Foundations of Computer Science
Lecture 21

Deviations from the Mean

How Good is the Expectation as a Summary of a Random Variable?
Variance: Uniform; Bernoulli; Binomial; Waiting Times.
Variance of a Sum
Law of Large Numbers: The 3-σ Rule
Last Time

1. Expected value of a Sum.
   - Sum of dice
   - Binomial
   - Waiting time
   - Coupon collecting.

2. Build-up expectation.

3. Expected value of a product.

4. Sum of Indicators.
   - Random arrangement of hats on heads.
Today: Deviations from the Mean

1. How well does the expected value (mean) summarize a random variable?

2. Variance.


4. Law of large numbers
   - The $3-\sigma$ rule.
Probability For Analyzing a Random Experiment.
Probability For Analyzing a Random Experiment.

Experiment (random) → Outcomes (complex)
Probability For Analyzing a Random Experiment.
Probability For Analyzing a Random Experiment.
Probability For Analyzing a Random Experiment.

Experiment (random) \rightarrow \text{Outcomes (complex)} \rightarrow \text{Measurement } X \text{ (random variable)} \rightarrow \text{Summary } E[X] \text{ (expectation)} \rightarrow \text{How good is } E[X]?
**Experiment.** Roll $n$ dice and compute $X$, the average of the rolls.
**Experiment.** Roll $n$ dice and compute $X$, the average of the rolls.

$E[\text{average}]$
**Experiment.** Roll $n$ dice and compute $X$, the average of the rolls.

$$
\mathbb{E}[\text{average}] = \mathbb{E}\left[ \frac{1}{n} \cdot \text{sum} \right]
$$
**Experiment.** Roll $n$ dice and compute $X$, the average of the rolls.

\[
\mathbb{E}[^\text{average}] = \mathbb{E} \left[ \frac{1}{n} \cdot \text{sum} \right] = \frac{1}{n} \cdot \mathbb{E}[^\text{sum}]
\]
**Experiment.** Roll $n$ dice and compute $X$, the average of the rolls.

$$
\mathbb{E}[\text{average}] = \mathbb{E}\left[\frac{1}{n} \cdot \text{sum}\right] = \frac{1}{n} \cdot \mathbb{E}[\text{sum}] = \frac{1}{n} \times n \times 3\frac{1}{2}
$$
Experiment. Roll \( n \) dice and compute \( X \), the average of the rolls.

\[
E[\text{average}] = E\left[\frac{1}{n} \cdot \text{sum}\right] = \frac{1}{n} \cdot E[\text{sum}] = \frac{1}{n} \times n \times 3\frac{1}{2} = 3\frac{1}{2}.
\]
Average of $n$ Dice
Average of $n$ Dice

- **Average of 4 Dice**
  - Number of dice: $n$
  - Average roll:
    - 1: 10
    - 2: 10
    - 3: 10
    - 4: 10
    - 5: 1

- **Average of 100 Dice**
  - Number of dice: $n$
  - Average roll:
    - 1: 23
    - 3: 3
    - 5: 4
    - 6: 0

**Deviations from the Mean:** 5 / 13

**Variance:**
$X = \text{sum of 2 dice. } \mathbb{E}[X] = 7 \leftarrow \mu(X)$
Variance: Size of the Deviations From the Mean

\( X = \text{sum of 2 dice. } \mathbb{E}[X] = 7 \leftarrow \mu(X) \)

<table>
<thead>
<tr>
<th>X</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta)</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Pop Quiz. What is \(\mathbb{E}[\Delta]\)?
**Variance: Size of the Deviations From the Mean**

\( X = \) sum of 2 dice. \( \mathbb{E}[X] = 7 \leftarrow \mu(X) \)

<table>
<thead>
<tr>
<th>X</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta )</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>( P_X )</td>
<td>( \frac{1}{36} )</td>
<td>( \frac{2}{36} )</td>
<td>( \frac{3}{36} )</td>
<td>( \frac{4}{36} )</td>
<td>( \frac{5}{36} )</td>
<td>( \frac{6}{36} )</td>
<td>( \frac{5}{36} )</td>
<td>( \frac{4}{36} )</td>
<td>( \frac{3}{36} )</td>
<td>( \frac{2}{36} )</td>
<td>( \frac{1}{36} )</td>
</tr>
</tbody>
</table>

**Pop Quiz.** What is \( \mathbb{E}[\Delta] \)?

**Variance,** \( \sigma^2 \), is the expected value of the squared deviations,

\[
\sigma^2 = \mathbb{E}[\Delta^2] = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[(X - \mathbb{E}[X])^2]
\]

\[
\sigma^2 = \mathbb{E}[\Delta^2] = \frac{1}{36} \cdot 25 +
\]
Variance: Size of the Deviations From the Mean

\( X = \text{sum of 2 dice. } \mathbb{E}[X] = 7 \leftarrow \mu(X) \)

\[
\begin{array}{c|cccccccccccc}
 X & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\
 \Delta & -5 & -4 & -3 & -2 & -1 & 0 & 1 & 2 & 3 & 4 & 5 \\
 P_X & \frac{1}{36} & \frac{2}{36} & \frac{3}{36} & \frac{4}{36} & \frac{5}{36} & \frac{6}{36} & \frac{5}{36} & \frac{4}{36} & \frac{3}{36} & \frac{2}{36} & \frac{1}{36} \\
\end{array}
\]

Pop Quiz. What is \( \mathbb{E}[\Delta] \)?

Variance, \( \sigma^2 \), is the expected value of the squared deviations,

\[
\sigma^2 = \mathbb{E}[\Delta^2] = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[(X - \mathbb{E}[X])^2]
\]

\[
\sigma^2 = \mathbb{E}[\Delta^2] = \frac{1}{36} \cdot 25 + \frac{2}{36} \cdot 16 +
\]
Variance: Size of the Deviations From the Mean

\( X = \) sum of 2 dice. \( \mathbb{E}[X] = 7 \leftarrow \mu(X) \)

<table>
<thead>
<tr>
<th>X</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta )</td>
<td>(-5)</td>
<td>(-4)</td>
<td>(-3)</td>
<td>(-2)</td>
<td>(-1)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>( P_x )</td>
<td>( \frac{1}{36} )</td>
<td>( \frac{2}{36} )</td>
<td>( \frac{3}{36} )</td>
<td>( \frac{4}{36} )</td>
<td>( \frac{5}{36} )</td>
<td>( \frac{6}{36} )</td>
<td>( \frac{5}{36} )</td>
<td>( \frac{4}{36} )</td>
<td>( \frac{3}{36} )</td>
<td>( \frac{2}{36} )</td>
<td>( \frac{1}{36} )</td>
</tr>
</tbody>
</table>

Pop Quiz. What is \( \mathbb{E}[\Delta] \)?

**Variance, \( \sigma^2 \), is the expected value of the squared deviations,**

\[
\sigma^2 = \mathbb{E}[\Delta^2] = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[(X - \mathbb{E}[X])^2]
\]

\[
\sigma^2 = \mathbb{E}[\Delta^2] = \frac{1}{36} \cdot 25 + \frac{2}{36} \cdot 16 + \frac{3}{36} \cdot 9 + \frac{4}{36} \cdot 4 + \frac{5}{36} \cdot 1 + \frac{6}{36} \cdot 0 + \frac{5}{36} \cdot 1 + \frac{4}{36} \cdot 4 + \frac{3}{36} \cdot 9 + \frac{2}{36} \cdot 16 + \frac{1}{36} \cdot 25
\]
Variation: Size of the Deviations From the Mean

\(X = \text{sum of 2 dice. } \mathbb{E}[X] = 7 \leftarrow \mu(X)\)

<table>
<thead>
<tr>
<th>X</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta)</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(P_{X})</td>
<td>(\frac{1}{36})</td>
<td>(\frac{2}{36})</td>
<td>(\frac{3}{36})</td>
<td>(\frac{4}{36})</td>
<td>(\frac{5}{36})</td>
<td>(\frac{6}{36})</td>
<td>(\frac{5}{36})</td>
<td>(\frac{4}{36})</td>
<td>(\frac{3}{36})</td>
<td>(\frac{2}{36})</td>
<td>(\frac{1}{36})</td>
</tr>
</tbody>
</table>

Pop Quiz. What is \(\mathbb{E}[\Delta]\)?

