Chapter 9
Normal Distribution

Normal (a.k.a. Gaussian) distribution

<u>Defn</u>: $X \sim \boldsymbol{Normal}(\mu, \sigma^2)$ if

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad -\infty < x < \infty$$

where $\sigma > 0$. The parameter μ is called the **mean**, and parameter $\sigma = \sqrt{\boldsymbol{Var}(X)}$ is called the **standard deviation**.

<u>Defn</u>: X follows a **standard Normal** distribution if $X \sim Normal(0,1)$, i.e.,

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \qquad -\infty < x < \infty$$

Bell shape of Normal distribution

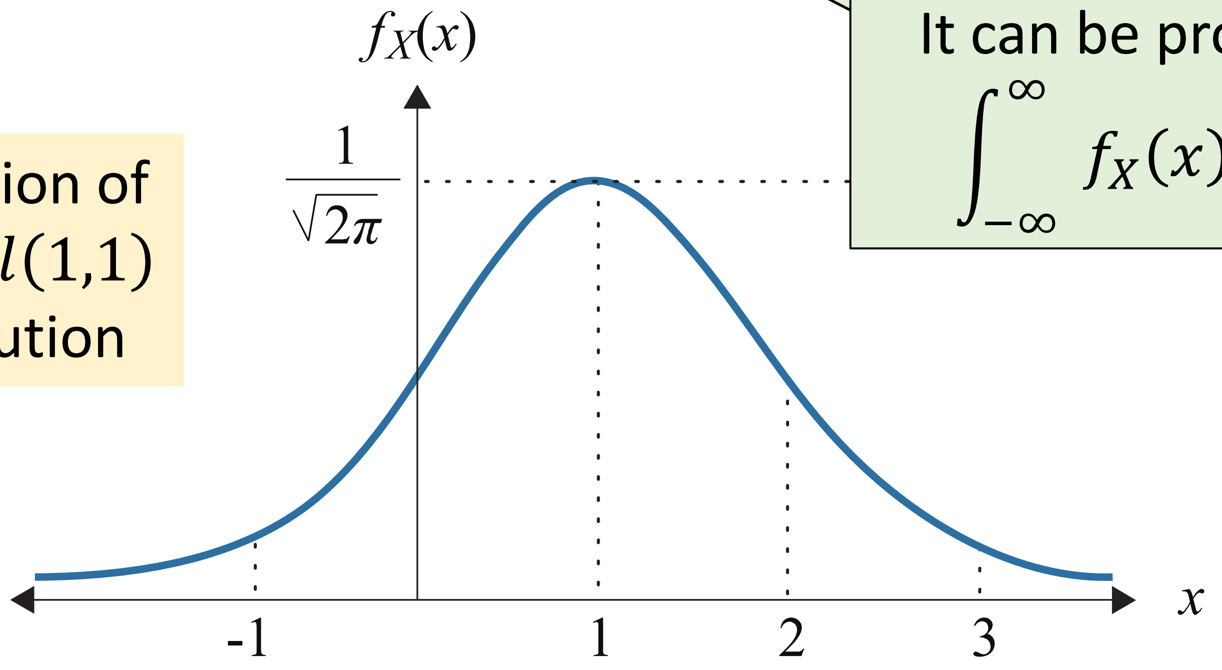
<u>Defn</u>: $X \sim Normal(\mu, \sigma^2)$ if

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad -\infty < x < \infty$$

It can be proven that

$$\int_{-\infty}^{\infty} f_X(x)dx = 1$$

Illustration of $Normal(1,1)$ distribution



Mean and Variance of Normal distribution

Theorem: Let $X \sim Normal(\mu, \sigma^2)$, then $\boldsymbol{E}[X] = \mu$ and $\boldsymbol{Var}(X) = \sigma^2$.

Proof – Part 1:
Given that $Normal(\mu, \sigma^2)$ is symmetric around μ, it follows that $\boldsymbol{E}[X] = \mu$.

Proof – Part 2:
Remains to show that $\boldsymbol{Var}(X) = \sigma^2$.

$$\boldsymbol{Var}(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f_{X(x)} dx \quad = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{1}{2}\cdot\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$y = \frac{x-\mu}{\sigma}$$

$$dx = \sigma\, dy$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}} dy$$

Variance of Normal distribution

Theorem: Let $X \sim Normal(\mu, \sigma^2)$, then $E[X] = \mu$ and $Var(X) = \sigma^2$.

