

Ministerul Educației și Cercetării
Universitatea Tehnică din Cluj-Napoca
Facultatea de Știința și Ingineria Materialelor
Catedra de Chimie

Programul:	Cercetare de Excelență
Modul: II	Proiecte de Dezvoltare a Resurselor Umane pentru Cercetare
Tipul proiectului:	Proiecte de cercetare de excelenta pentru tinerii cercetători
Cod proiect:	ET46/2006
Denumirea proiectului:	Distribuția Binomială: Modelare Statistică, Optimizare Numerică, cu Aplicații în Bioinformatică și Biochimie
Etapa:	Finală/2007

Lucrare în extenso

Cuprins

<i>Etape, obiective și activități</i>	<i>– pag. 02</i>
<i>Activități și rezultate</i>	<i>– pag. 03</i>
<i>Etapa 1, Activitatea 1</i>	<i>– pag. 04</i>
<i>Etapa 1, Activitatea 2</i>	<i>– pag. 08</i>
<i>Etapa 2, Activitatea 1</i>	<i>– pag. 12</i>
<i>Etapa 2, Activitatea 2</i>	<i>– pag. 17</i>
<i>Etapa 2, Activitatea 3</i>	<i>– pag. 24</i>
<i>Etapa 3, Activitatea 1</i>	<i>– pag. 36</i>
<i>Etapa 3, Activitatea 2</i>	<i>– pag. 49</i>
<i>Etapa 3, Activitatea 3</i>	<i>– pag. 58</i>
<i>Activitatea 4, Etapele 1,2, și 3</i>	<i>– pag. 75</i>
<i>Publicații</i>	<i>– pag. 78</i>
<i>Concluzii</i>	<i>– pag. 103</i>

Etape, obiective și activități

Pentru etapele anului 2006 obiectivele planificate au fost:

- *Proiectarea intervalelor de încredere pentru cazul celorlalte funcții de o singură variabilă binomială X (Etapa 1);*
- *Optimizarea intervalelor de încredere pentru cazul celorlalte funcții de o singură variabilă binomială X (Etapa 1);*
- *Optimizarea intervalelor de încredere pentru cazul funcțiilor $X(n-Y)/Y/(m-X)$ și $Xn/Y/m$ de două variabile binomiale X și Y (Etapa 2);*
- *Optimizarea intervalelor de încredere pentru cazul funcției $Y/n-X/m$ de două variabile binomiale X și Y (Etapa 2);*
- *Optimizarea intervalelor de încredere pentru cazul funcțiilor $|Y/n - X/m|$ și $1/|Y/n - X/m|$ de două variabile binomiale X și Y (Etapa 3 - finală 2007);*
- *Optimizarea intervalelor de încredere pentru cazul funcției $|1-Xn/Y/m|$ de două variabile binomiale X și Y (Etapa 3 - finală 2007).*

Activitățile prevăzute a se desfășura au fost:

- *Proiectarea și implementarea modului de calcul al intervalelor de încredere pentru $X/(n-X)$ (Etapa 1);*
- *Proiectarea și implementarea modului de calcul al intervalelor de încredere pentru $X/(2X-n)$ (Etapa 1);*
- *Optimizarea intervalelor de încredere pentru $X/(n-X)$ și $X/(2X-n)$ (Etapa 1);*
- *Proiectarea și implementarea modului de calcul al intervalelor de încredere și optimizarea intervalelor pentru $X(n-Y)/Y/(m-X)$ (Etapa 2);*
- *Proiectarea și implementarea modului de calcul al intervalelor de încredere și optimizarea intervalelor pentru $Xn/Y/m$ (Etapa 2);*
- *Proiectarea și implementarea modului de calcul al intervalelor de încredere și optimizarea intervalelor pentru $Y/n-X/m$ (Etapa 2);*
- *Proiectarea și implementarea modului de calcul al intervalelor de încredere și optimizarea intervalelor pentru $|Y/n - X/m|$ (Etapa 3 - finală 2007);*
- *Proiectarea și implementarea modului de calcul al intervalelor de încredere și optimizarea intervalelor pentru $1/|Y/n - X/m|$ (Etapa 3 - finală 2007);*

ET46/2006 – Et. Finală/2007 – Lucrare in extenso

- *Proiectarea și implementarea modului de calcul al intervalelor de încredere și optimizarea intervalelor pentru $|1-Xn/Y/m|$ (Etapa 3 - finală 2007);*
- *Participări la manifestări științifice și dobândirea de competențe complementare (Etapetele 1, 2, și 3).*

Activitățile au fost realizate și obiectivul planificat a fost atins.

Activități și rezultate

Etapa 1, Activitatea 1. Proiectarea și implementarea modului de calcul al intervalelor de încredere pentru $X/(n-X)$ și $\frac{1}{2}$ din Activitatea 3. Optimizarea intervalelor de încredere pentru $X/(n-X)$

► Funcția matematică $X/(n-X)$ este folosită în studiile medicale bazate pe tabela de contingență:

2×2	Caz	Test	Test
Succes	a	c	a+c
Eșec	b	d	b+d
Total	a+b	c+d	a+b+c+d

pentru studiu asupra eșantionului caz (când se folosește substituția $a=X$, $a+b=n$) sau eșantionului test (când se folosește substituția $c=X$, $c+d=n$), denumirea consacrată în limba engleză pentru acest studiu fiind *posttest odds* și respectiv *pretest odds* (prescurtarea fiind identică din motive lesne de înțeles, PTO).

► Notând cu CI_{PTO} intervalul de încredere pentru șansa post (pre) test (folosită în studii de tip diagnostic), și cu CI_L și respectiv CI_U capetele intervalelor de încredere pentru o variabilă ce urmează o distribuție binomială, următoarea formulă a fost folosită pentru conversia intervalelor de încredere ale variabilei distribuite binomial la șansa post (pre) test:

$$CI_{PTO}(X) = \left(\frac{1}{\frac{1}{CI_L(X)} - 1}, \frac{1}{\frac{1}{CI_U(X)} - 1} \right)$$

► În urma evaluării metodelor de calcul al intervalului de încredere pentru o variabilă distribuită binomial (ce însemnând corectat la continuitate):

Notăție	Metodă	Referință
Wald	Wald	Abraham WALD, <i>Contributions to the Theory of Statistical Estimation and Testing Hypothesis</i> , The Annals of Mathematical Statistics, p. 299-326, 1939.
Wald	Wald cc	Brown D.L., Cai T.T., DasGupta A., <i>Interval estimation for a binomial proportion</i> , Statistical Science, 16, p. 101-133, 2001.
AC	Agresti-Coull	Alan AGRESTI, Brent A. COULL, <i>Approximate is better than 'exact' for interval estimation of binomial proportions</i> , The American Statistician, 52, p. 119-126, 1998.

ET46/2006 – Et. Finală/2007 – Lucrare in extenso

ACC1	Agresti-Coull cc	Brown D.L., Cai T.T., DasGupta A., <i>Interval estimation for a binomial proportion</i> , Statistical Science, 16, p. 101-133, 2001.
ACC2	Agresti-Coull cc	Metodă nouă de corecție
Wilson	Wilson	Edwin Bidwell WILSON, <i>Probable Inference, the Law of Succession, and Statistical Inference</i> , Journal of the American Statistical Association, 22, p. 209-212, 1927.
WilsonC	Wilson cc	Robert G. NEWCOMBE, <i>Two-Sided Confidence Intervals for the Single Proportion: Comparison of Seven Methods</i> , Statistics in Medicine, 17, p. 857-872, 1998.
ArcS	ArcSine	J. R. Anderson, L. Bernstein, M. C. Pike, <i>Approximate Confidence Intervals for Probabilities of Survival and Quantiles in Life-Table Analysis</i> , Biometrics, 38(2), p. 407-416, 1982.
ArcSc1	ArcSine cc	Brown D.L., Cai T.T., DasGupta A., <i>Interval estimation for a binomial proportion</i> , Statistical Science, 16, p. 101-133, 2001.
ArcSc2	ArcSine cc	Brown D.L., Cai T.T., DasGupta A., <i>Interval estimation for a binomial proportion</i> , Statistical Science, 16, p. 101-133, 2001.
ArcSc3	ArcSine cc	Metodă nouă de corecție
Logit	Logit	Woolf B., <i>On estimating the relation between blood group and disease</i> , Annals of Human Genetics, 19, p. 251-253, 1955.
LogitC	Logit cc	Gart J. J., <i>Alternative analyses of contingency tables</i> , Journal of Royal Statistical Society, B28, p. 164-179, 1966.
BetaC11	Bayes (Fisher)	Fisher R. A., <i>Statistical Methods for Scientific Inference</i> , Oliver and Boyd, Edinburgh, 1956.
BetaC01	Clopper-Pearson	Clopper C., Pearson S., <i>The use of confidence or fiducial limits illustrated in the case of the binomial</i> , Biometrika, 26, p. 404-413, 1934.
BetaCJ0	Jeffreys	Jeffreys H., <i>Theory of Probability (3rd Ed)</i> , Clarendon Press, Oxford, 1961.
BetaC00	Beta cc	[New correction]
BetaC10	Beta cc	[New correction]
BetaCJ1	Beta cc	[New correction]
BetaCJ2	Beta cc	[New correction]

Optimized	OptB	Sorana D. BOLBOACA, Lorentz JÄNTSCHI, Optimized Confidence Intervals for Binomial Distributed Samples, International Journal of Pure and Applied Mathematics, accepted (August 2, 2007), in press (2007, Volume 40, Issue 3).
-----------	------	---

► S-au selectat următoarele metode pentru optimizarea intervalelor de încredere ale post(pre) test odds (șansa post/pre test):

Metoda	Formula
Wilson	$\frac{X + z^2/2 \pm z(z^2/4 + X(1-X/n))^{1/2}}{n + z^2}$
Logit	$1 - \left(1 + \frac{X}{n-X} \exp \left(\pm z \left(\frac{n}{X(n-X)} \right)^{1/2} \right) \right)^{-1}$
LogitC	$1 - \left(1 + \frac{X+0.5}{n-X+0.5} \exp \left(\pm z \left(\frac{(n+1)(n+2)}{n(X+0.5)(n-X+0.5)} \right)^{1/2} \right) \right)^{-1}$
BetaC11	$\text{BinI}(X, n, a, c_1, c_2) = \text{InvCDFBeta} \left(\frac{a}{2}, X + c_1, n - X + c_2 \right)$ $\text{BinS}(X, n, a, c_1, c_2) = \text{InvCDFBeta} \left(1 - \frac{a}{2}, X + c_2, n - X + c_1 \right)$ $\text{Bin}(X, n, a, c_1, c_2) = (\text{BinI}(X, n, a, c_1, c_2), \text{BinS}(X, n, a, c_1, c_2))$ $\text{BetaC11}(X, n, a) = \text{Bin}(X, n, a, 1, 1)$ $\text{BetaCJ0}(X, n, a) = \text{Bin}(X, n, a, 0.5, 0.5)$ $\text{BetaCJA}(X, n, a) = \text{Bin} \left(X, n, a, 1 - \frac{\sqrt{X(n-X)}}{n}, 1 - \frac{\sqrt{X(n-X)}}{n} \right)$
BetaCJ0	
BetaCJA	

Notă: funcția $\text{invCDFBeta}(\cdot, \cdot, \cdot)$ este funcția inversă a funcției cumulative de probabilitate (funcția de repartiție) a distribuției Beta, a cărei definiție este:

$$\text{CDFBeta}(X, u, v) = \frac{\text{IncBeta}(X, u, v)}{\text{IncBeta}(1, u, v)}$$

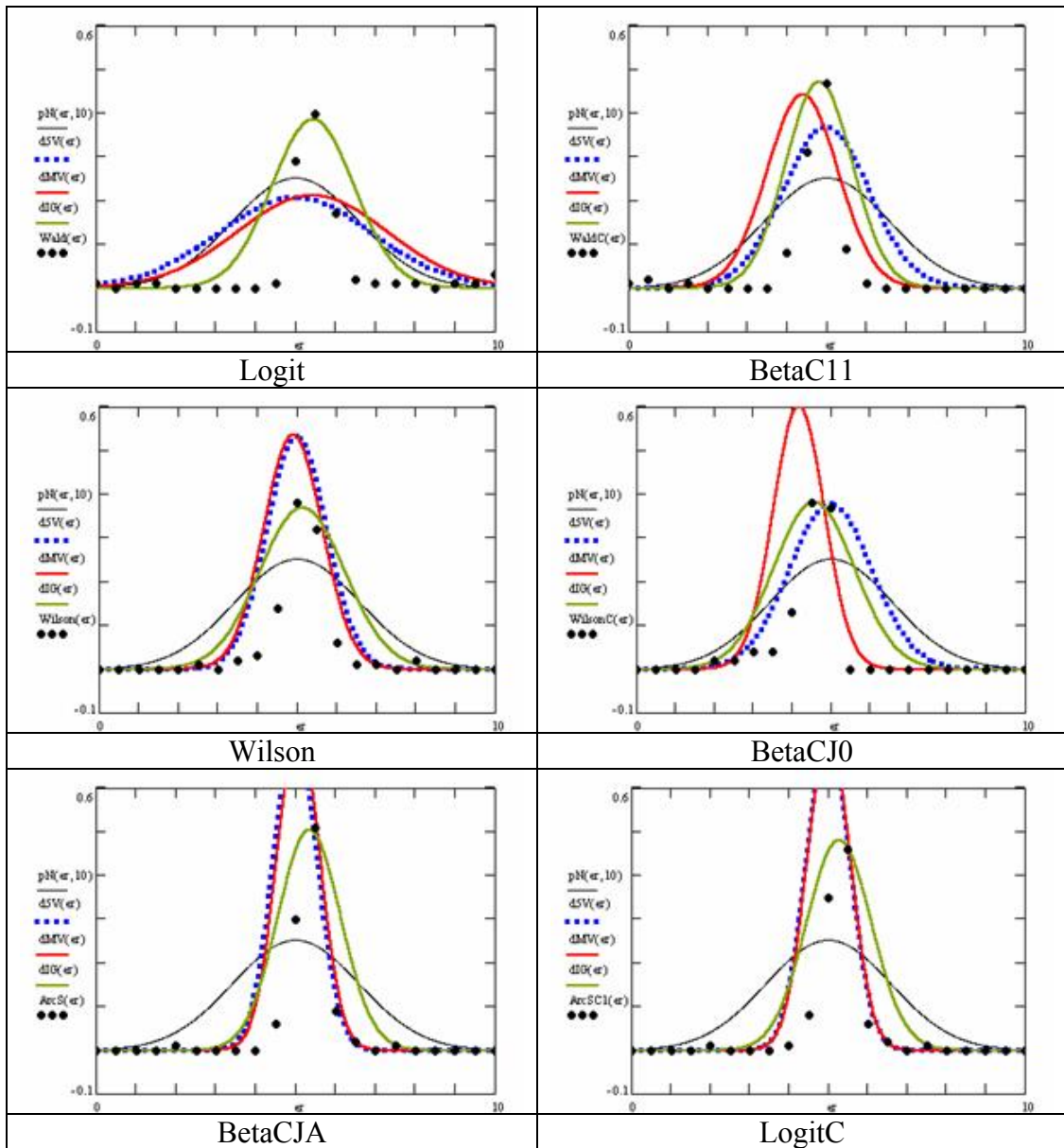
și funcția IncBeta este dată de:

$$\text{IncBeta}(X, u, v) = \int_0^X t^{u-1} (1-t)^{v-1} dt$$

► Evaluarea metodelor s-a realizat pentru volume ale eșantioanelor până la $n=1000$. Următorul tabel redă câteva din aceste rezultate (MErr fiind media erorii experimentale iar StdDev fiind deviația standard a acesteia):

n	Metoda	LogitC	BetaC11	Wilson	BetaCJA	Logit	BetaCJ0
10	MErr	4.22	4.22	4.22	3.84	1.81	3.84
	StdDev	2.45	2.45	2.45	2.17	0.90	2.17
50	MErr	4.94	5.13	5.13	4.88	4.35	5.03
	StdDev	1.13	1.31	1.31	1.12	1.23	1.47
100	MErr	5.20	5.05	5.05	4.96	4.69	4.96
	StdDev	0.81	0.91	0.98	0.80	0.93	1.03
300	MErr	5.03	4.99	4.98	5.00	4.85	5.02
	StdDev	0.73	0.75	0.76	0.65	0.74	0.79
1000	MErr	5.01	5.00	4.98	5.00	4.92	5.01
	StdDev	0.44	0.43	0.45	0.37	0.45	0.46

► Alegerea metodei BetaCJA s-a făcut așadar în consecință. Tabelul următor redă evaluarea comparativă a metodelor selectate, pentru eșantioane de volume aleatorii:



Etapa 1, Activitatea 2. Proiectarea și implementarea modului de calcul al intervalelor de încredere pentru $X/(2X-n)$ și ½ din Activitatea 3. Optimizarea intervalelor de încredere pentru $X/(2X-n)$

► Funcția matematică $X/(2X-n)$ este folosită în studiile medicale bazate pe tabela de contingență:

2×2	Caz	Test	Test
Succes	a	c	a+c
Eșec	b	d	b+d
Total	a+b	c+d	a+b+c+d

pentru studiu asupra eșantionului caz (când se folosește substituția $a=X$, $a+b=n$) sau eșantionului test (când se folosește substituția $c=X$, $c+d=n$), denumirea consacrată în limba engleză pentru acest studiu fiind *posttest probability* și respectiv *pretest probability* (prescurtarea fiind identică din motive lesne de înțeles, PTP).

► Notând cu CI_{PTP} intervalul de încredere pentru șansa post (pre) test (folosită în studii de tip diagnostic), și cu CI_L și respectiv CI_U capetele intervalelor de încredere pentru o variabilă ce urmează o distribuție binomială, următoarea formulă a fost folosită pentru conversia intervalelor de încredere ale variabilei distribuite binomial la șansa post (pre) test:

$$CI_{PTP}(X) = \left(\frac{0.5}{\frac{1}{CI_U(X)} - 1}, \frac{0.5}{\frac{1}{CI_L(X)} - 1} \right)$$

► În urma evaluării metodelor de calcul al intervalului de încredere pentru o variabilă distribuită binomial (cc însemnând corectat la continuitate):

Notație	Metodă	Referință
Wald	Wald	Abraham WALD, <i>Contributions to the Theory of Statistical Estimation and Testing Hypothesis</i> , The Annals of Mathematical Statistics, p. 299-326, 1939.
Wald	Wald cc	Brown D.L., Cai T.T., DasGupta A., <i>Interval estimation for a binomial proportion</i> , Statistical Science, 16, p. 101-133, 2001.
AC	Agresti-Coull	Alan AGRESTI, Brent A. COULL, <i>Approximate is better than 'exact' for interval estimation of binomial proportions</i> , The American Statistician, 52, p. 119-126, 1998.

ET46/2006 – Et. Finală/2007 – Lucrare in extenso

ACC1	Agresti-Coull cc	Brown D.L., Cai T.T., DasGupta A., <i>Interval estimation for a binomial proportion</i> , Statistical Science, 16, p. 101-133, 2001.
ACC2	Agresti-Coull cc	Metodă nouă de corecție
Wilson	Wilson	Edwin Bidwell WILSON, <i>Probable Inference, the Law of Succession, and Statistical Inference</i> , Journal of the American Statistical Association, 22, p. 209-212, 1927.
WilsonC	Wilson cc	Robert G. NEWCOMBE, <i>Two-Sided Confidence Intervals for the Single Proportion: Comparison of Seven Methods</i> , Statistics in Medicine, 17, p. 857-872, 1998.
ArcS	ArcSine	J. R. Anderson, L. Bernstein, M. C. Pike, <i>Approximate Confidence Intervals for Probabilities of Survival and Quantiles in Life-Table Analysis</i> , Biometrics, 38(2), p. 407-416, 1982.
ArcSc1	ArcSine cc	Brown D.L., Cai T.T., DasGupta A., <i>Interval estimation for a binomial proportion</i> , Statistical Science, 16, p. 101-133, 2001.
ArcSc2	ArcSine cc	Brown D.L., Cai T.T., DasGupta A., <i>Interval estimation for a binomial proportion</i> , Statistical Science, 16, p. 101-133, 2001.
ArcSc3	ArcSine cc	Metodă nouă de corecție
Logit	Logit	Woolf B., <i>On estimating the relation between blood group and disease</i> , Annals of Human Genetics, 19, p. 251-253, 1955.
LogitC	Logit cc	Gart J. J., <i>Alternative analyses of contingency tables</i> , Journal of Royal Statistical Society, B28, p. 164-179, 1966.
BetaC11	Bayes (Fisher)	Fisher R. A., <i>Statistical Methods for Scientific Inference</i> , Oliver and Boyd, Edinburgh, 1956.
BetaC01	Clopper-Pearson	Clopper C., Pearson S., <i>The use of confidence or fiducial limits illustrated in the case of the binomial</i> , Biometrika, 26, p. 404-413, 1934.
BetaCJ0	Jeffreys	Jeffreys H., <i>Theory of Probability (3rd Ed)</i> , Clarendon Press, Oxford, 1961.
BetaC00	Beta cc	[New correction]
BetaC10	Beta cc	[New correction]
BetaCJ1	Beta cc	[New correction]
BetaCJ2	Beta cc	[New correction]

Optimized	OptB	Sorana D. BOLBOACA, Lorentz JÄNTSCHI, Optimized Confidence Intervals for Binomial Distributed Samples, International Journal of Pure and Applied Mathematics, accepted (August 2, 2007), in press (2007, Volume 40, Issue 3).
-----------	------	---

► S-au selectat următoarele metode pentru optimizarea intervalelor de încredere ale post(pre) test odds (șansa post/pre test):

Metoda	Formula
Wilson	$\frac{X + z^2/2 \pm z(z^2/4 + X(1-X/n))^{1/2}}{n + z^2}$
Logit	$1 - \left(1 + \frac{X}{n-X} \exp \left(\pm z \left(\frac{n}{X(n-X)} \right)^{1/2} \right) \right)^{-1}$
LogitC	$1 - \left(1 + \frac{X+0.5}{n-X+0.5} \exp \left(\pm z \left(\frac{(n+1)(n+2)}{n(X+0.5)(n-X+0.5)} \right)^{1/2} \right) \right)^{-1}$
BetaC11	$\text{BinI}(X, n, a, c_1, c_2) = \text{InvCDFBeta} \left(\frac{a}{2}, X + c_1, n - X + c_2 \right)$ $\text{BinS}(X, n, a, c_1, c_2) = \text{InvCDFBeta} \left(1 - \frac{a}{2}, X + c_2, n - X + c_1 \right)$ $\text{Bin}(X, n, a, c_1, c_2) = (\text{BinI}(X, n, a, c_1, c_2), \text{BinS}(X, n, a, c_1, c_2))$ $\text{BetaC11}(X, n, a) = \text{Bin}(X, n, a, 1, 1)$ $\text{BetaCJ0}(X, n, a) = \text{Bin}(X, n, a, 0.5, 0.5)$ $\text{BetaCJA}(X, n, a) = \text{Bin} \left(X, n, a, 1 - \frac{\sqrt{X(n-X)}}{n}, 1 - \frac{\sqrt{X(n-X)}}{n} \right)$
BetaCJ0	
BetaCJA	

Notă: funcția $\text{invCDFBeta}(\cdot, \cdot, \cdot)$ este funcția inversă a funcției cumulative de probabilitate (funcția de repartiție) a distribuției Beta, a cărei definiție este:

$$\text{CDFBeta}(X, u, v) = \frac{\text{IncBeta}(X, u, v)}{\text{IncBeta}(1, u, v)}$$

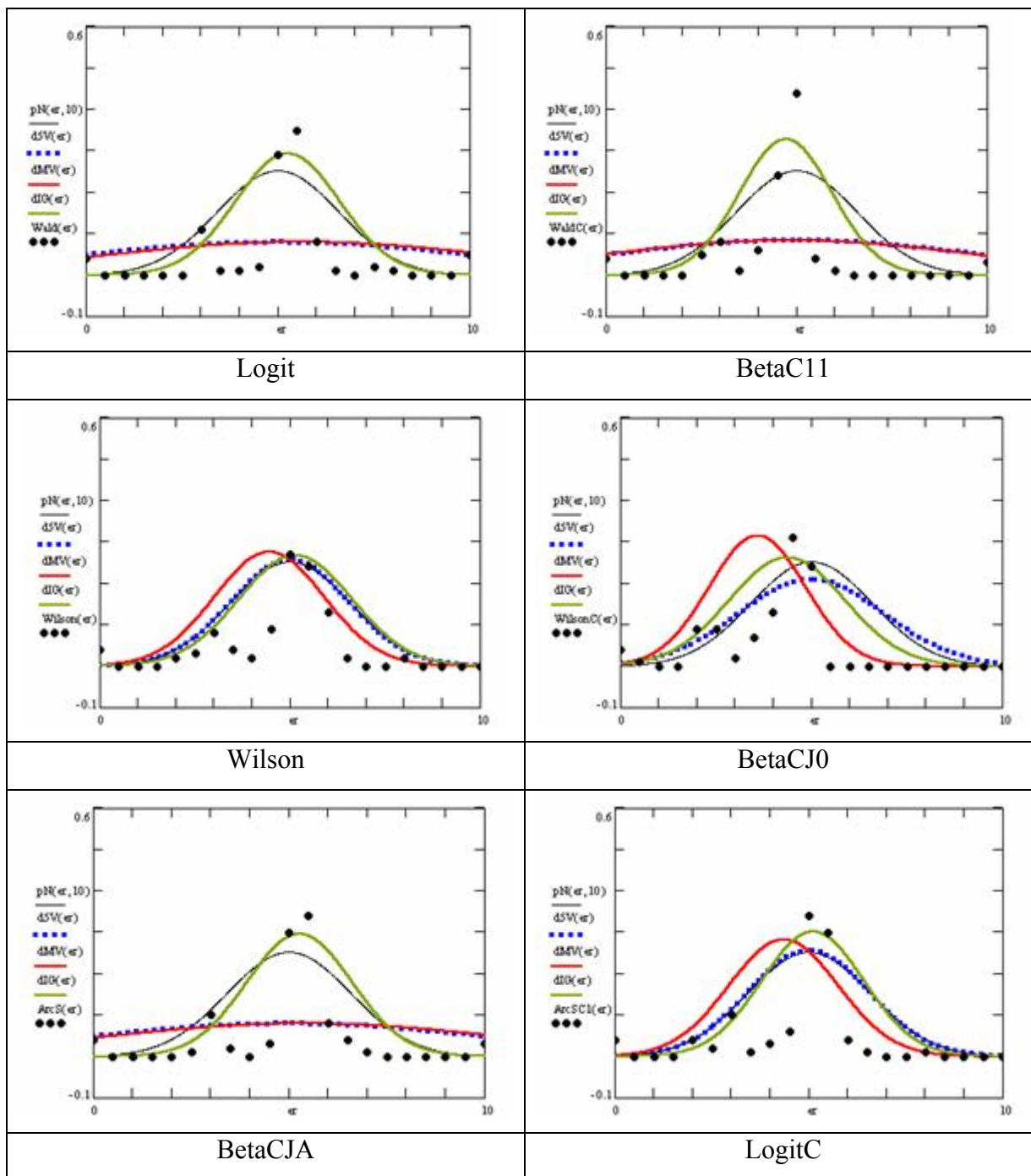
și funcția IncBeta este dată de:

$$\text{IncBeta}(X, u, v) = \int_0^X t^{u-1} (1-t)^{v-1} dt$$

► Evaluarea metodelor s-a realizat pentru volume ale eșantioanelor până la $n=1000$. Următorul tabel redă câteva din aceste rezultate (media erorii experimentale urmată în paranteză de deviația standard a acesteia):

n	LogitC	BetaC11	Wilson	BetaCJA	Logit	BetaCJ0
20	2.15 (2.19)	2.37 (2.10)	2.37 (2.10)	2.03 (1.13)	1.52 (1.11)	2.58 (1.24)
50	3.25 (2.19)	3.44 (1.79)	3.44 (1.79)	3.19 (1.44)	2.77 (1.37)	3.43 (1.76)
100	3.95 (1.88)	3.80 (1.83)	3.80 (1.84)	3.72 (1.63)	3.44 (1.59)	3.76 (1.80)
300	4.27 (1.64)	4.18 (1.54)	4.17 (1.54)	4.19 (1.50)	4.06 (1.49)	4.22 (1.54)
1000	4.53 (1.37)	4.50 (1.34)	4.48 (1.34)	4.49 (1.32)	4.43 (1.33)	4.51 (1.35)

► Alegerea metodei BetaCJA s-a făcut așadar în consecință. Tabelul următor redă evaluarea comparativă a metodelor selectate, pentru eșantioane de volume aleatorii:



Etapa 2, Activitatea 1. Proiectarea și implementarea modului de calcul al intervalelor de încredere și optimizarea intervalelor pentru $X(n-Y)/Y/(m-X)$

► Pornind de la expresia de definiție a funcției de probabilitate a distribuției normale standard:

$$P(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$

► Considerând α valoarea nivelului de semnificație (α este o probabilitate $0 \leq \alpha \leq 1$) a erorii impuse (uzual valoarea α se alege 0.05=5%; alte valori folosite: 0.1=10%, 0.01=1%), valoarea z se definește ca acel număr care verifică ecuația integrală:

$$\frac{\alpha}{2} = \int_z^{\infty} P(x)dx$$

► Următorul tabel redă valorile lui z pentru câteva valori uzuale ale lui α :

α	z
0.5	0.6745
0.1	1.645
0.05	1.96
0.01	2.576

► Funcția matematică $X(n-Y)/Y/(m-X)$ este folosită în studiile medicale bazate pe tabela de contingență:

2×2	Caz	Test	Test
Succes	a	b	a+b
Eșec	c	d	c+d
Total	a+c	b+d	a+b+c+d

pentru studiu asupra relației între eșantionul caz și eșantionul test (când se folosește substituția $a=X$, $a+c=m$, $b=Y$, $b+d=n$), denumirea consacrată în limba engleză pentru acest studiu fiind *odds ratio* (rata șansei, prescurtarea fiind OR).

► Notând cu CI_{OR} intervalul de încredere pentru rata șansei (folosită în studii de tip factor de risc), și cu CI_L și respectiv CI_U capetele intervalelor de încredere pentru o variabilă ce urmează o distribuție binomială, următoarele formule a fost folosite pentru conversia intervalelor de încredere ale variabilei distribuite binomial la rata șansei (α a fost notat cu a):

Metodă	Formulă
ORWald(X,m,Y,n,z)	$\frac{X \cdot (n - Y)}{Y \cdot (m - X)} \cdot \exp\left(\pm z \sqrt{\frac{1}{X} + \frac{1}{m - X} + \frac{1}{Y} + \frac{1}{n - Y}}\right)$ $\begin{cases} (0, \infty) & , Y + m - X = 0 \\ (0, a^{1/m+1/n}) & , X + n - Y = 0 \\ (0, Y \cdot a^{1/m}) & , X = 0, 0 < n - Y \\ (0, X \cdot a^{1/n}) & , 0 = n - Y, 0 < X \end{cases}$
ORAC(X,m,Y,n,z)	$\text{ORWald}\left(X + \frac{z^2}{4\sqrt{2}}, m + \frac{z^2}{2\sqrt{2}}, Y + \frac{z^2}{4\sqrt{2}}, n + \frac{z^2}{2\sqrt{2}}\right)$
ORBinomial(X,m,Y,n,a)	$\text{OR0}(\text{BetaCJAL}, \text{BetaCJAU}, X, m, Y, n, \sqrt{a/2}, \sqrt{a/2})$ <p>sau $(0, \infty)$ pentru $Y + m - X = 0$</p>
ORBinomialC(X,m,Y,n,a)	$\text{OR0}\left(\text{BetaCJAL}, \text{BetaCJAU}, X, m, Y, n, \frac{\sqrt{a/2}}{1 - \frac{1}{8 \ln(m)}}, \frac{\sqrt{a/2}}{1 - \frac{1}{8 \ln(n)}}\right)$ <p>sau $(0, \infty)$ pentru $Y + m - X = 0$</p>

► Notă: Expresile pentru $CI_{OR} = (CI_{ORL}, CI_{ORU})$ ale formulelor R2Binomial și R2BinomialC folosesc funcția OR0, definită astfel:

$$\text{OR0}(\text{FL}, \text{FU}, X, m, Y, n, a_1, a_2) =$$

$$= \left(\frac{1}{\frac{1}{\text{FL}(X, m, a_1)} - 1} \cdot \frac{1}{\frac{1}{\text{FL}(n - Y, Y, a_2)} - 1}, \frac{1}{\frac{1}{\text{FU}(X, m, a_1)} - 1} \cdot \frac{1}{\frac{1}{\text{FU}(n - Y, Y, a_2)} - 1} \right)$$

iar BetaCJAL și BetaCJAU sunt limitele inferioară și respectiv superioară ale intervalului de încredere calculat cu metoda BetaCJA:

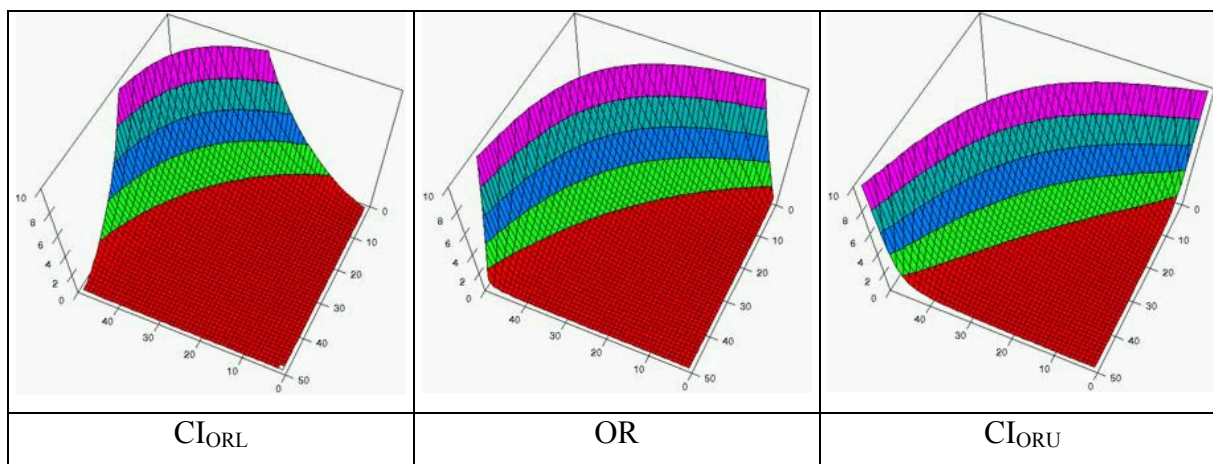
$$(\text{BetaCJAL}(X, n, a), \text{BetaCJAU}(X, n, a)) = \text{BetaCJA}(X, n, a)$$

$$\text{BetaCJA}(X, n, a) = \text{Bin}\left(X, n, a, 1 - \frac{\sqrt{X(n - X)}}{n}, 1 - \frac{\sqrt{X(n - X)}}{n}\right)$$

$$\text{Bin}(X, n, a, c_1, c_2) = (\text{BinI}(X, n, a, c_1, c_2), \text{BinS}(X, n, a, c_1, c_2))$$

și formulele rezultă din înmulțirea (așa cum rezultă din formulele din tabel) ale valorilor inferioare (pentru limita inferioară, funcția BinI) și superioare (pentru limita superioară, funcția BinS) ale intervalului de încredere produs de funcția Binomial.