**Variance,** \(\sigma^2\), is the expected value of the squared deviations,

\[
\sigma^2 = \mathbb{E}[\Delta^2] = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[(X - \mathbb{E}[X])^2]
\]

\[
\sigma^2 = \mathbb{E}[\Delta^2] = \frac{1}{36} \cdot 25 + \frac{2}{36} \cdot 16 + \frac{3}{36} \cdot 9 + \frac{4}{36} \cdot 4 + \frac{5}{36} \cdot 1 + \frac{6}{36} \cdot 0 + \frac{5}{36} \cdot 1 + \frac{4}{36} \cdot 4 + \frac{3}{36} \cdot 9 + \frac{2}{36} \cdot 16 + \frac{1}{36} \cdot 25
\]

\[
= \frac{5}{6}. 
\]
Variance: Size of the Deviations From the Mean

\( X = \text{sum of 2 dice.} \) \( \mathbb{E}[X] = 7 \leftarrow \mu(X) \)

<table>
<thead>
<tr>
<th>( X )</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta )</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>( P_X )</td>
<td>( \frac{1}{36} )</td>
<td>( \frac{2}{36} )</td>
<td>( \frac{3}{36} )</td>
<td>( \frac{4}{36} )</td>
<td>( \frac{5}{36} )</td>
<td>( \frac{6}{36} )</td>
<td>( \frac{5}{36} )</td>
<td>( \frac{4}{36} )</td>
<td>( \frac{3}{36} )</td>
<td>( \frac{2}{36} )</td>
<td>( \frac{1}{36} )</td>
</tr>
</tbody>
</table>

Pop Quiz. What is \( \mathbb{E}[\Delta] \)?

Variance, \( \sigma^2 \), is the expected value of the squared deviations,

\[
\sigma^2 = \mathbb{E}[\Delta^2] = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[(X - \mathbb{E}[X])^2]
\]

\[
\sigma^2 = \mathbb{E}[\Delta^2] = \frac{1}{36} \cdot 25 + \frac{2}{36} \cdot 16 + \frac{3}{36} \cdot 9 + \frac{4}{36} \cdot 4 + \frac{5}{36} \cdot 1 + \frac{6}{36} \cdot 0 + \frac{5}{36} \cdot 1 + \frac{4}{36} \cdot 4 + \frac{3}{36} \cdot 9 + \frac{2}{36} \cdot 16 + \frac{1}{36} \cdot 25
\]

\[
= 5 \frac{5}{6}.
\]

Standard Deviation, \( \sigma \), is the square-root of the variance,

\[
\sigma = \sqrt{\mathbb{E}[\Delta^2]} = \sqrt{\mathbb{E}[(X - \mu)^2]} = \sqrt{\mathbb{E}[(X - \mathbb{E}[X])^2]}
\]
Variance: Size of the Deviations From the Mean

\( X = \) sum of 2 dice. \( \mathbb{E}[X] = 7 \) \( \leftarrow \mu(X) \)

<table>
<thead>
<tr>
<th>( X )</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta )</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>( P_x )</td>
<td>\frac{1}{36}</td>
<td>\frac{2}{36}</td>
<td>\frac{3}{36}</td>
<td>\frac{4}{36}</td>
<td>\frac{5}{36}</td>
<td>\frac{6}{36}</td>
<td>\frac{5}{36}</td>
<td>\frac{4}{36}</td>
<td>\frac{3}{36}</td>
<td>\frac{2}{36}</td>
<td>\frac{1}{36}</td>
</tr>
</tbody>
</table>

Pop Quiz. What is \( \mathbb{E}[\Delta] \)?

**Variance, \( \sigma^2 \),** is the expected value of the squared deviations,

\[
\sigma^2 = \mathbb{E}[\Delta^2] = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[(X - \mathbb{E}[X])^2]
\]

\[
\sigma^2 = \mathbb{E}[\Delta^2] = \frac{1}{36} \cdot 25 + \frac{2}{36} \cdot 16 + \frac{3}{36} \cdot 9 + \frac{4}{36} \cdot 4 + \frac{5}{36} \cdot 1 + \frac{6}{36} \cdot 0 + \frac{5}{36} \cdot 1 + \frac{4}{36} \cdot 4 + \frac{3}{36} \cdot 9 + \frac{2}{36} \cdot 16 + \frac{1}{36} \cdot 25
\]

\[
= 5\frac{5}{6}.
\]

**Standard Deviation, \( \sigma \),** is the square-root of the variance,

\[
\sigma = \sqrt{\mathbb{E}[\Delta^2]} = \sqrt{\mathbb{E}[(X - \mu)^2]} = \sqrt{\mathbb{E}[(X - \mathbb{E}[X])^2]}
\]

\[
\sigma = \sqrt{\frac{5\frac{5}{6}}{6}} \approx 2.52
\]

sum of two dice rolls = 7 ± 2.52.
Variance: Size of the Deviations From the Mean

X = sum of 2 dice. \( \mathbb{E}[X] = 7 \leftarrow \mu(X) \)

<table>
<thead>
<tr>
<th>X</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta)</td>
<td>−5</td>
<td>−4</td>
<td>−3</td>
<td>−2</td>
<td>−1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(P_X)</td>
<td>(\frac{1}{36})</td>
<td>(\frac{2}{36})</td>
<td>(\frac{3}{36})</td>
<td>(\frac{4}{36})</td>
<td>(\frac{5}{36})</td>
<td>(\frac{6}{36})</td>
<td>(\frac{5}{36})</td>
<td>(\frac{4}{36})</td>
<td>(\frac{3}{36})</td>
<td>(\frac{2}{36})</td>
<td>(\frac{1}{36})</td>
</tr>
</tbody>
</table>

\[ \Delta = X - \mu \]

Pop Quiz. What is \(\mathbb{E}[\Delta]\)?

**Variance**, \(\sigma^2\), is the expected value of the squared deviations,

\[
\sigma^2 = \mathbb{E}[\Delta^2] = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[(X - \mathbb{E}[X])^2]
\]

\[
\sigma^2 = \mathbb{E}[\Delta^2] = \frac{1}{36} \cdot 25 + \frac{2}{36} \cdot 16 + \frac{3}{36} \cdot 9 + \frac{4}{36} \cdot 4 + \frac{5}{36} \cdot 1 + \frac{6}{36} \cdot 0 + \frac{5}{36} \cdot 1 + \frac{4}{36} \cdot 4 + \frac{3}{36} \cdot 9 + \frac{2}{36} \cdot 16 + \frac{1}{36} \cdot 25
\]

\[ = 5 \frac{5}{6}. \]

**Standard Deviation**, \(\sigma\), is the square-root of the variance,

\[
\sigma = \sqrt{\mathbb{E}[\Delta^2]} = \sqrt{\mathbb{E}[(X - \mu)^2]} = \sqrt{\mathbb{E}[(X - \mathbb{E}[X])^2]}
\]

\[ \sigma = \sqrt{5 \frac{5}{6}} \approx 2.52 \]

sum of two dice rolls = 7 \pm 2.52.

Practice. Exercise 21.2.
Variance is a Measure of Risk

Game 1

Game 2
Variance is a Measure of Risk

Game 1

\[ X_1 : \quad \text{win } \$2 \quad \text{probability} = \frac{2}{3}; \]
\[ \text{lose } \$1 \quad \text{probability} = \frac{1}{3}. \]

Game 2
Variance is a Measure of Risk

<table>
<thead>
<tr>
<th>Game 1</th>
<th>Game 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>(X_1)</td>
<td>(X_2)</td>
</tr>
<tr>
<td>win $2</td>
<td>win $102</td>
</tr>
<tr>
<td>probability = (\frac{2}{3});</td>
<td>probability = (\frac{2}{3});</td>
</tr>
<tr>
<td>lose $1</td>
<td>lose $201</td>
</tr>
<tr>
<td>probability = (\frac{1}{3}).</td>
<td>probability = (\frac{1}{3}).</td>
</tr>
</tbody>
</table>
Variance is a Measure of Risk

Game 1

\[ X_1 : \]
- win $2 \quad \text{probability} = \frac{2}{3};
- lose $1 \quad \text{probability} = \frac{1}{3}.

\[ \mathbb{E}[X_1] = $1 \]

Game 2

\[ X_2 : \]
- win $102 \quad \text{probability} = \frac{2}{3};
- lose $201 \quad \text{probability} = \frac{1}{3}.\]
Variance is a Measure of Risk

Game 1

\[ X_1: \]
- win $2 \hspace{1cm} \text{probability} = \frac{2}{3};
- lose $1 \hspace{1cm} \text{probability} = \frac{1}{3}.

\[ \mathbb{E}[X_1] = 1 \]

Game 2

\[ X_2: \]
- win $102 \hspace{1cm} \text{probability} = \frac{2}{3};
- lose $201 \hspace{1cm} \text{probability} = \frac{1}{3}.

\[ \mathbb{E}[X_2] = 1 \]
Variance is a Measure of Risk

<table>
<thead>
<tr>
<th>Game 1</th>
<th>Game 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>(X_1: )</td>
<td>(X_2: )</td>
</tr>
<tr>
<td>win $2 \quad \text{probability} = \frac{2}{3}; )</td>
<td>win $102 \quad \text{probability} = \frac{2}{3}; )</td>
</tr>
<tr>
<td>lose $1 \quad \text{probability} = \frac{1}{3}. )</td>
<td>lose $201 \quad \text{probability} = \frac{1}{3}. )</td>
</tr>
</tbody>
</table>

\[E[X_1] = 1\]

\[E[X_2] = 1\]

\[\sigma^2(X_1) = \frac{2}{3} \cdot (2 - 1)^2 + \frac{1}{3} \cdot (-1 - 1)^2 = 2\]
Variance is a Measure of Risk

Game 1

$X_1$: win $2$ probability $= \frac{2}{3}$; lose $1$ probability $= \frac{1}{3}$.

