Probability theory 2

Shyam Gopalakrishnan
16th March 2021

Recap of probability 1

- Probability theory vs Statistical inference

Recap of probability 1

- Probability theory vs Statistical inference
- Random variable

Recap of probability 1

- Probability theory vs Statistical inference
- Random variable
- Rules and concepts in probability

Recap of probability 1

- Probability theory vs Statistical inference
- Random variable
- Rules and concepts in probability
- Discrete and continuous random variables

Recap of probability 1

- Probability theory vs Statistical inference
- Random variable
- Rules and concepts in probability
- Discrete and continuous random variables
- Probability distributions

Recap of probability 1

- Probability theory vs Statistical inference
- Random variable
- Rules and concepts in probability
- Discrete and continuous random variables
- Probability distributions
- Expectation, variance, covariance, correlation

Recap of probability 1

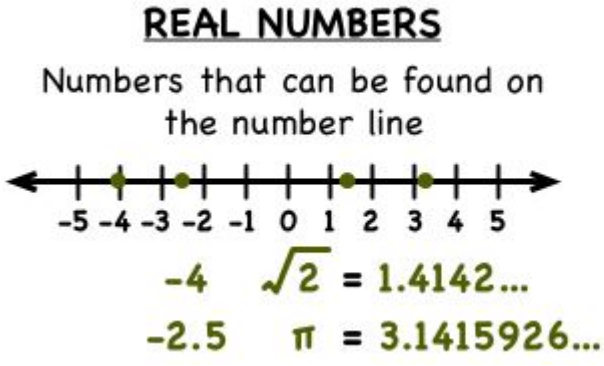
- Probability theory vs Statistical inference
- Random variable
- Rules and concepts in probability
- Discrete and continuous random variables
- Probability distributions
- Expectation, variance, covariance, correlation
- Examples of probability distributions for discrete random variables

Today's learning outcomes

- Difference between continuous and discrete random variables
- Defining probability distributions for continuous random variables (CRVs)
- Expectation and variance of CRVs
- Examples of continuous probability distributions
- If time allows, central limit theorem

Continuous random variables

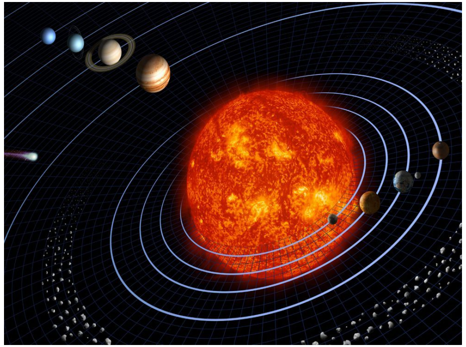
Defined over intervals on the real number line



Interlude: Why do we need continuous random variables?

Consider a specific case:

Let X be a random variable that represents the distance of a planet from its star.

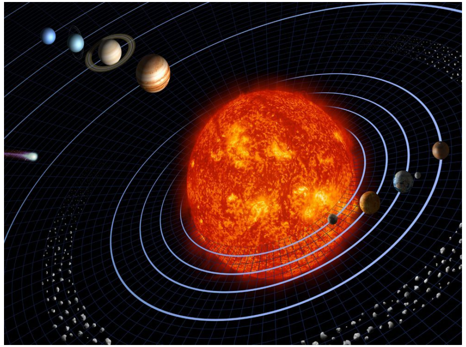


Interlude: Why do we need continuous random variables?

Consider a specific case:

Let X be a random variable that represents the distance of a planet from its star.

We can measure it in km, and then assign a probability to each possible value of the distance

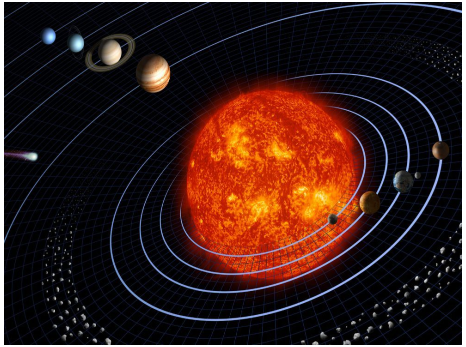


Interlude: Why do we need continuous random variables?

