Universidad Nacional Autónoma de México
UNAM

Facultad de Ciencias

Validation of supplier estimates using COSMIC method

IWSM October 2019, Haarlem, Netherlands

Dr. Francisco Valdés-Souto
Associate Professor
Department of Mathematics,
Science Faculty,
National Autonomous University of Mexico (UNAM)
fvaldes@Ciencias.unam.mx

CHAIRMAN COSMIC Practices Committee

Mexican Software Metrics Association (AMMS), Founder

COSMIC

AMMS

Facultad de Ciencias

AGENDA

01 Main Research Approach in Estimation is Software Projects

02 Validation Estimate Model Proposal

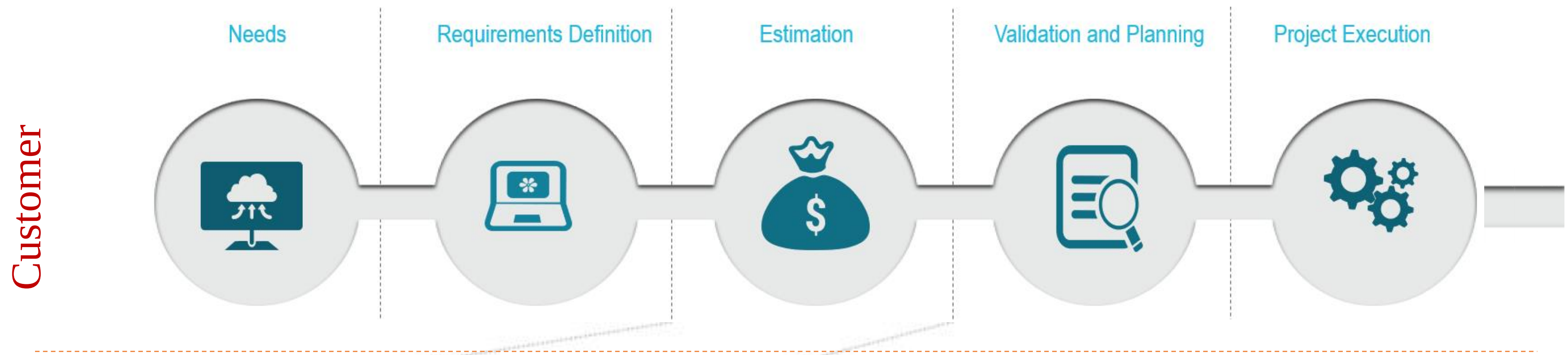
03 Case Study

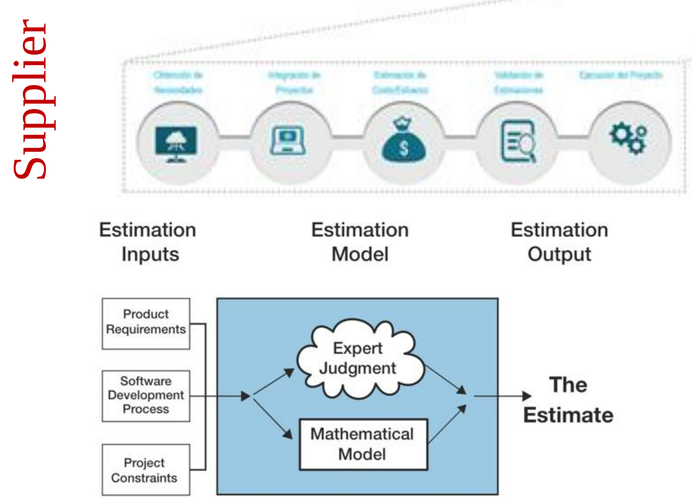
Software Projects

In the competitive software development industry, it is **well known** that software development organizations **need a better and formal** estimation approaches in order to **increase the success rate** of software projects

