Universiteit Utrecht

Universiteit Utrecht

[Faculty of Science
Information and Computing Sciences]

Algorithms for Decision Support

Output analysis

Stochastic variables occur in simulation at different places:

1. Input data are modeled as stochastic variables
   - E.g time until arrival of next customer
2. Generate random variables
   - When you schedule a new Arrival event you have to generate a random number for the time delay
3. Analysis of results

This lecture

This lecture

- ***Output:*** A simulation determines the value of some performance measures, e.g. production per hour, average queue length etc...
- If your model contains random input values (e.g. customers interarrival times), your output, i.e., performance measures, are stochastic variables as well
- *In this lecture you learn basic statistical principles to analyse the output values of a simulation*

Output analysis

- Quote from Law (simulation book):

`Simulation is computer-based statistical sampling experiment’

some statistics in general
Estimators (unbiased)

Assume $X_1,\ldots,X_n$ samples of a stochastic variable X modelling given entity (length of bachelor students)

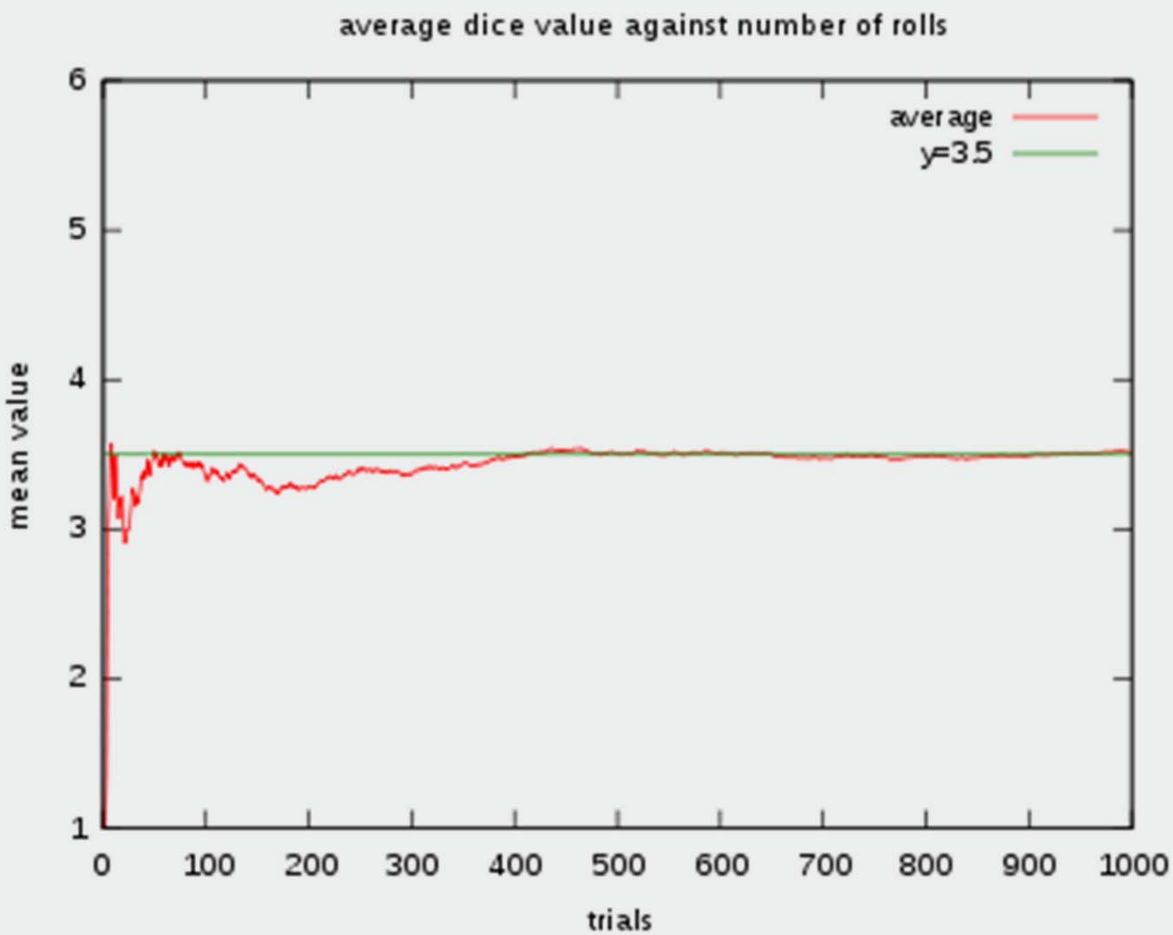
Sample mean: (estimates μ of the underlying distribution)

$$\overline{X}(n) = \frac{\sum_{i=1}^{n} x_i}{n}$$

Sample variance: (estimates σ^2 of the underlying distribution)

$$S^2(n) = \frac{\sum_{i=1}^{n} (X_i - \overline{X}(n))^2}{n-1}$$

Strong law of large numbers



Universiteit Utrecht

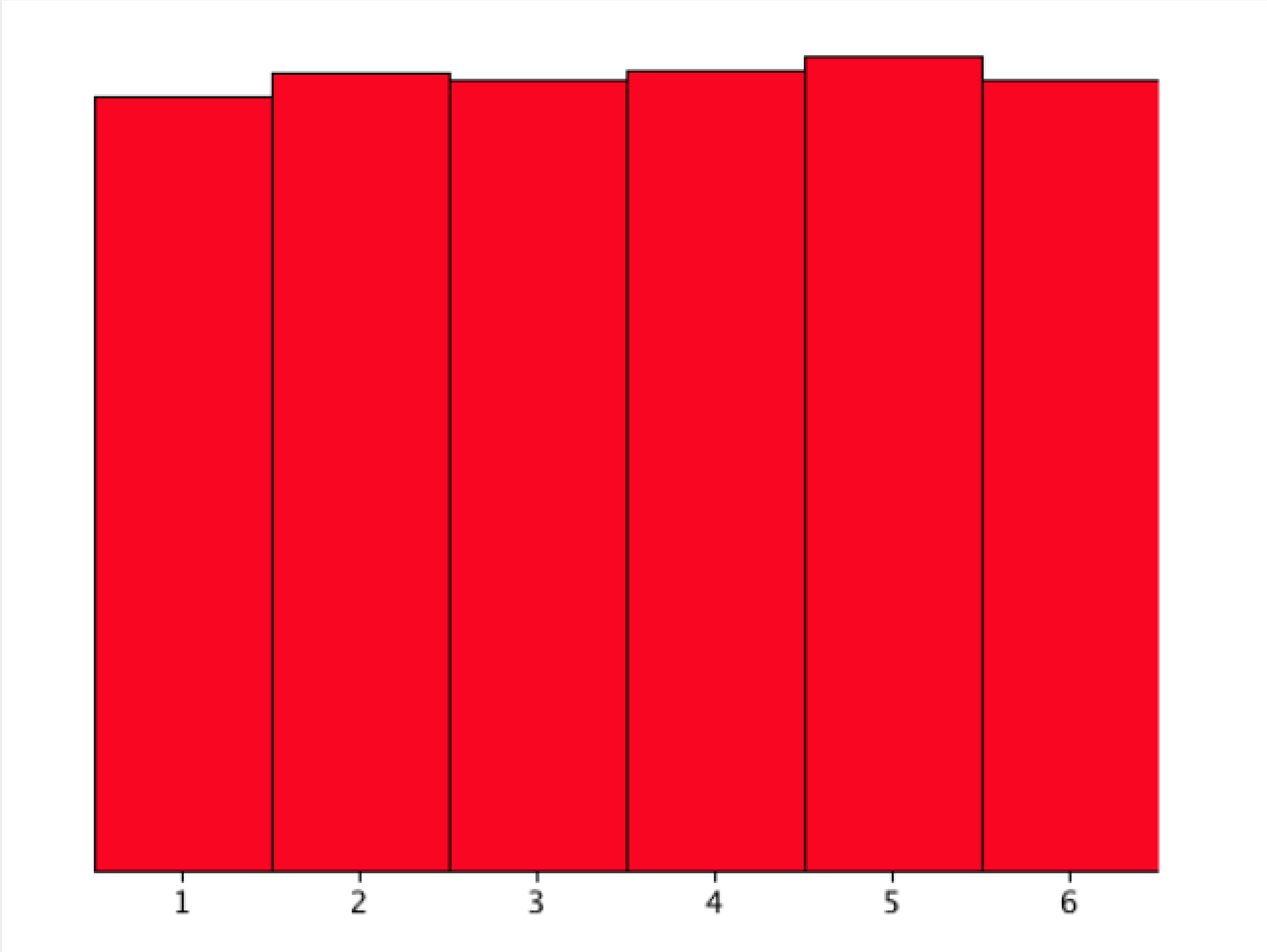
[Faculty of **Science**
Information and Computing Sciences]

Strong law of large numbers

$X_1,\ldots,X_n$ samples from a stochastic variable X with $E(X) = \mu$

$$\overline{X}(n) \rightarrow \mu \quad \text{with probability 1 if } n \rightarrow \infty$$

