# Discrete Event Simulation

# IN2045

Dr. Alexander Klein

Stephan Günther

Prof. Dr.-Ing. Georg Carle

**Chair for Network Architectures and Services**

**Department of Computer Science**
**Technische Universität München**
**http://www.net.in.tum.de**

Technische Universität München

# Topics

❑ System Initialization

❑ Estimator

- ▪ Consistent Estimator
- ▪ Unbiased Estimator
- ▪ Variance of an Estimator
- ▪ Efficient Calculation of an Estimator

❑ Confidence Interval

- ▪ Tschebyscheff Confidence Interval
- ▪ Central Limit Theorem
- ▪ t-Distribution Confidence Interval

❑ Evaluation of Simulation Results

- ▪ Replicate-Delete Method
- ▪ Batch Means Method
- ▪ Stationarity
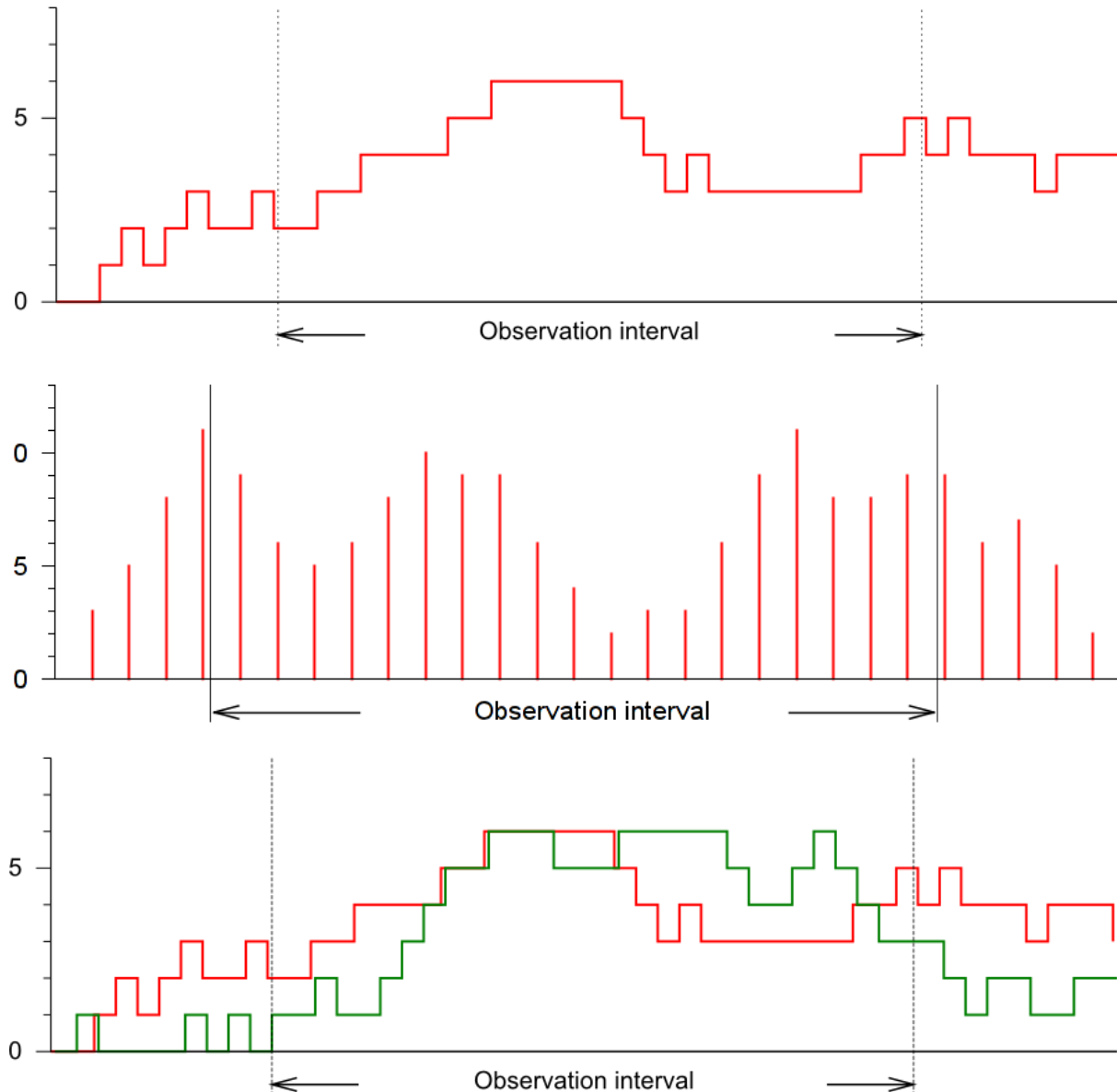
# Evaluation of Simulation Results

❑ **Simulation study**

- Goals:

  - Evaluation of system S

  - Impact of (manageable) input variables C

  - Impact of (unmanageable) input variables U

  - Evaluation of outcome (result) P

  - Performance parameter Y

- Problem:

  - Infinite number of different outcomes

  - Probability of a certain outcome cannot determined in advance

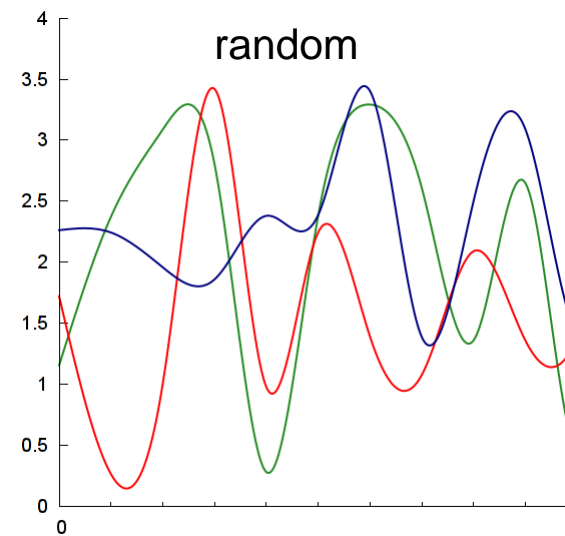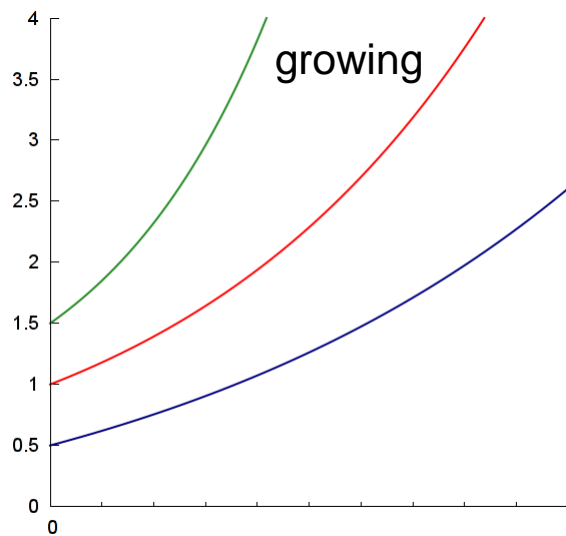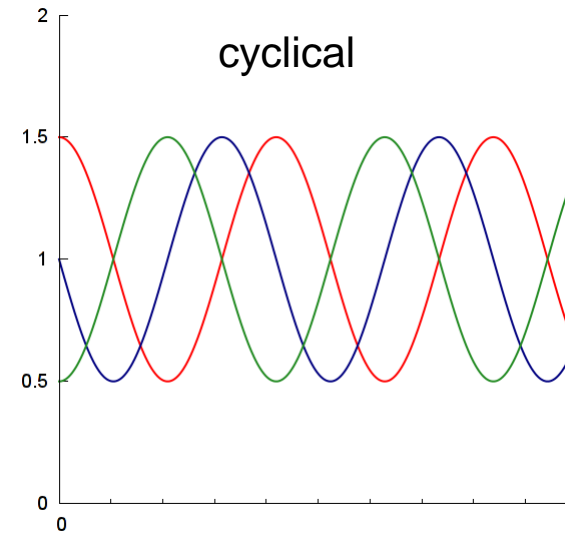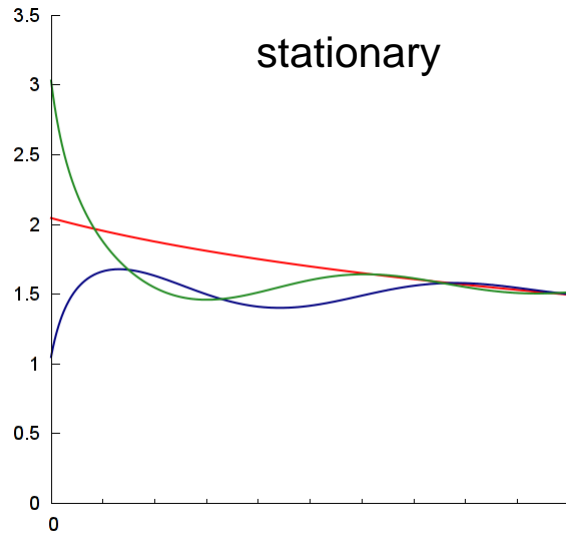  - Parameter of interest can be regarded as random variable