► Următoarele grafice redau (pe axa Z) valoarea ratei șansei încadrate de valorile intervalului de încredere (graficele alăturate stânga - pentru limita inferioară și dreapta - pentru limita superioară) pentru volume ale eșantioanelor egale ($m = n = 50$), în funcție de valorile variabilelor ce urmează distribuția binomială (X și $Y - 0 \leq X, Y \leq 50$), pentru formulele CI_{ORL} și CI_{ORU} calculate cu ajutorul funcției R2BinomialC:



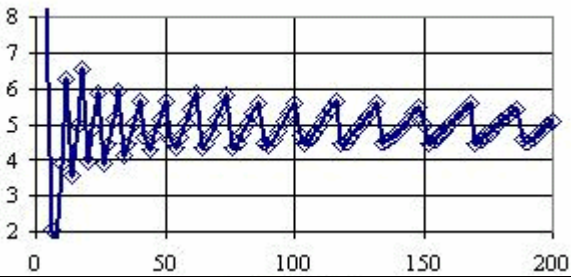
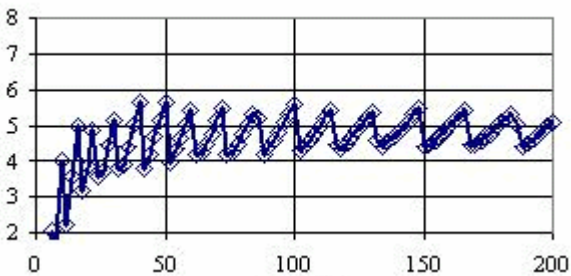
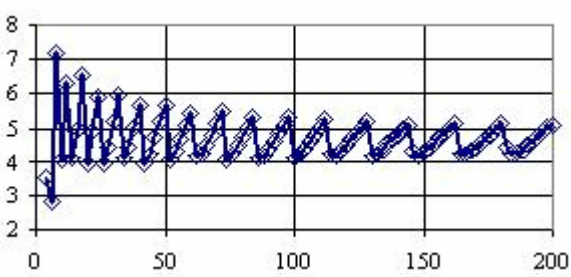
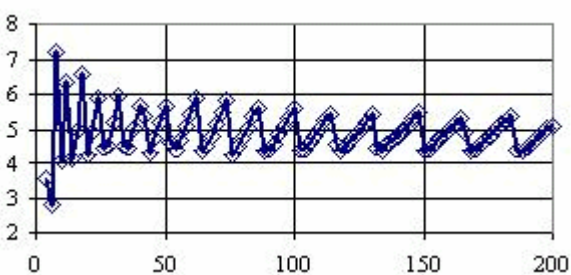
► Evaluarea metodelor s-a realizat pentru volume ale eșantioanelor până la $m=100$ și $n=100$. Următorul tabel redă câteva din aceste rezultate (media erorii experimentale urmată în paranteză de deviația standard a acesteia, pentru volume de eșantion egale $m = n$, pentru un nivel de semnificație impus $\alpha=5\%$):

n	ORWald	ORAC	ORBinomial	ORBinomialC
5	5.0 (6.1)	1.7 (1.5)	2.1 (1.2)	2.3 (1.4)
10	2.9 (2.0)	3.1 (1.6)	2.6 (0.8)	2.9 (1.0)
20	3.3 (1.2)	3.9 (1.4)	3.5 (1.0)	3.9 (1.1)
40	3.9 (1.0)	4.5 (1.2)	4.3 (0.8)	4.5 (0.6)

► În tabelul de mai sus se observă că metodele ORAC, ORBinomial și ORBinomialC sunt semnificativ mai bune decât metoda ORWald și în același timp prezintă diferențe nesemnificativ statistic una de cealaltă. Metoda ORBinomialC îmbunătățește metoda ORAC în ceea ce privește apropierea erorii experimentale de nivelul de semnificație impus de 5% (pentru $m=n=5$ și $m=n=10$) și în ceea ce privește abaterea erorii experimentale de la valoarea sa medie (pentru $m=n=20$ și $m=n=40$). Metoda ORBinomialC îmbunătățește metoda R2Binomial în ceea ce privește apropierea erorii experimentale de nivelul de semnificație impus de 5% (pentru $m=n=5$, $m=n=10$, $m=n=20$ și $m=n=40$) cumulând în același timp o abatere față de valoarea medie de maxim 0.2 puncte procentuale (la $m=n=5$ și $m=n=10$), de 0.1 puncte procentuale (la $m=n=20$) și cu creșterea volumului eșantionului îmbunătățind

metoda ORBinomial în ceea ce privește abaterea erorii experimentale de la valoarea sa medie (cu 0.2 puncte procentuale pentru $m=n=40$).

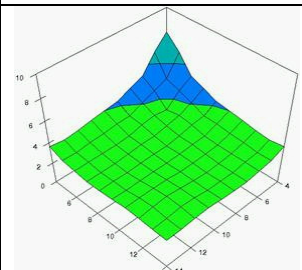
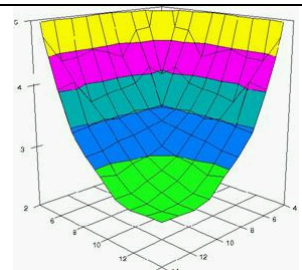
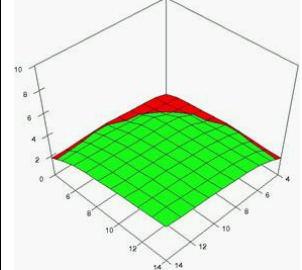
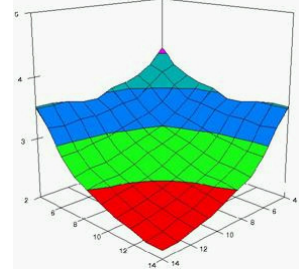
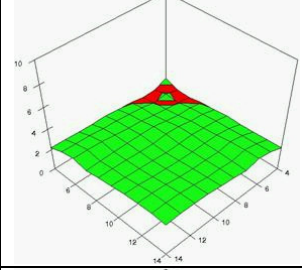
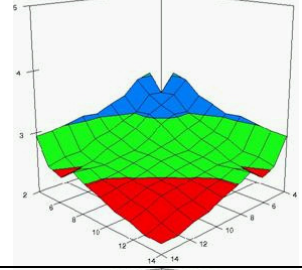
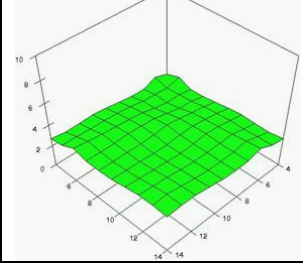
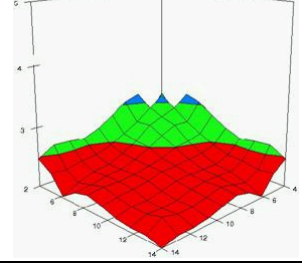
► Studiul punctului central ($m=n$, $X=Y$) ne oferă informații importante despre abilitatea predictivă a metodelor ORWald, ORAC, ORBinomial și ORBinomialC. Tabelul următor prezintă aceste rezultate (eroarea experimentală) pentru $4 \leq m=n \leq 200$:

	<p>ORWald</p> <p>Media erorii experimentale: 4.91 Abaterea standard: 0.76</p>
	<p>ORAC</p> <p>Media erorii experimentale: 4.60 Abaterea standard: 0.81</p>
	<p>ORBinomial</p> <p>Media erorii experimentale: 4.69 Abaterea standard: 0.62</p>
	<p>ORBinomialC</p> <p>Media erorii experimentale: 4.89 Abaterea standard: 0.60</p>

► Din nou metodele ORBinomial și ORBinomialC se dovedesc îmbunătățiri (atât ORBinomialC față de ORBinomial, cât și amândouă față de ORWald și ORAC). Argumentație: aparent ORWald este cea mai bună metodă când criteriul de comparație este media erorii experimentale, însă în realitate nu este așa, valoarea mare (și în același timp apropiată de valoarea impusă a erorii de 5%) este obținută pe seama depășirii (grosolane chiar în domeniul $m=n=4..25$) a erorii impuse, fapt care este reflectat și în abaterea standard (0.76).

ET46/2006 – Et. Finală/2007 – Lucrare in extenso

Dacă metoda ORWald depășește grosolan eroarea impusă, ORAC se situează la polul opus, variind chiar mai pronunțat decât ORWald (ORAC are o abatere de 0.81, ORWald are o abatere de 0.76). Metodele ORBinomial și ORBinomialC dovedesc o stabilitate mult mai mare (abateri de 0.62 și respectiv 0.60) cumulate cu o medie a erorii experimentale situate între ORWald și ORAC (metoda ORBinomial) și respectiv cu o diferență față de ORWald de ordinul abaterii față de medie (metoda ORBinomialC). Dacă aceste argumente nu sunt suficiente, atunci reprezentarea variației erorii experimentale (medii) pentru fiecare metodă pentru domeniul $4 \leq m=n \leq 14$ din tabelul anterior, și respectiv a deviației față de nivelul impus al erorii (5%, tabelul următor) sunt ilustrative.

Valori medii	Statistică descriptivă	Abateri medii
	<p>ORWald</p> <p>Media erorii: 3.33 Media abaterii: 3.90</p>	
	<p>ORAC</p> <p>Media erorii: 2.64 Media abaterii: 2.92</p>	
	<p>ORBinomial</p> <p>Media erorii: 2.57 Media abaterii: 2.70</p>	
	<p>ORBinomialC</p> <p>Media erorii: 2.95 Media abaterii: 2.41</p>	

Etapa 2, Activitatea 2. Proiectarea și implementarea modului de calcul al intervalelor de încredere și optimizarea intervalelor pentru $X_n/Y/m$

► Pornind de la expresia de definiție a funcției de probabilitate a distribuției normale standard:

$$P(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$

► Considerând α valoarea nivelului de semnificație (α este o probabilitate $0 \leq \alpha \leq 1$) a erorii impuse (uzual valoarea α se alege $0.05=5\%$; alte valori folosite: $0.1=10\%$, $0.01=1\%$), valoarea z se definește ca acel număr care verifică ecuația integrală:

$$\frac{\alpha}{2} = \int_z^{\infty} P(x)dx$$

► Următorul tabel redă valorile lui z pentru câteva valori uzuale ale lui α :

α	z
0.5	0.6745
0.1	1.645
0.05	1.96
0.01	2.576

► Funcția matematică $X_n/Y/m$ este folosită în studiile medicale bazate pe tabela de contingență:

2×2	Caz	Test	Test
Succes	a	b	a+b
Eșec	c	d	c+d
Total	a+c	b+d	a+b+c+d

pentru studiu asupra relației între eșantionul caz și eșantionul test, denumirea consacrată în limba engleză pentru acest studiu fiind *relative risk* (riscul relativ) în cazul studiilor de factor de risc și *likelihood ratio* (*positive and negative* - rată de succes sau rată de eșec) în cazul studiilor de tip diagnostic.

► Notând cu CI_{LR} intervalul de încredere pentru riscul relativ (folosită în studii de tip factor de risc), și respectiv rata de succes/eșec (folosită în studii de tip diagnostic), următoarele relații sunt relații de transformare de la formulele de definiție ale acestor parametrii medicali la funcția matematică $X_n/Y/m$:

Relația de transformare	Substituții
$LR+ = LR_+(a, b, c, d) = \frac{\frac{a}{a+c}}{1 - \frac{d}{b+d}} = \frac{\frac{X}{m}}{1 - \frac{n-Y}{n}} = \frac{X}{m} \cdot \frac{n}{Y} = LR^+(X, m, Y, n)$	$a = X$ $b = Y$ $c = m-X$ $d = n-Y$
$LR- = LR_-(a, b, c, d) = \frac{1 - \frac{a}{a+c}}{\frac{d}{b+d}} = \frac{1 - \frac{m-X}{m}}{\frac{Y}{n}} = \frac{X}{m} \cdot \frac{n}{Y} = LR^-(X, m, Y, n)$	$a = m-X$ $b = n-Y$ $c = X$ $d = Y$
$RR = RR_*(a, b, c, d) = \frac{\frac{a}{a+b}}{\frac{c}{c+d}} = \frac{\frac{m}{m}}{\frac{Y}{n}} = \frac{X}{m} \cdot \frac{n}{Y} = RR^*(X, m, Y, n)$	$a = X$ $b = m-X$ $c = Y$ $d = n-Y$

► Astfel, din punct de vedere matematic, acești trei parametri medicali se exprimă prin intermediul aceleiași funcții matematice, pe care în continuare o vom numi LR:

$$LR = LR(X, m, Y, n) = \frac{X}{m} \cdot \frac{n}{Y}$$

► Următoarele expresii au fost obținute pentru exprimarea intervalului de încredere al funcției LR (prima formulă este formula de bază, următoarele 2 sunt formule optimizate, următoarele 5 sunt soluții ale optimizării - doar acestea sunt tratate mai jos):

Nr	Metodă	Expresie de calcul
0	LRWald(X,m,Y,n,z)	$LR \cdot \exp\left(\pm z \sqrt{\frac{m-X}{X \cdot m} + \frac{n-Y}{Y \cdot n}}\right)$
0	ACType2(X,m,Y,n,c1,c2)	LRWald(X+c1,m+2c1,Y+c2,n+2c2,z)
0	ACType1(X,m,Y,n,c)	LRWald(X+c,m+2c,Y+c,n+2c,z)
1	LRWaldz(X,m,Y,n)	LRWald(X,m,Y,n,z)
2	LRAC(X,m,Y,n)	$ACType2\left(X, m, Y, n, \frac{1}{2\sqrt{m}}, \frac{1}{2\sqrt{n}}\right)$
3	LRAC0(X,m,Y,n)	$ACType1\left(X, m, Y, n, \sqrt{\frac{X}{m} \cdot \frac{Y}{n} \cdot \frac{1}{4}}\right)$
4	LRAC1(X,m,Y,n)	$ACType1\left(X, m, Y, n, \sqrt{\frac{X+1}{m} \cdot \frac{Y+1}{n} \cdot \frac{1}{4}}\right)$
5	LRAC2(X,m,Y,n)	$ACType1\left(X, m, Y, n, \sqrt{\frac{X+2}{m} \cdot \frac{Y+2}{n} \cdot \frac{1}{4}}\right)$

► Următoarele metode de evaluare au fost folosite pentru compararea rezultatelor metodelor de calcul al intervalului de încredere pentru funcția LR:

Nume	Metoda	Formula
AvErr	Av(Err)	$\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} \text{Err}(X, Y, m, n)}{(m-1)(n-1)}$
StdDev	StdDev(Err)	$\left(\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} (\text{Err}(X, Y, m, n) - \text{AvErr})^2}{(m-1)(n-1) - 1} \right)^{1/2}$
AvAD	AvAD(Err)	$\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} \text{Err}(X, Y, m, n) - \text{AvErr} }{(m-1)(n-1) - 1}$
AvADI	AvADI(Err)	$\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} \text{Err}(X, Y, m, n) - 100 \cdot \alpha }{(m-1)(n-1)}$
DevI	DevI(Err)	$\left(\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} (\text{Err}(X, Y, m, n) - 100 \cdot \alpha)^2}{(m-1)(n-1)} \right)^{1/2}$

► Funcția $\text{Err}(\cdot, \cdot, \cdot, \cdot)$ cumulează erori experimentale folosind distribuția binomială univariată pentru generarea distribuției binomiale bivariate:

$$d\text{Bin}(m, X, \Xi) := \frac{m!}{\Xi!(m-\Xi)!} \cdot \left(\frac{X}{m}\right)^\Xi \cdot \left(1 - \frac{X}{m}\right)^{m-\Xi}$$

$$d\text{Bin}(n, Y, \Psi) := \frac{n!}{\Psi!(n-\Psi)!} \cdot \left(\frac{Y}{n}\right)^\Psi \cdot \left(1 - \frac{Y}{n}\right)^{n-\Psi}$$

Notând capetele intervalului de încredere pentru funcția LR prin (LRL, LRU) - și aici nota bene LRL și LRU sunt funcții date de una din metodele de calcul al intervalului de încredere (1-5 în tabelul de mai sus), expresia pentru $\text{Err}(X, m, Y, n)$ este:

$$\frac{\sum_{\substack{\text{LRL}(\Xi, \Psi, m, n) \\ > \text{LRL}(X, Y, m, n)}} d\text{Bin}(m, X, \Xi) \cdot d\text{Bin}(n, Y, \Psi) + \sum_{\substack{\text{LRU}(\Xi, \Psi, m, n) \\ < \text{LRU}(X, Y, m, n)}} d\text{Bin}(m, X, \Xi) \cdot d\text{Bin}(n, Y, \Psi)}{\sum_{\Xi=1}^{m-1} \sum_{\Psi=1}^{n-1} d\text{Bin}(m, X, \Xi) \cdot d\text{Bin}(n, Y, \Psi)}$$

► Următoarele tabele sumarizează performanțele metodelor obținute prin prisma celor 5 criterii de evaluare pentru volume de eşantioane reprezentative:

Valori medii pentru $14 \leq m,n \leq 34$					
Metoda	AvErr	StdDev	AvAD	AVADI	DevI
LRWaldz	4.19	1.411	0.882	1.192	1.634
LRAC	4.22	1.262	0.874	1.132	1.485
LRAC0	4.16	1.222	0.864	1.141	1.485
LRAC1	4.17	1.226	0.870	1.140	1.484
LRAC2	4.18	1.229	0.876	1.137	1.481
“Best of”	5	0	0	0	0

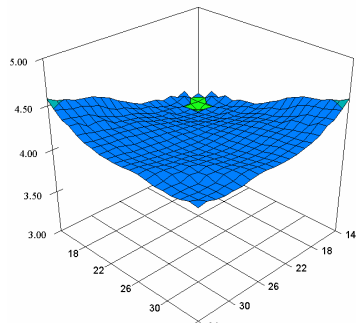
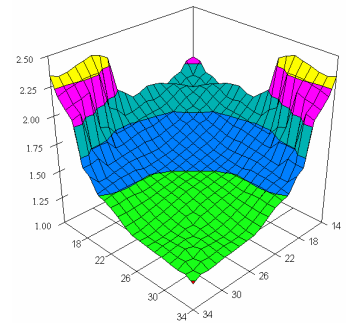
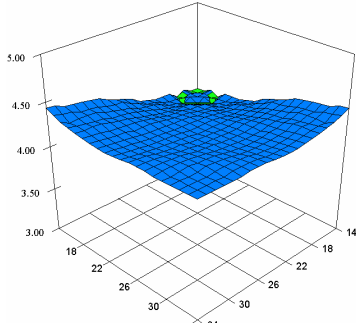
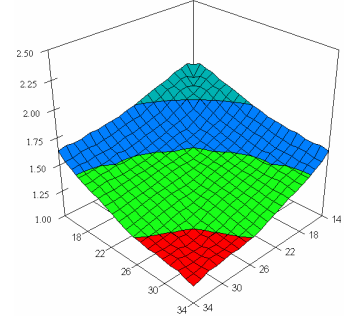
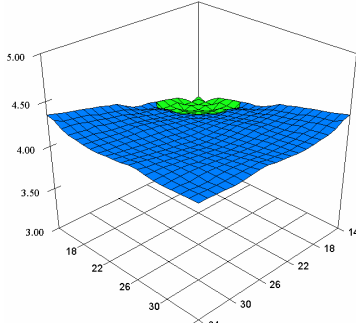
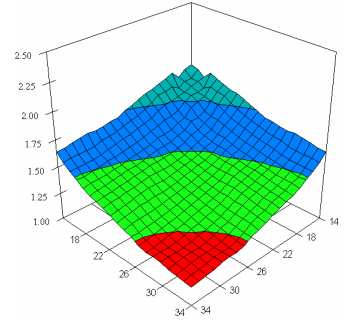
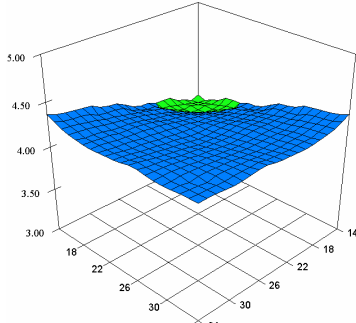
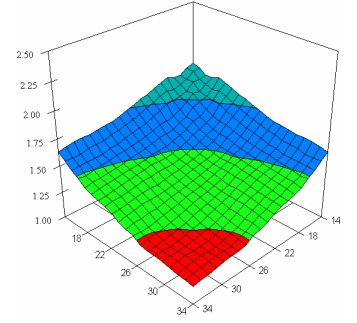
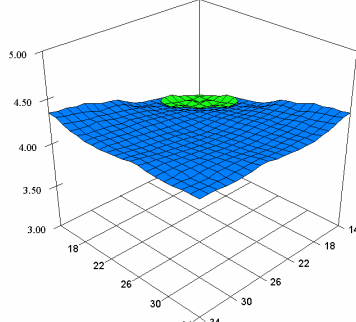
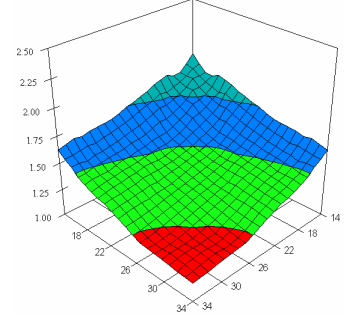
Valori medii pentru $90 \leq m,n \leq 100$					
Metoda	AvErr	StdDev	AvAD	AVADI	DevI
LRWaldz	4.61	0.590	0.387	0.462	0.705
LRAC	4.64	0.540	0.350	0.433	0.649
LRAC0	4.63	0.525	0.348	0.429	0.640
LRAC1	4.64	0.526	0.348	0.428	0.640
LRAC2	4.64	0.528	0.348	0.429	0.640
“Best of”	5	0	0	0	0

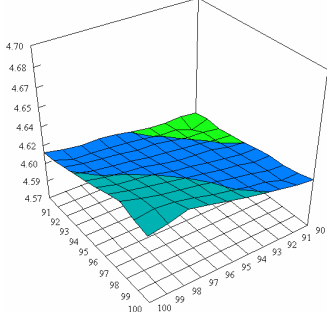
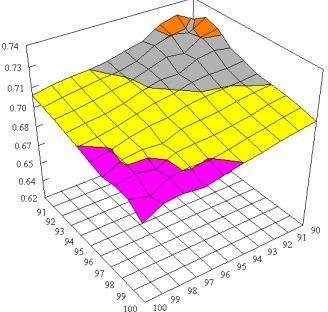
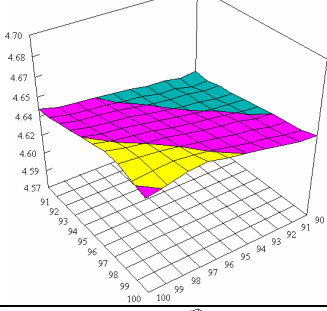
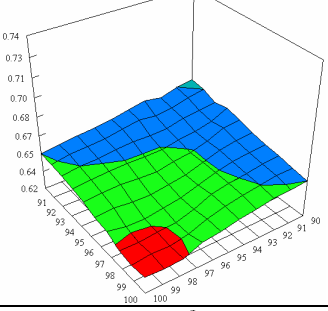
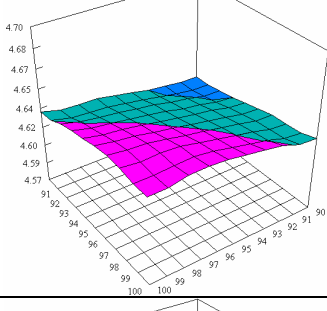
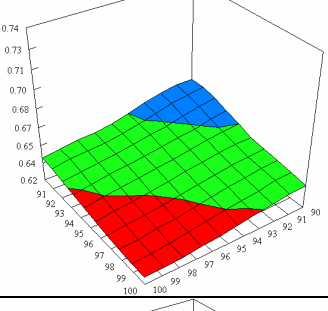
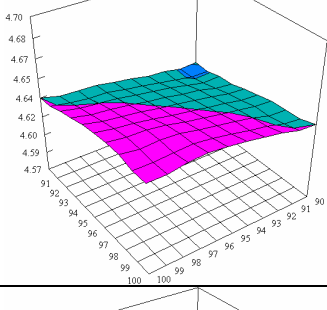
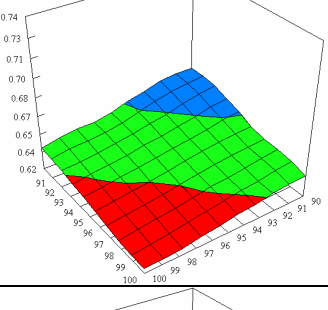
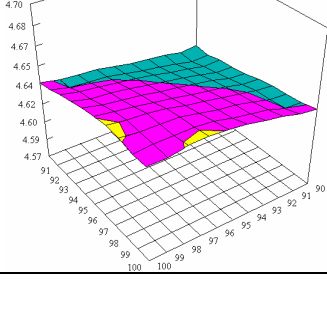
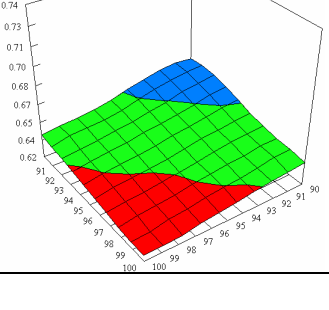
► Următorul tabel redă poziția metodei în ierarhia de performanță a metodelor pentru cele 5 criterii folosite, poziție obținută pentru cele 2 cazuri prezentate, precum și suma cumulată a pozițiilor, drept indicator de performanță (poziție mică, performanță mare; poziție mare, performanță mică) - 1..5 pentru $14 \leq m,n \leq 34$, 6..10 pentru $90 \leq m,n \leq 100$:

Met\Crit	1	2	3	4	5	6	7	8	9	10	Σ
LRWaldz	2	5	5	5	5	5	5	5	5	5	47
LRAC	1	4	3	1	3.5	2	4	4	4	4	30.5
LRAC0	5	1	1	4	3.5	4	1	2	2.5	2	26
LRAC1	4	2	2	3	2	2	2	2	1	2	22
LRAC2	3	3	4	2	1	2	3	2	2.5	2	24.5
“Best of”	1	1	1	1	1	1	1	1	1	1	min

► Tabelul de mai sus relevă apropierea în performanță a metodelor LRAC0, LRAC1 și LRAC2, toate formând un grup relativ compact și distant de celelalte două rămase LRWaldz și LRAC, toate reprezentând alegeri optimale în raport cu cele două. Din acest grup, una dintre ele se distanțează, și reprezintă cea mai bună optimizare a formulei de calcul pentru intervalul de încredere: LRAC1.

► Figurile din tabelul următor reprezintă distribuția erorii experimentale medii și a deviației față de eroarea impusă pentru cele 5 cazuri selectate pentru comparație pentru $14 \leq m,n \leq 34$:

	<p>LRWaldz</p> <p>Media erorii: 4.19</p> <p>Deviația față de 5%: 1.634</p>	
	<p>LRAC</p> <p>Media erorii: 4.22</p> <p>Deviația față de 5%: 1.485</p>	
	<p>LRAC0</p> <p>Media erorii: 4.16</p> <p>Deviația față de 5%: 1.485</p>	
	<p>LRAC1</p> <p>Media erorii: 4.17</p> <p>Deviația față de 5%: 1.484</p>	
	<p>LRAC2</p> <p>Media erorii: 4.18</p> <p>Deviația față de 5%: 1.481</p>	

	<p>LRWaldz</p> <p>Media erorii: 4.61</p> <p>Deviația față de 5%: 0.705</p>	
	<p>LRAC</p> <p>Media erorii: 4.64</p> <p>Deviația față de 5%: 0.649</p>	
	<p>LRAC0</p> <p>Media erorii: 4.63</p> <p>Deviația față de 5%: 0.640</p>	
	<p>LRAC1</p> <p>Media erorii: 4.64</p> <p>Deviația față de 5%: 0.640</p>	
	<p>LRAC2</p> <p>Media erorii: 4.64</p> <p>Deviația față de 5%: 0.640</p>	

► Figurile din tabelul anterior reprezintă distribuția erorii experimentale medii și a deviației față de eroarea impusă pentru cele 5 cazuri selectate pentru comparație pentru $90 \leq m, n \leq 100$.

► Evaluarea metodelor optimizate LRAC0, LRAC1, și LRAC2 în raport cu cele 2 metode folosite pentru comparație LRWaldz și LRAC s-a făcut prin generarea unui eșantion aleatoriu de 200 de valori (X,m,Y,n) , când $4 \leq m,n \leq 1000$, $0 \leq X \leq m$, $0 \leq Y \leq n$, calcularea intervalelor de încredere cu fiecare din metodele LRWaldz, LRAC, LRAC0, LRAC1, și LRAC2, calcularea erorii experimentale obținute prin utilizarea fiecărei metode, și compararea rezultatelor obținute folosind metodele de evaluare AvErr, StdDev, și DevI. Rezultatele sunt redată în tabelul următor:

Metoda	5-AvErr	StdDev	DevI
LRWaldz	0.150	2.210	2.209
LRAC	0.041	1.264	1.261
LRAC0	0.051	1.226	1.223
LRAC1	0.038	0.836	0.834
LRAC2	0.028	0.786	0.786
“Best of”	0	0	0

► Poziția metodelor în raport cu performanța obținută pentru fiecare criteriu în parte și poziția cumulativă este redată în tabelul următor:

Metoda	5-AvErr	StdDev	DevI	Σ
LRWaldz	5	5	5	15
LRAC	3	4	4	11
LRAC0	4	3	3	10
LRAC1	2	2	2	6
LRAC2	1	1	1	3
“Best of”	1	1	1	min

► Tabelul de mai sus arată în mod evident faptul că metoda LRAC2 se distanțează net de LRAC1, care la rândul său se distanțează net de grupul format de celelalte metode, LRAC2 fiind astfel metoda optimă pentru acest tip de funcție matematică $(X_n/Y/m)$ calculată pe tabela de contingență 2×2 .

Etapa 2, Activitatea 3. Proiectarea și implementarea modului de calcul al intervalelor de încredere și optimizarea intervalelor pentru Y/n-X/m

► Pornind de la expresia de definiție a funcției de probabilitate a distribuției normale standard:

$$P(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$

► Considerând α valoarea nivelului de semnificație (α este o probabilitate $0 \leq \alpha \leq 1$) a erorii impuse (uzual valoarea α se alege 0.05=5%; alte valori folosite: 0.1=10%, 0.01=1%), valoarea z se definește ca acel număr care verifică ecuația integrală:

$$\frac{\alpha}{2} = \int_z^{\infty} P(x)dx$$

► Următorul tabel redă valorile lui z pentru câteva valori uzuale ale lui α :

α	z
0.5	0.6745
0.1	1.645
0.05	1.96
0.01	2.576

► Funcția matematică Y/n-X/m este folosită în studiile medicale bazate pe tabela de contingență:

2x2	Caz	Test	Test
Succes	a	c	a+c
Eșec	b	d	b+d
Total	a+b	c+d	a+b+c+d

pentru studiu asupra relației între eșantionul caz și eșantionul test, denumirea consacrată în limba engleză pentru acest studiu fiind *excess risk* (riscul de exces) în cazul studiilor de factor de risc.

