Gaussian mixtures and financial returns

Carlos Cuevas Covarrubias,
Anahuac University, Mexico.

Jorge Rosales Contreras,
ITESM, Mexico.

CIEMA
CENTRO DE INVESTIGACIÓN EN ESTADÍSTICA Y MATEMÁTICAS APLICADAS

Motivation

ferent models in mathematical finance assume that the returns of ts are independent and normally distributed (*Markowitz and P-M, Black-Scholes, etc.*).

pirical evidence suggests that financial returns do not always ow a normal distribution. The reason:

- Heavier tails than the normal can capture,
- Skewed and leptokurtic distributions,
- Alternating periods of high and low volatility.

Evidence for a hyperbolic-like distribution of asset returns drawn from a simple economical financial markets model

Stefan Reimann

sk management and asset pricing benefit from simple nctional descriptions of the distribution of real asset returns...

e propose a simple model of price formation and return tribution ...

mulations of the price dynamics generate returns with fat-tail havior in line with that of a hyperbolic distribution, and also latility clustering, which is a mayor stylized fact of asset urns.

FINRISK
National Centre of Competence in Research
ncial Valuation and Risk Management

Working Paper No. 212
March 2005

odeling Marginal Distributions of Ten European Stock Market Index Returns

Andreas Behr

e model daily stock index returns for ten European countries ing flexible distributions: generalized hyperbolic distribution, ixtures of Gaussian normal distributions and the logF stribution…

ur results show that the empirical distribution of the daily ock index returns can be adequately modelled by either of e three flexible approaches but Gaussian mixtures are ghtly superior for all countries in almost all sub periods.

ernational Research Journal of Finance and Economics
SN 1450-2887 Issue 28 (2009)

p://www.eurojournals.com/finance.htm

NON-NORMALITY FACTS AND FALLACIES

David N. Esch

ntly, there has been an increasing trend in quantitative ce to call for models which are model returns with non- al probability distributions…

urvey the shortcomings of several popular non-normal cial modeling techniques, … we argue that it is still important haust the possibilities of normal models before switching to …

els with normal distributions can be extended through ods such as conditioning on other variables, inequality traints, mixtures…

JOURNAL OF INVESTMENT MANAGEMENT, Vol. 8, No. 1, (2010), pp. 49–61

TIMATION OF PORTFOLIO RETURN AND VALUE AT RISK ING A CLASS OF GAUSSIAN MIXTURE DISTRIBUTIONS

Kangrong Tan and Meifen Chu

t return distributions are frequently assumed to follow a al or lognormal distribution. ... Many empirical studies have n that return distributions are usually not normal...

ropose a class of Gaussian mixture distributions to oximate the return distributions of assets. The model is ed easily to estimate the return distribution of a portfolio, and aluate the VaR.

The International Journal of Business and Finance Research
Volume 6 ♦ Number 1 ♦ 2012

Finite Mixture Models

The random variable

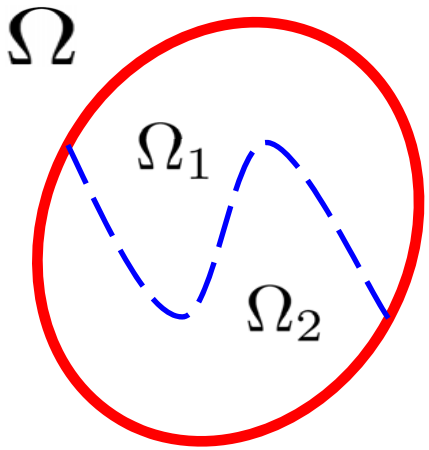
$$X : \Omega \rightarrow \mathbb{R}$$

is distributed according to a mixture of densities when

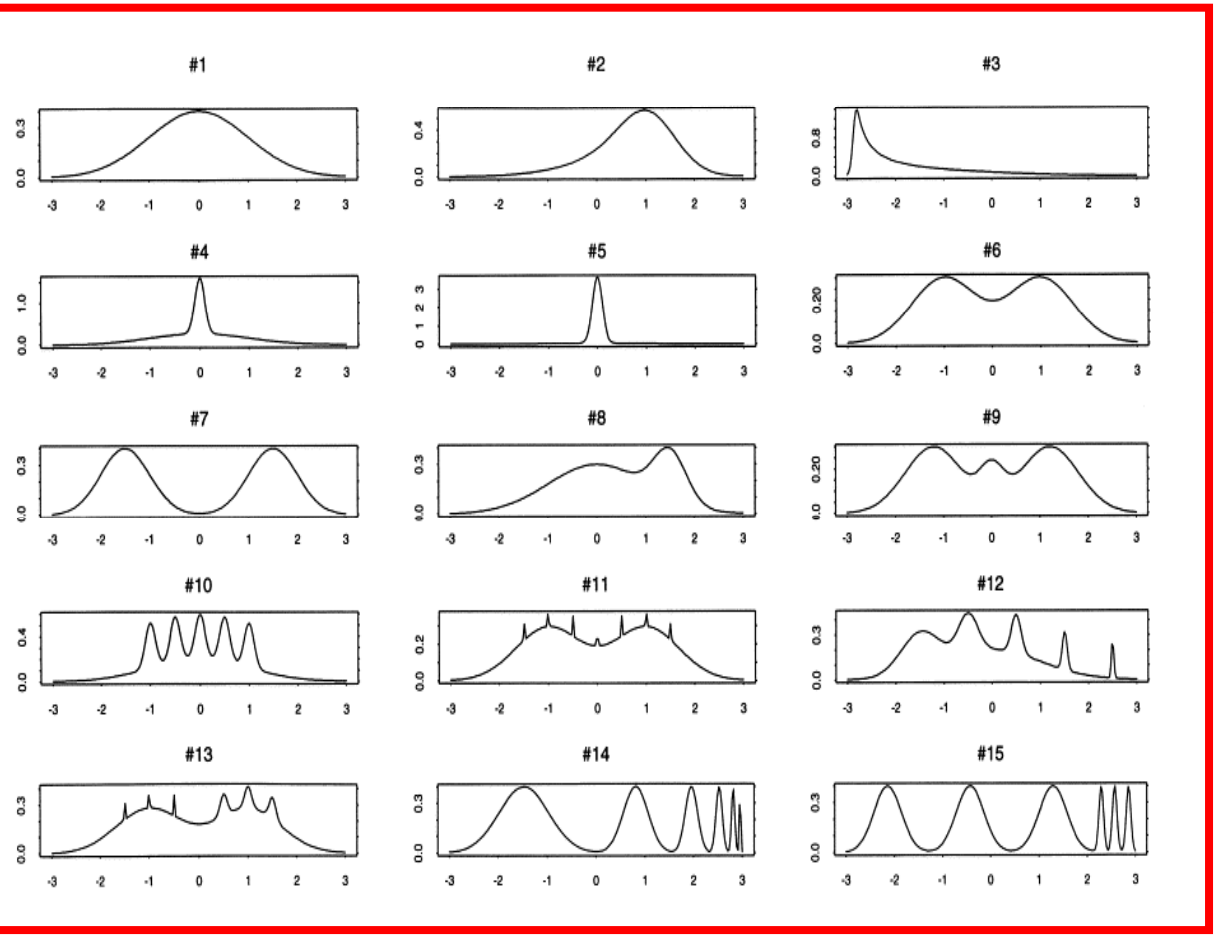
$$f_X(x) = \sum_{j=1}^{k} \alpha_j f_j(x)$$

Finite Mixture Models

1. $f_X(x) = \sum_{j=1}^{k} \alpha_j f_j(x)$
2. $\sum_{j=1}^{k} \alpha_j = 1; \quad \alpha_j \in (0,1)$
3. f_j is a density function.
4. $\Omega = \bigcup_{j=1}^{k} \Omega_j$
5. $\alpha_j = P(\Omega_j)$
6. $P[\Omega_j | X = x] = \frac{\alpha_j f_j(x)}{\sum_{h=1}^{k} \alpha_h f_h(x))}$



Finite Mixture Models



CIEMA

TECNOLOGICO

Estimation of parameters

Maximum Likelihood.

$$L(P) = \prod_{i=1}^{n} \sum_{j=1}^{k} \alpha_j f_j(x_i)$$

$$\pounds(P) = \ln L(P) = \sum_{i=1}^{n} \ln(\sum_{j=1}^{k} \alpha_j f_j(x_i))$$

Estimation of parameters

The log-likelihood is a function of at least *2k – 1* parameters.

$$\pounds(P) = \ln L(P) = \sum_{i=1}^{n} \ln(\sum_{j=1}^{k} \alpha_j f_j(x_i))$$

The estimated parameters are obtained by :

$$M\acute{a}x\ \pounds(P); \quad s.a. \sum_{j=1}^{k} \alpha_j = 1$$

Estimation of parameters

$$Máx\ \pounds(P); \quad s.a. \sum_{j=1}^{k} \alpha_j = 1$$

If we apply Lagrange multipliers…

$$\sum_{i=1}^{n} \frac{f_h'(x_i)}{\sum_{j=1}^{k} \alpha_j f_j(x_i)} = 0 \qquad \text{for} \qquad h = 1, 2, \ldots k$$

$$\sum_{i=1}^{n} \frac{f_h(x_i)}{\sum_{j=1}^{k} \alpha_j f_j(x_i)} = \lambda \qquad \text{for} \qquad h = 1, 2, \ldots k$$

$$\sum_{j=1}^{k} \alpha_j = 1$$

It is a system of *2k+1* equations

CIEMA

Expectation-Maximization

The system is solved by an iterative algorithm (EM).

1. *Expectation step*

$$\alpha_h = \frac{1}{n}\sum_{i=1}^{n} P[\Omega_h | X = x_i] \quad \text{par a} \quad h = 1, 2, \ldots k$$

2. *Maximization step*

$$\sum_{i=1}^{n} P[\Omega_h | X = x_i] \frac{\partial}{\partial \theta_h} \ln f_h(x_i) = 0 \quad \text{para} \quad h = 1, 2, \ldots k$$

Gaussian Mixtures

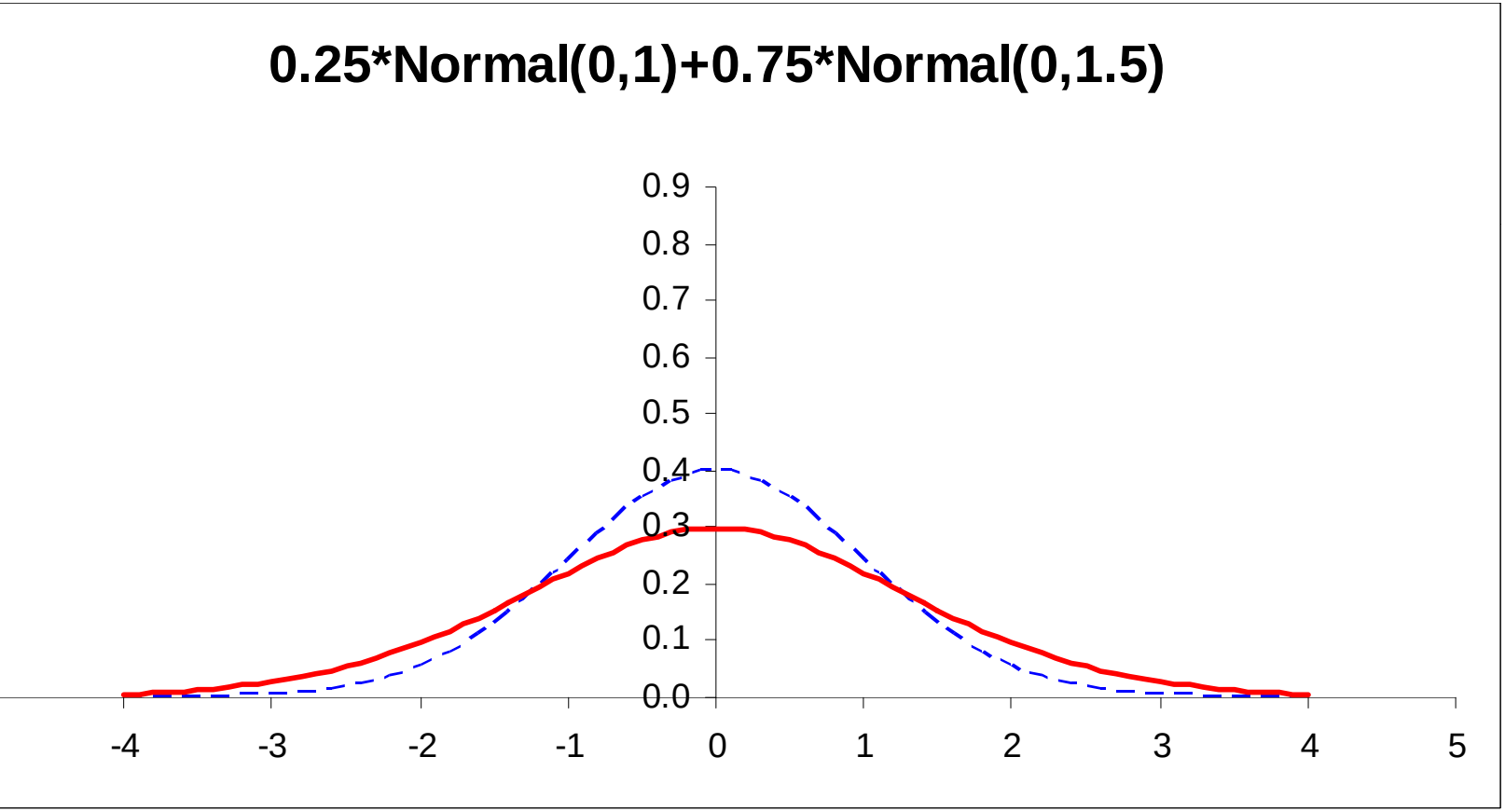
The random variable $X : \Omega \longrightarrow \mathbb{R}$

distributes according to a Gaussian mixture if

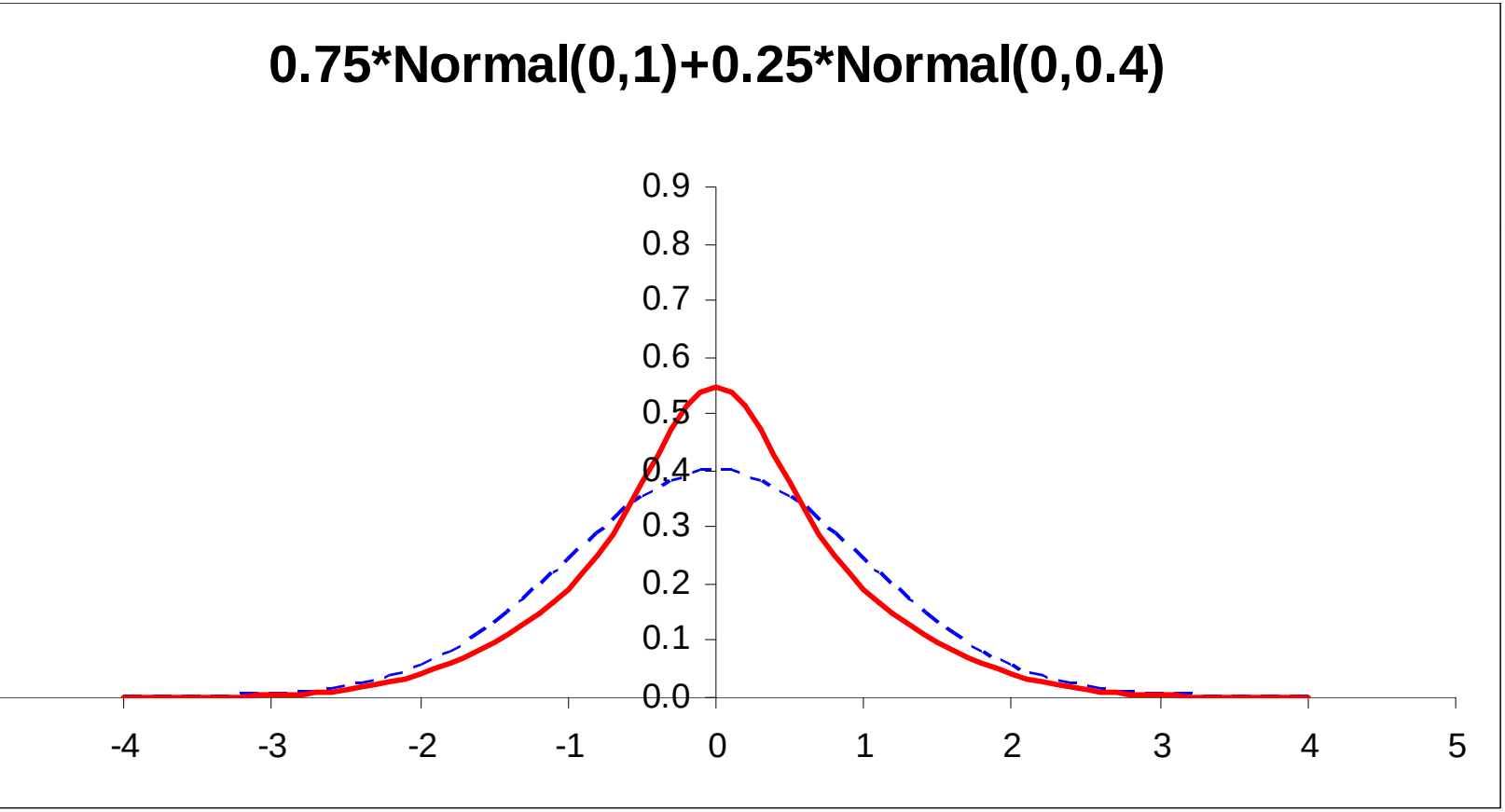
$$f_x(x, \psi) = \sum_{j=1}^{k} \frac{\alpha_j}{\sigma_j} \phi(\frac{x_i - \mu_j}{\sigma_j});$$

