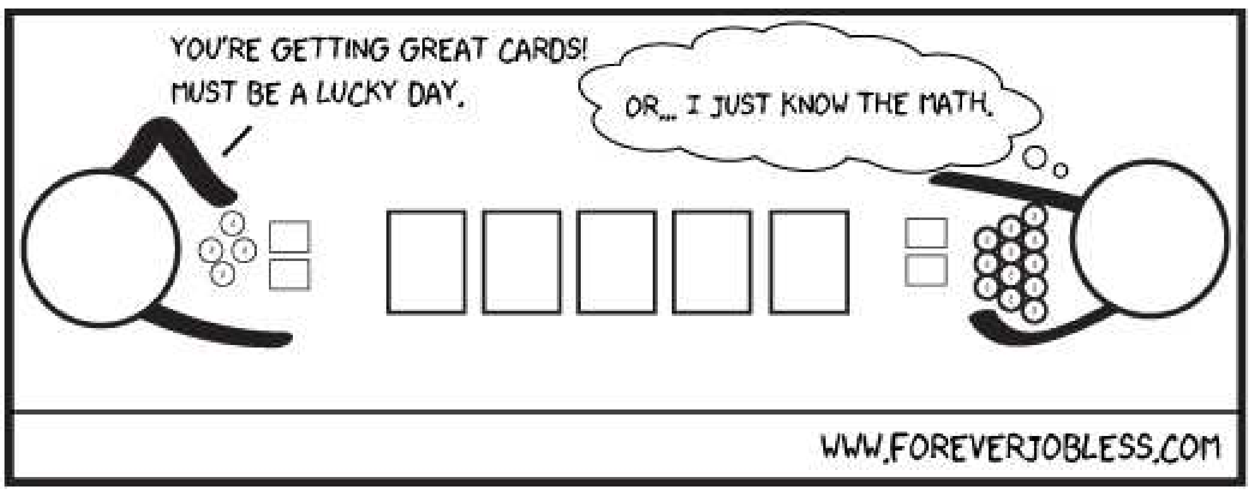
Foundations of Computer Science
Lecture 21

Deviations from the Mean

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



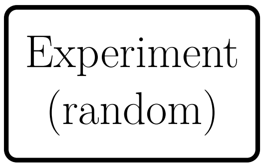
Last Time

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.

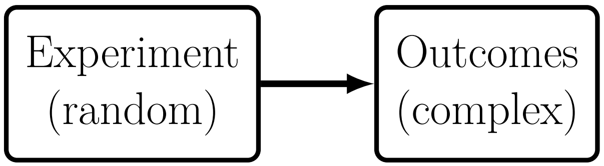
Today: Deviations from the Mean

1. How well does the expected value (mean) summarize a random variable?
2. Variance.
3. Variance of a sum.
4. Law of large numbers
   - The 3-σ rule.

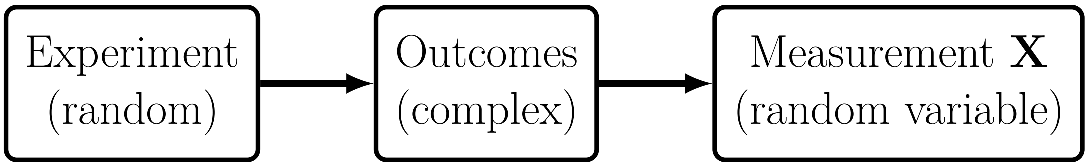
Probability For Analyzing a Random Experiment.



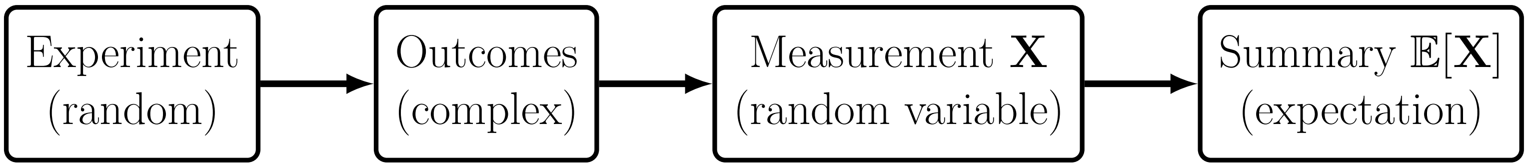
Probability For Analyzing a Random Experiment.



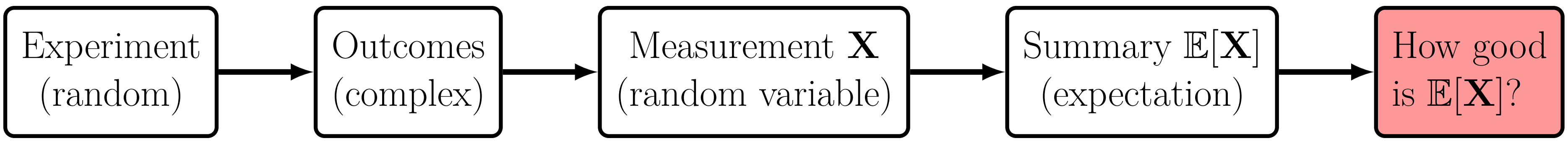
Probability For Analyzing a Random Experiment.



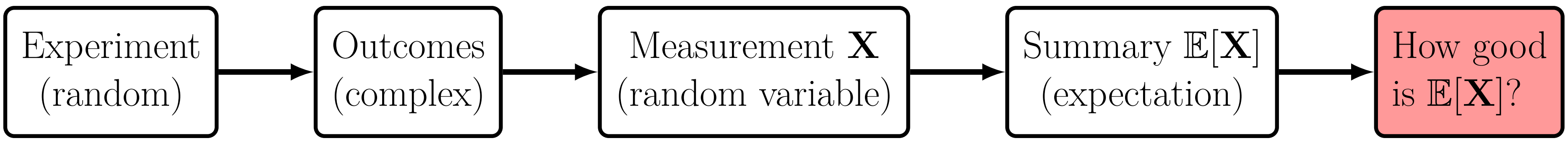
Probability For Analyzing a Random Experiment.



Probability For Analyzing a Random Experiment.

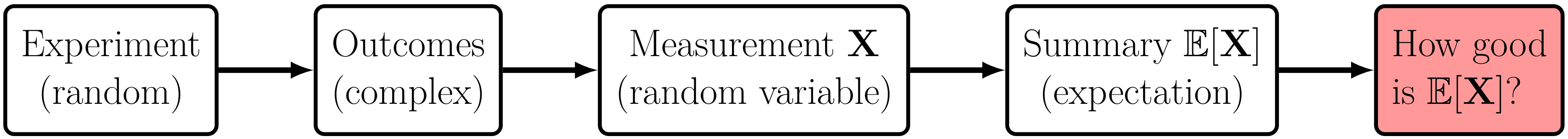


Probability For Analyzing a Random Experiment.



Experiment. Roll n dice and compute $\mathbf{X}$, the average of the rolls.

