Day 4

19 May 2024 18:46

Probability density function

- Continuous variables
  - Uniform distribution
  - Normal distribution

discrete

3 | 5 | 7 | 11 ....

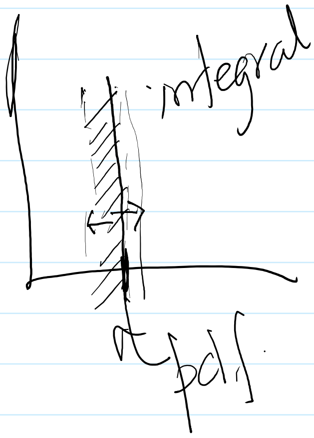
Continuous

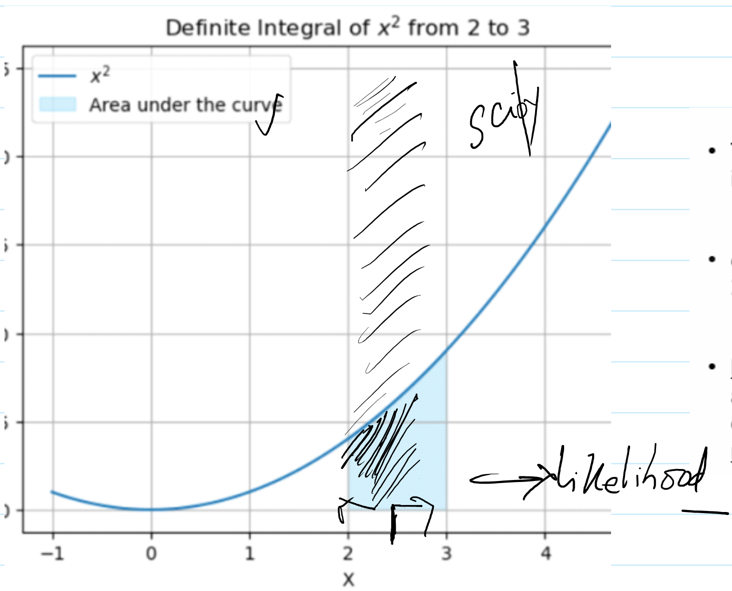
3.5

3.5000001

3.500001

integral calculus





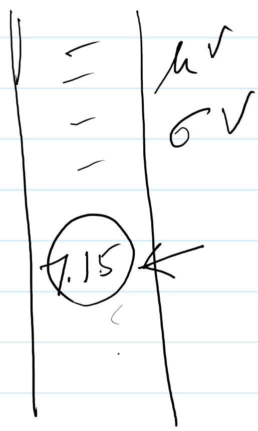
- The function $f(x)=x^2$ is plotted over the interval [-1, 5]
- compute the definite integral of x^2 from 2 to 3 → Area under the curve shaded
- Definite Integral - represents the signed area under the curve, and in this case, it corresponds to the area of the shaded region in the plot

- **Distribution Parameters:** Mean (μ), Standard Deviation (σ)
- **Objective:** Calculate the probability that a random variable X lies between two values a and b.
- **Probability Density Function (PDF) for Normal Distribution**

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

- **Probability Calculation**

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx$$



- **Distribution Parameters:** Mean (μ) = 0, Standard Deviation (σ) = 1
- **Objective:** Calculate the probability that a random variable X lies between two values a and b.
- **Probability Density Function (PDF) for Normal Distribution**

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

- **Probability Calculation -** Let's calculate the probability of the random variable X lying within the interval $-1 \le X \le 1$ using the PDF:

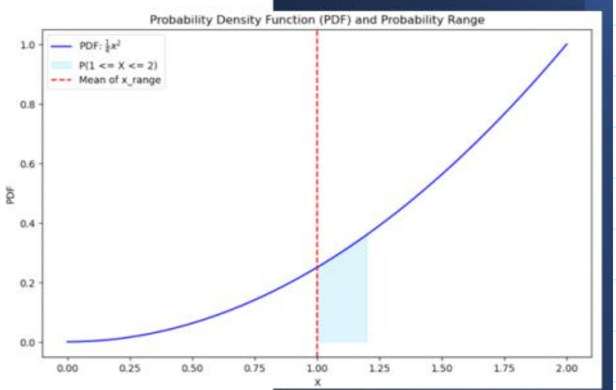
$$P(-1 \le X \le 1) = \int_{-1}^{1} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

