Initial Transients

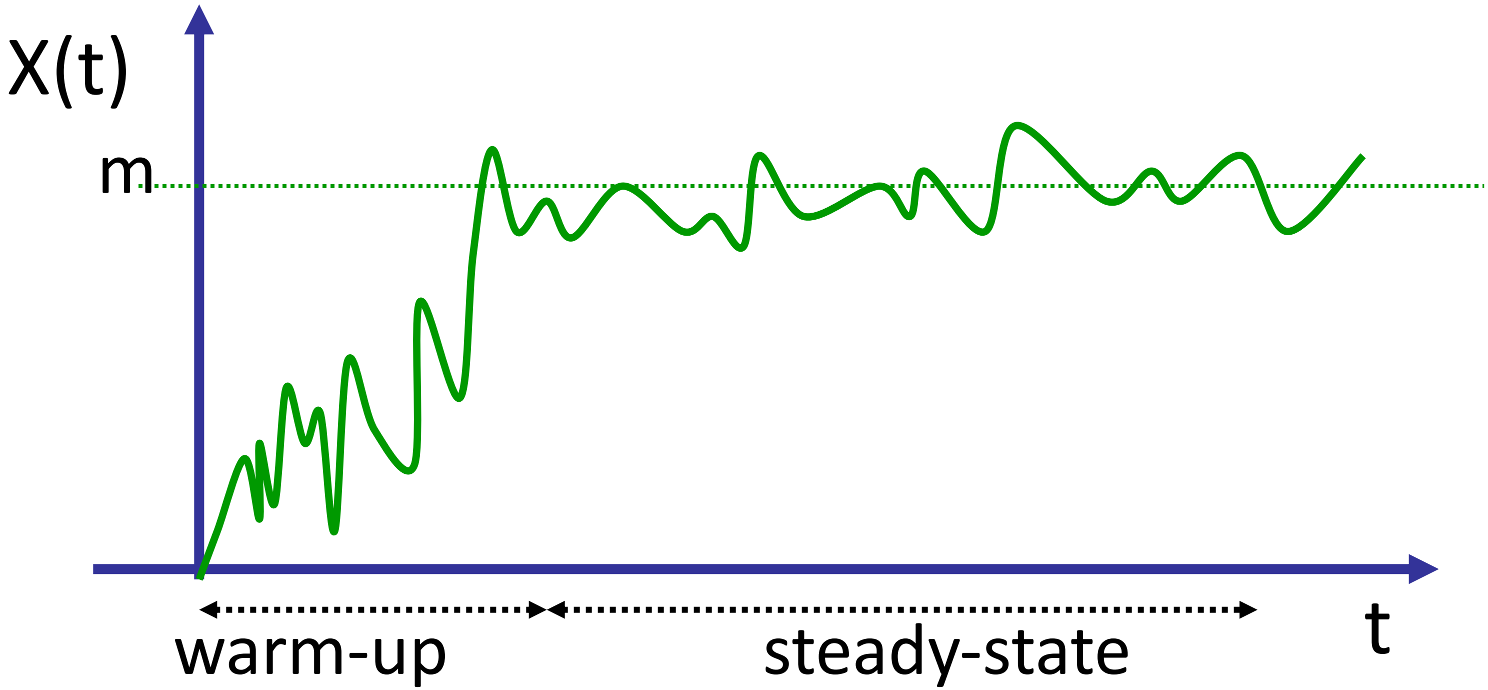
Independent observations

- To estimate the goodness of the results of simulations, confidence intervals are derived
- Confidence intervals theory is based on 2 assumptions
  - the process is stationary
  - the observations x_i are *independent*
- Issues:
  - Identify the stationarity conditions, removing the warm-up transient
  - Long simulation runs, to collect significant samples
  - Several independent simulation runs, to collect independent samples

