On project probabilistic cost analysis from LHC tender data

Ph. Lebrun

CERN, Geneva, Switzerland

TILC'09, Tsukuba, Japan

17-21 April 2009

Basis of probabilistic cost analysis

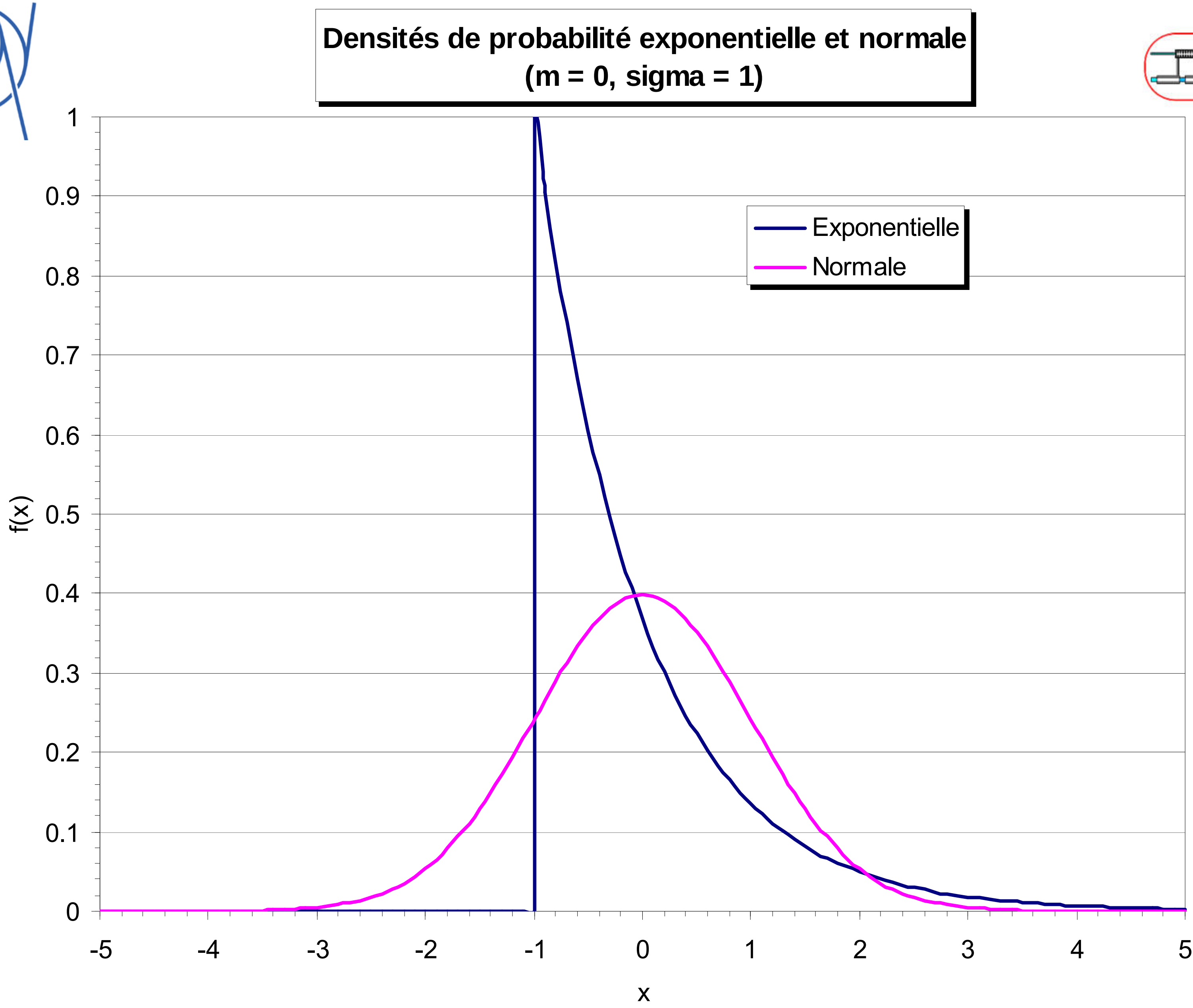
- Following the PBS, the project is split in i lots, the cost of which are random variables X_i with
  - mean value m_i
  - standard deviation σ_i
- The total cost of the project is a random variable $X = \Sigma X_i$
  - with mean value $m = \Sigma m_i$
- In the case when the X_i are statistically independant, $X = \Sigma X_i$ is characterized by
  - standard deviation $\sigma = (\Sigma \sigma_i^2)^{½}$
  - probability density function (PDF) asymptotically tending to Gaussian (central-limit theorem)
- Statistical independance or correlations between X_i is more important to probabilistic analysis of total cost, than detailed knowledge of the specific PDFs of X_i