Facultad de Ciencias

Estimation Context





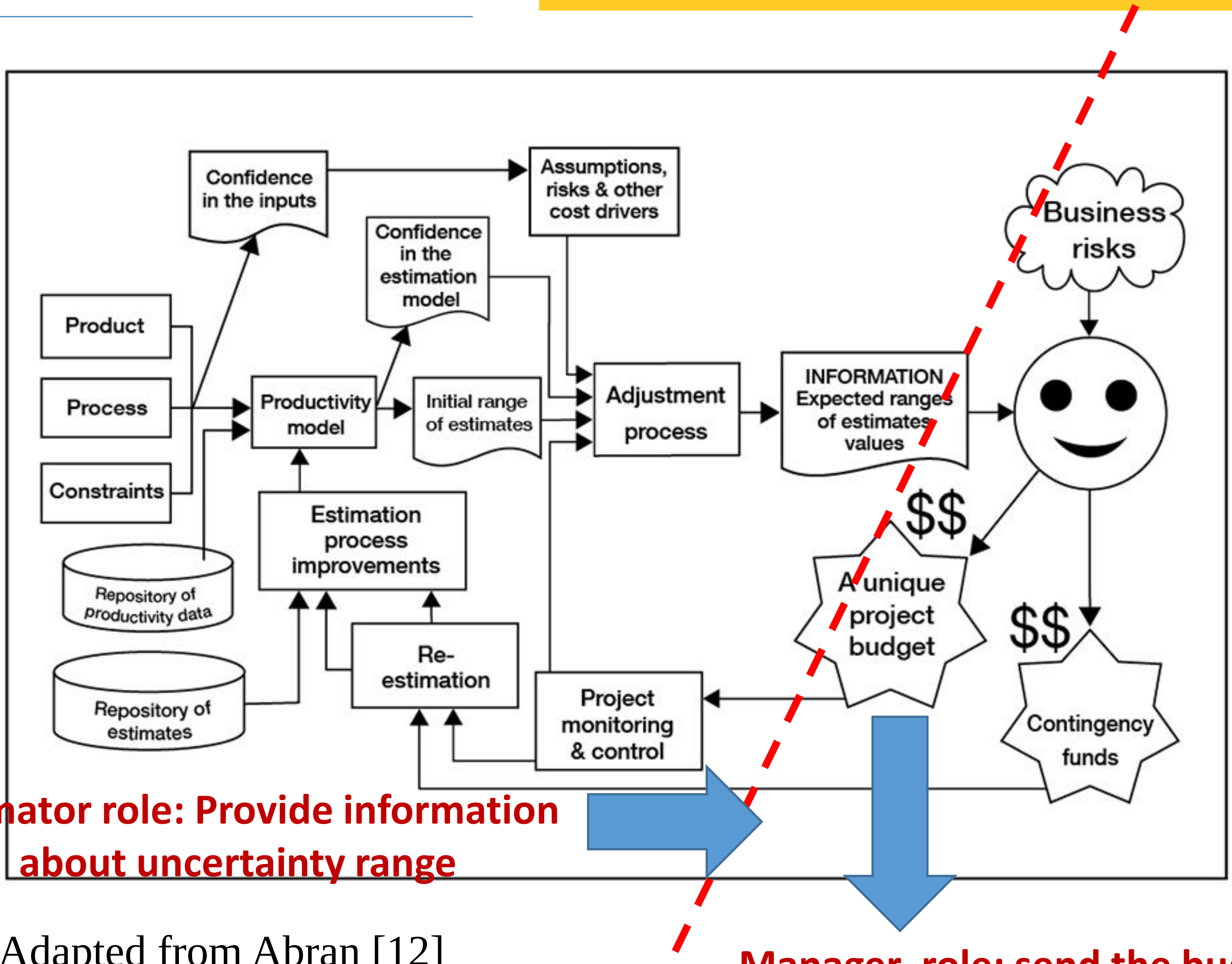
The common view of estimation process. Adapted from Abran [12]

In the literature **several techniques and models** have been proposed to improve the estimation capabilities in software projects over decades. [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12]

Most of the approaches and techniques have been proposed **considering a development organization point of view** (Supplier)**.**

