Warmup Problem

There's a random math fact we'll need later, namely that

$$\left(\sum_{i}^{n} x_i\right)^2 = \sum_{i}^{n} x_i^2 + 2\sum_{i<j} x_i x_j$$

Example: $(500 + a + b + c)^2 =$

Fill in the right hand side using the random math fact.

CC BY NC SA

Warmup Problem

There's a random math fact we'll need later, namely that

$$\left(\sum_i^n x_i\right)^2 = \sum_i^n x_i^2 + 2\sum_{i<j} x_i x_j$$

Example: $(500 + a + b + c)^2 = 500^2 + a^2 + b^2 + c^2$

$$+1000a + 1000b + 1000c + 2ab + 2ac + 2bc$$

CC BY NC SA

Coupon Collector and Expectation

Lecture 21, CS70 Fall 2026

Coupon Collector and Expectation

Variance

- Law of the Unconscious Statistician
- Derivation and Definition
- Properties of Variance
- Example Variance Calculations
- Sums of Random Variables

CC BY NC SA

Coupon Collector: Problem Statement and Empirical Result

We'll start today with the "Coupon Collector Problem".

Let's see a quick simulation:

Coupons, n **50** Boxes to buy, m **220** Time per box **10 ms**

Buy | Simulate | Reset

7	1	7	6		10	6	5	6	2
1	4	2	7	6	1	3	4	7	8
2	4	2	4	5	3	8	4	2	5
8	9	10	2	2	7	4	4	5	1
1	3	1	3	5	8	1	4	5	5

Each time we buy a box of cereal, we get one of the n coupons, chosen uniformly at random.

Boxes bought
220 / 220

Coupons collected
49 / 50

Missing 1 coupon.

Trials where we got every coupon
11 / 20 **= 55.0%**

CC BY NC SA

Coupon Collector: Problem Statement and Empirical Result

Now let's keep buying boxes until we have all n coupons:

Coupons, n **50**

Time per box **2 ms**

Simulate | **+10 trials** | **+100 trials** | **Reset**

4	4	3	3	4	4	2	4	4	4
5	2	3	5	7	4	3	4	3	5
4	3	4	2	5	4	1	5	3	1
3	5	3	3	1	5	3	4	2	1
4	3	3	2	3	3	8	3	2	3

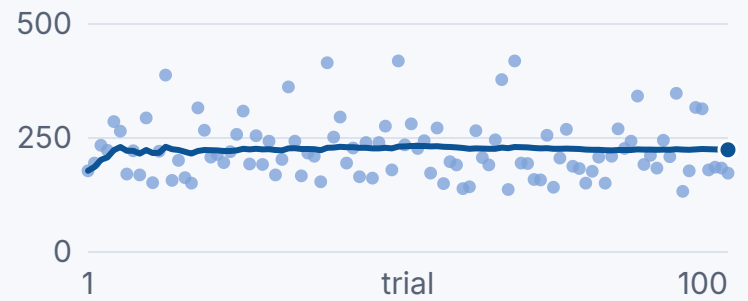
Each time we buy a box of cereal, we get one of the n coupons, chosen uniformly at random.

Boxes this trial
173
50 of 50 coupons

Trials
100

Average boxes bought per trial
224.15

Running average



Latest trials: **173** 184 186 180 314 317

CC BY NC SA

Coupon Collector: Problem Statement and Empirical Result

Let X be the number of cereal boxes we need to buy before we get all n coupons.

n	$E[X]$
25	95.40
50	224.96
75	367.60
100	518.74

Our goal is to compute the average number of boxes $E[X]$ we need to buy for a given n.

CC BY NC SA

Coupon Collector: Framed as a Random Variable

Let X be the number of cereal boxes we need to buy before we get all n coupons. Our goal is to find $E[X]$.

What is the domain of this random variable X? What is the range of X?

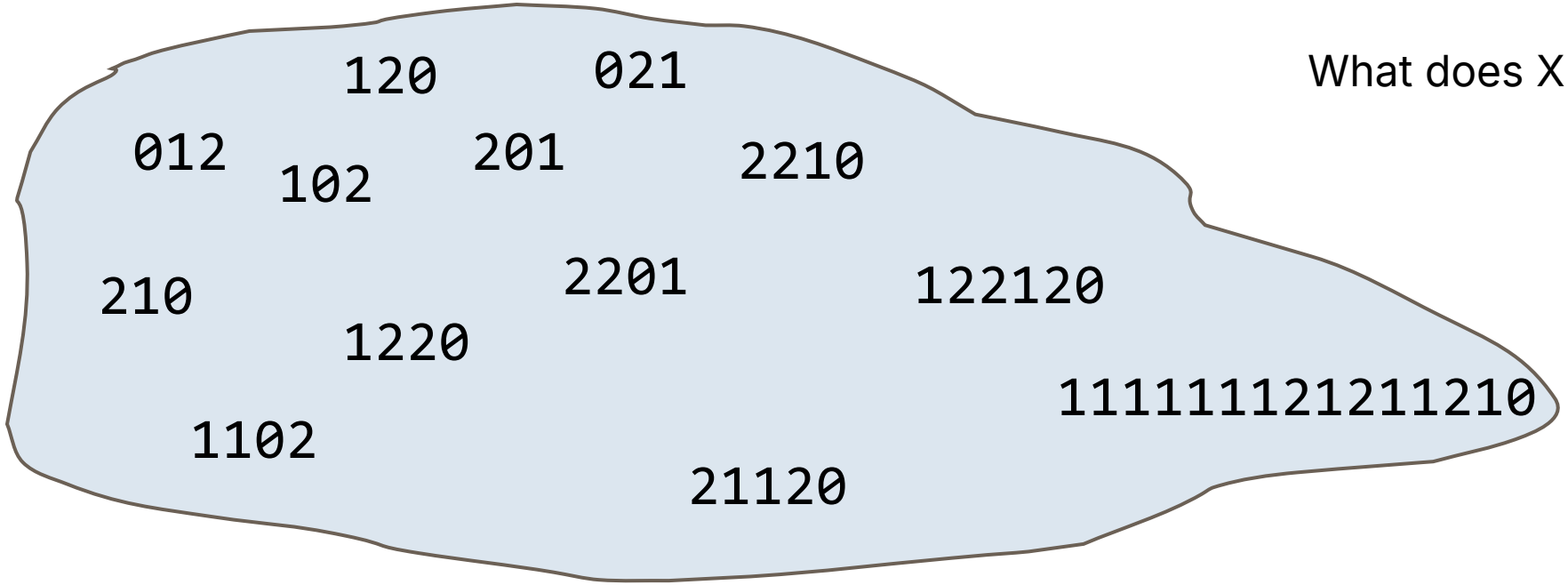
CC BY NC SA

Coupon Collector: Framed as a Random Variable

Let X be the number of cereal boxes we need to buy before we get all n coupons. What is the domain of this random variable X? What is the range of X?

- The domain of X is the set of sequences of purchases that end with last coupon.
- The range of X is the set of all integers greater than or equal to n.

Example: If $n = 3$, the sample space is infinite, and includes:



What does X "do"?

Coupon Collector: Framed as a Random Variable

Let X be the number of cereal boxes we need to buy before we get all n coupons. Our goal: Compute $E[X]$.

- Note: We often don't explicitly think about the domain of the RV, but it exists!
- Usually just think "X can take on any value greater than or equal to n"

How do we compute $E[X]$? This is a very interesting puzzle. If you're watching this online or reading these slides, maybe take 10 minutes to think about where you'd even start.

- Hint: The distribution of X isn't something we've discussed so far. However, you can decompose X into the sum of random variables that we have studied.

CC BY NC SA

Decomposing X as a Sum of N Random Variables

Let X be the number of cereal boxes we need to buy before we get all n coupons. Our goal: Compute $E[X]$.

CC BY NC SA

Decomposing X as a Sum of N Random Variables

Let X be the number of cereal boxes we need to buy before we get all n coupons. Our goal: Compute $E[X]$.

The first key insight is that we can define $X = X_1 + X_2 + \cdots + X_n$, where X_i is the number of cereal boxes we need to buy before we get the ith new coupon.

- Example: X_1 is always 1 because we always get a new coupon on the first purchase (we have no other coupons).

Again, if you're not in live lecture, consider taking 10 minutes to see if you can figure out $E[X]$ on your own starting from this first insight.

CC BY NC SA

Decomposing X as a Sum of N Random Variables

We define $X = X_1 + X_2 + \cdots + X_n$, where X_i is the number of cereal boxes we buy while trying to get the ith new coupon.

- Example: X_1 is always 1 because we always get a new coupon on the first purchase (we have no other coupons).

Then suppose we buy:

- 2 cereal boxes while trying to get the 2nd new coupon.
- 4 cereal boxes while trying to get the 3rd new coupon.
- 1 cereal box while trying to get the 4th new coupon (lucky!).
- 9 cereal boxes while trying to get the 5th new coupon (unlucky!).

Then $X = 1 + 2 + 4 + 1 + \cdots$

CC BY NC SA

Finding the Distribution of X_2

We define $X = X_1 + X_2 + \cdots + X_n$, where X_i is the number of cereal boxes we buy while trying to get the ith new coupon.

- Example: X_1 is always 1 because we always get a new coupon on the first purchase (we have no other coupons).

What is the distribution for X_2?

- joshh.ug/A



Feel free to discuss with your neighbor. This is a free response question. Your answer should be the name of a distribution and its parameter.

CC BY NC SA

Finding the Distribution of X_2

We define $X = X_1 + X_2 + \cdots + X_n$, where X_i is the number of cereal boxes we buy while trying to get the ith new coupon.

- Example: X_1 is always 1 because we always get a new coupon on the first purchase (we have no other coupons).

What is the distribution for X_2?

CC BY NC SA

Finding the Distribution of X_2

We define $X = X_1 + X_2 + \cdots + X_n$, where X_i is the number of cereal boxes we buy while trying to get the ith new coupon.

- Example: X_1 is always 1 because we always get a new coupon on the first purchase (we have no other coupons).

What is the distribution for X_2?

- X_2 is how many boxes we need to buy if $\frac{n-1}{n}$ chance of a new coupon.
- That is, each box we buy has an $\frac{n-1}{n}$ chance of giving us the new coupon. So it's basically like we're flipping a coin until we get our first heads, where the probability of heads is $\frac{n-1}{n}$.
- This is precisely the geometric distribution, i.e. $X_2 \sim \text{Geometric}\left(\frac{n-1}{n}\right)$

Finding the Expectation of X_i

We define $X = X_1 + X_2 + \cdots + X_n$, where X_i is the number of cereal boxes we buy while trying to get the ith new coupon.

- Example: X_1 is always 1 because we always get a new coupon on the first purchase (we have no other coupons).

What is $\mathbf{E}[X_i]$?

- $X_1 = 1$
- $X_2 \sim \textbf{Geometric}\left(\frac{n-1}{n}\right) \Rightarrow E[X_2] = \frac{n}{n-1}$

What is $\mathbf{E}[X_3]$? In other words, how many boxes do we expect to need to buy to get the third coupon? joshh.ug/B

Finding the Expectation of X_i

We define $X = X_1 + X_2 + \cdots + X_n$, where X_i is the number of cereal boxes we buy while trying to get the ith new coupon.

- Example: X_1 is always 1 because we always get a new coupon on the first purchase (we have no other coupons).

What is $\mathbf{E}[X_i]$?

- $X_1 = 1$
- $X_2 \sim \text{Geometric}\left(\frac{n-1}{n}\right) \Rightarrow E[X_2] = \frac{n}{n-1}$
- $X_3 \sim \text{Geometric}\left(\frac{n-2}{n}\right) \Rightarrow E[X_3] = \frac{n}{n-2}$
- $\ldots$

CC BY NC SA

The Expectation of X

We define $X = X_1 + X_2 + \cdots + X_n$, where X_i is the number of cereal boxes we buy while trying to get the ith new coupon.

$$\mathrm{E}[X] = E[X_1] + E[X_2] + E[X_3] + \cdots + E[X_n]$$

$$E[X_1] = 1 \qquad E[X_2] = \frac{n}{n-1} \qquad E[X_3] = \frac{n}{n-2} \qquad E[X_4] = \frac{n}{n-3} \qquad \ldots$$

CC BY NC SA

The Expectation of X

We define $X = X_1 + X_2 + \cdots + X_n$, where X_i is the number of cereal boxes we buy while trying to get the ith new coupon.

$$\mathrm{E}[X] = E[X_1] + E[X_2] + E[X_3] + \cdots + E[X_n]$$

$$E[X_1] = 1 \qquad E[X_2] = \frac{n}{n-1} \qquad E[X_3] = \frac{n}{n-2} \qquad E[X_4] = \frac{n}{n-3} \qquad \ldots$$

$$\begin{aligned}\mathrm{E}[X] &= \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{2} + \frac{n}{1} \\ &= n\left(\frac{1}{n} + \frac{1}{n-1} + \frac{1}{n-2} + \cdots + \frac{1}{2} + \frac{1}{1}\right) \\ &= n\sum_{i=1}^{n}\frac{1}{i}\end{aligned}$$

Example: $n = 50$, we have

$$50 \times \left(\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{50}\right) = 224.96$$

CC BY NC SA

Approximation of the Expectation of X

We define $X = X_1 + X_2 + \cdots + X_n$, where X_i is the number of cereal boxes we buy while trying to get the ith new coupon.