Continued:

$$\boldsymbol{Var}(X) = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}}\, dy = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y \cdot \left(y e^{-\frac{y^2}{2}} \right) dy$$

via integration by parts

$$= \frac{\sigma^2}{\sqrt{2\pi}} \left(-y e^{-\frac{y^2}{2}} \right) \Big|_{-\infty}^{\infty} + \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \sigma^2$$

WHY?

$$y = \frac{x - \mu}{\sigma}$$

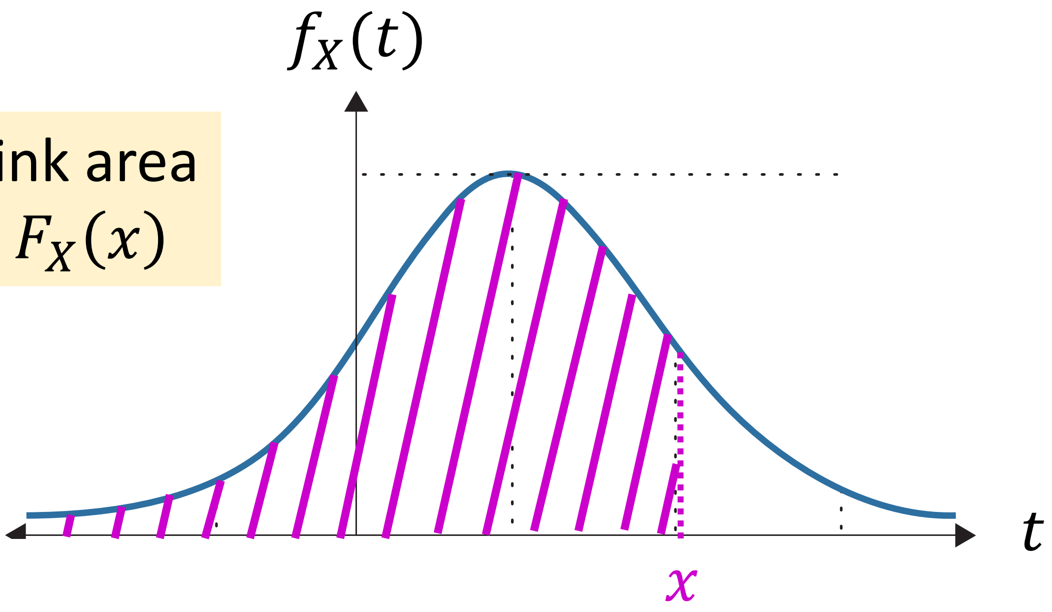
$$dx = \sigma\, dy$$

c.d.f. of Normal distribution

<u>Defn</u>: If $X \sim Normal(0,1)$, then the c.d.f. of X is denoted by

$$\mathbf{\Phi}(x) = F_X(x) = \boldsymbol{P}\{X \leq x\} = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt$$

Pink area is $F_X(x)$



Unfortunately, $\mathbf{\Phi}(x)$ is not known in closed form.

We compute it numerically, or look it up in a table of pre-computed values.

Linear Transformation Property

Theorem 9.5: (Linear Transformation Property) Let $X \sim Normal(\mu, \sigma^2)$. Let

$$Y = aX + b,$$

where $a > 0$ and $b \in \mathbb{R}$. Then $Y \sim Normal(a\mu + b, a^2\sigma^2)$.

Proof:

Clearly $\boldsymbol{E}[Y] = a\boldsymbol{E}[X] + b = a\mu + b$ ✓

Clearly $\boldsymbol{Var}(Y) = a^2\boldsymbol{Var}(X) = a^2\sigma^2$ ✓

All that's left: Show $f_Y(y)$ has Normal shape.

WTS: $$f_Y(y) = \frac{1}{\sqrt{2\pi}\,(a\sigma)} e^{-\frac{1}{2}\left(\frac{y-(a\mu+b)}{a\sigma}\right)^2}$$

Linear Transformation Property

Theorem 9.5: (Linear Transformation Property) Let $X \sim Normal(\mu, \sigma^2)$. Let

$$Y = aX + b,$$

where $a > 0$ and $b \in \mathbb{R}$. Then $Y \sim Normal(a\mu + b, a^2\sigma^2)$.

