdır:
  - Model varsayımları, kesin cevaplar elde etmek için matematiksel yöntemlerin kullanılabileceği şekilde basitse (analitik çözümler)

Terminology

- ***Sistem:*** *Mantıksal bir amaç doğrultusunda birlikte hareket eden ve etkileşimde bulunan nesneler topluluğu.*
  - *Örneğin. Bir süper mağazada 10 veya daha az ürüne sahip müşterilere ekspres servis sağlamak için gereken kasiyer sayısını belirlenmesi - sistem ekspres kasiyerlerden ve 10 veya daha az ürüne sahip müşterilerden oluşur.*
- ***Bir sistemin durumu:*** *Belirli bir zamanda bir sistemi karakterize etmek için gerekli değişkenlerin ve değerlerinin toplanması*
  - *İstenen hedeflere, performans ölçütlerine bağlı olabilir*
  - *SS Örneği: Ekspres kasiyer sayısı, 10 veya daha az ürüne sahip müşterilerin varış zamanı vb.*
- ***Olay:*** *Sistem durumunda bir değişiklik olmasıdır.*
  - *Müşterinin gelişi, hizmetin başlaması ve müşterinin ayrılması*

Application Areas

- Bilgisayar sistemlerinin tasarımı ve performans değerlendirmesi
  - İletişim ağları için donanım gereksinimlerinin veya protokollerin belirlenmesi
  - CPU Zamanlama algoritmalarını incelemek
  - Web önbelleğe alma politikalarının değerlendirilmesiİmalat sistemlerinin tasarımı ve analizi
- Bir üretim hattının işletilmesi
  - Hizmet organizasyonları için tasarımların değerlendirilmesi
- Çağrı merkezleri, fast food restoranları, hastaneler ve postanelerin çaışma verimliliğinin incelenmesi
- Askeri silah sistemlerinin veya lojistik gereksinimlerinin değerlendirilmesi
- Havaalanları, otoyollar, limanlar ve metrolar gibi ulaşım sistemlerinde yolcu erişmi ve transferindeki yoğunluğa göre tasarlamak ve işletmek
- Finansal veya ekonomik sistemleri analiz etmek

Simülasyonun Avantajları

- Her şeyi olduğu gibi modelleme esnekliği (dağınık ve karmaşık olsa bile)
- Modellemede belirsizliğe, durağanlığa izin verir
- Gerçek sistemin devam eden işleyişini kesintiye uğratmadan yeni politikalar, işletim prosedürleri keşfedilebilir.
- Yeni donanım tasarımları, fiziksel yerleşimler, sistemleri satın almaları için kaynak ayırmadan test edilebilir.
- Olguyu hızlandırmak veya yavaşlatmak için zaman sıkıştırılabilir veya genişletilebilir.
- Yazılım, bilgi işlem ve bilgi teknolojisindeki gelişmeler, simülasyonun popülaritesini artırıyor

Kötü Haber

- Kesin cevaplar alınmaz, yoruma açıktır; sadece tahminler ve kestirimler alınır.
- Model yapımı özel eğitim gerektirir.
- Simülasyon modelleme ve analizi zaman alıcı ve pahalı olabilir.
- Simülasyon sonuçlarının yorumlanması zor olabilir.
- Stokastik simülasyonlardan rastgele çıktı alınır.
- İstatistiksel tasarım, simülasyon deneylerinin analizi

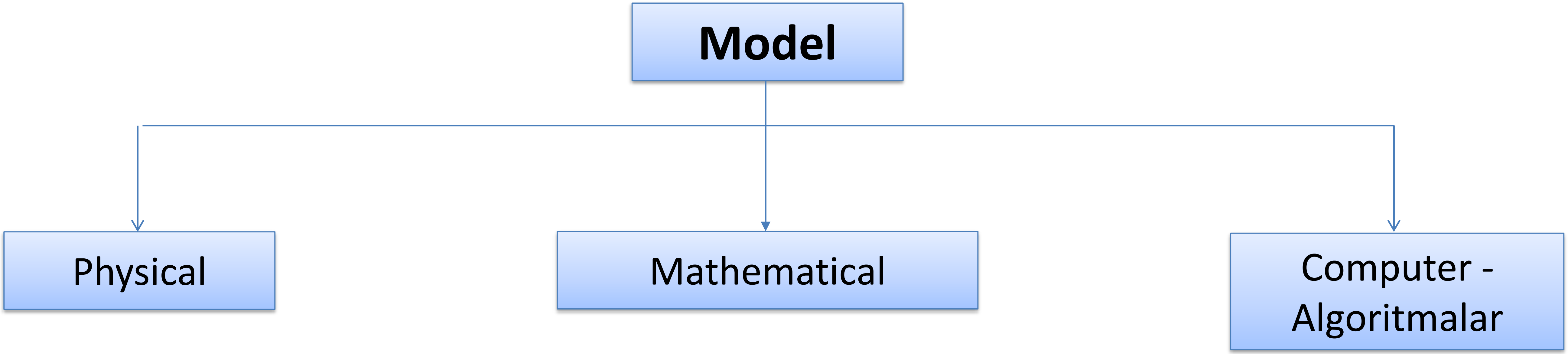
Simulation Model Classification

14

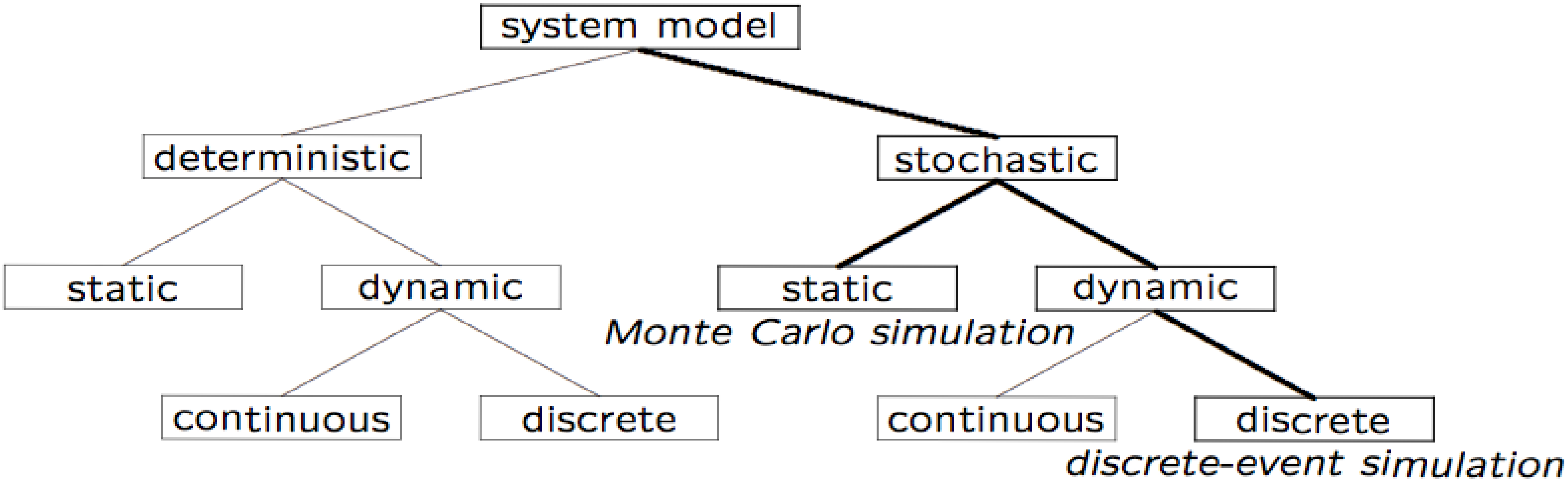
Simulation Model Classification

- Continuous-time - Discrete-time models
- Continuous-event - Discrete-event models
- Deterministic - Probabilistic (Stochastic) models
- Static - Dynamic models
- Linear - Non-linear models
- Open - Closed models
- Kararlı – Sistemler
- White box, black box and gray box

Types of Models



Simulation Model Taxonomy



What is Mathematical Model?

Mathematical Model is a set of mathematical equations (e.g., differential eqs.) that describes the input-output behavior of a system.

What is a mathematical model used for?

- Simulation
- Prediction/Forecasting
- Prognostics/Diagnostics
- Design/Performance Evaluation
- Control System Design

Discrete or Continuous models

- **Discrete model**: the state variables change only at a countable number of points in time. These points in time are the ones at which the event occurs/change in state.
- **Continuous**: the state variables change in a continuous way, and not abruptly from one state to another (infinite number of states).

Static or Dynamic models

- **Dynamic**: Durum değişkenleri zamanla değişir (System Dynamics, Discrete Event, Agent-Based, Econometrics?) IP adreslemede her bağlantıda farklı IP adresi.
- **Static**: Zamanda tek bir noktada anlık görüntü (Monte Carlo simulation, optimization models, etc.) IP adreslemede her bağlantıda tek size özel IP adresi.

Deterministic, Stochastic or Chaotic

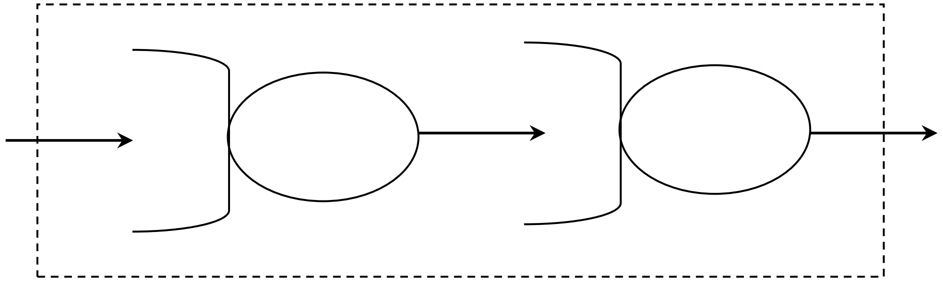
- Deterministik model, gelecekteki davranışı tamamen öngörülebilir olan bir modeldir. Sistem mükemmel bir şekilde anlaşılırsa, o zaman ne olacağını kesin olarak tahmin etmek mümkündür.
- Stokastik model, davranışı tamamen tahmin edilemeyen modeldir.
- Kaotik model, tamamen tahmin edilemeyen bir davranışa sahip deterministik bir modeldir. Sistem girişindeki çok küçük değişimlere, çok yüsek hassasiyetle bağımlı olmasıdır.

Deterministic and Stochastic Models

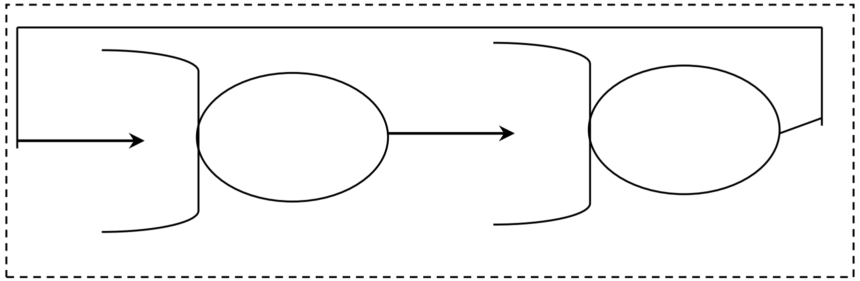
- Deterministik modeller deterministik sonuçlar üretir
- Stokastik veya olasılıksal modeller rastgele etkilere tabidir
  - Tipik olarak, bir veya daha fazla rasgele girdileri vardır (örneğin, müşterilerin gelişi, hizmet süresi vb.).
  - Stokastik modellerden elde edilen çıktılar, sistemin gerçek özelliklerinin “tahminleridir”.
  - Deneyleri birçok kez tekrarlama ihtiyacı
  - Sonuçlara güvenmek gerekiyor

More on models

- Static and dynamic models
  - – Static models – system state independent of time
  - – Dynamic models - system state change with time
- Linear and non-linear models
  - – Linear models – output is a linear function of input parameters
- Open and closed models



(a) Open Model



(b) Closed Model

Types of Simulation

24

Types of Simulation

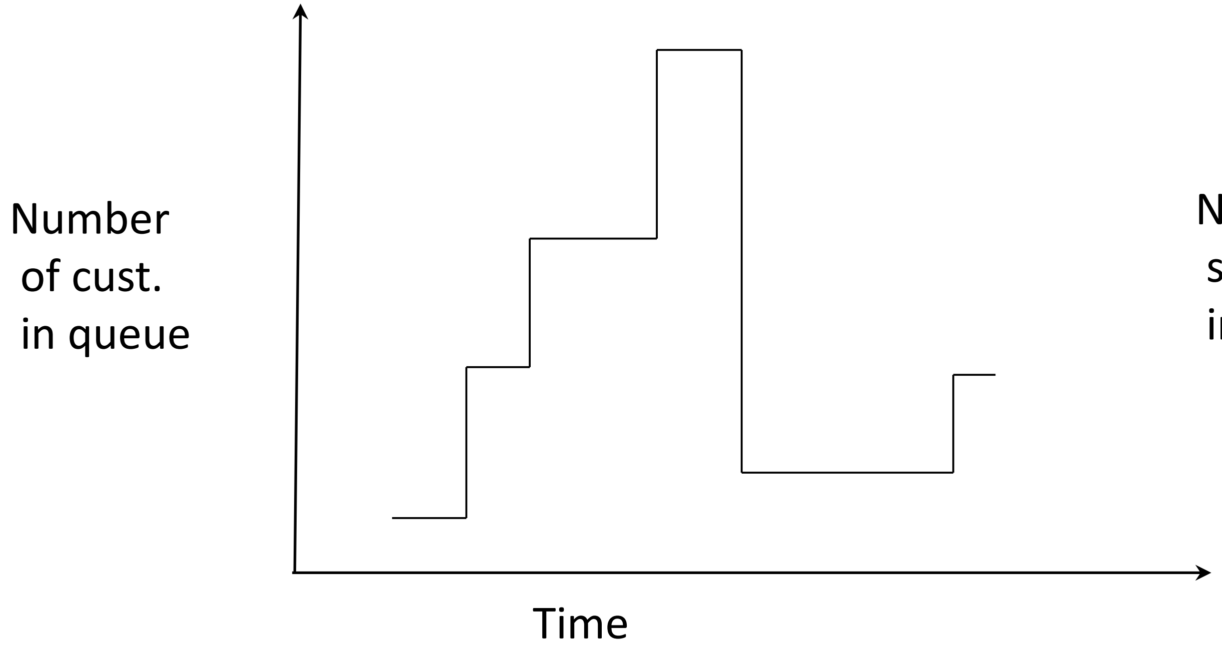
- ***Monte Carlo simulation***
- **Trace-driven simulation**
- **Discrete-event simulation:** a simulation using a discrete-event (also called discrete-state) model of the system
  - E.g., Widely used for studying computer systems
- **Continuous-event simulation:** uses a continuous-state models
  - E.g., Widely used in chemical/pharmaceutical studies
- Our focus will be on *discrete-event* systems.

Types of Simulation

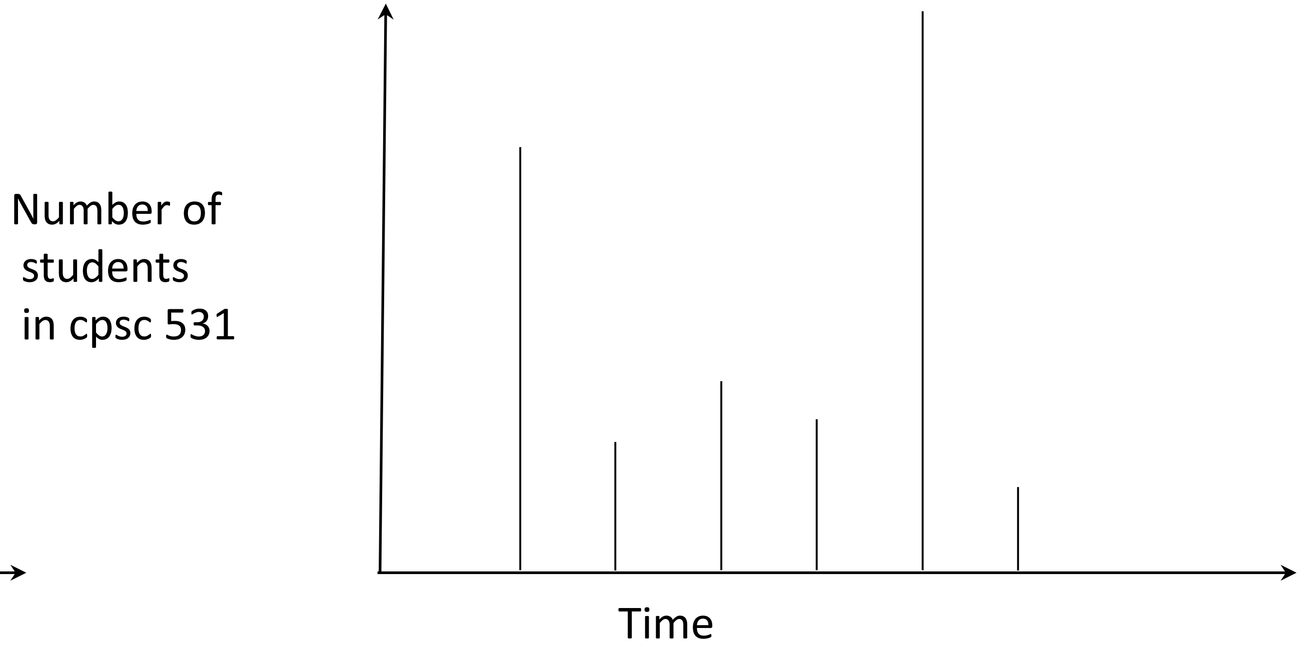
- ***Monte Carlo simulation***
  - Zaman öğesi yok (genellikle)
  - Olasılıksal olmayan ifadeleri (örneğin, bir integral) olasılıksal yöntemler kullanarak değerlendirmek için kullanılır
  - Çok çeşitli matematiksel problemler

- ***Trace-driven simulation***
  - Extensively used in computer systems performance evaluation; e.g., paging algorithms
  - Advantages: credibility, easy validation, less randomness, accurate workload
  - Disadvantages: complexity, only a snap-shot, representative?, single point of validation

Continuous and Discrete-time models

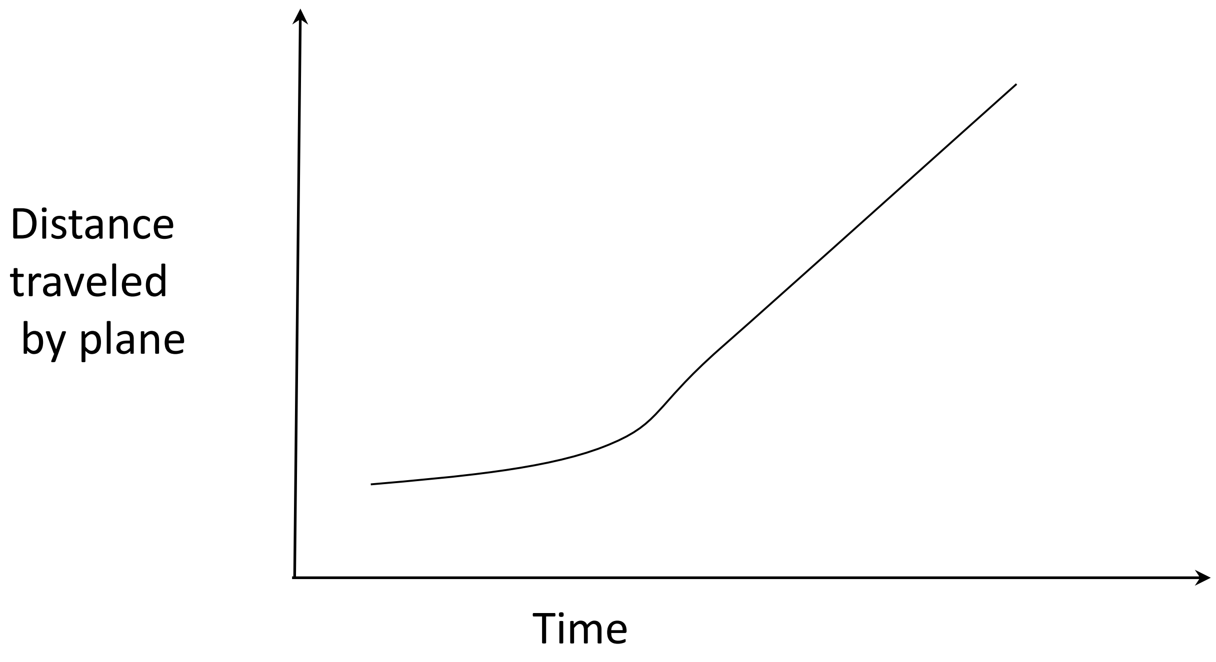


(a) Continuous-time

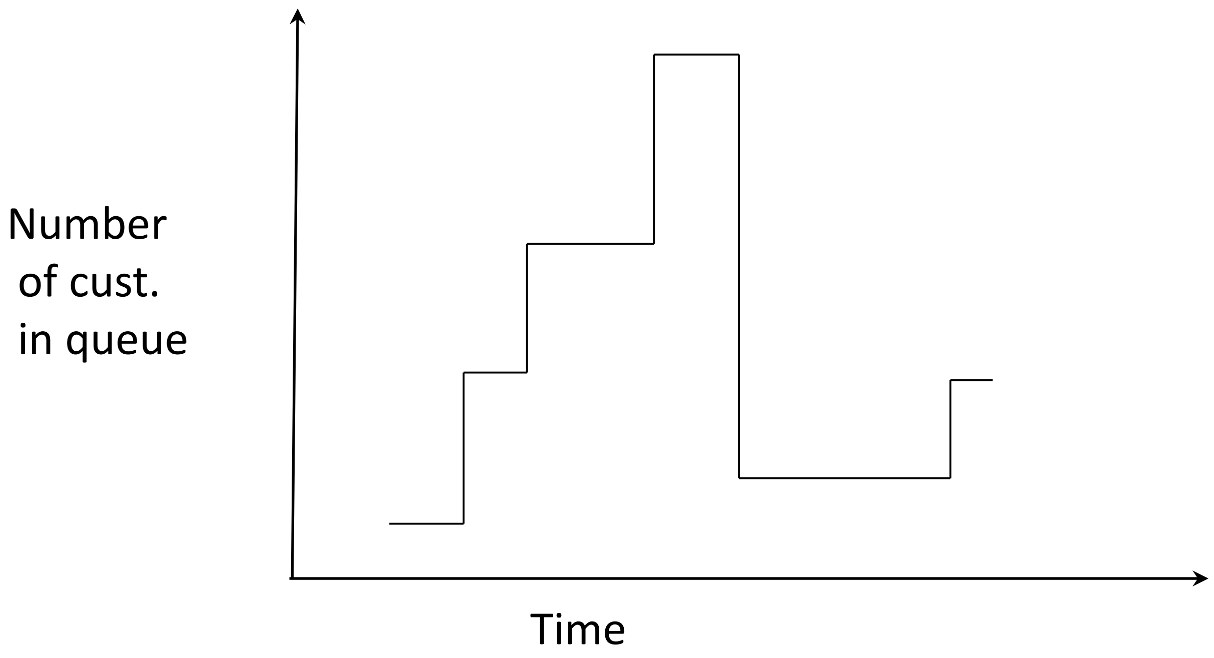


(b) Discrete-time

Continuous and Discrete-event models



(a) Continuous-event



(b) Discrete-event

Aggregate and Individual models

- **Aggregate model**: we look for a more distant position. Modeler is more distant. Policy model. This view tends to be more deterministic.
- **Individual model**: modeler is taking a closer look of the individual decisions. This view tends to be more stochastic.