We've shown that $\mathbf{E}[X] = n \sum_{i=1}^{n} \frac{1}{i}$

Example: $n = 50$, we have

$$50 \times \left(\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{50} \right) = 224.96$$

There exists a nice approximation: $\sum_{i=1}^{n} \frac{1}{i} \approx \ln n + \gamma_E$

where $\gamma_E = 0.5772\ldots$

Example: $E[X] \approx n(\ln n + \gamma_E) \approx 50(\ln 50 + 0.5772\ldots) = 224.46$. Not bad!

CC BY NC SA

The Law of the Unconscious Statistician

Lecture 21, CS70 Fall 2026

Coupon Collector and Expectation

Variance

- **Law of the Unconscious Statistician**
- Derivation and Definition
- Properties of Variance
- Example Variance Calculations
- Sums of Random Variables

CC BY NC SA

Expectation of a Function of a Random Variable

Suppose we flip a fair coin twice. Let X be the number of heads.

What is $E[\cos(X)]$?

CC BY NC SA

Expectation of a Function of a Random Variable

Suppose we flip a fair coin twice. Let X be the number of heads.

$$\overbrace{\text{new RV } Q = \cos(X)}$$

What is $E[\cos(X)]$?

$$X = \begin{cases} 0, & \text{with probability } 1/4 \\ 1, & \text{with probability } 1/2 \\ 2, & \text{with probability } 1/4 \end{cases}$$

$$\begin{aligned} E[Q] = E[\cos(X)] &= \frac{1}{4} \cdot \cos(0) \\ &+ \frac{1}{2} \cdot \cos(1) \\ &+ \frac{1}{4} \cdot \cos(2) \\ &\approx 0.416 \end{aligned}$$

$$Q = \begin{cases} \cos(0), & \text{with probability } 1/4 \\ \cos(1), & \text{with probability } 1/2 \\ \cos(2), & \text{with probability } 1/4 \end{cases}$$

CC BY NC SA

The Law of the Unconscious Statistician

The notes call this the Law of the Unconscious Statistician (LOTUS).

Suppose we have a random variable X and we define a new random variable $Q = f(X)$. Then we have that:

$$E[Q] = E[f(X)] = \sum_{x \in \text{range}(X)} f(x)\boldsymbol{P}(X = x)$$

Example: If $X \sim \text{Binomial}(2, 0.5)$, and $Q = \cos(X)$ then

$$E[Q] = \sum_{x=0}^{2} \cos(x) P(X = x) = \frac{1}{4} \cdot \cos(0) + \frac{1}{2} \cdot \cos(1) + \frac{1}{4} \cdot \cos(2)$$

CC BY NC SA

The Law of the Unconscious Statistician

The notes call this the Law of the Unconscious Statistician (LOTUS).

Suppose we have a random variable X and we define a new random variable $Q = f(X)$. Then we have that:

$$E[Q] = E[f(X)] = \sum_{x \in \text{range}(X)} f(x)\boldsymbol{P}(X = x)$$

Example: If $X \sim \text{Binomial}(50, 0.3)$, and $Q = \cos(X)$ then

$$E[Q] = \sum_{x=0}^{50} \cos(x) P(X = x) = \sum_{x=0}^{50} \cos(x) \binom{50}{x} 0.3^x 0.7^{50-x}$$

CC BY NC SA

The Law of the Unconscious Statistician

The notes call this the Law of the Unconscious Statistician (LOTUS).

Suppose we have a random variable X and we define a new random variable $Q = f(X)$. Then we have that:

$$E[Q] = E[f(X)] = \sum_{x \in \text{range}(X)} f(x)\boldsymbol{P}(X = x)$$

Example: If $X \sim \text{Binomial}(50, 0.3)$, and $Q = X^2$ then

$$E[Q] = \sum_{x=0}^{50} x^2 P(X = x) = \sum_{x=0}^{50} x^2 \binom{50}{x} 0.3^x 0.7^{50-x}$$

CC BY NC SA

The Law of the Unconscious Statistician

The notes call this the Law of the Unconscious Statistician (LOTUS).

Suppose we have a random variable X and we define a new random variable $Q = f(X)$. Then we have that:

$$E[Q] = E[f(X)] = \sum_{x \in \text{range}(X)} f(x)\boldsymbol{P}(X = x)$$

Important note: $f(E[X])$ is not generally equal to $E[f(X)]$

- Example: For two coins, $E[\cos(X)] \neq \cos(E[X])$.
- Example: $(E[X])^2 \neq E[X^2]$ in general.

CC BY NC SA

The Law of the Unconscious Statistician

$$E[Q] = E[f(X)] = \sum_{x \in \text{range}(X)} f(x)\boldsymbol{P}(X = x)$$

Why is this true?

$$E[Q] = \sum_{\omega \in \Omega} Q(\omega) \times \boldsymbol{P}(\omega) = Q(\text{HH}) \times \boldsymbol{P}(\text{HH}) + Q(\text{HT}) \times \boldsymbol{P}(\text{HT}) + Q(\text{TH}) \times \boldsymbol{P}(\text{TH}) + Q(\text{TT}) \times \boldsymbol{P}(\text{TT})$$

$$= \sum_{\omega \in \Omega} f(X(\omega)) \times \boldsymbol{P}(\omega) = \cos(X(\text{HH})) \times \boldsymbol{P}(\text{HH}) + \cos(X(\text{HT})) \times \boldsymbol{P}(\text{HT}) + \cos(X(\text{TH})) \times \boldsymbol{P}(\text{TH}) + \cos(X(\text{TT})) \times \boldsymbol{P}(\text{TT})$$

$$= \sum_{a \in \text{range}(X)} \sum_{\omega : X(\omega) = a} f(a) \times \boldsymbol{P}(\omega) = \cos(2) \times \boldsymbol{P}(\text{HH}) + \cos(1) \times \boldsymbol{P}(\text{HT}) + \cos(1) \times \boldsymbol{P}(\text{TH}) + \cos(0) \times \boldsymbol{P}(\text{TT})$$

$$= \sum_{a \in \text{range}(X)} f(a) \times \boldsymbol{P}(X = a) = \cos(2) \times \boldsymbol{P}(X = 2) + \cos(1) \times \boldsymbol{P}(X = 1) + \cos(0) \times \boldsymbol{P}(X = 0)$$

CC BY NC SA

Variance Derivation and Definition

Lecture 21, CS70 Fall 2026

Coupon Collector and Expectation

Variance

- Law of the Unconscious Statistician
- **Derivation and Definition**
- Properties of Variance
- Example Variance Calculations
- Sums of Random Variables

CC BY NC SA

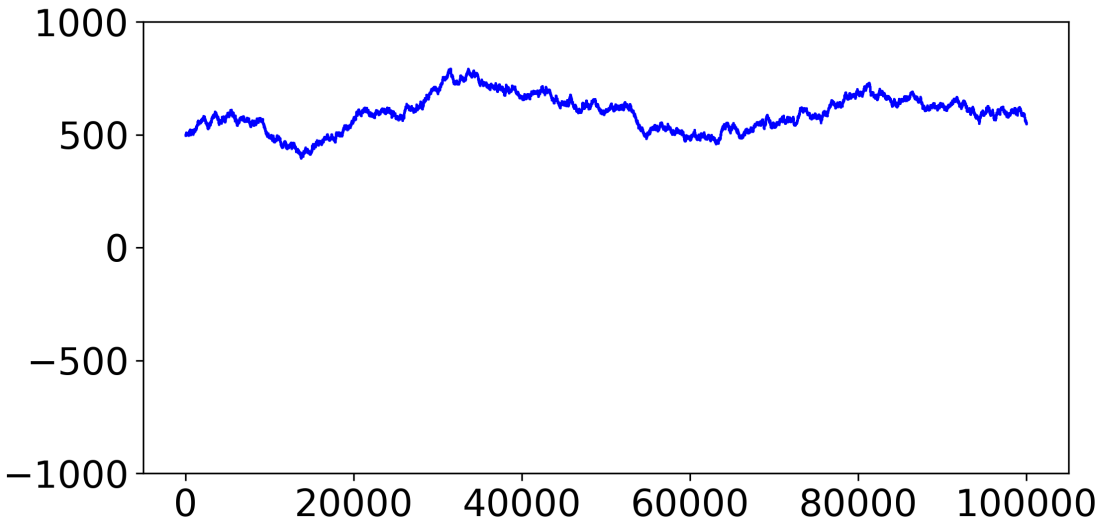
Motivation: Random Walks

Suppose we have a particle that starts at $\mathbf{y = 500}$. Every time step, it either:

- Moves up by 1 with probability ½.
- Moves down by 1 with probability ½.

$$P(\mathrm{M_i} = a) = \begin{cases} 1/2, & \text{if } a = +1 \\ 1/2, & \text{if } a = -1 \end{cases}$$

This is called a "random walk". For example, the plot below shows a random walk where the x-axis is time, and the y-axis is the position of the particle.



Motivation: Random Walks

Suppose we have a particle that starts at $\mathbf{y = 500}$. Every time step, it either:

- Moves up by 1 with probability ½.
- Moves down by 1 with probability ½.

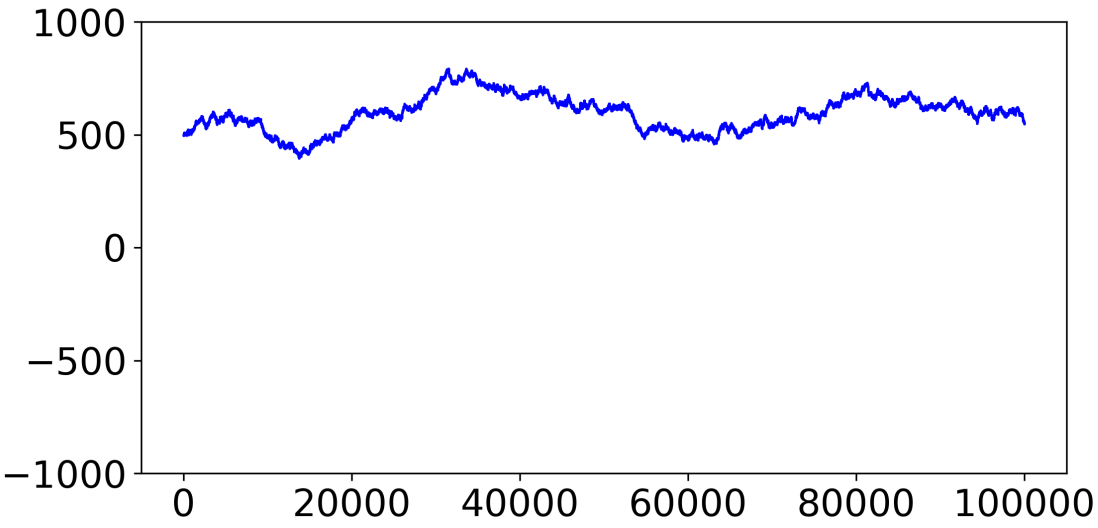
$$P(\mathrm{M_i} = a) = \begin{cases} 1/2, & \text{if } a = +1 \\ 1/2, & \text{if } a = -1 \end{cases}$$

Let Y_n be the position of the particle after n moves, i.e. Y_n is a sequence of RVs.

- (Note: Y_n is called a "random process", though we won't use this term in 70)
- $Y_n = 500 + M_1 + M_2 + \cdots + M_n$

We can easily show that $E[Y_n] = 500$:

- $\mathbb{E}[Y_n] = \mathbb{E}[500 + M_1 + M_2 + \cdots + M_n]$
  $= 500 + n\mathbb{E}[M_i]$
  $= 500 + n \times 0 = 500$



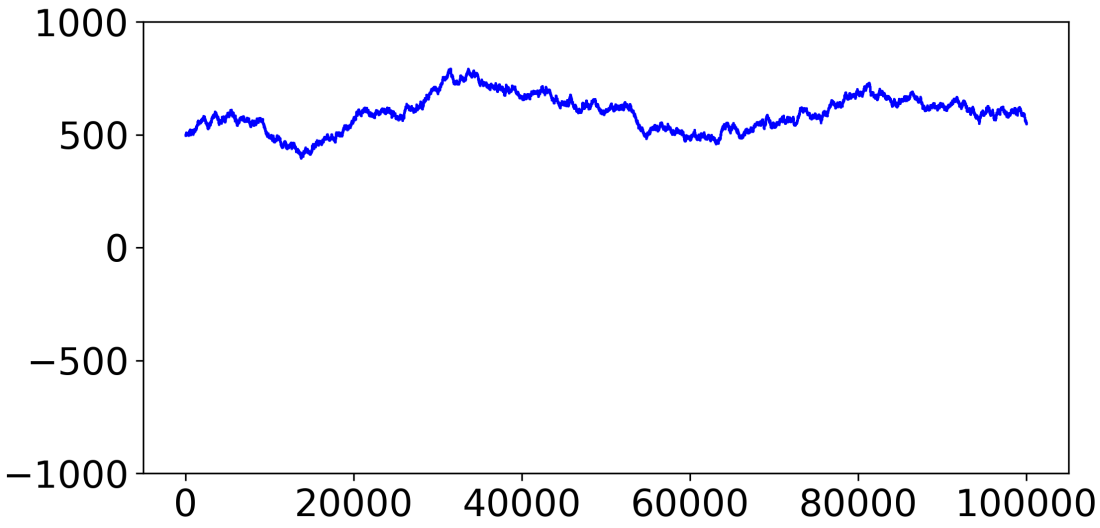
Motivation: Random Walks

Suppose we have a particle that starts at $\mathbf{y} = 500$. Every time step, it either:

- Moves up by 1 with probability ½.
- Moves down by 1 with probability ½.

$$P(\mathrm{M_i} = a) = \begin{cases} 1/2, & \text{if } a = +1 \\ 1/2, & \text{if } a = -1 \end{cases}$$

Let Y_n be the position of the particle after n moves. $\mathbb{E}[Y_n] = 500$.