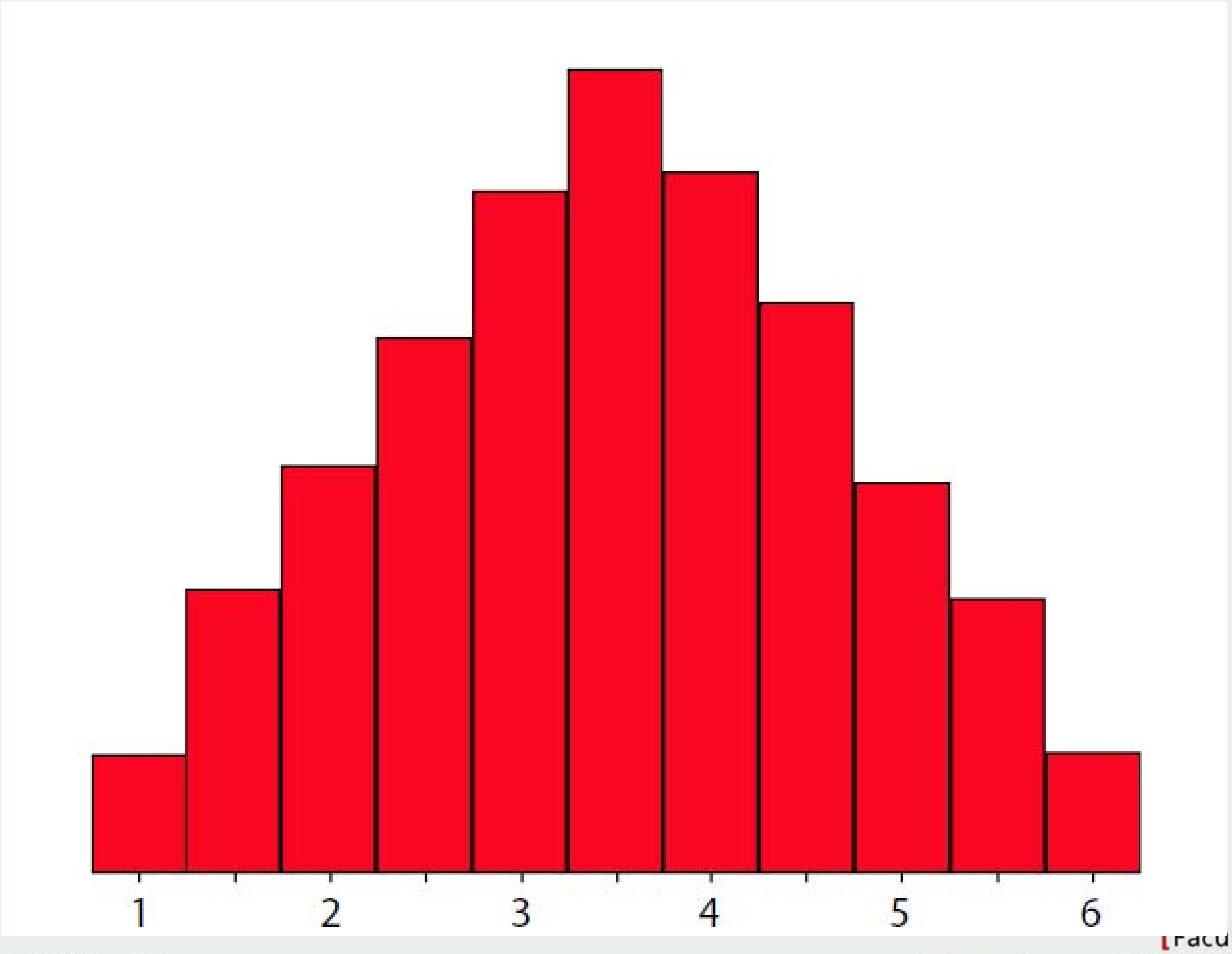
Histogram of throwing a die



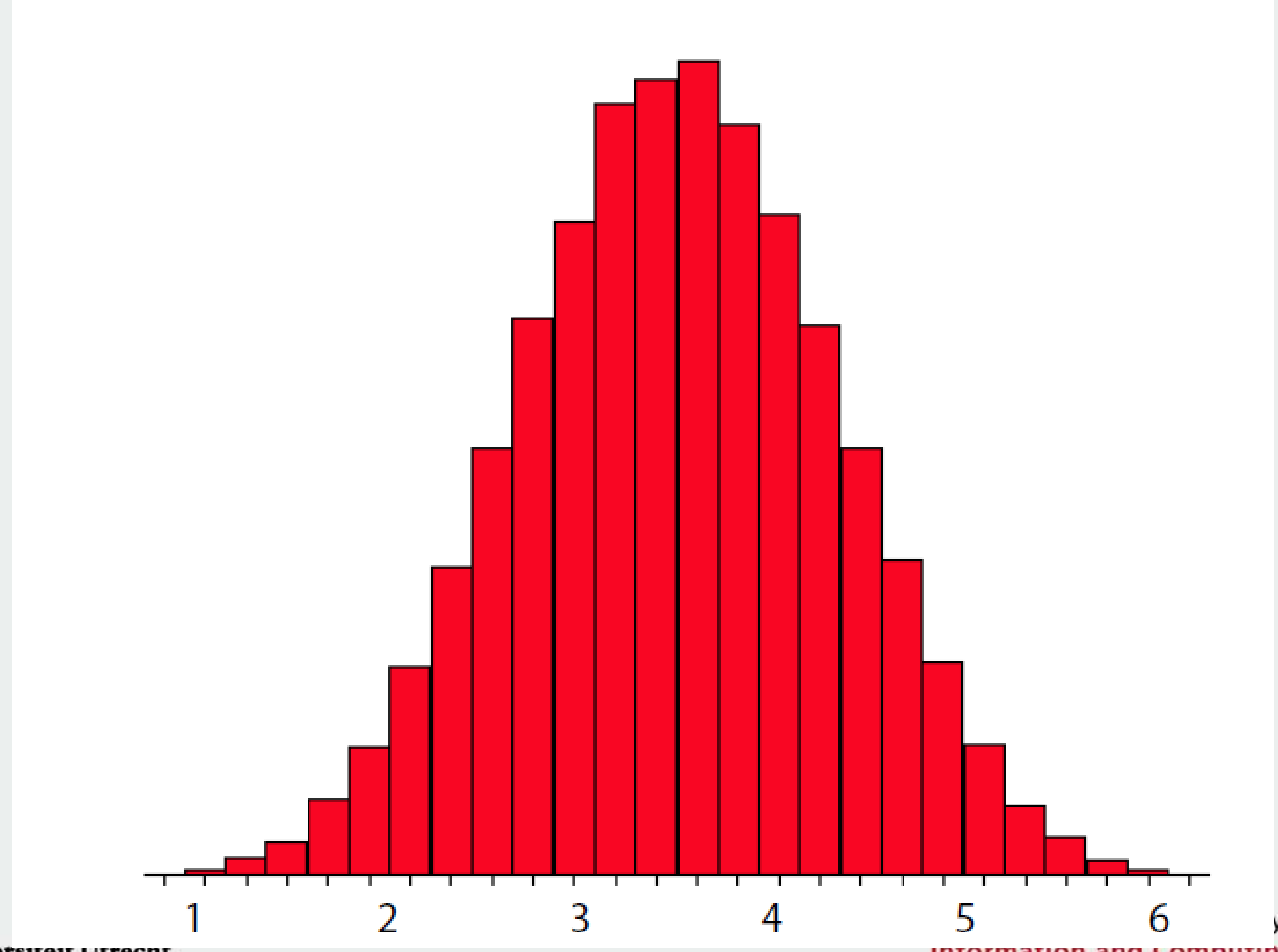
Universiteit Utrecht

[Faculty of Science
Information and Computing Sciences]

Historgram of throwing 2 dice, average nr. of eyes



Historgram of throwing 5 dice, average nr. of eyes



Central limit theorem

- $X_{1,},\ldots,X_n$ Independent Identically Distributed stochastic variables, average μ, variance σ^2

$$Z_n = \frac{\overline{X}(n) - \mu}{\sqrt{\sigma^2 / n}} = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}}$$

if n $\rightarrow\infty$ then Z_n normally distributed $N(0,1)$

Confidence interval: idea

- Amount from coffee machine have variance
  $\sigma^2 = 4$.
- Samples: 170, 171, 171, 172, 173, 175, 175, 176, 178, 179
- Find an interval for the real average μ using central limit theorem
- Such an interval is called a **confidence interval**
- What if σ^2 is unknown?

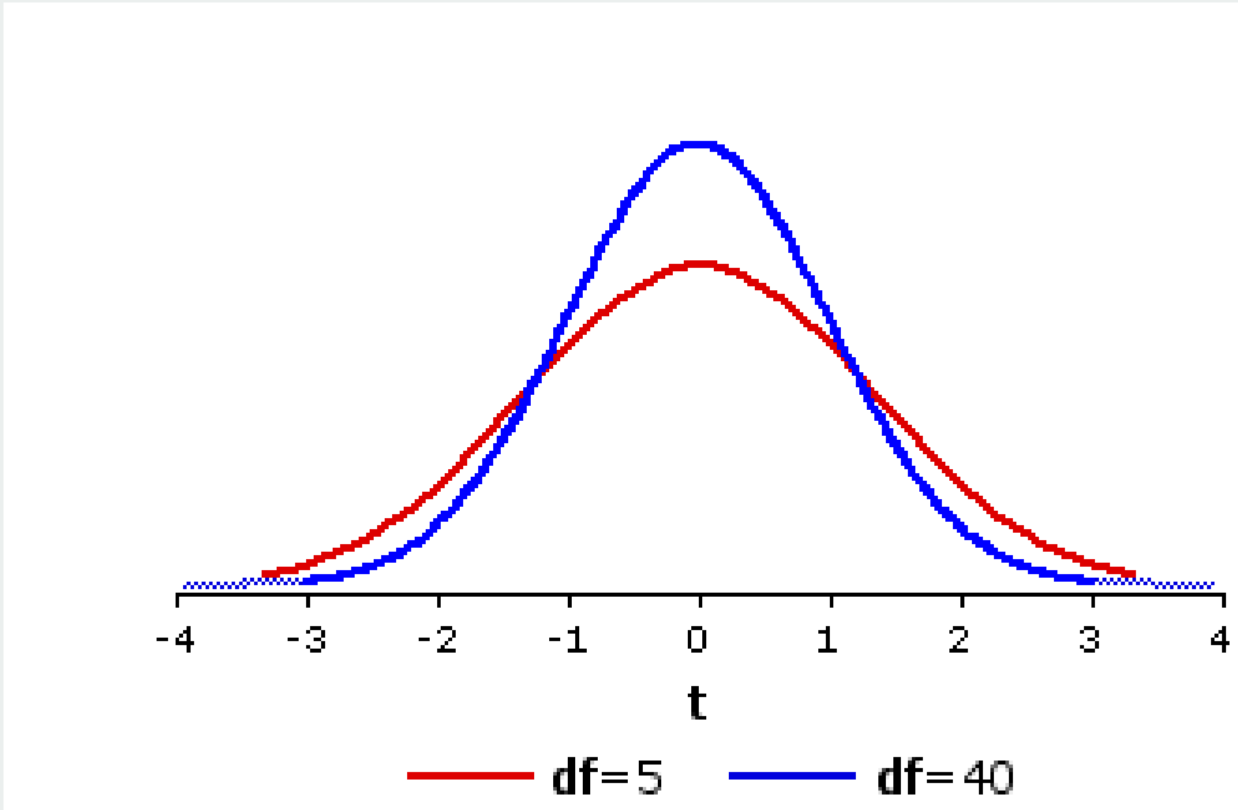
Confidence interval

- $X_1,\dots,X_n$ IID stochastic variables

$$t_n = \frac{\overline{X}(n) - \mu}{\sqrt{S^2(n)/n}}$$

- Follows student's t-distribution with n-1 degrees of freedom
- Note σ^2 replaced by estimate
- Assumption (not too strict): X_i are normally distributed