Estimation Process

Supplier

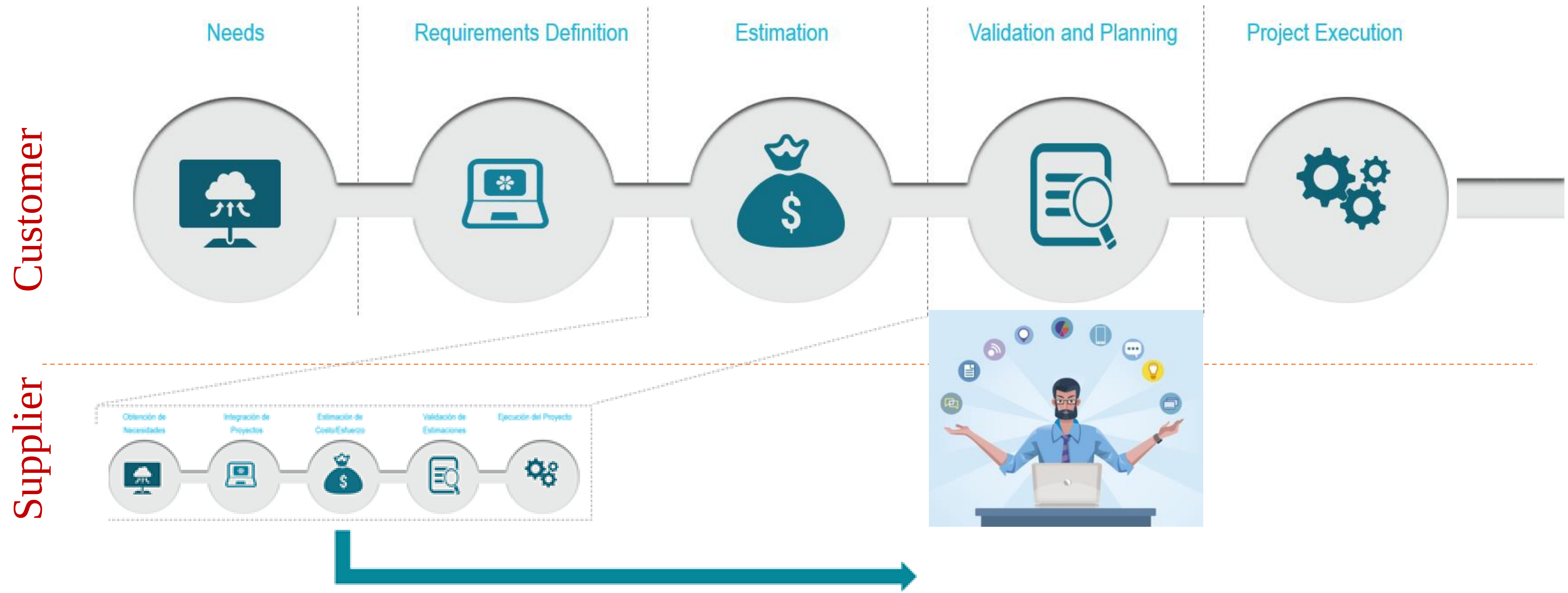


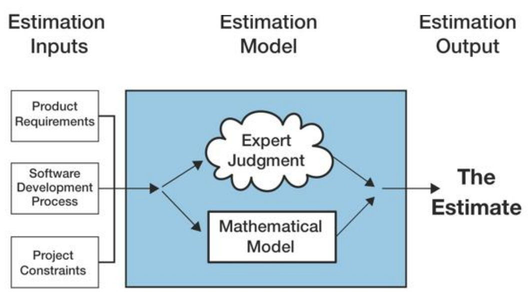
Estimator role: Provide information about uncertainty range

Adapted from Abran [12]

Manager role: send the budget to the Customer

Validation Estimates





The common view of estimation process [12]

The customer **facing the validation estimates problem, without knowing information used by suppliers to make their estimations**, a more useful solution is used "**the expert judgment**" but it is not formal and presenting several problems.

Estimation Model v.s. Validation Model

The estimation model will **predict a possible value, or range of values**, of the cost/effort to construct the software, considering the inputs and the historical projects **from the software development organization (supplier)** to generate a budget to be **proposed to the customer**.

The estimation model could be seen as a regression model as defined in [12] and is referenced to a local historical database that very often is not public.

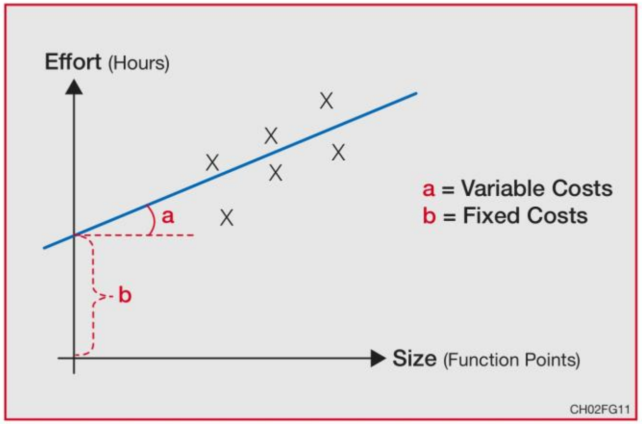
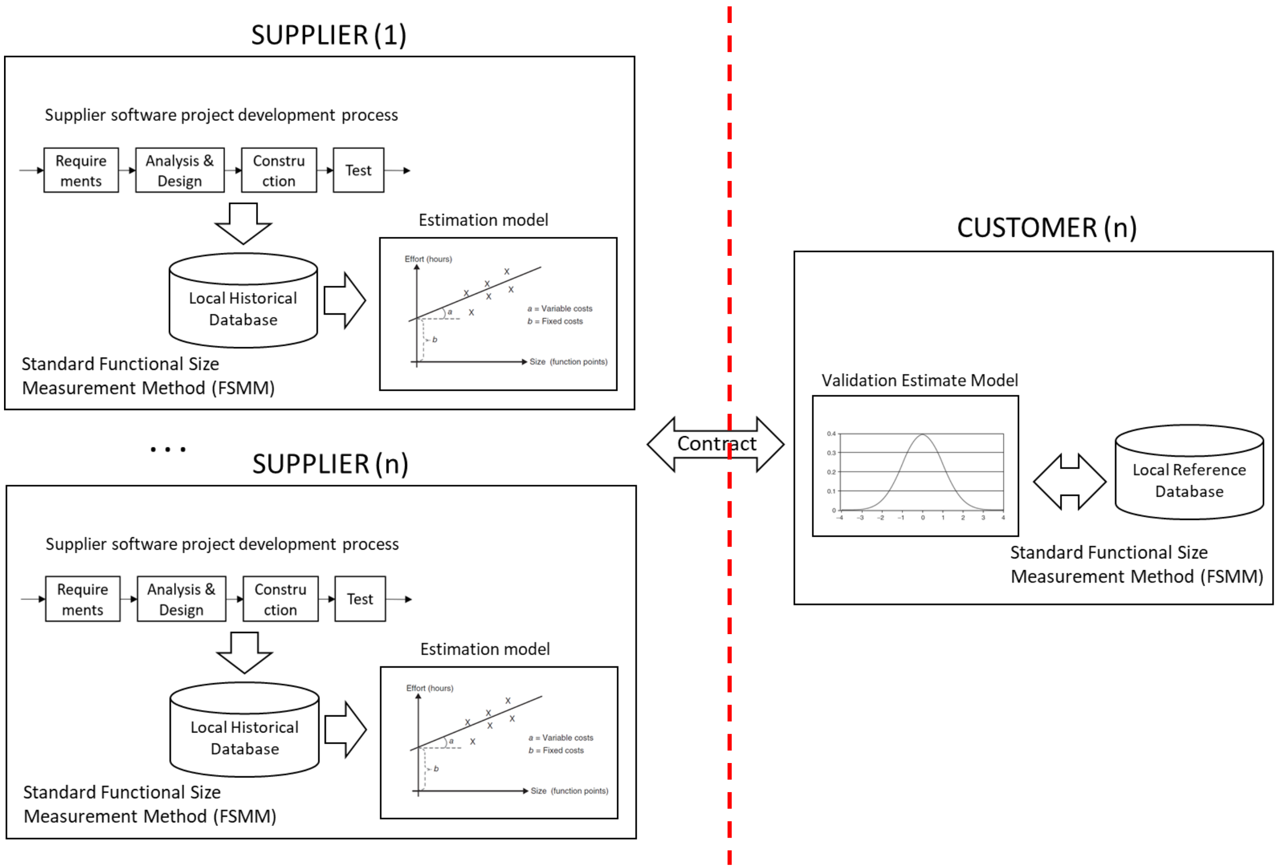


