Interactive Applications of MS Excel 2010

Gurmukh Singh, Ph.D.
Department of Computer and Information Sciences
State University of New York at Fredonia
Fredonia, NY 14063
singh@fredonia.edu

2nd Best Teaching Practices Conference
April 16, 2012

Main Objectives of Presentation

- MS Excel 2010 to develop interactive applications in computer science, information sciences, science education, natural sciences, genetics and bioinformatics.
- Also for computational physics/chemistry, and physics & chemistry education etc.
  - Actually, we will present computer simulations of rolling of nine dice [1], chances of winning of NY State Lotto [2]
  - Simulating Mendel's Laws in genetics and bioinformatics [3].
  - Simulating Bohr's quantum theory of H-atom in a virtual lab [4] (basis of quantum computing).
  - Simulating discrete X-ray spectral lines in quantum mechanics (being developed with MS Excel 2010 and MS Visual Studio .NET 2010 [5]).

Why Microsoft Excel 2010 for Scholarly Work?

- Development and advancement in high speed micro-computers such as IBM and Mac based PCs & Tablets
- Portable laptops/Tablets as versatile research, teaching and learning tools for natural and medical sciences
- Microcomputer machines employ several software systems such as MS Visual Studio .NET, MS Office Suite, Open Office, Open Source Unix/Linux based computers etc.
- Development of object oriented computing languages such as C++, C#, Visual Basic (VB), Java, Java Script, SQL etc.
- Such software systems are extensively used by researches in higher education, scientific labs, private companies, businesses and banks in the entire world.
- Next to MS Word, Excel is the most used software in USA.

Why Microsoft Excel 2010 in Research Work?

- Adoption of internet technologies in scholarly work done in science, engineering and medical informatics
- Use of Internet technologies to collaborate in local and global scholarly work in national and international institutes of higher education anywhere, anytime 24/7
- Internet technologies are very effective collaborative tools for natural, engineering, social and medical sciences
- Numerous International/National conferences are held to enhance and share the knowledge gathered with other educators and researchers using internet technologies
- Even online conferences and webinars can be held anywhere anytime 24/7

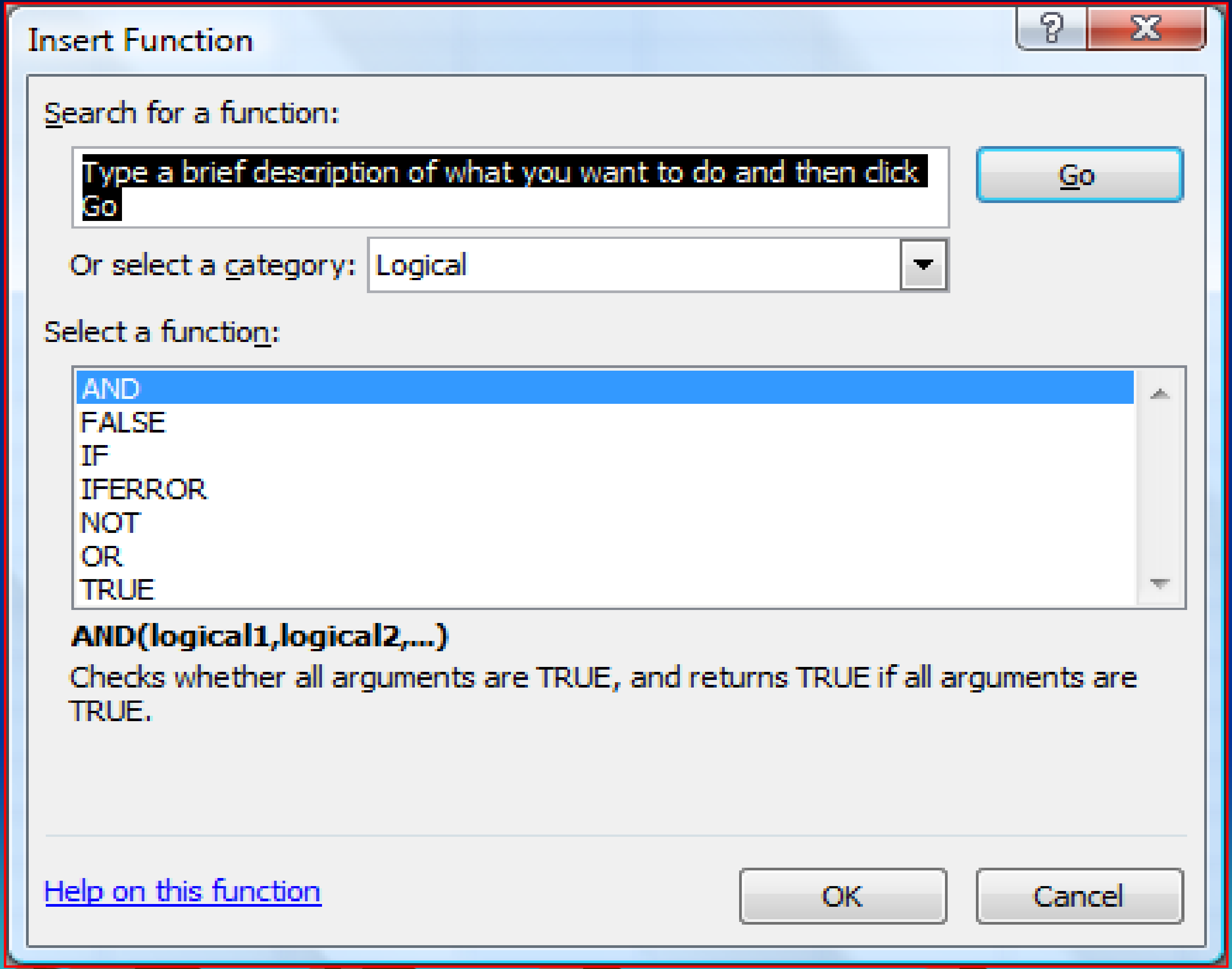
Assumptions Used in Interactive Simulations