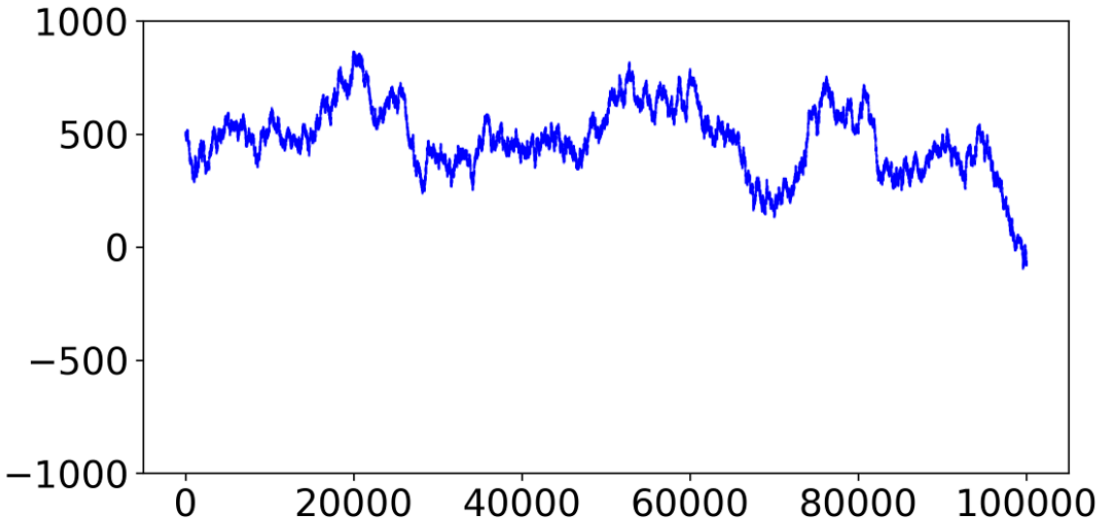
Motivation: Random Walks

Suppose we have a particle that starts at $\mathbf{y = 500}$. Every time step, it either:

- Moves up by **3** with probability ½.
- Moves down by **3** with probability ½.

$$P(\mathrm{M_i} = a) = \begin{cases} 1/2, & \text{if } a = +\mathbf{3} \\ 1/2, & \text{if } a = -\mathbf{3} \end{cases}$$

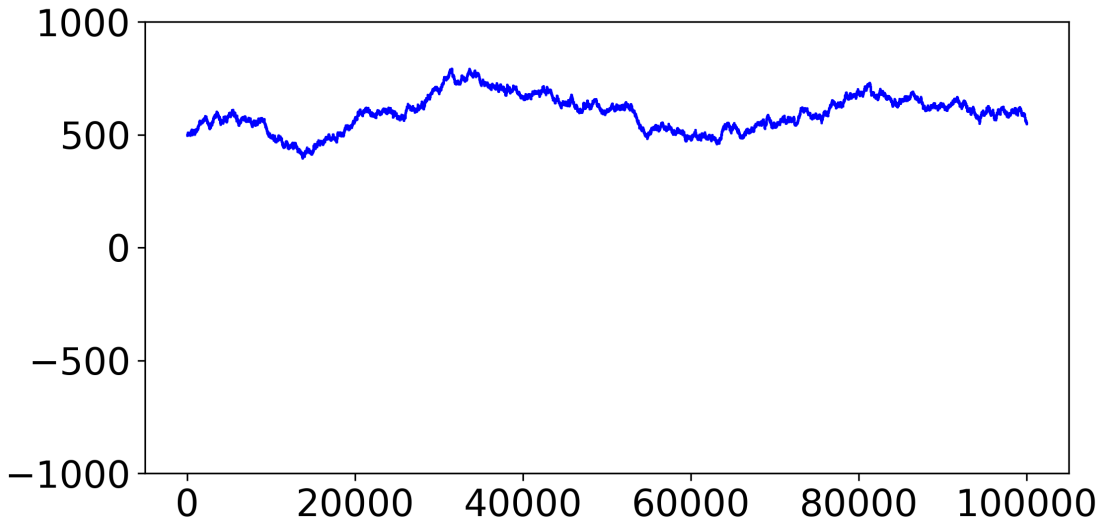
Let Y_n be the position of the particle after n moves. $\mathbb{E}[Y_n] = 500$.

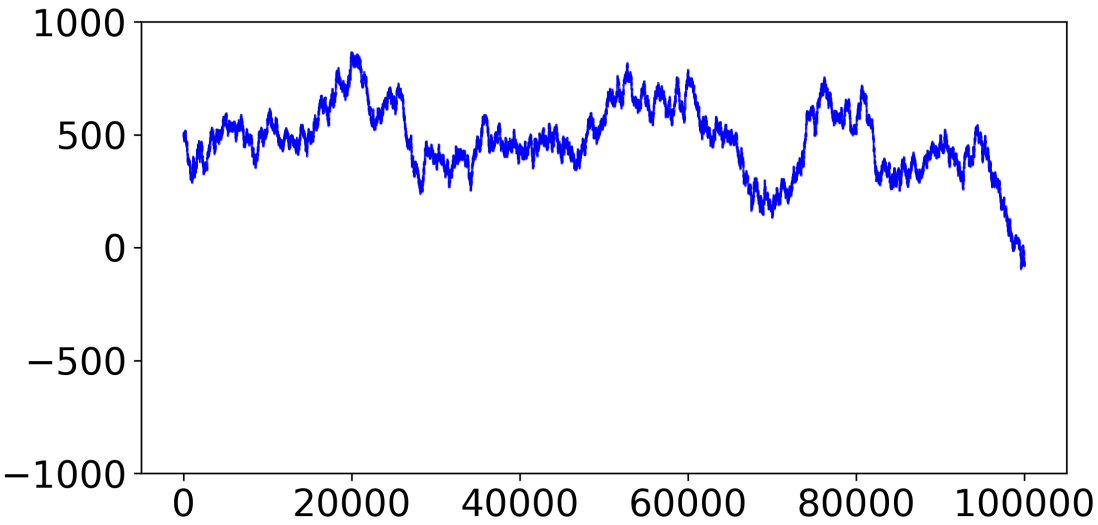


Motivation: Random Walks

Both of our random walks have an expected value of 500.

- ... but our random walk with a larger step size seems to swing more wildly.
- If we only use the expectation to summarize the random variable, we completely miss this important difference.



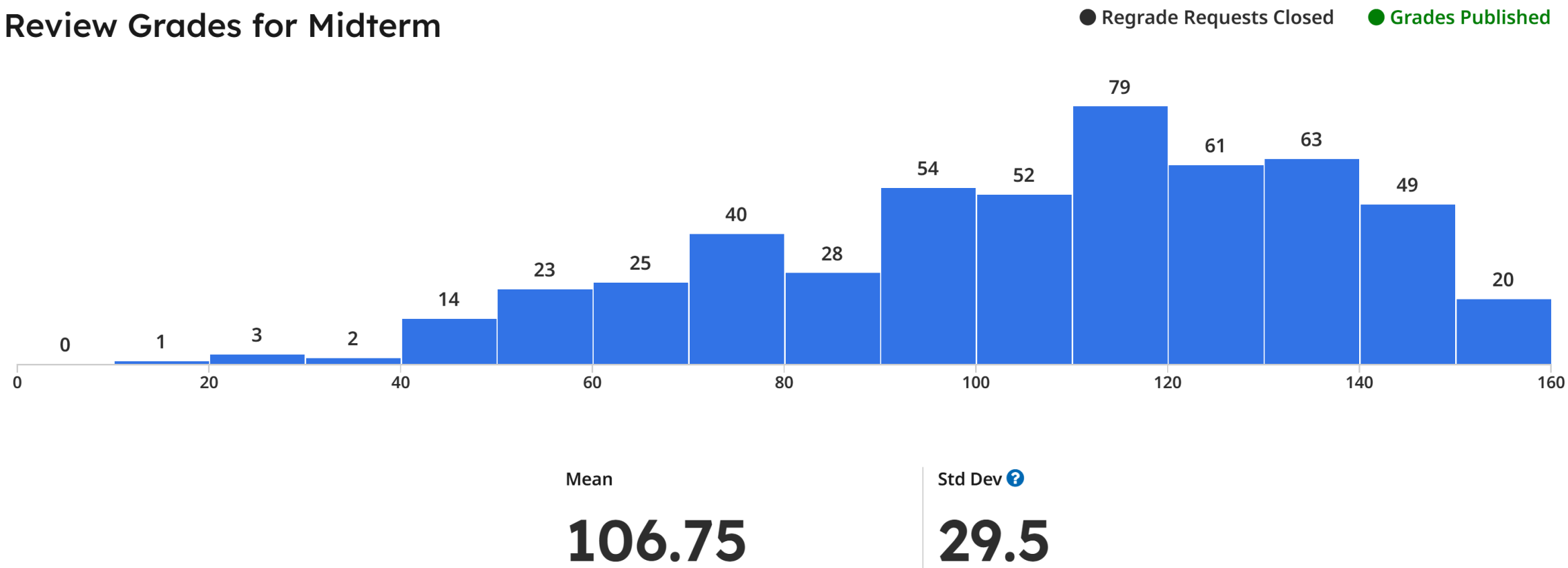


Measuring Dispersion

We'd like to measure the dispersion (or spread) of a distribution.

One summary statistic for dispersion is the standard deviation.

- But it's not the only choice.



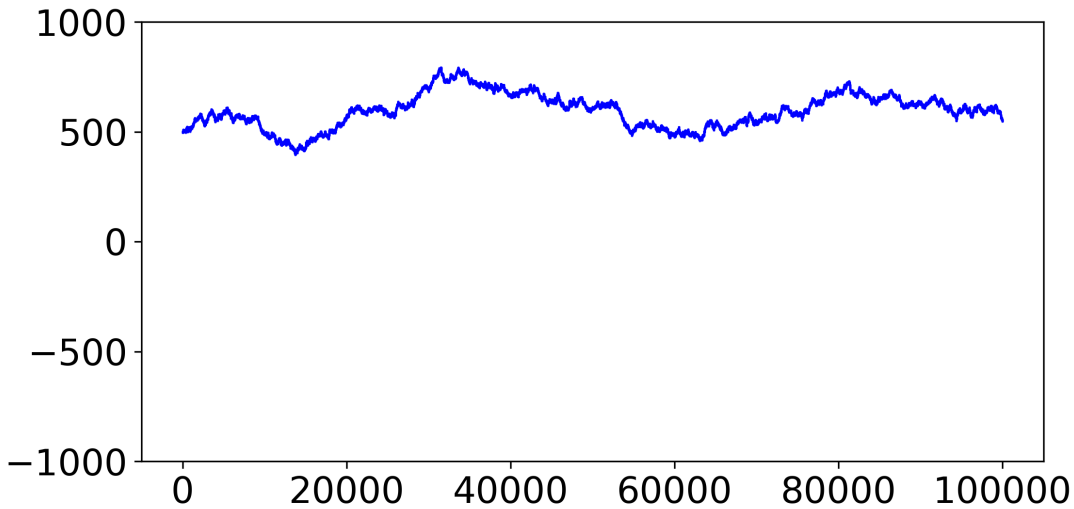
Attempt One: $\mathbb{E}[Y_n - \mathbb{E}[Y_n]]$

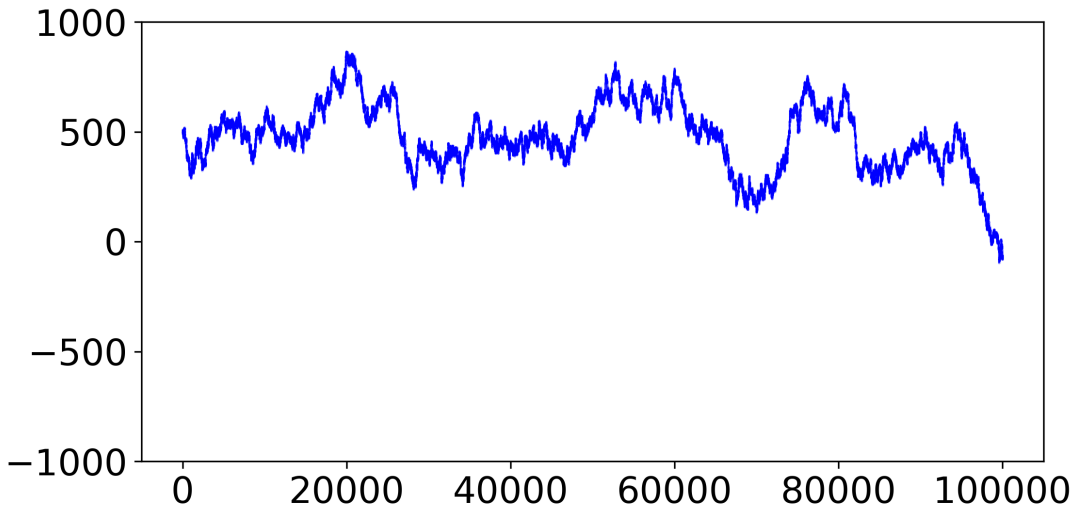
As a first attempt, imagine we attempt to define the "difference from the mean".