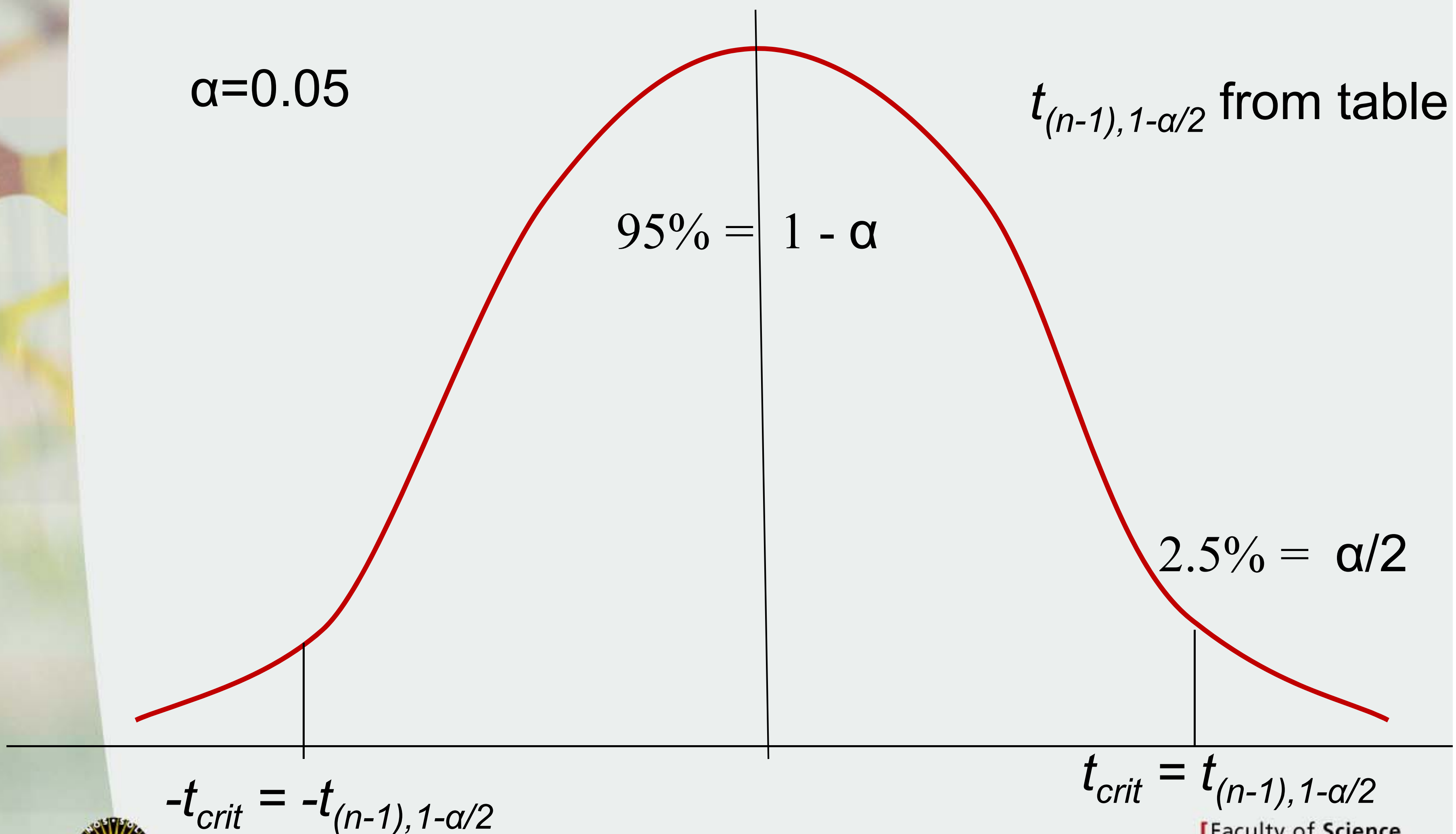
Student's t-distribution



Statistical table:

- book of Law, copy on course website
- http://www.statsoft.com/textbook/distribution-tables/#t

Student’s t-distribution



$$P(-t_{n-1,1-\alpha/2} \leq t_{n-1} \leq t_{n-1,1-\alpha/2}) \approx 1-\alpha$$

$$P\left(-t_{n-1,1-\alpha/2} \leq \frac{\bar{X}(n)-\mu}{\sqrt{S^2(n)/n}} \leq t_{n-1,1-\alpha/2}\right) \approx 1-\alpha$$

$$P\left(\bar{X}(n) - t_{n-1,1-\alpha/2}\sqrt{S^2(n)/n} \leq \mu \leq \bar{X}(n) + t_{n-1,1-\alpha/2}\sqrt{S^2(n)/n}\right) \approx 1-\alpha$$

Piared-t (1-α)100 % confidence interval

$$\left[\overline{X}(n) - t_{n-1,1-\alpha/2}\sqrt{\frac{S^2(n)}{n}}, \overline{X}(n) + t_{n-1,1-\alpha/2}\sqrt{\frac{S^2(n)}{n}}\right]$$

Confidence interval: example

- How many hours do computer science students spend on gaming?
- Sample: 18, 25, 28, 21, 23, 18, 18, 26, 25, 21
- 95% confidence interval?

$$\bar{X} = 22.3$$

$$t(9)_{0.025} = 2.262 \qquad S^2(10) = 13.34$$

95% confidence interval:

$$[22.3 - 2.262\sqrt{\tfrac{13.34}{10}}, 22.3 + 2.262\sqrt{\tfrac{13.34}{10}}] = [19.69, 24.91]$$

This means that with 95 % probabilitity the average number of hours that cs students spend on gaming is within the interval [19.69,24.91]

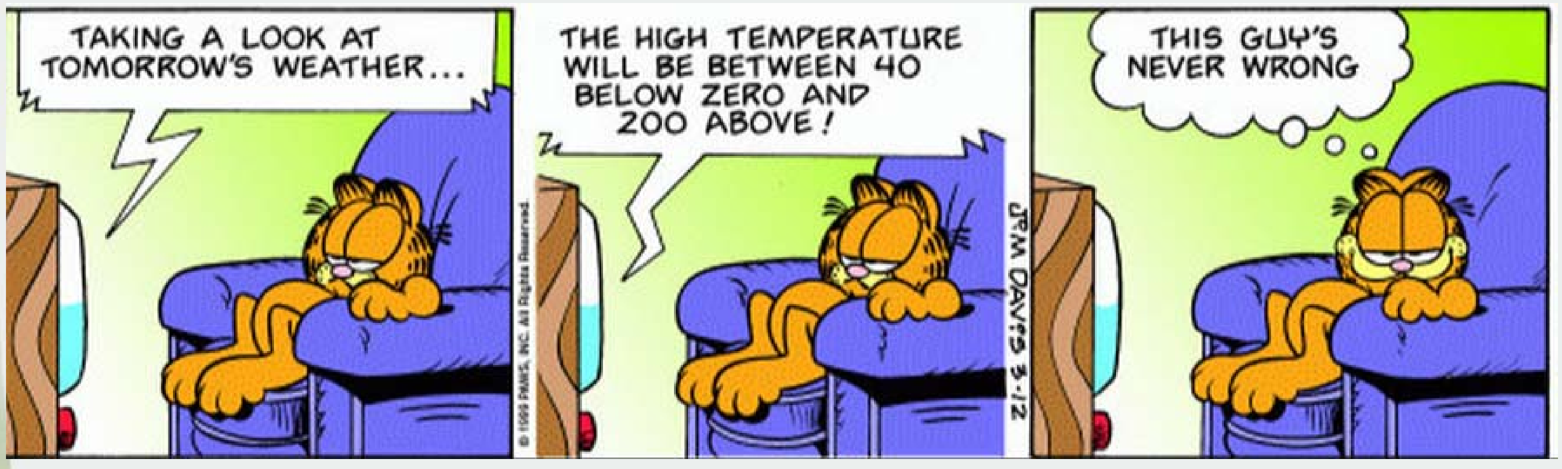
Paired-t Confidence interval (2)

$$\left[\overline{X}(n) - t_{n-1,1-\alpha/2}\sqrt{\frac{S^2(n)}{n}}, \overline{X}(n) + t_{n-1,1-\alpha/2}\sqrt{\frac{S^2(n)}{n}}\right]$$

What does this mean?

(1-α)100 % confidence, μ is in the interval with probability 1-α,

$t_{n-1,1-\alpha/2}$ converges to $z_{1-\alpha/2}$ for large n



End some statistics in general

Stochastic variables occur in simulation at different places:

1. Input data are modeled as stochastic variables
   - E.g time until arrival of next customer
2. Generate random variables
   - When you schedule a new Arrival event you have to generate a random number for the time delay
3. Analysis of results

This lecture

Universiteit Utrecht

[Faculty of Science
Information and Computing Sciences]

Types of simulation w.r.t. output analysis

- Terminating:
  - Endpoint of simulation run is defined by your model.
- Non-terminating

- Examples?
- Terminating or non-terminating for the same system?

Universiteit Utrecht

[Faculty of **Science**
Information and Computing Sciences]

Types of simulation w.r.t. output analysis: examples

- Terminating:
  - Bank open 9AM to 5PM, ends after departure last customer
  - Production line, time to produce 1000 aircraft

- Non-terminating
  - Continuous production line
  - Helpdesk for internet provider (if 24/7)
  - Emergency department

Analysis: first consider terminating simulation

- X_i output result, value of a certain performance measure, of simulation run *i*

  Example
  - Simulation of the orthopedia policlinic department in a hospital.
  - X_{ik} is waiting time of patient *k* in run *i.*
  - X_i average waiting time in run *i*

$$\text{run } 1: X_{11}, X_{12}, \ldots, X_{1j}, \ldots \qquad \text{avg} = X_1$$

$$\text{run } 2: X_{21}, X_{22}, \ldots, X_{2j}, \ldots \qquad \text{avg} = X_2$$

$$\vdots$$

$$\text{run n}: X_{n1}, X_{n2}, \ldots, X_{nj}, \ldots \qquad \text{avg} = X_n$$

Terminating simulation

- The X_i 's can be considered as Independent Identically Distributed (IID) stochastic variables
- We want to find the value μ = E(X)
  - the orthopedia policlinic department in a hospital.
  - *X* average waiting time in a simulation run
  - *What is the expected value of X?*

- Statistical theory from previous slides applies

Terminating simulation (3)