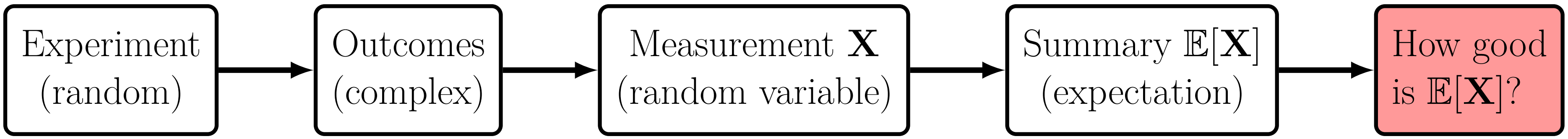
Probability For Analyzing a Random Experiment.



Experiment. Roll n dice and compute $\mathbf{X}$, the average of the rolls.

$\mathbb{E}[\text{average}]$

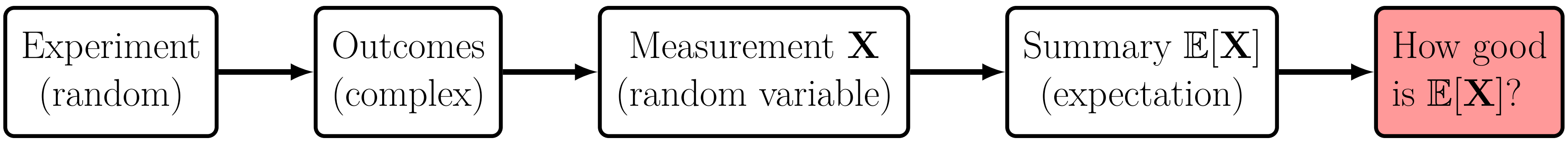
Probability For Analyzing a Random Experiment.



Experiment. Roll n dice and compute $\mathbf{X}$, the average of the rolls.

$$\mathbb{E}[\text{average}] = \mathbb{E}\left[\tfrac{1}{n} \cdot \text{sum}\right]$$

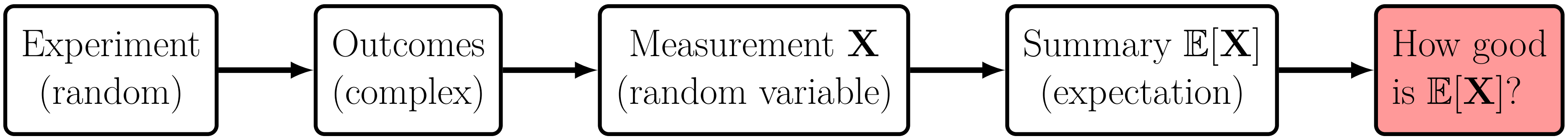
Probability For Analyzing a Random Experiment.



Experiment. Roll n dice and compute $\mathbf{X}$, the average of the rolls.

$$\mathbb{E}[\text{average}] \;=\; \mathbb{E}\left[\tfrac{1}{n}\cdot \text{sum}\right] \;=\; \tfrac{1}{n}\cdot \mathbb{E}\left[\text{sum}\right]$$

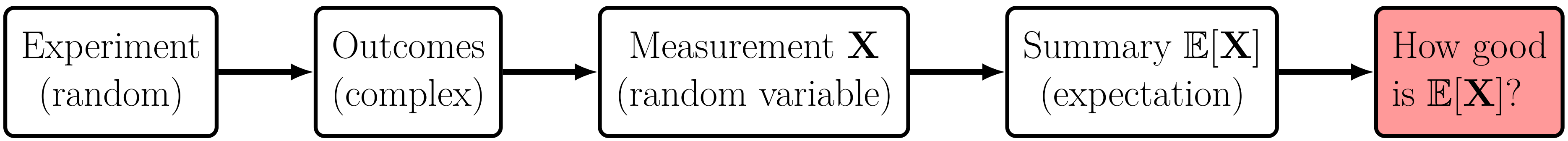
Probability For Analyzing a Random Experiment.



Experiment. Roll n dice and compute $\mathbf{X}$, the average of the rolls.

$$\mathbb{E}[\text{average}] \;=\; \mathbb{E}\left[\tfrac{1}{n}\cdot\text{sum}\right] \;=\; \tfrac{1}{n}\cdot\mathbb{E}\left[\text{sum}\right] \;=\; \tfrac{1}{n}\times n\times 3\tfrac{1}{2}$$

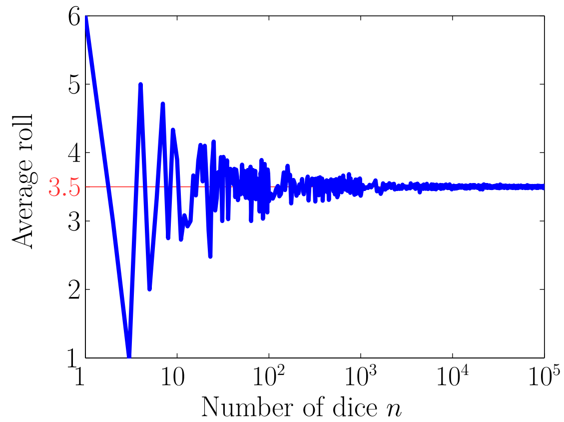
Probability For Analyzing a Random Experiment.



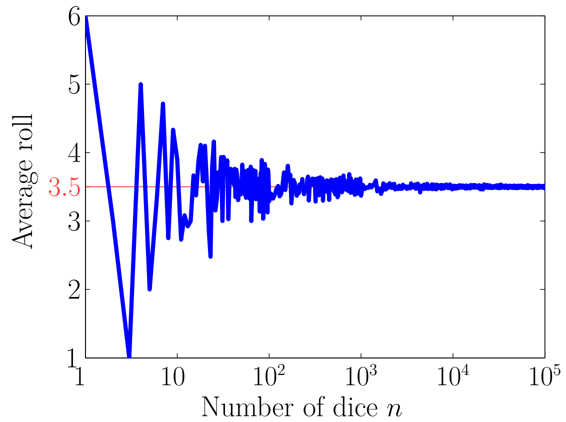
Experiment. Roll n dice and compute $\mathbf{X}$, the average of the rolls.

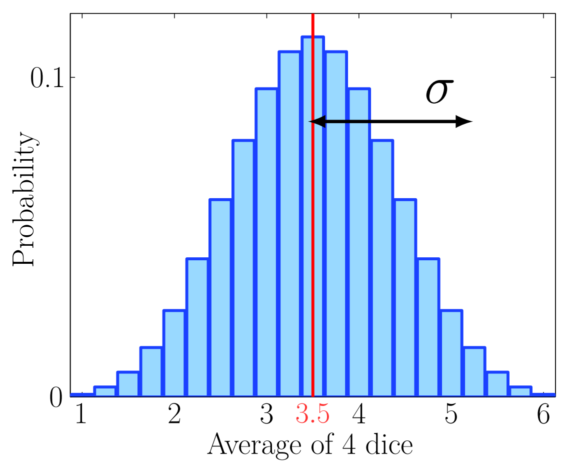
$$\mathbb{E}[\text{average}] \;=\; \mathbb{E}\left[\tfrac{1}{n}\cdot\text{sum}\right] \;=\; \tfrac{1}{n}\cdot\mathbb{E}\left[\text{sum}\right] \;=\; \tfrac{1}{n}\times n\times 3\tfrac{1}{2} \;=\; 3\tfrac{1}{2}.$$

