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

1

1

Review: Random Variable Examples

2

X_1, X_2: outcomes of two dice

- $R_1 = X_1 + X_2$
- $R_2 = \frac{(X_1+X_2)}{2}$
- $R_3 = I\{X_1 = 1\}$

I: Indicator function

Random variable induces a partition of the outcome space!

$$\{R_3 = 1\} \Leftrightarrow \{(1,1), (1,2), \dots, (1,6)\}$$
$$\{R_3 = 0\} \Leftrightarrow \{(2,1), (2,2), \dots, (2,6), (3,1), (3,2), \dots, (3,6), \dots (6,1), (6,2), \dots, (6,6)\}$$

Q: what distribution does R_3 follow with what parameter?

Bernoulli, PMF: $p(X = x) = \pi^x (1-\pi)^{1-x}, \pi = \frac{1}{6}$

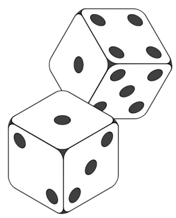
2

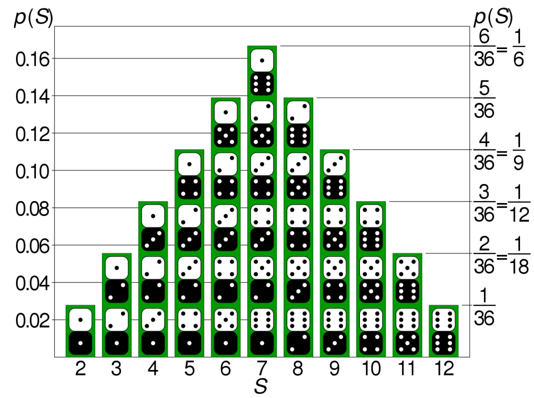
Review: Discrete Distribution

3

Another example.
- let S = sum of two dice;
- probability of S on different values:

$P(S = 2) = 1/36$
$P(S = 3) = 2/36$
$P(S = 4) = 3/36$
...
$P(S = 12) = 1/36$





PMF: $f_X(S) = \dfrac{\min(S-1, 13-S)}{36}$, for $S \in \{2,3,4,5,6,7,8,9,10,11,12\}$

3

Outline

4

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

4

Continuous Probability

5

(TV show spin the wheel)

5

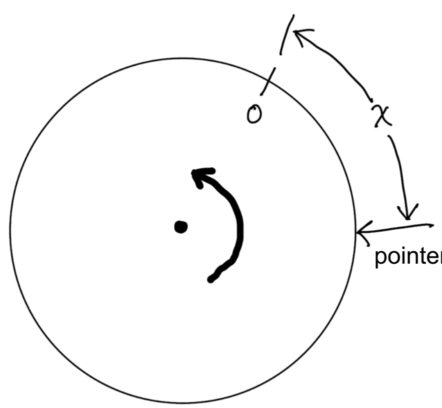
Continuous Probability

6

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≤ 0.5)

6

Continuous Probability

9

we could try to convince ourselves that it is sensible.

