ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Mathematical Analysis using Frequency and Cumulative Distribution functions for Mitogenic Pathway.

Shruti Jain*.

Department of Electronics and Communication Engineering, Jaypee University of Information Technology, Solan-173234, India.

ABSTRACT

In this paper we have used the mathematical analysis using frequency and cumulative distribution functions (Gaussian, weibull, triangular, normal, log normal, half normal etc) for JNK, MK2 and ERK which are the part of mitogenic pathways to make a best model of the survival/ death proteins. We have calculated KS, AD (statistics & p- value), chi-square (p- value and df) for frequency and cumulative distribution functions using kolmogrorov- smirrnov, chi- square and spapiro wilk tests. Results with half normal distribution function are the best as their AD and chi square values are the maximum.

Keywords: Mitogenic pathways, Kolmogorov Smirnov , Andesrson darling, chi-square tests.

Corresponding author

ISSN: 0975-8585

INTRODUCTION

MAPK is a family of kinase protein which found [1-4] in eukaryotic organism. Basically it is divided into three sub groups [5-7] : Extracellular-regulated kinase (ERK) [8], p38/high osmolarity glycerol (HOG)/Mitogen-activated protein kinase-activated protein kinase 2 (MK2) [8], c-Jun amino-terminal kinase / stress-activated protein kinase (JNK/SAPK) [8, 9]which leads to cell death/ cell survival. The ERK are mediated by differentiation and mitogenic signals while other two leads to inflammatory and stress cytokines.

ERK: EGF activates the ERK pathway through the binding of Grb2 or Shc, which in turn results in the recruitment of the son of sevenless (SOS). SOS then activates RAS leading to the activation of RAF 1[8]. RAF-1 phosphorylates MEK1 and MEK2 which activate ERK 1 and ERK 2 respectively. This pathway results in cell proliferation and in the increased transcription of Bcl2 family members and inhibitors of apoptosis proteins (IAPs), thereby promoting cell survival. It is one of the path-way of MAPK family.

JNK pathway: The second pathway for MAP kinase cascade is the JNK/SAPK [8, 9]. The JNK are activated when cells are exposed to heat shock, ultraviolet (UV) radiation, or inflammatory cytokines.

p38/ MK2 pathway: the p38 kinases are activated in response to inflammatory cytokines, endotoxins, and osmotic stress. It shares about 50% homology with the ERKs. Another known target of p38 is MAPK2 that is involved in the phosphorylation and activation of heat-shock proteins.

In this paper Anderson darling (AD), Kolmogorov Smirnov (K-S) and chi-square tests were done using different distribution functions i.e weibull, normal, log normal, triangular, Rayleigh, half normal and general pareto for mitogenic pathways proteins.

SYSTEM IMPLEMENTATION USING KOLMOGOROV–SMIRNOV, CHI- SQUARE AND SPAPIRO WILK TEST

In this paper we are using frequency and cumulative distribution functions on different MAPK pathways i.e ERK, MK2 and JNK. Ten different concentrations for TNF [11, 12], EGF [13-15] and Insulin [15-17] were used in ng/ml i.e 0-0-0, 5-0-0, 100-0-0, 0-100-0, 5-1-0, 100-100-0, 0-0-500, 0.2-0-1, 5-0-5, 100-0-500.

Cumulative distribution function specifies the distance of multivariate random variables *X*. The real valued random variables *X* with a given probability distance have a value equal to or less than *x*. There are different test performed on different distribution functions which are as follows AD test, KS test and chi-square test.

a) The Kolmogorov–Smirnov (K-S/ KS) test is an equality test using nonparametric. It can also be used as a goodness of fit test. One sample/one dimensional K-S test is used to compare a sample with a prob function while two sample/ 2-D test is used to compare two samples.

The K–S value for a cumulative distribution function *F*(*x*) is given as

$$D_n = \sup_x \left| F_n(x) - F(x) \right| \quad ..(1)$$

where sup *x* is the supremum values of the distances.

b) The Anderson Darling (AD) test/ Shapiro Wilk test is a statistical test and are based on the distance

$$A = n\int_{-\infty}^{\infty} \frac{\left(F_n(x) - F(x)\right)^2}{F(x)\left(1 - F(x)\right)} dF(x) \quad ..(2)$$

where $w(x) = \left[F(x)\left(1 - F(x)\right)\right]^{-1}$ is a weight function.

c) A chi-squared test (χ^2) test are used to determine the difference between observed and expected frequency in one or more categories.

Table 1, Table 2 and Table 3 shows the KS-d, KS, AD (stat and p-value), chi square (p-value and df) for different distribution functions for MK2. The frequency and cumulative distribution curves for different distribution techniques are shown in Fig 1 and Fig 2 respectively.