Figure 2.11 Model with a fixed and variale costs.

Adapted from Abran [12]

30

Estimation Model v.s. Validation Estimate Model



The validation estimate model **must to be used to validate if the estimation provided by the supplier** – budget- **accomplish the specific constraints defined by the customer (validation criteria)**.

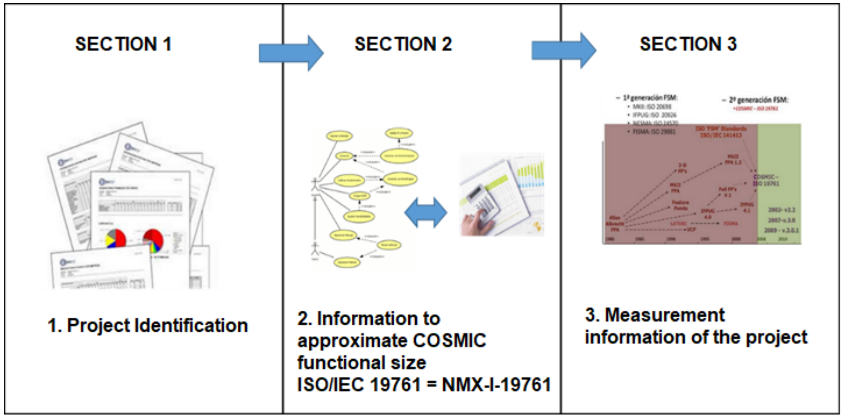
The validation estimate model is generated using the customer reference database, a customer perspective information.

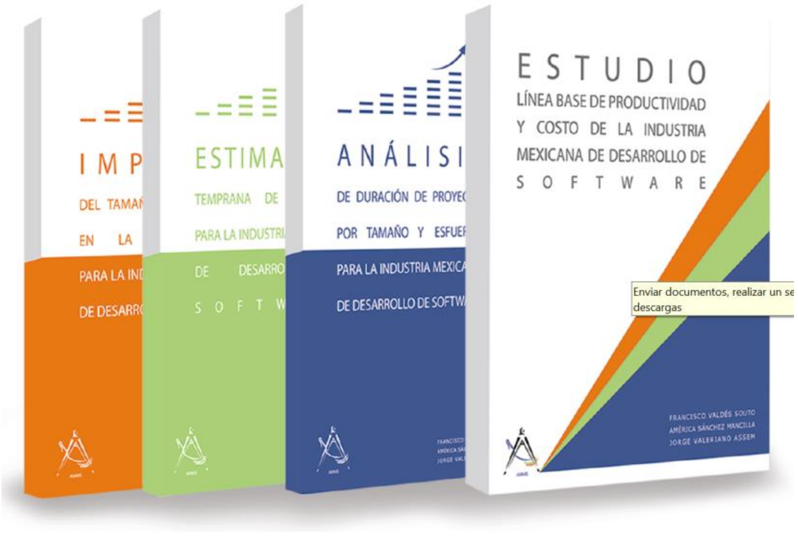
Defining a Validation Estimate Model

- In Mexico at the end of 2015, the **Mexican Software Metrics Association** (AMMS) collect information for the realization of the **Baseline Study of Productivity and Cost** of the Mexican Software Development Industry.