- Difference from the mean: $\mathbf{Y_n - \mathbb{E}[Y_n]}$

The difference from the mean is also a random variable.

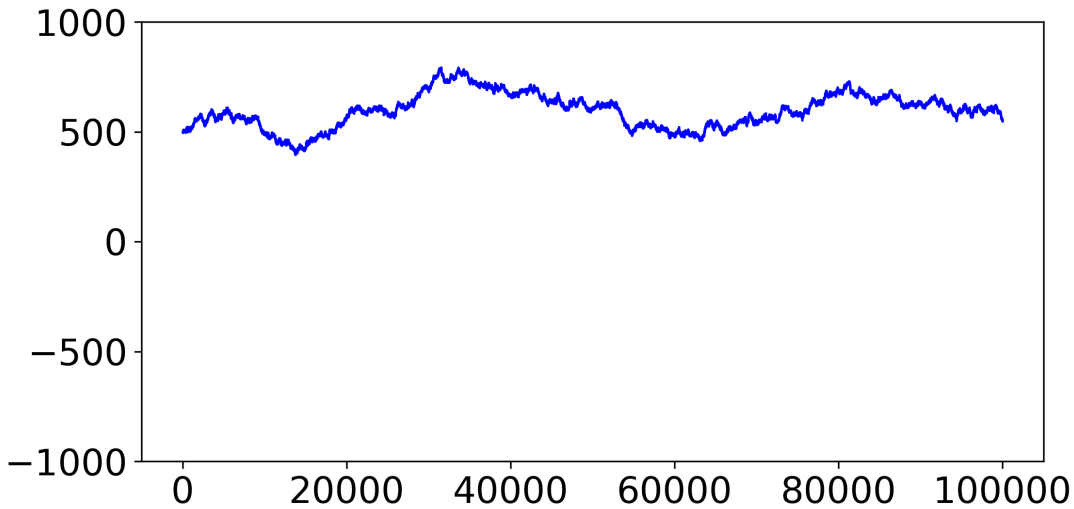
- A natural summary statistic would be the "expected difference from the mean" or $\mathbb{E}[Y_n - \mathbb{E}[Y_n]]$.
- Note: A summary statistic is a number which summarizes a distribution.

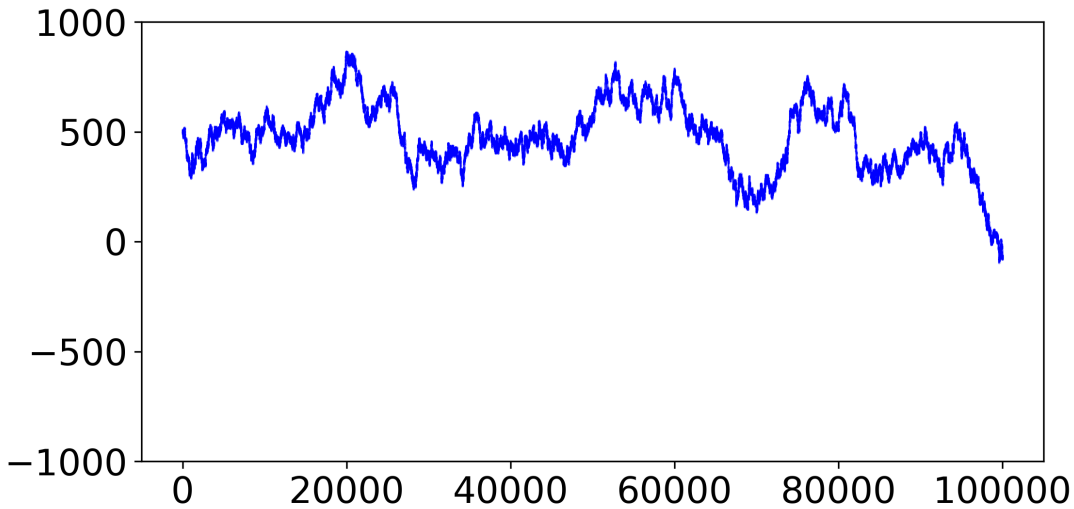




Attempt One: $\mathbb{E}[Y_n - \mathbb{E}[Y_n]]$

As a first attempt, imagine we attempt to define the "difference from the mean".





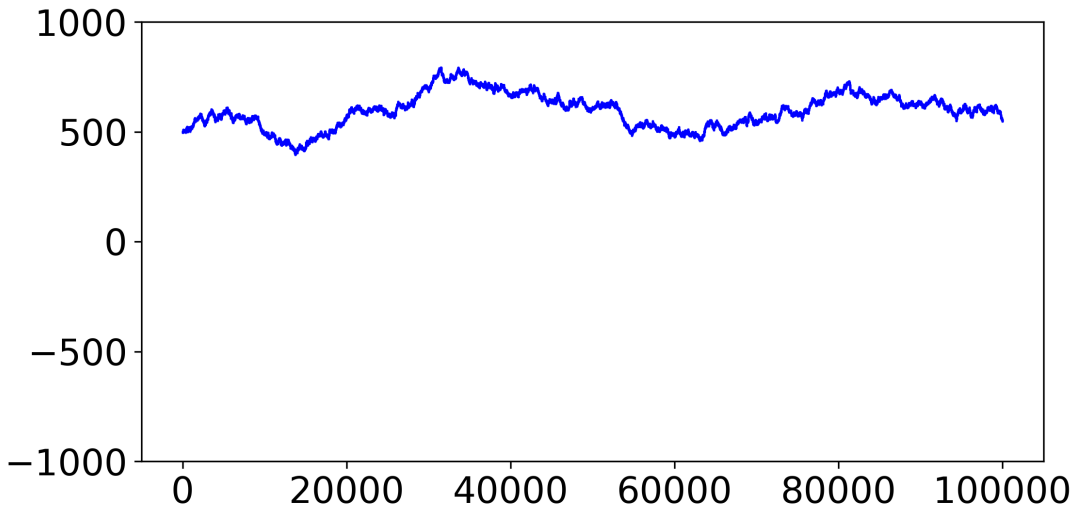
Attempt One: $\mathbb{E}[Y_n - \mathbb{E}[Y_n]]$

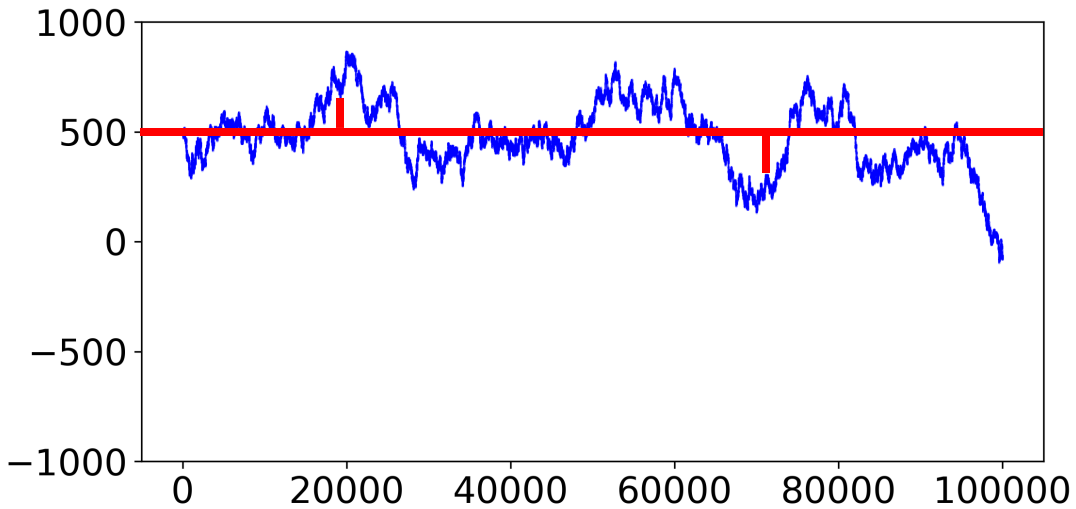
As a first attempt, imagine we attempt to define the "difference from the mean".

- Expected difference from the mean: $\mathbb{E}[\mathbf{Y_n} - \mathbb{E}[Y_n]]$

Using linearity of expectation, we get:

- $\mathbb{E}[Y_n - 500] = \mathbb{E}[Y_n] - \mathbb{E}[500] = E[Y_n] - 500 = 500 - 500 = 0.$
- Unfortunately, this quantity is on average zero.
- Positive and negative deviations cancel out.





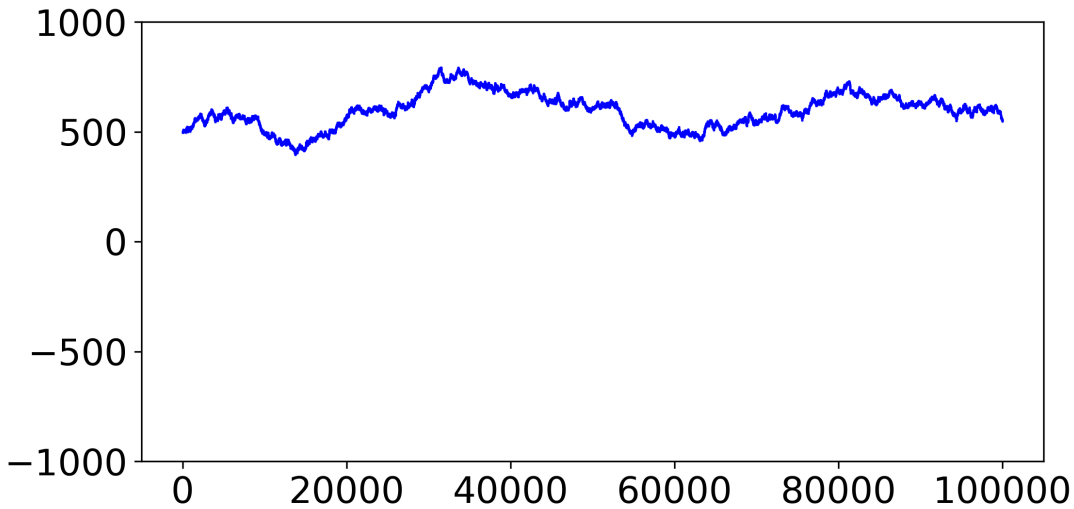
Attempt Two: $\mathbb{E}[|Y_n - \mathbb{E}[Y_n]|]$

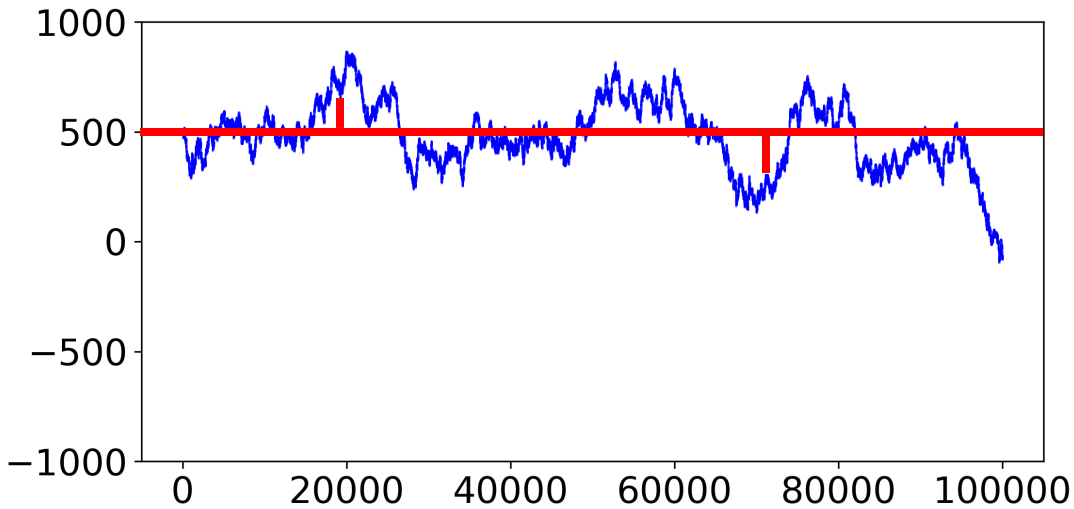
Our next idea: Compute the expected absolute difference from the mean.

- $\mathbb{E}[|Y_n - \mathbb{E}[Y_n]|] = \mathbb{E}[|Y_n - 500|]$

Now we're a bit stuck. The absolute value stops us from using linearity of expectation to distribute the operator.

