

Normal distribution + Central Limit Theorem

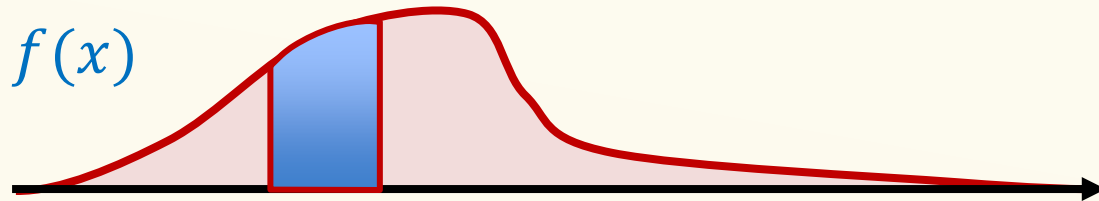
CSE 312 Spring 26
Lecture 17

Recap – Continuous RVs

Probability Density Function (PDF).

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

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



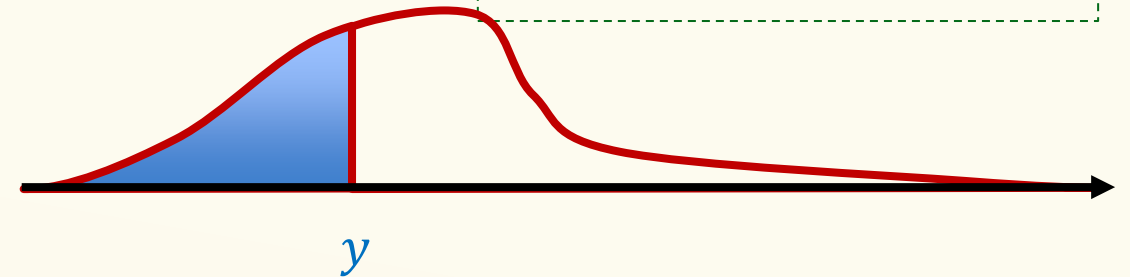
Density \neq Probability !

$$\begin{aligned} P(X \in [a, b]) &= \int_a^b f_X(x) dx \\ &= F_X(b) - F_X(a) \end{aligned}$$

Cumulative Distribution Function (CDF).

$$F(y) = \int_{-\infty}^y f(x) dx$$

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



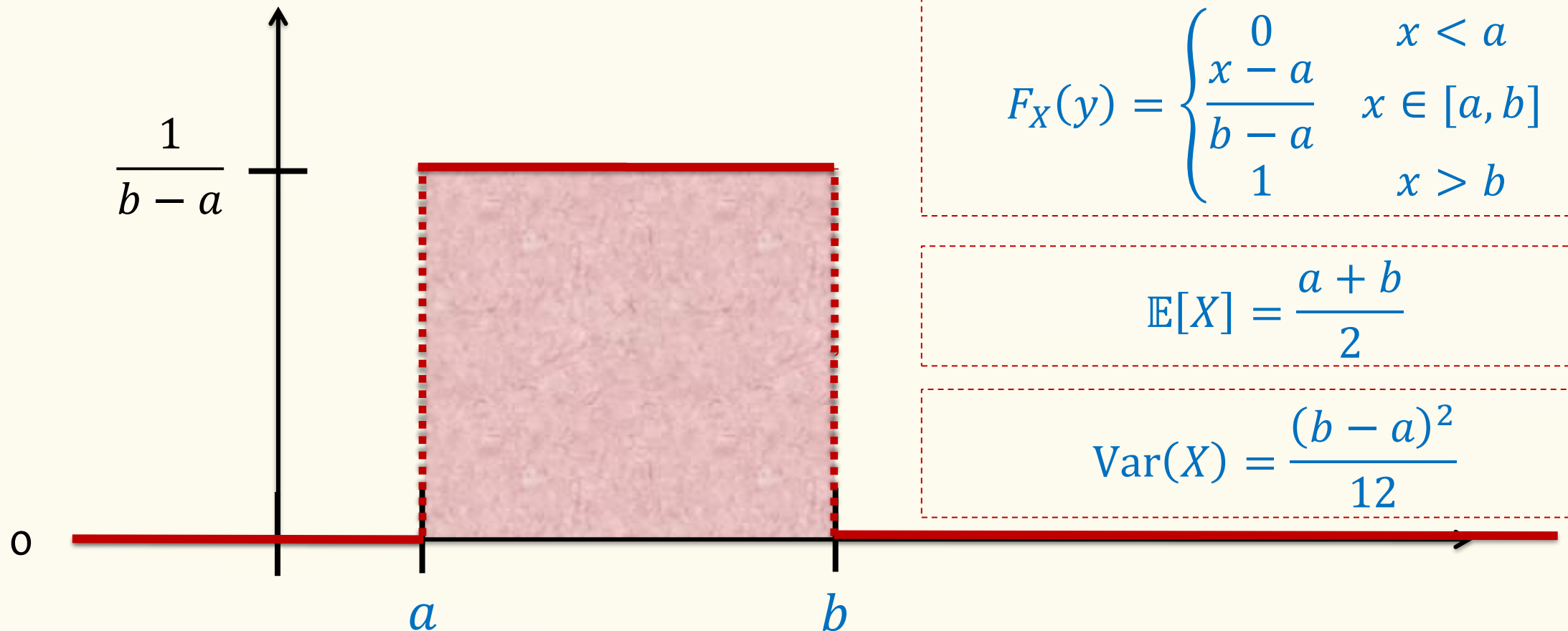
$$F_X(y) = P(X \leq y)$$

Recap: From Discrete to Continuous

	Discrete	Continuous
PMF/PDF	$p_X(x) = P(X = x)$	$f_X(x) \neq P(X = x) = 0$
CDF	$F_X(x) = \sum_{t \leq x} p_X(t)$	$F_X(x) = \int_{-\infty}^x f_X(t) dt$
Normalization	$\sum_x p_X(x) = 1$	$\int_{-\infty}^{\infty} f_X(x) dx = 1$
Expectation	$\mathbb{E}[g(X)] = \sum_x g(x) p_X(x)$	$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$

Uniform Distribution Summary

$X \sim \text{Unif}(a, b)$



$$f_X(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{else} \end{cases}$$

$$F_X(y) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a, b] \\ 1 & x > b \end{cases}$$

$$\mathbb{E}[X] = \frac{a+b}{2}$$

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

$$P(X > t) = e^{-t\lambda}$$

Exponential Distribution

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

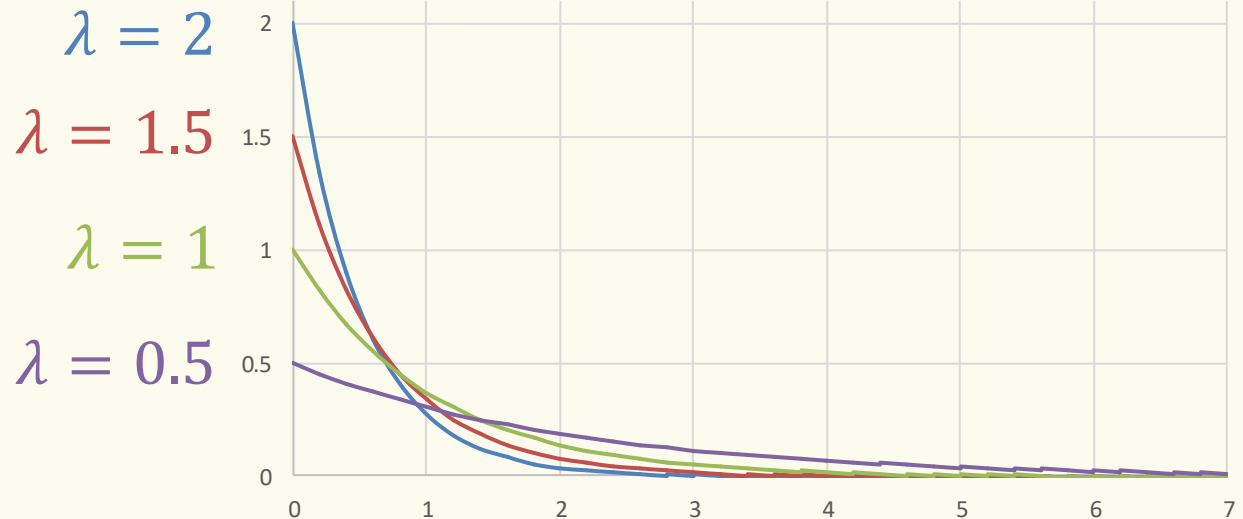
Definition. An **exponential random variable** X with parameter $\lambda \geq 0$ is follows the exponential density