Proof cont:

<u>WTS</u>: $f_Y(y) = \frac{1}{\sqrt{2\pi}\,(a\sigma)} e^{-\frac{1}{2}\left(\frac{y-(a\mu+b)}{a\sigma}\right)^2}$

<u>Attempt 1</u>:

$$f_Y(y) = \mathbf{P}\{Y = y\} = \mathbf{P}\{aX + b = y\} = \mathbf{P}\left\{X = \frac{y-b}{a}\right\} = f_X\left(\frac{y-b}{a}\right)$$

Q: Do you see the flaw?

Linear Transformation Property

Theorem 9.5: (Linear Transformation Property) Let $X \sim Normal(\mu, \sigma^2)$. Let

$$Y = aX + b,$$

where $a > 0$ and $b \in \mathbb{R}$. Then $Y \sim Normal(a\mu + b, a^2\sigma^2)$.

Proof cont:

WTS: $f_Y(y) = \frac{1}{\sqrt{2\pi}\,(a\sigma)} e^{-\frac{1}{2}\left(\frac{y-(a\mu+b)}{a\sigma}\right)^2}$

Attempt 1:

$$f_Y(y) = \boldsymbol{P}\{Y = y\} = \boldsymbol{P}\{aX + b = y\} = \boldsymbol{P}\left\{X = \frac{y-b}{a}\right\} = f_X\left(\frac{y-b}{a}\right)$$

FALSE!

Linear Transformation Property

Theorem 9.5: (Linear Transformation Property) Let $X \sim Normal(\mu, \sigma^2)$. Let

$$Y = aX + b,$$

where $a > 0$ and $b \in \mathbb{R}$. Then $Y \sim Normal(a\mu + b, a^2\sigma^2)$.

Proof cont:

WTS: $$f_Y(y) = \frac{1}{\sqrt{2\pi}\,(a\sigma)} e^{-\frac{1}{2}\left(\frac{y-(a\mu+b)}{a\sigma}\right)^2}$$

Correct solution requires going through c.d.f., which represents valid probability

$$F_Y(y) = \boldsymbol{P}\{Y \le y\} = \boldsymbol{P}\{aX + b \le y\} = \boldsymbol{P}\left\{X \le \frac{y-b}{a}\right\} = F_X\left(\frac{y-b}{a}\right)$$

$$\frac{d}{dy}F_Y(y) \quad \longleftrightarrow \quad \frac{d}{dy}F_X\left(\frac{y-b}{a}\right)$$

These two derivatives must be equal!

Linear Transformation Property

Theorem 9.5: (Linear Transformation Property) Let $X \sim Normal(\mu, \sigma^2)$. Let

$$Y = aX + b,$$

where $a > 0$ and $b \in \mathbb{R}$. Then $Y \sim Normal(a\mu + b, a^2\sigma^2)$.

Proof cont:

$$\frac{d}{dy} F_Y(y) \quad \longleftrightarrow \quad \frac{d}{dy} F_X\left(\frac{y-b}{a}\right)$$

These two derivatives must be equal!

$$\frac{d}{dy} F_Y(y) = \frac{d}{dy} \int_{-\infty}^{y} f_Y(t) dt \overset{\text{FTC}}{=} f_Y(y)$$

$$\frac{d}{dy} F_X\left(\frac{y-b}{a}\right) = \frac{d}{dy} \int_{-\infty}^{\frac{y-b}{a}} f_X(t) dt \overset{\text{FTC}}{=} f_X\left(\frac{y-b}{a}\right) \cdot \frac{d}{dy}\left(\frac{y-b}{a}\right) = \frac{1}{a} \cdot f_X\left(\frac{y-b}{a}\right)$$

Hence,

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

Linear Transformation Property

Theorem 9.5: (Linear Transformation Property) Let $X \sim Normal(\mu, \sigma^2)$. Let

$$Y = aX + b,$$

where $a > 0$ and $b \in \mathbb{R}$. Then $Y \sim Normal(a\mu + b, a^2\sigma^2)$.

