Article

Compedium model using frequency / cumulative distribution function for receptors of survival proteins: Epidermal growth factor and insulin

Shruti Jain

Department of Electronics and Communication Engineering, Jaypee University of Information Technology, Solan-173234, India
E-mail: jain.shruti15@gmail.com

Received 7 October 2016; *Accepted 14 October 2016*; *Published online 1 December 2016*

Abstract

In this paper I used the frequency and cumulative distribution functions to make a best model of the receptors of the survival proteins i.e. Epidermal Growth Factor Receptor (EGFR) and Insulin Receptor (IRS) using ten concentrations combination of TNF, EGF and Insulin. It has been revealed that survival and apoptosis signals induced by the receptors of EGF and insulin are temporarily separated and this is reflected in my model by the differences between the values of the parameters used. I conducted the analysis using KS-d, KS, AD stat, AD p-value, chi square, chi square p-value and chi square df for different distribution functions for EGFR and IRS. The frequency and cumulative distribution curves for different distribution techniques like exponential, log-normal, normal, gamma, chi-square etc are plotted using chi- square tests.

Keywords EGF; insulin; receptors; frequency distribution; cumulative distribution.

Network Biology
ISSN 2220-8879
URL: http://www.iaees.org/publications/journals/nb/online-version.asp
RSS: http://www.iaees.org/publications/journals/nb/rss.xml
E-mail: networkbiology@iaees.org
Editor-in-Chief: WenJun Zhang
Publisher: International Academy of Ecology and Environmental Sciences

1 Introduction

The programmed cell death (Jain, 2012; Suzzane, 2005; Weiss, 2001), Tumor Necrosis Factor α (TNF-α) (Brockhaus, 1990; Janes, 2005; Jain, 2011a, 2011b; Thoma, 1990), functions as apoptosis cues, whereas growth factors such as EGF (Arteaga, 2003; Libermann, 1984; Normano, 2006; Ullirich, 1990) and insulin (Lizcano, 2002; White, 2003; Jain, 2009, 2010, 2011a, 2010c, 2015a, 2015b) exert survival effects. These factors in single or in combination activate various key players in the network pertaining to cell survival/ apoptosis. Many proteins involved in this process that interact systematically regulating a specific pathway or cross talk with other proteins of different pathways. As a result many pathways activated simultaneously leading to many biochemical and physiological changes inside the cell. The final outcome of whether a cell dies or survives depends in the concentrations of key players among the pathways.

The epidermal growth factor (EGF) binds with its receptor to form EGF receptor (EGFR) which further binds with Src homology 2 (SH2) (Jain, 2015a; Toma, 1990) leading to the activation of RAS/extracellular signal regulated kinase (ERK) pathway (Kyriakis, 1996), the phosphatidylinositol 3 kinase(PI3K) pathway (Jain, 2014, 2015c, 2015d, 2016a, 2016b), and the activator of transcription (JAK/ STAT) pathway.

Insulin is a hormone that binds to its receptor (the insulin receptor) on cell membranes and initiates signal transduction leading to cell survival/ apoptosis. Binding of insulin to its receptor induces phosphorylation of tyrosine residue on the inner part of the receptor (White, 2003; Jain, 2015b). The phosphorylated tyrosine residues allow other intracellular proteins to bind to the intracellular domain of the receptor, and become phosphorylated.

In this paper I calculated KS, AD stat, AD p-value, chi square, chi square p-value and chi square df for different survival proteins. Later I plotted frequency distribution and cumulative distribution function curves for different distribution functions for EGFR and IRS.

2 Material and Methods

For mathematical modeling we have used frequency and cumulative distribution functions for cell survival/ cell death. Different parameters are calculated with the different concentrations of the TNF, EGF and Insulin (Jain, 2011b, 2012).

Cumulative distribution function specifies the distance of multivariate random variables X. The real valued random variables X with a given probability distance will found to have a value less than or equal to *x*. There are different test performed on different distribution functions: Anderson darling test, Kolmogorov–Smirnov test and chi- square test (Jain, 2016b).