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

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

$$\mathbb{E}[X] = \frac{1}{\lambda}$$

$$\text{Var}(X) = \frac{1}{\lambda^2}$$



Agenda

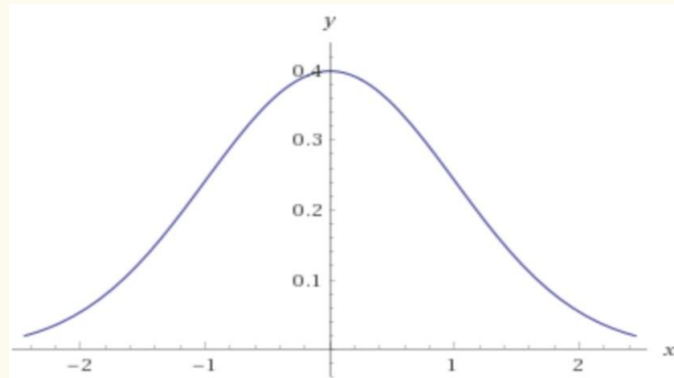
- Normal Distribution Practice with Normals
- Central Limit Theorem (CLT)

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



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



Carl Friedrich
Gauss

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

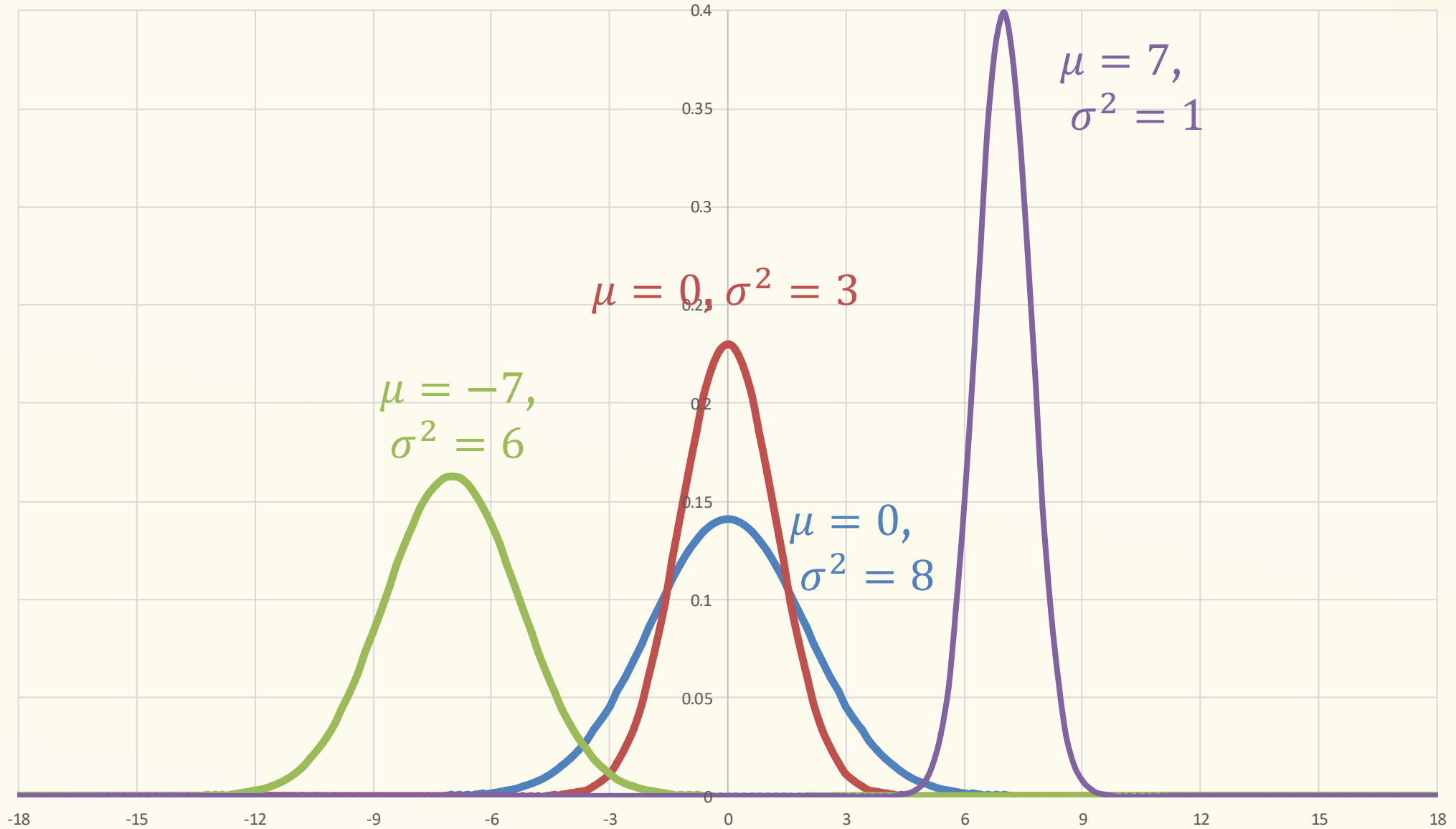
$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

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

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

The Normal Distribution

Aka a “Bell Curve” (imprecise name)



Standard normal distribution

Standard (unit) normal = $\mathcal{N}(0, 1)$

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

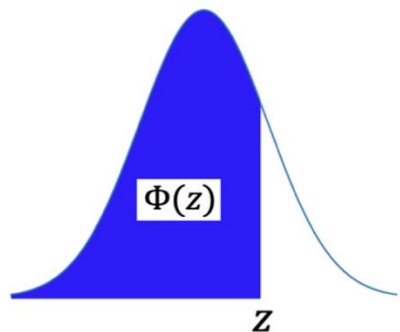
Note: $\Phi(z)$ has no closed form – generally given via tables

Review Table of $\Phi(z)$ CDF of Standard Normal

$$P(Z \leq 0.98) = \Phi(0.98) \approx 0.8365$$

For what a is

$$P(Z > a) \leq 0.01 ?$$



Φ 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

Review Table of $\Phi(z)$ CDF of Standard Normal

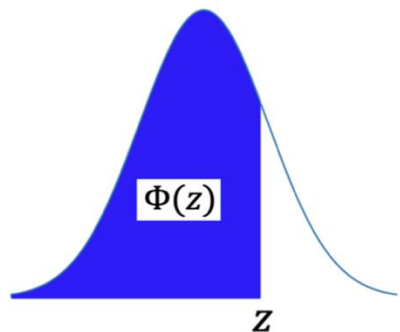
$$P(Z \leq 0.98) = \Phi(0.98) \approx 0.8365$$