$E[X_1] = \$1$

$\sigma^2(X_1) = \frac{2}{3} \cdot (2 - 1)^2 + \frac{1}{3} \cdot (-1 - 1)^2$

$= 2$

Game 2

$X_2$: win $102$ probability $= \frac{2}{3}$; lose $201$ probability $= \frac{1}{3}$.

$E[X_2] = \$1$

$\sigma^2(X_2) = \frac{2}{3} \cdot (102 - 1)^2 + \frac{1}{3} \cdot (-201 - 1)^2$

$\approx 2 \times 10^4$. 
Variance is a Measure of Risk

Game 1

$X_1$: win $2$ probability $= \frac{2}{3}$; lose $1$ probability $= \frac{1}{3}$.

$\mathbb{E}[X_1] = $ $1$

$\sigma^2(X_1) = \frac{2}{3} \cdot (2 - 1)^2 + \frac{1}{3} \cdot (-1 - 1)^2$

$= 2$

$X_1 = 1 \pm 1.41$

Game 2

$X_2$: win $102$ probability $= \frac{2}{3}$; lose $201$ probability $= \frac{1}{3}$.

$\mathbb{E}[X_2] = $ $1$

$\sigma^2(X_2) = \frac{2}{3} \cdot (102 - 1)^2 + \frac{1}{3} \cdot (-201 - 1)^2$

$\approx 2 \times 10^4$

$X_2 = 1 \pm 141$

For a small expected profit you might risk a small loss (Game 1), not a huge loss.
A More Convenient Formula for Variance

\[ \sigma^2 = \mathbb{E}[(X - \mu)^2] \]
A More Convenient Formula for Variance

\[ \sigma^2 = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2] \leftarrow \text{Expand } (X - \mu)^2 \]
A More Convenient Formula for Variance

\[
\sigma^2 = \mathbb{E}[(X - \mu)^2] \\
= \mathbb{E}[X^2 - 2\mu X + \mu^2] \quad \leftarrow \text{Expand } (X - \mu)^2 \\
= \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \quad \leftarrow \text{Linearity of expectation}
\]
A More Convenient Formula for Variance

\[ \sigma^2 = \mathbb{E}[(X - \mu)^2] \]
\[ = \mathbb{E}[X^2 - 2\mu X + \mu^2] \quad \leftarrow \text{Expand } (X - \mu)^2 \]
\[ = \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \quad \leftarrow \text{Linearity of expectation} \]
\[ = \mathbb{E}[X^2] - \mu^2. \quad \leftarrow \mathbb{E}[X] = \mu \]
A More Convenient Formula for Variance

\[
\sigma^2 = \mathbb{E}[(X - \mu)^2] \\
= \mathbb{E}[X^2 - 2\mu X + \mu^2] \quad \leftarrow \text{Expand } (X - \mu)^2 \\
= \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \quad \leftarrow \text{Linearity of expectation} \\
= \mathbb{E}[X^2] - \mu^2. \quad \leftarrow \mathbb{E}[X] = \mu
\]

**Variance:** \[\sigma^2 = \mathbb{E}[X^2] - \mu^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2.\]
A More Convenient Formula for Variance

\[ \sigma^2 = \mathbb{E}[(X - \mu)^2] \]

\[ = \mathbb{E}[X^2 - 2\mu X + \mu^2] \quad \leftarrow \text{Expand } (X - \mu)^2 \]

\[ = \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \quad \leftarrow \text{Linearity of expectation} \]

\[ = \mathbb{E}[X^2] - \mu^2. \quad \leftarrow \mathbb{E}[X] = \mu \]

**Variance:** \[ \sigma^2 = \mathbb{E}[X^2] - \mu^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2. \]

Sum of two dice,

\[ \mathbb{E}[X^2] = \sum_{x=2}^{12} P_X(x) \cdot x^2 \]
A More Convenient Formula for Variance

\[ \sigma^2 = \mathbb{E}[(X - \mu)^2] \]
\[ = \mathbb{E}[X^2 - 2\mu X + \mu^2] \quad \leftarrow \text{Expand } (X - \mu)^2 \]
\[ = \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \quad \leftarrow \text{Linearity of expectation} \]
\[ = \mathbb{E}[X^2] - \mu^2. \quad \leftarrow \mathbb{E}[X] = \mu \]

**Variance:** \[ \sigma^2 = \mathbb{E}[X^2] - \mu^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2. \]

Sum of two dice,
\[ \mathbb{E}[X^2] = \sum_{x=2}^{12} P_X(x) \cdot x^2 \]
\[ = \frac{1}{36} \cdot 2^2 + \]
A More Convenient Formula for Variance

\[ \sigma^2 = \mathbb{E}[(X - \mu)^2] \]
\[ = \mathbb{E}[X^2 - 2\mu X + \mu^2] \quad \leftarrow \text{Expand } (X - \mu)^2 \]
\[ = \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \quad \leftarrow \text{Linearity of expectation} \]
\[ = \mathbb{E}[X^2] - \mu^2. \quad \leftarrow \mathbb{E}[X] = \mu \]

**Variance:** \[ \sigma^2 = \mathbb{E}[X^2] - \mu^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2. \]

Sum of two dice,

\[ \mathbb{E}[X^2] = \sum_{x=2}^{12} P_X(x) \cdot x^2 \]
\[ = \frac{1}{36} \cdot 2^2 + \frac{2}{36} \cdot 3^2 + \]

Creator: Malik Magdon-Ismail

Deviations from the Mean: 8 / 13

Variance of Uniform and Bernoulli →
A More Convenient Formula for Variance

$$\sigma^2 = \mathbb{E}[(X - \mu)^2]$$

$$= \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2$$  ← Expand $(X - \mu)^2$

$$= \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2$$  ← Linearity of expectation

$$= \mathbb{E}[X^2] - \mu^2.$$  ← $\mathbb{E}[X] = \mu$

**Variance:**  $\sigma^2 = \mathbb{E}[X^2] - \mu^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$

Sum of two dice,

$$\mathbb{E}[X^2] = \sum_{x=2}^{12} P_X(x) \cdot x^2$$

$$= \frac{1}{36} \cdot 2^2 + \frac{2}{36} \cdot 3^2 + \frac{3}{36} \cdot 4^2 + \frac{4}{36} \cdot 5^2 + \frac{5}{36} \cdot 6^2 + \frac{6}{36} \cdot 7^2 + \frac{5}{36} \cdot 8^2 + \frac{4}{36} \cdot 9^2 + \frac{3}{36} \cdot 10^2 + \frac{2}{36} \cdot 11^2 + \frac{1}{36} \cdot 12^2$$
A More Convenient Formula for Variance

\[ \sigma^2 = \mathbb{E}[(X - \mu)^2] \]
\[ = \mathbb{E}[X^2 - 2\mu X + \mu^2] \quad \leftarrow \text{Expand } (X - \mu)^2 \]
\[ = \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \quad \leftarrow \text{Linearity of expectation} \]
\[ = \mathbb{E}[X^2] - \mu^2. \quad \leftarrow \mathbb{E}[X] = \mu \]

**Variance:** \( \sigma^2 = \mathbb{E}[X^2] - \mu^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2. \)

Sum of two dice,

\[ \mathbb{E}[X^2] = \frac{12}{x=2} P_{X}(x) \cdot x^2 \]
\[ = \frac{1}{36} \cdot 2^2 + \frac{2}{36} \cdot 3^2 + \frac{3}{36} \cdot 4^2 + \frac{4}{36} \cdot 5^2 + \frac{5}{36} \cdot 6^2 + \frac{6}{36} \cdot 7^2 + \frac{7}{36} \cdot 8^2 + \frac{8}{36} \cdot 9^2 + \frac{9}{36} \cdot 10^2 + \frac{10}{36} \cdot 11^2 + \frac{11}{36} \cdot 12^2 \]
\[ = \frac{545}{6} \]
A More Convenient Formula for Variance

\[ \sigma^2 = E[(X - \mu)^2] \]
\[ = E[X^2 - 2\mu X + \mu^2] \quad \leftarrow \text{Expand } (X - \mu)^2 \]
\[ = E[X^2] - 2\mu E[X] + \mu^2 \quad \leftarrow \text{Linearity of expectation} \]
\[ = E[X^2] - \mu^2. \quad \leftarrow E[X] = \mu \]