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 probability function while two sample/ 2-D test is used to compare two samples. The K–S statistic for a given cumulative distribution function $F(x)$ is

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

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

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

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

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

3 Results

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

Table 1 Different distribution function values using Kolmogorov-Smirnov Test for EGFR.

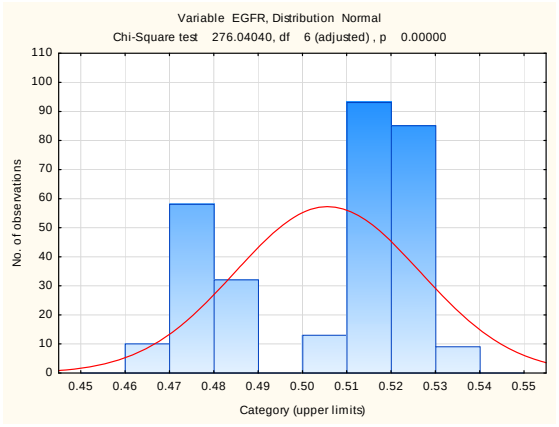
	K-S d	K-S	AD Stat	AD p-value	Chi-sq	Chi-sq p-value	Chi-sq df
Gaussian Mixture(Mixing.Coef.1,Mean 1, Std.Dev 1, Mixing Coef.2,...)	0 026799	0 978501	0 1503	0 998562	6 067	0 108415	3
Normal (location,scale)	0 183078	0 000000	18 5049	0 000000	193 533	0 000000	7
Log Normal (scale,shape)	0 231571	0 000000	22 6154	0 000000	266 733	0 000000	7
Half Normal (scale)	0 236718	0 000000	23 2578	0 000000	279 467	0 000000	7
Rayleigh (scale)	0 210974	0 000000	25 5090	0 000000	232 733	0 000000	6
Weibull (scale,shape)	0 568500	0 000000	117 2716	0 000000	1424 267	0 000000	8
General Pareto (scale,shape)	0 640743	0 000000	144 1607	0 000000	1907 267	0 000000	8
Triangular(min,max,mode)	0 867025	0 000000	532 6983	0 000000	1555 267	0 000000	7

Table 2 Different distribution function values using chi square test for EGFR.

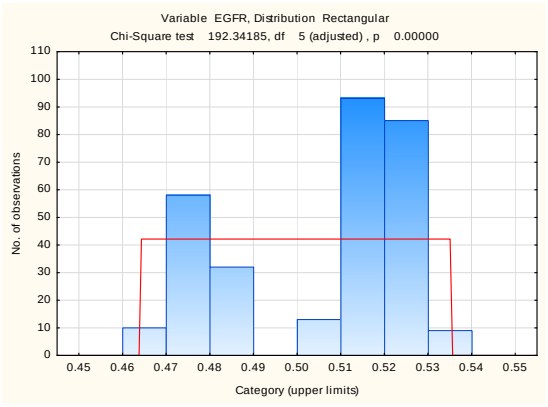
	K-S d	K-S	AD Stat	AD p-value	Chi-sq	Chi-sq p-value	Chi-sq df
Gaussian Mixture	0 026799	0 978501	0 1503	0 998562	6 067	0 108415	3
Normal (location,scale)	0 231571	0 000000	22 6154	0 000000	266 733	0 000000	7
Log Normal (scale,shape)	0 236718	0 000000	23 2578	0 000000	279 467	0 000000	7
Half Normal (scale)	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 Different distribution function values using Anderson darling test for EGFR.

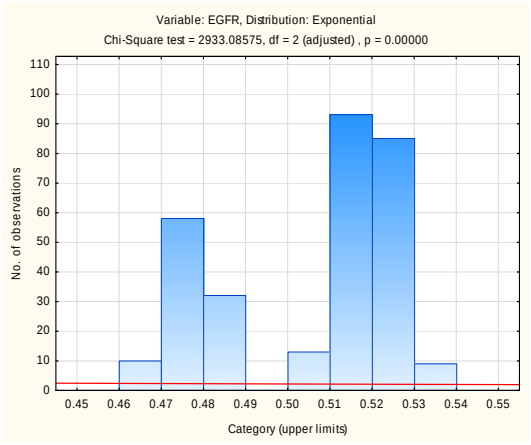
	K-S d	K-S	AD Stat	AD p-value	Chi-sq	Chi-sq p-value	Chi-sq df
Gaussian Mixture	0 026799	0 978501	0 1503	0 998562	6 067	0 108415	3
Weibull (scale,shape)	0 183078	0 000000	18 5049	0 000000	193 533	0 000000	7
Normal (location,scale)	0 231571	0 000000	22 6154	0 000000	266 733	0 000000	7
Log Normal (scale,shape)	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