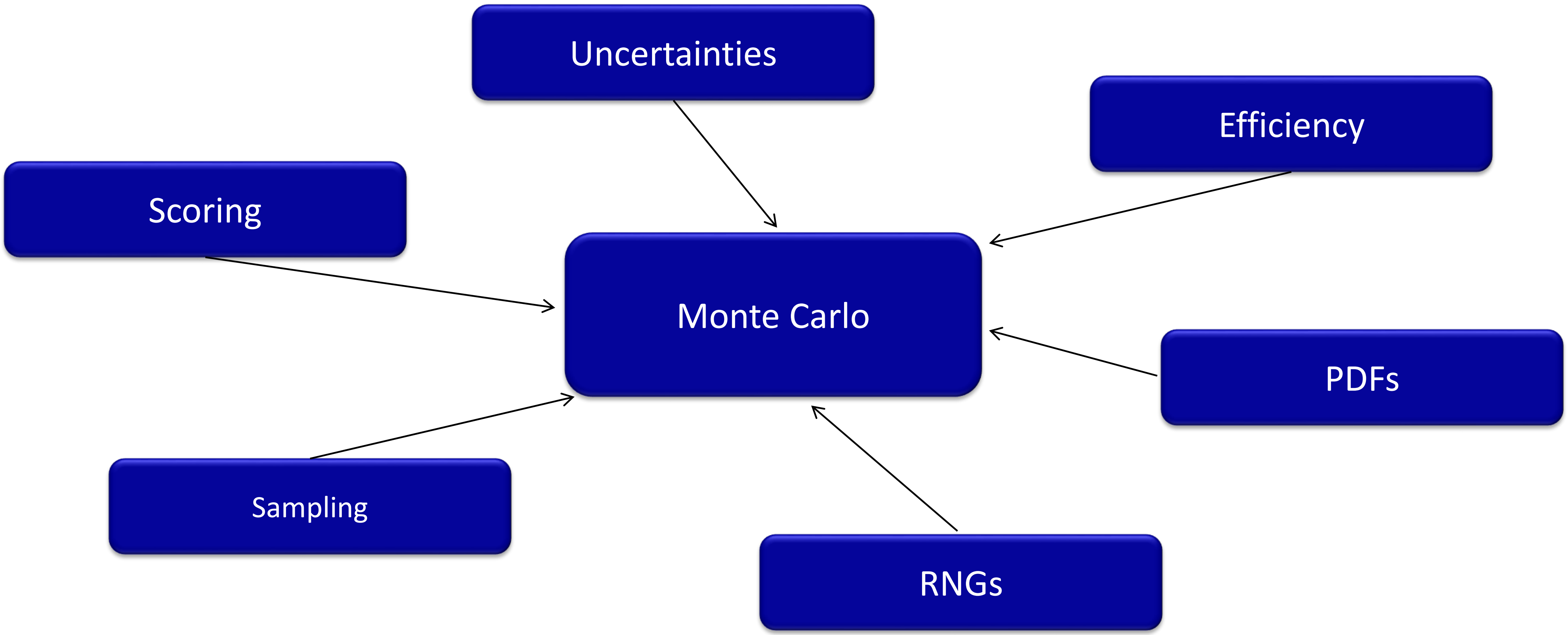
Monte Carlo çalışmaları

- Monte Carlo çalışmaları, 1940'lara dayanan en eski bilgisayar tekniklerinden biridir ve analitik bir çözüme doğrudan izlenemeyen bir dizi soruyu analiz etmemize izin verir.
- Bazen bu problemler o kadar karmaşık olabilir ki, büyük örneklemli (asimptotik) cevabı çözemeyiz. Diğer durumlarda, kolayca hesaplanamamayan küçük bir örneklem büyüklüğü veya sınırlı sayıda durum için yanıtla ilgilenebiliriz.
- Bu tekniğin muazzam bir uygulama yelpazesi, saf istatistikleri, proje değerlendirmesi, makroekonomik modelleri, türev fiyatlandırması vardır - stokastik süreçlerin veya zorlu matematiksel formüllerin olduğu hemen hemen her yerde.

Monte Carlo metotu nedir?

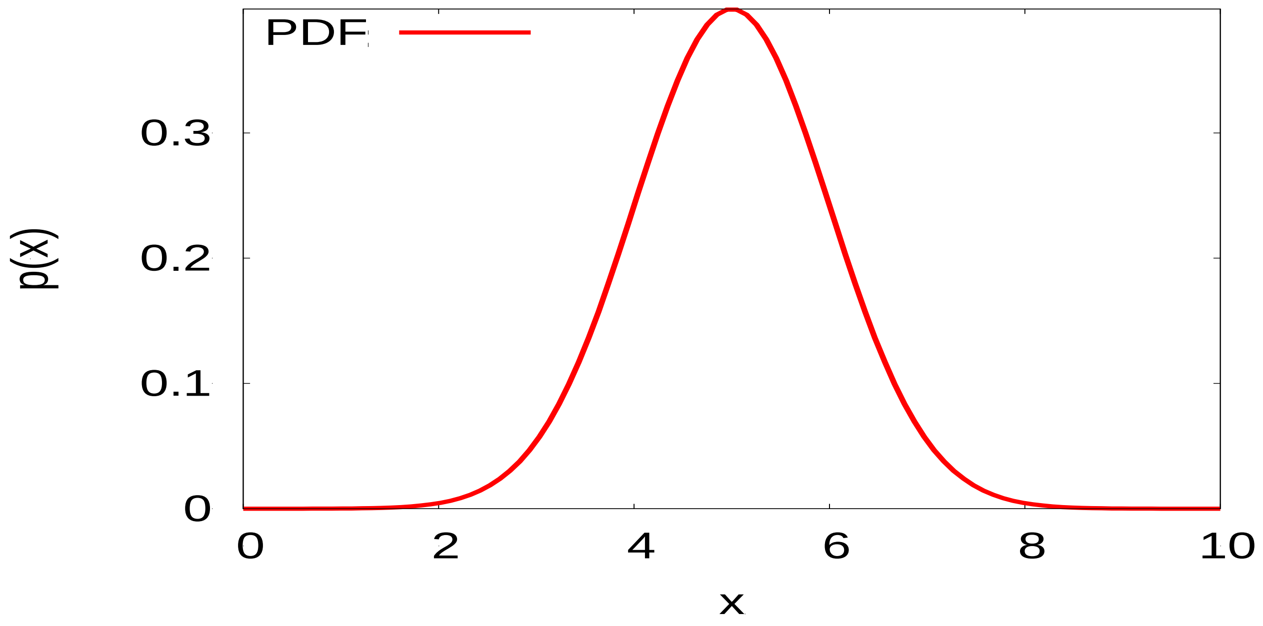
- Monte Carlo yöntemi, basit nesne-nesne veya nesne-çevre ilişkilerine dayalı olarak diğer nesnelerle veya çevreleriyle etkileşime giren nesneleri modelleyen sayısal bir çözümdür. Söz konusu sistemin temel dinamiklerinin doğrudan simülasyonu yoluyla doğayı modelleme girişimini temsil eder. Bu anlamda Monte Carlo yöntemi yaklaşımında esasen basittir - mikroskobik etkileşimlerinin simülasyonu yoluyla makroskopik bir sisteme bir çözüm.

Outline


Uncertainties
Efficiency
Scoring
Monte Carlo
PDFs
Sampling
RNGs

What is a PDF?

Probability distribution functions (PDFs)



Restrictions:

$$1 \geq p(x) \geq 0$$

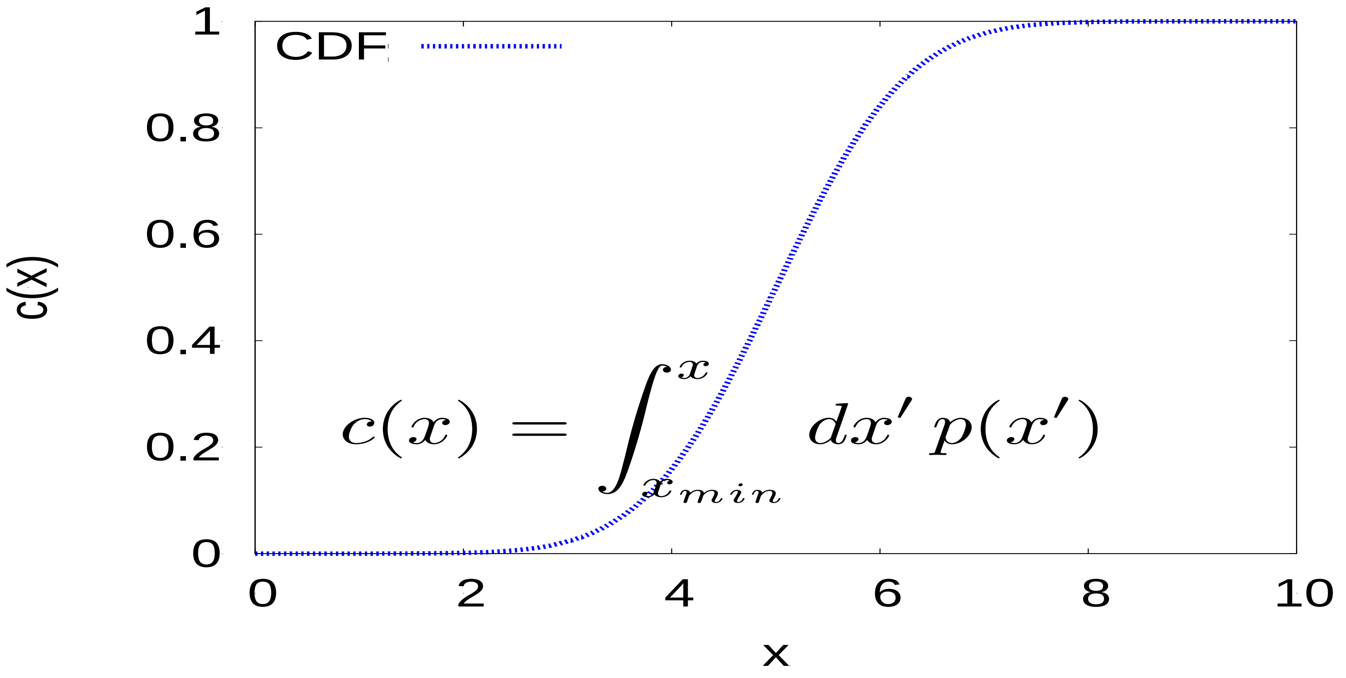
$$\int_{x_{min}}^{x_{max}} p(x)\, dx = 1$$

$$-\infty < x_{min} < x_{max} < \infty$$

- In Physics a PDF can represent the most likely spatial position, kinetic energy, momentum direction, etc., of a particle.
- The PDF can be obtained by either theoretical models and parameterization to experimental data, such as cross sections, etc.

But PDFs are hard to use…

- The domain of such functions is so diverse.
- Fortunately, associate to each PDF there exists a Cumulative Distribution Function (CDF)



Features:

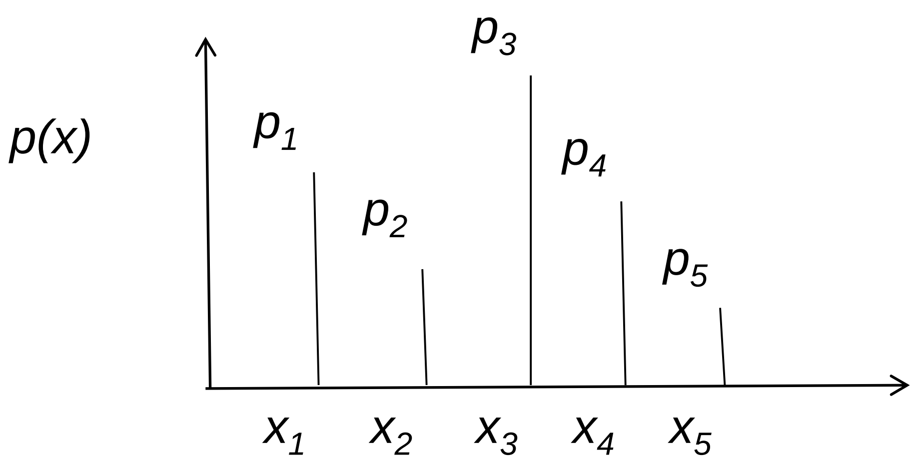
$$p(x) = \frac{dc(x)}{dx}$$

$$c(x_{max}) = 1$$

$$c(x_{min}) = 0$$

In the practice, we use the discrete form of PDF and CDF

Discrete PDF

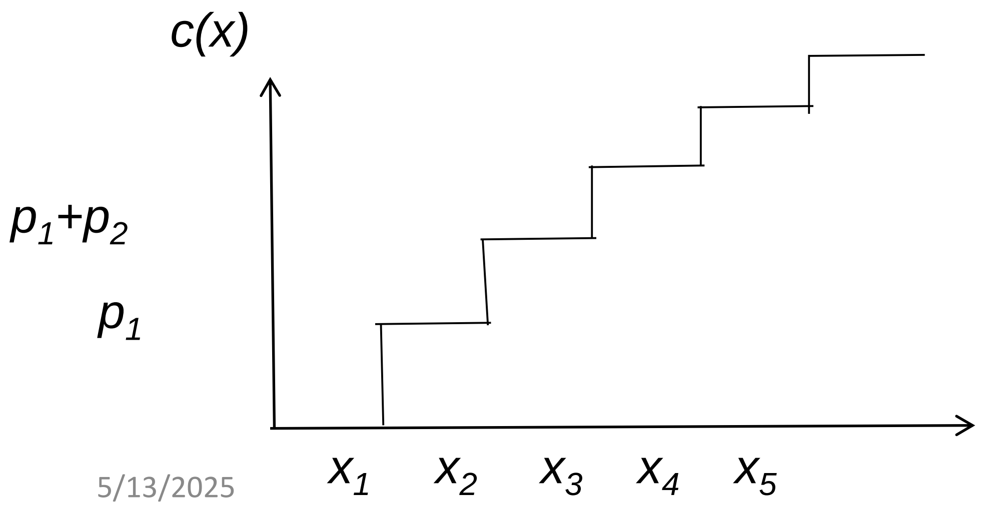


$$p(x_i) = p_i \delta(x - x_i)$$

$$\sum_i p(x_i)(\Delta x_i) = 1$$

$$p(x_i) \geq 0$$

Discrete CDF



$$c(x) = \sum_{x_i < x} p_i$$

$$0 \leq c(x) \leq 1$$

(Pseudo)Random numbers

- Non-correlated sequences of numbers generated by an iterative equation.
- Repeatability after a very long number of random numbers.
- Non-uniform sequence.
- Reproducible: "seed".

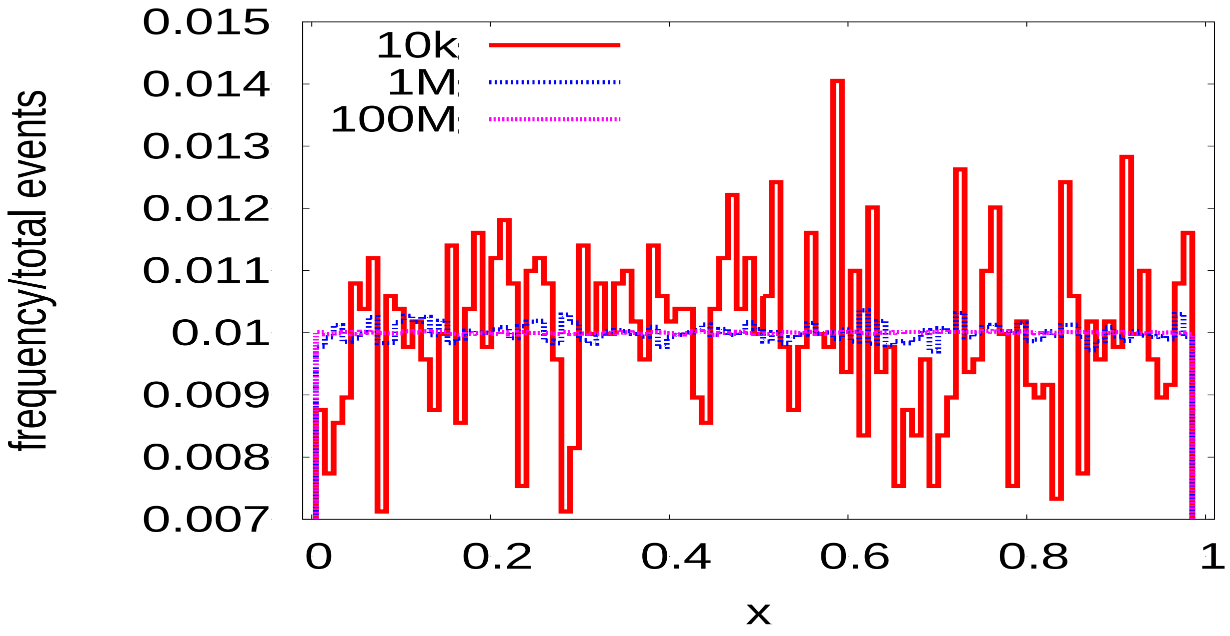
$$I_{j+1} = (aI_j + c) \bmod m$$

where:

$$a = 663608941$$

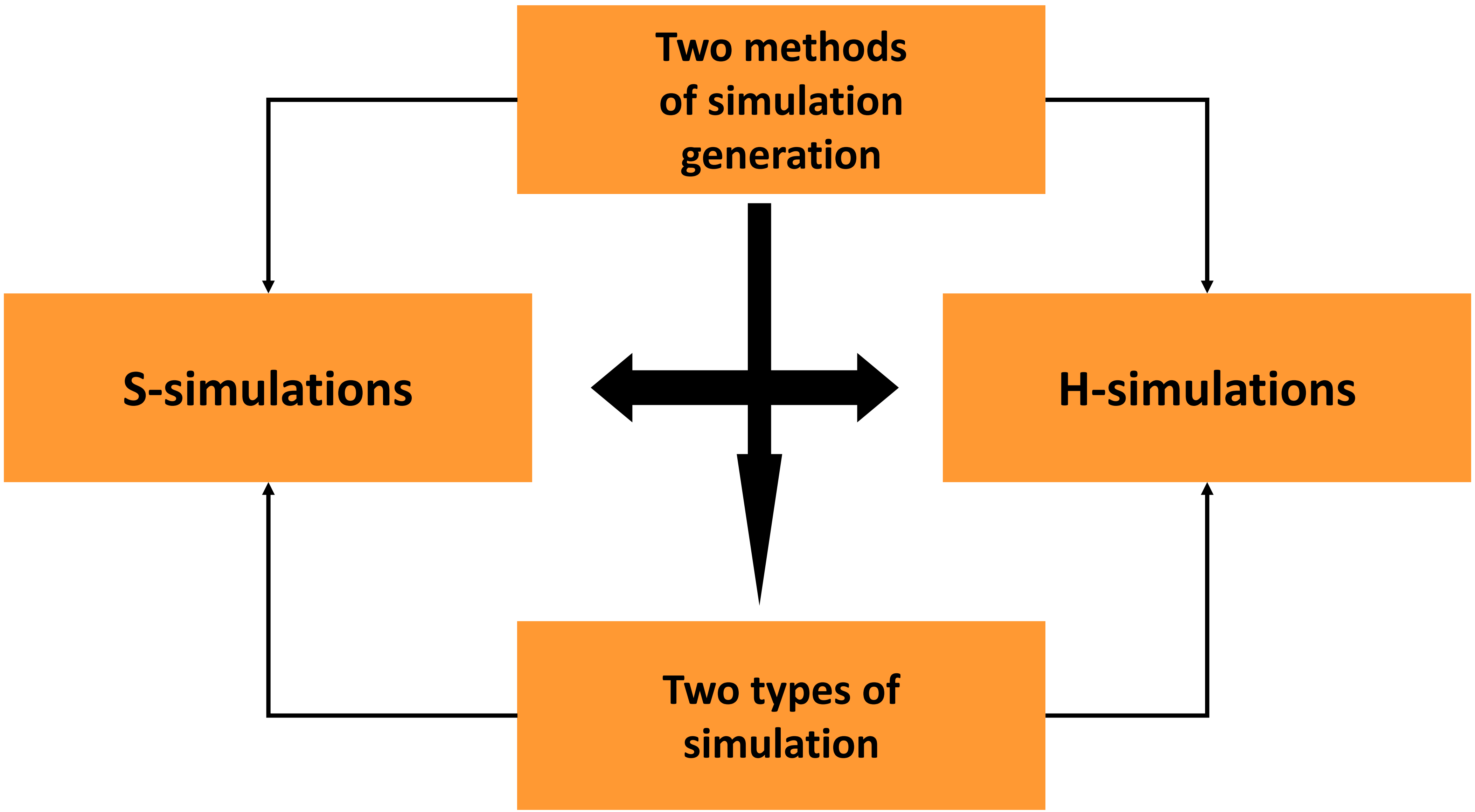
$$c = 0$$

$$m = 2^{32}$$



Where to use random numbers?

- In the practice, we can generate uniform random numbers in [0, 1].
- But we need random numbers that obey the PDF of the physical process we want to simulate. These numbers are not uniform!
- Sampling techniques allow to recover such numbers from:
  - Analytical distribution: theoretical models
  - Tabulated distribution: experimental data


Two methods of simulation generation
S-simulations
H-simulations
Two types of simulation

S-simulations:

- conscious lives generated by running software on computer
- (NOT brains!)

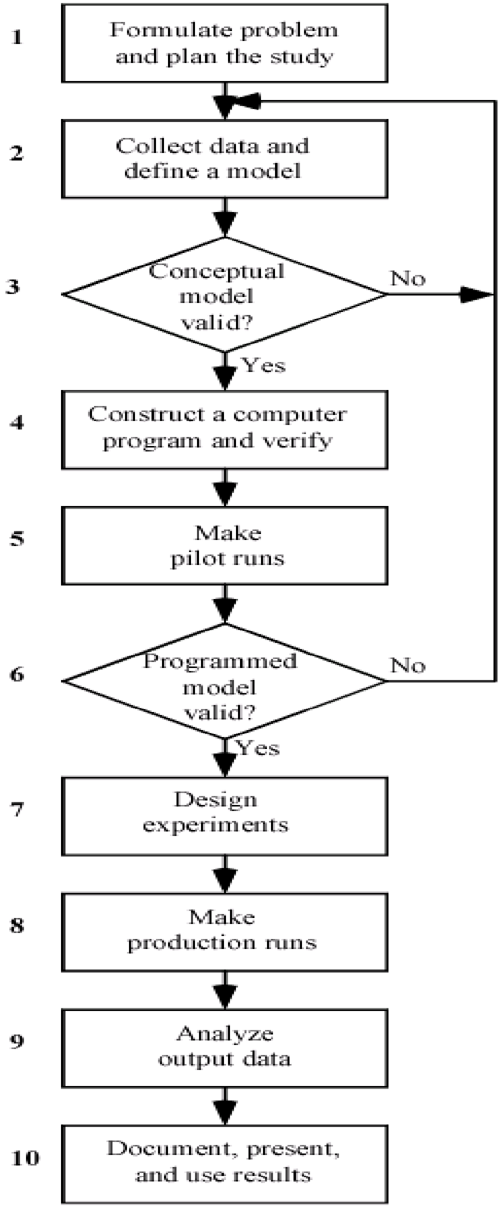
H-simulations:

- conscious lives produced by directly tampering with neural hardware
- Matrix-style
- vat-brain

COMPUTER
ALONE

COMPUTER
+ BRAIN

Steps in a Simulation Study



Conceptual: Kavramsal, kavrayan, anlayan

Monte Carlo Simulation

1. What is Monte Carlo? Basic Principles
2. Random Numbers
3. Sample Sizes
4. Distributions
5. Monte Carlo Simulation Examples
6. Monte Carlo Simulation Exercises

1 What is Monte Carlo? Basic Principles

2 Random Numbers

3 Sample Sizes

4 Distributions

5 Monte Carlo Simulation Examples

6 Monte Carlo Simulation Exercises

- History

Stanisław Ulam Polish-American scientist in the fields of mathematics and nuclear physics (worked for the Manhatan project – thermonuclear weapons)

While **playing solitaire** during his recovery from surgery (game which he kept losing) he wondered how many games he had to play to get a win.

So he thought about playing hundreds of games to estimate statistically the probability of a successful outcome and the availability of computers made such statistical methods very practical.

(mechanical simulation of random diffusion of neutrons)

Ulam had an uncle who went often to Monte Carlo casino to gamble so he and his colleagues named the technique "MONTE CARLO"

Monte Carlo simulation: Çıkarımsal istatistik ilkelerini (Aritmetiksel ortalama, Standart sapma) kullanarak bilinmeyen bir miktarın değerini tahmin etme yöntemi

Population: set of examples
Sample: a proper subset of the population
Key fact: a ***random*** sample tends to exhibit the same properties as the population from which it is drawn

- ***Monte Carlo simülasyonu, rastgele değişkenlerin müdahalesi nedeniyle kolayca tahmin edilemeyen bir süreçte farklı sonuçların olasılığını modellemeye izin veren bilgisayarlı bir matematiksel tekniktir.***
- ***Monte Carlo simülasyonu, simülasyon sırasında üretilen rastgele sayılar dizisine bağlıdır.***

- Genellikle ilgilendiğimiz fiziksel süreçler (sistem) karmaşıktır.
- Belirli bir süreci incelemek için, bir modelin (sistemin basitleştirilmiş gösterimi) kullanılmasını gerektirebilir.
- Sistemin davranışını taklit etmek için modeli uyguladığımızda, simülasyonları çalıştırdığımızı söylüyoruz.
  - - Modeller bilimin, mühendisliğin, ticaretin vb. temel araçlarıdır.
  - - Gerçekliğin soyutlanması bu nedenle her zaman güvenilirlik sınırlarına sahiptir.