❑ **Evaluation:**

- n simulation runs
- m samples per simulation run
- jth sample of the corresponding simulation run
- i – number of simulation run

Trajectory ⟹

$$
\begin{array}{ccccc}
y_{11}, & \cdots & y_{1j}, & \cdots & y_{1m} \\
\vdots & & \vdots & & \\
y_{i1}, & \cdots & y_{ij}, & \cdots & y_{im} \\
\vdots & & \vdots & & \vdots \\
y_{n1}, & \cdots & y_{nj}, & \cdots & y_{nm}
\end{array}
$$
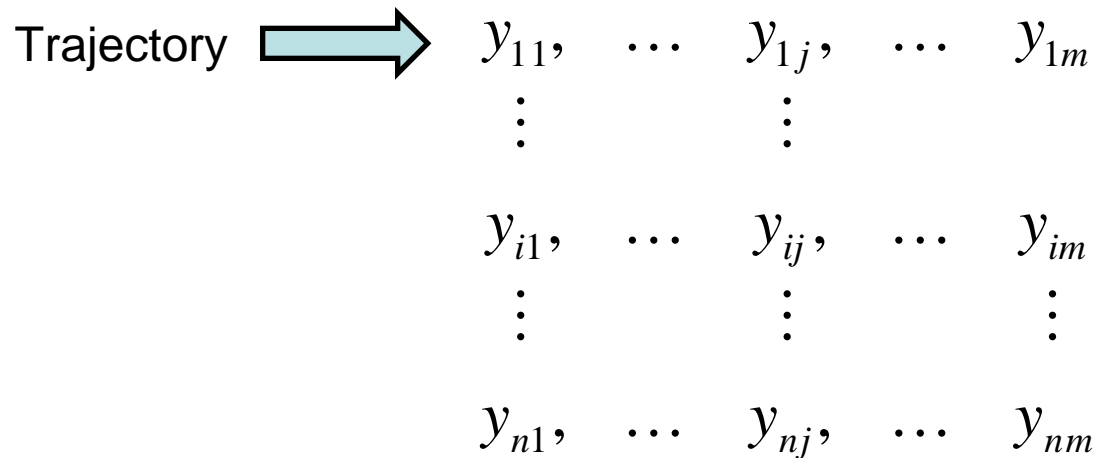
**Simulation Matrix**

# Evaluation of Simulation Results

❑ **System Initialization**

- Initialization state has to be chosen with respect to the expectation of the random variable

- The transient phase of the system depends on the initialization state

- The mean of the random variable (usually) converges to a certain level

- Subsequent measurements are often correlated
  (e.g. waiting queue length)
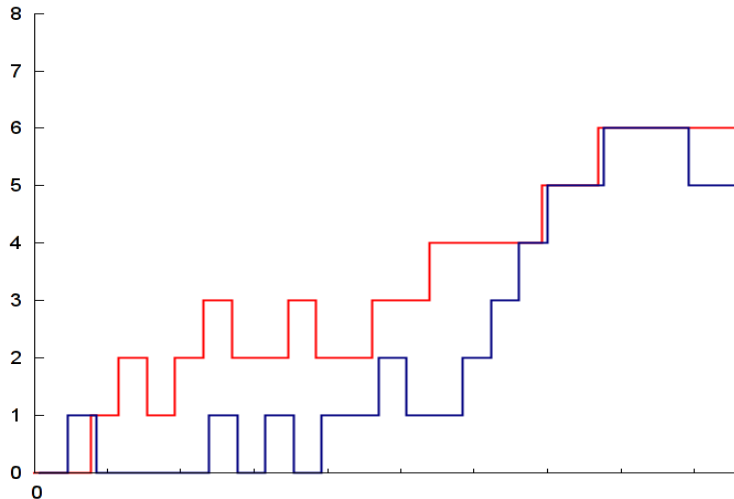
Random variable Y for different simulation runs
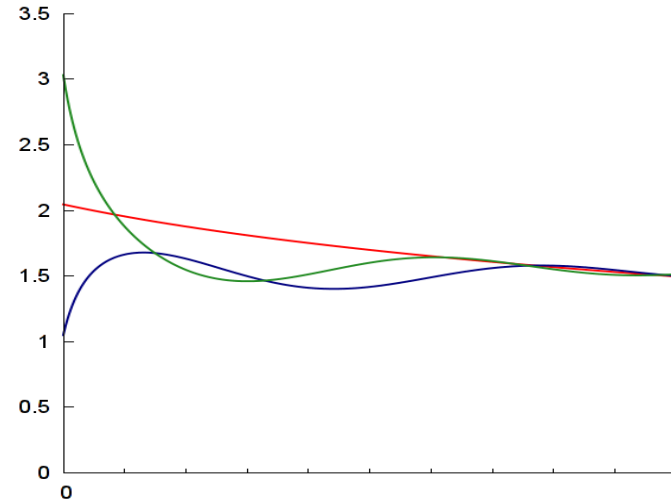
Mean of random variable Y depending on the initialization state

# Evaluation of Simulation Results

❑ **Estimator** (Schätzer)

- ▪ Definition:

  An estimator is a statistic which is used to infer the value of an unknown parameter (estimand) in a statistic model.

- ▪ Problem:

  - Estimate different characteristics (e.g. mean) of an observed parameter (e.g. delay or packet loss) with a small/certain number of samples.
  - Calculate the quality of the estimation

❑ **Consistent estimator** (konsistenter Schätzer)

- ▪ Definition:

  An estimator is called consistent if its precision increases with the number of samples

$$\lim_{n\to\infty} P[|\tilde{Y}_j - E(Y_j)| > \varepsilon] = 0 \qquad \forall \varepsilon > 0$$

❑ **Unbiased estimator** (erwartungstreuer Schätzer)

- ▪ Definition:

   An estimator is called unbiased if its mean equals the true mean of the estimation parameter.