► Notând cu CI_{ER} intervalul de încredere pentru excesul de risc (folosită în studii de tip factor de risc), următoarea relație este relație de transformare de la formulele de definiție a acestui parametru medical la funcția matematică Y/n-X/m:

$$ER = \frac{a}{a+b} - \frac{c}{c+d} = \frac{Y}{n} - \frac{X}{m}$$

► Astfel, din punct de vedere matematic, acest parametru medicali se exprimă prin intermediul funcției matematice, pe care în continuare o vom numi ER:

$$ER = ER(X, m, Y, n) = \frac{Y}{n} - \frac{X}{m}$$

► Următoarele expresii au fost obținute pentru exprimarea intervalului de încredere al funcției ER:

N	Metodă	Expresie de calcul
0	ER0(X,m,Y,n)	$\frac{X(m-X)}{m^3} + \frac{Y(n-Y)}{n^3}$
0	ERWL(X,m,Y,n,z)	$\frac{Y}{n} - \frac{X}{m} - z\sqrt{ER0(X, m, Y, n)}$
0	ERWU(X,m,Y,n,z)	$\frac{Y}{n} - \frac{X}{m} + z\sqrt{ER0(X, m, Y, n)}$
1	ERWald(X,m,Y,n,z)	$(\max(ERWL(X, m, Y, n, z), -1), \min(ERWU(X, m, Y, n, z), 1))$
2	ERAC(X,m,Y,n,z)	$ERWald\left(X + \frac{z^2}{4\sqrt{2}}, m + \frac{z^2}{2\sqrt{2}}, Y + \frac{z^2}{4\sqrt{2}}, n + \frac{z^2}{2\sqrt{2}}\right)$
3	ERAs(X,m,Y,n,z)	$\frac{Y}{n} - \frac{X}{m} \pm z\sqrt{ER0(X, m, Y, n) + a \sin \frac{\sqrt{Y/n - X/m}}{mn}}$
4	ERAsC(X,m,Y,n,z)	$\frac{Y}{n} - \frac{X}{m} \pm z\sqrt{ER0(X, m, Y, n) + a \sin \frac{\sqrt{Y/n - X/m}}{2mn} + \frac{0.25}{m^{3/2}} + \frac{0.25}{n^{3/2}}}$
0	ER1(FL,FU,X,m,Y,a ₁ ,a ₂)	$(FL(Y, n, a_2) - FU(X, m, a_1), FU(Y, n, a_2) - FL(X, m, a_1))$
5	ERBinomial(X,m,Y,n,a)	$ER1(\text{BetaCJAL}, \text{BetaCJAU}, X, m, Y, n, \sqrt{a/2}, \sqrt{a/2})$
6	ERBinomialC(X,m,Y,n,a)	$ER1(\text{BetaCJAL}, \text{BetaCJAU}, X, m, Y, n, \frac{\sqrt{a/2}}{1 - \frac{1}{8\ln(m)}}, \frac{\sqrt{a/2}}{1 - \frac{1}{8\ln(n)}})$

► Notă: Expresile pentru $CI_{ER} = (CI_{ERL}, CI_{ERU})$ ale formulelor ERBinomial și ERBinomialC folosesc BetaCJAL și BetaCJAU (limitele inferioară și respectiv superioară ale intervalului de încredere calculat cu metoda BetaCJA):

$$(\text{BetaCJAL}(X, n, a), \text{BetaCJAU}(X, n, a)) = \text{BetaCJA}(X, n, a)$$

$$\text{BetaCJA}(X, n, a) = \text{Bin}\left(X, n, a, 1 - \frac{\sqrt{X(n-X)}}{n}, 1 - \frac{\sqrt{X(n-X)}}{n}\right)$$

$$\text{Bin}(X, n, a, c_1, c_2) = (\text{BinI}(X, n, a, c_1, c_2), \text{BinS}(X, n, a, c_1, c_2))$$

► Următoarele metode de evaluare au fost folosite pentru compararea rezultatelor metodelor de calcul al intervalului de încredere pentru funcția LR:

Nume	Metoda	Formula
AvErr	Av(Err)	$\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} \text{Err}(X, Y, m, n)}{(m-1)(n-1)}$
StdDev	StdDev(Err)	$\left(\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} (\text{Err}(X, Y, m, n) - \text{AvErr})^2}{(m-1)(n-1) - 1} \right)^{1/2}$
AvAD	AvAD(Err)	$\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} \text{Err}(X, Y, m, n) - \text{AvErr} }{(m-1)(n-1) - 1}$
AvADI	AvADI(Err)	$\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} \text{Err}(X, Y, m, n) - 100 \cdot \alpha }{(m-1)(n-1)}$
DevI	DevI(Err)	$\left(\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} (\text{Err}(X, Y, m, n) - 100 \cdot \alpha)^2}{(m-1)(n-1)} \right)^{1/2}$

► Funcția $\text{Err}(\cdot, \cdot, \cdot, \cdot)$ cumulează erori experimentale folosind distribuția binomială univariată pentru generarea distribuției binomiale bivariate. Notând capetele intervalului de încredere pentru funcția ER prin (ERL, ERU) - și aici nota bene ERL și ERU sunt funcții date de una din metodele de calcul al intervalului de încredere (1-6 în tabelul de mai sus), funcția $\text{Err}(X, m, Y, n)$ se calculează din:

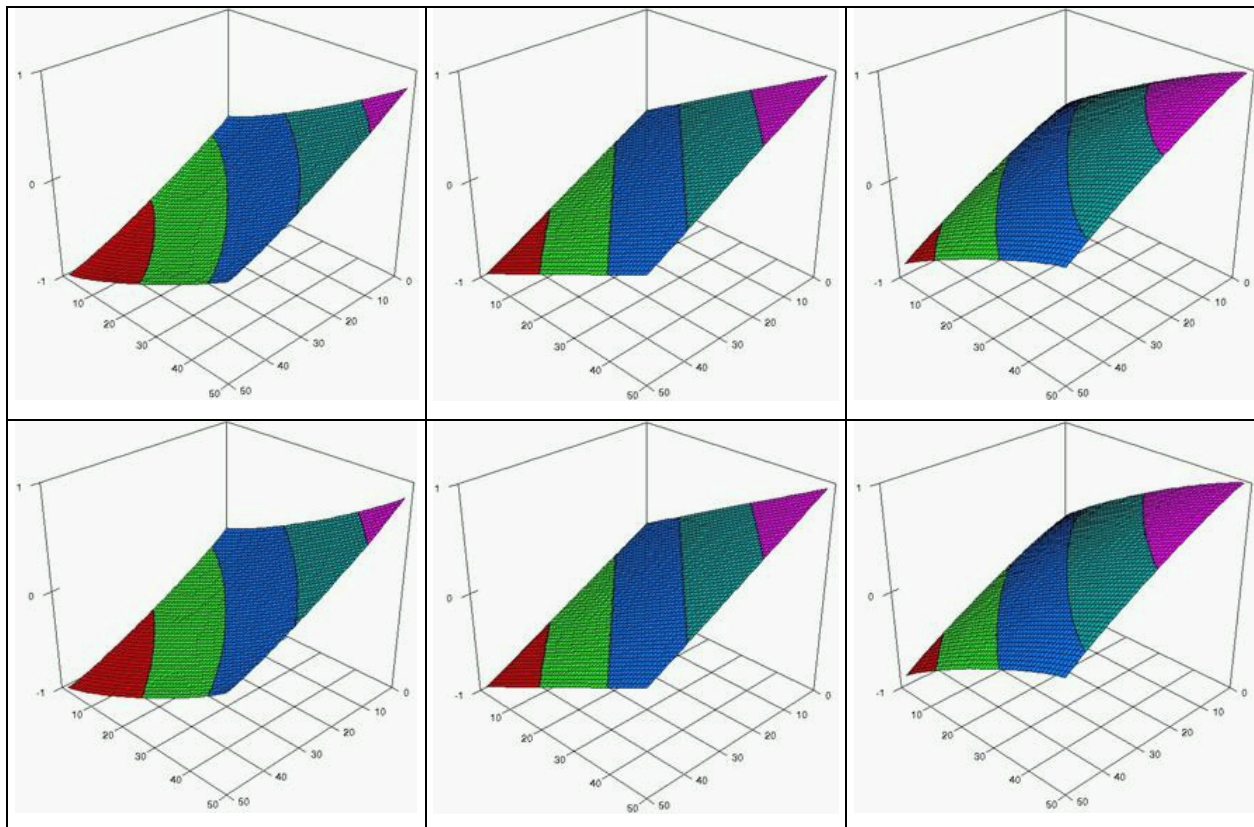
$$d\text{Bin}(m, X, \Xi) := \frac{m!}{\Xi!(m-\Xi)!} \cdot \left(\frac{X}{m}\right)^\Xi \cdot \left(1 - \frac{X}{m}\right)^{m-\Xi}$$

$$d\text{Bin}(n, Y, \Psi) := \frac{n!}{\Psi!(n-\Psi)!} \cdot \left(\frac{Y}{n}\right)^\Psi \cdot \left(1 - \frac{Y}{n}\right)^{n-\Psi}$$

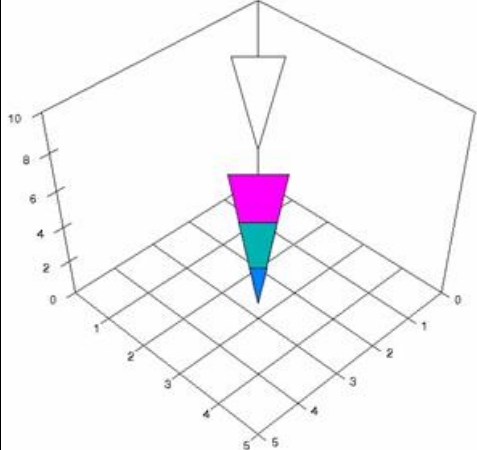
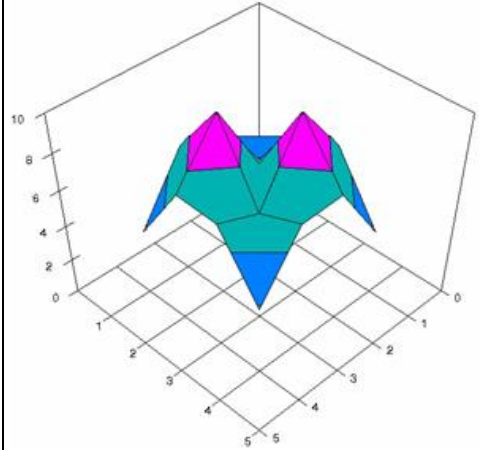
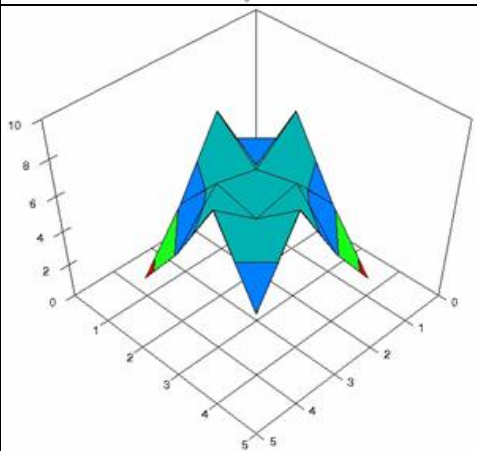
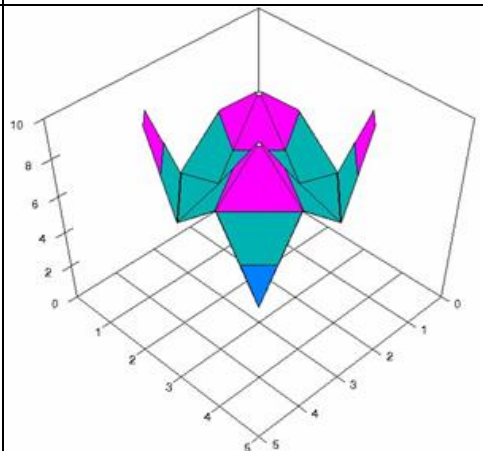
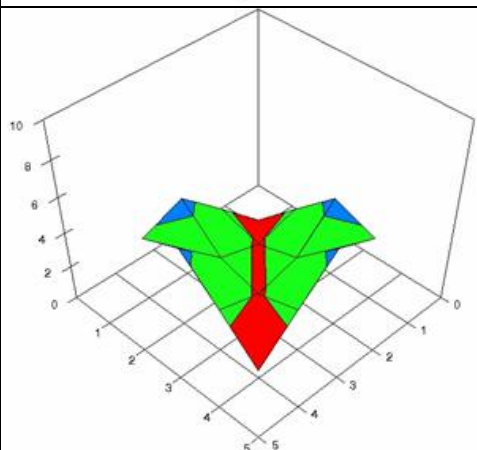
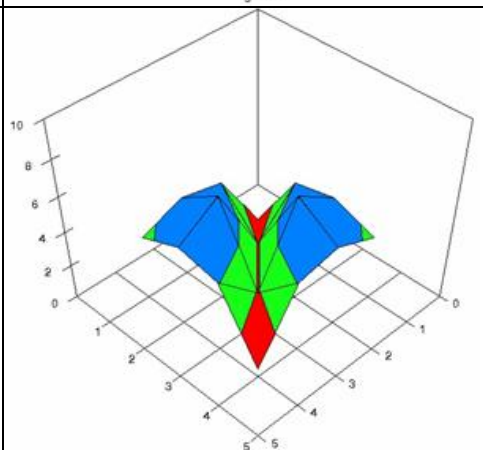
$$\sum_{\substack{\text{ERL}(\Xi, \Psi, m, n) \\ > \text{ER}(X, Y, m, n)}} d\text{Bin}(m, X, \Xi) \cdot d\text{Bin}(n, Y, \Psi) + \sum_{\substack{\text{ERU}(\Xi, \Psi, m, n) \\ < \text{ER}(X, Y, m, n)}} d\text{Bin}(m, X, \Xi) \cdot d\text{Bin}(n, Y, \Psi)$$

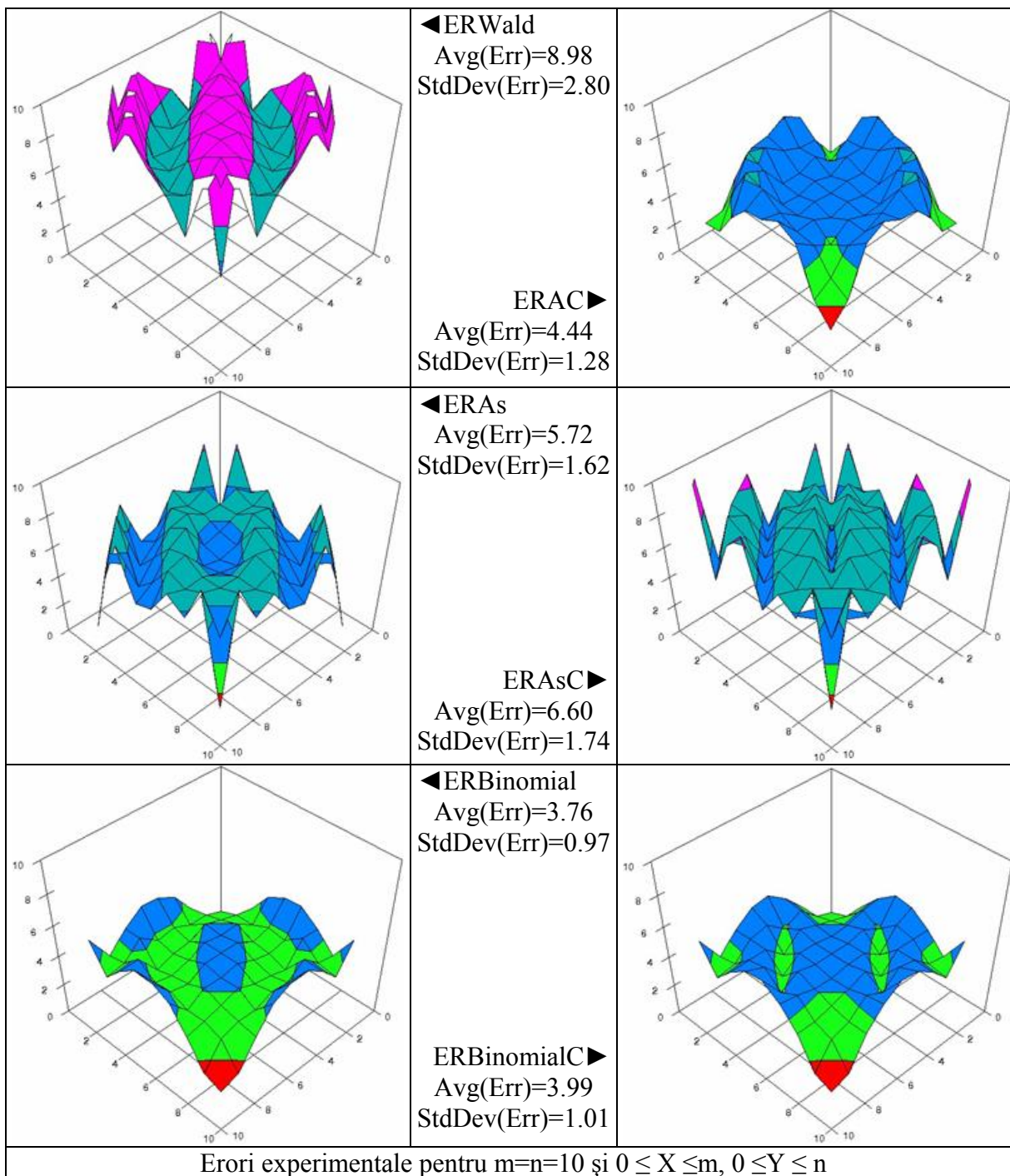
$$\sum_{\Xi=1}^{m-1} \sum_{\Psi=1}^{n-1} d\text{Bin}(m, X, \Xi) \cdot d\text{Bin}(n, Y, \Psi)$$

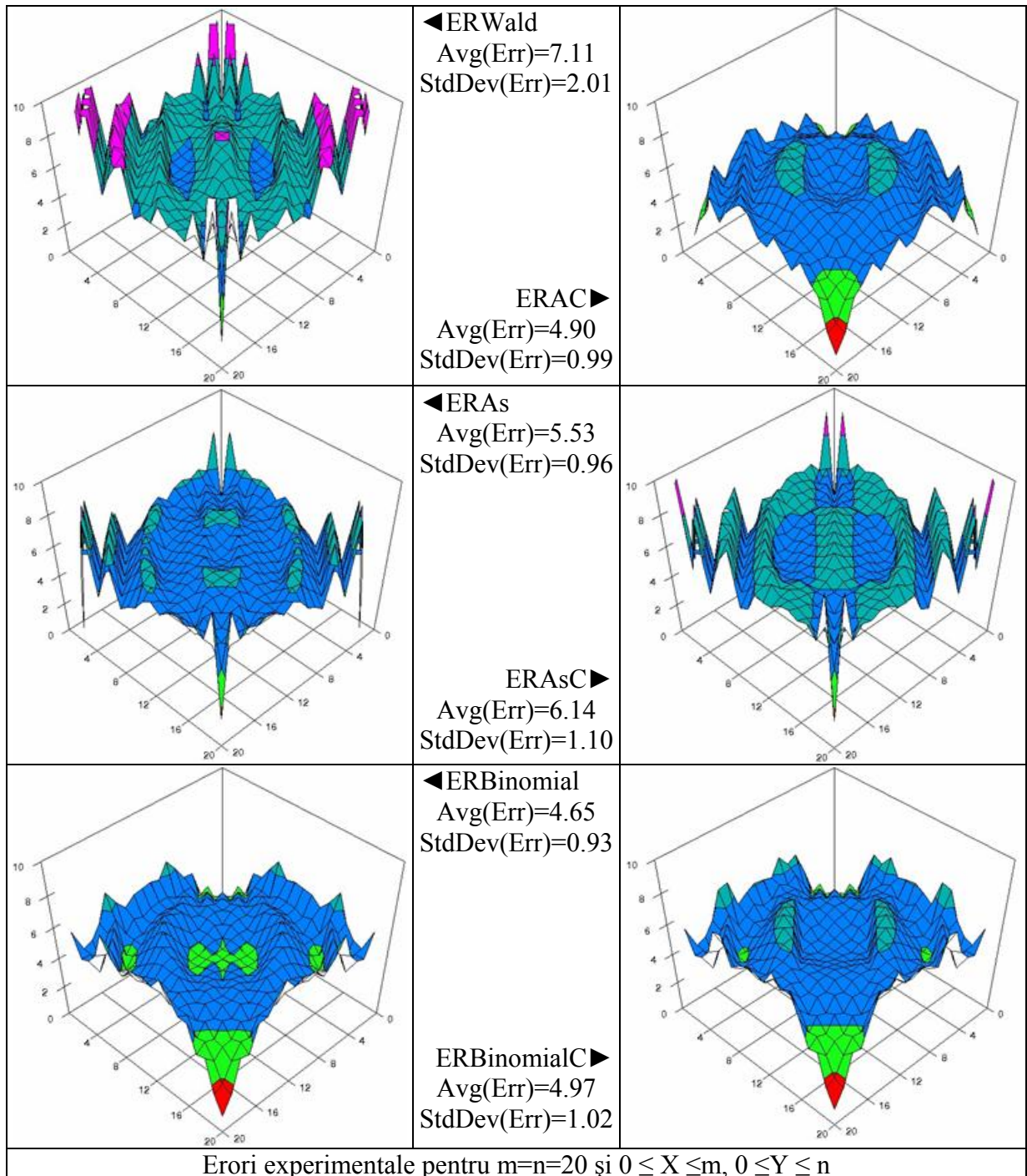
► Următorul tabel conține reprezentările grafice ale intervalelor de încredere (stânga - limita inferioară, centru - valoarea expresiei excesului de risc, dreapta - limita superioară) obținute cu metodele ERBinomialC (prima linie din tabel) și ERAC (a doua linie din tabel):

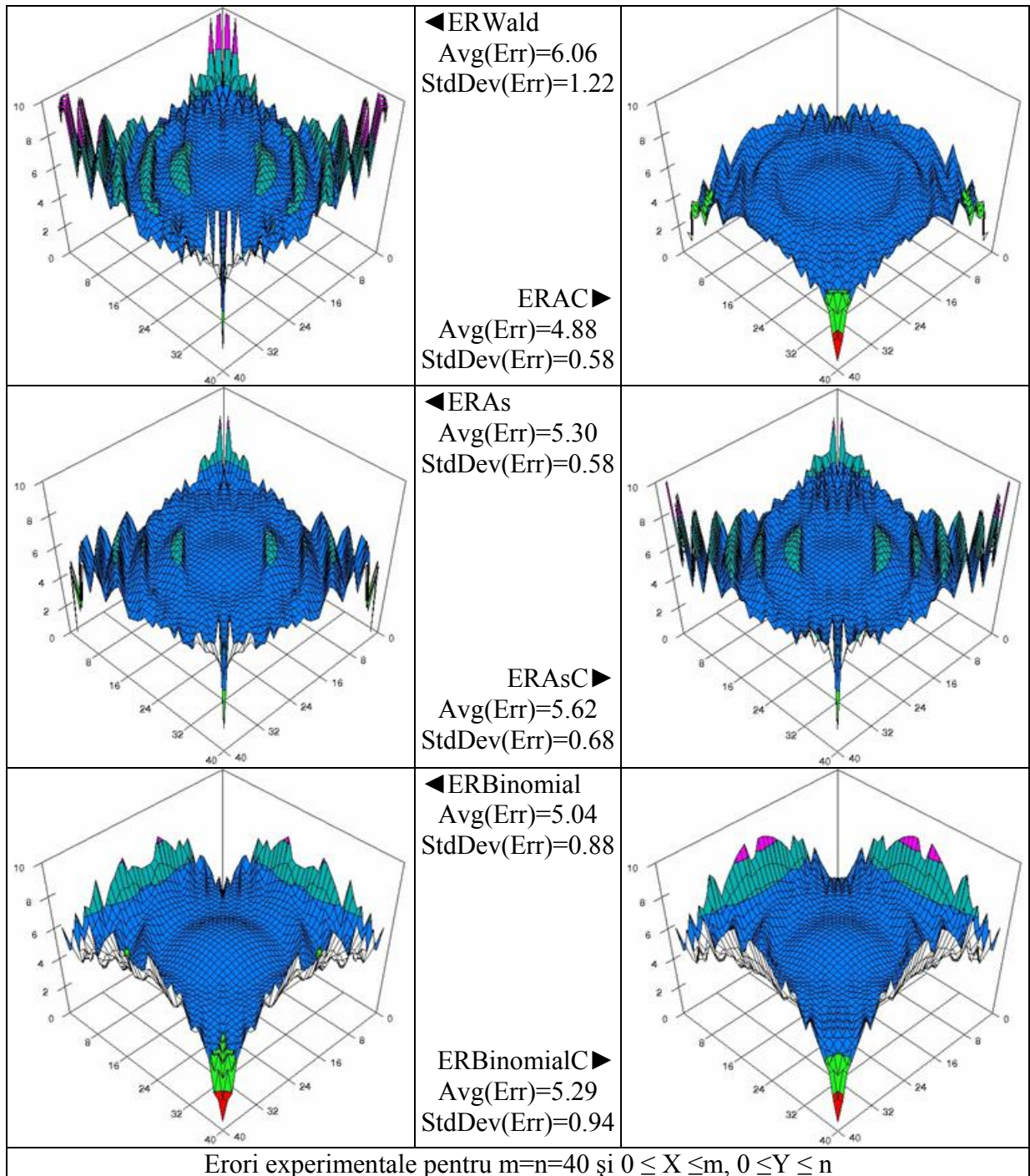


Următoarele tabele prezintă descriptiv rezultatele obținute (în termeni de eroare experimentală) cu fiecare dintre metodele ERWald, ERAC, ERAs, ERAsC, ERBinomial, și ERBinomialC.

	<p>◀ERWald Avg(Err)=12.32 StdDev(Err)=4.13</p>	
	<p>◀ERAs Avg(Err)=5.30 StdDev(Err)=2.30</p>	
	<p>◀ERBinomial Avg(Err)=2.87 StdDev(Err)=1.36</p>	
<p>ERBinomialC▶ Avg(Err)=3.70 StdDev(Err)=1.69</p>		
<p>Erori experimentale pentru $m=n=5$ și $0 \leq X \leq m$, $0 \leq Y \leq n$</p>		







► Următorul tabel sumarizează performanțele metodelor obținute prin prisma celor 2 criterii de evaluare pentru volume de eşantioane reprezentative:

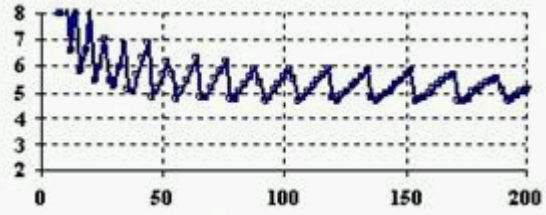
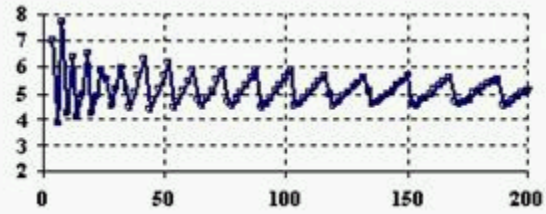
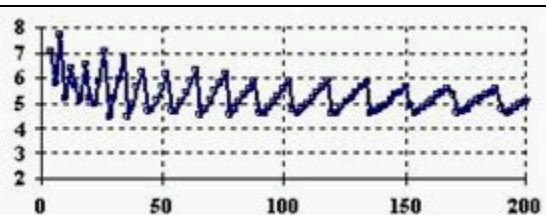
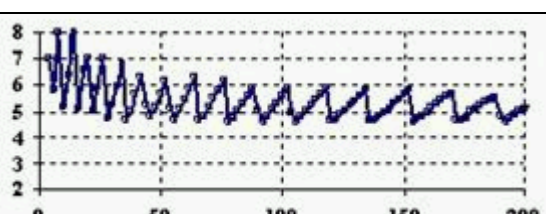
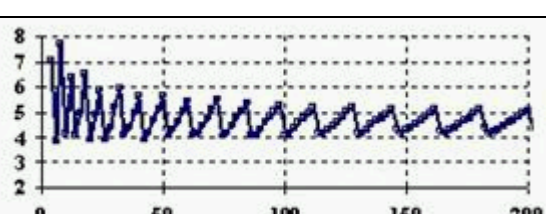
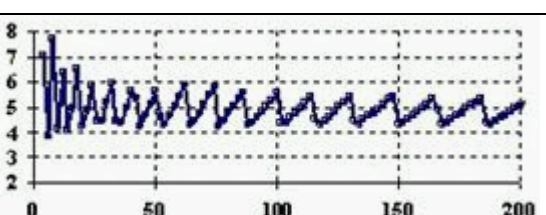
Domeniu	m=n=5		m=n=10		m=n=20		m=n=40	
Metoda	5-AvErr	StdDev	5-AvErr	StdDev	5-AvErr	StdDev	5-AvErr	StdDev
ERWald	7.32	4.13	3.98	2.80	2.11	2.01	1.06	1.22
ERAC	1.53	1.99	0.56	1.28	0.1	0.99	0.12	0.58
ERAs	0.3	2.30	0.72	1.62	0.53	0.96	0.3	0.58
ERAsC	2.65	2.37	1.6	1.74	1.14	1.10	0.62	0.68
ERBinomial	2.13	1.36	1.24	0.97	0.35	0.93	0.04	0.88
ERBinomialC	1.3	1.69	1.01	1.01	0.03	1.02	0.29	0.94
“Best of”	0	0	0	0	0	0	0	0

► Următorul tabel redă poziția metodei în ierarhia de performanță a metodelor pentru cele 2 criterii folosite, poziție obținută pentru cele 4 cazuri prezentate, precum și suma cumulată a pozițiilor, drept indicator de performanță (poziție mică, performanță mare; poziție mare, performanță mică) - 1..6:

Metoda	1	2	3	4	5	6	7	8	Σ
ERWald	6	6	6	6	6	6	6	6	48
ERAC	3	3	1	3	2	3	2	1.5	18.5
ERAs	1	4	2	4	4	2	4	1.5	22.5
ERAsC	5	5	5	5	5	5	5	3	38
ERBinomial	4	1	4	1	3	1	1	4	19
ERBinomialC	2	2	3	2	1	4	3	5	22
“Best of”	1	1	1	1	1	1	1	1	min

► Tabelul de mai sus relevă apropierea în performanță a metodelor ERAC și ERBinomial și de asemenea ERAs și ERBinomialC, formând două grupuri relativ compacte și distante de celelalte două rămase ERAsC și ERWald, toate reprezentând alegeri optime în raport cu cele două. Din acest grup, una dintre ele se distanțează, și reprezintă cea mai bună optimizare a formulei de calcul pentru intervalul de încredere: ERAC.

► Investigarea punctului central ($X=Y$) pentru eşantioane de volume pare și egale ($m=n=2k$) a constituit subiectul investigațiilor suplimentare, pentru a vedea dacă supoziția făcută (ERAC este “Best of”) este adevărată. Următorul tabel reprezintă variația erorii cu creșterea volumului eşantionului, acesta variind în intervalul $4 \leq m = n \leq 204$.

	<p>ERWald</p> <p>Media erorii experimentale: 5.67</p> <p>Deviația standard a erorii experimentale: 1.62</p>
	<p>ERAC</p> <p>Media erorii experimentale: 5.13</p> <p>Deviația standard a erorii experimentale: 0.60</p>
	<p>ERAs</p> <p>Media erorii experimentale: 5.29</p> <p>Deviația standard a erorii experimentale: 0.61</p>
	<p>ERAsC</p> <p>Media erorii experimentale: 5.38</p> <p>Deviația standard a erorii experimentale: 0.71</p>
	<p>ERBinomial</p> <p>Media erorii experimentale: 4.73</p> <p>Deviația standard a erorii experimentale: 0.64</p>
	<p>ERBinomialC</p> <p>Media erorii experimentale: 4.95</p> <p>Deviația standard a erorii experimentale: 0.61</p>

► Următorul tabel clasifică metodele după performanța în punctul central:

Metoda	ERWald	ERAC	ERAs	ERAsC	ERBinomial	ERBinomialC	“Best of”
5-Avg	6	2	4	5	3	1	1
StdDev	6	1	2.5	5	4	2.5	1
Σ	12	3	6.5	10	7	3.5	min

Tabelul arată că metoda ERAC este într-adevăr “Best of” (scor global 3), urmată foarte aproape de ERBinomialC (scor global 3.5).

ET46/2006 – Et. Finală/2007 – Lucrare in extenso

► Au fost generate 100 de valori (X,m,Y,n) respectând $4 \leq m, n \leq 1000$, $0 \leq X \leq m$, $0 \leq Y \leq n$ și a fost evaluată performanța metodelor prin prisma distribuției erorii experimentale. Tabelul de mai jos redă aceste împrăștieri pentru 4 metode:

	<p>ERWald</p> <p>DevI = 0.90</p> <p>$5\text{-AvgErr} = 0.30$</p> <p>StdDev = 0.85</p>
	<p>ERAC</p> <p>DevI = 0.32</p> <p>$5\text{-AvgErr} = 0.05$</p> <p>StdDev = 0.32</p>
	<p>ERAs</p> <p>DevI = 0.87</p> <p>$5\text{-AvgErr} = 0.23$</p> <p>StdDev = 0.84</p>
	<p>ERAsC</p> <p>DevI = 0.89</p> <p>$5\text{-AvgErr} = 0.27$</p> <p>StdDev = 0.85</p>

ET46/2006 – Et. Finală/2007 – Lucrare in extenso

► Următorul tabel redă performanța metodelor (aproprierea erorii experimentale față de nivelul impus) în raport cu eșantionul aleator și cu cele 3 criterii de comparație folosite și scorul global obținut de fiecare metodă în parte:

Metoda	DevI	5-AvgErr	StdDev	Σ
ERWald	4	4	4	12
ERAC	1	1	1	3
ERAs	2	2	2	6
ERAsC	3	3	3	9
“Best of”	1	1	1	min

► Se demonstrează prin tabelul de mai sus (și demonstrația rezidă prin toate criteriile) că metoda ERAC reprezintă metoda optimă de calcul al intervalului de încredere pentru funcția ER.

Etapa 3, Activitatea 1. Proiectarea și implementarea modului de calcul al intervalelor de încredere și optimizarea intervalelor pentru $|Y/n - X/m|$

► Pornind de la expresia de definiție a funcției de probabilitate a distribuției normale standard:

$$P(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$

► Considerând α valoarea nivelului de semnificație (α este o probabilitate $0 \leq \alpha \leq 1$) a erorii impuse (uzual valoarea α se alege 0.05=5%; alte valori folosite: 0.1=10%, 0.01=1%), valoarea z se definește ca acel număr care verifică ecuația integrală:

$$\frac{\alpha}{2} = \int_z^{\infty} P(x)dx$$

► Următorul tabel redă valorile lui z pentru câteva valori uzuale ale lui α :

α	z
0.5	0.6745
0.1	1.645
0.05	1.96
0.01	2.576

► Funcția matematică $|Y/n - X/m|$ este folosită în studiile medicale bazate pe tabela de contingență:

2x2	Caz	Test	Test
Succes	a	c	a+c
Eșec	b	d	b+d
Total	a+b	c+d	a+b+c+d

pentru studiu asupra relației între eșantionul caz și eșantionul test, denumirea consacrată în limba engleză pentru acest studiu fiind *absolute risk reduction/increase* (reducerea/creșterea riscului absolut) și *benefit increase* (creșterea beneficiului) și este folosită în cazul studiilor de terapie.