- Rolling of Nine Dice for CIS [1]
  - No face of nine dice has zero dots
  - Maximum number of dots on each dice faces is six
- Chance of Winning of NY State Lotto CIS [2]
  - Winning number must be more than zero
  - Six winning numbers must be in the range of 1 and 56
- Modeling Mendel's Laws in genetics and bioinformatics [3]
  - Dominant and recessive male and female genes participate in gamete formation
  - 50% probability of gamete formation from opposite sex chromosomes or genes
- Modeling quantum theory of H-atom for natural sciences [4]
  - Electronic transitions involve discrete amount of energy
  - Electrons do not emit radiation in ground states

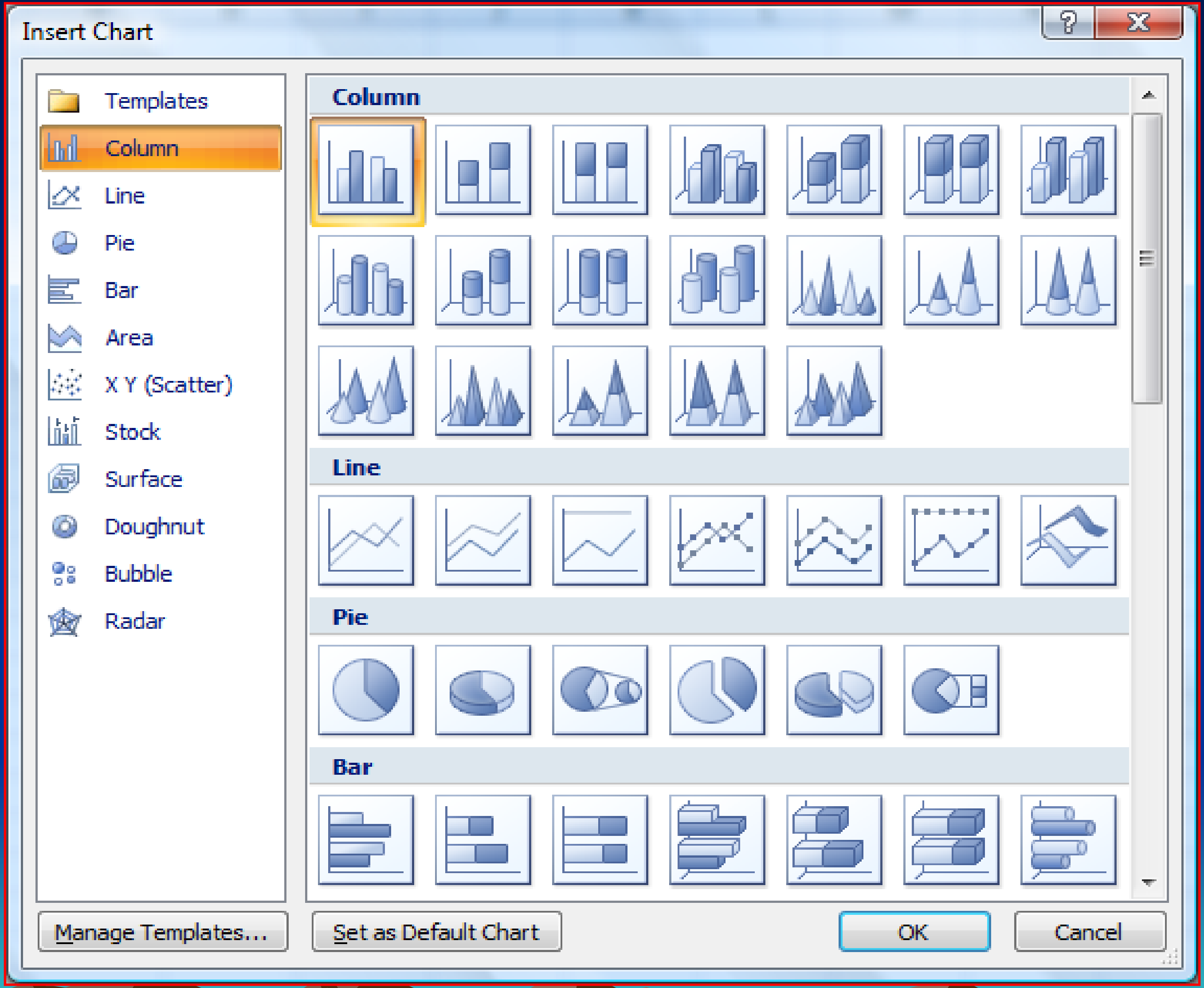
MS Excel 2010 Built-in Functions [6]

- Logical function:
  - ❖ IF(logical_test, [value_if_true], [value_if_false])
  - ❖ AND(logical1, [logical2],…)
  - ❖ AND returns a TRUE value if all arguments of the function are TURE
- Pseudo-random number generating function:
  - ❖ RAND( )
- RAND( ) function could be used to generate any kind of fractional pseudo-random number in a range between 0 and 1
- Performed up to 20,000 computer simulations in each interactive application for the current presentation