- ▪ Example:

   - • Assume a very large population of elements with a different characteristic (e.g. height of individuals) and μ being the mean of the characteristic

   - • Let $E[\hat{Y}_i]$ be the mean of n collected sample values and $\overline{X}$ the random variable which consists of these mean values.

$$\overline{X} = \frac{1}{n} \cdot (X_1 + X_2 + \ldots + X_n)$$

$$E(\overline{X}) = E\left[\frac{1}{n} \cdot (X_1 + X_2 + \ldots + X_n)\right] = \frac{1}{n} E\left[X_1 + X_2 + \ldots + X_n\right]$$

$$E(X_1) = E(X_2) = \ldots = E(X_n) = \mu$$

$$E(\overline{X}) = \mu$$

# Evaluation of Simulation Results

❑ **Estimation of E(Y)**

- Estimator $\tilde{Y}$

- Estimation (value) $\hat{Y}$

- Point Estimator of E(Y) $\quad \hat{Y}_j = \dfrac{1}{n} \cdot \sum_{i=1}^{n} y_{ij}$

- Random Variable $\quad \tilde{Y}_j$

- Outcome of $\quad RV \; \tilde{Y}_j \Longrightarrow \hat{Y}_j$

- Consistency of $\quad \tilde{Y} \Longrightarrow \lim_{n \to \infty} P[|\tilde{Y}_j - E(Y_j)| > \varepsilon] = 0 \quad \forall \varepsilon > 0$

**Quality of the estimator still unknown**

# Evaluation of Simulation Results

❑ Variance of the estimator $\tilde{Y}$ represents a first indicator of its quality

- Calculation of the variance of the estimator after n samples

$$\sigma^2(\tilde{Y}_j) = E\left(\left(\tilde{Y}_j - E(\tilde{Y}_j)\right)^2\right)$$

$$= E\left(\left(\frac{1}{n}\sum_{i=1}^{n} Y_{ij} - E(Y_j)\right)^2\right)$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} E\left((Y_{ij} - E(Y_j))^2\right) + \sum_{i=1}^{n}\sum_{k=1,k\neq i}^{n} E\left((Y_{ij} - E(Y_j))\cdot(Y_{kj} - E(Y_j))\right)\right)$$

Double sum represents the correlated part which is 0 if $y_{ij}$ are uncorrelated

$$\sigma^2(\tilde{Y}_j) = \frac{1}{n^2}\sum_{i=1}^{n} E\left((Y_{ij} - E(Y_j))^2\right) = \frac{1}{n^2}\cdot n \cdot \sigma^2(Y_j) = \frac{\sigma^2(Y_j)}{n}$$
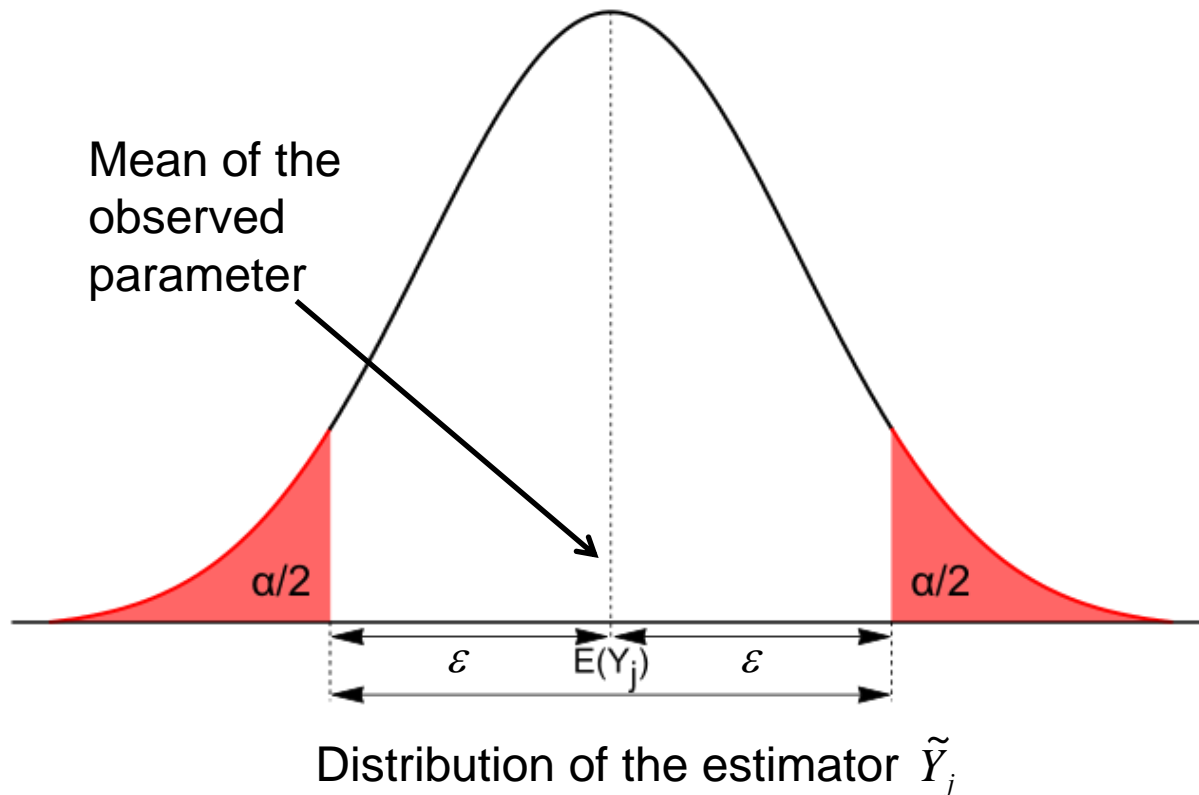
- Variances of the estimator increases
  - with the variance of the estimand
  - if the number of samples is reduced
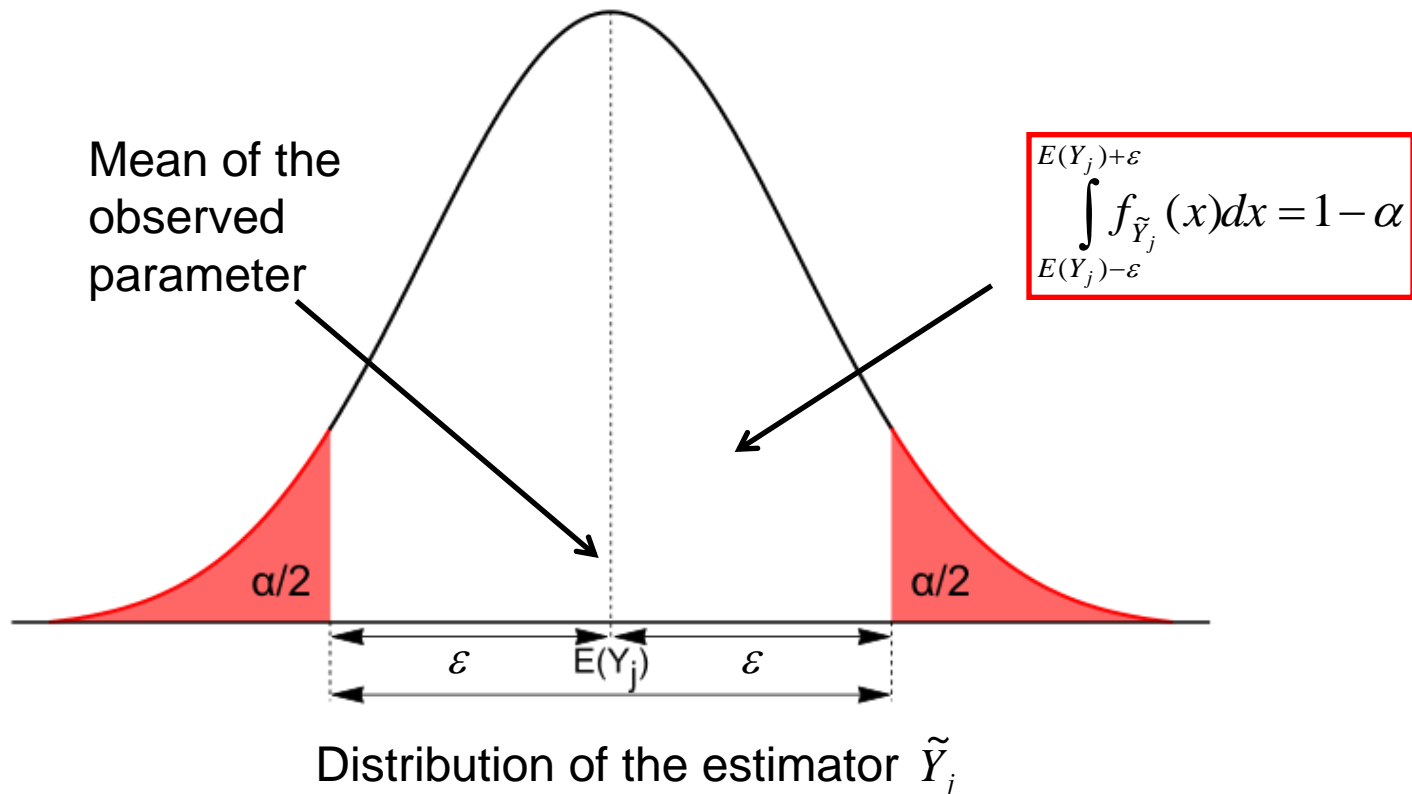
❏ **Unbiased estimator for** $\sigma^2(Y_j)$

$$\tilde{S}_j^2(Y_j) = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_{ij} - \tilde{Y}_j\right)^2$$

