Computer Science

CSC380: Principles of Data Science

Probability Primer 4

Outline:

- Continuous probability
- Continuous distribution
  - PDF
  - CDF
- Useful continuous distributions

Credit:

- Jason Pacheco,
- Kwang-Sung Jun,
- Chicheng Zhang
- Xinchen yu

Outline

- Continuous probability
- Continuous distribution
  - PDF
  - CDF
- Useful continuous distributions

Continuous Probability

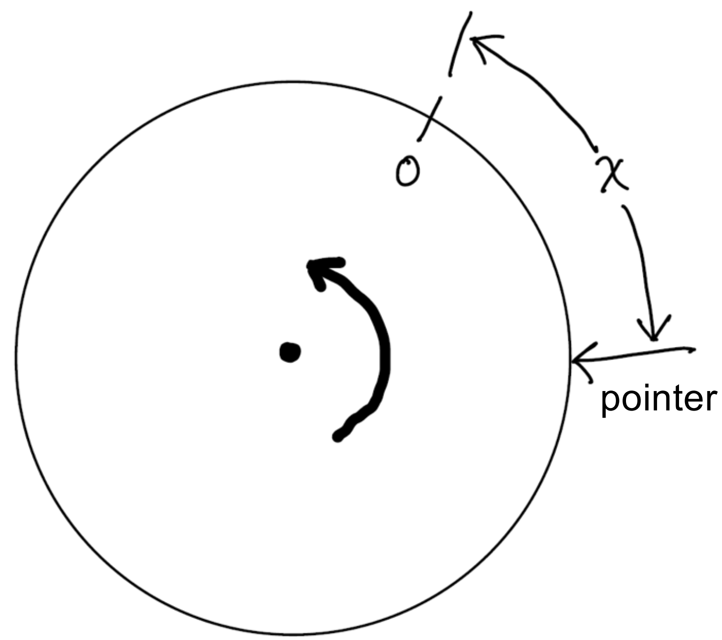
(TV show spin the wheel)

Continuous Probability

Experiment Spin continuous wheel and measure X displacement from 0

Say the circumference is 1.

Outcome space Ω is all points (real numbers) in $(0,1]$



Question Assuming uniform distribution,

what is $P(X = x)$?

A much better question: What is P(X$\leq x$).
(that is, that the wheel reached at most x.
Examples P(X$\leq 0.5)$

Continuous Probability

we could try to convince ourselves that it is sensible.

… or we could just accept this oddity…

Question that make sense in a discrete probability:

What is the probability that X=5

Question that make sense in a continuous probability:

What is the probability that $4 \leq X \leq 5$

Continuous Distributions

Continuous Probability Distributions

Definition *The <u>cumulative distribution function</u> (CDF) of a RV X is the function given by,*

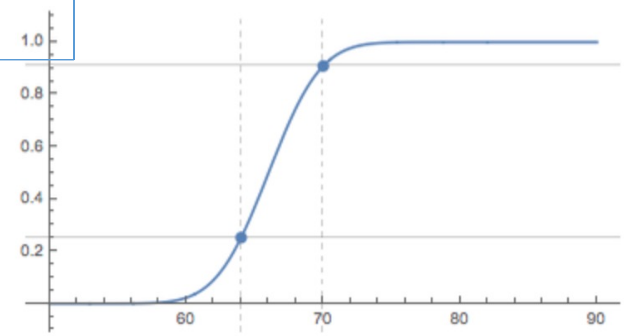
$$F(x) = P(X \leq x)$$

In words: F(0.3) is the probability of the event that causes the RV X to be ≤ 0.3

Key properties:

F is monotonically increasing
$F(\mathrm{x})$ goes to 0/1 if x goes to $-\infty/+\infty$

Example: Pick a number *p* uniformly between 0 and 1.
What is the probability that p< x