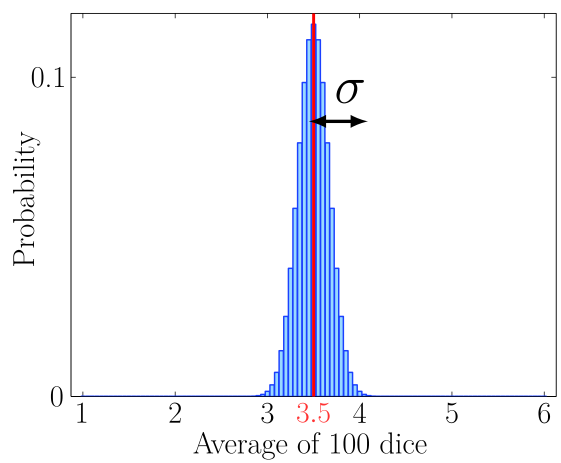
Average of n Dice



Average of n Dice







Variance: Size of the Deviations From the Mean

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

Variance: Size of the Deviations From the Mean

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

$\mathbf{X}$	2	3	4	5	6	7	8	9	10	11	12	
$\boldsymbol{\Delta}$	-5	-4	-3	-2	-1	0	1	2	3	4	5	$\leftarrow \mathbf{X} - \mu$
$P_{\mathbf{X}}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	

Pop Quiz. What is $\mathbb{E}[\boldsymbol{\Delta}]$?

Variance: Size of the Deviations From the Mean

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

$\mathbf{X}$	2	3	4	5	6	7	8	9	10	11	12	
$\boldsymbol{\Delta}$	-5	-4	-3	-2	-1	0	1	2	3	4	5	$\leftarrow \mathbf{X} - \mu$
$P_{\mathbf{X}}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	

Pop Quiz. What is $\mathbb{E}[\boldsymbol{\Delta}]$?

Variance, σ^2, is the expected value of the squared deviations,

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

$$\sigma^2 \ = \ \mathbb{E}[\boldsymbol{\Delta}^2] \ = \ \tfrac{1}{36}\cdot 25 +$$

Variance: Size of the Deviations From the Mean

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

$\mathbf{X}$	2	3	4	5	6	7	8	9	10	11	12	
$\boldsymbol{\Delta}$	-5	-4	-3	-2	-1	0	1	2	3	4	5	$\leftarrow \mathbf{X} - \mu$
$P_{\mathbf{X}}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	

Pop Quiz. What is $\mathbb{E}[\boldsymbol{\Delta}]$?

Variance, σ^2, is the expected value of the squared deviations,

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

$$\sigma^2 = \mathbb{E}[\boldsymbol{\Delta}^2] = \tfrac{1}{36} \cdot 25 + \tfrac{2}{36} \cdot 16 +$$

Variance: Size of the Deviations From the Mean

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

$\mathbf{X}$	2	3	4	5	6	7	8	9	10	11	12	
$\boldsymbol{\Delta}$	-5	-4	-3	-2	-1	0	1	2	3	4	5	$\leftarrow \mathbf{X} - \mu$
$P_{\mathbf{X}}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	

Pop Quiz. What is $\mathbb{E}[\boldsymbol{\Delta}]$?

Variance, σ^2, is the expected value of the squared deviations,

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

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

Variance: Size of the Deviations From the Mean

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

$\mathbf{X}$	2	3	4	5	6	7	8	9	10	11	12	
$\mathbf{\Delta}$	-5	-4	-3	-2	-1	0	1	2	3	4	5	$\leftarrow \mathbf{X} - \mu$
$P_{\mathbf{X}}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	

Pop Quiz. What is $\mathbb{E}[\mathbf{\Delta}]$?

Variance, σ^2, is the expected value of the squared deviations,

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

$$\begin{aligned}\sigma^2 = \mathbb{E}[\mathbf{\Delta}^2] &= \tfrac{1}{36}\cdot 25 + \tfrac{2}{36}\cdot 16 + \tfrac{3}{36}\cdot 9 + \tfrac{4}{36}\cdot 4 + \tfrac{5}{36}\cdot 1 + \tfrac{6}{36}\cdot 0 + \tfrac{5}{36}\cdot 1 + \tfrac{4}{36}\cdot 4 + \tfrac{3}{36}\cdot 9 + \tfrac{2}{36}\cdot 16 + \tfrac{1}{36}\cdot 25 \\ &= 5\tfrac{5}{6}.\end{aligned}$$

Variance: Size of the Deviations From the Mean

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

$\mathbf{X}$	2	3	4	5	6	7	8	9	10	11	12	
$\boldsymbol{\Delta}$	-5	-4	-3	-2	-1	0	1	2	3	4	5	$\leftarrow \mathbf{X} - \mu$
$P_{\mathbf{X}}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	

Pop Quiz. What is $\mathbb{E}[\boldsymbol{\Delta}]$?

Variance, σ^2, is the expected value of the squared deviations,

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

$$\begin{aligned}\sigma^2 = \mathbb{E}[\boldsymbol{\Delta}^2] &= \tfrac{1}{36}\cdot 25 + \tfrac{2}{36}\cdot 16 + \tfrac{3}{36}\cdot 9 + \tfrac{4}{36}\cdot 4 + \tfrac{5}{36}\cdot 1 + \tfrac{6}{36}\cdot 0 + \tfrac{5}{36}\cdot 1 + \tfrac{4}{36}\cdot 4 + \tfrac{3}{36}\cdot 9 + \tfrac{2}{36}\cdot 16 + \tfrac{1}{36}\cdot 25 \\ &= 5\tfrac{5}{6}.\end{aligned}$$

Standard Deviation, σ, is the square-root of the variance,

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

$\sigma = \sqrt{5\frac{5}{6}} \approx 2.52$

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

$$\mathbf{X}_1: \quad \begin{array}{ll} \text{win } \$2 & \text{probability} = \frac{2}{3}; \\ \text{lose } \$1 & \text{probability} = \frac{1}{3}. \end{array}$$

Game 2

Variance is a Measure of Risk

Game 1

$$\mathbf{X}_1 : \quad \begin{array}{rl} \text{win } \$2 & \text{probability} = \frac{2}{3}; \\ \text{lose } \$1 & \text{probability} = \frac{1}{3}. \end{array}$$

Game 2

$$\mathbf{X}_2 : \quad \begin{array}{rl} \text{win } \$102 & \text{probability} = \frac{2}{3}; \\ \text{lose } \$201 & \text{probability} = \frac{1}{3}. \end{array}$$

Variance is a Measure of Risk

Game 1

$$\mathbf{X}_1 : \begin{array}{ll} \text{win } \$2 & \text{probability} = \frac{2}{3}; \\ \text{lose } \$1 & \text{probability} = \frac{1}{3}. \end{array}$$

$$\mathbb{E}[\mathbf{X}_1] = \$1$$

Game 2

$$\mathbf{X}_2 : \begin{array}{ll} \text{win } \$102 & \text{probability} = \frac{2}{3}; \\ \text{lose } \$201 & \text{probability} = \frac{1}{3}. \end{array}$$

