Continuous Probability

CS70: Discrete Mathematics and Probability Theory

UC Berkeley – Summer 2026

Lecture 23

Ref: Note 20

Outline

Starting two days of continuous probability (one set of slides)

Uniformly at Random in $[0,2]$

Goal: Choose a real number X, uniformly distributed in $[0,2]$

Checking intuition:

What is the probability that $X = 0.3$?

What is $\mathbb{P}[X \leq \frac{1}{2}]$?

What is $\mathbb{P}[X \in [0.7, 0.9]]$?

Discrete Uniform Distribution

What would a *discrete* uniform distribution on $[0,2]$ look like?

$X \sim \text{Uniform}(\{0.1, 0.3, \ldots, 1.7, 1.9\})$

X	Probability
0.1	0.1
0.3	0.1
0.5	0.1
0.7	0.1
0.9	0.1
1.1	0.1
1.3	0.1
1.5	0.1
1.7	0.1
1.9	0.1

Interval: $\mathbb{P}[0 \le X \le 1] = \mathbb{P}[X = 0.1] + \mathbb{P}[X = 0.3] + \cdots + \mathbb{P}[X = 0.9] = \frac{1}{2}$

What if we had 20 points between 0 and 1? ... "density" stays same

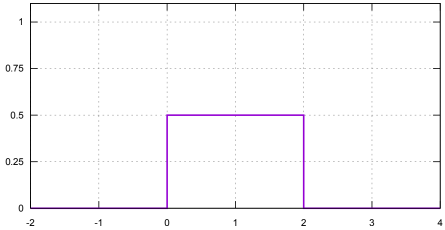
$\Rightarrow \quad \mathbb{P}[0 \le X \le 1] = \mathbb{P}[X = 0.05] + \mathbb{P}[X = 0.15] + \cdots + \mathbb{P}[X = 0.95] = \frac{1}{2}$

Idea: For finer-grained random variables, use intervals/density.

Continuous Uniform: Probability Density Function

Distribution is given by a *Probability Density Function*
(PDF or pdf or p.d.f.)

Density: Probability per unit length
Area under the curve (over a range) gives probability

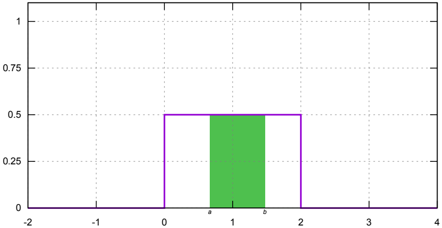


Total area under this curve? Width=2 and Height=0.5 $\rightarrow$ Area=1

Continuous Uniform: Probability Density Function

Distribution is given by a *Probability Density Function*
(PDF or pdf or p.d.f.)

Density: Probability per unit length
Area under the curve (over a range) gives probability



Total area under this curve? Width=2 and Height=0.5 $\rightarrow$ Area=1

Question: What is the area under this curve over $[a, b]$?

Area is $0.5 \cdot (b - a)$ (or $\frac{b-a}{2}$ as before)

Probability Density Function – Definition

Definition: A *Probability Density Function* (p.d.f.) for a real-valued random variable X is a function $f : \mathbb{R} \to \mathbb{R}$ such that

1. Non-negative: $f(x) \geq 0$ for all $x \in \mathbb{R}$
2. Total density one: $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$

Then

$$\mathbb{P}[a \leq X \leq b] = \int_a^b f(x)\,\mathrm{d}x \; .$$

Our "uniform over $[0,2]$" example has

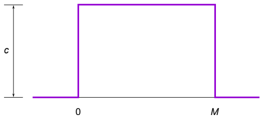
$$f(x) = \begin{cases} 0 & \text{if } x < 0; \\ \frac{1}{2} & \text{if } 0 \leq x \leq 2; \\ 0 & \text{if } x > 2. \end{cases}$$

Example probability calculation:

$$\mathbb{P}[0.7 \leq X \leq 0.9] = \int_{0.7}^{0.9} \tfrac{1}{2}\,\mathrm{d}x = \tfrac{1}{2}x\Big|_{0.7}^{0.9} = \tfrac{1}{2}(0.9 - 0.7) = 0.1$$

General Uniform

Uniform over $[0, M]$?



Uniform, so $f(x) = c$ when $0 \le x \le M$ (zero everywhere else)

Need

$$\int_0^M c\,\mathrm{d}x = 1, \quad \text{where} \quad \int_0^M c\,\mathrm{d}x = cx\Big|_0^M = c(M-0) = cM$$

So $cM = 1 \quad \Longrightarrow \quad c = \frac{1}{M}$

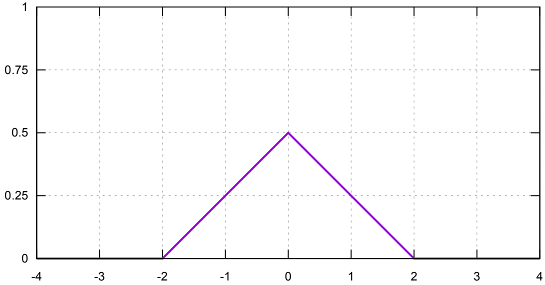
With $M = 2$, we have $f(x) = \frac{1}{2}$ for $0 \le x \le 2$ – as we had before

With $M = \frac{1}{2}$, we have $f(x) = 2$ for $0 \le x \le \frac{1}{2}$

Important: $f(x)$ is not a probability – it can be greater than 1!

Example 2: Non-Uniform

Consider this non-uniform density function:



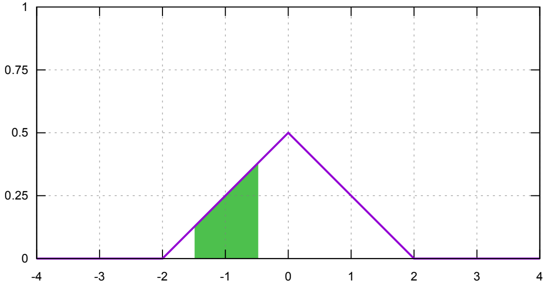
Is this a valid density function?

Area under $-2 \leq x \leq 0$ is a right triangle: $\frac{1}{2} \cdot w \cdot h = \frac{1}{2} \cdot 2 \cdot 0.5 = \frac{1}{2}$

Area under $0 \leq x \leq 2$ is the mirror image, so total area is 1. Good!

Example 2: Non-Uniform

Consider this non-uniform density function:



Is this a valid density function?

Area under $-2 \leq x \leq 0$ is a right triangle: $\frac{1}{2} \cdot w \cdot h = \frac{1}{2} \cdot 2 \cdot 0.5 = \frac{1}{2}$

Area under $0 \leq x \leq 2$ is the mirror image, so total area is 1. Good!

