

# Construction Operation Simulation

Lecture #9

## Simulation model verification and validation

Amin Alvanchi, PhD

Construction Engineering and Management

 [LinkedIn](#)

 [Instagram](#)

 [WebPage](#)



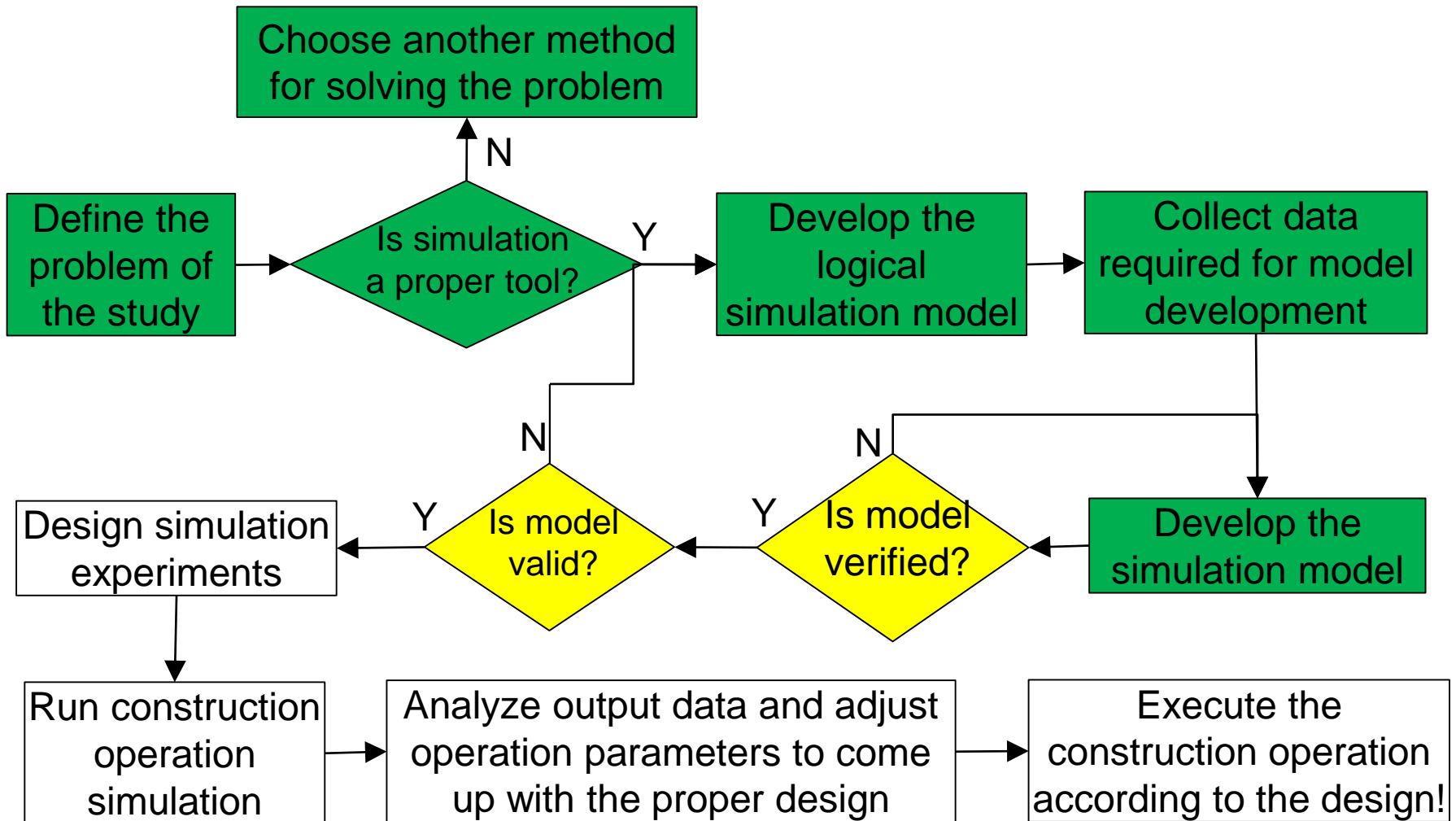
# Outline

2

- Introduction
- Verification of models
- Validation of models

# Introduction

3



# Introduction

4

- Model verification (راستنمایی): Verification is concerned with building the model correctly. Does model represent what we meant?

Could we implement the conceptual model correctly in the simulation software?

- Model validation (اعتبار سنجی): Validation is concerned with building the correct model. Is the developed model credible and valid?

Is the model accurate enough?

Making sure the simplifying assumptions do not have mal-effect on the model output.

# Verification of models

5

- Many common-sense suggestions can be given for use in the verification process:
  1. Have the operational model been checked by someone other than its developer, preferably an expert in the simulation software being used.
  2. Make event flowcharts, and follow the model logic for each activity for each event type.
  3. Closely examine the model output for reasonableness under a variety of settings of the input parameters.
  4. Check value of input parameters at the end of the simulation, to make sure that these parameter values have not been changed inadvertently.