For what a is

$$P(Z \geq a) \leq 0.01?$$

For any $a \geq 2.33$

$$P(Z > a) \leq 0.01.$$



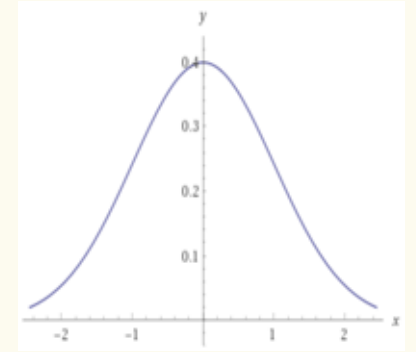
Φ Table: $\mathbb{P}(Z \leq z)$ when $Z \sim \mathcal{N}(0,1)$

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	0.85995	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

The Standard Normal CDF

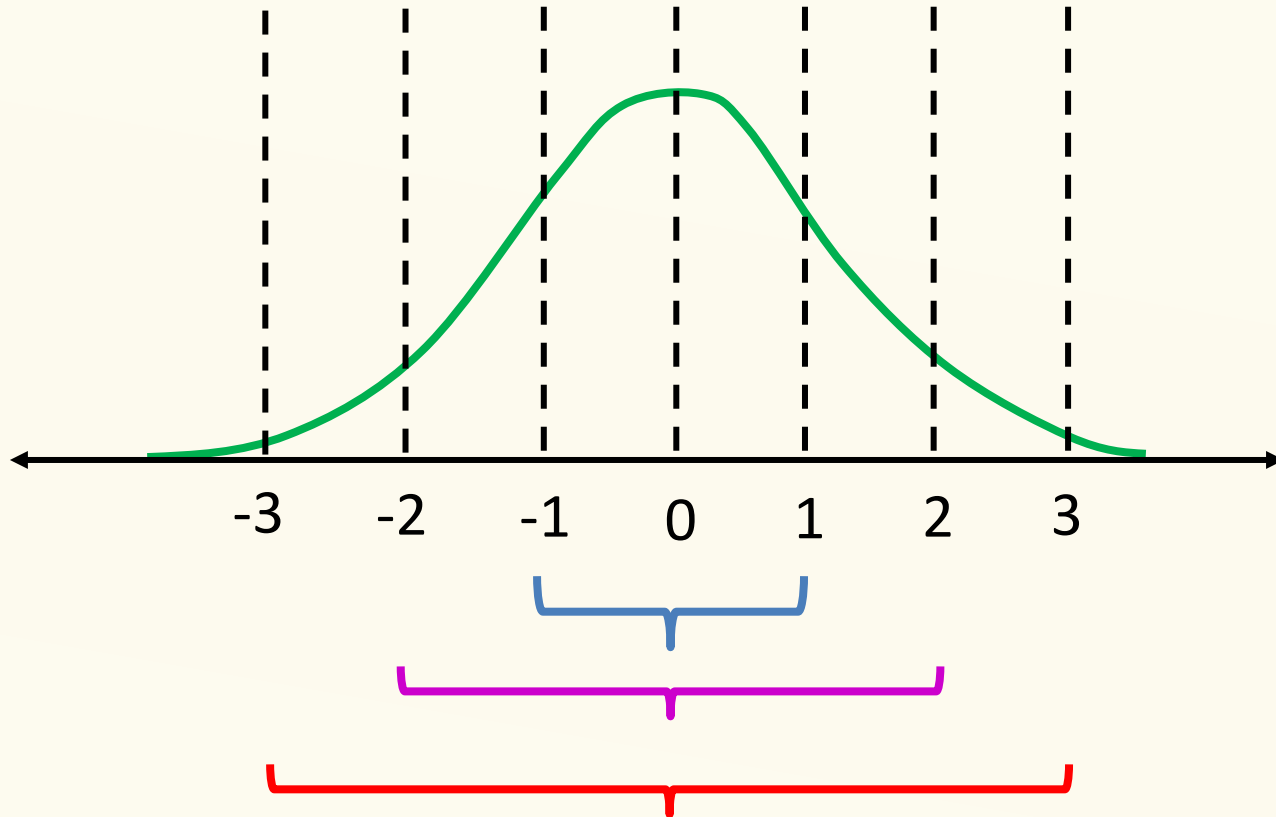
What is the probability that a standard Normal is within k standard deviations of its mean?

$$\begin{aligned}P(-k < Z < k) &= \Phi(k) - \Phi(-k) \\ &= \Phi(k) - (1 - \Phi(k)) \\ &= 2 \cdot \Phi(k) - 1\end{aligned}$$



Deviation from the Mean

If $Z \sim \mathcal{N}(\mu, \sigma^2)$, then $\mathbf{P}\{-k < Z < k\} = 2\Phi(k) - 1$



- w/prob 68%, Z is within 1 std of its mean
- w/prob 95%, Z is within 2 std of its mean
- w/prob 99.7%, Z is within 3 std of its mean

Closure of normal distribution – Under Shifting and Scaling

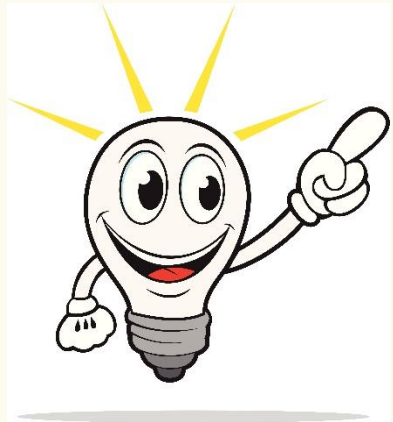
Fact. If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$

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

Mean and variance follow from properties you know! The fact that result of shifting and scaling still normal is not obvious, but not too difficult

But what if we don't have a standard Normal?

Bottom line: Deviations from mean we saw for a standard Normal holds for general Normal (provided it's phrased in terms of standard deviations).



$$X \sim \mathcal{N}(\mu, \sigma^2) \iff Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

$$\underbrace{P\{-k\sigma < X - \mu < k\sigma\}}_{\text{Prob. } X \text{ deviates from its mean by } k \text{ stds}} = P\left\{-k < \frac{X - \mu}{\sigma} < k\right\} = \underbrace{P\{-k < Z < k\}}_{\text{Prob. } Z \text{ deviates from its mean by } k \text{ stds}}$$

Prob. X deviates from its mean by k stds

Prob. Z deviates from its mean by k stds

Agenda

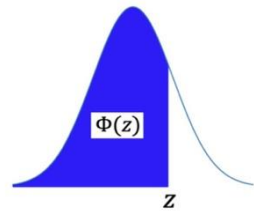
- Normal Distribution
- Practice with Normals ◀
- The Central Limit Theorem

Example

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

$$P(X \leq 1.2) =$$

Table of Standard Cumulative Normal Density



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

$$\begin{aligned} P(X \leq 1.2) &= P\left(\frac{X - 0.4}{2} \leq \frac{1.2 - 0.4}{2}\right) \\ &= P\left(\frac{X - 0.4}{2} \leq 0.4\right) = \Phi(0.4) \approx 0.6554 \end{aligned}$$