where $\phi(t) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{t^2}{2}}$

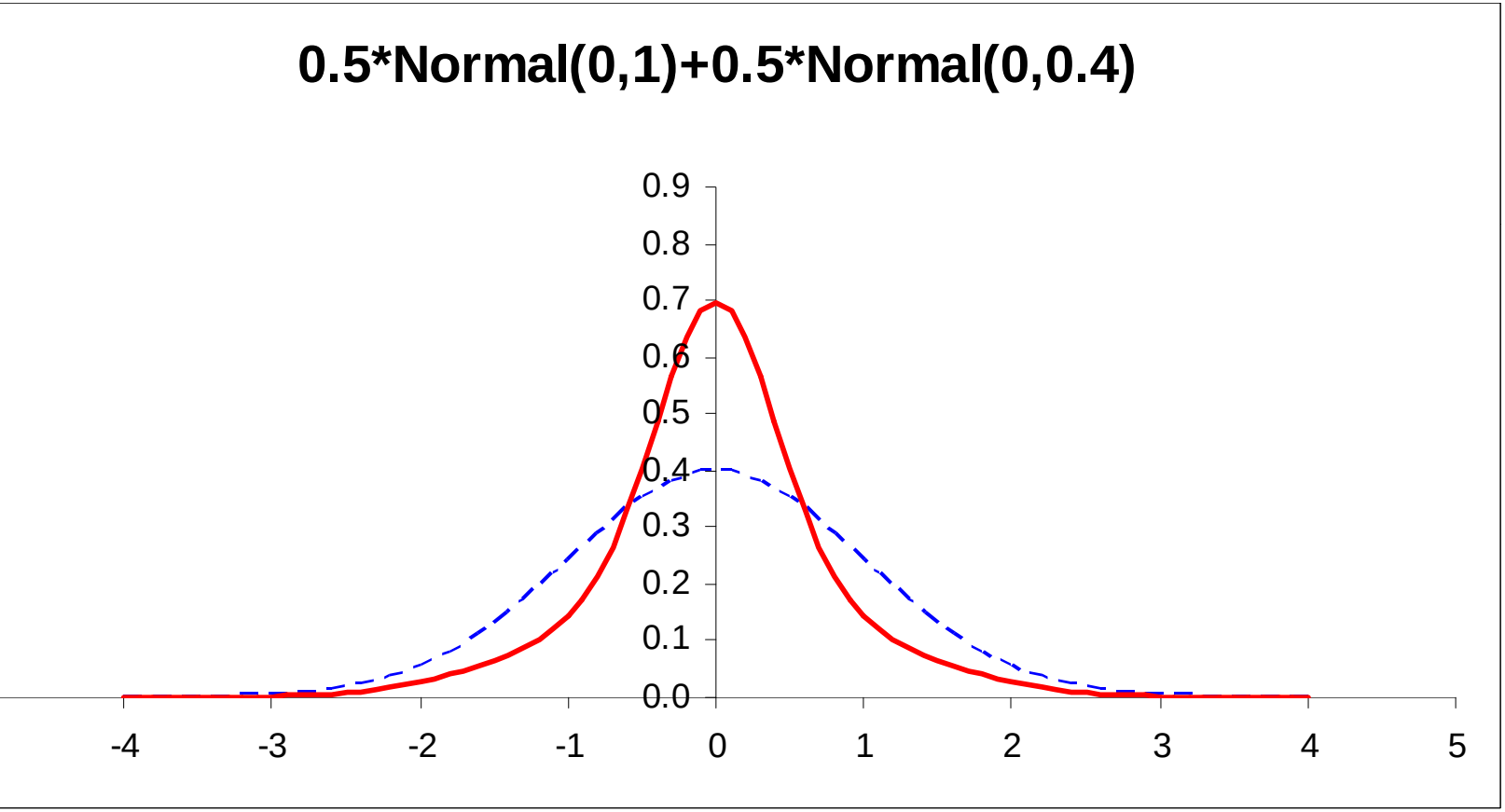
Gaussian Mixtures



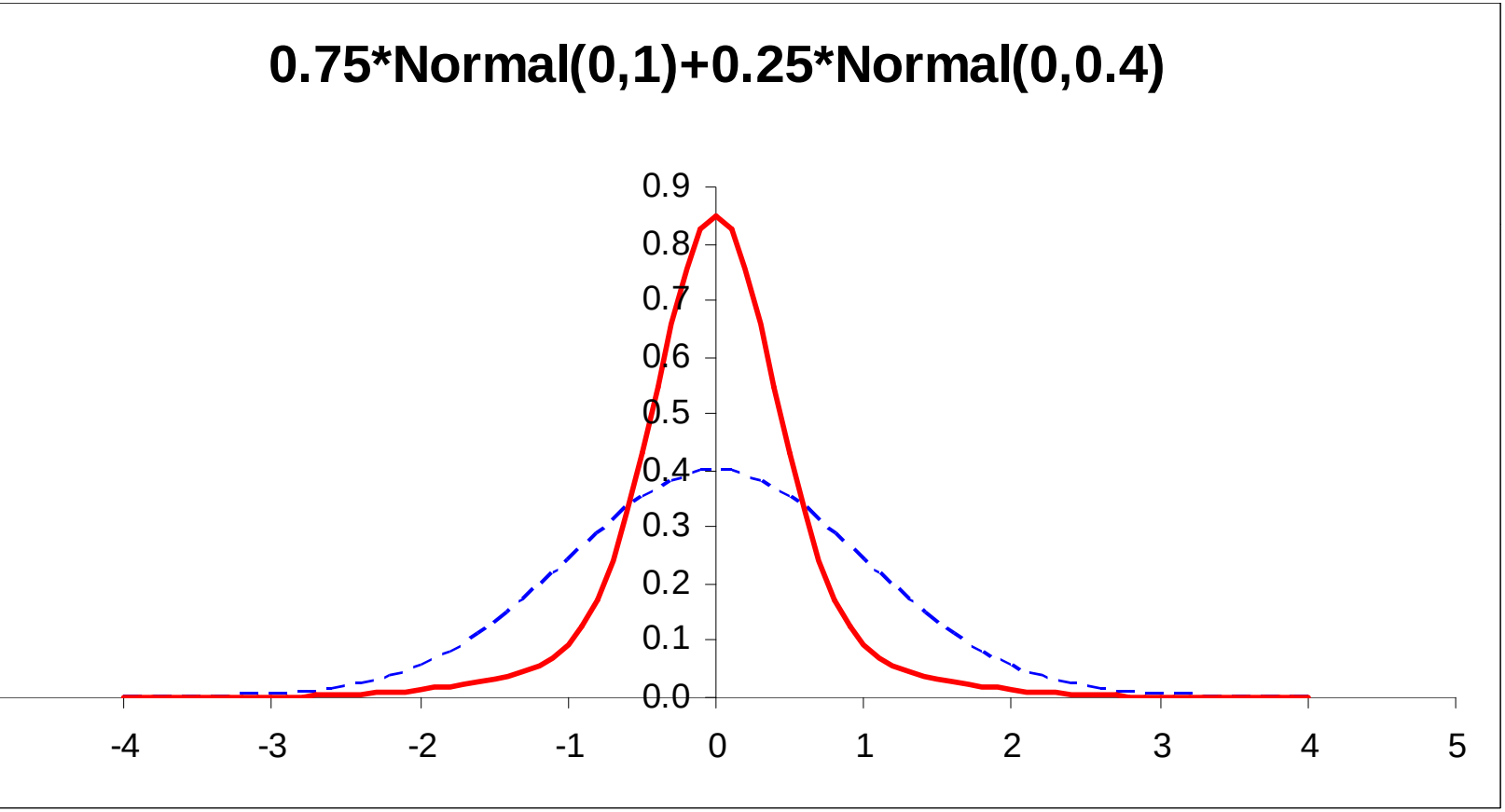
Gaussian Mixtures



Gaussian Mixtures



Gaussian Mixtures



Expectation - Maximization

iven a finite Gaussian mixture, the EM algorithm is deffined
s follows:

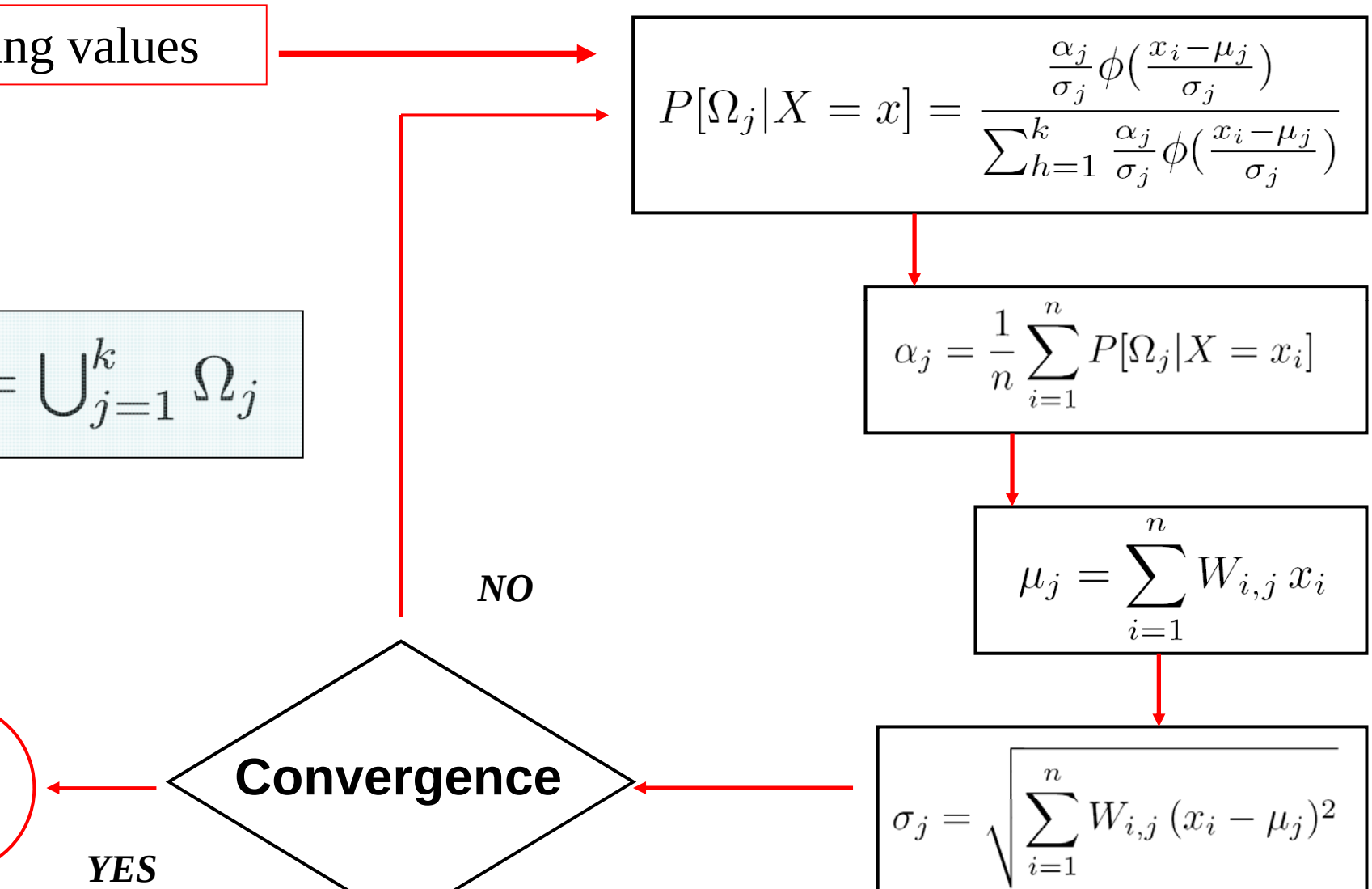
Expectation step

$$= \frac{1}{n}\sum_{i=1}^{n} P[\Omega_j|X=x_i] \qquad P[\Omega_j|X=x] = \frac{\frac{\alpha_j}{\sigma_j}\phi(\frac{x_i-\mu_j}{\sigma_j})}{\sum_{h=1}^{k}\frac{\alpha_j}{\sigma_j}\phi(\frac{x_i-\mu_j}{\sigma_j})}$$

Maximization step

$$\sum_{i=1}^{n}(\frac{P[\Omega_j|X=x_i]}{\sum_{h=1}^{n}P[\Omega_j|X=x_h]})x_i \qquad \sigma_j = \sqrt{\sum_{i=1}^{n}(\frac{P[\Omega_j|X=x_i]}{\sum_{h=1}^{n}P[\Omega_j|X=x_h]})(x_i-\mu_j)^2}$$

EM algorithm for Gaussian mixtures



IPC: stock market index in Mexico.

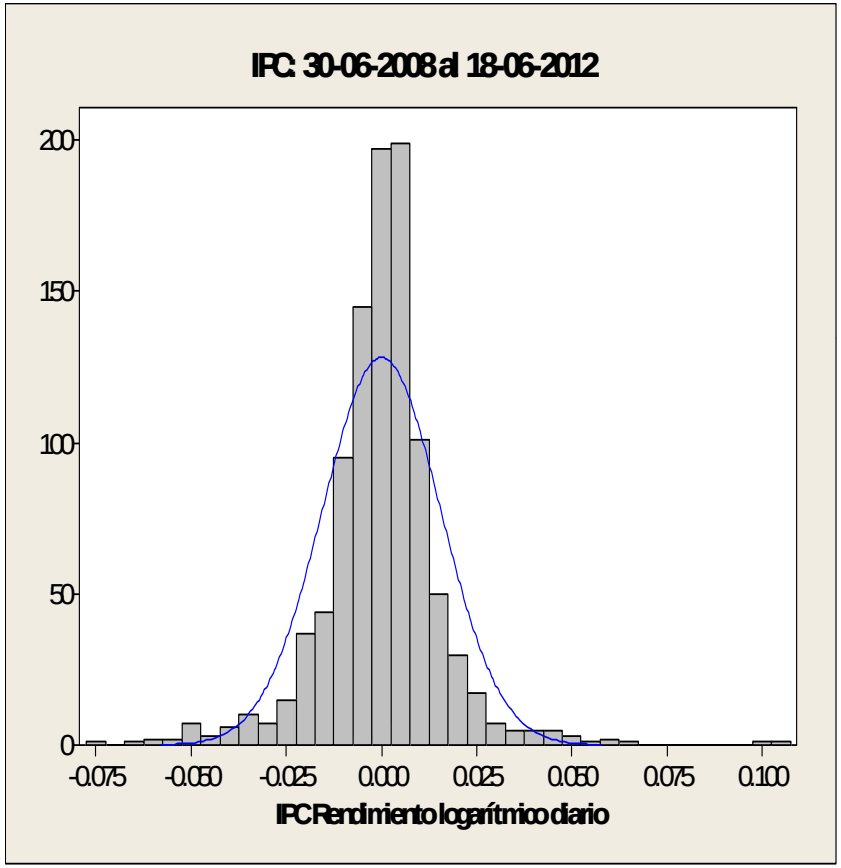
30-06-2008 to 18-06-2012

CIEMA

IPC: 30-06-2008 to 18-06-2012

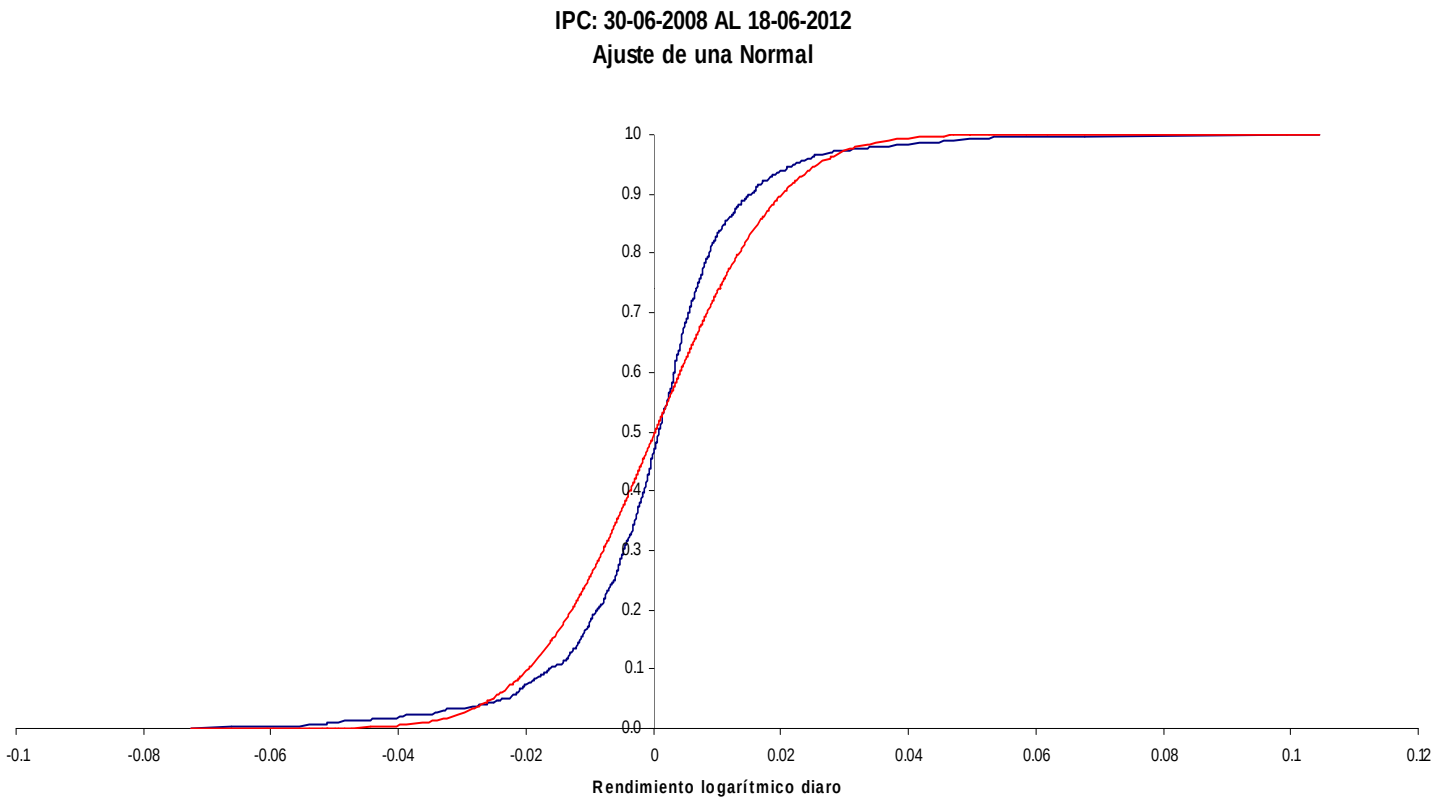
PC: Normal Distribution

```
riable       Mean      StDev
turn      0.000262  0.015553

S  = 0.09
V  < 0.01
```