**Variance:** \( \sigma^2 = E[X^2] - \mu^2 = E[X^2] - E[X]^2. \)

Sum of two dice,

\[ E[X^2] = \sum_{x=2}^{12} P_X(x) \cdot x^2 \]
\[ = \frac{1}{36} \cdot 2^2 + \frac{2}{36} \cdot 3^2 + \frac{3}{36} \cdot 4^2 + \frac{4}{36} \cdot 5^2 + \frac{5}{36} \cdot 6^2 + \frac{6}{36} \cdot 7^2 + \frac{5}{36} \cdot 8^2 + \frac{4}{36} \cdot 9^2 + \frac{3}{36} \cdot 10^2 + \frac{2}{36} \cdot 11^2 + \frac{1}{36} \cdot 12^2 \]
\[ = \frac{545}{6} \]

Since \( \mu = 7 \)

\[ \sigma^2 = \frac{545}{6} - 7^2 = \frac{55}{6} \]
A More Convenient Formula for Variance

\[ \sigma^2 = \mathbb{E}[(X - \mu)^2] \]
\[ = \mathbb{E}[X^2 - 2\mu X + \mu^2] \quad \leftarrow \text{Expand } (X - \mu)^2 \]
\[ = \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 \quad \leftarrow \text{Linearity of expectation} \]
\[ = \mathbb{E}[X^2] - \mu^2. \quad \leftarrow \mathbb{E}[X] = \mu \]

**Variance:** \( \sigma^2 = \mathbb{E}[X^2] - \mu^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2. \)

Sum of two dice,

\[ \mathbb{E}[X^2] = \sum_{x=2}^{12} P_X(x) \cdot x^2 \]
\[ = \frac{1}{36} \cdot 2^2 + \frac{2}{36} \cdot 3^2 + \frac{3}{36} \cdot 4^2 + \frac{4}{36} \cdot 5^2 + \frac{5}{36} \cdot 6^2 + \frac{6}{36} \cdot 7^2 + \frac{5}{36} \cdot 8^2 + \frac{4}{36} \cdot 9^2 + \frac{3}{36} \cdot 10^2 + \frac{2}{36} \cdot 11^2 + \frac{1}{36} \cdot 12^2 \]
\[ = \frac{545}{6} \]

Since \( \mu = 7 \)

\[ \sigma^2 = \frac{545}{6} - 7^2 = \frac{55}{6} \]

**Theorem.** Variance \( \geq 0 \) which means \( \mathbb{E}[X^2] \geq \mathbb{E}[X]^2 \) for any random variable \( X \).
Uniform. We saw earlier that $\mathbb{E}[X] = \frac{1}{2}(n + 1)$. 
Variance of Uniform and Bernoulli

**Uniform.** We saw earlier that $\mathbb{E}[X] = \frac{1}{2}(n + 1)$.

$$\mathbb{E}[X^2]$$
Uniform. We saw earlier that $\mathbb{E}[X] = \frac{1}{2}(n + 1)$.

$$\mathbb{E}[X^2] = \frac{1}{n}(1^2 + \cdots + n^2)$$
Uniform. We saw earlier that $\mathbb{E}[X] = \frac{1}{2}(n + 1)$.

$$\mathbb{E}[X^2] = \frac{1}{n}(1^2 + \cdots + n^2) = \frac{1}{n} \times \frac{n}{6}(n + 1)(2n + 1) = \frac{1}{6}(n + 1)(2n + 1)$$
Uniform. We saw earlier that \( \mathbb{E}[X] = \frac{1}{2}(n + 1) \).

\[
\mathbb{E}[X^2] = \frac{1}{n}(1^2 + \cdots + n^2) = \frac{1}{n} \times \frac{n}{6}(n + 1)(2n + 1) = \frac{1}{6}(n + 1)(2n + 1)
\]

so

\[
\sigma^2(\text{Uniform}) = \mathbb{E}[X^2] - \mathbb{E}[X]^2
\]
Uniform. We saw earlier that $E[X] = \frac{1}{2}(n + 1)$.

\[E[X^2] = \frac{1}{n}(1^2 + \cdots + n^2) = \frac{1}{n} \times \frac{n}{6}(n + 1)(2n + 1) = \frac{1}{6}(n + 1)(2n + 1)\]

so

\[\sigma^2(\text{Uniform}) = E[X^2] - E[X]^2 = \frac{1}{6}(n + 1)(2n + 1) - \left(\frac{1}{2}(n + 1)\right)^2\]
Uniform. We saw earlier that $E[X] = \frac{1}{2}(n + 1)$.

$$E[X^2] = \frac{1}{n}(1^2 + \cdots + n^2) = \frac{1}{n} \times \frac{n}{6}(n + 1)(2n + 1) = \frac{1}{6}(n + 1)(2n + 1)$$

so

$$\sigma^2(\text{Uniform}) = E[X^2] - E[X]^2 = \frac{1}{6}(n + 1)(2n + 1) - \left(\frac{1}{2}(n + 1)\right)^2 = \frac{1}{12}(n^2 - 1).$$
**Uniform.** We saw earlier that $\mathbb{E}[X] = \frac{1}{2}(n + 1)$.

$$
\mathbb{E}[X^2] = \frac{1}{n}(1^2 + \cdots + n^2) = \frac{1}{n} \times \frac{n}{6}(n + 1)(2n + 1) = \frac{1}{6}(n + 1)(2n + 1)
$$

so

$$
\sigma^2(\text{Uniform}) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{1}{6}(n + 1)(2n + 1) - \left(\frac{1}{2}(n + 1)\right)^2 = \frac{1}{12}(n^2 - 1).
$$

**Bernoulli.** We saw earlier that $\mathbb{E}[X] = p$.
Variance of Uniform and Bernoulli

**Uniform.** We saw earlier that $\mathbb{E}[X] = \frac{1}{2}(n + 1)$.

$$
\mathbb{E}[X^2] = \frac{1}{n}(1^2 + \cdots + n^2) = \frac{1}{n} \times \frac{n}{6}(n + 1)(2n + 1) = \frac{1}{6}(n + 1)(2n + 1)
$$

so

$$
\sigma^2(\text{Uniform}) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{1}{6}(n + 1)(2n + 1) - \left(\frac{1}{2}(n + 1)\right)^2 = \frac{1}{12}(n^2 - 1).
$$

**Bernoulli.** We saw earlier that $\mathbb{E}[X] = p$.

$$
\mathbb{E}[X^2]
$$
Uniform. We saw earlier that $E[X] = \frac{1}{2}(n + 1)$.

$$E[X^2] = \frac{1}{n}(1^2 + \cdots + n^2) = \frac{1}{n} \times \frac{n}{6}(n + 1)(2n + 1) = \frac{1}{6}(n + 1)(2n + 1)$$

so

$$\sigma^2(\text{Uniform}) = E[X^2] - E[X]^2 = \frac{1}{6}(n + 1)(2n + 1) - \left(\frac{1}{2}(n + 1)\right)^2 = \frac{1}{12}(n^2 - 1).$$

Bernoulli. We saw earlier that $E[X] = p$.

$$E[X^2] = p \cdot 1^2 + (1 - p) \cdot 0^2$$
**Uniform.** We saw earlier that $\mathbb{E}[X] = \frac{1}{2}(n + 1)$.

$$\mathbb{E}[X^2] = \frac{1}{n}(1^2 + \cdots + n^2) = \frac{1}{n} \times \frac{n}{6}(n + 1)(2n + 1) = \frac{1}{6}(n + 1)(2n + 1)$$

so

$$\sigma^2(\text{Uniform}) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{1}{6}(n + 1)(2n + 1) - \left(\frac{1}{2}(n + 1)\right)^2 = \frac{1}{12}(n^2 - 1).$$

**Bernoulli.** We saw earlier that $\mathbb{E}[X] = p$.

$$\mathbb{E}[X^2] = p \cdot 1^2 + (1 - p) \cdot 0^2 = p$$

so

$$\sigma^2(\text{Bernoulli}) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$
**Uniform.** We saw earlier that $\mathbb{E}[X] = \frac{1}{2}(n + 1)$.

$$\mathbb{E}[X^2] = \frac{1}{n}(1^2 + \cdots + n^2) = \frac{1}{n} \times \frac{n}{6}(n + 1)(2n + 1) = \frac{1}{6}(n + 1)(2n + 1)$$

So

$$\sigma^2(\text{Uniform}) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{1}{6}(n + 1)(2n + 1) - \left(\frac{1}{2}(n + 1)\right)^2 = \frac{1}{12}(n^2 - 1).$$

**Bernoulli.** We saw earlier that $\mathbb{E}[X] = p$.

$$\mathbb{E}[X^2] = p \cdot 1^2 + (1 - p) \cdot 0^2 = p$$

So

$$\sigma^2(\text{Bernoulli}) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = p - p^2 = p(1 - p).$$
Linearity of Variance?
Linearity of Variance?

Let $X$ be a Bernoulli and $Y = a + X$ ($a$ is a constant):

$$Y = \begin{cases} 
    a + 1 & \text{with probability } p; \\
    a & \text{with probability } 1 - p. 
\end{cases}$$
Linearity of Variance?