Warm-up transient

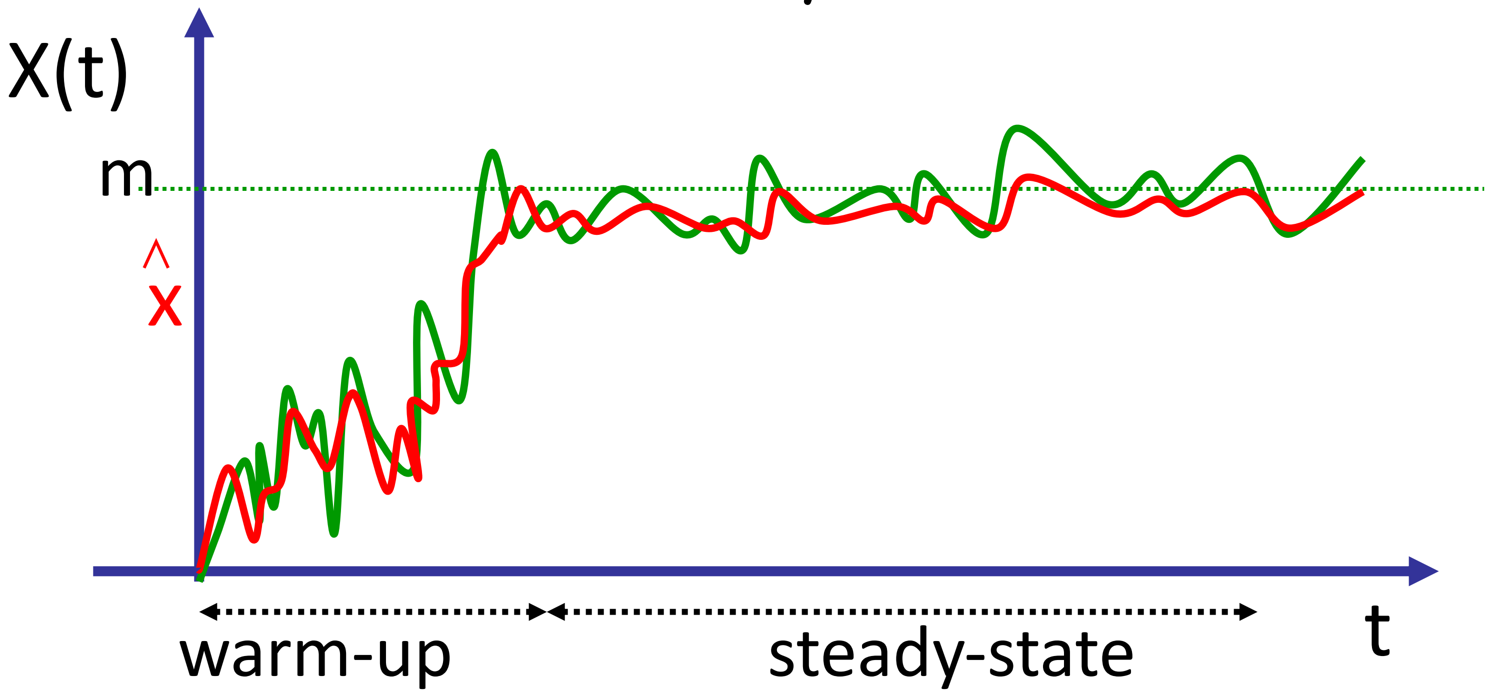
- In a steady-state simulation we suppose that
  - the system is continuously running
  - the system is stable and it is working at its steady-state conditions
  - its behavior does not depend on the initial conditions
- The *warm-up transient* is the time needed for the system to reach its steady-state conditions after starting from a given initial condition
- The warm-up transient must be removed

Warm-up transient



Warm-up transient

The estimated mean value is influenced by the initial transient



Warm-up transient

- It is not easy to identify the warm-up period and to distinguish it from the steady-state
- We use visual inspection and heuristics to identify the warm-up period
- Methods to remove the warm-up transient are usually based on the idea that the variance of the measures during the transient is higher than at steady-state
- The need for transient removal reduces the efficiency of the simulation, since we must throw away part of the collected data

A few simple methods

- Long runs: running a simulation for a long time, the effect of the warm-up phase over the performance measures is reduced
  - Quite costly
- Selection of initial conditions near to the steady-state conditions
  - It might help when previous simulation runs gave us information on the system status at steady-state
  - Sometime difficult to bring the system at the steady-state conditions

Initial data removal

- It is based on the removal of part of the data initially collected, and on the observation of the changes on the average of the remaining data
- The idea is that the removal of samples collected during the warm-up transient changes the average of the remaining data, while removing samples during the steady-state does not influence too much the average
- Given n observations of X ($x_1, x_2, \ldots, x_n$), we cancel the first k observations and we compute the average of the remaining ones