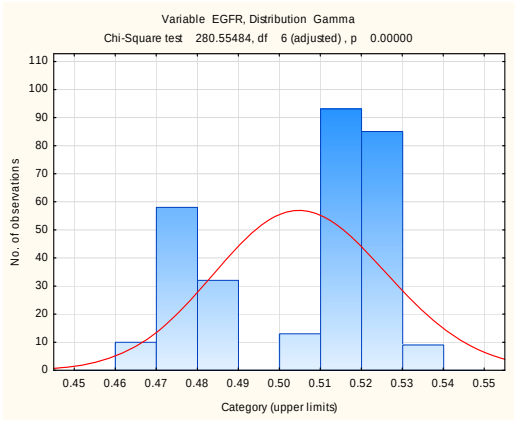
(a)



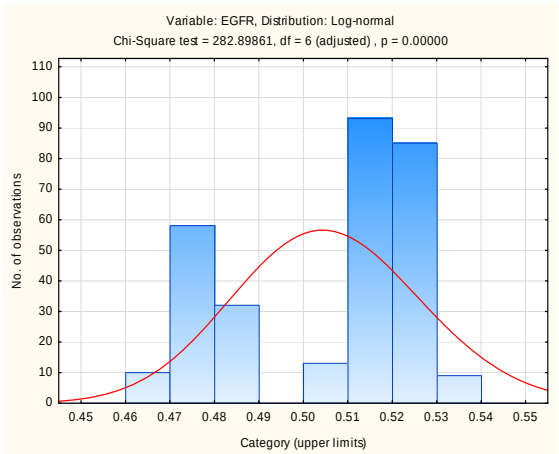
(b)



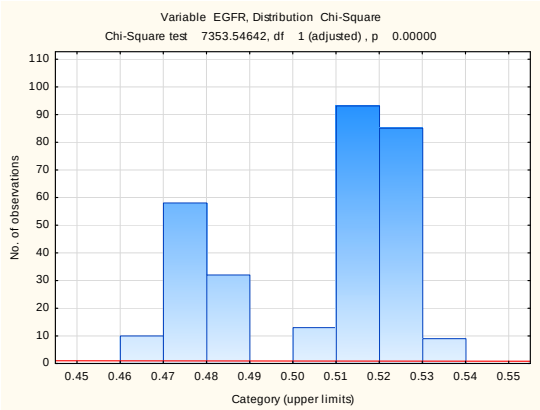
(c)



(d)

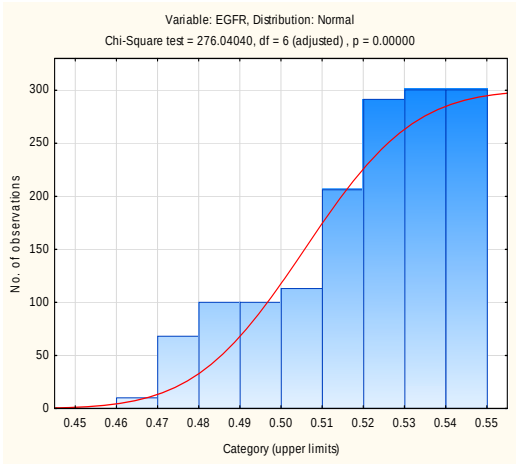


(e)