Estimate for average $\overline{X}(n) = \frac{1}{n}\sum_{i=1}^{n} X_i$

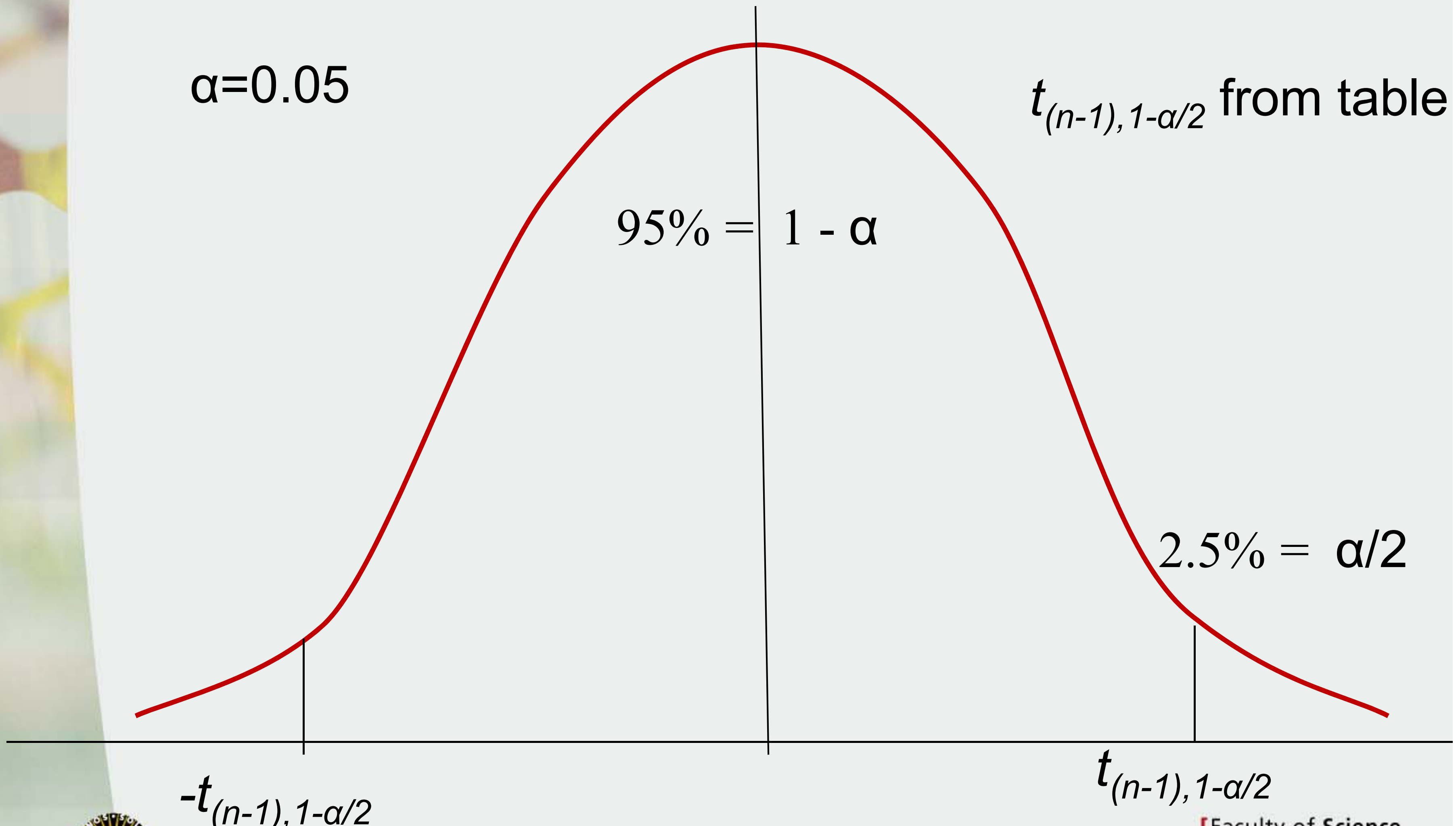
Sample variance $S^2(n) = \dfrac{\sum_{i=1}^{n}(X_i - \overline{X}(n))^2}{n-1}$

(1-α)100 % confidence interval;
μ=E(X) is in the interval with probability 1-α:

$$\left[\overline{X}(n) - t_{n-1,1-\alpha/2}\sqrt{\frac{S^2(n)}{n}}, \overline{X}(n) + t_{n-1,1-\alpha/2}\sqrt{\frac{S^2(n)}{n}}\right]$$

So from the simulation results X_1, X_2, …we can conclude that with probability $(1-\alpha)$ the average of the measure *X* is in the above interval

Student's t-distribution



Non-terminating simulation: Steady state (example)

- M|M|1 queue (single server queue with exponential inter arrival and service times) and ρ=0.9
- D_i: waiting time of customer i
- s number of customers present at time 0

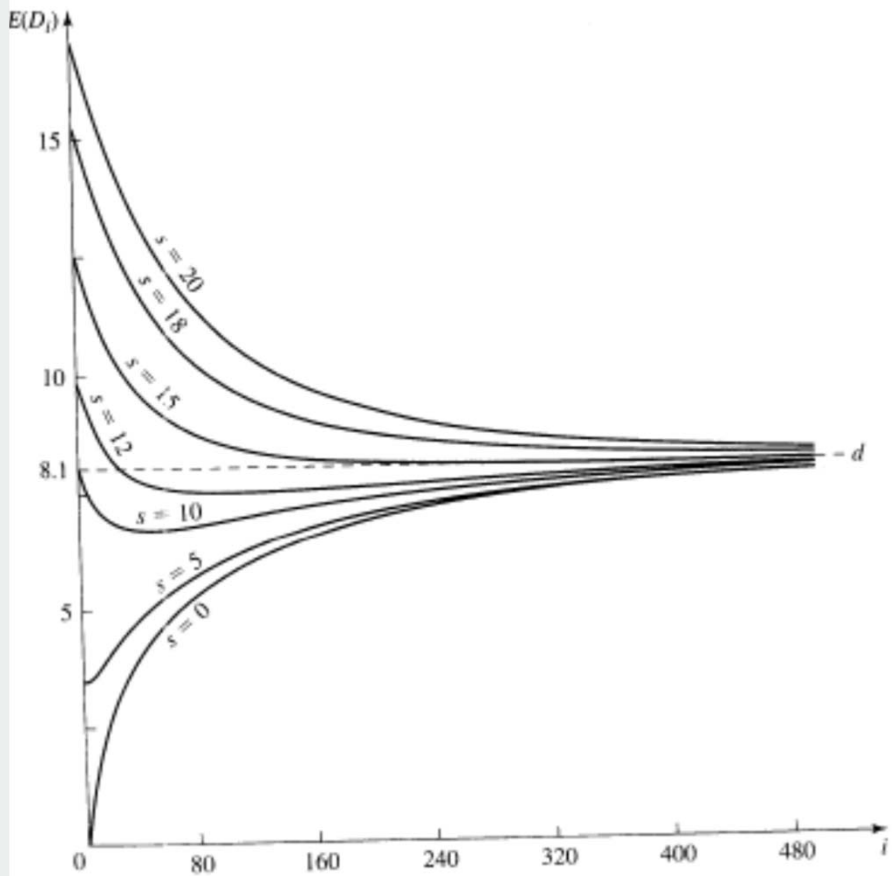


FIGURE 9.2
$E(D_i)$ as a function of i and the number in system at time 0, s, for the $M/M/1$ queue with $\rho = 0.9$

Universiteit Utrecht

[Faculty of Science Information and Computing Sciences]

Steady state

Y_i : i-th realization of performance measure Y within
a simulation run (e.g. the waiting time of the i-th customer)
I: initial conditions

Consider the conditional distribution function of Y_i given I
The simulation converges to a steady state if:

$$F_i(y \mid I) = P(Y_i \leq y \mid I)$$

$$F_i(y \mid I) \xrightarrow{i \to \infty} F(y) \text{ for all } y, I$$

In the steady state the *probability distribution* of *Y* is constant and independent from the initial conditions *I*
NB: Y itself is in general not constant

Steady state

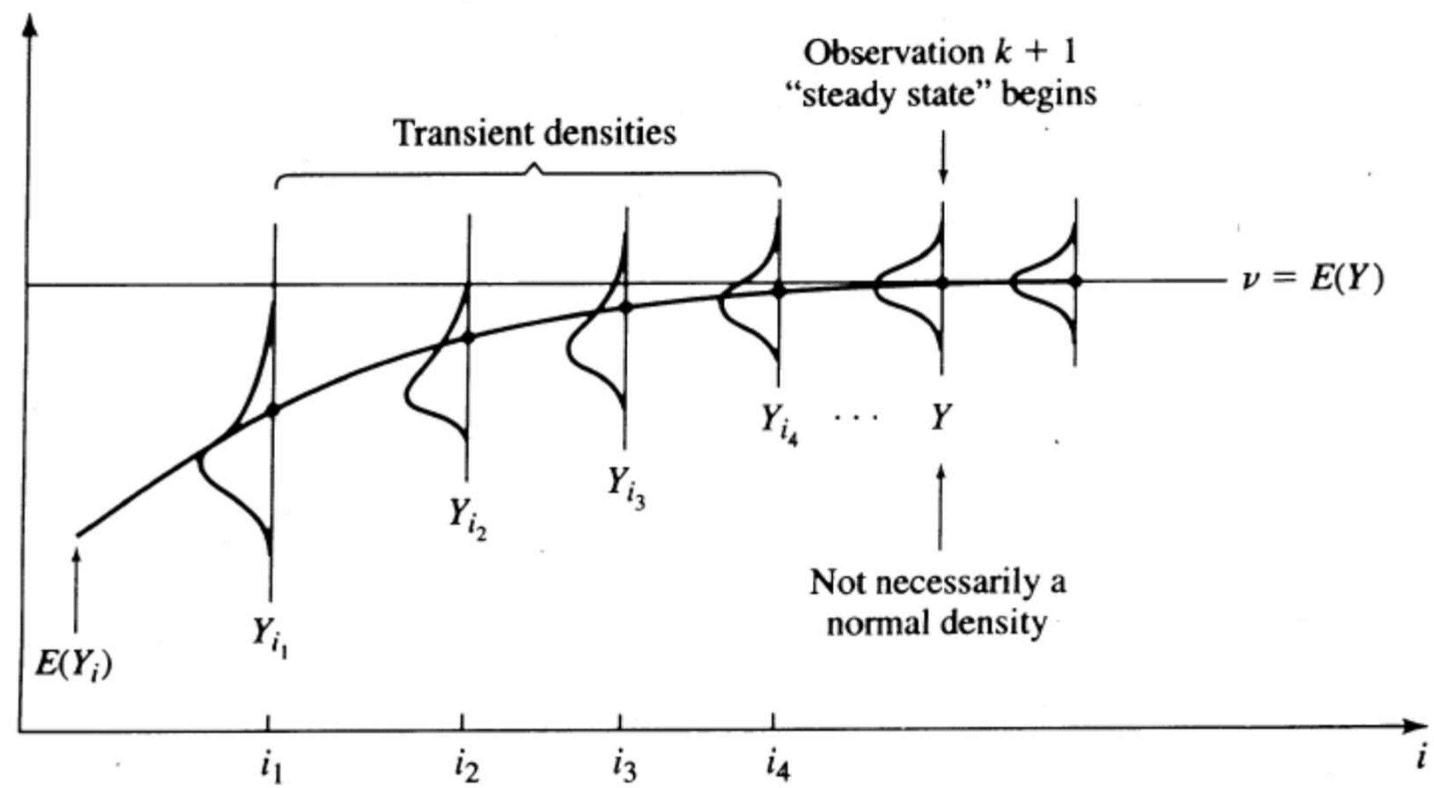


FIGURE 9.1
Transient and steady-state density functions for a particular stochastic process $Y_1, Y_2, \ldots$ and initial conditions I.