► Notând cu CI_{AR} intervalul de încredere pentru reducerea/creșterea riscului absolut (folosită în cazul studiilor de terapie), următoarea relație este relație de transformare de la formulele de definiție a acestui parametru medical la funcția matematică $|Y/n - X/m|$:

$$ARR = \left| \frac{a}{a+b} - \frac{c}{c+d} \right| = \left| \frac{Y}{n} - \frac{X}{m} \right|$$

ET46/2006 – Et. Finală/2007 – Lucrare in extenso

► Astfel, din punct de vedere matematic, acest parametrii medicali se exprimă prin intermediul funcției matematice, pe care în continuare o vom numi AR:

$$AR = AR(X, m, Y, n) = \left| \frac{X}{m} - \frac{Y}{n} \right|$$

► Următoarele expresii au fost obținute pentru exprimarea intervalului de încredere al funcției AR:

N	Metodă	Expresie de calcul
0	$V0(X, m, Y, n)$	$\frac{X(m - X)}{m^3} + \frac{Y(n - Y)}{n^3}$
0	$V1(X, m)$	$\frac{1}{2} - \frac{\sqrt{X(m - X)}}{m^2}$
0	$V2(X, m)$	$2W1(X, m)$
0	$ARWL(X, m, Y, n, z)$	$\left \frac{X}{m} - \frac{Y}{n} \right - z\sqrt{V0(X, m, Y, n)}$
0	$ARWU(X, m, Y, n, z)$	$\left \frac{X}{m} - \frac{Y}{n} \right + z\sqrt{V0(X, m, Y, n)}$
1	$ARWld(X, m, Y, n, z)$	$(\max(ARWL(X, m, Y, n, z), 0), \min(ARWU(X, m, Y, n, z), 1))$
2	$ARAC(X, m, Y, n, z)$	$ERWld\left(X + \frac{z^2}{4\sqrt{2}}, m + \frac{z^2}{2\sqrt{2}}, Y + \frac{z^2}{4\sqrt{2}}, n + \frac{z^2}{2\sqrt{2}}\right)$
3	$ARAC1(X, m, Y, n, z)$	$ARWld(X + V1(X, m), m + V2(X, m), Y + V1(Y, n), n + V2(Y, n))$
4	$ADAs0(X, m, Y, n, z)$	$\left(\begin{array}{l} \max \left(\left \frac{X}{m} - \frac{Y}{n} \right - z\sqrt{V0(X, m, Y, n)} + \left(1 + \frac{2}{\sqrt{mn}}\right) \frac{\arcsin \sqrt{\left \frac{X}{m} - \frac{Y}{n} \right }}{\pi \cdot m \cdot n / 2}, 0 \right), \\ \min \left(\left \frac{X}{m} - \frac{Y}{n} \right + z\sqrt{V0(X, m, Y, n)} + \left(1 + \frac{2}{\sqrt{mn}}\right) \frac{\arcsin \sqrt{\left \frac{X}{m} - \frac{Y}{n} \right }}{\pi \cdot m \cdot n / 2}, 1 \right) \end{array} \right)$

5	ADAs1(X,m,Y,n,z)	$\left(\begin{array}{l} \max \left(\left \frac{X}{m} - \frac{Y}{n} \right - z \sqrt{V0(X, m, Y, n) + \frac{\arcsin \sqrt{\left \frac{X}{m} - \frac{Y}{n} \right }}{m \cdot n}}, 0 \right), \\ \min \left(\left \frac{X}{m} - \frac{Y}{n} \right + z \sqrt{V0(X, m, Y, n) + \frac{\arcsin \sqrt{\left \frac{X}{m} - \frac{Y}{n} \right }}{m \cdot n}}, 1 \right) \end{array} \right)$
6	ADAs2(X,m,Y,n,z)	$\left(\begin{array}{l} \max \left(\left \frac{X}{m} - \frac{Y}{n} \right - z \sqrt{V0(X, m, Y, n) + \frac{\arcsin \sqrt{\left \frac{X}{m} - \frac{Y}{n} \right }}{2 \cdot m \cdot n} + \frac{m^{-3}}{4} + \frac{n^{-3}}{4}}, 0 \right), \\ \min \left(\left \frac{X}{m} - \frac{Y}{n} \right + z \sqrt{V0(X, m, Y, n) + \frac{\arcsin \sqrt{\left \frac{X}{m} - \frac{Y}{n} \right }}{2 \cdot m \cdot n} + \frac{m^{-3}}{4} + \frac{n^{-3}}{4}}, 1 \right) \end{array} \right)$
0	AR1(FL,FU, X,m,Y,a ₁ ,a ₂)	(FL(Y, n, a ₂) – FU(X, m, a ₁), FU(Y, n, a ₂) – FL(X, m, a ₁))
0	AR2((u,v))	$\begin{cases} (0, \max(u , v)), & u \cdot v \leq 0 \\ (\min(u , v), \max(u , v)), & u \cdot v > 0 \end{cases}$
7	ARJ(X,m,Y,n,a)	AR2(AR1(BetaCJ0L, BetaCJ0U, X, m, Y, n, $\sqrt{a/2}$, $\sqrt{a/2}$))
8	ARJC(X,m,Y,n,a)	AR2(AR1(BetaCJ0L, BetaCJ0U, X, m, Y, n, $\frac{\sqrt{a/2}}{1 - \frac{1}{8 \ln(m)}}$, $\frac{\sqrt{a/2}}{1 - \frac{1}{8 \ln(n)}}$))
9	ARB(X,m,Y,n,a)	AR2(AR1(BetaCJAL, BetaCJAU, X, m, Y, n, $\sqrt{a/2}$, $\sqrt{a/2}$))
10	ARBC(X,m,Y,n,a)	AR2(AR1(BetaCJAL, BetaCJAU, X, m, Y, n, $\frac{\sqrt{a/2}}{1 - \frac{1}{8 \ln(m)}}$, $\frac{\sqrt{a/2}}{1 - \frac{1}{8 \ln(n)}}$))

► Notă: Expresile pentru $CI_{AR} = (CI_{ARL}, CI_{ARU})$ ale formulelor ARJ și ARJC folosesc BetaCJ0L și BetaCJ0U (limitele inferioară și respectiv superioară ale intervalului de încredere calculat cu metoda BetaCJ0), și respectiv ale formulelor ARB și ARBC folosesc BetaCJAL și BetaCJAU (limitele inferioară și respectiv superioară ale intervalului de încredere calculat cu metoda BetaCJA):

$$(BetaCJ0L(X, n, a), BetaCJ0U(X, n, a)) = BetaCJ0(X, n, a)$$

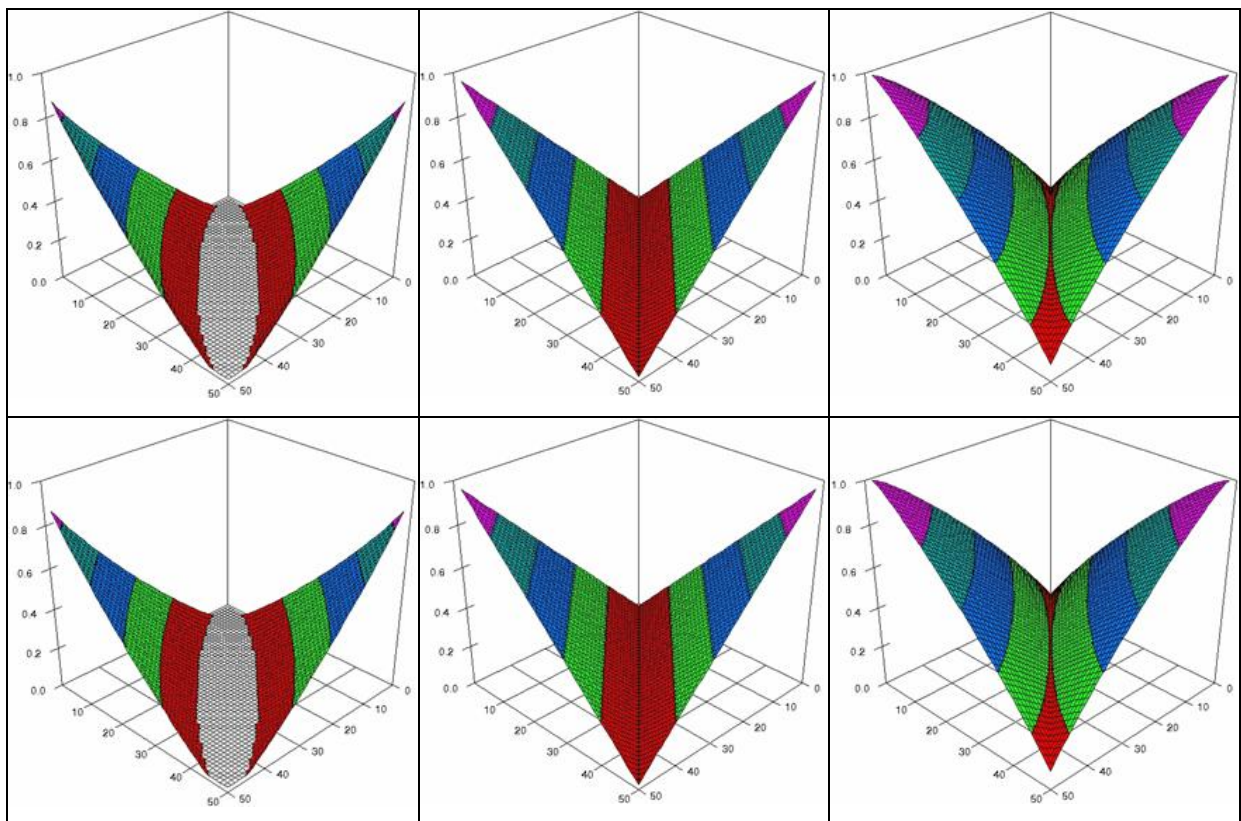
$$(BetaCJAL(X, n, a), BetaCJAU(X, n, a)) = BetaCJA(X, n, a)$$

$$BetaCJ0(X, n, a) = Bin(X, n, a, \frac{1}{2}, \frac{1}{2})$$

$$BetaCJA(X, n, a) = Bin(X, n, a, 1 - \frac{\sqrt{X(n-X)}}{n}, 1 - \frac{\sqrt{X(n-X)}}{n})$$

$$Bin(X, n, a, c_1, c_2) = (BinI(X, n, a, c_1, c_2), BinS(X, n, a, c_1, c_2))$$

► Următorul tabel conține reprezentările grafice ale intervalelor de încredere (stânga - limita inferioară, centru - valoarea expresiei creșterii beneficiului, dreapta - limita superioară) obținute cu metodele ARBC (prima linie din tabel) și ARAC (a doua linie din tabel):



► Următoarele metode de evaluare au fost folosite pentru compararea rezultatelor metodelor de calcul al intervalului de încredere pentru funcția LR:

Nume	Metoda	Formula
AvErr	Av(Err)	$\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} \text{Err}(X, Y, m, n)}{(m-1)(n-1)}$
StdDev	StdDev(Err)	$\left(\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} (\text{Err}(X, Y, m, n) - \text{AvErr})^2}{(m-1)(n-1) - 1} \right)^{1/2}$
AvAD	AvAD(Err)	$\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} \text{Err}(X, Y, m, n) - \text{AvErr} }{(m-1)(n-1) - 1}$
AvADI	AvADI(Err)	$\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} \text{Err}(X, Y, m, n) - 100 \cdot \alpha }{(m-1)(n-1)}$
DevI	DevI(Err)	$\left(\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} (\text{Err}(X, Y, m, n) - 100 \cdot \alpha)^2}{(m-1)(n-1)} \right)^{1/2}$

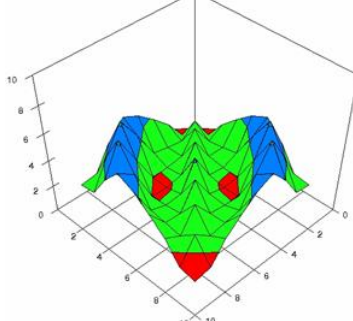
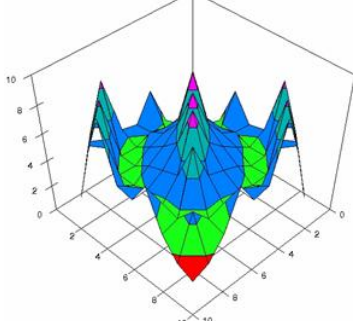
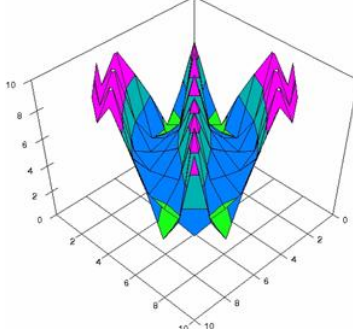
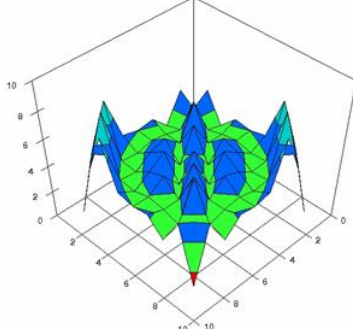
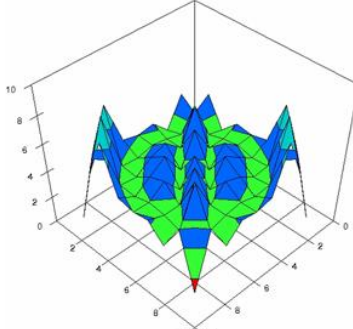
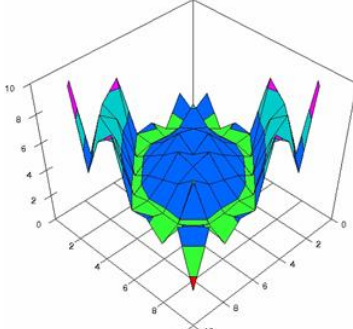
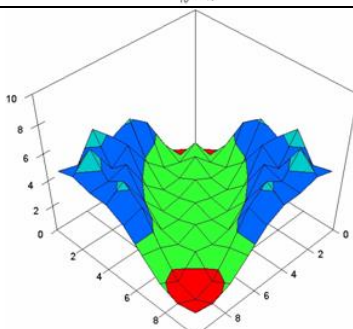
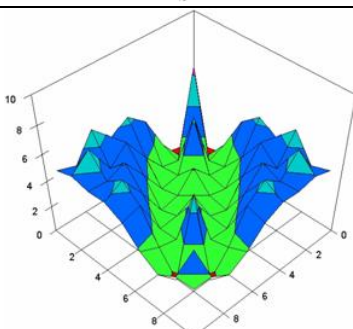
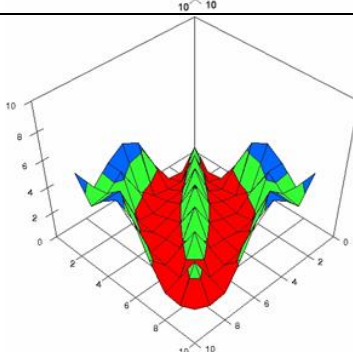
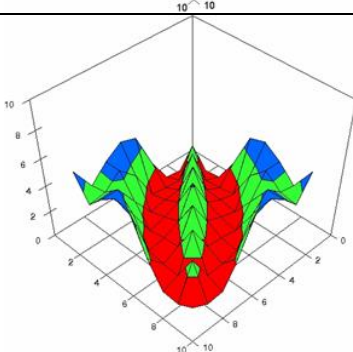
► Funcția $\text{Err}(\cdot, \cdot, \cdot, \cdot)$ cumulează erori experimentale folosind distribuția binomială univariată pentru generarea distribuției binomiale bivariate. Notând capetele intervalului de încredere pentru funcția ER prin (ERL, ERU) - și aici nota bene ERL și ERU sunt funcții date de una din metodele de calcul al intervalului de încredere (1-6 în tabelul de mai sus), funcția $\text{Err}(X, m, Y, n)$ se calculează din:

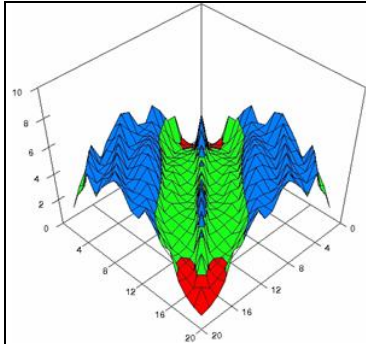
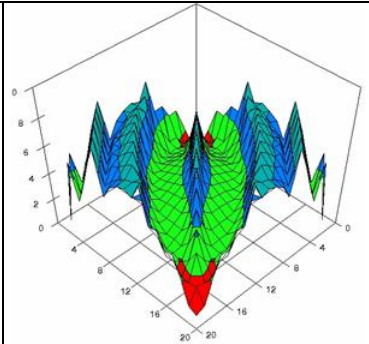
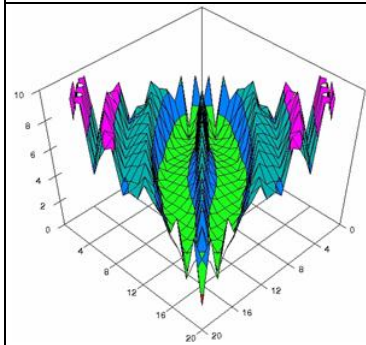
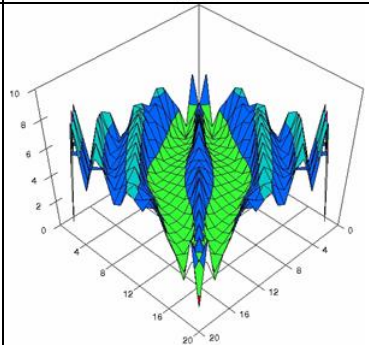
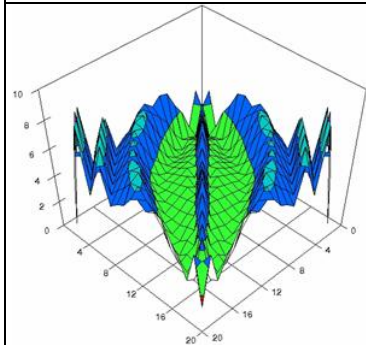
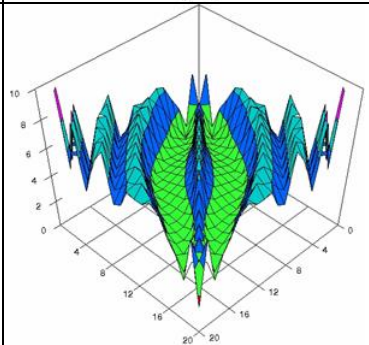
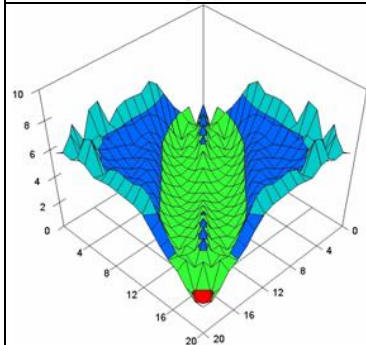
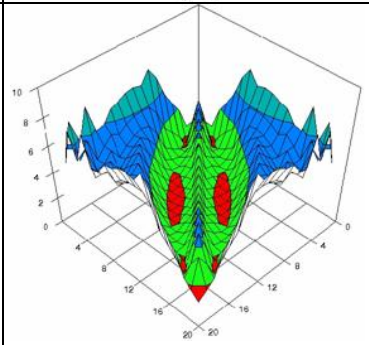
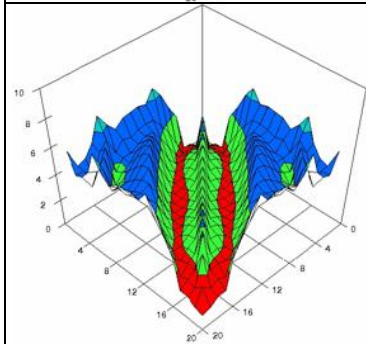
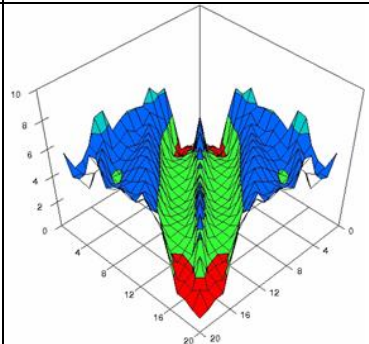
$$d\text{Bin}(m, X, \Xi) := \frac{m!}{\Xi!(m-\Xi)!} \cdot \left(\frac{X}{m}\right)^\Xi \cdot \left(1 - \frac{X}{m}\right)^{m-\Xi}$$

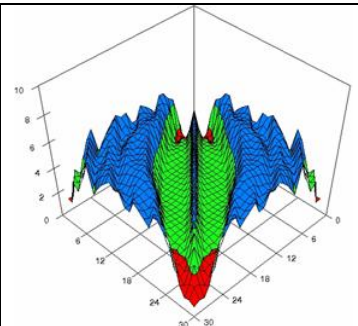
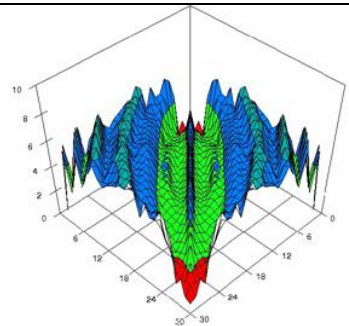
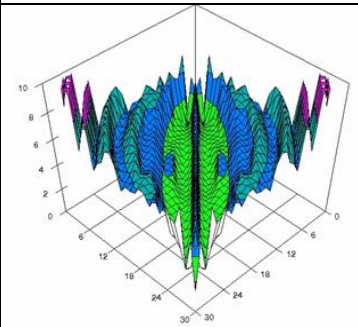
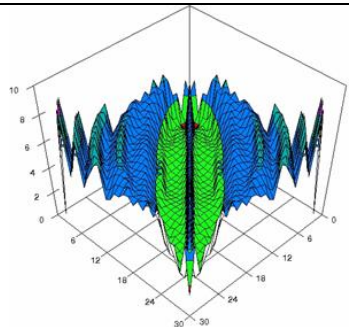
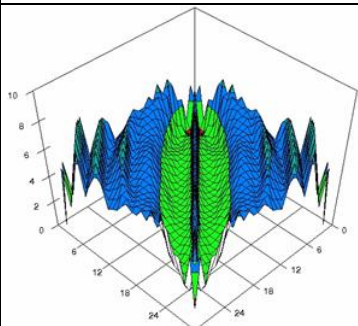
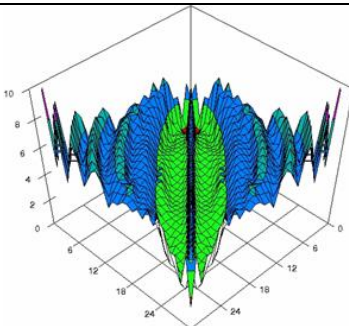
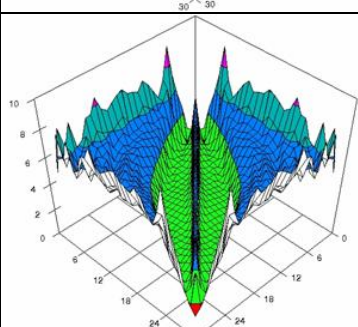
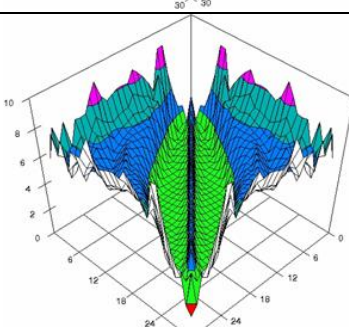
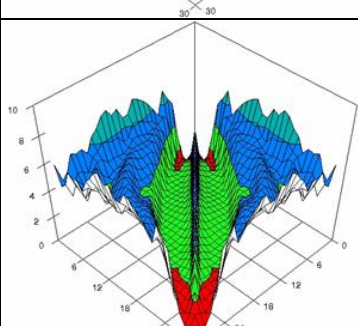
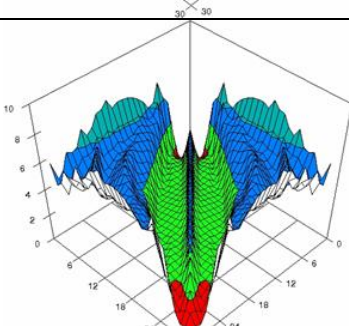
$$d\text{Bin}(n, Y, \Psi) := \frac{n!}{\Psi!(n-\Psi)!} \cdot \left(\frac{Y}{n}\right)^\Psi \cdot \left(1 - \frac{Y}{n}\right)^{n-\Psi}$$

$$\frac{\sum_{\substack{\text{ERL}(\Xi, \Psi, m, n) \\ > \text{ER}(X, Y, m, n)}} d\text{Bin}(m, X, \Xi) \cdot d\text{Bin}(n, Y, \Psi) + \sum_{\substack{\text{ERU}(\Xi, \Psi, m, n) \\ < \text{ER}(X, Y, m, n)}} d\text{Bin}(m, X, \Xi) \cdot d\text{Bin}(n, Y, \Psi)}{\sum_{\Xi=1}^{m-1} \sum_{\Psi=1}^{n-1} d\text{Bin}(m, X, \Xi) \cdot d\text{Bin}(n, Y, \Psi)}$$

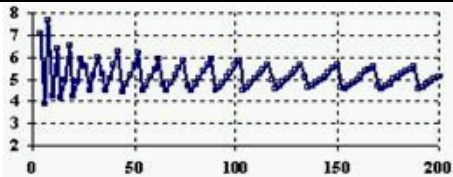
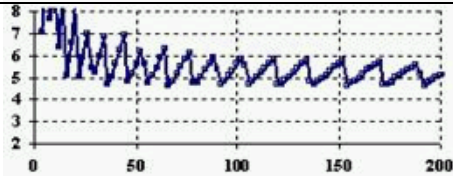
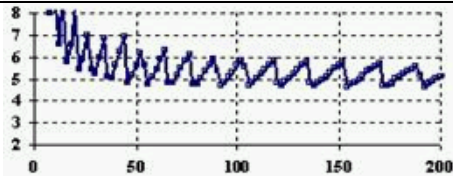
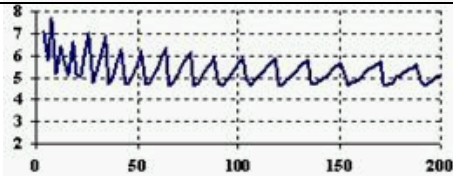
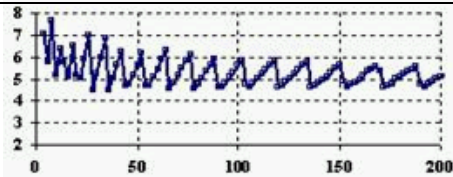
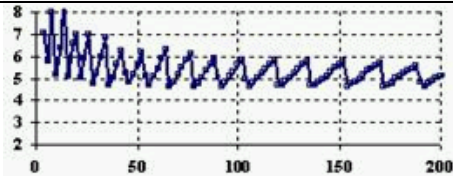
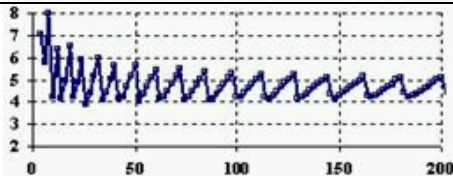
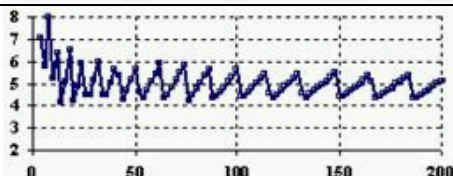
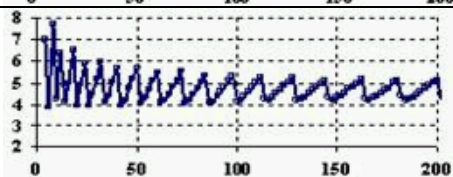
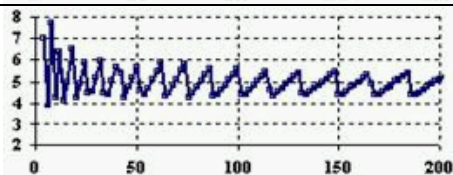
► Următoarele tabele prezintă descriptiv rezultatele obținute (în termeni de eroare experimentală) cu fiecare dintre cele 10 metode pentru $m=n=10, 20, 30$.

	<p>◀ARAC Avg(Err)=3.10 StdDev(Err)=1.36</p>	
	<p>◀ARWld Avg(Err)=7.00 StdDev(Err)=3.60</p>	
	<p>◀ARAs1 Avg(Err)=4.24 StdDev(Err)=1.40</p>	
	<p>◀ARJ Avg(Err)=3.78 StdDev(Err)=1.95</p>	
	<p>◀ARB Avg(Err)=2.43 StdDev(Err)=1.58</p>	
<p>Erori experimentale pentru $m=n=10$ și $0 \leq X \leq m$, $0 \leq Y \leq n$</p>		

	<p>◀ARAC Avg(Err)=3.75 StdDev(Err)=1.38</p>	
	<p>◀ARWld Avg(Err)=5.74 StdDev(Err)=2.60</p>	
	<p>◀ARAs1 Avg(Err)=4.36 StdDev(Err)=1.44</p>	
	<p>◀ARJ Avg(Err)=4.44 StdDev(Err)=1.72</p>	
	<p>◀ARB Avg(Err)=3.60 StdDev(Err)=1.61</p>	
<p>Ereri experimentale pentru $m=n=20$ și $0 \leq X \leq m$, $0 \leq Y \leq n$</p>		

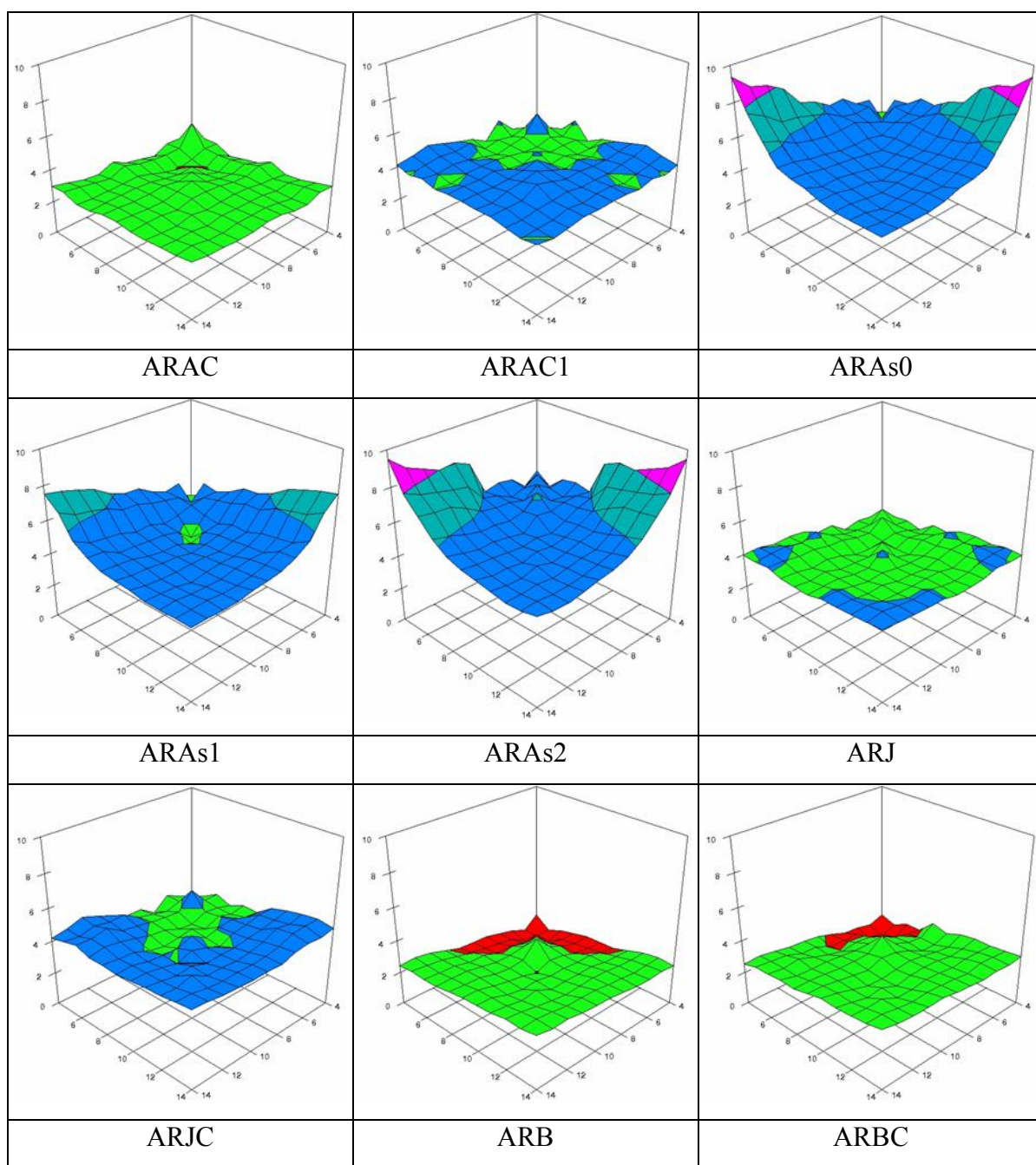
	<p>◀ ARAC Avg(Err)=5.28 StdDev(Err)=2.04</p> <p>ARAC1 ▶ Avg(Err)=4.42 StdDev(Err)=1.45</p>	
	<p>◀ ARWld Avg(Err)=5.28 StdDev(Err)=2.04</p> <p>ARAs0 ▶ Avg(Err)=3.84 StdDev(Err)=1.37</p>	
	<p>◀ ARAs1 Avg(Err)=4.62 StdDev(Err)=1.40</p> <p>ARAs2 ▶ Avg(Err)=4.41 StdDev(Err)=1.37</p>	
	<p>◀ ARJ Avg(Err)=4.74 StdDev(Err)=1.48</p> <p>ARJC ▶ Avg(Err)=4.47 StdDev(Err)=1.73</p>	
	<p>◀ ARB Avg(Err)=4.84 StdDev(Err)=1.84</p> <p>ARBC ▶ Avg(Err)=3.98 StdDev(Err)=1.63</p>	
<p>Erori experimentale pentru $m=n=30$ și $0 \leq X \leq m$, $0 \leq Y \leq n$</p>		

► Următorul tabel conține dependența erorii în punctul central ($X=Y$) pentru $4 \leq m=n \leq 204$:

	<p>ARAC Media erorii experimentale: 5.14 Deviația standard a erorii experimentale: 0.60</p>
	<p>ARAC1 Media erorii experimentale: 5.53 Deviația standard a erorii experimentale: 1.07</p>
	<p>ARWld Media erorii experimentale: 5.67 Deviația standard a erorii experimentale: 1.61</p>
	<p>ARAs0 Media erorii experimentale: 5.31 Deviația standard a erorii experimentale: 0.60</p>
	<p>ARAs1 Media erorii experimentale: 5.29 Deviația standard a erorii experimentale: 0.61</p>
	<p>ARAs2 Media erorii experimentale: 5.38 Deviația standard a erorii experimentale: 0.71</p>
	<p>ARJ Media erorii experimentale: 4.76 Deviația standard a erorii experimentale: 0.65</p>
	<p>ARJC Media erorii experimentale: 4.98 Deviația standard a erorii experimentale: 0.61</p>
	<p>ARB Media erorii experimentale: 4.73 Deviația standard a erorii experimentale: 0.64</p>
	<p>ARBC Media erorii experimentale: 4.95 Deviația standard a erorii experimentale: 0.61</p>

ET46/2006 – Et. Finală/2007 – Lucrare in extenso

► Tabela următoare prezintă eroarea experimentală în domeniul $4 \leq m, n \leq 14$ (metoda ARWld a fost exclusă, obținând cele mai slabe performanțe până în acest moment):



► Tabelul de mai jos dă valorile statistice pentru tabelul anterior:

Metoda	ARAC	ARAC1	ARAs0	ARAs1	ARAs2	ARJ	ARJC	ARB	ARBC
AvgErr	2.79	2.27	5.43	4.89	5.95	3.72	4.18	2.29	2.61
DevI	2.58	2.58	2.58	2.41	2.73	2.34	2.23	3.19	3.00

► Tabela următoare cumulează performanțele metodelor raportat la diferența absolută între media erorii experimentale și eroarea impusă ($|5-AvErr|$, coloana 5AE) și deviația standard a

ET46/2006 – Et. Finală/2007 – Lucrare in extenso

erorii experimentale (StdDev, coloana SD), pentru cele 5 situații investigate ($m=n=10$, $m=n=20$, $m=n=30$, $X=Y$ & $4 \leq m=n \leq 204$, $4 \leq m,n \leq 14$):

Domeniu	m=n=10		m=n=20		m=n=30		X=Y		4 ≤ m,n ≤ 14	
Metoda	5AE	SD	5AE	SD	5AE	SD	5AE	SD	5AE	DevI
ARAC	1.9	1.36	1.25	1.38	0.28	2.04	0.14	0.6	2.21	2.58
ARAC1	0.37	1.95	0.65	1.77	0.58	1.45	0.53	1.07	2.73	2.58
ARWld	2	3.6	0.74	2.6	0.28	2.04	0.67	1.61	N/A	N/A
ARAs0	0.76	1.4	0.29	1.45	1.16	1.37	0.31	0.6	0.43	2.58
ARAs1	0.76	1.4	0.64	1.44	0.38	1.4	0.29	0.61	0.11	2.41
ARAs2	0.13	1.99	0.05	1.68	0.59	1.37	0.38	0.71	0.95	2.73
ARJ	1.22	1.95	0.56	1.72	0.26	1.48	0.24	0.65	1.28	2.34
ARJC	0.71	2.2	0.28	1.72	0.53	1.73	0.02	0.61	0.82	2.23
ARB	2.57	1.58	1.4	1.61	0.16	1.84	0.27	0.64	2.71	3.19
ARBC	2.47	1.59	1.13	1.65	1.02	1.63	0.05	0.61	2.39	3.00
“Best of”	0	0	0	0	0	0	0	0	0	0

► Tabela următoare clasifică metodele în funcție de performanță raportat la diferența absolută între media erorii experimentale și eroarea impusă și deviația standard a erorii experimentale, pentru cele 5 situații investigate ($m=n=10$, $m=n=20$, $m=n=30$, $X=Y$ & $4 \leq m=n \leq 204$, $4 \leq m,n \leq 14$), și obține scorul global al acestora:

Metoda	1	2	3	4	5	6	7	8	9	10	Σ
ARAC	7	1	9	1	3.5	9.5	3	1.5	6	5	46.5
ARAC1	2	6.5	6	9	7	4	9	9	9	5	66.5
ARWld	8	10	7	10	3.5	9.5	10	10	10	10	88
ARAs0	4.5	2.5	3	3	10	1.5	7	1.5	2	5	40
ARAs1	4.5	2.5	5	2	5	3	6	4	1	3	36
ARAs2	1	8	1	6	8	1.5	8	8	4	7	52.5
ARJ	6	6.5	4	7.5	2	5	4	7	5	2	49
ARJC	3	9	2	7.5	6	7	1	4	3	1	43.5
ARB	10	4	10	4	1	8	5	6	8	9	65
ARBC	9	5	8	5	9	6	2	4	7	8	63
“Best of”	1	1	1	1	1	1	1	1	1	1	min

► Din tabelul de mai sus rezultă că metode optime sunt ARAs1 (36), ARAs0 (40), ARJC (43.5), ARAC (46.5), și ARJ (49), cea mai bună optimizare fiind ARAs1 (36).