Variance is a Measure of Risk

Game 1

$$\mathbf{X}_1: \quad \begin{array}{ll} \text{win } \$2 & \text{probability} = \frac{2}{3}; \\ \text{lose } \$1 & \text{probability} = \frac{1}{3}. \end{array}$$

$$\mathbb{E}[\mathbf{X}_1] = \$1$$

Game 2

$$\mathbf{X}_2: \quad \begin{array}{ll} \text{win } \$102 & \text{probability} = \frac{2}{3}; \\ \text{lose } \$201 & \text{probability} = \frac{1}{3}. \end{array}$$

$$\mathbb{E}[\mathbf{X}_2] = \$1$$

Variance is a Measure of Risk

<u>Game 1</u>

$$\mathbf{X}_1 : \quad \begin{array}{ll} \text{win } \$2 & \text{probability} = \frac{2}{3}; \\ \text{lose } \$1 & \text{probability} = \frac{1}{3}. \end{array}$$

$$\mathbb{E}[\mathbf{X}_1] = \$1$$

$$\begin{aligned} \sigma^2(\mathbf{X}_1) &= \tfrac{2}{3} \cdot (2-1)^2 + \tfrac{1}{3} \cdot (-1-1)^2 \\ &= 2 \end{aligned}$$

<u>Game 2</u>

$$\mathbf{X}_2 : \quad \begin{array}{ll} \text{win } \$102 & \text{probability} = \frac{2}{3}; \\ \text{lose } \$201 & \text{probability} = \frac{1}{3}. \end{array}$$

$$\mathbb{E}[\mathbf{X}_2] = \$1$$

Variance is a Measure of Risk

<u>Game 1</u>

$$\mathbf{X}_1 : \begin{array}{ll} \text{win \$2} & \text{probability} = \frac{2}{3}; \\ \text{lose \$1} & \text{probability} = \frac{1}{3}. \end{array}$$

$$\mathbb{E}[\mathbf{X}_1] = \$1$$

$$\begin{aligned} \sigma^2(\mathbf{X}_1) &= \tfrac{2}{3} \cdot (2-1)^2 + \tfrac{1}{3} \cdot (-1-1)^2 \\ &= 2 \end{aligned}$$

<u>Game 2</u>

$$\mathbf{X}_2 : \begin{array}{ll} \text{win \$102} & \text{probability} = \frac{2}{3}; \\ \text{lose \$201} & \text{probability} = \frac{1}{3}. \end{array}$$

$$\mathbb{E}[\mathbf{X}_2] = \$1$$

$$\begin{aligned} \sigma^2(\mathbf{X}_2) &= \tfrac{2}{3} \cdot (102-1)^2 + \tfrac{1}{3} \cdot (-201-1)^2 \\ &\approx 2 \times 10^4. \end{aligned}$$

Variance is a Measure of Risk

Game 1

$$\mathbf{X}_1 : \begin{array}{ll} \text{win } \$2 & \text{probability} = \frac{2}{3}; \\ \text{lose } \$1 & \text{probability} = \frac{1}{3}. \end{array}$$

$$\mathbb{E}[\mathbf{X}_1] = \$1$$

$$\begin{aligned} \sigma^2(\mathbf{X}_1) &= \tfrac{2}{3} \cdot (2-1)^2 + \tfrac{1}{3} \cdot (-1-1)^2 \\ &= 2 \end{aligned}$$

Game 2

$$\mathbf{X}_2 : \begin{array}{ll} \text{win } \$102 & \text{probability} = \frac{2}{3}; \\ \text{lose } \$201 & \text{probability} = \frac{1}{3}. \end{array}$$

$$\mathbb{E}[\mathbf{X}_2] = \$1$$

$$\begin{aligned} \sigma^2(\mathbf{X}_2) &= \tfrac{2}{3} \cdot (102-1)^2 + \tfrac{1}{3} \cdot (-201-1)^2 \\ &\approx 2 \times 10^4. \end{aligned}$$

$$\begin{aligned} \mathbf{X}_1 &= 1 \pm 1.41 \\ \mathbf{X}_2 &= 1 \pm \mathbf{141} \end{aligned}$$

For a small expected profit you might risk a small loss (Game 1), not a huge loss.

A More Convenient Formula for Variance

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

A More Convenient Formula for Variance

$$\begin{aligned} \sigma^2 &= \mathbb{E}[(\mathbf{X} - \mu)^2] \\ &= \mathbb{E}[\mathbf{X}^2 - 2\mu\mathbf{X} + \mu^2] \end{aligned}$$

← Expand $(\mathbf{X} - \mu)^2$

A More Convenient Formula for Variance

$$\begin{aligned}
\sigma^2 &= \mathbb{E}[(\mathbf{X}-\mu)^2] \\
&= \mathbb{E}[\mathbf{X}^2 - 2\mu\mathbf{X} + \mu^2] && \leftarrow \text{Expand } (\mathbf{X}-\mu)^2 \\
&= \mathbb{E}[\mathbf{X}^2] - 2\mu\,\mathbb{E}\,[\mathbf{X}] + \mu^2 && \leftarrow \text{Linearity of expectation}
\end{aligned}$$

A More Convenient Formula for Variance

$$\begin{aligned}
\sigma^2 &= \mathbb{E}[(\mathbf{X}-\mu)^2] \\
&= \mathbb{E}[\mathbf{X}^2 - 2\mu\mathbf{X} + \mu^2] && \leftarrow \text{Expand } (\mathbf{X}-\mu)^2 \\
&= \mathbb{E}[\mathbf{X}^2] - 2\mu\,\mathbb{E}[\mathbf{X}] + \mu^2 && \leftarrow \text{Linearity of expectation} \\
&= \mathbb{E}[\mathbf{X}^2] - \mu^2. && \leftarrow \mathbb{E}[\mathbf{X}] = \mu
\end{aligned}$$

A More Convenient Formula for Variance

$$
\begin{aligned}
\sigma^2 &= \mathbb{E}[(\mathbf{X}-\mu)^2] \\
&= \mathbb{E}[\mathbf{X}^2 - 2\mu\mathbf{X} + \mu^2] && \leftarrow \text{Expand } (\mathbf{X}-\mu)^2 \\
&= \mathbb{E}[\mathbf{X}^2] - 2\mu\,\mathbb{E}\,[\mathbf{X}] + \mu^2 && \leftarrow \text{Linearity of expectation} \\
&= \mathbb{E}[\mathbf{X}^2] - \mu^2. && \leftarrow \mathbb{E}[\mathbf{X}] = \mu
\end{aligned}
$$

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

A More Convenient Formula for Variance

$$\begin{aligned}
\sigma^2 &= \mathbb{E}[(\mathbf{X}-\mu)^2] \\
&= \mathbb{E}[\mathbf{X}^2 - 2\mu\mathbf{X} + \mu^2] && \leftarrow \text{Expand } (\mathbf{X}-\mu)^2 \\
&= \mathbb{E}[\mathbf{X}^2] - 2\mu\,\mathbb{E}\,[\mathbf{X}] + \mu^2 && \leftarrow \text{Linearity of expectation} \\
&= \mathbb{E}[\mathbf{X}^2] - \mu^2. && \leftarrow \mathbb{E}[\mathbf{X}] = \mu
\end{aligned}$$