Screenshot of Excel Built-in Functions



Screenshot of Graph Plotting Capabilities



Examples of Interactive Simulations performed with MS Excel 2010

- Rolling of Nine Dice (Computer and Information Sciences and probability problem)
- Chances of Winning of NY State Lotto (Computer and Information Sciences and probability problem)
- Simulating Mendel's Laws in genetics and bioinformatics
- Simulating quantum theory of H-atom for natural sciences
- Simulating discrete X-ray spectral lines in quantum mechanics

Table 1: Simulated values of number of dots on six faces of each dice in rolling process

Dice 1	Dice 2	Dice 3	Dice 4	Dice 5	Dice 6	Dice 7	Dice 8	Dice 9	Total	Total/Max
3	4	1	3	5	3	3	6	3	31	0.70
4	6	3	6	2	3	3	3	1	31	0.70
3	4	2	5	4	4	4	6	2	34	0.77
6	5	1	1	1	3	5	3	1	26	0.59
5	1	3	6	1	6	5	1	3	31	0.70
4	1	2	2	2	6	2	1	6	26	0.59
4	4	6	4	2	2	3	5	2	32	0.73
3	6	1	6	6	1	2	3	5	33	0.75
2	1	3	4	5	5	4	6	4	34	0.77
2	6	1	4	4	3	4	4	2	30	0.68
2	6	1	1	5	6	5	6	5	37	0.84
6	1	6	6	5	1	6	4	6	41	0.93
4	1	6	5	5	4	2	4	3	34	0.77
3	6	5	5	4	3	3	5	6	40	0.91
6	5	4	3	1	3	2	6	1	31	0.70
1	1	6	2	3	2	3	5	6	29	0.66
1	5	2	2	2	1	6	1	6	26	0.59
5	2	3	1	2	5	4	2	4	28	0.64
1	3	4	1	6	6	5	4	2	32	0.73
3	5	3	5	6	4	6	4	4	40	0.91
5	1	6	6	4	6	3	6	3	40	0.91
6	3	6	1	3	4	6	2	2	33	0.75
1	1	1	6	2	5	2	2	6	26	0.59
4	3	3	1	4	3	2	2	6	28	0.64

Fig. 1: A plot of ratio of total score in one row to the maximum score as a function of number of trials