► Metodele ARAC, ARAC1, ARAs2, ARAs1, ARAs0, și ARWld (pentru comparație) au fost selectate pentru investigarea pe eșantion generat aleatoriu. Au fost astfel generate 100 de valori (X,m,Y,n) din domeniul 4..1000, rezultatele fiind redată în tabelul următor:

	<p>ARWld $5\text{-AvgErr} = 0.48$ DevI = 0.72</p>
	<p>ARAC $5\text{-AvgErr} = 0.18$ DevI = 0.80</p>
	<p>ARAC1 $5\text{-AvgErr} = 0.37$ DevI = 0.80</p>
	<p>ARAs0 $5\text{-AvgErr} = 0.42$ DevI = 0.87</p>
	<p>ARAs1 $5\text{-AvgErr} = 0.37$ DevI = 0.95</p>
	<p>ARAs2 $5\text{-AvgErr} = 0.43$ DevI = 0.83</p>

ET46/2006 – Et. Finală/2007 – Lucrare in extenso

► Clasificarea metodelor în funcție de performanța obținută pe eșantionul aleator este redată în tabelul următor:

Metoda	$ 5-AvgErr $	DevI	Σ
ARWld	6	1	7
ARAC	1	2.5	3.5
ARAC1	2.5	2.5	5
ARAs0	4	5	9
ARAs1	2.5	6	8.5
ARAs2	5	4	9
“Best of”	1	1	min

► Concluzia ce rezultă din tabelul de mai sus este că metoda ARAC este cea mai bună optimizare pentru funcția AR.

Etapa 3, Activitatea 2. Proiectarea și implementarea modului de calcul al intervalelor de încredere și optimizarea intervalelor pentru $1/|Y/n - X/m|$

► Pornind de la expresia de definiție a funcției de probabilitate a distribuției normale standard:

$$P(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$

► Considerând α valoarea nivelului de semnificație (α este o probabilitate $0 \leq \alpha \leq 1$) a erorii impuse (uzual valoarea α se alege $0.05=5\%$; alte valori folosite: $0.1=10\%$, $0.01=1\%$), valoarea z se definește ca acel număr care verifică ecuația integrală:

$$\frac{\alpha}{2} = \int_z^{\infty} P(x)dx$$

► Următorul tabel redă valorile lui z pentru câteva valori uzuale ale lui α :

α	z
0.5	0.6745
0.1	1.645
0.05	1.96
0.01	2.576

► Funcția matematică $|Y/n - X/m|$ este folosită în studiile medicale bazate pe tabela de contingență:

2x2	Caz	Test	Test
Succes	a	c	a+c
Eșec	b	d	b+d
Total	a+b	c+d	a+b+c+d

pentru studiu asupra relației între eșantionul caz și eșantionul test, denumirea consacrată în limba engleză pentru acest studiu fiind *number needed to treat/harm* (numărul necesar pentru a trata/dăuna) și este folosită în cazul studiilor de terapie.

► Notând cu CI_{NN} intervalul de încredere pentru numărul necesar pentru a trata/dăuna (folosită în cazul studiilor de terapie), următoarea relație este relație de transformare de la formulele de definiție a acestui parametru medical la funcția matematică $1/|Y/n - X/m|$:

$$NNT = \frac{1}{\left| \frac{a}{a+b} - \frac{c}{c+d} \right|} = \frac{1}{\left| \frac{Y}{n} - \frac{X}{m} \right|}$$

► Astfel, din punct de vedere matematic, acest parametru medical se exprimă prin intermediul funcției matematice, pe care în continuare o vom numi NN:

$$NN = NN(X, m, Y, n) = 1 / \left| \frac{X}{m} - \frac{Y}{n} \right|$$

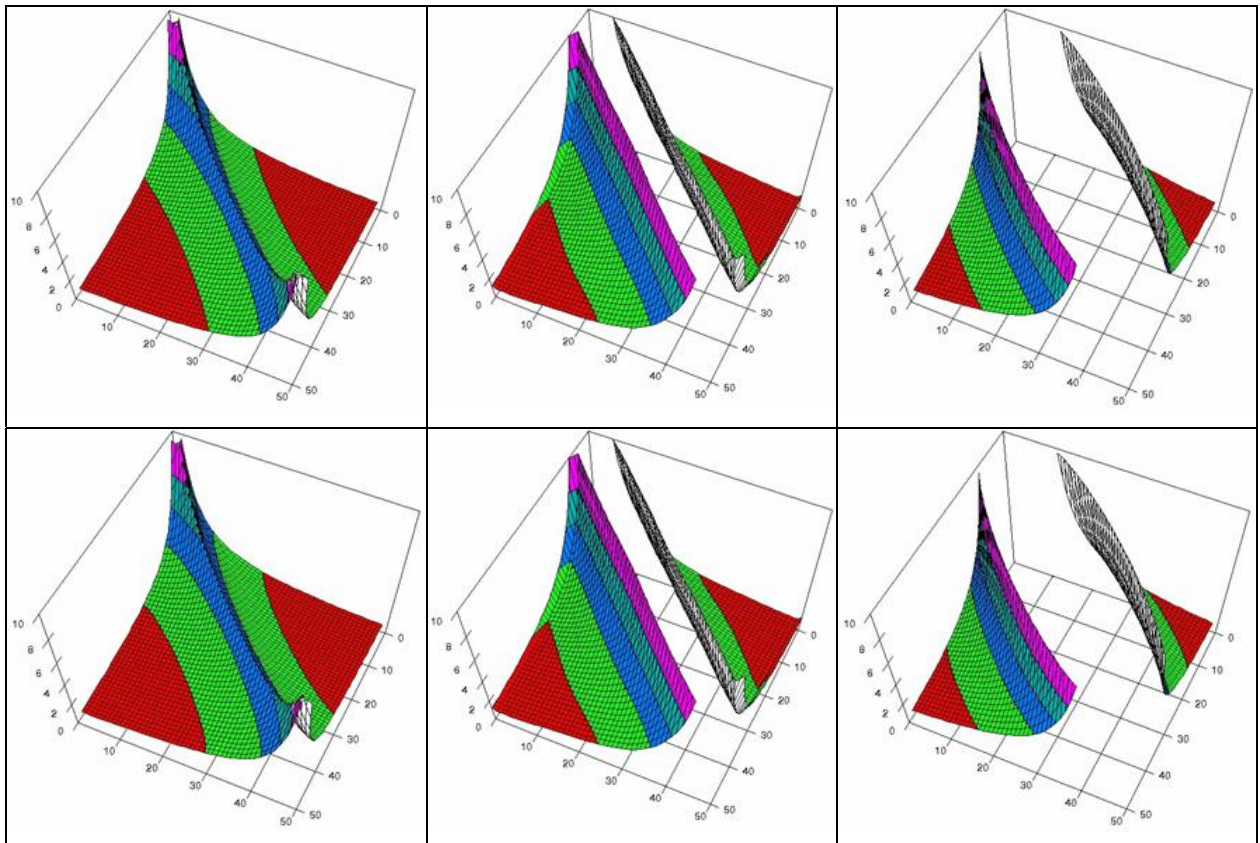
► Următoarele expresii au fost obținute pentru exprimarea intervalului de încredere al funcției NN:

N	Metodă	Expresie de calcul
1	NNWald(X,m,Y,n,z)	1/ARWld(X,m,Y,n,z)
2	NNAC(X,m,Y,n,z)	1/ARAC(X,m,Y,n,z)
3	NNAs0(X,m,Y,n,z)	1/ARAs0(X,m,Y,n,z)
4	NNJ(X,m,Y,n,a)	1/ARJ(X,m,Y,n,a)

► Notă: Expresile pentru $CI_{NN} = (CI_{NNL}, CI_{NNU})$ folosesc formulele de definiție ale intervalelor de încredere pentru funcția AR. Relația prin care acestea rezultă este:

$$CI_{NN} = (CI_{NNL}, CI_{NNU}) = (1/CI_{ARU}, 1/CI_{ARL})$$

► Următorul tabel conține reprezentările grafice ale intervalelor de încredere (stânga - limita inferioară, centru - valoarea expresiei creșterii beneficiului, dreapta - limita superioară) obținute cu metodele NNWald (prima linie din tabel) și NNJ (a doua linie din tabel):



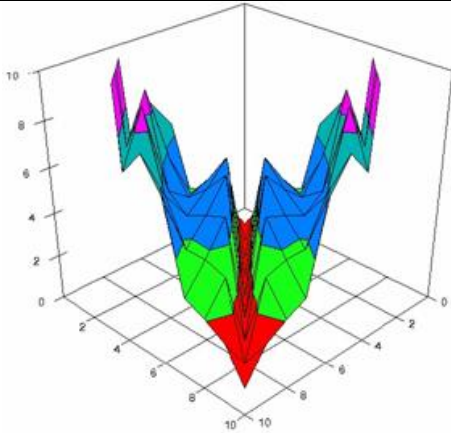
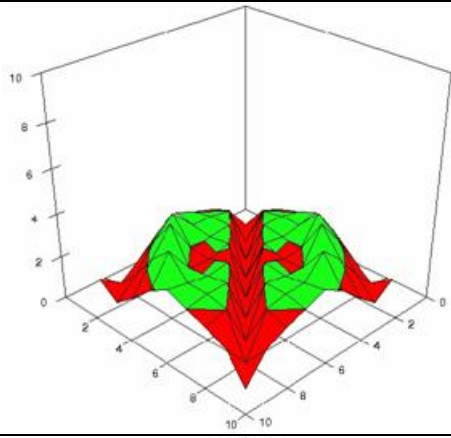
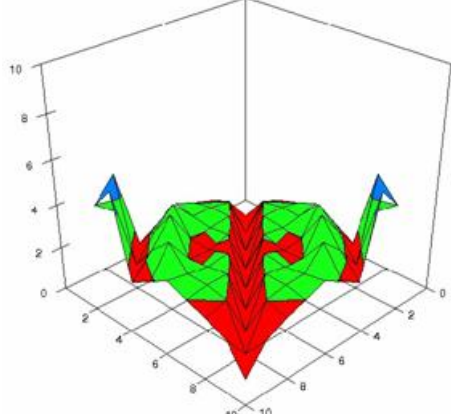
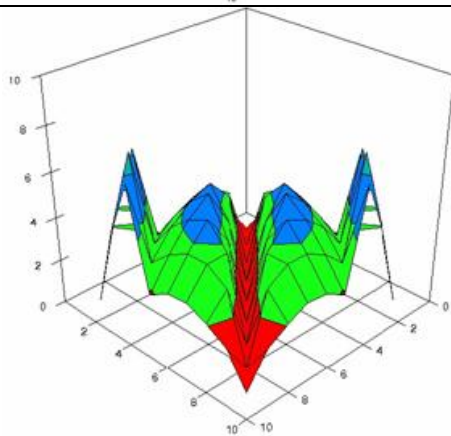
► Următoarele metode de evaluare au fost folosite pentru compararea rezultatelor metodelor de calcul al intervalului de încredere pentru funcția NN:

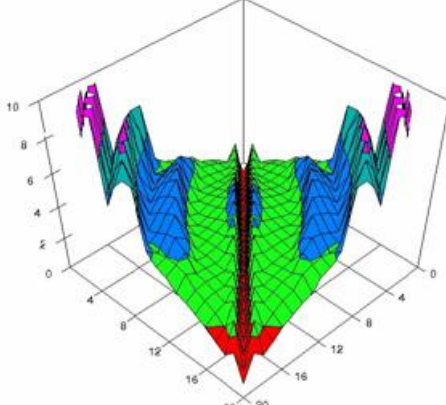
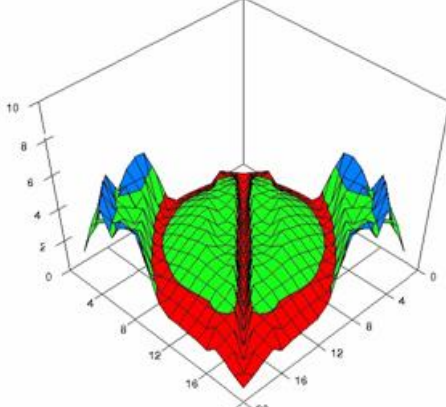
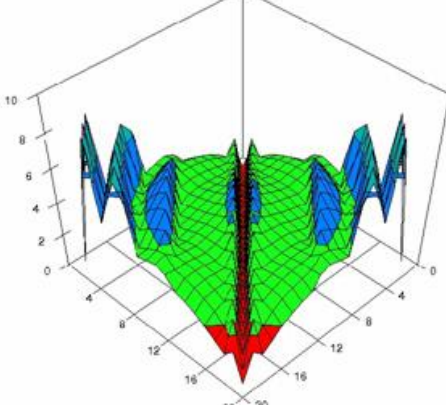
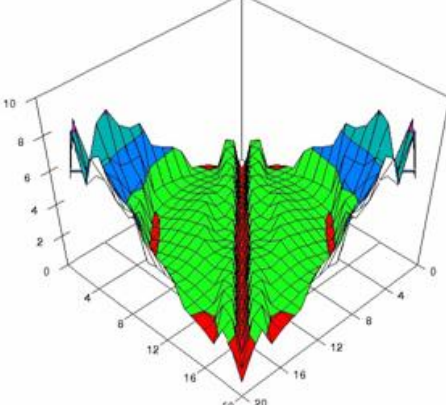
Nume	Metoda	Formula
AvErr	Av(Err)	$\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} \text{Err}(X, Y, m, n)}{(m-1)(n-1)}$
StdDev	StdDev(Err)	$\left(\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} (\text{Err}(X, Y, m, n) - \text{AvErr})^2}{(m-1)(n-1) - 1} \right)^{1/2}$
AvAD	AvAD(Err)	$\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} \text{Err}(X, Y, m, n) - \text{AvErr} }{(m-1)(n-1) - 1}$
AvADI	AvADI(Err)	$\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} \text{Err}(X, Y, m, n) - 100 \cdot \alpha }{(m-1)(n-1)}$
DevI	DevI(Err)	$\left(\frac{\sum_{X=1}^{m-1} \sum_{Y=1}^{n-1} (\text{Err}(X, Y, m, n) - 100 \cdot \alpha)^2}{(m-1)(n-1)} \right)^{1/2}$

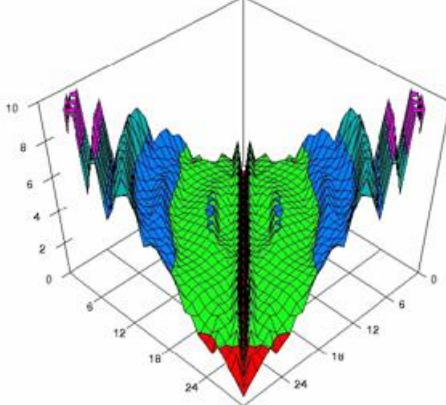
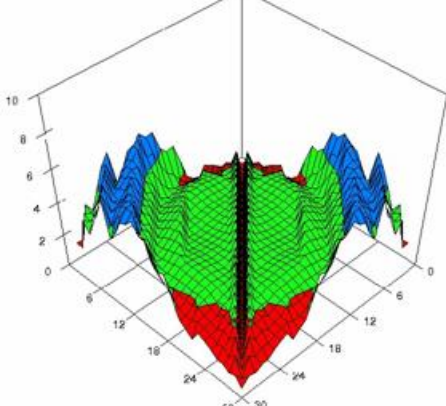
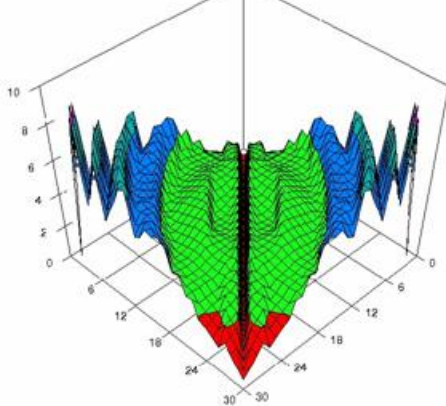
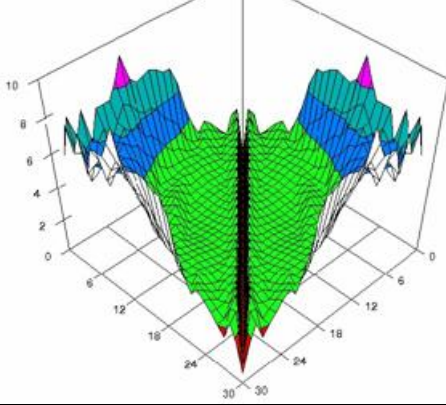
► Funcția $\text{Err}(\cdot, \cdot, \cdot, \cdot)$ cumulează erori experimentale folosind distribuția binomială univariată pentru generarea distribuției binomiale bivariate. Notând capetele intervalului de încredere pentru funcția ER prin (ERL, ERU) - și aici nota bene ERL și ERU sunt funcții date de una din metodele de calcul al intervalului de încredere (1-6 în tabelul de mai sus), funcția $\text{Err}(X, m, Y, n)$ se calculează din:

$$\begin{aligned}
 d\text{Bin}(m, X, \Xi) &:= \frac{m!}{\Xi!(m-\Xi)!} \cdot \left(\frac{X}{m}\right)^\Xi \cdot \left(1 - \frac{X}{m}\right)^{m-\Xi} \\
 d\text{Bin}(n, Y, \Psi) &:= \frac{n!}{\Psi!(n-\Psi)!} \cdot \left(\frac{Y}{n}\right)^\Psi \cdot \left(1 - \frac{Y}{n}\right)^{n-\Psi} \\
 \sum_{\substack{\text{ERL}(\Xi, \Psi, m, n) \\ > \text{ER}(X, Y, m, n)}} d\text{Bin}(m, X, \Xi) \cdot d\text{Bin}(n, Y, \Psi) &+ \sum_{\substack{\text{ERU}(\Xi, \Psi, m, n) \\ < \text{ER}(X, Y, m, n)}} d\text{Bin}(m, X, \Xi) \cdot d\text{Bin}(n, Y, \Psi) \\
 \hline
 \sum_{\Xi=1}^{m-1} \sum_{\Psi=1}^{n-1} d\text{Bin}(m, X, \Xi) \cdot d\text{Bin}(n, Y, \Psi) &
 \end{aligned}$$

► Următoarele tabele prezintă descriptiv rezultatele obținute (în termeni de eroare experimentală) cu fiecare dintre cele 4 metode pentru $m=n=10, 20, 30$.

	<p>◀NNWald Avg(Err)=5.16 StdDev(Err)=3.89</p>
	<p>◀NNAC Avg(Err)=1.69 StdDev(Err)=0.90</p>
	<p>◀NNAs0 Avg(Err)=2.97 StdDev(Err)=1.71</p>
	<p>◀NNJ Avg(Err)=2.02 StdDev(Err)=1.11</p>
<p>Erori experimentale pentru $m=n=10$ și $0 \leq X \leq m$, $0 \leq Y \leq n$</p>	

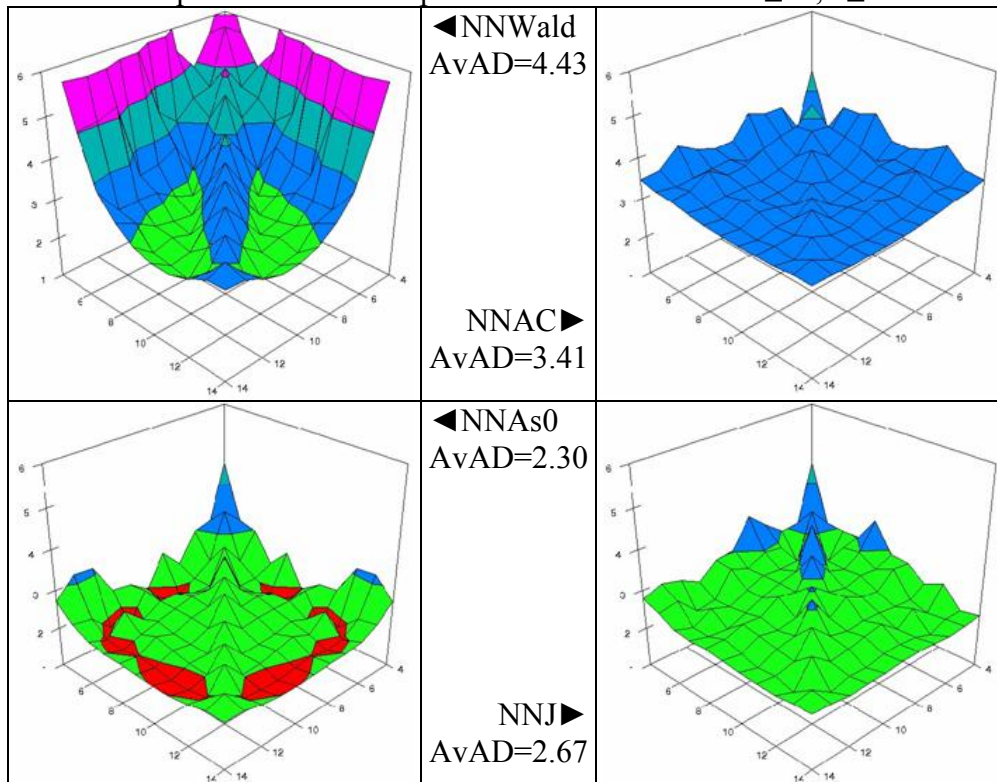
	<p>◀NNWald Avg(Err)=4.39 StdDev(Err)=2.77</p>
	<p>◀NNAC Avg(Err)=2.40 StdDev(Err)=1.20</p>
	<p>◀NNAs0 Avg(Err)=3.55 StdDev(Err)=1.53</p>
	<p>◀NNJ Avg(Err)=3.19 StdDev(Err)=1.69</p>
<p>Erori experimentale pentru $m=n=20$ și $0 \leq X \leq m$, $0 \leq Y \leq n$</p>	

	<p>◀NNWald Avg(Err)=4.24 StdDev(Err)=2.27</p>
	<p>◀NNAC Avg(Err)=2.71 StdDev(Err)=1.25</p>
	<p>◀NNAs0 Avg(Err)=3.54 StdDev(Err)=1.52</p>
	<p>◀NNJ Avg(Err)=3.57 StdDev(Err)=1.80</p>
<p>Erori experimentale pentru $m=n=30$ și $0 \leq X \leq m$, $0 \leq Y \leq n$</p>	

► Următorul tabel conține dependența erorii în punctul ($X=3m/4$, $Y=n/4$) pentru $4 \leq m=n \leq 204$:

	<p>NNWald Media erorii experimentale: 5.57 Deviația standard a erorii experimentale: 0.91</p>
	<p>NNAC Media erorii experimentale: 4.55 Deviația standard a erorii experimentale: 1.20</p>
	<p>NNAs0 Media erorii experimentale: 5.04 Deviația standard a erorii experimentale: 0.90</p>
	<p>NNJ Media erorii experimentale: 4.47 Deviația standard a erorii experimentale: 1.12</p>

► Tabela următoare prezintă eroarea experimentală în domeniul $4 \leq m,n \leq 14$:



ET46/2006 – Et. Finală/2007 – Lucrare in extenso

► Tabela următoare cumulează performanțele metodelor de calcul al intervalului de încredere raportat la metodele de evaluare folosite ($m=n=10$; $m=n=20$; $m=n=30$; $X=3m/4$, $Y=n/4$, $4 \leq m = n \leq 204$, $4 \leq m, n \leq 14$):

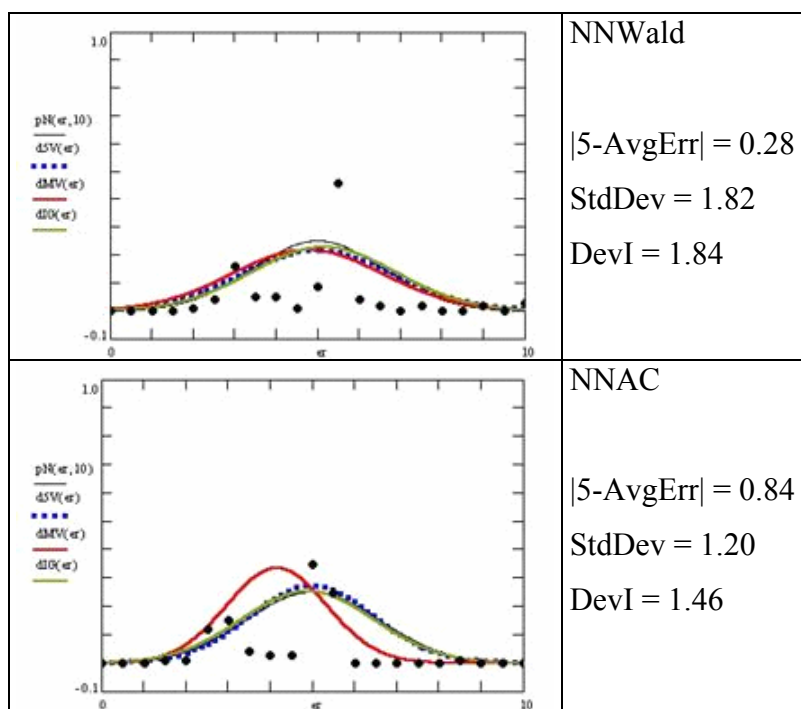
Domeniu	m=n=10		m=n=20		m=n=30		X=Y		4 ≤ m,n ≤ 14
Metoda	5AE	SD	5AE	SD	5AE	SD	5AE	SD	AvAD
NNWald	0.16	3.89	0.61	2.77	0.76	2.27	0.57	0.91	4.43
NNAC	3.31	0.90	2.60	1.20	2.29	1.25	0.45	1.20	3.41
NNAs0	2.03	1.71	1.45	1.53	1.46	1.52	0.04	0.90	2.30
NNJ	2.98	1.11	1.81	1.69	1.43	1.80	0.53	1.12	2.67
“Best of”	0	0	0	0	0	0	0	0	0

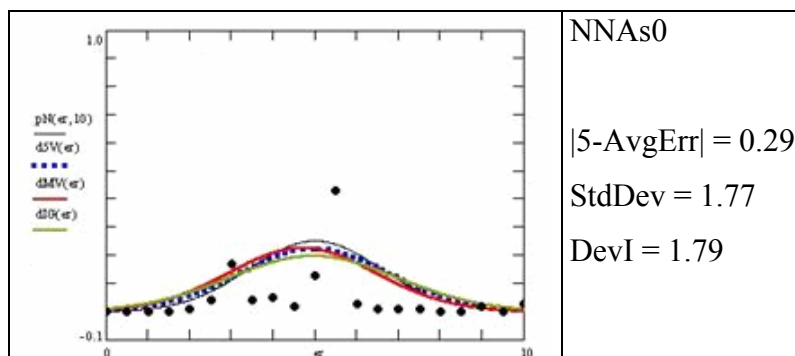
► Tabela următoare clasifică metodele în funcție de performanță raportat la diferența absolută între media erorii experimentale și eroarea impusă și deviația standard a erorii experimentale, pentru cele 5 situații investigate ($m=n=10$, $m=n=20$, $m=n=30$, $X=Y$ & $4 \leq m=n \leq 204$, $4 \leq m, n \leq 14$), și obține scorul global al acestora:

Metoda	1	2	3	4	5	6	7	8	9	Σ
NNWald	1	4	1	4	1	4	4	2	4	25
NNAC	4	1	4	1	4	1	2	4	3	24
NNAs0	2	3	2	2	3	2	1	1	1	17
NNJ	3	2	3	3	2	3	3	3	2	24
“Best of”	1	1	1	1	1	1	1	1	1	min

► Din tabelul de mai sus rezultă că metodă optimă este ARAs0 (17).

► Metodele au fost investigate pe eșantion generat aleatoriu. Au fost astfel generate 100 de valori (X, m, Y, n) din domeniul 4..1000, rezultatele fiind redată în tabelul următor:





► Clasificarea metodelor în funcție de performanța obținută pe eșantionul aleator este redată în tabelul următor:

Metoda	5-AvgErr	StdDev	DevI	Σ
NNWald	1	3	3	7
NNAC	3	1	1	5
NNAs0	2	2	2	6
“Best of”	1	1	1	min

► Concluzia ce rezultă din tabelul de mai sus este că metoda NNAC este cea mai bună optimizare pentru funcția NN (fapt care rezultă pentru întreg domeniul de volum de eșantion de la 4 la 1000) urmată de metoda NNAs0 (metodă care merge mai bine pentru volume de eșantion mici, așa cum rezultă din investigația anterioară), ambele metode reprezentând optimizări față de metoda uzuală NNWald.

Etapa 3, Activitatea 3. Proiectarea și implementarea modului de calcul al intervalelor de încredere și optimizarea intervalelor pentru $|1-X_n/Y/m|$

► Pornind de la expresia de definiție a funcției de probabilitate a distribuției normale standard:

$$P(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$

► Considerând α valoarea nivelului de semnificație (α este o probabilitate $0 \leq \alpha \leq 1$) a erorii impuse (uzual valoarea α se alege $0.05=5\%$; alte valori folosite: $0.1=10\%$, $0.01=1\%$), valoarea z se definește ca acel număr care verifică ecuația integrală:

$$\frac{\alpha}{2} = \int_z^{\infty} P(x)dx$$

► Următorul tabel redă valorile lui z pentru câteva valori uzuale ale lui α :

α	z
0.5	0.6745
0.1	1.645
0.05	1.96
0.01	2.576

► Funcția matematică $|1-X_n/Y/m|$ este folosită în studiile medicale bazate pe tabela de contingență:

2×2	Caz	Test	Test
Succes	a	c	a+c
Eșec	b	d	b+d
Total	a+b	c+d	a+b+c+d

pentru studiu asupra relației între eșantionul caz și eșantionul test, denumirea consacrată în limba engleză pentru acest studiu fiind *relative risk reduction/increase* (reducerea/creșterea riscului relativ) și *benefit increase* (creșterea beneficiului) și este folosită în cazul studiilor de terapie.