Proof cont:

$$f_Y(y) = \frac{1}{a} \cdot f_X\left(\frac{y-b}{a}\right) = \frac{1}{a\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}\cdot\left(\frac{y-b}{a}-\mu\right)^2}$$

$$= \frac{1}{\sqrt{2\pi}\,(a\sigma)} e^{-\frac{1}{2a^2\sigma^2}\cdot(y-b-a\mu)^2}$$

$$= \frac{1}{\sqrt{2\pi}\,(a\sigma)} e^{-\frac{1}{2}\left(\frac{y-(a\mu+b)}{a\sigma}\right)^2}$$

⇒ $Y \sim Normal(a\mu + b, a^2\sigma^2)$ ■

Back to $\Phi(x)$

Defn: If $Y \sim Normal(0,1)$, then the c.d.f. of X is denoted by

$$\mathbf{\Phi}(y) = F_Y(y) = \boldsymbol{P}\{Y \leq y\} = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{y} e^{-\frac{t^2}{2}}\,dt$$

Unfortunately, $\mathbf{\Phi}(y)$ is not known in closed form.

We compute it numerically, or look it up in a table of pre-computed values.

y	0.5	1.0	1.5	2.0	2.5	3.0
$\mathbf{\Phi}(y)$	0.69	0.84	0.93	0.97	0.99	0.999

Values are rounded.

Q: If $Y \sim Normal(0,1)$, what is the probability that Y is within 1 std of its mean?

y	0.5	1.0	1.5	2.0	2.5	3.0
$\boldsymbol{\Phi}(y)$	0.69	0.84	0.93	0.97	0.99	0.999

A:

$$\boldsymbol{P}\{-1 < Y < 1\} = \boldsymbol{P}\{Y < 1\} - \boldsymbol{P}\{Y < -1\}$$

$$= \boldsymbol{P}\{Y < 1\} - \boldsymbol{P}\{Y > 1\}$$ (by symmetry)

