Normal Distribution

CS109
Review
Two Continuous RVs: Uniform and Exponential

### Uniform Random Variable

**Notation:** $X \sim \text{Uni}(\alpha, \beta)$

**Description:** A continuous random variable that takes on values, with equal likelihood, between $\alpha$ and $\beta$.

**Parameters:**
- $\alpha \in \mathbb{R}$, the minimum value of the variable.
- $\beta \in \mathbb{R}$, $\beta > \alpha$, the maximum value of the variable.

**Support:** $x \in [\alpha, \beta]$.

**PDF equation:**
- $f(x) = \frac{1}{\beta - \alpha}$ for $x \in [\alpha, \beta]$.
- $0$ else.

**CDF equation:**
- $F(x) = \frac{x - \alpha}{\beta - \alpha}$ for $x \in [\alpha, \beta]$.
- $0$ for $x < \alpha$.
- $1$ for $x > \beta$.

**Expectation:** $E[X] = \frac{1}{2} (\alpha + \beta)$.

**Variance:** $\text{Var}(X) = \frac{1}{12} (\beta - \alpha)^2$.

**PDF graph:**
- Parameter $\alpha$: 0
- Parameter $\beta$: 1

### Exponential Random Variable

**Notation:** $X \sim \text{Exp}(\lambda)$

**Description:** Time until next events if (a) the events occur with a constant mean rate and (b) they occur independently of time since last event.

**Parameters:**
- $\lambda \in \{0, 1, \ldots\}$, the constant average rate.

**Support:** $x \in \mathbb{R}^+$.

**PDF equation:** $f(x) = \lambda e^{-\lambda x}$.

**CDF equation:** $F(x) = 1 - e^{-\lambda x}$.

**Expectation:** $E[X] = 1/\lambda$.

**Variance:** $\text{Var}(X) = 1/\lambda^2$.

**PDF graph:**
- Parameter $\lambda$: 5
$X \sim \text{Uniform}(0,1)$: A Continuous Random Variable

Possible values are between 0 and 1

All values are equally likely

$P(0 \leq X \leq 1) = 1$

$P(0.5 \leq X \leq 1) = 0.5$

$P(0.5 \leq X \leq 0.6) = 0.1$

Finding the probability of a range of values is straightforward!
$X \sim \text{Uniform}(0,1)$: A Continuous Random Variable

All values are equally likely

Possible values are between 0 and 1

- $P(0 \leq X \leq 1) = 1$
- $P(0.5 \leq X \leq 1) = 0.5$
- $P(0.5 \leq X \leq 0.6) = 0.1$
- $P(X = 0.5) = 0$

Because of infinitely many outcomes, the probability of any exact outcome is zero

No PMFs!
The probability density function (PDF) of a continuous random variable represents the relative likelihood of various values.

Units: probability divided by units of $X$, or the derivative of the probability of $x$. **Integrate it** to get probabilities!
The probability density function (PDF) of a continuous random variable represents the relative likelihood of various values. Units: probability divided by units of \( X \), or the derivative of the probability of \( x \). **Integrate it** to get probabilities!

\[
P(a < X < b) = \int_{x=a}^{b} f(X = x) \, dx
\]
PDFs - \( f(X = x) \) vs. PMFs - \( P(X = x) \)

\[ P(X = x) \quad \text{“The probability that a \textit{discrete} random variable X takes on the value } x \text{.”} \]

\[ f(X = x) \quad \text{“The \textit{derivative} of the probability that a \textit{continuous} random variable X takes at the value } x \text{.”} \]
PDFs - $f(X = x)$ vs. PMFs - $P(X = x)$

$P(X = x)$

“The probability that a **discrete** random variable $X$ takes on the value $x$.”

$f(X = x)$

“The **derivative** of the probability that a **continuous** random variable $X$ takes at the value $x$.”

What do you get if you integrate over a probability **density function**?

A probability!

They are *both* measures of how **likely** $X$ is to take on the value $x$. 
A cumulative density function (CDF) is a “closed-form” equation for the probability that a continuous random variable is less than a given value.

\[ F(x) = P(X < x) \]

\[ P(X < x) = \int_{y=-\infty}^{x} f(y) \, dy \]
A cumulative density function (CDF) is a “closed-form” equation for the probability that a continuous random variable is less than a given value.

\[ F(x) = P(X < x) \]

\[ P(X < x) = \int_{y=-\infty}^{x} f(y) \, dy \]

For random variables that have cumulative density functions, we can avoid integrals!

Piech & Cain, CS109, Stanford University
Example: $X \sim \text{Exp}(\lambda = 1)$

**Probability Density Function**

$f(x) = \lambda e^{-\lambda x}$

$P(X < 2) = \int_{x=-\infty}^{2} f(x) \, dx$
Example: $X \sim \text{Exp}(\lambda = 1)$

**Probability Density Function**

$$f(x) = \lambda e^{-\lambda x}$$

**Cumulative Density Function**

$$F_X(x) = P(X < x) = \int_{-\infty}^{x} f(y) \, dy$$

$$F(x) = 1 - e^{-\lambda x}$$

$$P(X < 2) = \int_{x=-\infty}^{2} f(x) \, dx$$

or

$$F(2) = 1 - e^{-2} \approx 0.84$$
Example: $X \sim \text{Exp}(\lambda = 1)$

**Probability Density Function**

$$f(x) = \lambda e^{-\lambda x}$$

$$P(1 < X < 2)$$

$$= \int_{x=1}^{2} f(x) \, dx$$
Example: $X \sim \text{Exp}(\lambda = 1)$

*Probability Density Function*

$$f(x) = \lambda e^{-\lambda x}$$

*Cumulative Density Function*

$$F_X(x) = P(X < x) = \int_{y=-\infty}^{x} f(y) \, dy$$

$$F(x) = 1 - e^{-\lambda x}$$

Probability Density Function

Cumulative Density Function

$$F(2) - F(1) = (1 - e^{-2}) - (1 - e^{-1}) \approx 0.23$$

$$P(1 < X < 2) = \int_{x=1}^{2} f(x) \, dx$$

or

$$F(2) - F(1)$$

Piech & Cain, CS109, Stanford University
How Long Until the Next Big Earthquake?

Based on historical data, major earthquakes (with magnitude 8.0+) happen at a rate of 0.002 per year*. What is the probability of a major earthquake in the next 30 years?

Let $Y$ be years until the next earthquake of magnitude 8.0+.