# Verification of models

6

5. Use comments to make the operation model as self-documenting as possible. Give a precise definition of every model element used and a general description of the purpose of each sub-model, procedure (or code executions), component, or other model subdivision.
6. If the operational model is animated, verify that what is seen in the animation imitates the actual system. Examples of errors that can be observed through animation are automated guided vehicles (AGVs) that pass through one another on a unidirectional path or at an intersection and entities that disappear (unintentionally) during a simulation.
7. The Interactive Run Controller (IRC) or debugger is an essential component of successful simulation model building. Even the best simulation analysts makes mistakes or commits logical errors when building a model. The IRC assists in finding and correcting those errors in the following ways:

# Verification of models

7

- (a) The simulation can be monitored as it progresses. This can be accomplished by advancing the simulation until a desired time has elapsed, then displaying model information at that time. Another possibility is to advance the simulation until a particular condition is in effect, and then display information.
- (b) Attention can be focused on a particular entity line, of code, or procedure. For instance, every time that an entity enters a specified procedure, the simulation will pause so that information can be gathered. As another example, every time that a specified entity becomes active, the simulation will pause.
- (c) Values of selected model components can be observed when the simulation has paused, the current value or status of variables, attributes, queues, resources, counters, and so on can be observed.
- (d) The simulation can be temporarily suspended, or paused, not only to view information, but also to reassign values or redirect entities.

# Verification of models

8

8. Graphical interfaces (e.g., AnyLogic modeling environment) are recommended for accomplishing verification and validation. The graphical representation of the model is essentially a form of self-documentation. It simplifies the task of understanding the model.



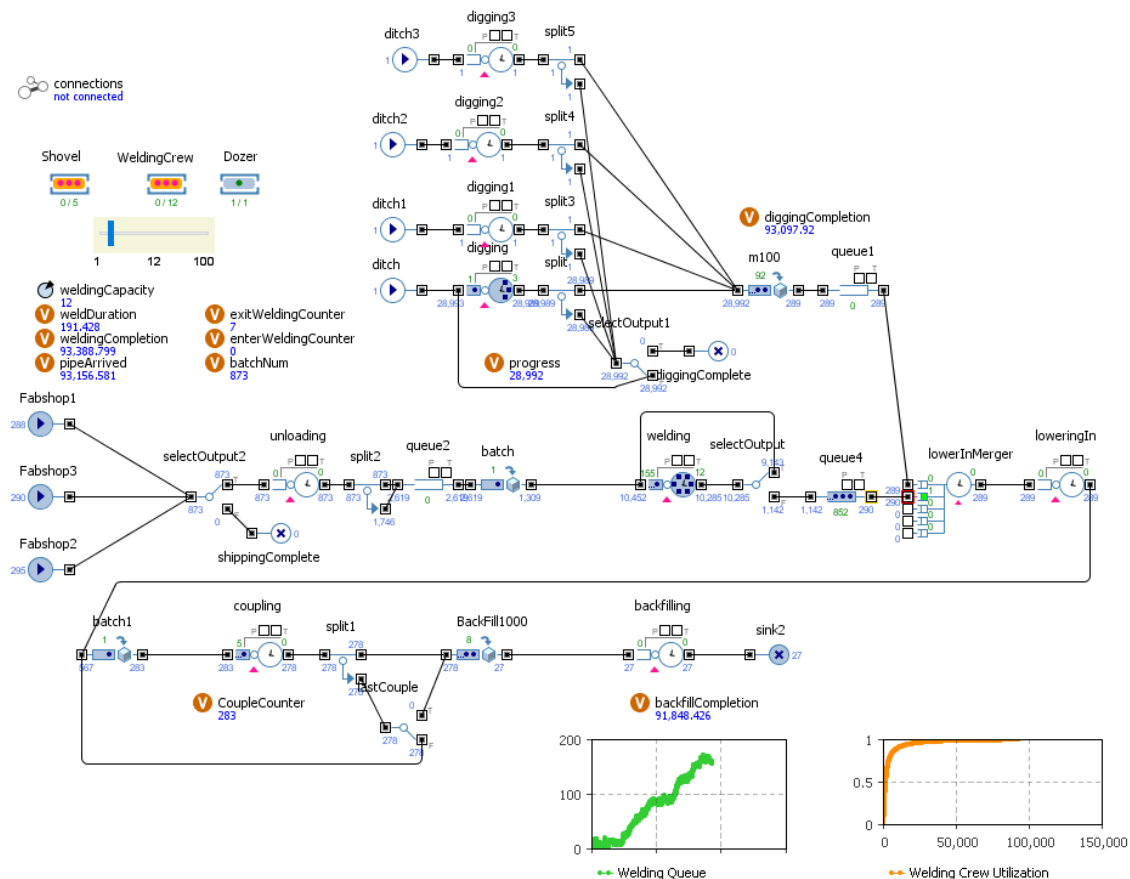
# Verification of models-Example

9

Hands on AnyLogic:



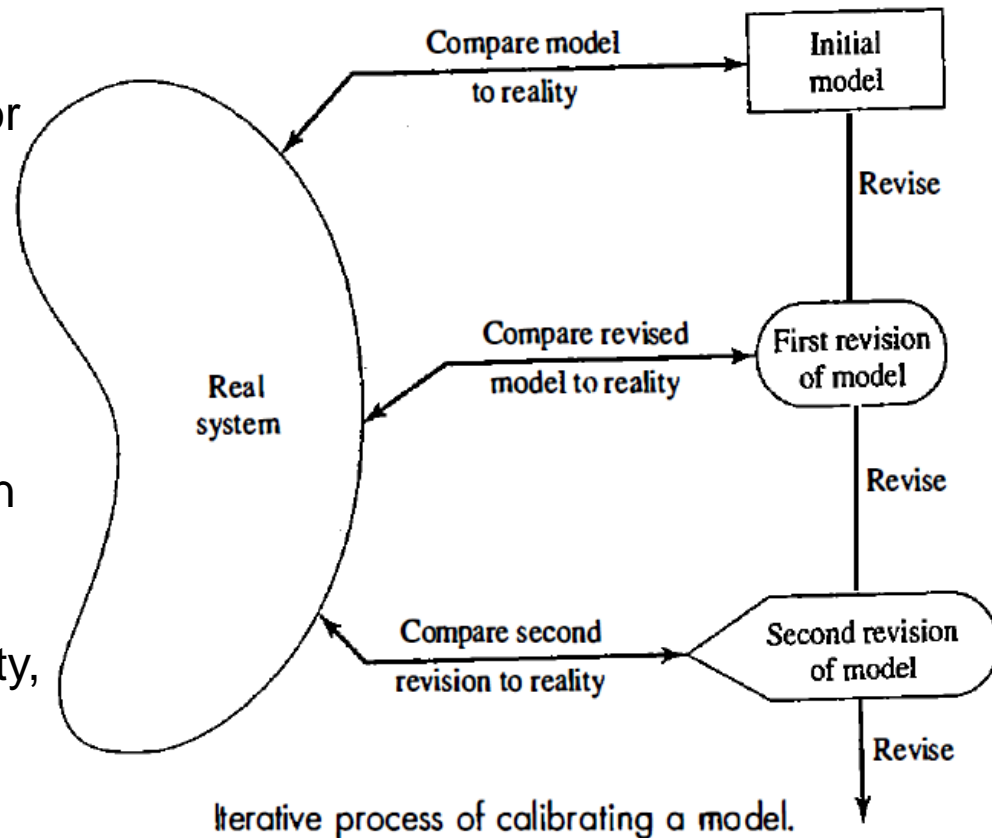
- Output reasonableness: Number of entities; Resource utilization; Duration; etc.
- Commenting and proper naming
- Interactive run control (adjusting parameter values; tracing run time)



# Validation of models

10

- Verification and validation, although conceptually distinct, usually are conducted simultaneously by the modeller.
- Validation is the overall process of comparing the model and its behavior to the real system and its behavior.
- Calibration is the iterative process of comparing the model to the real system, making adjustments (or even major changes) to the model, comparing the revised model to reality, making additional adjustments, comparing again, and so on.



# Validation of models

11

- As an aid in the validation process, Naylor and Finger [1967] formulated a three-step approach that has been widely followed:
  1. Build a model that has high face validity.
  2. Validate model assumptions.
  3. Compare the model input-output transformations to corresponding input-output transformations for the real system.

# Validation of models - face validity

12

- Involve model users at the model development process
  - Dose model appear reasonable on its face to model users and others who are knowledgeable about the real system being simulated?
  - Involve model users form the conceptualization to the implementation step.
  - Involve user for the calibration process to identify model deficiencies!
  - By involving system user, you are going to gain credit from the model user and the system manager can more confidently accept model results.
- Do sensitivity analysis:
  - Whether the model behaves in the expected way when one or more input variables is changed.
  - There are usually many input variables, Choose the most critical input variables for testing!

# Validation of models - assumptions

13

- Are our model structure assumption valid?
  - Is there only one spot for dumping for all truck types or two small trucks can dump concurrently?
  - Do two loaders work in one station for loading trucks or they can load two trucks concurrently?
  - Check validity of your model assumptions with the operation manager and experienced individuals,
- Is our model data valid?
  - Have you selected right people for data collection?
  - Have you done steps required for goodness of fit correctly

# Validation of models - output

14

- Simulation model is used for predicting an operation's result with a new setting in future; however, past historical results of the operation (which has achieved with the old setting of the operation) can be used for validation test!



Is this type of validation feasible for every simulation experiment?

# Validation of models - output

15

- We can only do this type of validation when we have an existing operation, in case of designing new construction operations, we should depend on data from similar existing systems.
- This is especially true when new setting of the operation is not radically changed from the old setting. Examples: To simulate new earthmoving operation when:
  - Two new trucks are added to the operation!
  - Dumping site is moved to a new location!
  - New operation in new location is going to be held for a new project with a whole new set of equipment!

# Validation of models - output

16

- Level of simulation model validation by using historical data from old operation depends on the level of changes we are going to make to the new operation:
  - When minor changes are made to the new operation, new model is validated more carefully. E.g., minor changes made to the single numerical parameters such as number of resources, working hour duration; minor changes made to the physical specification of the system such as length of the road, speed of the truck.
  - When major changes are made to the new operation, new model is partially validated. E.g., structural changes are made to the operation such as setting priority when waiting in a queue, adding new weighing station; different types of tools or systems are used such as adding conveyor for moving dirt instead of trucks!



# Validation of models - output

17

- In output validation the main question is:
  - Can simulation model generate output with the same distribution as the current operation result?
  - In other word we are facing two Hypothesis:
    - H0: Generated results have the same distribution as the current operation
    - H1: Generated results do not have the same distribution as the current operation

**First step:** we do not want to easily reject H0 if we do have a robust reason!

So, accept equality of the result distributions if you can not reject their equality!

=> Can we prove generated result are unequal?