Consider a specific case:

Let X be a random variable that represents the distance of a planet from its star.

We can measure it in km, and then assign a probability to each possible value of the distance - Quantization



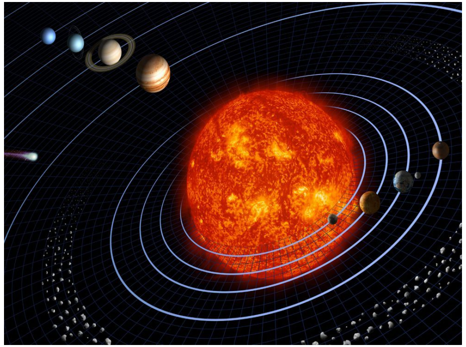
Interlude: Why do we need continuous random variables?

Consider a specific case:

Let X be a random variable that represents the distance of a planet from its star.

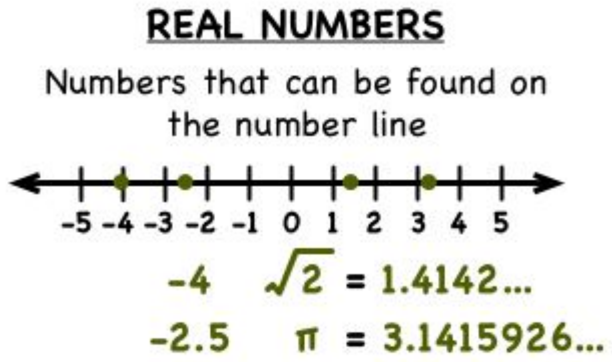
We can measure it in km, and then assign a probability to each possible value of the distance - Quantization

But the distance is a *real* number - so why limit to *integers*



Continuous random variables

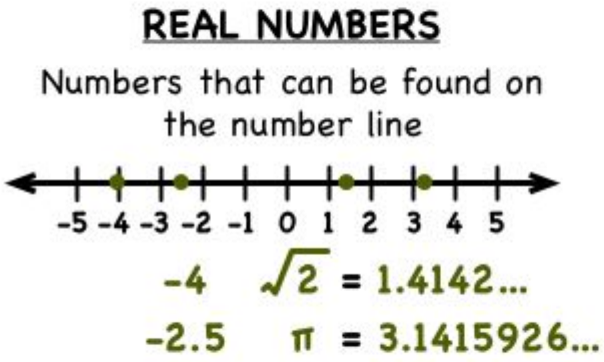
Defined over intervals on the real number line



Intervals are notated as (a, b) where both a and b are real numbers, a < b. The interval represents the part of the real number line between a and b.

Continuous random variables

Defined over intervals on the real number line



Intervals are notated as (a, b) where both a and b are real numbers, a < b. The interval represents the part of the real number line between a and b.

Reminder: Discrete random variable - Coin toss

- Assuming **X ~ Bernoulli (0.5)** then:
  - $P[X = 1] = 0.5$
  - $P[X = 0] = 0.5$

Reminder: Discrete random variable - Coin toss

- Assuming **X ~ Bernoulli (0.5)** then:
  - $P[\, X = 1 \,] = 0.5$
  - $P[\, X = 0 \,] = 0.5$

Here 0 and 1 are the values the random variable can take.
Do you remember what this was called?

Reminder: Discrete random variable - Coin toss

- Assuming **X ~ Bernoulli (0.5)** then:
  - $P[\,X = 1\,] = 0.5$
  - $P[\,X = 0\,] = 0.5$

Here 0 and 1 are the values the random variable can take.
Sample space or **Sample set**

Reminder: Discrete random variable - Coin toss

- Assuming **X ~ Bernoulli (0.5)** then:
  - P[X = 1] = 0.5
  - P[X = 0] = 0.5

Here 0 and 1 are the values the random variable can take. **Discrete random variables** have **countably many points** in their sample set, while **continuous random variables** have **uncountably many points**.