► Notând cu CI_{BI} intervalul de încredere pentru reducerea/creșterea riscului relativ (folosită în cazul studiilor de terapie), următoarea relație este relație de transformare de la formulele de definiție a acestui parametru medical la funcția matematică $|1-X_n/Y/m|$:

$$BI = \frac{\left| \frac{a}{a+b} - \frac{c}{c+d} \right|}{\frac{c}{c+d}} = \frac{\left| \frac{X}{m} - \frac{Y}{n} \right|}{\frac{Y}{n}} = \left| \frac{Xn}{Ym} - 1 \right|$$

► Astfel, din punct de vedere matematic, acest parametru medicali se exprimă prin intermediul funcției matematice, pe care în continuare o vom numi BI:

$$BI = BI(X, m, Y, n) = \left| 1 - \frac{Xn}{Ym} \right|$$

► Următoarele expresii au fost obținute pentru exprimarea intervalului de încredere al funcției BI:

N	Metodă	Expresie de calcul
1	BIWaldC0(X,m,Y,n,z)	AD1(LRWald(X,m,Y,n,z))
2	BIWaldC1(X,m,Y,n,z)	AD1(LRWaldC1(X,m,Y,n,z))
3	BIWaldC2(X,m,Y,n,z)	AD1(LRWaldC2(X,m,Y,n,z))
4	BIWaldC3(X,m,Y,n,z)	AD1(LRWaldC3(X,m,Y,n,z))
5	BIAC(X,m,Y,n)	AD1(LRAC(X,m,Y,n,z))

unde funcția AD1 este definită de:

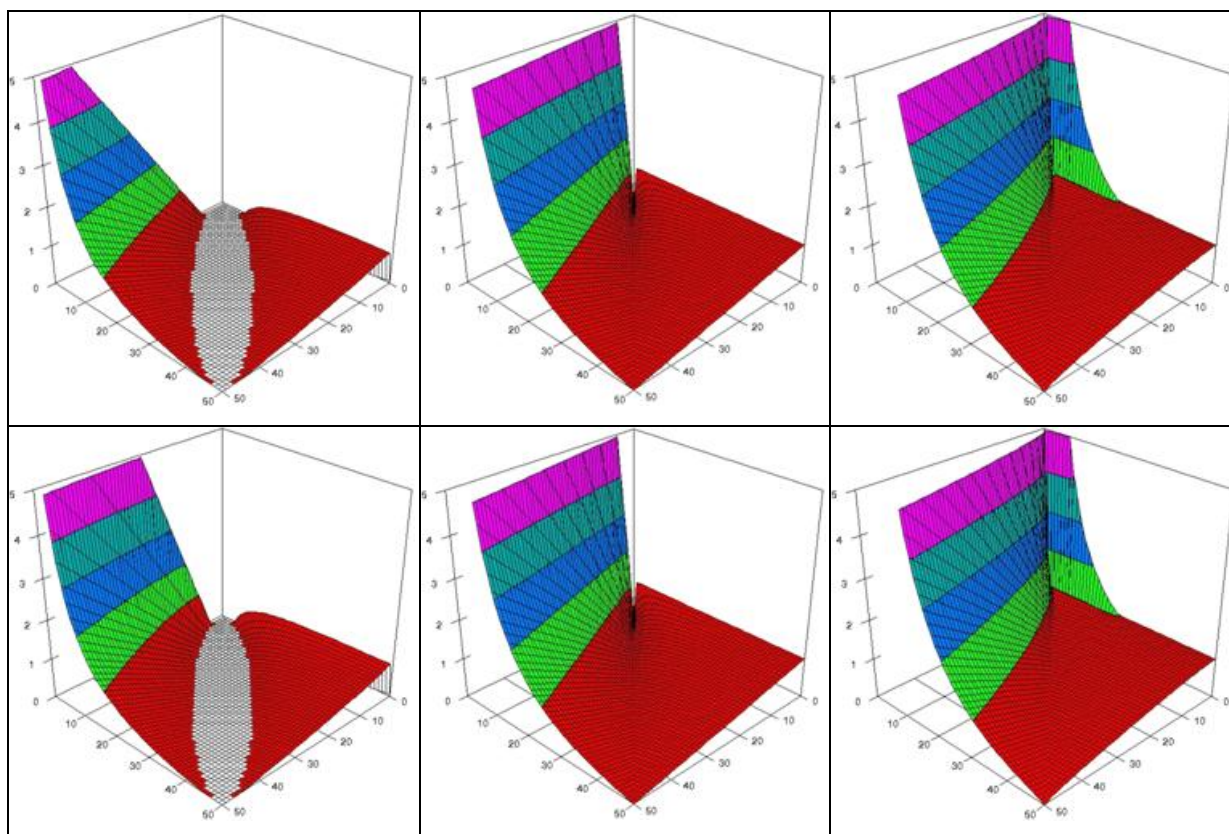
$$AD1((u, v)) = \begin{cases} (0, \max(|1-u|, |1-v|)), & (1-u)(1-v) \leq 0 \\ (\min(|1-u|, |1-v|), \max(|1-u|, |1-v|)), & (1-u)(1-v) > 0 \end{cases}$$

și funcțiile LRWald, LRWaldC1, LRWaldC2, LRWaldC3, și LRAC sunt definite de:

Metodă	Expresie de calcul
LRWald(X,m,Y,n,z)	$\frac{X}{m} \cdot \frac{n}{Y} \cdot \exp\left(\pm z \sqrt{\frac{m-X}{m \cdot X} + \frac{n-Y}{n \cdot Y}}\right)$
LRWaldC1(X,m,Y,n,z)	$\frac{X}{m} \cdot \frac{n}{Y} \cdot \exp\left(\pm z \sqrt{\frac{m-X}{m \left(X + \sqrt[4]{\frac{m-X}{m}} \right)} + \frac{n-Y}{n \cdot \left(Y + \sqrt[4]{\frac{n-Y}{n}} \right)}}\right)$

LRWaldC2(X,m,Y,n,z)	$\frac{X}{m} \cdot \frac{n}{Y} \cdot \exp \left(\pm z \sqrt{\frac{m-X}{m \left(X + \sqrt{1 - \frac{X}{mY}} \right)} + \frac{n-Y}{n \cdot \left(Y + \sqrt{1 - \frac{Y}{nX}} \right)}} \right)$
LRWaldC3(X,m,Y,n,z)	$\frac{X}{m} \cdot \frac{n}{Y} \cdot \exp \left(\pm z \sqrt{\frac{m-X}{m \left(X + \sqrt{1 - \frac{X}{mY}} \right)} + \frac{n-Y}{n \cdot \left(Y + \sqrt{1 - \frac{Y}{nX}} \right)}} \right)$
LRAC(X,m,Y,n,z)	$\text{LRWald} \left(X + \frac{1}{2\sqrt{m}}, m + \frac{1}{\sqrt{m}}, Y + \frac{1}{2\sqrt{n}}, n + \frac{1}{\sqrt{n}}, z \right)$

► Următorul tabel conține reprezentările grafice ale intervalelor de încredere (stânga - limita inferioară, centru - valoarea expresiei creșterii beneficiului, dreapta - limita superioară) obținute cu metodele BIWaldC0 (prima linie din tabel) și BIWaldC1 (a doua linie din tabel):



Evaluarea erorii are loc cu formulele:

÷ Probabilitatea de apariție a lui Ξ în eșantionul de volum m din care s-au extras X succese este:

$$d\text{Bin}(m, X, \Xi) := \frac{m!}{\Xi!(m-\Xi)!} \cdot \left(\frac{X}{m}\right)^\Xi \cdot \left(1 - \frac{X}{m}\right)^{m-\Xi}$$

÷ Probabilitatea de apariție a lui Ψ în eșantionul de volum n din care s-au extras Y succese este:

$$d\text{Bin}(N, Y, \Psi) := \frac{n!}{\Psi!(n-\Psi)!} \cdot \left(\frac{Y}{n}\right)^\Psi \cdot \left(1 - \frac{Y}{n}\right)^{n-\Psi}$$

Fie expresia funcției de cele două variabile binomiale pentru care se calculează intervalul de încredere F , dată de:

$$F(X, m, Y, n)$$

iar intervalul de încredere pentru funcția F dat de (unde P parametrizează metoda de calcul a intervalului de încredere):

$$\text{CIP}(X, m, Y, n) = (\text{CIPL}(X, m, Y, n), \text{CIPU}(X, m, Y, n))$$

De notat că între funcțiile CIPL , CIPU și F există relația de ordine:

$$\text{CIPL}(X, m, Y, n) \leq F(X, m, Y, n) \leq \text{CIPU}(X, m, Y, n)$$

Eroarea experimentală pentru funcția F și intervalul de încredere CI calculat prin metoda P se cumulează cu formula:

$$\varepsilon(P, X, M, Y, N) = 1 - \frac{\sum_{\text{CIPL}(\Xi, \Psi, m, n) \leq F(X, Y, m, n) \leq \text{CIPU}(\Xi, \Psi, m, n)} d\text{Bin}(m, X, \Xi) \cdot d\text{Bin}(n, Y, \Psi)}{\sum_{\Xi=0}^m \sum_{\Psi=0}^n d\text{Bin}(m, X, \Xi) \cdot d\text{Bin}(n, Y, \Psi)}$$

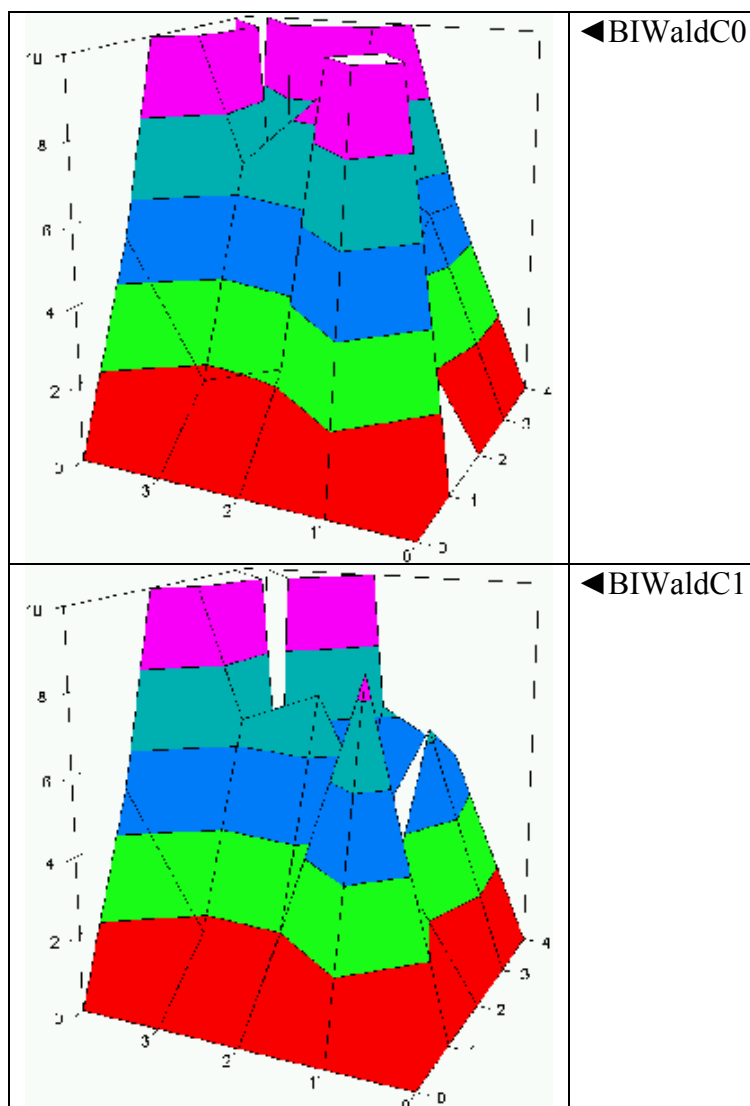
► Următoarele metode de evaluare au fost folosite pentru compararea rezultatelor metodelor de calcul al intervalului de încredere pentru funcția BI :

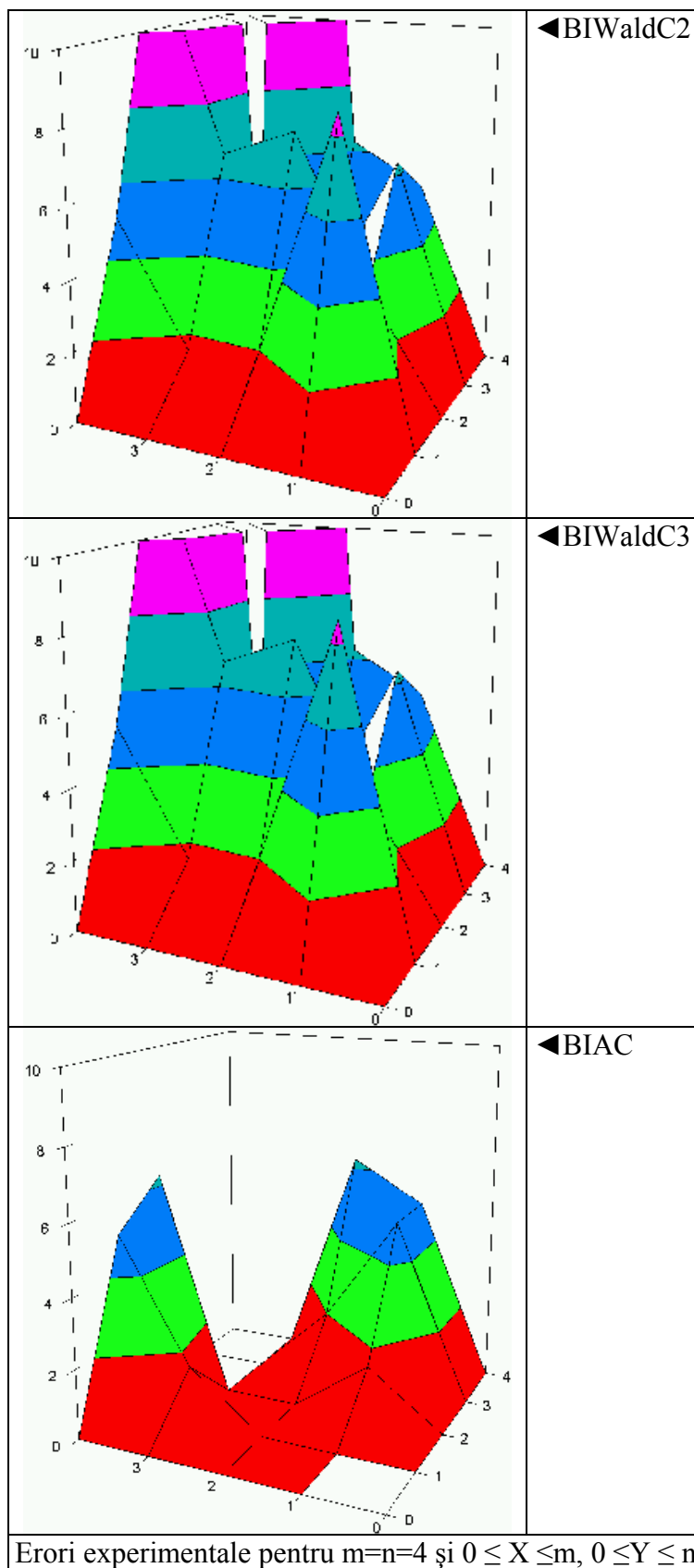
Nr	Metoda	Formula
0 (A=0), 1(A=1)	IAvOE(P, A)	$\left \alpha - \sum_{X=A}^{m-A} \sum_{Y=A}^{n-A} \frac{\varepsilon(P, X, m, Y, n)}{(m+1-2A)(n+1-2A)} \right $
2 (A=0), 3(A=1)	StDOE(P, A)	$\sqrt{\sum_{X=A}^{m-A} \sum_{Y=A}^{n-A} \frac{(\varepsilon(P, X, m, Y, n) - \text{AvgOEA}(P))^2}{(m-2A)(n-2A)}}$
4 (A=0), 5(A=1)	SiDOE(P, A)	$\sqrt{\sum_{X=A}^{m-A} \sum_{Y=A}^{n-A} \frac{(\varepsilon(P, X, m, Y, n) - 100\alpha)^2}{(m+1-2A)(n+1-2A)}}$
6 (A=0), 7(A=1)	AvADA(P, A)	$\sum_{X=A}^{m-A} \sum_{Y=A}^{n-A} \frac{ \varepsilon(P, X, m, Y, n) - \text{AvgOEA}(P) }{(m-2A)(n-2A)}$

8 (A=0), 9(A=1)	AvADS(P, A)	$\sum_{X=A}^{m-A} \sum_{Y=A}^{n-A} \frac{ \varepsilon(P, X, m, Y, n) - 100\alpha }{(m+1-2A)(n+1-2A)}$
10 (A=0), 11(A=1)	S8DOE(P, A)	$\sqrt[8]{\sum_{X=A}^{m-A} \sum_{Y=A}^{n-A} \frac{(\varepsilon(P, X, m, Y, n) - 100\alpha)^8}{(m+1-2A)(n+1-2A)}}$

unde $P \in \{\text{BIWaldC0, BIWaldC1, BIWaldC2, BIWaldC3, BIAC}(X,m,Y,n,z)\}$

► Următorul tabel prezintă descriptiv rezultatele obținute (în termeni de eroare experimentală) cu fiecare dintre cele 5 metode pentru $m=n=4$ și $\alpha=5\%$.

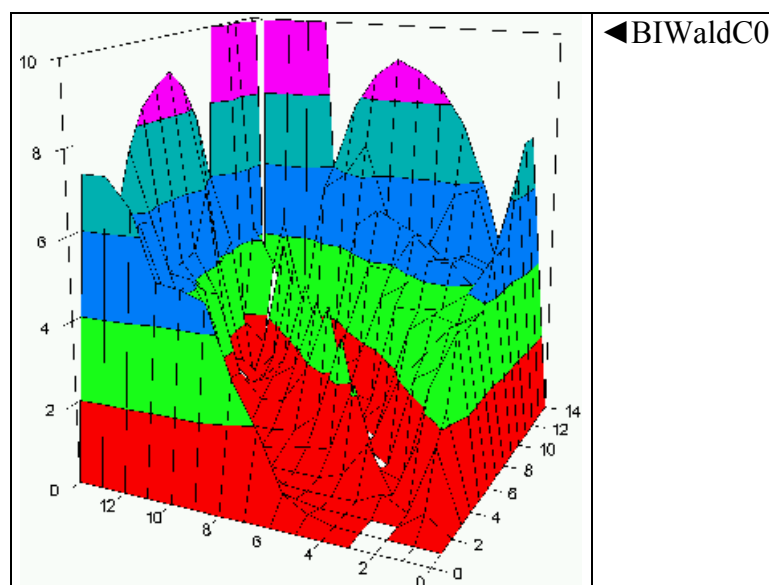


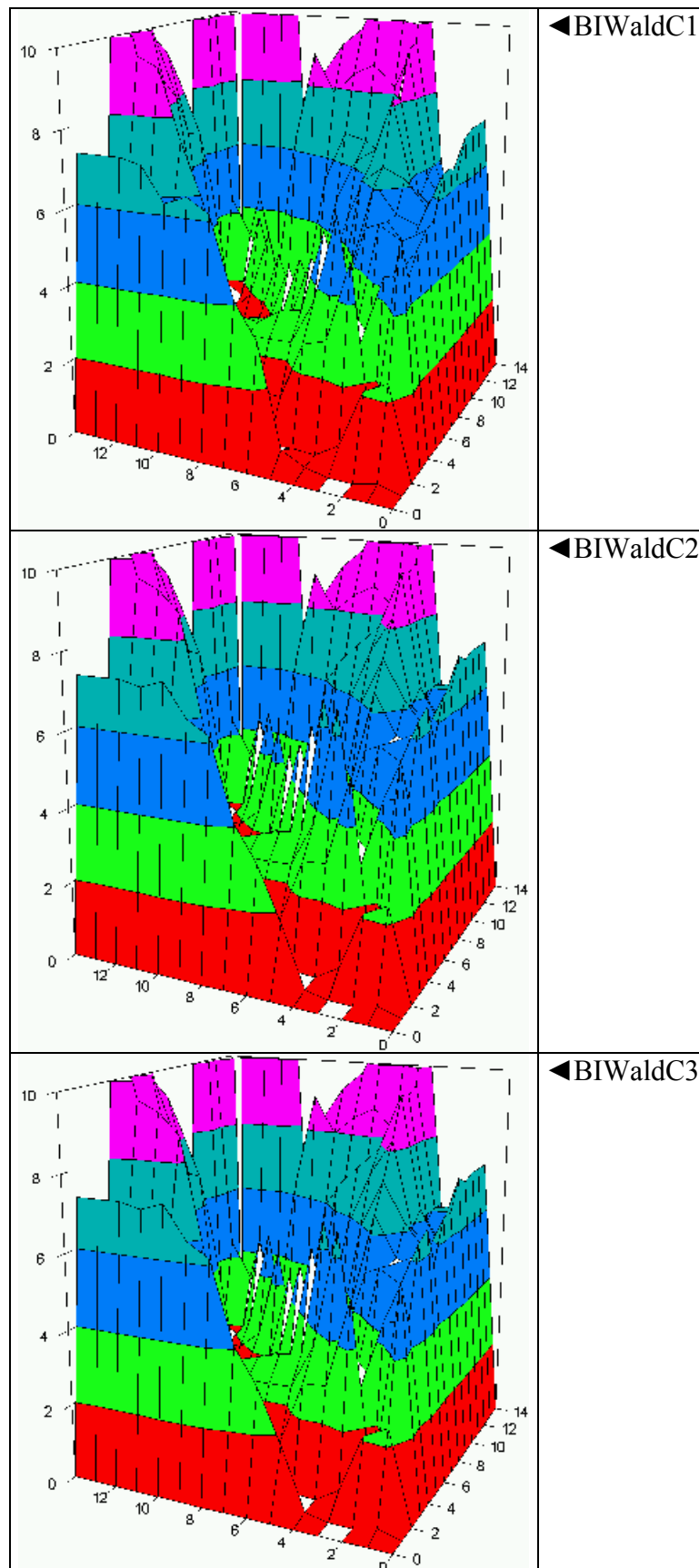


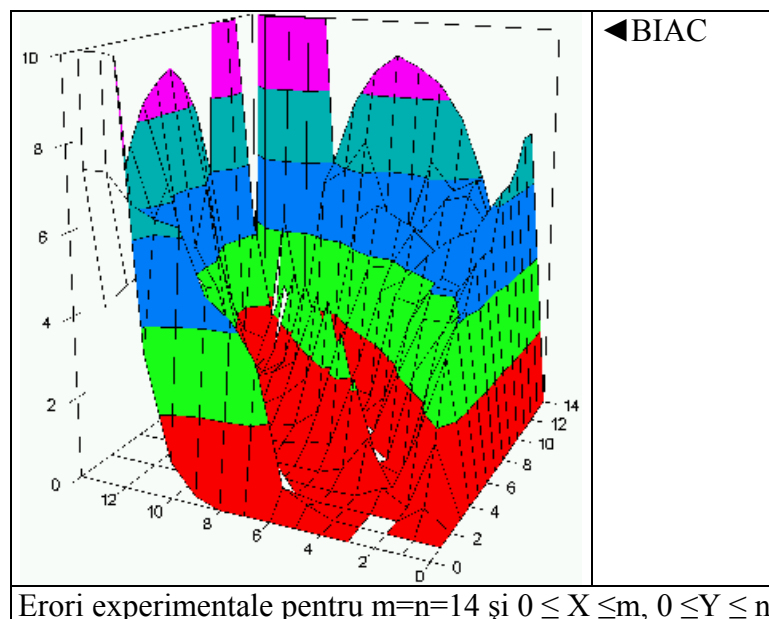
► Următorul tabel centralizează rezultatele obținute prin evaluarea cu fiecare din cele 12 metode pentru $m=n=4$ și $0 \leq X \leq m$, $0 \leq Y \leq n$ și $\alpha = 5\%$:

	BIAC	BIWaldC0	BIWaldC1	BIWaldC2	BIWaldC3	“Best of”
IAvOE0	3.678	1.065	0.362	0.362	0.362	0
IAvOE1	3.196	1.330	0.410	0.410	0.410	0
StDOE0	2.258	9.097	8.663	8.663	8.663	0
StDOE1	2.361	8.877	8.628	8.628	8.628	0
SiDOE0	4.292	8.977	8.496	8.496	8.496	0
SiDOE1	3.915	8.603	8.271	8.271	8.271	0
AvADA0	1.870	6.607	5.721	5.721	5.721	0
AvADA1	2.045	5.943	5.315	5.315	5.315	0
AvADS0	3.908	6.004	5.378	5.378	5.378	0
AvADS1	3.442	5.080	4.752	4.752	4.752	0
S8ADS0	4.767	19.692	19.548	19.548	19.548	0
S8ADS1	4.592	19.791	19.498	19.498	19.498	0

► Următoarul tabel prezintă descriptiv rezultatele obținute (în termeni de eroare experimentală) cu fiecare dintre cele 5 metode pentru $m=n=14$ și $\alpha=5\%$.



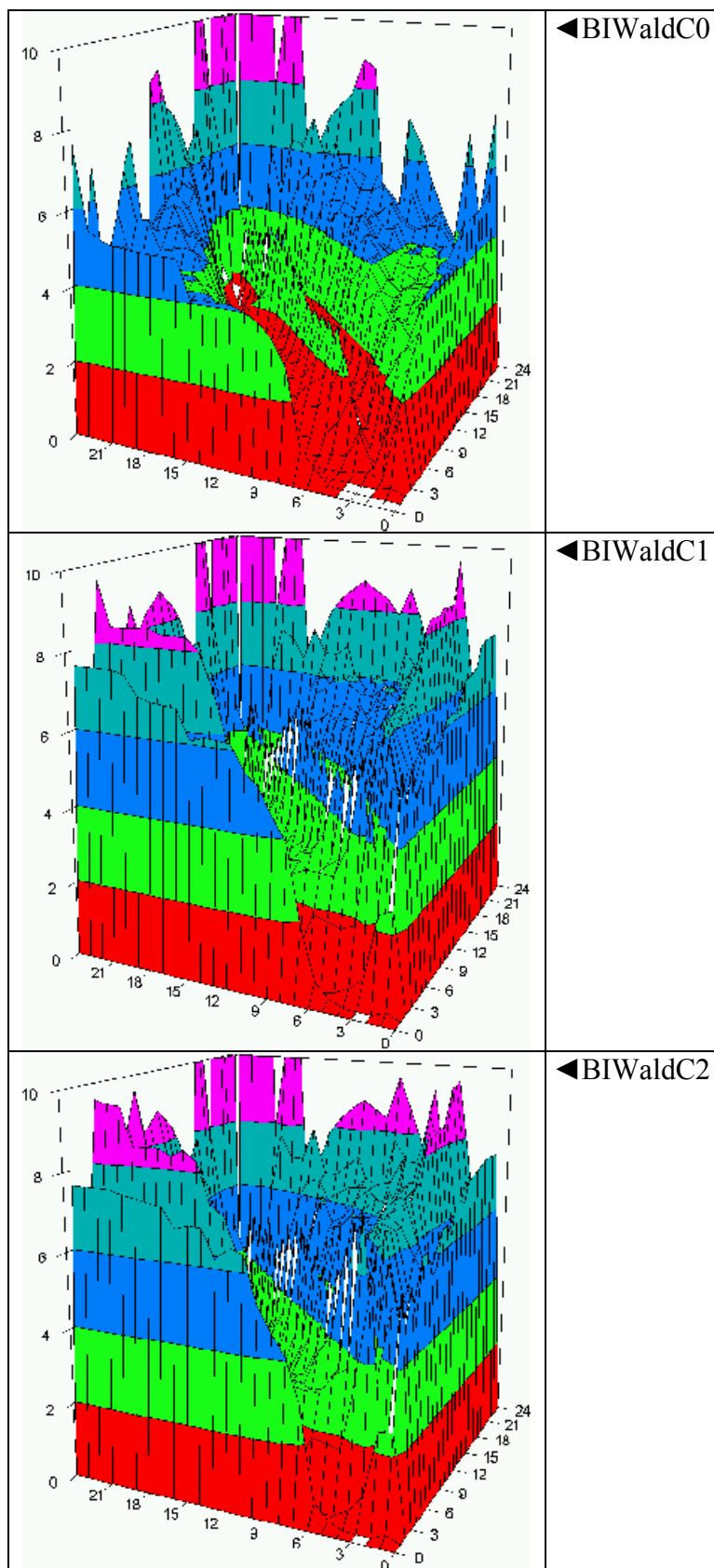


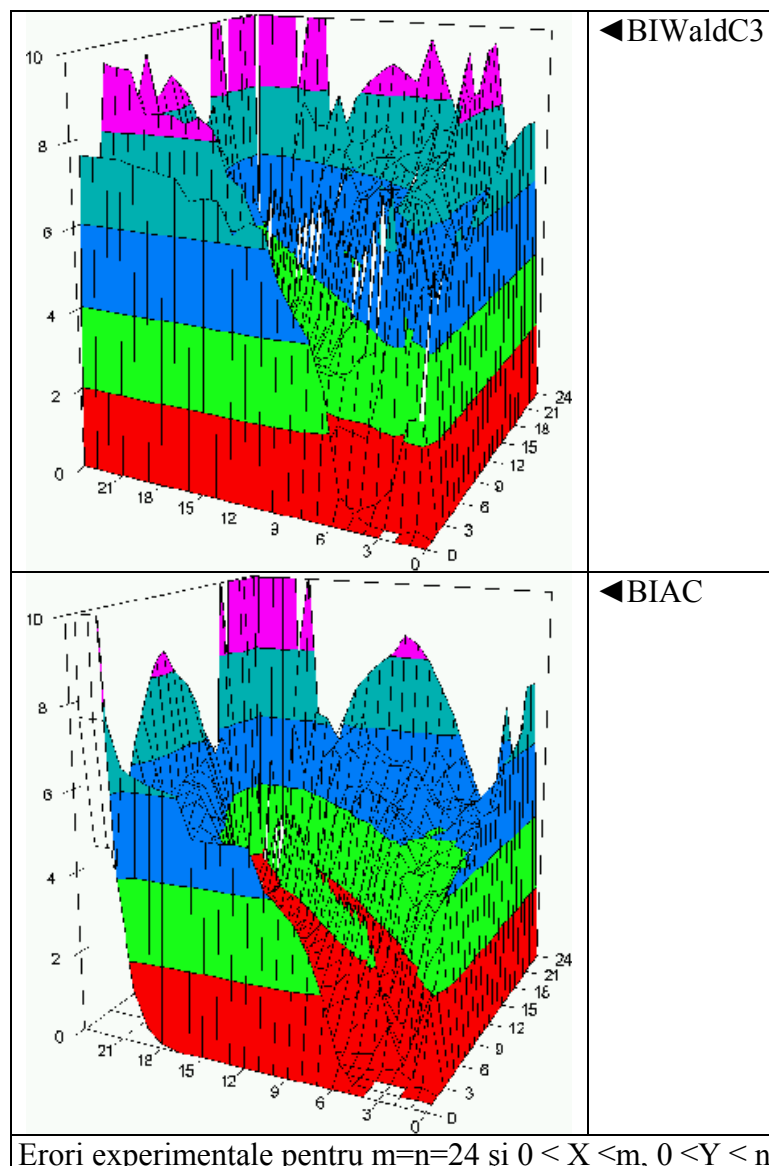


► Următorul tabel centralizează rezultatele obținute prin evaluarea cu fiecare din cele 12 metode pentru $m=n=14$ și $0 \leq X \leq m$, $0 \leq Y \leq n$ și $\alpha = 5\%$:

	BIAC	BIWaldC0	BIWaldC1	BIWaldC2	BIWaldC3	“Best of”
IAvOE0	1.216	1.847	0.531	0.301	0.301	0
IAvOE1	1.568	2.020	0.545	0.311	0.311	0
StDOE0	7.737	4.219	4.398	4.402	4.402	0
StDOE1	4.360	3.546	3.767	3.770	3.770	0
SiDOE0	7.815	4.597	4.421	4.403	4.403	0
SiDOE1	4.622	4.072	3.796	3.772	3.772	0
AvADA0	3.249	2.611	2.958	3.011	3.011	0
AvADA1	2.647	2.297	2.638	2.687	2.687	0
AvADS0	3.662	3.289	3.002	3.012	3.012	0
AvADS1	3.079	2.997	2.656	2.681	2.681	0
S8ADS0	48.274	16.846	16.847	16.847	16.847	0
S8ADS1	17.298	15.863	15.863	15.863	15.863	0

► Următoarul tabel prezintă descriptiv rezultatele obținute (în termeni de eroare experimentală) cu fiecare dintre cele 5 metode pentru $m=n=24$ și $\alpha=5\%$.