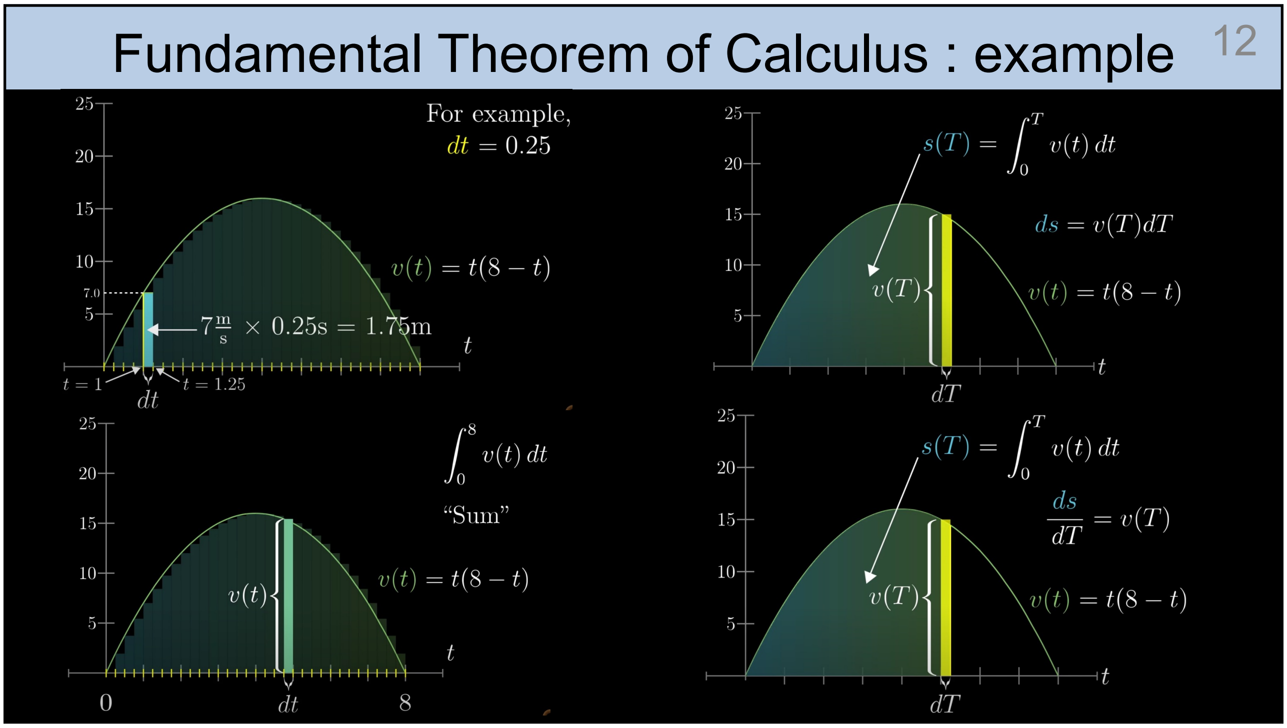
… or we could just accept this oddity…

9

Continuous Distributions

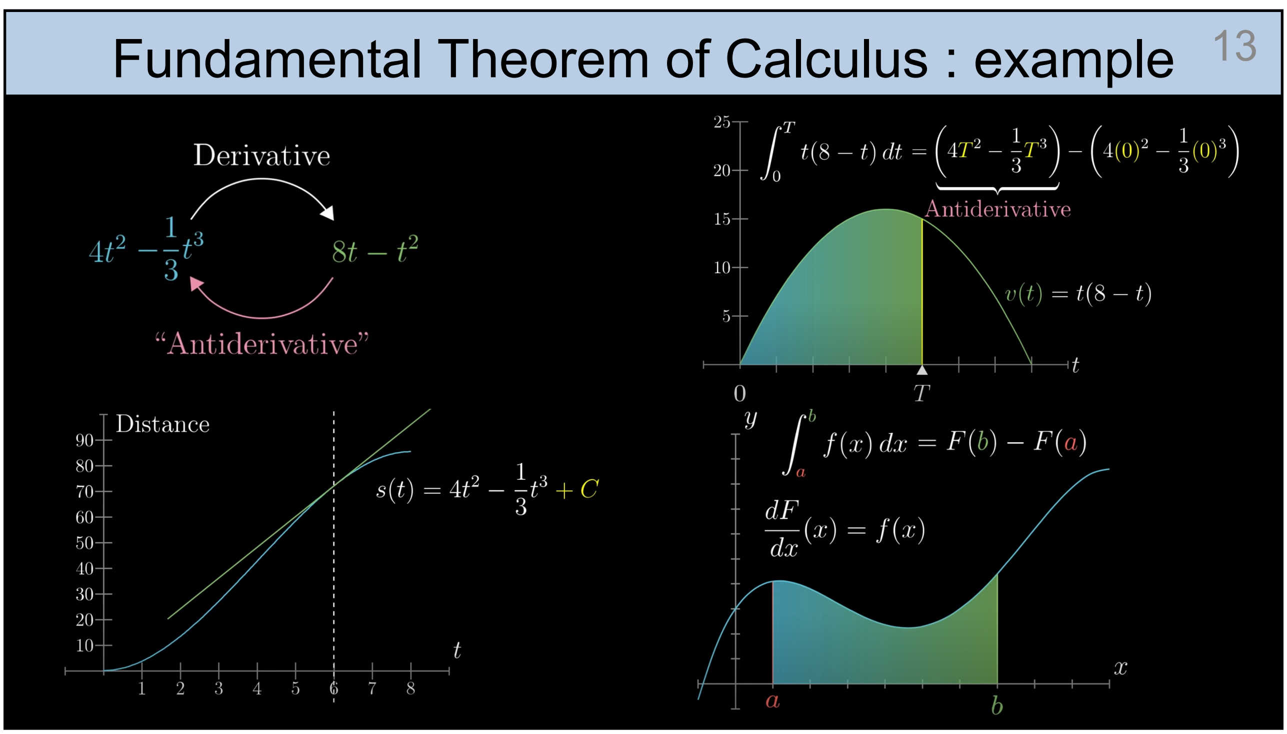
10

10



Fundamental Theorem of Calculus : example
12
For example,
$dt = 0.25$
$v(t) = t(8-t)$
$7\frac{\mathrm{m}}{\mathrm{s}} \times 0.25\mathrm{s} = 1.75\mathrm{m}$
$t = 1$
$t = 1.25$
dt
$\int_0^8 v(t)\,dt$
"Sum"
$v(t)$
$s(T) = \int_0^T v(t)\,dt$
$ds = v(T)dT$
$\frac{ds}{dT} = v(T)$
$v(T)$
dT

12

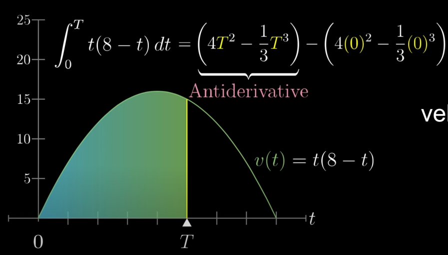


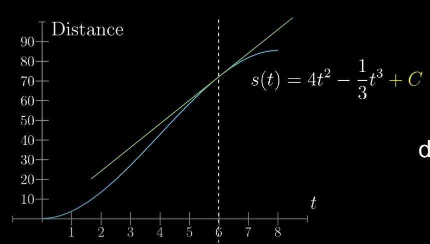
Fundamental Theorem of Calculus : example
13
Derivative
$4t^2 - \frac{1}{3}t^3$
$8t - t^2$
