CSE 312

# Foundations of Computing II

## Lecture 17: The Normal Distribution

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

**Aleks Jovcic**

Slide Credit: Based on Stefano Tessaro's slides for 312 19au
incorporating ideas from Anna Karlin, Alex Tsun, Rachel Lin, Hunter Schafer & myself ☺

# Agenda

- Law of Total Expectation (LTE) Practice
- Chebyshev's Inequality
- The Normal Distribution
- Practice with Normals

# Example: Flipping Coins

LTE: $E[X] = \sum_{i=1}^{n} E[X|A_i] \mathbb{P}(A_i)$

$Y = 2$

Suppose wanted to analyze flipping a random number of coins. Suppose someone gave us $Y \sim Poi(5)$ fair coins and we wanted to compute the expected number of heads $X$ from flipping those coins.

$E[X] = 0.5 \cdot$

$E[X] =$

$E[X|Y=2] = 0.5 \cdot 2$

# Agenda

- Law of Total Expectation (LTE) Practice
- Chebyshev's Inequality
- The Normal Distribution
- Practice with Normals

# Using variance

- If we know more about the random variable, e.g. its variance, we can get a better bound!

# Chebyshev's Inequality

Markov's inequality
$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t}$.

**Theorem.** Let $X$ be a random variable. Then, for any $t > 0$,

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq t) \leq \frac{\mathrm{Var}(X)}{t^2}.$$

**Proof:** Define $Z = X - \mathbb{E}(X)$

Definition of Variance

$$\mathbb{P}(|Z| \geq t) = \mathbb{P}(Z^2 \geq t^2) \leq \frac{\mathbb{E}(Z^2)}{t^2} = \frac{\mathrm{Var}(X)}{t^2}$$

$|Z| \geq t$ iff $Z^2 \geq t^2$

Markov's inequality ($Z^2 \geq 0$)

# Example – Binomial Random Variable

Chebychev's Inequality
$\mathbb{P}(|X - \mathbb{E}(X)| \geq t) \leq \frac{\mathrm{Var}(X)}{t^2}$.

Let $X$ be Binomial RV with parameters. $n, p = 0.5$

$$\mathbb{E}(X) = \frac{n}{2} \qquad Var(X) =$$

*What is the probability that* $X \geq \frac{3n}{4}$?

Chebychev's inequality: $\mathbb{P}\left(X \geq \frac{3n}{4}\right) \leq$

Markov's inequality: $\mathbb{P}\left(X \geq \frac{3n}{4}\right) \leq \frac{4}{3n} \cdot \frac{n}{2} = \frac{2}{3}$

# Tail Bounds

Useful for approximations of complex systems. How good the approximation is depends on the actual distribution and the context you are using it in.

- Usually loose upper-bounds are okay when designing for worst-case

Generally, the more you know about your random variable the better tail bounds you can get.

# Agenda

- Law of Total Expectation (LTE) Practice
- Chebyshev's Inequality
- The Normal Distribution
- Practice with Normals

# Review – Continuous RVs

**Probability Density Function (PDF).**

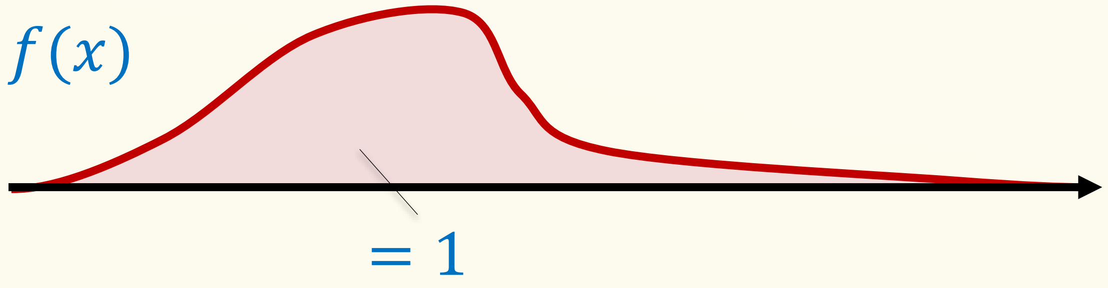
$f: \mathbb{R} \to \mathbb{R}$ s.t.