**Exponential PDF:**

$$f_Y(y) = \lambda e^{-\lambda y}$$

**Exponential CDF:**

$$F_Y(y) = 1 - e^{-\lambda y}$$

$$P(Y < 30) = \int_0^{30} \frac{1}{500} e^{-\frac{y}{500}} dy$$

$$= \left[-e^{-\frac{y}{500}}\right]_0^{30} = -e^{\frac{30}{500}} + e^0 \approx 0.058$$

$$P(Y < 30) = 1 - e^{-\frac{30}{500}} \approx 0.058$$

*In California, according to the USGS, 2015*
End Review
The most famous continuous random variable
Normal (Gaussian) Random Variable

Support: $(-\infty, \infty)$

$X \sim \mathcal{N}(\mu, \sigma^2)$
Normal (Gaussian) Random Variable

Support: \((-\infty, \infty)\)

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

PDF:

\[ f(X = x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x - \mu)^2}{2\sigma^2}} \]
Normal (Gaussian) Random Variable

Support: \((-\infty, \infty)\)

PDF:

\[
f(X = x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x - \mu)^2}{2\sigma^2}}
\]
Anatomy of a The Normal PDF

\[ f(X = x) = \frac{1}{\sigma \sqrt{2\pi}} e^{\frac{-(x - \mu)^2}{2\sigma^2}} \]
Anatomy of a The Normal PDF

\[ f(X = x) = \frac{1}{\sigma \sqrt{2\pi}} e^{\frac{-(x - \mu)^2}{2\sigma^2}} \]

- A constant:
  - makes the integral over all possible outcomes sum to 1

Piech & Cain, CS109, Stanford University
Anatomy of a The Normal PDF

\[ f(X = x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x - \mu)^2}{2\sigma^2}} \]

- **distance to the mean**: makes the PDF symmetric around the mean.
- **a constant**: makes the integral over all possible outcomes sum to 1.
Anatomy of a The Normal PDF

\[ f(X = x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x - \mu)^2}{2\sigma^2}} \]

- Distance to the mean (makes the PDF symmetric around the mean)
- A constant: makes the integral over all possible outcomes sum to 1
- ...normalized by the variance

Piech & Cain, CS109, Stanford University
Carl Friedrich Gauss (1777-1855)

- German mathematician
- Sort-of invented the normal distribution
- Also astronomer, geologist, physicist
- Super influential in a lot of fields
Carl Friedrich Gauss (1777-1855)

- German mathematician
- Sort-of invented the normal distribution
- Also astronomer, geologist, physicist
- Super influential in a lot of fields

Looks like Robin Williams
Why the Normal?
Why the Normal?

- Common for natural phenomena: human height, weight, shoe sizes, etc.

(random example from Kelly’s research)
Why the Normal?

• Common for natural phenomena: human height, weight, shoe sizes, etc.
• A lot of noise in the world is Normal
  • E.g. random errors in measurements, residuals in linear regression
Why the Normal?

- Common for natural phenomena: human height, weight, shoe sizes, etc.
- A lot of noise in the world is Normal
  - E.g. random errors in measurements, residuals in linear regression
- The sum of many random variables often looks Normal (spoilers)
- Sample means are distributed normally – important for statistics
Why the Normal?

- Common for natural phenomena: human height, weight, shoe sizes, etc.
- A lot of noise in the world is Normal
  - E.g. random errors in measurements, residuals in linear regression
- The sum of many random variables often looks Normal (spoilers)
- Sample means are distributed normally – important for statistics
- Even things that aren’t Normal might fit a normal-related distribution
Why the Normal?

- Common for natural phenomena: human height, weight, shoe sizes, etc.
- A lot of noise in the world is Normal
  - E.g. random errors in measurements, residuals in linear regression
- The sum of many random variables often looks Normal (spoilers)
- Sample means are distributed normally – important for statistics
- Even things that aren’t Normal might fit a normal-related distribution

People also just assume things are normally distributed a lot.

- They can do this in part because the Normal is so common
- But there’s a deeper reason to it...
“The simplest explanation is usually the best one”
When We Fit Models To Data, We Try To Keep It Simple
This curve fits the data well, but does it really represent the distribution? Or is it “overfit”, so that the curve captures too much of the noise?
This curve fits the data about as well, but appears to overfit less. We could say that this simpler distribution makes fewer assumptions. The formal concept for this idea is entropy.
This curve fits the data about as well, but appears to overfit less.

We could say that this simpler distribution makes fewer assumptions.

The formal concept for this idea is entropy.

When we fit models to data, we try to keep it simple.

The Normal distribution is the simplest distribution, that makes the fewest assumptions (has maximum entropy), for a given mean and variance.
Let’s Try It Out: Cybertruck Manufacturing

Your team is tasked with producing the side panels for cybertrucks. Elon Musk requires all panels to be built “accurate within 10 microns”. You check how precise your manufacturing is, and find these stats:

• Average panel thickness: $\mu = 500$ microns
• Variance of thickness: $\sigma^2 = 36$ microns$^2$

What fraction of the panels you manufacture will meet Elon’s standards?
Let’s Try It Out: Cybertruck Manufacturing

Your team is tasked with producing the side panels for cybertrucks. Elon Musk requires all panels to be built “accurate within 10 microns”. You check how precise your manufacturing is, and find these stats:

- Average panel thickness: $\mu = 500$ microns
- Variance of thickness: $\sigma^2 = 36$ microns$^2$

What fraction of the panels you manufacture will meet Elon’s standards?

$$X \sim \mathcal{N}(\mu = 500, \sigma^2 = 36)$$
Let’s Try It Out: Cybertruck Manufacturing

Your team is tasked with producing the side panels for cybertrucks. Elon Musk requires all panels to be built "accurate within 10 microns". You check how precise your manufacturing is, and find these stats:

- Average panel thickness: \( \mu = 500 \) microns
- Variance of thickness: \( \sigma^2 = 36 \) microns\(^2\)

What fraction of the panels you manufacture will meet Elon’s standards?

\[
X \sim \mathcal{N}(\mu = 500, \sigma^2 = 36)
\]

\[
P(490 \leq X \leq 510) = \int_{490}^{510} f(X = x)dx = \int_{490}^{510} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x - \mu)^2}{2\sigma^2}} dx
\]
Let’s Try It Out: Cybertruck Manufacturing

Your team is tasked with producing the side panels for cybertrucks. Elon Musk requires all panels to be built ”accurate within 10 microns”. You check how precise your manufacturing is, and find these stats:

- Average panel thickness: $\mu = 500$ microns
- Variance of thickness: $\sigma^2 = 36$ microns$^2$

What fraction of the panels you manufacture will meet Elon’s standards?

$$X \sim \mathcal{N} (\mu = 500, \sigma^2 = 36)$$

$$P(490 \leq X \leq 510) = \int_{490}^{510} f(X = x)dx = \int_{490}^{510} \frac{1}{\sigma \sqrt{2\pi}} e^{- \frac{(x - \mu)^2}{2\sigma^2}} dx$$
There is no closed form for the integral of this PDF
There is no closed form for the integral of this PDF

So no CDF???
The Standard Normal: \( Z \sim N(\mu = 0, \sigma^2 = 1) \)
The Standard Normal: $Z \sim N(\mu = 0, \sigma^2 = 1)$

For the Standard Normal, we have a CDF!

$$F(x) = \phi(x)$$

$\phi(0.5) = P(Z < 0.5)$
What Does The Phi Function Look Like? Oh

Standard Normal Cumulative Probability Table

Cumulative probabilities for POSITIVE z-values are shown in the following table:

<table>
<thead>
<tr>
<th>z</th>
<th>0.00</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.06</th>
<th>0.07</th>
<th>0.08</th>
<th>0.09</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.5000</td>
<td>0.5040</td>
<td>0.5080</td>
<td>0.5120</td>
<td>0.5160</td>
<td>0.5199</td>
<td>0.5239</td>
<td>0.5279</td>
<td>0.5319</td>
<td>0.5359</td>
</tr>
<tr>
<td>0.1</td>
<td>0.5398</td>
<td>0.5438</td>
<td>0.5478</td>
<td>0.5517</td>
<td>0.5557</td>
<td>0.5596</td>
<td>0.5636</td>
<td>0.5675</td>
<td>0.5714</td>
<td>0.5753</td>
</tr>
<tr>
<td>0.2</td>
<td>0.5793</td>
<td>0.5832</td>
<td>0.5871</td>
<td>0.5910</td>
<td>0.5948</td>
<td>0.5987</td>
<td>0.6026</td>
<td>0.6064</td>
<td>0.6103</td>
<td>0.6141</td>
</tr>
<tr>
<td>0.3</td>
<td>0.6179</td>
<td>0.6217</td>
<td>0.6255</td>
<td>0.6293</td>
<td>0.6331</td>
<td>0.6368</td>
<td>0.6406</td>
<td>0.6443</td>
<td>0.6480</td>
<td>0.6517</td>
</tr>
<tr>
<td>0.4</td>
<td>0.6554</td>
<td>0.6591</td>
<td>0.6628</td>
<td>0.6664</td>
<td>0.6700</td>
<td>0.6736</td>
<td>0.6772</td>
<td>0.6808</td>
<td>0.6844</td>
<td>0.6879</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6915</td>
<td>0.6950</td>
<td>0.6985</td>
<td>0.7019</td>
<td>0.7054</td>
<td>0.7088</td>
<td>0.7123</td>
<td>0.7157</td>
<td>0.7190</td>
<td>0.7224</td>
</tr>
<tr>
<td>0.6</td>
<td>0.7257</td>
<td>0.7291</td>
<td>0.7324</td>
<td>0.7357</td>
<td>0.7389</td>
<td>0.7422</td>
<td>0.7454</td>
<td>0.7486</td>
<td>0.7517</td>
<td>0.7549</td>
</tr>
<tr>
<td>0.7</td>
<td>0.7580</td>
<td>0.7611</td>
<td>0.7642</td>
<td>0.7673</td>
<td>0.7704</td>
<td>0.7734</td>
<td>0.7764</td>
<td>0.7794</td>
<td>0.7823</td>
<td>0.7852</td>
</tr>
<tr>
<td>0.8</td>
<td>0.7881</td>
<td>0.7910</td>
<td>0.7939</td>
<td>0.7967</td>
<td>0.7995</td>
<td>0.8023</td>
<td>0.8051</td>
<td>0.8078</td>
<td>0.8106</td>
<td>0.8133</td>
</tr>
<tr>
<td>0.9</td>
<td>0.8159</td>
<td>0.8186</td>
<td>0.8212</td>
<td>0.8238</td>
<td>0.8264</td>
<td>0.8289</td>
<td>0.8315</td>
<td>0.8340</td>
<td>0.8365</td>
<td>0.8389</td>
</tr>
<tr>
<td>1.0</td>
<td>0.8413</td>
<td>0.8438</td>
<td>0.8461</td>
<td>0.8485</td>
<td>0.8508</td>
<td>0.8531</td>
<td>0.8554</td>
<td>0.8577</td>
<td>0.8599</td>
<td>0.8621</td>
</tr>
<tr>
<td>1.1</td>
<td>0.8643</td>
<td>0.8665</td>
<td>0.8688</td>
<td>0.8708</td>
<td>0.8729</td>
<td>0.8749</td>
<td>0.8770</td>
<td>0.8790</td>
<td>0.8810</td>
<td>0.8830</td>
</tr>
<tr>
<td>1.2</td>
<td>0.8849</td>
<td>0.8869</td>
<td>0.8888</td>
<td>0.8907</td>
<td>0.8925</td>
<td>0.8944</td>
<td>0.8962</td>
<td>0.8980</td>
<td>0.8997</td>
<td>0.9015</td>
</tr>
<tr>
<td>1.3</td>
<td>0.9032</td>
<td>0.9049</td>
<td>0.9066</td>
<td>0.9082</td>
<td>0.9099</td>
<td>0.9115</td>
<td>0.9131</td>
<td>0.9147</td>
<td>0.9162</td>
<td>0.9177</td>
</tr>
<tr>
<td>1.4</td>
<td>0.9192</td>
<td>0.9207</td>
<td>0.9222</td>
<td>0.9236</td>
<td>0.9251</td>
<td>0.9265</td>
<td>0.9279</td>
<td>0.9292</td>
<td>0.9306</td>
<td>0.9319</td>
</tr>
</tbody>
</table>
What Does The Phi Function Look Like? Oh

Standard Normal Cumulative Probability Table

$\Phi(0.54) = 0.7054$

Cumulative probabilities for POSITIVE z-values are shown in the following table:

<table>
<thead>
<tr>
<th>z</th>
<th>0.00</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.06</th>
<th>0.07</th>
<th>0.08</th>
<th>0.09</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.5000</td>
<td>0.5040</td>
<td>0.5080</td>
<td>0.5120</td>
<td>0.5160</td>
<td>0.5199</td>
<td>0.5239</td>
<td>0.5279</td>
<td>0.5319</td>
<td>0.5359</td>
</tr>
<tr>
<td>0.1</td>
<td>0.5358</td>
<td>0.5438</td>
<td>0.5518</td>
<td>0.5597</td>
<td>0.5676</td>
<td>0.5754</td>
<td>0.5832</td>
<td>0.5910</td>
<td>0.5987</td>
<td>0.6064</td>
</tr>
<tr>
<td>0.2</td>
<td>0.5793</td>
<td>0.5832</td>
<td>0.5871</td>
<td>0.5910</td>
<td>0.5948</td>
<td>0.5987</td>
<td>0.6026</td>
<td>0.6064</td>
<td>0.6103</td>
<td>0.6141</td>
</tr>
<tr>
<td>0.3</td>
<td>0.6179</td>
<td>0.6217</td>
<td>0.6255</td>
<td>0.6293</td>
<td>0.6331</td>
<td>0.6369</td>
<td>0.6406</td>
<td>0.6443</td>
<td>0.6480</td>
<td>0.6517</td>
</tr>
<tr>
<td>0.4</td>
<td>0.6554</td>
<td>0.6591</td>
<td>0.6628</td>
<td>0.6664</td>
<td>0.6700</td>
<td>0.6736</td>
<td>0.6772</td>
<td>0.6808</td>
<td>0.6844</td>
<td>0.6879</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6915</td>
<td>0.6950</td>
<td>0.6985</td>
<td>0.7019</td>
<td>0.7054</td>
<td>0.7088</td>
<td>0.7123</td>
<td>0.7157</td>
<td>0.7190</td>
<td>0.7224</td>
</tr>
<tr>
<td>0.6</td>
<td>0.7257</td>
<td>0.7291</td>
<td>0.7324</td>
<td>0.7357</td>
<td>0.7389</td>
<td>0.7422</td>
<td>0.7454</td>
<td>0.7486</td>
<td>0.7517</td>
<td>0.7549</td>
</tr>
<tr>
<td>0.7</td>
<td>0.7580</td>
<td>0.7611</td>
<td>0.7642</td>
<td>0.7673</td>
<td>0.7704</td>
<td>0.7734</td>
<td>0.7764</td>
<td>0.7794</td>
<td>0.7823</td>
<td>0.7852</td>
</tr>
<tr>
<td>0.8</td>
<td>0.7881</td>
<td>0.7910</td>
<td>0.7939</td>
<td>0.7967</td>
<td>0.7995</td>
<td>0.8023</td>
<td>0.8051</td>
<td>0.8078</td>
<td>0.8106</td>
<td>0.8133</td>
</tr>
<tr>
<td>0.9</td>
<td>0.8159</td>
<td>0.8186</td>
<td>0.8212</td>
<td>0.8238</td>
<td>0.8264</td>
<td>0.8289</td>
<td>0.8315</td>
<td>0.8340</td>
<td>0.8365</td>
<td>0.8389</td>
</tr>
<tr>
<td>1.0</td>
<td>0.8413</td>
<td>0.8438</td>
<td>0.8461</td>
<td>0.8485</td>
<td>0.8508</td>
<td>0.8531</td>
<td>0.8554</td>
<td>0.8577</td>
<td>0.8599</td>
<td>0.8621</td>
</tr>
<tr>
<td>1.1</td>
<td>0.8643</td>
<td>0.8665</td>
<td>0.8686</td>
<td>0.8708</td>
<td>0.8729</td>
<td>0.8749</td>
<td>0.8770</td>
<td>0.8790</td>
<td>0.8810</td>
<td>0.8830</td>
</tr>
<tr>
<td>1.2</td>
<td>0.8849</td>
<td>0.8869</td>
<td>0.8888</td>
<td>0.8907</td>
<td>0.8925</td>
<td>0.8944</td>
<td>0.8962</td>
<td>0.8980</td>
<td>0.8997</td>
<td>0.9015</td>
</tr>
<tr>
<td>1.3</td>
<td>0.9032</td>
<td>0.9049</td>
<td>0.9066</td>
<td>0.9082</td>
<td>0.9099</td>
<td>0.9115</td>
<td>0.9131</td>
<td>0.9147</td>
<td>0.9162</td>
<td>0.9177</td>
</tr>
<tr>
<td>1.4</td>
<td>0.9192</td>
<td>0.9207</td>
<td>0.9222</td>
<td>0.9236</td>
<td>0.9251</td>
<td>0.9265</td>
<td>0.9279</td>
<td>0.9292</td>
<td>0.9306</td>
<td>0.9319</td>
</tr>
</tbody>
</table>
The Standard Normal: $Z \sim N(\mu = 0, \sigma^2 = 1)$

For the Standard Normal, we have a CDF!

$$F(x) = \phi(x)$$

A function that has been solved for us numerically.

Our probability ancestors did the work of solving for the CDF of the standard normal.

How do we use this for any normal distribution?
Fun Fact: The Linear Transform of a Normal Is...Normal

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

is also Normal.
Fun Fact: The Linear Transform of a Normal Is...Normal

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

$Y = aX + b$

is also Normal.

What would the mean and variance of $Y$ be?
Fun Fact: The Linear Transform of a Normal Is...Normal

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

$Y = aX + b$

is also Normal.

What would the mean and variance of $Y$ be?

$E[Y] = E[aX + b]$

$= aE[X] + b$

$= a\mu + b$

Linearity property of expectation!
Let $X \sim \mathcal{N}(\mu, \sigma^2)$ and $Y = aX + b$. $Y$ is also Normal.

What would the mean and variance of $Y$ be?

$$E[Y] = E[aX + b] = aE[X] + b = a\mu + b$$

$$Var(Y) = Var(aX + b) = a^2Var(X) = a^2\sigma^2$$
Let $X \sim \mathcal{N} (\mu, \sigma^2)$ be a Normal distribution. Then the linear transform $Y = aX + b$ is also Normal. 

What would the mean and variance of $Y$ be?

$$E[Y] = E[aX + b] = aE[X] + b = a\mu + b$$

$$Var(Y) = Var(aX + b) = a^2 Var(X) = a^2 \sigma^2$$

So, $Y \sim \mathcal{N} (a\mu + b, a^2 \sigma^2)$.
Let’s Linear-Transform $X$ into $Z$, The Standard Normal

Let $X \sim \mathcal{N}(\mu, \sigma^2)$  $\rightarrow$  $Y = aX + b$  $\rightarrow$  $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$ is also Normal

What linear transform of $X$ would get us to $Z$?
Let’s Linear-Transform $X$ into $Z$, The Standard Normal

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

is also Normal

What linear transform of $X$ would get us to $Z$?

$$Z = \frac{X - \mu}{\sigma} = \frac{1}{\sigma}X - \frac{\mu}{\sigma}$$
Let’s Linear-Transform X into Z, The Standard Normal

Let $X \sim \mathcal{N}(\mu, \sigma^2) \implies Y = aX + b \implies Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$
is also Normal

What linear transform of $X$ would get us to $Z$?

$$Z = \frac{X - \mu}{\sigma} = \frac{1}{\sigma} X - \frac{\mu}{\sigma}$$

$$a = \frac{1}{\sigma} \quad b = -\frac{\mu}{\sigma}$$
Let’s Linear-Transform $X$ into $Z$, The Standard Normal

Let $X \sim \mathcal{N}(\mu, \sigma^2)$ \implies $Y = aX + b$ \implies $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$ is also Normal

What linear transform of $X$ would get us to $Z$?

$$Z = \frac{X - \mu}{\sigma} = \frac{1}{\sigma}X - \frac{\mu}{\sigma}$$

$$a = \frac{1}{\sigma} \quad b = -\frac{\mu}{\sigma}$$

If we plug in these values for $a$ and $b$, we get the standard normal:

$$Z \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$$

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

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

Piech & Cain, CS109, Stanford University
Let’s Linear-Transform X into Z, The Standard Normal

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

is also Normal

What linear transform of \( X \) would get us to \( Z \)?

\[
Z = \frac{X - \mu}{\sigma} = \frac{1}{\sigma}X - \frac{\mu}{\sigma} \quad a = \frac{1}{\sigma} \quad b = -\frac{\mu}{\sigma}
\]

If we plug in these values for \( a \) and \( b \), we get the standard normal:

\[
Z \sim \mathcal{N}(a\mu + b, a^2\sigma^2) \sim \mathcal{N}\left(\frac{\mu}{\sigma} - \frac{\mu}{\sigma}, \frac{\sigma^2}{\sigma^2}\right) \sim \mathcal{N}(0, 1)
\]

Piech & Cain, CS109, Stanford University
How Do We Use This?

Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Use the fact that $Z = \frac{X - \mu}{\sigma}$ to compute the CDF for $X$.

$$F_X(x) = P(X \leq x)$$
Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Use the fact that $Z = \frac{X - \mu}{\sigma}$ to compute the CDF for $X$.

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

$$= P(X - \mu \leq x - \mu)$$

$$= P \left( \frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma} \right)$$

Apply linear transform to both sides
Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Use the fact that $Z = \frac{X - \mu}{\sigma}$ to compute the CDF for $X$.