Mean of the observed parameter

$\alpha/2$        $\alpha/2$

$\varepsilon$   $E(Y_j)$   $\varepsilon$

Distribution of the estimator $\tilde{Y}_j$

- ❑ Assume that the probability density function of the estimator $f_{\tilde{Y}_j}(y)$ is known in advance

    - ▪ The probability that $\tilde{Y}_j$ lies within the interval is $1-\alpha$



Mean of the observed parameter

$$\int_{E(Y_j)-\varepsilon}^{E(Y_j)+\varepsilon} f_{\tilde{Y}_j}(x)dx = 1-\alpha$$

$\alpha/2$       $\alpha/2$

$\varepsilon$    $E(Y_j)$    $\varepsilon$

Distribution of the estimator $\tilde{Y}_j$

# Evaluation of Simulation Results

❑ **Biased estimator for the sample variance** $\sigma^2(Y_j)$

$$\widetilde{S}_{nj}^2(Y_j) = \frac{1}{n}\sum_{i=1}^{n}\left(Y_{ij} - \widetilde{Y}_j\right)^2 = \frac{\sum_{i=1}^{n}x_i^2}{n} - \frac{\left(\sum_{i=1}^{n}x_i^2\right)}{n^2}$$

$\widetilde{S}_{nj}^2$ is a biased estimator of the sample variance since it systematically underestimates it.

❑ **Bessel's correction**

The biased estimator can be transformed in an unbiased estimator of the sample variance by applying Bessels's correction.

$$\frac{n}{n-1}$$

❑ **Unbiased estimator for the sample variance** $\sigma^2(Y_j)$

$$\widetilde{S}_j^2(Y_j) = \frac{n}{n-1}\cdot\widetilde{S}_{nj}^2(Y_j) = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_{ij} - \widetilde{Y}_j\right)^2$$

# Evaluation of Simulation Results

❑ **Efficient calculation of the estimator**

- ▪ 1. Problem

  Number of samples $y_{ij}$ can become high which results in high memory consumption

  - • Solution

    Recursion $\quad \hat{Y}_j(k) = \dfrac{k-1}{k}\hat{Y}_j(k-1) + \dfrac{y_{kj}}{k}$

- ▪ 2. Problem

  The calculation of the sample variance requires the direct evaluation of the estimation of the variance

  $$\hat{S}_j^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_{ij} - \hat{Y}_j\right)^2$$

  ⟹ Every sample $y_{ij}$ has to be stored

  - • Solution

    – Store the sums of $y_{ij}$ and $y_{ij}^2$

    – Calculate $\quad \hat{S}_j^2 = \dfrac{1}{n-1}\left(\sum_{i=1}^{n}\left(y_{ij}\right)^2 - n\cdot\left(\hat{Y}_j\right)^2\right)$

The size of n can be reduced if means are used instead of single sample values.

❑ **Confidence interval**

- Definition: $P[|\tilde{Y}_j - E(Y_j)| \geq \varepsilon] = \alpha$

  Calculate an interval 2ε around the $E(Y_j)$ such that a sample of $\tilde{Y}_j$ lies in the interval with a probability of $1 - \alpha$

  Larger $\alpha$ ⟹ Smaller $\varepsilon$

  Smaller $\alpha$ ⟹ Larger $\varepsilon$

- $\hat{Y}_j \pm \varepsilon$ is called $(1-\alpha) \cdot 100\%$ confidence interval of $E(Y_j)$

- $[\hat{Y}_j - \varepsilon, \hat{Y}_j + \varepsilon]$ is called interval estimator

⟹ Interval estimators are more important than point estimators since they make probability based assumptions which consider the variance of the estimator.

□ **Confidence interval according to Tschebyscheff**

- Let X be a random variable with mean $E(X)$ and variance $\sigma^2(X)$

$$P[|X - E(X)| \geq c] \leq \frac{\sigma^2(X)}{c^2} \qquad \forall c > 0$$

- Replace c by ε and X by $\tilde{Y}_j$. Variance of $\tilde{Y}_j$ is $\frac{\sigma^2(Y_j)}{n}$

$$\Longrightarrow \qquad P[|\tilde{Y}_j - E(Y_j)| \geq c] \leq \frac{\sigma^2(X)}{n \cdot \varepsilon^2}$$

❑ **Confidence interval according to Tschebyscheff**

- ▪ Example:
  - • Calculate the 90% confidence interval of

  $$\Longrightarrow \quad \alpha = \frac{\sigma^2(Y_j)}{n \cdot \varepsilon^2} = 0.1$$

  - • Now assume a sample size n = 10

  $$\Longrightarrow \quad \frac{\sigma^2(Y_j)}{10 \cdot \varepsilon^2} = 0.1 \quad \Longleftrightarrow \quad \varepsilon^2 = \sigma^2(Y_j) \quad \Longleftrightarrow \quad \varepsilon = \sigma(Y_j)$$

  - • Thus, a confidence interval of half size requires four times the number of samples

  $$\Longrightarrow \quad \frac{\sigma^2(Y)}{n \cdot \varepsilon^2} = \frac{\sigma^2(Y)}{m \cdot (\varepsilon/2)^2} \quad \Longleftrightarrow \quad m = 4 \cdot n$$

# Evaluation of Simulation Results

❑ **Confidence interval according to Tschebyscheff**

- ▪ Disadvantages:
  - • The calculation requires knowledge of the variance $\sigma^2(Y)$ of the estimand which is typically unknown and must thus be replaced by the estimator $\hat{S}_j^2$ .

    > Note that this breaks the pre-condition of Tschebyscheff which makes the calculated bounds invalid.

  - • Tschebyscheff intervals are very large / pessimistic since they are valid for any given distribution.

    > The pessimistic characteristic of Tschebyscheff is often used as justification for replacing the variance $\sigma^2(Y)$ with the estimator $\hat{S}_j^2$ .

- ❑ **Confidence interval according to Tschebyscheff**
  - ▪ Example:
    - Flipping a coin. RV Y => Y ∈ {0-head, 1-tail}
    - $E(Y) = 0.5$ , $\sigma^2(Y) = 0.25$
    - Flipping the coin 10 times after another  ⟹  n = 10 (samples)

    - Calculate 90% confidence interval  ⟹  $\alpha = 0.1$

    - Experiment 1:  0000101001

      ⟹  $\hat{Y} = 0.3$  ∧  $\hat{S}^2 = 0.23333$

      ⟹  [-0.183, 0.783]  ⟹  [0, 0.783]

    - Experiment 2:  0110111001

      ⟹  $\hat{Y} = 0.6$  ∧  $\hat{S}^2 = 0.26667$

      ⟹  [0.084, 1.116]  ⟹  [0.084, 1.000]

# Evaluation of Simulation Results

❑ **Confidence interval according to Tschebyscheff**

- ▪ Example:
  - • Flipping a coin. RV Y => Y ∈ {0-head, 1-tail}
  - • $E(Y) = 0.5$ , $\sigma^2(Y) = 0.25$
  - • Flipping the coin 20 times after another ⟹ n = **20** (samples)
  - • Calculate 90% confidence interval ⟹ $\alpha = 0.1$
  - • Concatenation of Experiment 1 and 2:  00001 01001    01101 11001

    ⟹ $\hat{Y} = 0.3$ ∧ $\hat{S}^2 = 0.23333$

    ⟹ [0.089, 0.811] ⟹

  - • Summary:
    - – The true mean lies within the interval with a probability of 90%
    - – 1 of 10 experiments will not contain the true mean

## ❑ Central limit theorem

The distribution of the (normalized and centralized) sum of a large number of independent and identical distributed random variables can be approximated by the (standard) normal distribution.