Let $X$ be a Bernoulli and $Y = a + X$ ($a$ is a constant):

\[
Y = \begin{cases} 
  a + 1 & \text{with probability } p; \\
  a & \text{with probability } 1 - p.
\end{cases}
\]

\[
\mathbb{E}[Y] = p \cdot (a + 1) + (1 - p) \cdot a = a + p = a + \mathbb{E}[X]
\]

(as expected)
Linearity of Variance?

Let $X$ be a Bernoulli and $Y = a + X$ ($a$ is a constant):

$$Y = \begin{cases} a + 1 & \text{with probability } p; \\ a & \text{with probability } 1 - p. \end{cases}$$

$$\mathbb{E}[Y] = p \cdot (a + 1) + (1 - p) \cdot a = a + p = a + \mathbb{E}[X] \quad \text{(as expected)}$$

Deviations from the mean $\mu = a + p$:

$$\Delta_Y = \begin{cases} 1 - p & \text{with probability } p; \\ -p & \text{with probability } 1 - p, \end{cases}$$
Linearity of Variance?

Let $X$ be a Bernoulli and $Y = a + X$ ($a$ is a constant):

$$Y = \begin{cases} a + 1 & \text{with probability } p; \\ a & \text{with probability } 1 - p. \end{cases}$$

$$\mathbb{E}[Y] = p \cdot (a + 1) + (1 - p) \cdot a = a + p = a + \mathbb{E}[X]$$  \hspace{1cm} \text{(as expected)}

Deviations from the mean $\mu = a + p$:

$$\Delta_Y = \begin{cases} 1 - p & \text{with probability } p; \\ -p & \text{with probability } 1 - p, \end{cases}$$  \hspace{1cm} \text{(deviations independent of } a!\text{)}
Linearity of Variance?

Let $X$ be a Bernoulli and $Y = a + X$ ($a$ is a constant):

$$Y = \begin{cases} 
  a + 1 & \text{with probability } p; \\
  a & \text{with probability } 1 - p.
\end{cases}$$

$$\mathbb{E}[Y] = p \cdot (a + 1) + (1 - p) \cdot a = a + p = a + \mathbb{E}[X] \quad \text{(as expected)}$$

Deviations from the mean $\mu = a + p$:

$$\Delta_Y = \begin{cases} 
  1 - p & \text{with probability } p; \\
  -p & \text{with probability } 1 - p,
\end{cases} \quad \text{(deviations independent of } a!\text{)}$$

Therefore $\sigma^2(Y) = \sigma^2(X)$. 
Let $X$ be a Bernoulli and $Y = a + X$ ($a$ is a constant):

$$Y = \begin{cases} 
  a + 1 & \text{with probability } p; \\
  a & \text{with probability } 1 - p.
\end{cases}$$

$$\mathbb{E}[Y] = p \cdot (a + 1) + (1 - p) \cdot a = a + p = a + \mathbb{E}[X]$$

(as expected)

Deviations from the mean $\mu = a + p$:

$$\Delta_Y = \begin{cases} 
  1 - p & \text{with probability } p; \\
  -p & \text{with probability } 1 - p,
\end{cases}$$

(deviations independent of $a$!)

Therefore $\sigma^2(Y) = \sigma^2(X)$.

**Pop Quiz.** $Y = bX$. Compute $\mathbb{E}[Y]$ and $\sigma^2(Y)$. 


Linearity of Variance?

Let $X$ be a Bernoulli and $Y = a + X$ ($a$ is a constant):

$$Y = \begin{cases} a + 1 & \text{with probability } p; \\ a & \text{with probability } 1 - p. \end{cases}$$

$$\mathbb{E}[Y] = p \cdot (a + 1) + (1 - p) \cdot a = a + p = a + \mathbb{E}[X] \quad \text{(as expected)}$$

Deviations from the mean $\mu = a + p$:

$$\Delta_Y = \begin{cases} 1 - p & \text{with probability } p; \\ -p & \text{with probability } 1 - p, \end{cases} \quad \text{(deviations independent of } a\text{!)}$$

Therefore $\sigma^2(Y) = \sigma^2(X)$.

**Pop Quiz.** $Y = bX$. Compute $\mathbb{E}[Y]$ and $\sigma^2(Y)$.

**Theorem.** Let $Y = a + bX$. Then,

$$\sigma^2(Y) = b^2 \sigma^2(X).$$
Variance of a Sum

\[ X = X_1 + X_2 \]
Variance of a Sum

\[ X = X_1 + X_2 \]
\[ \mathbb{E}[X]^2 \]
Variance of a Sum

\[ X = X_1 + X_2 \]
\[ \mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 \]
Variance of a Sum

\[ X = X_1 + X_2 \]

\[ \mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 \overset{(*)}{=} (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 \]
\[ X = X_1 + X_2 \]

\[
\mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 \overset{(*)}{=} (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2];
\]
Variance of a Sum

\[ X = X_1 + X_2 \]

\[ \mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 = (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2\mathbb{E}[X_1]\mathbb{E}[X_2]; \]

\[ \mathbb{E}[X^2] \]
Variance of a Sum

\[ X = X_1 + X_2 \]

\[ \mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 \stackrel{(*)}{=} (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2]; \]

\[ \mathbb{E}[X^2] = \mathbb{E}[(X_1 + X_2)^2] \]
Variance of a Sum

\[ X = X_1 + X_2 \]

\[ \mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 \overset{(*)}{=} (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2]; \]

\[ \mathbb{E}[X^2] = \mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2 + X_2^2 + 2X_1X_2] \]
Variance of a Sum

\[ X = X_1 + X_2 \]

\[ \mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 \overset{(*)}{=} (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2]; \]

\[ \mathbb{E}[X^2] = \mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2 + X_2^2 + 2X_1X_2] \overset{(*)}{=} \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2]. \]

In both derivations above, we use linearity in \((*)\).
Variance of a Sum

\[ X = X_1 + X_2 \]

\[ \mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 \overset{(*)}{=} (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2]; \]

\[ \mathbb{E}[X^2] = \mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2 + X_2^2 + 2X_1X_2] \overset{(*)}{=} \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2]. \]

In both derivations above, we use linearity in 

\[ \sigma^2(X) \]
Variance of a Sum

\[ X = X_1 + X_2 \]

\[ E[X]^2 = E[X_1 + X_2]^2 \overset{(*)}{=} (E[X_1] + E[X_2])^2 = E[X_1]^2 + E[X_2]^2 + 2E[X_1]E[X_2]; \]

\[ E[X^2] = E[(X_1 + X_2)^2] = E[X_1^2 + X_2^2 + 2X_1X_2] \overset{(*)}{=} E[X_1^2] + E[X_2^2] + 2E[X_1X_2]. \]

In both derivations above, we use linearity in \((*)\).

\[ \sigma^2(X) = E[X^2] - E[X]^2 \]
Variance of a Sum

\[ \mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 \]

\[ \mathbb{E}[\mathbf{X}]^2 = (\mathbb{E}[\mathbf{X}_1] + \mathbb{E}[\mathbf{X}_2])^2 = \mathbb{E}[\mathbf{X}_1]^2 + \mathbb{E}[\mathbf{X}_2]^2 + 2 \mathbb{E} [\mathbf{X}_1] \mathbb{E} [\mathbf{X}_2]; \]

\[ \mathbb{E}[\mathbf{X}^2] = \mathbb{E}[(\mathbf{X}_1 + \mathbf{X}_2)^2] = \mathbb{E}[\mathbf{X}_1^2 + \mathbf{X}_2^2 + 2\mathbf{X}_1\mathbf{X}_2] = \mathbb{E} [\mathbf{X}_1^2] + \mathbb{E} [\mathbf{X}_2^2] + 2 \mathbb{E} [\mathbf{X}_1\mathbf{X}_2]. \]

In both derivations above, we use linearity in (\textasteriskcentered).

\[ \sigma^2(\mathbf{X}) = \mathbb{E}[\mathbf{X}^2] - \mathbb{E}[\mathbf{X}]^2 \]
\[ = \mathbb{E}[\mathbf{X}_1^2] + \mathbb{E}[\mathbf{X}_2^2] + 2 \mathbb{E} [\mathbf{X}_1\mathbf{X}_2] \]
Variance of a Sum

\[ X = X_1 + X_2 \]

\[
\mathbb{E}[X]^2 = \mathbb{E}[(X_1 + X_2)^2] = (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2];
\]

\[
\mathbb{E}[X^2] = \mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2 + X_2^2 + 2X_1X_2] \overset{(*)}{=} \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2].
\]

In both derivations above, we use linearity in (\(*)\).

\[
\sigma^2(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2
\]

\[
= \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]^2 - \mathbb{E}[X_2]^2 - 2 \mathbb{E}[X_1] \mathbb{E}[X_2]
\]
Variance of a Sum

\[ X = X_1 + X_2 \]

\[ \mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 = (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2]; \]

\[ \mathbb{E}[X^2] = \mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2 + X_2^2 + 2X_1X_2] = \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2]. \]