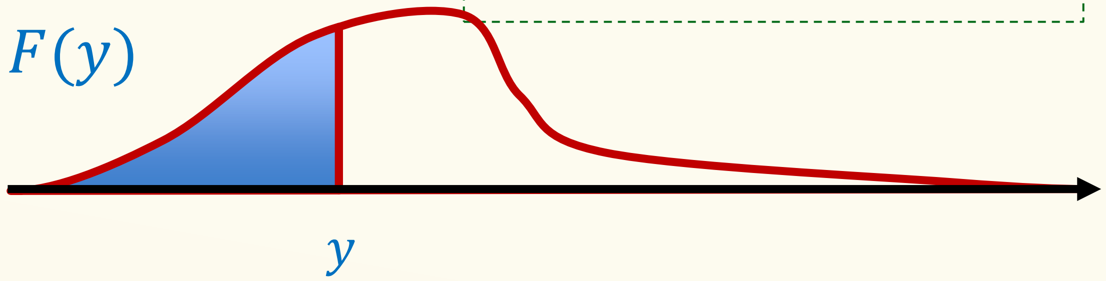
- $f(x) \geq 0$ for all $x \in \mathbb{R}$
- $\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1$

**Cumulative Density Function (CDF).**

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

**Theorem.** $f(x) = \frac{dF(x)}{dx}$

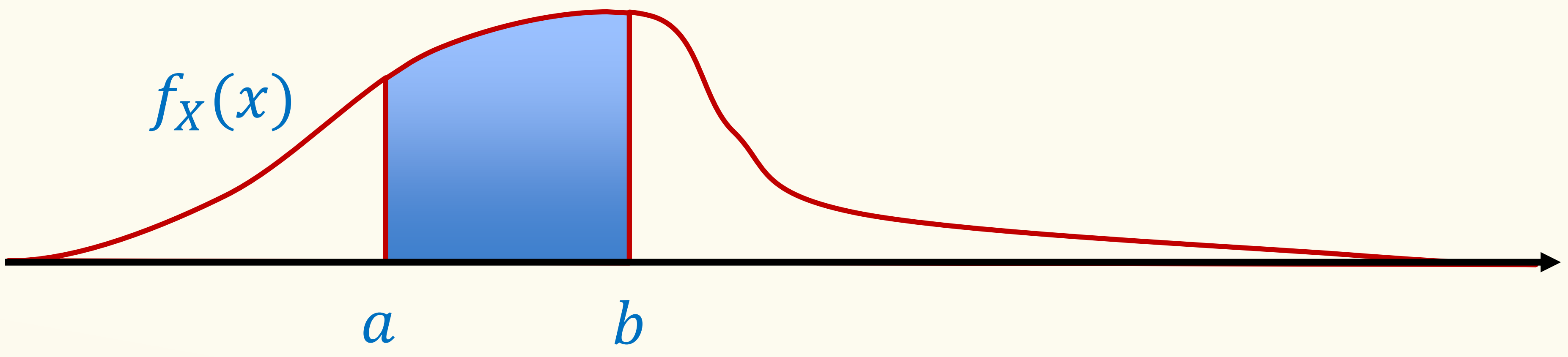



Density ≠ Probability !