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

Apply linear transform to both sides

$$= P(X - \mu \leq x - \mu)$$

Recognize that left-hand side is $Z$

$$= P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right)$$

$$= P\left(Z \leq \frac{x - \mu}{\sigma}\right)$$
How Do We Use This?

Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Use the fact that $Z = \frac{X - \mu}{\sigma}$ to compute the CDF for $X$.

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

Apply linear transform to both sides

$$= P(X - \mu \leq x - \mu)$$

Recognize that left-hand side is $Z$

$$= P \left( \frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma} \right)$$

Recognize that the whole expression is the CDF

$$= P \left( Z \leq \frac{x - \mu}{\sigma} \right)$$

$$= \Phi \left( \frac{x - \mu}{\sigma} \right)$$
General CDF For Any Normal Random Variable

The cumulative density function of any normal, $X \sim \mathcal{N} (\mu, \sigma^2)$:

$$F(x) = \Phi \left( \frac{x - \mu}{\sigma} \right)$$

To calculate $P(X < x)$, for any normally distributed $X$, we transform $X$ to the standard normal, $Z$, and then use phi.
General CDF For Any Normal Random Variable

The cumulative density function of any normal, $X \sim \mathcal{N}(\mu, \sigma^2)$:

$$F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

To calculate $P(X < x)$, for any normally distributed $X$, we transform $X$ to the standard normal, $Z$, and then use phi.

Piech & Cain, CS109, Stanford University
Do We Have To Use The Table??

Standard Normal Cumulative Probability Table

Cumulative probabilities for POSITIVE z-values are shown in the following table:

<table>
<thead>
<tr>
<th>z</th>
<th>0.00</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.06</th>
<th>0.07</th>
<th>0.08</th>
<th>0.09</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.5000</td>
<td>0.5040</td>
<td>0.5080</td>
<td>0.5120</td>
<td>0.5160</td>
<td>0.5199</td>
<td>0.5239</td>
<td>0.5279</td>
<td>0.5319</td>
<td>0.5359</td>
</tr>
<tr>
<td>0.1</td>
<td>0.5398</td>
<td>0.5438</td>
<td>0.5478</td>
<td>0.5517</td>
<td>0.5557</td>
<td>0.5596</td>
<td>0.5636</td>
<td>0.5675</td>
<td>0.5714</td>
<td>0.5753</td>
</tr>
<tr>
<td>0.2</td>
<td>0.5793</td>
<td>0.5832</td>
<td>0.5871</td>
<td>0.5910</td>
<td>0.5948</td>
<td>0.5987</td>
<td>0.6026</td>
<td>0.6064</td>
<td>0.6103</td>
<td>0.6141</td>
</tr>
<tr>
<td>0.3</td>
<td>0.6179</td>
<td>0.6217</td>
<td>0.6255</td>
<td>0.6293</td>
<td>0.6331</td>
<td>0.6368</td>
<td>0.6406</td>
<td>0.6443</td>
<td>0.6480</td>
<td>0.6517</td>
</tr>
<tr>
<td>0.4</td>
<td>0.6554</td>
<td>0.6591</td>
<td>0.6628</td>
<td>0.6664</td>
<td>0.6700</td>
<td>0.6736</td>
<td>0.6772</td>
<td>0.6808</td>
<td>0.6844</td>
<td>0.6879</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6915</td>
<td>0.6950</td>
<td>0.6985</td>
<td>0.7019</td>
<td>0.7054</td>
<td>0.7088</td>
<td>0.7123</td>
<td>0.7157</td>
<td>0.7190</td>
<td>0.7224</td>
</tr>
<tr>
<td>0.6</td>
<td>0.7267</td>
<td>0.7291</td>
<td>0.7324</td>
<td>0.7357</td>
<td>0.7389</td>
<td>0.7422</td>
<td>0.7454</td>
<td>0.7486</td>
<td>0.7517</td>
<td>0.7549</td>
</tr>
<tr>
<td>0.7</td>
<td>0.7580</td>
<td>0.7611</td>
<td>0.7642</td>
<td>0.7673</td>
<td>0.7704</td>
<td>0.7734</td>
<td>0.7764</td>
<td>0.7794</td>
<td>0.7823</td>
<td>0.7852</td>
</tr>
<tr>
<td>0.8</td>
<td>0.7881</td>
<td>0.7910</td>
<td>0.7939</td>
<td>0.7967</td>
<td>0.7995</td>
<td>0.8023</td>
<td>0.8051</td>
<td>0.8078</td>
<td>0.8106</td>
<td>0.8133</td>
</tr>
<tr>
<td>0.9</td>
<td>0.8159</td>
<td>0.8186</td>
<td>0.8212</td>
<td>0.8238</td>
<td>0.8264</td>
<td>0.8289</td>
<td>0.8315</td>
<td>0.8340</td>
<td>0.8365</td>
<td>0.8389</td>
</tr>
<tr>
<td>1.0</td>
<td>0.8413</td>
<td>0.8438</td>
<td>0.8461</td>
<td>0.8485</td>
<td>0.8508</td>
<td>0.8531</td>
<td>0.8554</td>
<td>0.8577</td>
<td>0.8599</td>
<td>0.8621</td>
</tr>
<tr>
<td>1.1</td>
<td>0.8643</td>
<td>0.8665</td>
<td>0.8688</td>
<td>0.8708</td>
<td>0.8729</td>
<td>0.8749</td>
<td>0.8770</td>
<td>0.8790</td>
<td>0.8810</td>
<td>0.8830</td>
</tr>
<tr>
<td>1.2</td>
<td>0.8849</td>
<td>0.8869</td>
<td>0.8888</td>
<td>0.8907</td>
<td>0.8925</td>
<td>0.8944</td>
<td>0.8962</td>
<td>0.8980</td>
<td>0.8997</td>
<td>0.9015</td>
</tr>
<tr>
<td>1.3</td>
<td>0.9032</td>
<td>0.9049</td>
<td>0.9066</td>
<td>0.9082</td>
<td>0.9099</td>
<td>0.9115</td>
<td>0.9131</td>
<td>0.9147</td>
<td>0.9162</td>
<td>0.9177</td>
</tr>
<tr>
<td>1.4</td>
<td>0.9192</td>
<td>0.9207</td>
<td>0.9222</td>
<td>0.9236</td>
<td>0.9251</td>
<td>0.9265</td>
<td>0.9279</td>
<td>0.9292</td>
<td>0.9306</td>
<td>0.9319</td>
</tr>
</tbody>
</table>
Do We Have To Use The Table??