- The purpose was to **obtain information related to software projects** carried out in Mexico (already concluded) from the customer perspective.

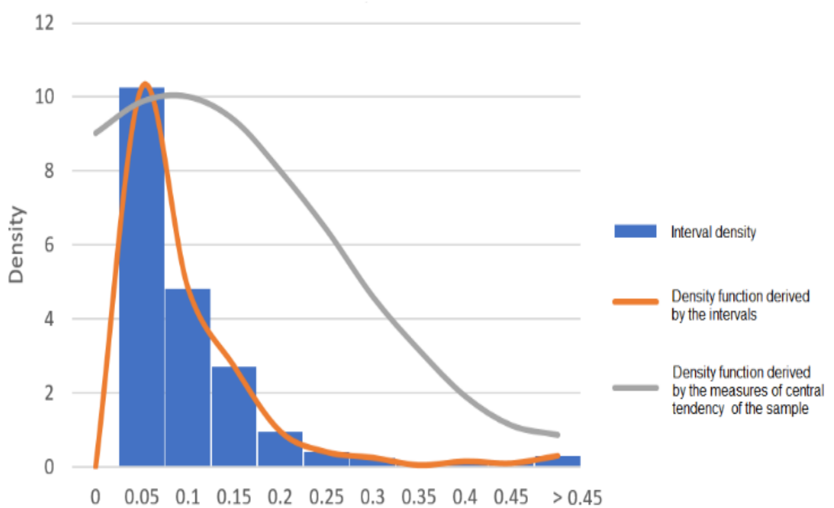
- This study enables defining the baseline of productivity and cost of the IMDS, aiming to improve the knowledge of the IMDS from different points of view, such as the technical aspect and the economic aspect.



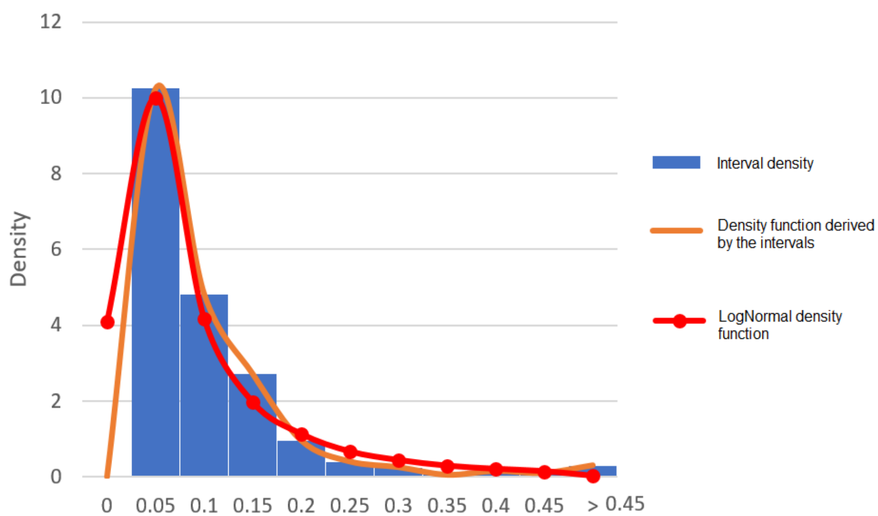


Finding a Probability Distribution for Productivity

Having the customer, a reference database from their perspective, a probability distribution model could be developed.