Properties of steady state formulations

- The theoretical utilization degree is less than 100%
- All input distributions are constant

- Examples steady state:
  - Helpdesk with interarrival times exp(3 mins), service times exp(2 mins)
- Examples without steady state:
  - Helpdesk with interarrival times exp(2 mins), service times exp(3 mins)
  - Uithoflijn

Question

- Suppose you consider the Uithoflijn simulation as non-terminating and we use runs of length 3 months.
- Does it have a steady state? Explain.
  - No, passenger arrival rates vary
- If not, how could you change the model to obtain a steady state simulation?
  - X_{ij} : average passenger waiting time on day *j* in run *i*
  - X_{ij} : average passenger waiting time during daily peak hour 8:00-9:00 on day *j* in run *i*

If you do not have a steady state you might have:

- Time axis can be divided into time interval *cycles.*
  - One week in call center
  - One week in an emergency department
  - The daily peak hour 8:00-9:00 for the Uithoflijn
- Y^C_i random variable on *i-th* cycle
  - e.g. number of calls with a waiting time longer than 15 minutes in week *i* a call center
- $Y^C_1 Y^C_2 Y^C_3$ ... has a steady state distribution F^C.
- **Steady cycle**

Non-terminating simulation

- Steady state
- Steady cycle
- Others

Non terminating simulation

- We assume a steady state or steady cycle

Methods

- Separate runs
- Batch means

Non-terminating simulation

- Replication/deletion approach, i.e. separate runs of length K:
  - Initialization effect i.e. warm-up period (K_0)
  - Either very large runs or
  - Known confidence interval from:

$$X_i = \frac{\sum_{j=K_0+1}^{K} X_{i,j}}{K - K_0}$$

  - Where X_{ij} is the *j-th* observation in run *i*

Non-terminating simulation (2)

- Batch means method (sub runs):
  - Correlation
  - X_j is observation j
  - Either very large runs or
  - Assume
    - Covariance stationary: $\text{cov}(X_j, X_{j+k})$ independent of j
    - Weak independence: $\text{cov}(X_j, X_{j+n}) \to 0 \ (\text{n} \to \infty)$
  - Confidence interval from

$$Y_i = \frac{\sum_{j=1}^{K} X_{(i-1)K+j}}{K}$$

Non-terminating simulation (3)

$$\frac{S^2(n)}{n} \text{ is replaced by } \frac{S^2(n)}{n} + 2\frac{C(n)}{n^2}$$

$$\text{with } C(n) = \sum_{i=1}^{n-1} (Y_i - \bar{Y})(Y_{i+1} - \bar{Y})$$

$$\left[\bar{Y}(n) - t_{n-1,1-\alpha/2}\sqrt{\frac{S^2(n)}{n} + \frac{2C(n)}{n^2}}, \bar{Y}(n) + t_{n-1,1-\alpha/2}\sqrt{\frac{S^2(n)}{n} + \frac{2C(n)}{n^2}}\right]$$

Warm up period ->Steady state

$$\text{run } 1: X_{11}, X_{12}, \ldots, X_{1j}, \ldots$$

$$\text{run } 2: X_{21}, X_{22}, \ldots, X_{2j}, \ldots$$

$$\vdots$$

$$\underline{\text{run } n: X_{n1}, X_{n2}, \ldots, \quad X_{nj}, \ldots}$$

$$\text{average }: \overline{X}_1, \overline{X}_2, \ldots, \quad \overline{X}_n, \ldots$$

Warm up period ->Steady state

- X_{ij} *j-th* observation in run *i*

$$\overline{X}_j = \frac{\sum_{i=1}^{n} X_{ij}}{n}$$

- Moving average should converge

$$\overline{X}_j(w) = \frac{\sum_{s=-w}^{w} \overline{X}_{j+s}}{2w+1}$$

Warm-up period: example

- Exponential interarrival times with mean 1 minute
- Machine processing times uniform [0.65,0.7] minutes
- Inspection times uniform [0.75,0.8] minutes
- Machine: lifetime exp(6 hours) and repair times uniform 8 to 12 minutes.

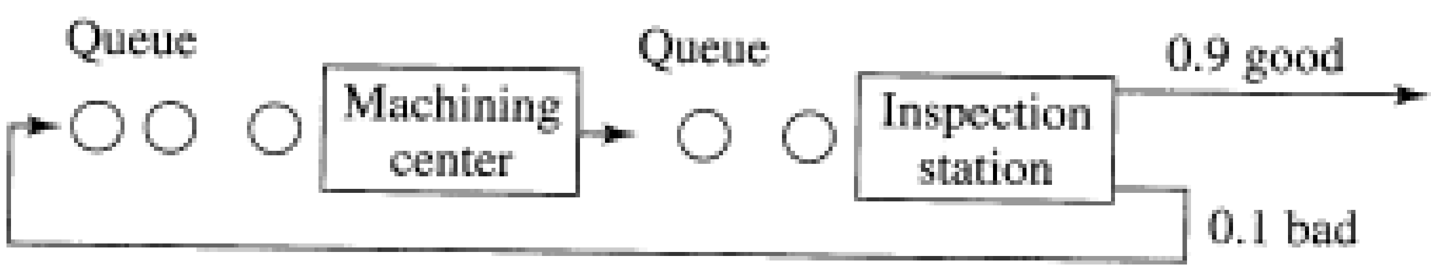


FIGURE 9.9
Small factory consisting of a machining center and an inspection station.

N_i production in hour i

$\overline{N}_i$ Average over 10 runs:

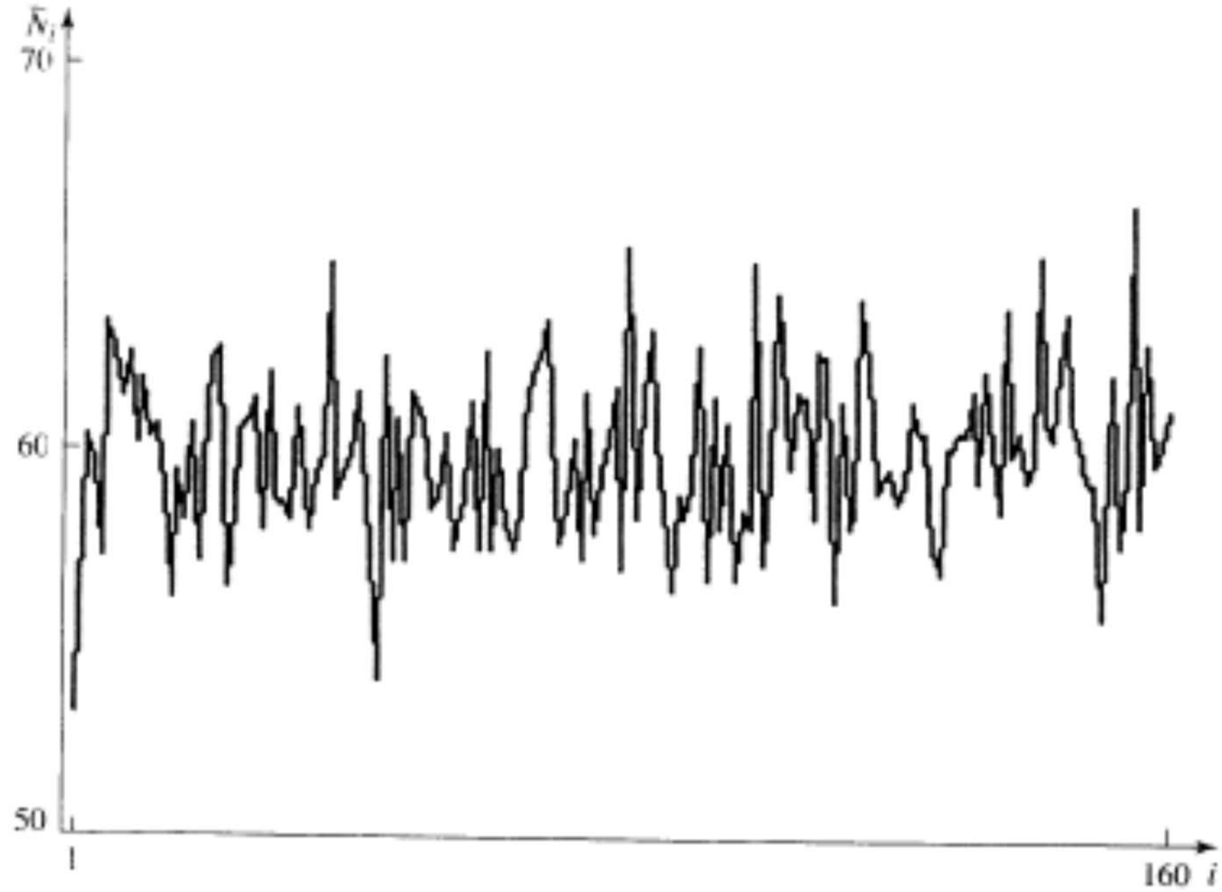


FIGURE 9.10
Averaged process for hourly throughputs, small factory.