# Validation of models - output

18

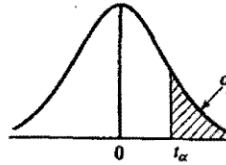
- *t-Student test* is our main tool for the decision:

1- Calculate t-Student test measure as: 
$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

Where  $\bar{x}$  is the sample (or historical data) mean,  $s$  is the standard deviation of the sample and  $n$  is the sample size.  $\mu_0$  is the average of the model output results (since we can get many runs from our computer simulation model we determine average of the model result with reasonably high accuracy)

- 2- Pick an acceptable **error level of type 1** (or  $\alpha$ ). Type 1 error: Acceptable level of error for rejecting  $H_0$  or quality of results distribution by mistake!  $\alpha$  is usually 10% or below!
- 3- Accept  $H_0$  if the absolute value of  $t$  ( $|t|$ ) value is less than corresponding value in the t-Student table:  $t(\alpha/2, n-1)$  where  $n-1 = v$  or degree of freedom.

**Note:** t-Student is a symmetric distribution. The error percentage fall on both left or right side of distribution. So, at one-sided t-Student table we divide  $\alpha$  by 2.

**Table A.5** Percentage Points of The Student's  $t$  Distribution with  $\nu$  Degrees of Freedom

$\nu$	$t_{0.005}$	$t_{0.01}$	$t_{0.025}$	$t_{0.05}$	$t_{0.10}$
1	63.66	31.82	12.71	6.31	3.08
2	9.92	6.92	4.30	2.92	1.89
3	5.84	4.54	3.18	2.35	1.64
4	4.60	3.75	2.78	2.13	1.53
5	4.03	3.36	2.57	2.02	1.48
6	3.71	3.14	2.45	1.94	1.44
7	3.50	3.00	2.36	1.90	1.42
8	3.36	2.90	2.31	1.86	1.40
9	3.25	2.82	2.26	1.83	1.38
10	3.17	2.76	2.23	1.81	1.37
11	3.11	2.72	2.20	1.80	1.36
12	3.06	2.68	2.18	1.78	1.36
13	3.01	2.65	2.16	1.77	1.35
14	2.98	2.62	2.14	1.76	1.34
15	2.95	2.60	2.13	1.75	1.34
16	2.92	2.58	2.12	1.75	1.34
17	2.90	2.57	2.11	1.74	1.33
18	2.88	2.55	2.10	1.73	1.33
19	2.86	2.54	2.09	1.73	1.33
20	2.84	2.53	2.09	1.72	1.32
21	2.83	2.52	2.08	1.72	1.32
22	2.82	2.51	2.07	1.72	1.32
23	2.81	2.50	2.07	1.71	1.32
24	2.80	2.49	2.06	1.71	1.32
25	2.79	2.48	2.06	1.71	1.32
26	2.78	2.48	2.06	1.71	1.32
27	2.77	2.47	2.05	1.70	1.31
28	2.76	2.47	2.05	1.70	1.31
29	2.76	2.46	2.04	1.70	1.31
30	2.75	2.46	2.04	1.70	1.31
40	2.70	2.42	2.02	1.68	1.30
60	2.66	2.39	2.00	1.67	1.30
120	2.62	2.36	1.98	1.66	1.29
$\infty$	2.58	2.33	1.96	1.645	1.28

Source: Robert E. Shannon, *Systems Simulation: The Art and Science*, © 1975, p. 375. Reprinted by permission of Prentice Hall, Upper Saddle River, NJ.

# Validation of models - output

20

- **Example:** Remember our earthmoving example with 3 x 10-tonne trucks, 1 loader, no limitation in number of dumping sites, working hours from 7 am to 7 pm and following activity durations estimated:

Loading: Triang(8,10,13)(minutes)

Trip to dumping site: U(3, 7) (minutes)

Dumping 2 minutes

Trip from dumping: Uniform (3, 6) (minutes)

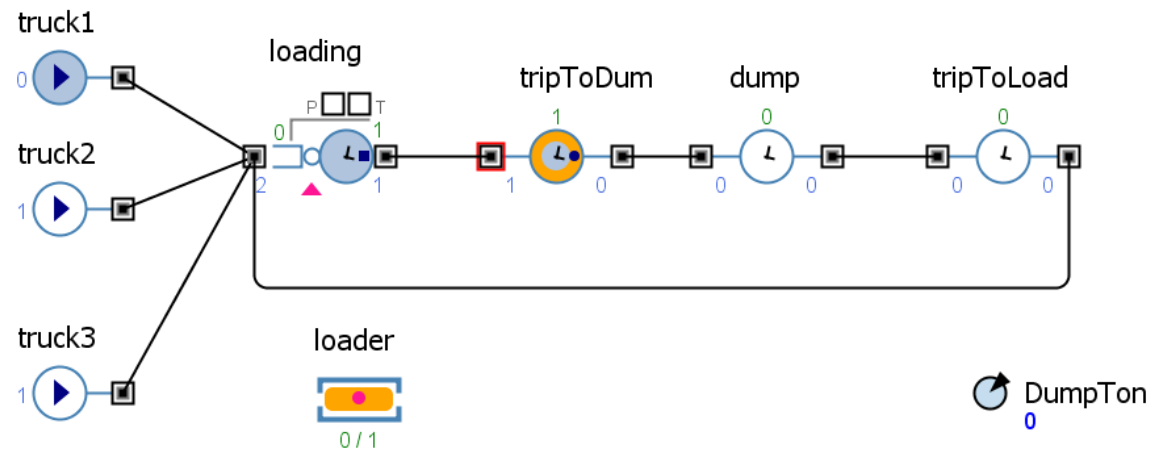
We are going to add another 10-tonne truck and one more loader to the operation. Develop the model, verify the model and validate the model with the historical productivity recorded in different dates as in below:

<b>Productivity (tonne)</b>	760	605	740	630	610	670	625	720	610	750
	730	655	680	675	690	635	635	640	625	694

# Validation of models - output

21

## Example (cont'd):



$$H0: E(x) = \mu_0$$

$$H1: E(x) \neq \mu_0$$

$$\mu_0 = 688.2$$

$$\text{Sample average} = \bar{x} = 668.95 \quad s = 49.9$$

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = -1.727$$

$\alpha = 10\%$      $t(5\%, 19) = 1.729 > 1.727 \Rightarrow$  Cannot accept H1 so accept H0 (equality)

# Validation of models - output

22

We did our **first step** to avoid rejecting H0 (or equality of result distributions) in cases we have no proof of inequality with accepting **type 1 error** at the level of  $\alpha$ !!!!

What if we do not reject inequality but accept unequal distributions!

**Second step:** We want to decrease our error in accepting H0 (or equality of result distributions) by mistake when distributions are not equal (**error type 2**) with accepting error level of  $\beta$ !

1- Set **E(Y)** (where Y is not equal variable to X as our model result variable) as the furthest actual average value of the operation to represent *critical mean difference* that you can tolerate when results are not exactly equal!

2- Set a value for  $\beta$  or **error type 2** (e.g., 10%). This is maximum error level that you set for accepting equality of result distribution by mistake (when it is not equal!)

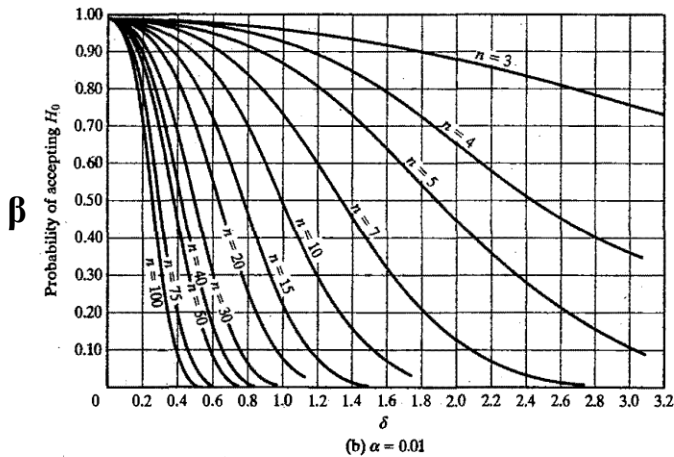
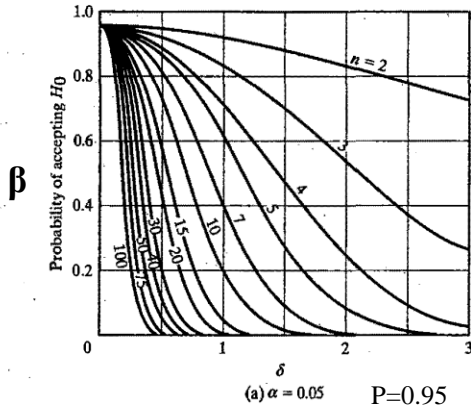
3- Calculate delta measure as:

$$\delta = \frac{|E(Y) - \mu_0|}{\sigma}$$

Where  $\sigma = s$  or sample standard deviation

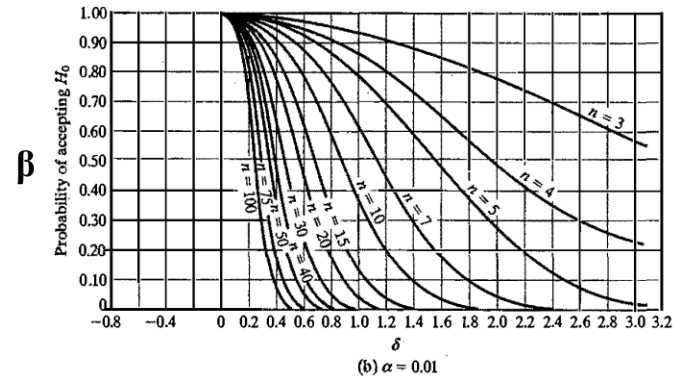
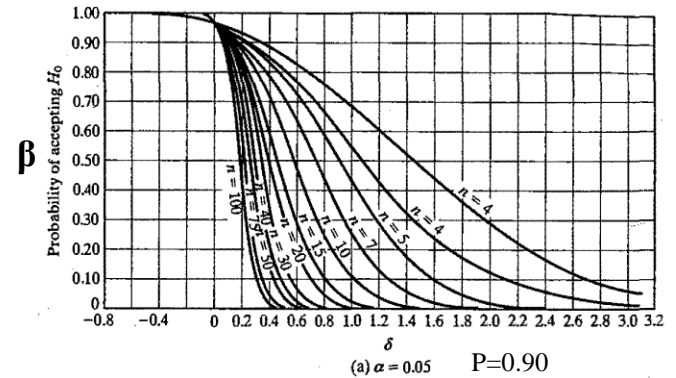
# Validation of models - output

**Table A.10** Operating Characteristic Curves for The Two-Sided  $t$  Test for Different Values of Sample Size  $n$



Source: C. L. Ferris, P. E. Grubbs, and C. L. Weaver, "Operating Characteristics for the Common Statistical Tests of Significance," *Annals of Mathematical Statistics*, June 1946. Reproduced with permission of The Institute of Mathematical Statistics.