Monte Carlo simulation: technique that combines ***distributions*** with ***random number generation***

Random numbers can be generated in different ways

Any variable has a probability distribution for its occurrence

Best way to relate random number to a variable is to use **cumulative probability distribution**

(probability density functions – pdf)

Bahar aylarında klon paketleri (80 fide) için günlük talep incelenmiştir ve olasılıklar şunlardır:

Relative frequencies (probability)

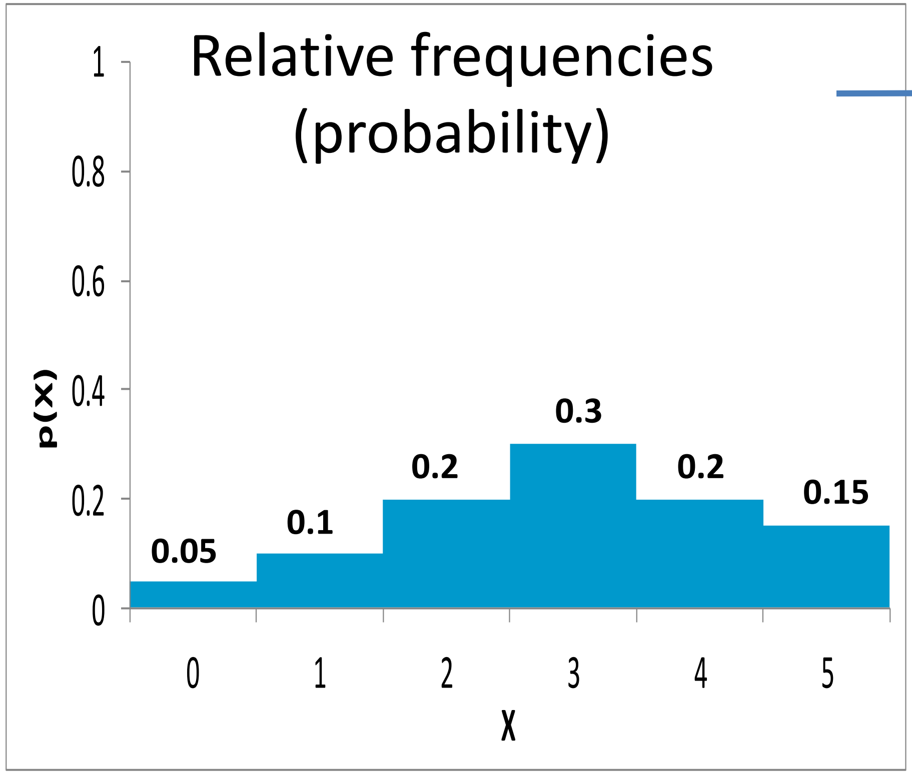
Nr packs ordered	probability
0	0.05
1	0.1
2	0.15
3	0.3
4	0.25
5	0.15

If the distribution is known , **WHY** do we use **random numbers** to simulate it?

BECAUSE, although the **probability is known** (the relative frequency of each demand level), **the order of occurrence is not**

It is **the order of occurrence** (which is assumed random) **which we want to simulate**

Assume that the demand/day is given by:



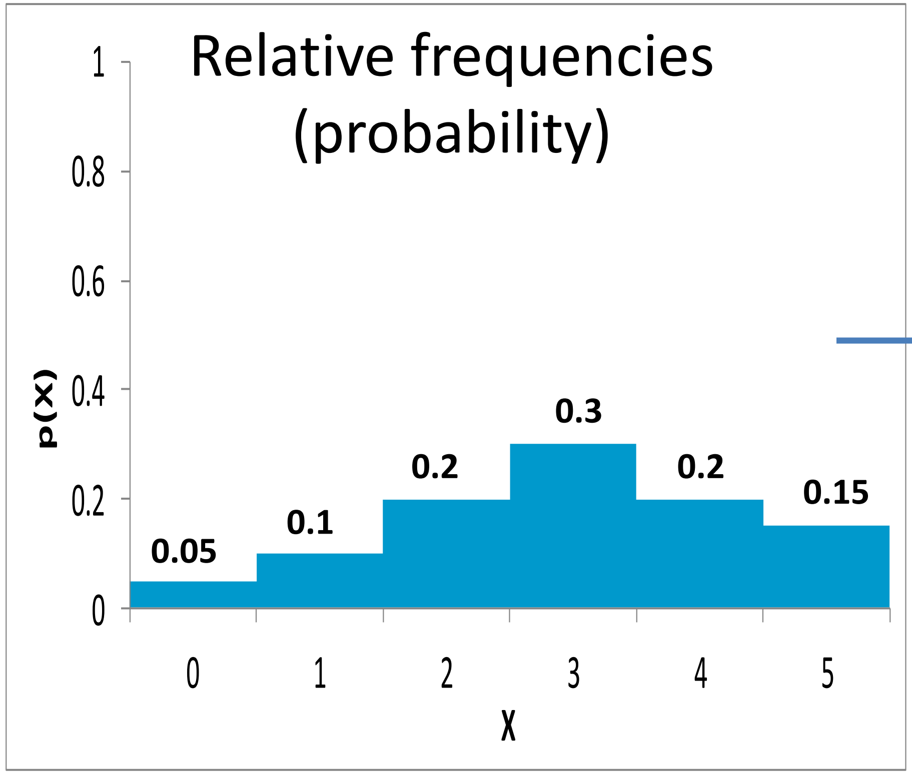
Sample: a proper subset of the population

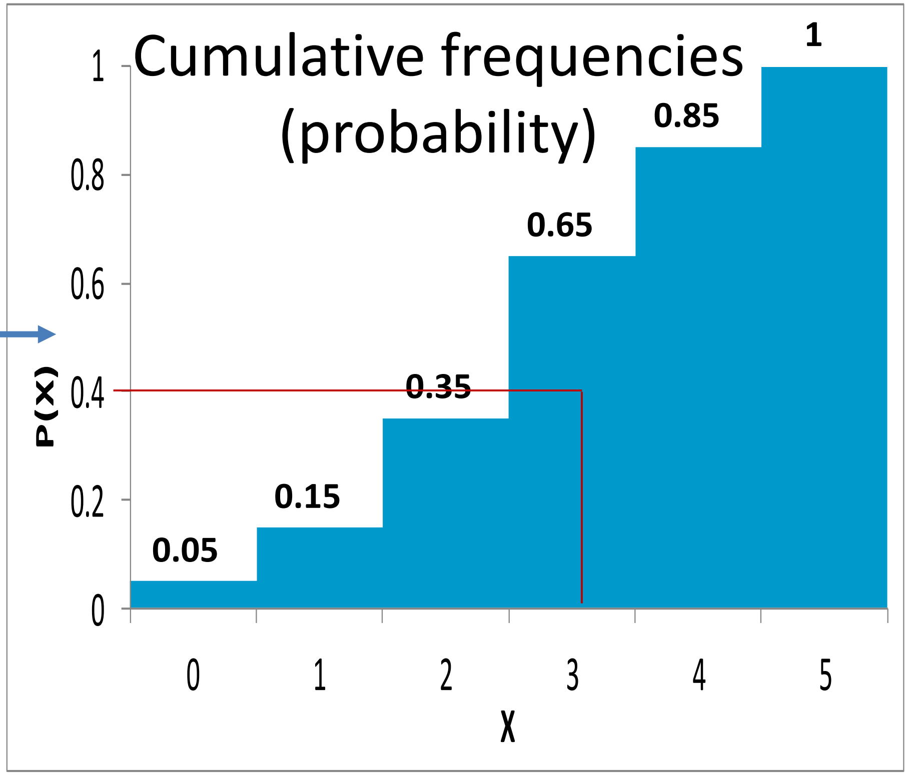
If the distribution is known , **WHY** do we use **random numbers** to simulate it?

BECAUSE, although the **probability is known** (the relative frequency of each demand level), **the order of occurrence is not**

It is **the order of occurrence** (which is assumed random) **which we want to simulate**

Assume that the demand/day is given by:

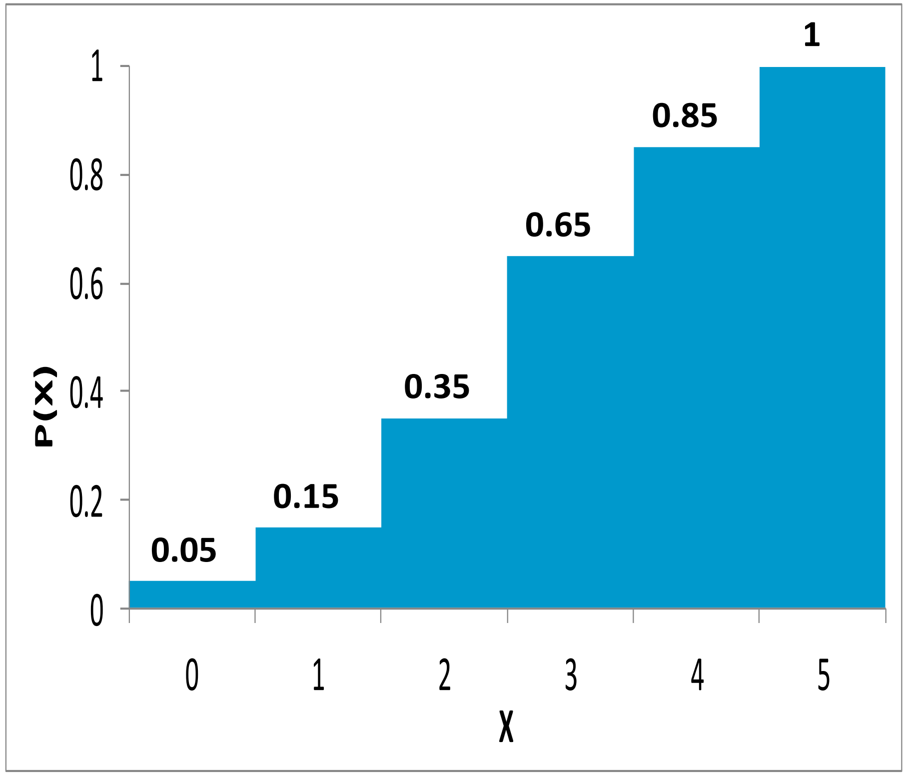




Assume that the demand/day is given by:

Demand (x)	Cumulative frequencies	Interval for random numbers
0	0.05	0 – 4
1	0.15	5 – 14
2	0.35	15 – 34
3	0.65	35 – 64
4	0.85	65 – 84
5	1	85 – 99

Cumulative frequencies (probability)



Assume that the demand/day is given by:

Simulate the demand for 10 days

Demand (x)	Cumulative frequencies	Interval for random numbers
0	0.05	0 – 4
1	0.15	5 – 14
2	0.35	15 – 34
3	0.65	35 – 64
4	0.85	65 – 84
5	1	85 – 100

day	Random number	demand
1	14	1
2		
3		
4		
5		
6		
7		
8		
9		
10		

Assume that the demand/day is given by:

Simulate the demand for 10 days

Demand (x)	Cumulative frequencies	Interval for random numbers
0	0.05	0 – 4
1	0.15	5 – 14
2	0.35	15 – 34
3	0.65	35 – 64
4	0.85	65 – 84
5	1	85 – 100

day	Random number	demand
1	14	1
2	74	4
3		
4		
5		
6		
7		
8		
9		
10		

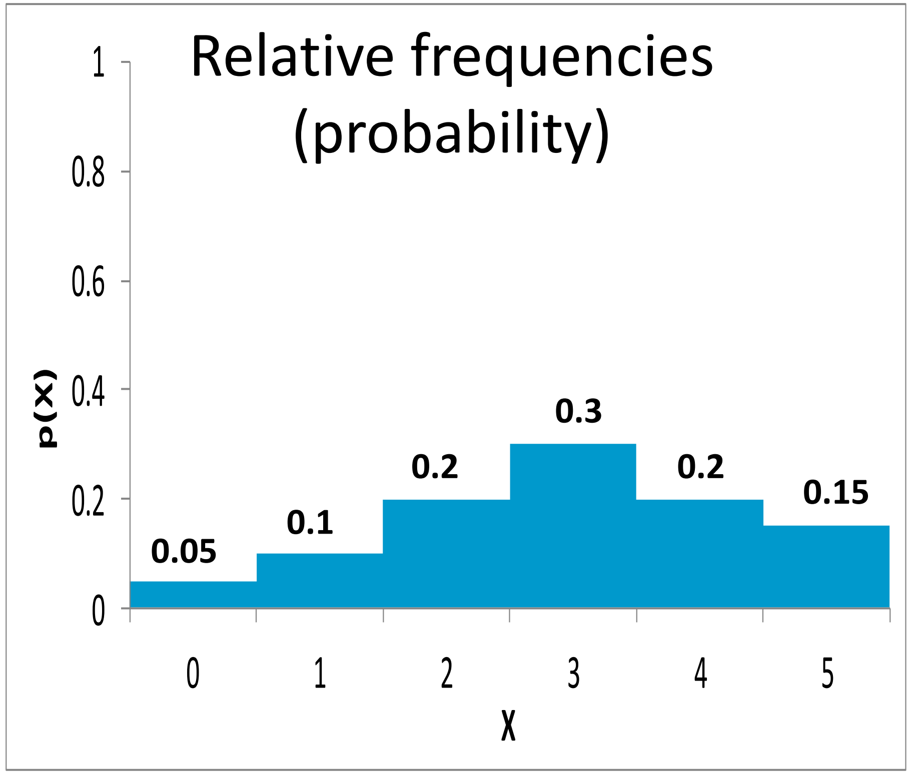
Assume that the demand/day is given by:

Demand (x)	Cumulative frequencies	Interval for random numbers
0	0.05	0 – 4
1	0.15	5 – 14
2	0.35	15 – 34
3	0.65	35 – 64
4	0.85	65 – 84
5	1	85 – 100

Simulate the demand for 10 days

day	Random number	demand
1	14	1
2	74	4
3	24	2
4	87	5
5	7	1
6	45	3
7	26	2
8	66	4
9	26	2
10	94	5

Assume that the demand/day is given by:



If 10000 **random numbers** were drawn it would be expected that the number of observations per class would be:

Demand (x)	frequencies	observations
0	0.05	500
1	0.1	1000
2	0.2	2000
3	0.3	3000
4	0.2	2000
5	0.15	1500

Some Advantages of Simulation

- Often **simulation** is the **only type of model possible** for complex systems (e.g assess the impact of a certain silvicultural practices, fire)
- Process of building simulators can **clarify the understanding** of real systems and sometimes can be more useful than the implementation of the results itself
- Allows for sensitivity analysis and optimization of real system **without need to operate real system** (e.g no need to burn all stands to infer about fire behavior or its economic impacts)
- Can maintain **better control over experimental conditions** than real system

Some Disadvantages of Simulation

- May be very **expensive and time consuming** to build simulation
- **Easy to misuse simulation** by “stretching” it beyond the limits of credibility

  - when using commercial simulation packages due to ease of use and lack of familiarity with underlying assumptions and restrictions
  - Slick graphics, animation, tables, etc. may tempt user to assign unwarranted credibility to output

- Monte Carlo simulation usually **requires several (perhaps many) runs** at given input values, whereas analytical solutions provide exact values

1. What is Monte Carlo? Basic Principles
2. Random Numbers
3. Sample Sizes
4. Distributions
5. Monte Carlo Simulation Examples
6. Monte Carlo Simulation Exercises

Random numbers are used to obtain random observations from a probability distribution. **How are they generated?**

Manually – laborious and not practical

Random Tables – random physical process but difficult to implement in a computer

Computers – generate pseudo random numbers (pseudo because they're obtained with a deterministic mathematical process EXCEL ***RAND(), RANDBETWEEN(min, max)***

What is Monte Carlo? Basic Principles 1

Random Numbers 2

Sample Sizes 3

Distributions 4

Monte Carlo Simulation Examples 5

Monte Carlo Simulation Exercises 6

Sample Sizes

Determining sample size (*number of simulation runs*) is important and relies on 2 parameters: mean and standard deviation. **However**, the process is rather ciclic and based on **assumptions**:

1. **Population mean has a normal distribution**
2. **A large sample will be taken**
3. **Etc...**

$$n = \left[\frac{Z_{\alpha/2}\, \sigma}{E} \right]^2$$

Standard deviation of the population. If not known , we have to use an **estimate** (but the calculus requires sample size)

Normal deviate for a confidence interval of 1-alfa, but should be replaced by $t_{alfa/2,n-1}$ (also requires sample size)

Difference between sample mean and real mean (requires sample size to be known so that simulations can be run and the sample mean calculated)

Therefore, other options are commonly used:

Option 1

Choose a big sample size and then establish confidence intervals

Option 2

Compute the average value after each trial so it approaches a limit and stop simulating when the difference between population mean and sample mean reaches an acceptable value

Option 3

Make several pilot simulation runs to have an idea of the mean and standard deviation and then use it to calculate the sample size

1. What is Monte Carlo? Basic Principles
2. Random Numbers
3. Sample Sizes
4. Distributions
5. Monte Carlo Simulation Examples
6. Monte Carlo Simulation Exercises

Where do we obtain the probability distribution for one variable?

Empirical distributions

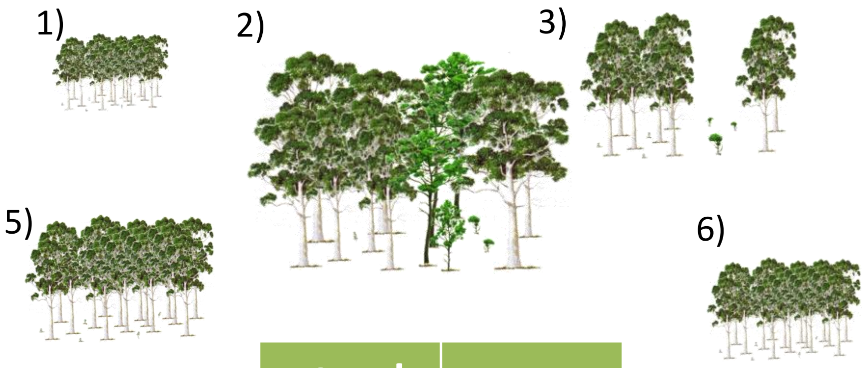
Theoretical distributions

(Poisson, Normal, Exponential, Weibull)

(Historic) Observed Data → Estimates

Empirical distributions

1) **Rely on a considerable amount of real (historic) data for a given variable**
   - Forest inventory data on stand age for eucalyptus



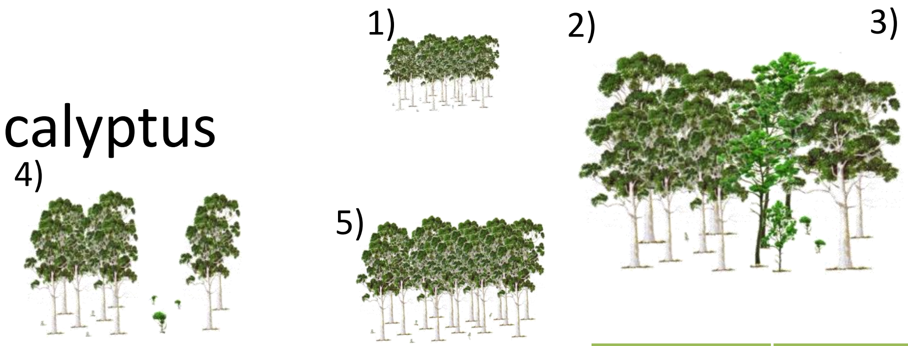
2) **Require classes to be defined**

3) **Counting observations/class**

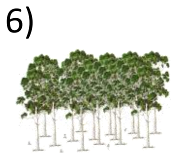
stand	age
1	*3*
2	*16*
3	*8*
4	*12*
5	*11*
6	*9*

Empirical distributions

1) **Rely on a considerable amount of real (historic) data for a given variable**
   - Forest inventory data on stand age for eucalyptus



2) **Require classes to be defined**
   - Age classes with amplitude 4 (e.g using IF clause in EXCEL)



stand	age
1	*3*
2	*16*
3	*8*
4	*12*
5	*11*
6	*9*

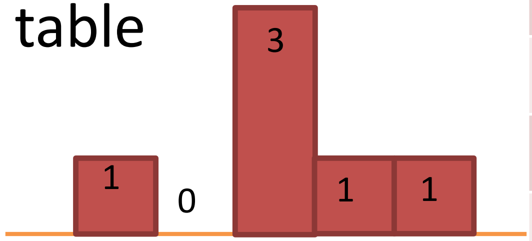
3) **Counting observations/class**

Empirical distributions

1) **Rely on a considerable amount of real (historic) data for a given variable**
   - Forest inventory data on stand age for eucalyptus

2) **Require classes to be defined**
   - Age classes with amplitude 4

3) **Counting observations/class**
   - Histogram or pivot table



class	freq
0 -3	1
4 - 7	0
8 - 11	3
12 - 15	1
>=16	1
	6

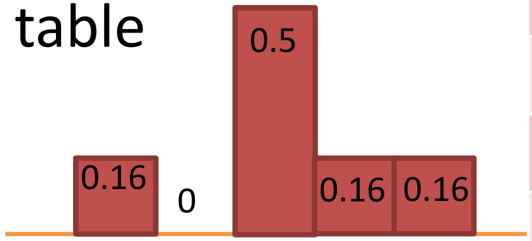
stand	age	class
1	*3*	*0 -3*
2	*16*	*>=16*
3	*8*	*8 - 11*
4	*12*	*12 - 15*
5	*11*	*8 - 11*
6	*9*	*8 - 11*

Empirical distributions

1) **Rely on a considerable amount of real (historic) data for a given variable**
   - Forest inventory data on stand age for eucalyptus

2) **Require classes to be defined**
   - Age classes with amplitude 4

3) **Counting observations/class**
   - Histogram or pivot table



class	freq	pdf
0 -3	1	(1/6) 0.16
4 - 7	0	0
8 - 11	3	(3/6) 0.5
12 - 15	1	(1/6) 0.16
>=16	1	(1/6) 0.16
	6	1

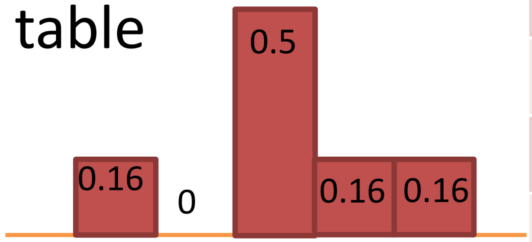
stand	age	class
1	3	0 -3
2	16	>=16
3	8	8 - 11
4	12	12 - 15
5	11	8 - 11
6	9	8 - 11

Empirical distributions

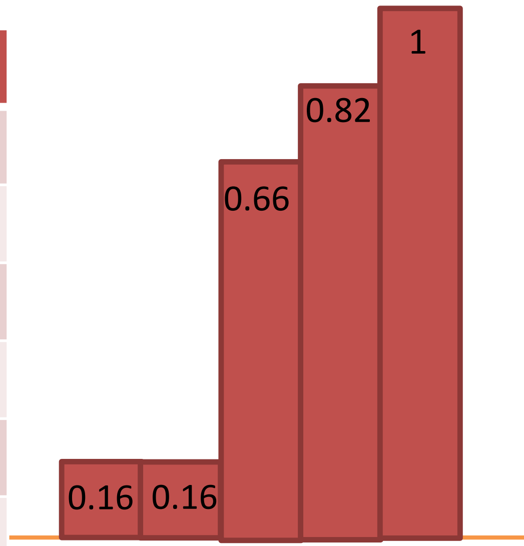
1) **Rely on a considerable amount of real (historic) data for a given variable**
   - Forest inventory data on stand age for eucalyptus

2) **Require classes to be defined**
   - Age classes with amplitude 4

3) **Counting observations/class**
   - Histogram or pivot table



class	freq	pdf	cpdf
0 -3	1	(1/6) 0.16	0.16
4 - 7	0	0	0.16
8 - 11	3	(3/6) 0.5	0.66
12 - 15	1	(1/6) 0.16	0.82
>=16	1	(1/6) 0.16	1
	6	1	



1. What is Monte Carlo? Basic Principles
2. Distributions
3. Random Numbers
4. Sample Sizes
5. Monte Carlo Simulation Examples
6. Monte Carlo Simulation Exercises

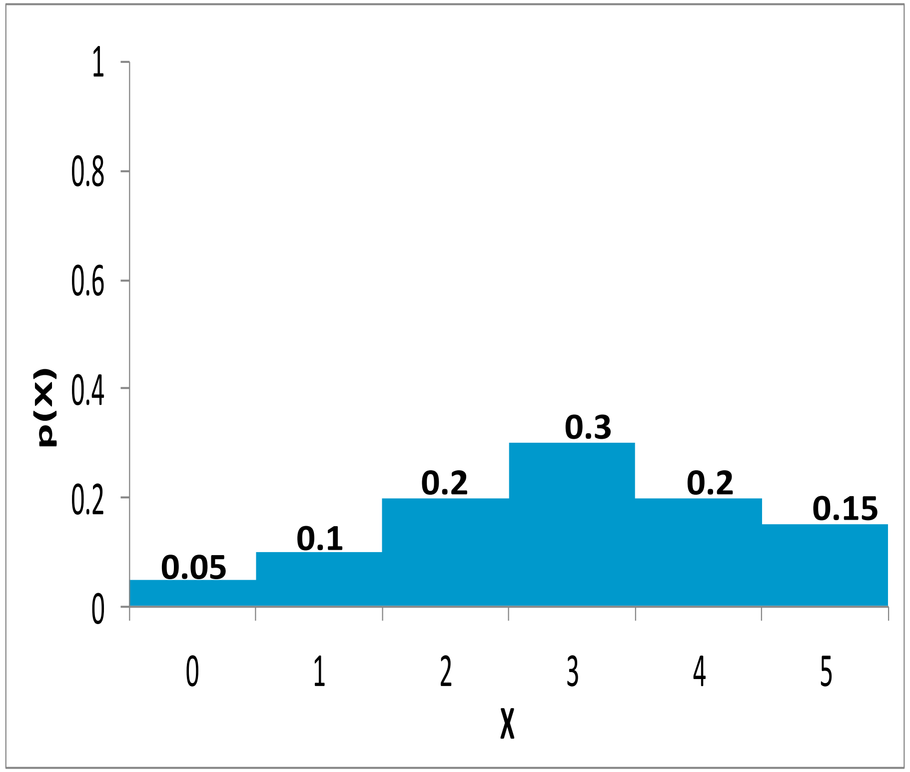
- In general, Monte Carlo Simulation is roughly composed of **five steps**:

1. Set up probability distribution that will be considered in the simulation
2. Build cumulative probability distribution
3. Establish an interval of random numbers for each variable
4. Generate random numbers
5. Simulate trials

- **Example 1** – Simulating with a distribution provided (empirical)
- **Example 2 –** Setting a distribution based on know distributions of other variables (independent variables)
- **Example 3 –** Simulating using dependent variables

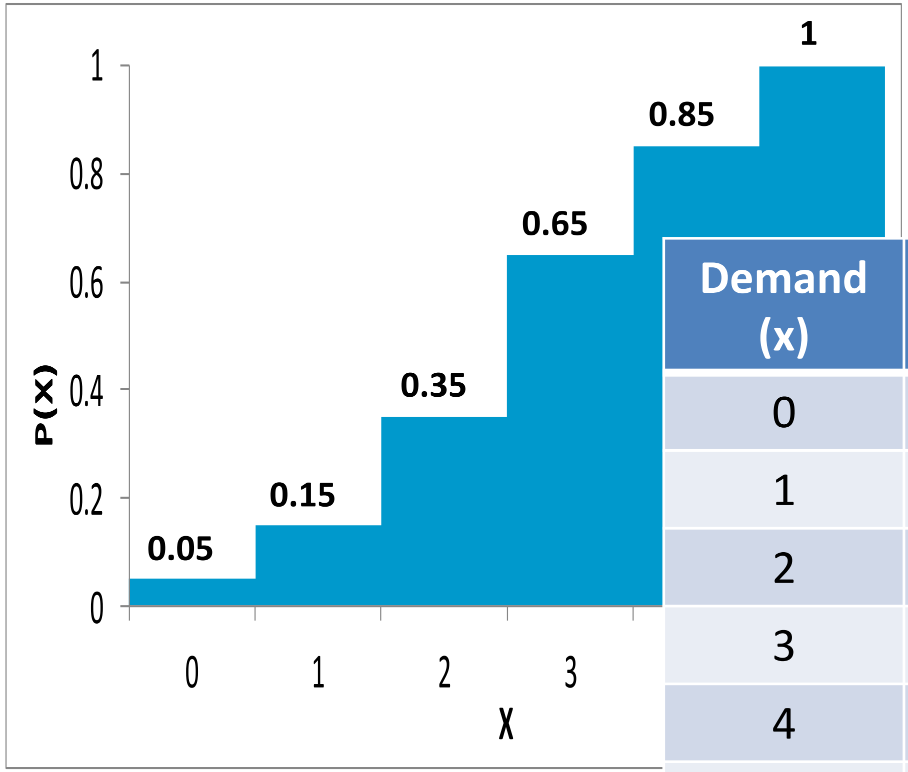
- **Example 1** - Simulate a 10 days demand based on the distribution below assuming the following random numbers: 14, 74, 24, 87, 7, 45, 26, 66, 26, 94
  (previously presented)

  1. Set up probability distributions that will be considered in the simulation



- **Example 1** - Simulate a 10 days demand based on the distribution below assuming the following random numbers: 14, 74, 24, 87, 7, 45, 26, 66, 26, 94 *(previously presented)*

1. Set up probability distributions that will be considered in the simulation
2. Build cumulative probability distributions



Demand (x)	Cumulative frequencies
0	0.05
1	0.15
2	0.35
3	0.65
4	0.85
5	1

- **Example 1** - Simulate a 10 days demand based on the distribution below assuming the following random numbers: 14, 74, 24, 87, 7, 45, 26, 66, 26, 94 *(previously presented)*

1. Set up probability distributions that will be considered in the simulation
2. Build cumulative probability distributions
3. Establish an interval of random numbers for each variable

Demand (x)	Cumulative frequencies	interval
0	0.05	0 – 4
1	0.15	5 – 14
2	0.35	15 – 34
3	0.65	35 – 64
4	0.85	65 – 84
5	1	85 – 99

- **Example 1** - Simulate a 10 days demand based on the distribution below assuming the following random numbers: 14, 74, 24, 87, 7, 45, 26, 66, 26, 94 *(previously presented)*

1. Set up probability distributions that will be considered in the simulation
2. Build cumulative probability distributions
3. Establish an interval of random numbers for each variable
4. Generate random numbers

day	Random number
1	14
2	74
3	24
4	87
5	7
6	45
7	26
8	66
9	26
10	94

- **Example 1** - Simulate a 10 days demand based on the distribution below assuming the following random numbers: 14, 74, 24, 87, 7, 45, 26, 66, 26, 94
  (previously presented)

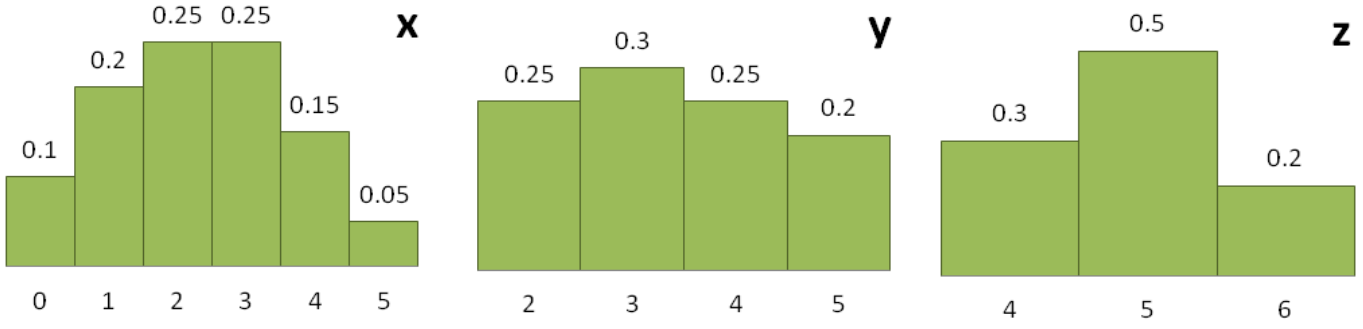
1. Set up probability distributions that will be considered in the simulation
2. Build cumulative probability distributions
3. Establish an interval of random numbers for each variable
4. Generate random numbers
5. Simulate trials

Demand	interval
0	0 – 4
1	5 – 14
2	15 – 34
3	35 – 64
4	65 – 84
5	85 – 1

day	Random number	demand
1	14	1
2	74	4
3	24	2
4	87	5
5	7	1
6	45	3
7	26	2
8	66	4
9	26	2
10	94	5

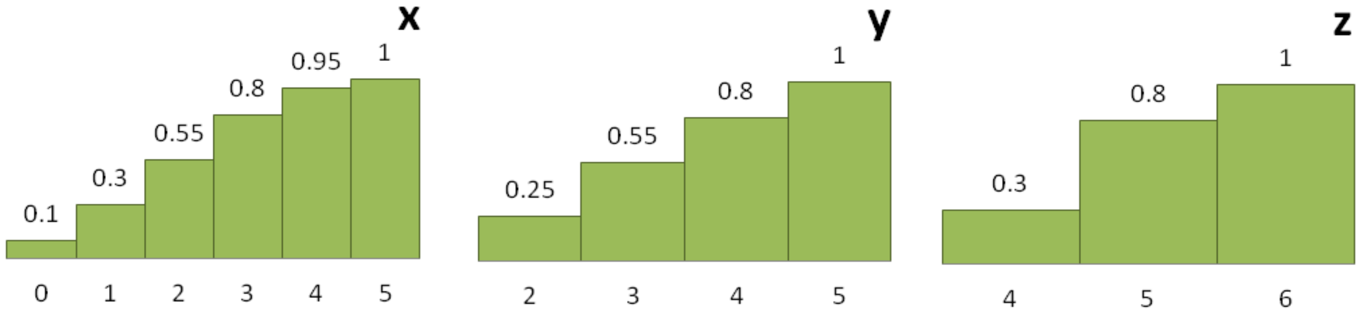
- **Example 2** - Assume the effectiveness function for a system is W = 5x + 2y + z, where the variables x, y and z are **independent** and described by the probabilities below. Run 18 trials

  1. Set up probability distributions that will be considered in the simulation



- **Example 2** - Assume the effectiveness function for a system is W = 5x + 2y + z, where the variables x, y and z are **independent** and described by the probabilities below. Run 18 trials

1. Set up probability distributions that will be considered in the simulation
2. Build cumulative probability distributions



- **Example 2** - Assume the effectiveness function for a system is W = 5x + 2y + z, where the variables x, y and z are **independent** and described by the probabilities below. Run 18 trials

1. Set up probability distributions that will be considered in the simulation
2. Build cumulative probability distributions
3. Establish an interval of random numbers for each variable

x	distribution	cumulative distribution	aux	lower lim	upper lim	interval
0	0.1	0.1	10	0	9	0 - 9
1	0.2	0.3	30	10	29	10 - 29
2	0.25	0.55	55	30	54	30 - 54
3	0.25	0.8	80	55	79	55 - 79
4	0.15	0.95	95	80	94	80 - 94
5	0.05	1	100	95	99	95 - 99
y	distribution	cumulative distribution	aux	lower lim	upper lim	interval
2	0.25	0.25	25	0	24	0 - 24
3	0.3	0.55	55	25	54	25 - 54
4	0.25	0.8	80	55	79	55 - 79
5	0.2	1	100	80	99	80 - 99
z	distribution	cumulative distribution	aux	lower lim	upper lim	interval
4	0.3	0.3	30	0	29	0 - 29
5	0.5	0.8	80	30	79	30 - 79
6	0.2	1	100	80	99	80 - 99

- **Example 2** - Assume the effectiveness function for a system is W = 5x + 2y + z, where the variables x, y and z are **independent** and described by the probabilities below. Run 18 trials

1. Set up probability distributions that will be considered in the simulation
2. Build cumulative probability distributions
3. Establish an interval of random numbers for each variable
4. Generate random numbers

trial nr	rand_x	rand_y	rand_z
1	43	22	1
2	74	9	8
3	84	10	82
4	42	38	65
5	83	16	34
6	25	1	27
7	21	67	62
8	25	38	58
9	83	65	42
10	76	25	32
11	74	27	63
12	68	73	55
13	3	7	96
14	60	53	29
15	35	34	31
16	56	25	17
17	71	83	83
18	15	72	49

- **Example 2** - Assume the effectiveness function for a system is W = 5x + 2y + z, where the variables x, y and z are **independent** and described by the probabilities below. Run 18 trials

1. Set up probability distributions that will be considered in the simulation
2. Build cumulative probability distributions
3. Establish an interval of random numbers for each variable
4. Generate random numbers
5. Simulate trials

trial nr	rand_x	rand_y	rand_z	x	y	z	w
1	43	22	1	2	2	4	18
2	74	9	8	3	2	4	23
3	84	10	82	4	2	6	30
4	42	38	65	2	3	5	21
5	83	16	34	4	2	5	29
6	25	1	27	1	2	4	13
7	21	67	62	1	4	5	18
8	25	38	58	1	3	5	16
9	83	65	42	4	4	5	33
10	76	25	32	3	3	5	26
11	74	27	63	3	3	5	26
12	68	73	55	3	4	5	28
13	3	7	96	0	2	6	10
14	60	53	29	3	3	4	25
15	35	34	31	2	3	5	21
16	56	25	17	3	3	4	25
17	71	83	83	3	5	6	31
18	15	72	49	1	4	5	18

- **Example 2** - Assume the effectiveness function for a system is W = 5x + 2y + z, where the variables x, y and z are **independent** and described by the probabilities below. Run 18 trials

1. Set up probability distributions that will be considered in the simulation
2. Build cumulative probability distributions
3. Establish an interval of random numbers for each variable
4. Generate random numbers
5. Simulate trials
6. Set probability distribution for w

trial nr	w
1	18
2	23
3	30
4	21
5	29
6	13
7	18
8	16
9	33
10	26
11	26
12	28
13	10
14	25
15	21
16	25
17	31
18	18

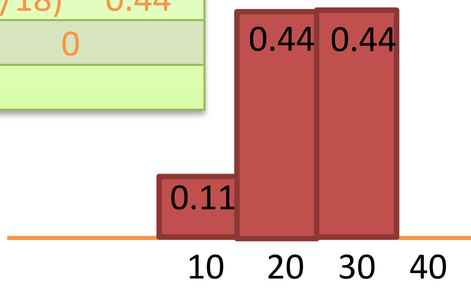
- **Example 2** - Assume the effectiveness function for a system is W = 5x + 2y + z, where the variables x, y and z are **independent** and described by the probabilities below. Run 18 trials

1. Set up probability distributions that will be considered in the simulation
2. Build cumulative probability distributions
3. Establish an interval of random numbers for each variable
4. Generate random numbers
5. Simulate trials
6. Set probability distribution for w

trial nr	w	class
1	18	15 – 24 (20)
2	23	15 – 24 (20)
3	30	25 – 34 (30)
4	21	15 – 24 (20)
5	29	15 – 24 (20)
6	13	5 – 14 (10)
7	18	15 – 24 (20)
8	16	15 – 24 (20)
9	33	25 – 34 (30)
10	26	25 – 34 (30)
11	26	25 – 34 (30)
12	28	25 – 34 (30)
13	10	5 – 14 (10)
14	25	25 – 34 (30)
15	21	15 – 24 (20)
16	25	25 – 34 (30)
17	31	25 – 34 (30)
18	18	15 – 24 (20)

7. Define classes for w with amplitude of 10
8. Count observations per class with pivot table

class	freq	prob
5 – 14 (10)	2	(2/18) 0.11
15 – 24 (20)	8	(8/18) 0.44
25 – 34 (30)	8	(8/18) 0.44
35 – 44 (40)	0	0
	18	



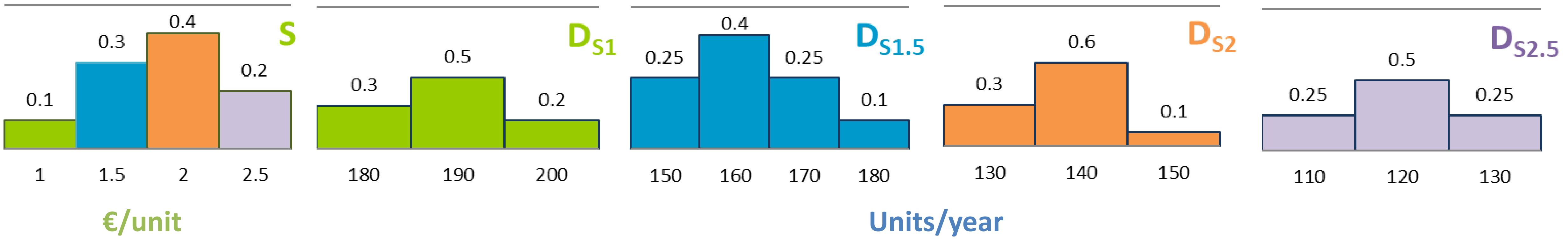
- **Example 3**

The gross income/year = selling price/unit (S) x sales/year(D).

In general, as the selling price decreases, sales increase (**dependent variables**). Calculate the gross income assuming:

- the distributions bellow
- the random number for sales (D) = 73 and
- the random number for selling price (S) = 22

1. Set up probability distributions
2. Build cumulative probability distributions



- **Example 3** – The gross income/year = sales/year (D) x selling price/unit (S). In general, as the selling price decreases, sales increase (**dependent variables**). Calculate the gross income assuming:
  - the distributions bellow
  - the random number for sales (D) = **73**
  - the random number for selling price (S) =**22**

3. Establish an interval and simulate

S	distribution	cumulative distribution	interval
1	0.1	0.1	0 - 9
1.5	0.3	0.4	10 - 39
2	0.4	0.8	40 - 79
2.5	0.2	1	80 - 99

D1	distribution	cumulative distribution	interval
180	0.3	0.3	0 - 29
190	0.5	0.8	30 - 79
200	0.2	1	80 - 99
D1.5	**distribution**	**cumulative distribution**	**interval**
150	0.25	0.25	0 – 24
160	0.4	0.65	25 - 64
170	0.25	0.9	65 - 89
180	0.1	1	90 - 99
D2	**distribution**	**cumulative distribution**	**interval**
130	0.3	0.3	0 - 29
140	0.6	0.9	30 - 89
150	0.1	1	90 - 99
D2.5	**distribution**	**cumulative distribution**	**interval**
110	0.25	0.25	0 - 24
120	0.5	0.75	25 - 74
130	0.25	1	75 - 99

gross income/year = 170 x 1.5 = **255 €**

- **Example 4** – Assign site index (S) values to the NFI plots missing that information

Even-aged stands
S = f (hdom, t, 10)

279 NFI plots:

139 plots with S
137 plots without S

Uneven-aged stands
S = f (?, ??, ???)

69 NFI plots:

Without S

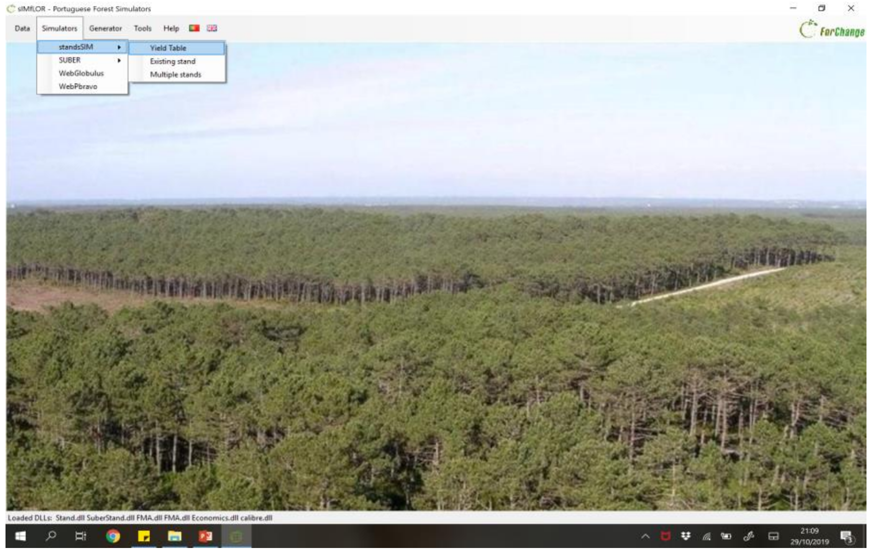
Suppose you need to prepare inputs to run some eucalyptus simulations using StandsSIM.md simulator. This tool requires information about site index (S), but S estimates are not available for all NFI plots.

The data in spreadsheet Ex_4 shows that only 139 of the 348 plots have been assigned an S value.

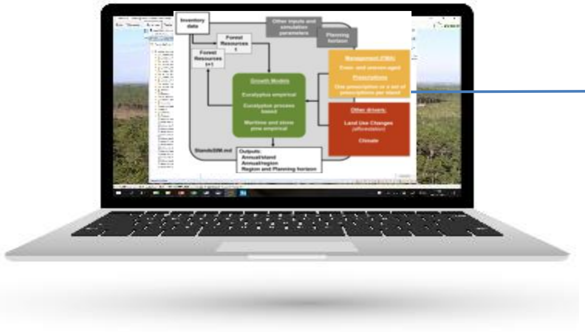
Use plots with S to build the distribution of NFI plots by S class and using Monte Carlo simulation assign S values to the remaining plots taking into consideration that **S values lower than 8 and greater than 26 are to be disregarded**. Consider S classes with range=1

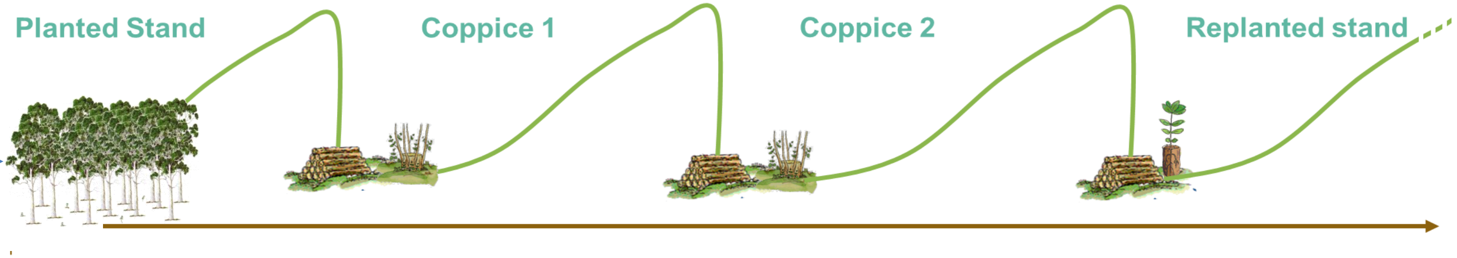
- **Example 4** – Assign site index (S) values to the NFI plots missing that information

Even-aged stands
S = f (hdom, t, 10)

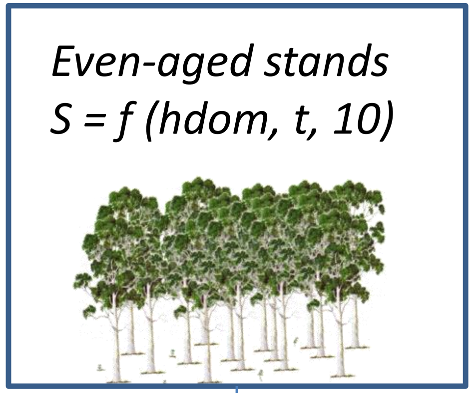


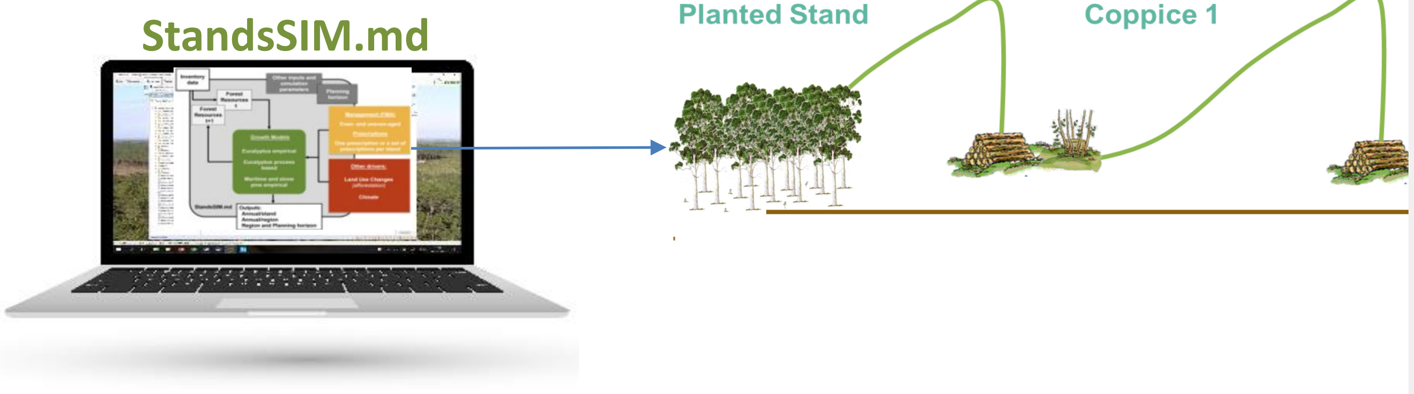
StandsSIM.md

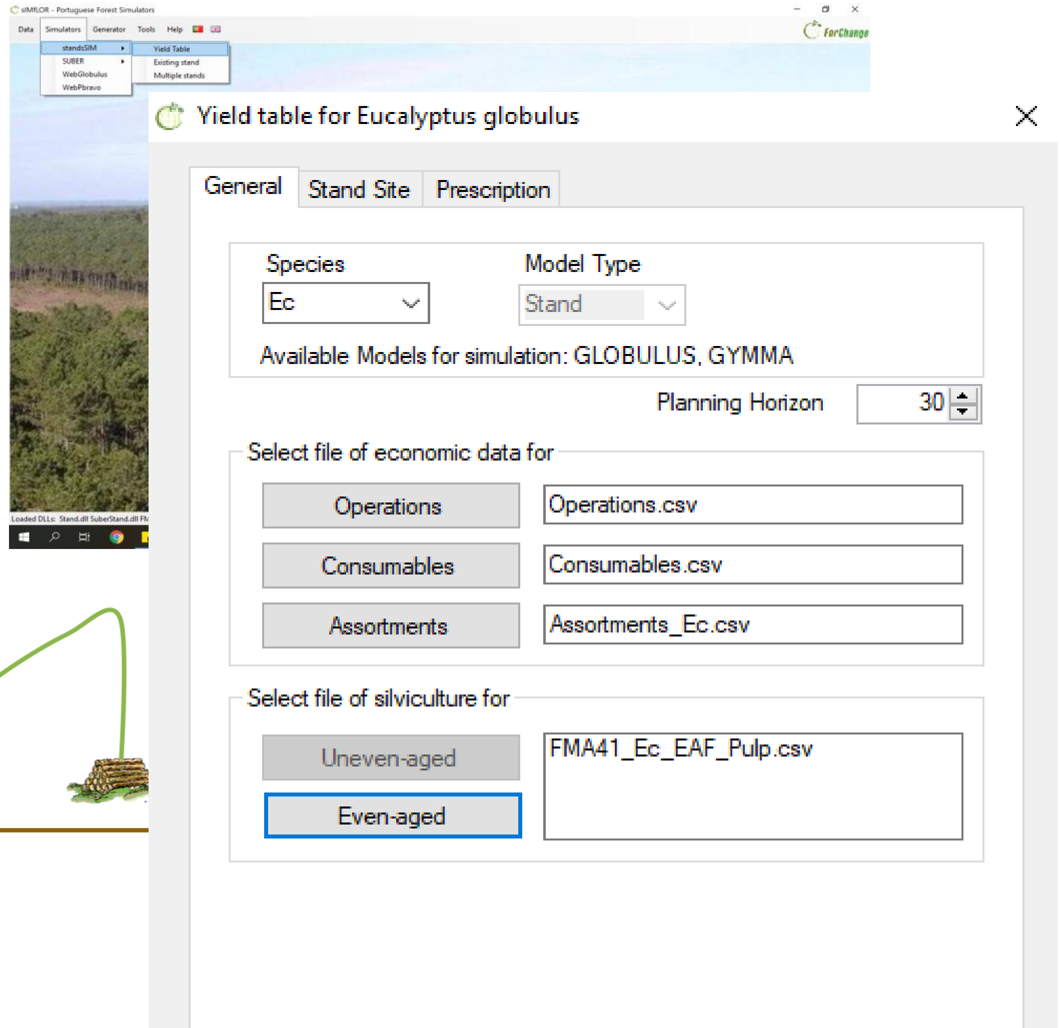




- **Example 4** – Assign site index (S) values to the NFI plots missing that information