- We could start trying to break this into cases, e.g. when $Y_n \leq 500$ vs. $Y_n > 500$, but for many reasons we won't pursue this approach.





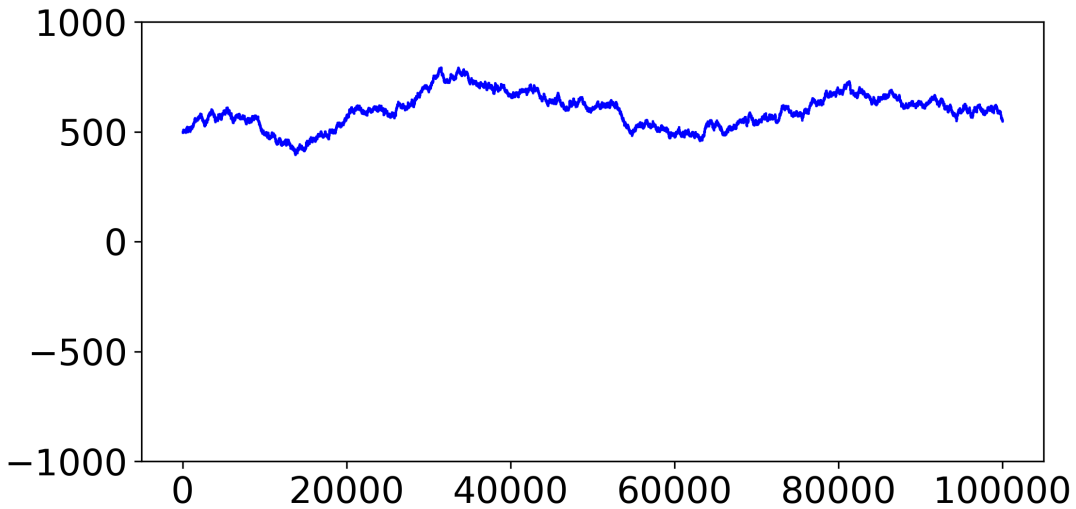
Attempt Three: $\mathbb{E}[(Y_n - \mathbb{E}[Y_n])^2]$

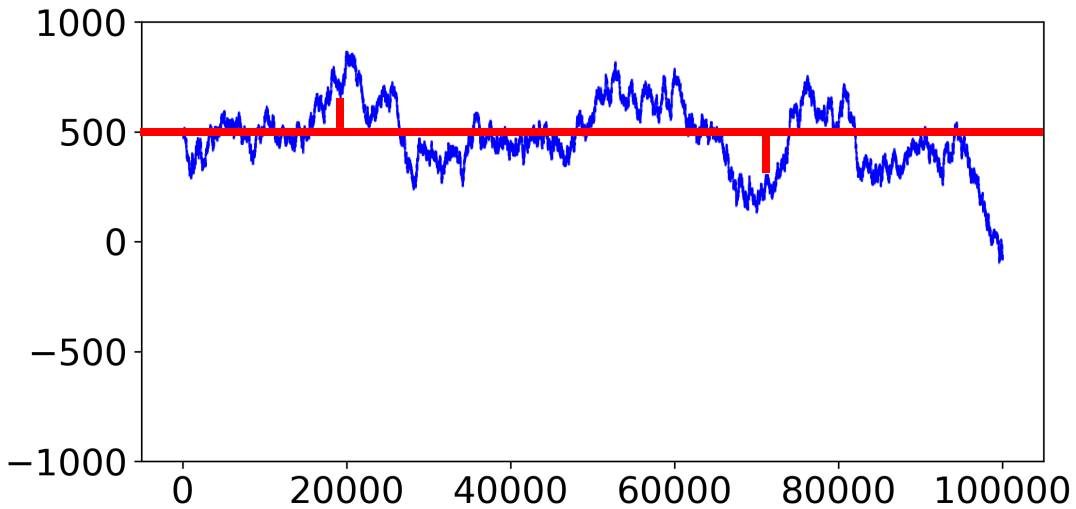
An alternate way to convert the distance into a positive value is to square it.

- That is, let's compute the expected squared difference between Y_n and $\mathbb{E}[Y_n]$.

Using linearity of expectation, we can try to compute this quantity.

- $\mathbb{E}[(Y_n - \mathbb{E}[Y_n])^2] = \mathbb{E}[(Y_n - 500)^2] = \mathbb{E}[Y_n^2 - 1000Y_n + 250000]$





Attempt Three: $\mathbb{E}[(Y_n - \mathbb{E}[Y_n])^2]$

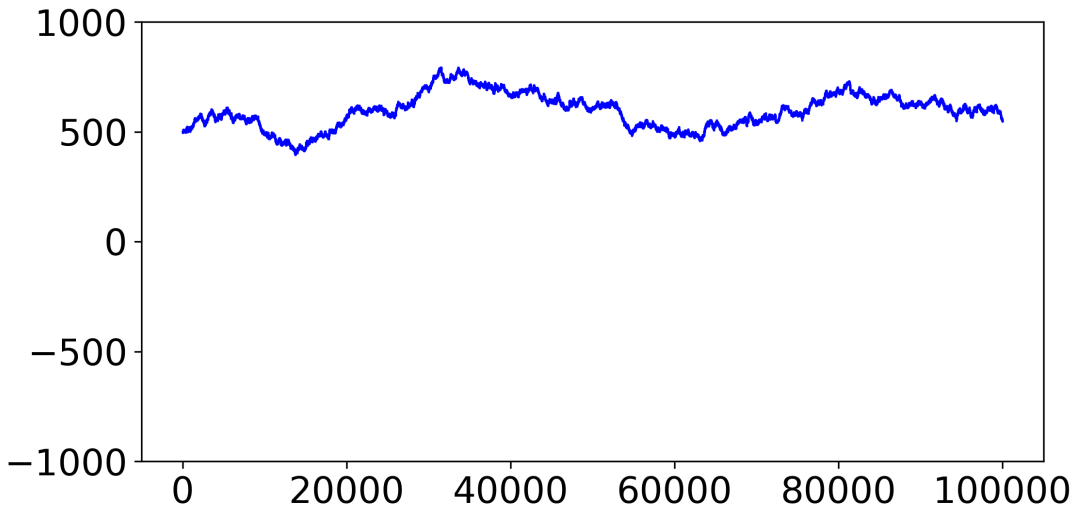
An alternate way to convert the distance into a positive value is to square it.

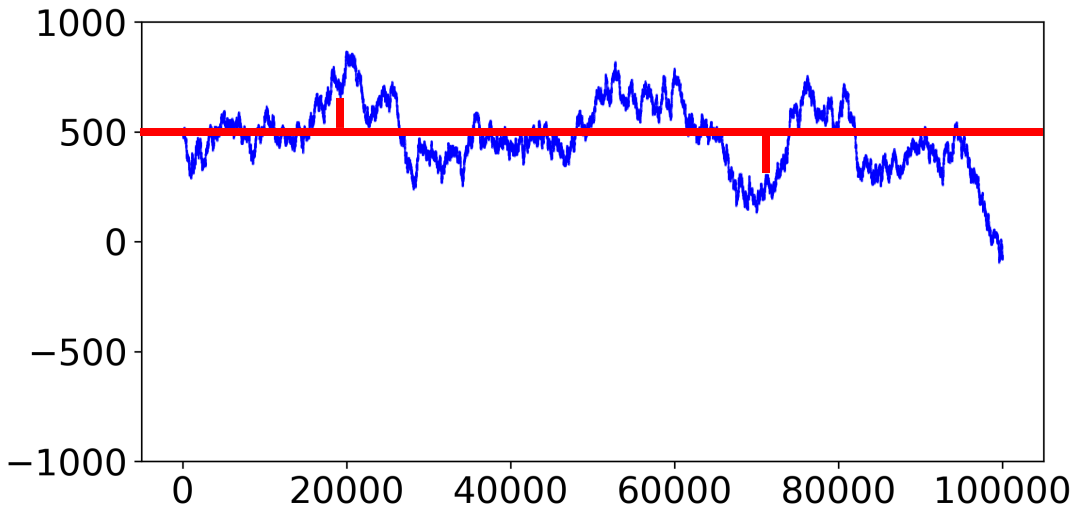
- That is, let's compute the expected squared difference between Y_n and $\mathbb{E}[Y_n]$.

Using linearity of expectation, we can try to compute this quantity.

- $$\mathbb{E}[(Y_n - \mathbb{E}[Y_n])^2] = \mathbb{E}[(Y_n - 500)^2] = \mathbb{E}[Y_n^2 - 1000Y_n + 250000]$$
$$= \mathbb{E}[Y_n^2] - 1000\mathbb{E}[Y_n] + 250000$$
$$= \mathbb{E}[Y_n^2] - 500000 + 250000 = \mathbb{E}[Y_n^2] - 250000$$

Need to compute this somehow.





CC BY NC SA

Attempt Three: $\mathbb{E}[(Y_n - \mathbb{E}[Y_n])^2]$

Using linearity of expectation, compute expected squared difference from mean.

- $\mathbb{E}[(Y_n - \mathbb{E}[Y_n])^2] = \mathbb{E}[Y_n^2] - 250000$

To compute $\mathbb{E}[Y_n^2]$, we can write $\mathbb{E}[Y_n^2] =$

CC BY NC SA

Attempt Three: $\mathbb{E}[(Y_n - \mathbb{E}[Y_n])^2]$

Using linearity of expectation, compute expected squared difference from mean.

- $\mathbb{E}[(Y_n - \mathbb{E}[Y_n])^2] = \mathbb{E}[Y_n^2] - 250000$

To compute $\mathbb{E}[Y_n^2]$, we can write $\mathbb{E}[Y_n^2] = \mathbb{E}[(500 + M_1 + M_2 + \cdots + M_n)^2]$

$$= \mathbb{E}[500^2 + M_1^2 + M_2^2 + \cdots + M_n^2$$

$$+1000M_1 + 1000M_2 + \cdots + 1000M_n$$

$$+2M_1M_2 + \cdots + 2M_{n-1}M_n]$$

Need to compute these somehow.

CC BY NC SA

Math Interlude

What are $\mathbb{E}[M_1^2]$, $E[M_1]$ and $\mathbb{E}[M_1 M_2]$?

- joshh.ug/C



$$P(\mathrm{M_i} = a) = \begin{cases} 1/2, & \text{if } a = +\mathbf{3} \\ 1/2, & \text{if } a = -\mathbf{3} \end{cases}$$

Math Interlude

What are $\mathbb{E}[M_1^2]$, $E[M_1]$ and $\mathbb{E}[M_1 M_2]$?

- M_1 is always 3 or -3, so M_1^2 is always 9. Thus, $\mathbb{E}[M_1^2] = 9$.
- M_1 is -3 or 3 with probability ½, so $\mathbb{E}[M_1] = -3 \times \frac{1}{2} + 3 \times \frac{1}{2} = 0$.
- $M_1 M_2$ can be:

  $3 \times 3 = 9$ with probability ¼

  $3 \times -3 = -9$ with probability ¼

  $-3 \times 3 = -9$ with probability ¼

  $-3 \times -3 = 9$ with probability ¼

Thus $\mathbb{E}[M_1 M_2] = 9 \times \frac{1}{4} - 9 \times \frac{1}{4} - 9 \times \frac{1}{4} + 9 \times \frac{1}{4} = 0$.

$$P(\mathrm{M_i} = a) = \begin{cases} 1/2, & \text{if } a = +3 \\ 1/2, & \text{if } a = -3 \end{cases}$$

Attempt Three: $\mathbb{E}[(Y_n - \mathbb{E}[Y_n])^2]$

Using linearity of expectation, we can try to compute this quantity.

- $\mathbb{E}[(Y_n - \mathbb{E}[Y_n])^2] = \mathbb{E}[Y_n^2] - 250000$

To compute $\mathbb{E}[Y_n^2]$, we can write $\mathbb{E}[Y_n^2] = \mathbb{E}[(500 + M_1 + M_2 + \cdots + M_n)^2]$

$$\mathbb{E}[Y_n^2] = \mathbb{E}[500^2 + M_1^2 + M_2^2 + \cdots + M_n^2] \quad E[M_i^2] = 9$$

$$+1000M_1 + 1000M_2 + \cdots + 1000M_n \quad E[M_i] = 0$$

$$+2M_1M_2 + \cdots + 2M_{n-1}M_n] \quad E[M_iM_j] = 0$$

$$= 250000 + 9n$$

CC BY NC SA

Attempt Three: $\mathbb{E}[(Y_n - \mathbb{E}[Y_n])^2]$

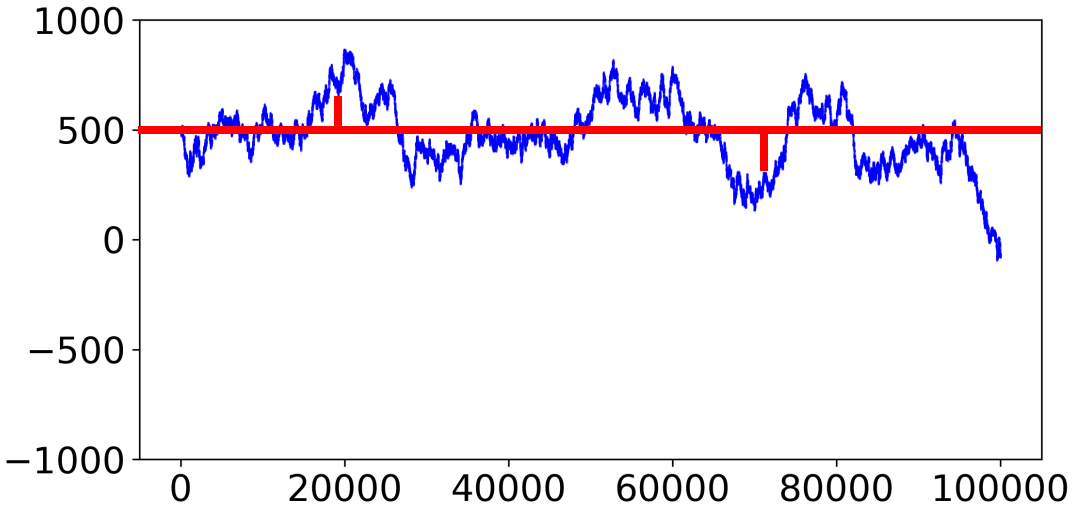
An alternate way to convert the distance into a positive value is to square it.