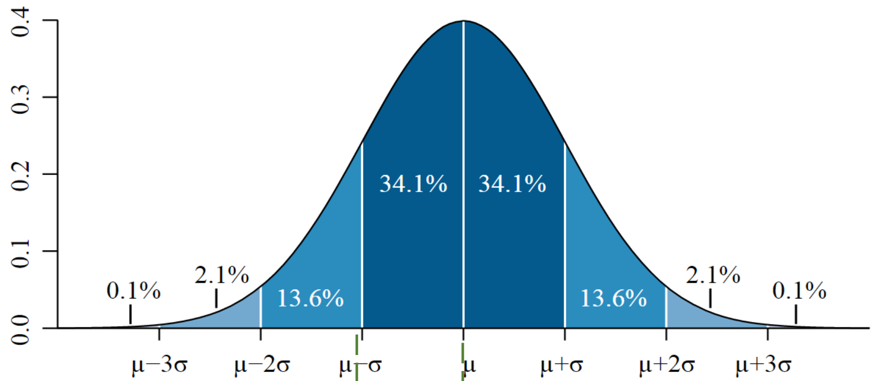


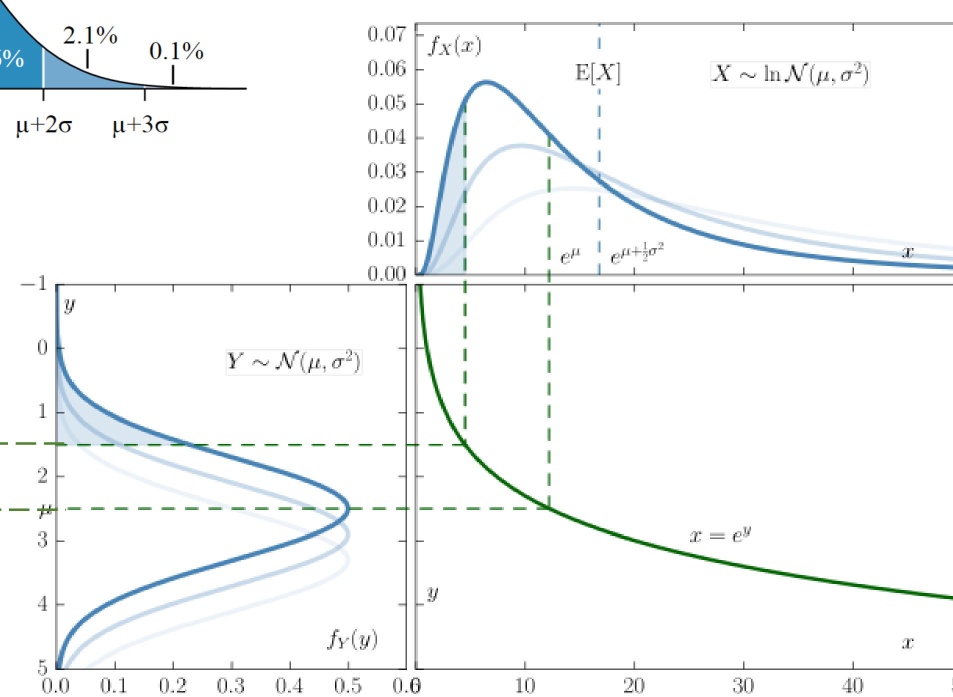
Frequency analysis by productivity Effort/CFP, Adapted from [32]



LogNornal density function by productivity [CFP/Effort], Adapted from [32]

Lognormal – Normal density functions

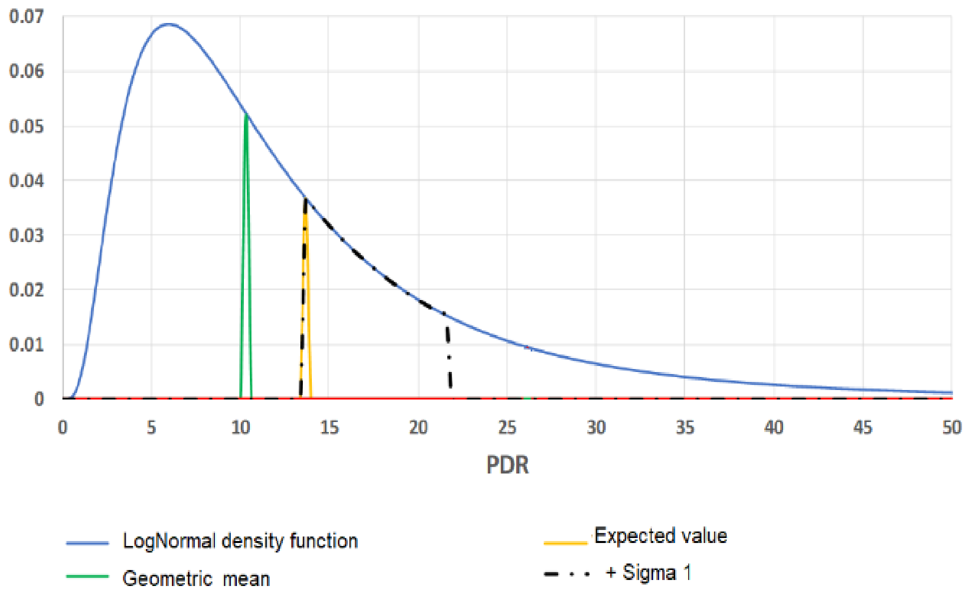




LogNormal Model element	Probability %
- σ3	0.1%
- σ2	2.3%
- σ1	15.8%
Geometric Mean	50%
Expected Value	64.4%
σ1	84.1%
σ2	97.7%
σ3	99.9%

Defining the Validation Criteria

After the explanation of the density function, the customer decides to expect a range between more or equal to 64.4% (Expected Value) and less or equal to 84.1% (σ1) of probability that the supplier estimate is met.