Table 1 : Frequency and Cumulative function using KS Test for MK2

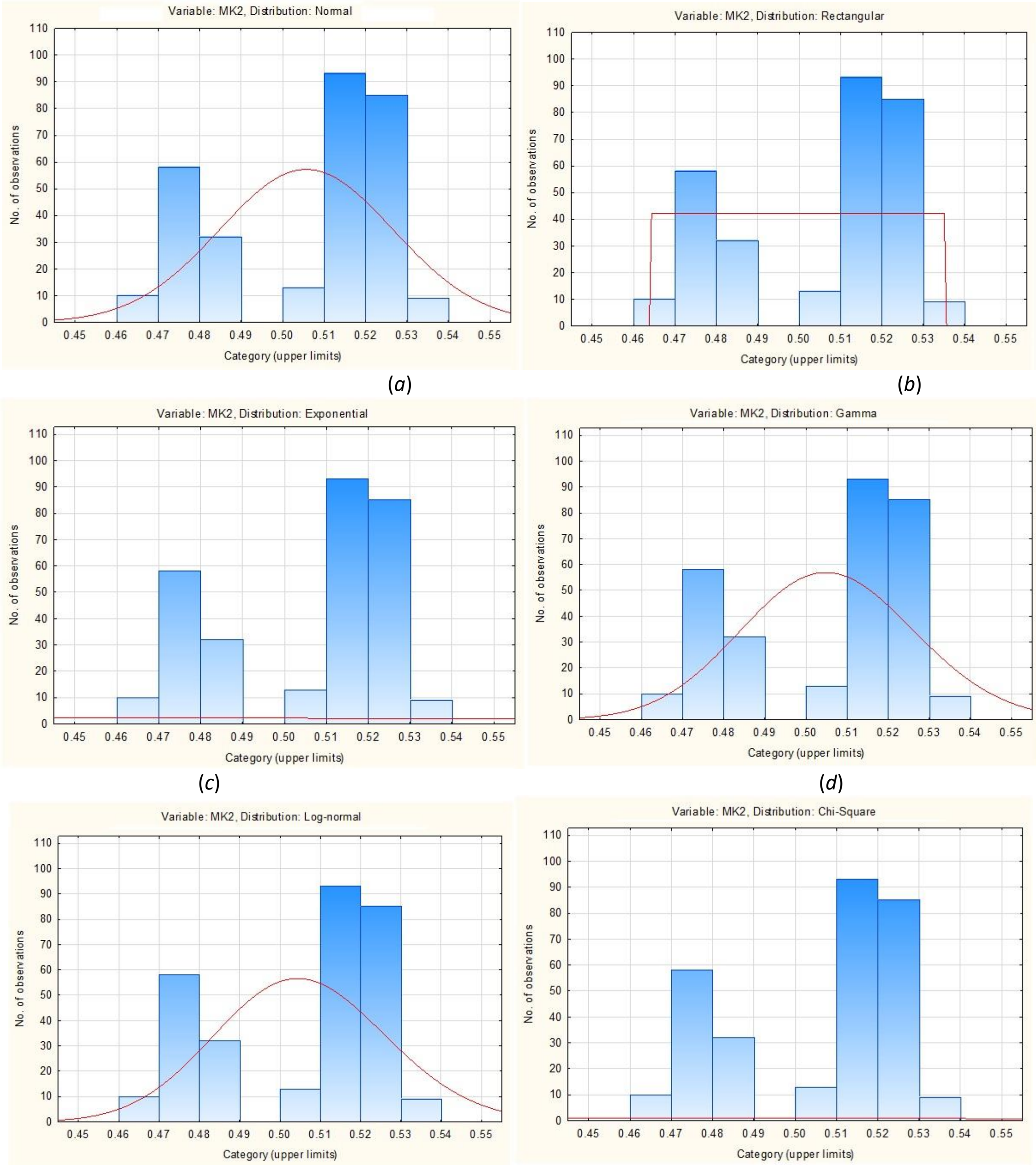
	K-S d	K-S	AD Stat	AD p-value	Chi-sq	Chi-sq p-value	Chi-sq df
Gaussian Mixture	0.026799	0.978501	0.1503	0.998562	6.067	0.108415	3
Weibull (scale,shape)	0.183078	0.000000	18.5049	0.000000	193.533	0.000000	7
Triangular (min,max,mode)	0.210974	0.000000	25.5090	0.000000	232.733	0.000000	6
Normal (location,scale)	0.231571	0.000000	22.6154	0.000000	266.733	0.000000	7
Log Normal (scale,shape)	0.236718	0.000000	23.2578	0.000000	279.467	0.000000	7
Rayleigh (scale)	0.568500	0.000000	117.2716	0.000000	1424.267	0.000000	8
Half Normal (scale)	0.640743	0.000000	144.1607	0.000000	1907.267	0.000000	8
General Pareto (scale, shape)	0.867025	0.000000	532.6983	0.000000	1555.267	0.000000	7

Table 2 : Frequency and Cumulative function for chi square test for MK2

	K-S d	K-S	AD Stat	AD p-value	Chi-sq	Chi-sq p-value	Chi-sq df
Gaussian Mixture(Mixing.Coef.1,Mean 1, Std.Dev 1, Mixing Coef.2,...)	0.026799	0.978501	0.1503	0.998562	6.067	0.108415	3
Normal (location,scale)	0.231571	0.000000	22.6154	0.000000	266.733	0.000000	7
Log Normal (scale,shape)	0.236718	0.000000	23.2578	0.000000	279.467	0.000000	7
Half Normal (scale)	0.640743	0.000000	144.1607	0.000000	1907.267	0.000000	8
Rayleigh (scale)	0.568500	0.000000	117.2716	0.000000	1424.267	0.000000	8
Weibull (scale,shape)	0.183078	0.000000	18.5049	0.000000	193.533	0.000000	7
General Pareto (scale,shape)	0.867025	0.000000	532.6983	0.000000	1555.267	0.000000	7
Triangular(min,max,mode)	0.210974	0.000000	25.5090	0.000000	232.733	0.000000	6

Table 3 : Frequency and Cumulative function for AD test for MK2

	K-S d	K-S	AD Stat	AD p-value	Chi-sq	Chi-sq p-value	Chi-sq df
Gaussian Mixture(Mixing.Coef.1,Mean 1, Std.Dev 1, Mixing Coef.2,...)	0.026799	0.978501	0.1503	0.998562	6.067	0.108415	3
Weibull (scale,shape)	0.183078	0.000000	18.5049	0.000000	193.533	0.000000	7
Normal (location,scale)	0.231571	0.000000	22.6154	0.000000	266.733	0.000000	7
Log Normal (scale,shape)	0.236718	0.000000	23.2578	0.000000	279.467	0.000000	7
Triangular(min,max,mode)	0.210974	0.000000	25.5090	0.000000	232.733	0.000000	6
Rayleigh (scale)	0.568500	0.000000	117.2716	0.000000	1424.267	0.000000	8
Half Normal (scale)	0.640743	0.000000	144.1607	0.000000	1907.267	0.000000	8
General Pareto (scale,shape)	0.867025	0.000000	532.6983	0.000000	1555.267	0.000000	7