Defining the probabilities for CRVs

Defined over intervals on the real number line

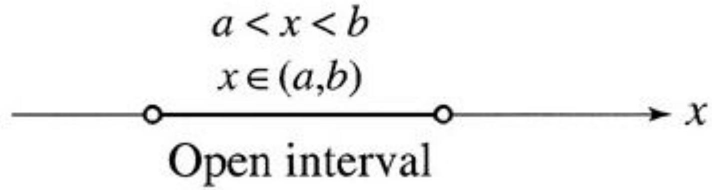
$$P(X \in (a, b)) = P(a < X < b)$$

"Belongs to"

Interval (a, b)

$$x \in (a, b)$$

$$a < x < b$$



Computing probabilities for CRVs

Reminder: Probability mass function

Defines the probability of each point in the sample set

$$P[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$$

$$: P[T = t] = \overbrace{(1-p)^{t-1}}^{\text{t-1 failures}} \overbrace{p}^{\text{1 success}}$$

$$P[X = k] = \frac{\lambda^k e^{-\lambda}}{k!}$$

Computing probabilities for CRVs

Let us try to do the same for CRVs - but how do we do this since probabilities for CRVs are defined over intervals on the real line?

How can we compute $P(a < X < b)$?

Let us define two functions.

Cumulative distribution function

Cumulative distribution function (CDF): Usually denoted by F

$$F(a) = P(X \leq x)$$

For a discrete random variable,

$$F(a) = \sum_{x \leq a} P(X = x)$$

We can do this because there are a **countable** number of points.

How can we do this for CRVs?

Probability density function

Equivalent to PMF for Discrete random variables.