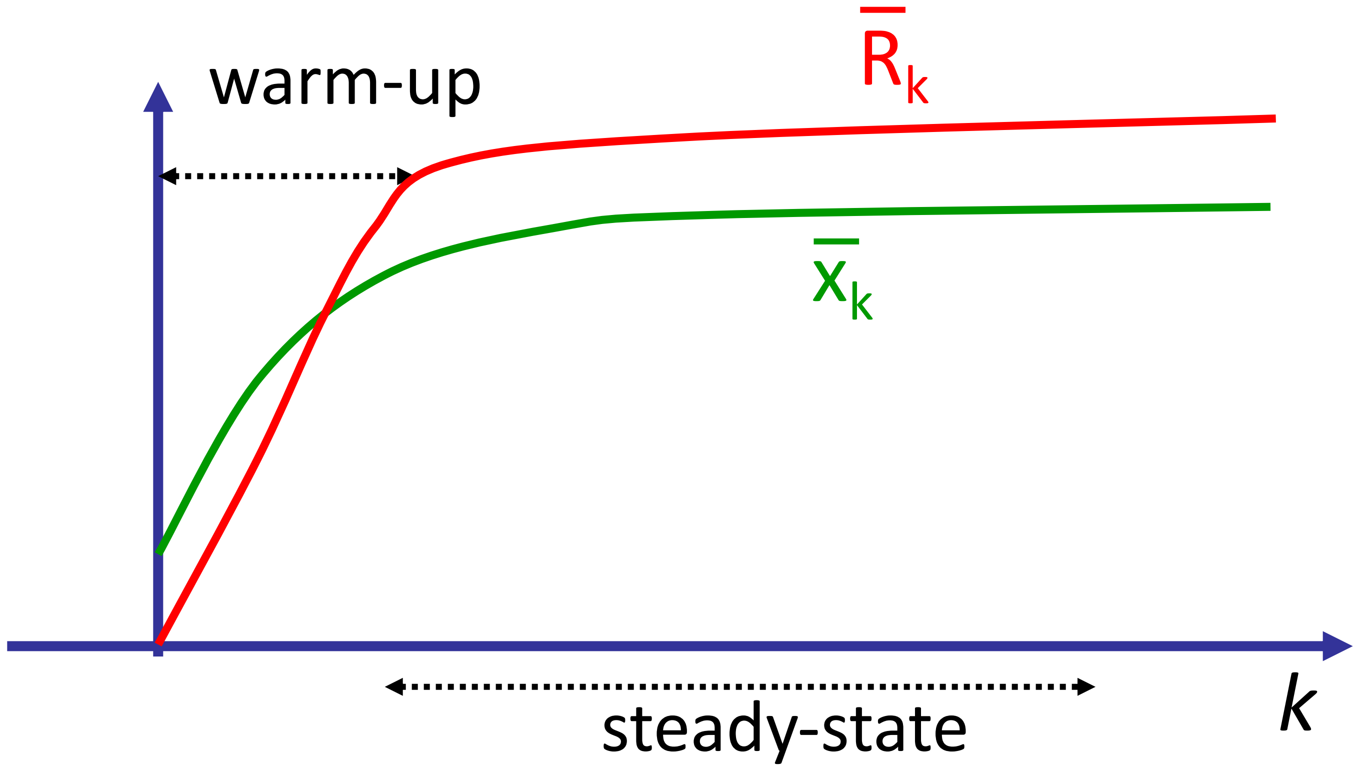
Initial data removal

- We study the variations of the average when the number *k* of cancelled data changes, using a graph of the average as a function of *k*
- After some critical value of *k* the average of the remaining samples starts reducing its changes at further increases of *k:* this critical value identifies the warm-up period
- To reduce the randomness effects, the samples x_i can be obtained averaging corresponding samples of different replications of the same simulation (e.g. with different seeds)

Initial data removal

1. Compute the average $\overline{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$
2. Compute $\overline{x}_k = \frac{1}{n-k}\sum_{j=k+1}^{n} x_j$
3. Compute the relative variation $R_k = \frac{\overline{x}_k - \overline{x}}{\overline{x}}$
4. After drawing a graph of x_k or $\overline{R}_k$ as a function of k, choose the value of k identifying the knee in the curve