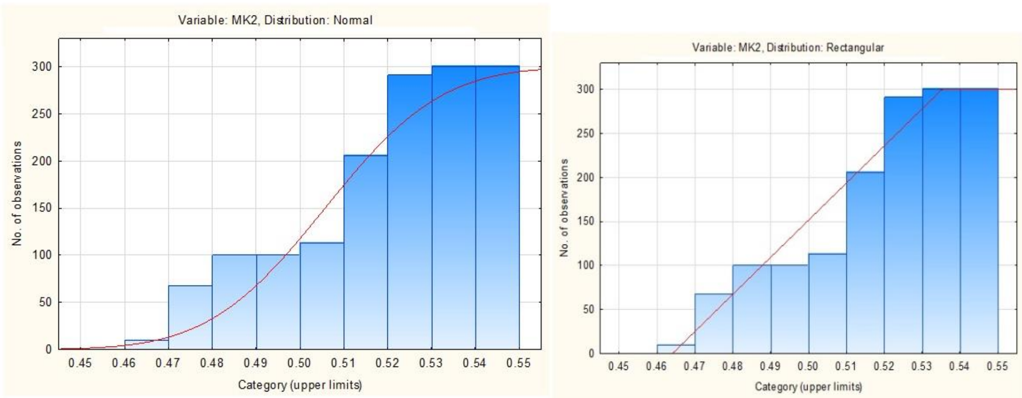


(*a*) (*b*)

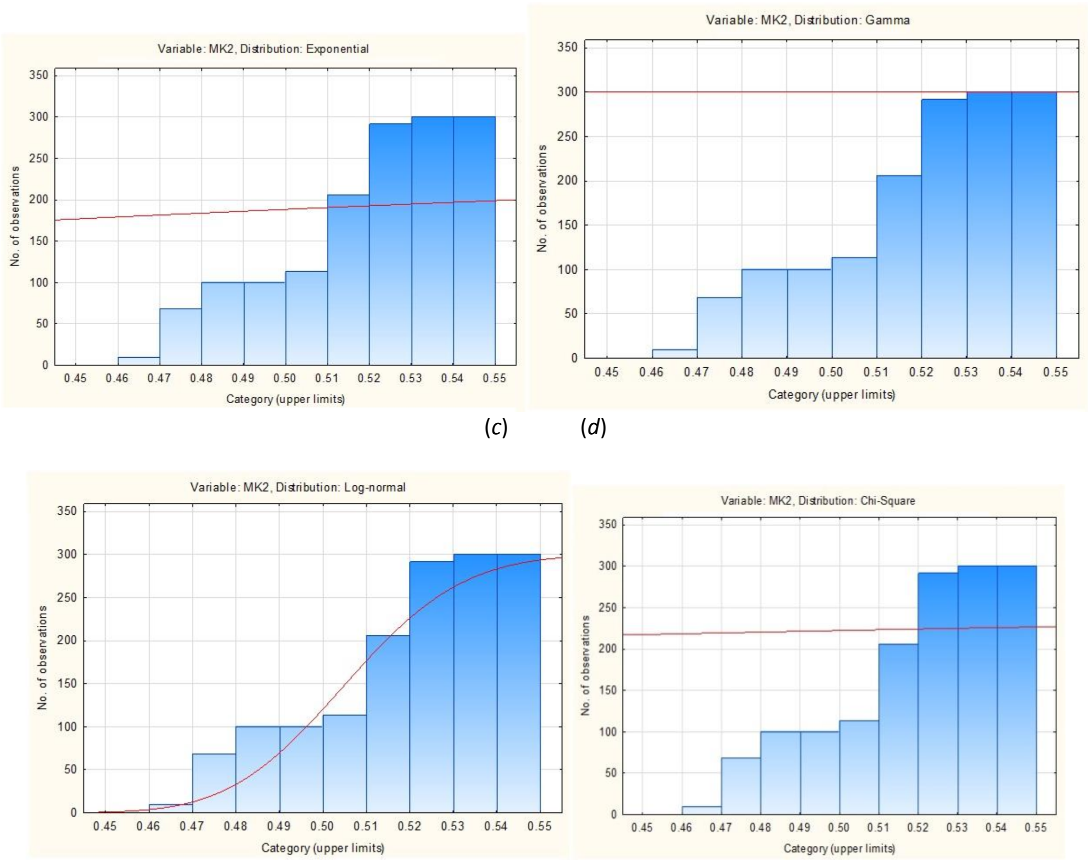
(*c*) (*d*)

(*e*) (*f*)

Fig 1 : Frequency Distribution curves for MK2 using different distribution techniques (*a*) Normal, (*b*) Rectangular , (*c*) Exponential, (*d*) Gamma, (*e*) Log-normal, (*f*) Chi-square.



(*a*) (*b*)



(*c*) (*d*)

(*e*) (*f*)

Fig 2 : Cumulative Distribution curves For MK2 using different distribution techniques (*a*) Normal, (*b*) Rectangular , (*c*) Exponential, (*d*) Gamma, (*e*) Log-normal, (*f*) Chi-square.

Table 4, table 5 and Table 6 shows the KS-d, KS, AD (stat and p-value), chi square (p-value and df) for different distribution functions for JNK. The frequency and cumulative distribution curves for different distribution techniques are shown in Fig 3 and Fig 4 respectively.