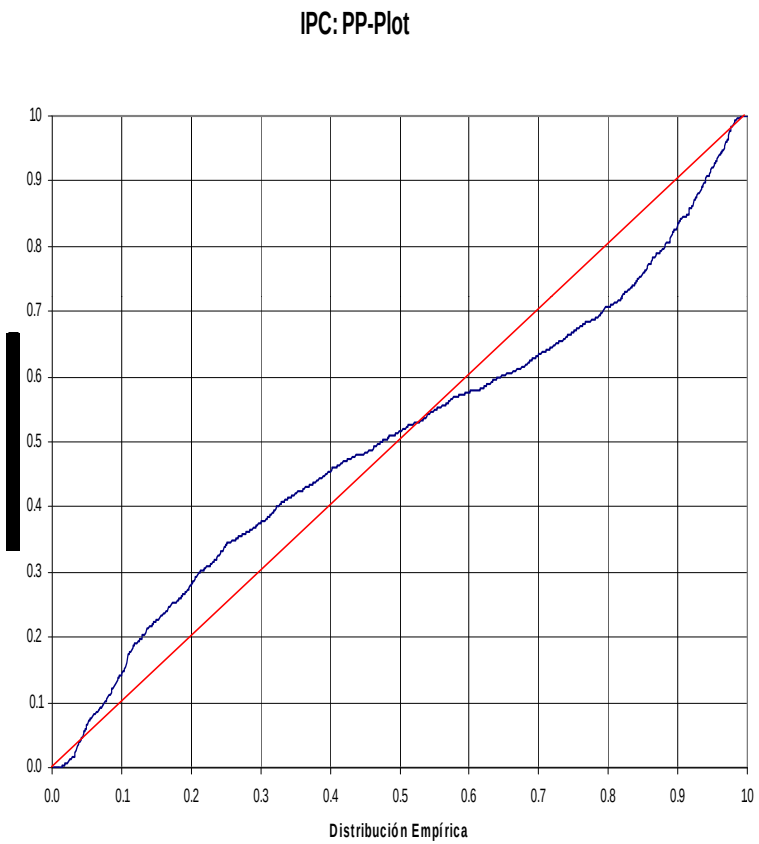
CIEMA

IPC: 30-06-2008 to 18-06-2012



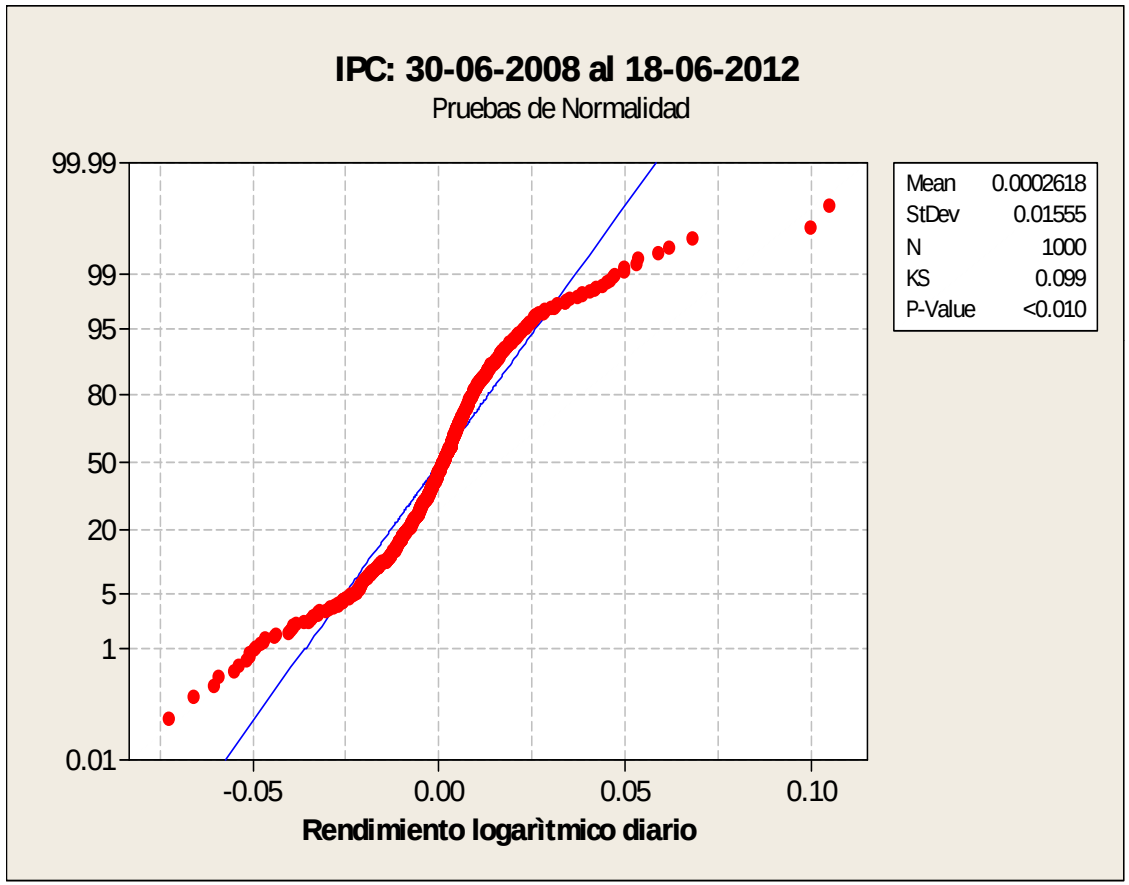
CIEMA

IPC: 30-06-2008 to 18-06-2012

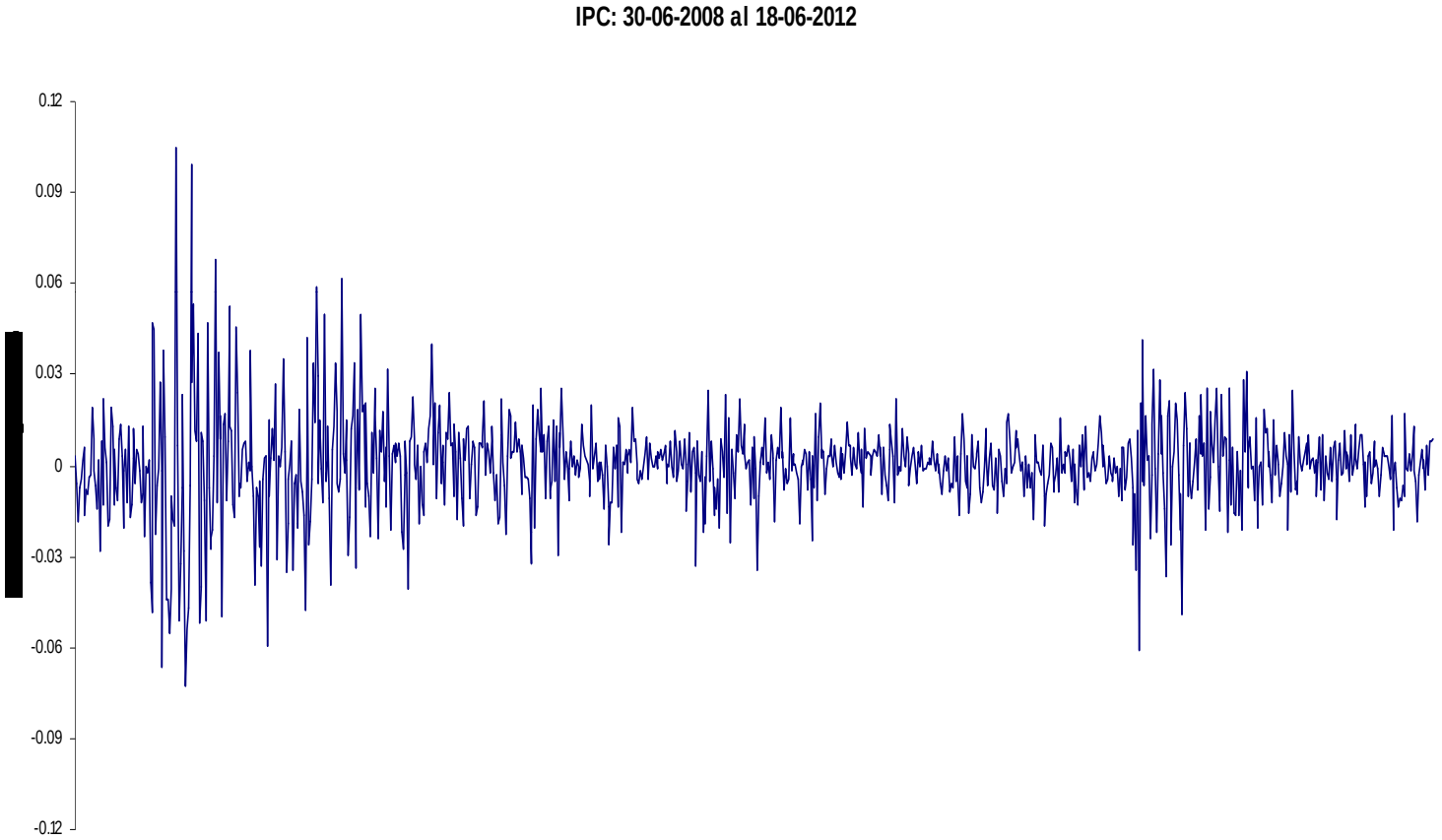


CIEMA

IPC: 30-06-2008 al 18-06-2012



IPC: 30-06-2008 to 18-06-2012



CIEMA

IPC: 30-06-2008 to 18-06-2012

Mixture of two normal distributions with different means and different variances.

$$f(x) = p\phi((x-\mu_1)/\sigma_1)/\sigma_1 + (1-p)\phi((x-\mu_2)/\sigma_2)/\sigma_2$$

Maximum Likelihood estimates:

- p = 0.2411
- S1 = 0.0084
- S0 = 0.0279
- M1 = 0.0008
- M0 = -0.0016
- K-S = 0.02
- P-Val > 0.1

CIEMA

IPC: 30-06-2008 to 18-06-2012

Mixture of two normal distributions with different means and different variances.

$$f(x) = p\phi((x-\mu_1)/\sigma_1)/\sigma_1 + (1-p)\phi((x-\mu_2)/\sigma_2)/\sigma_2$$

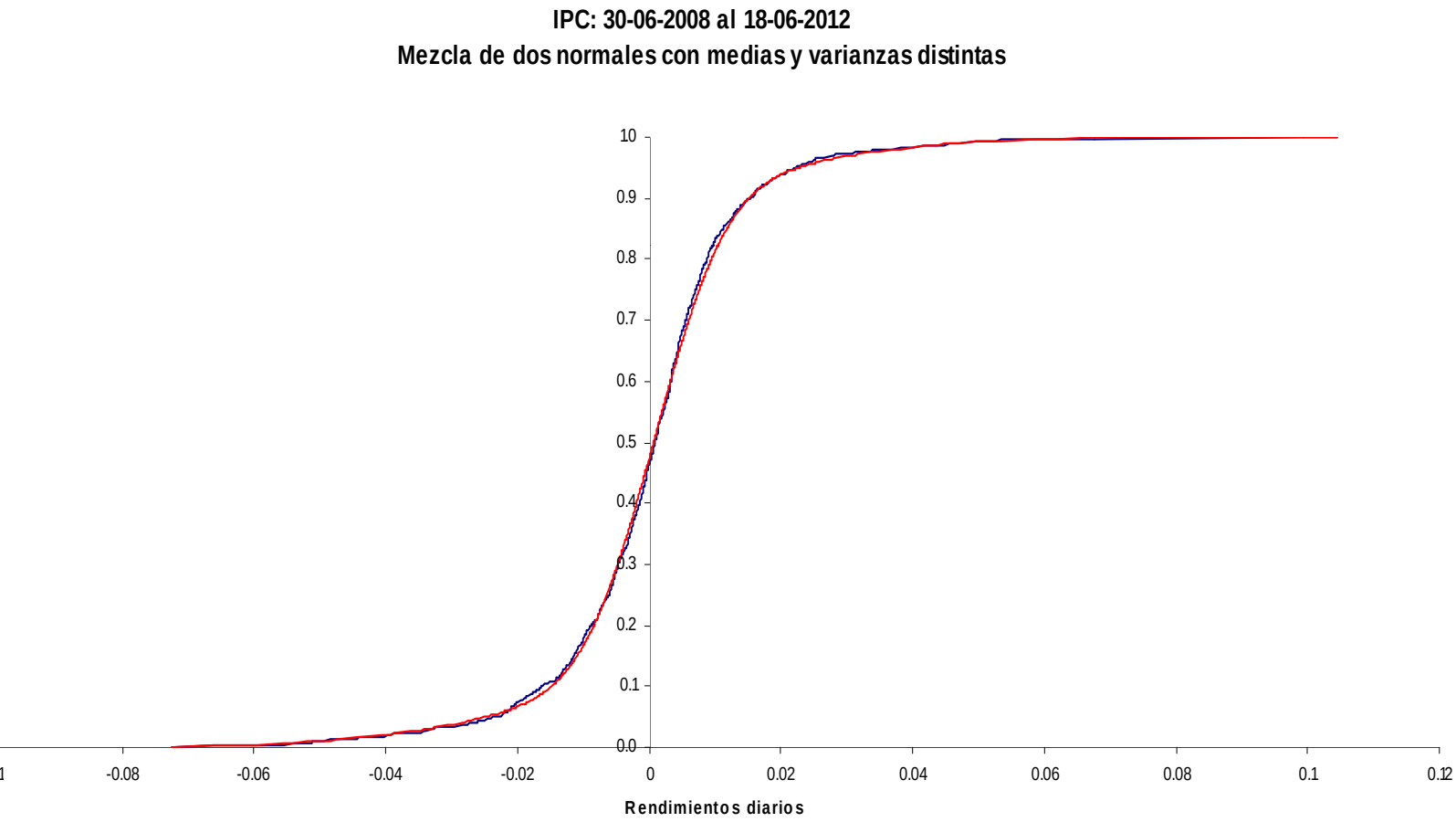
Maximum Likelihood estimates:

- p = 0.2411
- S1 = 0.0084
- S0 = 0.0279
- M1 = 0.0008
- M0 = -0.0016
- K-S = 0.02
- P-Val > 0.1

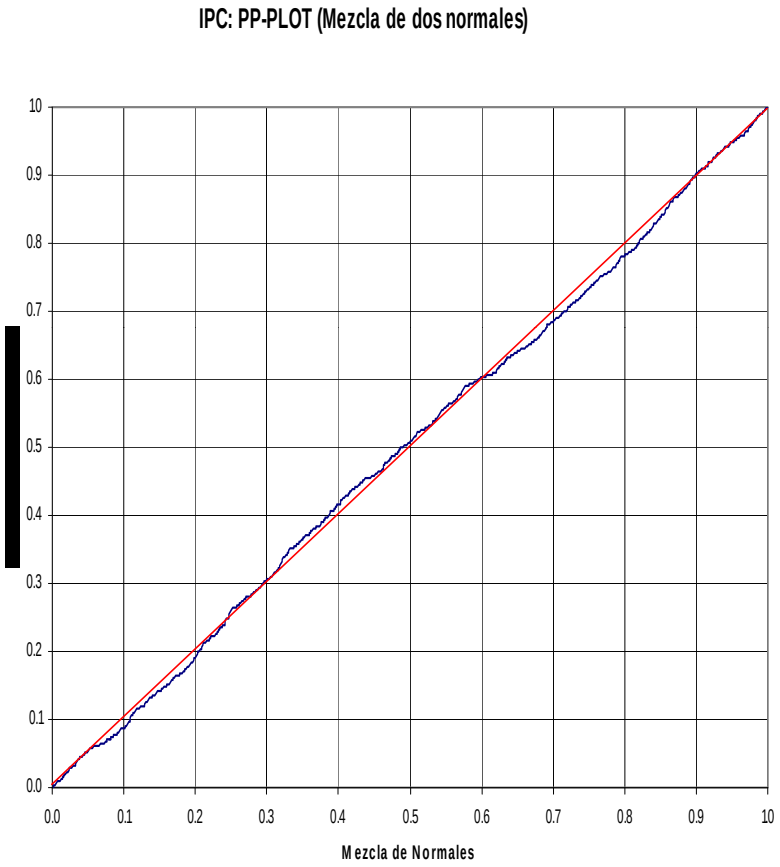
An expected loss comes with a larger volatility