Statistical modeling of component costs

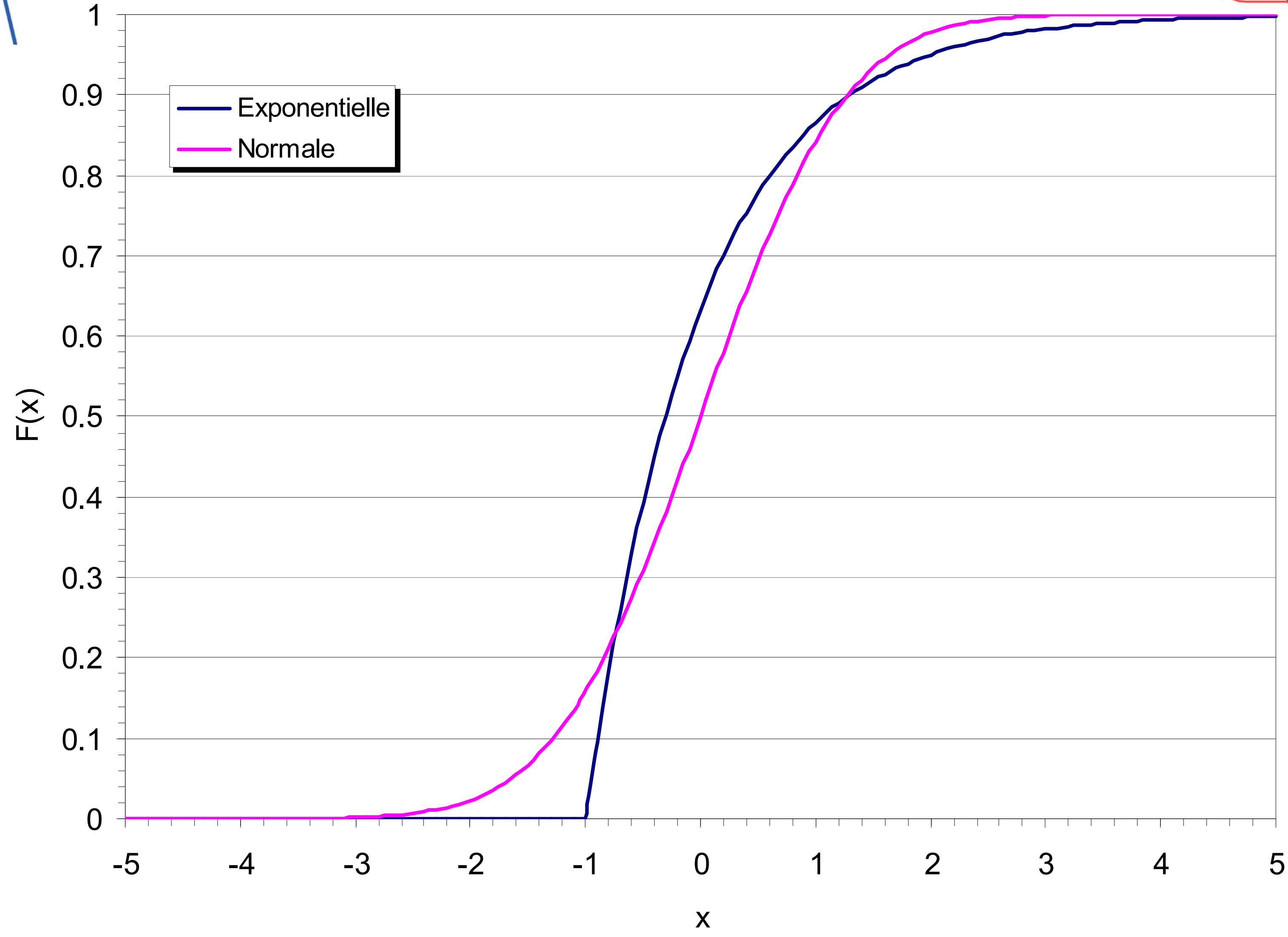
- Heuristic considerations
  - things tend to cost more rather than less ⇒ statistical distributions of X_i are strongly skew
  - PDFs $f_i(x_i)$ are equal to zero for x_i below threshold values b_i equal to the lowest market prices available
  - commercial competition tends to crowd prices close to lowest ⇒ PDFs $f_i(x_i)$ are likely to be monotonously decreasing above threshold values b_i
- The exponential PDF is a simple mathematical law satisfying these conditions

$$\boxed{\begin{array}{ll} f(x) = 0 & \text{for } x < b \\ f(x) = a \exp[-a(x-b)] & \text{for } x \geq b \end{array}}$$

- Characteristics of the exponential law
  - only two parameters a and b
  - threshold b
  - mean value $m = 1/a + b$
  - standard deviation $\sigma = 1/a = m - b$
  - « mean value = threshold + one standard deviation »


Densités de probabilité exponentielle et normale
(m = 0, sigma = 1)
Exponentielle
Normale
f(x)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
-5
-4
-3
-2
-1
0
1
2
3
4
5
x

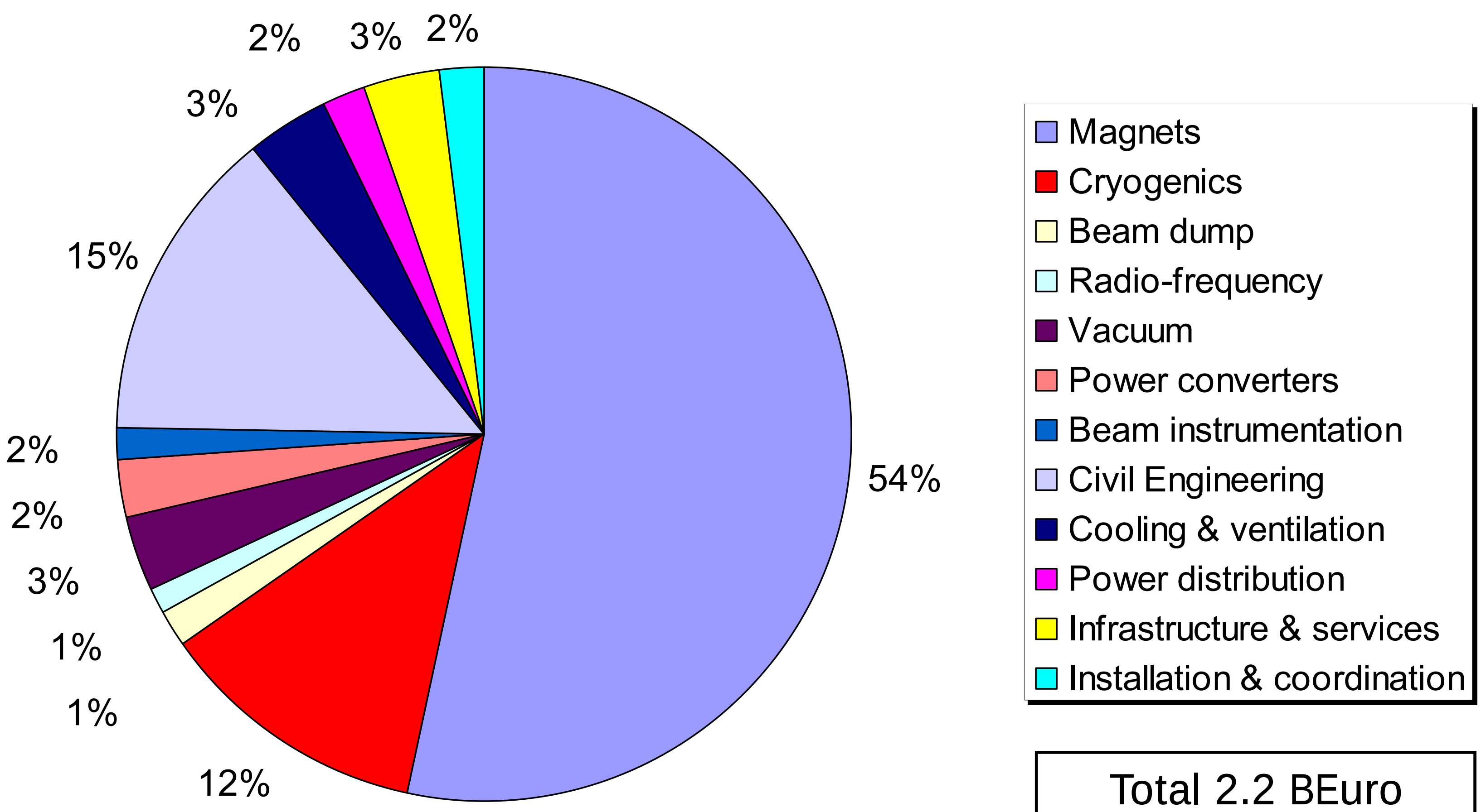
Fonctions de distribution exponentielle et normale (m = 0, sigma = 1)