Table 4 : Frequency and Cumulative function for Kolmogorov-Smirnov tests for JNK

	K-S d	K-S	AD Stat	AD p-value	Chi-sq	Chi-sq p-value	Chi-sq df
Gaussian Mixture	0.026799	0.978501	0.1503	0.998562	6.067	0.108415	3
Weibull (scale,shape)	0.183078	0.000000	18.5049	0.000000	193.533	0.000000	7
Triangular(min,max,mode)	0.210974	0.000000	25.5090	0.000000	232.733	0.000000	6
Normal (location,scale)	0.231571	0.000000	22.6154	0.000000	266.733	0.000000	7
Log Normal (scale,shape)	0.236718	0.000000	23.2578	0.000000	279.467	0.000000	7
Rayleigh (scale)	0.568500	0.000000	117.2716	0.000000	1424.267	0.000000	8
Half Normal (scale)	0.640743	0.000000	144.1607	0.000000	1907.267	0.000000	8
General Pareto (scale,shape)	0.867025	0.000000	532.6983	0.000000	1555.267	0.000000	7

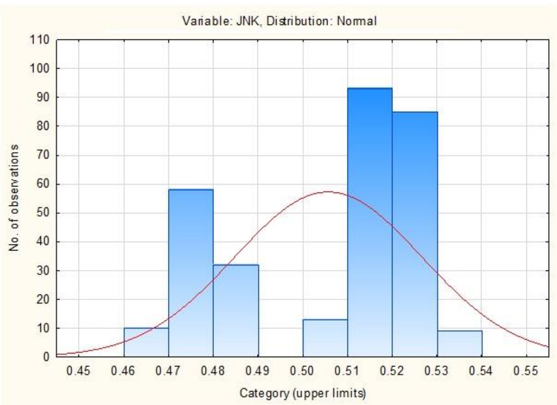
ISSN: 0975-8585

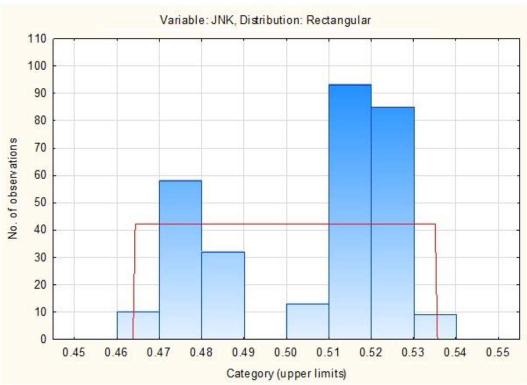
Table 5 : Frequency and Cumulative function for chi square tests for JNK

	K-S d	K-S	AD Stat	AD p-value	Chi-sq	Chi-sq p-value	Chi-sq df
Gaussian Mixture(Mixing.Coef.1,Mean 1, Std.Dev 1, Mixing Coef.2,...)	0.026799	0.978501	0.1503	0.998562	6.067	0.108415	3
Normal (location,scale)	0.231571	0.000000	22.6154	0.000000	266.733	0.000000	7
Log Normal (scale,shape)	0.236718	0.000000	23.2578	0.000000	279.467	0.000000	7
Half Normal (scale)	0.640743	0.000000	144.1607	0.000000	1907.267	0.000000	8
Rayleigh (scale)	0.568500	0.000000	117.2716	0.000000	1424.267	0.000000	8
Weibull (scale,shape)	0.183078	0.000000	18.5049	0.000000	193.533	0.000000	7
General Pareto (scale,shape)	0.867025	0.000000	532.6983	0.000000	1555.267	0.000000	7
Triangular(min,max,mode)	0.210974	0.000000	25.5090	0.000000	232.733	0.000000	6

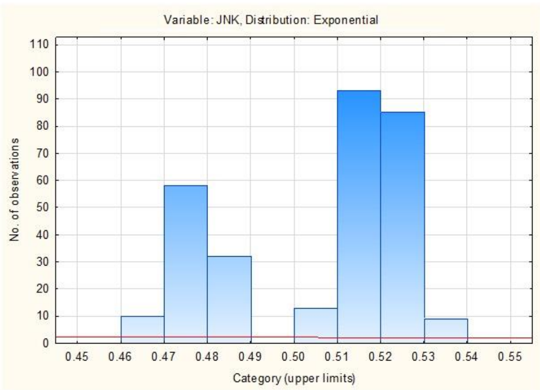
Table 6 : Frequency and Cumulative function for Anderson darling test for JNK

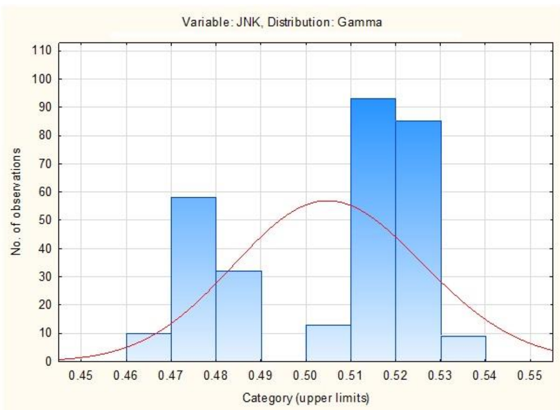
	K-S d	K-S	AD Stat	AD p-value	Chi-sq	Chi-sq p-value	Chi-sq df
Gaussian Mixture(Mixing.Coef.1,Mean 1, Std.Dev 1, Mixing Coef.2,...)	0.026799	0.978501	0.1503	0.998562	6.067	0.108415	3
Weibull (scale,shape)	0.183078	0.000000	18.5049	0.000000	193.533	0.000000	7
Normal (location,scale)	0.231571	0.000000	22.6154	0.000000	266.733	0.000000	7
Log Normal (scale,shape)	0.236718	0.000000	23.2578	0.000000	279.467	0.000000	7
Triangular(min,max,mode)	0.210974	0.000000	25.5090	0.000000	232.733	0.000000	6
Rayleigh (scale)	0.568500	0.000000	117.2716	0.000000	1424.267	0.000000	8
Half Normal (scale)	0.640743	0.000000	144.1607	0.000000	1907.267	0.000000	8
General Pareto (scale,shape)	0.867025	0.000000	532.6983	0.000000	1555.267	0.000000	7