$F(y) = \mathbb{P}(X \leq y)$

# Review – Continuous RVs




$$\mathbb{P}(X \in [a, b]) = \int_a^b f_X(x)\mathrm{d}x = F_X(b) - F_X(a)$$

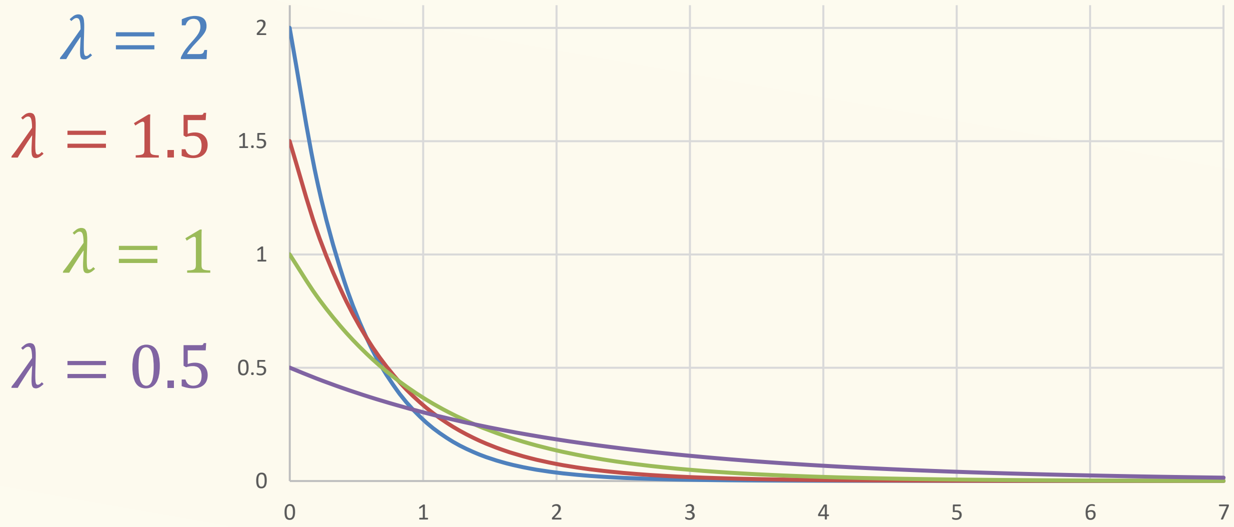
# Exponential Distribution

**Definition.** An **exponential random variable** $X$ with parameter $\lambda \geq 0$ is follows the exponential density

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

We write $X \sim \mathrm{Exp}(\lambda)$ and say $X$ that follows the exponential distribution.

CDF: For $y \geq 0$,
$$F_X(y) = 1 - e^{-\lambda y}$$



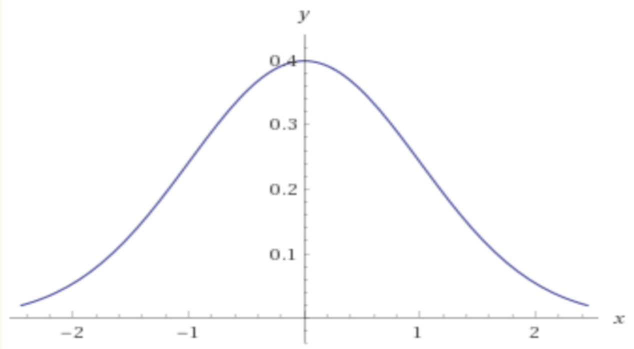
# The Normal Distribution

**Carl Friedrich Gauss**

**Definition.** A **Gaussian (or normal) random variable** with parameters $\mu \in \mathbb{R}$ and $\sigma \geq 0$ has density

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

(We say that $X$ follows the Normal Distribution, and write $X \sim \mathcal{N}(\mu, \sigma^2)$)



# The Normal Distribution

**Definition.** A **Gaussian (or normal) random variable** with parameters $\mu \in \mathbb{R}$ and $\sigma \geq 0$ has density

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

(We say that $X$ follows the Normal Distribution, and write $X \sim \mathcal{N}(\mu, \sigma^2)$)