Comparing Gaussian and exponential PDFs

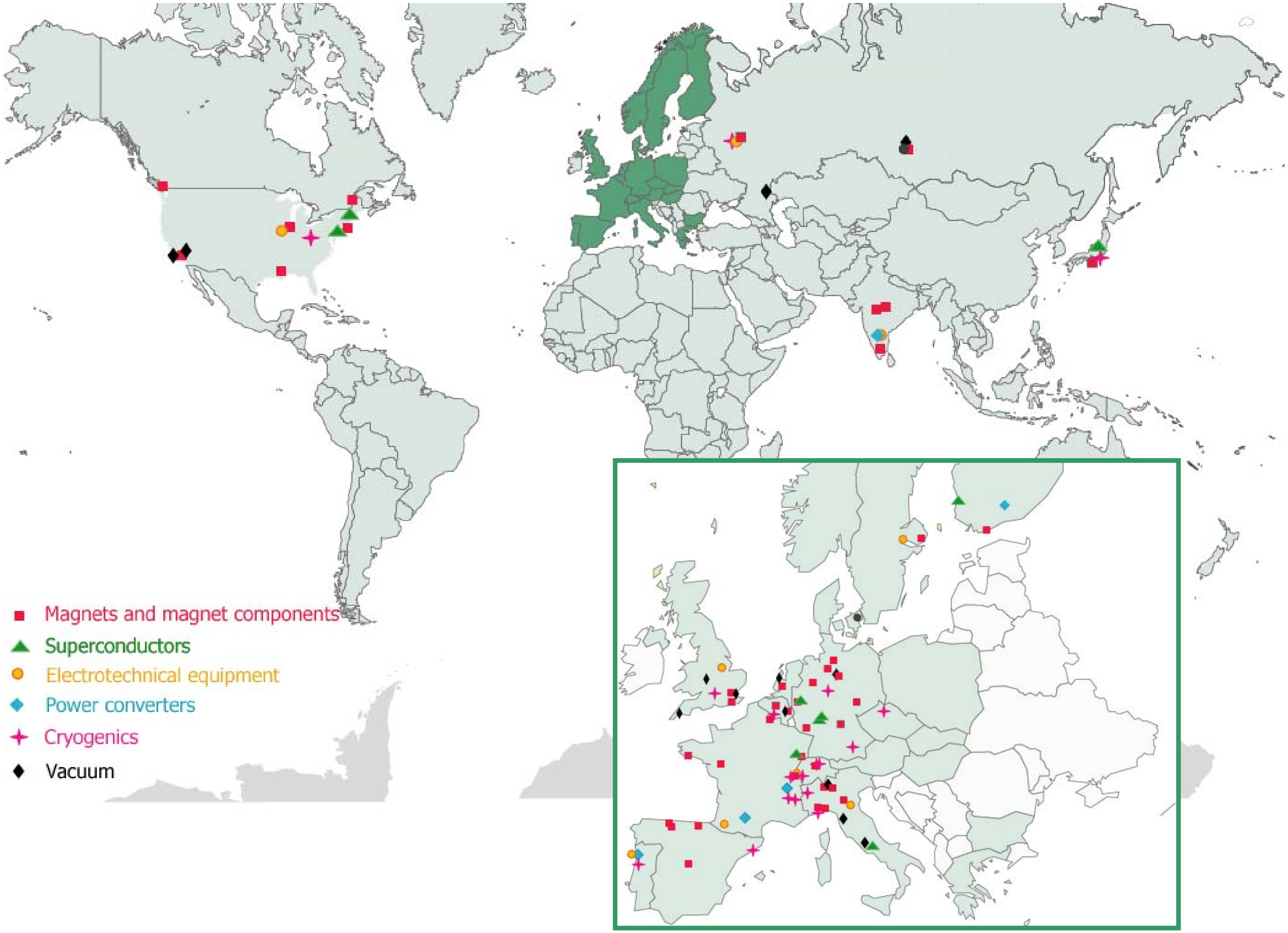
- Gaussian
  - $X \leq m - \sigma$ at confidence level 15,9%
  - $X \leq m$ at confidence level 50%
  - $X \leq m + 1{,}28\,\sigma$ at confidence level 90%
  - $X \leq m + 1{,}65\,\sigma$ at confidence level 95%
  - $X \leq m + 2{,}06\,\sigma$ at confidence level 98%
- Exponential
  - $X \leq m - \sigma$ at confidence level 0
  - $X \leq m$ at confidence level 63,2%
  - $X \leq m + 1{,}30\,\sigma$ at confidence level 90%
  - $X \leq m + 2{,}00\,\sigma$ at confidence level 95%
  - $X \leq m + 2{,}91\,\sigma$ at confidence level 98%

LHC cost structure



Total 2.2 BEuro

90 main contracts in advanced technology



Cost variance analysis

Cost variance factor	Evolution of configuration	Technical risk in execution	Evolution of market	Commercial strategy of vendor	Industrial price index	Exchange rates, taxes, custom duties
Lot 1						
Lot 2						
Lot 3						
...						
Lot N						
Total						