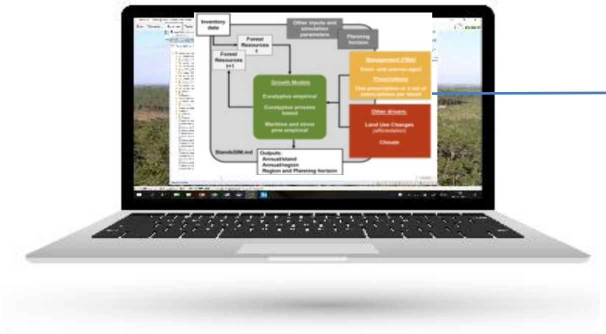


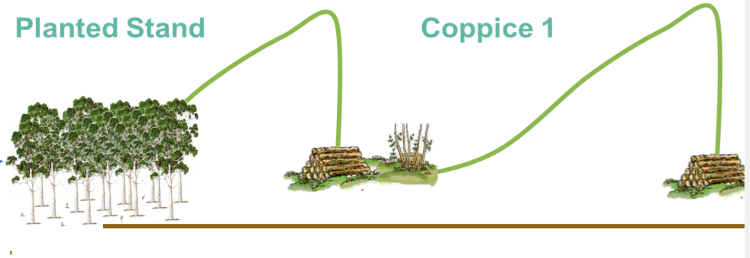


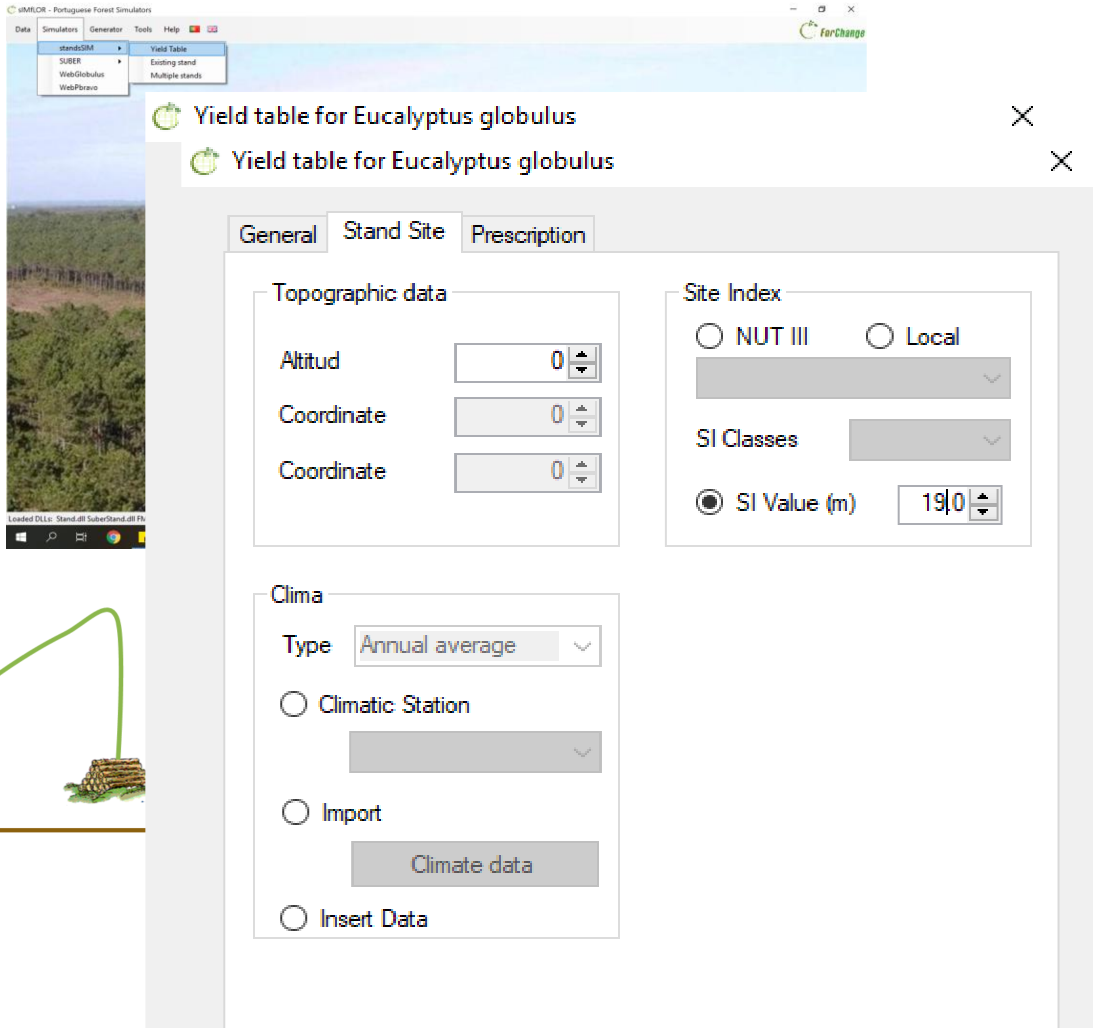
- **Example 4** – Assign site index (S) values to the NFI plots missing that information

Even-aged stands
S = f (hdom, t, 10)

StandsSIM.md







- **Example 4** – Assign site index (S) values to the NFI plots missing that information

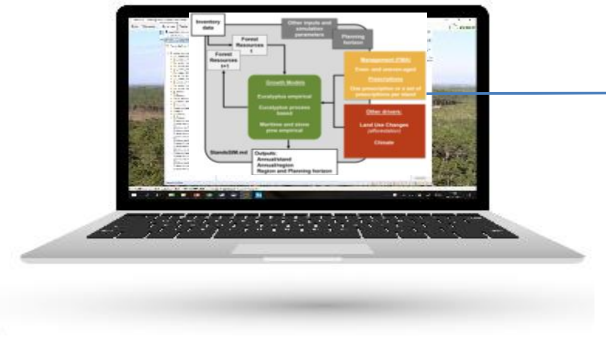
Even-aged stands
S = f (hdom, t, 10)

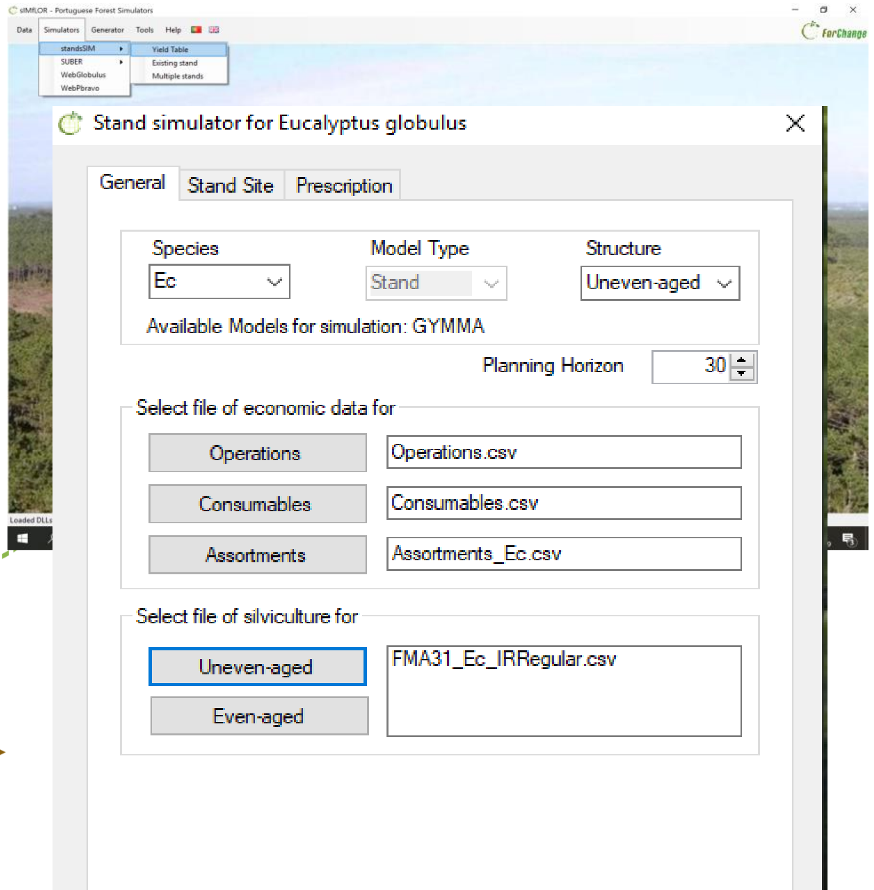
Uneven-aged stands
S = f (?, ??, ???)

CONVERT: Uneven- to Even-aged

Hdom = ?

StandsSIM.md





- **Example 4** – Assign site index (S) values to the NFI plots missing that information

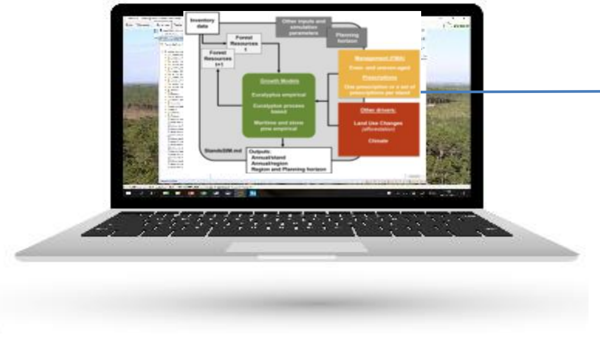
Even-aged stands
S = f (hdom, t, 10)

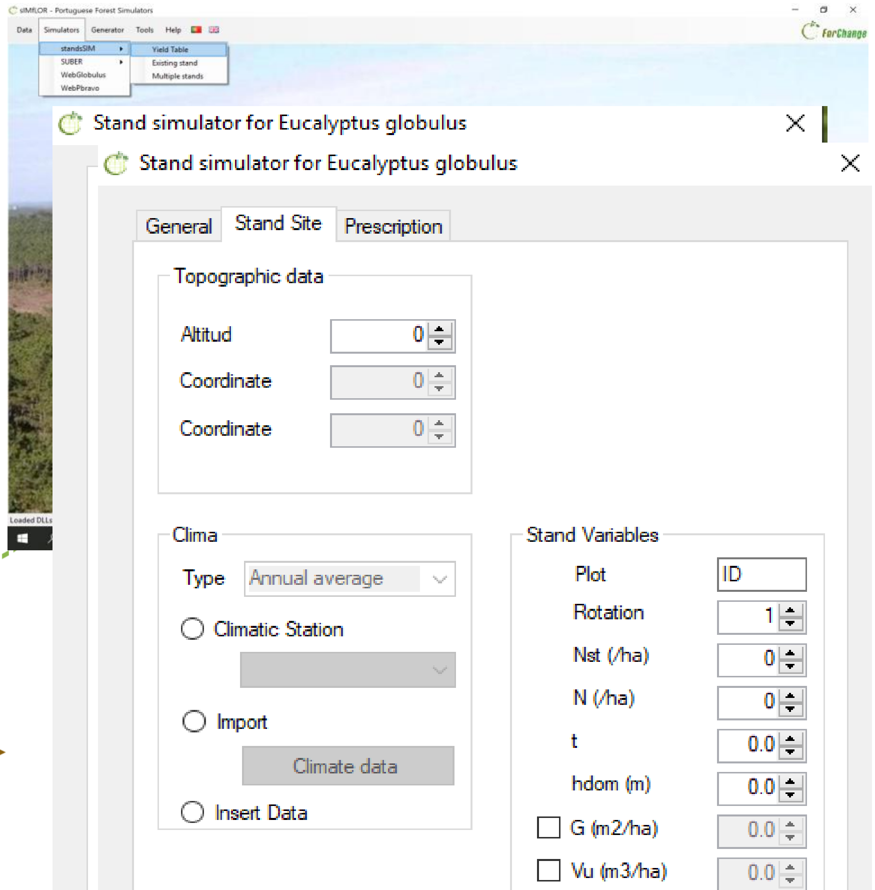
Uneven-aged stands
S = f (?, ??, ???)

CONVERT: *Uneven- to Even-aged*

Hdom = ?

StandsSIM.md





- **Example 4** – Assign site index (S) values to the NFI plots missing that information

Even-aged stands
S = f (hdom, t, 10)

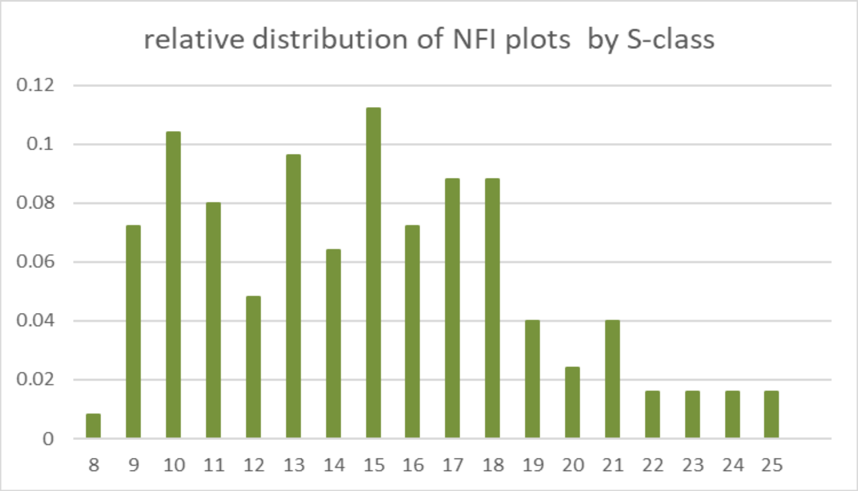
279 NFI plots:

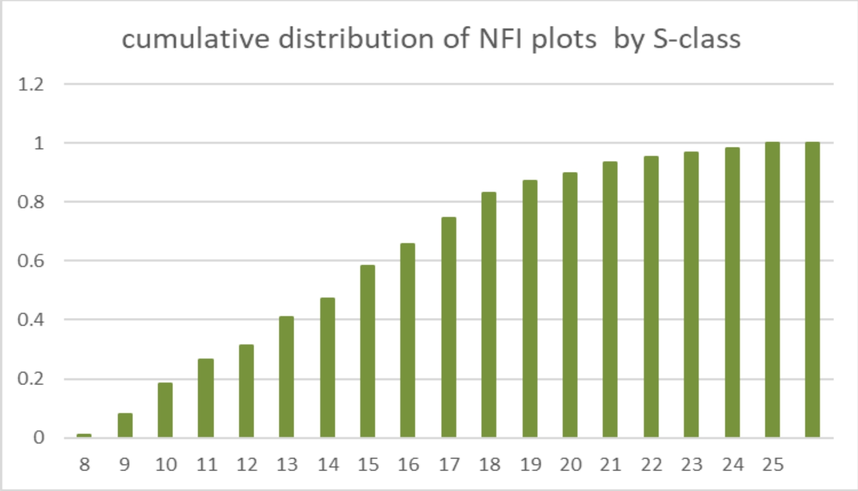
139 plots with S
137 plots without S
Because:
tree heights were not measured or stand age not recorded (e.g. recently harvested stands)

Uneven-aged stands
S = f (?, ??, ???)

69 NFI plots:

Without S





What is Monte Carlo? Basic Principles 1

Random Numbers 2

Sample Sizes 3

Distributions 4

Monte Carlo Simulation Examples 5

Monte Carlo Simulation Exercises 6

Where do we obtain the probability distribution for one variable?

Empirical distributions

Theoretical distributions

(Poisson, Normal, Exponential, Weibull)

(Historic) Observed Data

Estimates

Theoretical distributions – Probability distributions:

Normal → often used in natural and social sciences to represent real-valued random variables whose distributions are not known

Poisson → Number of events that occur in an interval of time

Exponential → Time taken between 2 events occurring

Weibull → Commonly used for generating diameter distributions in forest plots

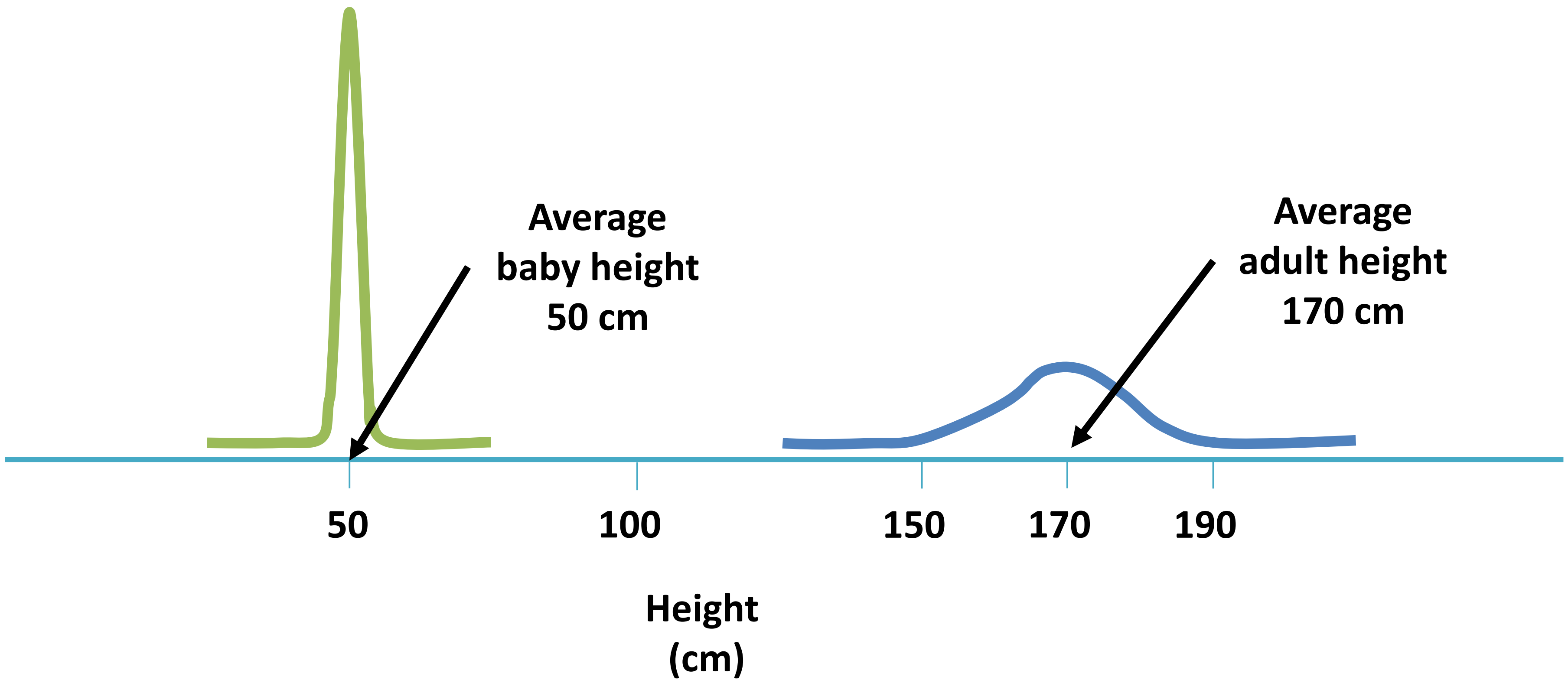
Theoretical distributions – Probability distributions:

Normal

often used in natural and social sciences to represent real-valued random variables whose distributions are not known

2 Normal distributions of the heights of male humans: **at birth** and **as adults**

Normal distributions are **always centered in the average value**

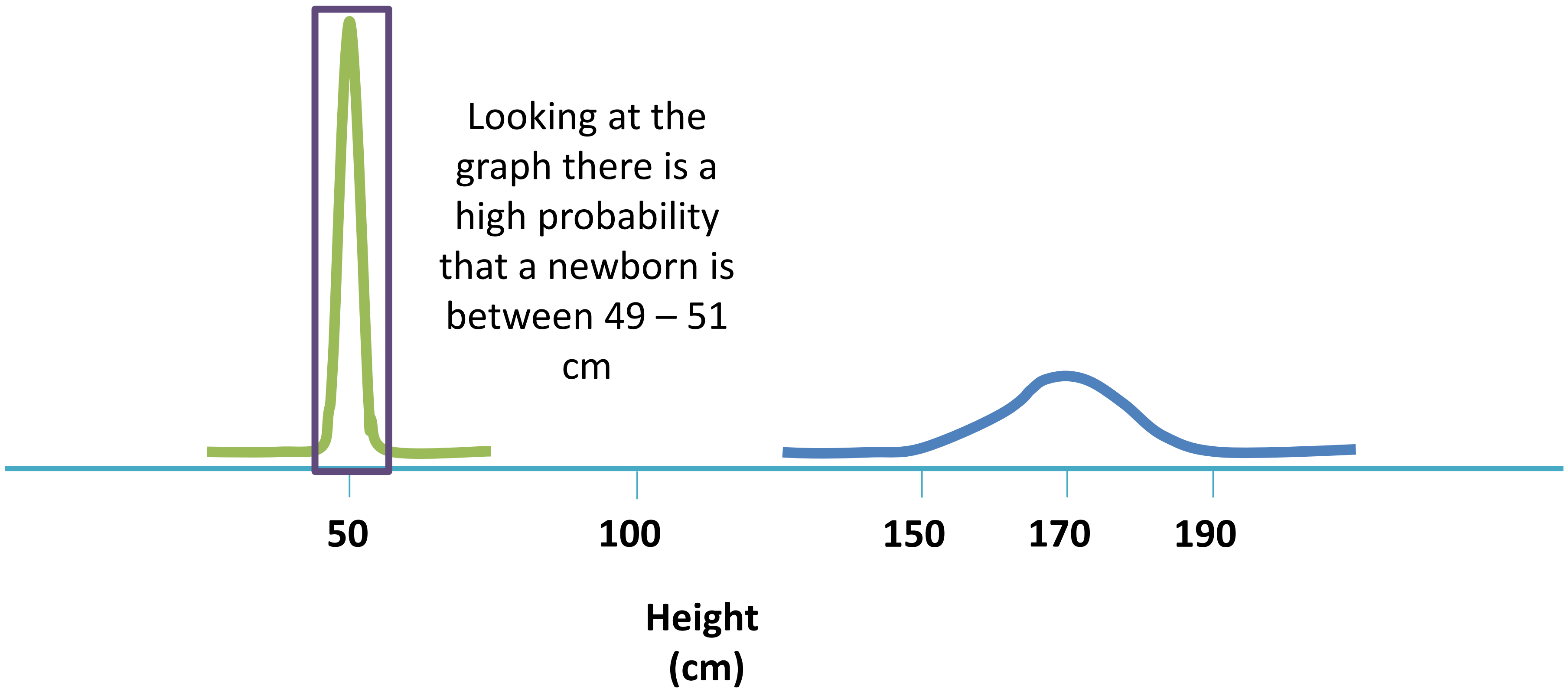


Theoretical distributions – Probability distributions:

Normal

often used in natural and social sciences to represent real-valued random variables whose distributions are not known