$$= \boldsymbol{P}\{Y < 1\} - (1 - \boldsymbol{P}\{Y < 1\})$$

$$= 2\boldsymbol{P}\{Y < 1\} - 1$$

$$= 2\boldsymbol{\Phi}(1) - 1$$

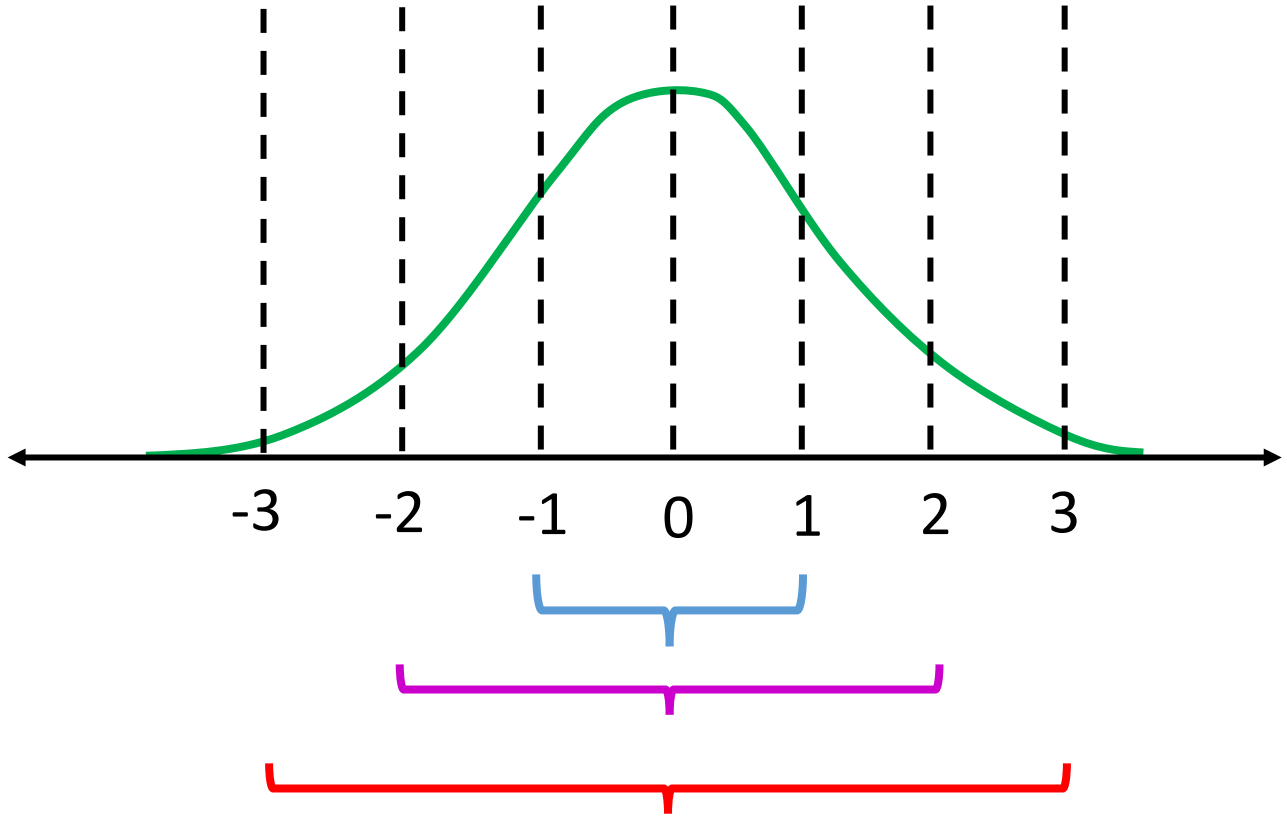
$$\approx 2 \cdot 0.84 - 1 = 0.68$$

Q: What's the probability that Y is within k std of its mean?

A: $2\boldsymbol{\Phi}(k) - 1$

Deviation from mean

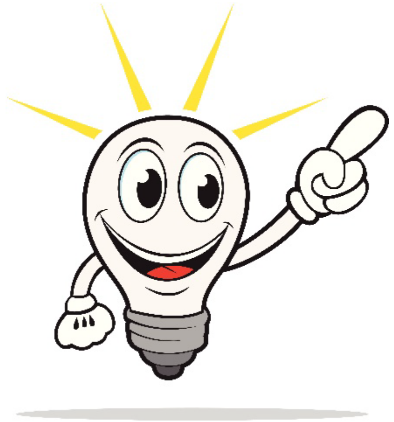
If $Y \sim Normal(0,1)$, then $\mathbf{P}\{-k < Y < k\} = 2\boldsymbol{\Phi}(k) - 1$



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

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

Bottom line: Everything that you saw for a standard Normal holds for general Normal (provided it's phrased in terms of stds).



Key Idea:

$$X \sim Normal(\mu, \sigma^2) \iff Y = \frac{X - \mu}{\sigma} \sim Normal(0, 1)$$

$$\boldsymbol{P}\{-k\sigma < X - \mu < k\sigma\} = \boldsymbol{P}\left\{-k < \frac{X - \mu}{\sigma} < k\right\} = \boldsymbol{P}\{-k < Y < k\}$$

Prob. X deviates from its mean by k stds

Prob. Y deviates from its mean by k stds

Example: Gifted Folks

Human intelligence (IQ) is thought to be Normally distributed with mean 100 and std 15.

The "gifted cutoff" is 130.

Q: What fraction of people have IQ greater than the gifted cutoff?

A: Phrased in terms of stds, we're asking what fraction of people have IQ which is more than 2 stds above the mean.

$$1 - \mathbf{\Phi}(2) \approx 0.023$$

So about 2.3%

Sum of two independent Normals

Theorem 9.7: (Sum of two indpt Normals)

Let $X \sim Normal(\mu_X, \sigma_X^2)$. Let $Y \sim Normal(\mu_Y, \sigma_Y^2)$. Assume $X \perp Y$.

Let

$$W = X + Y.$$

Then:

$$W \sim Normal(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2).$$

The proof depends on Laplace transforms. See Exercise 11.10 for the proof.

Central Limit Theorem (CLT)

The CLT is about what happens when we sum up a large number of i.i.d. random variables.

The common example is many i.i.d. sources of noise that occur at once.

CLT (at a high level) says that the distribution of the average tends towards Normal, even though the original distributions are NOT Normal.

Central Limit Theorem (CLT)

Let $X_1, X_2, X_3, \ldots, X_n$ be i.i.d. random variables with mean μ and variance σ^2.

Let

$$S_n = X_1 + X_2 + \cdots + X_n$$

Q: What is the mean and std of S_n?

A: $\mathbf{E}[S_n] = n\mu$; $\mathbf{Var}(S_n) = n\,\sigma^2$; $\mathbf{Std}(S_n) = \sigma\sqrt{n}$

Let

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

Z_n is a normalized version of S_n

Q: What is the mean and std of Z_n?