➢ Can easily measure probability of closed intervals,

$$P(a < X \leq b) = F(b) - F(a)$$

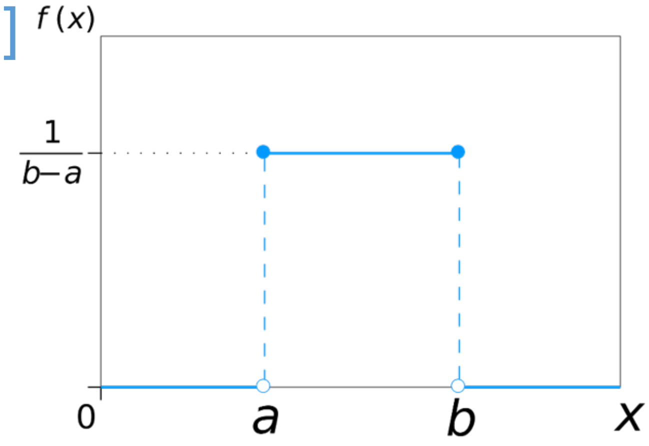
e.g. $a = 64,\ \mathrm{b} = 70$

https://demonstrations.wolfram.com/ConnectingTheCDFAndThePDF/

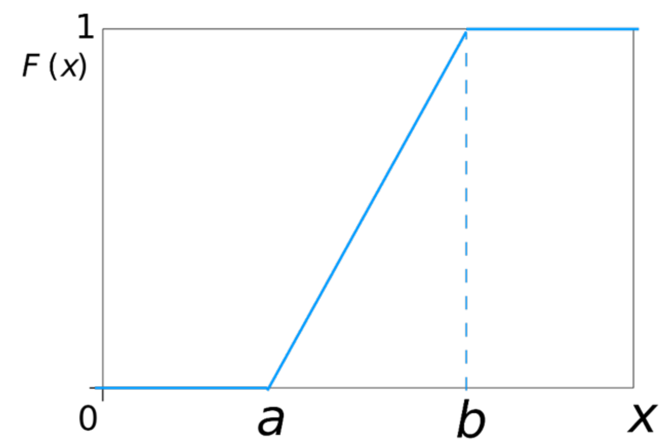
Uniform Continuous Distribution

Uniform distribution on interval $[a, b]$: Uniform[a,b]

$$p(x) = \begin{cases} 0 & \text{if } x \le a, \\ \frac{1}{b-a} & \text{if } a \le x \le b, \\ 0 & \text{if } b \le x \end{cases} \qquad P(X \le x) = \begin{cases} 0 & \text{if } x \le a, \\ \frac{x-a}{b-a} & \text{if } a \le x \le b, \\ 1 & \text{if } b \le x \end{cases}$$



$$\boxed{\mathrm{P}(X \le x) = \int_{-\infty}^{x} p(t)dt}$$

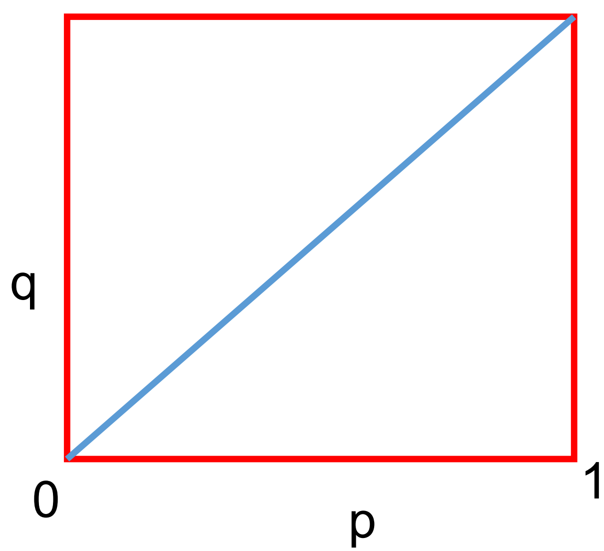


Notation:
$p(x)$ for the PDF function at location x
$\mathrm{P}(A)$ for the probability of event A

Again, PDF function $\neq$ probability

Another exampels

- Pick two numbers p,q from in the range [0,1] uniformly independently.
- What is the probability that $p \leq q$?