Lindeberg-Lévy theorem

Let $X_1, X_2, \ldots, X_n$ be a sequence of random variables within the same probability space which are independent and follow the same distribution. The mean of each random variable is μ and the standard variation is σ.

In the following we take a closer look at the nth sum of the sequence.

$$S_n = X_1 + X_2 + \ldots + X_n \quad \Longrightarrow \quad E[S_n] = n\mu \quad \Longrightarrow \quad \sigma^2[S_n] = n\sigma$$

Introduce a new standardized random variable $Z_n$

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

❑ **Central limit theorem**

The distribution of the random variable $Z_n$ converges against the (standard) normal distribution according to the central limit theorem if the number of summands n increases.

$$\Longrightarrow \quad \lim_{n \to \infty} P(Z_n \leq z) = \Phi(z) \qquad z \in R$$

With $\Phi(z)$ representing the (standard) normal distribution $N(0,1)$

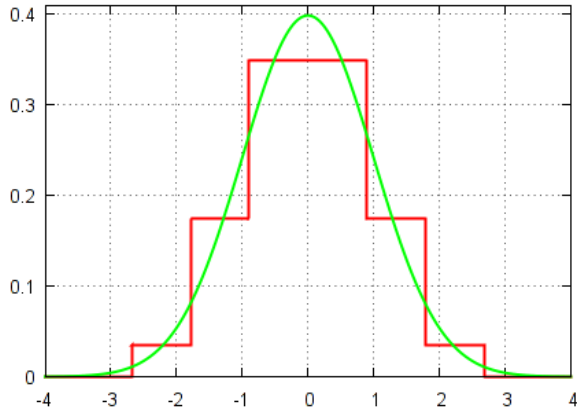$$\Longrightarrow \quad \lim_{n \to \infty} P(\frac{\overline{X}_n - \mu}{\sigma / \sqrt{n}} \leq z) = \Phi(z)$$

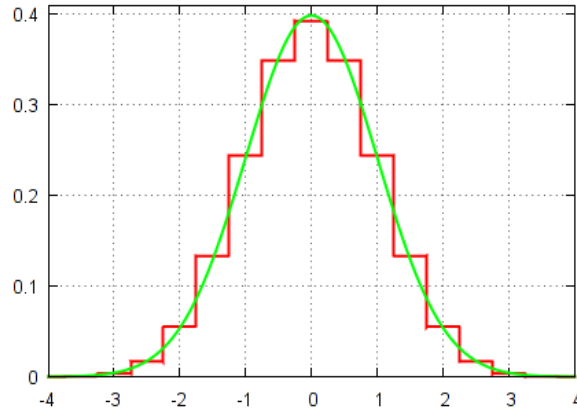$$\overline{X}_n = \frac{S_n}{n} = \frac{X_1 + \ldots + X_n}{n}$$
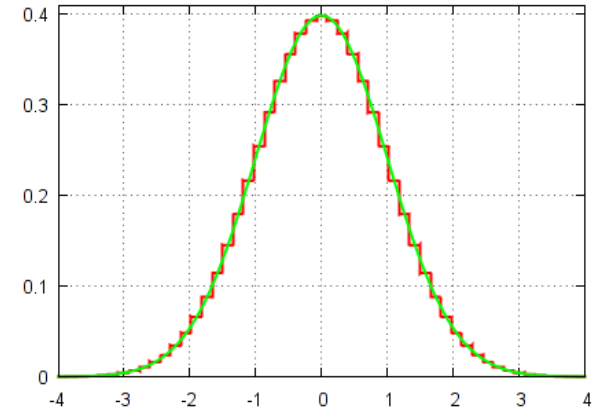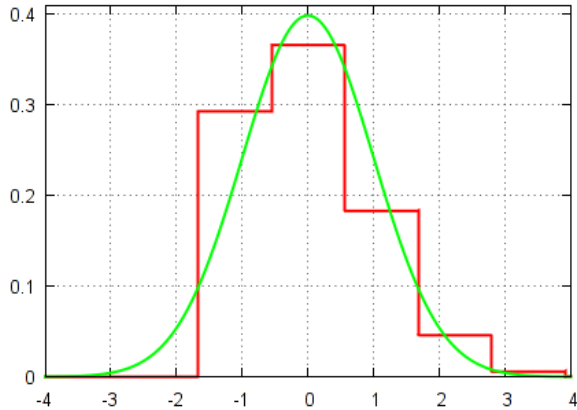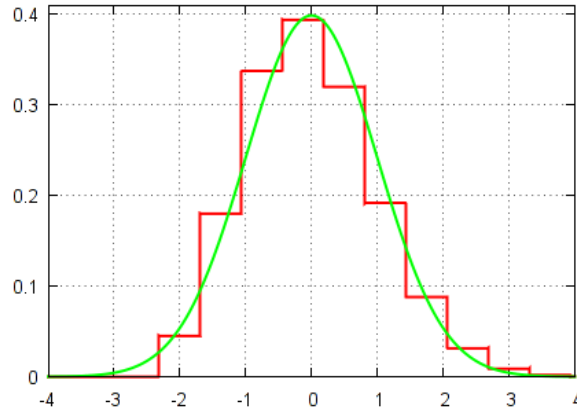
## ❑ Central limit theorem
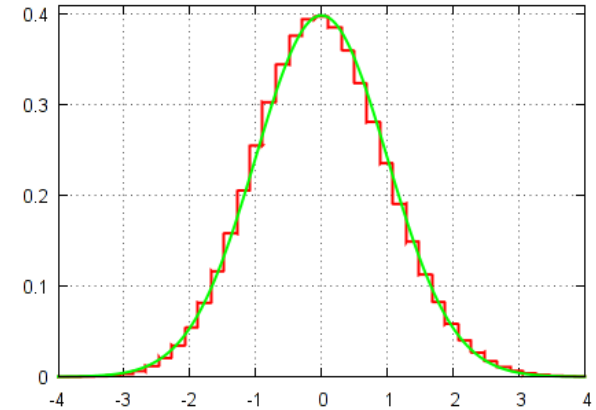


Sum of binomial distributed random variables

□ **Confidence interval according to the central limit theorem**

- Idea: The central limit theorem is still valid if $\sigma^2$ is replaced by $\tilde{S}^2$. Thus, it is possible to calculate the critical values out of the normal distribution.

- Recapitulate the "flipping of a coin example" with $\tilde{Y}$ representing the distribution of the estimator and $Y$ being the distribution of the estimand. Then we can calculate the confidence interval as follows:

$$\Longrightarrow \quad P[|Z| \geq \varepsilon] = P\left[\left|\frac{\tilde{Y} - E(Y)}{\tilde{S}/\sqrt{n}}\right| \geq \varepsilon\right] = \alpha$$

$$\Longrightarrow \quad P[|\tilde{Y} - E(Y)| \geq \varepsilon \cdot \tilde{S}/\sqrt{n}] = \alpha$$

$$\Longrightarrow \quad \boxed{\tilde{Y} \pm z_\alpha \cdot \tilde{S}/\sqrt{n}}$$

$$\Longrightarrow \quad z_\alpha \text{ is the } \alpha/2 \text{ percentile of } N(0,1)$$
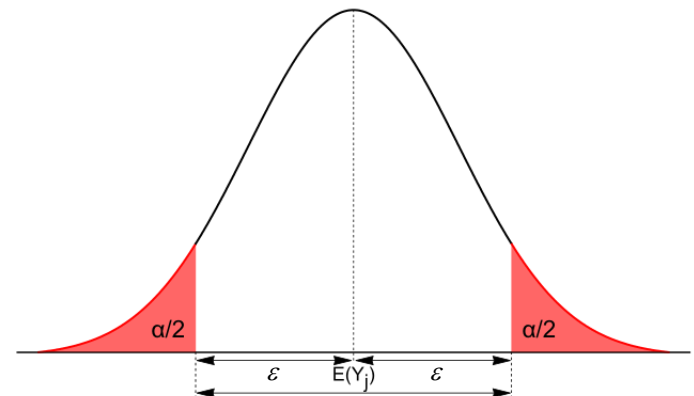
❑ **Confidence interval according to the central limit theorem**

The central limit theorem generates smaller confidence intervals

➡ Tschebyscheff [0.089, 0.811]

➡ Central Limit [0.262, 0.638]

▪ Question: What is the minimum value for n to allow the assumption that the estimator is normally distributed?