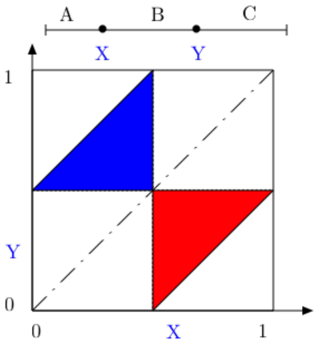
Interval probability example (right trapezoid):

$\mathbb{P}[-1.5 \leq X \leq -0.5] = \frac{0.125+0.375}{2} = 0.25$

Example Application: Breaking a Stick

Context: Pick two points X and Y independently uniformly on a stick. Break at those points.

Question: What is the probability you can make a triangle with the three pieces?



Can make a triangle iff

$$A < B + C, \quad B < A + C, \text{ and } C < A + B.$$

If $X < Y$, then $A = X$, $B = Y - X$, $C = 1 - Y$

$A < B + C \Leftrightarrow X < (Y - X) + (1 - Y) = 1 - X \Leftrightarrow X < \frac{1}{2}$

$\Rightarrow$ Overall: $X < \frac{1}{2}$, $Y < X + \frac{1}{2}$, $Y > \frac{1}{2}$.

$\Rightarrow$ This is the blue triangle.

If $X > Y$, get red triangle, by symmetry.

One fourth of each side $\Rightarrow$ $\mathbb{P}[\text{make triangle}] = 1/4$

Cumulative Distribution Function

Definition: For continuous random variable X, the *cumulative distribution function* (c.d.f.) is $F(x) = \mathbb{P}[X \leq x]$.

Relation between p.d.f. and c.d.f.:

$$F(x) = \mathbb{P}[X \leq x] = \int_{-\infty}^{x} f(z)\,\mathrm{d}z$$

and

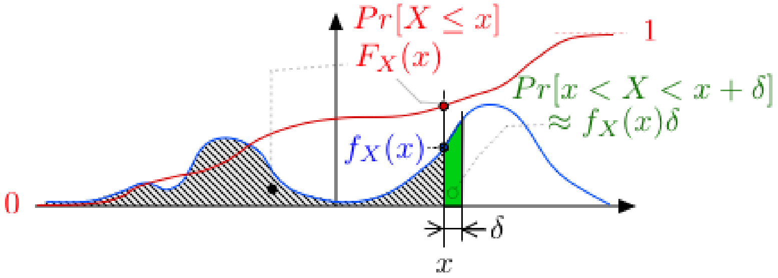
$$\underbrace{f(x) = \frac{\mathrm{d}F(x)}{\mathrm{d}x}}_{\text{Leibniz notation}} \quad \text{or} \quad \underbrace{f(x) = F'(x)}_{\text{Lagrange Notation}}$$

Convention: p.d.f. is lowercase ($f(x)$) and c.d.f. is uppercase ($F(x)$)

Can use subscripts for multiple random variables, say X and Y:
p.d.f. of X is $f_X(x)$, p.d.f. of Y is $f_Y(x)$, etc.

Some conditions:
$F(x)$ must be non-decreasing (since $f(x)$ is non-negative)
$F(x)$ goes from 0 (as $x \to -\infty$) to 1 (as $x \to +\infty$)

A Picture



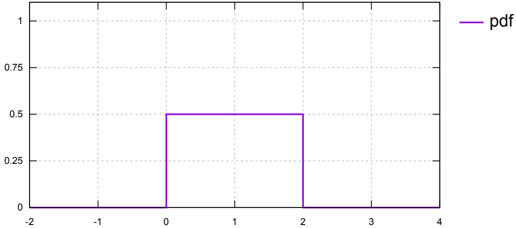
The pdf $f_X(x)$ is a non-negative function that integrates to 1.

The cdf $F_X(x)$ is the integral of f_X.

Area: $\mathbb{P}[x < X < x + \delta] \approx f_X(x)\delta$

As integral: $\mathbb{P}[X \leq x] = F_x(x) = \int_{-\infty}^{x} f_X(u)\,du$

Example 1: Uniform CDF

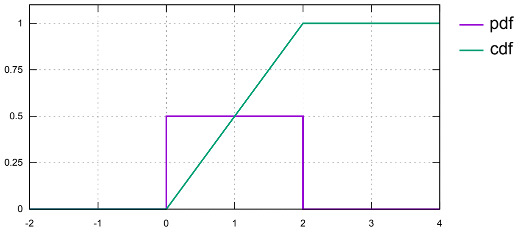


$$f(x) = \begin{cases} 0 & \text{if } x < 0; \\ \frac{1}{2} & \text{if } 0 \leq x \leq 2; \\ 0 & \text{if } x > 2. \end{cases}$$

When $x \in [0, 2]$:

$$F(x) = \int_{-\infty}^{x} f(z)\,dz = \int_{0}^{x} \frac{1}{2}\,dz = \frac{1}{2} \cdot z \Big|_{0}^{x} = \frac{1}{2}x - \frac{1}{2} \cdot 0 = \frac{1}{2}x.$$

Example 1: Uniform CDF



$$f(x) = \begin{cases} 0 & \text{if } x < 0; \\ \frac{1}{2} & \text{if } 0 \le x \le 2; \\ 0 & \text{if } x > 2. \end{cases}$$

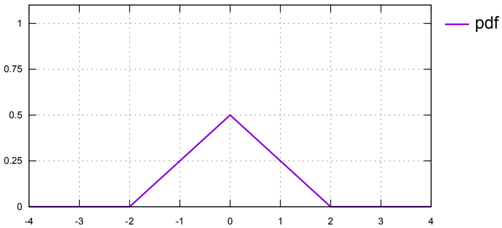
When $x \in [0, 2]$:

$$F(x) = \int_{-\infty}^{x} f(z)\,\mathrm{d}z = \int_{0}^{x} \frac{1}{2}\,\mathrm{d}z = \frac{1}{2} \cdot z\Big|_{0}^{x} = \frac{1}{2}x - \frac{1}{2} \cdot 0 = \frac{1}{2}x.$$

Full specification:

$$F(x) = \begin{cases} 0 & \text{if } x < 0; \\ \frac{1}{2}x & \text{if } 0 \le x \le 2; \\ 1 & \text{if } x > 2. \end{cases}$$

Example 2: Triangle CDF

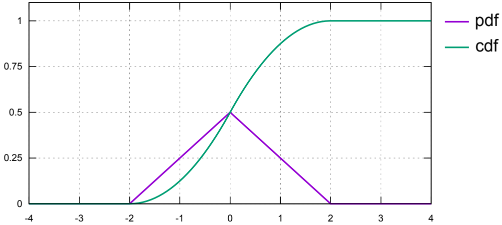


$$f(x) = \begin{cases} 0 & \text{if } x < -2; \\ \frac{1}{4}x + \frac{1}{2} & \text{if } -2 \le x < 0; \\ -\frac{1}{4}x + \frac{1}{2} & \text{if } 0 \le x < 2; \\ 0 & \text{if } x \ge 2. \end{cases}$$