Probability density function (PDF): Usually denoted by f

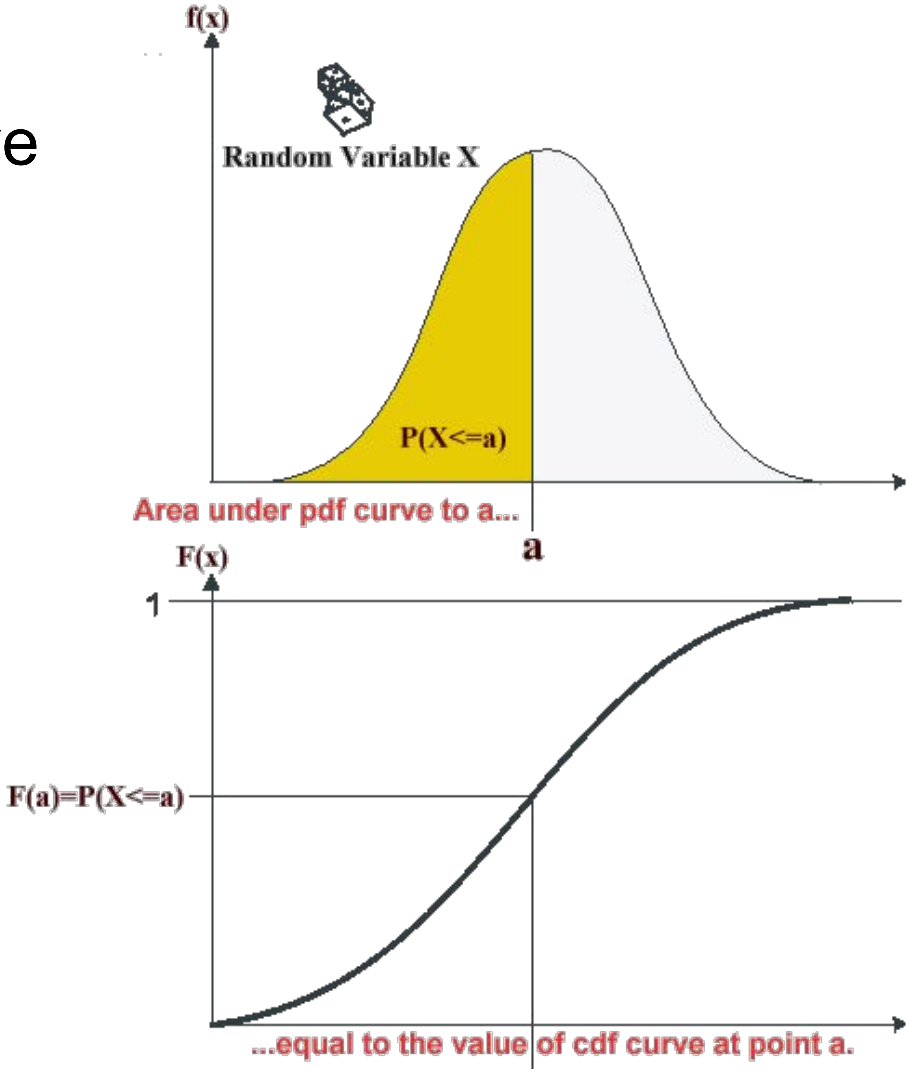
$$f(x) = F'(x)$$

In reverse, we can define the CDF using the PDF

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

Interlude: Area under a curve

Integration is the "calculus" way of computing area under a curve



Computing probabilities for CRVs

How can we compute $P(a < X < b)$?

$$\boxed{P(a < X < b) = F(b) - F(a)}$$

Prove this!

Hint : Use definition of F(x).

Computing probabilities for CRVs

How can we compute $P(a < X < b)$?

$$\boxed{P(a < X < b) = F(b) - F(a)}$$

Prove this!

Hint : Use definition of F(x).

But what about $P(X=a)$? Is $f(a) = P(X=a)$?

Expectation of CRVs

Let X be a CRV with CDF F, and PDF f

$$E[X] = \int_{-\infty}^{+\infty} (1 - F(x))dx$$

Expectation of CRVs

Let X be a CRV with CDF F, and PDF f

$$E[X] = \int_{-\infty}^{+\infty} (1 - F(x))dx$$

Or the more familiar

$$E[X] = \int_{-\infty}^{+\infty} xf(x)dx$$

Expectation of CRVs

Let X be a CRV with CDF F, and PDF f

$$E[X] = \int_{-\infty}^{+\infty} (1 - F(x))dx$$

Or the more familiar

$$E[X] = \int_{-\infty}^{+\infty} xf(x)dx$$

The discrete case:

$$E[X] = \sum_i x_i P[X = x_i]$$

Expectation of CRVs

Let X be a CRV with CDF F, and PDF f

$$E[X] = \int_{-\infty}^{+\infty} (1 - F(x))dx$$

Or the more familiar

$$E[X] = \int_{-\infty}^{+\infty} x f(x) dx$$

Common observation: Sums in discrete case become integrals in continuous case.

The discrete case:

$$E[X] = \sum_i x_i P[X = x_i]$$

Differences: discrete and continuous random variables

Discrete RV

Defined on countable sample set

Defined by Probability Mass Function

Expectation is sum over sample set

Continuous RV

Defined on uncountable sample space (real number line).

Defined by Probability density function

Expectation is integration over sample space

Common probability distributions for CRVs

Many processes are modeled using CRVs

- Human height
- Temperature in a room
- Waiting time for the next bus
- Prevalence of a disease

Common probability distributions for CRVs

Many processes are modeled using CRVs

- Human height
- Temperature in a room
- Waiting time for the next bus
- Prevalence of a disease

We will look at some common probability distributions for CRVs.