LogNormal Model element	Productivity [CFP/WH]	PDR [WH/CFP]	Probability %
- σ3	0.01	1.12	0.1%
- σ2	0.02	2.35	2.3%
- σ1	0.04	4.93	15.8%
Geometric Mean	0.97	10.34	50%
Expected Value	0.074	13.60	64.4%
σ1	0.20	21.68	84.1%
σ2	0.42	45.45	97.7%
σ3	0.89	95.28	99.9%

LogNornal density function by productivity [CFP/Effort], Adapted from [32]

Facultad de Ciencias

Case Study Description

- A government entity in Mexico has a contract with a software development organization to develop nine (9) projects in 2018.
- The amount for that contract was $33,928,580.00 MNX equivalent approximately to $1,785,714.74 USD in twelve months.
- CONTRACT OPERATION
  - The **customer** request projects to be developed for the supplier.
  - The **supplier** estimates the projects using the technique they use. In this case uses Expert Judgment; the **supplier** must present the **effort estimated** for each project and the assumptions.
  - The **customer** validated the estimates using the model defined previously

- To develop the validation, the **customer** gathers the functional size in COSMIC units (CFP) using the EPCU approximation technique [23,24] for each project and obtains the **PDR** considering the effort estimated by the **supplier**.

Facultad de Ciencias

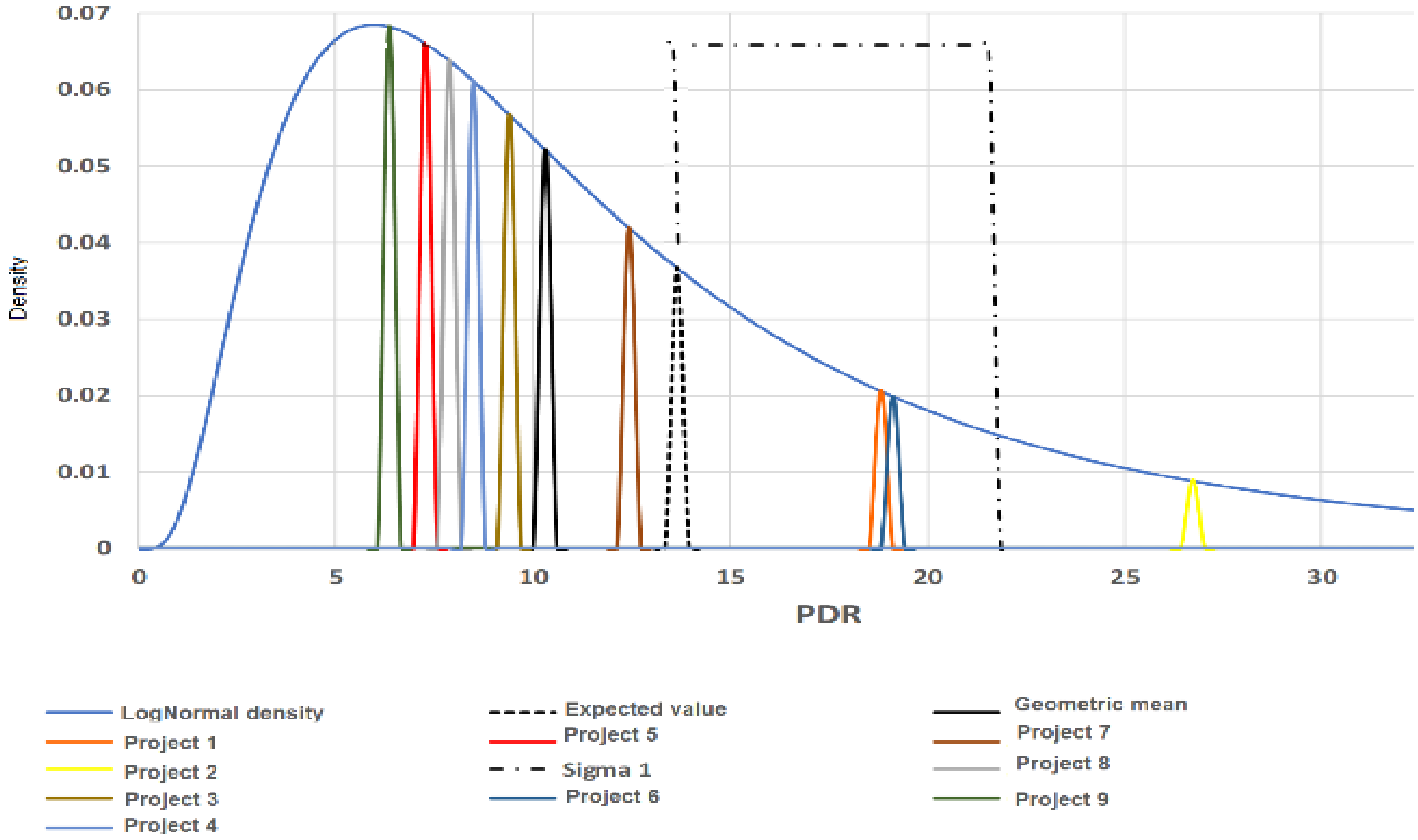
Case Study Data

Project	Number of UC or FP	Type (UC / FP)	Total effort [WH]	Approximated Functional Size (CFP)	PDR [HH/CFP]	Validation estimate range [13.60 , 21.68]	Success Probability [%]
1	35	CU	12,994	686.6	18.92	✓	79.3%
2	12	CU	3,370	126.2	26.71	✗	90.0%
3	21	CU	8,634	907.5	9.51	✗	45.5%
4	97	PF	7,648	910.4	8.40	✗	38.9%
5	26	CU	4,294	580.0	7.40	✗	32.6%
6	26	CU	8,665	456.9	18.96	✓	79.4%
7	19	CU	5,897	476.1	12.39	✗	59.6%
8	73	PF	6,280	807.4	7.78	✗	35.0%
9	134	PF	11,460	1,745.9	6.56	✗	27.0%
TOTAL			69,242	6,697.0			

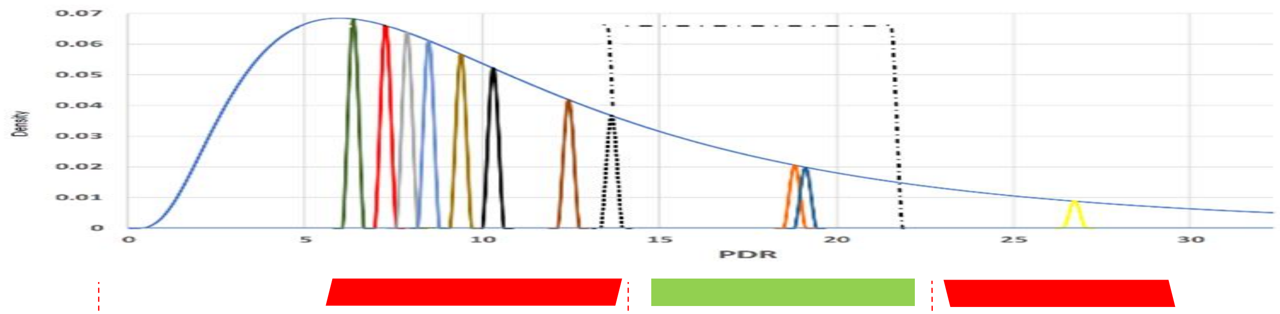
- (light green) 2 projects (22.2%) that accomplished the validation criteria (P1 & P6)
- (light grey) 5 projects presenting less than 50% of probability of success. (P9L & P3H)
- (grey) P7 presenting more than 50% of probability of success, but less than min validation criteria
- (orange) P2 presents the highest probability with a higher PDR, more the max validation criteria

Facultad de Ciencias

Case Study Graphically



Case Study Results



Theoretical decision	Reject six projects because were underestimated, the problem, in this case, is not the cost because will be lower, the problem arises because the low probability of the estimations will be met.	Accept two projects because they are between the validation criteria.	Reject one project because was overestimated, this project could be 23.2% more expensive than the limit, of course, the probability success was very high (90%).