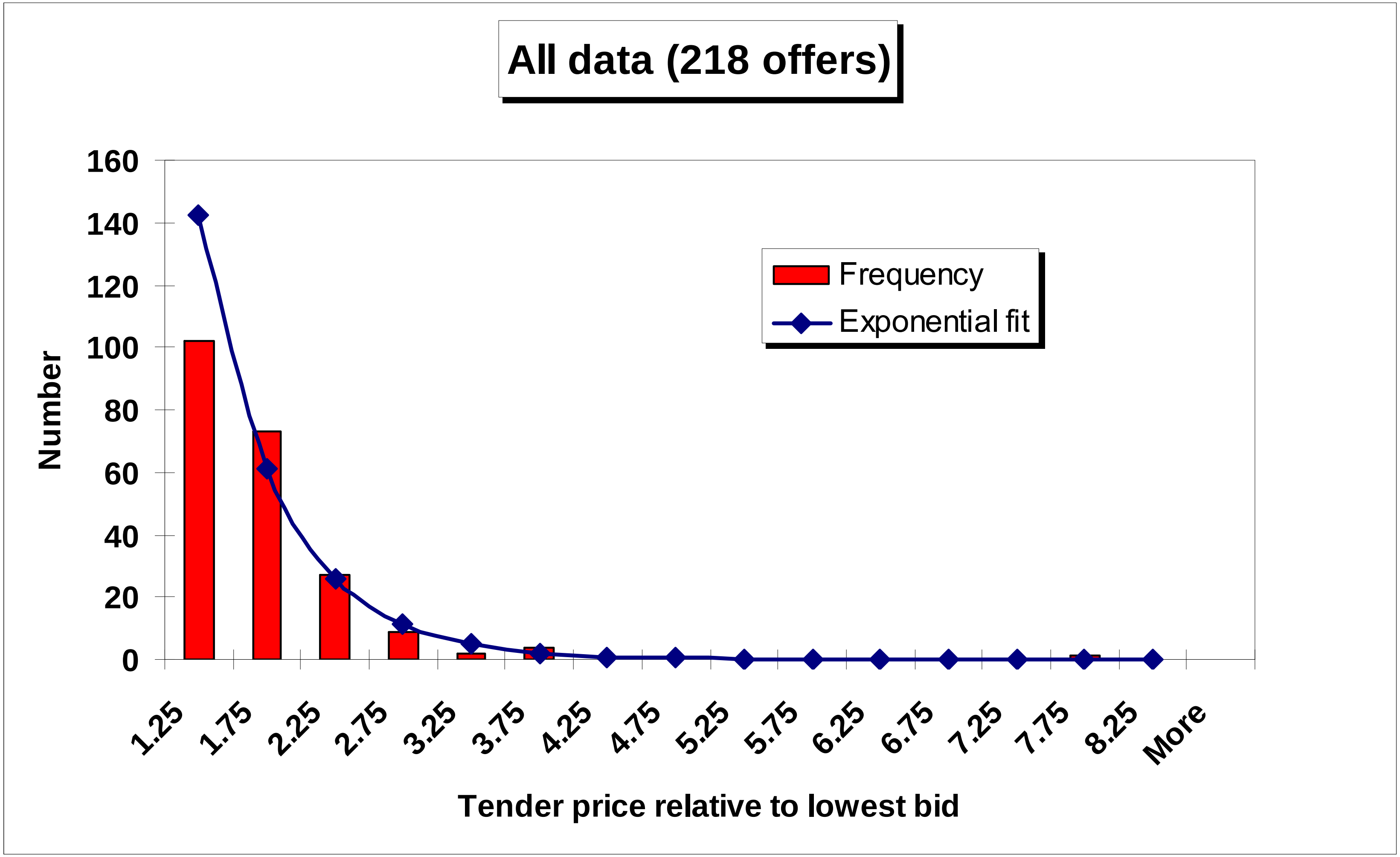
Not addressed here

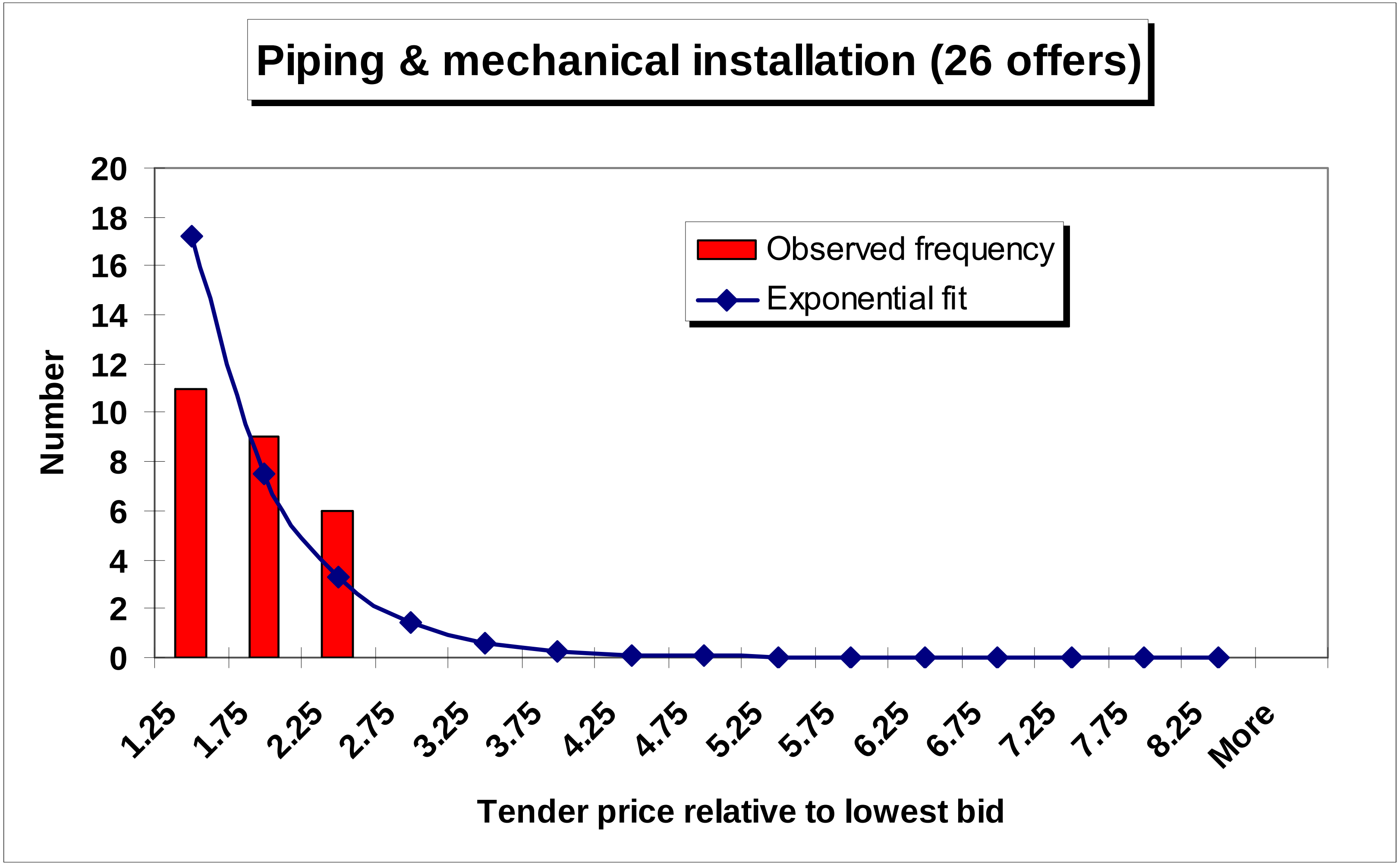
Coped for in tender price variance

