CSE 312

Foundations of Computing II

Lecture 15: Exponential and Normal Distribution

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

Anna R. Karlin

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

Review – Continuous RVs

Probability Density Function (PDF).

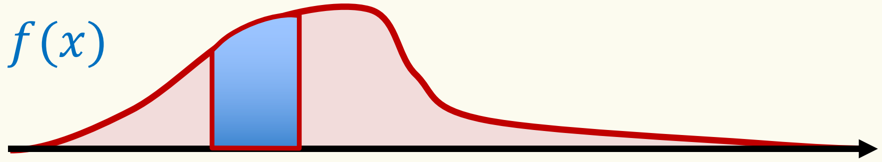
$f: \mathbb{R} \to \mathbb{R}$ s.t.

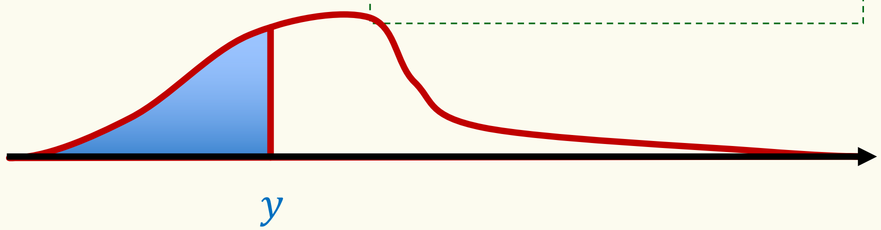
- $f(x) \geq 0$ for all $x \in \mathbb{R}$
- $\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1$

Cumulative Density Function (CDF).

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

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





Density ≠ Probability !

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

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

Expectation of a Continuous RV

Definition. The **expected value** of a continuous RV X is defined as

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} f_X(x) \cdot x \, \mathrm{d}x$$

Fact. $\mathbb{E}(aX + bY + c) = a\mathbb{E}(X) + b\mathbb{E}(Y) + c$

Definition. The **variance** of a continuous RV X is defined as

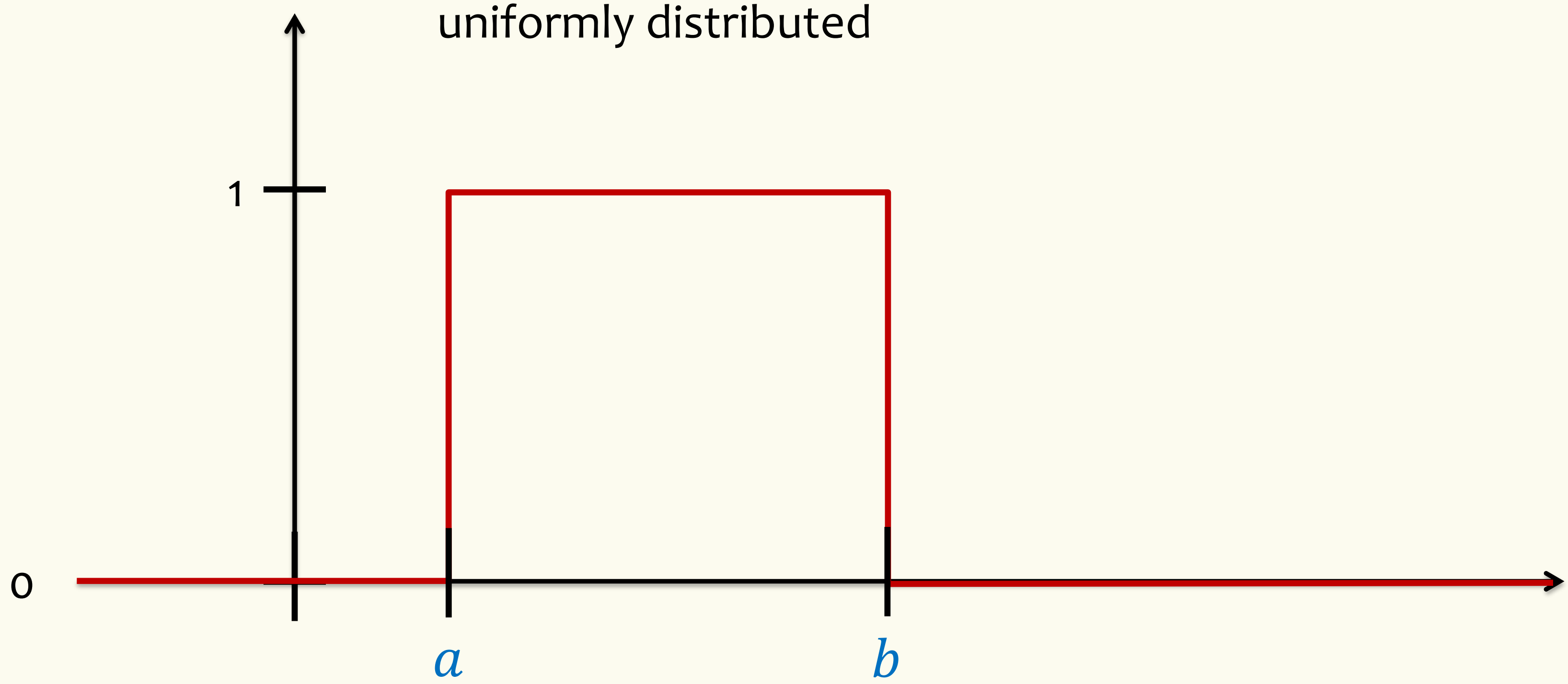
$$\mathrm{Var}(X) = \int_{-\infty}^{+\infty} f_X(x) \cdot \left(x - \mathbb{E}(X)\right)^2 \mathrm{d}x = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