► Următorul tabel centralizează rezultatele obținute prin evaluarea cu fiecare din cele 12 metode pentru $m=n=24$ și $0 \leq X \leq m$, $0 \leq Y \leq n$ și $\alpha = 5\%$:

	BIAC	BIWaldC0	BIWaldC1	BIWaldC2	BIWaldC3	“Best of”
IAvOE0	1.455	1.766	0.568	0.218	0.218	0
IAvOE1	1.636	1.823	0.587	0.216	0.216	0
StDOE0	5.057	2.951	3.149	3.224	3.224	0
StDOE1	2.900	2.474	2.712	2.794	2.794	0
SiDOE0	5.258	3.437	3.197	3.229	3.229	0
SiDOE1	3.328	3.071	2.772	2.800	2.800	0
AvADA0	2.175	1.884	2.213	2.322	2.322	0
AvADA1	1.859	1.701	2.017	2.120	2.120	0
AvADS0	2.606	2.449	2.198	2.301	2.301	0
AvADS1	2.270	2.262	2.001	2.100	2.100	0
S8ADS0	42.487	15.120	15.120	15.120	15.120	0
S8ADS1	15.356	14.082	14.082	14.082	14.082	0

ET46/2006 – Et. Finală/2007 – Lucrare in extenso

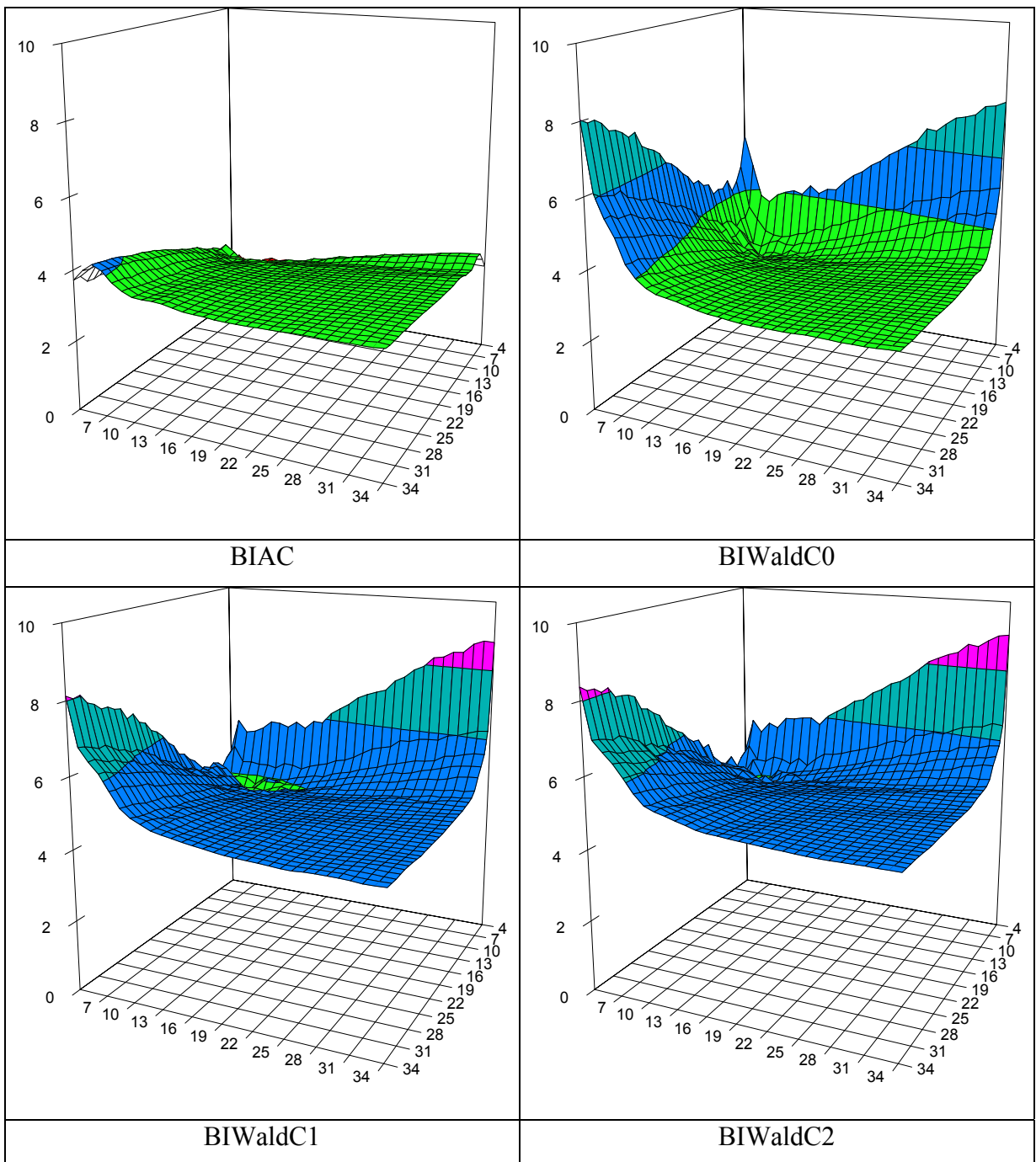
► Următorul tabel centralizează rezultatele obținute prin evaluarea cu fiecare din cele 12 metode pentru $m=n=34$ și $0 \leq X \leq m$, $0 \leq Y \leq n$ și $\alpha = 5\%$:

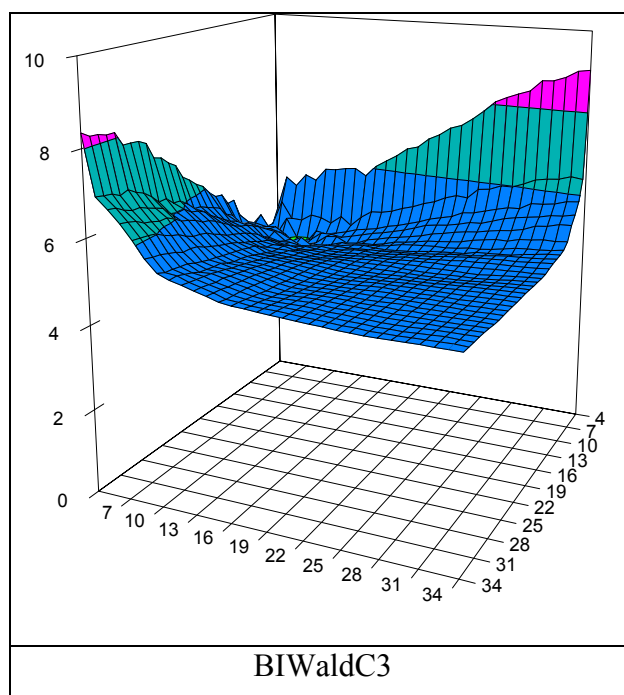
	BIAC	BIWaldC0	BIWaldC1	BIWaldC2	BIWaldC3	“Best of”
IAvOE0	1.576	1.616	0.620	0.280	0.280	0
IAvOE1	1.569	1.630	0.590	0.244	0.244	0
StDOE0	2.294	2.338	2.572	2.650	2.650	0
StDOE1	2.000	2.002	2.272	2.347	2.347	0
SiDOE0	2.783	2.842	2.644	2.663	2.663	0
SiDOE1	2.541	2.581	2.346	2.358	2.358	0
AvADA0	1.610	1.572	1.855	1.955	1.955	0
AvADA1	1.505	1.449	1.738	1.828	1.828	0
AvADS0	1.972	2.023	1.811	1.912	1.912	0
AvADS1	1.859	1.907	1.686	1.789	1.789	0
S8ADS0	14.005	14.007	14.007	14.007	14.007	0
S8ADS1	12.984	12.986	12.986	12.986	12.986	0

► Următorul tabel centralizează rezultatele obținute prin evaluarea cu fiecare din cele 12 metode pentru $m=n=54$ și $0 \leq X \leq m$, $0 \leq Y \leq n$ și $\alpha = 5\%$:

	BIAC	BIWaldC0	BIWaldC1	BIWaldC2	BIWaldC3	“Best of”
IAvOE0	1.315	1.368	0.577	0.280	0.280	0
IAvOE1	1.288	1.346	0.540	0.234	0.234	0
StDOE0	1.768	1.779	2.021	2.103	2.103	0
StDOE1	1.583	1.571	1.828	1.914	1.914	0
SiDOE0	2.203	2.244	2.101	2.122	2.122	0
SiDOE1	2.041	2.069	1.906	1.928	1.928	0
AvADA0	1.286	1.257	1.473	1.554	1.554	0
AvADA1	1.219	1.185	1.397	1.473	1.473	0
AvADS0	1.513	1.556	1.381	1.480	1.480	0
AvADS1	1.443	1.482	1.305	1.408	1.408	0
S8ADS0	12.574	12.577	12.577	12.577	12.577	0
S8ADS1	11.610	11.613	11.613	11.613	11.613	0

► Următoarul tabel prezintă descriptiv rezultatele obținute (în termeni de eroare experimentală) cu fiecare dintre cele 5 metode pentru $4 \leq m, n \leq 34$ și $\alpha=5\%$.





► Următorul tabel centralizează rezultatele obținute prin evaluarea cu fiecare din cele 12 metode pentru $4 \leq m, n \leq 34$ și $0 \leq X \leq m$, $0 \leq Y \leq n$ și $\alpha = 5\%$:

	BIAC	BIWaldC0	BIWaldC1	BIWaldC2	BIWaldC3	“Best of”
IAvOE0	1.583	1.624	0.494	0.180	0.180	0
IAvOE1	1.397	1.442	0.248	0.083	0.083	0
StDOE0	5.009	3.479	3.634	3.701	3.701	0
StDOE1	5.060	3.469	3.562	3.610	3.610	0
SiDOE0	5.253	3.839	3.668	3.705	3.705	0
SiDOE1	5.250	3.757	3.571	3.611	3.611	0
AvADA0	2.252	2.126	2.404	2.500	2.500	0
AvADA1	2.186	2.054	2.273	2.351	2.351	0
AvADS0	2.736	2.662	2.425	2.496	2.496	0
AvADS1	2.608	2.532	2.281	2.355	2.355	0
S8ADS0	42.121	15.751	15.765	15.774	15.774	0
S8ADS1	42.350	15.841	15.855	15.864	15.865	0

ET46/2006 – Et. Finală/2007 – Lucrare in extenso

► Performanța metodelor BIAC, BIWaldC0, BIWaldC1, BIWaldC2, BIWaldC3 (prezentate în tabelele de mai sus) este dată în următoarele tabele ce conțin metodele ordonate:

m=n=4	BIAC	BIWaldC0	BIWaldC1	BIWaldC2	BIWaldC3	“Best of”
IAvOE0	5	4	2	2	2	1
IAvOE1	5	4	2	2	2	1
StDOE0	1	5	3	3	3	1
StDOE1	1	5	3	3	3	1
SiDOE0	1	5	3	3	3	1
SiDOE1	1	5	3	3	3	1
AvADA0	1	5	3	3	3	1
AvADA1	1	5	3	3	3	1
AvADS0	1	5	3	3	3	1
AvADS1	1	5	3	3	3	1
S8ADS0	1	5	3	3	3	1
S8ADS1	1	5	3	3	3	1
Σ	20	58	34	34	34	min

m=n=14	BIAC	BIWaldC0	BIWaldC1	BIWaldC2	BIWaldC3	“Best of”
IAvOE0	4	5	3	1.5	1.5	1
IAvOE1	4	5	3	1.5	1.5	1
StDOE0	5	1	4	2.5	2.5	1
StDOE1	5	1	2	3.5	3.5	1
SiDOE0	5	4	3	1.5	1.5	1
SiDOE1	5	4	3	1.5	1.5	1
AvADA0	5	1	4	2.5	2.5	1
AvADA1	3	1	2	4.5	4.5	1
AvADS0	5	4	1	2.5	2.5	1
AvADS1	5	4	1	2.5	2.5	1
S8ADS0	5	1	3	3	3	1
S8ADS1	5	2.5	2.5	2.5	2.5	1
Σ	56	33.5	31.5	29.5	29.5	min

m=n=24	BIAC	BIWaldC0	BIWaldC1	BIWaldC2	BIWaldC3	“Best of”
IAvOE0	4	5	3	1.5	1.5	1
IAvOE1	4	5	3	1.5	1.5	1
StDOE0	5	1	2	3.5	3.5	1
StDOE1	5	1	2	3.5	3.5	1
SiDOE0	5	4	1	2.5	2.5	1
SiDOE1	5	4	1	2.5	2.5	1
AvADA0	2	1	3	4.5	4.5	1
AvADA1	2	1	3	4.5	4.5	1
AvADS0	5	4	1	2.5	2.5	1
AvADS1	5	4	1	2	2	1
S8ADS0	5	2.5	2.5	2.5	2.5	1
S8ADS1	5	2.5	2.5	2.5	2.5	1
Σ	52	35	25	33.5	33.5	min

m=n=34	BIAC	BIWaldC0	BIWaldC1	BIWaldC2	BIWaldC3	“Best of”
IAvOE0	4	5	3	1.5	1.5	1
IAvOE1	4	5	3	1.5	1.5	1
StDOE0	1	2	3	4.5	4.5	1
StDOE1	1	2	3	4.5	4.5	1
SiDOE0	4	5	1	2.5	2.5	1
SiDOE1	4	5	1	2.5	2.5	1
AvADA0	2	1	3	4.5	4.5	1
AvADA1	2	1	3	4.5	4.5	1
AvADS0	4	5	1	2.5	2.5	1
AvADS1	4	5	1	2.5	2.5	1
S8ADS0	1	3.5	3.5	3.5	3.5	1
S8ADS1	1	3.5	3.5	3.5	3.5	1
Σ	32	43	29	38	38	min

m=n=54	BIAC	BIWaldC0	BIWaldC1	BIWaldC2	BIWaldC3	“Best of”
IAvOE0	4	5	3	1.5	1.5	1
IAvOE1	4	5	3	1.5	1.5	1
StDOE0	1	2	3	4.5	4.5	1
StDOE1	1	2	3	4.5	4.5	1
SiDOE0	4	5	1	2.5	2.5	1
SiDOE1	4	5	1	2.5	2.5	1
AvADA0	2	1	3	4.5	4.5	1
AvADA1	2	1	3	4.5	4.5	1
AvADS0	4	5	1	2.5	2.5	1
AvADS1	4	5	1	2.5	2.5	1
S8ADS0	1	3.5	3.5	3.5	3.5	1
S8ADS1	1	3.5	3.5	3.5	3.5	1
Σ	32	43	29	38	38	min

3<m,n<35	BIAC	BIWaldC0	BIWaldC1	BIWaldC2	BIWaldC3	“Best of”
IAvOE0	4	5	3	1.5	1.5	1
IAvOE1	4	5	3	1.5	1.5	1
StDOE0	5	1	2	3.5	3.5	1
StDOE1	5	1	2	3.5	3.5	1
SiDOE0	5	4	1	2.5	2.5	1
SiDOE1	5	4	1	2.5	2.5	1
AvADA0	2	1	3	4.5	4.5	1
AvADA1	2	1	3	4.5	4.5	1
AvADS0	5	4	1	2.5	2.5	1
AvADS1	5	4	1	2.5	2.5	1
S8ADS0	5	1	2	3.5	3.5	1
S8ADS1	5	1	2	3	4	1
Σ	52	32	24	35.5	36.5	min

► Tabelul următor centralizează performanțele obținute de metode:



Σ	BIAC	BIWaldC0	BIWaldC1	BIWaldC2	BIWaldC3	“Best of”
m=n=4	20	58	34	34	34	min
m=n=14	56	33.5	31.5	29.5	29.5	min
m=n=24	52	35	25	33.5	33.5	min
m=n=34	32	43	29	38	38	min
m=n=54	32	43	29	38	38	min
3<m,n<35	52	32	24	35.5	36.5	min
Σ	244	244.5	172.5	208.5	209.5	min

► Rezultatele expuse în tabelul de mai sus permit următoarele concluzii:



- ÷ Metoda clasică (notată BIWaldC0) este surclasată de toate metodele optimizate propuse (chiar dacă numai cu 0.5 unități de rang, cum este cazul metodei BIAC).
- ÷ Fiecare metodă propusă are modul său de cumulare a erorii experimentale (chiar dacă metodele BIWaldC2 și BIWaldC3 diferă doar cu 1 unitate de rang - 208.5 pentru BIWaldC2 față de 209.5 pentru BIWaldC3).
- ÷ Metoda BIWaldC1 se distanțează semnificativ de următoarele (cu rang cumulat de 172.5 față de 208.5 pentru următoarea) fiind deci cea mai bună optimizare a intervalului de încredere pentru funcția BI.

Etapa 1, Activitatea 4 & Etapa 2, Activitatea 4 & Etapa 3 (finală) Activitatea 4.
 Participări la manifestări științifice și dobândirea de competențe complementare

Nr	Conferință	Contribuții
1	 <p>Workshop on Ecomaterials and Processes: Characterization and Metrology April 19-21, 2007 St. Kirik, Plovdiv, Bulgaria</p>	÷ Sorana Daniela BOLBOACĂ, Lorentz JÄNTSCHI, <i>Modeling the Property of Compounds from Structure: Statistical Methods for Models Validation</i> , Institute of General and Inorganic Chemistry, Bulgarian Academy of Sciences [http://sizemat.igic.bas.bg], FP6: EC-INCO-CT-2005-016414 Specific Support Action, Plovdiv, Bulgaria, p. 71, April 19-21, 2007.
2	 <p>Monmouth University, New Jersey, USA, Internet</p>	÷ Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI, Data Mining on Structure-Activity/Property Relationships Models, ECC11- The 11th Electronic Computational Chemistry Conference, online, Monmouth University, New Jersey, USA, N/A, www, Internet, paper #29, April 2-30, 2007.
3	 <p>27-30 May 2007, Bucharest, ROMANIA An Event of the CAPE Working Party of the European Federation of Chemical Engineering</p>	÷ Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI, Modelling the Inhibitory Activity on Carbonic Anhydrase I of Some Substituted Thiadiazole- and Thiadiazoline-Disulfonamides: Integration of Structure Information, ESCAPE17 - 17th European Symposium on Computer Aided Process Engineering, CAPE Working Party of the European Federation of Chemical Engineering then Elsevier Netherlands & UK, ISBN 978-0-444-53157-5 & eISBN 0-444-53158-2, Bucharest, Romania then Amsterdam, Netherlands, T4-212 (oral presentation), May 27-30 then p. 965-970, 2007.

<p>4</p>	 <p>22 - 24 June 2007 Bayview Beach Resort, Penang, Malaysia</p> <p>INTERNATIONAL MANAGEMENT EDUCATION CONFERENCE 'Comparative Perspectives On Best Practices'</p>	<p>÷ Carmen E. STOENOIU, Lorentz JÄNTSCHI, Ioan ABRUDAN, Sorana D. BOLBOACĂ, Romanian Higher Education: Modelling Evolution Tendencies, The International Management Education Conference 2007 Proceedings, Faculty of Business and Economics, University Pendidikan Sultan Idris, ISBN 978-983-3759-19-4, Penang, Malaysia, paper #10§02 (6 pages), 2007.</p> <p>÷ Lorentz JÄNTSCHI, Carmen E. STOENOIU, Sorana D. BOLBOACĂ, Research Policy via Funding Allocation Analysis, The International Management Education Conference 2007 Proceedings, Faculty of Business and Economics, University Pendidikan Sultan Idris, ISBN 978-983-3759-19-4, Penang, Malaysia, paper #10§03 (7 pages), 2007.</p>
<p>5</p>	<p>Welcome to <i>International Conference of Applied Mathematics</i></p>  <p>Fourth International Conference of Applied Mathematics and Computing August 12 - 18, 2007 Plovdiv, Bulgaria</p>	<p>÷ Sorana-Daniela BOLBOACĂ, Lorentz JÄNTSCHI, Are confidence intervals for binomial distributed samples an optimization meters?, Fourth International Conference of Applied Mathematics and Computing, August 12-18, 2007, University of Chemical Technology and Metallurgy Sofia & Technical University of Plovdiv, Invited lecture, presented on August 13, from 18.00 to 18.30, Plovdiv, Bulgaria, p. 47, 2007.</p> <p>÷ Lorentz JÄNTSCHI, Sorana-Daniela BOLBOACĂ, Cristina Maria FURDUI, Characteristic and Counting Polynomials on Modeling Nonane Isomers Properties, Fourth International Conference of Applied</p>

ET46/2006 – Et. Finală/2007 – Lucrare in extenso

		Mathematics and Computing, August 12-18, 2007, University of Chemical Technology and Metallurgy Sofia & Technical University of Plovdiv, Invited lecture, presented on August 15, from 12.20 to 13.00, Plovdiv, Bulgaria, p. 234, 2007.
6	  The IEEE Region 8 Eurocon 2007 Conference September 9-12, 2007 Warsaw, Poland	÷ Lorentz JÄNTSCHI, Carmen Elena STOENOIU, Sorana Daniela BOLBOACĂ, Linking Assessment to e-Learning in Microbiology and Toxicology for Undergraduate Students, Eurocon 2007 Conference, IEEE Region 8, Poster 258, Warsaw, Poland, September 9-12, 2007. ÷ Lorentz JÄNTSCHI, Mugur BĂLAN, Margareta Emilia PODAR, Sorana Daniela BOLBOACĂ, Thermal Energy Efficiency Analysis for Residential Buildings, Eurocon 2007 Conference, IEEE Region 8, Poster 143, Warsaw, Poland, September 9-12, 2007.
7	Ninth Annual Conference YUCOMAT 2007 September 10-14, 2007 Herceg Novi, Montenegro	÷ Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI, Similarities Analysis on Hydroxyapatite-Zirconia Composites, 9th Annual Conference of the Yugoslav Materials Research Society, Serbian Academy of Sciences and Arts, ISBN 978-86-80321-11-0, Herceg Novi, Montenegro, OSE10, p. 64, September 10-14, 2007.

Publicații

Nr	Tip	Publicație
1	Articol indexat în baze de date internaționale	Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI , <i>Virtual Environment for Continuing Medical Education</i> , Electronic Journal of Biomedicine, Red UniNet Spain, ISSN 1697-090X, www, Internet, 2007(2), p. 19-28, 2007 .
2	Articol indexat în baze de date internaționale	Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI , <i>Mapping Cigarettes Similarities using Cluster Analysis Methods</i> , International Journal of Environmental Research and Public Health, Molecular Diversity Preservation International, ISSN 1660-4601, www & Basel, Internet & Switzerland, 4(3), p. 233-242, 2007 .
3	Articol indexat în baze de date internaționale	Carmen E. STOENOIU, Ioan ABRUDAN, Lorentz JÄNTSCHI , <i>Time Series of Agricultural Sciences Higher Education in Romania</i> , Bulletin of University of Agricultural Sciences and Veterinary Medicine - Horticulture, AcademicPres, ISSN 1843-5254, eISSN 1843-5394, Cluj-Napoca, Romania, 64(1-2), p. 543-547, 2007 .
4	Articol indexat în baze de date internaționale	Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI , <i>Amino Acids Sequences Analysis on Collagen</i> , Bulletin of University of Agricultural Sciences and Veterinary Medicine - Animal Sciences and Biotechnologies, AcademicPres, ISSN 1843-5262, eISSN 1843-536X, Cluj-Napoca, Romania, 63-64, p. 311-316, 2007 .
5	Articol indexat în baze de date internaționale	Dan ILINA, Monica Palaghia FODOR, Lorentz JÄNTSCHI , <i>The Control of the Infrastructure of the Rural Water Wells in the Plain Area</i> , Bulletin of University of Agricultural Sciences and Veterinary Medicine - Horticulture, AcademicPres, ISSN 1843-5254, eISSN 1843-5394, Cluj-Napoca, Romania, 64(1-2), p. 628-633, 2007 .
7	Articol indexat în baze de date internaționale	Lorentz JÄNTSCHI, Sorana D. BOLBOACĂ, Carmen E. STOENOIU , <i>National Trends on Agricultural Crops Production: Cluster Analysis</i> , Bulletin of University of Agricultural Sciences and Veterinary Medicine - Agriculture, AcademicPres, ISSN 1843-5246, eISSN 1843-5386, Cluj-Napoca, Romania, 63-64, p. 194-202, 2007 .
8	Conferință internațională	Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI , <i>Similarities Analysis on Hydroxyapatite-Zirconia Composites</i> , 9th Annual Conference of the Yugoslav Materials Research Society, Serbian Academy of Sciences and Arts, ISBN 978-86-80321-11-0, Herceg

		Novi, Montenegro, OSE10, p. 64, September 10-14, 2007 .
9	Articol indexat în baze de date internaționale	Lorentz JÄNTSCHI, Carmen Elena STOENOIU, Sorana Daniela BOLBOACĂ , <i>Linking Assessment to e-Learning in Microbiology and Toxicology for Undergraduate Students</i> , Eurocon 2007 Conference, IEEE Region 8, Poster 258, Warsaw, Poland, September 9-12, 2007 , IEEE Region 8 Eurocon 2007: The International Conference on Computer as a Tool - Proceedings (ISBN: 1-4244-0813-X, IEEE: 07EX1617C, LoC: 2006937182) , 2447-2452, 2007 .
10	Articol indexat în baze de date internaționale	Lorentz JÄNTSCHI, Mugur BĂLAN, Margareta Emilia PODAR, Sorana Daniela BOLBOACĂ , <i>Thermal Energy Efficiency Analysis for Residential Buildings</i> , Eurocon 2007 Conference, IEEE Region 8, Poster 143, Warsaw, Poland, September 9-12, 2007 , IEEE Region 8 Eurocon 2007: The International Conference on Computer as a Tool - Proceedings (ISBN: 1-4244-0813-X, IEEE: 07EX1617C, LoC: 2006937182) , 2009-2014, 2007 .
11	Articol Thompson ISI	Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI , <i>Data Mining on Structure-Activity/Property Relationships Models</i> , World Applied Sciences Journal, IDOSI Publications, ISSN 1818-4952, www, Internet & Faisalabad, Pakistan, 2(4), p. 323-332, 2007 .
12	Viitor articol indexat în baze de date internaționale	Sorana-Daniela BOLBOACĂ, Lorentz JÄNTSCHI , <i>Are confidence intervals for binomial distributed samples an optimization meters?</i> , Fourth International Conference of Applied Mathematics and Computing, August 12-18, 2007, University of Chemical Technology and Metallurgy Sofia & Technical University of Plovdiv, Invited lecture, presented on August 13, from 18.00 to 18.30, Plovdiv, Bulgaria, p. 47, 2007 , International Journal of Pure and Applied Mathematics, 40(3), p. XX-YY, 2007 .
13	Viitor articol indexat în baze de date internaționale	Lorentz JÄNTSCHI, Sorana-Daniela BOLBOACĂ, Cristina Maria FURDUI , <i>Characteristic and Counting Polynomials on Modeling Nonane Isomers Properties</i> , Fourth International Conference of Applied Mathematics and Computing, August 12-18, 2007, University of Chemical Technology and Metallurgy Sofia & Technical University of Plovdiv, Invited lecture, presented on August 15, from 12.20 to 13.00, Plovdiv, Bulgaria, p. 234, 2007 , International Journal of Pure and Applied Mathematics, 40(3), p. XX-YY, 2007 .
14	Conferință internațională	Carmen E. STOENOIU, Lorentz JÄNTSCHI, Ioan ABRUDAN, Sorana D. BOLBOACĂ , <i>Romanian Higher Education: Modelling Evolution Tendencies</i> , The International Management Education Conference 2007 Proceedings, Faculty of Business and Economics, University Pendidikan Sultan Idris, ISBN 978-983-3759-19-4, Penang, Malaysia, paper #10§02 (6 pages), 2007 .
15	Conferință internațională	Lorentz JÄNTSCHI, Carmen E. STOENOIU, Sorana D. BOLBOACĂ , <i>Research Policy via Funding Allocation Analysis</i> , The International Management Education Conference 2007

		Proceedings, Faculty of Business and Economics, University Pendidikan Sultan Idris, ISBN 978-983-3759-19-4, Penang, Malaysia, paper #10§03 (7 pages), 2007 .
16	Articol indexat în baze de date internaționale	Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI , <i>Modelling the Inhibitory Activity on Carbonic Anhydrase I of Some Substituted Thiadiazole- and Thiadiazoline-Disulfonamides: Integration of Structure Information</i> , ESCAPE17 - 17th European Symposium on Computer Aided Process Engineering, CAPE Working Party of the European Federation of Chemical Engineering then Elsevier Netherlands & UK, ISBN 978-0-444-53157-5 & eISBN 0-444-53158-2, Bucharest, Romania then Amsterdam, Netherlands, T4-212 (oral presentation), May 27-30 then p. 965-970, 2007 .
17	Articol indexat în baze de date internaționale	Lorentz JÄNTSCHI, Sorana D. BOLBOACĂ , <i>The Jungle of Linear Regression Revisited</i> , Leonardo Electronic Journal of Practices and Technologies, AcademicDirect, ISSN 1583-1078, www, Internet, 6(10), p. 169-187, 2007 .
18	Articol indexat în baze de date internaționale	Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI , <i>Communication of Results on Risk Factors Studies: Confidence Intervals</i> , Leonardo Journal of Sciences, AcademicDirect, ISSN 1583-0233, www, Internet, 6(10), p. 179-187, 2007 .
19	Articol Thompson ISI SCI Expanded	Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI , <i>How Good the Characteristic Polynomial Can Be for Correlations?</i> , International Journal of Molecular Sciences, Molecular Diversity Preservation International , ISSN 1422-0067, www & Basel, Internet & Switzerland, 8(4), p. 335-345, 2007 .
20	Conferință internațională	Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI , <i>Data Mining on Structure-Activity/Property Relationships Models</i> , ECC11- The 11th Electronic Computational Chemistry Conference , online, Monmouth University, New Jersey, USA, N/A, www, Internet, paper #29, Presentation located here , April 2-30, 2007 .
21	viitor articol Thompson ISI SCI Expanded	Sorana Daniela BOLBOACĂ, Lorentz JÄNTSCHI , <i>Modeling the Property of Compounds from Structure: Statistical Methods for Models Validation</i> , Institute of General and Inorganic Chemistry, Bulgarian Academy of Sciences [http://sizemat.igic.bas.bg], FP6: EC-INCO-CT-2005-016414 Specific Support Action, Plovdiv, Bulgaria, p. 71, April 19-21, 2007 , <i>Environmental Chemistry Letters</i> , DOI 10.1007/s10311-007-0119-9 .
22	Articol indexat în baze de date internaționale	Sorana-Daniela BOLBOACĂ, Lorentz JÄNTSCHI , <i>Computer-based testing on physical chemistry topic: A case study</i> , International Journal of Education and Development using Information and Communication Technology, The University of the West Indies, Barbados, ISSN 1814-0556, www, Internet, id=242 , 3(1), 2007 .

Sorana D. BOLBOACĂ, Lorentz JĂNTSCHI, *Virtual Environment for Continuing Medical Education*, Electronic Journal of Biomedicine, Red UniNet Spain, ISSN 1697-090X, www, Internet, 2007(2), p. 19-28, **2007**.

- [id] => 160
- [recorddate] => 2007:10:12:16:51:44
- [lastupdate] => 2007:10:12:16:51:44
- [type] => article
- [place] => www, Internet
- [subject] => informatics - applied; informatics - databases; medicine - evidence based; medicine - informatics
- [relatedworks] =>
 - 4 (some):
 - Creating Diagnostic Critical Appraised Topics. CATRom Original Software for Romanian Physicians, [?id=36](#)
 - Creating Therapy Studies Critical Appraised Topics. CATRom Original Software for Romanian Physicians, [?id=37](#)
 - Relational Information in Medicine: A Challenge, [?id=43](#)
 - PHP and MySQL Medical Application Based on Tanner Whitehouse Standard, [?id=77](#)
 - Interactive Web Application for Evidence-Based Medicine Training, [?id=97](#)

- [file] => [?f=160](#)
- [mime] => application/pdf
- [size] => 188893
- [pubname] => Electronic Journal of Biomedicine
- [pubinfo] => Red UniNet Spain
- [pubkey] => ISSN 1697-090X
- [workinfo] => 2007(2), p. 19-28
- [year] => 2007
- [title] => Virtual Environment for Continuing Medical Education
- [authors] => Sorana D. BOLBOACĂ, Lorentz JĂNTSCHI
- [abstract] =>

The rapid development of communication and information technologies lead to the changes in continuing medical education by offering the possibility to move up-to-date medical information through Internet to the physicians. The main goal of this study was to create a virtual space for continuing medical education. In this context, a number of computer-assisted tools for instruction, evaluation and utilization in daily activity have been developed and integrated into a unitary system. The main imposed specifications of the system were accessibility, integrity, availability, and security.

This report describes the characteristics of tables design and organization, and of system integration. The security level was imposed for assuring the accessibility of each physician to medical information useful in his or her activity and the knowledge database development.

- [keywords] => Continuing Professional Education; Computer-Assisted Instruction and Evaluation; Medical Applications; Database Architecture

Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI, *Mapping Cigarettes Similarities using Cluster Analysis Methods*, International Journal of Environmental Research and Public Health, Molecular Diversity Preservation International, ISSN 1660-4601, www & Basel, Internet & Switzerland, 4(3), p. 233-242, **2007**.

- [id] => 159
- [recorddate] => 2007:10:06:18:02:54
- [lastupdate] => 2007:10:06:18:02:54
- [type] => article
- [place] => www & Basel, Internet & Switzerland
- [subject] => medicine - epidemiology; medicine - evidence based; medicine - informatics
- [relatedworks] =>
 - 3 (low):
 - Binomial Distribution Sample Confidence Intervals Estimation 1. Sampling and Medical Key Parameters Calculation, [?id=32](#)
- [file] => [?f=159](#)
- [mime] => application/pdf
- [size] => 226010
- [pubname] => International Journal of Environmental Research and Public Health
- [pubinfo] => Molecular Diversity Preservation International
- [pubkey] => ISSN 1660-4601
- [workinfo] => 4(3), p. 233-242
- [year] => 2007
- [title] => Mapping Cigarettes Similarities using Cluster Analysis Methods
- [authors] => Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI
- [abstract] =>

The aim of the research was to investigate the relationship and/or occurrences in and between chemical composition information (tar, nicotine, carbon monoxide), market information (brand, manufacturer, price), and public health information (class, health warning) as well as clustering of a sample of cigarette data. A number of thirty cigarette brands have been analyzed. Six categorical (cigarette brand, manufacturer, health warnings, class) and four continuous (tar, nicotine, carbon monoxide concentrations and package price) variables were collected for investigation of chemical composition, market information and public health information. Multiple linear regression and two clusterization techniques have been applied. The study revealed interesting remarks. The carbon monoxide concentration proved to be linked with tar and nicotine concentration. The applied clusterization methods identified groups of cigarette brands that shown similar characteristics. The tar and carbon monoxide concentrations were the main criteria used in clusterization. An analysis of a largest sample could reveal more relevant and useful information regarding the similarities between cigarette brands.
- [keywords] => Tobacco tar, nicotine and carbon monoxide; heath warnings; explanatory health messages; multidimensional analysis; cluster analysis

Carmen E. STOENOIU, Ioan ABRUDAN, Lorentz JÄNTSCHI, *Time Series of Agricultural Sciences Higher Education in Romania*, Bulletin of University of Agricultural Sciences and Veterinary Medicine - Horticulture, AcademicPres, ISSN 1843-5254, eISSN 1843-5394, Cluj-Napoca, Romania, 64(1-2), p. 543-547, **2007**.