CIEMA

IPC: 30-06-2008 to 18-06-2012



IPC: 30-06-2008 to 18-06-2012



USDMXN

01-01-2007 to 14-05-2012

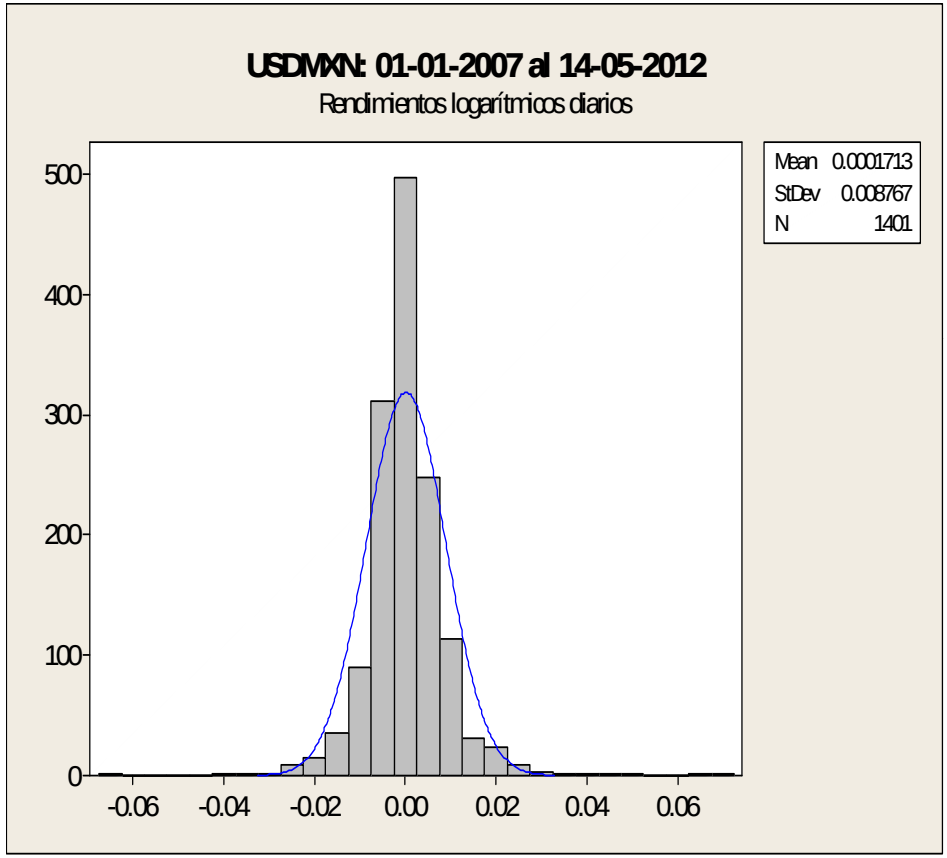
USDMXN: 01-01-2007 to 14-05-2012

al Distribution,
num Likelihood fit

= -0.0002

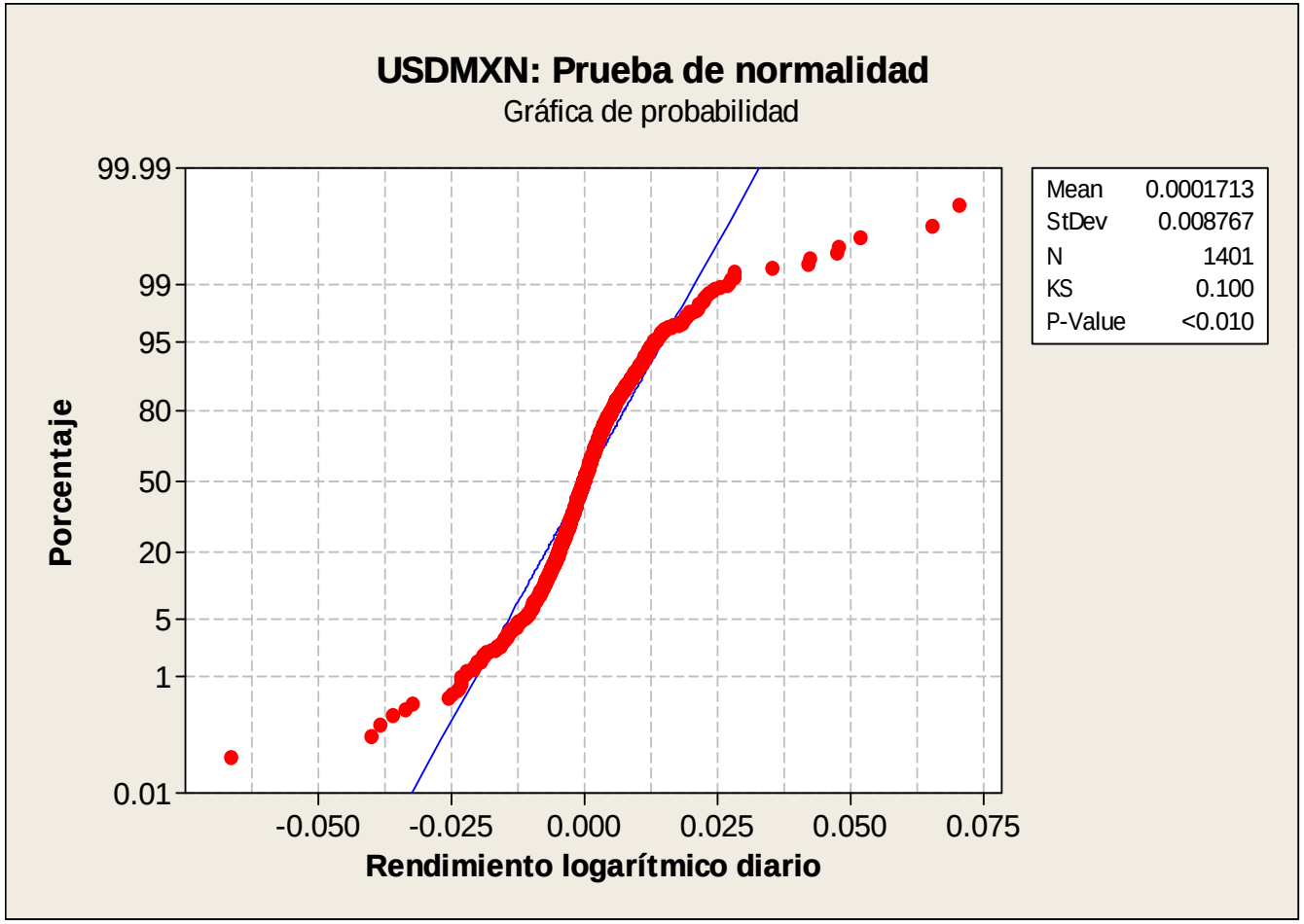
= 0.0087

= 0.10
< 0.01



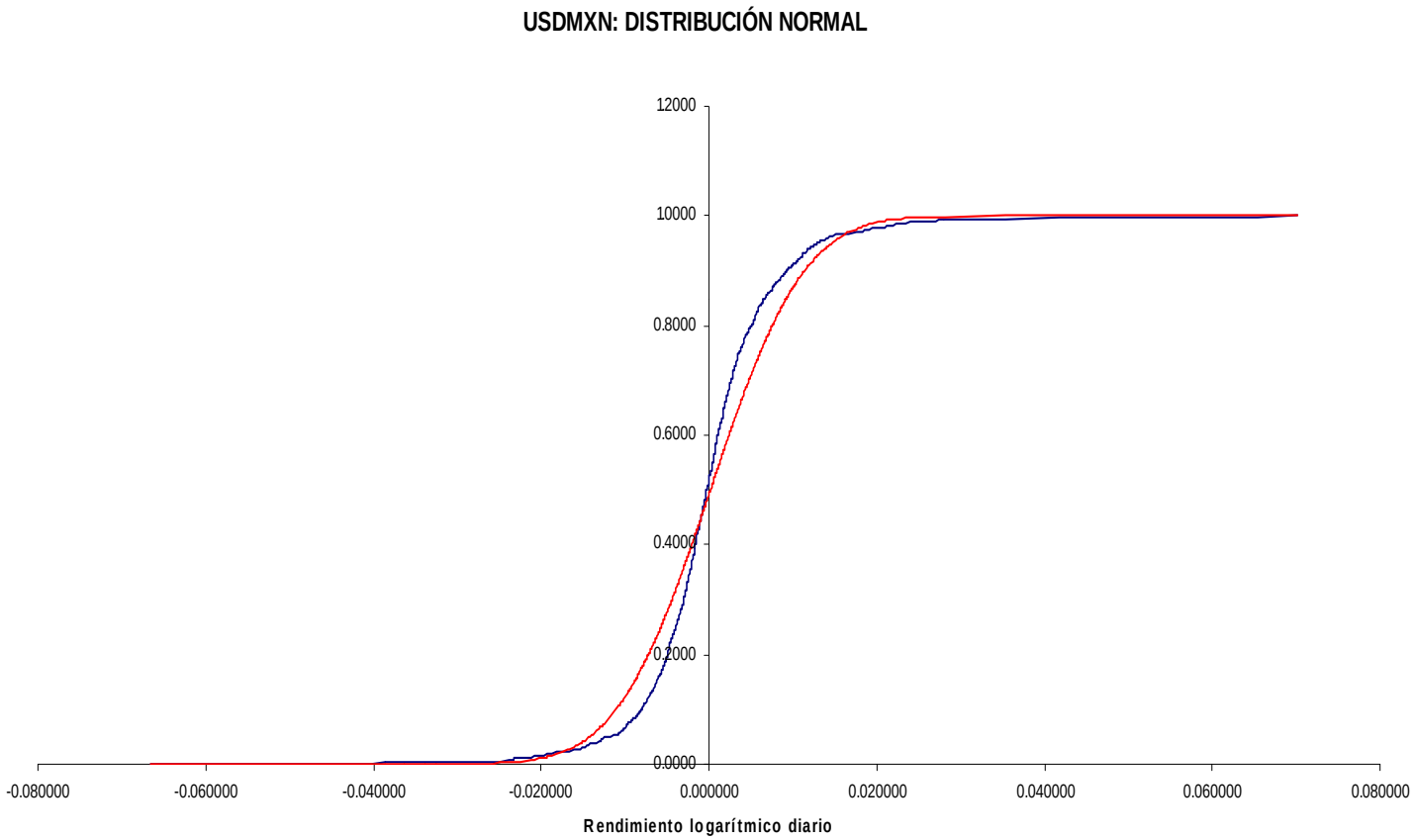
CIEMA

USDMXN: 01-01-2007 to 14-05-2012



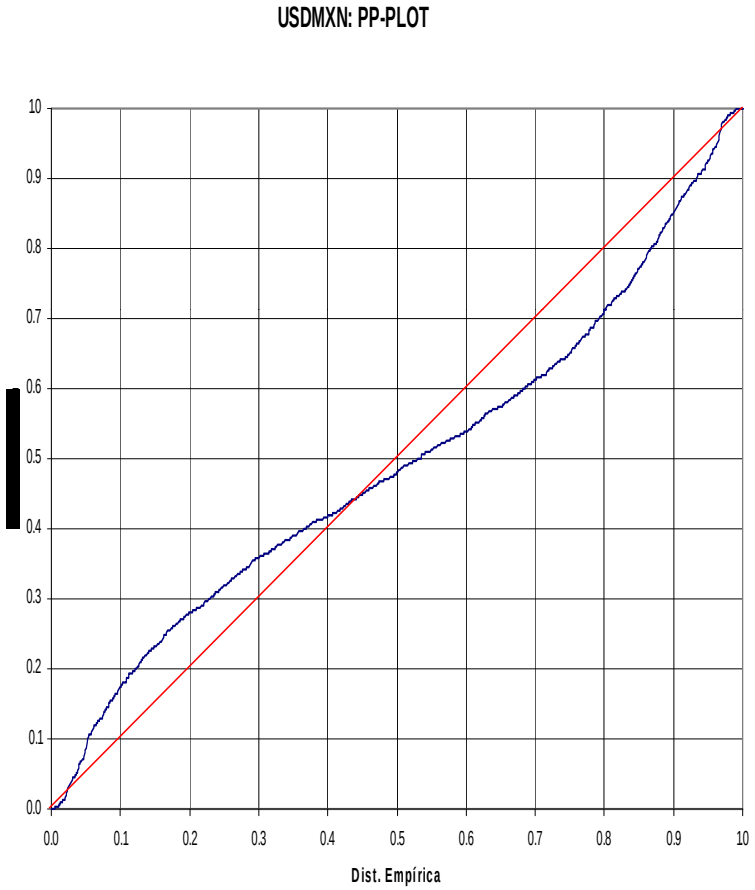
CIEMA

USDMXN: 01-01-2007 to 14-05-2012

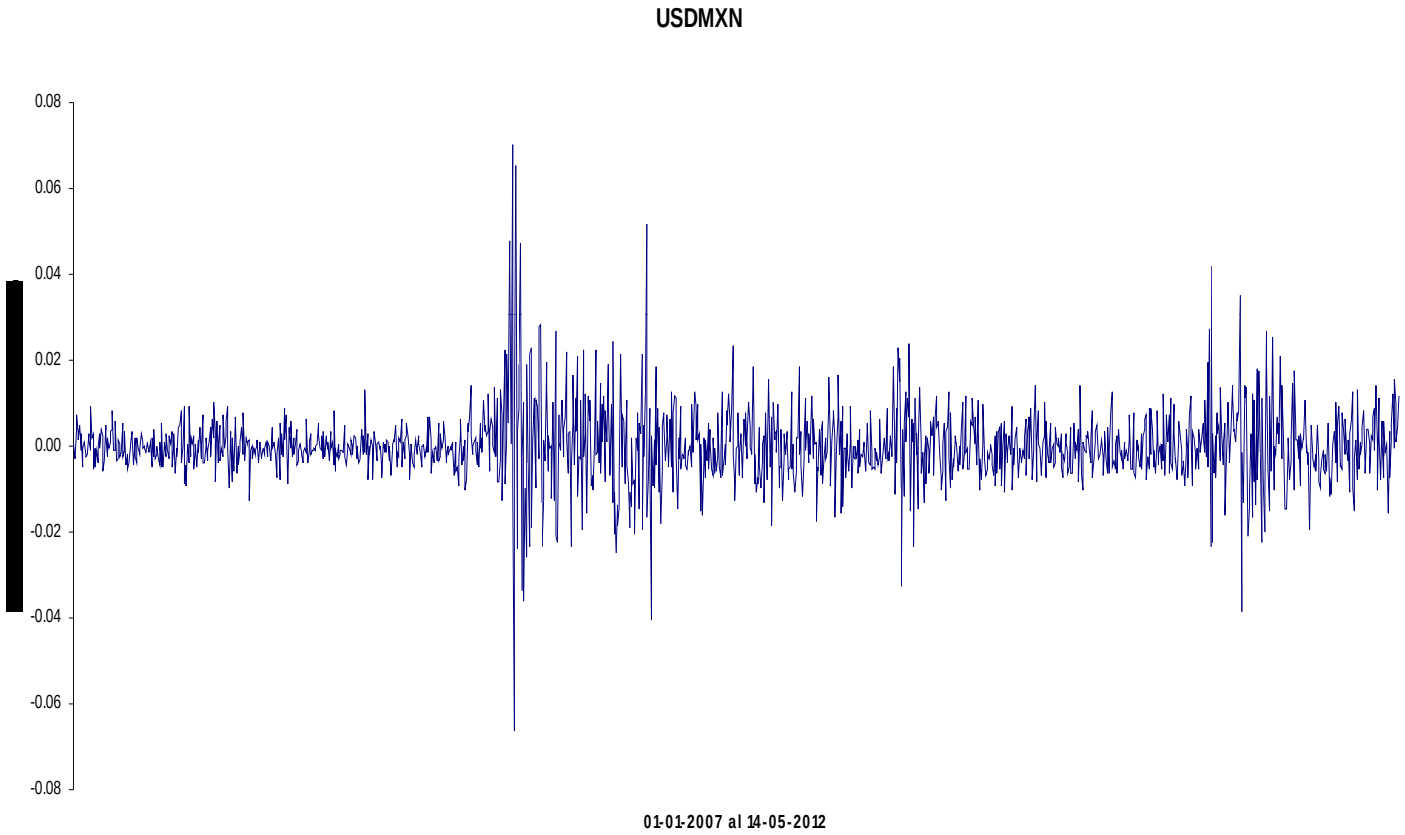


CIEMA

USDMXN: 01-01-2007 to 14-05-2012



USDMXN: 01-01-2007 to 14-05-2012



USDMXN: 01-01-2007 to 14-05-2012

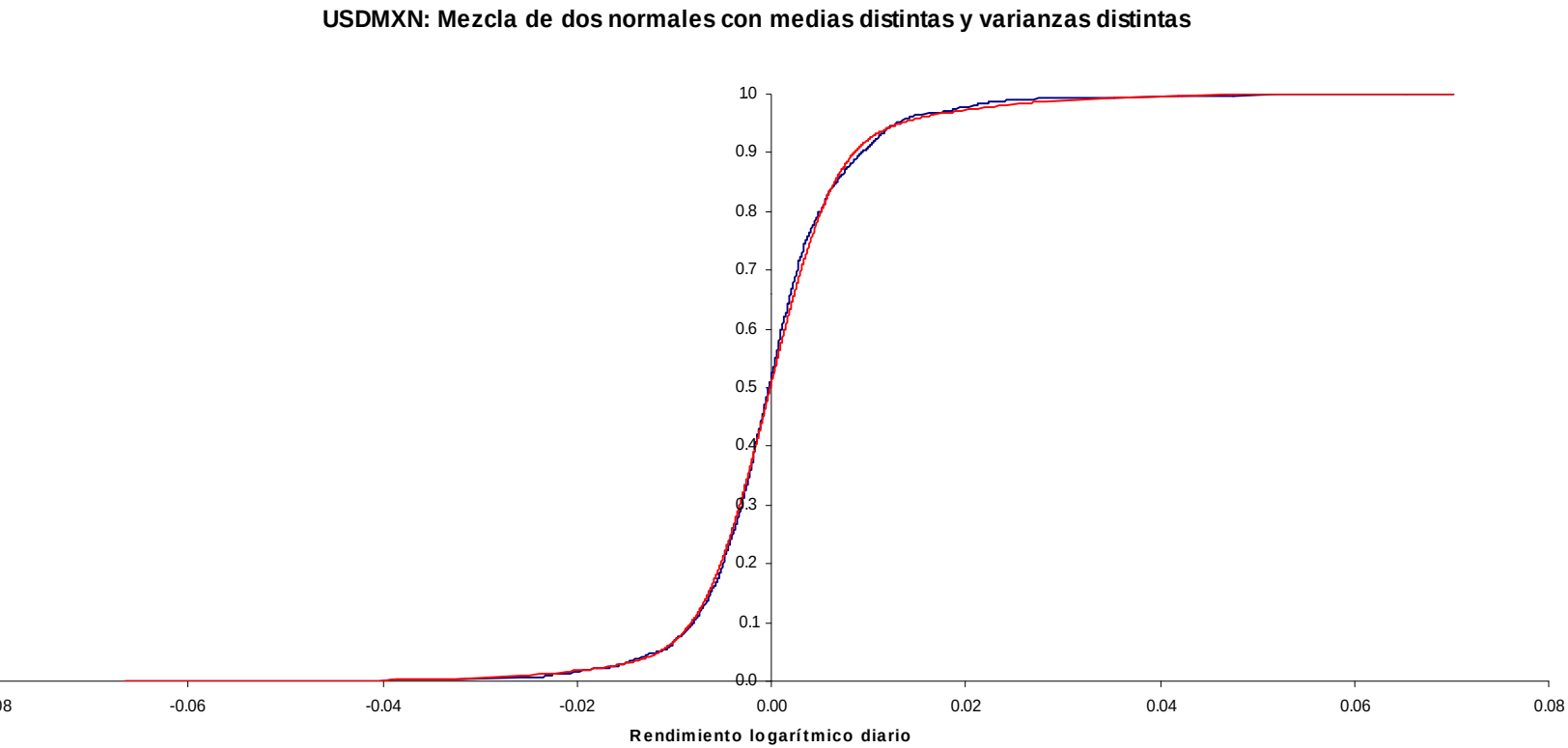
Ajuste de una mixtura de dos normales con esperanzas distintas y varianzas distintas.

$$f(x) = p\phi((x-\mu_1)/\sigma_1)/\sigma_1 + (1-p)\phi((x-\mu_2)/\sigma_2)/\sigma_2$$

Modelo ajustado por Máxima Verosimilitud

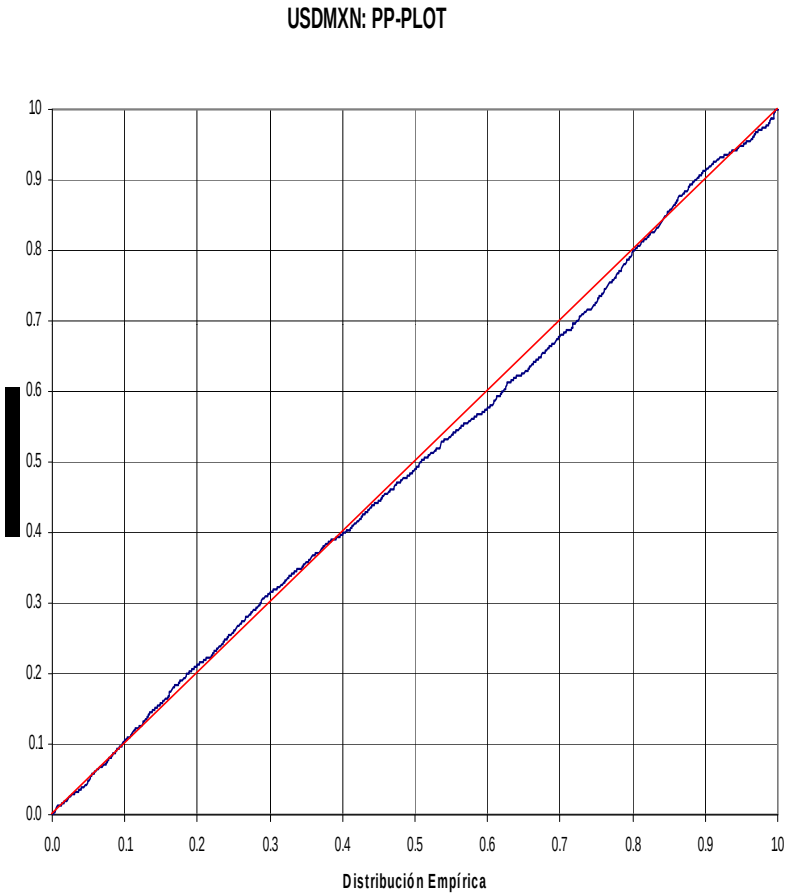
- p = 0.1676
- S1 = 0.0108
- S0 = 0.0052
- M1 = 0.0023
- M0 = -0.0003
- K-S = 0.0269
- P-Val > 0.1

USDMXN: 01-01-2007 to 14-05-2012



CIEMA

USDMXN: 01-01-2007 to 14-05-2012



TELMEX-L

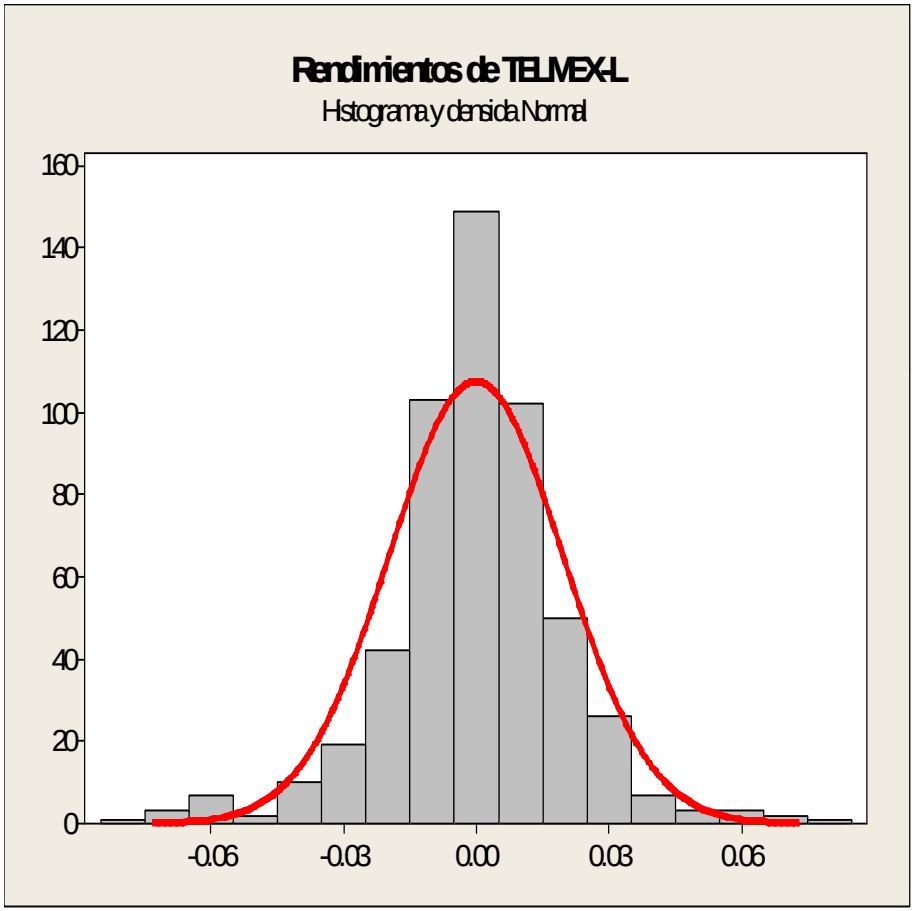
04-03-2008 to 15-04-2010

CIEMA

TELMEX-L: 04-03-2008 to 15-04-2010

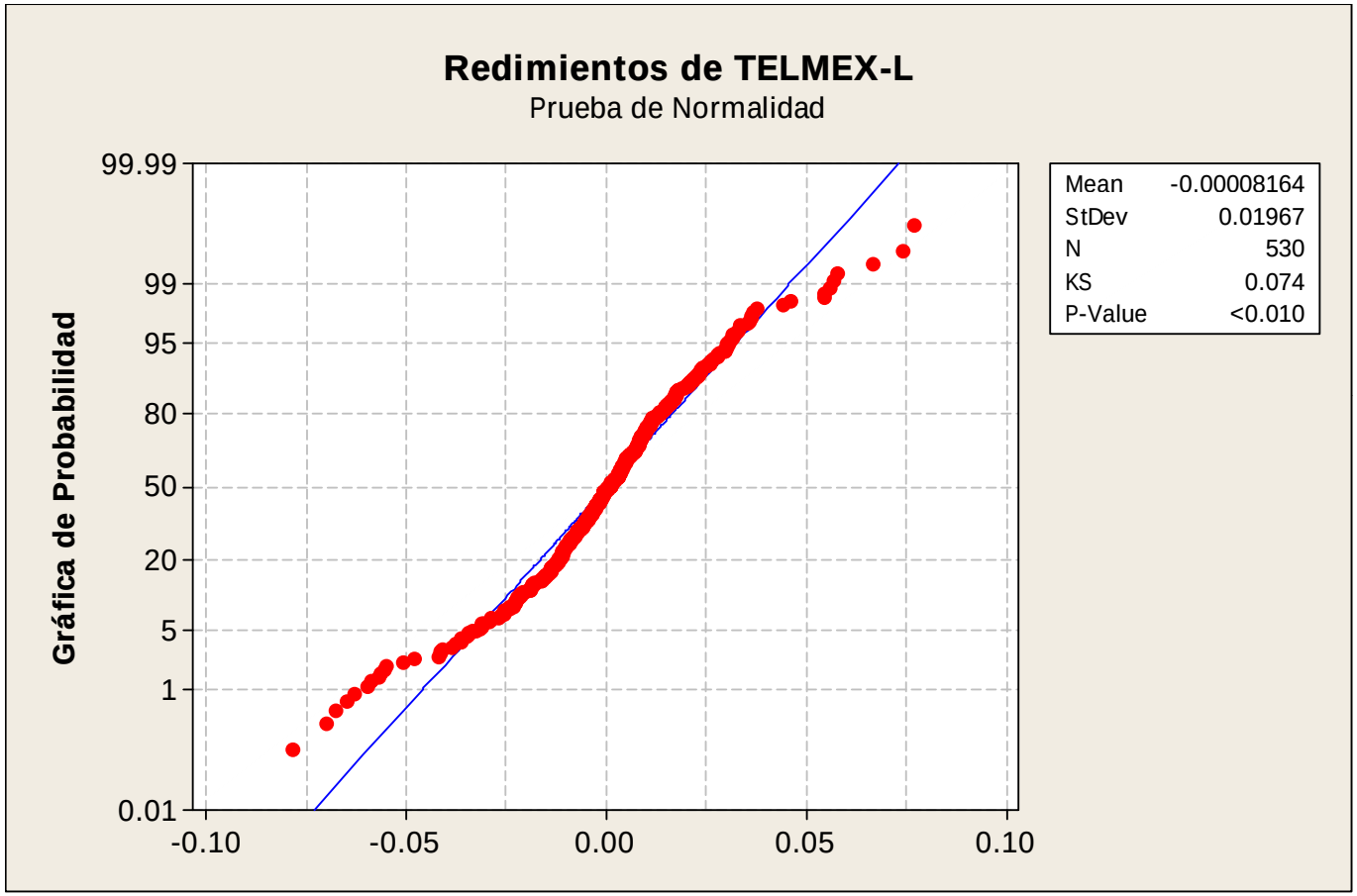
ormal Distribution;
aximum likelihood fit.

```
      = -0.000082

      =  0.019672

S     =  0.07
val < 0.01
```