Uniform distribution (continuous)

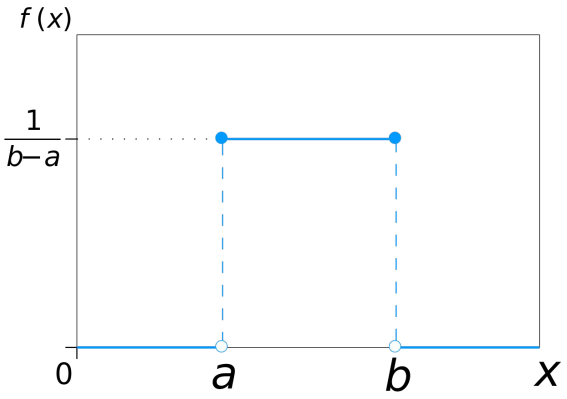
PDF: $f(x) = 1/(b-a)$

Support: (a,b)

Parameters: a, b

$E[X] = (a+b)/2$

$Var[X] = (b-a)^2/12$



Equivalent to the Uniform distribution (discrete), extended to a continuous interval.

Uniform distribution (continuous)

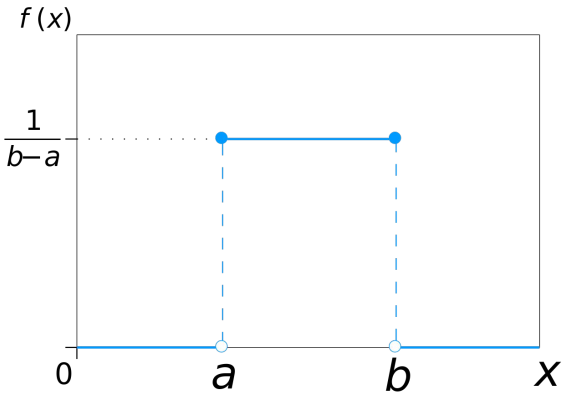
PDF: $f(x) = 1/(b-a)$

Support: (a,b)

Parameters: a, b

$E[X] = (a+b)/2$

$Var[X] = (b-a)^2/12$



What does the CDF look like?

Exponential distribution

Time to the next solar flare

Time between Geiger counter clicks

Time till you see the next yellow car on the highway

Exponential distribution

Time to the next solar flare

Time between Geiger counter clicks

Time till you see the next yellow car on the highway

Memoryless property! - similar to geometric distribution.

Exponential distribution

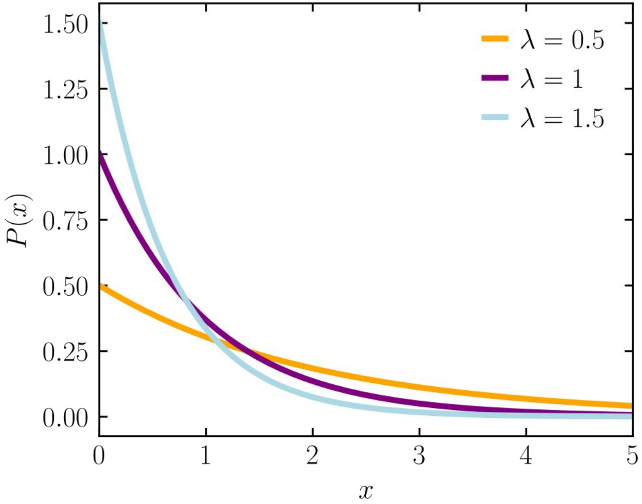
PDF: $f(x) = \lambda e^{-\lambda x}$

Support: (0, +inf)