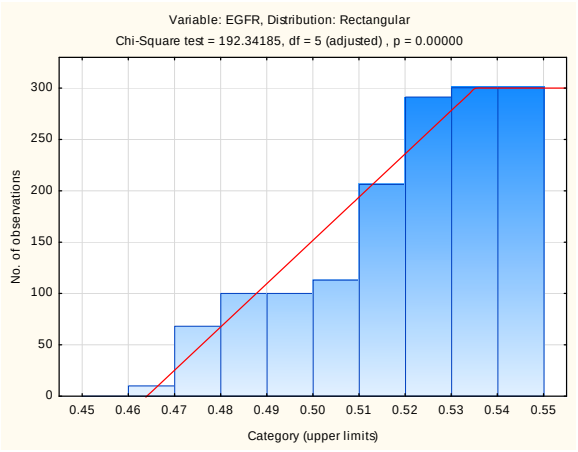


(f)

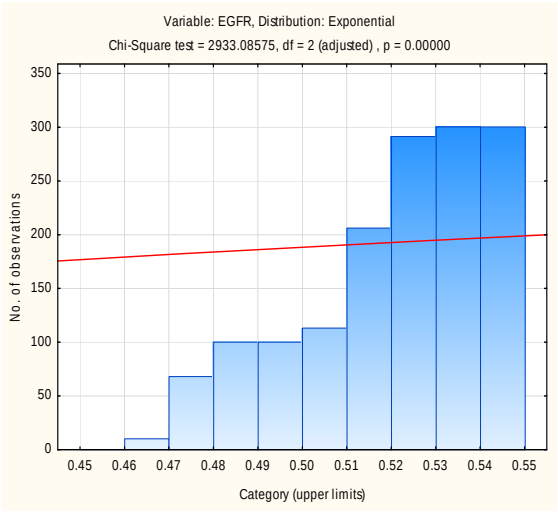
Fig. 1 Frequency Distribution curves for EGFR using different distribution techniques (a) Normal, (b) Rectangular, (c) Exponential, (d) Gamma, (e) Log-normal, (f) Chi-square.



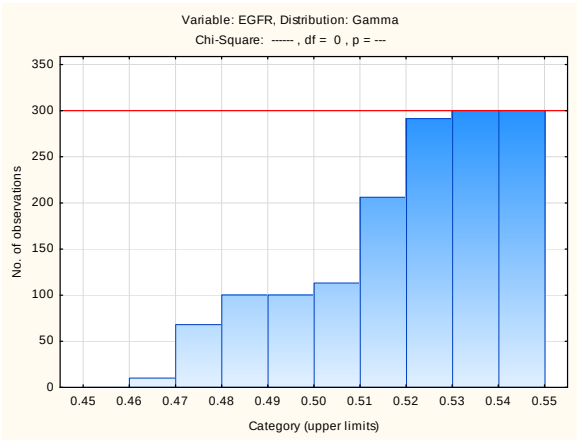
(a)



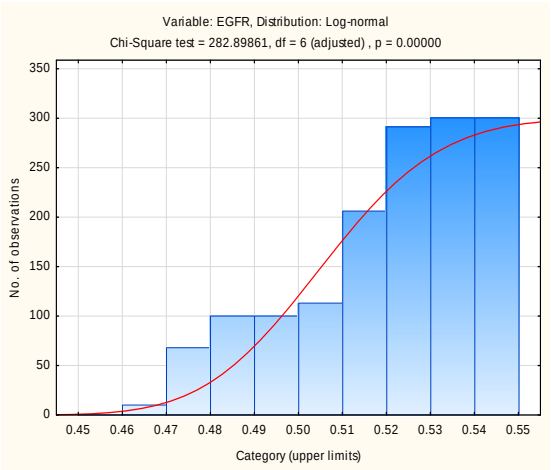
(b)



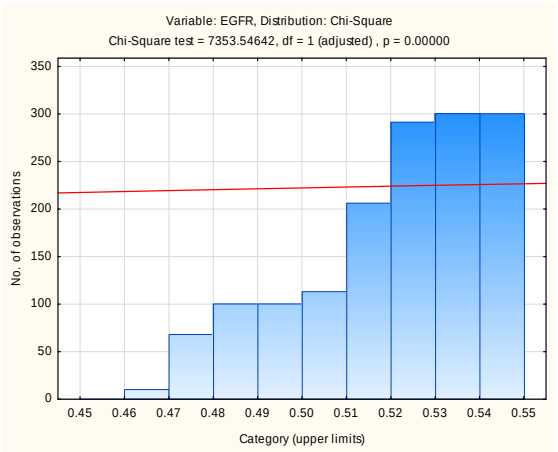
(c)