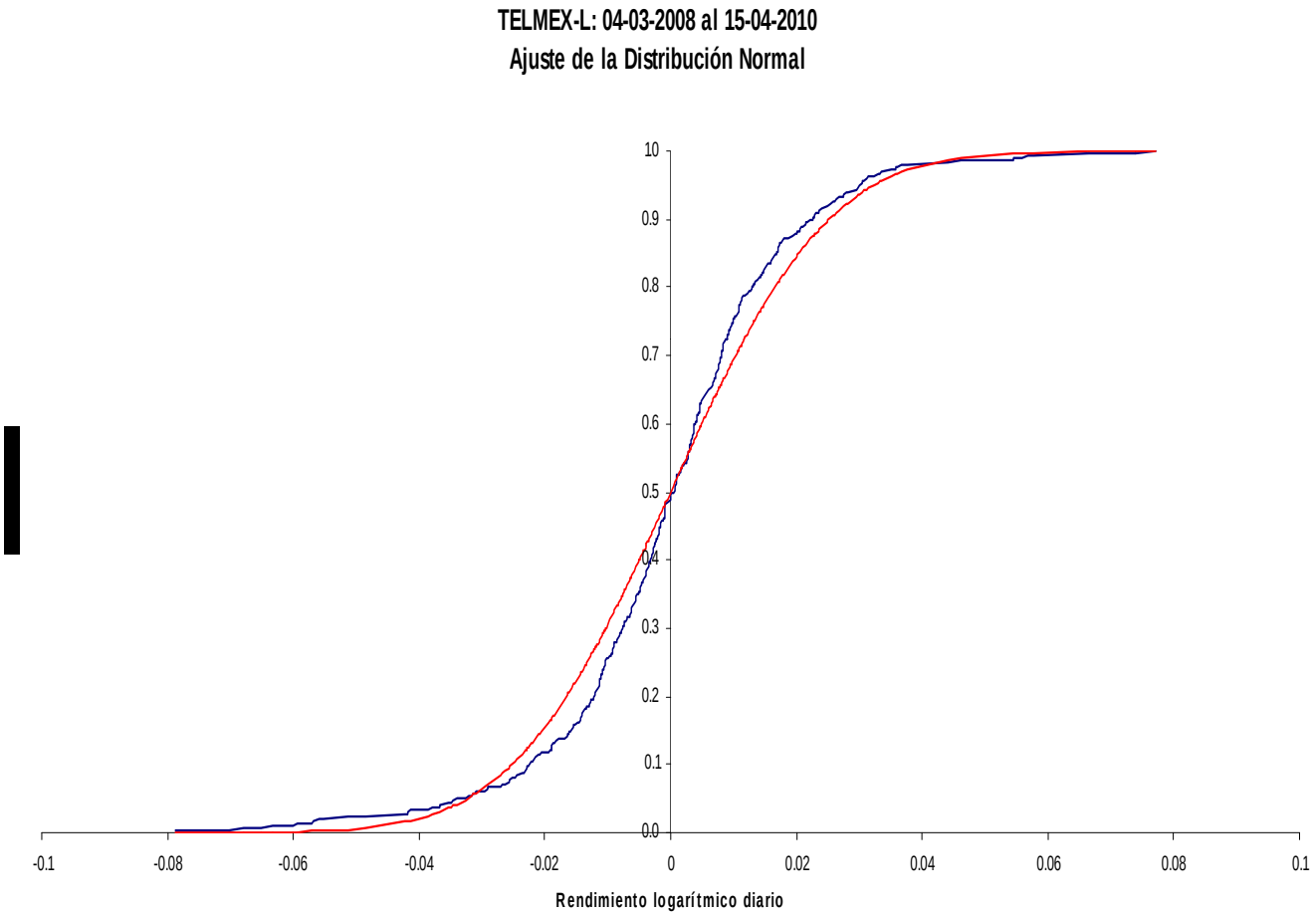
CIEMA

TELMEX-L: 04-03-2008 to 15-04-2010



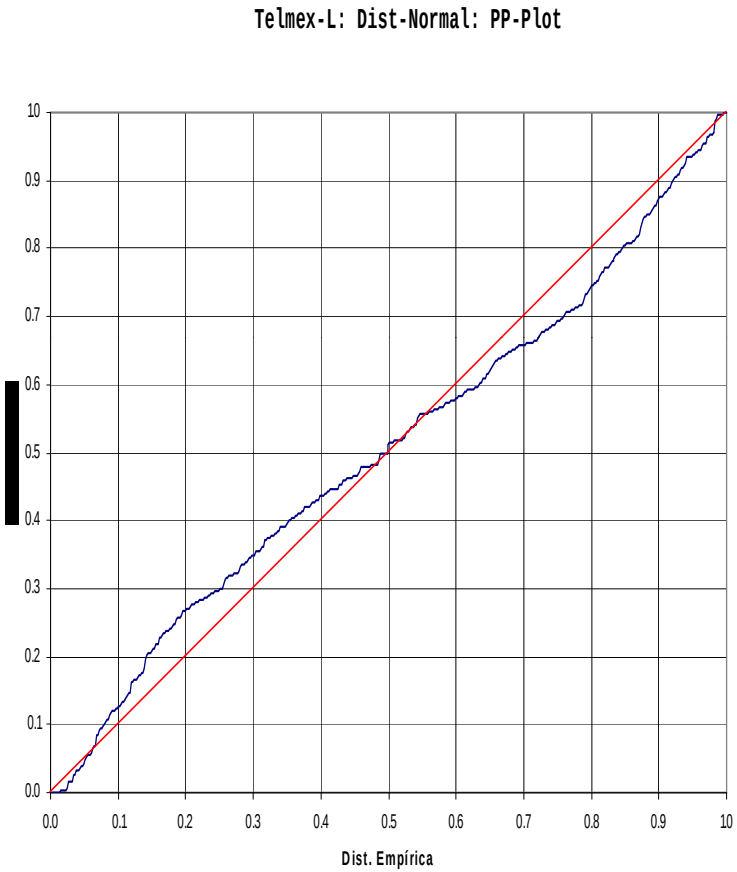
CIEMA

TELMEX-L: 04-03-2008 to 15-04-2010



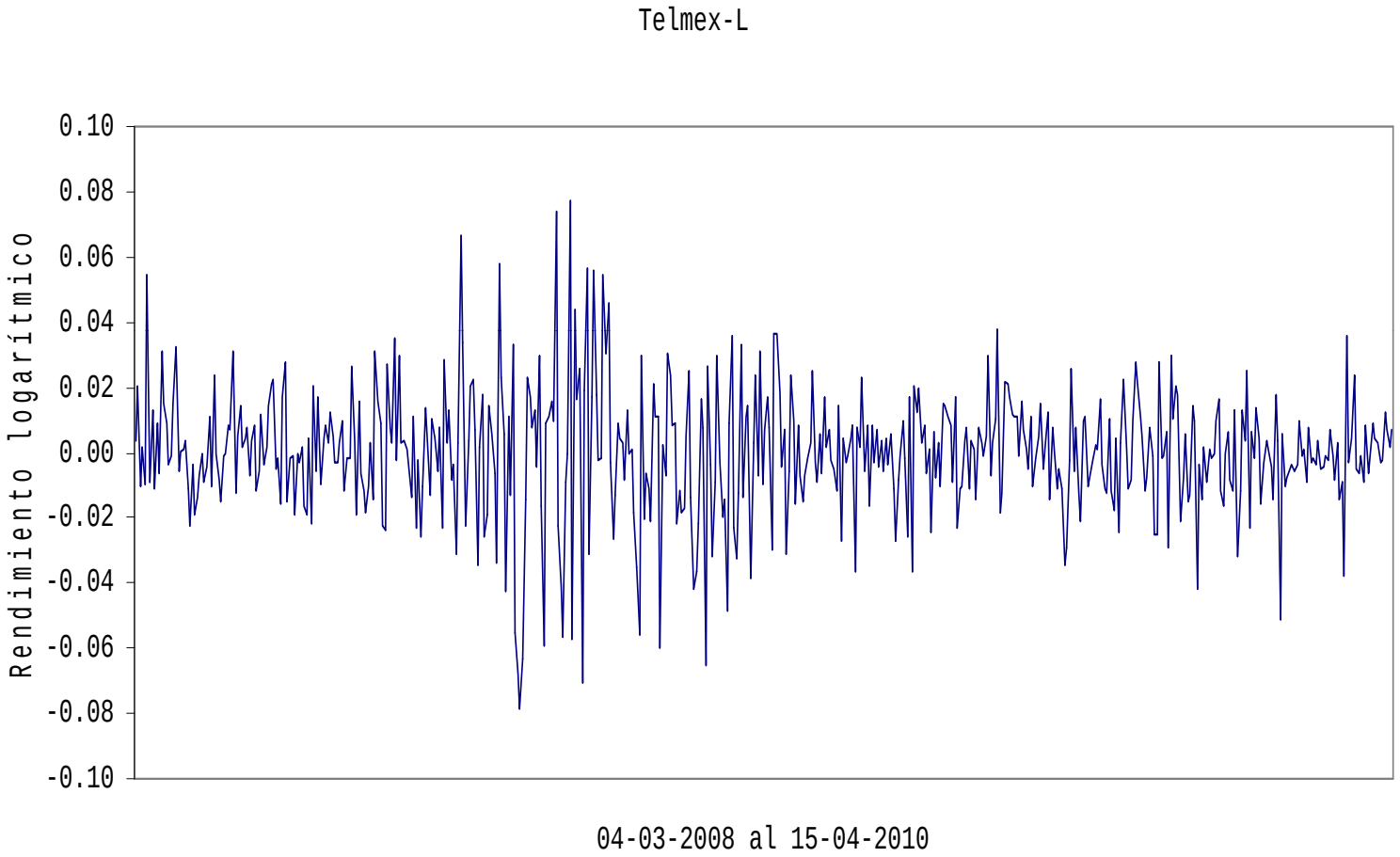
CIEMA

TELMEX-L: 04-03-2008 to 15-04-2010



CIEMA

TELMEX-L: 04-03-2008 to 15-04-2010



TELMEX-L: 04-03-2008 to 15-04-2010

Mixture of two normal distributions
with zero mean and different variances.

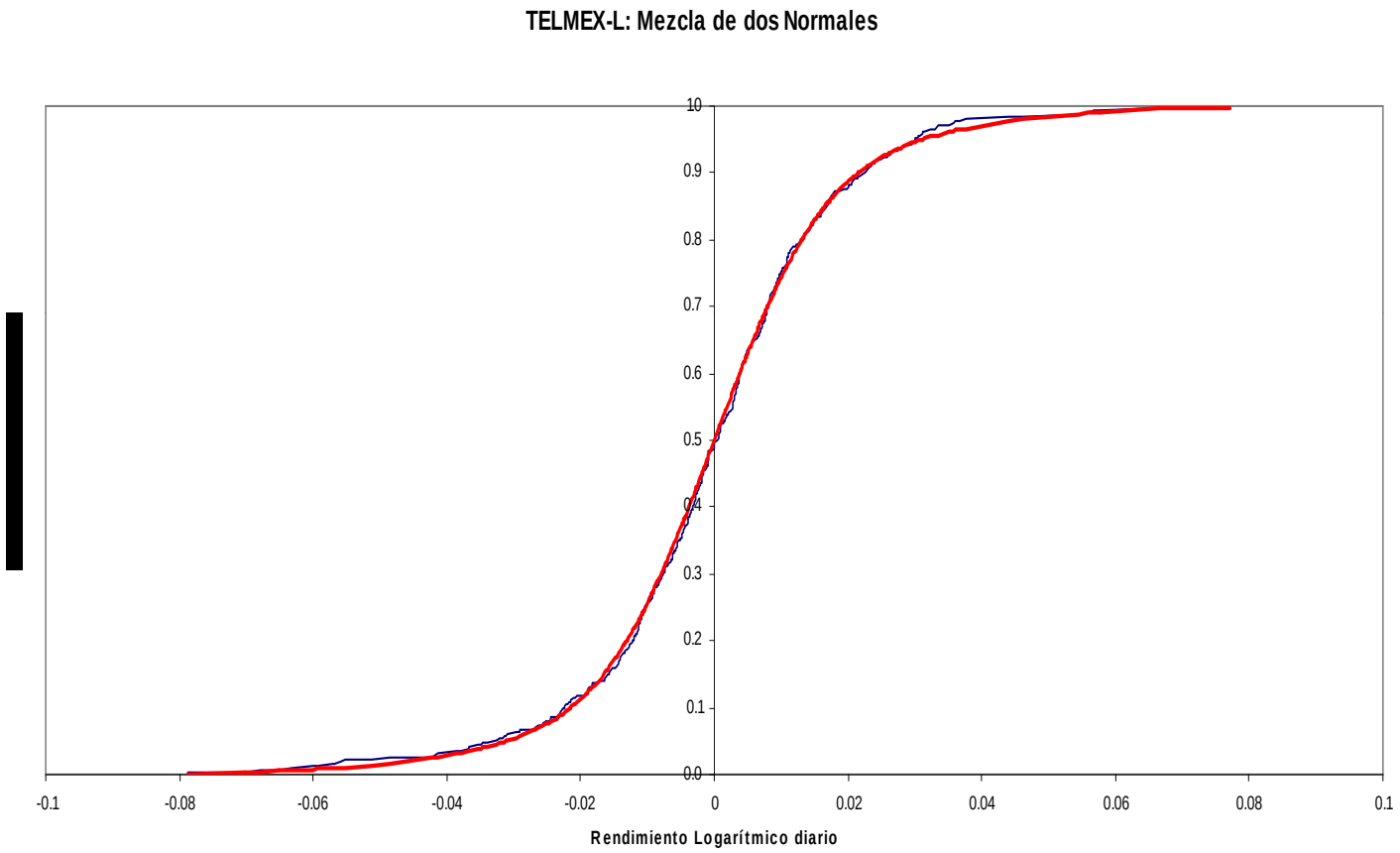
$$f(x) = p\phi(x/s_1)/s_1 + (1-p)\phi(x/s_2)/s_2$$