Deterministic & compensated, not addressed here

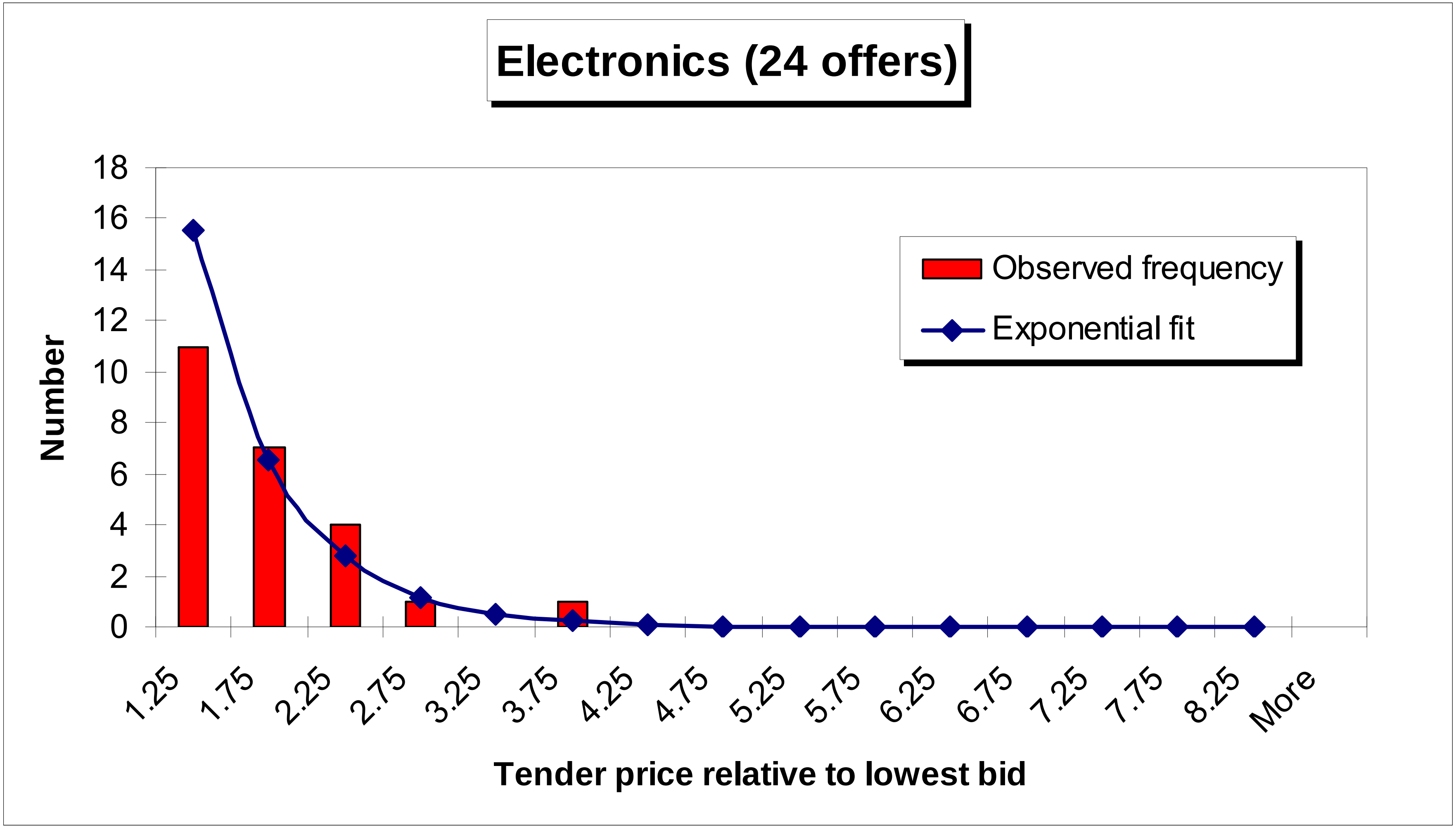
Scatter of LHC offers as a measure of cost variance

