**1. Theoretical**

**Welford’s Recursion**

Welford’s recursion is a practical and stable method to calculate the running mean and variance of a dataset incrementally. Unlike the standard approach, which recalculates mean and variance from scratch after each new data point, Welford’s method updates these quantities using only a few variables. This makes it suitable for streaming data or large datasets where storing all observations is impractical.

**Variables used:**

* n: number of data points processed so far
* ​μn: running mean after nnn data points
* M2n​: running sum of squared differences from the current mean

**Step-by-step update:**

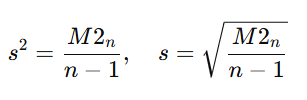
1. **Initialization**



1. **For each new data point Xn ​:**

* Increment the counter: n←n+1
* Compute the difference from the previous mean: δ=Xn−μn−1​
* Update the mean: μn = μn-1 + δ/n
* Adjust the difference: δ2=Xn−μn
* Update the sum of squared differences: M2n=M2n−1+δ⋅ δ2

1. **Compute final variance and standard deviation (for n>1):**

****

**Why this method is effective:**

* **Efficient:** Each data point is processed only once.
* **Memory-friendly:** Only three variables are required, regardless of dataset size.
* **Stable:** Incremental updates reduce rounding errors.
* **Flexible:** Works with any dataset without modification.

In short, Welford’s recursion provides a simple, robust, and elegant way to track running mean and variance.

## ****2. Applications****

**Euler–Maruyama Simulator Refinement**

To better study stochastic processes, I extended the Euler–Maruyama simulator with the following features:

### **A. Random Walk with Jumps**

* Each step can increase or decrease the current state by 1, with probability p for a +1 and 1−p for a −1.
* This models a discrete-time random walk and allows analysis of systems where outcomes can both succeed and fail.

### **B. Absolute and Relative Frequency Trajectories**

* **Absolute frequency (Bernoulli):** Counts total successes, always increasing with each success.
* **Absolute frequency (Random Walk):** Oscillates because outcomes can be positive or negative.
* **Relative frequency:** Normalized by the number of trials.
  + Bernoulli relative frequency approaches ppp over time.
  + Random walk relative frequency stabilizes around 2p−1, the expected drift.

### **C. Final and Intermediate Distributions**

* The simulator can show the distribution of outcomes either at the end or at an intermediate time step selected by the user.
* For each case, it calculates the **mean** and **variance**, making it easy to compare different processes.
* A single interface controls all variants (absolute/relative frequency, Bernoulli/random walk, step number, probability p), allowing for flexible experimentation.

### **Observations on Mean and Variance**

* **Bernoulli model:**
  + Absolute frequency grows steadily; mean increases linearly (np) and variance as np(1−p).
  + Relative frequency converges to p, variance decreases with time.
* **Random walk model:**
  + Absolute frequency fluctuates above and below zero; mean grows as n(2p−1).
  + Relative frequency remains near the expected drift; variance decreases with normalization but is higher than in the Bernoulli case.

**Differences between absolute and relative frequencies:**

* Absolute frequencies are integers and can have a wide range of values (0 to n in Bernoulli, −n to n in random walk).
* Relative frequencies are normalized, real-valued, and more concentrated around the expected value.

**Summary:**

* Bernoulli processes are predictable and steadily growing, both in absolute and relative terms.
* Random walks show oscillations in absolute frequencies but stabilized relative frequencies.
* Variance grows faster in random walks but relative normalization reduces its effect over time.

These observations give a clear understanding of how mean and variance evolve in different stochastic processes and highlight the value of parametric simulation tools for exploring such behaviors.