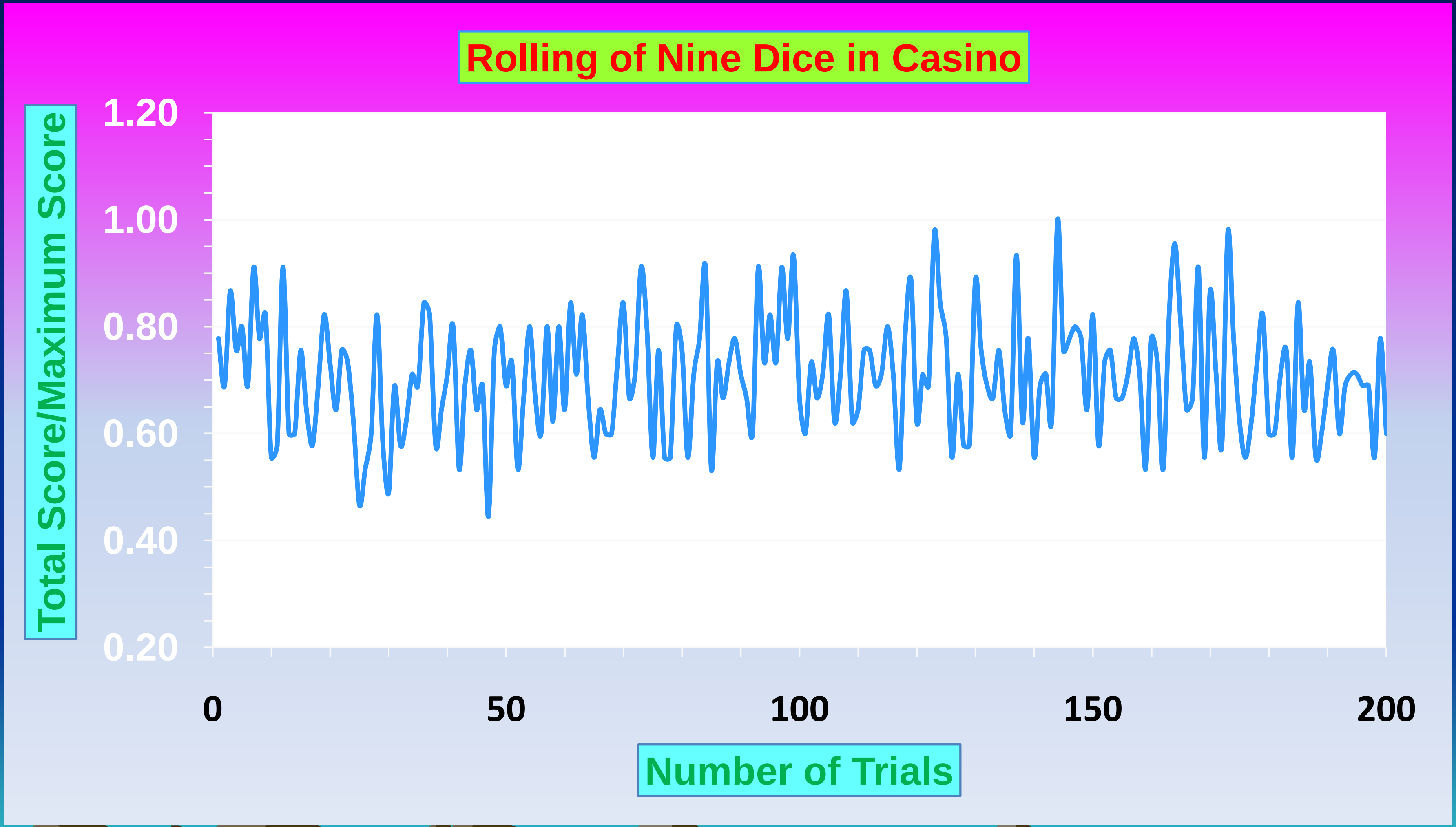


Table 2: Simulated two sets of six NY State Lotto numbers and chances of victory

No # 1	No # 2	No # 3	No # 4	No # 5	No # 6		No # 1	No # 2	No # 3	No # 4	No # 5	No # 6		Win/Loose
25	35	41	27	52	27		26	13	10	16	23	25		You Loose
49	10	15	30	43	26		23	28	25	30	36	37		You Loose
11	14	14	38	46	33		2	21	53	3	2	20		You Loose
37	27	7	23	24	42		47	38	18	24	14	27		You Loose
23	41	45	54	49	15		35	14	23	26	15	20		You Loose
39	38	5	27	20	35		7	21	25	19	24	49		You Loose
42	13	53	35	1	12		9	36	9	28	19	53		You Loose
33	23	29	20	11	7		21	38	28	11	44	8		You Loose
51	20	4	5	55	27		38	8	43	44	41	42		You Loose
53	47	7	21	20	23		42	14	26	46	44	9		You Loose
45	55	42	50	27	35		2	49	6	11	37	12		You Loose
48	52	4	27	26	1		14	36	23	51	33	9		You Loose
8	17	30	12	38	26		25	2	53	50	13	31		You Loose
1	1	10	26	32	30		17	20	14	27	24	15		You Loose
16	10	33	30	34	41		54	19	24	44	45	32		You Loose
31	22	21	12	9	24		19	6	49	29	19	20		You Loose
30	44	10	30	39	2		37	34	49	11	12	7		You Loose
16	41	46	42	7	43		35	50	46	25	1	10		You Loose
35	47	52	49	15	4		18	14	15	17	3	49		You Loose
40	39	11	36	19	40		40	46	38	31	38	41		You Loose
9	49	39	3	25	35		1	54	41	37	39	33		You Loose
55	47	54	52	43	51		46	31	39	25	7	46		You Loose
9	5	22	10	3	10		7	41	14	9	4	55		You Loose
43	11	18	8	13	52		11	47	22	3	46	51		You Loose

Fig. 2: Computer Model Terminology for Mendel's Laws

Code for Dominant gene = 0

Code for Recessive gene = 1

D + DR + RD + R

Dominant = D

Recessive = R

Multiplier = $(2^n - 1)$

n = Generation

$$P_n(D) = K*(2^n - 1)*P_n(R)$$

where n = 1, 2, 3, 4, 5, 6, 7 … etc., and K = constant

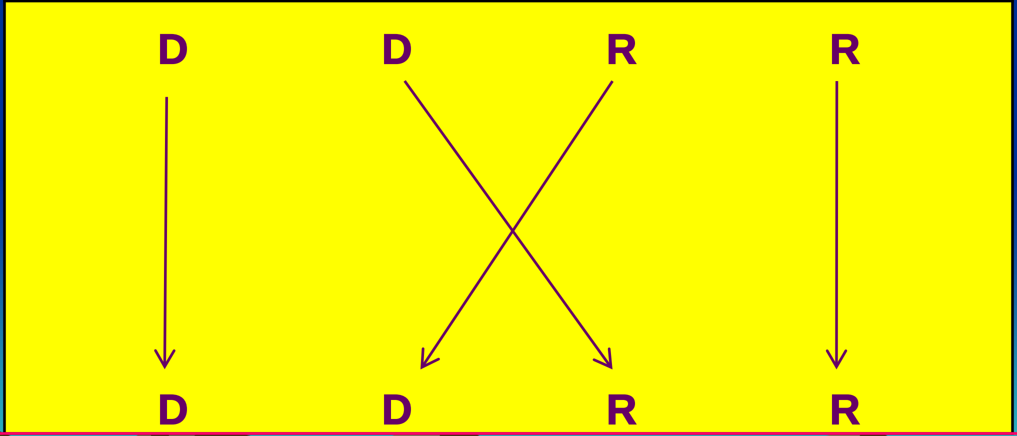
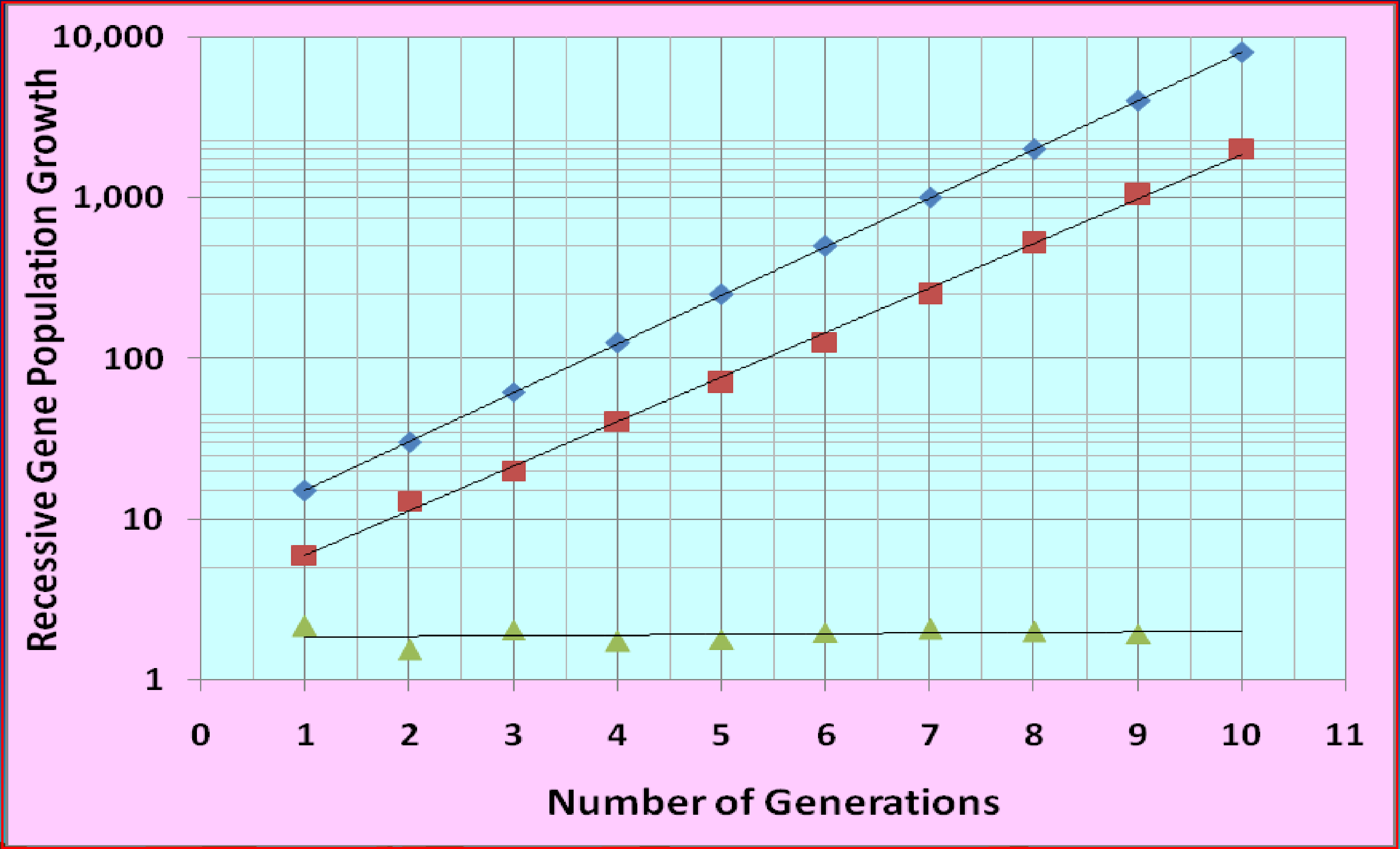