"Antiderivative"
Distance
$s(t) = 4t^2 - \frac{1}{3}t^3 + C$
$\int_0^T t(8-t)\,dt = \left(4T^2 - \frac{1}{3}T^3\right) - \left(4(0)^2 - \frac{1}{3}(0)^3\right)$
Antiderivative
$v(t) = t(8-t)$
$\int_a^b f(x)\,dx = F(b) - F(a)$
$\frac{dF}{dx}(x) = f(x)$

13

Mapping example to continuous probability

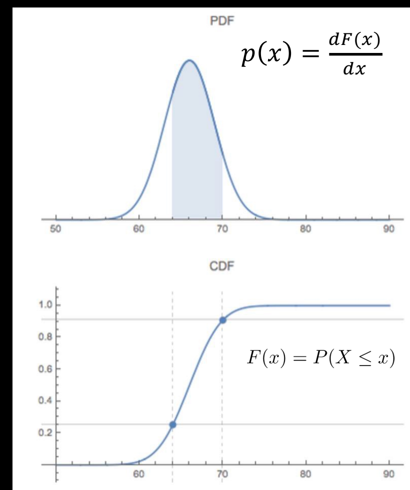
14





velocity → probability density

distance → cumulative distribution



14

Continuous Probability Distributions

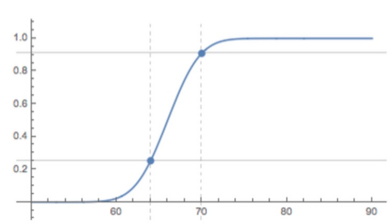
15

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

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

Key properties:

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



➢ 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/

15

Continuous Probability Distributions

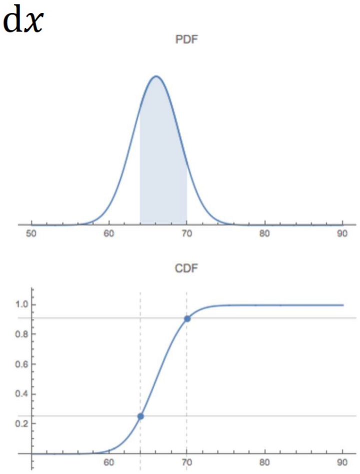
- If $F(X)$ is differentiable then,

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)**

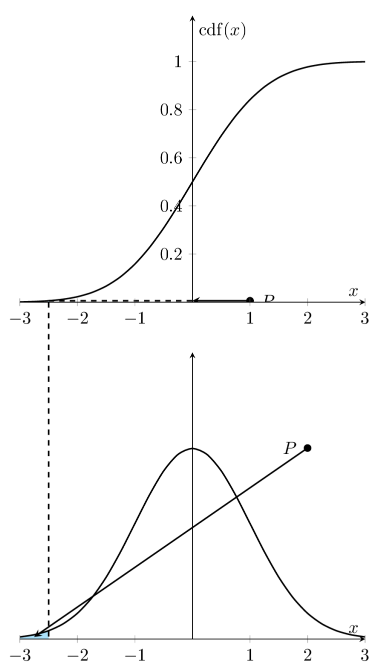
$\approx \frac{F(x)-F(x-\epsilon)}{x-(x-\epsilon)} = \frac{P(X\in(x-\epsilon,x])}{\epsilon}$ when $\epsilon \to 0$



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

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

16



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

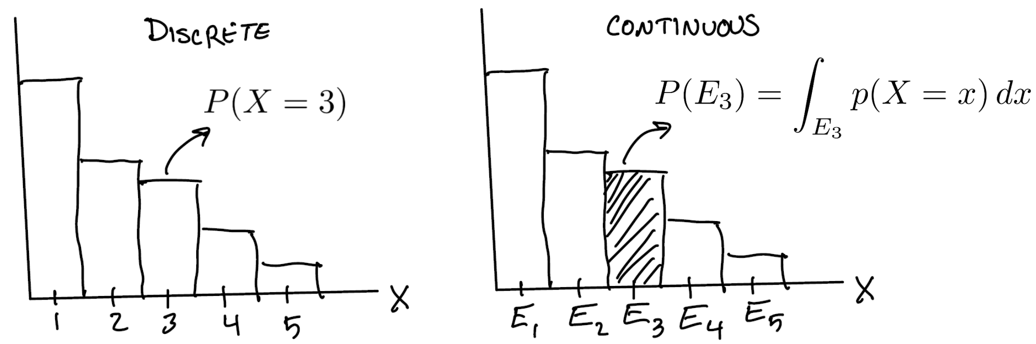
$$F(t) = \int_{-\infty}^{t} p(x)\,\mathrm{d}x$$

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

17

Continuous Probability

18



- 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* $p(X = x) > 0$

18

Notation

19

- 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"

19

Continuous Probability Distributions

20

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:

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

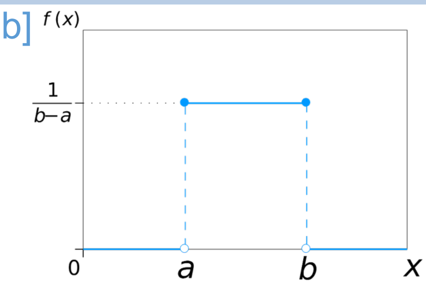
20

Uniform Continuous Distribution

21

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

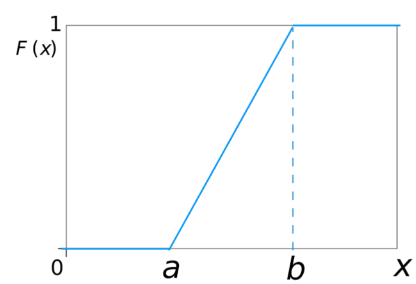
$$p(x) = \begin{cases} 0 & \text{if } x \leq a, \\ \frac{1}{b-a} & \text{if } a \leq x \leq b, \\ 0 & \text{if } b \leq x \end{cases} \qquad P(X \leq x) = \begin{cases} 0 & \text{if } x \leq a, \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b, \\ 1 & \text{if } b \leq x \end{cases}$$



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

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

Again, PDF function ≠ probability



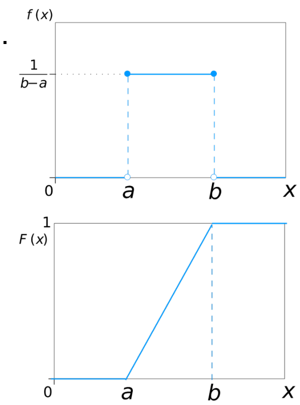
21

Uniform Continuous Distribution

22

Example: Let $X = \text{length}$ 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 $0 \le \mathrm{X} \le 23$.