When $x \in [-2, 0]$:

$$F(x) = \int_{-\infty}^{x} f(z)\,\mathrm{d}z = \int_{-2}^{x} \left(\frac{1}{4}z + \frac{1}{2}\right)\mathrm{d}z = \left(\frac{1}{8}z^2 + \frac{1}{2}z\right)\Big|_{-2}^{x} = \frac{1}{8}x^2 + \frac{1}{2}x + \frac{1}{2}.$$

Example 2: Triangle CDF



$$f(x) = \begin{cases} 0 & \text{if } x < -2; \\ \frac{1}{4}x + \frac{1}{2} & \text{if } -2 \le x < 0; \\ -\frac{1}{4}x + \frac{1}{2} & \text{if } 0 \le x < 2; \\ 0 & \text{if } x \ge 2. \end{cases}$$

When $x \in [-2, 0]$:

$$F(x) = \int_{-\infty}^{x} f(z)\,\mathrm{d}z = \int_{-2}^{x} \left(\frac{1}{4}z + \frac{1}{2}\right) \mathrm{d}z = \left(\frac{1}{8}z^2 + \frac{1}{2}z\right)\Big|_{-2}^{x} = \frac{1}{8}x^2 + \frac{1}{2}x + \frac{1}{2}.$$

Full specification:

$$F(x) = \begin{cases} 0 & \text{if } x < -2; \\ \frac{1}{8}x^2 + \frac{1}{2}x + \frac{1}{2} & \text{if } -2 \le x < 0; \\ -\frac{1}{8}x^2 + \frac{1}{2}x + \frac{1}{2} & \text{if } 0 \le x < 2; \\ 0 & \text{if } x \ge 2. \end{cases}$$

Concept Check

True or False for p.d.f. $f_X(x)$ and c.d.f. $F_X(x)$?

1. $F_X(x) = \int_{-\infty}^{\infty} f_X(y)\,\mathrm{d}y$

2. $\int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x = 1$

3. $F_X(x) = \int_{-\infty}^{x} f(y)\,\mathrm{d}y$

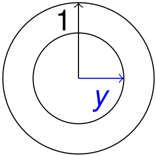
4. $f(x) = F'_X(x)$

5. $\int_{-\infty}^{\infty} F_X(x)\,\mathrm{d}x = 1$

Example 3: Dartboard

Setting: You have a dartboard with radius 1 and uniform probability of anywhere on dartboard.

Question: What is the probability you hit within distance y of center?



Check your intuition. Is this

(A) 1

(B) 0

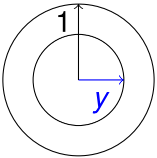
(C) $\int_0^y (2\pi y)\,dy$

(D) y^2

Example 3: Dartboard

Setting: You have a dartboard with radius 1 and uniform probability of anywhere on dartboard.

Question: What is the probability you hit within distance y of center?



Random Variable: Y distance from center.

$$\mathbb{P}[Y \leq y] =$$

Hence,

$$F_Y(y) = \mathbb{P}[Y \leq y] = \begin{cases} 0 & \text{for } y < 0 \\ y^2 & \text{for } 0 \leq y \leq 1 \\ 1 & \text{for } y > 1 \end{cases}$$

Example 3: Dartboard

Cumulative distribution function (c.d.f.):

$$F_Y(y) = \mathbb{P}[Y \le y] = \begin{cases} 0 & \text{for } y < 0 \\ y^2 & \text{for } 0 \le y \le 1 \\ 1 & \text{for } y > 1 \end{cases}$$

Bullseye is "within $\frac{1}{10}$ of the center": $\mathbb{P}[Y \le \frac{1}{10}] =$

Probability between .5 and .6 of center?

$$\begin{aligned} \mathbb{P}[0.5 < Y \le 0.6] &= \mathbb{P}[Y \le 0.6] - \mathbb{P}[Y \le 0.5] \\ &= F_Y(0.6) - F_Y(0.5) \\ &= .36 - .25 \\ &= .11 \end{aligned}$$

Example 3: Dartboard

Cumulative distribution function (c.d.f.):

$$F_Y(y) = \mathbb{P}[Y \leq y] = \begin{cases} 0 & \text{for } y < 0 \\ y^2 & \text{for } 0 \leq y \leq 1 \\ 1 & \text{for } y > 1 \end{cases}$$

What is probability density function (p.d.f.)? Take the derivative!

$$f_Y(y) = F_Y'(y) = \begin{cases} 0 & \text{for } y < 0 \\ 2y & \text{for } 0 \leq y \leq 1 \\ 0 & \text{for } y > 1 \end{cases}$$

The c.d.f. and p.d.f. both give full information.

$\Rightarrow$ Use whichever is convenient.

Expectation and Variance

Definition: The *expectation* of a continuous random variable X with pdf f is

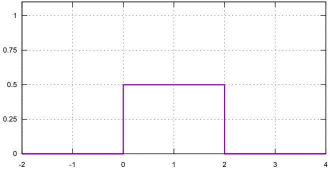
$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x\, f(x)\, \mathrm{d}x\ .$$

Definition: The *variance* of continuous random variable X with pdf f is

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\ &= \int_{-\infty}^{\infty} x^2 f(x)\, \mathrm{d}x - \left(\int_{-\infty}^{\infty} x\, f(x)\, \mathrm{d}x\right)^2\ .\end{aligned}$$

Example 1: Uniform

Recall: X uniform over $[0,2]$



$$f(x) = \begin{cases} 0 & \text{if } x < 0; \\ \frac{1}{2} & \text{if } 0 \le x \le 2; \\ 0 & \text{if } x > 2. \end{cases}$$

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x\, f(x)\, \mathrm{d}x = \int_0^2 \frac{1}{2} x\, \mathrm{d}x = \frac{1}{4} x^2 \Big|_0^2 = \frac{1}{4} 2^2 - \frac{1}{4} 0^2 = 1 \ .$$

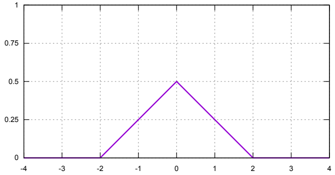
Middle of the range. Makes sense....