Uniform Distribution

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

We also say that X follows the uniform distribution / is uniformly distributed

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



Uniform Density – Expectation

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

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

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} f_X(x) \cdot x \, \mathrm{d}x$$

$$= \frac{1}{b-a}\int_a^b x \, \mathrm{d}x \quad = \frac{1}{b-a}\left(\frac{x^2}{2}\right)\Bigg|_a^b = \frac{1}{b-a}\left(\frac{b^2-a^2}{2}\right)$$

$$= \frac{(b-a)(a+b)}{2(b-a)} \quad = \frac{a+b}{2}$$

Uniform Density – Variance

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

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

$$\mathbb{E}(X^2) = \int_{-\infty}^{+\infty} f_X(x) \cdot x^2 \, \mathrm{d}x$$

$$= \frac{1}{b-a} \int_a^b x^2 \, \mathrm{d}x = \frac{1}{b-a} \left(\frac{x^3}{3} \right) \Bigg|_a^b = \frac{b^3 - a^3}{3(b-a)}$$

$$= \frac{(b-a)(b^2 + ab + a^2)}{3(b-a)} = \frac{b^2 + ab + a^2}{3}$$

Uniform Density – Variance

$$\mathbb{E}(X^2) = \frac{b^2 + ab + a^2}{3} \qquad \mathbb{E}(X) = \frac{a+b}{2}$$