Continuous Probability Distributions

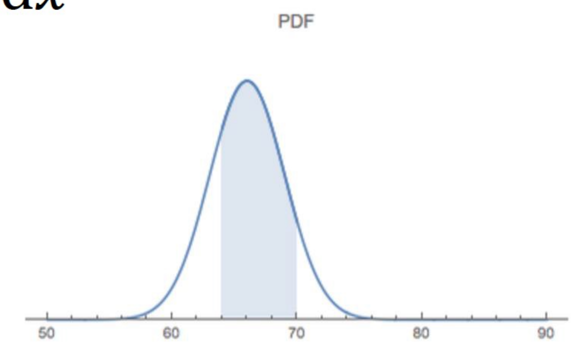
- If $F(X)$ is differentiable then,

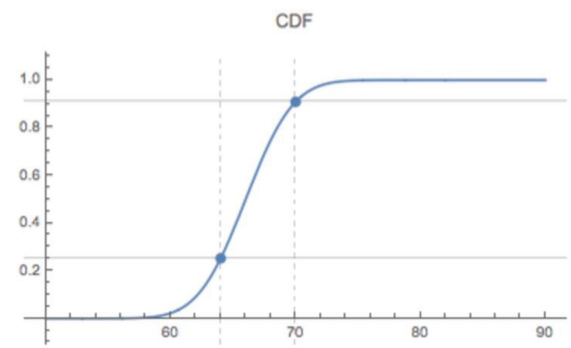
Fundamental Theorem of Calculus

$$p(x) = \frac{dF(x)}{dx} \quad \text{and} \quad F(t) = \int_{-\infty}^{t} p(x)\,\mathrm{d}x$$

$p(x)$ is called X's ***probability density function* (PDF)**

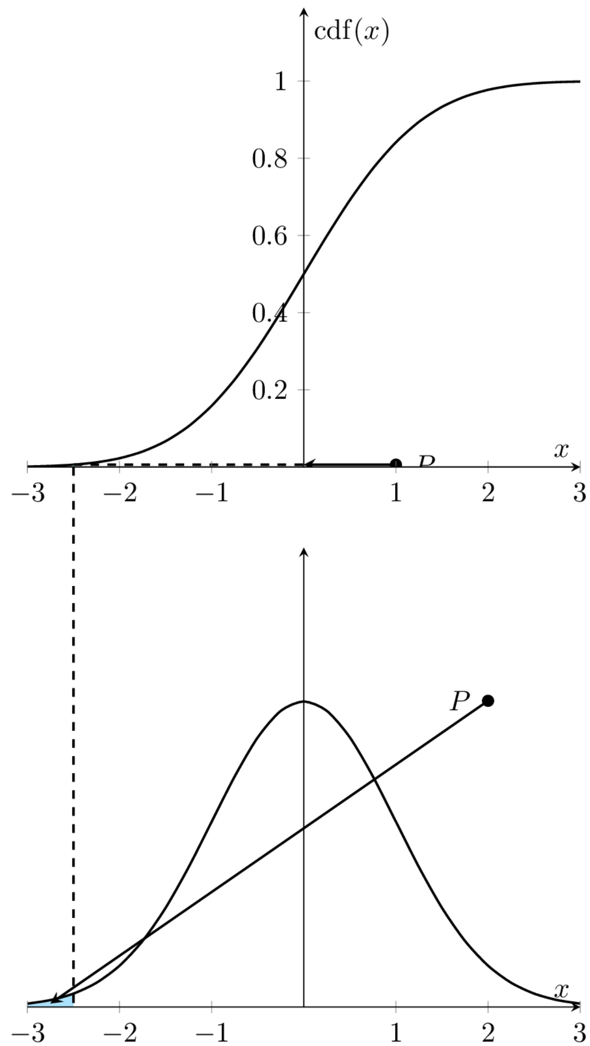
$$\approx \frac{F(x) - F(x-\epsilon)}{x - (x-\epsilon)} = \frac{P(X \in (x-\epsilon, x])}{\epsilon} \text{ when } \epsilon \to 0$$