(d)



(e)



(f)

Fig. 2 Cumulative Distribution curves for EGFR 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, AD p-value, chi square, chi square p-value and chi square df for different distribution functions for IRS. The frequency distribution curves and cumulative distribution curves for different distribution techniques are shown in Fig. 3 and Fig. 4 respectively.

Table 4 Different distribution function values using Kolmogorov-Smirnov Test for IRS.

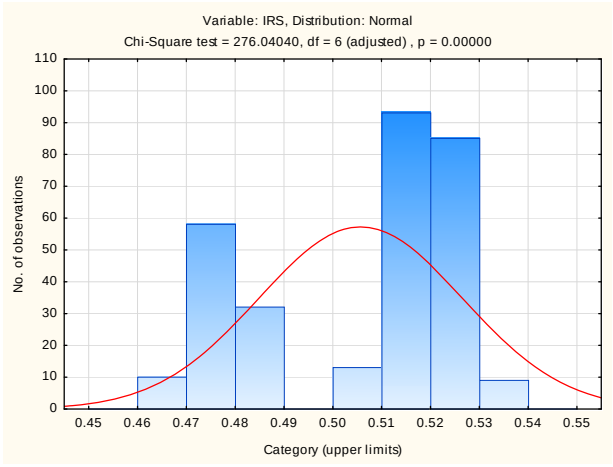
	K-S d	K-S	AD Stat	AD p-value	Chi-sq	Chi-sq p-value	Chi-sq df
Gaussian Mixture(Mixing.Coef.1,Mean 1, Std.Dev 1, Mixing Coef.2,...)	0 026799	0 978501	0 1503	0 998562	6 067	0 108415	3
Normal (location,scale)	0 183078	0 000000	18 5049	0 000000	193 533	0 000000	7
Log Normal (scale,shape)	0 210974	0 000000	25 5090	0 000000	232 733	0 000000	6
Half Normal (scale)	0 231571	0 000000	22 6154	0 000000	266 733	0 000000	7
Rayleigh (scale)	0 236718	0 000000	23 2578	0 000000	279 467	0 000000	7
Weibull (scale,shape)	0 568500	0 000000	117 2716	0 000000	1424 267	0 000000	8
General Pareto (scale,shape)	0 640743	0 000000	144 1607	0 000000	1907 267	0 000000	8
Triangular(min,max,mode)	0 867025	0 000000	532 6983	0 000000	1555 267	0 000000	7

Table 5 Different distribution function values using chi square test for IRS.

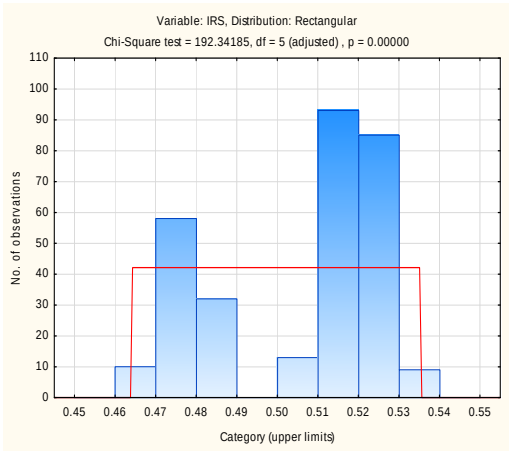
	K-S d	K-S	AD Stat	AD p-value	Chi-sq	Chi-sq p-value	Chi-sq df
Gaussian Mixture	0 026799	0 978501	0 1503	0 998562	6 067	0 108415	3
Normal (location,scale)	0 231571	0 000000	22 6154	0 000000	266 733	0 000000	7
Log Normal (scale,shape)	0 236718	0 000000	23 2578	0 000000	279 467	0 000000	7
Half Normal (scale)	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 Different distribution function values using Anderson darling test for IRS.