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

Sum of two dice,

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

A More Convenient Formula for Variance

$$
\begin{aligned}
\sigma^2 &= \mathbb{E}[(\mathbf{X}-\mu)^2] \\
&= \mathbb{E}[\mathbf{X}^2 - 2\mu\mathbf{X} + \mu^2] && \leftarrow \text{Expand } (\mathbf{X}-\mu)^2 \\
&= \mathbb{E}[\mathbf{X}^2] - 2\mu\,\mathbb{E}\,[\mathbf{X}] + \mu^2 && \leftarrow \text{Linearity of expectation} \\
&= \mathbb{E}[\mathbf{X}^2] - \mu^2. && \leftarrow \mathbb{E}[\mathbf{X}] = \mu
\end{aligned}
$$

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

Sum of two dice,

$$
\begin{aligned}
\mathbb{E}[\mathbf{X}^2] &= \sum_{x=2}^{12} P_{\mathbf{X}}(x)\cdot x^2 \\
&= \tfrac{1}{36}\cdot 2^2 +
\end{aligned}
$$

A More Convenient Formula for Variance

$$\begin{aligned}
\sigma^2 &= \mathbb{E}[(\mathbf{X}-\mu)^2] \\
&= \mathbb{E}[\mathbf{X}^2 - 2\mu\mathbf{X} + \mu^2] && \leftarrow \text{Expand } (\mathbf{X}-\mu)^2 \\
&= \mathbb{E}[\mathbf{X}^2] - 2\mu\,\mathbb{E}\,[\mathbf{X}] + \mu^2 && \leftarrow \text{Linearity of expectation} \\
&= \mathbb{E}[\mathbf{X}^2] - \mu^2. && \leftarrow \mathbb{E}[\mathbf{X}] = \mu
\end{aligned}$$

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

Sum of two dice,

$$\begin{aligned}
\mathbb{E}[\mathbf{X}^2] &= \sum_{x=2}^{12} P_{\mathbf{X}}(x)\cdot x^2 \\
&= \tfrac{1}{36}\cdot 2^2 + \tfrac{2}{36}\cdot 3^2 +
\end{aligned}$$

A More Convenient Formula for Variance

$$\begin{aligned}
\sigma^2 &= \mathbb{E}[(\mathbf{X}-\mu)^2] \\
&= \mathbb{E}[\mathbf{X}^2 - 2\mu\mathbf{X} + \mu^2] && \leftarrow \text{Expand } (\mathbf{X}-\mu)^2 \\
&= \mathbb{E}[\mathbf{X}^2] - 2\mu\,\mathbb{E}[\mathbf{X}] + \mu^2 && \leftarrow \text{Linearity of expectation} \\
&= \mathbb{E}[\mathbf{X}^2] - \mu^2. && \leftarrow \mathbb{E}[\mathbf{X}] = \mu
\end{aligned}$$

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

Sum of two dice,

$$\begin{aligned}
\mathbb{E}[\mathbf{X}^2] &= \sum_{x=2}^{12} P_{\mathbf{X}}(x) \cdot x^2 \\
&= \tfrac{1}{36}\cdot 2^2 + \tfrac{2}{36}\cdot 3^2 + \tfrac{3}{36}\cdot 4^2 + \tfrac{4}{36}\cdot 5^2 + \tfrac{5}{36}\cdot 6^2 + \tfrac{6}{36}\cdot 7^2 + \tfrac{5}{36}\cdot 8^2 + \tfrac{4}{36}\cdot 9^2 + \tfrac{3}{36}\cdot 10^2 + \tfrac{2}{36}\cdot 11^2 + \tfrac{1}{36}\cdot 12^2
\end{aligned}$$

A More Convenient Formula for Variance

$$\begin{aligned}
\sigma^2 &= \mathbb{E}[(\mathbf{X}-\mu)^2] \\
&= \mathbb{E}[\mathbf{X}^2 - 2\mu\mathbf{X} + \mu^2] && \leftarrow \text{Expand } (\mathbf{X}-\mu)^2 \\
&= \mathbb{E}[\mathbf{X}^2] - 2\mu\,\mathbb{E}\,[\mathbf{X}] + \mu^2 && \leftarrow \text{Linearity of expectation} \\
&= \mathbb{E}[\mathbf{X}^2] - \mu^2. && \leftarrow \mathbb{E}[\mathbf{X}] = \mu
\end{aligned}$$

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

Sum of two dice,

$$\begin{aligned}
\mathbb{E}[\mathbf{X}^2] &= \sum_{x=2}^{12} P_{\mathbf{X}}(x)\cdot x^2 \\
&= \tfrac{1}{36}\cdot 2^2 + \tfrac{2}{36}\cdot 3^2 + \tfrac{3}{36}\cdot 4^2 + \tfrac{4}{36}\cdot 5^2 + \tfrac{5}{36}\cdot 6^2 + \tfrac{6}{36}\cdot 7^2 + \tfrac{5}{36}\cdot 8^2 + \tfrac{4}{36}\cdot 9^2 + \tfrac{3}{36}\cdot 10^2 + \tfrac{2}{36}\cdot 11^2 + \tfrac{1}{36}\cdot 12^2 \\
&= 54\tfrac{5}{6}
\end{aligned}$$

A More Convenient Formula for Variance

$$\begin{aligned}
\sigma^2 &= \mathbb{E}[(\mathbf{X}-\mu)^2] \\
&= \mathbb{E}[\mathbf{X}^2 - 2\mu\mathbf{X} + \mu^2] && \leftarrow \text{Expand } (\mathbf{X}-\mu)^2 \\
&= \mathbb{E}[\mathbf{X}^2] - 2\mu\,\mathbb{E}\,[\mathbf{X}] + \mu^2 && \leftarrow \text{Linearity of expectation} \\
&= \mathbb{E}[\mathbf{X}^2] - \mu^2. && \leftarrow \mathbb{E}[\mathbf{X}] = \mu
\end{aligned}$$

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

Sum of two dice,

$$\begin{aligned}
\mathbb{E}[\mathbf{X}^2] &= \sum_{x=2}^{12} P_{\mathbf{X}}(x)\cdot x^2 \\
&= \tfrac{1}{36}\cdot 2^2 + \tfrac{2}{36}\cdot 3^2 + \tfrac{3}{36}\cdot 4^2 + \tfrac{4}{36}\cdot 5^2 + \tfrac{5}{36}\cdot 6^2 + \tfrac{6}{36}\cdot 7^2 + \tfrac{5}{36}\cdot 8^2 + \tfrac{4}{36}\cdot 9^2 + \tfrac{3}{36}\cdot 10^2 + \tfrac{2}{36}\cdot 11^2 + \tfrac{1}{36}\cdot 12^2 \\
&= 54\tfrac{5}{6}
\end{aligned}$$