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

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

$$P(2 < X < 5) =$$

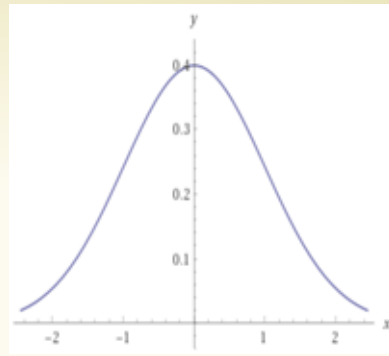
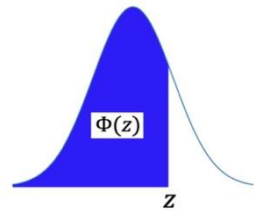


Table of Standard Cumulative Normal Density



Φ 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 $X \sim \mathcal{N}(3, 16)$.

$$\begin{aligned} P(2 < X < 5) &= P\left(\frac{2 - 3}{4} < \frac{X - 3}{4} < \frac{5 - 3}{4}\right) \\ &= P\left(-\frac{1}{4} < Z < \frac{1}{2}\right) \\ &= \Phi\left(\frac{1}{2}\right) - \Phi\left(-\frac{1}{4}\right) \\ &= \Phi\left(\frac{1}{2}\right) - \left(1 - \Phi\left(\frac{1}{4}\right)\right) \approx 0.29017 \end{aligned}$$

Summary so far

- Normal distributions stay normal under shifting and scaling.
- To “standardize” a normal random variable $X \sim \mathcal{N}(\mu, \sigma^2)$, you subtract the mean and divide by the standard deviation, i.e.,

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

- This allows you to use the standard normal tables (showing $\Phi(z) = P(Z \leq z)$ for $Z \sim \mathcal{N}(0, 1)$) to do calculations for any normal distribution.

Another important property: closure 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 are **not** surprising.

Why not surprising?

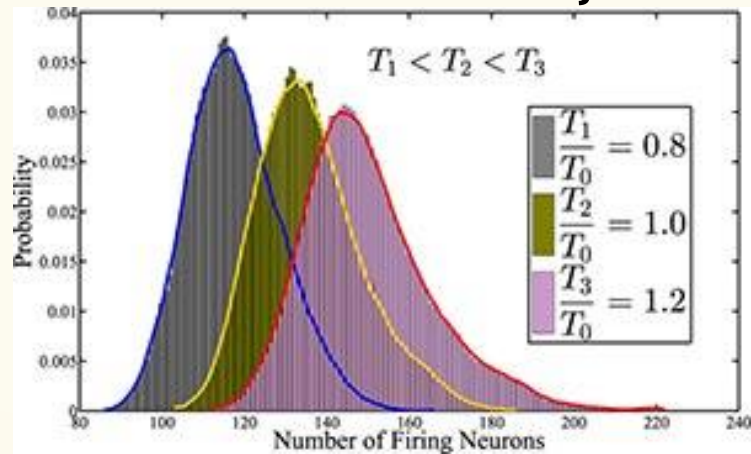
- Linearity of expectation (always true)
- When X and Y are independent, $\text{Var}(aX + bY) = a^2\text{Var}(X) + b^2\text{Var}(Y)$

Agenda

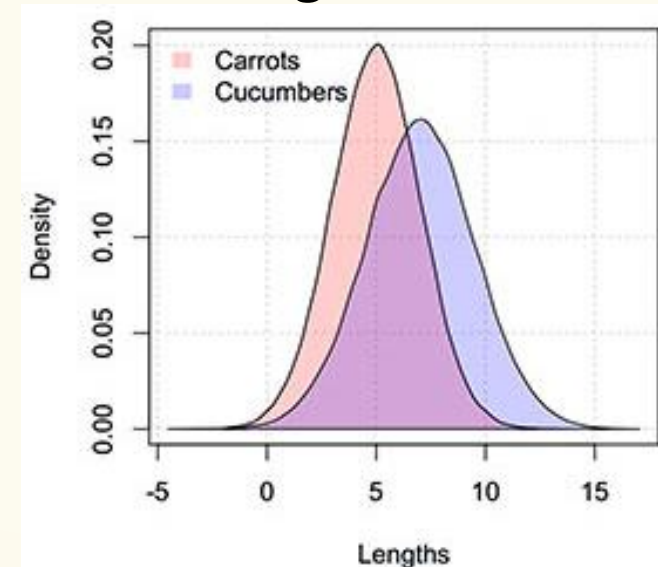
- Normal Distribution
- Practice with Normals
- Central Limit Theorem (CLT) ◀

Normal Distributions EVERYWHERE – why?

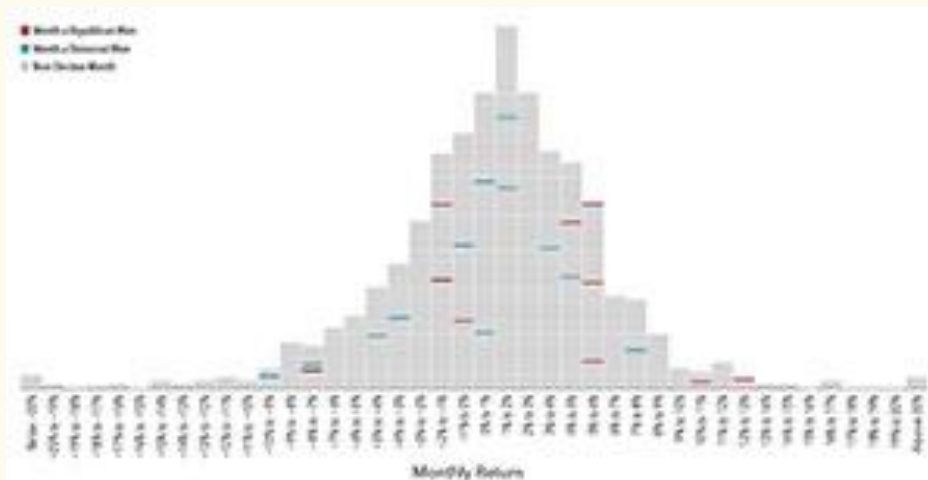
Neuron Activity



Vegetables



S&P 500 Returns after Elections



Examples from:
<https://galtonboard.com/probabilityexamplesinlife>

Sums of i.i.d. RVs look normal!

i.i.d. = independent and identically distributed

X_1, \dots, X_n i.i.d. with expectation μ and variance σ^2

Consider

$$S_n = X_1 + \dots + X_n$$

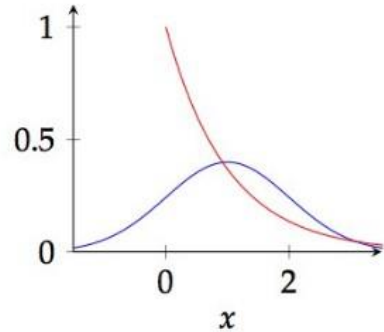
$$\mathbb{E}[S_n] =$$

$$\text{Var}(S_n) =$$

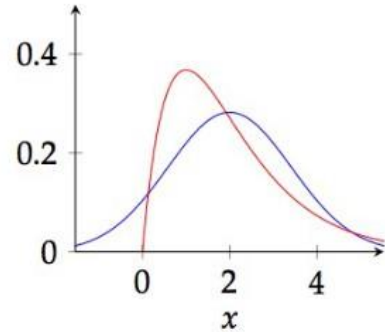
Empirical observation:

S_n looks like a normal RV as n grows.