A: $\mathbf{E}[Z_n] = 0$; $\mathbf{Var}(S_n) = 1$; $\mathbf{Std}(S_n) = 1$

Central Limit Theorem (CLT)

Central Limit Theorem: Let $X_1, X_2, , \ldots, X_n$ be a sequence of i.i.d. r.v.s with common mean μ and finite variance σ^2 . Define

$$S_n = \sum_{i=1}^{n} X_i \quad \text{and} \quad Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

Then the distribution of Z_n converges to $Normal(0,1)$ as $n \to \infty$.
That is,

$$\lim_{n\to\infty} \boldsymbol{P}\{Z_n \le z\} = \boldsymbol{\Phi}(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{x^2}{2}} dx$$

for every z.

Proof uses Laplace transforms so it is deferred to Chpt 11.

Central Limit Theorem (CLT)

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} \to Normal(0,1)$$

Q: What does this say about the distribution of $S_n = \sum_{i=1}^{n} X_i$?

A: By the Linear Transformation Property, S_n should also be getting closer to a Normal distribution as $n \to \infty$.

But there are some caveats:

- $S_n \to Normal(n\mu, n\sigma^2)$. This is well-defined for finite n but not for infinite n.
- There are problems that come from looking at a sum, rather than an average. For example, if the X_i are all discrete, then S_n will also be discrete (but with a bell shape).

Central Limit Theorem (CLT)

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} \to Normal(0,1)$$

Q: What does this say about the distribution of $A_n = \frac{1}{n}\sum_{i=1}^{n} X_i$?

A: By the Linear Transformation Property, A_n should also be getting closer to a Normal distribution as $n \to \infty$.

$$A_n \to Normal\left(\mu\ , \frac{\sigma^2}{n}\right)$$

Note this doesn't have the issues of S_n

What happens to $\boldsymbol{Var}(A_n)$ as $n \to \infty$?

Central Limit Theorem (CLT)

$$A_n = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad\qquad A_n \to Normal\left(\mu\ , \frac{\sigma^2}{n}\right)$$

How can the above be correct?
Suppose that the X_i are people's heights. They can't be negative!

With extremely high probability, the value of A_n is near 5.5 ft,
where the shape looks Normal.

CLT Example

Problem:

We're trying to transmit a signal.

During transmission there are 100 indpt sources of noise, each $\sim Uniform(-1, 1)$.

If

$$|Total\ Noise| < 10$$

then the signal is not corrupted.

What is the probability that the signal is not corrupted?

CLT Example

Problem: We're trying to transmit a signal. During transmission there are 100 indpt sources of noise, each $\sim Uniform(-1, 1)$. If

$$|Total\ Noise| < 10$$

then the signal is not corrupted. What is the probability of this?

Let X_i denote the noise from source i.

Q: What is $\boldsymbol{E}[X_i]$? $\boldsymbol{E}[X_i] = 0$

Q: What is $\boldsymbol{Var}(X_i)$? $\boldsymbol{Var}(X_i) = \frac{(\mathrm{b} - \mathrm{a})^2}{12} = \frac{1}{3}$

Q: What is σ_{X_i}? $\sigma_{X_i} = \frac{1}{\sqrt{3}}$

CLT Example

Problem: We're trying to transmit a signal. During transmission there are 100 indpt sources of noise, each $\sim Uniform(-1,1)$. If

$$|Total\ Noise| < 10$$

then the signal is not corrupted. What is the probability of this?

Let $S_{100} = X_1 + X_2 + \cdots + X_{100}$

Want: $\boldsymbol{P}\{-10 < S_{100} < 10\}$

$$\boldsymbol{P}\{-10 < S_{100} < 10\} = \boldsymbol{P}\left\{-\frac{10}{10/\sqrt{3}} < \frac{S_{100} - 0}{10/\sqrt{3}} < \frac{10}{10/\sqrt{3}}\right\}$$

$$\approx \boldsymbol{P}\{-\sqrt{3} < Normal(0,1) < \sqrt{3}\}$$

$$= 2\boldsymbol{\Phi}(\sqrt{3}) - 1 \approx 0.91$$

$$\boldsymbol{E}[X_i] = 0$$

$$\boldsymbol{Var}(X_i) = \frac{1}{3}$$

$$\boldsymbol{E}[S_{100}] = 0$$

$$\boldsymbol{Var}(S_{100}) = \frac{100}{3}$$

$$\sigma_{S_{100}} = \frac{10}{\sqrt{3}}$$