2 Normal distributions of the heights of male humans: **at birth** and **as adults**



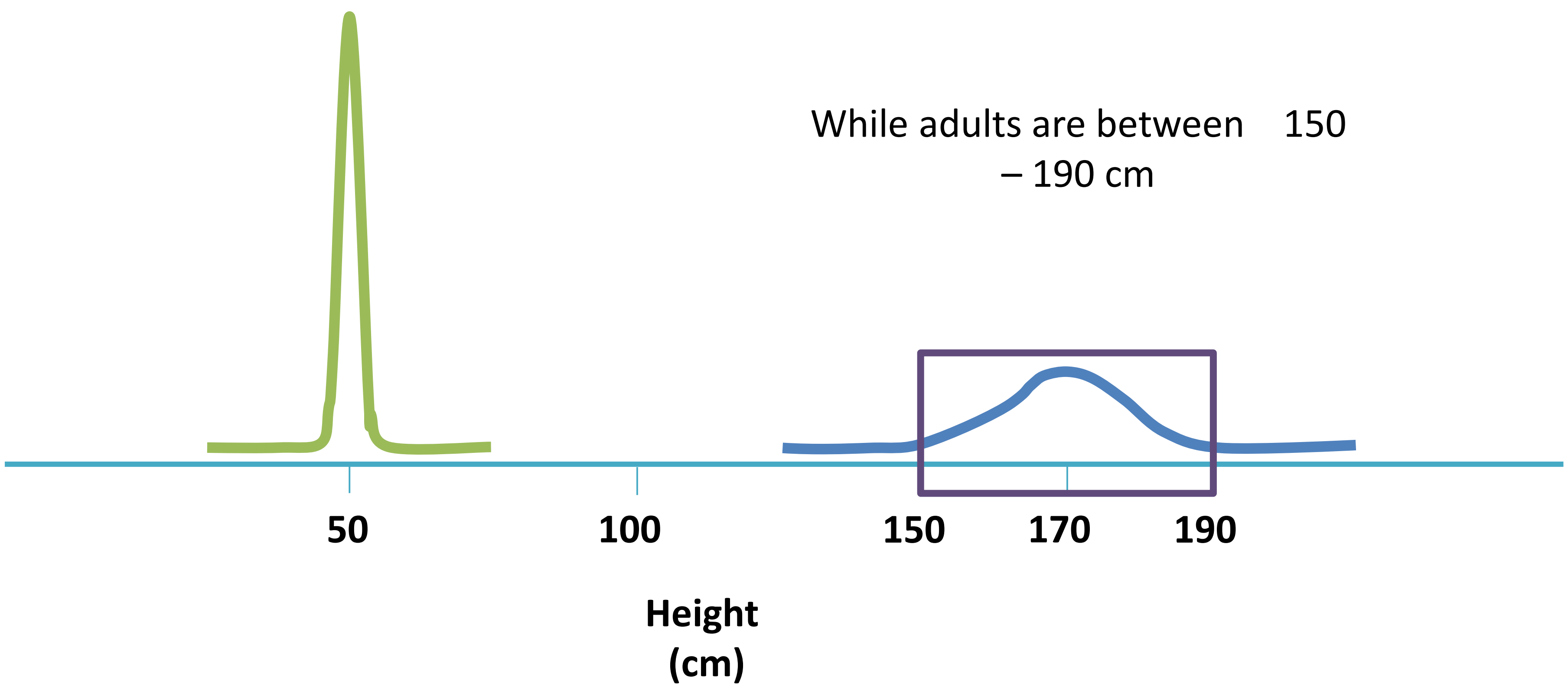
Theoretical distributions – Probability distributions:

Normal

↓

often used in natural and social sciences to represent real-valued random variables whose distributions are not known

2 Normal distributions of the heights of male humans: **at birth** and **as adults**

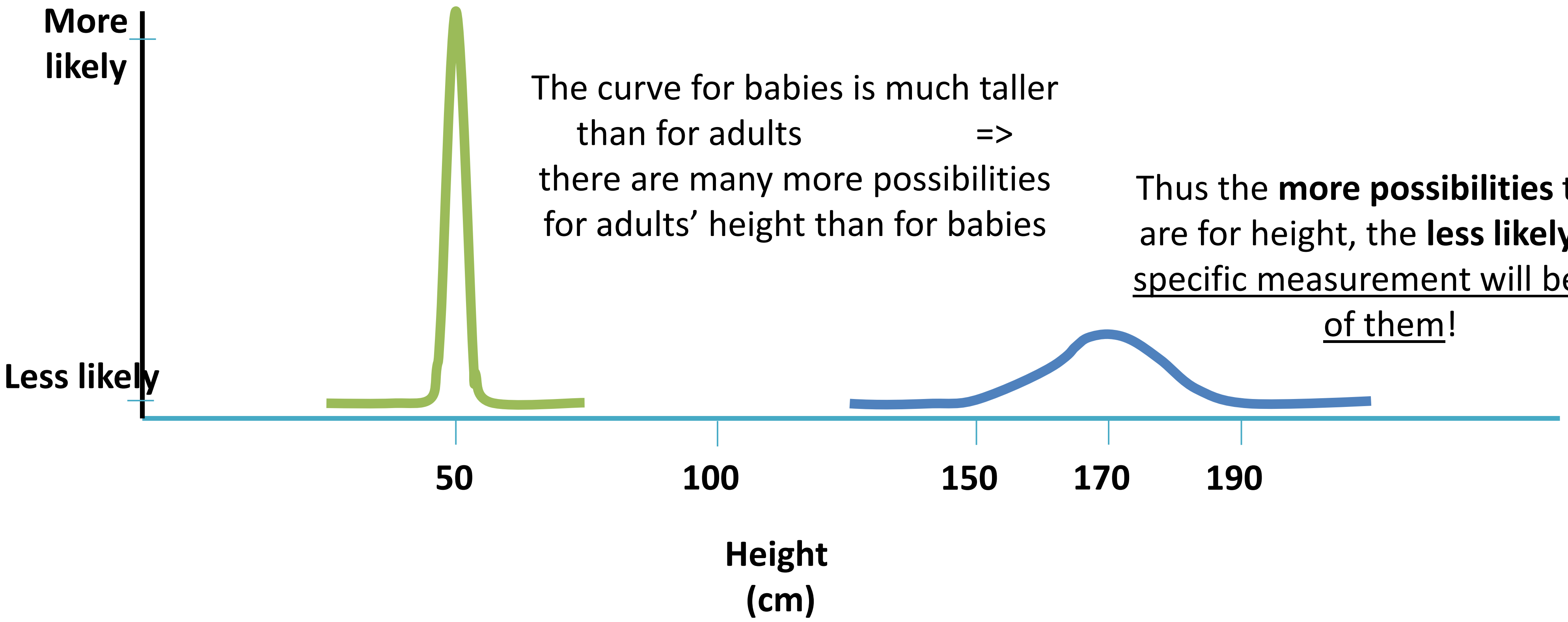


Theoretical distributions – Probability distributions:

Normal

often used in natural and social sciences to represent real-valued random variables whose distributions are not known

2 Normal distributions of the heights of male humans: **at birth** and **as adults**



The curve for babies is much taller than for adults => there are many more possibilities for adults' height than for babies

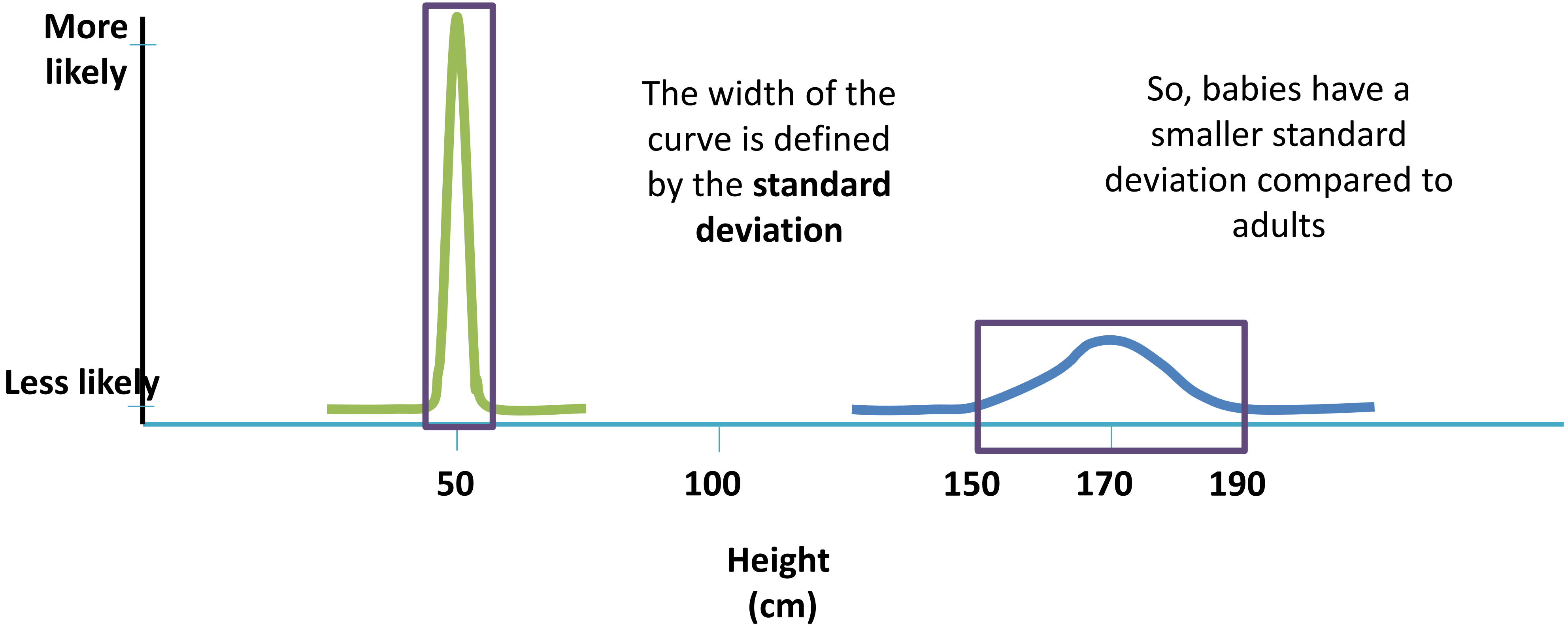
Thus the **more possibilities** there are for height, the **less likely** one <u>specific measurement will be one of them</u>!

Theoretical distributions – Probability distributions:

Normal

often used in natural and social sciences to represent real-valued random variables whose distributions are not known

2 Normal distributions of the heights of male humans: at birth and as adults



The width of the curve is defined by the **standard deviation**

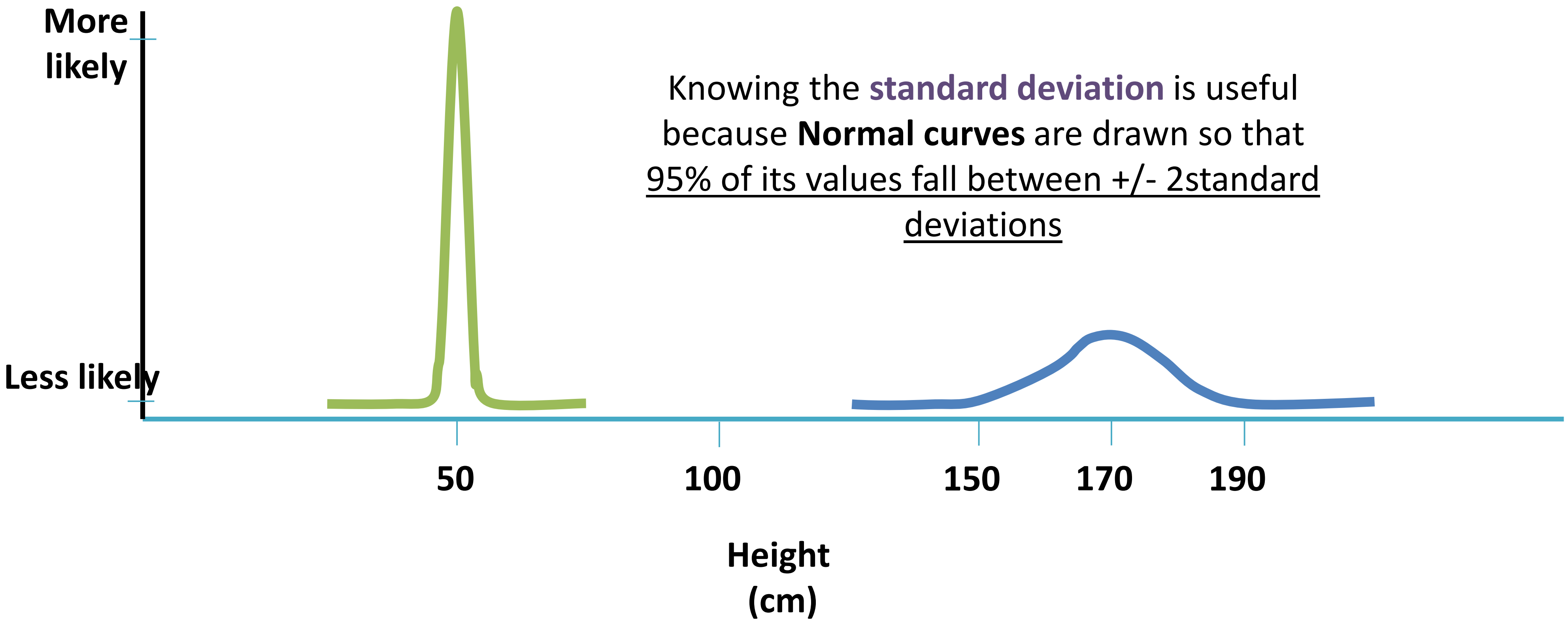
So, babies have a smaller standard deviation compared to adults

Theoretical distributions – Probability distributions:

Normal

often used in natural and social sciences to represent real-valued random variables whose distributions are not known

2 Normal distributions of the heights of male humans: **at birth** and **as adults**



Knowing the **standard deviation** is useful because **Normal curves** are drawn so that 95% of its values fall between +/- 2standard deviations

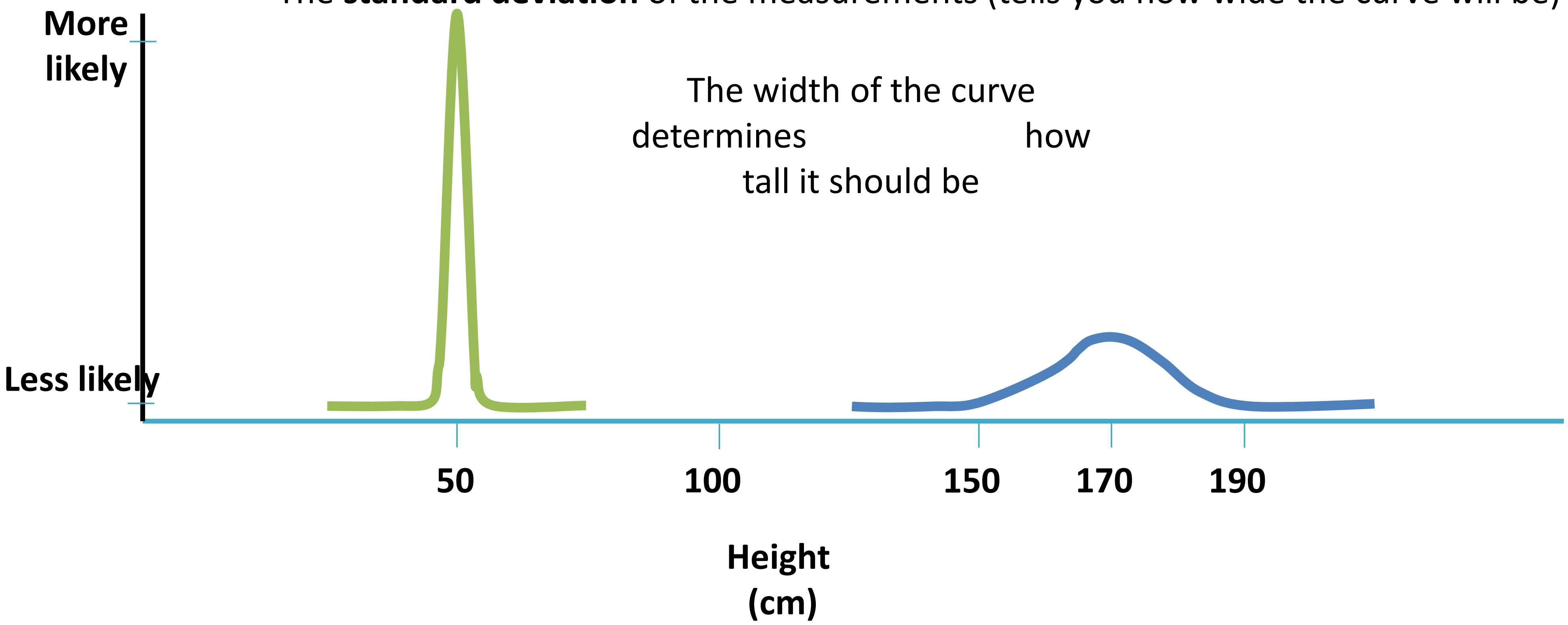
Theoretical distributions – Probability distributions:

Normal

often used in natural and social sciences to represent real-valued random variables whose distributions are not known

To draw a Normal distribution you need to know:

The **average** measurement (tells you where the curve is centered)
The **standard deviation** of the measurements (tells you how wide the curve will be)



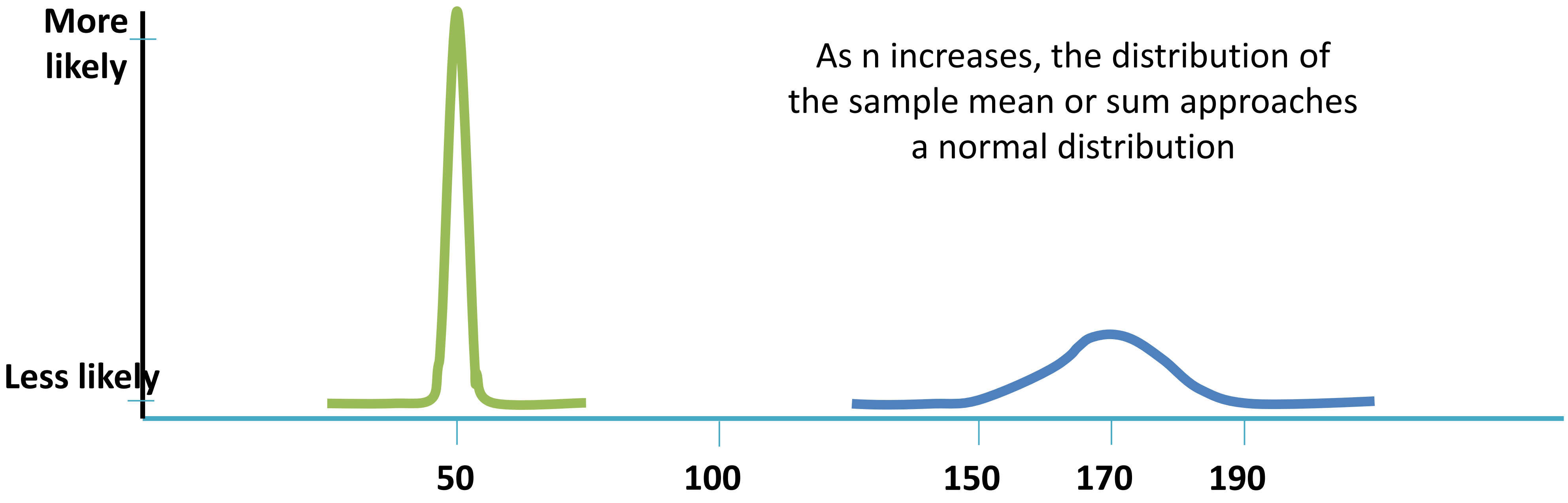
Theoretical distributions – Probability distributions:

Normal

↓

often used in natural and social sciences to represent real-valued random variables whose distributions are not known

Are found a lot in Nature and there is a reason for that that make it quite useful:
CENTRAL LIMIT THEOREM



As n increases, the distribution of the sample mean or sum approaches a normal distribution

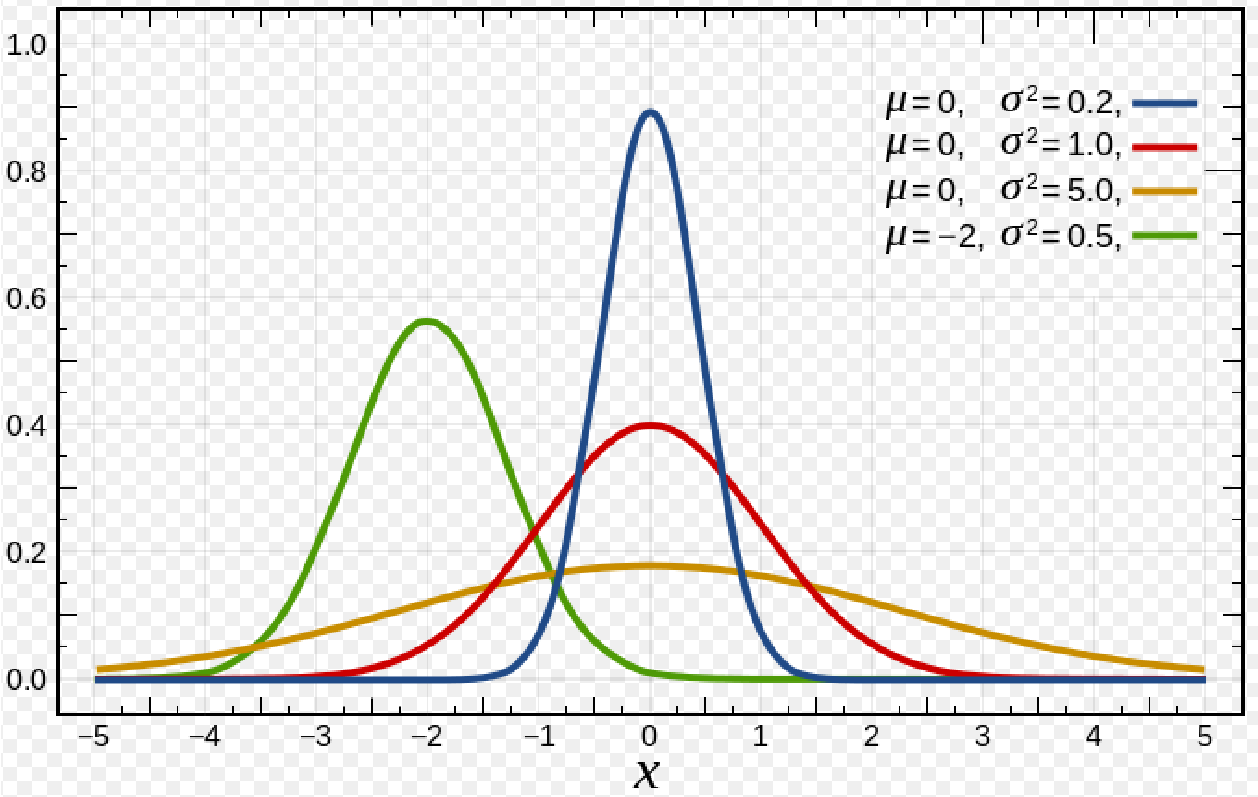
Many other variables!