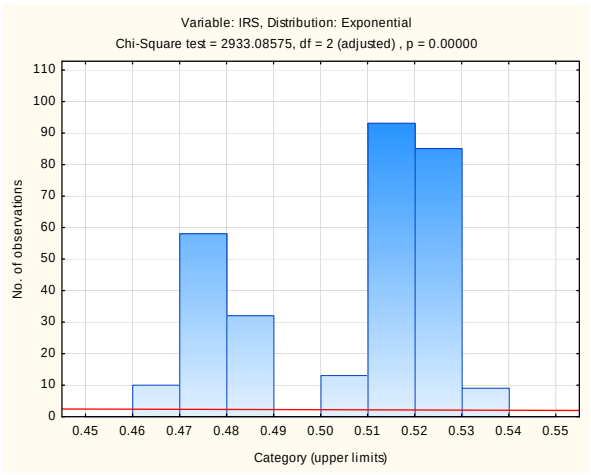
	K-S d	K-S	AD Stat	AD p-value	Chi-sq	Chi-sq p-value	Chi-sq df
Gaussian Mixture	0 026799	0 978501	0 1503	0 998562	6 067	0 108415	3
Weibull (scale,shape)	0 183078	0 000000	18 5049	0 000000	193 533	0 000000	7
Normal (location,scale)	0 231571	0 000000	22 6154	0 000000	266 733	0 000000	7
Log Normal (scale,shape)	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



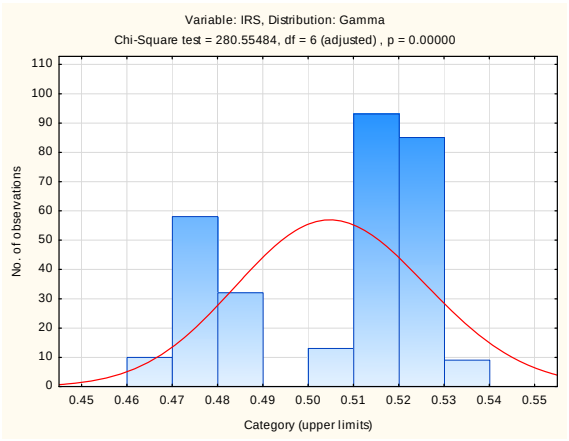
(a)



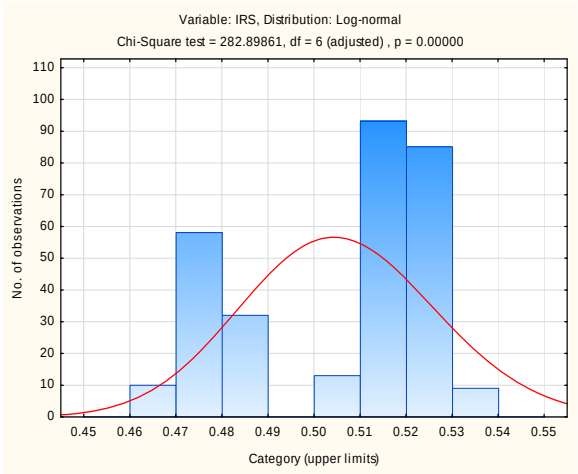
(b)



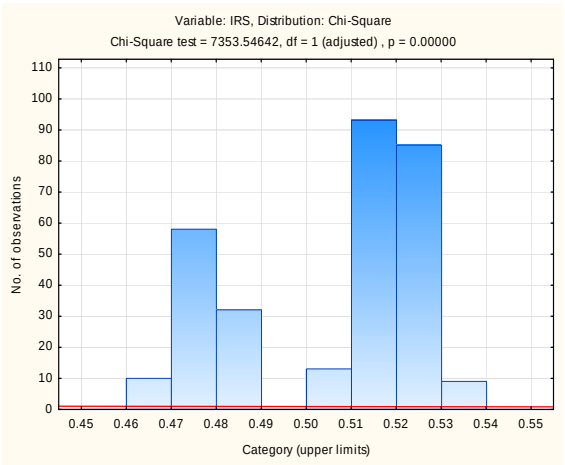
(c)