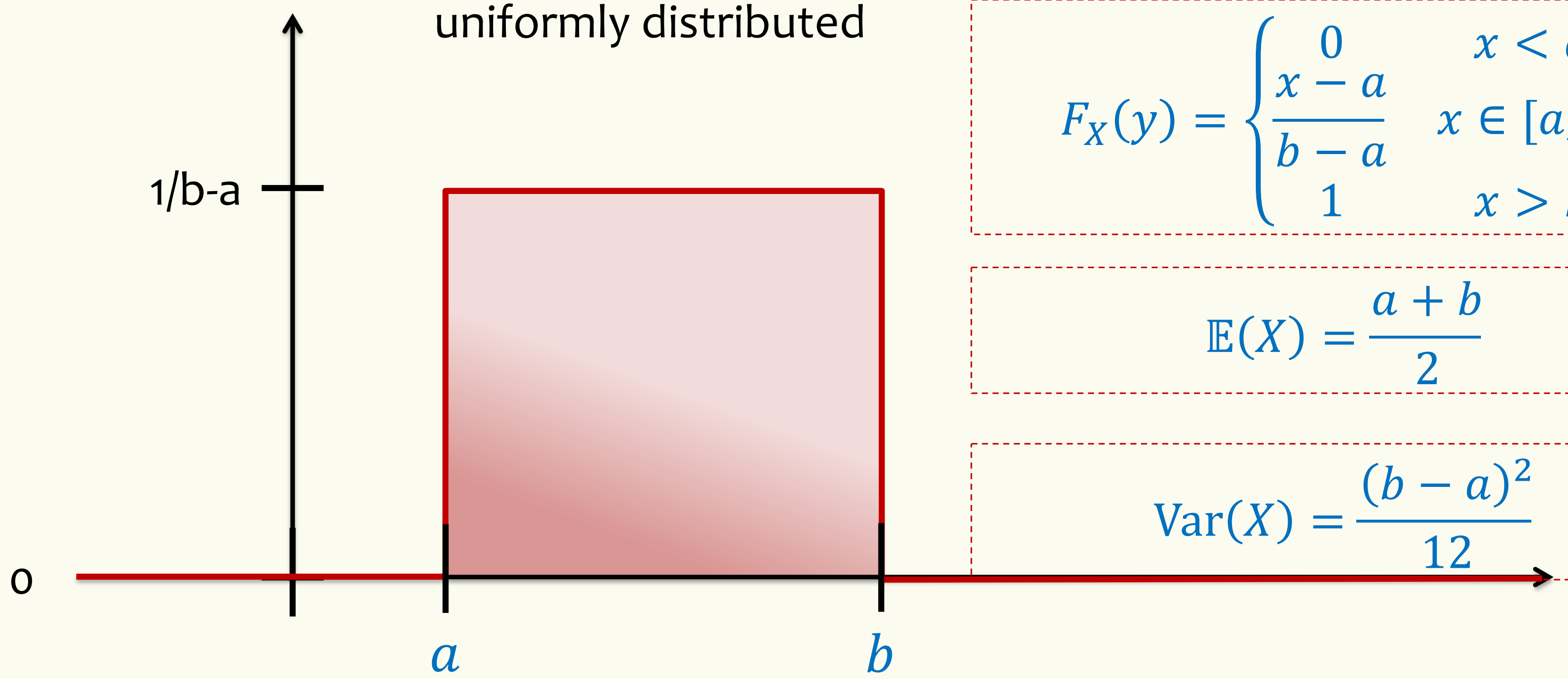
$X \sim \text{Unif}(a, b)$

$$\begin{aligned}\text{Var}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\ &= \frac{b^2 + ab + a^2}{3} - \frac{a^2 + 2ab + b^2}{4} \\ &= \frac{4b^2 + 4ab + 4a^2}{12} - \frac{3a^2 + 6ab + 3b^2}{12} \\ &= \frac{b^2 - 2ab + a^2}{12} = \frac{(b-a)^2}{12}\end{aligned}$$

Uniform Distribution

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

We also say that X follows the uniform distribution / is uniformly distributed



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

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

$$\mathbb{E}(X) = \frac{a+b}{2}$$

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

Exponential Density

Assume expected # of occurrences of an event per unit of time is λ

- Cars going through intersection
- Number of lightning strikes
- Requests to web server
- Patients admitted to ER

Numbers of occurrences of event: Poisson distribution

$$\mathbb{P}(X = i) = e^{-\lambda}\frac{\lambda^i}{i!} \qquad \text{(Discrete)}$$

How long to wait until next event? Exponential density!

Let's define it and then derive it!

The Exponential PDF/CDF

Assume expected # of occurrences of an event per unit of time is λ

Numbers of occurrences of event: Poisson distribution

How long to wait until next event? <u>Exponential density!</u>

- The exponential RV has range $[0, \infty]$, unlike Poisson with range $\{0,1,2,\dots\}$
- Let $Y \sim Exp(\lambda)$ be the time till the first event. We will compute $F_Y(t)$ and $f_Y(t)$

The Exponential PDF/CDF

Assume expected # of occurrences of an event per unit of time is λ

Numbers of occurrences of event: Poisson distribution

How long to wait until next event? Exponential density!

