

# Foundations of Computer Science

## Lecture 21

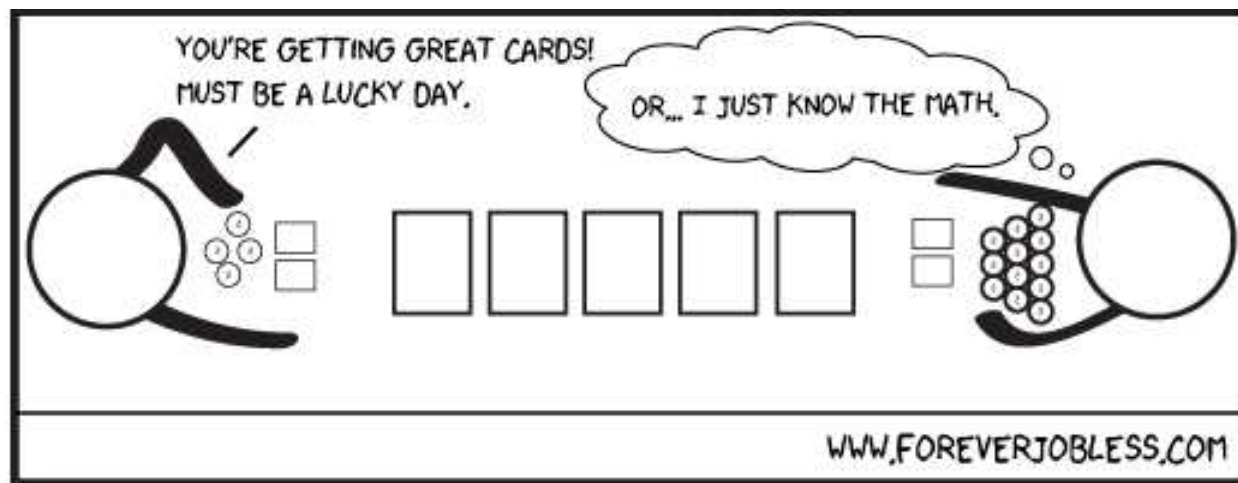
### Deviations from the Mean

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  $3\text{-}\sigma$  Rule



- 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.

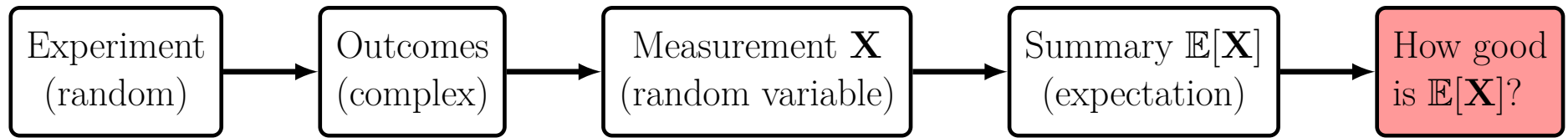
# Today: Deviations from the Mean

---

- 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  $3\text{-}\sigma$  rule.

# Probability For Analyzing a Random Experiment.

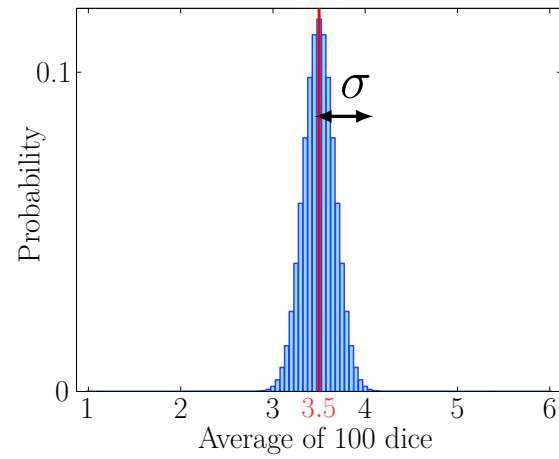
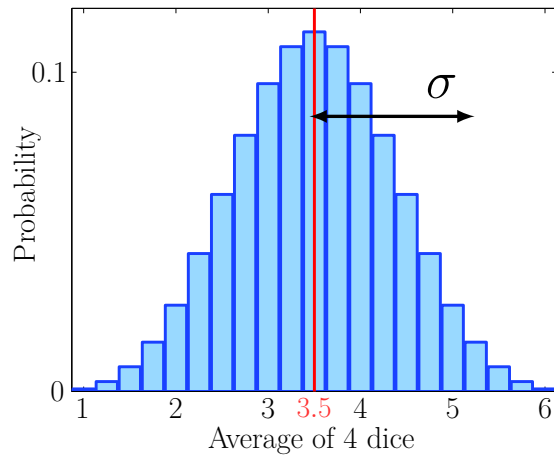
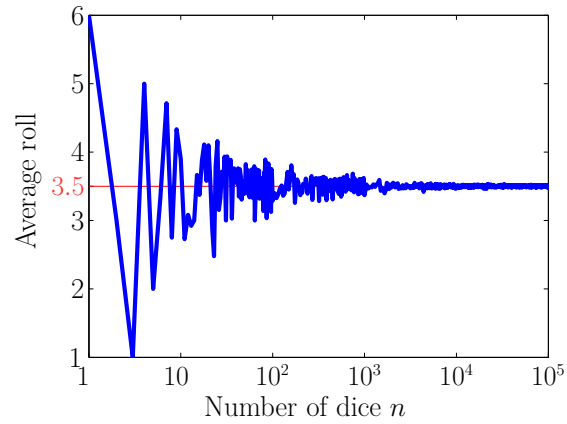
---



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

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

# Average of $n$ Dice



# Variance: Size of the Deviations From the Mean

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

$\mathbf{X}$	2	3	4	5	6	7	8	9	10	11	12	
$\Delta$	-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  $\mathbb{E}[\Delta]$ ?

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

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

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

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

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

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

sum of two dice rolls =  $7 \pm 2.52$ .

**Practice.** Exercise 21.2.

# Variance is a Measure of Risk

---

## Game 1

$\mathbf{X}_1$  : win \$2    probability =  $\frac{2}{3}$ ;  
lose \$1    probability =  $\frac{1}{3}$ .

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

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

## Game 2

$\mathbf{X}_2$  : win \$102    probability =  $\frac{2}{3}$ ;  
lose \$201    probability =  $\frac{1}{3}$ .

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

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

$$\mathbf{X}_1 = 1 \pm 1.41$$

$$\mathbf{X}_2 = 1 \pm \mathbf{141}$$

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

# A More Convenient Formula for Variance

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

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

Sum of two dice,

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

Since  $\mu = 7$

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

**Theorem.** Variance  $\geq 0$ , 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  $\mathbb{E}[\mathbf{X}] = \frac{1}{2}(n + 1)$ .

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

so

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

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

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

so

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

# Linearity of Variance?

---

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

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

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

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

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

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

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

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

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

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

---

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

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

**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).**

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

*Proof.*

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

In (a) we used Markov's Inequality. ■

To get the 3- $\sigma$  rule, use Chebyshev's Inequality with  $t = 3$ .

# Law of Large Numbers

Expectation of the average of  $n$  dice:

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

Variance of the average of  $n$  dice:

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