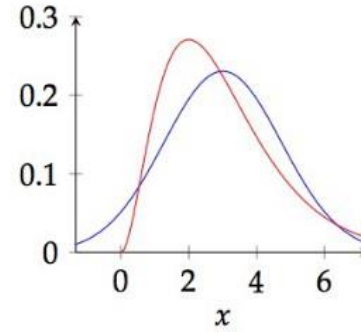
Example: Sum of n i.i.d. $\text{Exp}(1)$ random variables



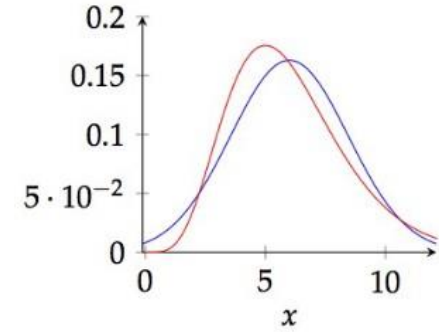
(a) $n = 1$



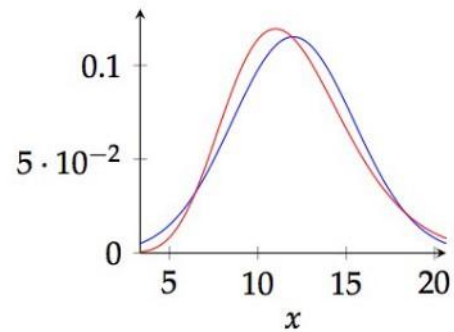
(b) $n = 2$



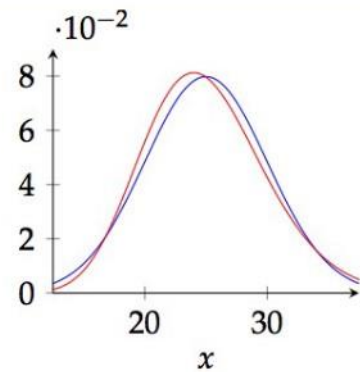
(c) $n = 3$



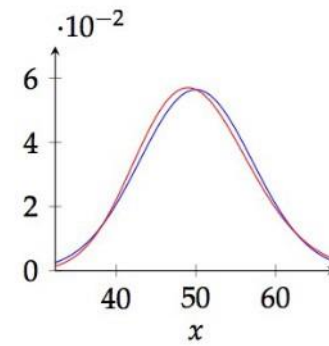
(d) $n = 6$



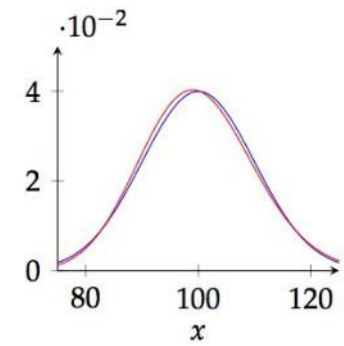
(e) $n = 12$



(f) $n = 25$

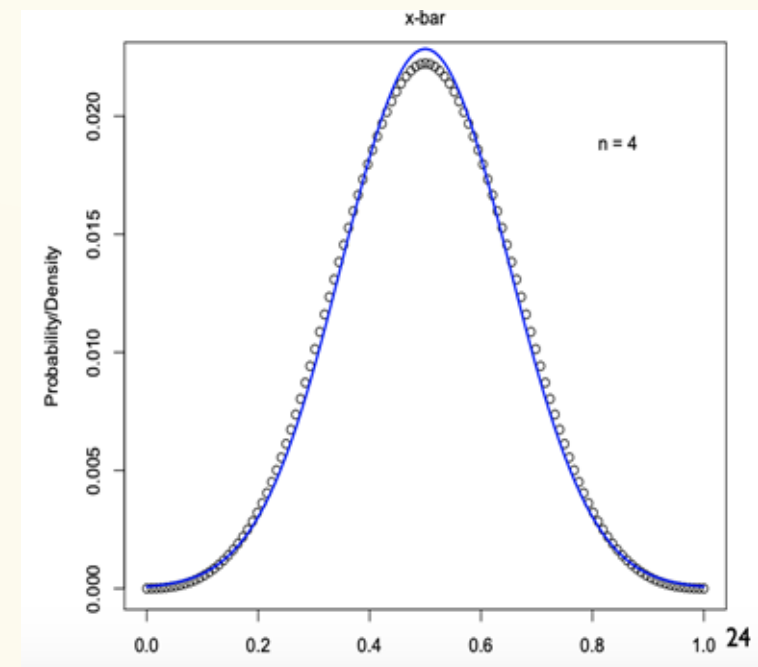
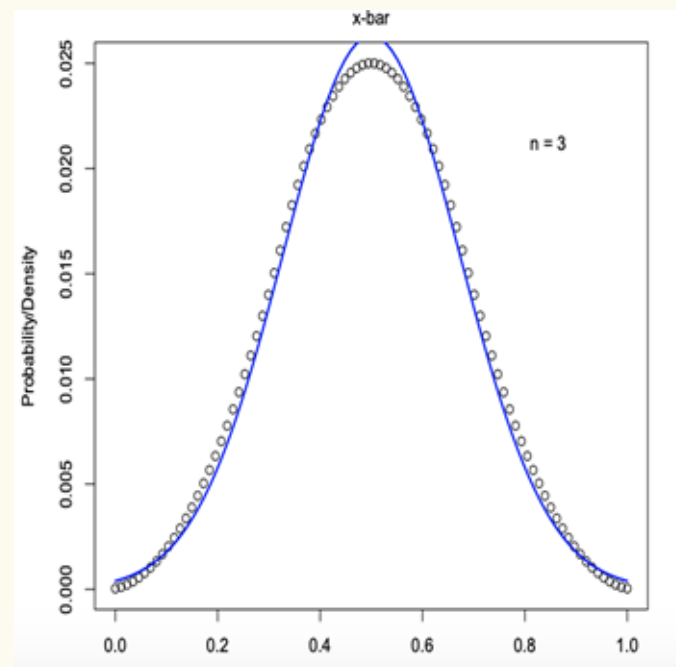
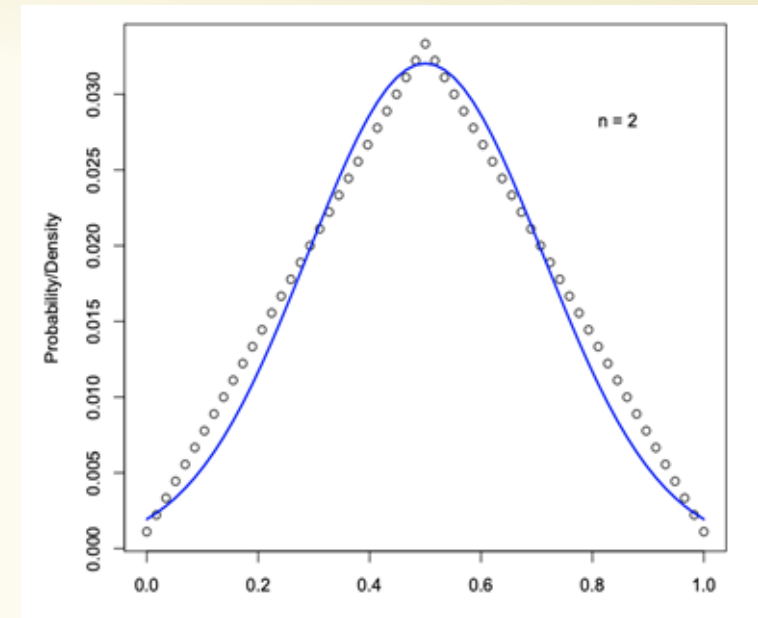
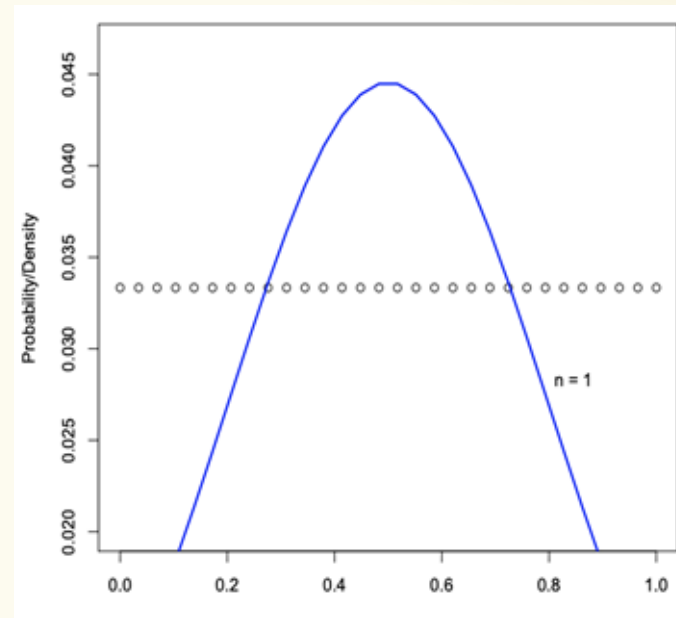