- The exponential RV has range [0, ∞], unlike Poisson with range {0,1,2,…}
- Let $Y \sim Exp(\lambda)$ be the time till the first event. We will compute $F_Y(t)$ and $f_Y(t)$
- Let $\mathrm{X} \sim Poi(t\lambda)$ be the # of events in the first t units of time, for $t \geq 0$.
- $\mathrm{P(Y > t)} = P(no\ event\ in\ the\ first\ t\ units) = P(X = 0) = e^{-t\lambda}\frac{t\lambda^0}{0!} = e^{-t\lambda}$
- $\mathrm{F_Y(t)} = 1 - P(Y > t) = 1 - e^{-t\lambda}$
- $\mathrm{f_Y(t)} = \frac{d}{dt}F_Y(t) = \lambda e^{-t\lambda}$

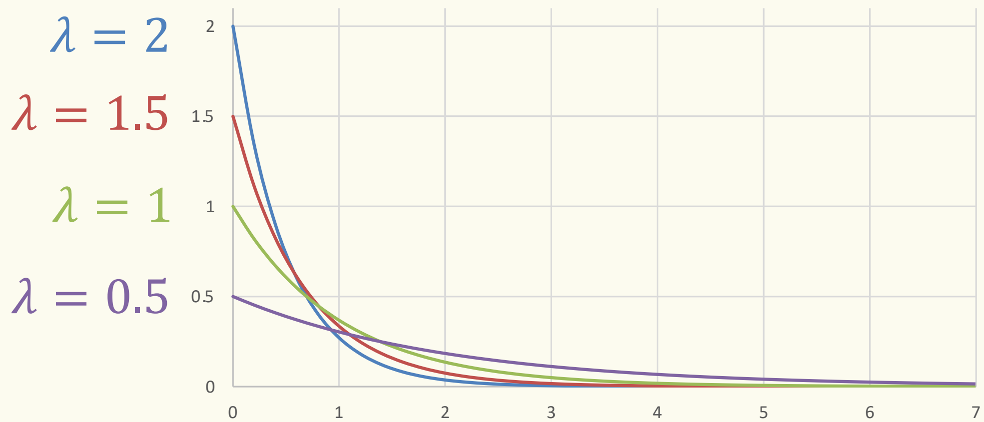
Exponential Distribution

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

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



Expectation

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

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} f_X(x) \cdot x \, \mathrm{d}x$$

Expectation

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

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} f_X(x) \cdot x \, \mathrm{d}x$$

$$= \int_{0}^{+\infty} \lambda e^{-\lambda x} \cdot x \, \mathrm{d}x$$

$$= \left(-(x + \frac{1}{\lambda}) e^{-\lambda x} \right) \Big|_0^{\infty} = \frac{1}{\lambda}$$

Somewhat complex calculation
use integral by parts

$$\mathbb{E}(X) = \frac{1}{\lambda}$$

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

Memorylessness

Definition. A random variable is **memoryless** if for all $s, t > 0$,

$$\mathbb{P}(X > s + t \mid X > s) = \mathbb{P}(X > t).$$

Fact. $X \sim \mathrm{Exp}(\lambda)$ is memoryless.

Assuming exp distr, if you've waited s minutes,
prob of waiting t more is exactly same as $s = 0$

Assuming exp distr, if you've waited s minutes, prob of waiting t more is exactly same as $s = 0$

Memorylessness of Exponential

Fact. $X \sim \mathrm{Exp}(\lambda)$ is memoryless.

Proof.

$$\mathbb{P}(X > s + t \mid X > s)$$

Memorylessness of Exponential

Assuming exp distr, if you've waited s minutes, prob of waiting t more is exactly same as $s = 0$

Fact. $X \sim \mathrm{Exp}(\lambda)$ is memoryless.