For variance:

$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2\, f(x)\, \mathrm{d}x = \int_0^2 \frac{1}{2} x^2\, \mathrm{d}x = \frac{1}{6} x^3 \Big|_0^2 = \frac{1}{6} 2^3 - \frac{1}{6} 0^3 = \frac{8}{6} = \frac{4}{3} \ .$$

So $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{4}{3} - 1 = \frac{1}{3} \quad ; \quad \sigma = \frac{1}{\sqrt{3}} \approx 0.577$

Example 2: Triangle



$$f(x) = \begin{cases} 0 & \text{if } x < -2; \\ \frac{1}{4}x + \frac{1}{2} & \text{if } -2 \leq x < 0; \\ -\frac{1}{4}x + \frac{1}{2} & \text{if } 0 \leq x < 2; \\ 0 & \text{if } x \geq 2. \end{cases}$$

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x\, f(x)\,\mathrm{d}x = \int_{-2}^{0} \left(\tfrac{1}{4}x^2 + \tfrac{1}{2}x\right)\mathrm{d}x + \int_{0}^{2} \left(-\tfrac{1}{4}x^2 + \tfrac{1}{2}x\right)\mathrm{d}x$$

$$= \left(\tfrac{1}{12}x^3 + \tfrac{1}{4}x^2\right)\Big|_{-2}^{0} + \left(-\tfrac{1}{12}x^3 + \tfrac{1}{4}x^2\right)\Big|_{0}^{2} = -\left(\tfrac{1}{12}\cdot(-8) + \tfrac{1}{4}\cdot 4\right) + \left(-\tfrac{1}{12}\cdot 8 + \tfrac{1}{4}\cdot 4\right) = 0$$

Symmetric around zero, so this makes sense.

$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2\, f(x)\,\mathrm{d}x = \int_{-2}^{0} \left(\tfrac{1}{4}x^3 + \tfrac{1}{2}x^2\right)\mathrm{d}x + \int_{0}^{2} \left(-\tfrac{1}{4}x^3 + \tfrac{1}{2}x^2\right)\mathrm{d}x$$

$$= \left(\tfrac{1}{16}x^4 + \tfrac{1}{6}x^3\right)\Big|_{-2}^{0} + \left(-\tfrac{1}{16}x^4 + \tfrac{1}{6}x^3\right)\Big|_{0}^{2} = -\left(\tfrac{1}{16}\cdot 16 + \tfrac{1}{6}\cdot(-8)\right) + \left(-\tfrac{1}{16}\cdot 16 + \tfrac{1}{6}\cdot 8\right)$$

$$= -(1 - \tfrac{4}{3}) + (-1 + \tfrac{4}{3}) = \tfrac{1}{3} + \tfrac{1}{3} = \tfrac{2}{3}\ .$$

So $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{2}{3} - 0 = \frac{2}{3}$; $\sigma = \sqrt{\frac{2}{3}} \approx 0.8165$

Problem: Expected Squared Distance

Problem 1: Independently pick points X and Y uniformly from $[0,1]$.
What is $\mathbb{E}[(X-Y)^2]$?

Analysis: Can calculate $\mathbb{E}[X] = \frac{1}{2}$ and $\mathbb{E}[X^2] = \frac{1}{3}$. Then

$$\begin{aligned}
\mathbb{E}[(X-Y)^2] &= \mathbb{E}[X^2 + Y^2 - 2XY] \\
&= \mathbb{E}[X^2] + \mathbb{E}[Y^2] - 2\mathbb{E}[X]\mathbb{E}[Y] \\
&= \frac{1}{3} + \frac{1}{3} - 2 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}.
\end{aligned}$$

Problem 2: What about in a unit square?

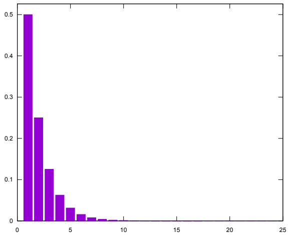
Analysis:
$$\begin{aligned}
\mathbb{E}[\| \mathbf{X} - \mathbf{Y} \|^2] &= \mathbb{E}[(X_1 - Y_1)^2] + \mathbb{E}[(X_2 - Y_2)^2] \\
&= 2 \times \frac{1}{6} = \frac{2}{6}.
\end{aligned}$$

Problem 3: What about in n dimensions? $\frac{n}{6}$.

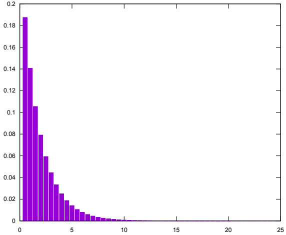
Geometric Distribution Review

Recall $X \sim \text{Geometric}(p) \quad \Longrightarrow \quad \mathbb{P}[X = k] = (1-p)^{k-1}p$

Time unit δ with $p = \lambda\delta$: $\quad \mathbb{P}[X = t \text{ sec}] = \mathbb{P}[X = t/\delta] = \delta\lambda(1-\delta\lambda)^{t/\delta-1}$



Time unit: "1 second"
$X \sim \text{Geometric}(1/2)$
$\mathbb{E}[X] = 2$ (2 seconds)
$\mathbb{P}[X > t \text{ sec}] = (1/2)^t$



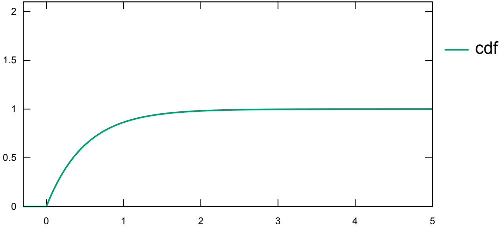
Time unit: "0.5 second"
$X \sim \text{Geometric}(1/4)$
$\mathbb{E}[X] = 4$ (2 seconds)
$\mathbb{P}[X > t \text{ sec}] = (3/4)^{2t} = (9/16)^t$

In the limit? $\quad \lim_{\delta\to 0} \mathbb{P}[X_\lambda > t \text{ sec}] = \lim_{\delta\to 0}(1-\lambda\delta)^{\lambda t/(\lambda\delta)} = e^{-\lambda t}$

$\Rightarrow$ Proposed c.d.f.: $F_X(t) = \mathbb{P}[X \le t] = 1 - \mathbb{P}[X > t] = 1 - e^{-\lambda t}$

Exponential Distribution

Proposed c.d.f.: $F(x) = 1 - e^{-\lambda x}$ for $x \geq 0$

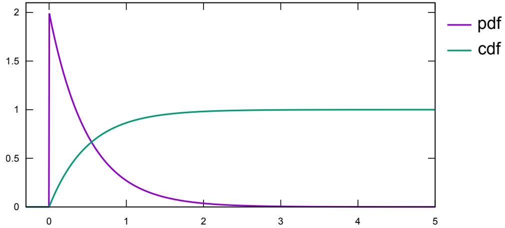


Question: Is this a valid c.d.f.?

Probability Density Function (p.d.f.)? Take the derivative!

Exponential Distribution

Proposed c.d.f.: $F(x) = 1 - e^{-\lambda x}$ for $x \geq 0$



Question: Is this a valid c.d.f.?

Starts at 0, limit as $x \to \infty$ is 1, always non-decreasing... ✓

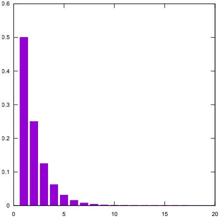
Probability Density Function (p.d.f.)?