<table>
<thead>
<tr>
<th>z</th>
<th>0.00</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.06</th>
<th>0.07</th>
<th>0.08</th>
<th>0.09</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.5000</td>
<td>0.5040</td>
<td>0.5080</td>
<td>0.5119</td>
<td>0.5159</td>
<td>0.5196</td>
<td>0.5236</td>
<td>0.5279</td>
<td>0.5319</td>
<td>0.5359</td>
</tr>
<tr>
<td>0.1</td>
<td>0.5398</td>
<td>0.5438</td>
<td>0.5478</td>
<td>0.5517</td>
<td>0.5554</td>
<td>0.5590</td>
<td>0.5625</td>
<td>0.5659</td>
<td>0.5692</td>
<td>0.5724</td>
</tr>
<tr>
<td>0.2</td>
<td>0.5793</td>
<td>0.5832</td>
<td>0.5869</td>
<td>0.5905</td>
<td>0.5939</td>
<td>0.5972</td>
<td>0.6004</td>
<td>0.6036</td>
<td>0.6067</td>
<td>0.6103</td>
</tr>
<tr>
<td>0.3</td>
<td>0.6179</td>
<td>0.6217</td>
<td>0.6253</td>
<td>0.6288</td>
<td>0.6321</td>
<td>0.6353</td>
<td>0.6385</td>
<td>0.6416</td>
<td>0.6446</td>
<td>0.6476</td>
</tr>
<tr>
<td>0.4</td>
<td>0.6554</td>
<td>0.6591</td>
<td>0.6626</td>
<td>0.6659</td>
<td>0.6691</td>
<td>0.6722</td>
<td>0.6752</td>
<td>0.6781</td>
<td>0.6810</td>
<td>0.6838</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6915</td>
<td>0.6950</td>
<td>0.6984</td>
<td>0.7017</td>
<td>0.7049</td>
<td>0.7081</td>
<td>0.7112</td>
<td>0.7142</td>
<td>0.7171</td>
<td>0.7200</td>
</tr>
<tr>
<td>0.6</td>
<td>0.7257</td>
<td>0.7291</td>
<td>0.7324</td>
<td>0.7355</td>
<td>0.7385</td>
<td>0.7415</td>
<td>0.7444</td>
<td>0.7472</td>
<td>0.7500</td>
<td>0.7527</td>
</tr>
<tr>
<td>0.7</td>
<td>0.7580</td>
<td>0.7611</td>
<td>0.7641</td>
<td>0.7670</td>
<td>0.7699</td>
<td>0.7727</td>
<td>0.7755</td>
<td>0.7782</td>
<td>0.7809</td>
<td>0.7835</td>
</tr>
<tr>
<td>0.8</td>
<td>0.7881</td>
<td>0.7910</td>
<td>0.7938</td>
<td>0.7965</td>
<td>0.7991</td>
<td>0.8016</td>
<td>0.8041</td>
<td>0.8065</td>
<td>0.8089</td>
<td>0.8113</td>
</tr>
<tr>
<td>0.9</td>
<td>0.8159</td>
<td>0.8186</td>
<td>0.8212</td>
<td>0.8237</td>
<td>0.8261</td>
<td>0.8284</td>
<td>0.8306</td>
<td>0.8328</td>
<td>0.8349</td>
<td>0.8370</td>
</tr>
<tr>
<td>1.0</td>
<td>0.8413</td>
<td>0.8438</td>
<td>0.8461</td>
<td>0.8485</td>
<td>0.8508</td>
<td>0.8531</td>
<td>0.8554</td>
<td>0.8577</td>
<td>0.8599</td>
<td>0.8621</td>
</tr>
<tr>
<td>1.1</td>
<td>0.8643</td>
<td>0.8665</td>
<td>0.8686</td>
<td>0.8708</td>
<td>0.8729</td>
<td>0.8749</td>
<td>0.8770</td>
<td>0.8790</td>
<td>0.8810</td>
<td>0.8830</td>
</tr>
<tr>
<td>1.2</td>
<td>0.8849</td>
<td>0.8869</td>
<td>0.8888</td>
<td>0.8907</td>
<td>0.8925</td>
<td>0.8944</td>
<td>0.8962</td>
<td>0.8980</td>
<td>0.8997</td>
<td>0.9015</td>
</tr>
<tr>
<td>1.3</td>
<td>0.9032</td>
<td>0.9049</td>
<td>0.9066</td>
<td>0.9082</td>
<td>0.9099</td>
<td>0.9115</td>
<td>0.9131</td>
<td>0.9147</td>
<td>0.9162</td>
<td>0.9177</td>
</tr>
<tr>
<td>1.4</td>
<td>0.9192</td>
<td>0.9207</td>
<td>0.9222</td>
<td>0.9236</td>
<td>0.9251</td>
<td>0.9265</td>
<td>0.9279</td>
<td>0.9292</td>
<td>0.9306</td>
<td>0.9319</td>
</tr>
</tbody>
</table>
We Are Computer Scientists!

Every modern programming language has phi stored in a library:

```python
from scipy import stats
stats.norm.cdf(x, mean, std)
```

\[ = P(X < x) \text{ where } X \sim \mathcal{N}(\mu, \sigma^2) \]
We Are Computer Scientists!

Every modern programming language has phi stored in a library:

```python
from scipy import stats
stats.norm.cdf(x, mean, std) = P(X < x) where X \sim \mathcal{N}(\mu, \sigma^2)
```

not variance!!!
Every modern programming language has phi stored in a library:

```python
from scipy import stats
stats.norm.cdf(x, mean, std)
```

= \( P(X < x) \) where \( X \sim \mathcal{N}(\mu, \sigma^2) \)

The course reader also has a calculator:

---

Piech & Cain, CS109, Stanford University
Fun Ways To Use Phi

\[ P(c < Z < d) = \phi(d) - \phi(c) \]
Fun Ways To Use Phi

\[ P(c < Z < d) = \phi(d) - \phi(c) \]

\[ \phi(-a) = 1 - \phi(a) \]
Practice: Cybertruck Manufacturing

Your team is tasked with producing the side panels for cybertrucks. Elon Musk requires all panels to be built "accurate within 10 microns". You check how precise your manufacturing is, and find these stats:

- Average panel thickness: $\mu = 500$ microns
- Variance of thickness: $\sigma^2 = 36$ microns²

What fraction of the panels you manufacture will meet Elon’s standards?

$X \sim \mathcal{N}(\mu = 500, \sigma^2 = 36)$

$$P(490 \leq X \leq 510) = \int_{490}^{510} f(X = x)dx$$
Your team is tasked with producing the side panels for cybertrucks. Elon Musk requires all panels to be built "accurate within 10 microns". You check how precise your manufacturing is, and find these stats:

- Average panel thickness: $\mu = 500$ microns
- Variance of thickness: $\sigma^2 = 36$ microns$^2$

What fraction of the panels you manufacture will meet Elon’s standards?

$$X \sim \mathcal{N}(\mu = 500, \sigma^2 = 36)$$

$$P(490 \leq X \leq 510) = ?$$
Practice: Cybertruck Manufacturing

Your team is tasked with producing the side panels for cybertrucks. Elon Musk requires all panels to be built "accurate within 10 microns". You check how precise your manufacturing is, and find these stats:

- Average panel thickness: $\mu = 500$ microns
- Variance of thickness: $\sigma^2 = 36$ microns$^2$

What fraction of the panels you manufacture will meet Elon’s standards?

$$X \sim \mathcal{N}(\mu = 500, \sigma^2 = 36)$$

$$P(490 \leq X \leq 510) = P(X < 510) - P(X < 490) = \Phi\left(\frac{510 - 500}{6}\right) - \Phi\left(\frac{490 - 500}{6}\right)$$

subtract mean, divide by std. dev.
Your team is tasked with producing the side panels for cybertrucks. Elon Musk requires all panels to be built "accurate within 10 microns". You check how precise your manufacturing is, and find these stats:

- Average panel thickness: $\mu = 500$ microns
- Variance of thickness: $\sigma^2 = 36$ microns$^2$

What fraction of the panels you manufacture will meet Elon’s standards?

$$X \sim \mathcal{N}(\mu = 500, \sigma^2 = 36)$$

$$P(490 \leq X \leq 510) = P(X < 510) - P(X < 490) = \Phi \left( \frac{510 - 500}{6} \right) - \Phi \left( \frac{490 - 500}{6} \right)$$

$$= \Phi \left( \frac{5}{3} \right) - \left( 1 - \Phi \left( \frac{5}{3} \right) \right) = 2 \Phi \left( \frac{5}{3} \right) - 1 \approx 0.904$$
Get your Gaussian On

Let $X \sim \mathcal{N}(\mu = 3, \sigma^2 = 16)$. Std deviation $\sigma = 4$.