Proof.

$$\mathbb{P}(X > s + t \mid X > s) = \frac{\mathbb{P}(\{X > s + t\} \cap \{X > s\})}{\mathbb{P}(X > s)}$$

$$= \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)}$$

$$= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = \mathbb{P}(X > t)$$

The only memoryless RVs are the geometric RV (discrete) and Exp RV (continuous)

example

- Time it takes to check someone out at a grocery store is exponential with an expected value of 10 mins.
- Independent for different customers
- If you are the second person in line, what is the probability that you will have to wait between 10 and 20 mins.

example

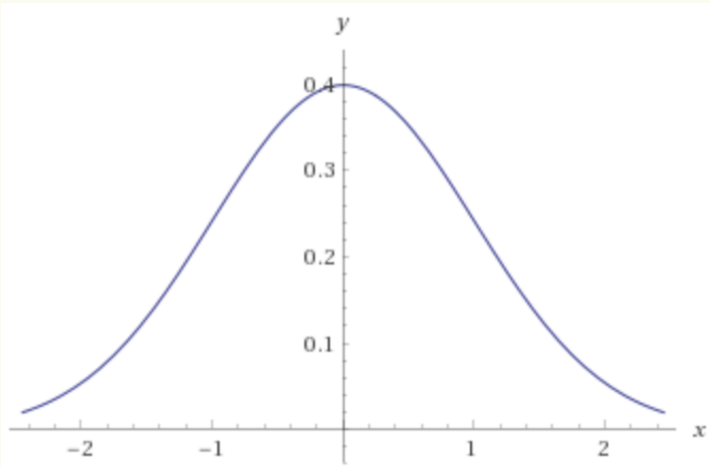
- Time it takes to check someone out at a grocery store is exponential with an expected value of 10 mins.
- Independent for different customers
- If you are the second person in line, what is the probability that you will have to wait between 10 and 20 mins.

$$T \sim Exp(\frac{1}{10})$$

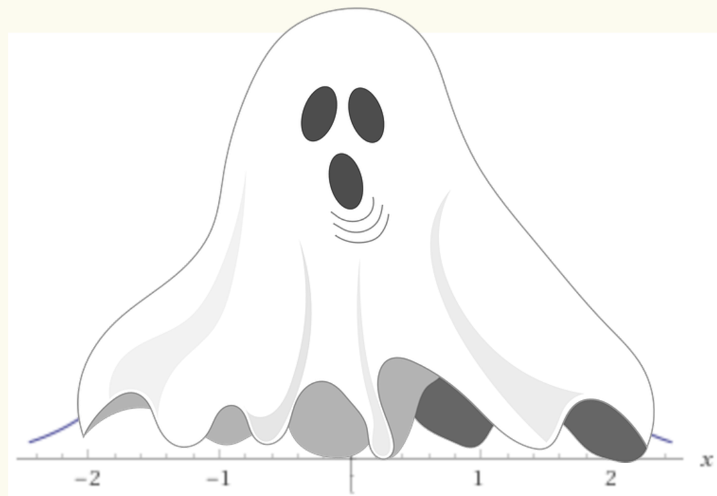
$$P(10 \leq T \leq 20) = \int_{10}^{20} \frac{1}{10} e^{-\frac{x}{10}} \, dx$$

$$y = \frac{x}{10}, dy = \frac{dx}{10}$$

$$P(10 \leq T \leq 20) = \int_{1}^{2} e^{-y} \, dy = -e^{-y} \Big|_{1}^{2} = e^{-1} - e^{-2}$$