- That is, let's compute the expected squared difference between Y_n and $\mathbb{E}[Y_n]$.

Using linearity of expectation, we can try to compute this quantity.

- $$\begin{aligned}\mathbb{E}[(Y_n - \mathbb{E}[Y_n])^2] &= \mathbb{E}[(Y_n - 500)^2] = \mathbb{E}[Y_n^2 - 1000Y_n + 250000] \\ &= \mathbb{E}[Y_n^2] - 1000\mathbb{E}[Y_n] + 250000 \\ &= \mathbb{E}[Y_n^2] - 500000 + 250000 = \mathbb{E}[Y_n^2] - 250000 \\ &= 250000 + 9n - 250000 \\ &= 9n\end{aligned}$$

Took some deeper thought. This is $250000 + 9n$



Attempt Three: $\mathbb{E}[(Y_n - \mathbb{E}[Y_n])^2]$

An alternate way to convert the distance into a positive value is to square it.

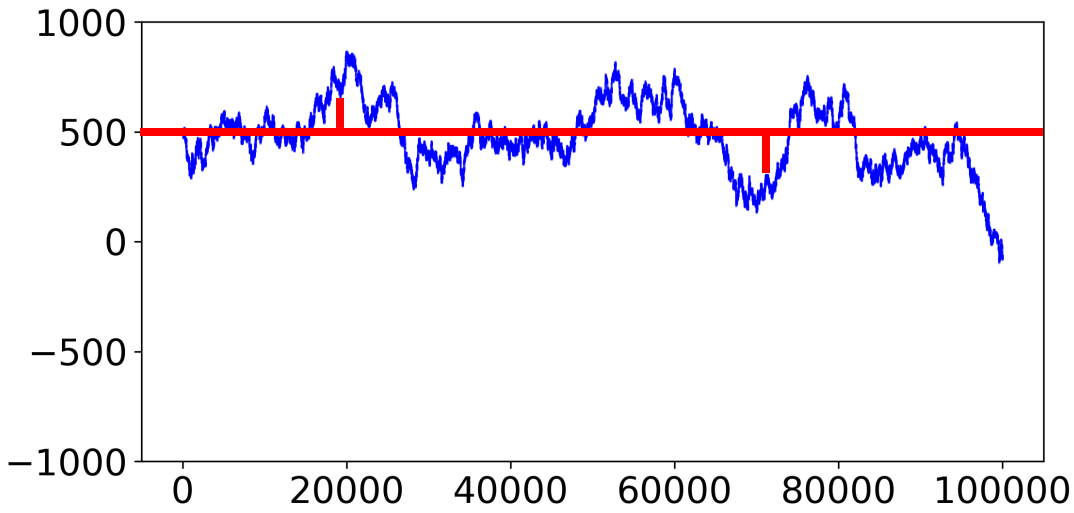
- That is, let's compute the expected squared difference between Y_n and $\mathbb{E}[Y_n]$.

Using linearity of expectation, we can try to compute this quantity.

- $\mathbb{E}[(Y_n - \mathbb{E}[Y_n])^2] = 9n$

That is, on average, we can expect that by time step n, the squared distance of the particle from 500 will be around $9n$.

- We called this the "expected squared difference from the mean".
- This quantity is commonly known as the "variance" of the random variable.



Variance

The **variance** of a random variable X is defined as:

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

Note, we sometimes use $\mu = \mathbb{E}[X]$ to avoid messy nested brackets, i.e.

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

CC BY NC SA

Properties of Variance

Lecture 21, CS70 Fall 2026

Coupon Collector and Expectation

Variance

- Law of the Unconscious Statistician
- Derivation and Definition
- **Properties of Variance**
- Example Variance Calculations
- Sums of Random Variables

CC BY NC SA

Variance Property 1

Variance is defined as $\mathbf{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$

Useful Property 1: Can write variance as: $\mathbf{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

CC BY NC SA

Variance Property 1

Variance is defined as $\mathbf{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$

Useful Property 1: Can write variance as: $\mathbf{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

Proof: Let $\mu = \mathbb{E}[X]$, then

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

CC BY NC SA

Variance Property 2

Suppose we have a random variable X whose variance is $\mathbf{Var}(X)$.

- Then $\mathbf{Var}(cX) = c^2\mathbf{Var}(X)$.

Proof: $\mathbf{Var}(cX) = \mathbb{E}[(cX)^2] - \mathbb{E}[cX]^2$ (by useful property 1)

CC BY NC SA

Variance Property 2

Suppose we have a random variable X whose variance is $\mathbf{Var}(X)$.

- Then $\mathbf{Var}(cX) = c^2\mathbf{Var}(X)$.

Proof: $\mathbf{Var}(cX) = \mathbb{E}[(cX)^2] - \mathbb{E}[cX]^2$ (by useful property 1)

$$= \mathbb{E}[(cX)^2] - (\mathbb{E}[cX])^2$$

$$= \mathbb{E}[(cX)^2] - (c\mathbb{E}[X])^2$$

$$= \mathbb{E}[(cX)^2] - c^2\mathbb{E}[X]^2$$

$$= \mathbb{E}[c^2X^2] - c^2\mathbb{E}[X]^2$$

$$= c^2\mathbb{E}[X^2] - c^2\mathbb{E}[X]^2$$

$$= c^2\mathbf{Var}(X)$$

cc BY NC SA

Variance Property 3

Suppose we have a random variable X whose variance is $\mathbf{Var}(X)$.

- Then $\mathbf{Var}(c + X) = \mathbf{Var}(X)$.

CC BY NC SA

Variance Property 3

Suppose we have a random variable X whose variance is $\mathbf{Var}(X)$.

- Then $\mathbf{Var}(c+X) = \mathbf{Var}(X)$.

Proof: $\mathbf{Var}(c+X) = \mathbb{E}[(c+X-\mathbb{E}[c+X])^2]$ (by basic definition)

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

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

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

$$= \mathbf{Var}(X)$$

Recall basic definition: $\mathbf{Var}(X) = \mathbb{E}[(X-\mathbb{E}[X])^2]$

CC BY NC SA

Variance and Standard Deviation

Variance is not in the same units as the original random variable.

- Example: If X is the score for the midterms, then $\mathbf{Var}(X)$ is given in points squared.

The most natural adjustment is just to compute the square root of the variance. This is called the standard deviation (stdev for short).

- Variance of 2007.1 points squared: less intuitive than stdev of 44.8 points.

Often use σ as symbol for stdev.

- $\sigma_X = \sqrt{\mathbf{Var}(X)}$

Mean	Variance
106.75	870.25

CC BY NC SA

Useful Properties So Far

Basic definition of variance: $\mathbf{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$

Useful property 1: Can rewrite as $\mathbf{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

Useful property 2: $\mathbf{Var}(cX) = c^2\mathbf{Var}(X)$

Useful property 3: $\mathbf{Var}(c + X) = \mathbf{Var}(X)$

Units of variance are squared units of X. Can use square root (standard deviation) of variance instead.

CC BY NC SA

Example Variance Calculations

Lecture 21, CS70 Fall 2026

Coupon Collector and Expectation

Variance

- Law of the Unconscious Statistician
- Derivation and Definition
- Properties of Variance
- **Example Variance Calculations**
- Sums of Random Variables

CC BY NC SA

Computing the Variance of a Random Variable

When computing the variance of a random variable X, we need two quantities:

- The expectation $\mathbb{E}[X]$.
- The expectation of X^2, i.e. $\mathbb{E}[X^2]$

Often, but not always, $\mathbb{E}[X^2]$ requires some cleverness to compute.

CC BY NC SA

Example 1: Bernoulli RV

Suppose $X \sim \text{Bernoulli}(p)$. Recall that $\mathbb{E}[X] = p$.

CC BY NC SA

Example 1: Bernoulli RV

Suppose $X \sim \mathbf{Bernoulli}(p)$. Recall that $\mathbb{E}[X] = p$.

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

- Since X is always 0 or 1, we have that $X^2 = X$, therefore $\mathbb{E}[X^2] = \mathbb{E}[X] = p$.
- The variance is therefore $\mathbf{Var}(X) = p - \mathbb{E}[X]^2 = p - p^2$

CC BY NC SA

Example 1: Bernoulli RV

Suppose $X \sim \mathbf{Bernoulli}(p)$. Recall that $\mathbb{E}[X] = p$.

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

- Since X is always 0 or 1, we have that $X^2 = X$, therefore $\mathbb{E}[X^2] = \mathbb{E}[X] = p$.
- The variance is therefore $\mathbf{Var}(X) = p - \mathbb{E}[X]^2 = p - p^2$

Often this is written in the form $p(1-p)$

Observation, if we pick $p = 1$ (always heads), then the variance is 0.

- This makes sense! If the RV isn't actually random, there should be no variance.
- Similarly, if p is close to 0 or 1, the variance is low, usually get same flip.

CC BY NC SA

Example 2: Rolling a Fair Six Sided Die

Recall that if X is the result of rolling a fair six-sided die, we have that $\mathbb{E}[X] = 7/2$.

To compute variance, we just need $\mathbb{E}[X^2]$. This is straightforward:

CC BY NC SA

Example 2: Rolling a Fair Six Sided Die

Recall that if X is the result of rolling a fair six-sided die, we have that $\mathbb{E}[X] = 7/2$.

To compute variance, we just need $\mathbb{E}[X^2]$. This is straightforward:

$$\begin{aligned}\mathbb{E}[X^2] &= \frac{1}{6}\times 1^2 + \frac{1}{6}\times 2^2 + \frac{1}{6}\times 3^2 + \frac{1}{6}\times 4^2 + \frac{1}{6}\times 5^2 + \frac{1}{6}\times 6^2 \\ &= \frac{1}{6}\times (1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) \\ &= \frac{91}{6}\end{aligned}$$

CC BY NC SA

Example 2: Rolling a Fair Six Sided Die

Recall that if X is the result of rolling a fair six-sided die, we have that $\mathbb{E}[X] = 7/2$. We also just computed $\mathbb{E}[X^2] = 91/6$

The variance of a six sided die roll is given by $\mathbf{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

$$= \frac{91}{6} - \left(\frac{7}{2}\right)^2$$

$$= \frac{35}{12} = 2.92$$

2.92 will come up again later today.

CC BY NC SA

Example 3: Fixed Points of Permutation

Let X_n be the number of students who receive their own homework if we take homeworks from every student and hand them back randomly.

- Last time we showed that $\mathbb{E}[X_n] = 1$.
- How dispersed is this? Do we expect to sometimes get $X_n = 10$, $X_n = 20$?

As before let $X_n = I_1 + I_2 + \cdots + I_n$, where I_i means student i got their homework back.

To compute variance, we already have $\mathbb{E}[X_n]$. We just need $\mathbb{E}[X_n^2]$.

CC BY NC SA

Example 3: Fixed Points of Permutation

$X_n = I_1 + I_2 + \cdots + I_n$, where I_i means student i got their homework back.

We need $\mathbb{E}[X_n^2]$. To get this, we compute $\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2]$. Using the same formula from before, we have that:

$$\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2] = \sum_{i}^{n} \mathbb{E}[I_i^2] + 2\sum_{i<j} \mathbb{E}[I_i I_j]$$

What is $\mathbb{E}[I_i^2]$?

- joshh.ug/D



CC BY NC SA

Example 3: Fixed Points of Permutation

$X_n = I_1 + I_2 + \cdots + I_n$, where I_i means student i got their homework back.

We need $\mathbb{E}[X_n^2]$. To get this, we compute $\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2]$. Using the same formula from before, we have that:

$$\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2] = \sum_{i}^{n} \mathbb{E}[I_i^2] + 2 \sum_{i<j} \mathbb{E}[I_i I_j]$$

Since I_i is just a Bernoulli indicator variable $I_i^2 = I_i$. Since the chance a student gets back their homework is $1/n$, $I_i \sim \text{Bernoulli}(1/n)$. Together these facts give us:

$$\mathbb{E}[I_i^2] = \mathbb{E}[I_i] = 1/n$$

CC BY NC SA

Example 3: Fixed Points of Permutation

$X_n = I_1 + I_2 + \cdots + I_n$, where I_i means student i got their homework back.

