

## Welford's algorithm in batches

### Recursive Batch Mean

$$N \bar{x}_N - (N-k)\bar{x}_{N-k} = \sum_{i=1}^N x_i - \sum_{i=1}^{N-k} x_i = \sum_{i=N-k+1}^N x_i =$$
$$\bar{x}_N = \frac{\sum_{i=N-k+1}^N x_i + (N-k)\bar{x}_{N-k}}{N} = \bar{x}_{N-k} + \frac{\sum_{i=N-k+1}^N (x_i - \bar{x}_{N-k})}{N}$$

### Recursive Variance

$$\begin{aligned} N\sigma_N^2 - (N-k)\sigma_{N-k}^2 &= \sum_{i=1}^N (x_i - \bar{x}_{N-k})^2 - \sum_{i=1}^{N-k} (x_i - \bar{x}_{N-k})^2 = \\ &= \sum_{i=1}^{N-k} (x_i - \bar{x}_N)^2 - \sum_{i=1}^{N-k} (x_i - \bar{x}_N)^2 - (x_i - \bar{x}_{N-k})^2 \\ &= \sum_{i=1}^{N-k} (x_i - \bar{x}_N)^2 - \sum_{i=1}^{N-k} (x_i + x_i - \bar{x}_{N-k} - \bar{x}_N)(\bar{x}_{N-k} - \bar{x}_N) \\ &= \sum_{i=1}^{N-k} (x_i - \bar{x}_N)^2 - (\bar{x}_{N-k} - \bar{x}_N) \sum_{i=1}^{N-k} (x_i - \bar{x}_N) = \\ &= \sum_{i=1}^{N-k} (x_i - \bar{x}_N)(x_i - \bar{x}_N + \bar{x}_N - \bar{x}_{N-k}) = \sum_{i=1}^{N-k} (x_i - \bar{x}_N)(x_i - \bar{x}_{N-k}) \\ &= \frac{\sum_{i=1}^{N-k} (x_i - \bar{x}_N)(x_i - \bar{x}_{N-k}) + (N-k)\sigma_{N-k}^2}{N} = \\ &= \sigma_{N-k}^2 + \frac{\sum_{i=1}^{N-k} (x_i - \bar{x}_N)(x_i - \bar{x}_{N-k}) + k\sigma_{N-k}^2}{N} \end{aligned}$$