1. $P(X > 0)$

- If $X \sim \mathcal{N}(\mu, \sigma^2)$, then
  $$F(x) = \Phi \left( \frac{x - \mu}{\sigma} \right)$$
- Symmetry of the PDF of Normal RV implies
  $$\Phi(-z) = 1 - \Phi(z)$$
Get your Gaussian On

Let $X \sim \mathcal{N} (\mu = 3, \sigma^2 = 16)$. Note standard deviation $\sigma = 4$.

How would you write each of the below probabilities as a function of the standard normal CDF, $\Phi$?

1. $P(X > 0)$ (we just did this)
2. $P(2 < X < 5)$
3. $P(|X - 3| > 6)$

• If $X \sim \mathcal{N} (\mu, \sigma^2)$, then $F(x) = \Phi \left( \frac{x-\mu}{\sigma} \right)$
• Symmetry of the PDF of Normal RV implies $\Phi(-z) = 1 - \Phi(z)$

Piech & Cain, CS109, Stanford University
Get your Gaussian On

Let $X \sim \mathcal{N}(\mu = 3, \sigma^2 = 16)$. Std deviation $\sigma = 4$.

1. $P(X > 0)$
2. $P(2 < X < 5)$

- If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $F(x) = \Phi \left( \frac{x-\mu}{\sigma} \right)$
- Symmetry of the PDF of Normal RV implies $\Phi(-z) = 1 - \Phi(z)$
Get your Gaussian On

Let $X \sim \mathcal{N} (\mu = 3, \sigma^2 = 16)$. Std deviation $\sigma = 4$.

1. $P(X > 0)$
2. $P(2 < X < 5)$
3. $P(|X - 3| > 6)$

Compute $z = \frac{x - \mu}{\sigma}$

$P(X < -3) + P(X > 9)$

$= F(-3) + (1 - F(9))$

$= \Phi \left( \frac{-3 - 3}{4} \right) + \left( 1 - \Phi \left( \frac{9 - 3}{4} \right) \right)$

- If $X \sim \mathcal{N} (\mu, \sigma^2)$, then $F(x) = \Phi \left( \frac{x - \mu}{\sigma} \right)$
- Symmetry of the PDF of Normal RV implies $\Phi(-x) = 1 - \Phi(x)$

Look up $\Phi(z)$ in table

Piech & Cain, CS109, Stanford University
Get your Gaussian On

Let $X \sim \mathcal{N}(\mu = 3, \sigma^2 = 16)$. Std deviation $\sigma = 4$.

1. $P(X > 0)$
2. $P(2 < X < 5)$
3. $P(|X - 3| > 6)$

Compute $z = \frac{x - \mu}{\sigma}$

$P(X < -3) + P(X > 9)$

$= F(-3) + (1 - F(9))$

$= \Phi\left(\frac{-3 - 3}{4}\right) + (1 - \Phi\left(\frac{9 - 3}{4}\right))$

Look up $\Phi(z)$ in table

$= \Phi\left(-\frac{3}{2}\right) + \left(1 - \Phi\left(\frac{3}{2}\right)\right)$

$= 2 \left(1 - \Phi\left(\frac{3}{2}\right)\right)$

$\approx 0.1337$

• If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$

• Symmetry of the PDF of Normal RV implies $\Phi(-x) = 1 - \Phi(x)$

Piech & Cain, CS109, Stanford University
The Normal can also approximate the Binomial
Poisson Approximates Binomial, With Extreme $n$ and $p$
Normal Approximates Binomial, With Moderate $p$

The shapes are the same!

Just set the normal’s $\mu, \sigma^2$ to be the mean and variance of the binomial.

Piech & Cain, CS109, Stanford University
Two Ways To Approximate The Binomial

\[ X \sim \text{Bin}(n, p) \]

\[ E[X] = np \]
\[ \text{Var}(X) = np(1 - p) \]

\[ Y \sim \text{Poi}(\lambda) \]
\[ \lambda = np \]

\[ Y \sim \mathcal{N}(\mu, \sigma^2) \]
\[ \mu = np \]
\[ \sigma^2 = np(1 - p) \]

Poisson approximation for big \( n \), small \( p \).
Normal approximation for big \( n \), medium \( p \).

Piech & Cain, CS109, Stanford University
Website Testing

A new website design is tested out on 100 users.

- Let $X$ be the number of users whose time on the site increases with the new design.
- The CEO will endorse the new design if $X \geq 65$.

What is $P(\text{CEO endorses change} | \text{it has no effect})$?
Website Testing

A new website design is tested out on 100 users.

- Let $X$ be the number of users whose time on the site increases with the new design.
- The CEO will endorse the new design if $X \geq 65$.

What is $P(\text{CEO endorses change}|\text{ it has no effect})$?

Without approximation: $X \sim \text{Bin}(n = 100, p = 0.5)$

$$P(X \geq 65) = \sum_{i=65}^{100} \binom{100}{i} (0.5)^i (1 - 0.5)^{100-i} \approx 0.0018$$
A new website design is tested out on 100 users.

- Let $X$ be the number of users whose time on the site increases with the new design.
- The CEO will endorse the new design if $X \geq 65$.

What is $P(\text{CEO endorses change} | \text{it has no effect})$?

Without approximation: $X \sim \text{Bin}(n = 100, p = 0.5)$
Website Testing

A new website design is tested out on 100 users.

- Let $X$ be the number of users whose time on the site increases with the new design.
- The CEO will endorse the new design if $X \geq 65$.

What is $P(\text{CEO endorses change}| \text{ it has no effect})$?

With approximation: $Y \sim \mathcal{N}(\mu, \sigma^2)$
Website Testing

A new website design is tested out on 100 users.

- Let $X$ be the number of users whose time on the site increases with the new design.
- The CEO will endorse the new design if $X \geq 65$.

What is $P(\text{CEO endorses change| it has no effect})$?

With approximation: $Y \sim \mathcal{N}(\mu, \sigma^2)$

\[
\begin{align*}
\mu &= np = 50 \\
\sigma^2 &= np(1 - p) = 25 \\
\sigma &= \sqrt{25} = 5
\end{align*}
\]
Website Testing

A new website design is tested out on 100 users.

- Let $X$ be the number of users whose time on the site increases with the new design.
- The CEO will endorse the new design if $X \geq 65$.

What is $P(\text{CEO endorses change} \mid \text{it has no effect})$?