Since $\mu = 7$

$$\sigma^2 = 54\tfrac{5}{6} - 7^2 = 5\tfrac{5}{6}$$

A More Convenient Formula for Variance

$$\begin{aligned}
\sigma^2 &= \mathbb{E}[(\mathbf{X}-\mu)^2] \\
&= \mathbb{E}[\mathbf{X}^2 - 2\mu\mathbf{X} + \mu^2] && \leftarrow \text{Expand } (\mathbf{X}-\mu)^2 \\
&= \mathbb{E}[\mathbf{X}^2] - 2\mu\,\mathbb{E}\,[\mathbf{X}] + \mu^2 && \leftarrow \text{Linearity of expectation} \\
&= \mathbb{E}[\mathbf{X}^2] - \mu^2. && \leftarrow \mathbb{E}[\mathbf{X}] = \mu
\end{aligned}$$

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

Sum of two dice,

$$\begin{aligned}
\mathbb{E}[\mathbf{X}^2] &= \sum_{x=2}^{12} P_{\mathbf{X}}(x)\cdot x^2 \\
&= \tfrac{1}{36}\cdot 2^2 + \tfrac{2}{36}\cdot 3^2 + \tfrac{3}{36}\cdot 4^2 + \tfrac{4}{36}\cdot 5^2 + \tfrac{5}{36}\cdot 6^2 + \tfrac{6}{36}\cdot 7^2 + \tfrac{5}{36}\cdot 8^2 + \tfrac{4}{36}\cdot 9^2 + \tfrac{3}{36}\cdot 10^2 + \tfrac{2}{36}\cdot 11^2 + \tfrac{1}{36}\cdot 12^2 \\
&= 54\tfrac{5}{6}
\end{aligned}$$

Since $\mu = 7$

$$\sigma^2 = 54\tfrac{5}{6} - 7^2 = 5\tfrac{5}{6}$$

Theorem. Variance ≥ 0, which means $\mathbb{E}[\mathbf{X}^2] \geq \mathbb{E}[\mathbf{X}]^2$ for any random variable $\mathbf{X}$.

Variance of Uniform and Bernoulli

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

Variance of Uniform and Bernoulli

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

$$\mathbb{E}[\mathbf{X}^2]$$

Variance of Uniform and Bernoulli

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

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

Variance of Uniform and Bernoulli

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

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

Variance of Uniform and Bernoulli

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

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

so

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

Variance of Uniform and Bernoulli

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

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

so

$$\sigma^2(\text{Uniform}) \;=\; \mathbb{E}[\mathbf{X}^2] - \mathbb{E}[\mathbf{X}]^2 \;=\; \tfrac{1}{6}(n+1)(2n+1) - (\tfrac{1}{2}(n+1))^2$$

Variance of Uniform and Bernoulli

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

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

so

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

Variance of Uniform and Bernoulli

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

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

so

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

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

Variance of Uniform and Bernoulli

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

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

so

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

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

$$\mathbb{E}[\mathbf{X}^2]$$

Variance of Uniform and Bernoulli

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

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

so

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

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

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

Variance of Uniform and Bernoulli

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

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

so

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

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

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

so

$$\sigma^2(\text{Bernoulli}) \;=\; \mathbb{E}[\mathbf{X}^2] - \mathbb{E}[\mathbf{X}]^2$$

Variance of Uniform and Bernoulli

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

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

so

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

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

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

so

$$\sigma^2(\text{Bernoulli}) \;=\; \mathbb{E}[\mathbf{X}^2] - \mathbb{E}[\mathbf{X}]^2 \;=\; p - p^2 \;=\; p(1-p).$$

Linearity of Variance?

Linearity of Variance?

Let $\mathbf{X}$ be a Bernoulli and $\mathbf{Y} = a + \mathbf{X}$ (a is a constant):

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

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

Linearity of Variance?

Let $\mathbf{X}$ be a Bernoulli and $\mathbf{Y} = a + \mathbf{X}$ (a is a constant):

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

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

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

$$\boldsymbol{\Delta}_{\mathbf{Y}} = \begin{cases} 1-p & \text{with probability } p; \\ -p & \text{with probability } 1-p, \end{cases} \qquad \text{(deviations independent of } a\text{!)}$$

Linearity of Variance?

Let $\mathbf{X}$ be a Bernoulli and $\mathbf{Y} = a + \mathbf{X}$ (a is a constant):

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

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

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

$$\mathbf{\Delta}_{\mathbf{Y}} = \begin{cases} 1-p & \text{with probability } p; \\ -p & \text{with probability } 1-p, \end{cases} \qquad \text{(deviations independent of } a\text{!)}$$

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

Linearity of Variance?

Let $\mathbf{X}$ be a Bernoulli and $\mathbf{Y} = a + \mathbf{X}$ (a is a constant):

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

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

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

$$\boldsymbol{\Delta}_{\mathbf{Y}} = \begin{cases} 1-p & \text{with probability } p; \\ -p & \text{with probability } 1-p, \end{cases} \qquad \text{(deviations independent of } a\text{!)}$$

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

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

Linearity of Variance?

Let $\mathbf{X}$ be a Bernoulli and $\mathbf{Y} = a + \mathbf{X}$ (a is a constant):

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

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

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

$$\boldsymbol{\Delta}_{\mathbf{Y}} = \begin{cases} 1-p & \text{with probability } p; \\ -p & \text{with probability } 1-p, \end{cases} \qquad \text{(deviations independent of } a\text{!)}$$

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

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

Theorem. Let $\mathbf{Y} = a + b\mathbf{X}$. Then,

$$\sigma^2(\mathbf{Y}) = b^2\sigma^2(\mathbf{X}).$$

Variance of a Sum

$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$

$$\mathbb{E}[\mathbf{X}]^2 = \mathbb{E}[\mathbf{X}_1 + \mathbf{X}_2]^2$$

Variance of a Sum

$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$

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

$(*)$ is by linearity of expectation.

Variance of a Sum

$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$

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

$(*)$ is by linearity of expectation.

Variance of a Sum

$$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$$

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

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

$(*)$ is by linearity of expectation.

Variance of a Sum

$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$

$$\mathbb{E}[\mathbf{X}]^2 \;=\; \mathbb{E}[\mathbf{X}_1 + \mathbf{X}_2]^2 \overset{(*)}{=} (\mathbb{E}[\mathbf{X}_1] + \mathbb{E}[\mathbf{X}_2])^2 \;=\; \mathbb{E}[\mathbf{X}_1]^2 + \mathbb{E}[\mathbf{X}_2]^2 + 2\,\mathbb{E}\left[\mathbf{X}_1\right]\mathbb{E}\left[\mathbf{X}_2\right];$$

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

$(*)$ is by linearity of expectation.

Variance of a Sum

$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$

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

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

$(*)$ is by linearity of expectation.

Variance of a Sum

$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$

$$\mathbb{E}[\mathbf{X}]^2 \;=\; \mathbb{E}[\mathbf{X}_1 + \mathbf{X}_2]^2 \stackrel{(*)}{=} (\mathbb{E}[\mathbf{X}_1] + \mathbb{E}[\mathbf{X}_2])^2 \;=\; \mathbb{E}[\mathbf{X}_1]^2 + \mathbb{E}[\mathbf{X}_2]^2 + 2\,\mathbb{E}\left[\mathbf{X}_1\right]\mathbb{E}\left[\mathbf{X}_2\right];$$

$$\mathbb{E}\left[\mathbf{X}^2\right] \;=\; \mathbb{E}[(\mathbf{X}_1 + \mathbf{X}_2)^2] \;=\; \mathbb{E}[\mathbf{X}_1^2 + \mathbf{X}_2^2 + 2\mathbf{X}_1\mathbf{X}_2] \stackrel{(*)}{=} \mathbb{E}\left[\mathbf{X}_1^2\right] + \mathbb{E}[\mathbf{X}_2^2] + 2\,\mathbb{E}\left[\mathbf{X}_1\mathbf{X}_2\right].$$

$(*)$ is by linearity of expectation.