Theoretical distributions – Probability distributions

Normal - real-valued random variables whose distributions are not known

μ —The mean σ —The standard deviation $X \approx N\ (\mu, \sigma)$

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



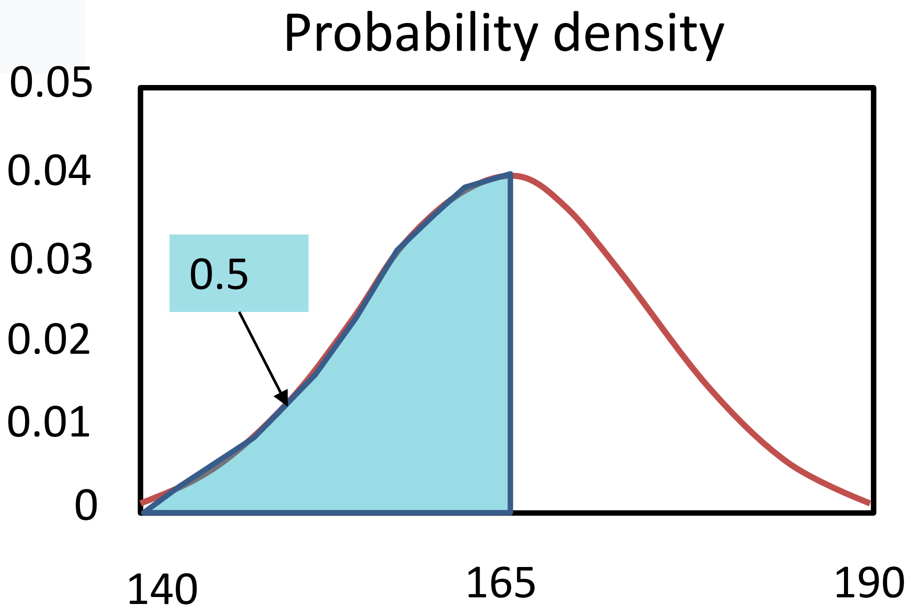
Theoretical distributions – Probability distributions

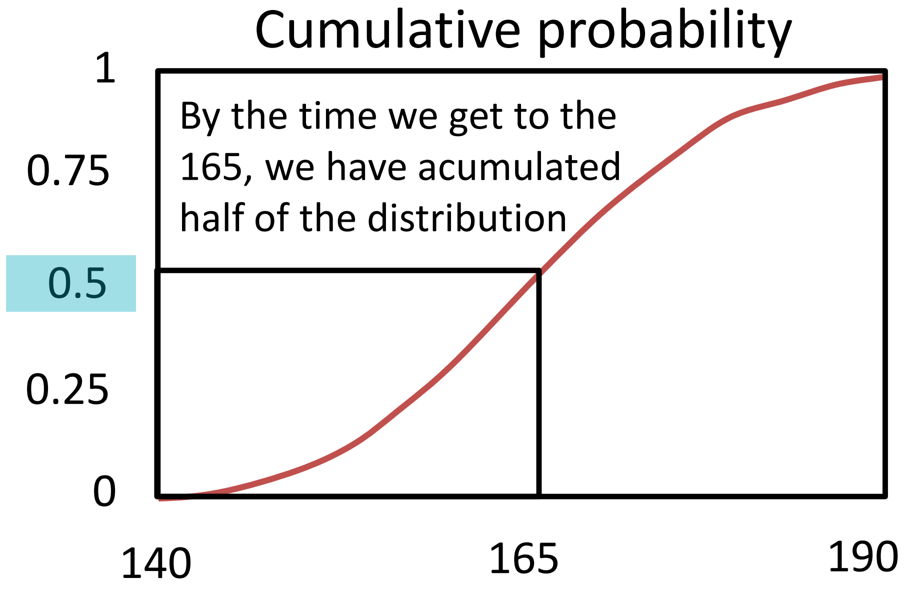
Normal - real-valued random variables whose distributions are not known

μ —The mean σ —The standard deviation $X \approx N(\mu, \sigma)$

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

Probabilities are **AREAS**, so if I take the mean value (165), the proportion to the left of the mean will be 50%





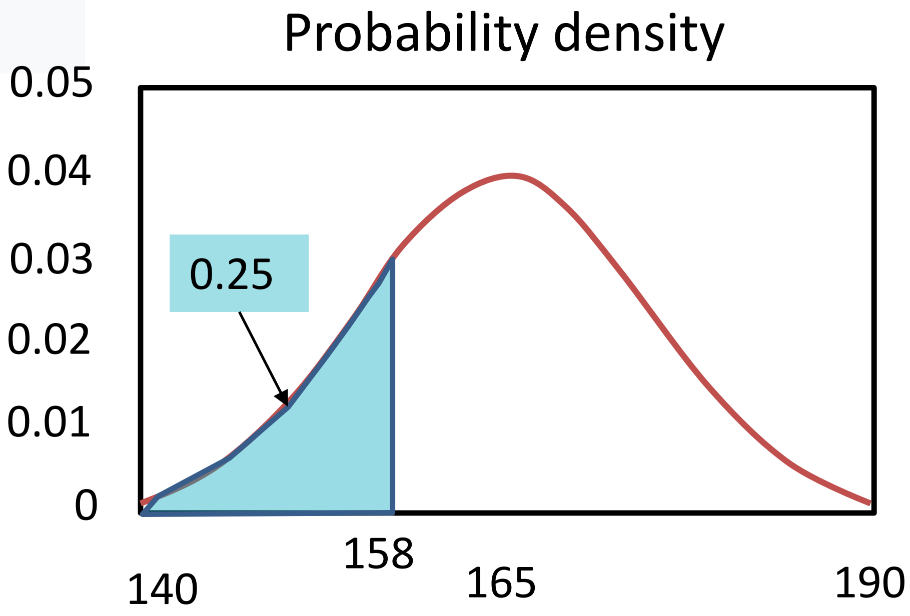
Theoretical distributions – Probability distributions

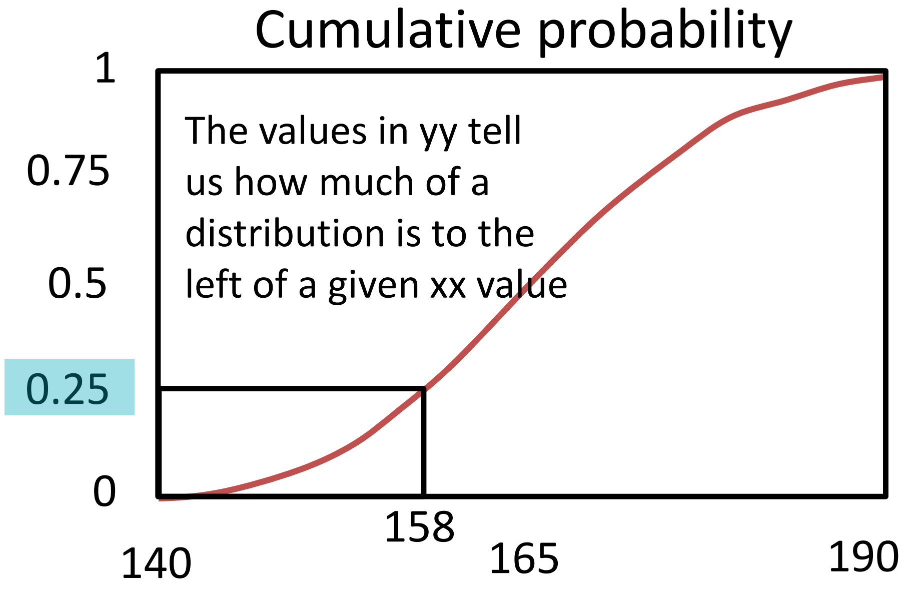
Normal - real-valued random variables whose distributions are not known

μ —The mean σ —The standard deviation $X \approx N(\mu, \sigma)$

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

so if we had chosen a values left to 165, let's say the one that represents 25% of the region (assume it is 158)





Theoretical distributions – Probability distributions

Normal - real-valued random variables whose distributions are not known

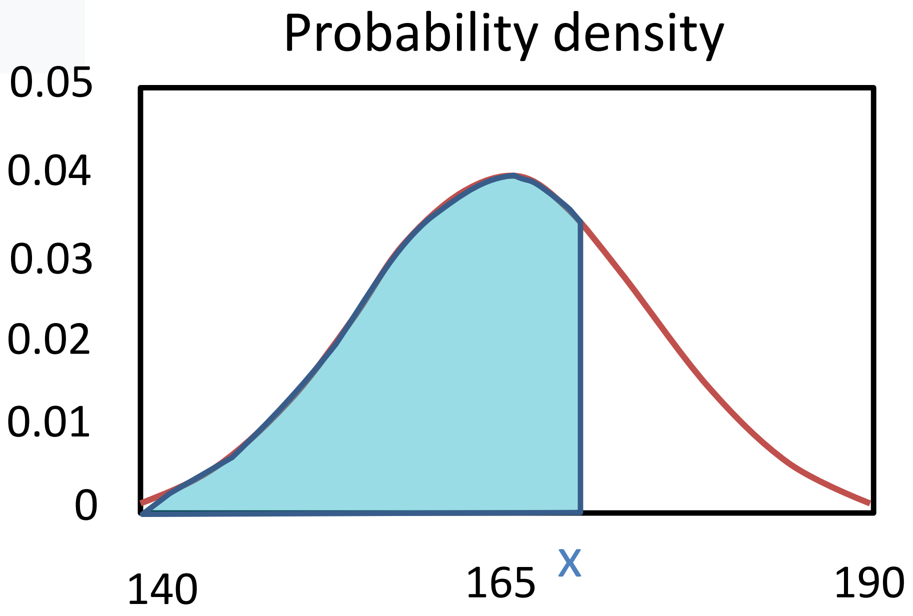
μ —The mean σ —The standard deviation $X \approx N(\mu, \sigma)$

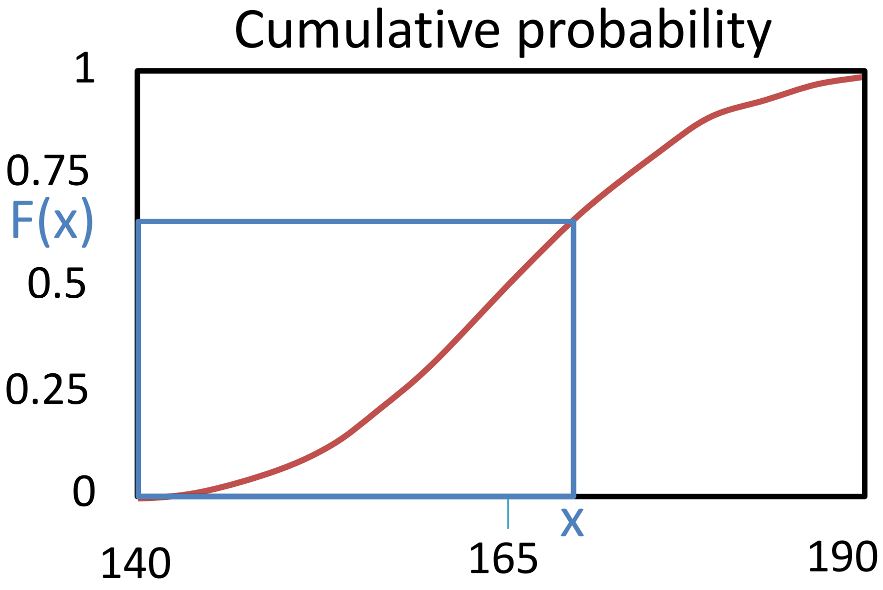
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

If probabilities are **AREAS** to move from the probability density to the cumulative probability, we need to…

$$\int_{-\infty}^{x} f(x)\,dx = F(x)$$

Tedious and time consuming





Theoretical distributions – Probability distributions

Normal - real-valued random variables whose distributions are not known

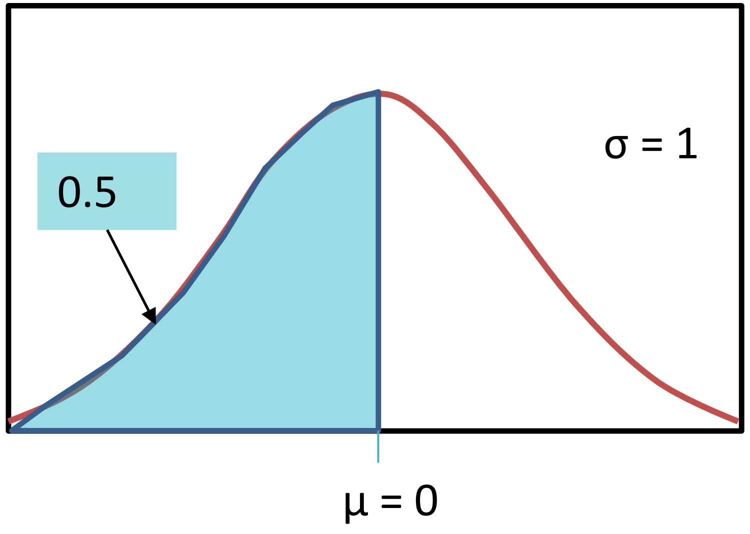
μ —The mean σ —The standard deviation $X \approx N\,(\mu, \sigma)$

Standard Normal distribution $X \approx N\,(0,1)$

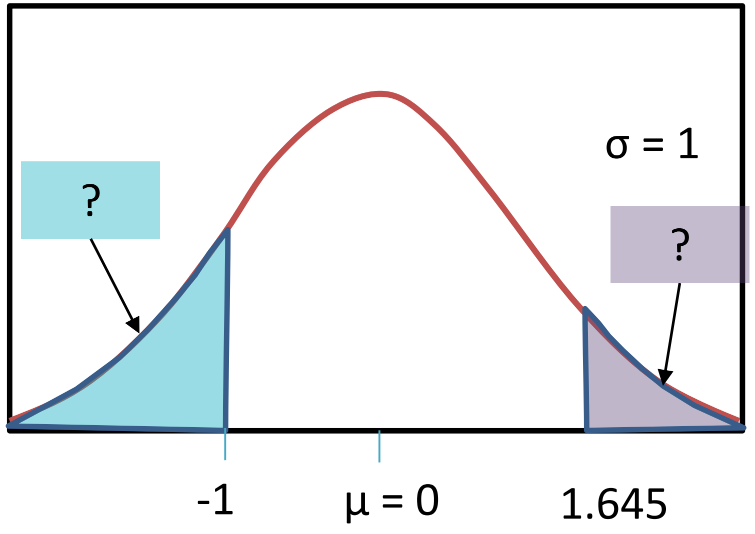
There is a 50% probability that a given selection from this distribution being less than zero

But what's the probability of being less than -1? or greater than 1.645?

Probability density



Probability density



Theoretical distributions – Probability distributions

Normal - real-valued random variables whose distributions are not known

μ —The mean σ —The standard deviation $X \approx N(\mu, \sigma)$

Standard Normal distribution $X \approx N(0,1)$

has been tabulated

Standard Normal Probabilities

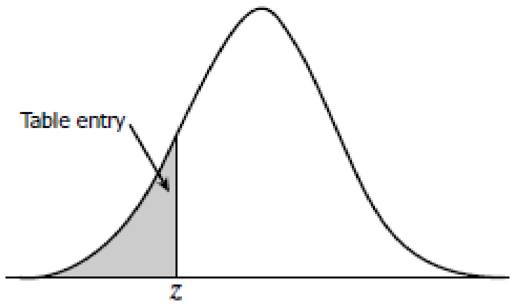
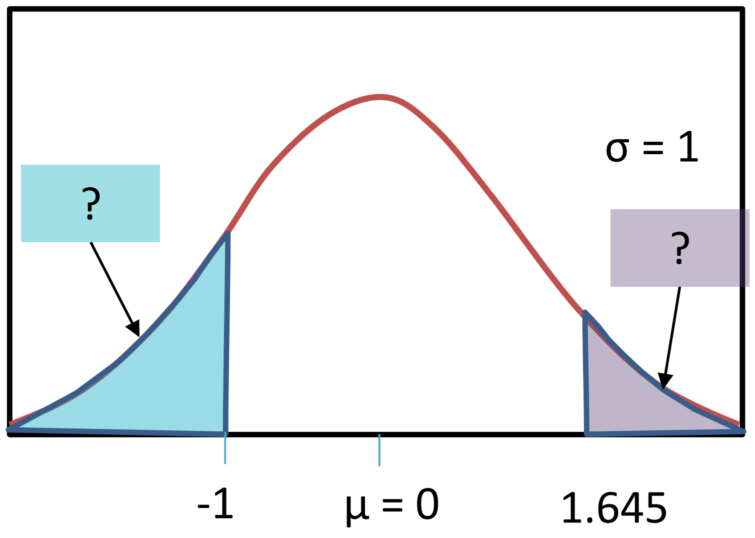


Table entry for z is the area under the standard normal curve to the left of z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019

Probability density



Standard Normal Probabilities

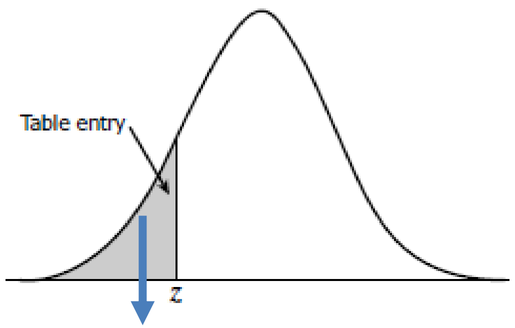


Table entry for z is the area under the standard normal curve to the left of z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867

tions

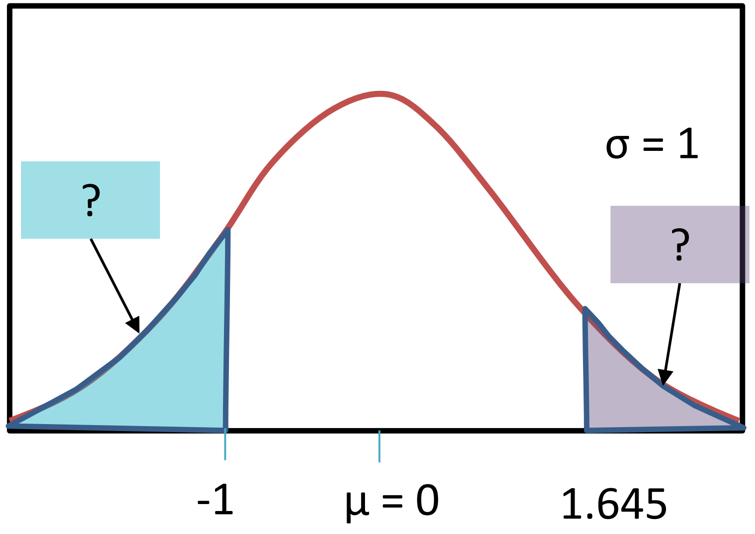
se distributions are not known

eviation

$$X \approx N(\mu, \sigma)$$

$$X \approx N(0,1)$$

Probability density



Standard Normal Probabilities

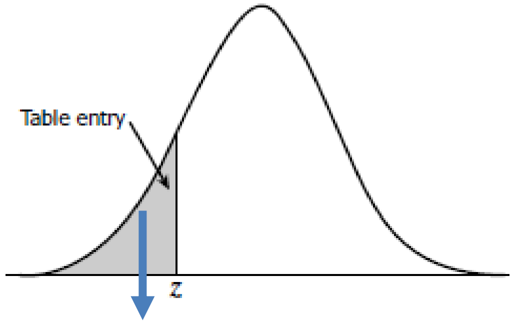


Table entry for z is the area under the standard normal curve to the left of z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867

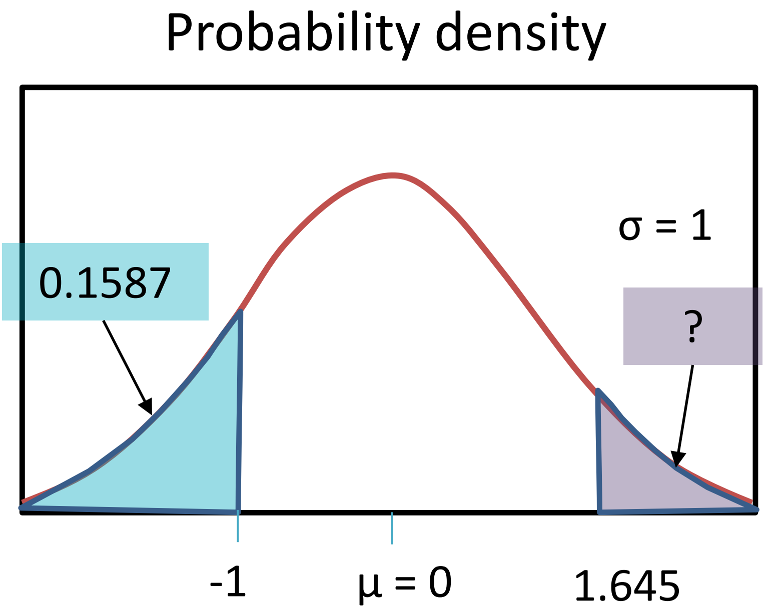
tions

se distributions are not known

eviation

$$X \approx N\,(\mu,\sigma)$$

$$X \approx N\,(0,1)$$



Standard Normal Probabilities

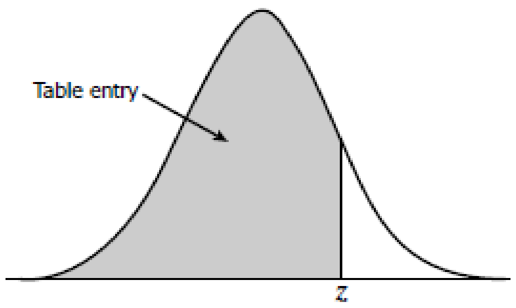


Table entry for z is the area under the standard normal curve to the left of z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964

itions

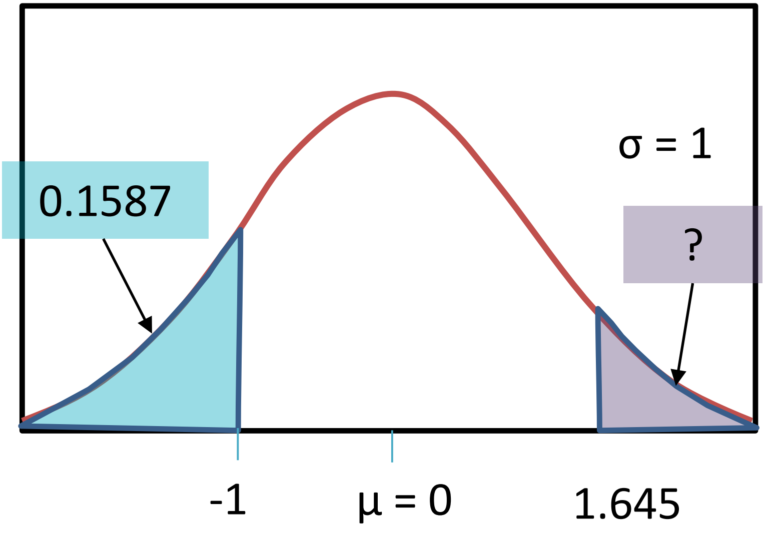
se distributions are not known

deviation

$$X \approx N\ (\mu, \sigma)$$

$$X \approx N\ (0,1)$$

Probability density



Standard Normal Probabilities

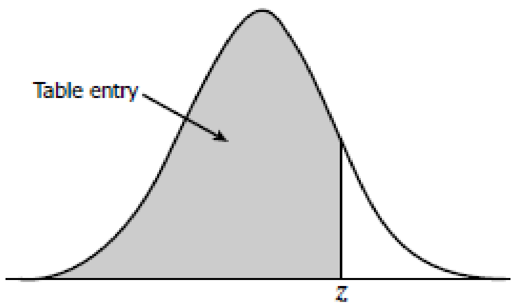


Table entry for z is the area under the standard normal curve to the left of z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964

ıtions

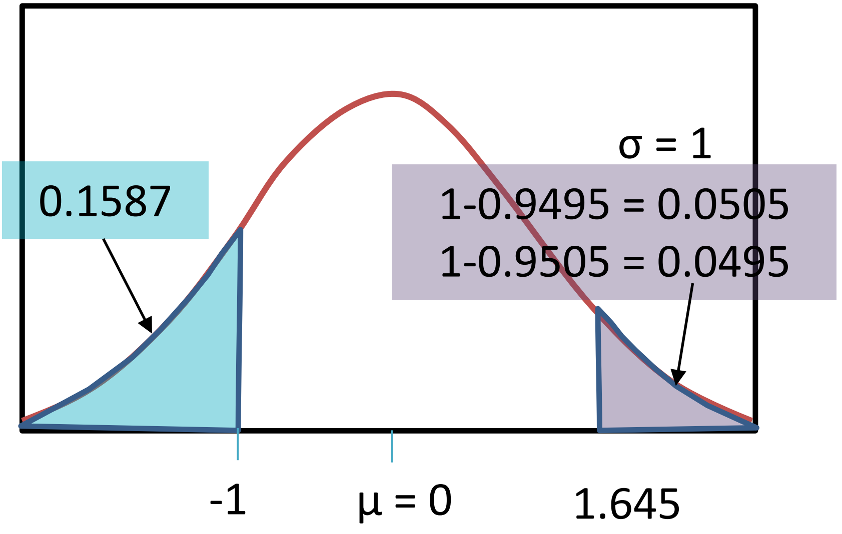
ɔse distributions are not known

ɔdeviation

$$X \approx N\ (\mu, \sigma)$$

$$X \approx N\ (0,1)$$

Probability density



Standard Normal Probabilities

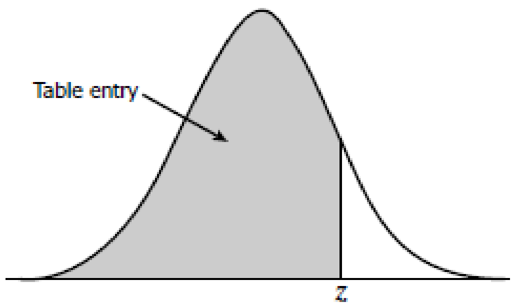


Table entry for z is the area under the standard normal curve to the left of z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964