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0; \\ 0 & \text{otherwise} \end{cases}$$

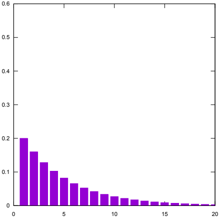
This is the *exponential distribution*, and we write $X \sim \text{Exp}(\lambda)$

Exponential Distribution: Comparison with Geometric

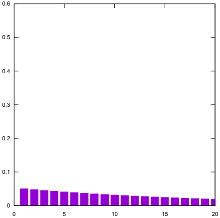
Geometric:



p=1/2

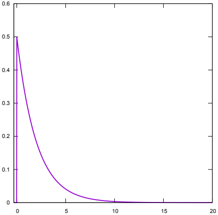


p=1/5

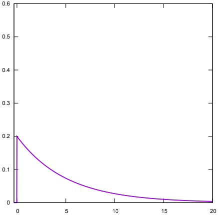


p=1/20

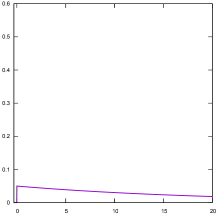
Exponential:



λ=1/2



λ=1/5



λ=1/20

Exponential Distribution: Expectation

Reference: $X \sim \text{Exp}(\lambda)$; $f_X(x) = \lambda e^{-\lambda x}$ for $x \geq 0$ (0 otherwise)

Calculus review: the integration by parts formula is

$$\int_a^b u(x) \cdot v'(x)\,\mathrm{d}x = u(x) \cdot v(x)\Big|_a^b - \int_a^b u'(x) \cdot v(x)\,\mathrm{d}x$$

Using this:

$$\begin{aligned}
\mathbb{E}[X] = \int_0^\infty \underbrace{x}_{u} \cdot \underbrace{\lambda e^{-\lambda x}}_{v'(x)} \mathrm{d}x &= x \cdot (-e^{-\lambda x})\Big|_0^\infty - \int_0^\infty 1 \cdot (-e^{-\lambda x})\,\mathrm{d}x \\
&= \qquad 0 \quad - \quad \frac{e^{-\lambda x}}{\lambda}\Big|_0^\infty \\
&= \frac{1}{\lambda}
\end{aligned}$$

Can also use integration by parts to calculate (see notes): $\text{Var}(X) = \frac{1}{\lambda^2}$

Problem: Minimum of Two Exponentials

Context: Independent random variables $X \sim \text{Exp}(\lambda)$ and $Y \sim \text{Exp}(\mu)$

Question: Let $Z = \min(X, Y)$; What is the distribution of Z?

For $z \geq 0$, we have $Z > z$ iff $X > z$ and $Y > z$, so

$$\mathbb{P}[Z > z] = \mathbb{P}[X > z, Y > z] = \mathbb{P}[X > z] \cdot \mathbb{P}[Y > z] = e^{-\lambda z} e^{-\mu z} = e^{-(\lambda+\mu)z}$$

Therefore,

$$F_Z(z) = \mathbb{P}[Z \leq z] = 1 - e^{-(\lambda+\mu)z},$$

but this is just the c.d.f. of an exponential random variable with parameter $\lambda + \mu$

Therefore, $Z \sim \text{Exp}(\lambda + \mu)$.

Problem: Maximum of Two Exponentials

Context: Independent random variables $X \sim \text{Exp}(\lambda)$ and $Y \sim \text{Exp}(\mu)$

Question: Let $Z = \max(X, Y)$; What is $\mathbb{E}[Z]$?

Strategy: Compute $F_Z(z)$, then $f_Z(z)$, then $\mathbb{E}[Z]$.

For $F_Z(z)$ (with $z \geq 0$), we have $Z \leq z$ iff $X \leq z$ and $Y \leq z$, so

$$\begin{aligned} F_Z(z) = \mathbb{P}[Z \leq z] = \mathbb{P}[X \leq z, Y \leq z] &= \mathbb{P}[X \leq z] \cdot \mathbb{P}[Y \leq z] \\ = (1 - e^{-\lambda z})(1 - e^{-\mu z}) &= 1 - e^{-\lambda z} - e^{-\mu z} + e^{-(\lambda+\mu)z} \end{aligned}$$

Thus, $f_Z(z) = F_Z'(z) = \lambda e^{-\lambda z} + \mu e^{-\mu z} - (\lambda + \mu) e^{-(\lambda+\mu)z}$.

Since, $\int_0^\infty x \lambda e^{-\lambda x} \, dx = \mathbb{E}[X] = \frac{1}{\lambda}$, generalizing we get

$$\mathbb{E}[Z] = \int_0^\infty z f_Z(z) \, dz = \frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda + \mu}.$$

Exponential Distribution: Some Properties

1. Exp(λ) is memoryless. Let $X \sim \text{Exp}(\lambda)$. Then, for $s, t > 0$,

$$\begin{aligned}
\mathbb{P}[X > t+s \mid X > s] &= \frac{\mathbb{P}[X > t+s]}{\mathbb{P}[X > s]} \\
&= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t} \\
&= \mathbb{P}[X > t].
\end{aligned}$$

2. Scaling Exp(λ). Let $X \sim \text{Exp}(\lambda)$ and $Y = aX$ for some $a > 0$. Then

$$\mathbb{P}[Y > t] = \mathbb{P}[aX > t] = \mathbb{P}[X > t/a] = e^{-\lambda(t/a)} = e^{-(\lambda/a)t}$$

Thus $X \sim \text{Exp}(\lambda) \implies aX \sim \text{Exp}(\lambda/a)$.

So any exponential random variable is a scaling of Exp(1):

$$\text{Exp}(\lambda) = \tfrac{1}{\lambda}\text{Exp}(1)$$

Multiple Continuous Random Variables

Concepts related to multiple discrete random variables:

- Joint distribution
- Marginal distribution
- Conditional distribution
- Independence
- Total probability

All of these concepts carry over to continuous random variables (with "distribution" replaced by "density") in the obvious way.

Explore these on the following slides...

Joint Density

For continuous random variables X and Y, the joint density function is $F_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$ with:

$$\mathbb{P}[a \le X \le b, c \le Y \le d] = \int_c^d \int_a^b f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

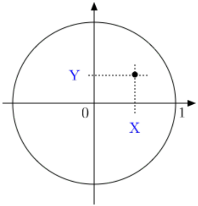
Example: Choose a point (X, Y) uniformly in the set $A \subseteq \mathbb{R}^2$. Then

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{\text{Area}(A)} & \text{if } (x,y) \in A; \\ 0 & \text{otherwise.} \end{cases}$$

Discrete Version: Think of (X, Y) as being discrete on a grid with mesh size ε and $\mathbb{P}[X = m\varepsilon, Y = n\varepsilon] = f_{X,Y}(m\varepsilon, n\varepsilon)\varepsilon^2$.

Example: Uniform in the Unit Circle

Pick a point (X, Y) uniformly in the unit circle.