With approximation:

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

$$\mu = np = 50$$
$$\sigma^2 = np(1 - p) = 25$$
$$\sigma = \sqrt{25} = 5$$

$$P(X \geq 65) \approx P(Y \geq 65) = 1 - F_Y(65)$$
Website Testing

A new website design is tested out on 100 users.

- Let $X$ be the number of users whose time on the site increases with the new design.
- The CEO will endorse the new design if $X \geq 65$.

What is $P(\text{CEO endorses change} | \text{it has no effect})$?

With approximation:

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

$$P(X \geq 65) \approx P(Y \geq 65) = 1 - F_Y(65)$$

$$= 1 - \Phi \left( \frac{65 - 50}{5} \right) = 1 - \Phi(3) \approx 0.0013$$
Website Testing, With **Continuity Correction**

\( Y \sim \mathcal{N}(50, 25) \) approximates \( X \sim \text{Bin}(100, 0.5) \), but \( P(X \geq 65) \neq P(Y \geq 65) \)?

\[
P(X \geq 65) \quad \text{Binomial} \\
\approx P(Y \geq 64.5) \quad \text{Normal} \\
\approx 0.0018 \quad \text{✅ the better Approach 2}
\]

*We have to **continuity correct** when we approximate a Binomial using a Normal.*

---

Piech & Cain, CS109, Stanford University
Continuity Correction Practice

$Y \sim \mathcal{N}(np, np(1 - p))$ approximates $X \sim \text{Bin}(n, p)$.

How do we approximate the following probabilities?

<table>
<thead>
<tr>
<th>Discrete (e.g., Binomial) probability question</th>
<th>Continuous (Normal) probability question</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X = 6)$</td>
<td></td>
</tr>
<tr>
<td>$P(X \geq 6)$</td>
<td></td>
</tr>
<tr>
<td>$P(X &gt; 6)$</td>
<td></td>
</tr>
<tr>
<td>$P(X &lt; 6)$</td>
<td></td>
</tr>
<tr>
<td>$P(X \leq 6)$</td>
<td></td>
</tr>
</tbody>
</table>

Piech & Cain, CS109, Stanford University
Continuity Correction Practice

$Y \sim \mathcal{N}(np, np(1-p))$ approximates $X \sim \text{Bin}(n, p)$. How do we approximate the following probabilities?

<table>
<thead>
<tr>
<th>Discrete (e.g., Binomial) probability question</th>
<th>Continuous (Normal) probability question</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X = 6)$</td>
<td>$P(5.5 \leq Y \leq 6.5)$</td>
</tr>
<tr>
<td>$P(X \geq 6)$</td>
<td>$P(Y \geq 5.5)$</td>
</tr>
<tr>
<td>$P(X &gt; 6)$</td>
<td>$P(Y \geq 6.5)$</td>
</tr>
<tr>
<td>$P(X &lt; 6)$</td>
<td>$P(Y \leq 5.5)$</td>
</tr>
<tr>
<td>$P(X \leq 6)$</td>
<td>$P(Y \leq 6.5)$</td>
</tr>
</tbody>
</table>
Practice: Stanford Admissions

Stanford accepts 2480 students.
• Each admitted student independently matriculates with probability 0.68.
• Let $X$ be the number of students who will attend.

What is $P(X > 1745)$? Give a numerical approximation.

Strategy:  
A. Just Binomial  
B. Poisson  
C. Normal  
D. None/other
Practice: Stanford Admissions

Stanford accepts 2480 students.
• Each admitted student independently matriculates with probability 0.68.
• Let $X$ be the number of students who will attend.

What is $P(X > 1745)$? Give a numerical approximation.

Strategy:  
A. Just Binomial  
B. Poisson  
C. Normal  
D. None/other
Practice: Stanford Admissions

Stanford accepts 2480 students.
- Each admitted student independently matriculates with probability 0.68.
- Let $X$ be the number of students who will attend.

What is $P(X > 1745)$? Give a numerical approximation.

$$X \sim \mathcal{N}(n = 2480, p = 0.68)$$

Step 1: define binomial, like you normally would
Practice: Stanford Admissions

Stanford accepts 2480 students.
- Each admitted student independently matriculates with probability 0.68.
- Let $X$ be the number of students who will attend.

What is $P(X > 1745)$? *Give a numerical approximation.*

$$X \sim \mathcal{N}(n = 2480, p = 0.68)$$

Let $Y \sim \mathcal{N}(E[X], \text{Var}(X))$

**Step 1:** define binomial, like you normally would

**Step 2:** define the normal that will approximate $X$
Practice: Stanford Admissions

Stanford accepts 2480 students.

- Each admitted student independently matriculates with probability 0.68.
- Let $X$ be the number of students who will attend.

What is $P(X > 1745)$? *Give a numerical approximation.*

$$X \sim \mathcal{N}(n = 2480, p = 0.68) \quad \Rightarrow \quad \text{Let } Y \sim \mathcal{N}(E[X], \text{Var}(X))$$

**Step 3: find parameters for the normal**

$$E[X] = np = 1686$$

$$\text{Var}(X) = np(1 - p) \approx 540 \rightarrow \sigma = 23.3$$
Practice: Stanford Admissions

Stanford accepts 2480 students.
- Each admitted student independently matriculates with probability 0.68.
- Let $X$ be the number of students who will attend.

What is $P(X > 1745)$? *Give a numerical approximation.*

$$X \sim \mathcal{N}(n = 2480, p = 0.68)$$

Let $Y \sim \mathcal{N}(E[X], \text{Var}(X))$

$$E[X] = np = 1686$$

$$\text{Var}(X) = np(1-p) \approx 540 \rightarrow \sigma = 23.3$$

Step 4: figure out what probability you want, then continuity correct

$$P(X > 1745) \approx P(Y \geq 1745.5)$$
Practice: Stanford Admissions

Stanford accepts 2480 students.
- Each admitted student independently matriculates with probability 0.68.
- Let $X$ be the number of students who will attend.

What is $P(X > 1745)$? *Give a numerical approximation.*

\[ X \sim \mathcal{N}(n = 2480, p = 0.68) \quad \text{Let } Y \sim \mathcal{N}(E[X], \text{Var}(X)) \]

\[
E[X] = np = 1686
\]
\[
\text{Var}(X) = np(1 - p) \approx 540 \rightarrow \sigma = 23.3
\]

\[
P(X > 1745) \approx P(Y \geq 1745.5)
\]

\[
P(Y \geq 1745.5) = 1 - F(1745.5) = 1 - \Phi\left(\frac{1745.5 - 1686}{23.3}\right) = 1 - \Phi(2.54) \approx 0.0055
\]

*Step 5: solve!*

Piech & Cain, CS109, Stanford University
Challenge Problem
11.11 SINGLE'S DAY SALE
How Many Servers Is Enough?

At the busiest minute of the shopping rush, your website receives $R$ pings:

$$R \sim N(\mu = 10^6, \sigma = 10^4)$$

To anticipate the rush, you plan to buy $N$ servers. Each server can handle 10,000 pings per minute, but if it receives any more, it will drop customers. What is the smallest value of $N$ such that $P(\text{drop}) < 0.0001$?
Ponder Before Wednesday!