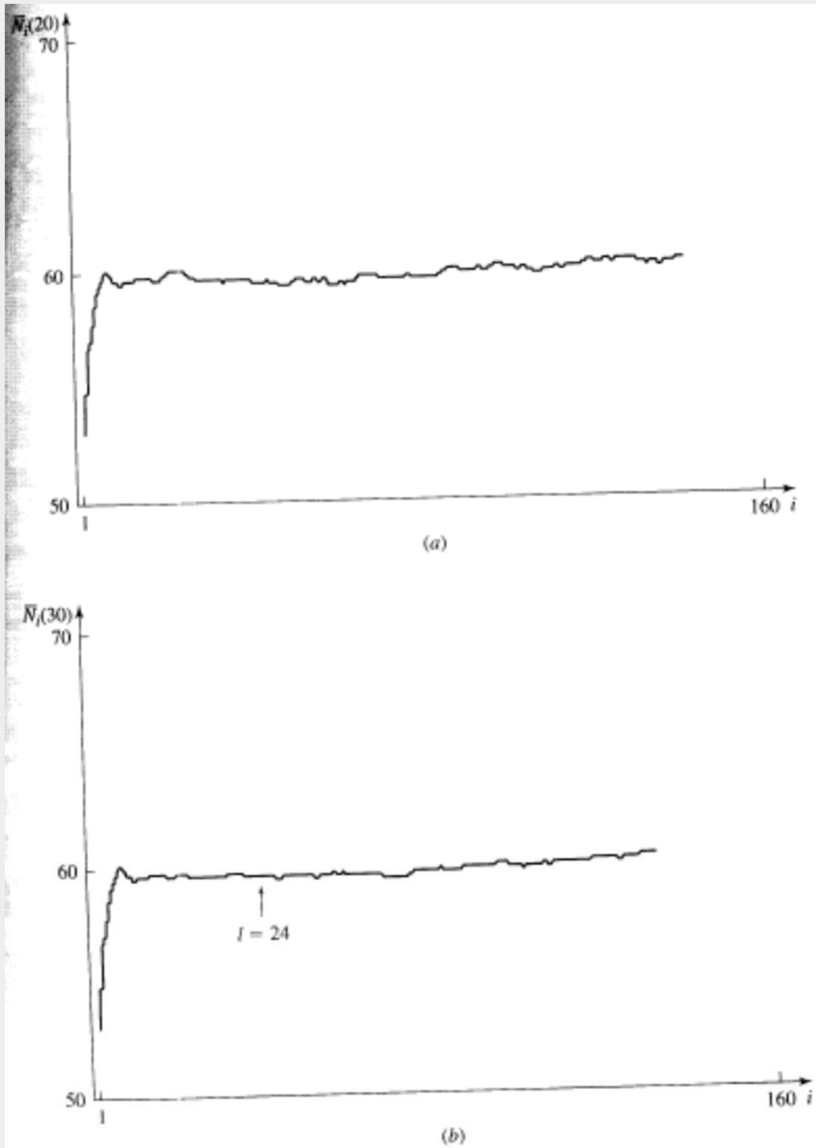


FIGURE 9.11
Moving averages for hourly throughputs, small factory: (a) $w = 20$; (b) $w = 30$.

$\overline{N}_i(20)$

$\overline{N}_i(30)$

Universiteit

[Faculty of **Science**
ormation and Computing Sciences]

Multiple measures of performance

- k performance measures, $s=1,2,\ldots,k$
- I_s confidence interval for μ_s with confidence level $1-\alpha_s$
- Then (Bonferroni inequality)

$$P(\mu_s \in I_s \text{ for all } s = 1,2,\ldots,k) \geq 1 - \sum_{s=1}^{k} \alpha_k$$

- Example:
  - $k=2$, performance measures
    - average waiting time
    - busy factor of server
  - If we have 95% confidence intervals for each single measures we overall have 90% confidence

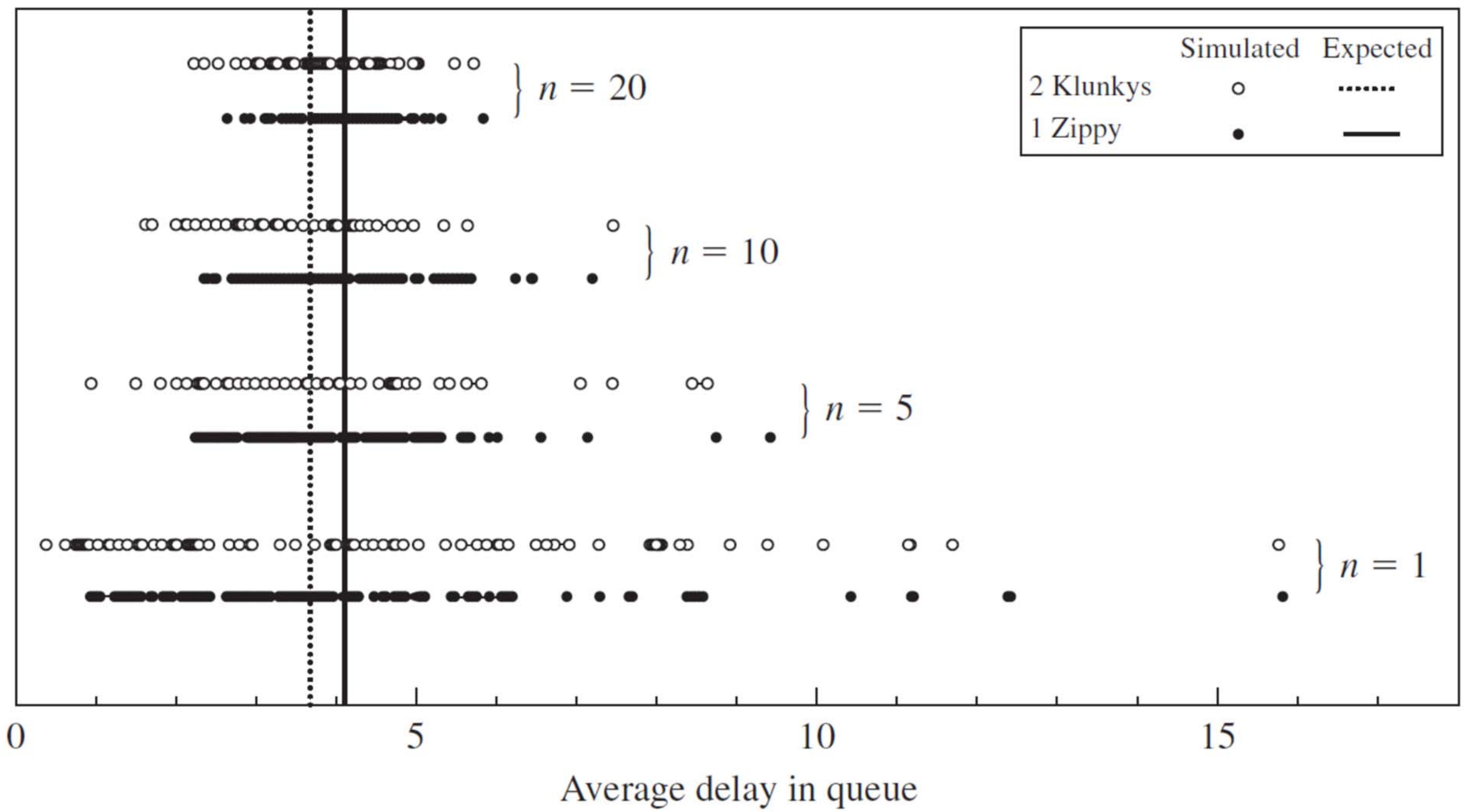
Comparing systems: example

- Arrival: Poisson 1 per minute
- Two types of ATM's:
  - Zippy: service time exp(0.9 min)
  - Klunky: service time exp(1.8 min)
- One Zippy or 2 Klunkies?
  - Cost are equal
  - Average customer waiting time matters
  - Since waiting is more annoying then being served, we consider pure waiting time and exclude service time

Comparing systems (2)

- Zippy: X_{1j} $(j=1,..,n)$ average delay run j
- 2 Klunkies: X_{2j} $(j=1,..,n)$ average delay run j
- Compare: perform the following experiment **100** times to collect 100 votes:
  - Compare average delay of **n** runs with Zippy to average delay of **n** runs with 2 Klunkies. **Vote** (Zippy/2Klunkies)

n (# runs)	% Zippy
1	52
5	43
10	38
20	34



Simulated
Expected
2 Klunkys
1 Zippy
n = 20
n = 10
n = 5
n = 1
0
5
10
15
Average delay in queue

Comparing two systems: use paired t-confidence interval

$Z_j = X_{1j} - X_{2j}$ $\mu = \mathrm{E(Z)}$

Assume X_{1j} and X_{2j} follow normal distribution.

$$\bar{Z}(n) = \frac{\sum_{j=1}^{n} Z_j}{n} \qquad S_Z^{\,2}(n) = \frac{\sum_{j=1}^{n} [Z_j - \bar{Z}(n)]^2}{n-1}$$

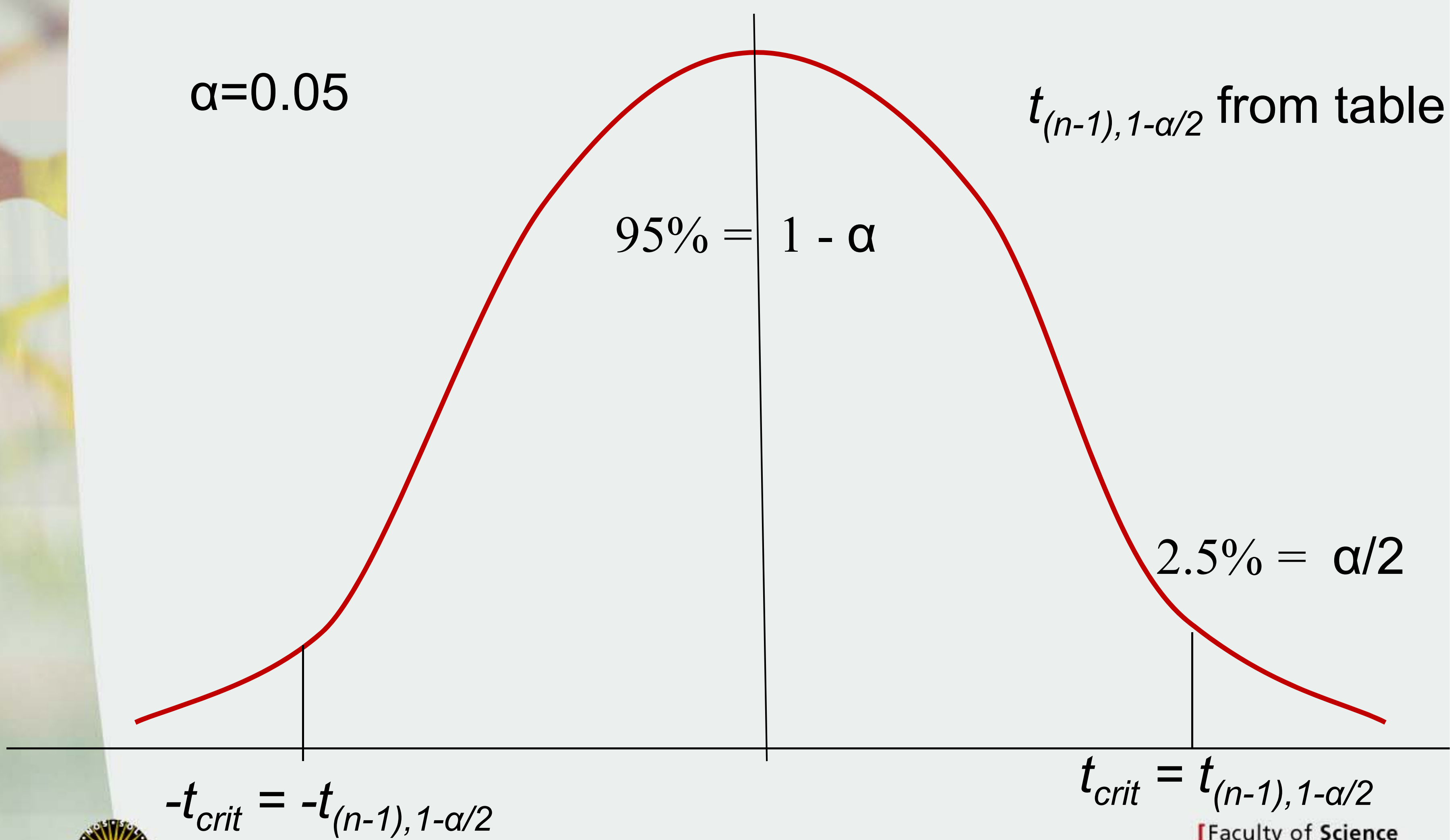
$\dfrac{\bar{Z}(n) - \mu}{\sqrt{\dfrac{S_Z^{\,2}(n)}{n}}}$ follows t-distribution with $n-1$ df