In both derivations above, we use linearity in (\(\ast\)).

\[ \sigma^2(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \]
\[ = \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]^2 - \mathbb{E}[X_2]^2 - 2 \mathbb{E}[X_1] \mathbb{E}[X_2] \]
\[ = \frac{\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2}{\sigma^2(X_1)} \]
**Variance of a Sum**

\[
X = X_1 + X_2
\]

\[
\mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 \overset{(*)}{=} (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2];
\]

\[
\mathbb{E}[X^2] = \mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2 + X_2^2 + 2X_1X_2] \overset{(*)}{=} \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2].
\]

In both derivations above, we use linearity in \((*)\).

\[
\sigma^2(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2
\]

\[
= \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]^2 - \mathbb{E}[X_2]^2 - 2 \mathbb{E}[X_1] \mathbb{E}[X_2]
\]

\[
= \frac{\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2}{\sigma^2(X_1)} + \frac{\mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2}{\sigma^2(X_2)}
\]
Variance of a Sum

\[ X = X_1 + X_2 \]

\[ \mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 = (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2]; \]

\[ \mathbb{E}[X^2] = \mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2 + X_2^2 + 2X_1X_2] = \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2]. \]

In both derivations above, we use linearity in (*)..

\[ \sigma^2(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \]

\[ = \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]^2 - \mathbb{E}[X_2]^2 - 2 \mathbb{E}[X_1] \mathbb{E}[X_2] \]

\[ = \frac{\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2}{\sigma^2(X_1)} + \frac{\mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2}{\sigma^2(X_2)} + 2 \left( \frac{\mathbb{E}[X_1X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2]}{0 \text{ if } X_1 \text{ and } X_2 \text{ are independent}} \right) \]
Variance of a Sum

\[ X = X_1 + X_2 \]

\[ \mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 \overset{(*)}{=} (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2]; \]

\[ \mathbb{E}[X^2] = \mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2 + X_2^2 + 2X_1X_2] \overset{(*)}{=} \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2]. \]

In both derivations above, we use linearity in (\( \ast \)).

\[ \sigma^2(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \]

\[ = \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]^2 - \mathbb{E}[X_2]^2 - 2 \mathbb{E}[X_1] \mathbb{E}[X_2] \]

\[ = \frac{\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2}{\sigma^2(X_1)} + \frac{\mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2}{\sigma^2(X_2)} + 2 \left( \frac{\mathbb{E}[X_1X_2] - \mathbb{E}[X_1] \mathbb{E}[X_2]}{\sigma^2(X_1) \sigma^2(X_2)} \right) \]

\[ 0 \text{ if } X_1 \text{ and } X_2 \text{ are independent} \]

**Variance of a Sum.** For independent random variables, the variance of the sum is a sum of the variances.
Variance of a Sum

\[ X = X_1 + X_2 \]

\[ \mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 = (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2]; \]

\[ \mathbb{E}[X^2] = \mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2 + X_2^2 + 2X_1X_2] \]

In both derivations above, we use linearity in (*)

\[ \sigma^2(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \]

\[ = \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]^2 - \mathbb{E}[X_2]^2 - 2 \mathbb{E}[X_1] \mathbb{E}[X_2] \]

\[ = \underbrace{\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2}_{\sigma^2(X_1)} + \underbrace{\mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2}_{\sigma^2(X_2)} + 2 \left( \underbrace{\mathbb{E}[X_1X_2] - \mathbb{E}[X_1] \mathbb{E}[X_2]}_{0 \text{ if } X_1 \text{ and } X_2 \text{ are independent}} \right) \]

**Variance of a Sum.** For independent random variables, the variance of the sum is a sum of the variances.

**Practice.** Compute the variance of 1 dice roll. Compute the variance of the sum of \( n \) dice rolls.
Variance of a Sum

\[ X = X_1 + X_2 \]
\[ \mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 \overset{(*)}{=} (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2]; \]
\[ \mathbb{E}[X^2] = \mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2 + X_2^2 + 2X_1X_2] \overset{(*)}{=} \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2]. \]

In both derivations above, we use linearity in (\( \ast \)).

\[ \sigma^2(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \]
\[ = \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]^2 - \mathbb{E}[X_2]^2 - 2 \mathbb{E}[X_1] \mathbb{E}[X_2] \]
\[ = \left( \frac{\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2}{\sigma^2(X_1)} + \frac{\mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2}{\sigma^2(X_2)} \right) + 2 \left( \frac{\mathbb{E}[X_1X_2] - \mathbb{E}[X_1] \mathbb{E}[X_2]}{\sigma^2(X_1)} \right) \]
\[ 0 \text{ if } X_1 \text{ and } X_2 \text{ are independent} \]

**Variance of a Sum.** For independent random variables, the variance of the sum is a sum of the variances.

**Practice.** Compute the variance of 1 dice roll. Compute the variance of the sum of \( n \) dice rolls.

**Example.** The Variance of the Binomial (sum of independent Bernoullis)
Variance of a Sum

\[ X = X_1 + X_2 \]

\[ E[X]^2 = E[X_1 + X_2]^2 \overset{(*)}{=} (E[X_1] + E[X_2])^2 = E[X_1]^2 + E[X_2]^2 + 2 E[X_1] E[X_2]; \]

\[ E[X]^2 = E[(X_1 + X_2)^2] = E[X_1^2 + X_2^2 + 2X_1X_2] \overset{(*)}{=} E[X_1^2] + E[X_2^2] + 2 E[X_1X_2]. \]

In both derivations above, we use linearity in \((*)\).

\[
\sigma^2(X) = E[X^2] - E[X]^2 \\
= E[X_1^2] + E[X_2^2] + 2 E[X_1X_2] - E[X_1]^2 - E[X_2]^2 - 2 E[X_1] E[X_2] \\
= \frac{E[X_1^2] - E[X_1]^2}{\sigma^2(X_1)} + \frac{E[X_2^2] - E[X_2]^2}{\sigma^2(X_2)} + 2 \left( \frac{E[X_1X_2] - E[X_1] E[X_2]}{0 \text{ if } X_1 \text{ and } X_2 \text{ are independent}} \right)
\]

**Variance of a Sum.** For *independent* random variables, the variance of the sum is a sum of the variances.

**Practice.** Compute the variance of 1 dice roll. Compute the variance of the sum of \(n\) dice rolls.

**Example.** The Variance of the Binomial (sum of *independent* Bernoullis)

\[ X = X_1 + \cdots + X_n \text{ (sum of *independent* Bernoullis)} \]
**Variance of a Sum**

\[ X = X_1 + X_2 \]

\[ \mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 = (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2]; \]

\[ \mathbb{E}[X]^2 = \mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2 + X_2^2 + 2X_1X_2] \stackrel{(*)}{=} \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2]. \]

In both derivations above, we use linearity in (\( \ast \)).

\[ \sigma^2(X) = \mathbb{E}[X]^2 - \mathbb{E}[X]^2 \]

\[ = \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]^2 - \mathbb{E}[X_2]^2 - 2 \mathbb{E}[X_1] \mathbb{E}[X_2] \]

\[ = \frac{\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2}{\sigma^2(X_1)} + \frac{\mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2}{\sigma^2(X_2)} + 2 \left( \mathbb{E}[X_1X_2] - \mathbb{E}[X_1] \mathbb{E}[X_2] \right) \]

0 if \( X_1 \) and \( X_2 \) are independent

**Variance of a Sum.** For *independent* random variables, the variance of the sum is a sum of the variances.

**Practice.** Compute the variance of 1 dice roll. Compute the variance of the sum of \( n \) dice rolls.

**Example.** The Variance of the Binomial (sum of *independent* Bernoullis)

\[ X = X_1 + \cdots + X_n \] (sum of *independent* Bernoullis), and \( \sigma^2(X_i) = p(1 - p) \)
Variance of a Sum

\[ X = X_1 + X_2 \]

\[
\mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 \overset{(*)}{=} (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2];
\]

\[
\mathbb{E}[X^2] = \mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2 + X_2^2 + 2X_1X_2] \overset{(*)}{=} \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2].
\]

In both derivations above, we use linearity in (\(\ast\)).

\[
\sigma^2(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2
\]

\[
= \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]^2 - \mathbb{E}[X_2]^2 - 2 \mathbb{E}[X_1] \mathbb{E}[X_2]
\]

\[
= \underbrace{\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2}_{\sigma^2(X_1)} + \underbrace{\mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2}_{\sigma^2(X_2)} + 2 \left( \mathbb{E}[X_1X_2] - \mathbb{E}[X_1] \mathbb{E}[X_2] \right) \tag{0 \text{ if } X_1 \text{ and } X_2 \text{ are independent}}
\]

Variance of a Sum. For independent random variables, the variance of the sum is a sum of the variances.

Practice. Compute the variance of 1 dice roll. Compute the variance of the sum of \(n\) dice rolls.