**Table A.11** Operating Characteristic Curves for the One-Sided  $t$  Test for Different Values of Sample Size  $n$



Source: A. H. Bowker and G. J. Lieberman, *Engineering Statistics*, 2d ed., © 1972, p. 203. Reprinted by permission of Prentice Hall, Upper Saddle River, NJ.

# Validation of models - output

24

4- Use tables A10 (for two-sided  $\alpha$  values) and A11 (for one-sided  $\alpha$  values) (namely operating characteristics curves) which are graphs of the probability of a Type 2 error.

Select the graph corresponding to your  $\alpha$ ; mark value of  $\delta$  on the x-Axis; draw a perpendicular line from  $\delta$  to intersect the n-curve ( $n$  = number of historical data); find the corresponding value, so called  $\beta(\delta)$ , on the y-Axis;!

5- We cannot reject  $H_0$  (and so will accept  $H_0$ ) if  $\beta(\delta) \leq \alpha$ !!!

- **Example:** In our last earthmoving example with the error type 1 or  $\alpha$  of 10%, we can accept 10% of error type 2 with the maximum mean critical difference of **100** tonnes in average productivity:  $H_0: E(x) = 688.2$        $H_1: E(x) \neq 688.2$

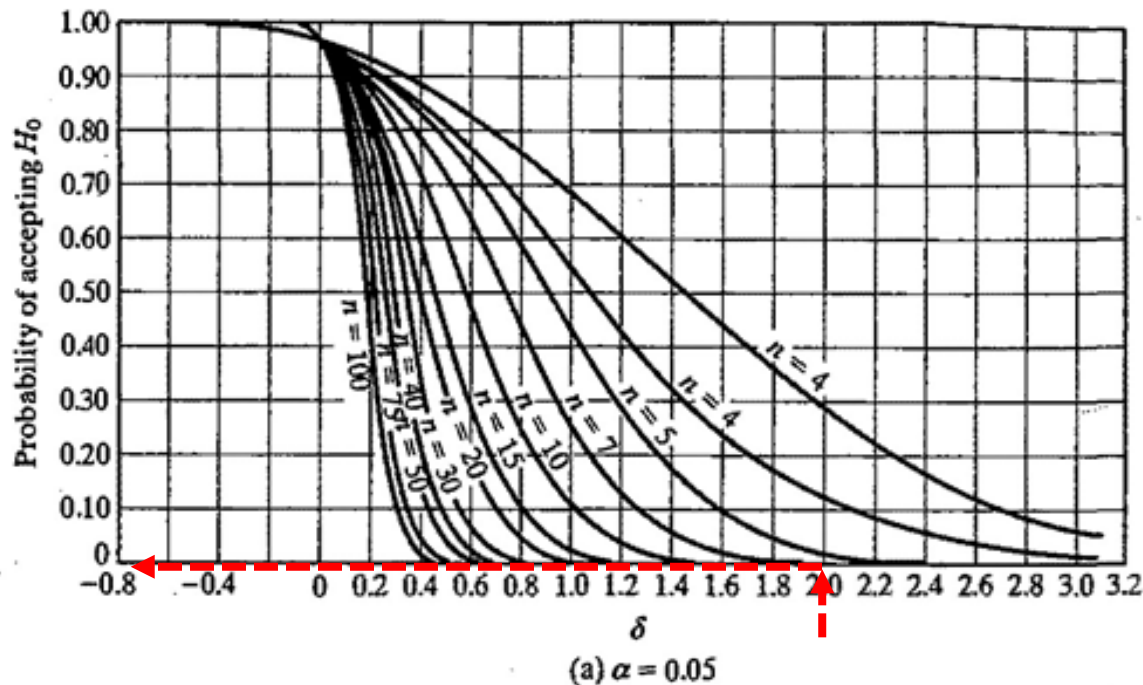
$$\delta = \frac{|E(Y) - \mu_0|}{\sigma} = 100 / 49.9 = 2$$