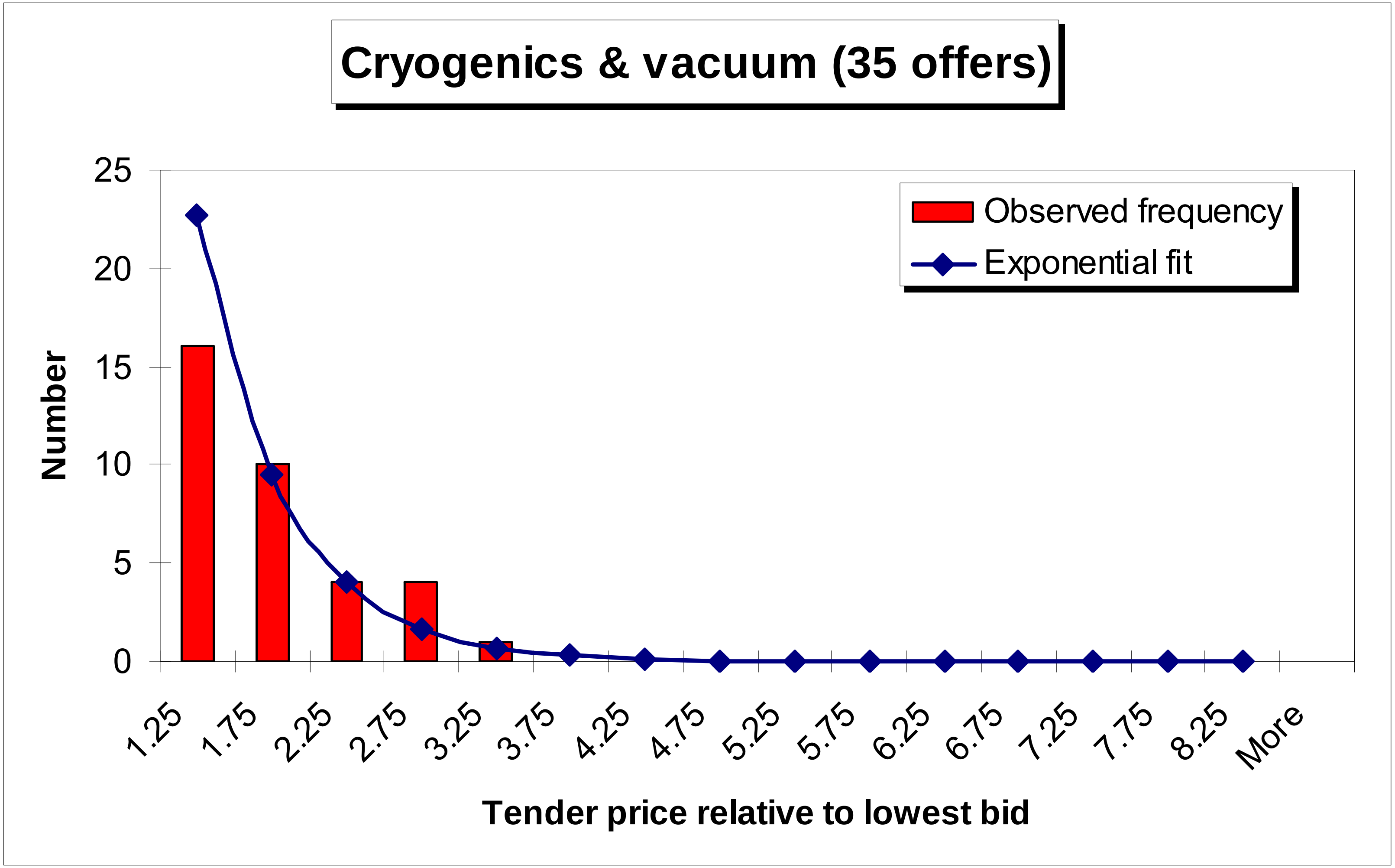
- Available data: CERN purchasing rules impose to procure on the basis of lowest valid offer ⇒ offers ranked by price with reference to lowest for adjudication by FC
- Postulate: scatter of (valid) offers received for procurement of LHC components is a measure of their cost variance due to technical, manufacturing and commercial aspects
- Survey of 218 offers for LHC machine components, grouped in classes of similar equipment
- Prices normalized to that of lowest valid offer, i.e. value of contract
- Exponential PDFs fitted to observed frequency distributions with same mean and standard deviation