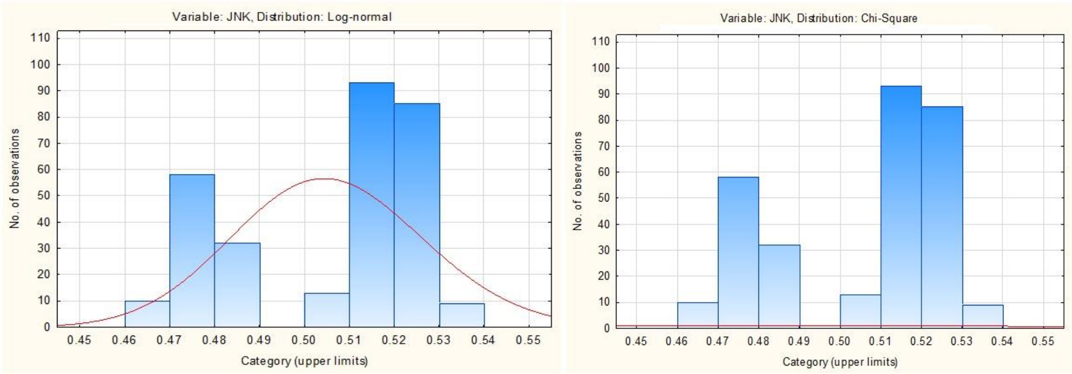
(*a*) (*b*)





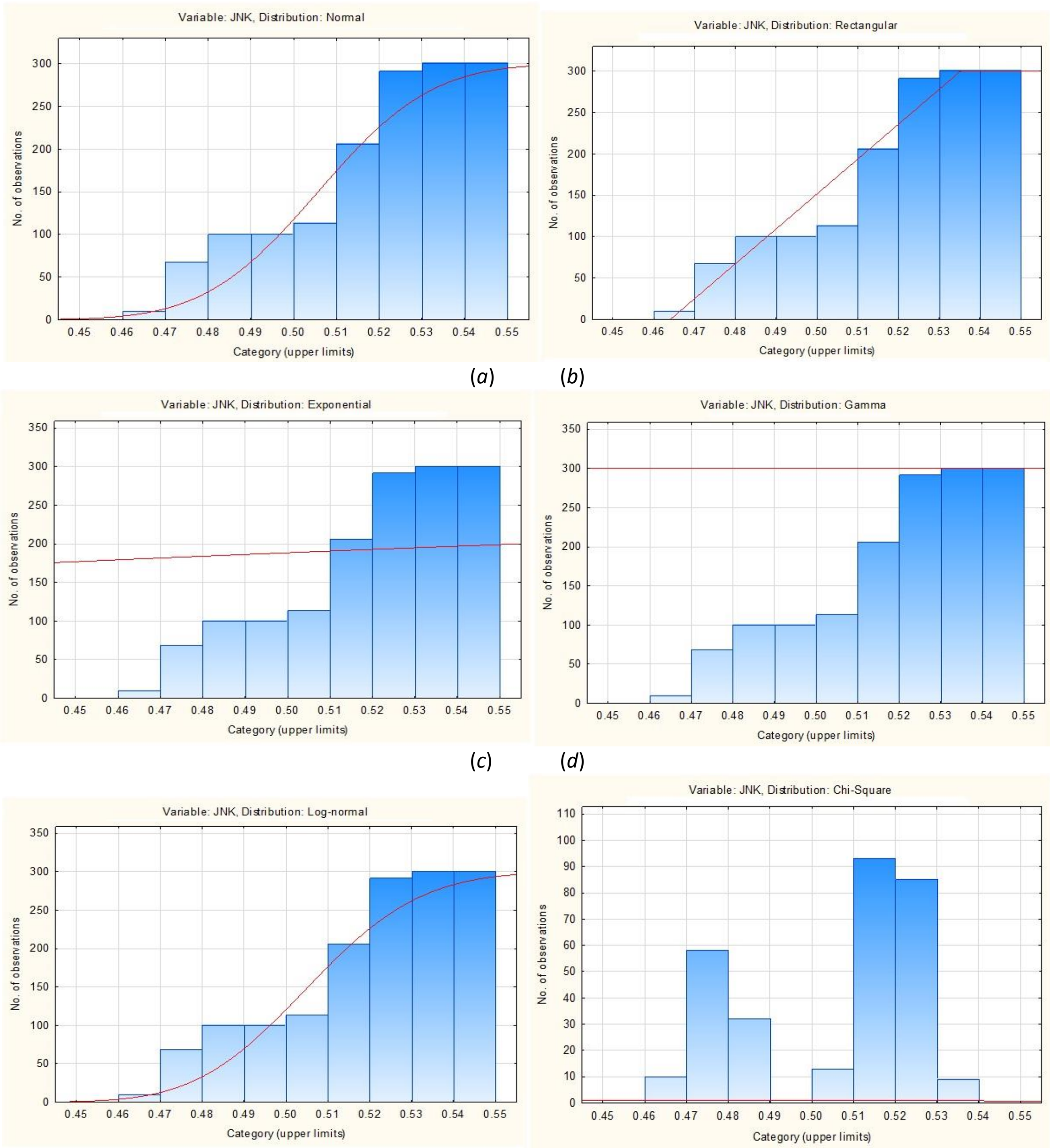
(*c*) (*d*)

ISSN: 0975-8585



(*e*) (*f*)

Fig 3 : Frequency Distribution curves For JNK using different distribution techniques (*a*) Normal, (*b*) Rectangular , (*c*) Exponential, (*d*) Gamma, (*e*) Log-normal, (*f*) Chi-square.



(*a*) (*b*) (*c*) (*d*) (*e*) (*f*)

Fig 4 : Cumulative Distribution curves For JNK using different distribution techniques (*a*) Normal, (*b*) Rectangular , (*c*) Exponential, (*d*) Gamma, (*e*) Log-normal, (*f*) Chi-square.