– The minimum value of n depends on the distribution of the estimand $Y$. In worst case scenarios, the mean of the estimand $E[Y]$ may be outside the confidence interval with a probability which is significantly higher than $\alpha$.

# Evaluation of Simulation Results

❑ **Empirical evaluation of the confidence interval calculation according to the (student) t-distribution**

- ▪ Problem:

  Only a view results for different distributions are known.

- ▪ In the following we assume that $Y$ is already normal distributed

$$\implies \frac{\tilde{Y} - E[Y]}{\tilde{S} / \sqrt{n}} \quad \text{follows a t-distribution with n-1 degrees of freedom}$$

- ▪ Critical / popular values of the t-distribution can be taken from tables

$$\implies \hat{Y} \pm t_{n-1,1-\alpha/2} \cdot \frac{\widehat{S}}{\sqrt{n}}$$

❑ **Empirical evaluation of the confidence interval calculation according to the (student) t-distribution**

- Idea:

  Apply a known distribution and calculate the confidence interval. Then repeat the experiment k times and estimate the probability with which the outcome of the experiment remains within the calculated boundaries.
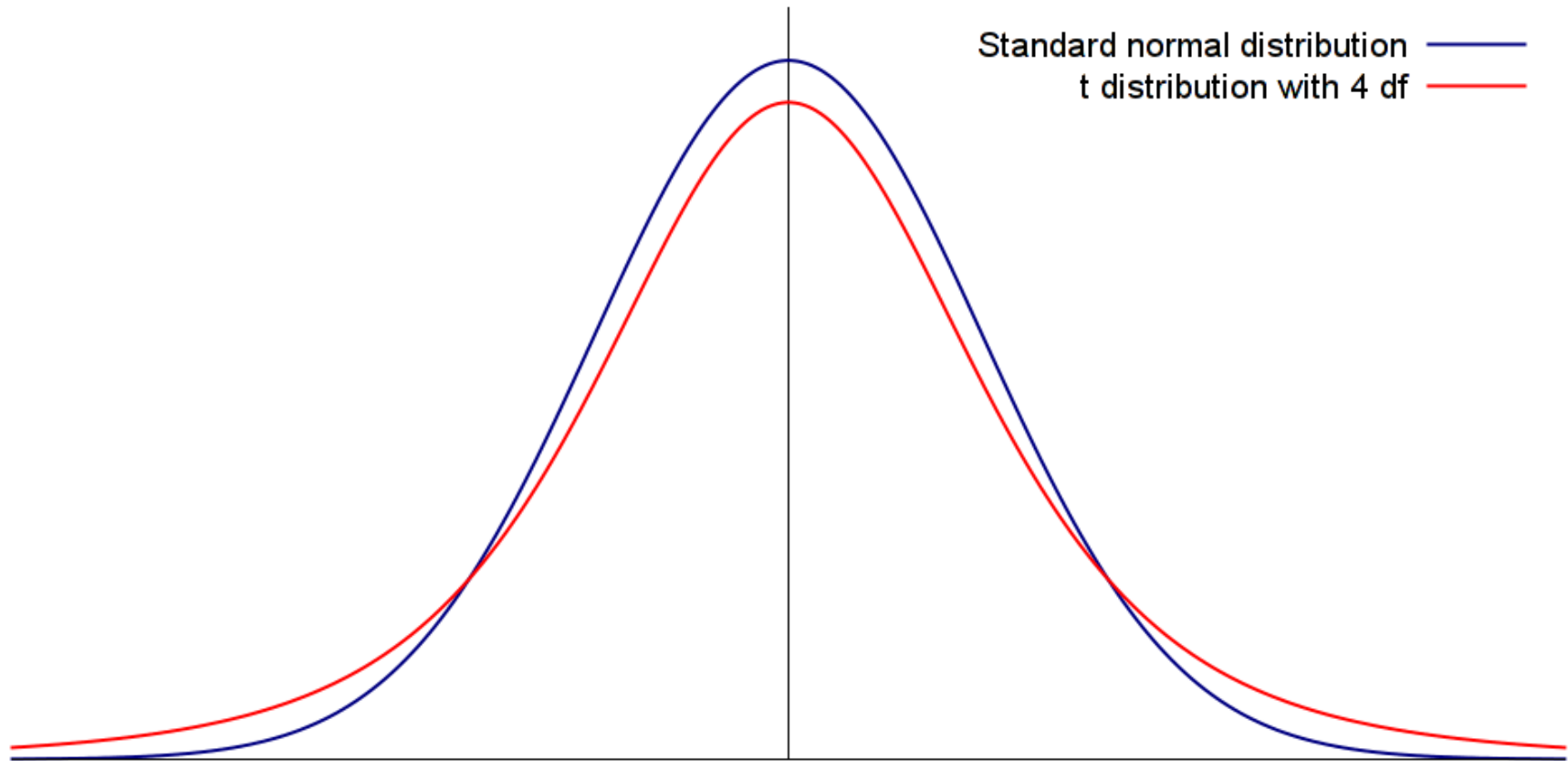
- Example: 90% confidence interval, k = 500 repetitions

| Verteilung | n=5 | n=10 | n=20 | n=40 |
|---|---|---|---|---|
| Normal | 0.910 | 0.902 | 0.898 | 0.900 |
| Exponential | 0.854 | 0.878 | 0.870 | 0.890 |
| Lognormal | 0.758 | 0.768 | 0.842 | 0.852 |
| Hyperexp. | 0.584 | 0.586 | 0.682 | 0.774 |

Table taken from Law: Simulation Modeling and Analysis, 4th Edition

## ❑ Student-t distribution

Standard normal distribution ——
t distribution with 4 df ——

Student-t distribution converges against the normal distribution with increasing numbers of degrees of freedom.

# Evaluation of Simulation Results

❑ **Breakdown point**

- The Breakdown point represents a metric for the robustness of an estimator since it defines the percentage of samples which are required to falsify the result of the estimator.

How to get useful simulation results out of a simulation

❑ **Replicate-Delete Method (LK 9.5.2)**

- Estimate the duration of the transient phase

- Replicate – Simulate a large number of runs

- Delete – Remove the transient phase since it does not contain meaningful results

- The duration of the simulation has to be a much longer than the duration of the transient phase

- Calculate the confidence intervals by using the mean values of the individual simulation runs
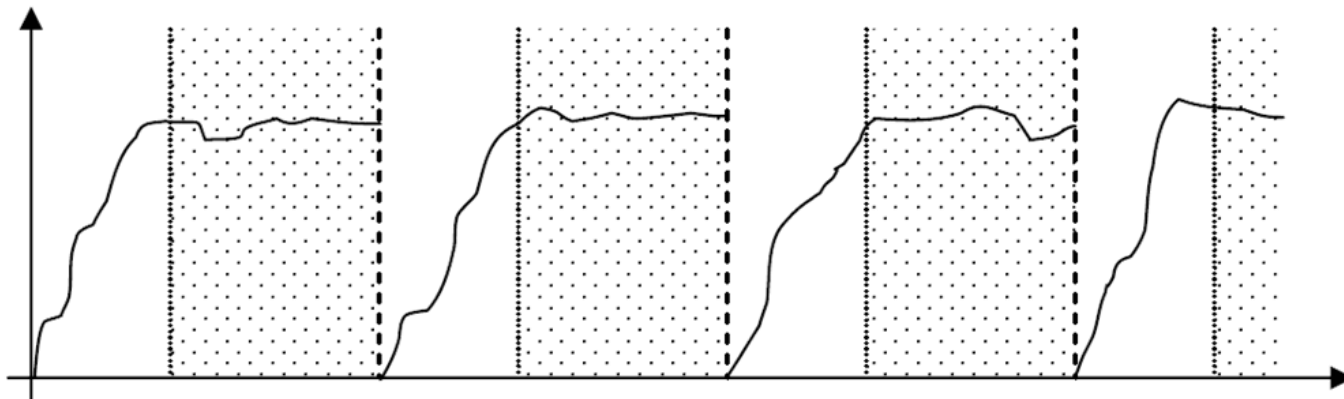
❑ **Replicate-Delete Method**