Confidence interval: $\bar{Z}(n) \pm t_{n-1,1-\alpha/2}\sqrt{\dfrac{S_Z^{\,2}(n)}{n}}$

Universiteit Utrecht

[Faculty of **Science**
Information and Computing Sciences]

Student's t-distribution



Universiteit Utrecht

[Faculty of Science
Information and Computing Sciences]

Comparing two systems: use paired t-confidence interval

$$Z_j = X_{1j} - X_{2j}$$

$$\text{confidence interval}: \overline{Z}(n) \pm t_{n-1,1-\alpha/2}\sqrt{\frac{S_Z^{\,2}(n)}{n}}$$

If 0 in confidence interval, no significant difference

$$\text{If left side of interval } \overline{Z}(n) - t_{n-1,1-\alpha/2}\sqrt{\frac{S_Z^{\,2}(n)}{n}} > 0, \text{ then } X_1 \text{ larger than } X_2$$

$$\text{If right side interval } \overline{Z}(n) + t_{n-1,1-\alpha/2}\sqrt{\frac{S_Z^{\,2}(n)}{n}} < 0 \text{ then } X_1 \text{ smaller than } X_2$$

Background: Paired t-test (two-sided)

Used in course Evolutionary Computing

$$Z_j = X_{1j} - X_{2j}$$

$$\mu = E(Z)$$

$$H_0 : \mu = 0$$

$$H_1 : \mu \neq 0$$

$$\bar{Z}(n) = \frac{\sum_{j=1}^{n} Z_j}{n}$$

$$S_Z^{\,2}(n) = \frac{\sum_{j=1}^{n} [Z_j - \bar{Z}(n)]^2}{n-1}$$

Paired t-test

If H_0 is true:

$$t_{obs} = \frac{\overline{Z}(n)}{\sqrt{\frac{S_Z^{\,2}(n)}{n}}}$$ follows a t-distribution with $n-1$ df

We want confidence level $1-\alpha$

So we accept H_0 when $-t_{n-1,1-\alpha/2} \le t_{obs} \le t_{n-1,1-\alpha/2}$

and reject otherwise.

Relation to paired-*t* confidence interval

Accept H$_0$: $\mu = E(Z) = 0$

if and only if

$$-t_{n-1,1-\alpha/2} \leq \frac{\bar{Z}(n)}{\sqrt{\frac{S_Z{}^2(n)}{n}}} \leq t_{n-1,1-\alpha/2}$$

if and only if

0 is in the paired t-confidence interval

A confidence interval gives more information

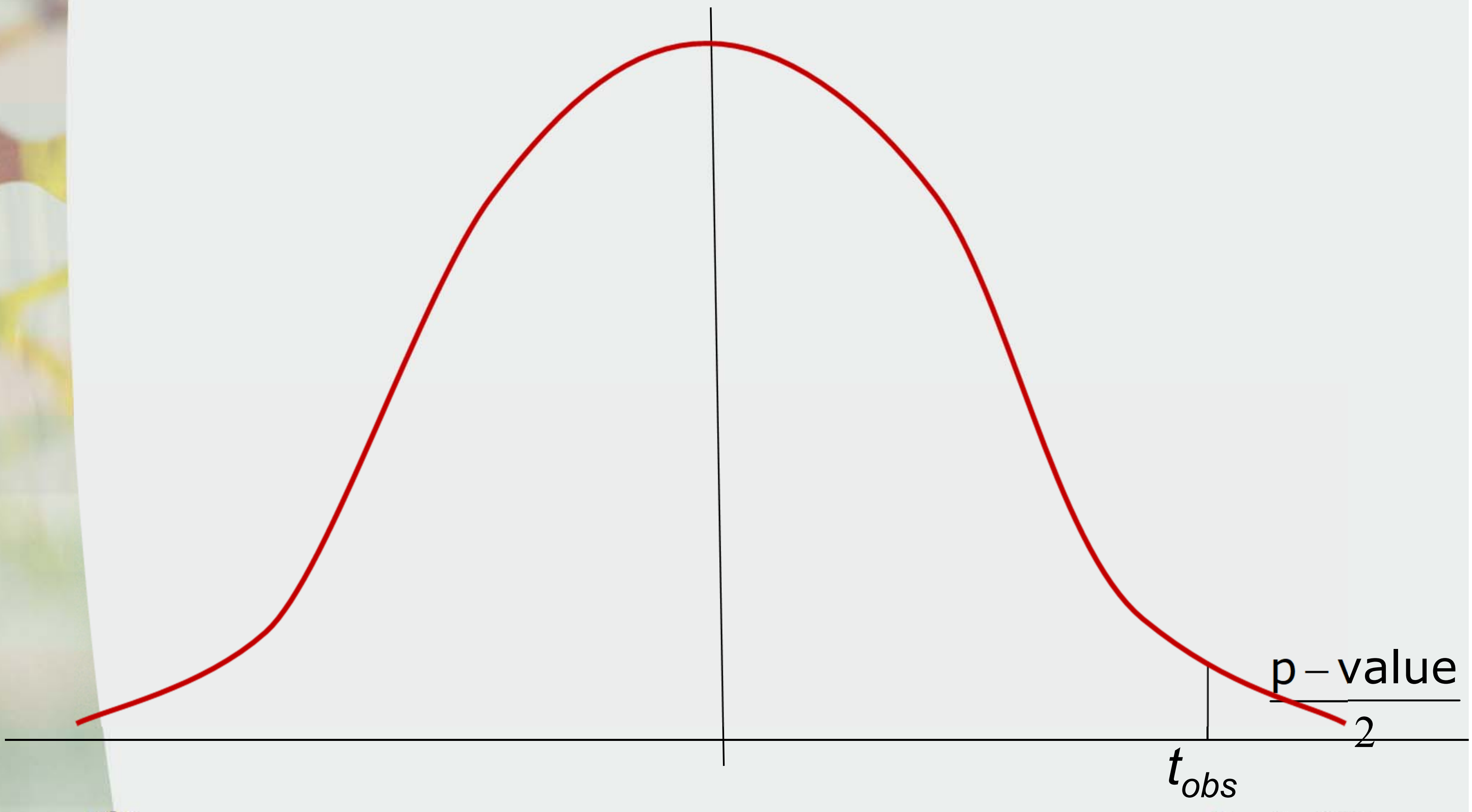
Paired t-test: p-value

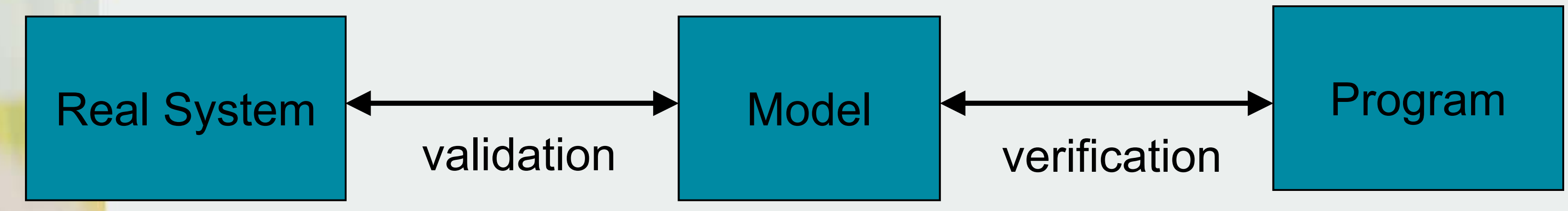
- p-value (or significance):
  - Indicates: how extreme is t_{obs} ?

$$p\text{-value} = 2\min(P(T \geq t_{obs}), P(T \leq t_{obs})),$$
$$\text{where T follows a t-distribution with } n-1 \text{ df.}$$

$$\text{We reject } H_0 \text{ if } p < 0.05$$

Student's t-distribution





Real System
validation
Model
verification
Program

Validate output

μ_s : system

μ_m : model

$\hat{\mu}_m$: result of simulation

$$|\hat{\mu}_m - \mu_s| \leq |\hat{\mu}_m - \mu_m| + |\mu_m - \mu_s|$$

$|\hat{\mu}_m - \mu_m|$: good experimentation

$|\mu_m - \mu_s|$: validation

Wrap-up example

- Service desk of CoolGreen is opened 8AM-9PM
- 5 simulation runs to measure average waiting time

run	Avg waiting time
X_1	10
X_2	7
X_3	9
X_4	12
X_5	12

- Find 95% confidence interval?

Recall: Terminating simulations

Estimate for average $\overline{X}(n) = \frac{1}{n}\sum_{i=1}^{n} X_i$

Sample variance $S^2(n) = \frac{\sum_{i=1}^{n}(X_i - \overline{X}(n))^2}{n-1}$

(1-α)100 % confidence interval;
μ=E(X) is in the interval with probability 1-α:

$$\left[\overline{X}(n) - t_{n-1,1-\alpha/2}\sqrt{\frac{S^2(n)}{n}}, \overline{X}(n) + t_{n-1,1-\alpha/2}\sqrt{\frac{S^2(n)}{n}}\right]$$

So from the simulation results X_1, X_2, …we can conclude that with probability $(1-\alpha)$ the average of the measure *X* is in the above interval

Wrap up example

- The estimate of the average $\bar{X}(5) = 10$.
- Sample variance $S^2(5) = \frac{0+9+1+4+4}{4} = \frac{9}{2}$
- From the statistical table http://www.cs.uu.nl/docs/vakken/mads/tabelTandNormalDistribution.pdf you obtain $t_{4\ 0.975} = 2.776$
- The 95% confidence interval is

$$\left[10 - 2.776\sqrt{\frac{9/2}{5}}, 10 + 2.776\sqrt{\frac{9/2}{5}}\right] = [7.366, 12.634]$$

- So from these simulation results we can conclude that with probability 95% the average waiting time is in the interval $[7.366, 12.634]$

Wrap-up example (2)

- CoolGreen considers new computer system.
- Simulation is performed again

Old		**New**	
run	Avg waiting time	run	Avg waiting time
X_{11}	10	X_{21}	6
X_{12}	7	X_{22}	5
X_{13}	9	X_{23}	7
X_{14}	12	X_{24}	10
X_{15}	12	X_{25}	11

- Is the new situation better?