(g) $n = 50$

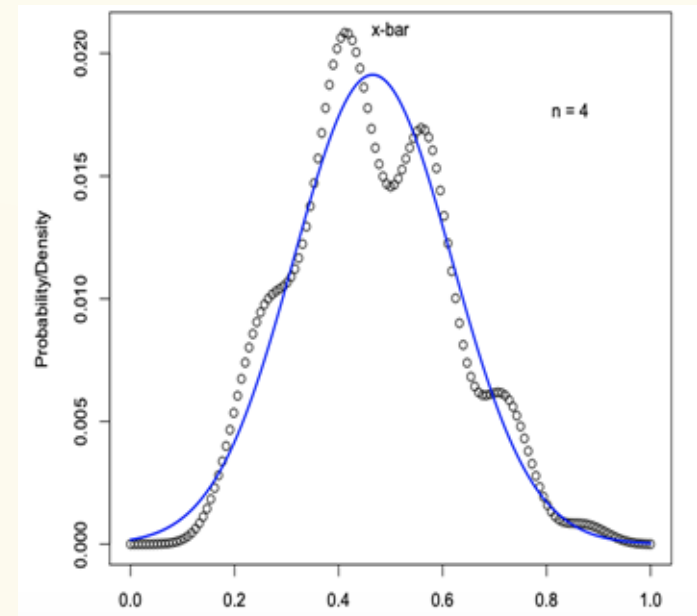
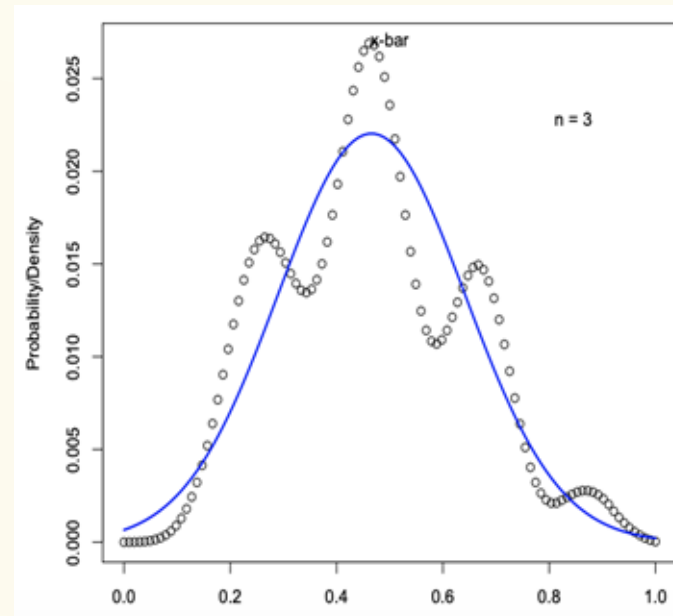
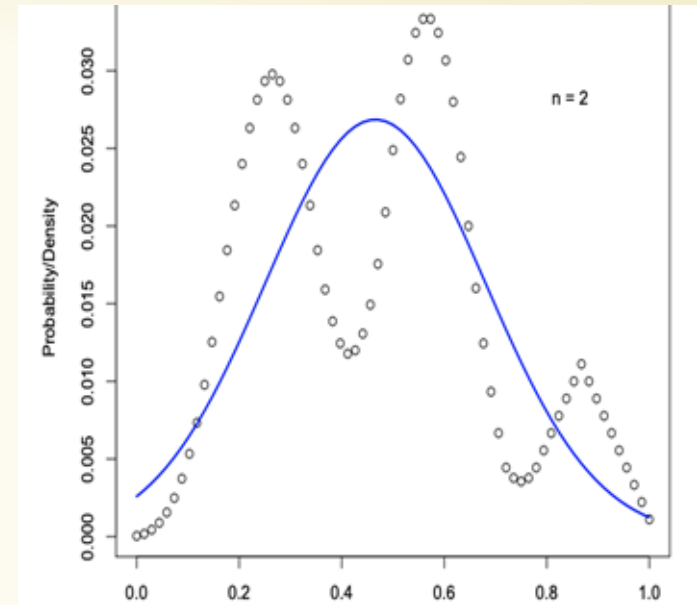
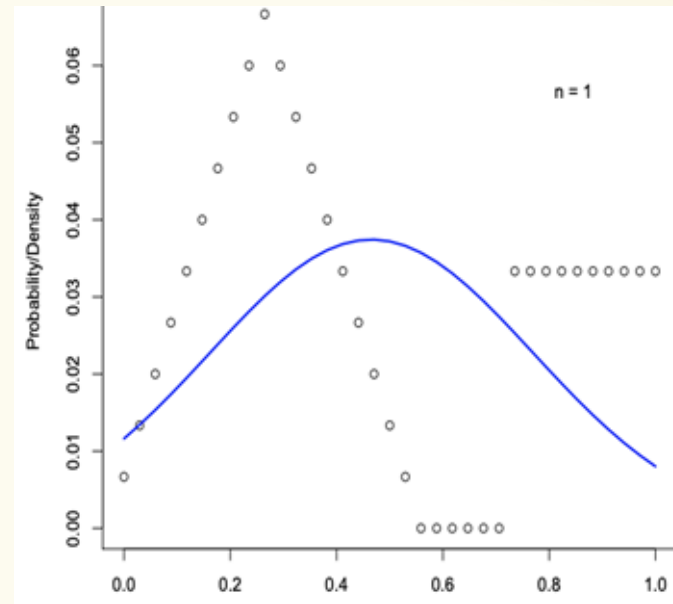