- ▪ Advantage:
  - • Most simple approach
  - • Less affected by correlation
  - • Typically supported by all simulation tools

- ▪ Disadvantage:
  - • Requires correct estimation of the duration of the transient phase.
  - • Underestimation of the duration of transient phase results in falsified simulation results.
  - • Requires more time compared to Batch-Means since the transient phase has to be simulated several times.

❑ **Covariance**

Covariance is a measure which describes how two variables change together

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X] \cdot E[Y]$$

▪ Special Case: $\quad Cov(X,X) = VAR[X]$

▪ Other Characteristics:

- $\quad Cov(X,a) = 0$

- $\quad Cov(X,Y) = Cov(Y,X)$

- $\quad Cov(aX,bY) = abCov(X,Y)$

- $\quad Cov(X+a,Y+b) = Cov(X,Y)$

## Correlation function

Correlation function describes how two random variable tend to derivate from their expectation

$$Cor(X,Y) = \frac{Cov(X,Y)}{\sqrt{VAR(X) \cdot VAR(Y)}}$$

- Characteristics:

  - $Y = X$ $\implies$ $Cor(X,Y) = 1$ (Maximum positive)

  - $Y = -X$ $\implies$ $Cor(X,Y) = -1$ (Maximum negative)

  - $Cor(X,Y) > 0$    Both random variable tend to have either high or low values (difference to their expectation)

  - $Cor(X,Y) < 0$    The random variables differ from each other such that one has high values while the other has low values and vice versa (difference to their expectation)

## Autocorrelation (LK 4.9)

- Autocorrelation is the cross-correlation of a signal with itself. In the context of statistics it represents a metric for the similarity between observations of a stochastic process. From a mathematical point of view, autocorrelation can be regarded as a tool for finding repeating patterns of a stochastic process.

Definition:

- Correlation of two samples with distance k from a stochastic process X is given by:

$$Cor(X, Y) \quad \text{with} \quad Y_i = X_{i+j}$$

Use case:

- Test of random number generators (remember spectral test)
- Evaluation of simulation results (c.f. Batch-Means)

# Evaluation of Simulation Results

❑ **Batch-Means Method (LK 9.5.3)**

- Estimate the duration of the transient phase
- Perform a long simulation run
- Remove the transient phase
- Divide the gathered results in n intervals of equal length (Batches) which hold m samples
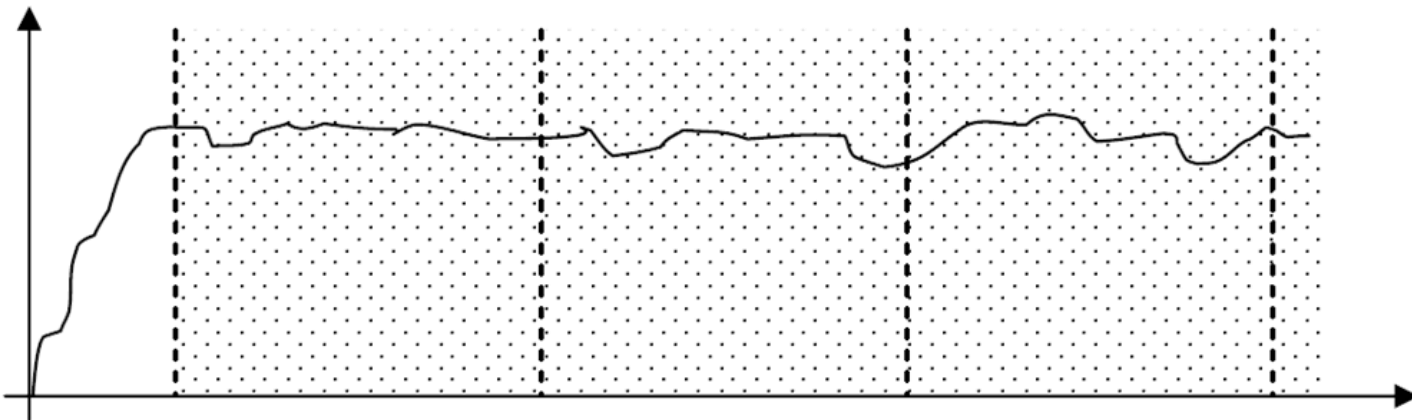
➡ Assure that the mean of subsequent batches is uncorrelated (calculate the empirical autocorrelation)

➡ Number of batches $n \geq 10$    Batch size $m \geq 10 \cdot x$

❑ **Batch-Means Method**

- Calculate the confidence intervals by using the mean values of the batches
- Minimize the absolute and relative error by increasing the number of batches (longer simulation run)

Optional approach:

- Estimate the duration of the transient phase
- Choose a sufficient value for m
- Simulate until the confidence interval has the desired size

# Evaluation of Simulation Results

❑ **Batch-Means Method**

- ▪ Advantage:
  - • Minimizes the time to get meaningful results since only a single transient phase has to be simulated
  - • Errors of the estimation of the duration of the transient phase decrease with increasing number of batches