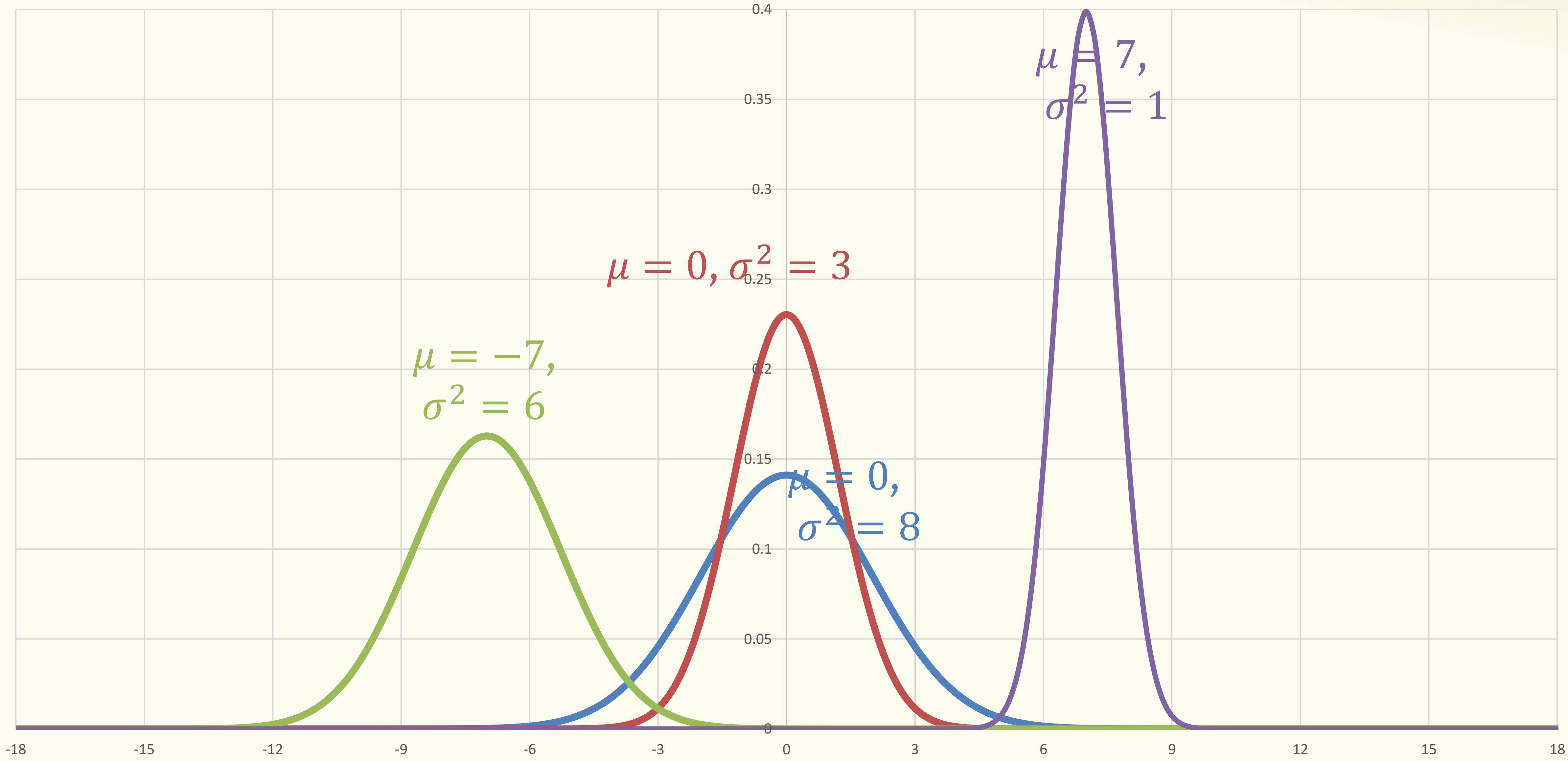
**Fact.** If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $\mathbb{E}(X) = \mu$, and $\mathrm{Var}(X) = \sigma^2$