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):

$$X \sim U(8,23)$$
$$\mathrm{p}(x) = \frac{1}{23-8} = \frac{1}{15}$$
$$\mathrm{P}(23 > x > 12)$$
$$= \frac{(23-12)}{15}$$
$$\approx 0.7333$$

Method 2 (bayes rule):

$$\mathrm{P}(x > 12 \mid x > 8)$$
$$= \frac{P(x > 12 \text{ and } x > 8)}{P(x > 8)} = \frac{P(x > 12)}{P(x > 8)}$$
$$= \frac{(23-12)\times\frac{1}{23}}{(23-8)\times\frac{1}{23}} \approx 0.7333$$

22

Uniform Continuous Distribution

23

numpy.random.uniform

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

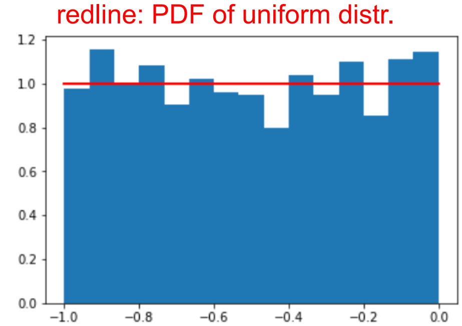
Draw samples from a uniform distribution.

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

23

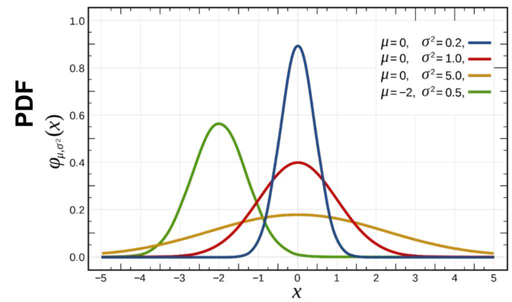
Gaussian/Normal Distribution

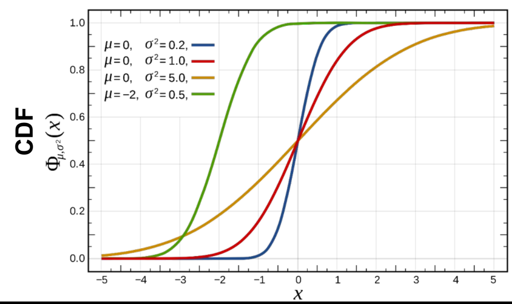
24

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





Observations:

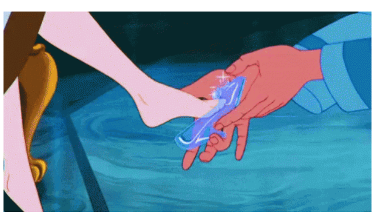
- Larger σ^2 : $p(x)$ more "spread out"
- Larger μ : $p(x)$'s center shifts to the right more

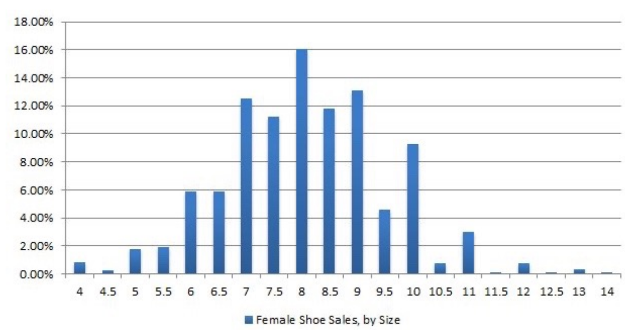
24

Things that follow Gaussian

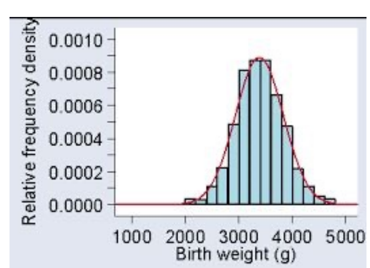
25

Female shoe size





Birth Weight



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

25

numpy.random

26

numpy.random.normal

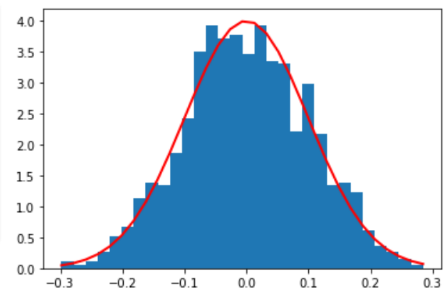
scale = $\sqrt{\sigma^2}$

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.

26

Recap

27

Useful discrete distributions

- Bernoulli → "Coinflip Distribution"
- Binomial → Multiple Bernoulli draws

Continuous probability

- P(X=x) = 0 <u>does not mean you won't see x</u>
- 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…

27