Table 3: Partial Data of Interactive Simulations for Recessive- Gene Progenies

Female Gene	Male Gene	Baby	Recessive/ dominant	Recessive Genes (Variables)			
dominant	dominant	dominant	0	**Generation**	**Trials**	**Babies**	**Ratio**
recessive	recessive	recessive	1	1	15	4	2
recessive	dominant	dominant	0	2	30	9	2
dominant	recessive	dominant	0	3	62	17	2
dominant	recessive	dominant	0	4	125	34	2
recessive	recessive	recessive	1	5	250	63	2
recessive	dominant	dominant	0	6	500	123	2
dominant	recessive	dominant	0	7	1000	245	2
recessive	dominant	dominant	0	8	2000	508	2
recessive	recessive	recessive	1	9	4000	1014	2
recessive	dominant	dominant	0	10	8000	2042	2
recessive	recessive	recessive	1	11	16000	4089	

Fig. 3: Overtime Recessive Genes Population Growth: (i) Diamonds (♦): starting dominant/recessive gene pairs, (ii) Squares (■): population growth due to recessive gene pairs, and (iii) Triangle (▲): ratio of two successive progenies for recessive genes population



Analysis of Recessive Gene Progenies

- Best-fit curve through theoretical data points (♦, diamonds) for sum total of simulations for each starting generation:

$$y = C_1\, e^{0.696x} \qquad (1)$$

  where constant, C_1 = 7.585 and exponent, e_1 = 0.696
- Best-fit curve through simulated data points for recessive gene progenies (■, squares) corresponding to each generation:

$$y = C_2\, e^{0.695x} \qquad (2)$$

  where constant, C_2 = 1.937 and exponent, e_2 = 0.695
- Ratio $C_2/C_1 \approx 25\%$, which proves that only a quarter of total population in a given generation comes from recessive gene contribution, and rest is from the contribution of dominant genes in each generation, which is an interesting result
- Exponent, $e_1 = e_2$, means no growth as expected from experiments

Modeling of Dominant Gene Populations

- Following assumptions of Ref. [4] that growth of dominant gene progeny, $P_n(D)$, in a given generation could be expressed in terms of progeny produced from recessive genes, $P_n(R)$, we wrote an empirical correlation between the two progenies for the first generation with $n = 1$:

$$P_1(D) = \Sigma P(\mathrm{D}) - \Sigma P(\mathrm{R}), \tag{3}$$

where $\Sigma P(\mathrm{D})$ and $\Sigma \mathrm{P(R)}$ is sum total population in the first generation due to the contribution of dominant and recessive genes, respectively

- For second and subsequent ($n \geq 2$) generation, correlation between growth of dominant and recessive gene progenies is expressed as:

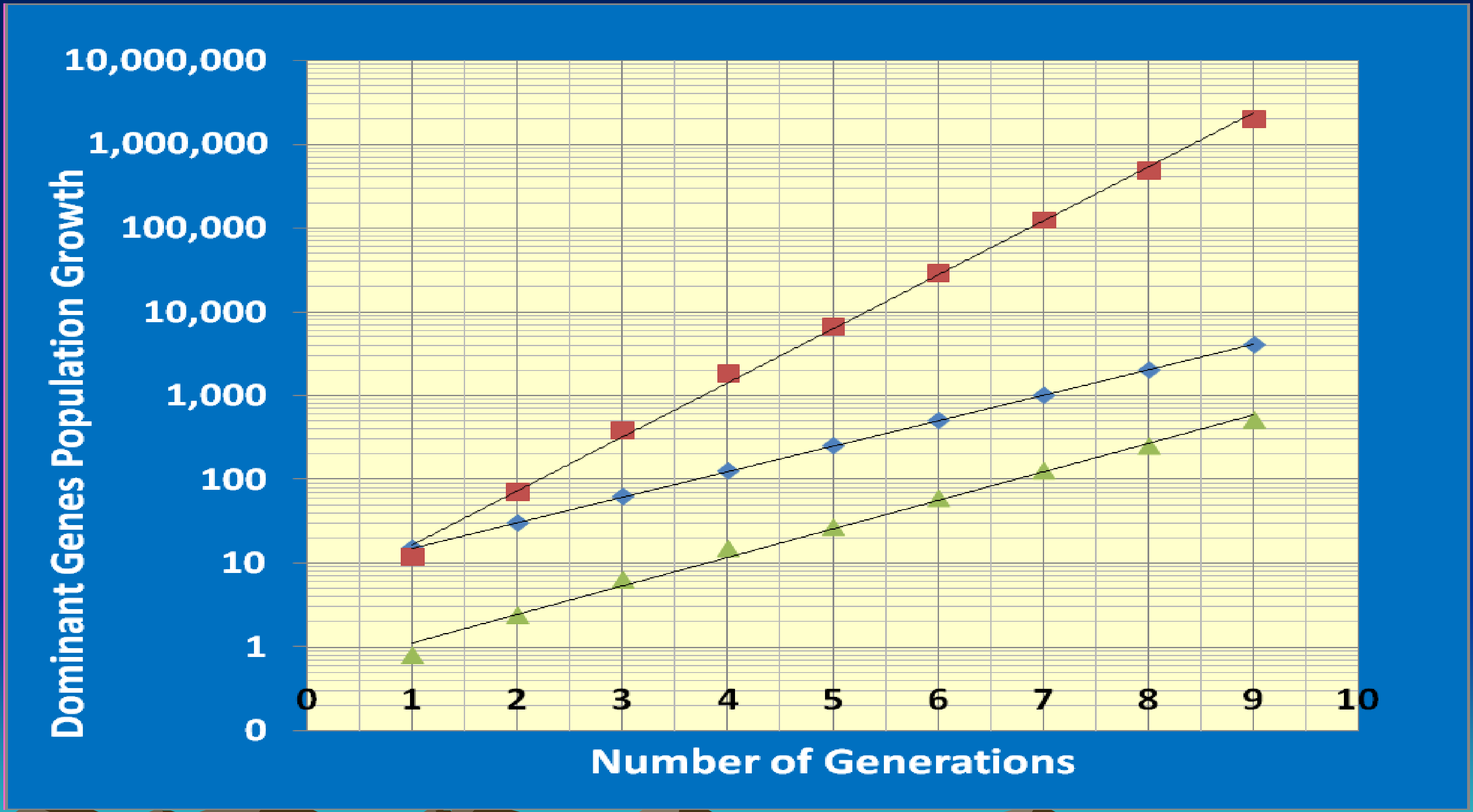
$$P_n(D) = K^*(2^n - 1)^*P_n(R), \tag{4}$$

where $K = 4$ and $n = 2, 3, 4, 5$, …….etc. is a running index for 2^{nd}, 3^{rd}, 4^{th},… etc. generation, respectively. Multiplier factor, $(2^n – 1)$ used in Eq. (4) has odd values, which are given in the 2^{nd} column of next slide

Table 4: Interactive Simulations of Dominant-Gene Progenies

Dominant Genes (Constants)				
Generation	**Multiplier**	**Trials**	**Progeny**	**Ratio**
1	1	15	12	0.8
2	3	30	96	3.2
3	7	62	448	7.2
4	15	125	2040	16.3
5	31	250	8308	33.2
6	63	500	30240	60.5
7	127	1000	130048	130.0
8	255	2000	520200	260.1
9	511	4000	2062396	515.6
10	1023	8000	8335404	1041.9
11	2047	16000	32956700	2059.8

Fig. 4: Dominant Genes Population Growth: (i) Diamonds (♦): starting dominant/recessive gene pairs, (ii) Squares (■): population growth due to dominant genes and (iii) Triangles (▲): ratio of progeny due to dominant genes to the starting pairs of dominant/recessive genes.