Uniform in area π, so

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{\pi} & \text{if } x^2 + y^2 \leq 1; \\ 0 & \text{otherwise.} \end{cases}$$

So:

$\mathbb{P}[X > 0, Y > 0] =$

$\mathbb{P}[X < 0, Y > 0] =$

$\mathbb{P}[X^2 + Y^2 \leq r^2] =$

$\mathbb{P}[X > Y] =$

Marginal Density

Given joint density function $f_{X,Y}(x,y)$ the marginal density for X (and similarly for Y) is defined as

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y.$$

Example: $f_{X,Y}(x,y)$ is $\frac{1}{4}$ on the rectangle $([0,2],[0,2])$, so for $x \in [0,2]$:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y = \int_0^2 \frac{1}{4}\,\mathrm{d}y = \frac{1}{4}y\Big|_0^2 = \frac{2}{4} = \frac{1}{2}.$$

$\Rightarrow$ Single-variable uniform on $[0,2]$.

Unit circle: For $x \in [0,1]$, the y values on the unit circle are $\pm\sqrt{1-x^2}$.

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}\,\mathrm{d}y = \frac{1}{\pi}\cdot x\Big|_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} = \frac{2\sqrt{1-x^2}}{\pi}$$

$\Rightarrow$ Proportional to the height of circle at x.

Conditional Density

Discrete case: Random variables X and Y and sets of values A and B:

Conditional Probability: $\mathbb{P}[X \in A \,|\, Y \in B] = \frac{\mathbb{P}[X \in A, Y \in B]}{\mathbb{P}[Y \in B]}$

Generalize to intervals for continuous random variables X and Y:

$$\mathbb{P}[X \in [x, x+\mathrm{d}x] \,|\, Y \in [y, y+\mathrm{d}y]] = \frac{f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y}{f_Y(y)\,\mathrm{d}y}$$

Or in the limit:

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_{X,Y}(x,y)}{\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,\mathrm{d}x}$$

Independent Continuous Random Variables

Recall: Discrete RVs X and Y independent if and only if for all A, B,

$$\mathbb{P}[X \in A, Y \in B] = \mathbb{P}[X \in A] \cdot \mathbb{P}[Y \in B].$$

Continuous version: Random variables X and Y independent if and only if, for all $a \le b$ and $c \le d$

$$\mathbb{P}[a \le X \le b, c \le Y \le d] = \mathbb{P}[a \le X \le b] \cdot \mathbb{P}[c \le Y \le d].$$

Theorem: Continuous RVs X and Y independent if and only if

$$f_{X,Y}(x,y) = f_X(x) f_Y(y).$$

Note: $f_X(x)$ (resp. $f_Y(y)$) is marginal distribution of X (resp. Y).

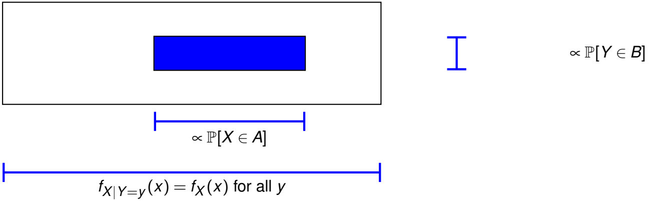
Proof: Intervals: $A = [x, x + \mathrm{d}x]$, $B = [y, y + \mathrm{d}y]$.

$$\begin{aligned} Pr[X \in A, Y \in B] &= \mathbb{P}[X \in A] \cdot \mathbb{P}[Y \in B] \\ &\approx f_X(x)\,\mathrm{d}x \cdot f_Y(y)\,\mathrm{d}y \\ &= f_X(x) f_Y(y)\ \mathrm{d}x\,\mathrm{d}y. \end{aligned}$$

Thus, $f_{X,Y}(x,y) = f_X(x) f_Y(y)$. $\square$

Independent Random Variables

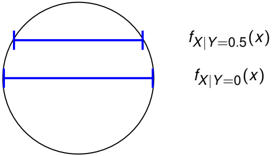
Uniform on a rectangle? Independent?



$\mathbb{P}[X \in A, Y \in B] \propto$ Area of rectangle $\propto \mathbb{P}[X \in A] \cdot \mathbb{P}[Y \in B]$.

$\Rightarrow$ Independent!

Uniform on a circle? Independent?



Not independent!

Total Probability

Discrete version:

$$\mathbb{P}[A] = \sum_x \mathbb{P}[A \mid X = x] \cdot \mathbb{P}[X = x]$$

Continuous version:

$$\mathbb{P}[A] = \int_{-\infty}^{\infty} \mathbb{P}[A \mid X = x] \, f_X(x) \, \mathrm{d}x$$

With event A being $y \leq Y \leq y + \mathrm{d}y$:

$$f_Y(y)\,\mathrm{d}y \approx \mathbb{P}[y \leq Y \leq y + \mathrm{d}y] \approx \int_{-\infty}^{\infty} (f_{Y|X=x}(y)\,\mathrm{d}y) f_X(x)\,\mathrm{d}x$$

Or in the limit:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{Y|X=x}(y) \, f_X(x) \, \mathrm{d}x$$

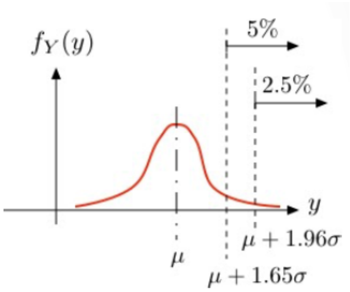
Relation to marginal and conditional probabilities:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{Y|X=x}(y) \, f_X(x) \, \mathrm{d}x = \int_{-\infty}^{\infty} \frac{f_{X,Y}(x,y)}{f_X(x)} \, f_X(x) \, \mathrm{d}x = \int_{-\infty}^{\infty} f_{X,Y}(x,y) \, \mathrm{d}x$$

One More Distribution: Normal (Gaussian)

For any μ and σ, a **normal** (aka **Gaussian**) random variable Y, which we write as $Y \sim N(\mu, \sigma^2)$, has pdf

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2}.$$



Note: $\mathbb{P}[|Y - \mu| > 1.65\sigma] \approx 10\%$; $\mathbb{P}[|Y - \mu| > 2\sigma] \approx 5\%$

Standard Normal: Expectation

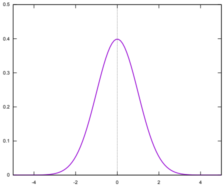
Standard normal ($N(0,1)$) has $\mu = 0$ and $\sigma = 1$. If $X \sim N(0,1)$,

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \quad \rightarrow \quad f_X(x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}$$

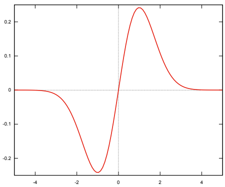
Expected value?

$$\mathbb{E}[X] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x\, e^{-x^2/2}\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}} \left(\underbrace{\int_{-\infty}^{0} x\, e^{-x^2/2}\,\mathrm{d}x}_{A} + \underbrace{\int_{0}^{\infty} x\, e^{-x^2/2}\,\mathrm{d}x}_{B} \right)$$

Distribution is symmetric around $x = 0$



Times x makes flipped mirror images



Integrals A and B cancel out, giving $\mathbb{E}[X] = 0$.