Initial data removal



Independent observations

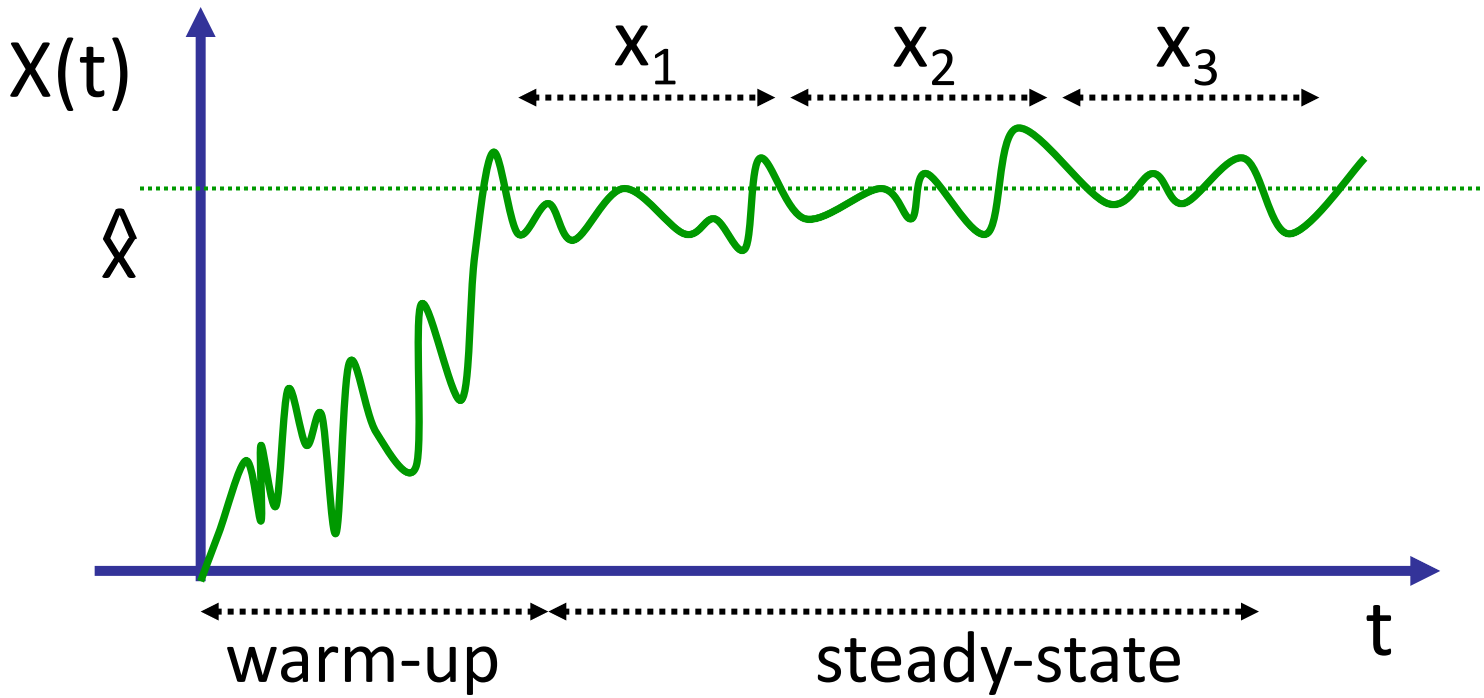
- Various methods can be used to have independent simulations:
  - Use different seeds for the sequences generated by a single generator: we need to verify that the seeds are distant in the sequence, to avoid using overlapping sequences
  - Partition a single sequence (run) in non-overlapping subsequences (*batch means* method)

Batch means

- The following steps
  - Run a very long simulation
  - Split the simulation in intervals (batches)
  - For each interval, compute an estimation of the quantity under study
- Samples are correlated, but only at the boundaries of the intervals
- The pseudorandom sequences do not overlap
  - We do not need to choose different seeds

Batch means

- Need to remove the warm-up transient only once



Batch means

- The number of batches should be between 10 and 30
  - Using more than 30 batches usually does not produce significant improvements of the confidence interval
  - Less than 10 batches might produce too wide intervals
- Often, we choose the number of batches computing on-the-fly the width of the obtained confidence interval and deciding if further batches are needed

Batch means

1. Remove the warm-up transient
2. Collect n=10 batches
3. Compute the confidence interval I_n=[x-z,x+z]
4. If (2z/x)> P,
   i. n=n+1
   ii. Collect a further batch
   iii. Go to step 3

   otherwise return *x* and I_n