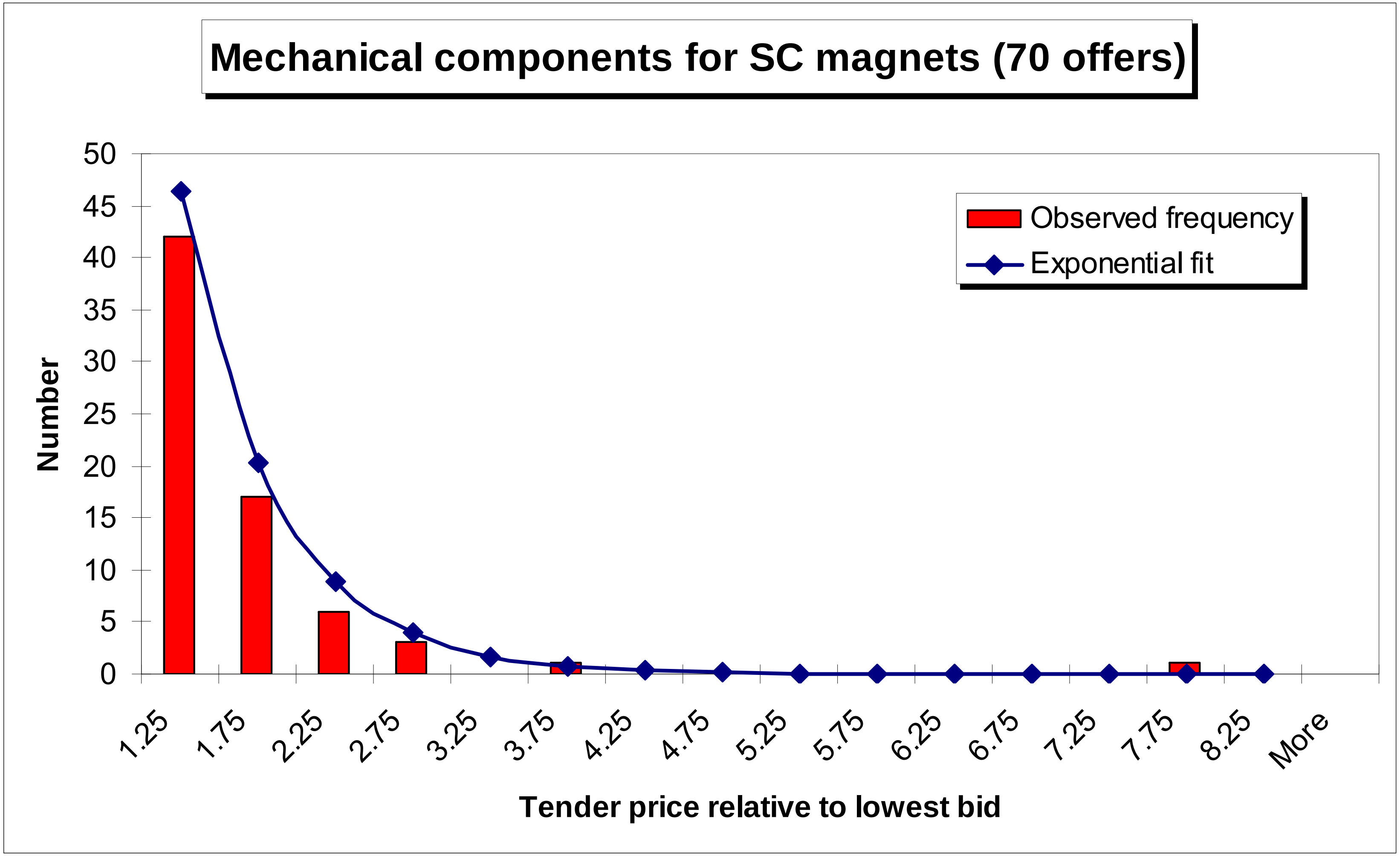

All data (218 offers)
160
140
120
100
80
60
40
20
0
Number
Frequency
Exponential fit
1.25
1.75
2.25
2.75
3.25
3.75
4.25
4.75
5.25
5.75
6.25
6.75
7.25
7.75
8.25
More
Tender price relative to lowest bid



Piping & mechanical installation (26 offers)
Number
20
18
16
14
12
10
8
6
4
2
0
Observed frequency
Exponential fit
1.25
1.75
2.25
2.75
3.25
3.75
4.25
4.75
5.25
5.75
6.25
6.75
7.25
7.75
8.25
More
Tender price relative to lowest bid


Electronics (24 offers)
Number
18
16
14
12
10
8
6
4
2
0
1.25
1.75
2.25
2.75
3.25
3.75
4.25
4.75
5.25
5.75
6.25
6.75
7.25
7.75
8.25
More
Tender price relative to lowest bid
Observed frequency
Exponential fit


Cryogenics & vacuum (35 offers)
Observed frequency
Exponential fit
Number
25
20
15
10
5
0
1.25
1.75
2.25
2.75
3.25
3.75
4.25
4.75
5.25
5.75
6.25
6.75
7.25
7.75
8.25
More
Tender price relative to lowest bid



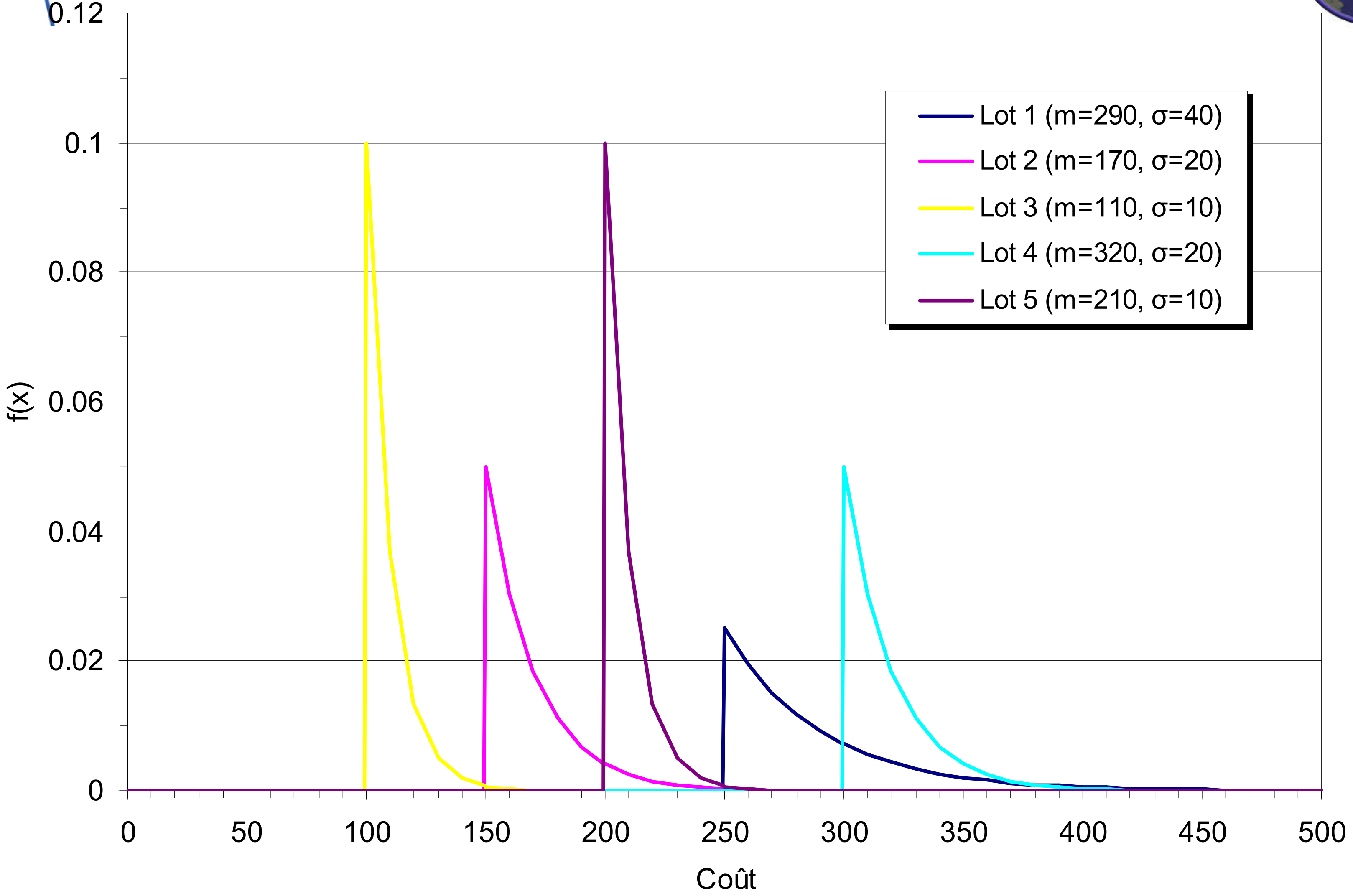
Mechanical components for SC magnets (70 offers)
Observed frequency
Exponential fit
Number
50
45
40
35
30
25
20
15
10
5
0
1.25
1.75
2.25
2.75
3.25
3.75
4.25
4.75
5.25
5.75
6.25
6.75
7.25
7.75
8.25
More
Tender price relative to lowest bid