Table 7, Table 8 and Table 9 shows the KS-d, KS, AD stat, AD (stat and p-value), chi square (p-value and df) for different distribution functions for MK2. The frequency and cumulative distribution curves for different distribution techniques are shown in Fig 5 and Fig 6 respectively.

Table 7: Frequency and Cumulative function for Kolmogorov-Smirnov test for ERK

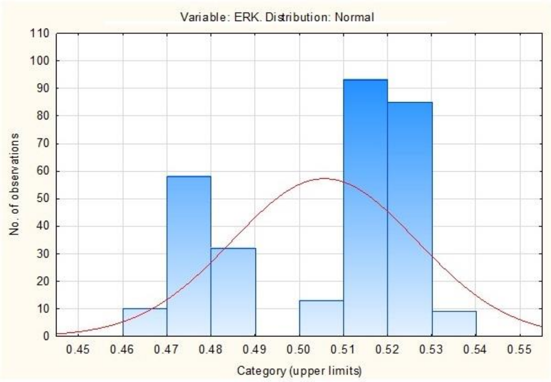
	K-S d	K-S	AD Stat	AD p-value	Chi-sq	Chi-sq p-value	Chi-sq df
Gaussian Mixture	0.026799	0.978501	0.1503	0.998562	6.067	0.108415	3
Weibull (scale,shape)	0.183078	0.000000	18.5049	0.000000	193.533	0.000000	7
Triangular(min,max,mode)	0.210974	0.000000	25.5090	0.000000	232.733	0.000000	6
Normal (location,scale)	0.231571	0.000000	22.6154	0.000000	266.733	0.000000	7
Log Normal (scale,shape)	0.236718	0.000000	23.2578	0.000000	279.467	0.000000	7
Rayleigh (scale)	0.568500	0.000000	117.2716	0.000000	1424.267	0.000000	8
Half Normal (scale)	0.640743	0.000000	144.1607	0.000000	1907.267	0.000000	8
General Pareto (scale,shape)	0.867025	0.000000	532.6983	0.000000	1555.267	0.000000	7

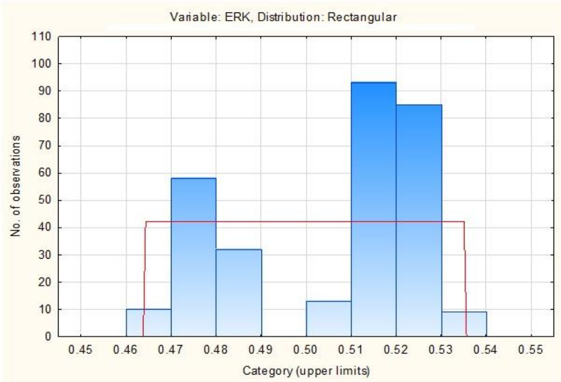
Table 8 : Frequency and Cumulative function for Chi square test for ERK

	K-S d	K-S	AD Stat	AD p-value	Chi-sq	Chi-sq p-value	Chi-sq df
Gaussian Mixture(Mixing.Coef.1,Mean 1, Std.Dev 1, Mixing Coef.2,...)	0.026799	0.978501	0.1503	0.998562	6.067	0.108415	3
Normal (location,scale)	0.231571	0.000000	22.6154	0.000000	266.733	0.000000	7
Log Normal (scale,shape)	0.236718	0.000000	23.2578	0.000000	279.467	0.000000	7
Half Normal (scale)	0.640743	0.000000	144.1607	0.000000	1907.267	0.000000	8
Rayleigh (scale)	0.568500	0.000000	117.2716	0.000000	1424.267	0.000000	8
Weibull (scale,shape)	0.183078	0.000000	18.5049	0.000000	193.533	0.000000	7
General Pareto (scale,shape)	0.867025	0.000000	532.6983	0.000000	1555.267	0.000000	7
Triangular(min,max,mode)	0.210974	0.000000	25.5090	0.000000	232.733	0.000000	6

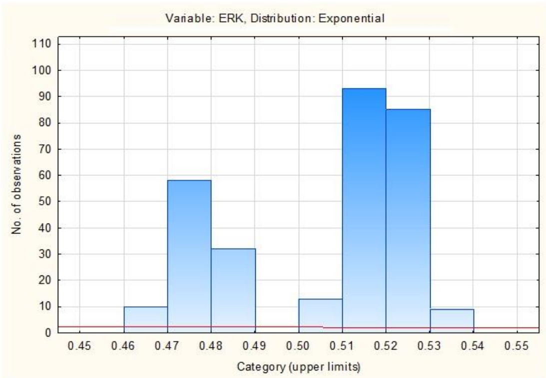
Table 9 : Frequency and Cumulative function for Anderson darling test for ERK

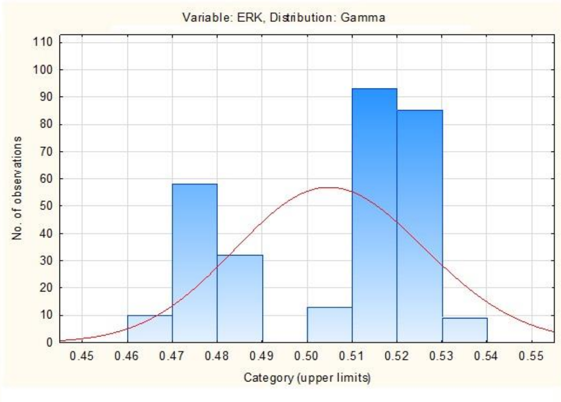
	K-S d	K-S	AD Stat	AD p-value	Chi-sq	Chi-sq p-value	Chi-sq df
Gaussian Mixture(Mixing.Coef.1,Mean 1, Std.Dev 1, Mixing Coef.2,...)	0.026799	0.978501	0.1503	0.998562	6.067	0.108415	3
Weibull (scale,shape)	0.183078	0.000000	18.5049	0.000000	193.533	0.000000	7
Normal (location,scale)	0.231571	0.000000	22.6154	0.000000	266.733	0.000000	7
Log Normal (scale,shape)	0.236718	0.000000	23.2578	0.000000	279.467	0.000000	7
Triangular(min,max,mode)	0.210974	0.000000	25.5090	0.000000	232.733	0.000000	6
Rayleigh (scale)	0.568500	0.000000	117.2716	0.000000	1424.267	0.000000	8
Half Normal (scale)	0.640743	0.000000	144.1607	0.000000	1907.267	0.000000	8
General Pareto (scale,shape)	0.867025	0.000000	532.6983	0.000000	1555.267	0.000000	7