Analysis of Recessive Genes Progenies

- First best-fit through starting dominant/recessive gene-pair data points (♦, diamonds) for all generations as given below:

$$y = C_1\, e^{0.696x} \qquad (5)$$

Once again, $C_1 = 7.585$ and exponent value is $e_1 = 0.696$. It is exactly the same as in Eq. (1) since Eq. (1) and Eq. (5) represent same number of starting dominant/recessive gene pairs for each generation

- Second best-fit through dominant gene-pairs data (■) for all generations:

$$y = C_1(D)\, e^{1.482x} \qquad (6)$$

where, $C_1(D) = 3.801$ and $e_1(D) = 1.482$. Comparing constants through their ratio: $C_1/C_1(D) = 7.585/3.801 = 1.996$, i.e., we find that intercept of dominant progeny graph is two times more than the starting dominant recessive gene-pairs plot. Similarly, ratio of exponents, $e_1(D)/e_1 = 1.482/0.696 = 2.129$, indicates that slope of dominant gene progeny population plot is a little more than starting dominant/recessive genes

Bohr's Model of H-atom Assumptions

Three basic assumptions used:

- Electron always stays in a stationary state unless it is excited by an external energy source to jump to a higher energy state
- Electron does not emit radiations in a given stationary state
- Angular momentum of the electron in a stationary state is always quantized

1. To calculate magnitude of the Bohr radius and energy of H-atom in a stationary state, we use the following equations:

$$R_n = \frac{\varepsilon_0 h^2 n^2}{\pi \mu_e Z e^2}$$

$$E_n = -\frac{\mu_e Z^2 e^4}{8\varepsilon_0^2 h^2 n^2}$$

2. Actual formula to compute wavelength of spectral lines for H-atom was derived by Bohr in 1913 using postulates based on quantum theory:

$$\frac{1}{\lambda} = \frac{m_e Z^2 e^4}{8\varepsilon_0^2 h^3 c}\left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right)$$

Table 5: Computed Bohr's Radius, R_n and Energy, E_n

Quantum number, n	Radius, Rn (m)	Energy, En(J)	Energy, En(eV)
1	5.29567E-11	-2.1779E-18	-13.611855
2	2.11827E-10	-5.44474E-19	-3.402964
3	4.7661E-10	-2.41989E-19	-1.512428
4	8.47307E-10	-1.36119E-19	-0.850741
5	1.32392E-09	-8.71159E-20	-0.544474
6	1.90644E-09	-6.04971E-20	-0.378107
7	2.59488E-09	-4.44469E-20	-0.277793
8	3.38923E-09	-3.40296E-20	-0.212685
9	4.28949E-09	-2.68876E-20	-0.168048
10	5.29567E-09	-2.1779E-20	-0.136119
11	6.40776E-09	-1.79991E-20	-0.112495
12	7.62576E-09	-1.51243E-20	-0.094527
13	8.94968E-09	-1.2887E-20	-0.080544
14	1.03795E-08	-1.11117E-20	-0.069448
15	1.19153E-08	-9.67954E-21	-0.060497
16	1.35569E-08	-8.50741E-21	-0.053171
17	1.53045E-08	-7.53598E-21	-0.047100
18	1.7158E-08	-6.7219E-21	-0.042012
19	1.91174E-08	-6.03296E-21	-0.037706
20	2.11827E-08	-5.44474E-21	-0.034030

Table 6: Spectral Series Wavelengths, λ

Energy, E_n (eV)	Lyman (nm)	Balmer (nm)	Paschen (nm)	Brackett (nm)	Pfund (nm)	Humphreys (nm)
-13.611855	121.611	656.700	1876.287	4053.707	7462.505	12376.278
-3.402964	102.609	486.445	1282.618	2626.802	4655.429	7505.148
-1.512428	97.289	434.326	1094.501	2166.890	3741.883	5910.304
-0.850741	95.009	410.438	1005.573	1945.779	3298.161	5130.472
-0.544474	93.814	397.263	955.201	1818.555	3040.280	4674.162
-0.378107	93.109	389.156	923.485	1737.303	2874.015	4378.003
-0.277793	92.656	383.786	902.061	1681.709	2759.246	4172.270
-0.212685	92.349	380.035	886.839	1641.751	2676.080	4022.290
-0.168048	92.130	377.307	875.601	1611.945	2613.574	3908.931
-0.136119	91.968	375.257	867.050	1589.053	2565.236	3820.803
-0.112495	91.846	373.678	860.383	1571.054	2526.986	3750.720
-0.094527	91.751	372.434	855.079	1556.623	2496.139	3693.940
-0.080544	91.676	371.437	850.786	1544.863	2470.863	3647.213
-0.069448	91.616	370.625	847.261	1535.144	2449.868	3608.244
-0.060497	91.566	369.954	844.329	1527.014	2432.224	3575.369
-0.053171	91.525	369.394	841.864	1520.140	2417.242	3547.355
-0.047100	91.491	368.921	839.770	1514.274	2404.404	3523.271
-0.042012	91.462	368.519	837.977	1509.226	2393.316	3502.403
-0.037706	91.437	368.173	836.429	1504.850	2383.668	3484.192
-0.034030	91.416	367.874	835.083	1501.030	2375.219	3468.199
-0.030866	91.397	367.613	833.905	1497.675	2367.776	3454.074