Example. The Variance of the Binomial (sum of independent Bernoullis)

\[ X = X_1 + \cdots + X_n \text{ (sum of independent Bernoullis), and } \sigma^2(X_i) = p(1 - p) \]

\[ \sigma^2(\text{Binomial}) = \sigma^2(X_1) + \cdots + \sigma^2(X_n) \]
## Variance of a Sum

\[ X = X_1 + X_2 \]

\[
\mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 \overset{(*)}{=} (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2];
\]

\[
\mathbb{E}[X^2] = \mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2 + X_2^2 + 2X_1X_2] \overset{(*)}{=} \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2].
\]

In both derivations above, we use linearity in \((*)\).

\[
\sigma^2(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2
\]

\[
= \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]^2 - \mathbb{E}[X_2]^2 - 2 \mathbb{E}[X_1] \mathbb{E}[X_2]
\]

\[
= \underbrace{\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2}_{\sigma^2(X_1)} + \underbrace{\mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2}_{\sigma^2(X_2)} + 2 \mathbb{E}[X_1X_2] - \mathbb{E}[X_1] \mathbb{E}[X_2] - \sigma^2(X_1) - \sigma^2(X_2)
\]

\[
0 \text{ if } X_1 \text{ and } X_2 \text{ are independent}
\]

### Variance of a Sum. For independent random variables, the variance of the sum is a sum of the variances.

**Practice.** Compute the variance of 1 dice roll. Compute the variance of the sum of \(n\) dice rolls.

### Example. The Variance of the Binomial (sum of independent Bernoullis)

\[ X = X_1 + \cdots + X_n \] (sum of independent Bernoullis), and \(\sigma^2(X_i) = p(1-p)\)

\[
\sigma^2(\text{Binomial}) = \sigma^2(X_1) + \cdots + \sigma^2(X_n) = p(1-p) + \cdots + p(1-p)
\]
Variance of a Sum

\[ X = X_1 + X_2 \]

\[ \mathbb{E}[X]^2 = \mathbb{E}[X_1 + X_2]^2 = (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2 = \mathbb{E}[X_1]^2 + \mathbb{E}[X_2]^2 + 2 \mathbb{E}[X_1] \mathbb{E}[X_2]; \]

\[ \mathbb{E}[X]^2 = \mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2 + X_2^2 + 2X_1X_2] = \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2]. \]

In both derivations above, we use linearity in (*)

\[ \sigma^2(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \]

\[ = \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2 \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]^2 - \mathbb{E}[X_2]^2 - 2 \mathbb{E}[X_1] \mathbb{E}[X_2] \]

\[ = \underbrace{\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2}_{\sigma^2(X_1)} + \underbrace{\mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2}_{\sigma^2(X_2)} + 2 \left( \mathbb{E}[X_1X_2] - \mathbb{E}[X_1] \mathbb{E}[X_2] \right) \]

0 if \( X_1 \) and \( X_2 \) are independent

**Variance of a Sum.** For *independent* random variables, the variance of the sum is a sum of the variances.

**Practice.** Compute the variance of 1 dice roll. Compute the variance of the sum of \( n \) dice rolls.

**Example.** The Variance of the Binomial (sum of *independent* Bernoullis)

\[ X = X_1 + \cdots + X_n \] (sum of *independent* Bernoullis), and \( \sigma^2(X_i) = p(1 - p) \)

\[ \sigma^2(\text{Binomial}) = \sigma^2(X_1) + \cdots + \sigma^2(X_n) = p(1 - p) + \cdots + p(1 - p) = np(1 - p). \]
3-σ Rule: \( X = \mu(X) \pm \sigma(X) \)

**3-σ Rule.** For *any* random variable \( X \), the chances are at least (about) 90% that
\[
\mu - 3\sigma < X < \mu + 3\sigma \quad \text{or} \quad X = \mu \pm 3\sigma.
\]
3-σ Rule: \( X = \mu(X) \pm \sigma(X) \)

**3-σ Rule.** For any random variable \( X \), the chances are at least (about) 90% that

\[
\mu - 3\sigma < X < \mu + 3\sigma \quad \text{or} \quad X = \mu \pm 3\sigma.
\]

**Lemma (Markov Inequality).** For a positive random variable \( X \),

\[
P[X \geq \alpha] \leq \frac{E[X]}{\alpha}.
\]
3-σ Rule: \( X = \mu(X) \pm \sigma(X) \)

**3-σ Rule.** For any random variable \( X \), the chances are at least (about) 90% that
\[
\mu - 3\sigma < X < \mu + 3\sigma \quad \text{or} \quad X = \mu \pm 3\sigma.
\]

**Lemma (Markov Inequality).** For a positive random variable \( X \),
\[
P[X \geq \alpha] \leq \frac{\mathbb{E}[X]}{\alpha}.
\]

**Proof.** \( \mathbb{E}[X] \)
3-σ Rule: \( X = \mu(X) \pm \sigma(X) \)

**3-σ Rule.** For *any* random variable \( X \), the chances are at least (about) 90% that

\[
\mu - 3\sigma < X < \mu + 3\sigma \quad \text{or} \quad X = \mu \pm 3\sigma.
\]

**Lemma (Markov Inequality).** For a positive random variable \( X \),

\[
P[X \geq \alpha] \leq \frac{\mathbb{E}[X]}{\alpha}.
\]

*Proof.* \( \mathbb{E}[X] = \sum_{x \geq 0} x \cdot P_X(x) \)
3-σ Rule: \( X = \mu(X) \pm \sigma(X) \)

**3-σ Rule.** For *any* random variable \( X \), the chances are at least (about) 90% that
\[
\mu - 3\sigma < X < \mu + 3\sigma \quad \text{or} \quad X = \mu \pm 3\sigma.
\]

**Lemma (Markov Inequality).** For a positive random variable \( X \),
\[
P[X \geq \alpha] \leq \frac{E[X]}{\alpha}.
\]

*Proof.* \( E[X] = \sum_{x \geq 0} x \cdot P_X(x) \geq \sum_{x \geq \alpha} x \cdot P_X(x) \)
3-σ Rule: \( X = \mu(X) \pm \sigma(X) \)

**3-σ Rule.** For *any* random variable \( X \), the chances are at least (about) 90% that

\[
\mu - 3\sigma < X < \mu + 3\sigma \quad \text{or} \quad X = \mu \pm 3\sigma.
\]

**Lemma (Markov Inequality).** For a positive random variable \( X \),

\[
P[X \geq \alpha] \leq \frac{\mathbb{E}[X]}{\alpha}.
\]

**Proof.** \( \mathbb{E}[X] = \sum_{x \geq 0} x \cdot P_X(x) \geq \sum_{x \geq \alpha} x \cdot P_X(x) \geq \sum_{x \geq \alpha} \alpha \cdot P_X(x) \)
### 3-σ Rule

**3-σ Rule.** For *any* random variable $X$, the chances are at least (about) 90% that

$$\mu - 3\sigma < X < \mu + 3\sigma \quad \text{or} \quad X = \mu \pm 3\sigma.$$

---

**Lemma (Markov Inequality).** For a positive random variable $X$,

$$\mathbb{P}[X \geq \alpha] \leq \frac{\mathbb{E}[X]}{\alpha}.$$

**Proof.**

$$\mathbb{E}[X] = \sum_{x \geq 0} x \cdot P_X(x) \geq \sum_{x \geq \alpha} x \cdot P_X(x) \geq \sum_{x \geq \alpha} \alpha \cdot P_X(x) = \alpha \cdot \mathbb{P}[X \geq \alpha].$$
3-σ Rule: \( X = \mu(X) \pm \sigma(X) \)

**3-σ Rule.** For any random variable \( X \), the chances are at least (about) 90% that
\[
\mu - 3\sigma < X < \mu + 3\sigma \quad \text{or} \quad X = \mu \pm 3\sigma.
\]

**Lemma (Markov Inequality).** For a positive random variable \( X \),
\[
P[X \geq \alpha] \leq \frac{\mathbb{E}[X]}{\alpha}.
\]

**Proof.**
\[
\mathbb{E}[X] = \sum_{x \geq 0} x \cdot P_X(x) \geq \sum_{x \geq \alpha} x \cdot P_X(x) \geq \sum_{x \geq \alpha} \alpha \cdot P_X(x) = \alpha \cdot P[X \geq \alpha].
\]

**Lemma (Chebyshev Inequality).**
\[
P[|\Delta| \geq t\sigma] \leq \frac{1}{t^2}.
\]
3-σ Rule: \( X = \mu(X) \pm \sigma(X) \)

**3-σ Rule.** For *any* random variable \( X \), the chances are at least (about) 90% that
\[
\mu - 3\sigma < X < \mu + 3\sigma \quad \text{or} \quad X = \mu \pm 3\sigma.
\]