(h) $n = 100$

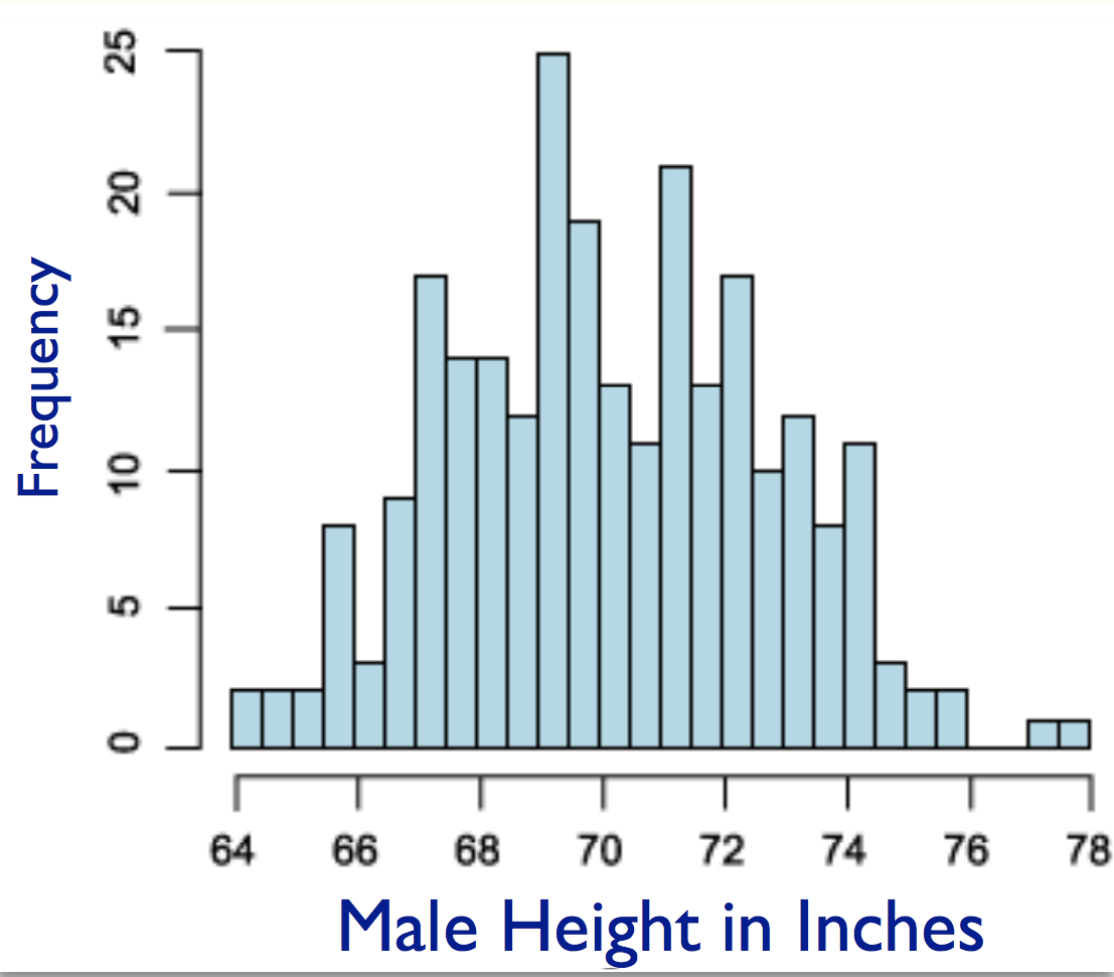
Example: avg of uniform r.v.s



CLT : Avg of some other weird i.i.d. r.v.s



Suppose that what we see in nature results from combining (summing) many independent random observations...



Then distribution might look normal.
e.g. Height distribution resembles Gaussian.

R.A.Fisher (1918) observed that the height is likely the outcome of the sum of many independent random parameters, i.e., can be written as

$$X = X_1 + \dots + X_n$$

Sums of i.i.d. RVs

i.i.d. = independent and identically distributed

X_1, \dots, X_n i.i.d. with expectation μ and variance σ^2

Define $S_n = X_1 + \dots + X_n$

$$\mathbb{E}[S_n] = \mathbb{E}[X_1] + \dots + \mathbb{E}[X_n] = n\mu$$

$$\text{Var}(S_n) = \text{Var}(X_1) + \dots + \text{Var}(X_n) = n\sigma^2$$

Empirical observation:

S_n looks like a normal RV as n grows, but...

Central Limit Theorem

X_1, \dots, X_n i.i.d., each with expectation μ and variance σ^2

Define $S_n = X_1 + \dots + X_n$ and

$$Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

$$\mathbb{E}[S_n] = \mathbb{E}[X_1] + \dots + \mathbb{E}[X_n] = n\mu$$

$$\text{Var}(S_n) = \text{Var}(X_1) + \dots + \text{Var}(X_n) = n\sigma^2$$

$$\mathbb{E}[Y_n] =$$

$$\text{Var}(Y_n) =$$

Central Limit Theorem

X_1, \dots, X_n i.i.d., each with expectation μ and variance σ^2

Define $S_n = X_1 + \dots + X_n$ and

$$\mathbb{E}[S_n] = \mathbb{E}[X_1] + \dots + \mathbb{E}[X_n] = n\mu$$

$$Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

$$\text{Var}(S_n) = \text{Var}(X_1) + \dots + \text{Var}(X_n) = n\sigma^2$$

$$\mathbb{E}[Y_n] = \frac{1}{\sigma\sqrt{n}} (\mathbb{E}[S_n] - n\mu) = \frac{1}{\sigma\sqrt{n}} (n\mu - n\mu) = 0$$

$$\text{Var}(Y_n) = \frac{1}{\sigma^2 n} (\text{Var}(S_n - n\mu)) = \frac{\text{Var}(S_n)}{\sigma^2 n} = \frac{\sigma^2 n}{\sigma^2 n} = 1$$

Central Limit Theorem

X_1, \dots, X_n i.i.d., each with expectation μ and variance σ^2

Define $S_n = X_1 + \dots + X_n$,

$$Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

Then distribution of $Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$ converges to that of a normal distribution with mean 0 and variance 1 as $n \rightarrow \infty$.

Central Limit Theorem

$$Y_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

Theorem. (Central Limit Theorem) The CDF of Y_n converges to the CDF of the standard normal $\mathcal{N}(0,1)$, i.e.,

$$\lim_{n \rightarrow \infty} P(Y_n \leq y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-x^2/2} dx$$

Central Limit Theorem

$$Y_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

Theorem. (Central Limit Theorem) The CDF of Y_n converges to the CDF of the standard normal $\mathcal{N}(0,1)$, i.e.,

$$\lim_{n \rightarrow \infty} P(Y_n \leq y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-x^2/2} dx$$

Also stated as:

- $\lim_{n \rightarrow \infty} Y_n \rightarrow \mathcal{N}(0,1)$
- $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i \rightarrow \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$ for $\mu = \mathbb{E}[X_i]$ and $\sigma^2 = \text{Var}(X_i)$

Summary Central Limit Theorem

X_1, \dots, X_n i.i.d., each with expectation μ and variance σ^2

Define $S_n = X_1 + \dots + X_n$ and $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ and $Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$

mean

variance

CLT:

S_n

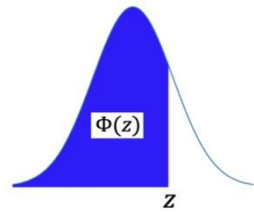
\bar{X}

Y_n

CLT application

- Suppose lightbulbs have a lifetime that is exponential with a mean of 5 hours.
- You buy a pack of 10 light bulbs. As soon as one burns out you replace it with another. What is the probability you still have a working light bulb after 70 hours?
- Estimate using the Central Limit Theorem.

Table of Standard Cumulative Normal Density



Φ 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

Outline of how CLT is used

- Write the event you are interested in, in terms of a sum of i.i.d. random variables.
- Normalize RV to have mean 0 and standard deviation 1.
- Write event in terms of Φ , the CDF of a $\mathcal{N}(0,1)$.
- Look up in table.