Fig. 5: H-atom Energy Level Diagram

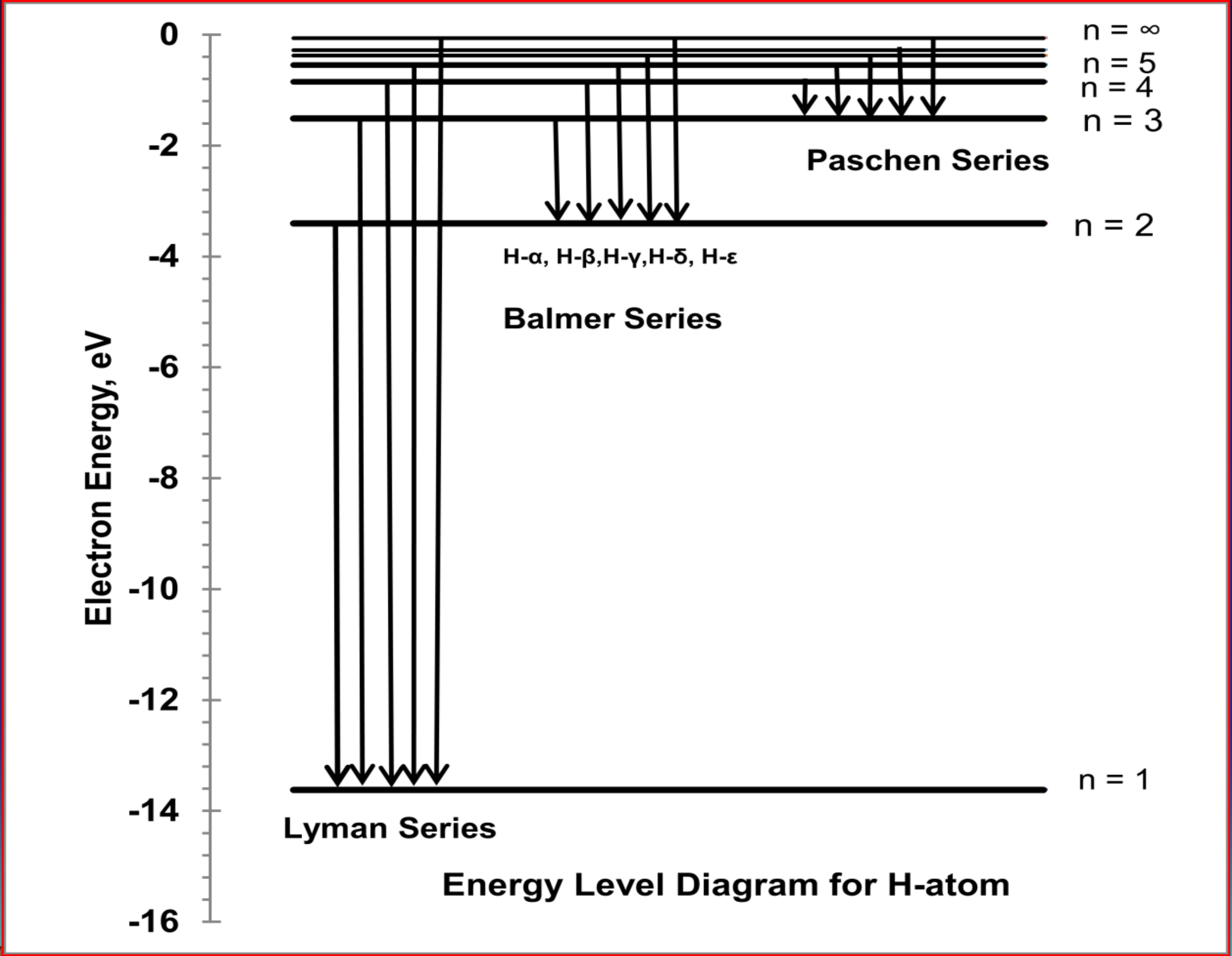


Table 7: Computed and Experimental Spectral Lines of H-atom

	Lyman Series	Balmer Series	Paschen Series	Brackett Series	Pfund Series	Experimental Balmer Series
λ_α	121.611	656.700	1876.287	4053.707	7462.505	656.285 (red)
λ_β	102.609	486.445	1282.618	2626.802	4655.429	486.133 (Blue-green)
λ_γ	97.289	434.326	1094.501	2166.890	3741.883	434.047 (Violet)
λ_δ	95.009	410.438	1005.573	1945.779	3298.161	410.174 (Violet)
λ_ε	93.814	397.363	955.201	1818.555	3040.280	397.007 (Violet)
λ_∞	91.208	364.834	820.876	1459.334	2280.210	-
	$n_f = 2, 3, 4..$ $n_i = 1$	$n_f = 3, 4, 5$ $n_i = 2$	$n_f = 4, 5, 6$ $n_i = 3$	$n_f = 5, 6, 7$ $n_i = 4$	$n_f = 6, 7, 8..$ $n_i = 5$	$n_f = 3, 4, 5..$ $n_i = 2$

Concluding Remarks

In conclusion, we may emphasize that the present talk has quite important implications both in computer science, information systems, natural science as well as in medical science, genetics, bioinformatics etc.

(i) In computer science, information systems and educational technology, researches could visualize the real time application of rolling of nine dice in casino and chances of winning of the NY State Lotto in a virtual laboratory.

(ii) Additionally, medical instructors can present an idea of regular heart-beat of human beings, which has been proven with the help of a plot of normalized total score as a function of the number of trials from the simulations of nine rolling dice.

(iii) Simulation of Mendel's laws of heredity definitely sheds important light to present this important aspect of genetics and bioinformatics in a virtual lab.

(iv) Simulation of quantum theory of H-atom can be assigned as a research project to the graduate students of natural sciences.

References

1. G. Singh and K. Siddiqui, J. Edu. Tech. Sys. 37(4), 405-417 (2008-09).
2. G. Singh, Unpublished work (2012).
3. G. Singh, K. Siddiqui, M. Singh and S. Singh, J. Edu. Tech. Sys., 39(1), 31-46 (2010-2011).
4. G. Singh, J. Edu. Tech. Sys., 40(3), 273-286 (2011-2012).
5. G. Singh, Unpublished work (2012).
6. R. Grauer, K. Mulbery, K. Mast, and M. A. Poasty, The Pearson Custom Program for CIS, *Introduction to Information Systems*, Pearson Learning Solutions Inc. (2011); R. Grauer and J. Scheeren, *Microsoft Office, Excel 2007 (Comprehensive Ed),* Prentice Hall Inc. (2008).