**Lemma (Markov Inequality).** For a positive random variable \( X \),
\[
P[X \geq \alpha] \leq \frac{E[X]}{\alpha}.
\]

*Proof.*
\[
E[X] = \sum_{x \geq 0} x \cdot P_X(x) \geq \sum_{x \geq \alpha} x \cdot P_X(x) \geq \sum_{x \geq \alpha} \alpha \cdot P_X(x) = \alpha \cdot P[X \geq \alpha].
\]

**Lemma (Chebyshev Inequality).**
\[
P[|\Delta| \geq t\sigma] \leq \frac{1}{t^2}.
\]

*Proof.*
\[
P[|\Delta| \geq t\sigma]
\]
3-σ Rule: \( X = \mu(X) \pm \sigma(X) \)

**3-σ Rule.** For *any* random variable \( X \), the chances are at least (about) 90% that
\[
\mu - 3\sigma < X < \mu + 3\sigma \quad \text{or} \quad X = \mu \pm 3\sigma.
\]

**Lemma (Markov Inequality).** For a positive random variable \( X \),
\[
\mathbb{P}[X \geq \alpha] \leq \frac{\mathbb{E}[X]}{\alpha}.
\]

*Proof.* \( \mathbb{E}[X] = \sum_{x \geq 0} x \cdot P_X(x) \geq \sum_{x \geq \alpha} x \cdot P_X(x) \geq \sum_{x \geq \alpha} \alpha \cdot P_X(x) = \alpha \cdot \mathbb{P}[X \geq \alpha]. \)

**Lemma (Chebyshev Inequality).**
\[
\mathbb{P}[|\Delta| \geq t\sigma] \leq \frac{1}{t^2}.
\]

*Proof.*
\[
\mathbb{P}[|\Delta| \geq t\sigma] = \mathbb{P}[\Delta^2 \geq t^2\sigma^2].
\]
3-σ Rule: \( X = \mu(X) \pm \sigma(X) \)

3-σ Rule. For any random variable \( X \), the chances are at least (about) 90% that
\[
\mu - 3\sigma < X < \mu + 3\sigma \quad \text{or} \quad X = \mu \pm 3\sigma.
\]

Lemma (Markov Inequality). For a positive random variable \( X \),
\[
P[X \geq \alpha] \leq \frac{E[X]}{\alpha}.
\]

Proof. \( E[X] = \sum_{x \geq 0} x \cdot P_X(x) \geq \sum_{x \geq \alpha} x \cdot P_X(x) \geq \sum_{x \geq \alpha} \alpha \cdot P_X(x) = \alpha \cdot P[X \geq \alpha]. \)

Lemma (Chebyshev Inequality).
\[
P[|\Delta| \geq t\sigma] \leq \frac{1}{t^2}.
\]

Proof.
\[
P[|\Delta| \geq t\sigma] = P[\Delta^2 \geq t^2\sigma^2] \overset{(a)}{=} \frac{E[\Delta^2]}{t^2\sigma^2}
\]
3-σ Rule: $X = \mu(X) \pm \sigma(X)$

3-σ Rule. For any random variable $X$, the chances are at least (about) 90% that

$$\mu - 3\sigma < X < \mu + 3\sigma$$

or

$$X = \mu \pm 3\sigma.$$

Lemma (Markov Inequality). For a positive random variable $X$,

$$P[X \geq \alpha] \leq \frac{E[X]}{\alpha}.$$

Proof. $E[X] = \sum_{x \geq 0} x \cdot P_X(x) \geq \sum_{x \geq \alpha} x \cdot P_X(x) \geq \sum_{x \geq \alpha} \alpha \cdot P_X(x) = \alpha \cdot P[X \geq \alpha].$

Lemma (Chebyshev Inequality).

$$P[|\Delta| \geq t\sigma] \leq \frac{1}{t^2}.$$

Proof.

$$P[|\Delta| \geq t\sigma] = P[\Delta^2 \geq t^2\sigma^2] \leq \frac{E[\Delta^2]}{t^2\sigma^2} = \frac{\sigma^2}{t^2\sigma^2}.$$
**3-σ Rule.** For any random variable $X$, the chances are at least (about) 90% that

$$\mu - 3\sigma < X < \mu + 3\sigma$$

or

$$X = \mu \pm 3\sigma.$$

**Lemma (Markov Inequality).** For a positive random variable $X$,

$$\Pr[X \geq \alpha] \leq \frac{\mathbb{E}[X]}{\alpha}.$$

**Proof.**

$$\mathbb{E}[X] = \sum_{x \geq 0} x \cdot P_X(x) \geq \sum_{x \geq \alpha} x \cdot P_X(x) \geq \sum_{x \geq \alpha} \alpha \cdot P_X(x) = \alpha \cdot \Pr[X \geq \alpha].$$

**Lemma (Chebyshev Inequality).**

$$\Pr[|\Delta| \geq t\sigma] \leq \frac{1}{t^2}.$$

**Proof.**

$$\Pr[|\Delta| \geq t\sigma] = \Pr[\Delta^2 \geq t^2\sigma^2] \overset{(a)}{=} \frac{\mathbb{E}[\Delta^2]}{t^2\sigma^2} = \frac{\sigma^2}{t^2\sigma^2} = \frac{1}{t^2}.$$ 

In (a) we used Markov’s Inequality.

To get the 3-σ rule, use Chebyshev’s Inequality with $t = 3$. 
Law of Large Numbers

Expectation of the average of $n$ dice:

$$\mathbb{E}[\text{average}] = \mathbb{E}[\frac{1}{n} \times \text{sum}] = \frac{1}{n} \times \mathbb{E}[\text{sum}] = \frac{1}{n} \times n \times 3\frac{1}{2}$$
Law of Large Numbers

Expectation of the average of $n$ dice:
$$\mathbb{E}[\text{average}] = \mathbb{E}\left[\frac{1}{n} \times \text{sum}\right] = \frac{1}{n} \times \mathbb{E}[\text{sum}] = \frac{1}{n} \times n \times 3\frac{1}{2}$$

Variance of the average of $n$ dice:
$$\sigma^2(\text{average})$$
Law of Large Numbers

Expectation of the average of $n$ dice:

$$\mathbb{E}[\text{average}] = \mathbb{E}[\frac{1}{n} \times \text{sum}] = \frac{1}{n} \times \mathbb{E}[\text{sum}] = \frac{1}{n} \times n \times \frac{3}{2}$$

Variance of the average of $n$ dice:

$$\sigma^2(\text{average}) = \sigma^2(\frac{1}{n} \times \text{sum})$$
Law of Large Numbers

Expectation of the average of $n$ dice:

$$\mathbb{E}[\text{average}] = \mathbb{E}[\frac{1}{n} \times \text{sum}] = \frac{1}{n} \times \mathbb{E}[\text{sum}] = \frac{1}{n} \times n \times 3\frac{1}{2}$$

Variance of the average of $n$ dice:

$$\sigma^2(\text{average}) = \sigma^2(\frac{1}{n} \times \text{sum}) = \frac{1}{n^2} \times \sigma^2(\text{sum})$$
Law of Large Numbers

Expectation of the average of $n$ dice:
\[
\mathbb{E}[\text{average}] = \mathbb{E}\left[\frac{1}{n} \times \text{sum}\right] = \frac{1}{n} \times \mathbb{E}[\text{sum}] = \frac{1}{n} \times n \times 3\frac{1}{2}
\]

Variance of the average of $n$ dice:
\[
\sigma^2(\text{average}) = \sigma^2\left(\frac{1}{n} \times \text{sum}\right) = \frac{1}{n^2} \times \sigma^2(\text{sum}) = \frac{1}{n^2} \times n \times \sigma^2(\text{one die})
\]
Law of Large Numbers

Expectation of the average of $n$ dice:

$$\mathbb{E}[\text{average}] = \mathbb{E}[\frac{1}{n} \times \text{sum}] = \frac{1}{n} \times \mathbb{E}[\text{sum}] = \frac{1}{n} \times n \times 3\frac{1}{2}$$

Variance of the average of $n$ dice:

$$\sigma^2(\text{average}) = \sigma^2(\frac{1}{n} \times \text{sum}) = \frac{1}{n^2} \times \sigma^2(\text{sum}) = \frac{1}{n^2} \times n \times \sigma^2(\text{one die}) = \frac{1}{n} \times \sigma^2(\text{one die})$$
Law of Large Numbers

Expectation of the average of \( n \) dice:

\[
\mathbb{E}[\text{average}] = \mathbb{E}\left[ \frac{1}{n} \times \text{sum} \right] = \frac{1}{n} \times \mathbb{E}[\text{sum}] = \frac{1}{n} \times n \times 3\frac{1}{2}
\]

Variance of the average of \( n \) dice:

\[
\sigma^2(\text{average}) = \sigma^2\left( \frac{1}{n} \times \text{sum} \right) = \frac{1}{n^2} \times \sigma^2(\text{sum}) = \frac{1}{n^2} \times n \times \sigma^2(\text{one die}) = \frac{1}{n} \times \sigma^2(\text{one die})
\]