Real results	Present problems because the extra effort was required, for 4 of them, more effort was obtained through several changes request, for two projects the cost was assumed by the supplier.	For the projects recommended to be accepted were that the project finished according to the estimated effort.	The supplier reports the consumption of all the effort, there is no way to check if the supplier uses less effort.

Facultad de Ciencias

Threats to validity

- The use of COSMIC as FSMM was established because in Mexico COSMIC is the National Standard: NMX-I-19761-NYCE-2017. Intuitively, the proposed approach in this paper could be generalized for other FSMM like NESMA or IFPUG but must be proven.

- It could be assumed that if the customer and the supplier use the same FSMM as a “understanding” element, the application of the approach could be simplified, but must to be proven.

- Because the use of validation estimates model in this case study was only for test-ing and no decisions were taken, the results need to be validated against real context. Hence taking the decisions and comparing with the real results of the projects, after the application of the validation criteria defined

Conclusion

- The validation estimate model was based in a Lognormal density function that fits better to the data gathering in the Baseline Study.
- The validation criteria was defined by the customer expecting a range between more or equal to 64.4% and less or equal to 84.1% of probability, this probability is higher than several studies in the industry about software projects success.
- Another relevant issue is that both under-estimation as well as over-estimation are considered as disqualifiers.
- With this formal way to validate the supplier estimation, a customer could improve the decision making about the feasibility of the project because the model was defined by a reference database and the probability that the supplier estimate is known.
- Using the validation estimates model defined with their validation criteria could be useful for the customer to avoid some problems in the project's execution for the underestimated projects and could help to save resources and control the supplier for the overestimated project.

Facultad de Ciencias

QUESTIONS?

Dr. Francisco Valdés-Souto

Associate Professor
Department of Mathematics,
Science Faculty,
National Autonomous University of Mexico (UNAM)
fvaldes@Ciencias.unam.mx

CHAIRMAN
COSMIC Practices Committee

COSMIC