Exercise 1: PDF Calculation

- Given the probability density function (PDF) for a continuous random variable X:

$$f(x) = \frac{1}{4}x^2, \quad \text{for } 0 \le x \le 2$$

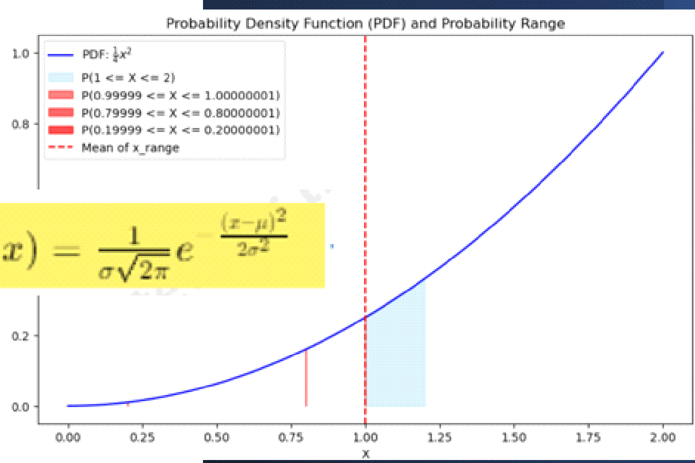
- Calculate the probability that X lies between 1 and 2 .
- Use SciPy library to compute the area under the curve




```
spi.quad(pdf_function, 0.19999, 0.20000001)
```

PDFs at various points

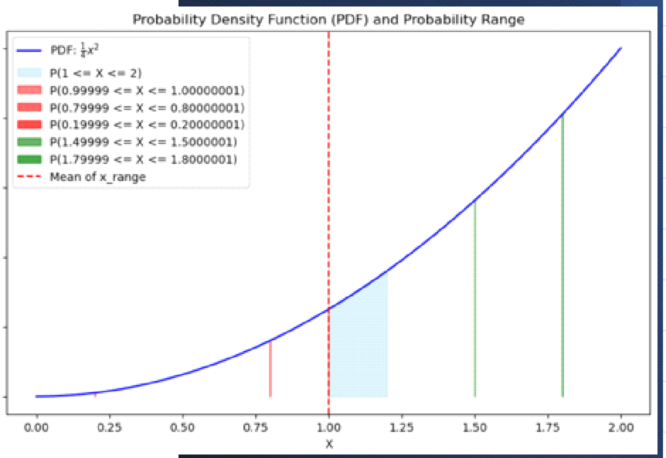
- **Around 1 (the mean):**
  - Result: 2.502475×10−6
  - Interpretation: probability density around the mean of 1 is low within the extremely narrow interval [0.99999, 1.00000001]. While the density is small, it's not zero, indicating that there is some probability mass within this tiny range.
- **Around 0.8:**
  - Result: 1.60158×10−6
  - Interpretation: Similarly, the probability density around 0.8 is low within the narrow interval [0.79999, 0.80000001]. The small value suggests a concentration of probability density around this point.
- **Around 0.2:**
  - Result: 1.00095×10−7
  - Interpretation: The probability density around 0.2 is also low within the narrow interval [0.19999, 0.20000001]. The very small value indicates a concentration of probability density around this point.



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

PDFs at points after the mean

- Around 1.5:
  - Result: 5.6812125039030244×10−06
  - Interpretation: The probability density around 1.5 is relatively higher within the narrow interval [1.49999, 1.5000001] compared to the previous examples. This suggests a concentration of probability density aroun
    1.5.
- Around 1.8:
  - Result: 8.180955004683693×10−06
  - Interpretation: Similarly, the probability density arour
    1.8 is relatively higher within the narrow interval [1.79999, 1.8000001]. The larger value indicates a hig
    concentration of probability density around 1.8 compared to the previous examples.
- In summary
  - the probability density is still small, but it's larger thar
    the values we obtained for intervals around 1, 0.8, an
    0.2.
  - aligns with the general expectation that the probability density is highest around the mean (1) and decreases as you move away from it.



Cumulative density function (CDF)

The Cumulative Distribution Function (CDF) provides the cumulative probability of a random variable being less than or equal to a specific value.

CDF is denoted as *F*(*x*) and is defined as the integral of the PDF up to that value.

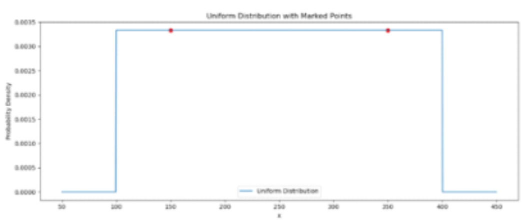
$$F(x) = \int_{-\infty}^{x} f(t)dt$$

Uniform probability distribution

Example

- between a and b is a horizontal line, indicating a constant probability density within the specified interval.
- For illustration, let's take a uniform distribution between a=100 and b=400
- The probability density is the same for all values within the range [a, b].

$$f(x) = \frac{1}{b-a}$$



pdf value at x1 (150): 0.0033333333333333335
pdf value at x2 (350): 0.0033333333333333335

Explanation

- PDF(20.2) = 0.0106
- PDF(30.2) = 0.0134
- PDF(60.2) = 0.0118