Recall: Comparing two systems: use paired t-confidence interval

$Z_j = X_{1j} - X_{2j}$ $\qquad \mu = \text{E(Z)}$

Assume X_{1j} and X_{2j} follow normal distribution.

$$\bar{Z}(n) = \frac{\sum_{j=1}^{n} Z_j}{n} \qquad S_Z^{\,2}(n) = \frac{\sum_{j=1}^{n} [Z_j - \bar{Z}(n)]^2}{n-1}$$

$\dfrac{\bar{Z}(n) - \mu}{\sqrt{\dfrac{S_Z^{\,2}(n)}{n}}}$ follows t-distribution with $n-1$ df

Confidence interval: $\bar{Z}(n) \pm t_{n-1,1-\alpha/2}\sqrt{\dfrac{S_Z^{\,2}(n)}{n}}$

Wrap up example (2)

- We obtain $Z_1 = 4, Z_2 = 2, Z_3 = 2, Z_4 = 2, Z_5 = 1$
- The estimate of the average $\overline{Z}\,(5) = 2.2.$
- Sample variance $S^2(5) = \frac{1.8^2 + 0.02^2 + 0.02^2 + 0.02^2 + 1.2^2}{4} = \frac{4.8}{4} = 1.2$
- From the statistical table http://www.cs.uu.nl/docs/vakken/mads/tabelTandNormalDistribution.pdf you obtain $t_{4\;0.975} = 2.776$
- The 95% confidence interval is

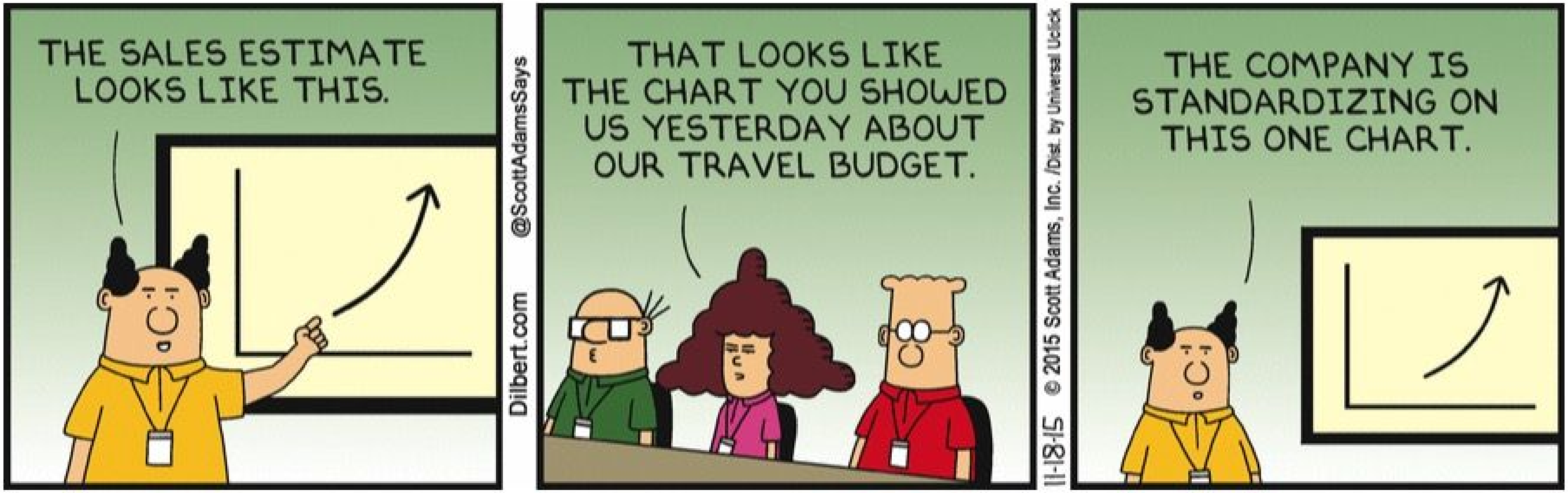
$$\left[2.2 - 2.776\sqrt{\frac{1.2}{5}}, 2.2 + 2.776\sqrt{\frac{1.2}{5}}\right] = [0.840, 3.559]$$

- Since this interval contains only positive values we can conclude with confidence 95% that the average waiting time is scenario 1 is larger

Elements of output analysis required in simulation assignment

- Questions to be answered by the experiments
- Description of the investigated scenarios including all relevant parameter settings and performance measures
- Number of runs
- Tables (at least the most interesting ones)
- Graphs
- Observations from your tables and graphs
- Statistical analysis.
  - The minimum requirement is to find confidence intervals for comparing two different scenarios. You can make a selection of the most interesting combinations (select at least 10).
  - Additional analysis such as Comparisons with a standard', All pairwise comparisons, or Ranking and selection are optional.

There is no standard output analysis



Universiteit Utrecht

[Faculty of **Science**
Information and Computing Sciences]

Learning some Dutch on graphs

Tuurplaatje:

- Picture at which you stare for a long time wondering what is going on

Stuurplaatje:

- Graph that gives you insight

Last lecture:

- My estimate: Some of you found last lecture a bit abstract and difficult

- But you have to learn the material

- What can you do:
  - Self-study of material is a regular course activity
  - Do not hesitate to ask questions:
    - The most stupid question is the one you do not ask
    - You are never the only one with that question. Your fellow students will be grateful to you.

- What I will do to help:
  - Add reading material
  - Add some remarks (small examples) to the slides

Welch confidence interval (modified two sample-t confidence interval)

- n_1 observations X_{1j}, n_2 observations X_{2j}
- Not paired, independent
- If both samples do not have the same variance
- Assume Normally distributed

- Examples
  - CoolGreen has 5 runs for the old situation and 20 simulation runs for the new situation
  - Airport wants to build new runway and want to compare the simulation of extended airport to real world observations. Case 1 gives a few real-worlds observations, e.g. (n_1=5) but the number of simulations for the new situation is larger, e.g. n_2=50

Welch confidence interval

$$\frac{\bar{X}_1(n_1) - \bar{X}_2(n_2) - \mu_{12}}{S_{\bar{X}_1-\bar{X}_2}} \text{ t−distribution with } q \text{ df}$$

$$S_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{S^2_{X_1}}{n_1} + \frac{S^2_{X_2}}{n_2}}$$

confidence interval for $\mu_{12} = E\ (X_1 - X_2)$:

$$\bar{X}_1(n_1) - \bar{X}_2(n_2) \pm t_{q,1-\alpha/2} S_{\bar{X}_1-\bar{X}_2}$$

$$\text{with } q = \frac{[S^2_{X_1}/n_1 + S^2_{X_2}/n_2]^2}{\frac{[S^2_{X_1}/n_1]^2}{n_1 - 1} + \frac{[S^2_{X_2}/n_2]^2}{n_2 - 1}}$$

More than two systems: general idea

Based on the Bonferroni inequality:

- k systems, s=1,2,…,k
- I_s confidence interval for μ_s with confidence level $1-\alpha_s$ (s=1,2,…,k)
- Then

$$P(\mu_s \in I_s \text{ for all } s = 1,2,\dots,k) \geq 1 - \sum_{s=1}^{k} \alpha_k$$

So

If we have *c* confidence interval with confidence level $1-\frac{\alpha}{c}$ their combintation has confidence level $1-\alpha$

Comparison with a Standard

- Let *1* be the standard systems and *2,3,…,k* the other variants.
- Construct *k-1* confidence intervals for

$$\mu_2 - \mu_1,\ \mu_3 - \mu_1, \dots ,\mu_k - \mu_1$$

Or alternatively $\mu_1 - \mu_2,\ \mu_1 - \mu_3, \dots ,\mu_1 - \mu_k$

each with confidence level

$$1 - \frac{\alpha}{k-1}$$

- Overall confidence level $1 - \alpha$
- Example
  - CoolGreen has as other option: dedicated employee. This results in $X_{31}, X_{32}, \dots X_{35}$. If [1.5;6] is 95% confidence interval for $\mu_1 - \mu_3$, then overall confidence is 90%. Note that in the example we computed interval for $\mu_1 - \mu_2$.

TABLE 10.5

Average total cost per month for five independent replications of each of the five inventory policies, with sample means and variances

j	X_{1j}	X_{2j}	X_{3j}	X_{4j}	X_{5j}
1	126.97	118.21	120.77	131.64	141.09
2	124.31	120.22	129.32	137.07	143.86
3	126.68	122.45	120.61	129.91	144.30
4	122.66	122.68	123.65	129.97	141.72
5	127.23	119.40	127.34	131.08	142.61
Mean	125.57	120.59	124.34	131.93	142.72
Variance	4.00	3.76	15.23	8.79	1.87

TABLE 10.6

Individual 97.5 percent confidence intervals for all comparisons with the standard system ($\mu_i - \mu_1, i = 2, 3, 4, 5$); * denotes a significant difference

		Paired-t		Welch	
i	$\bar{X}_i - \bar{X}_1$	Half-length	Interval	Half-length	Interval
2	−4.98	5.45	(−10.44, 0.48)	3.54	(−8.52, −1.44)*
3	−1.23	7.58	(−8.80, 6.34)	6.21	(−7.44, 4.97)
4	6.36	6.08	(0.27, 12.46)*	4.55	(1.82, 10.91)*
5	17.15	3.67	(13.48, 20.81)*	6.15	(14.07, 20.22)*

All pairwise comparisons

- k alternatives
- Construct $\frac{k(k-1)}{2}$ confidence intervals for all pairs

$$\mu_{i_2} - \mu_{i_1}$$

each with confidence level

$$1 - \frac{\alpha}{[k(k-1)]/2}$$

- Overall confidence level $1 - \alpha$

Selecting the best of *k* systems

- Section 10.4.1 of Law (see course website).
- Optional challenge

Variance reduction

- Use common random numbers for X_1 and X_2
- Apply standard (paired-t) confidence interval

Wrap up

- ***Output:*** A simulation determines the value of some performance measures, e.g. production per hour, average queue length etc...
- In this lecture you learned basic statistical principles to analyse the output values of a simulation
- After this lecture you understand:
  - Terminating and non-terminating simulations
  - Steady state
  - Confidence intervals
  - Comparison of different systems