Agenda

- Applying CLT to discrete random variables – continuity correction

Using CLT to estimate Binomial probabilities

We flip n independent coins, heads with probability $p = 0.75$.

$$X = \# \text{ heads} \quad \mu = \mathbb{E}(X) = 0.75n \quad \sigma^2 = \text{Var}(X) = p(1 - p)n = 0.1875n$$

$$\mathbb{P}(X \leq 0.7n)$$

n	exact	$\mathcal{N}(\mu, \sigma^2)$ approx
10	0.4744072	0.357500327
20	0.38282735	0.302788308
50	0.25191886	0.207108089
100	0.14954105	0.124106539
200	0.06247223	0.051235217
1000	0.00019359	0.000130365

Example – Naive Approximation

Fair coin flipped (independently) **40** times. Probability of **20** or **21** heads?

Exact. $\mathbb{P}(X \in \{20,21\}) = \left[\binom{40}{20} + \binom{40}{21} \right] \left(\frac{1}{2} \right)^{40} \approx \boxed{0.2448}$

Approx. $X = \# \text{ heads}$ $\mu = \mathbb{E}(X) = 0.5n = 20$ $\sigma^2 = \text{Var}(X) = 0.25n = 10$

$$\mathbb{P}(20 \leq X \leq 21) = \Phi \left(\frac{20 - 20}{\sqrt{10}} \leq \frac{X - 20}{\sqrt{10}} \leq \frac{21 - 20}{\sqrt{10}} \right)$$

$$\approx \Phi \left(0 \leq \frac{X - 20}{\sqrt{10}} \leq 0.32 \right)$$

$$= \Phi(0.32) - \Phi(0) \approx \boxed{0.1241}$$



Example – Even Worse Approximation

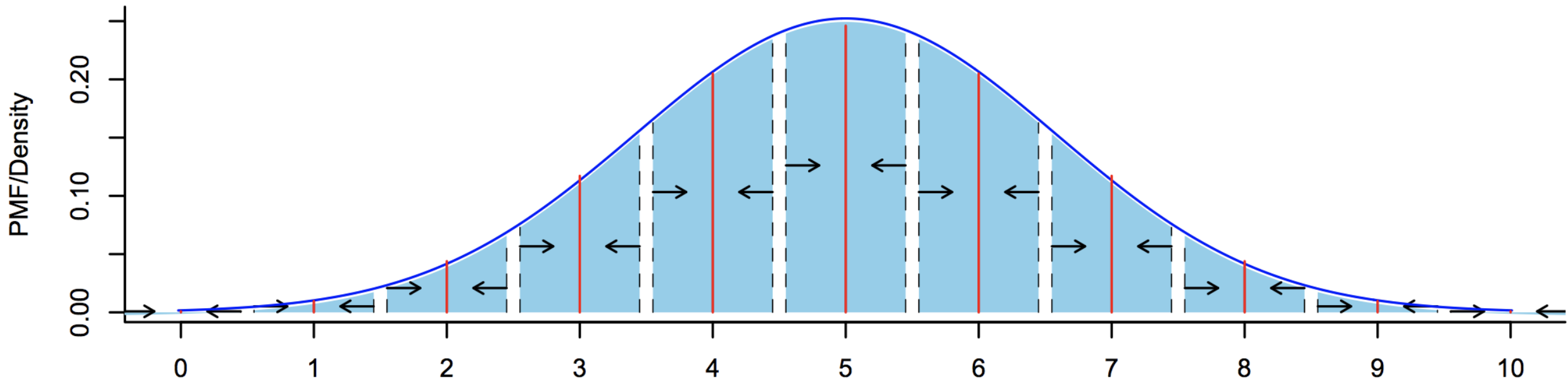
Fair coin flipped (independently) **40** times. Probability of **20** heads?

Exact. $\mathbb{P}(X = 20) = \binom{40}{20} \left(\frac{1}{2}\right)^{40} \approx \boxed{0.1254}$

Approx. $\mathbb{P}(20 \leq X \leq 20) = 0$ 

Solution – Continuity Correction

Probability estimate for i : Probability for all x that round to i !



To estimate probability that discrete RV lands in (integer) interval $\{a, \dots, b\}$, compute probability continuous approximation lands in interval $[a - \frac{1}{2}, b + \frac{1}{2}]$

Example – Continuity Correction

Fair coin flipped (independently) **40** times. Probability of **20** or **21** heads?

Exact. $\mathbb{P}(X \in \{20,21\}) = \left[\binom{40}{20} + \binom{40}{21} \right] \left(\frac{1}{2} \right)^{40} \approx \boxed{0.2448}$

Approx. $X = \# \text{ heads}$ $\mu = \mathbb{E}(X) = 0.5n = 20$ $\sigma^2 = \text{Var}(X) = 0.25n = 10$

$$\begin{aligned} \mathbb{P}(19.5 \leq X \leq 21.5) &= \Phi\left(\frac{19.5 - 20}{\sqrt{10}} \leq \frac{X - 20}{\sqrt{10}} \leq \frac{21.5 - 20}{\sqrt{10}}\right) \\ &\approx \Phi\left(-0.16 \leq \frac{X - 20}{\sqrt{10}} \leq 0.47\right) \\ &= \Phi(0.47) - \Phi(-0.16) \approx \boxed{0.2452} \end{aligned}$$

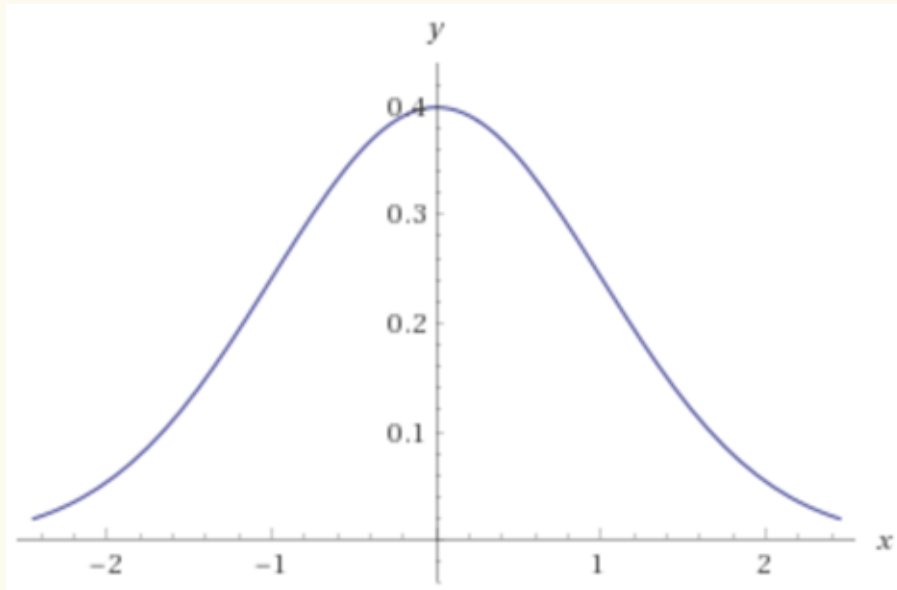


Example – Continuity Correction

Fair coin flipped (independently) **40** times. Probability of **20** heads?

Exact. $\mathbb{P}(X = 20) = \binom{40}{20} \left(\frac{1}{2}\right)^{40} \approx \boxed{0.1254}$

Approx.
$$\mathbb{P}(19.5 \leq X \leq 20.5) = \Phi\left(\frac{19.5 - 20}{\sqrt{10}} \leq \frac{X - 20}{\sqrt{10}} \leq \frac{20.5 - 20}{\sqrt{10}}\right)$$
$$\approx \Phi\left(-0.16 \leq \frac{X - 20}{\sqrt{10}} \leq 0.16\right)$$
$$= \Phi(0.16) - \Phi(-0.16) \approx \boxed{0.1272}$$



Normal Distribution



Paranormal Distribution