A. **Expectation and Independent Random Variables.** Suppose we have two random variables X and Y on a probability space and we want to compute the expected value of their product. Unfortunately:

$\mathbb{E}[XY] \neq$ except in special cases

As an easy example why not, let X and Y each be uniform on $\{0, 1\}$. But assume they are highly dependent: X and Y always have opposite values. Then $\mathbb{E}[XY] = 0$ but $\mathbb{E}[X]\mathbb{E}[Y] = \frac{1}{2} \cdot \frac{1}{2}$.

One of those special cases is when the events are **independent**. We will prove this fact later, when we talk about joint distributions, but at the moment let's state it, because it fits nicely into the flow of our discussion.

> **Expectation of a Product of Indpendent Random Variables.** Let X and Y be **independent** random variables on a common probability space. Then:
>
> $\mathbb{E}[XY] =$

For example, if X and Y are independent rolls of a fair 6–sided die, then:

$\mathbb{E}[XY] =$

Remember: if X is uniform on $\{1, \ldots, n\}$, then its expected value is:

$$\mathbb{E}[X] = \frac{n+1}{2}$$

B. **Variance.** Let X be a random variable with **mean** (expected value) $\mathbb{E}[X] = \mu$. The **variance** of X measures the average **distance–squared** between X and the mean:

$\mathrm{Var}(X) =$

> **Variance.** Let X be a random variable with expected value $\mathbb{E}[X] = \mu$. Then the **variance** of X is:
>
> $\mathrm{Var}(X) = \mathbb{E}\left[(X-\mu)^2\right] =$
>
> The **standard deviation** of X is the average **distance** between X and the mean:
>
> $\mathrm{SD}(X) =$

Variance and standard deviation both measure "spread" of the distribution. To model this:

$$X = \begin{cases} 1 \text{ with probability } 1/2 \\ -1 \text{ with probability } 1/2 \end{cases} \implies \mathbb{E}[X] = 0 \text{ and } \mathrm{Var}(X) =$$

$$Y = \begin{cases} 100 \text{ with probability } 1/2 \\ -100 \text{ with probability } 1/2 \end{cases} \implies \mathbb{E}[Y] = 0 \text{ and } \mathrm{Var}(Y) =$$

To find variance, we need $\mathbb{E}[X^2]$, which we call the **second moment**, in other words we need the expected value of a **function** of X, in this case, the squaring function. The following result tells us how to compute the expected value of a function $g(X)$ of X, using the probability mass function of X. It's proof amounts to rearranging terms in a sum.

The **first moment** is the expectation $\mathbb{E}[X]$. The **third moment** is $\mathbb{E}[X^3]$. And the **nth moment** is $\mathbb{E}[X^n]$.

> **Expectation of Function of a Random Variable.** If X is a random variable and $g(X)$ is a function of X, then:
>
> $\mathbb{E}[g(X)] =$

C. **Variance of a Discrete Uniform Random Variable.** Let X be a uniform random variable on $\{1, 2, \ldots, n\}$. Then:

first moment: $\mathbb{E}[X] = \dfrac{n+1}{2}$

second moment: $\mathbb{E}[X^2] =$

We need the sum formula:

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$

variance: $\mathrm{Var}(X) =$

Uniform Variance. If X is uniform on $\{1, 2, \ldots, n\}$ then:

$\mathrm{Var}(X) =$

D. **Variance of a Bernoulli Random Variable.** Let X be a Bernoulli random variable with parameter p. Then:

first moment: $\mathbb{E}[X] = p$

second moment: $\mathbb{E}[X^2] =$

variance: $\mathrm{Var}(X) =$

Bernoulli Variance. If $X \sim \mathrm{Bernoulli}(p)$ then:

$\mathrm{Var}(X) =$

As a special case, if $\mathbb{1}_E$ is the indicator function of an event E then:

$\mathrm{Var}(\mathbb{1}_E) =$

E. **Variance of Sums.** Suppose we have two random variables X and Y on a probability space and we want to compute the variance of their sum. Unfortunately:

$\mathrm{Var}(X+Y) \neq$ except in special cases

In particular, this says variance is **not** linear.

As an easy example why not, let X and Y each be uniform on $\{0, 1\}$. But assume they are highly dependent: X and Y always have opposite values. Then by direct calculation, $\mathrm{Var}(X) + \mathrm{Var}(Y) = \frac{1}{4} + \frac{1}{4}$. But $X + Y$ always equals 1, so does not vary at all: $\mathrm{Var}(X + Y) = 0$.

One of those special cases is when the events are **independent**. Let's verify this: assume X and Y are independent, and let's calculate:

if $X \perp Y$: $\mathrm{Var}(X+Y) =$

Variance of a Sum of Indpendent Random Variables. Let X and Y be **independent** random variables on a common probability space. Then:

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

F. **Expectation of a Binomial Random Variable.** Let X be a binomial random variable, counting the number of success after n independent trials of an experiment, each with success probability p. We had noted that X is a sum of **independent** Bernoulli random variables:

$X =$

and consequently:

$\mathrm{Var}(X) =$

Binomial Expectation. If $X \sim \mathrm{Binom}(n, p)$ then:

$\mathrm{Var}(X) =$

Example 1. Alice is applying to 10 clubs. She thinks she has a $\frac{1}{3}$ chance of getting into each club, independent of all others. Let X be the number of clubs to which Alice is accepted. Find the second moment, $\mathbb{E}[X^2]$.