Intuition: $p(x)$ characterizes how likely X takes values in the neighborhood of x

- $p(x) \geq 0$ for all x
- $P(a < X \leq b) = F(b) - F(a) = \int_a^b p(x)\mathrm{d}x$
- $\int_{-\infty}^{+\infty} p(x)\mathrm{d}x = 1$

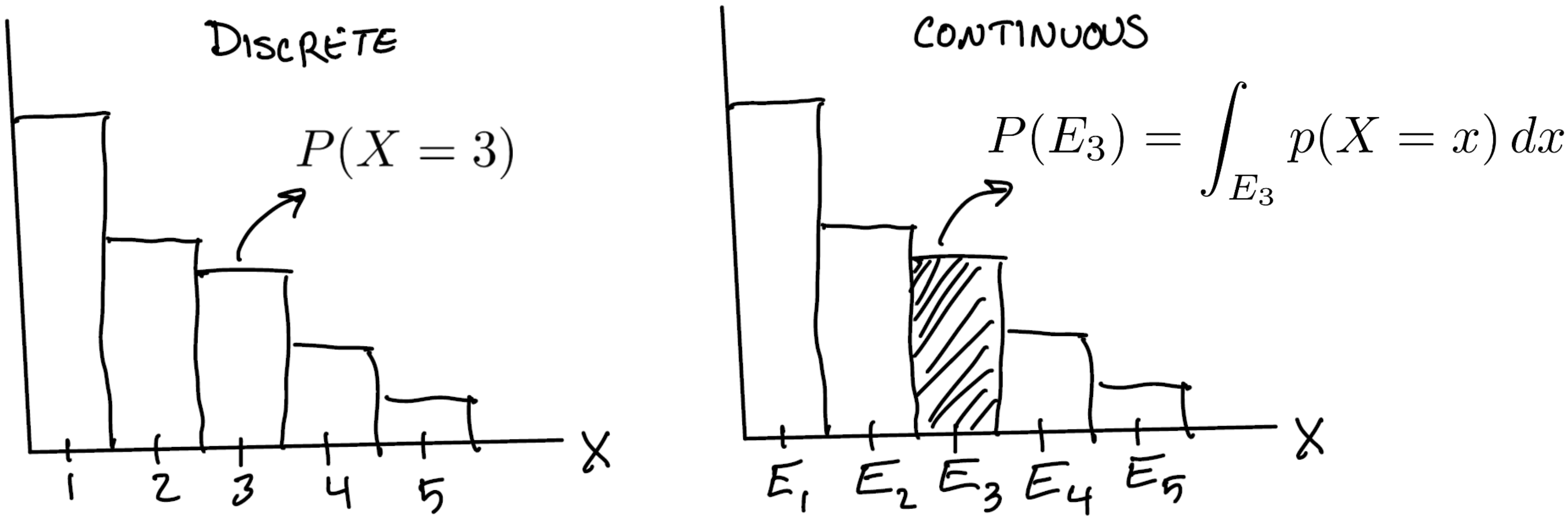


$$P(a < X \leq b) = F(b) - F(a)$$

$$F(t) = \int_{-\infty}^{t} p(x)\,\mathrm{d}x$$

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

Continuous Probability



- Events represented as intervals $a \leq X < b$ with probability,

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

- Specific outcomes have zero probability $P(X = x) = 0$
- But may have nonzero *probability density* $p(X = x) > 0$

Notation

- For continuous RV X, use $p(X = x), p(x), pX(x)$ to denote its PDF (probability density function)
  - Recall: $\mathrm{P}(X = x)$ is not its PDF value (in fact, always 0)

- For **discrete** RV X, use $p(X = x), p(x), pX(x)$ to denote its PMF (probability **mass** function)
  - In this case, $p(X = x) = P(X = x)$

- General suggestions for HW / exams: to be extra safe, you can explicitly declare "we use $p(X = x)$ to denote the PDF of continuous RV X"