Expectation follows from density being symmetric around $\mu$, $f_X(\mu - x) = f_X(\mu + x)$

Carl Friedrich Gauss

# The Normal Distribution

Aka a "Bell Curve" (imprecise name)



# Shifting and Scaling – turning one normal dist into another

**Fact.** If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

**Proof.** $\mathbb{E}(Y) = a\,\mathbb{E}(X) + b = a\mu + b$

$\mathrm{Var}(Y) = a^2\,\mathrm{Var}(X) = a^2\sigma^2$

Can show with algebra that the PDF of $Y = aX + b$ is still normal.

Note: $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$

# CDF of normal distribution

**Fact.** If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

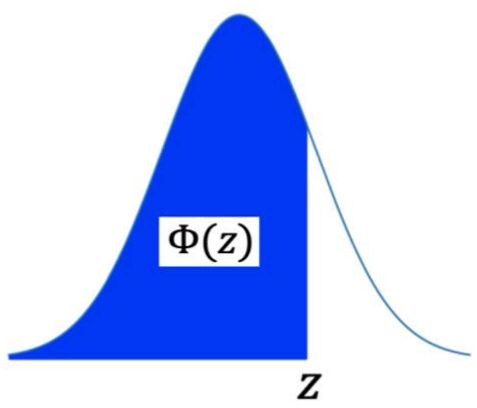
**Standard (unit) normal** $\boldsymbol{Z} \sim \mathcal{N}(0, 1)$

**CDF.** $\Phi(z) = \mathbb{P}(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} \mathrm{d}x$ for $Z \sim \mathcal{N}(0, 1)$

Note: $\Phi(z)$ has no closed form – generally given via tables

# Table of $\Phi(z)$ CDF of Standard Normal Distn

Make sure to use the one linked on the site!