We need $\mathbb{E}[X_n^2]$. To get this, we compute $\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2]$. Using the same formula from before, we have that:

$$\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2] = \sum_i^n \frac{1}{n} + 2\sum_{i<j} \mathbb{E}[I_i I_j]$$

Since $\sum_i^n \frac{1}{n} = 1$, we can simplify by replacing $\sum_i^n \frac{1}{n}$ by 1.

CC BY NC SA

Example 3: Fixed Points of Permutation

$X_n = I_1 + I_2 + \cdots + I_n$, where I_i means student i got their homework back.

We need $\mathbb{E}[X_n^2]$. To get this, we compute $\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2]$. Using the same formula from before, we have that:

$$\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2] = 1 + 2\sum_{i<j} \mathbb{E}[I_i I_j]$$

Since $\sum_i^n \frac{1}{n} = 1$, we can simplify by replacing $\sum_i^n \frac{1}{n}$ by 1.

- Now we just need to know $2\sum_{i<j} \mathbb{E}[I_i I_j]$

cc BY NC SA

Example 3: Fixed Points of Permutation

$X_n = I_1 + I_2 + \cdots + I_n$, where I_i means student i got their homework back.

We need $\mathbb{E}[X_n^2]$. To get this, we compute $\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2]$. Using the same formula from before, we have that:

$$\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2] = 1 + 2\sum_{i<j} \mathbb{E}[I_i I_j]$$

What is $\mathbb{E}[I_i I_j]$?

- joshh.ug/E

We'll take a 4 minute break. Feel free to answer.



CC BY NC SA

Example 3: Fixed Points of Permutation

$X_n = I_1 + I_2 + \cdots + I_n$, where I_i means student i got their homework back.

We need $\mathbb{E}[X_n^2]$. To get this, we compute $\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2]$. Using the same formula from before, we have that:

$$\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2] = 1 + 2\sum_{i<j} \mathbb{E}[I_i I_j]$$

$$\mathbb{E}[I_i I_j] = 0\times 0\times P(I_i = 0,\ I_j = 0) + 0\times 1\times P(I_i = 0,\ I_j = 1)$$
$$+1\times 0\times P(I_i = 1,\ I_j = 0) + 1\times 1\times P(I_i = 1,\ I_j = 1)$$

But only one of these terms is not trivially zero: $1\times 1\times P(I_i = 1,\ I_j = 1)$

CC BY NC SA

Example 3: Fixed Points of Permutation

$X_n = I_1 + I_2 + \cdots + I_n$, where I_i means student i got their homework back.

We need $\mathbb{E}[X_n^2]$. To get this, we compute $\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2]$. Using the same formula from before, we have that:

$$\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2] = 1 + 2\sum_{i<j} \mathbb{E}[I_i I_j]$$

Since both I_i and I_j are indicators, only one term is not trivially zero:

$$\begin{aligned}
\mathbb{E}[I_i I_j] &= P(I_i I_j = 1) \\
&= P(i \text{ and } j \text{ get their homework back}) \\
&= \frac{1}{n(n-1)}
\end{aligned}$$

Example 3: Fixed Points of Permutation

$X_n = I_1 + I_2 + \cdots + I_n$, where I_i means student i got their homework back.

We need $\mathbb{E}[X_n^2]$. To get this, we compute $\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2]$. Using the same formula from before, we have that:

$$\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2] = 1 + 2\sum_{i<j} \frac{1}{n(n-1)}$$

Since both I_i and I_j are indicators, only one term is not trivially zero:

$$\begin{aligned}
\mathbb{E}[I_i I_j] &= P(I_i I_j = 1) \\
&= P(i \text{ and } j \text{ get their homework back}) \\
&= \frac{1}{n(n-1)}
\end{aligned}$$

Example 3: Fixed Points of Permutation

$X_n = I_1 + I_2 + \cdots + I_n$, where I_i means student i got their homework back.

We need $\mathbb{E}[X_n^2]$. To get this, we compute $\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2]$. Using the same formula from before, we have that:

$$\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2] = 1 + 2\sum_{i<j} \frac{1}{n(n-1)}$$

What is $2\sum_{i<j} \frac{1}{n(n-1)}$? There are $\binom{n}{2}$ ways that i is less than j, so

$$2\sum_{i<j} \frac{1}{n(n-1)} = 2\binom{n}{2} \times \frac{1}{n(n-1)}$$

$$= 2\frac{n(n-1)}{2} \times \frac{1}{n(n-1)} \quad = 1$$

CC BY NC SA

Example 3: Fixed Points of Permutation

$X_n = I_1 + I_2 + \cdots + I_n$, where I_i means student i got their homework back.

We need $\mathbb{E}[X_n^2]$. To get this, we compute $\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2]$. Using the same formula from before, we have that:

$$\mathbb{E}[X_n^2] = \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2] = 1 + 1 = 2$$

After all that work, we've learned that $\mathbb{E}[X_n^2] = 2$

CC BY NC SA

Example 3: Fixed Points of Permutation

Let X_n be the number of students who receive their own homework if we take homeworks from every student and hand them back randomly.

- Last time we showed that $\mathbb{E}[X_n] = 1$.
- How dispersed is this? Do we expect to sometimes get $X_n = 10$, $X_n = 20$?

To compute variance, we know $\mathbb{E}[X_n] = 1$. We also know $\mathbb{E}[X_n^2] = 2$.

What is $\mathbf{Var}(X_n)$?

- joshh.ug/F

Example 3: Fixed Points of Permutation

Let X_n be the number of students who receive their own homework if we take homeworks from every student and hand them back randomly.

- Last time we showed that $\mathbb{E}[X_n] = 1$.
- How dispersed is this? Do we expect to sometimes get $X_n = 10$, $X_n = 20$?

To compute variance, we know $\mathbb{E}[X_n] = 1$. We just need $\mathbb{E}[X_n^2] = 2$.

$$\mathrm{Var}(X_n) = \mathbb{E}[X_n^2] - \mu^2$$

$$= 2 - 1$$

$$= 1$$

Note: If you weren't sure it was 1 or -1, keep in mind variance must be positive.

CC BY NC SA

Example 3: Fixed Points of Permutation

Let X_n be the number of students who receive their own homework if we take homeworks from every student and hand them back randomly.

- Last time we showed that $\mathbb{E}[X_n] = 1$.
- And we also know that $\mathbf{Var}(X_n) = 1$.

How dispersed is this? Do we expect to sometimes get $X_n = 10$, $X_n = 20$?

- We don't expect to get a value near 10, no matter how big n gets.
- On average, the number of students who get their homework back is 1.
- And the variance (and standard deviation) is 1.
- If we were to run a bunch of experiments, the usual deviation between our observation and the mean (1) would be around 1.

CC BY NC SA

Sums of Independent Random Variables

Lecture 21, CS70 Fall 2026

Coupon Collector and Expectation

Variance

- Law of the Unconscious Statistician
- Derivation and Definition
- Properties of Variance
- Example Variance Calculations
- **Sums of Random Variables**

CC BY NC SA

Sums of Independent Random Variables

In a previous lecture we already saw that if we have independent random variables $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\mu)$, and $\text{S} = X + Y$, then $S \sim \text{Poisson}(\lambda + \mu)$.

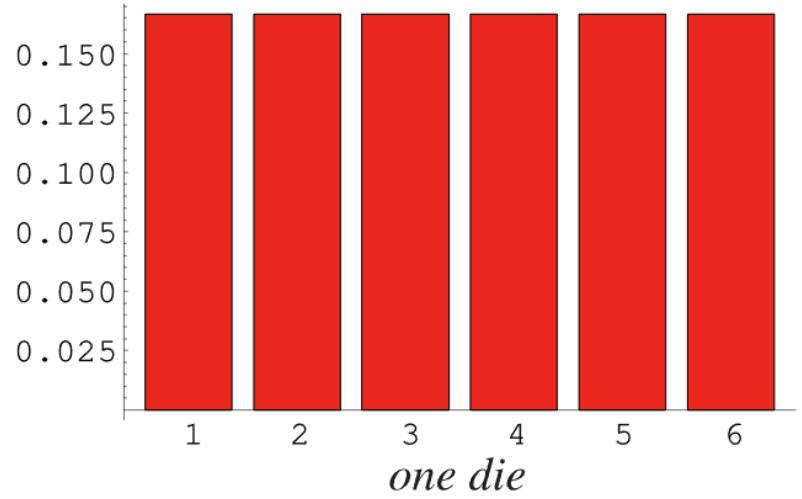
In general, summing random variables is a complicated business.

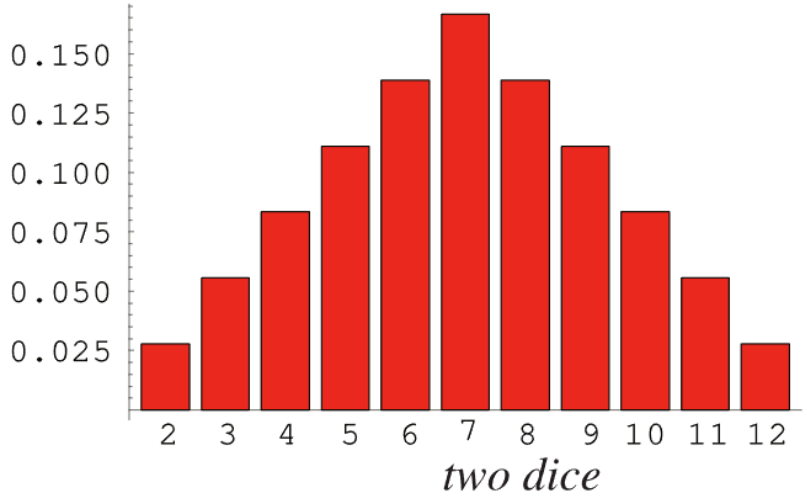
- It's a notable property that two Poisson RVs sum to be Poisson as well.
- Not (always) true for the sum of Bernoulli, Binomial, Geometric.

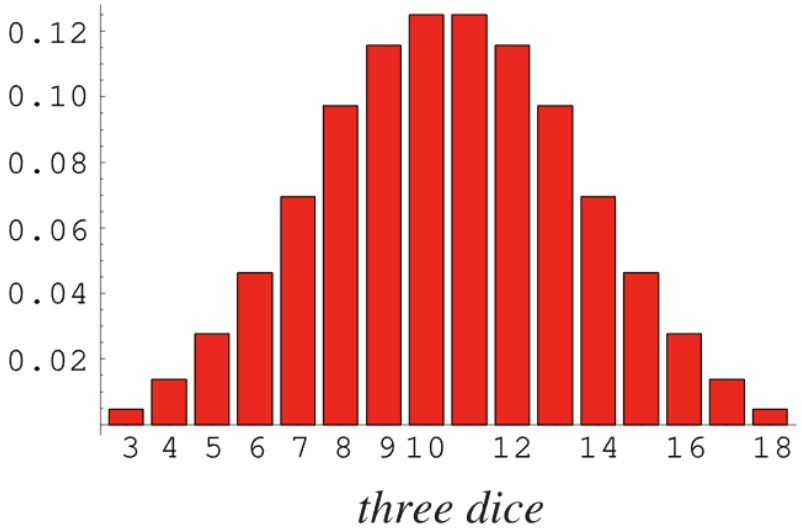
As an example, consider summing n rolls of fair six-sided dice (next slide).