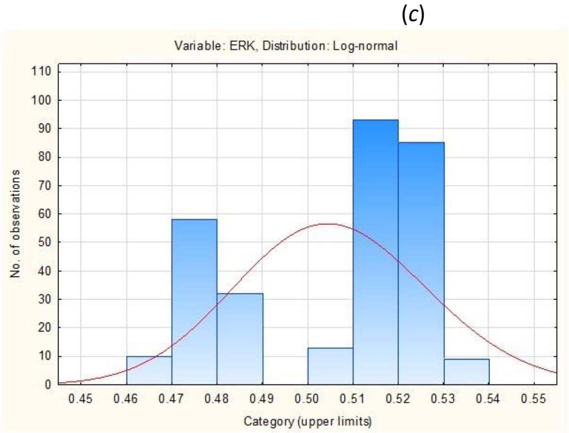


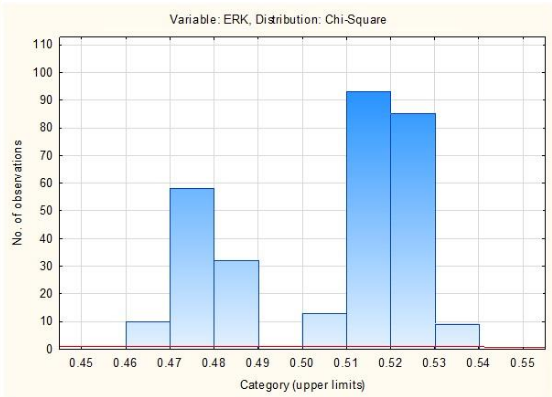
(*a*) (*b*)





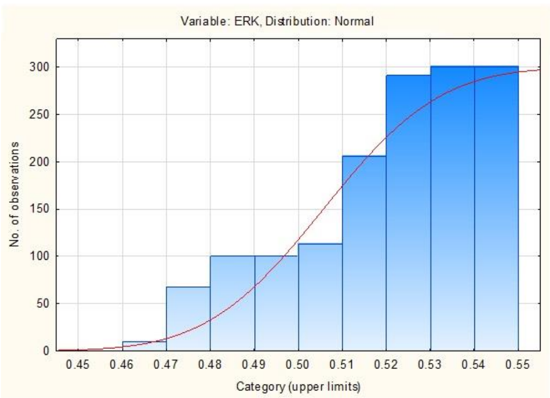
(*c*) (*d*)

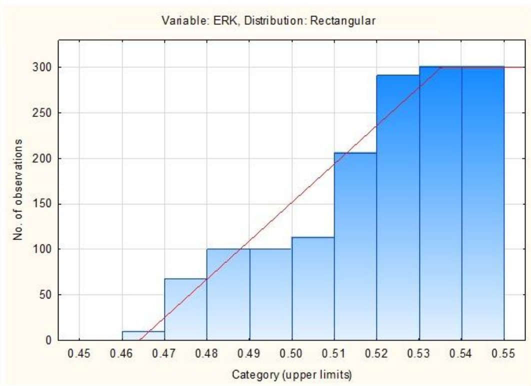




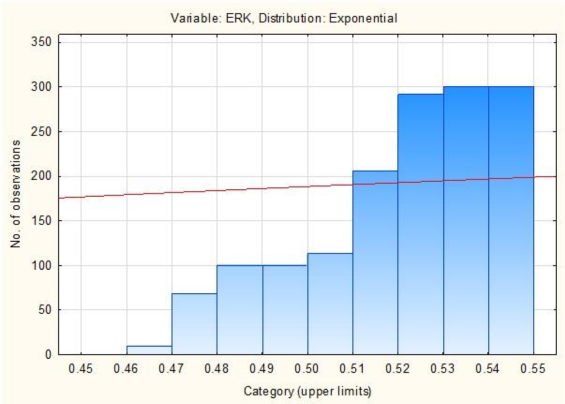
(*e*) (*f*)

Fig 5 : Frequency Distribution curves For ERK using different distribution techniques (*a*) Normal, (*b*) Rectangular , (*c*) Exponential, (*d*) Gamma, (*e*) Log-normal, (*f*) Chi-square.

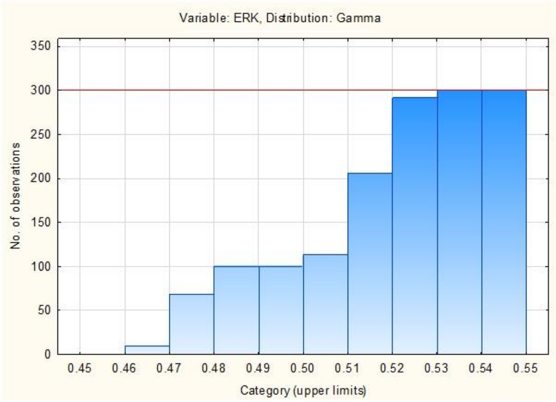




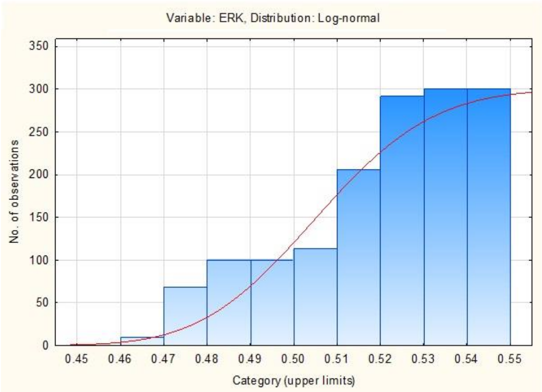
(*a*) (*b*)