A simple worked-out example

- Consider a project made of 5 lots according to the table below

Lot	Seuil	Ecart-type	Moyenne	Variance
1	250	40	290	1600
2	150	20	170	400
3	100	10	110	100
4	300	20	320	400
5	200	10	210	100
Somme	1000	100	1100	2600
Sigma				50.9901951

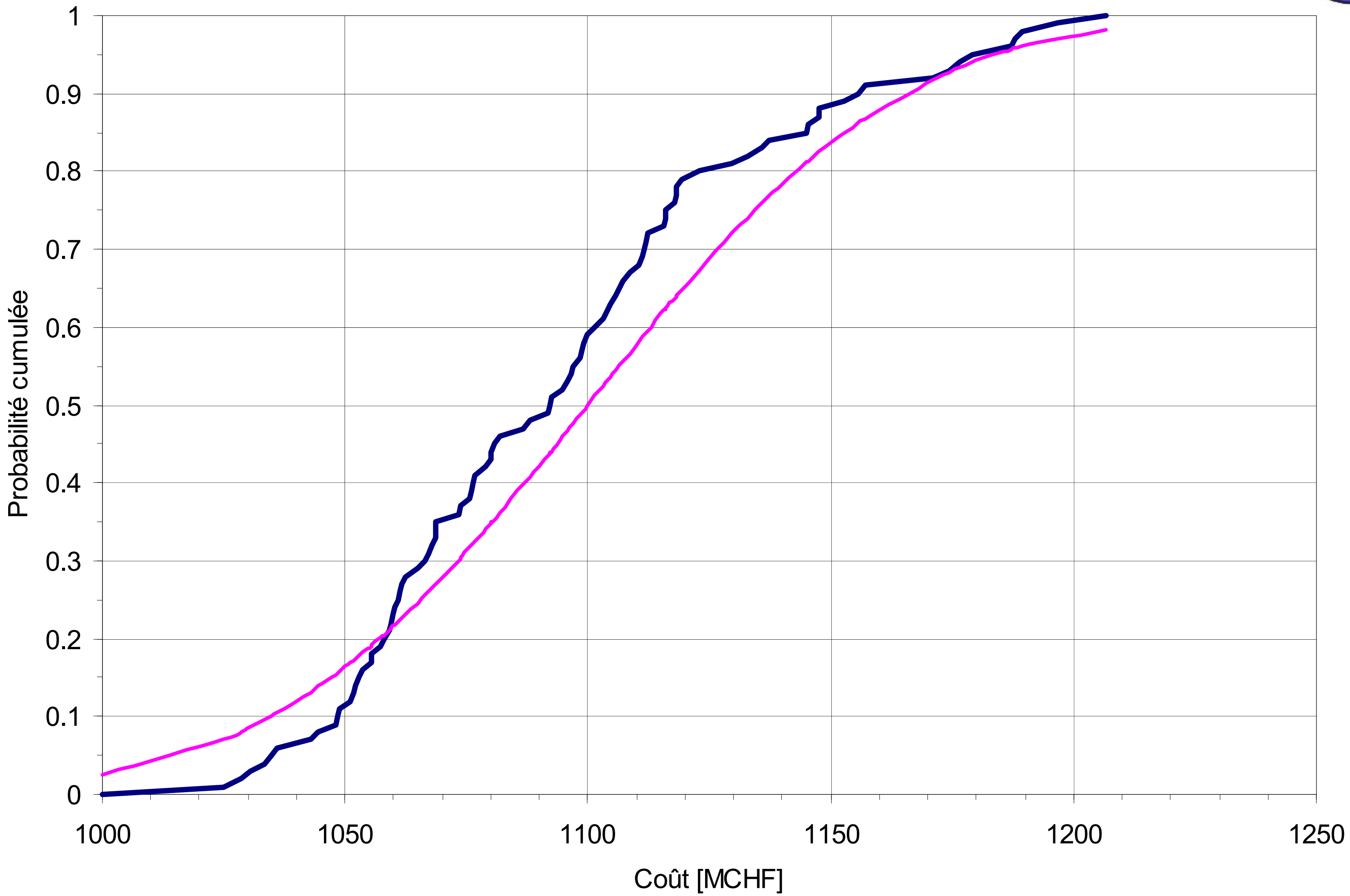
Densités de probabilité du coût des lots (lois exponentielles)



Application of central-limit theorem

- In case the elementary costs are statistically independent, the total cost is a random variable with
  - mean value $m = \Sigma\, m_i = 1100$
  - standard deviation $\sigma = (\Sigma\, \sigma_i^2)^{½} \approx 51$
- Its PDF tends towards a Gaussian law [1100, 51]
  - $X \leq 1165$ at confidence level 90%
  - $X \leq 1184$ at confidence level 95%
  - $X \leq 1205$ at confidence level 98%
- This law can be compared to the result of a Monte-carlo simulation based on exponential PDFs for elementary costs, treated as independent

Fonction de distribution du coût total du projet
(simulation Monte Carlo n = 100, comparée à une loi normale [1100, 51])



Conclusion: proposed procedure for probabilistic cost analysis

- Identify sources of cost variance and separate deterministic effects
- Identify correlated random effects and estimate their standard deviations (not to be added quadratically!)
- Estimate mean value and standard deviation of independant elementary costs and modelize by simple skew law, e.g. exponential
- Apply central-limit theorem and/or Monte-Carlo on sum of independant elementary costs
- Apply uncertainty due to correlated random effects on previous result
- Apply compensation of deterministic effects by established factors (e.g. currency exchange rates & industrial price indices)