Parameters: rate λ

$E[X] = 1/\lambda$

$Var[X] = 1/\lambda^2$



Normal distribution

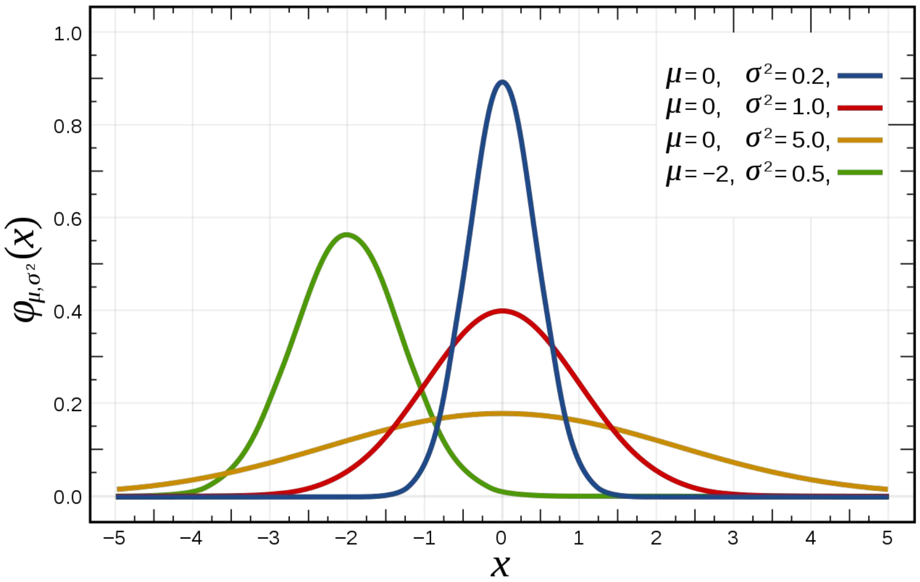
PDF: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

Support: (-inf, +inf)

Parameters: mean μ, variance σ^2

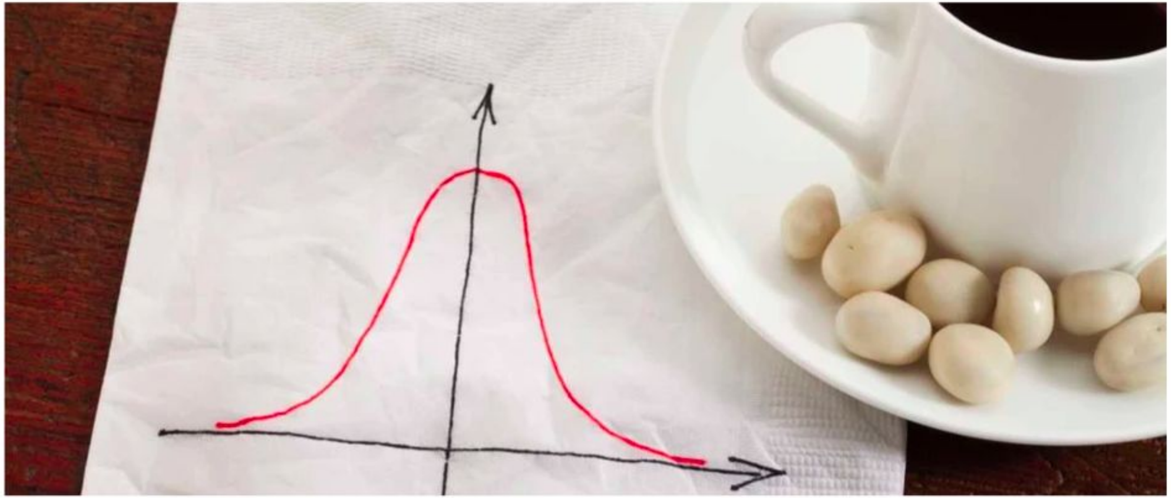
E[X] = μ

Var[X] = σ^2



Normal distribution

Home > Everyday science > **Why is the bell curve so ubiquitous?**



Why is the bell curve so ubiquitous?

Save 50% when you subscribe to *BBC Science Focus Magazine*

Also known, as the Gaussian curve or Normal distribution, the bell curve demonstrates a common mathematical pattern.

Central Limit Theorem

The **central limit theorem** states that if you have a population with mean μ and standard deviation σ and take sufficiently large random samples from the population with replacement, then the distribution of the sample means will be approximately normally distributed.

In short, averages of things tend to be normally distributed, even if the original observations are not!

Exercise time