Normal Distribution



Paranormal Distribution

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

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

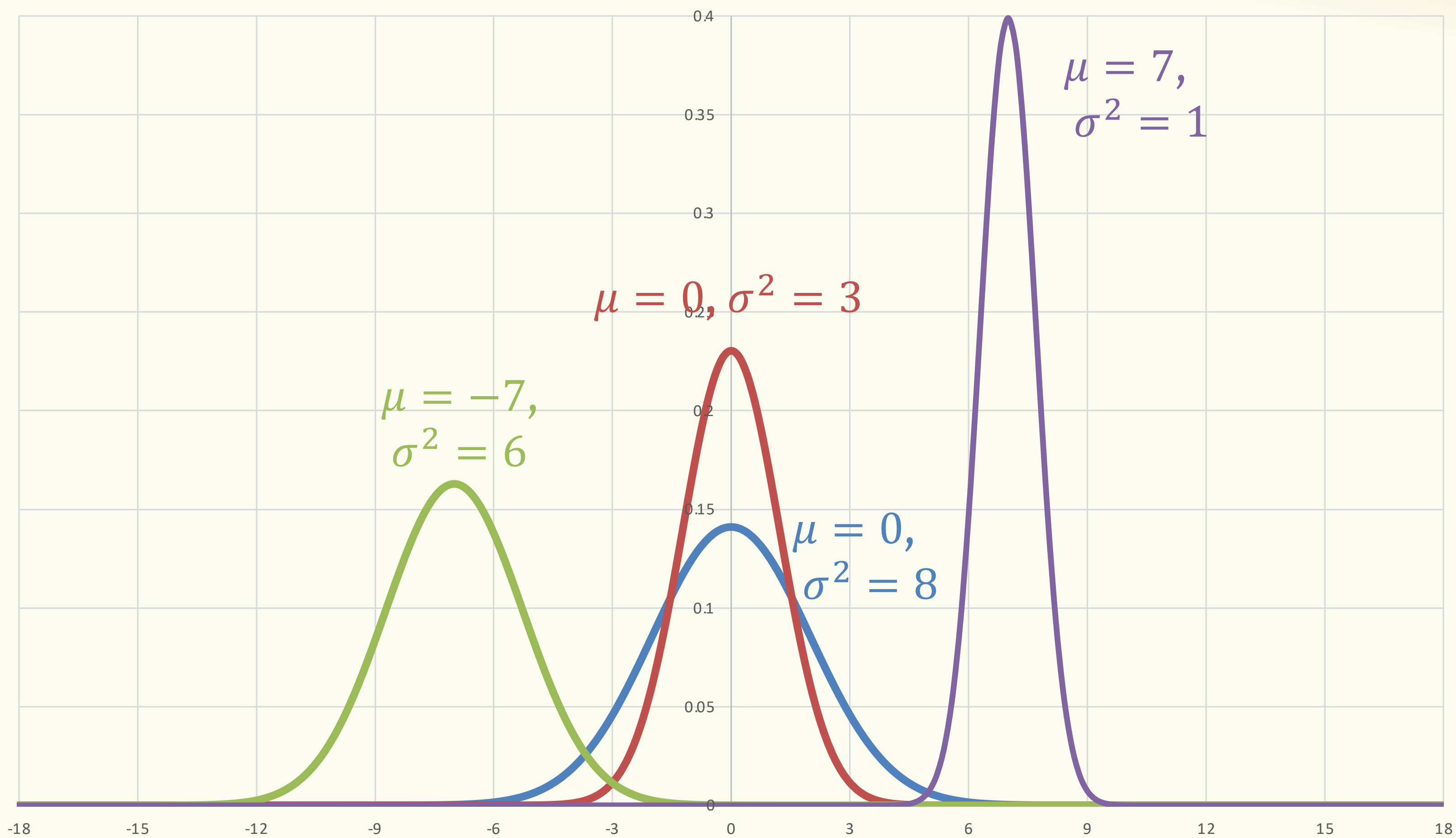
Fact. If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $\mathbb{E}(X) = \mu$, and $\mathrm{Var}(X) = \sigma^2$

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

We will see next time why the normal distribution is (in some sense) the most important distribution.

The Normal Distribution

Aka a "Bell Curve" (imprecise name)



Shifting and Scaling the Normal Distribution

Suppose $X \sim \mathcal{N}(\mu, \sigma^2)$ and $Y = aX + b$

What is $\mathbb{E}(Y) =$

$\mathrm{Var}(Y) =$

What is mean and variance of $\frac{X - \mu}{\sigma}$ **?**

Closure of normal distribution – Under Shifting and Scaling

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

We know: $\mathbb{E}(Y) = a\ \mathbb{E}(X) + b = a\mu + b$

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

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

Closure of the normal -- under addition

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

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.