...tions

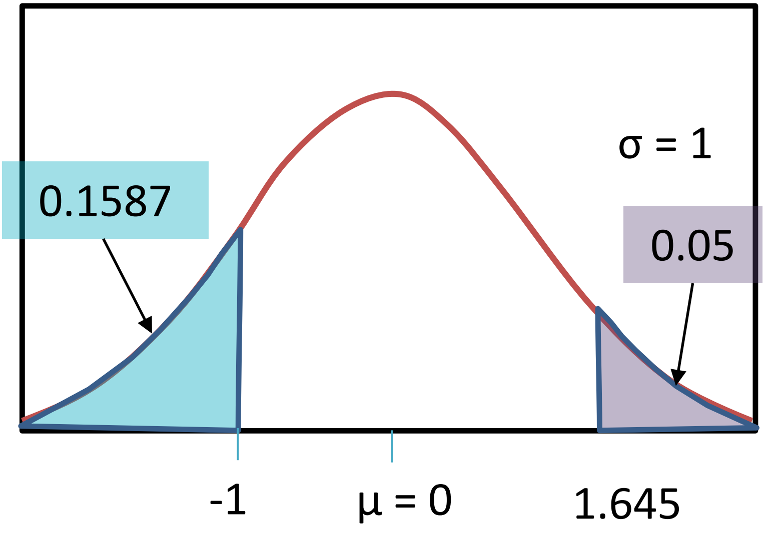
...se distributions are not known

...deviation

$$X \approx N(\mu, \sigma)$$

$$X \approx N(0,1)$$

Probability density

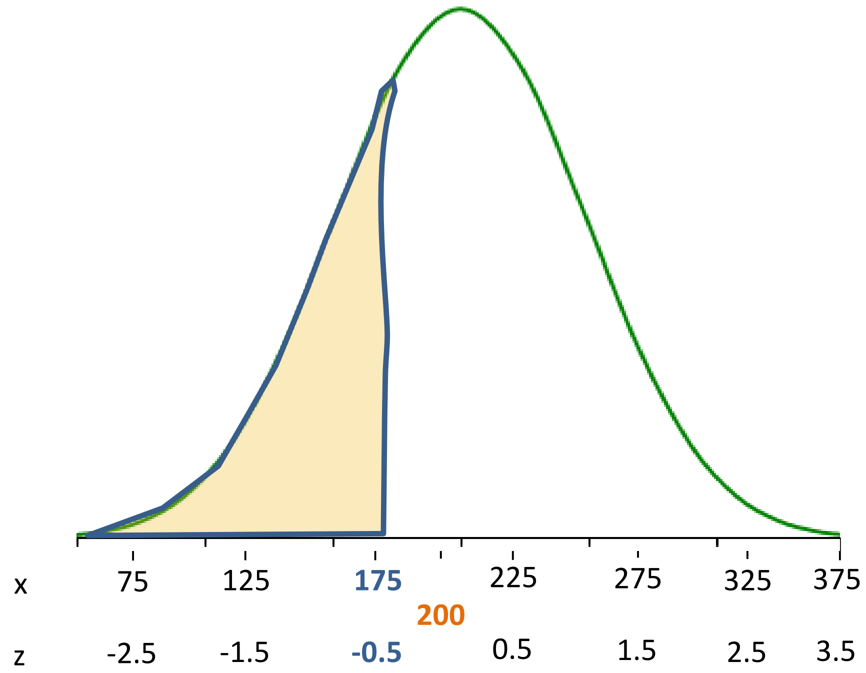


Theoretical distributions – Probability distributions

Standard Normal distribution

- Programmed (e.g. EXCEL: cpdf = Normsdist (*(upperLimit-mean)/std)*)

$N(200,50)$ $$f(x \leq 175) = f\left(\frac{x-\mu}{\sigma} \leq \frac{175-200}{50}\right)$$ $$f(x \leq 175) = f(z \leq -0.5)$$



x	Upper limit	z=(xup-200)/50
50	75	-2.5
100	125	-1.5
150	175	-0.5
200	225	0.5
250	275	1.5
300	325	2.5
350	375	3.5
400	∞	∞

Theoretical distributions – Probability distributions

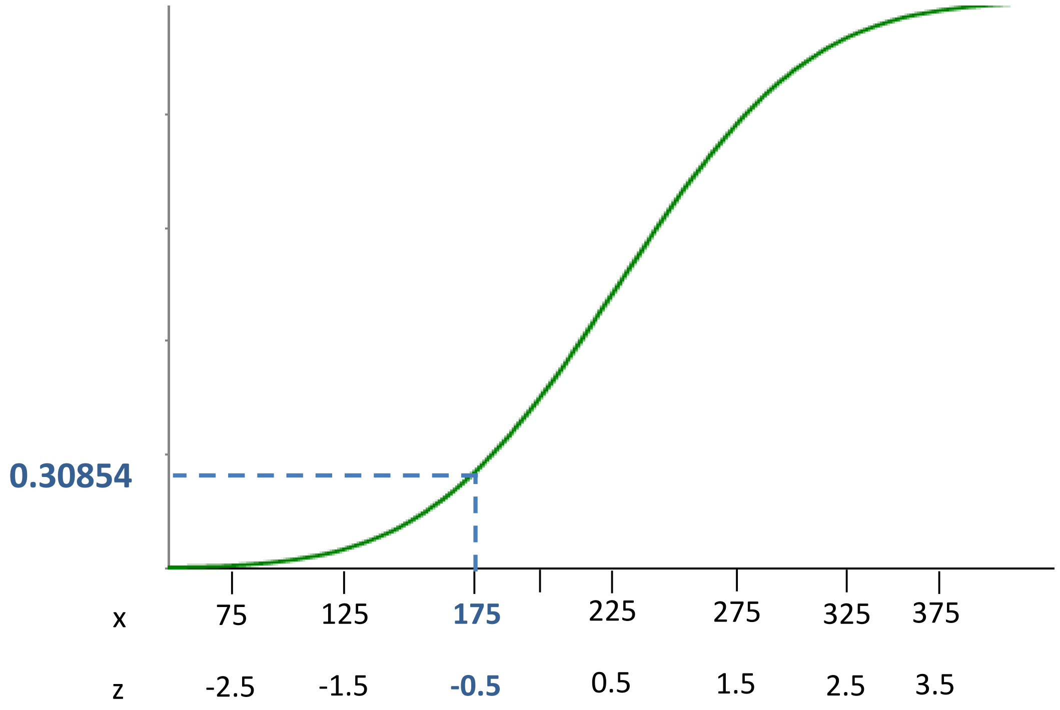
Standard Normal distribution

- Programmed (e.g. EXCEL: cpdf = Normsdist (*(upperLimit-mean)/std)*)

$N\ (200{,}50)$

$$f(x \le 175) = f\left(\frac{x-\mu}{\sigma} \le \frac{175-200}{50}\right)$$

$$f(x \le 175) = f(z \le -0.5)$$



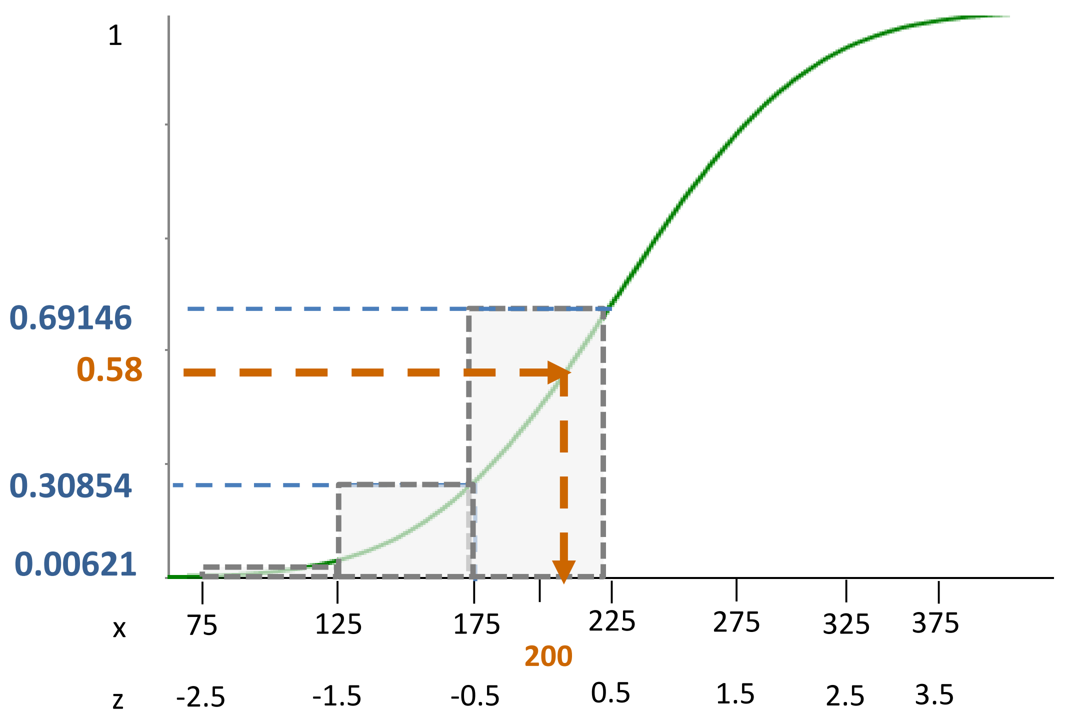
x	Upper limit	z=(xup-200)/50	cpdf
50	75	-2.5	0.00621
100	125	-1.5	0.06681
150	175	-0.5	0.30854
200	225	0.5	0.69146
250	275	1.5	0.93319
300	325	2.5	0.99379
350	375	3.5	0.99977
400	∞	∞	1

Theoretical distributions – Probability distributions

Standard Normal distribution

- Programmed (e.g. EXCEL: cpdf = Normsdist (*(upperLimit-mean)/std)*)

$N\ (200{,}50)$



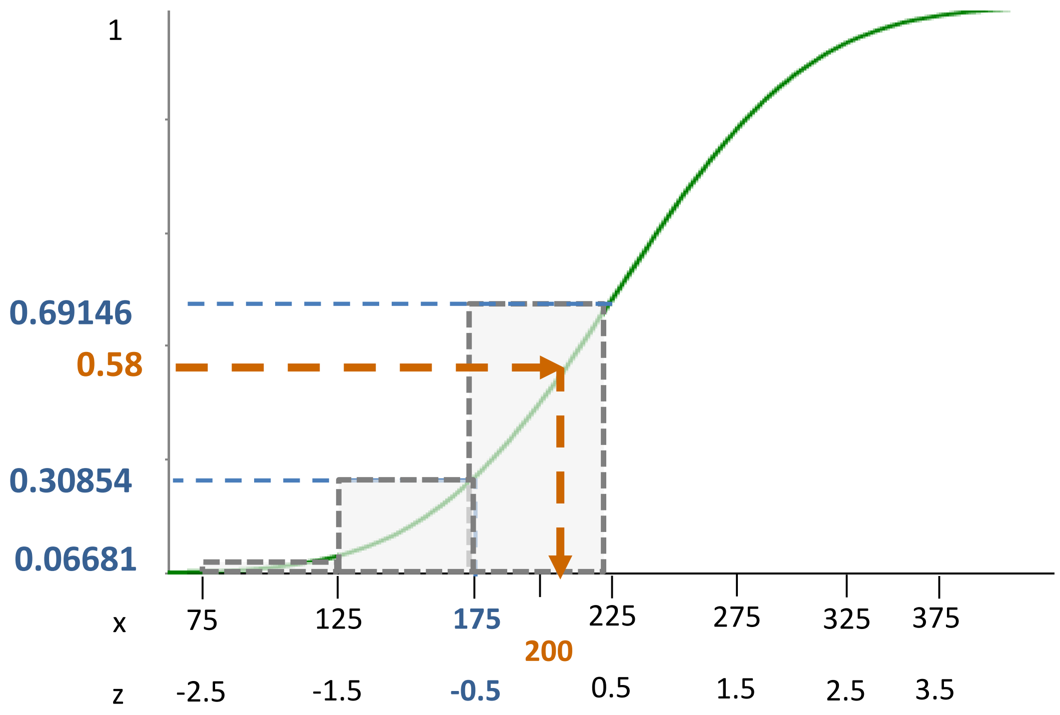
Draw random number (0-1) = 0.58

x	Upper limit	z=(xup-200)/50	cpdf
50	75	-2.5	0.00621
100	125	-1.5	0.06681
150	175	-0.5	0.30854
200	225	0.5	0.69146
250	275	1.5	0.93319
300	325	2.5	0.99379
350	375	3.5	0.99977
400	∞	∞	1

Theoretical distributions – Probability distributions

Normal N (200,50) cpdf = Normsdist (*(upperLimit-mean)/std)*)

Draw random number (0-1) = 0.58



x	Upper limit	z=(xup-200)/50	cpdf
50	75	-2.5	0.00621
100	125	-1.5	0.06681
150	175	-0.5	0.30854
200	225	0.5	0.69146
250	275	1.5	0.93319
300	325	2.5	0.99379
350	375	3.5	0.99977
400			1

Theoretical distributions – Probability distributions:

Poisson

↓

Number of events that occur in an interval of time

Only has one parameter λ
Bounded between 0 and infinity

Assumptions:

- Rate at which events occur is CONSTANT
- Events are independent, ie one event does not affect the subsequent

In some cases it may be applied even if the rate is not constant (e.g. nr of purchaces during a day , not likely to have it happening at 3 in the morning…)

Theoretical distributions – Probability distributions:

Exponential

↓

Time taken between 2 events occurring

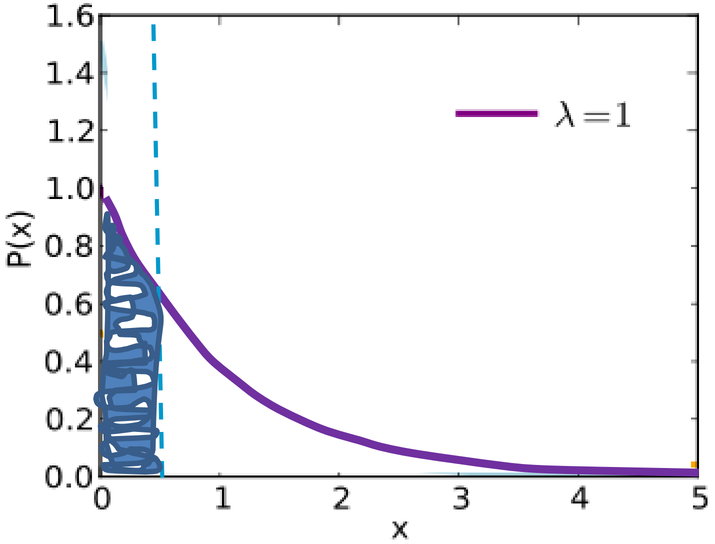
Theoretical distributions – Probability distributions:

Exponential - Time taken between 2 events occurring

λ — Average nr of events occurring in one time unit (min., days, weeks…)

$P(x) = \lambda e^{-\lambda x}$

Probability density function



P (X <= 0.5)

Probabilities are AREAS

$\int$ integrate!

$P(X <= x) = 1 - e^{-\lambda x}$

Theoretical distributions – Probability distributions

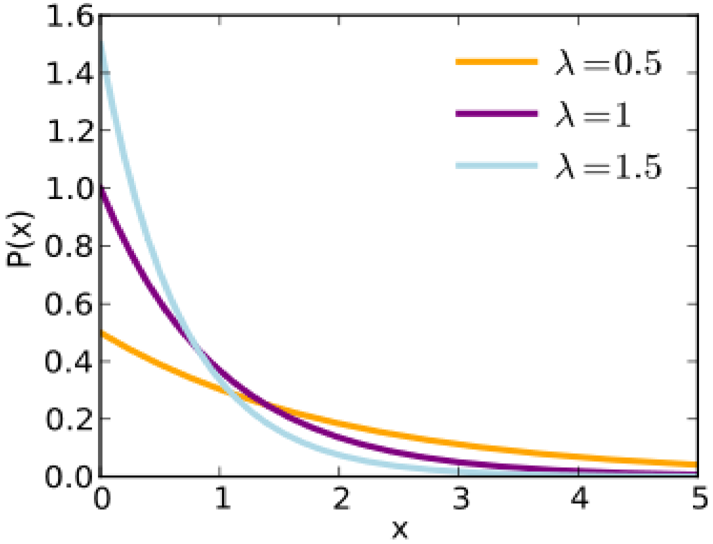
Exponential - Time taken between 2 events occurring

λ — Average nr of events occurring in one time unit (min., days, weeks…)

$$P(x) = \lambda e^{-\lambda x}$$

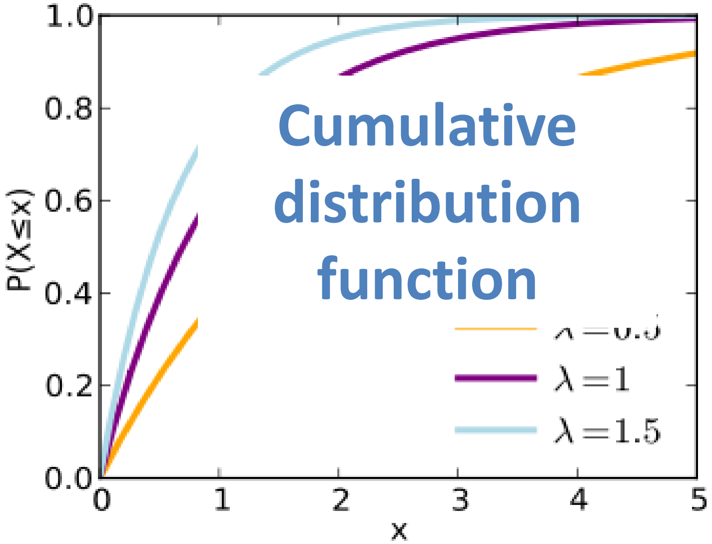
$$P(X <= x) = 1 - e^{-\lambda x}$$

Probability density function



Probabilities are AREAS

∫ int egrate!



Cumulative distribution function

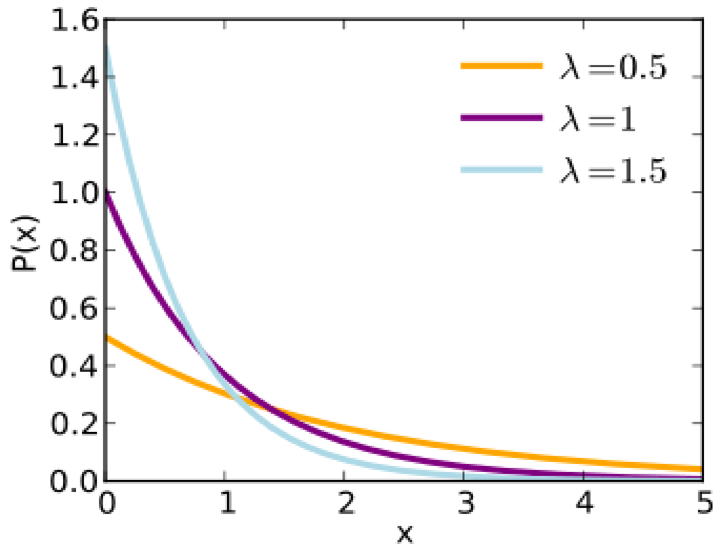
Theoretical distributions – Probability distributions

Exponential - Time taken between 2 events occurring

λ — Average nr of events occurring in one time unit (min., days, weeks…)

Probability density function

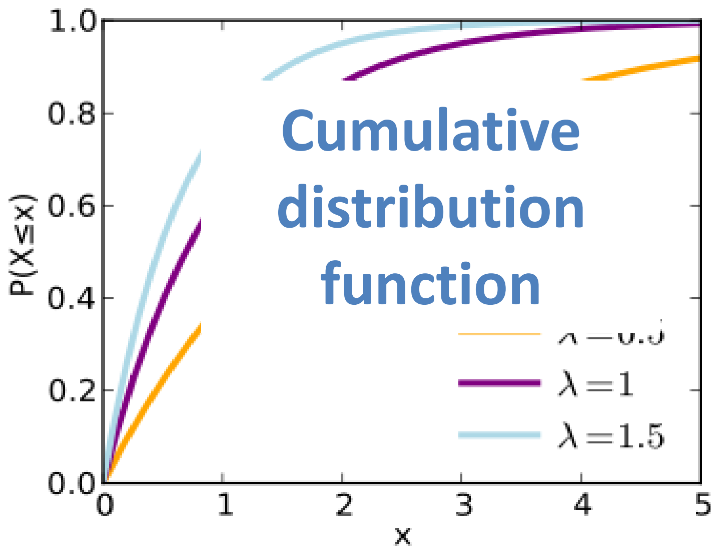
$$P(x) = \lambda e^{-\lambda x}$$



Probabilities are AREAS

∫ integrate!

$$P(X <= x) = 1 - e^{-\lambda x}$$



Cumulative distribution function

Theoretical distributions:

Weibull

is a 3-parameter pdf, used in diameter distribution modelling

$$f(x) = \frac{c}{b}\left(\frac{x-a}{b}\right)^{c-1} exp\left[-\left(\frac{x-a}{b}\right)^{c}\right] \quad (a <= x < \infty)$$

$$= 0 \quad \text{otherwise}$$

a – location parameter (related to the d_{min})

b – scale parameter (>0)

c – shape parameter (>0; if c>1 implies a inverse J shape; if c=3.6 is close to Normal; c<3.6 is right skewed; if c>3.6 is left skewed)

a+b is close to percentile 63% (P_{63}) of the distribution

1. What is Monte Carlo? Basic Principles
2. Distributions
3. Random Numbers
4. Sample Sizes
5. Monte Carlo Simulation Examples
6. Monte Carlo Simulation Exercises

Monte Carlo Simulation Examples

5

- **Example 4** - Demand for paper (units/week) and the lead time for paper production (weeks) are given by theoretical distributions: **demand** has a **normal dist**. (200,50) and the **lead time an exponential dist** (1).

  Simulate *The Old Library* stock assuming that:

  the initial stock (units)= 600,

  the order point (units)= 200 and
  the quantity ordered (units)= 600

- **Example 4** - Demand for paper (units/week) has a **normal dist**. (200; 50)

1. Set up the probability distribution
2. Build cumulative probability distribution considering a 50 units interval for demand

demand	upper xup	z=(xup-200)/50	cummulative distribution
50	75	-2.5	0.00621
100	125	-1.5	0.06681
150	175	-0.5	0.30854
200	225	0.5	0.69146
250	275	1.5	0.93319
300	325	2.5	0.99379
350	375	3.5	0.99977
400	¥	¥	1.00000

Excel Function gives the cumulative distribution

Normsdist *(z)*

- **Example 4** - Demand for paper (units/week) has a **normal dist**. (200; 50)

1. Set up the probability distribution
2. Build cumulative probability distribution considering a 50 units interval for demand
3. Establish an interval of random numbers

demand	upper xup	z=(xup-200)/50	cummulative distribution	aux	lower lim	upper lim	interval
50	75	-2.5	0.00621	6	0	5	0 - 5
100	125	-1.5	0.06681	67	6	66	6 - 66
150	175	-0.5	0.30854	309	67	308	67 - 308
200	225	0.5	0.69146	691	309	690	309 - 690
250	275	1.5	0.93319	933	691	932	691 - 932
300	325	2.5	0.99379	994	933	993	933 - 993
350	375	3.5	0.99977	1000	994	999	994 - 999
400	¥	¥	1.00000	1000	1000	999	1000 - 999

- **Example 4** - Lead time for pulp production (weeks) has **an exponential dist (1)**

1. Set up the probability distribution
2. Build cumulative probability distribution considering a 1 week interval for lead time

The Exponential cumulative function

1-EXP(-xup/mean)

Lead time	upper xup	cumulative distribution
1	1.5	0.78
2	2.5	0.92
3	3.5	0.97
4	4.5	0.99
5	5.5	1.00

- **Example 4** - Lead time for pulp production (weeks) has **an exponential dist (1)**

1. Set up the probability distribution
2. Build cumulative probability distribution considering a 1 week interval for lead time
3. Establish an interval of random numbers

Lead time	upper xup	cumulative distribution	aux	lower lim	upper lim	interval
1	1.5	0.78	78	0	77	0 - 77
2	2.5	0.92	92	78	91	78 - 91
3	3.5	0.97	97	92	96	92 - 96
4	4.5	0.99	99	97	98	97 - 98
5	5.5	1.00	100	99	99	99 - 99

- **Example 4** - Simulate the library paper *stock for 16 weeks*

the initial stock (units)= 600,
the order point (units)= **200** and
the quantity ordered (units)= 600

week	r_demand	Paper demand	Stock	r_lead-time	lead-time	orders	receive	
0			600					
1	201	150	450 *(600-150)*					
2	765	250	**200** *(450-250)*	52	1	600		order 600 units
3	648	200	600 (200-200+600)				600	receive 600 units
4	196	150	450 *(600-150)*					
5	93	150	300 *(450-150)*					
6	705	250	**50** *(300-250)*	82	2	600		order 600 units
7	10	100	-50 *(50-100)*					
8	20	100	450 *(-50-100+600)*				600	receive 600 units
9	149	150	300 *(450-150)*					
10	398	200	**100** *(300-200)*	35	1	600		order 600 units
11	865	250	450 *(100-250+600)*				600	receive 600 units
12	875	250	**200** *(450-250)*	79	2	600		order 600 units
13	174	150	50 *(200-150)*					
14	975	300	350 *(50-300+600)*				600	receive 600 units
15	269	150	**200** *(350-150)*	43	1	600		order 600 units
16	361	200	600 *(200-200+600)*				600	receive 600 units

1 What is Monte Carlo? Basic Principles

2 Distributions

3 Random Numbers

4 Sample Sizes

5 Monte Carlo Simulation Examples

6 Monte Carlo Simulation Exercises

Monte Carlo Simulation Examples

- **Example 1** – Simulating with a distribution provided (empirical)
- **Example 2** – Setting a distribution based on know distributions of other variables (independent variables)
- **Example 3** – Simulating using dependent variables
- **Example 4** – Simulating using theoretical distributions

Usage Notes

- A lot of slides are adopted from the presentations and documents published on internet by experts who know the subject very well.
- I would like to thank who prepared slides and documents.
- Also, these slides are made publicly available on the web for anyone to use
- If you choose to use them, I ask that you alert me of any mistakes which were made and allow me the option of incorporating such changes (with an acknowledgment) in my set of slides.

Sincerely,

Dr. Cahit Karakuş

cahitkarakus@gmail.com