# Validation of models - output

25

**Table A.11** Operating Characteristic Curves for the One-Sided  $t$  Test for Different Values of Sample Size  $n$




for  $n=20$ ,  $\beta(\delta) = \beta(2) \approx 0 < \beta=0.1$

***We cannot reject  $H_0$ :***

**$H_0$  is accepted!!!!**

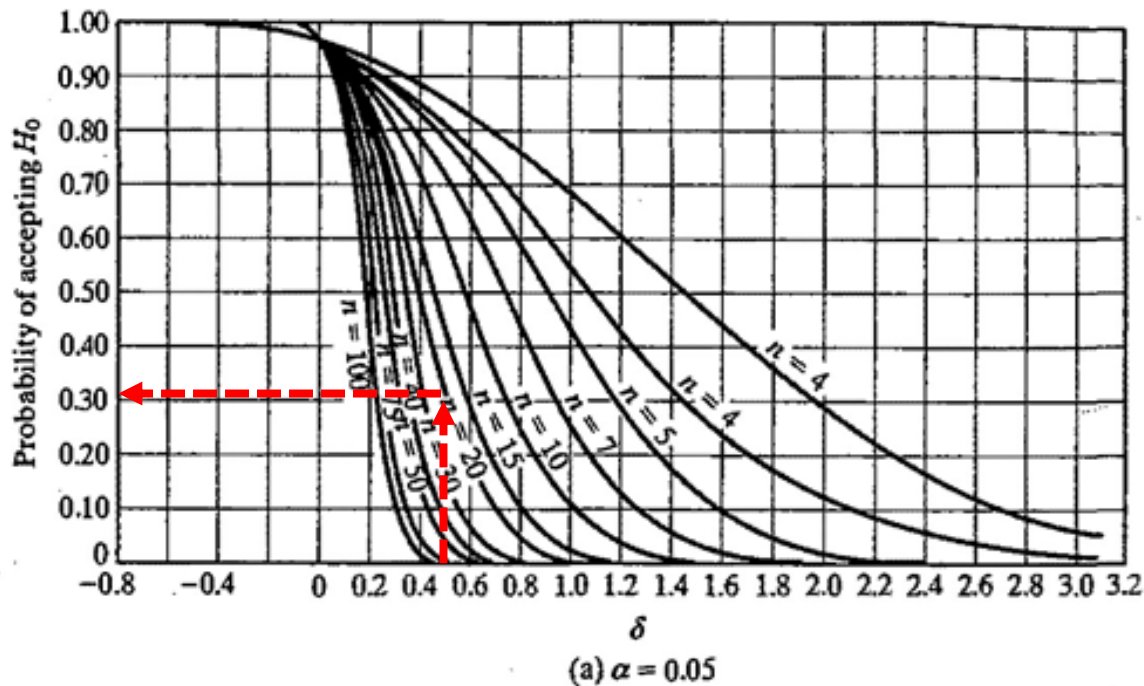
This means chance that the actual average of earthmoving productivity has a difference more than 100 tonnes is almost zero, so the model average and the actual system average are not critically different!

 **What if our mean critical difference is not more than 25 tonnes?**

# Validation of models - output

26

**Table A.11** Operating Characteristic Curves for the One-Sided  $t$  Test for Different Values of Sample Size  $n$



$$\delta = 25 / 49.9 = 0.50$$

for  $n=20$ ,  $\beta(0.5) \approx 0.3 > \beta=0.1$

**We cannot accept  $H_0$ :**

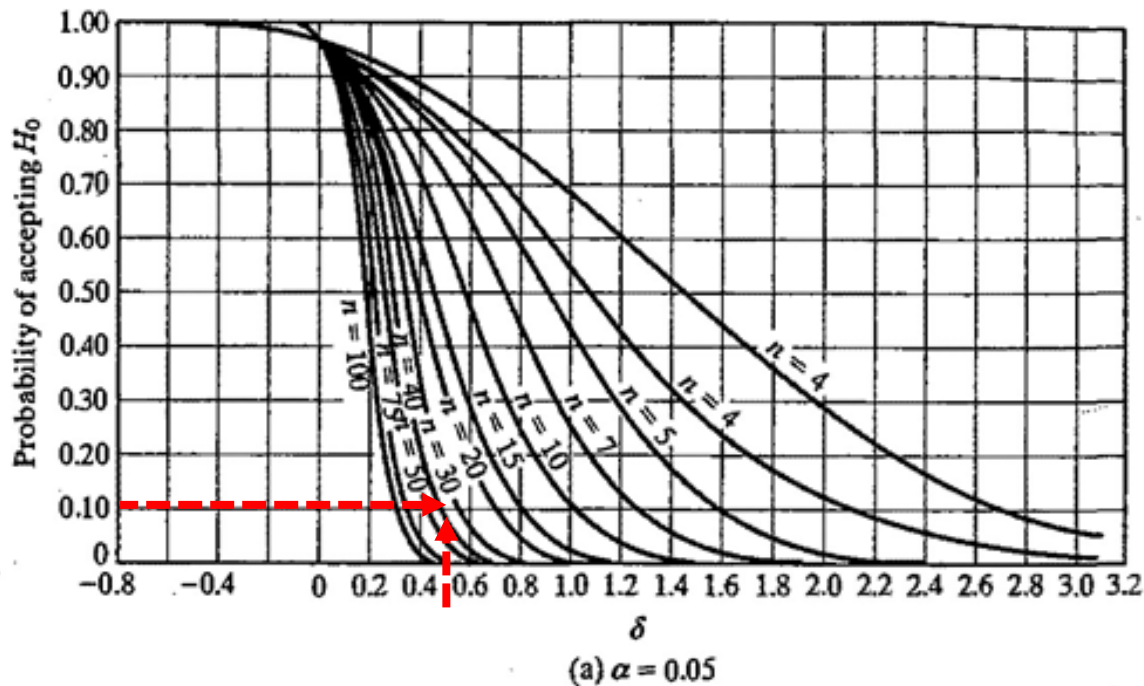
**$H_0$  is rejected!**

This means the chance that the actual average of earthmoving productivity has a difference more than 25 tonnes from 688.2 tonnes of our simulation model is almost **30%** which is more than our acceptable chance for error type 2 (or accepting the equality when they are not equal) of 10%!