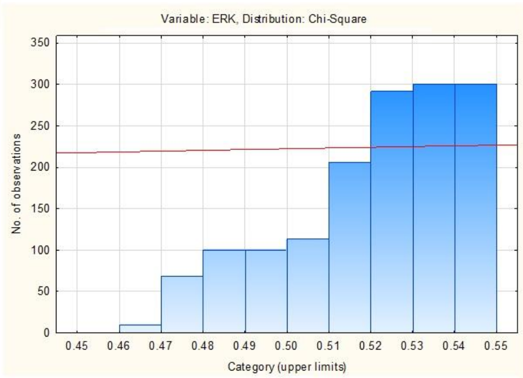
(*c*)



(*d*)



(*e*)



(*f*)

Fig 6 : Cumulative Distribution curves For ERK using different distribution techniques (*a*) Normal, (*b*) Rectangular , (*c*) Exponential, (*d*) Gamma, (*e*) Log-normal, (*f*) Chi-square.

CONCLUSION

A best fit model for survival/ death of mitogenic proteins a mathematical analysis was done using frequency and cumulative distribution functions. In this paper we have calculated KS, AD (statistics and p-value), chi-squrae (p- value and df) for frequency and cumulative distribution functions using KS, chi- square and spapiro wilk (AD) tests. Results with half normal distribution function are the best as their AD and chi square values are the maximum. In future we will find the pdf all proteins using different distribution functions.

REFERENCES

[1]. Jain S., "Communication of signals and responses leading to cell survival / cell death using Engineered Regulatory Networks". PhD Thesis, Jaypee University of Information Technology, Solan, Himachal Pradesh, India, 2012.

[2]. Weiss, R.,"Cellular computation and communications using engineered genetic regulatory networks". PhD Thesis, MIT, 2001.

[3]. Gaudet Suzanne, Janes Kevin A., Albeck John G., Pace Emily A., Lauffenburger Douglas A, and Sorger Peter K. (2005) A compendium of signals and responses trigerred by prodeath and prosurvival cytokines Manuscript M500158-MCP200.

[4]. Janes Kevin A, Albeck John G, Gaudet Suzanne, Sorger Peter K, Lauffenburger Douglas A, Yaffe Michael B. Dec.9, 2005 A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis; Science 310, 1646-1653.

[5]. Pearson, G., F. Robinson, G.T. Beers, B.E. Xu, M. Karandikar, K. Berman and M.H. Cobb, 2001. Mitogen-activated protein (MAP) kinase pathways: regulation and physiological functions. Endocrine Rev., 22(2):153-83.

[6]. Kyriakis, J.M. and J. Avruch, 1996. Sounding the alarm: protein kinase cascades activated by stress and inflammation. J. Biol. Chem., 271: 24313-24316.

[7]. Jain S, Naik P.K. , Sharma R, A Computational Modeling of cell survival/ death using VHDL and MATLAB Simulator, Digest Journal of Nanomaterials and Biostructures (DJNB), 4 (4): 863- 879, 2009, (ISSN 1842 – 3582).

[8]. Jain S, Bhooshan S. V., Naik P. K., “Model of Mitogen Activated Protein Kinases for Cell Survival/Death and its Equivalent Bio-Circuit”, Current Research Journal of Biological Sciences (CRJBS), 2(1):59-71, 2010.

[9]. Jain S, “Implementation of Fuzzy System using Operational Transconductance Amplifier for ERK pathway of EGF/ Insulin leading to Cell Survival/ Death”, J Pharm Biomed Sci, 2014, 4(8), 701-707.

[10]. Jain S, Chauhan D. S., “Implementation of fuzzy system using different voltages of OTA for JNK pathway leading to cell survival/ death”, Network Biology, 5(2), 62-70 : 2015.

[11]. Brockhaus M, Schoenfeld HJ, Schlaeger EJ, Hunziker W, Lesslauer W, and Loetscher H (1990) Identification of two types of tumor necrosis factor receptors on human cell lines by monoclonal antibodies. Proc Natl Acad Sci USA 87, 3127-3131.

[12]. Shruti Jain, Sunil V. Bhooshan, Pradeep K. Naik, “Mathematical modeling deciphering balance between cell survival and cell death using Tumor Necrosis Factor α”, *Research Journal of Pharmaceutical, Biological and Chemical Sciences* (*RJPBCS*), 574-583, 2(3): , 2011.

[13]. Libermann TA , Razon TA., Bartal AD, Yarden Y., Schlessinger J and Soreq H 1984 Expression of epidermal growth factor receptors in human brain tumors Cancer Res. 44,753-760.

[14]. Normanno N, De Luca A, Bianco C, Strizzi L, Mancino M, Maiello MR,, Carotenuto A, De Feo G, Caponiqro F, Salomon DS. 2006 Epidermal growth factor receptor (EGFR) signaling in cancer Gene 366, 2–16.

[15]. Jain S, Chauhan D. S., “Mathematical Analysis of Receptors For Survival Proteins”, International Journal of Pharma and Bio Sciences (IJPBS),6(3), 164-176 : 2015.

[16]. Lizcano J. M. Alessi D. R. 2002 The insulin signalling pathway. Curr Biol. 12, 236-238.

[17]. Shruti Jain, Sunil V. Bhooshan, Pradeep K. Naik, “Mathematical modeling deciphering balance between cell survival and cell death using insulin”, Network Biology, 1(1):46-58, 2011.