Standard Normal: Variance

If $X \sim N(0,1)$,

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mathbb{E}[X^2]$ (since we just calculated $\mathbb{E}[X] = 0$)

$$\mathbb{E}[X^2] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 \cdot e^{-x^2/2}\, \mathrm{d}x$$

Integrate by parts with $u(x) = x$ and $v'(x) = xe^{-x^2/2}$

$$= \frac{1}{\sqrt{2\pi}} (x)(-e^{-x^2/2}) \Big|_{-\infty}^{\infty} - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (1) \cdot (-e^{-x^2/2})\, \mathrm{d}x$$

$$= \frac{1}{\sqrt{2\pi}} \cdot \frac{-x}{e^{x^2/2}} \Big|_{-\infty}^{\infty} \qquad + \int_{-\infty}^{\infty} f_X(x)\, \mathrm{d}x$$

The first term is 0 (limit at both ends is 0 – more formally, use l'Hôpital)

The second term is 1 (the integral of the pdf)

Conclusion: $\mathrm{Var}(X) = 1$

Shifting and Scaling a Continuous RV

Let X be *any* continuous random variable, and let $Y = a + bX$

$\Rightarrow$ Y is scaled by b and shifted by a

What is the c.d.f. of Y?

$$F_Y(y) = \mathbb{P}[Y \le y] = \mathbb{P}[a + bX \le y] = \mathbb{P}[X \le \tfrac{y-a}{b}] = F_X(\tfrac{y-a}{b})$$

What is the p.d.f. of Y?

$$f_Y(y) = \frac{\mathrm{d}F_Y(y)}{\mathrm{d}y} = \frac{\mathrm{d}F_X(\frac{y-a}{b}))}{\mathrm{d}y} = \frac{1}{b} \cdot f_X(\tfrac{y-a}{b})$$

From basic properties of expectation and variance:

$\mathbb{E}[Y] = \mathbb{E}[a + bX] = a + b\mathbb{E}[X]$

$\mathrm{Var}(Y) = \mathrm{Var}(a + bX) = \mathrm{Var}(bX) = b^2\mathrm{Var}(X)$

Shifting and Scaling

Summary: Let X be *any* continuous random variable, and let $Y = a + bX$

$f_Y(y) = \frac{1}{b} \cdot f_X(\frac{y-a}{b})$

$F_Y(y) = F_X(\frac{y-a}{b})$

$\mathbb{E}[Y] = a + b\mathbb{E}[X]$

$\mathrm{Var}(Y) = b^2 \mathrm{Var}(X)$

Let $X \sim N(0,1)$ and let $Y = \mu + \sigma X$ (so use above with $a = \mu$ and $b = \sigma$)

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \Rightarrow f_Y(y) = \frac{1}{\sigma} \cdot \frac{1}{\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)}$$

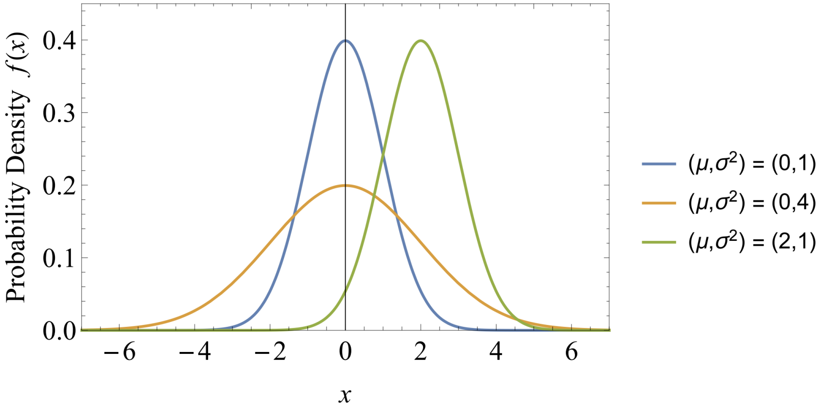
So Y is Gaussian! Specifically, $Y \sim N(\mu, \sigma)$

We already saw that $\mathbb{E}[X] = 0$ and $\mathrm{Var}(X) = 1$, giving the following.

Theorem: Let $X \sim N(0,1)$ and $Y = \mu + \sigma X$. Then $Y = N(\mu, \sigma^2)$, $\mathbb{E}[Y] = \mu$ and $\mathrm{Var}(Y) = \sigma^2$.

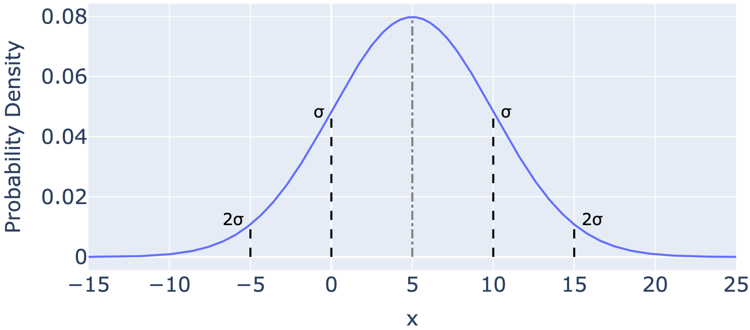
Visualizing the Gaussian

Gaussian with different μ and σ:



Visualizing the Gaussian

For $X \sim N(5, 25)$, what does the PDF look like?



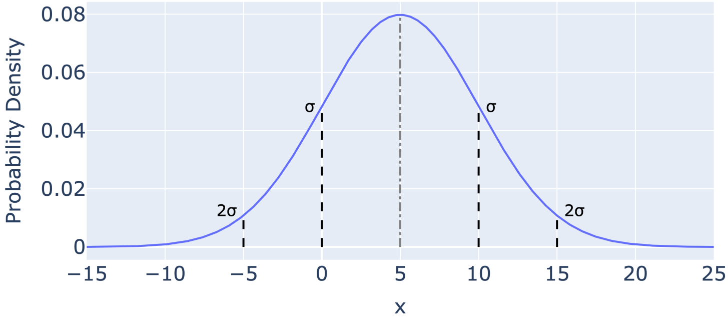
Visualizing the Gaussian

For $X \sim N(5, 25)$, what is $\mathbb{P}[X \leq 0]$?

$$\int_{-\infty}^{0} \frac{1}{\sqrt{50\pi}} e^{(x-5)^2/50} \approx 0.1587$$

Note that this is also $\mathbb{P}[Y \leq -1]$ for $Y \sim N(0, 1)$.