$\Phi$ Table: $\mathbb{P}(Z \leq z)$ when $Z \sim \mathcal{N}(0,1)$

| $z$ | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.5 | 0.50399 | 0.50798 | 0.51197 | 0.51595 | 0.51994 | 0.52392 | 0.5279 | 0.53188 | 0.53586 |
| 0.1 | 0.53983 | 0.5438 | 0.54776 | 0.55172 | 0.55567 | 0.55962 | 0.56356 | 0.56749 | 0.57142 | 0.57535 |
| 0.2 | 0.57926 | 0.58317 | 0.58706 | 0.59095 | 0.59483 | 0.59871 | 0.60257 | 0.60642 | 0.61026 | 0.61409 |
| 0.3 | 0.61791 | 0.62172 | 0.62552 | 0.6293 | 0.63307 | 0.63683 | 0.64058 | 0.64431 | 0.64803 | 0.65173 |
| 0.4 | 0.65542 | 0.6591 | 0.66276 | 0.6664 | 0.67003 | 0.67364 | 0.67724 | 0.68082 | 0.68439 | 0.68793 |
| 0.5 | 0.69146 | 0.69497 | 0.69847 | 0.70194 | 0.7054 | 0.70884 | 0.71226 | 0.71566 | 0.71904 | 0.7224 |
| 0.6 | 0.72575 | 0.72907 | 0.73237 | 0.73565 | 0.73891 | 0.74215 | 0.74537 | 0.74857 | 0.75175 | 0.7549 |
| 0.7 | 0.75804 | 0.76115 | 0.76424 | 0.7673 | 0.77035 | 0.77337 | 0.77637 | 0.77935 | 0.7823 | 0.78524 |
| 0.8 | 0.78814 | 0.79103 | 0.79389 | 0.79673 | 0.79955 | 0.80234 | 0.80511 | 0.80785 | 0.81057 | 0.81327 |
| 0.9 | 0.81594 | 0.81859 | 0.82121 | 0.82381 | 0.82639 | 0.82894 | 0.83147 | 0.83398 | 0.83646 | 0.83891 |
| 1.0 | 0.84134 | 0.84375 | 0.84614 | 0.84849 | 0.85083 | 0.85314 | 0.85543 | 0.85769 | 0.85993 | 0.86214 |
| 1.1 | 0.86433 | 0.8665 | 0.86864 | 0.87076 | 0.87286 | 0.87493 | 0.87698 | 0.879 | 0.881 | 0.88298 |
| 1.2 | 0.88493 | 0.88686 | 0.88877 | 0.89065 | 0.89251 | 0.89435 | 0.89617 | 0.89796 | 0.89973 | 0.90147 |
| 1.3 | 0.9032 | 0.9049 | 0.90658 | 0.90824 | 0.90988 | 0.91149 | 0.91309 | 0.91466 | 0.91621 | 0.91774 |
| 1.4 | 0.91924 | 0.92073 | 0.9222 | 0.92364 | 0.92507 | 0.92647 | 0.92785 | 0.92922 | 0.93056 | 0.93189 |
| 1.5 | 0.93319 | 0.93448 | 0.93574 | 0.93699 | 0.93822 | 0.93943 | 0.94062 | 0.94179 | 0.94295 | 0.94408 |
| 1.6 | 0.9452 | 0.9463 | 0.94738 | 0.94845 | 0.9495 | 0.95053 | 0.95154 | 0.95254 | 0.95352 | 0.95449 |
| 1.7 | 0.95543 | 0.95637 | 0.95728 | 0.95818 | 0.95907 | 0.95994 | 0.9608 | 0.96164 | 0.96246 | 0.96327 |
| 1.8 | 0.96407 | 0.96485 | 0.96562 | 0.96638 | 0.96712 | 0.96784 | 0.96856 | 0.96926 | 0.96995 | 0.97062 |
| 1.9 | 0.97128 | 0.97193 | 0.97257 | 0.9732 | 0.97381 | 0.97441 | 0.975 | 0.97558 | 0.97615 | 0.9767 |
| 2.0 | 0.97725 | 0.97778 | 0.97831 | 0.97882 | 0.97932 | 0.97982 | 0.9803 | 0.98077 | 0.98124 | 0.98169 |
| 2.1 | 0.98214 | 0.98257 | 0.983 | 0.98341 | 0.98382 | 0.98422 | 0.98461 | 0.985 | 0.98537 | 0.98574 |
| 2.2 | 0.9861 | 0.98645 | 0.98679 | 0.98713 | 0.98745 | 0.98778 | 0.98809 | 0.9884 | 0.9887 | 0.98899 |
| 2.3 | 0.98928 | 0.98956 | 0.98983 | 0.9901 | 0.99036 | 0.99061 | 0.99086 | 0.99111 | 0.99134 | 0.99158 |
| 2.4 | 0.9918 | 0.99202 | 0.99224 | 0.99245 | 0.99266 | 0.99286 | 0.99305 | 0.99324 | 0.99343 | 0.99361 |
| 2.5 | 0.99379 | 0.99396 | 0.99413 | 0.9943 | 0.99446 | 0.99461 | 0.99477 | 0.99492 | 0.99506 | 0.9952 |
| 2.6 | 0.99534 | 0.99547 | 0.9956 | 0.99573 | 0.99585 | 0.99598 | 0.99609 | 0.99621 | 0.99632 | 0.99643 |
| 2.7 | 0.99653 | 0.99664 | 0.99674 | 0.99683 | 0.99693 | 0.99702 | 0.99711 | 0.9972 | 0.99728 | 0.99736 |
| 2.8 | 0.99744 | 0.99752 | 0.9976 | 0.99767 | 0.99774 | 0.99781 | 0.99788 | 0.99795 | 0.99801 | 0.99807 |
| 2.9 | 0.99813 | 0.99819 | 0.99825 | 0.99831 | 0.99836 | 0.99841 | 0.99846 | 0.99851 | 0.99856 | 0.99861 |
| 3.0 | 0.99865 | 0.99869 | 0.99874 | 0.99878 | 0.99882 | 0.99886 | 0.99889 | 0.99893 | 0.99896 | 0.999 |