Maximum Likelihood estimation

- p = 0.29
- S1 = 0.0308
- S0 = 0.0123
- K-S = 0.0274
- P-Val > 0.1

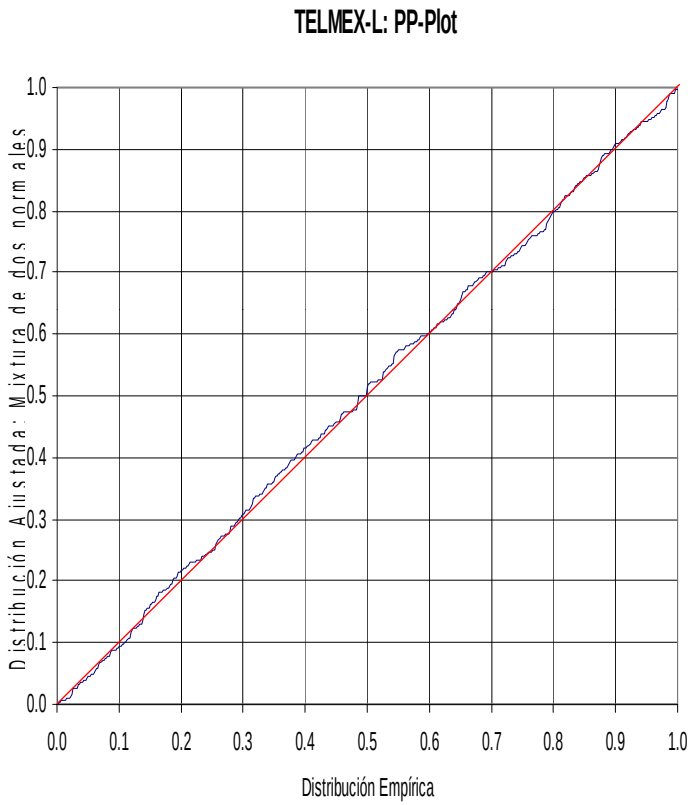
CIEMA

TELMEX-L: 04-03-2008 to 15-04-2010



CIEMA

TELMEX-L: 04-03-2008 to 15-04-2010



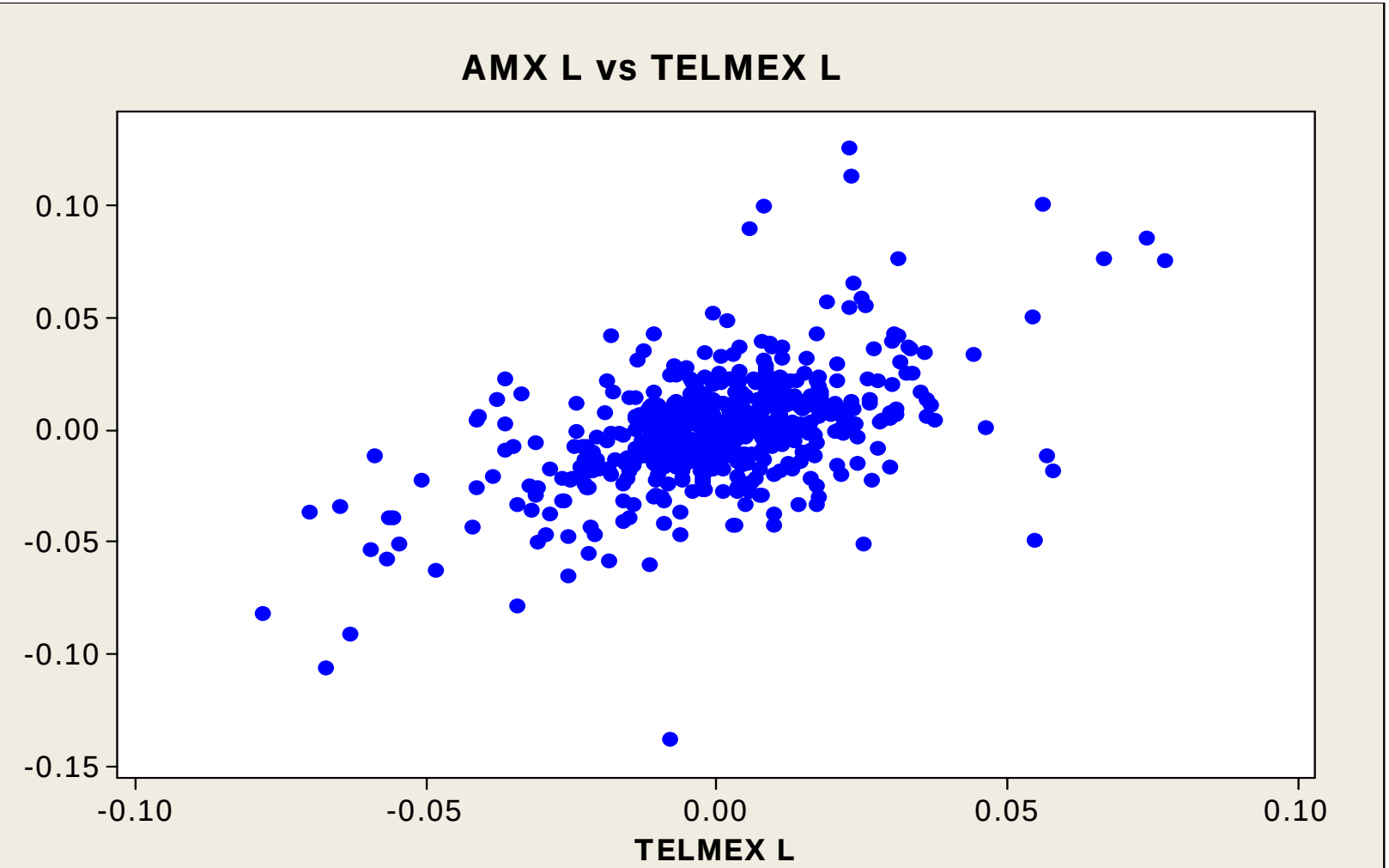
CIEMA

TELMEX-L & AMX-L
bivariate fit

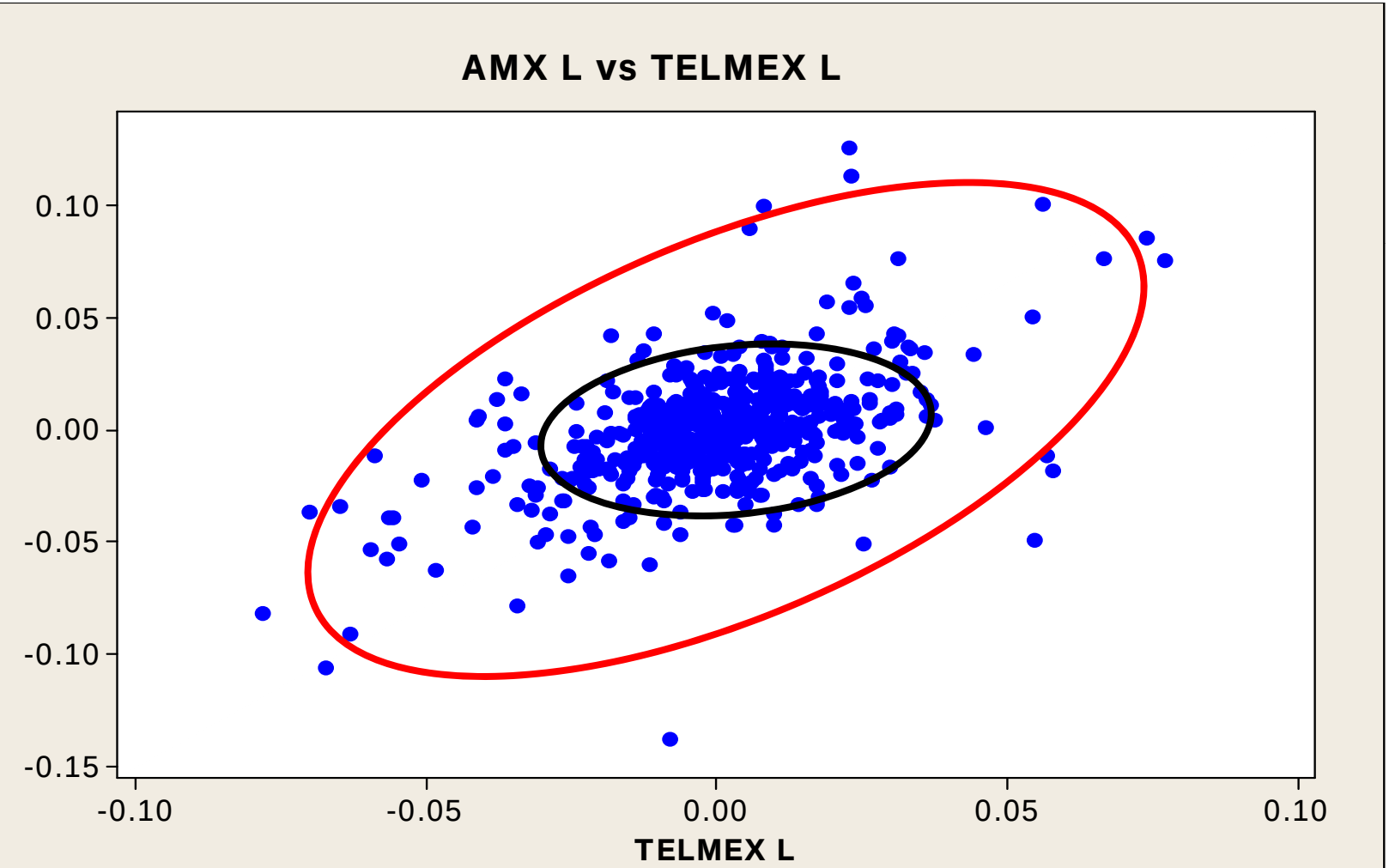
From 04-03-2008 to 15-04-2010

CIEMA

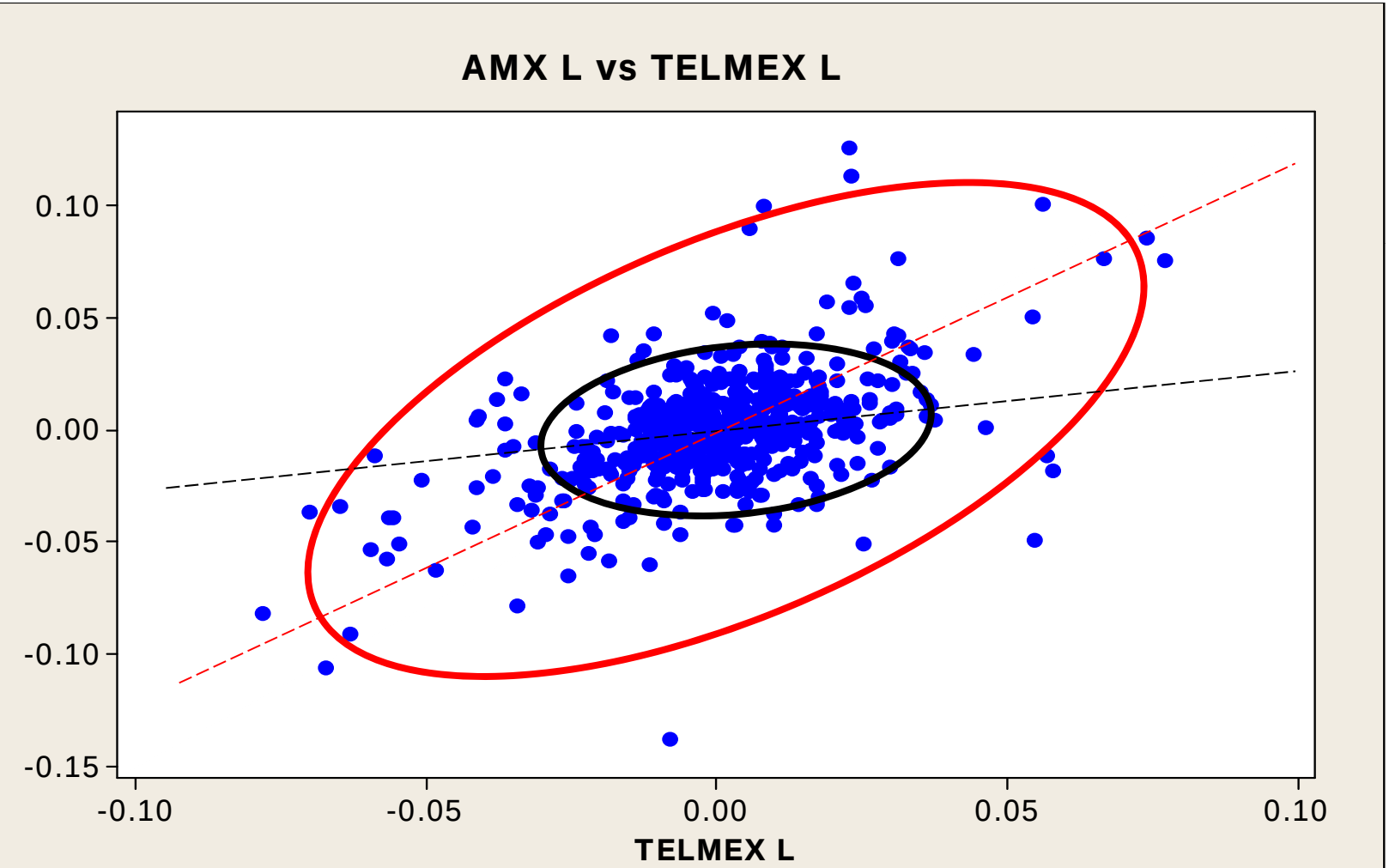
Bivariate fit: Telmex L & AMX L



Bivariate fit: Telmex L & AMX L



Bivariate fit: Telmex L & AMX L



Bivariate fit: Telmex L & AMX L

ixture of two bivariate normal distributions, with
ifferent mean vectors and different covariance matrices.

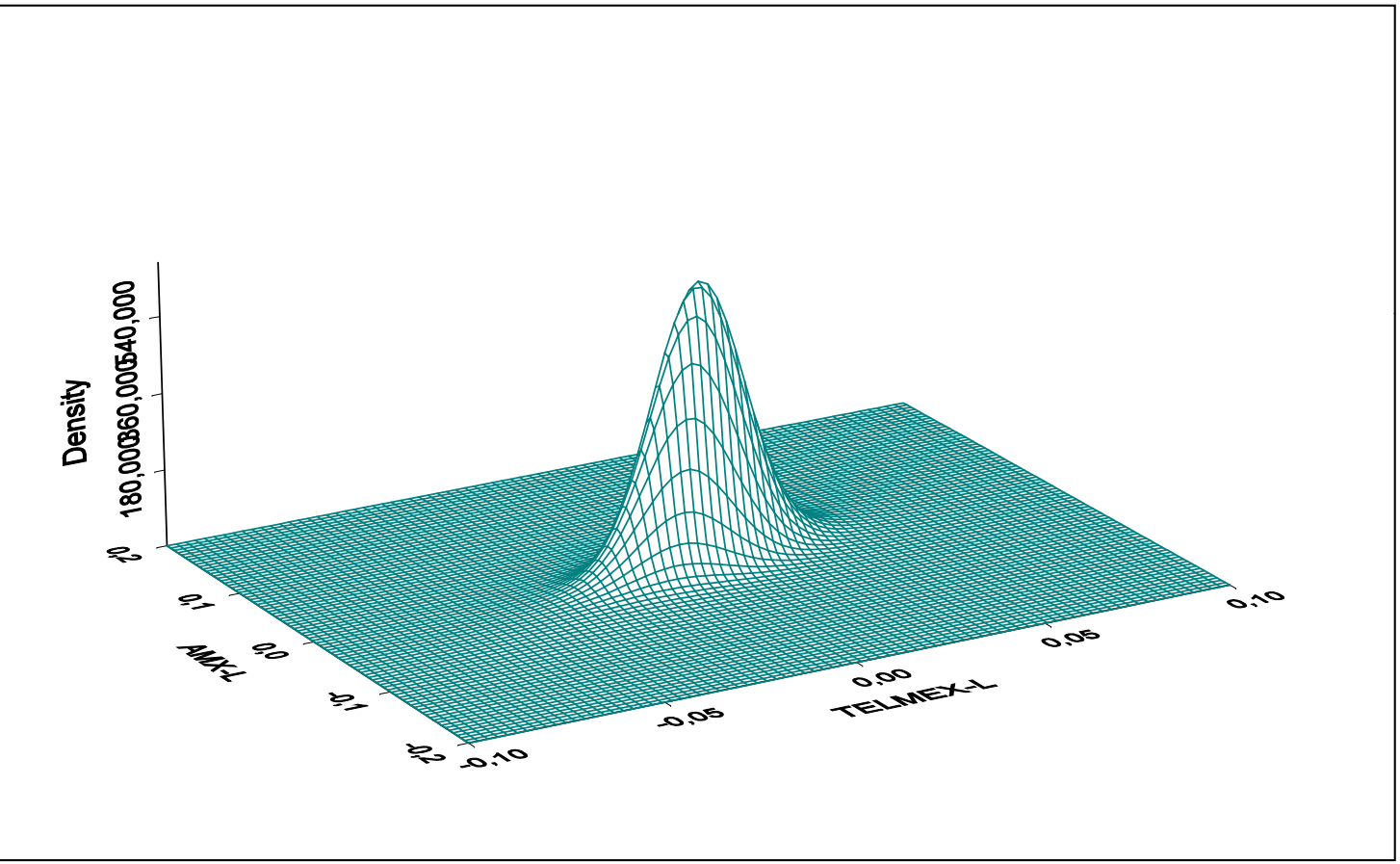
$$f(x,y) = p * N(x,y \mid \mu_o, \Sigma_o) + (1-p) * N(x,y \mid \mu_1, \Sigma_1)$$

aximum Likelihood Estimates (EM Algorithm)

p	= 0.2356		
x0	= 0.0008	Mx1	= -0.0327
y0	= 0.0004	My1	= -0.0014
x0	= 0.0008	Sx1	= 0.0327
y0	= 0.0160	Sy1	= 0.0450
0	= 0.4159	R1	= 0.5678