(d)



(e)



(f)

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

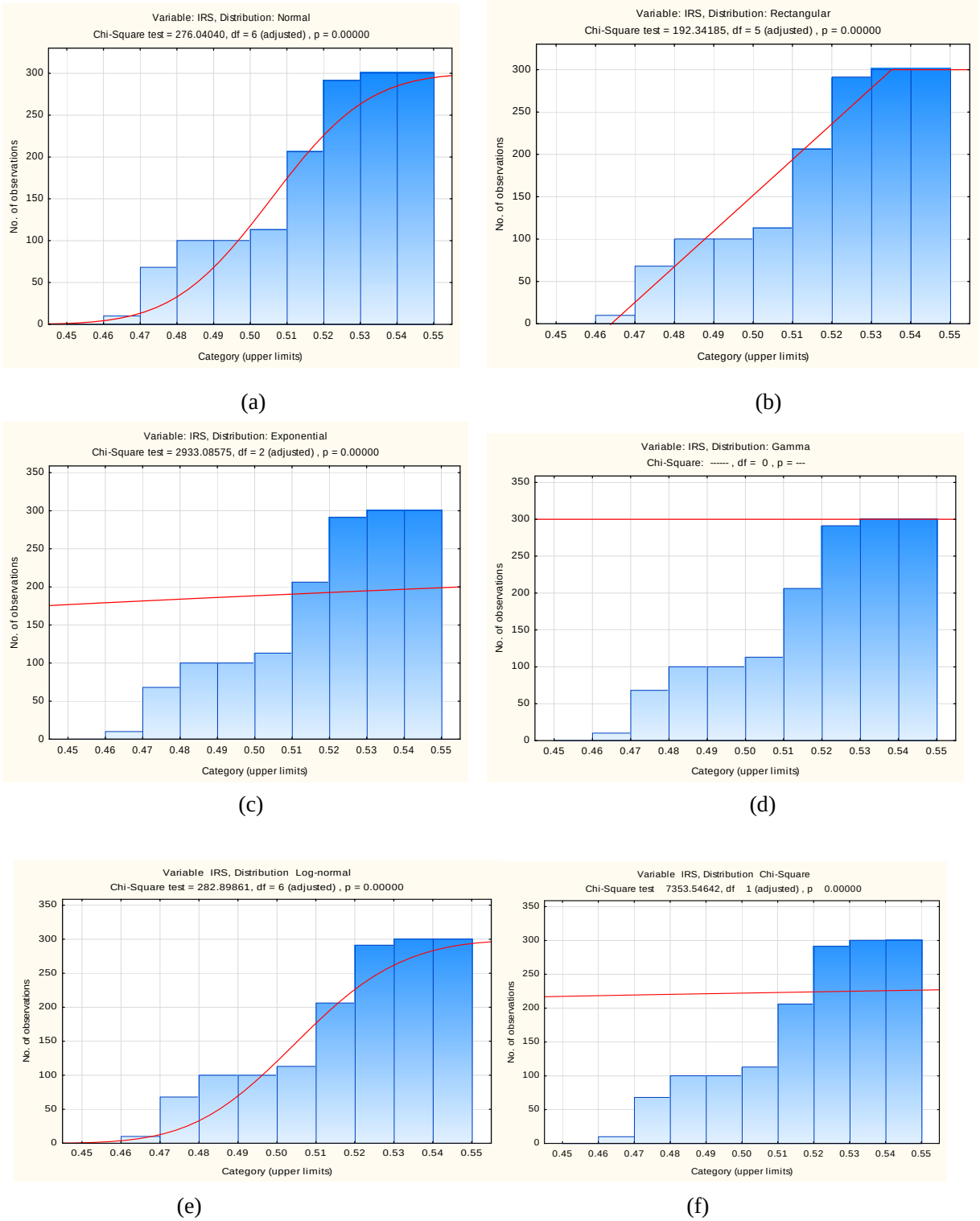


Fig 4 Cummulative Distribution curves for IRS using different distribution techniques (a) Normal, (b) Rectangular, (c) Exponential, (d) Gamma, (e) Log-normal, (f) Chi-square.