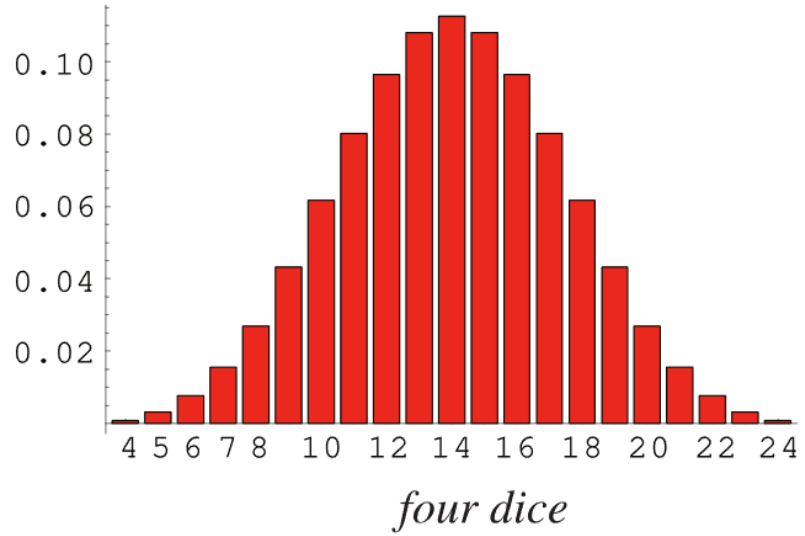
CC BY NC SA

From https://mathworld.wolfram.com/Dice.html



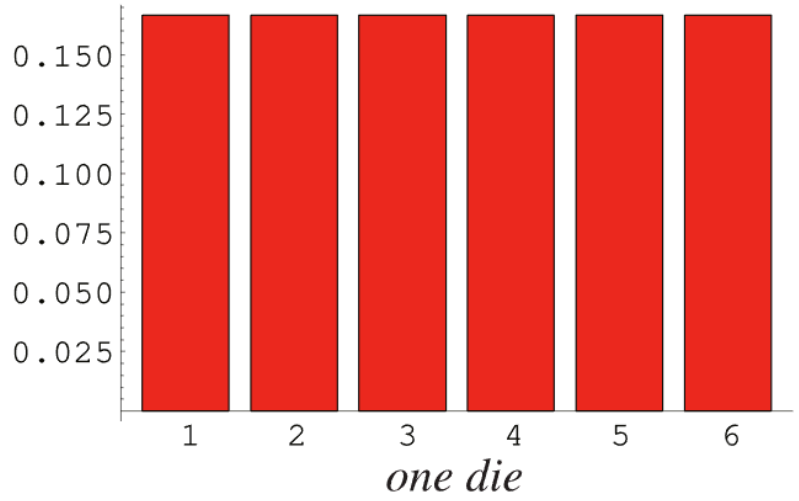


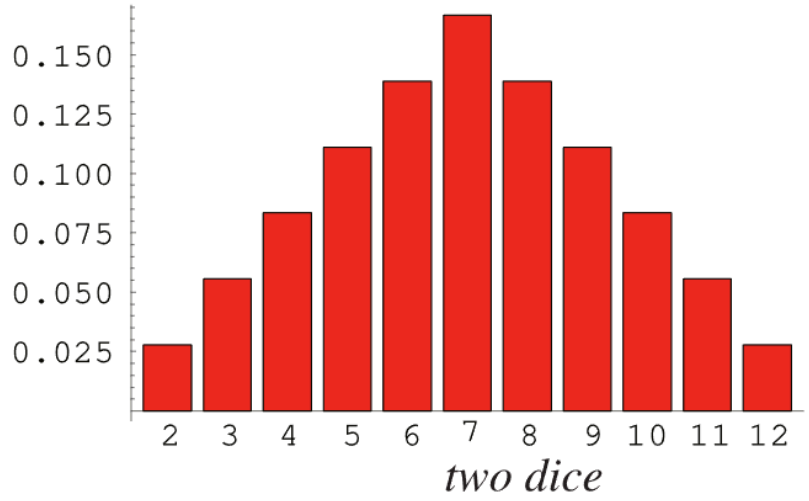


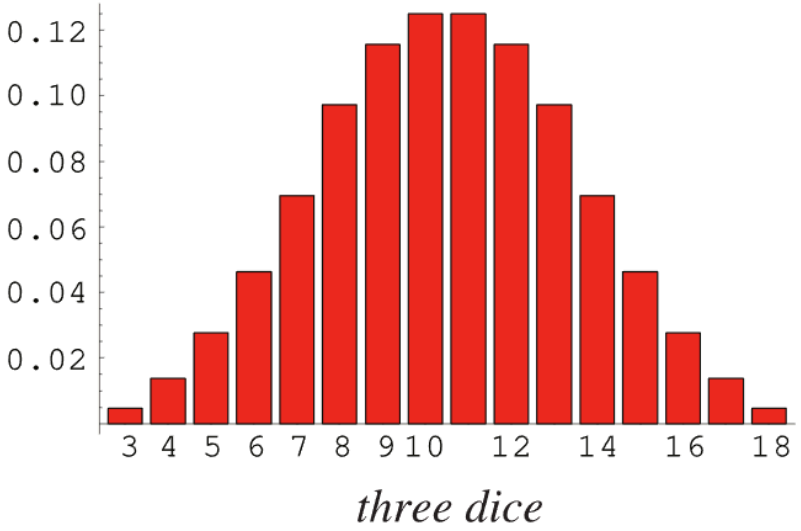


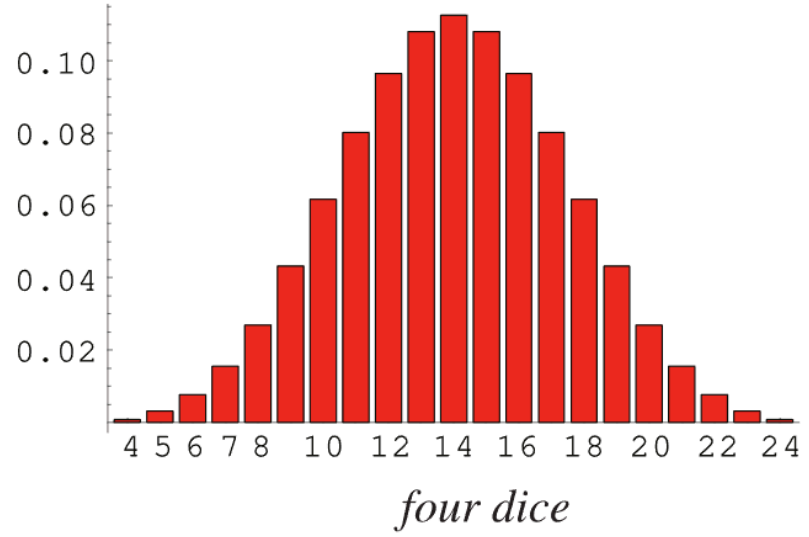
Note: The operation that is happening here is called “convolution”. We won’t discuss today.

Expected Value and Standard Deviation vs. N



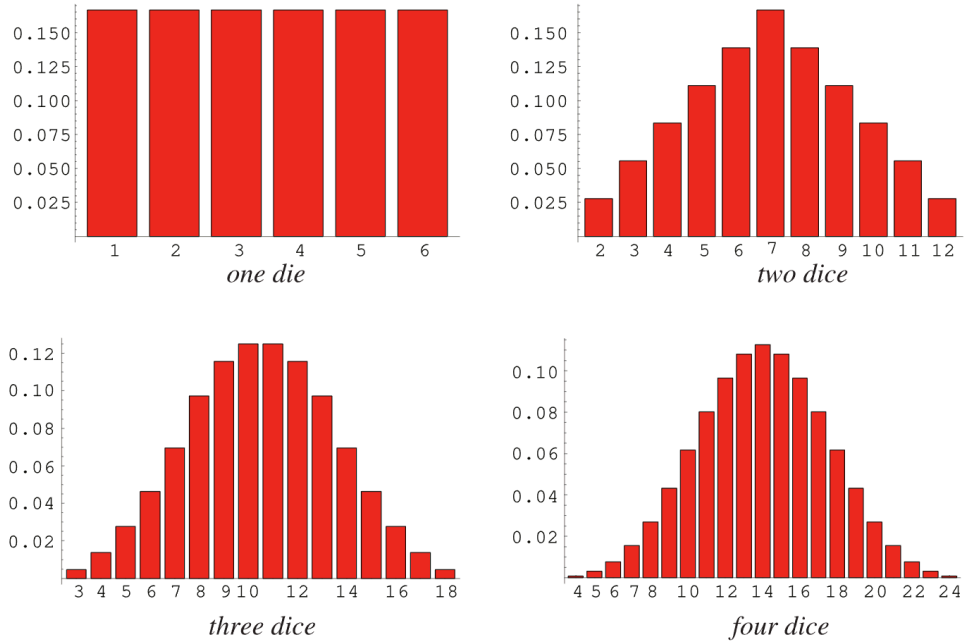






n	$E[X]$	$\sigma(X)$
1	3.5	1.71
2	7	2.41
3	10.5	2.96
4	14	3.42
5	17.5	3.82

Expected Value and Standard Deviation vs. N



n	$E[X]$	$\sigma(X)$
1	3.5	1.71
2	7	2.41
3	10.5	2.96
4	14	3.42
5	17.5	3.82

Two interesting observations as n, the number of random variables summed together, increases:

- The mean increases (linearly).
- The distribution concentrates around the mean, i.e. the standard deviation grows relatively slowly.
- In fact, $\sigma(n) \approx 1.7\sqrt{n}$, or equivalently $\mathrm{Var}(n) \approx 2.9n$

Expectation and Sums of Independent Random Variables

While we can't concisely give the distribution of the sum of two independent random variables, we can reason about the expectation and variance of the sum.

We already know that $\mathbb{E}[X + Y] = \mathbb{E}[X] + E[Y]$.

- Example: Flipping 100 fair coins yields an average of 50 heads.

But what about their variance?

- Example: We saw empirically with dice that rolling 4 dice gives 4 times the variance (or 2 times the standard deviation).
- Let's investigate further.

CC BY NC SA

Expectation and Sums of Independent Random Variables

Claim: If X and Y are independent, then $\mathbf{Var}(X+Y) = \mathbf{Var}(X) + \mathbf{Var}(Y)$.

- We need to consider $\mathbf{Var}(X+Y) = \mathbb{E}[(X+Y)^2] - (\mathbb{E}[X+Y])^2$

$$\mathbb{E}[(X+Y)^2] = \mathbb{E}[X^2 + 2XY + Y^2]$$
$$= \mathbb{E}[X^2] + 2\mathbb{E}[XY] + \mathbb{E}[Y^2]$$

$$(\mathbb{E}[X+Y])^2 = (\mathbb{E}[X] + E[Y])^2$$
$$= \mathbb{E}[X]^2 + 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[Y]^2$$

With these two quantities expanded, we can write:

$$\mathbf{Var}(X+Y) = \mathbb{E}[(X+Y)^2] - (\mathbb{E}[X+Y])^2$$
$$= \mathbb{E}[X^2] - \mathbb{E}[X]^2 + 2\mathbb{E}[XY] - 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[Y^2] - \mathbb{E}[Y]^2$$
$$= \mathbf{Var}(X) + \underbrace{2\mathbb{E}[XY] - 2\mathbb{E}[X]\mathbb{E}[Y]} + \mathbf{Var}(Y)$$

To deal with this, let's take a detour.

CC BY NC SA

Expectation and Sums of Independent Random Variables

Lemma: If X and Y are independent, then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

Proof:

$$\mathbb{E}[XY] = \sum_a \sum_b ab \times P(X = a, Y = b)$$

$$= \sum_a \sum_b ab \times P(X = a) \times P(Y = b)$$

$$= \left(\sum_a a \times P(X = a) \right) \times \left(\sum_b b \times P(Y = b) \right)$$

$$= \mathbb{E}[X]\mathbb{E}[Y]$$

Example term for rolling a four sided and six sided die:

$$2 \times 5 \times \mathrm{P}(\mathrm{D}_4 = 2, \mathrm{D}_6 = 5)$$

$$= 2 \times 5 \times \mathrm{P}(\mathrm{D}_4 = 2) \times \mathrm{P}(\mathrm{D}_6 = 5)$$

$$= 2 \times 5 \times \frac{1}{4} \times \frac{1}{6}$$

$$= \left(2 \times \frac{1}{4}\right) \times \left(5 \times \frac{1}{6}\right)$$

CC BY NC SA

Expectation and Sums of Independent Random Variables

Claim: If X and Y are independent, then $\mathbf{Var}(X+Y) = \mathbf{Var}(X) + \mathbf{Var}(Y)$.

- We need to consider $\mathbf{Var}(X+Y) = \mathbb{E}[(X+Y)^2] - (\mathbb{E}[X+Y])^2$

$$\mathbf{Var}(X+Y) = \mathbb{E}[(X+Y)^2] - (\mathbb{E}[X+Y])^2$$

$$= \mathbb{E}[X^2] - \mathbb{E}[X]^2 + 2\mathbb{E}[XY] - 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[Y^2] - \mathbb{E}[Y]^2$$

$$= \mathbf{Var}(X) + \underbrace{2\mathbb{E}[XY] - 2\mathbb{E}[X]\mathbb{E}[Y]}_{\text{We just showed that if } X \text{ and } Y \text{ are independent, then } \mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]} + \mathbf{Var}(Y)$$

$$= \mathbf{Var}(X) + \mathbf{Var}(Y)$$

CC BY NC SA

Summing i.i.d. Random Variables

If we have n random variables X_i that are independent and identically distributed, and $S = X_1 + \cdots + X_n$, then we have that:

- $\mathbb{E}[S] = n\mathbb{E}[X_1]$
- $\mathbf{Var}(S) = n\mathbf{Var}(X_1)$

Example: For rolling n six-sided dice, we have that:

- $\mathbb{E}[S] = 3.5n$
- $\mathbf{Var}(S) = 2.92n$
- $\sigma(S) = 1.7\sqrt{n}$

In the notes, they use the example of flipping n coins (rather than dice).

CC BY NC SA

Summary

Key Discoveries of Slides/Notes 21:

- For the coupon collector problem $X = X_1 + X_2 + \cdots + X_n$ where $X_i \sim \text{Geometric}$
  - $\mathbb{E}[X] \approx n(\ln n + \gamma_E)$
- $\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$
- $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$
- $\sigma(X) = \sqrt{\text{Var}(X)}$
- $\text{Var}(X + Y) = \text{Var}(X) + 2\mathbb{E}[XY] - 2\mathbb{E}[X]\mathbb{E}[Y] + \text{Var}(Y)$
- If X and Y independent:
  - $\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$
  - $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$

In the next set of slides/notes, we'll talk about the variance of two random variables that are not independent.

CC BY NC SA