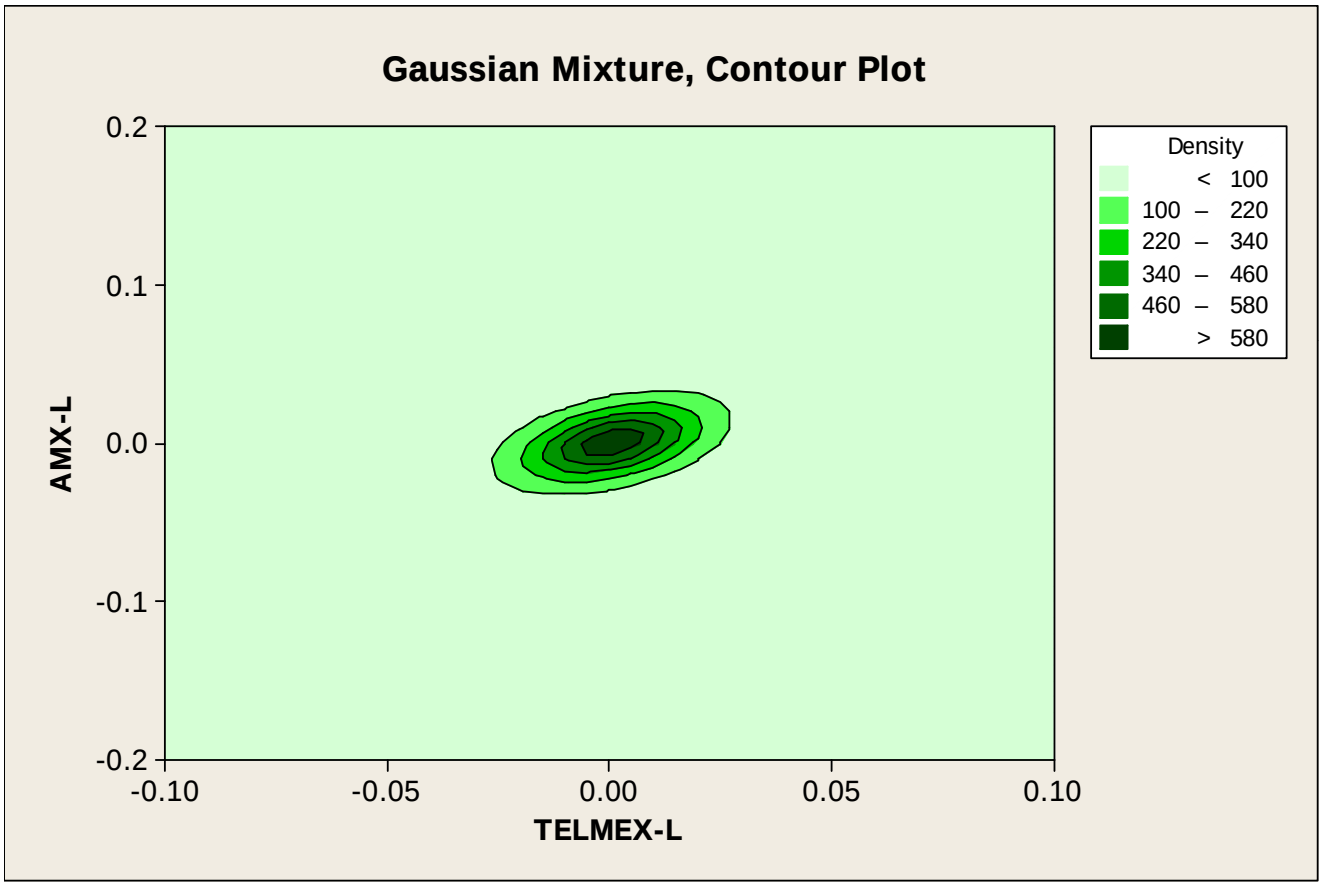
CIEMA

Bivariate fit: Telmex L & AMX L

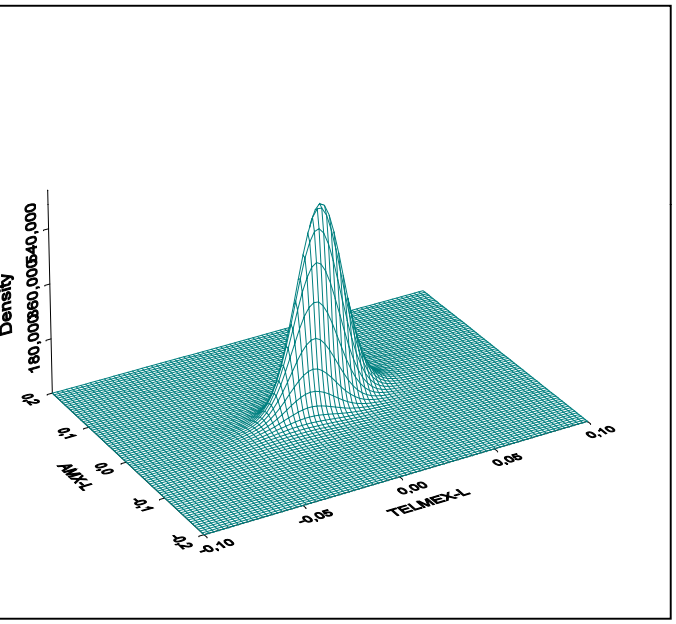


CIEMA

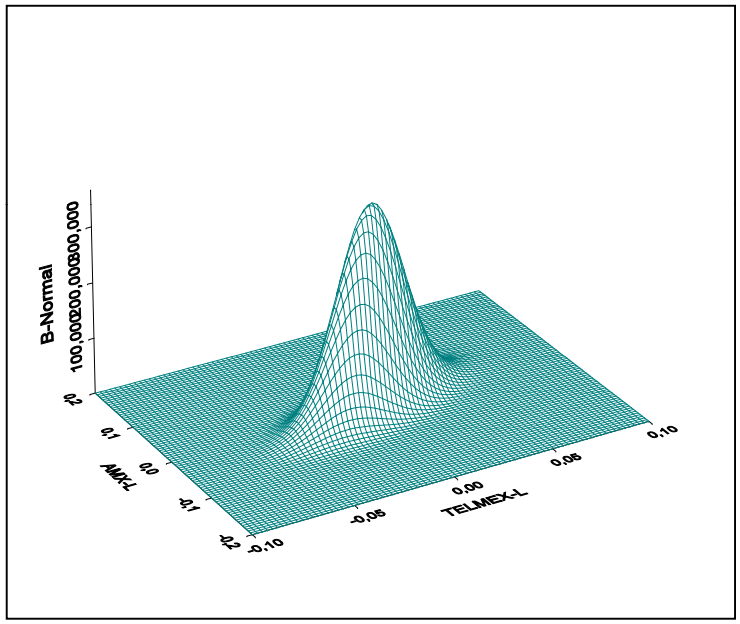
Bivariate fit: Telmex L & AMX L



Bivariate fit: Telmex L & AMX L



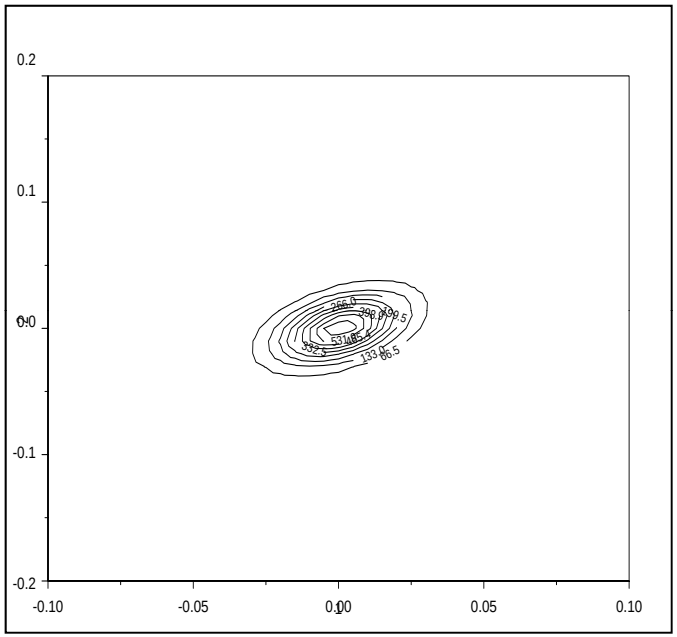
Gaussian Mixture



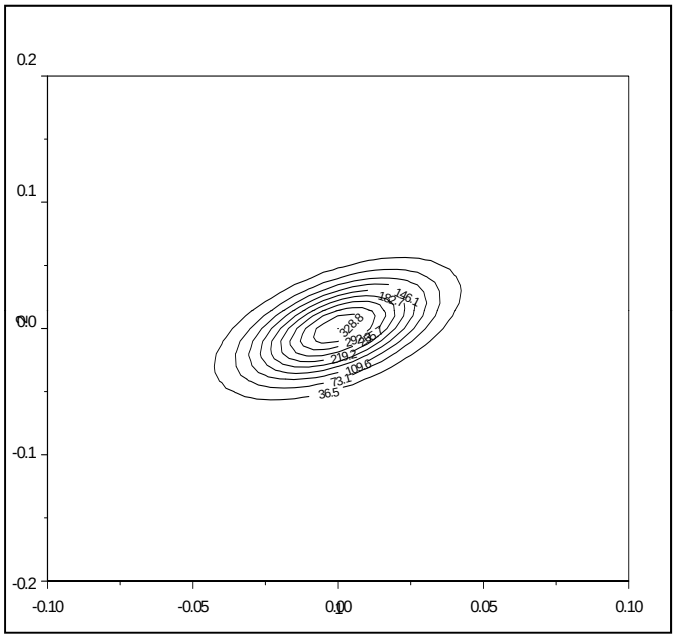
Bivariate Normal

CIEMA

Bivariate fit: Telmex L & AMX L



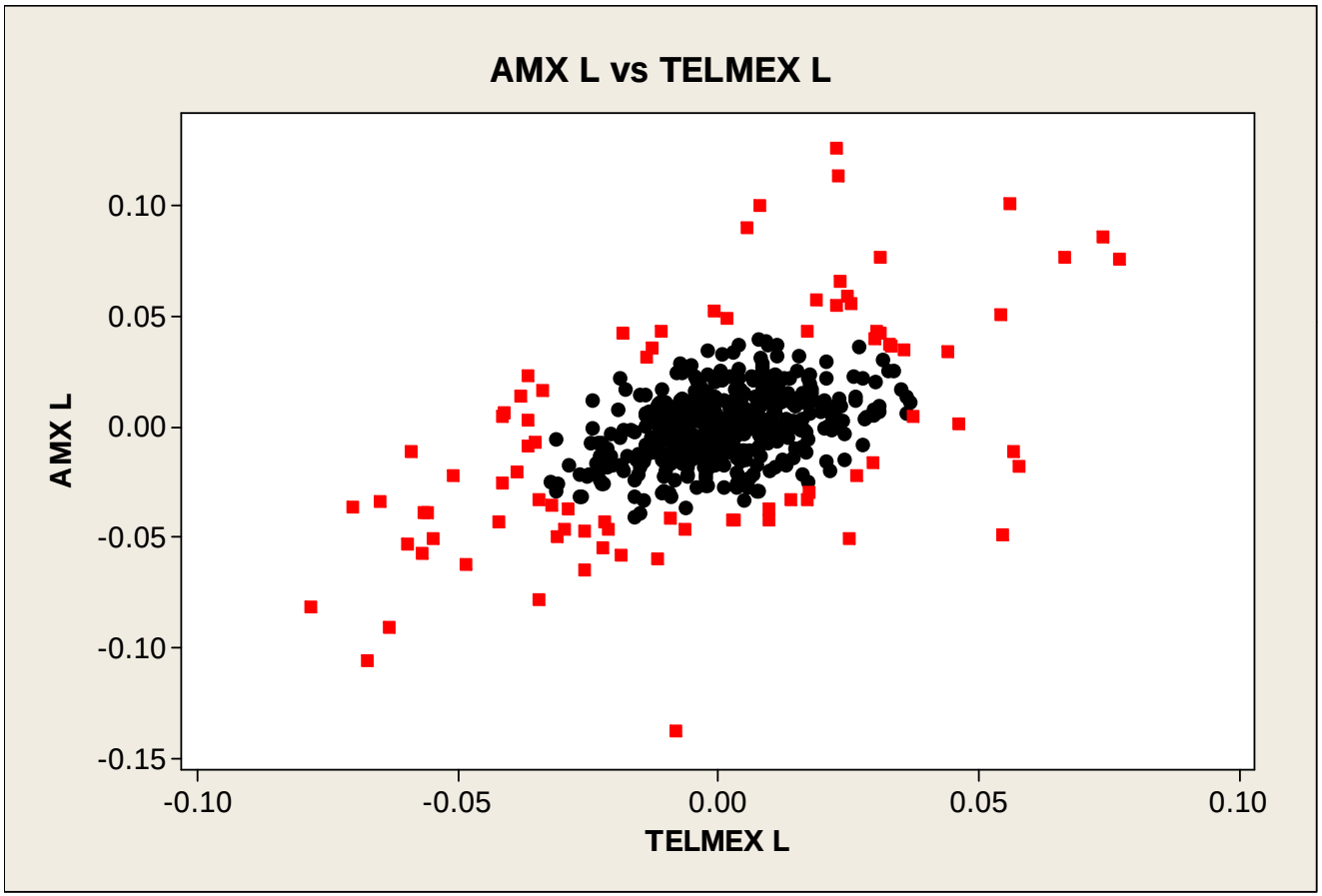
Gaussian Mixture



Bivariate Normal

CIEMA

Bivariate fit: Telmex L & AMX L



CIEMA

Conclusions and future work

Conclusion

evidence suggests that financial returns follow a Gaussian
ture (… or something very similar).

we say that the assumption of normality is totally wrong?

ssian Mixture models seem to work well in financial modeling.
y are parsimonious, accurate, and easy to work with.

en their relation with the normal distribution, it is important to
y their implications and possible contributions to standard
ncial theory.

Questions and comments.

Thank you very much!

CIEMA

Examples

AMX-L

04-03-2008 al 15-04-2010

CIEMA

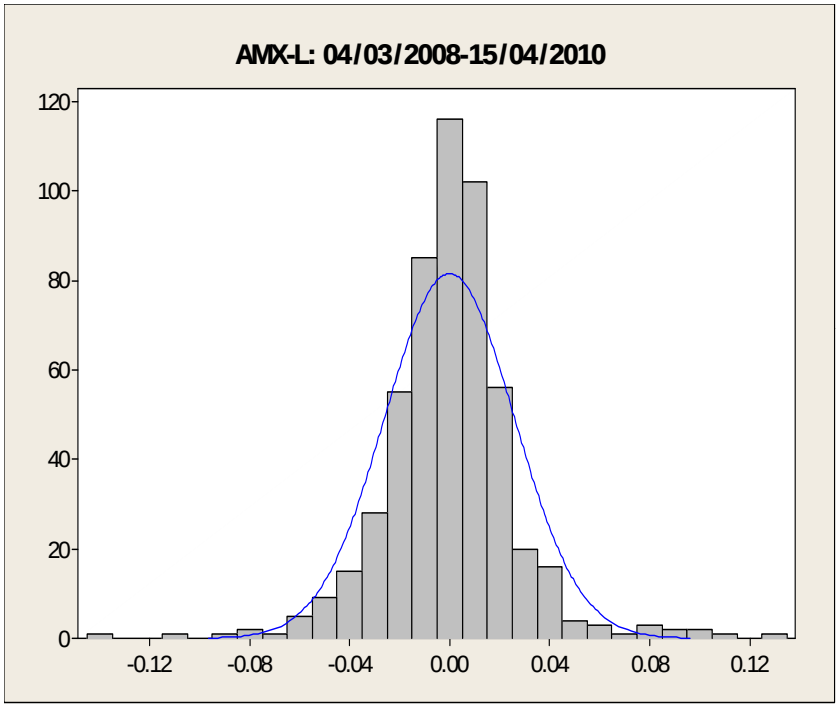
AMX-L: 04-03-2008 al 15-04-2010

irbución Normal
ma Verosimilitud

= -0.00005

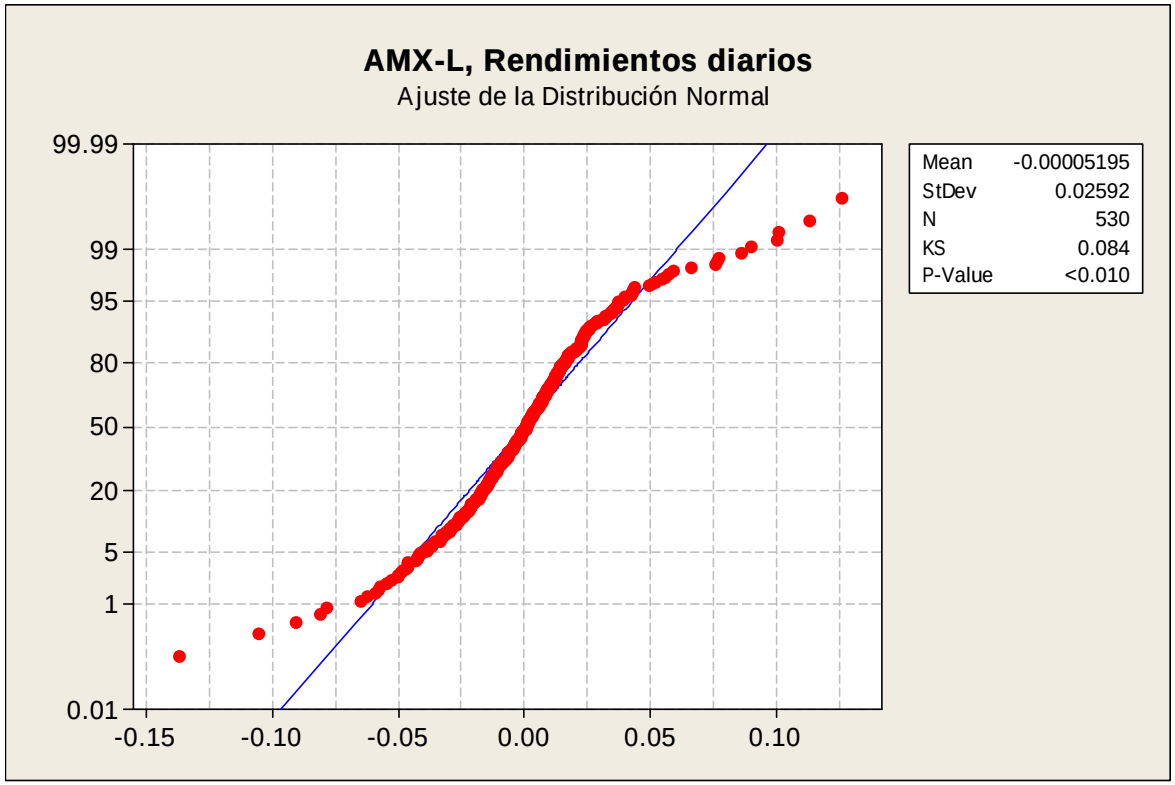
= 0.02592

= 0.08
l < 0.01

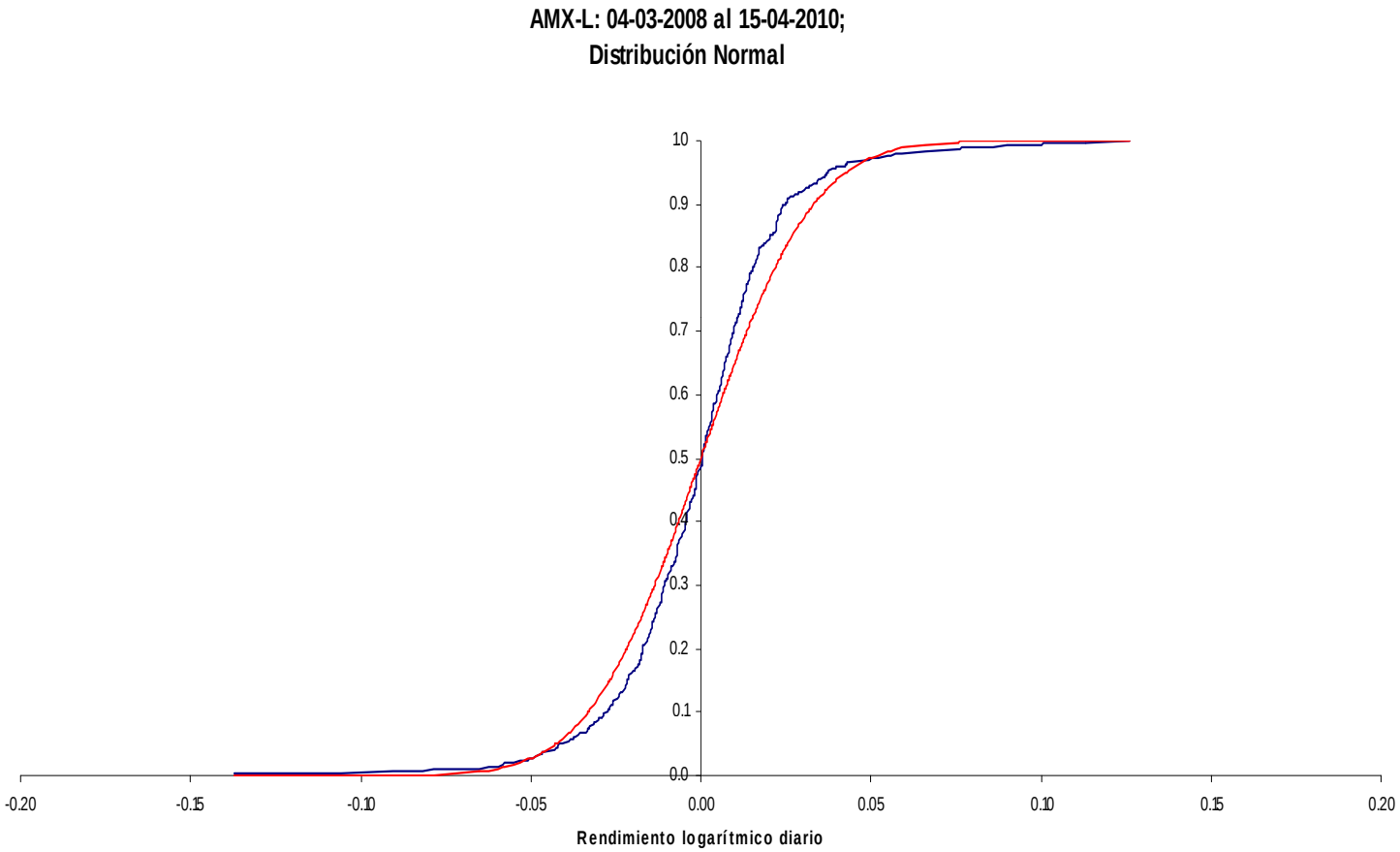


CIEMA

AMX-L: 04-03-2008 al 15-04-2010

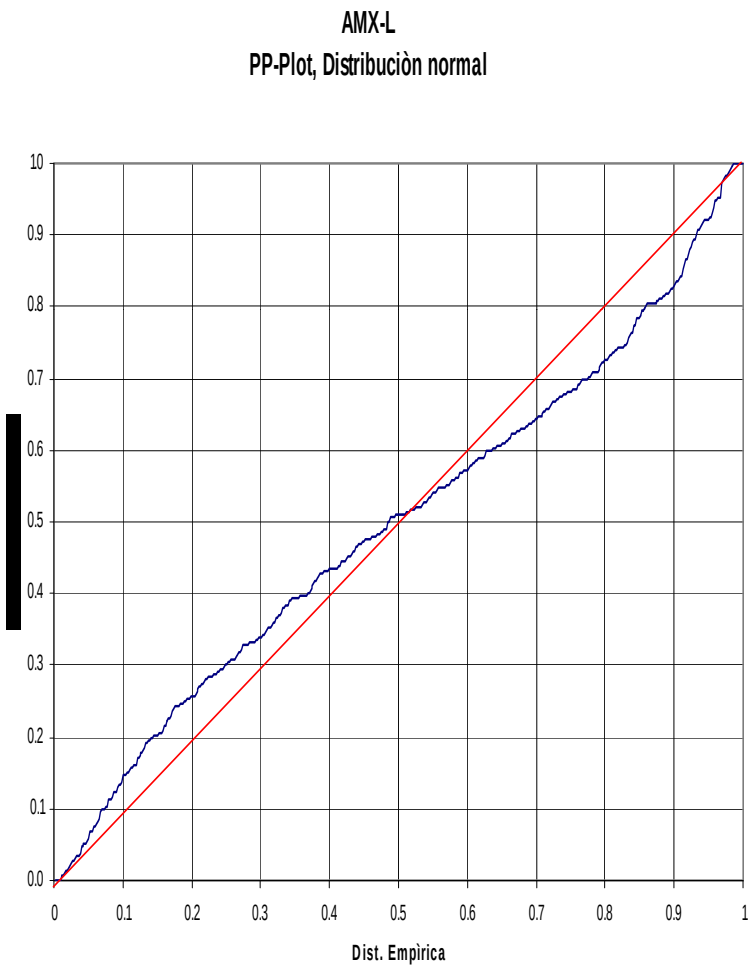


AMX-L: 04-03-2008 al 15-04-2010

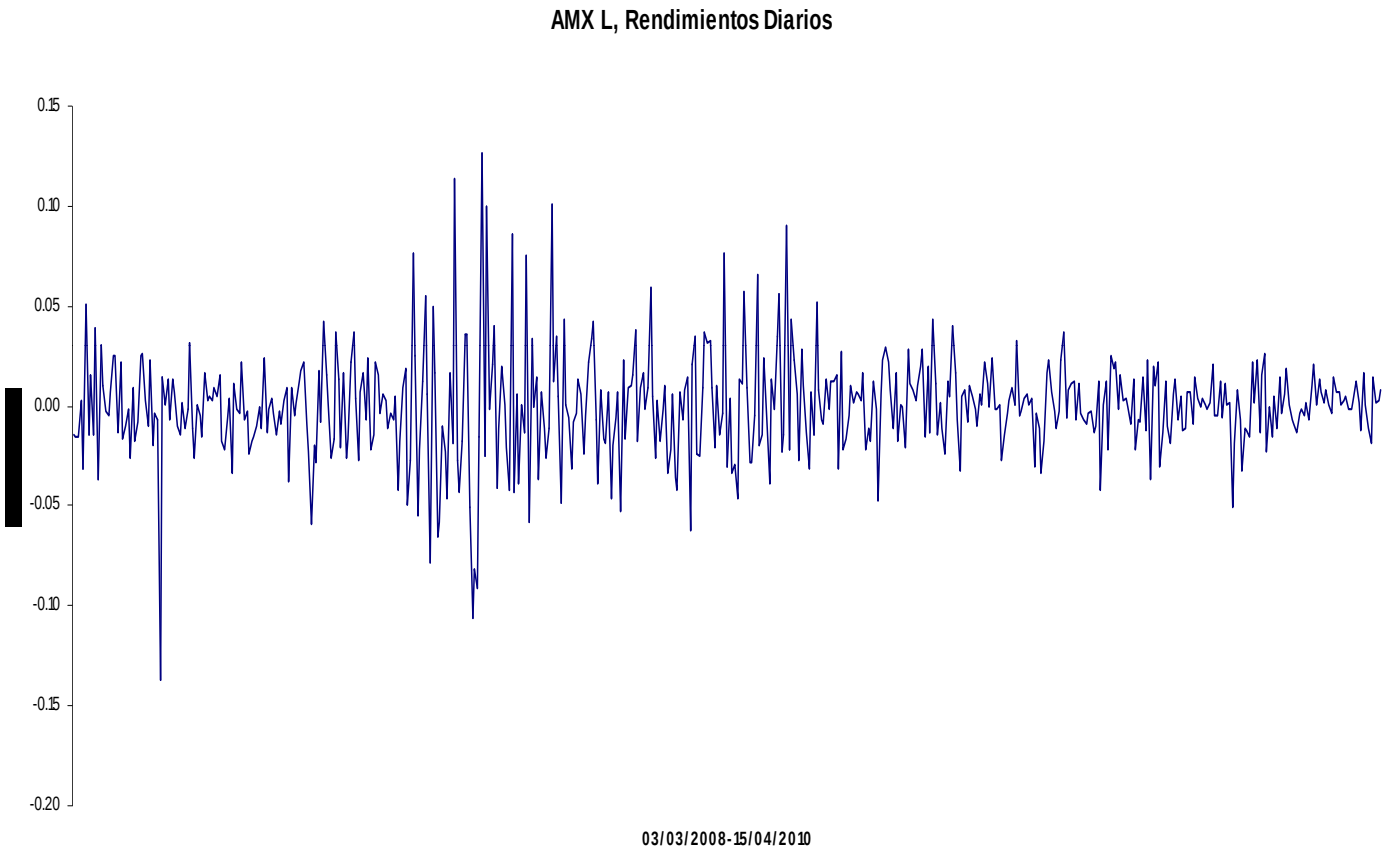


CIEMA

AMX-L: 04-03-2008 al 15-04-2010



AMX-L: 04-03-2008 al 15-04-2010



AMX-L: 04-03-2008 al 15-04-2010

Ajuste de una mixtura de dos normales centradas en cero y con varianzas distintas.