CDF of normal distribution

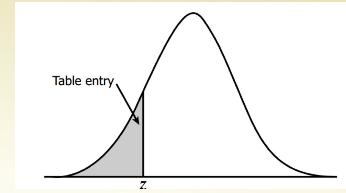
Fact. If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

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

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

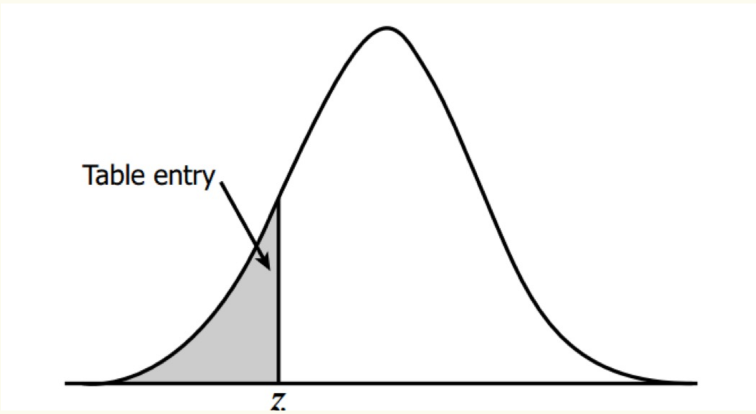
Note: $\Phi(z)$ has no closed form – generally given via tables

Table of Standard Cumulative Normal Density



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

The Standard Normal CDF



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

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

Example

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

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

Example

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

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

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

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

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 – Off by Standard Deviations

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

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

Example – Off by Standard Deviations

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

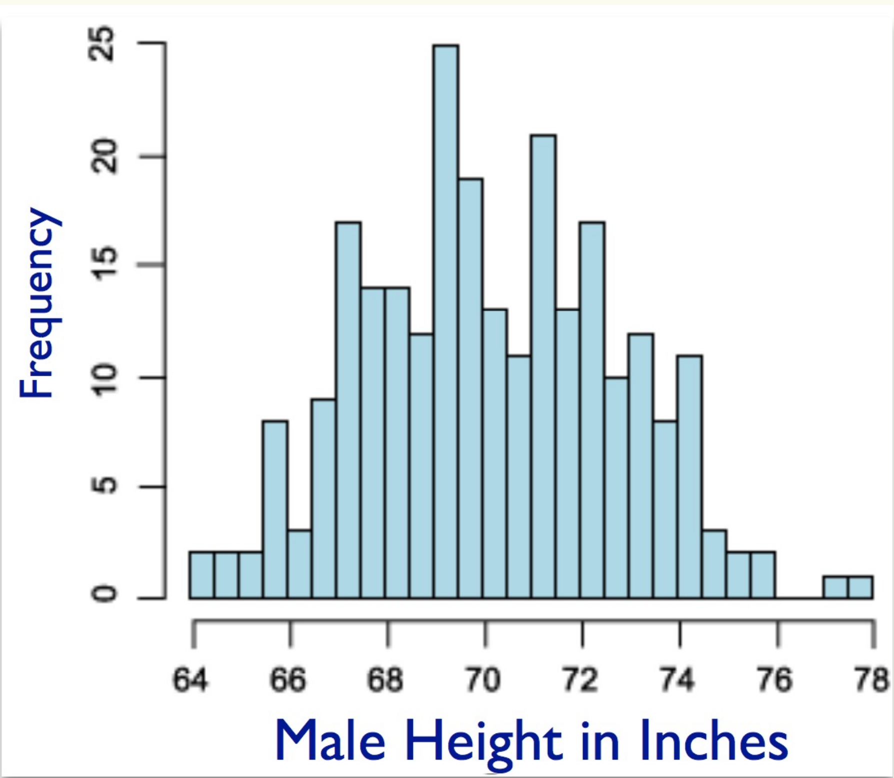
$$\mathbb{P}(|X-\mu| < k\sigma) = \mathbb{P}\left(\frac{|X-\mu|}{\sigma} < k\right) =$$

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

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

Gaussian in Nature

Empirical distribution of collected data often resembles a Gaussian …



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 written as

$$X = X_1 + \cdots + X_n$$

Next time: The Central Limit Theorem! Sum of independent and identical RVs is close to the normal distribution!