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

Variance of a Sum

$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$

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

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

$(*)$ is by linearity of expectation.

$$\begin{aligned}\sigma^2(\mathbf{X}) &= \mathbb{E}[\mathbf{X}^2] - \mathbb{E}[\mathbf{X}]^2 \\ &= (\mathbb{E}[\mathbf{X}_1^2] + \mathbb{E}[\mathbf{X}_2^2] + 2\,\mathbb{E}[\mathbf{X}_1\mathbf{X}_2]) - (\mathbb{E}[\mathbf{X}_1]^2 + \mathbb{E}[\mathbf{X}_2]^2 + 2\,\mathbb{E}[\mathbf{X}_1]\,\mathbb{E}[\mathbf{X}_2])\end{aligned}$$

Variance of a Sum

$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$

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

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

$(*)$ is by linearity of expectation.

$$\begin{aligned}
\sigma^2(\mathbf{X}) &= \mathbb{E}[\mathbf{X}^2] - \mathbb{E}[\mathbf{X}]^2 \\
&= (\mathbb{E}[\mathbf{X}_1^2] + \mathbb{E}[\mathbf{X}_2^2] + 2\,\mathbb{E}[\mathbf{X}_1\mathbf{X}_2]) - (\mathbb{E}[\mathbf{X}_1]^2 + \mathbb{E}[\mathbf{X}_2]^2 + 2\,\mathbb{E}[\mathbf{X}_1]\,\mathbb{E}[\mathbf{X}_2]) \\
&= \underbrace{\mathbb{E}[\mathbf{X}_1^2] - \mathbb{E}[\mathbf{X}_1]^2}_{\sigma^2(\mathbf{X}_1)} + \underbrace{\mathbb{E}[\mathbf{X}_2^2] - \mathbb{E}[\mathbf{X}_2]^2}_{\sigma^2(\mathbf{X}_2)} + 2\underbrace{(\mathbb{E}[\mathbf{X}_1\mathbf{X}_2] - \mathbb{E}[\mathbf{X}_1]\,\mathbb{E}[\mathbf{X}_2])}_{0 \text{ if } \mathbf{X}_1 \text{ and } \mathbf{X}_2 \text{ are independent}}
\end{aligned}$$

Variance of a Sum

$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$

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

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

$(*)$ is by linearity of expectation.

$$\begin{aligned}
\sigma^2(\mathbf{X}) &= \mathbb{E}[\mathbf{X}^2] - \mathbb{E}[\mathbf{X}]^2 \\
&= (\mathbb{E}[\mathbf{X}_1^2] + \mathbb{E}[\mathbf{X}_2^2] + 2\,\mathbb{E}[\mathbf{X}_1\mathbf{X}_2]) - (\mathbb{E}[\mathbf{X}_1]^2 + \mathbb{E}[\mathbf{X}_2]^2 + 2\,\mathbb{E}[\mathbf{X}_1]\,\mathbb{E}[\mathbf{X}_2]) \\
&= \underbrace{\mathbb{E}[\mathbf{X}_1^2] - \mathbb{E}[\mathbf{X}_1]^2}_{\sigma^2(\mathbf{X}_1)} + \underbrace{\mathbb{E}[\mathbf{X}_2^2] - \mathbb{E}[\mathbf{X}_2]^2}_{\sigma^2(\mathbf{X}_2)} + 2\underbrace{(\mathbb{E}[\mathbf{X}_1\mathbf{X}_2] - \mathbb{E}[\mathbf{X}_1]\,\mathbb{E}[\mathbf{X}_2])}_{0 \text{ if } \mathbf{X}_1 \text{ and } \mathbf{X}_2 \text{ are independent}}
\end{aligned}$$

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

$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$

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

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

$(*)$ is by linearity of expectation.

$$\begin{aligned}\sigma^2(\mathbf{X}) &= \mathbb{E}[\mathbf{X}^2] - \mathbb{E}[\mathbf{X}]^2 \\ &= (\mathbb{E}[\mathbf{X}_1^2] + \mathbb{E}[\mathbf{X}_2^2] + 2\,\mathbb{E}[\mathbf{X}_1\mathbf{X}_2]) - (\mathbb{E}[\mathbf{X}_1]^2 + \mathbb{E}[\mathbf{X}_2]^2 + 2\,\mathbb{E}[\mathbf{X}_1]\,\mathbb{E}[\mathbf{X}_2]) \\ &= \underbrace{\mathbb{E}[\mathbf{X}_1^2] - \mathbb{E}[\mathbf{X}_1]^2}_{\sigma^2(\mathbf{X}_1)} + \underbrace{\mathbb{E}[\mathbf{X}_2^2] - \mathbb{E}[\mathbf{X}_2]^2}_{\sigma^2(\mathbf{X}_2)} + 2\underbrace{(\mathbb{E}[\mathbf{X}_1\mathbf{X}_2] - \mathbb{E}[\mathbf{X}_1]\,\mathbb{E}[\mathbf{X}_2])}_{0 \text{ if } \mathbf{X}_1 \text{ and } \mathbf{X}_2 \text{ are independent}}\end{aligned}$$

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)

$\mathbf{X} = \mathbf{X}_1 + \cdots + \mathbf{X}_n$ (sum of *independent* Bernoullis), and $\sigma^2(\mathbf{X}_i) = p(1-p)$

Variance of a Sum

$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$

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

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

$(*)$ is by linearity of expectation.

$$\begin{aligned}\sigma^2(\mathbf{X}) &= \mathbb{E}[\mathbf{X}^2] - \mathbb{E}[\mathbf{X}]^2 \\ &= (\mathbb{E}[\mathbf{X}_1^2] + \mathbb{E}[\mathbf{X}_2^2] + 2\,\mathbb{E}[\mathbf{X}_1\mathbf{X}_2]) - (\mathbb{E}[\mathbf{X}_1]^2 + \mathbb{E}[\mathbf{X}_2]^2 + 2\,\mathbb{E}[\mathbf{X}_1]\,\mathbb{E}[\mathbf{X}_2]) \\ &= \underbrace{\mathbb{E}[\mathbf{X}_1^2] - \mathbb{E}[\mathbf{X}_1]^2}_{\sigma^2(\mathbf{X}_1)} + \underbrace{\mathbb{E}[\mathbf{X}_2^2] - \mathbb{E}[\mathbf{X}_2]^2}_{\sigma^2(\mathbf{X}_2)} + 2\underbrace{(\mathbb{E}[\mathbf{X}_1\mathbf{X}_2] - \mathbb{E}[\mathbf{X}_1]\,\mathbb{E}[\mathbf{X}_2])}_{0 \text{ if } \mathbf{X}_1 \text{ and } \mathbf{X}_2 \text{ are independent}}\end{aligned}$$

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)