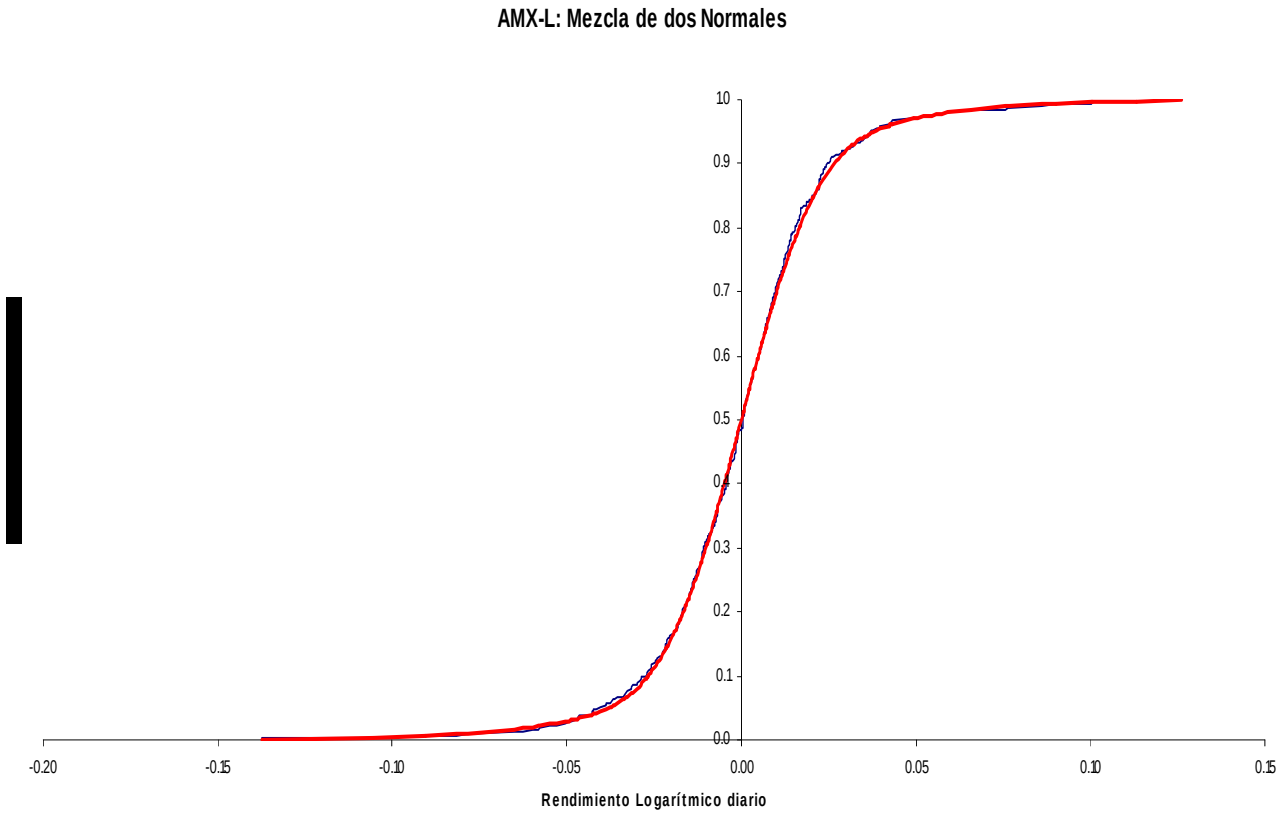
$$f(x) = p\phi(x/s_1)/s_1 + (1-p)\phi(x/S2)/s_2$$

Ajuste por Máxima Verosimilitud

- p = 0.18
- S1 = 0.0491
- S0 = 0.0169
- K-S = 0.02
- P-val > 0.1

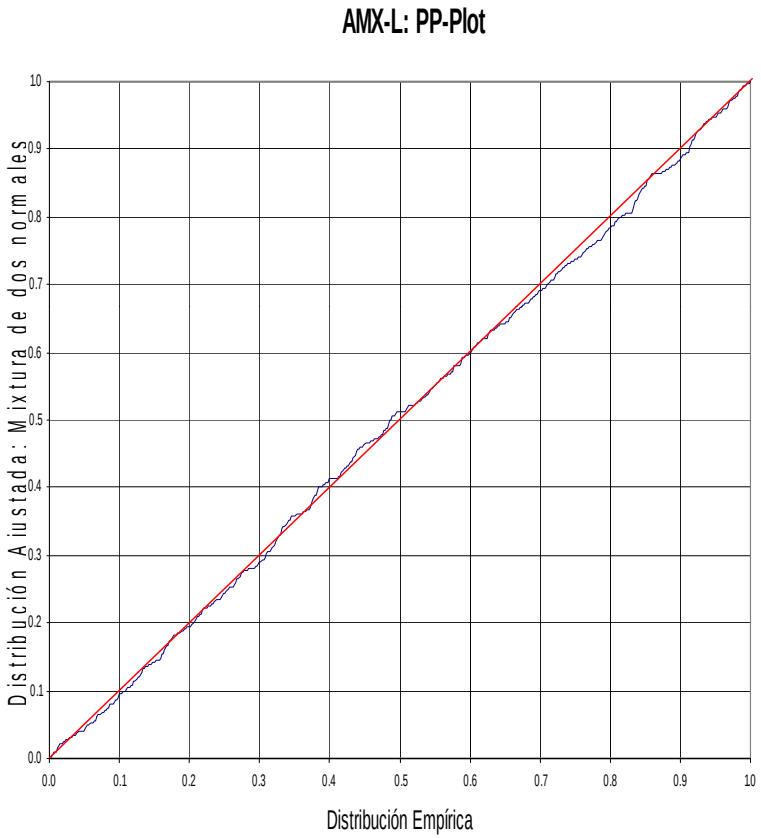
CIEMA

AMX-L: 04-03-2008 al 15-04-2010



CIEMA

AMX-L: 04-03-2008 al 15-04-2010



CIEMA

CETES 28 DÌAS

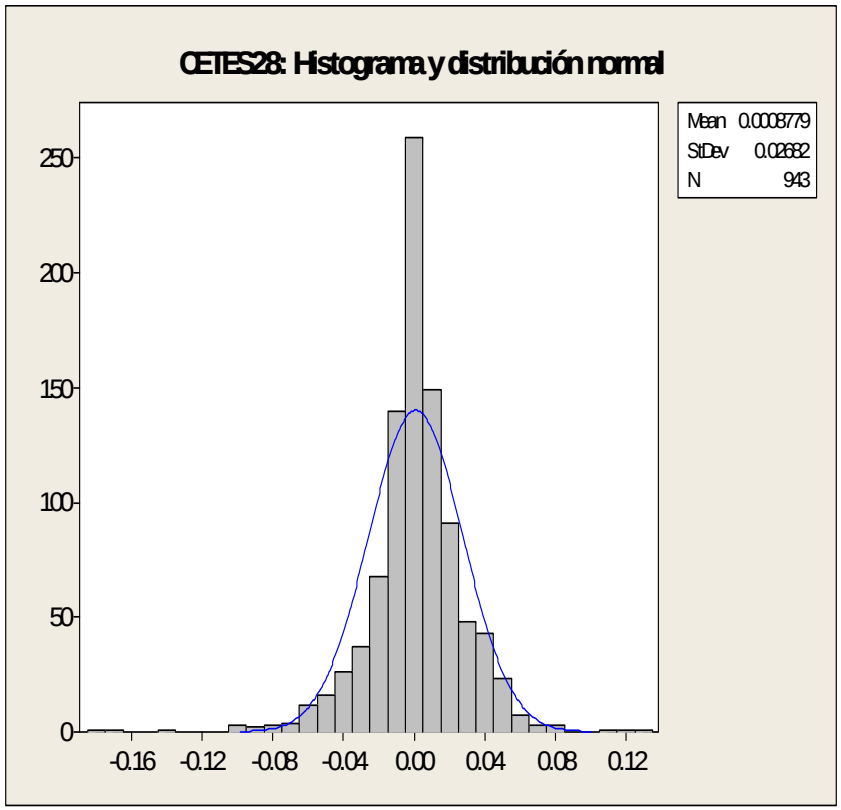
02-01-2001 al 29-09-2004

CIEMA

CETES28: 02-01-01 al 29-09-04

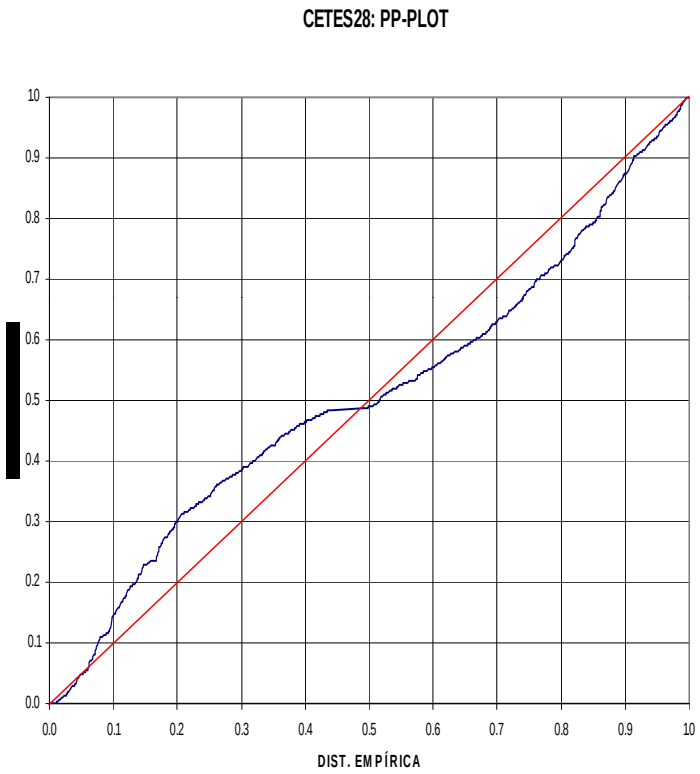
C: Estadísticas Descriptivas

```
riable          Mean      StDev
ndimiento   0.000878   0. 026823


S   = 0.10
V   < 0.01
```

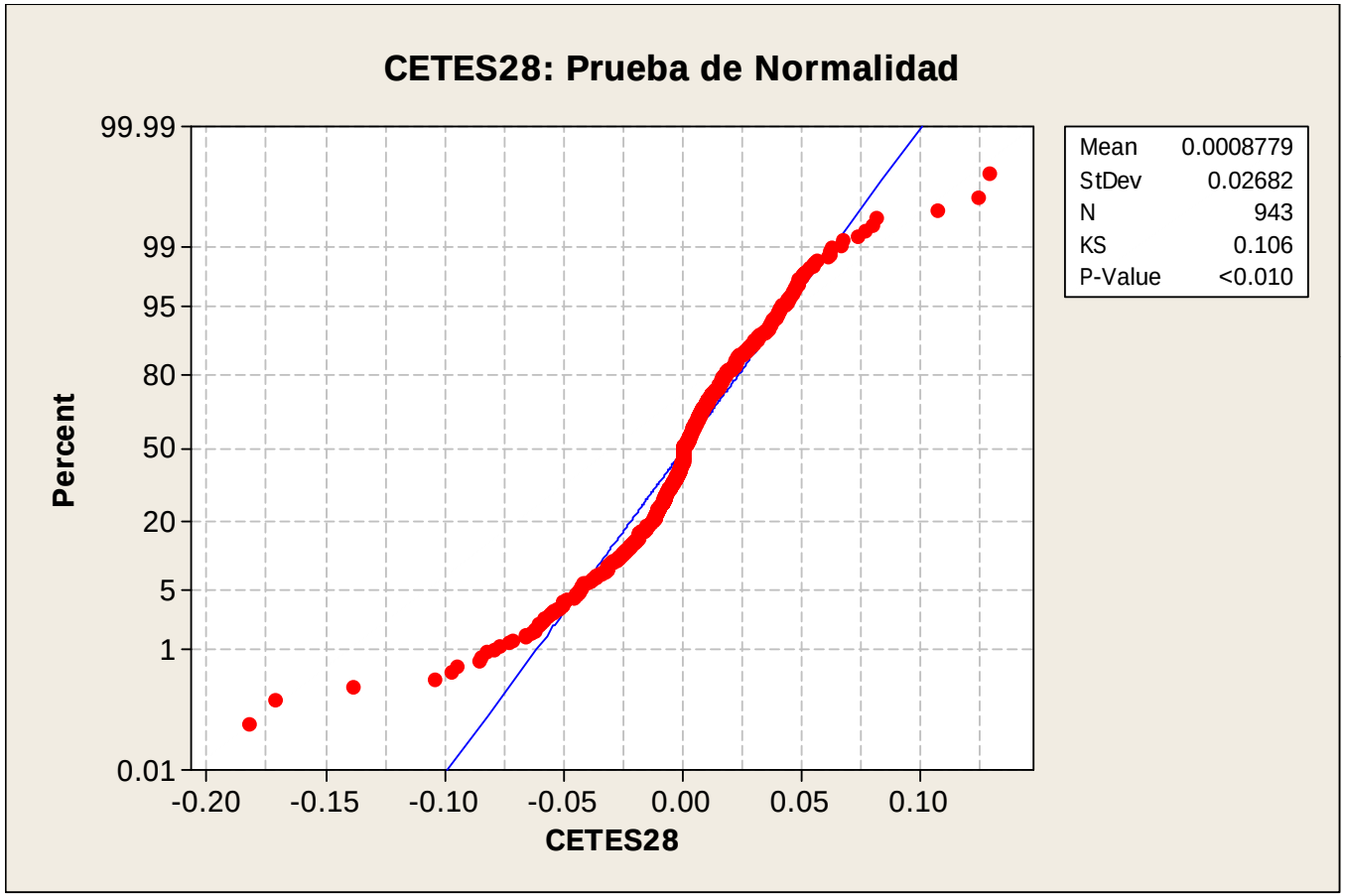


CIEMA

CETES28: 02-01-01 al 29-09-04

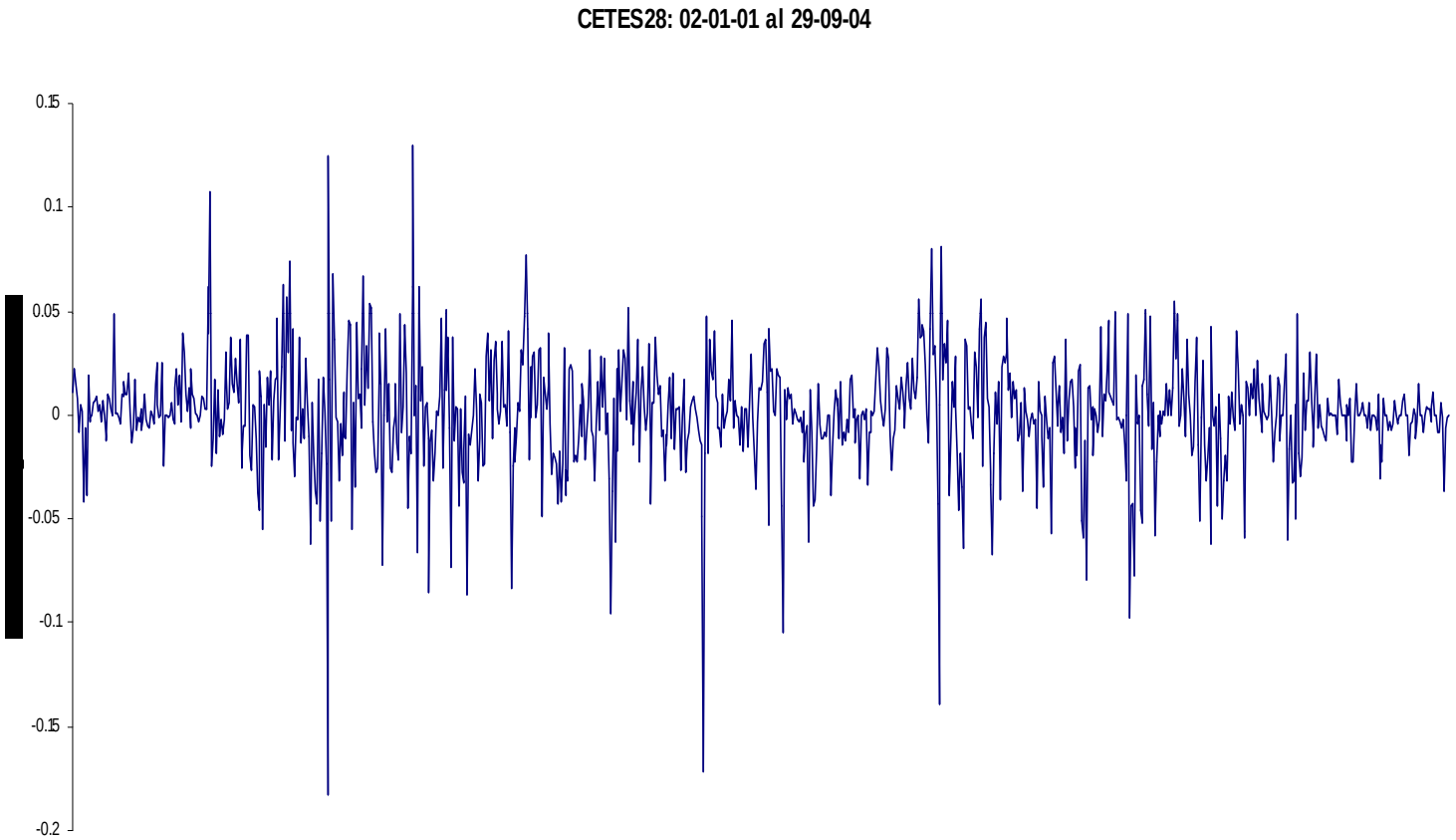


CETES28: 02-01-01 al 29-09-04



CIEMA

CETES28: 02-01-01 al 29-09-04



CIEMA

CETES28: 02-01-01 al 29-09-04

Ajuste de una mixtura de tres normales con esperanzas distintas y varianzas distintas, más una masa de probabilidad en cero. La funciòn de distribución queda definida como:

$$F(t) = \begin{cases} (1-\theta)\sum_{(j=1)}^{3} p_j \Phi\left(\frac{t-\mu_j}{\sigma_j}\right) & \text{si } t < 0, \\ (1-\theta)\sum_{(j=1)}^{3} p_j \Phi\left(\frac{t-\mu_j}{\sigma_j}\right) + \theta & \text{si } t \geq 0. \end{cases}$$

El parámetro θ representa la probabilidad de observar un rendimiento nulo.

CIEMA

CETES28: 02-01-01 al 29-09-04

Ajuste del Modelo

ámetros Estimados M.V.

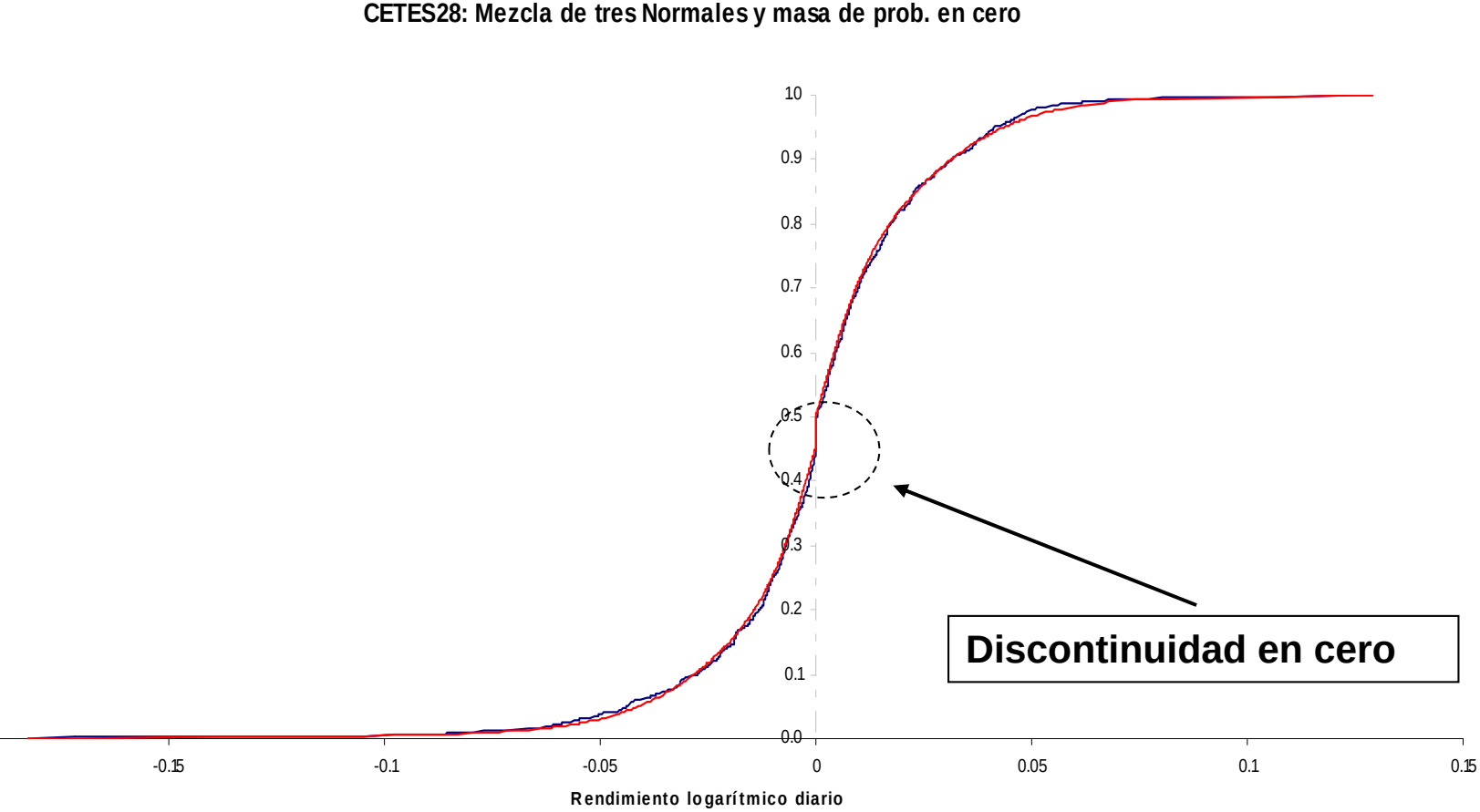
	0.3089
	0.6449
	0.0462
a 1=	0.0081
a 2=	0.0274
a 3=	0.0725
	0.0004
	0.0021
	-0.0121
aciones =	124
en cero =	0.0476

Bondad de Ajuste

P-val.	0.1	0.05	0.02	0.01
K-S	943	943	943	943
0.0219	0.0397	0.0443	0.0495	0.0531

CIEMA

CETES28: 02-01-01 al 29-09-04



CIEMA

CETES28: 02-01-01 al 29-09-04