Continuous Probability Distributions

Most definitions for discrete RVs hold, replacing sum with integral…

Law of Total Probability for continuous distributions,

$$p(x) = \int_{\mathcal{Y}} p(x, y)\, dy$$

Recall: for discrete X

$$P(X = x) = \sum_{y} P(Y = y, X = x)$$

All the rules apply when replacing PMF with PDF…

Conditional PDF:

$$p(X \mid Y) = \frac{p(X,Y)}{p(Y)} = \frac{p(X,Y)}{\int p(x,Y)\, dx}$$

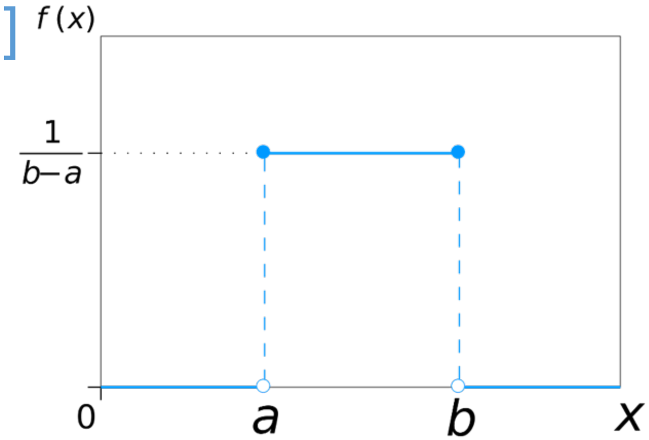
Probability Chain Rule:

$$p(X, Y) = p(Y)p(X \mid Y)$$

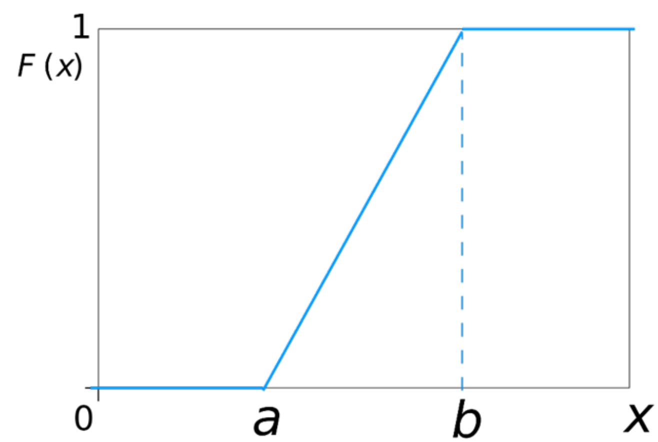
Uniform Continuous Distribution

Uniform distribution on interval $[a, b]$: Uniform[a,b]

$$p(x) = \begin{cases} 0 & \text{if } x \le a, \\ \frac{1}{b-a} & \text{if } a \le x \le b, \\ 0 & \text{if } b \le x \end{cases} \qquad P(X \le x) = \begin{cases} 0 & \text{if } x \le a, \\ \frac{x-a}{b-a} & \text{if } a \le x \le b, \\ 1 & \text{if } b \le x \end{cases}$$



$$\boxed{\mathrm{P}(X \le x) = \int_{-\infty}^{x} p(t)dt}$$



Notation:

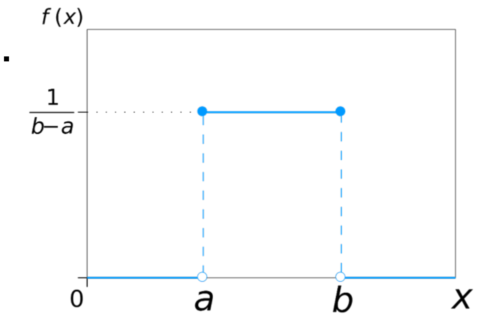
$p(x)$ for the PDF function at location x

$\mathrm{P}(A)$ for the probability of event A

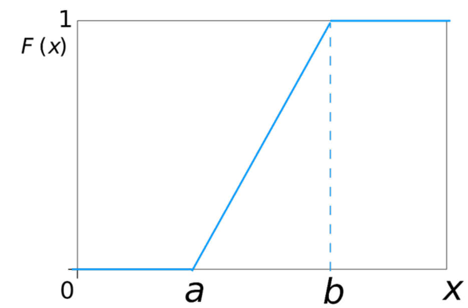
Again, PDF function ≠ probability

Uniform Continuous Distribution

Example: Let $X = \mathrm{length}$ of an eight-week-old baby's smile ($X \sim \mathrm{U}(0, 23)$).
The probability density function is $\mathrm{p}(x) = \frac{1}{23-0} = \frac{1}{23}$ for $0 \leq \mathrm{X} \leq 23$.