$\mathbf{X} = \mathbf{X}_1 + \cdots + \mathbf{X}_n$ (sum of *independent* Bernoullis), and $\sigma^2(\mathbf{X}_i) = p(1-p)$

$$\sigma^2(\text{Binomial}) = \sigma^2(\mathbf{X}_1) + \cdots + \sigma^2(\mathbf{X}_n) = p(1-p) + \cdots + p(1-p) = np(1-p).$$

3-σ Rule: $\mathbf{X} = \mu(\mathbf{X}) \pm \sigma(\mathbf{X})$

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

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

3-σ Rule: $\mathbf{X} = \mu(\mathbf{X}) \pm \sigma(\mathbf{X})$

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

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

Lemma (Markov Inequality). For a positive random variable $\mathbf{X}$,

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

3-σ Rule: $\mathbf{X} = \mu(\mathbf{X}) \pm \sigma(\mathbf{X})$

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

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

Lemma (Markov Inequality). For a positive random variable $\mathbf{X}$,

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

Proof. $\mathbb{E}[\mathbf{X}]$

3-σ Rule: $\mathbf{X} = \mu(\mathbf{X}) \pm \sigma(\mathbf{X})$

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

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

Lemma (Markov Inequality). For a positive random variable $\mathbf{X}$,

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

Proof. $\mathbb{E}[\mathbf{X}] = \sum_{x \geq 0} x \cdot P_{\mathbf{X}}(x)$

3-σ Rule: $\mathbf{X} = \mu(\mathbf{X}) \pm \sigma(\mathbf{X})$

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

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

Lemma (Markov Inequality). For a positive random variable $\mathbf{X}$,

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

Proof. $\mathbb{E}[\mathbf{X}] \;=\; \sum_{x\geq 0} x \cdot P_{\mathbf{X}}(x) \;\geq\; \sum_{x\geq\alpha} x \cdot P_{\mathbf{X}}(x)$

3-σ Rule: $\mathbf{X} = \mu(\mathbf{X}) \pm \sigma(\mathbf{X})$

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

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

Lemma (Markov Inequality). For a positive random variable $\mathbf{X}$,

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

Proof. $\mathbb{E}[\mathbf{X}] \;=\; \sum_{x \geq 0} x \cdot P_{\mathbf{X}}(x) \;\geq\; \sum_{x \geq \alpha} x \cdot P_{\mathbf{X}}(x) \;\geq\; \sum_{x \geq \alpha} \alpha \cdot P_{\mathbf{X}}(x)$

3-σ Rule: $\mathbf{X} = \mu(\mathbf{X}) \pm \sigma(\mathbf{X})$

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

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

Lemma (Markov Inequality). For a positive random variable $\mathbf{X}$,

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

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

3-σ Rule: $\mathbf{X} = \mu(\mathbf{X}) \pm \sigma(\mathbf{X})$

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

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

Lemma (Markov Inequality). For a positive random variable $\mathbf{X}$,

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

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

Lemma (Chebyshev Inequality).

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

3-σ Rule: $\mathbf{X} = \mu(\mathbf{X}) \pm \sigma(\mathbf{X})$

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

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

Lemma (Markov Inequality). For a positive random variable $\mathbf{X}$,

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

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

Lemma (Chebyshev Inequality).

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

Proof.

$$\mathbb{P}[|\mathbf{\Delta}| \geq t\sigma]$$

3-σ Rule: $\mathbf{X} = \mu(\mathbf{X}) \pm \sigma(\mathbf{X})$

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

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

Lemma (Markov Inequality). For a positive random variable $\mathbf{X}$,

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

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

Lemma (Chebyshev Inequality).

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

Proof.

$$\mathbb{P}[|\mathbf{\Delta}| \geq t\sigma] = \mathbb{P}[\mathbf{\Delta}^2 \geq t^2\sigma^2]$$

3-σ Rule: $\mathbf{X} = \mu(\mathbf{X}) \pm \sigma(\mathbf{X})$

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

Lemma (Markov Inequality). For a positive random variable $\mathbf{X}$,

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

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

Lemma (Chebyshev Inequality).

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

Proof.

$$\mathbb{P}[|\mathbf{\Delta}| \geq t\sigma] = \mathbb{P}[\mathbf{\Delta}^2 \geq t^2\sigma^2] \overset{(a)}{\leq} \frac{\mathbb{E}[\mathbf{\Delta}^2]}{t^2\sigma^2}$$

3-σ Rule: $\mathbf{X} = \mu(\mathbf{X}) \pm \sigma(\mathbf{X})$

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

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

Lemma (Markov Inequality). For a positive random variable $\mathbf{X}$,

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

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

Lemma (Chebyshev Inequality).

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

Proof.

$$\mathbb{P}[|\boldsymbol{\Delta}| \geq t\sigma] = \mathbb{P}[\boldsymbol{\Delta}^2 \geq t^2\sigma^2] \overset{(a)}{\leq} \frac{\mathbb{E}[\boldsymbol{\Delta}^2]}{t^2\sigma^2} = \frac{\cancel{\sigma^2}}{t^2\cancel{\sigma^2}}$$

3-σ Rule: $\mathbf{X} = \mu(\mathbf{X}) \pm \sigma(\mathbf{X})$

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

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

Lemma (Markov Inequality). For a positive random variable $\mathbf{X}$,

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

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

Lemma (Chebyshev Inequality).

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

Proof.

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

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:

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

Law of Large Numbers

Expectation of the average of n dice:

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

Variance of the average of n dice:

$\sigma^2(\text{average})$

Law of Large Numbers

Expectation of the average of n dice:

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

Variance of the average of n dice:

$$\sigma^2(\text{average}) = \sigma^2(\tfrac{1}{n} \times \text{sum})$$

Law of Large Numbers

Expectation of the average of n dice:

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

Variance of the average of n dice:

$$\sigma^2(\text{average}) = \sigma^2(\tfrac{1}{n} \times \text{sum}) = \tfrac{1}{n^2} \times \sigma^2(\text{sum})$$

Law of Large Numbers

Expectation of the average of n dice:

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

Variance of the average of n dice:

$$\sigma^2(\text{average}) = \sigma^2(\tfrac{1}{n} \times \text{sum}) = \tfrac{1}{n^2} \times \sigma^2(\text{sum}) = \tfrac{1}{n^2} \times n \times \sigma^2(\text{one die})$$

Law of Large Numbers

Expectation of the average of n dice:

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

Variance of the average of n dice:

$$\sigma^2(\text{average}) = \sigma^2(\tfrac{1}{n} \times \text{sum}) = \tfrac{1}{n^2} \times \sigma^2(\text{sum}) = \tfrac{1}{n^2} \times n \times \sigma^2(\text{one die}) = \tfrac{1}{n} \times \sigma^2(\text{one die})$$

Law of Large Numbers

Expectation of the average of n dice:

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

Variance of the average of n dice:

$$\sigma^2(\text{average}) = \sigma^2(\tfrac{1}{n} \times \text{sum}) = \tfrac{1}{n^2} \times \sigma^2(\text{sum}) = \tfrac{1}{n^2} \times n \times \sigma^2(\text{one die}) = \tfrac{1}{n} \times \sigma^2(\text{one die})$$