

# Is it possible to improve the accuracy of expert judgment when estimating the effort of a software product?

Gabriela Robiolo

UA  
Buenos Aires, Argentina


Oscar Castillo y  
Bibiana Rossi

UADE  
Buenos Aires, Argentina


Silvana Santos

UNLP  
La Plata, Argentina

# Contents of this presentation

- Motivations
  - Goal
  - Method
  - Case Study
  - Conclusion
  - Future work
- 

# Motivations

- ▶ Estimation based on expert judgment is the most common method used to estimate effort in the software industry today
  - ▶ The use of formal estimation methods is not invalid, but those methods have been unable to exceed the human capacity to synthesize several complex variables.
  - ▶ Experts' estimations have the same or better accuracy than those obtained when using a formal estimation model.
- 


# Goal

- ▶ Our research questions are:
  - When working on a complex software product, is it possible to reduce the error in effort estimation based on experts' judgment, by applying a formal estimation method which uses one variable, i.e. size or complexity?
  - When working on a complex software product, is it possible to reduce the effort estimation error by obtaining a sequence of successive estimations either by using the analogy evaluated by experts method or the analogy based on size or complexity method?

# Method

- ▶ Comparison of expert judgment vs two formal methods:
  - Linear regression
  - Analogy
  
- ▶ Measures:
  - COSMIC
  - Paths

# Case Study

- ▶ The company was Argentinean, public, large and complex.
  - ▶ The application studied was a traffic violation processing system.
- 

# Data

Version	ID UC	AE	CFP	P	Expert (UC)	Expert (successive versions)
7	1	264	20	3	240	240
7	2	32	5	3	24	24
7	3	248	15	23	208	208
7	4	112	37	15	88	88
9	5	104	16	5	80	80
9	6	136	15	9	96	96
10	7	56	10	3	64	64
10	8	184	7	25	112	120
10	9	416	93	63	328	344
10	10	8	11	2	8	16
10	11	16	4	2	8	16
11	12	208	90	58	176	200
11	13	24	5	2	16	16
11	14	144	149	16	120	144
11	15	96	54	5	96	88
12	16	520	46	43	504	504
11	17	112	97	6	136	144
11	18	40	71	30	40	40

# Experts' descriptions

Skills	Description
Training	Engineering degree
Years of experience in software development	12
Years of experience in leadership	8
Specialty, Knowledge	Development. NET, Java patterns, team leadership, development methodologies

Capabilities	High	Medium	Low
General Experience	x		
Knowledge of the performance profiles of the development group		x	
Technology used	x		
Knowledge		x	



# Analogy

- ▶ Find a UC similar to the UC to be estimated, based on its size or Complexity
- ▶ The Productivity of the "most similar" UC –in size or complexity– is used to calculate EE:

$$Pr = AE / X_{[CFP,P]}$$
$$EE = X_{[CFP,P]} * Pr_{UCA}$$

# Data

Version	ID UC	AE	CFP	P	Expert (UC)	Expert (successive versions)
7	1	264	20	3	240	240
7	2	32	5	3	24	24
7	3	248	15	23	208	208
7	4	112	37	15	88	88
9	5	104	16	5	80	80
9	6	136	15	9	96	96
10	7	56	10	3	64	64
10	8	184	7	25	112	120
10	9	416	93	63	328	344
10	10	8	11	2	8	16
10	11	16	4	2	8	16
11	12	208	90	58	176	200
11	13	24	5	2	16	16
11	14	144	149	16	120	144
11	15	96	54	5	96	88
12	16	520	46	43	504	504
11	17	112	97	6	136	144
11	18	40	71	30	40	40

$$Pr_9 = 416/93$$

$$EE_{12} = 90 * Pr_6$$

# Analogy applied to successive estimations

- ▶ The analogy estimation method, i.e. using CFP and P as variables, was applied in the way explained above, with the particularity that the set used to find the analogue UC was the subset of the corresponding UC version plus the UC previous versions.

# Estimation methods: comparison

<b>Estimation Method</b>	<b>MMRE</b>	<b>MeMRE</b>	<b>PQ(0.25)</b>	<b>ErR (min..max)</b>
Linear Regression (P)	1.45	0.34	0.38	-8.38..0.79
Analogy (CFP)	1.53	0.83	0.06	-6.7..0.87
Analogy (P)	0.94	0.46	0.39	-4.52..0.89
Expert	0.19	0.18	0.72	-0.21..0.5

# Comparing successive versions estimates

Estimate method	MMRE	MeMRE	PQ(0,25)	Relative errors (min..max)
COSMIC v7	1.35	0.92	0.25	-3.36.. 0.2
COSMIC v9	1.18	0.9	0	-3.36.. 0.5
COSMIC v10	1.43	0.75	0	-6.7..0.87
COSMIC v11	1.29	0.79	0	-6.7..0.87
COSMIC v12	1.53	0.83	0.06	-6.7..0.87
Paths v7	2.22	0.66	0	-7.25..0.88
Paths v9	1.86	0.95	0	-7.25..0.88
Paths v10	1.01	0.83	0.1	-4..0.83
Paths v11	0.76	0.46	0.33	-4..0.83
Paths v12	0.94	0.46	0.39	-4.52..0.89
Expert (V7)	0.18	0.19	0.75	0.09..0.25
Expert (V9)	0.20	0.22	0.67	0.09..0.29
Expert (V10)	0.26	0.21	0.64	-1..0.35
Expert (V11)	0.22	0.17	0.67	-1..0.35
Expert (V12)	0.20	0.17	0.63	-1..0.35

Wilcoxon  
p-value

Alternative  
hypothesis

0.009

0.000

0.006

0.000

0.000

- ▶ The MRE mean of expert estimate is lower than the MRE mean of the estimation obtained when using:
  - H1a: the linear regression model and the independent variable P.
  - H1b: the analogy method and the measure size (CFP).
  - H1c: the analogy method and the measure complexity(P)
  - H1d: the analogy method and the measure size (CFP), in a history of successive estimates.
  - H1e: the analogy method and the measure complexity (P), in a history of successive estimates.

# Threats to validity


- ▶ AE registration was made by the company:
  - It was validated with other records of hours worked
- ▶ The amount of use cases used was limited and this could have affected the results.
  - statistical testing of the hypothesis was significant.
- ▶ The expert's knowledge of the actual hours of this case study.
  - The expert joined the company at a later stage, after the actual development of the versions presented in this case study and had no knowledge of the actual hours.
  - The group that developed the versions had changed at the time the expert was consulted.

# Conclusion

- ▶ In this case study, it was not possible to exceed the experts' estimates by using formal estimation methods (linear regression and Analogy), in an estimate for all UC of the application as an estimate for successive versions .
- ▶ The participation of only one expert limits the conclusion of the case study, but it does not rule it out; it highlights the value of every case study in a real environment.
- ▶ The expert was typified, as it is expected to obtain different accuracies when the expert's profile varies .



# Final conclusions

- ▶ The expert's characteristics favored the accuracy of his estimates.
  - ▶ Surprising that in the subsequent history of estimates we did not obtain an improvement in accuracy.
  - ▶ Only one variable was used in the estimation models using formal methods, which affected the accuracy of the models.
  - ▶ To generalize the results, various other domains, different experts' profiles and other formal methods must be analyzed .
- 

# Future work

- ▶ Work with models that include other variables that may affect effort estimation
- ▶ Change the profile of the experts to those who have a high knowledge of:
  - Group development performance
  - Domain development.

[grobiolo@austral.edu.ar](mailto:grobiolo@austral.edu.ar)