Q: find the probability that a random eight-week-old baby smiles more than 12 seconds knowing the baby smiles more than 8 seconds.



Method 1 (write a new PDF):

$X \sim U(8, 23)$

$$\mathrm{p}(x) = \frac{1}{23-8} = \frac{1}{15}$$

$\mathrm{P}(23 > x > 12\,)$

$$= \frac{(23-12)}{15}$$

≈ 0.7333

Method 2 (bayes rule):

$\mathrm{P}(x > 12 \mid x > 8\,)$

$$= \frac{P(x > 12 \; and \; x > 8)}{P(x > 8)} = \frac{P(x > 12)}{P(x > 8)}$$

$$= \frac{(23-12)\times\frac{1}{23}}{(23-8)\times\frac{1}{23}} \approx 0.7333$$

Uniform Continuous Distribution

numpy.random.uniform

numpy.random.uniform(*low=0.0, high=1.0, size=None*)

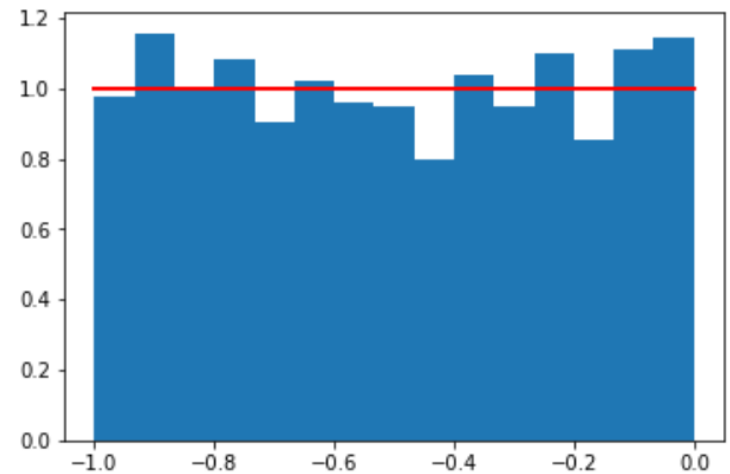
Draw samples from a uniform distribution.

Samples are uniformly distributed over the half-open interval `[low, high)` (includes low, but excludes high). In other words, any value within the given interval is equally likely to be drawn by **uniform**.

Example Draw 1,000 samples from a uniform on [-1,0),

redline: PDF of uniform distr.

```
a = -1
b = 0
N = 1000
X = np.random.uniform(a,b,N)
count, bins, ignored = plt.hist(X, 15, density=True)
plt.plot(bins, np.ones_like(bins), linewidth=2, color='r')
plt.show()
```