# Validation of models - output

27

**Table A.11** Operating Characteristic Curves for the One-Sided  $t$  Test for Different Values of Sample Size  $n$



What would be the minimum sample size that with this amount of acceptable difference in sample average for daily productivity we could still validate the equality?


At the same level of error type 1 (of 10%), the same critical difference of 25 and sample variance if we have  $n=40$  for sample size error Type 2 will be equal to 10%!

# Validation of models - output

28

Summary of type I and type II error in model validation:

<i>Statistical Terminology</i>	<i>Modeling Terminology</i>	<i>Associated Risk</i>
Type I: rejecting $H_0$ when $H_0$ is true	Rejecting a valid model	$\alpha$
Type II: failure to reject $H_0$ when $H_1$ is true	Failure to reject an invalid model	$\beta$

 Use of error type 2 test is critical for validation. We are not interested in accepting equality of system and model results when they really are not!

# After class practice 1



29



In our asphalt plant example suppose two types of trucks with the capacity of 10 tonne (60% chance) and 12 tonne (40% chance) arrive to the plant in exponential distribution with the mean of 15 minutes. Loading duration of 10 tonne truck is  $N(11, 2)$  and for 12 tonne truck is  $N(14, 2)$ . Daily working hours is from 7 am to 7 pm.

- Verify the model by reporting number of output trucks in each sizes (10 and 12 tonne trucks)
- Data in below are historical data from daily tonnage of the asphalt sale!

Validate your model result with this sample output data!

<b>Productivity (tonne)</b>	455	519	536	529	479	465	530	501	455	503
	542	455	487	503	511	513	520	541	517	489

# Validation of models – input / output

30

- In this type of validation we are going to use real system's historical data as our model's input.



What can we expect from our model's output result when using historical data as the model input?

# Validation of models – input / output

31

- In this case we expect same results in both the real system and the model.
- So, our validation test is to compare system's actual result with model result:

$T_j: \{t_{1j}, t_{2j}, \dots, t_{nj}\}$  is  $j^{\text{th}}$  set of input data where  $j= 1, \dots, K$

$Z_{ij}$  is  $i^{\text{th}}$  corresponding real system output to  $T_j$  set of input data

$W_{ij}$  is  $i^{\text{th}}$  corresponding model output result to  $T_j$  set of input data

Input Data Set	System Output, $Z_{ij}$	Model Output, $W_{ij}$	Observed Difference, $d_j$	Squared Deviation from Mean, $(d_j - \bar{d})^2$
1	$Z_{i1}$	$W_{i1}$	$d_1 = Z_{i1} - W_{i1}$	$(d_1 - \bar{d})^2$
2	$Z_{i2}$	$W_{i2}$	$d_2 = Z_{i2} - W_{i2}$	$(d_2 - \bar{d})^2$
3	$Z_{i3}$	$W_{i3}$	$d_3 = Z_{i3} - W_{i3}$	$(d_3 - \bar{d})^2$
⋮	⋮	⋮	⋮	⋮
K	$Z_{iK}$	$W_{iK}$	$d_K = Z_{iK} - W_{iK}$	$(d_K - \bar{d})^2$
$\bar{d} = \frac{1}{K} \sum_{j=1}^K d_j$				$S_d^2 = \frac{1}{K-1} \sum_{j=1}^K (d_j - \bar{d})^2$

$$t_0 = \frac{\bar{d} - \mu_d}{S_d / \sqrt{K}}$$

(Expectedly  $\mu_d=0$ )

If  $|t_0| \leq t(\alpha/2, K-1)$   
 can not reject  $H_0 \Rightarrow$   
 Accept  $H_0$

# Validation of models – Turing test

32

- In Turing test we create 5 reports from real output of the system and 5 reports from model output. We randomly mix 10 reports and present them to an expert working in real system and ask him/her to say which reports are fake and which reports are real. If expert could identify significant amount of fake reports (e.g., 5 or 4 of them) we are going to change the model using expert's suggestions!



# After class practice 2

33




In an asphalt plant with one type of trucks with the capacity of 10 ton and arrival time with exponential distribution with the mean of 15 minutes, historical data for duration of loading trucks during a day has been achieved as in below. Simulate the model using historical data!

**Loading duration historical data**

12	12.4	13.3	12.4	8.1	8.6	12.5	8.6	9	13.6
12.6	10.7	13.6	13.7	12.7	10.8	10.3	10.3	12.9	10.1

# Home assignment 10

34

 We are going to do a simulation study on a structural steel installation job where 1) structural steel elements arrive at the site from steel fabrication shop 2) structural steel elements are stored in order based on their arrival time by iron worker crew 3) a tower crane moves the elements to their erection location 4) iron worker crew first temporarily stabilizes the steel element on its location to let the crane off and can serve other steel element installation 5) Iron worker crew finishes the installation by bolting and welding the steel elements. There is one crane and two iron worker crews for the operation!

Historical input and output data are presented in the [excel sheet attached](#). Develop a simulation model for the operation. Do verification steps to make sure that you have correctly built the model. Do input/ output validation test to make sure that your model is correct. Explain and submit your steps and the model developed!

**Due in one week**

# Reference

- Banks, J., Carson, J.S., Nelson, B.L. And Nicol D.M. (2004) “Discrete event simulation” Prentice Hall, ISBN: 0131446797., chapter 10.



**Thank you!**