- ▪ Disadvantage:
  - • Calculation of n and m is complicated and usually requires detailed knowledge of the simulation
  - • Calculation of the autocorrelation of the intervals have to be calculated in order to assure that the corresponding means are not correlated
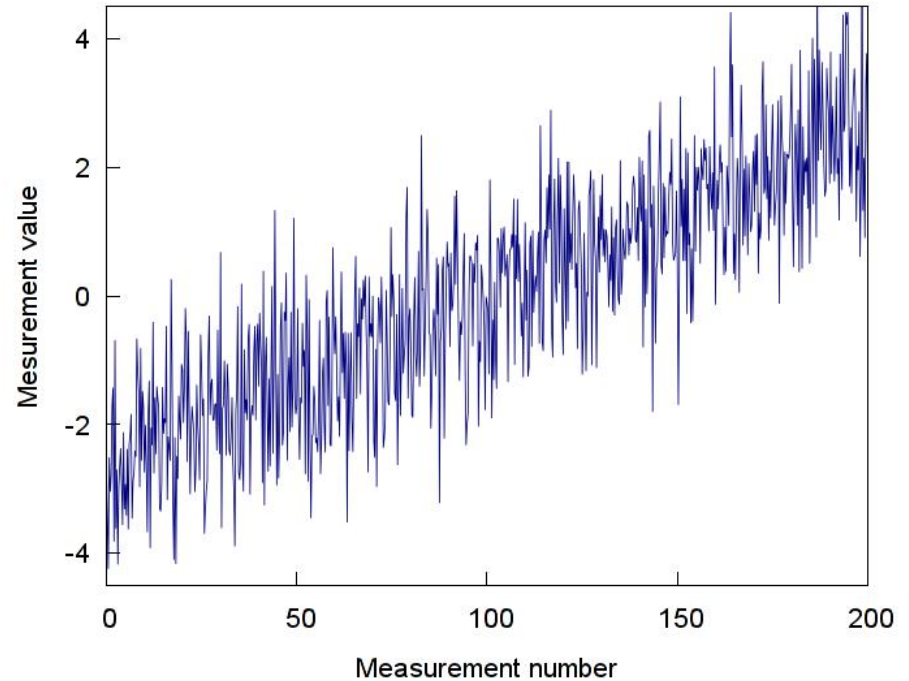
❑ **What's stationarity? – An intuitive graphical explanation:**
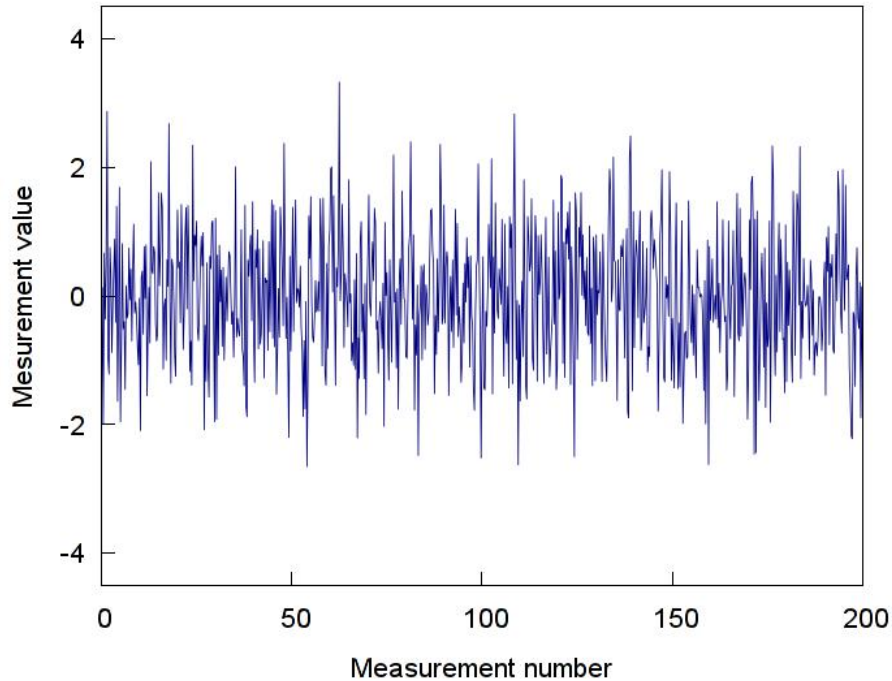
❑ **Its not just trends**
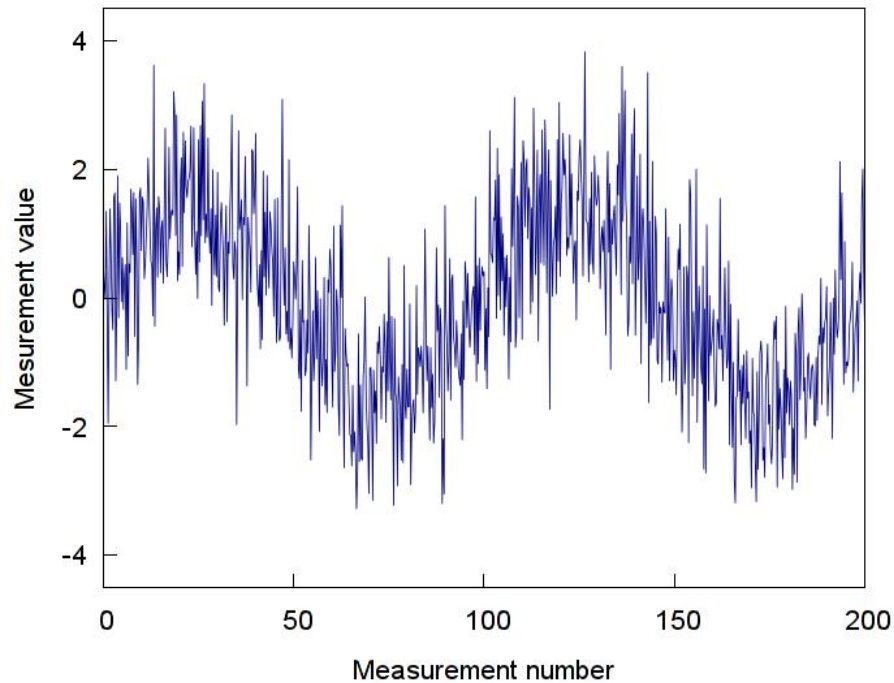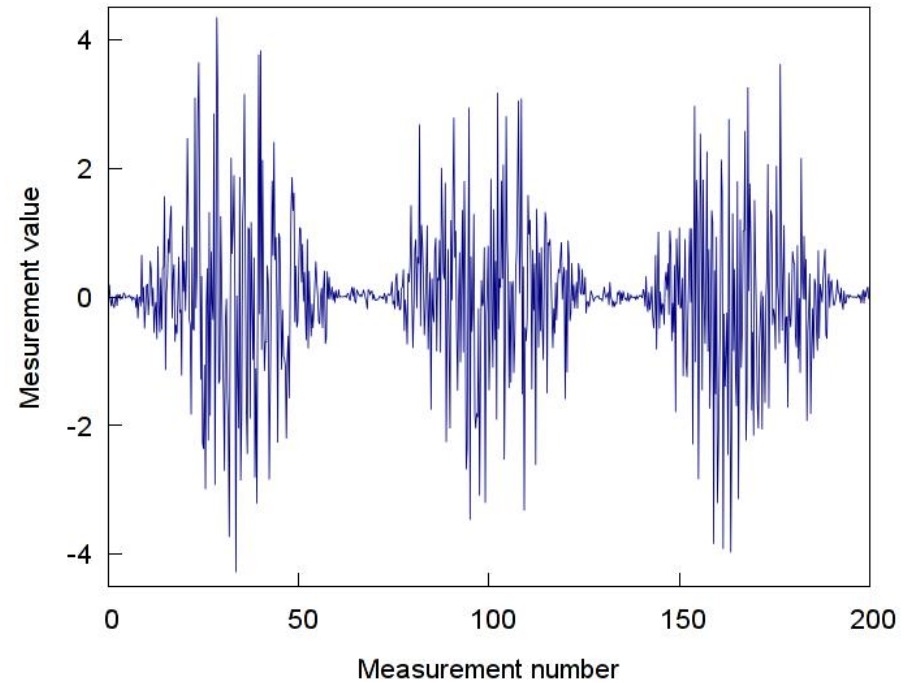
# Mathematical definitions of stationarity

❑ Strong stationarity:

- All samples $X_i$ are drawn from exactly the same underlying distribution

- In practice, this is hard or impossible to prove

❑ Other types of stationarity:

- Mean stationary: $\mu(X_i) = const \; \forall i$

- Variance stationary: $Var(X_i) = const \; \forall i$

- Covariance stationary: $Cov(X_i, X_{i+k}) = const(k) \; \forall i$ (only dependent on lag)

- Weakly stationary: The $X_i$ are mean stationary and covariance stationary

- In practice, weak stationarity is most commonly used

Assume Xi and Yi are [weakly] stationary processes. Then…:

❑ You can shift a stationary process:
$\alpha$ + Xi is stationary, too

❑ You can scale a stationary process:
$\beta \cdot$ Xi is stationary, too

❑ You can add stationary processes together:
Xi + Yi is stationary, too

# Why…?

- ❑ Important term in statistics
  - ▪ Many methods, algorithms, mechanisms assume that all samples come from the same distribution
    - • Warning: We experience phenomena such as convergence phases at the beginning of simulations, etc. – this means it's not stationary [yet]!
    - • Often would need strong stationarity, but often weak can do the trick
  - ▪ May be interesting to analyse if the output of a simulator / experiment / … is [weakly] stationary or not
- ❑ How to test for [weak] stationarity?
  - ▪ Tests usually built into statistics packages
  - ▪ Parametric tests for stationarity
    - • Make assumptions about underlying data (e.g., normally distributed)
  - ▪ Nonparametric tests for stationarity
    - • Need more measurements (usually 5%–35% more samples)

# Example

- Calculating confidence intervals
  - Assumption: All samples are drawn from the same population
  - But what if you take measurements from a process that has not converged yet?
  - Solution: Check the time series of measurements for stationarity