# Example

Let $X \sim \mathcal{N}(0.4, 4)$.

$$\mathbb{P}(X \leq 1.2)$$

# Example

Let $X \sim \mathcal{N}(0.4, 4 = 2^2)$.

$$\mathbb{P}(X \le 1.2) = \mathbb{P}\left(\frac{X - 0.4}{2} \le \frac{1.2 - 0.4}{2}\right)$$

$$= \mathbb{P}\left(\frac{X - 0.4}{2} \le 0.4\right) = \Phi(0.4) \approx 0.6554$$

$\sim \mathcal{N}(0, 1)$

| | | |
|---|---|---|
| 0.1 | 0.5398 | 0.5438 |
| 0.2 | 0.5793 | 0.5832 |
| 0.3 | 0.6179 | 0.6217 |
| 0.4 | 0.6554 | 0.6591 |
| 0.5 | 0.6915 | 0.6950 |
| 0.6 | 0.7257 | 0.7291 |
| 0.7 | 0.7580 | 0.7611 |

# Example

Let $X \sim \mathcal{N}(3, 16)$.

$$\mathbb{P}(2 < X < 5)$$

# Example

Let $X \sim \mathcal{N}(3, 16)$.

$$\mathbb{P}(2 < X < 5) = \mathbb{P}\left(\frac{2 \ - 3}{4} < \frac{X - 3}{4} < \frac{5 \ - 3}{4}\right)$$

$$= \mathbb{P}\left(-\frac{1}{4} < Z < \frac{1}{2}\right)$$

$$= \Phi\left(\frac{1}{2}\right) - \Phi\left(-\frac{1}{4}\right)$$

$$= \Phi\left(\frac{1}{2}\right) - \left(1 - \ \Phi\left(\frac{1}{4}\right)\right) \ \approx 0.29017$$

# Example – Off by Standard Deviations

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

$$\mathbb{P}(|X - \mu| < k\sigma) =$$

# Example – Off by Standard Deviations

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

$$\mathbb{P}(|X-\mu| < k\sigma) = \mathbb{P}\left(\frac{|X-\mu|}{\sigma} < k\right) =$$

$$= \mathbb{P}\left(-k < \frac{X-\mu}{\sigma} < k\right) = \Phi(k) - \Phi(-k)$$

e.g. $k = 1$: 68%, $k = 2$: 95%, $k = 3$: 99%

# Summary of procedure for doing calculations with normal r.v.

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$

Therefore,

$$F_X(z) = \mathbb{P}(X \leq z) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \leq \frac{z-\mu}{\sigma}\right) = \Phi\left(\frac{z-\mu}{\sigma}\right)$$

# CDF of normal distribution

**Fact.** If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

**Standard (unit) normal** $\boldsymbol{Z} \sim \mathcal{N}(0,1)$

**CDF.** $\Phi(z) = \mathbb{P}(Z \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} \mathrm{d}x$ for $Z \sim \mathcal{N}(0,1)$

Note: $\Phi(z)$ has no closed form – generally given via tables

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $F_X(z) = \mathbb{P}(X \le z) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \le \frac{z-\mu}{\sigma}\right) = \Phi(\frac{z-\mu}{\sigma})$

# Closure of the normal -- under addition


**Fact.** If $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ (both independent normal RV) then $aX + bY + c \sim \mathcal{N}(a\mu_X + b\mu_Y + c, a^2\sigma_X^2 + b^2\sigma_Y^2)$

Note: The special thing is that the sum of normal **RVs is still a normal RV.**

The values of the expectation and variance is not surprising.

- Linearity of expectation (always true)
- When $X$ and $Y$ are independent, $Var(aX + bY) = a^2Var(X) + b^2Var(Y)$