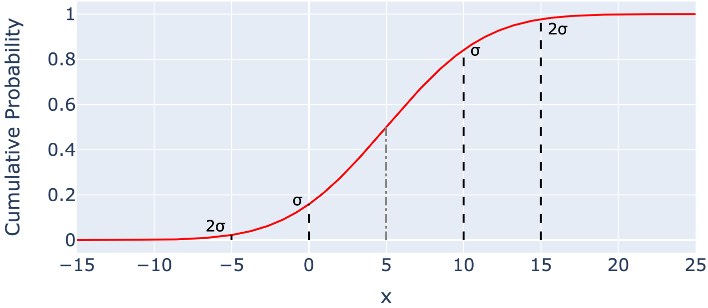
$\Rightarrow$ Important point: "one standard deviation below the mean"



Visualizing the Gaussian

For $X \sim N(5,25)$, the CDF is given below.

- $\Rightarrow$ Note: There is no closed-form expression for the CDF.
- $\Rightarrow$ As on prev slide: $\mathbb{P}[X \leq 0] \approx 0.1587$
- $\Rightarrow$ Other quantities: $F_X(\mu) = 0.5$, $F_X(\mu + 2\sigma) \approx 0.977$
- $\Rightarrow$ Distance: $\mathbb{P}[|X - \mu| \leq 2\sigma] = F_X(X + 2\mu) - F_X(x - 2\mu) \approx 0.9545$



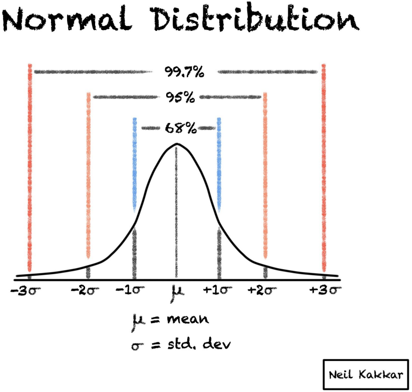
The 68–95–99.7 Rule

Memorable rule: 68–95–99.7:

68% of mass is within in 1 std-dev

95% of mass is within in 2 std-dev

99.7% of mass is within in 3 std-dev



https://www.freecodecamp.org/news/normal-distribution-explained/

“Z-score” is the number of standard deviations you are from the mean.

Review: Law of Large Numbers

Theorem: Independent identically distributed random variables, X_i,

$A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Each X_i, has $\mu = \mathbb{E}[X_i]$ and $\sigma^2 = \text{Var}(X_i)$.

Mean of A_n is μ, and variance is σ^2/n.

Using Chebyshev:

$$\mathbb{P}[|A_n - \mu| > \varepsilon] \leq \frac{\text{Var}(A_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon} \to 0$$

Central Limit Theorem

Central Limit Theorem: Let $X_1, X_2, \ldots$ be i.i.d. with $\mathbb{E}[X_1] = \mu$ and $\text{Var}(X_1) = \sigma^2$. Define

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then

$$S_n \to N(0,1) \quad \text{as} \quad n \to \infty.$$

That is,

$$\mathbb{P}[S_n \leq \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2}\,\mathrm{d}x.$$

Note:

$$\mathbb{E}[S_n] = \frac{\mathbb{E}[A_n] - \mu}{\sigma/\sqrt{n}} = 0$$

$$\text{Var}(S_n) = \frac{\text{Var}(A_n)}{\sigma^2/n} = 1$$

Central Limit Theorem: Example

Fixed points: 1000 schools randomly return assignments and report how many students got their own.

$X_1, \ldots, X_{1000}$: numbers reported from 1000 schools
 Think: 1000 independent experiments with same μ and σ
 We get $S = X_1 + X_2 + \cdots + X_{1000} = 1231$ – how likely is this?

From previous lectures: $\mathbb{E}[X_i] = 1$ and $\text{Var}(X_i) = 1$
 $\Rightarrow$ $\mathbb{E}[S] = 1000$ and $\text{Var}(S) = 1000$ (and so $\sigma = \sqrt{1000} \approx 31.6$)
 $\Rightarrow$ Then $1231 \approx \mu + \frac{231}{31.6}\sigma \approx \mu + 7.31\sigma$
 $\Rightarrow$ Z-score is $+7.31$

Remember the 68-95-99.7 rule: probability of Z-score of > 3 is 0.3%
 $\Rightarrow$ Z-score of $+7.31$ is *extraordinarily* unlikely
 $\Rightarrow$ Specifically: probability is 1.33×10^{-13}

Want to learn more? Learn about p-values.

Confidence Intervals

For iid X_i with $\mu = \mathbb{E}[X_i]$ and $\mathrm{Var}(X_i) = \sigma^2$, estimate $\mathbb{E}[X_i]$ with $A_n = \frac{1}{n} X_i$.

$\Rightarrow$ Mean of A_n is μ, and variance is σ^2/n.

$\Rightarrow$ Chebyshev gives $\mathbb{P}[|A_n - \mu| > \varepsilon] \le \frac{\mathrm{Var}(A_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$

Using Chebyshev, to get confidence $1-\delta$ we need $\frac{\sigma^2}{n\varepsilon^2} \le \delta$ or $n \ge \frac{\sigma^2}{\varepsilon^2}\frac{1}{\delta}$

Central Limit Thm: $\mathbb{P}[|A_n - \mu| > \varepsilon] \le C \int_{x \ge \varepsilon}^{\infty} e^{-\frac{x^2}{2\mathrm{Var}(A_n)}}\,dx \le C\, e^{-\frac{\varepsilon^2}{2\mathrm{Var}(A_n)}}$

for $\varepsilon > \sqrt{\mathrm{Var}(A_n)}$ (C is roughly $2/\sqrt{2\pi}$)

Implies to get confidence $1 - C\delta$ we need

$$e^{-\frac{\varepsilon^2}{2\,\mathrm{Var}(A_n)}} = e^{-\left(\frac{\varepsilon^2}{2}\frac{n}{\sigma^2}\right)} \le \delta \quad \Longrightarrow \quad -\frac{n\varepsilon^2}{2\sigma^2} \le \ln\delta \quad \Longrightarrow \quad n \ge \frac{\sigma^2}{\varepsilon^2}\left(2\ln\frac{1}{\delta}\right).$$

Comparing for $\delta = .05$: $\frac{1}{\delta} = 20$ and $2\ln\frac{1}{\delta} = 6$.

Summary

Two days on continuous probability

Continuous Random Variables

- Idea and Probability Density
- Cumulative Density Function
- Examples

Important Distribution 1: Exponential

Probability Concepts in a Continuous Setting

- Expectation and Variance
- Joint and Marginal Distributions
- Conditional distribution, Independence, and Total Probability

Important Distribution 2: Normal (Gaussian)

- Definition
- Shifting and Scaling
- Standard deviation bands
- Central Limit Theorem