`bins`: length 16, consisting of boundary points

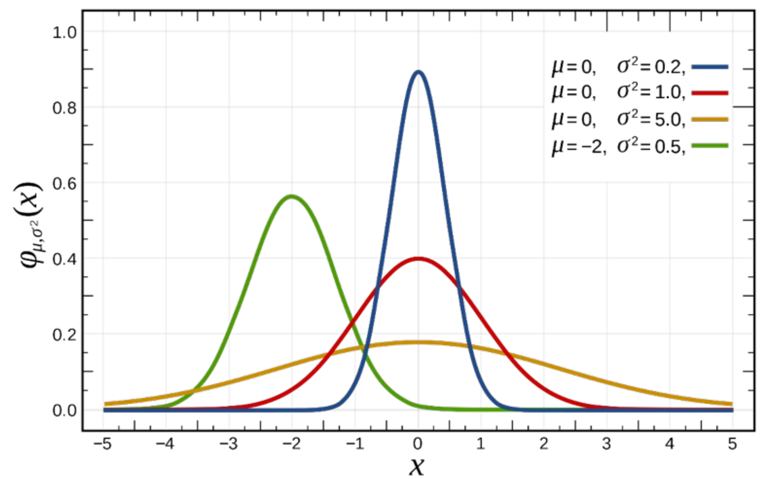
Gaussian/Normal Distribution

Gaussian (a.k.a. Normal) distribution with mean mean (location) μ and variance (scale) σ^2 parameters,

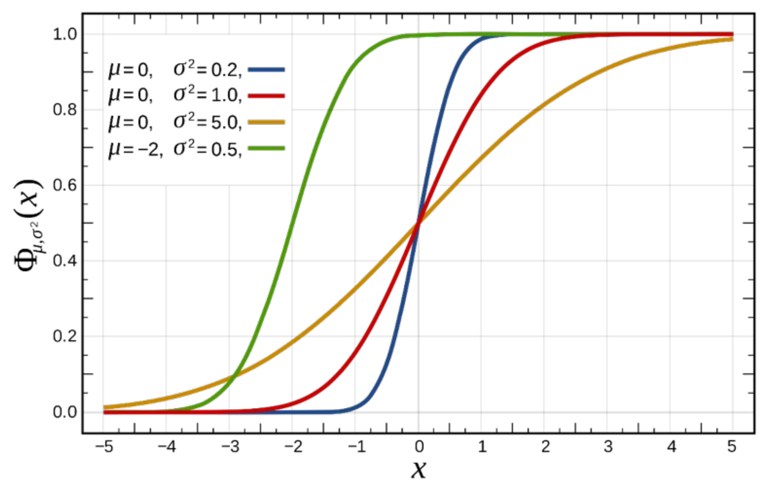
$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{(x-\mu)^2}{2\sigma^2})$$

Compactly, $X \sim \mathcal{N}(\mu, \sigma^2)$

PDF



CDF

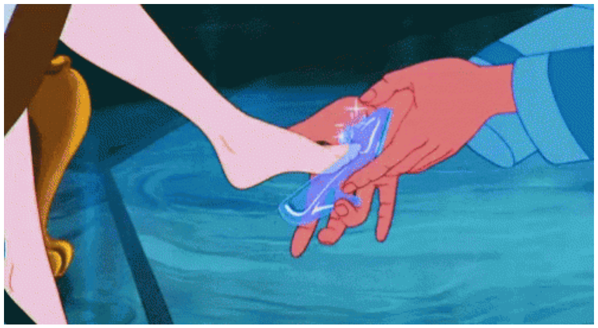


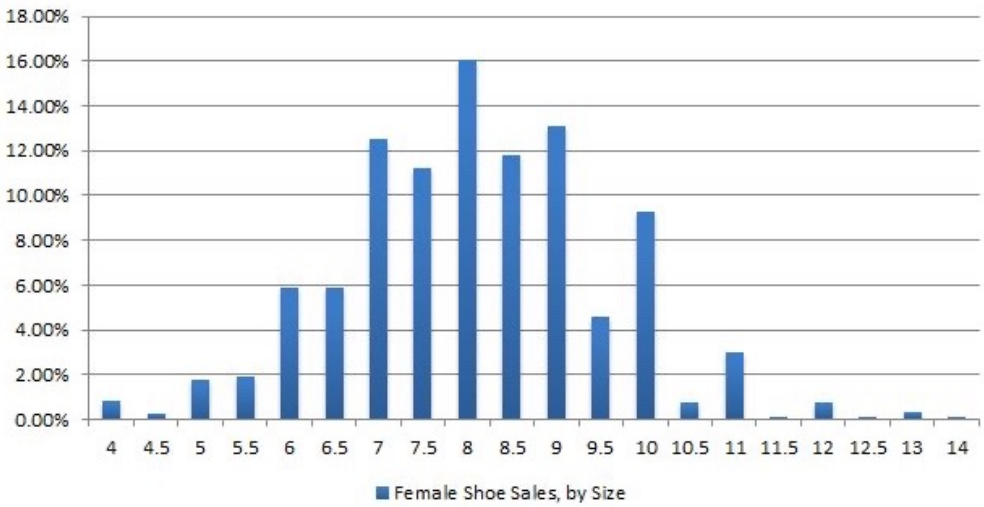
Observations:

- Larger σ^2 : $p(x)$ more “spread out”
- Larger μ : $p(x)$ ’s center shifts to the right more

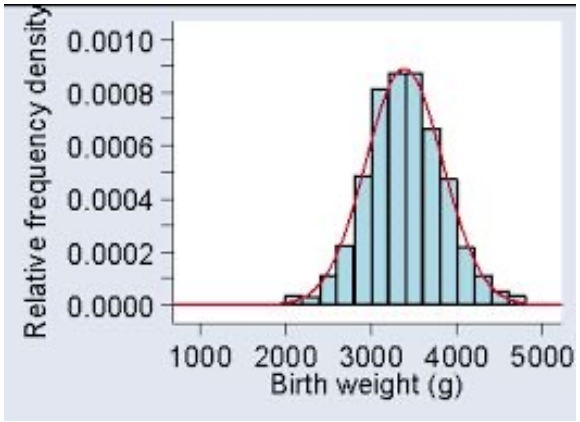
Things that follow Gaussian

Female shoe size





Birth Weight



(From https://studiousguy.com/real-life-examples-normal-distribution/)

numpy.random

numpy.random.normal

scale = $\sqrt{\sigma^2}$

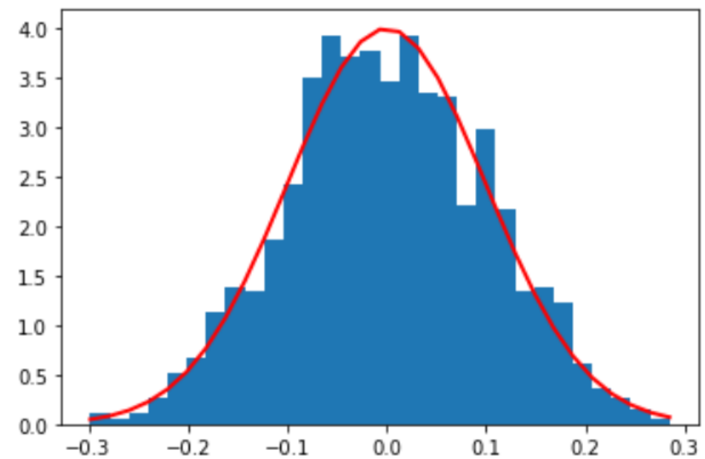
numpy.random.normal(*loc=0.0, scale=1.0, size=None***)**

Draw random samples from a normal (Gaussian) distribution.

Example Sample zero-mean gaussian with scale 0.1,

```
mu, sigma = 0, 0.1 # mean and standard deviation
X = np.random.normal(mu, sigma, 1000)
count, bins, ignored = plt.hist(X, 30, density=True)
plt.plot(bins, 1/(sigma * np.sqrt(2 * np.pi)) *
         np.exp( - (bins - mu)**2 / (2 * sigma**2) ),
         linewidth=2, color='r')
plt.show()
```

`bins`: length 31, consisting of boundary points



redline: PDF of gaussian distr.

Recap

Useful discrete distributions

- Bernoulli → “Coinflip Distribution”
- Binomial → Multiple Bernoulli draws

Continuous probability

- $P(X=x) = 0$ does not mean you won’t see x
- Probabilities assigned to *intervals* via CDF $P(X > x)$
- PDF measures probability *density* of single points $p(X=x) >= 0$

Useful continuous distributions

- Exponential → waiting time.
- Univariate / Multivariate Gaussian → Probably most ubiquitous distribution
- There are a lot more we will touch on later in the course…