4 Conclusion

In this paper we have find the KS-d, KS, AD stat, AD p-value, chi square, chi square p-value and chi square df for different distribution functions for EGFR and IRS. The frequency and cumulative distribution curves for different distribution techniques like exponential, log-normal, normal, gamma, chi-square etc are plotted. More generally, these models are flexible, able to incorporate qualitative and noisy data, and powerful enough to produce quantitative predictions and new biological insights about the operation of signaling networks. In future we will design a CAD system that will help radiologists using wavelet transforms.

References

Arteaga C. 2003. Targeting HER1/EGFR: a molecular approach to cancer therapy. Seminars in Oncology, 30: 314

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. Proceedings of the National Academy of Sciences of USA, 87: 3127-3131

Jain S, Naik PK, Sharma R 2009 A computational modeling of cell survival/ death using VHDL and MATLAB simulator. Digest Journal of Nanomaterials and Biostructures, 4(4): 863- 879

Jain S, Bhooshan SV, Naik PK 2010 Model of Mitogen Activated Protein Kinases for Cell Survival/Death and its Equivalent Bio-Circuit. Current Research Journal of Biological Sciences, 2(1): 59-71

Jain S, Bhooshan SV, Naik PK. 2011a Mathematical modeling deciphering balance between cell survival and cell death using insulin. Network Biology, 1(1): 46-58

Jain S, Bhooshan SV, Naik PK. 2011b. Mathematical modeling deciphering balance between cell survival and cell death using Tumor Necrosis Factor α. Research Journal of Pharmaceutical, Biological and Chemical Sciences, 2(3): 574-583

Jain S, 2012 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

Jain S. 2014. Implementation of fuzzy system using operational transconductance amplifier for ERK pathway of EGF/ Insulin leading to cell survival/ death. Journal of Pharmaceutical and Biomedical Sciences, 4(8): 701-707

Jain S, Chauhan DS. 2015a. Mathematical analysis of receptors for survival proteins. International Journal of Pharma and Bio Sciences, 6(3): 164-176

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

Jain S. 2015c. Mathematical Analysis and Probability Density Function of FKHR pathway for Cell Survival /Death. Control System and Power Electronics – CSPE 2015. 84-93, Banglore, India

Jain S, Chauhan DS. 2015d, Linear and Non Linear Modeling of Protein Kinase B/ AkT, International Conference on Information and Communication Technology for Sustainable Development (ICT4SD-2015). 81-88, Ahmedabad, India

Jain S. 2016a, Regression analysis on different mitogenic pathways. Network Biology, 6(2): 40-46

Jain S. 2016b. Mathematical analysis using frequency and cumulative distribution functions for mitogenic pathway. Research Journal of Pharmaceutical, Biological and Chemical Sciences, 7(3): 262-272

Janes KA, John AG, Suzanne G, Peter SK, Douglas LA, Michael YB. 2005 A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis. Science, 310: 1646-1653

Kyriakis JM., Avruch J. 1996. Sounding the alarm: protein kinase cascades activated by stress and inflammation. Journal of Biological Chemistry, 271: 24313-24316

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

Lizcano J. M. Alessi D. R. 2002. The insulin signalling pathway. Current Biology, 12: 236-238

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

Suzanne G, Janes KA, John AG, Emily PA, Douglas LA, Peter SK. 2005. A compendium of signals and responses trigerred by prodeath and prosurvival cytokines. Manuscript M500158-MCP200. MCP,USA

Thoma B, Grell M, Pfizenmaier K, and Scheurich P 1990, Identification of a 60-kD tumor necrosis factor (TNF) receptor as the major signal transducing component in TNF responses. Journal of Experimental Medicine, 172: 1019-1023

Ullrich A, Schlessinger J. 1990. Signal transduction by receptors with tyrosine kinase activity. Cell, 61: 203-211

Weiss R. 2001. Cellular computation and communications using engineered genetic regulatory networks. PhD Thesis, MIT, USA

White MF. 2003. Insulin signaling in health and disease Science, 302(5651): 1710-1711