- [id] => 157
- [recorddate] => 2007:10:06:17:44:37
- [lastupdate] => 2007:10:06:17:44:37
- [type] => article
- [place] => Cluj-Napoca, Romania
- [subject] => education - policy; engineering - agriculture
- [relatedworks] =>
 - 3 (low):
 - National Trends on Agricultural Crops Production: Cluster Analysis, [?id=155](#)
 - The Control of the Infrastructure of the Rural Water Wells in the Plain Area, [?id=156](#)
- [file] => [?f=157](#)
- [mime] => application/pdf
- [size] => 465354
- [pubname] => Bulletin of University of Agricultural Sciences and Veterinary Medicine - Horticulture
- [pubinfo] => AcademicPres
- [pubkey] => ISSN 1843-5254, eISSN 1843-5394
- [workinfo] => 64(1-2), p. 543-547
- [year] => 2007
- [title] => Time Series of Agricultural Sciences Higher Education in Romania
- [authors] => Carmen E. STOENOIU, Ioan ABRUDAN, Lorentz JÄNTSCHI
- [abstract] =>

In the last two decades in Romania a series of changes on both higher education and economy occurred. These changes had determined variations at the option to follow a profile at first level of higher education. The present study includes data about the evolution of the number of students in higher education at the agronomical profiles in the period 1991 to 2003. In order to analyze the trends referring to the agronomical profiles of Romanian higher education a mathematical model has been developed. The model integrates the following parameters: (1) the mean annual variation ratio (as absolute and relative values), (2) the increasing mean annual ratio (as absolute and relative values), (3) the number of students' estimation for 1989 (as absolute value), (4) the numbers of students prediction for 2008, (5) the correlation coefficient, (6) the linear trend, (7) the number of students estimation for 1989 and (8) for 2008 obtained by the model. The mathematical model has been integrated into an online application, and here are analyzed and discussed. Using the descriptive information of public funds allocated on the last two years to four universities with agricol profile from Romania, a discussion regarding the evolution of the composite index of unitary students on each university was open.
- [keywords] => Higher Education, Mathematical Model, Agronomical Profiles
- [acknowledgment] => This research was partly supported by UEFISCSU Romania through project ET46/2006.

Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI, *Amino Acids Sequences Analysis on Collagen*, Bulletin of University of Agricultural Sciences and Veterinary Medicine - Animal Sciences and Biotechnologies, AcademicPres, ISSN 1843-5262, eISSN 1843-536X, Cluj-Napoca, Romania, 63-64, p. 311-316, **2007**.

- [id] => 158
- [recorddate] => 2007:10:06:17:47:01
- [lastupdate] => 2007:10:06:17:47:01
- [type] => article
- [place] => Cluj-Napoca, Romania
- [subject] => engineering - agriculture; medicine - evidence based; medicine - informatics
- [relatedworks] =>
 - 4 (some):
 - National Trends on Agricultural Crops Production: Cluster Analysis, [?id=155](#)
 - The Control of the Infrastructure of the Rural Water Wells in the Plain Area, [?id=156](#)
- [file] => [?f=158](#)
- [mime] => application/pdf
- [size] => 421293
- [pubname] => Bulletin of University of Agricultural Sciences and Veterinary Medicine - Animal Sciences and Biotechnologies
- [pubinfo] => AcademicPres
- [pubkey] => ISSN 1843-5262, eISSN 1843-536X
- [workinfo] => 63-64, p. 311-316
- [year] => 2007
- [title] => Amino Acids Sequences Analysis on Collagen
- [authors] => Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI
- [abstract] =>

Starting from available information about amino acids properties and sequences on collagen type I chains, the aims of the study were to identify the principal property component and to analyze the similarities within and between collagens on five species. The principal component analysis applied on twentyfour amino acids properties revealed that the hydrophobic or hydrophilic character measured by Wealling et al. is more stable comparing with the other investigated properties. Similarity analysis identified similar and dissimilar within and between studied species from the viewpoint of amino acids sequences on collagen type I alpha 1 and 2 chains.
- [keywords] => Amino Acids; Collagen Type I; Properties; Multivariate Analysis
- [acknowledgment] => The research was partly supported by CNCSIS Romania through project AT93/2007 and UEFISCSU Romania through project ET46/2006.

Dan ILINA, Monica Palaghia FODOR, Lorentz JÄNTSCHI, *The Control of the Infrastructure of the Rural Water Wells in the Plain Area*, Bulletin of University of Agricultural Sciences and Veterinary Medicine - Horticulture, AcademicPres, ISSN 1843-5254, eISSN 1843-5394, Cluj-Napoca, Romania, 64(1-2), p. 628-633, **2007**.

- [id] => 156
- [recorddate] => 2007:10:06:17:42:59
- [lastupdate] => 2007:10:06:17:42:59
- [type] => article
- [place] => Cluj-Napoca, Romania
- [subject] => engineering - agriculture
- [relatedworks] =>
 - 0 (highest):
 - National Trends on Agricultural Crops Production: Cluster Analysis, [?id=155](#)
 - 2 (average):
 - Mechanical Milling: Evolution of Crystal Parameter of Iron Powder, [?id=150](#)
 - Superficial Distortion Influence on Characteristics of the Iron-Based Materials, [?id=151](#)
 - 3 (low):
 - Time Series of Agricultural Sciences Higher Education in Romania, [?id=157](#)
 - 4 (some):
 - Amino Acids Sequences Analysis on Collagen, [?id=158](#)
- [file] => [?f=156](#)
- [pubname] => Bulletin of University of Agricultural Sciences and Veterinary Medicine - Horticulture
- [pubinfo] => AcademicPres
- [pubkey] => ISSN 1843-5254, eISSN 1843-5394
- [workinfo] => 64(1-2), p. 628-633
- [year] => 2007
- [title] => The Control of the Infrastructure of the Rural Water Wells in the Plain Area
- [authors] => Dan ILINA, Monica Palaghia FODOR, Lorentz JÄNTSCHI
- [abstract] =>

In the Romanian rural area, especially in the plain areas, water well infrastructure is spread in the entire landscape. However, very little is known about the condition or distribution of the water well infrastructure on which much of the rural population depends on. Groundwater is the preferred water supply source in most rural areas, because it is readily available and accessible over large parts of the Romanian Plain and West Plain which have the tendency, regarding the climatic changes, to turn into semi-arid area. A current review is a necessary way in ongoing efforts to know climate adaptation challenges and to develop instruments that will allow water wells to endure drought conditions without disruption at these critical times. A medium strategy at a local level and a long term strategy at a national level, as well as creating a database concerning the water wells and permanently upgrading the data regarding these wells will constitute the premises for some viable alternatives for water supplies.
- [keywords] => water reserves; geophysical investigations; water monitoring

Lorentz JÄNTSCHI, Sorana D. BOLBOACĂ, Carmen E. STOENOIU, *National Trends on Agricultural Crops Production: Cluster Analysis*, Bulletin of University of Agricultural Sciences and Veterinary Medicine - Agriculture, AcademicPres, ISSN 1843-5246, eISSN 1843-5386, Cluj-Napoca, Romania, 63-64, p. 194-202, **2007**.

- [id] => 155
- [recorddate] => 2007:10:06:17:39:39
- [lastupdate] => 2007:10:06:17:39:39
- [type] => article
- [place] => Cluj-Napoca, Romania
- [subject] => engineering - agriculture
- [relatedworks] =>
 - 0 (highest):
 - The Control of the Infrastructure of the Rural Water Wells in the Plain Area, [?id=156](#)
 - 2 (average):
 - Mechanical Milling: Evolution of Crystal Parameter of Iron Powder, [?id=150](#)
 - Superficial Distortion Influence on Characteristics of the Iron-Based Materials, [?id=151](#)
 - 3 (low):
 - Time Series of Agricultural Sciences Higher Education in Romania, [?id=157](#)
 - 4 (some):
 - Amino Acids Sequences Analysis on Collagen, [?id=158](#)
- [file] => [?f=155](#)
- [mime] => application/pdf
- [size] => 797600
- [pubname] => Bulletin of University of Agricultural Sciences and Veterinary Medicine - Agriculture
- [pubinfo] => AcademicPres
- [pubkey] => ISSN 1843-5246, eISSN 1843-5386
- [workinfo] => 63-64, p. 194-202
- [year] => 2007
- [title] => National Trends on Agricultural Crops Production: Cluster Analysis
- [authors] => Lorentz JÄNTSCHI, Sorana D. BOLBOACĂ, Carmen E. STOENOIU
- [abstract] =>

Starting from descriptive data on crop production and cultivated area at national level during on fifteen years, the aim of this study is to reveal the trends on crops cultivation. The cluster analysis reveals linkages between crops classes as well as between different crops, which can be partly assigned to crops rotation. Time series analysis reveals dramatically reducing of production of some crops, such as flax, hemp, and sugar beet, and increasing of production, such at sunflower, and increasing of productivity, such at potatoes and field vegetables.
- [keywords] => Resources Management; Crops Production; National Trends; Cluster Analysis
- [acknowledgment] => This research was partly supported by UEFISCSU Romania through project ET46/2006 and CNCSIS Romania through project AT93/2007.

Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI, *Similarities Analysis on Hydroxyapatite-Zirconia Composites*, 9th Annual Conference of the Yugoslav Materials Research Society, Serbian Academy of Sciences and Arts, ISBN 978-86-80321-11-0, Herceg Novi, Montenegro, OSE10, p. 64, September 10-14, 2007.

- [id] => 152
- [recorddate] => 2007:10:06:16:18:28
- [lastupdate] => 2007:10:06:16:18:28
- [type] => conference
- [place] => Herceg Novi, Montenegro
- [subject] => engineering - materials science; mathematics - statistics
- [relatedworks] =>
 - 3 (low):
 - Property Investigations with an Automat Correlation Routine and Applications for a Set of Alloys, [?id=11](#)
 - Mechanical Milling: Evolution of Crystal Parameter of Iron Powder, [?id=150](#)
 - Superficial Distortion Influence on Characteristics of the Iron-Based Materials, [?id=151](#)
- [pubname] => 9th Annual Conference of the Yugoslav Materials Research Society
- [pubinfo] => Serbian Academy of Sciences and Arts
- [pubkey] => ISBN 978-86-80321-11-0
- [workinfo] => OSE10, p. 64, September 10-14
- [year] => 2007
- [title] => Similarities Analysis on Hydroxyapatite-Zirconia Composites
- [authors] => Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI
- [abstract] =>

Hydroxyapatite ($\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$) is one of the implants materials with medical applications due to its higher biocompatibility. The hydroxyapatite found complete utilization after proper preparation of composite. The influence of zirconia (ZrO_2) on the phase composition and on mechanical properties of hydroxyapatite-zirconia composites has been investigated and reported by Rapacz-Kmita et al [1]. Starting with the experimental data reported in [1], hierarchical cluster analysis methods were applied in order to assess the similarities of four different types of composites. Four classes of composites: HAp (hydroxyapatite), HAp-CGz (CGz = Coarse-Grained Zirconia), HAp-FGZ (FGZ = Fine-Grained Zirconia), and HAp-NGZ (NGZ = Needle-Grained Zirconia) cumulating a total number of sixteen experiments were analyzed. A number of nine quantitative variables were included into analysis: sintering temperature ($^{\circ}\text{C}$), Vickers hardness (GPa), bending strength (MPa), characteristic strength (MPa), Weibull modulus, anisotropy (%), Young's modulus (GPa), rigidity modulus (GPa), and Poisson ratio. Data were analyzed with SPSS software (average linkage between groups as cluster method and squared Euclidean distance as measure of analyzed variables). The analysis revealed interesting information regarding similarities between studied hydroxyapatite-zirconia composites.
- [keywords] => Hydroxyapatite (HAp); Zirconium dioxide (zirconia, ZrO_2); Hierarchical cluster analysis
- [acknowledgment] => The research was partly supported by UEFISCSU Romania through project ET108/2006.

Lorentz JÄNTSCHI, Carmen Elena STOENOIU, Sorana Daniela BOLBOACĂ, *Linking Assessment to e-Learning in Microbiology and Toxicology for Undergraduate Students*, Eurocon 2007 Conference, IEEE Region 8, Poster 258, Warsaw, Poland, September 9-12, **2007**.

- [id] => 149
- [recorddate] => 2007:10:06:15:44:39
- [lastupdate] => 2007:10:06:15:44:39
- [type] => conference
- [place] => Warsaw, Poland
- [subject] => education - evaluation; education - management; education - methodology; education - training
- [relatedworks] => N/A
- [file] => [?f=149](#)
- [mime] => application/pdf
- [size] => 227462
- [pubname] => Eurocon 2007 Conference
- [pubinfo] => IEEE Region 8
- [pubkey] => Poster 258
- [workinfo] => September 9-12
- [year] => 2007
- [title] => Linking Assessment to e-Learning in Microbiology and Toxicology for Undergraduate Students
- [authors] => Lorentz JÄNTSCHI, Carmen Elena STOENOIU, Sorana Daniela BOLBOACĂ
- [abstract] =>

Development of communication and information technology opens the possibility to create new learning and assessments tools. Beyond the world wide access to education across the country and globe, the opportunity of running virtual experiments and assisting processes modelling, the communication and information technology facilitate implementation of collaborative learning, promoting active implication of students in educational process. Regarding the assessment process, its design had also been changed, the concept of computer-aided assessment being more frequently used at university and post-university level. The students' knowledge assessment is necessary to be as objective as possible. Starting with experiences obtained by creation of online assessment systems for general chemistry, and from the necessity of a valid and reliable assessment, an auto-calibrated system has been developed. The aim of present research was to assess the microbiology and toxicology knowledge of fourth year students at the Faculty of Materials Science and Engineering from the Technical University of Cluj-Napoca, Romania by using the developed knowledge evaluation system.
- [keywords] => online evaluation; online training

Lorentz JÄNTSCHI, Mugur BĂLAN, Margareta Emilia PODAR, Sorana Daniela BOLBOACĂ, *Thermal Energy Efficiency Analysis for Residential Buildings*, Eurocon 2007 Conference, IEEE Region 8, Poster 143, Warsaw, Poland, September 9-12, **2007**.

- [id] => 148
- [recorddate] => 2007:10:06:15:24:40
- [lastupdate] => 2007:10:06:15:24:40
- [type] => conference
- [place] => Warsaw, Poland
- [subject] => education - training; informatics - applied
- [relatedworks] =>
 - 4 (some):
 - Computer Aided System for Student's Knowledge Assessment, [?id=91](#)
- [file] => [?f=148](#)
- [mime] => application/pdf
- [size] => 358642
- [pubname] => Eurocon 2007 Conference
- [pubinfo] => IEEE Region 8
- [pubkey] => Poster 143
- [workinfo] => September 9-12
- [year] => 2007
- [title] => Thermal Energy Efficiency Analysis for Residential Buildings
- [authors] => Lorentz JÄNTSCHI, Mugur BĂLAN, Margareta Emilia PODAR, Sorana Daniela BOLBOACĂ
- [abstract] =>

The concern of energy conservation, the reduction of green house gases and sustainability was continuously growing in last years. The concept of green building has been introduced, and refers the practice of increasing the efficiency with which buildings and their sites use and harvest energy, water, and materials, and reducing building impacts on human health and on environment, through better design, construction, operation, maintenance, and removal. Starting from the national and international trends in development of environmental performance of new and existing home buildings, an interactive system for assisting the calculation of home energy efficiency has been created and validated, and its performances are presented.
- [keywords] => online application; energetics; civil constructions

Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI, *Data Mining on Structure-Activity/Property Relationships Models*, World Applied Sciences Journal, IDOSI Publications, ISSN 1818-4952, www, Internet & Faisalabad, Pakistan, 2(4), p. 323-332, **2007**.

- [id] => 147
- [recorddate] => 2007:08:20:19:17:54
- [lastupdate] => 2007:08:20:19:17:54
- [type] => article
- [place] => www, Internet & Faisalabad, Pakistan
- [subject] => chemistry - computational; informatics - applied; mathematics - statistics; research - evaluation
- [relatedworks] =>
 - 2 (average):
 - MDF - A New QSAR/QSPR Molecular Descriptors Family, [?id=33](#)
- [file] => [?f=147](#)
- [mime] => application/pdf
- [size] => 220742
- [pubname] => World Applied Sciences Journal
- [pubinfo] => IDOSI Publications
- [pubkey] => ISSN 1818-4952
- [workinfo] => 2(4), p. 323-332
- [year] => 2007
- [title] => Data Mining on Structure-Activity/Property Relationships Models
- [authors] => Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI
- [abstract] =>

Molecular descriptors family on structure-activity/property relationships studies were carried out in order to identify the link between compounds structure and their activity/property. A number of fifty-five classes of properties or activities of different compounds sets were investigated. Single and multi-varied linear regression models using molecular descriptors as variables were identified. The models with estimation and prediction abilities and associated characteristics were stored into a database. A data mining analysis using classification and clustering were applied on the obtained database for searching and extracting useful information. The methodology applied in searching and extracting for information and the obtained results are presented.
- [keywords] => Knowledge-Discovery in Database (KDD); cluster analysis; Structure-Activity/Property Relationships (SAR/SPR); Molecular Descriptors Family (MDF)

Sorana-Daniela BOLBOACĂ, Lorentz JÄNTSCHI, Are confidence intervals for binomial distributed samples an optimization meters?, Fourth International Conference of Applied Mathematics and Computing, August 12-18, 2007, University of Chemical Technology and Metallurgy Sofia & Technical University of Plovdiv, Invited lecture, presented on August 13, from 18.00 to 18.30, Plovdiv, Bulgaria, p. 47, **2007**.

- [id] => 146
- [recorddate] => 2007:08:19:19:20:08
- [lastupdate] => 2007:08:19:19:20:08
- [type] => conference
- [place] => Plovdiv, Bulgaria
- [subject] => informatics - fundamentals; informatics - models implementation; informatics - numerical optimization; mathematics - applied; mathematics - modeling; mathematics - statistics; medicine - informatics
- [relatedworks] => N/A
- [pubname] => Fourth International Conference of Applied Mathematics and Computing, August 12-18, 2007
- [pubinfo] => University of Chemical Technology and Metallurgy Sofia & Technical University of Plovdiv
- [pubkey] => Invited lecture, presented on August 13, from 18.00 to 18.30
- [workinfo] => p. 47
- [year] => 2007
- [title] => Are confidence intervals for binomial distributed samples an optimization meters?
- [authors] => Sorana-Daniela BOLBOACĂ, Lorentz JÄNTSCHI
- [abstract] =>

The aim of the research was to develop an optimization procedure of computing confidence intervals for binomial distributed samples based. An inductive algorithm is proposed method used to solve the problem of confidence intervals estimation for binomial proportions. The implemented optimization procedure uses two triangulations (varying simultaneously two pairs of three variables).

The optimization method was assessed in a simulation study for a significance level of 5%, and sample sizes that vary from six to one thousand and associated possible proportions. The obtained results are available online at the following address: http://l.academicdirect.org/Statistics/binomial_distribution/

Overall, the optimization method performed better, the values of cumulative error function decreasing in average with 10%, depending on the sample sizes and the confidence intervals method with which it is compared.

The performances of the optimization method increase with increasing of the sample size, surprisingly because it is well known that the confidence interval methods that use the normal approximation hypothesis for a binomial distribution obtain good results with increasing of sample sizes.
- [keywords] => Optimization; Confidence interval; Binomial distribution; Contingency table
- [acknowledgment] => This research was partly supported by UEFISCSU Romania through project ET46/2006 and CNCSIS Romania through project AT93/2007.

Lorentz JÄNTSCHI, Sorana-Daniela BOLBOACĂ, Cristina Maria FURDUI,
Characteristic and Counting Polynomials on Modeling Nonane Isomers Properties, Fourth
International Conference of Applied Mathematics and Computing, August 12-18, 2007,
University of Chemical Technology and Metallurgy Sofia & Technical University of Plovdiv,
Invited lecture, presented on August 15, from 12.20 to 13.00, Plovdiv, Bulgaria, p. 234, **2007**.

- [id] => 144
- [recorddate] => 2007:08:19:19:09:57
- [lastupdate] => 2007:08:19:19:09:57
- [type] => conference
- [place] => Plovdiv, Bulgaria
- [subject] => chemistry - computational; chemistry - physical; informatics - fundamentals; mathematics - applied; mathematics - modeling; mathematics - statistics
- [relatedworks] =>
 - 2 (average):
 - Structure-Property Based Model Estimation of Alkanes Boiling Points, [?id=145](#)
- [file] => [?f=144](#)
- [mime] => application/pdf
- [size] => 244136
- [pubname] => Fourth International Conference of Applied Mathematics and Computing, August 12-18, 2007
- [pubinfo] => University of Chemical Technology and Metallurgy Sofia & Technical University of Plovdiv
- [pubkey] => Invited lecture, presented on August 15, from 12.20 to 13.00
- [workinfo] => p. 234
- [year] => 2007
- [title] => Characteristic and Counting Polynomials on Modeling Nonane Isomers Properties
- [authors] => Lorentz JÄNTSCHI, Sorana-Daniela BOLBOACĂ, Cristina Maria FURDUI
- [abstract] =>

The major goal of this study was to investigate the broad application of graph polynomials to the analysis of Henry's law constants of nonane isomers. In this context, Henry's law constants of nonane isomers were modeled using characteristic and counting polynomials and the characteristic and counting polynomials on the distance matrix, on the maximal fragments matrix, on the complement of maximal fragments matrix, and on the Szeged matrix were calculated for each compound.

One of included nonane isomers, 4-methyloctane, was identified as an outlier and was withdrawn from further analysis. This report describes the performance and characteristics of top five significant models. The results show that Henry's law constants of nonane isomers can be modeled by applying characteristic and counting polynomials.
- [keywords] => Characteristic polynomial; Counting polynomials; Nonane isomers; Henry's law constant
- [acknowledgment] => The research was partly supported by UEFISCSU Romania through project ET46/2006.

Carmen E. STOENOIU, Lorentz JÄNTSCHI, Ioan ABRUDAN, Sorana D. BOLBOACĂ,
Romanian Higher Education: Modelling Evolution Tendencies, The International Management Education Conference 2007 Proceedings, Faculty of Business and Economics, University Pendidikan Sultan Idris, ISBN 978-983-3759-19-4, Penang, Malaysia, paper #10§02 (6 pages), **2007**.

- [id] => 141
- [recorddate] => 2007:07:07:18:34:04
- [lastupdate] => 2007:07:07:18:34:04
- [type] => conference
- [place] => Penang, Malaysia
- [subject] => education - evaluation; education - management; education - policy; informatics - applied; mathematics - statistics
- [relatedworks] => N/A
- [file] => [?f=141](#)
- [size] => 872833
- [pubname] => The International Management Education Conference 2007 Proceedings
- [pubinfo] => Faculty of Business and Economics, University Pendidikan Sultan Idris
- [pubkey] => ISBN 978-983-3759-19-4
- [workinfo] => paper #10§02 (6 pages)
- [year] => 2007
- [title] => Romanian Higher Education: Modelling Evolution Tendencies
- [authors] => Carmen E. STOENOIU, Lorentz JÄNTSCHI, Ioan ABRUDAN, Sorana D. BOLBOACĂ
- [abstract] =>

After communist regime downfall (in 1989, December) the Romania Educational System was continuous changed. In 1995 was adopted Law of Education, and each university develops respecting the law, and personal management strategies taking into account the national educational necessities. Starting from the data regarding the number of students from higher education from 1992 to 2005 the aim of the present research is to develop and analyze a mathematical model useful in prediction of students' number for a given year (in our case is 2008). The higher education was analyzed after geographical clusterization of the forty-two Romanian counties. Eight clusters included into analysis: Central, Capital, West, South, South West, South East, North West, and North East (called developing regions). In order to analyze the trends of Romanian higher education a mathematical model has been developed. The model integrates the following parameters: (1) the mean annual variation ratio (as absolute and relative values), (2) the increasing mean annual ratio (as absolute and relative values), (3) the number of students estimation for 1989 (as absolute value), (4) the numbers of students prediction for 2008, (5) the correlation coefficient, (6) the linear trend, (7) the number of students estimation for 1989 and (8) for 2008 obtained by the model. The mathematical model has been integrated into an online program and is available at:
http://vl.academicdirect.ro/applied_statistics/management&policy/education/romania/e_regions/.

The above-described parameters were computed for each item, cluster, and globally. The obtained results regarding the evolutions and predictions are analyzed and discussed. The plan of future development is highlighted.
- [keywords] => Higher Education; Mathematical Model; Romania Regions

Lorentz JÄNTSCHI, Carmen E. STOENOIU, Sorana D. BOLBOACĂ, *Research Policy via Funding Allocation Analysis*, The International Management Education Conference 2007 Proceedings, Faculty of Business and Economics, University Pendidikan Sultan Idris, ISBN 978-983-3759-19-4, Penang, Malaysia, paper #10§03 (7 pages), **2007**.

- [id] => 140
- [recorddate] => 2007:07:07:18:31:56
- [lastupdate] => 2007:07:07:18:31:56
- [type] => conference
- [place] => Penang, Malaysia
- [subject] => informatics - applied; mathematics - statistics; research - evaluation; research - management; research - policy
- [relatedworks] => N/A
- [file] => [?f=140](#)
- [mime] => application/pdf
- [size] => 842190
- [pubname] => The International Management Education Conference 2007 Proceedings
- [pubinfo] => Faculty of Business and Economics, University Pendidikan Sultan Idris
- [pubkey] => ISBN 978-983-3759-19-4
- [workinfo] => paper #10§03 (7 pages)
- [year] => 2007
- [title] => Research Policy via Funding Allocation Analysis
- [authors] => Lorentz JÄNTSCHI, Carmen E. STOENOIU, Sorana D. BOLBOACĂ
- [abstract] =>

The National Council of Scientific Research in Higher Education, created in 1994 as a component of the Romanian educational reform, has the aim to ensure development and consolidation of the scientific research in higher education. Since 1995, an annual grant competition is opened and the research projects are received and analyzed. The eligible projects enter in evaluation process and according with the obtained score, a number of projects receive funds. The purpose of our research was to identify and analyse the research policy in higher education between 1995 and 2004. The allocation of the research funds according with institutions and their geographical clusterization, project type and research field were analyzed. All research projects financed by the National Council of Scientific Research in Higher Education between 1995 and 2004 were included into analysis (a total number of eleven thousand seven hundred and eighty five projects). The following variables were collected and stored for each project: the project type, the commission where the project was evaluated, the title of the project, the institution or university name, the number of points obtained after evaluation, and the value of the funding expressed as USD. The data were summarized and statistical analyzed. The averages of the funds according with the commission, year, university or institution, and domain were compared at a significance level of 5%. According with the obtained results the research funding policy between 1995 and 2004 was discussed.
- [keywords] => Research Policy; Romania National Council of Scientific Research in Higher Education Analysis

Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI, *Modelling the Inhibitory Activity on Carbonic Anhydrase I of Some Substituted Thiadiazole- and Thiadiazoline-Disulfonamides: Integration of Structure Information*, ESCAPE17 - 17th European Symposium on Computer Aided Process Engineering, CAPE Working Party of the European Federation of Chemical Engineering then Elsevier Netherlands & UK, ISBN 978-0-444-53157-5 & eISBN 0-444-53158-2, Bucharest, Romania then Amsterdam, Netherlands, T4-212 (oral presentation), May 27-30 then p. 965-970, **2007**.

- [id] => 131
- [recorddate] => 2007:05:07:17:08:31
- [lastupdate] => 2007:05:30:13:24:00
- [type] => conference
- [place] => Bucharest, Romania then Amsterdam, Netherlands
- [subject] => chemistry - biochemistry; chemistry - computational; informatics - applied; informatics - models implementation; mathematics - modeling; mathematics - statistics; medicine - immunology; medicine - informatics
- [relatedworks] =>
 - 0 (highest):
 - Modelling the Inhibitory Activity on Carbonic Anhydrase I of Some Substituted Thiadiazole and Thiadiazoline-Disulfonamides: Integration of Structure Information, [?id=133](#)
 - 4 (some):
 - Antiallergic Activity of Substituted Benzamides: Characterization, Estimation and Prediction, [?id=137](#)
- [file] => [?f=131](#)
- [mime] => application/pdf
- [size] => 1335735
- [pubname] => ESCAPE17 - 17th European Symposium on Computer Aided Process Engineering
- [pubinfo] => CAPE Working Party of the European Federation of Chemical Engineering then Elsevier Netherlands & UK
- [pubkey] => ISBN 978-0-444-53157-5 & eISBN 0-444-53158-2
- [workinfo] => T4-212 (oral presentation), May 27-30 then p. 965-970
- [year] => 2007
- [title] => Modelling the Inhibitory Activity on Carbonic Anhydrase I of Some Substituted Thiadiazole- and Thiadiazoline-Disulfonamides: Integration of Structure Information
- [authors] => Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI
- [abstract] =>

Aim: The paper presents the abilities in estimation and prediction of the inhibition on carbonic anhydrase I of some substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides through the integration of complex structures information by using of an original molecular descriptors family on the structure-activity relationships approach. Material and Method: The proposed approach uses the complex information obtained from substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides structure in order to generate and calculate the molecular descriptors family. The structure-activity relationship models were built based on the generated descriptors. The obtained multivariate models (the models with two, respectively four descriptors) were validated by computing the cross-validation leave-one-out score (r^2_{cv-loo}), and analyzed through assessment of the squared correlation coefficients (r^2), and the models stability ($r^2 - r^2_{cv}$).

loo). The prediction ability of the multivariate MDF-SAR model with four descriptors was analyzed in training versus test sets.

Results: The best performing MDF-SAR model proved to be the model with four descriptors ($r^2 = 0.9175$). The MDF-SAR model with four descriptors shown that the inhibition on carbonic anhydrase I of substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides is likely to be of geometry and topology nature, being in relation with the partial charge and relative atomic mass of compounds. The estimation ability of this model is sustained by the multiple correlation coefficient ($r = 0.9579$, $95\%CI = [0.9212, 0.9776]$) and by the significance of the model ($F = 97$, $p < 0.001$). The prediction ability is sustained by the cross validation leave-one-out score ($r^2_{cv-loo} = 0.8911$), the model stability ($r^2 - r^2_{cv-loo} = 0.0264$), and by the results on training versus test analysis. The MDF-SAR model with four descriptors proved to render higher value of the correlation coefficient comparing with previously reported model ($p < 0.01$).

Conclusion: Modelling the inhibition activity on carbonic anhydrase I of substituted 1,3,4-thiadiazole- and 1,3,4-thiadiazoline-disulfonamides by integration of complex structure information provide stable MDF-SAR models, revealing that there is a relationship between the compounds structure and their inhibition activity on carbonic anhydrase I.

- [keywords] => Molecular Descriptors Family on Structure-Activity Relationships (MDF-SAR); Substituted 1,3,4-Thiadiazole- and 1,3,4-Thiadiazoline-Disulfonamides; Carbonic Anhydrase I (CA I); Inhibition Activity
- [acknowledgment] => The research was partly supported by UEFISCSU Romania through project ET46/2006.

Lorentz JÄNTSCHI, Sorana D. BOLBOACĂ, *The Jungle of Linear Regression Revisited*, Leonardo Electronic Journal of Practices and Technologies, AcademicDirect, ISSN 1583-1078, www, Internet, 6(10), p. 169-187, **2007**.

- [id] => 130
- [recorddate] => 2007:05:06:17:30:20
- [lastupdate] => 2007:05:20:17:30:20
- [type] => article
- [place] => www, Internet
- [subject] => mathematics - modeling; mathematics - statistics; research - evaluation; research - management
- [relatedworks] => N/A
- [file] => [?f=130](#)
- [mime] => application/pdf
- [size] => 166849
- [pubname] => Leonardo Electronic Journal of Practices and Technologies
- [pubinfo] => AcademicDirect
- [pubkey] => ISSN 1583-1078
- [workinfo] => 6(10), p. 169-187
- [year] => 2007
- [title] => The Jungle of Linear Regression Revisited
- [authors] => Lorentz JÄNTSCHI, Sorana D. BOLBOACĂ
- [abstract] =>
Simple linear regression is reviewed. Some well known facts are analyzed from different approaches. Some new formulas and equations are posted and discussed.
- [keywords] => Simple Linear Regression; Least Squares Method; Independent and Dependent Variables

Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI, *Communication of Results on Risk Factors Studies: Confidence Intervals*, Leonardo Journal of Sciences, AcademicDirect, ISSN 1583-0233, www, Internet, 6(10), p. 179-187, **2007**.

- [id] => 129
- [recorddate] => 2007:05:06:17:26:39
- [lastupdate] => 2007:05:06:17:26:39
- [type] => article
- [place] => www, Internet
- [subject] => mathematics - statistics; medicine - evidence based; medicine - informatics
- [relatedworks] =>
 - 4 (some):
 - Binomial Distribution Sample Confidence Intervals Estimation 1. Sampling and Medical Key Parameters Calculation, [?id=32](#)
- [file] => [?f=129](#)
- [pubname] => Leonardo Journal of Sciences
- [pubinfo] => AcademicDirect
- [pubkey] => ISSN 1583-0233
- [workinfo] => 6(10), p. 179-187
- [year] => 2007
- [title] => Communication of Results on Risk Factors Studies: Confidence Intervals
- [authors] => Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI
- [abstract] =>

Purpose: Starting from the hypothesis that confidence intervals are used in medical research as a criterion of trustworthiness and robustness of findings, the aim of the research was to determine whether the medical parameters communicated as results in abstracts of risk factors studies published in PubMed database and in Romania online journals are accompanied by confidence intervals.

Method: The search strategy included four keywords and some limitations (publication data, the language of publication, and studies on human). Four inclusion criteria were imposed. The obtained results were summarized and analyzed with Statistica.

Results: A number of 3191 were identified after applying the search strategy on PubMed database. Almost three and a half percent (n = 110) remained after inclusion the confidence intervals as keyword. Sixty articles out of one hundred and one accomplished the inclusion criteria. A number of sixty-five articles from nine Romanian online journals were identified. After applying the inclusion criteria, five articles out of sixty-five were included into the analysis. Testing the null hypothesis that there are not significant differences between the number of articles indexed in PubMed database and the number of articles published in Romanian online journals which refer as outcome the relative risk or odds ratio with associated confidence intervals, a p-value less than 0.0001 was obtained.

Conclusion: It can be conclude that the publication standards in Romanian medical journals must by arise in order to become aligned with the international trends and standards.

- [keywords] => Risk Factors; Cohort Study; Relative Risk; Odds Ratio; Confidence Intervals (CIs)
- [acknowledgment] => Research was partly supported through project ET/46/2006 by UEFISCSU Romania.

Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI, *How Good the Characteristic Polynomial Can Be for Correlations?*, International Journal of Molecular Sciences, [Molecular Diversity Preservation International](#), ISSN 1422-0067, www & Basel, Internet & Switzerland, 8(4), p. 335-345, **2007**.

- [id] => 128
- [recorddate] => 2007:05:03:11:48:36
- [lastupdate] => 2007:05:03:11:49:36
- [type] => article
- [place] => www & Basel, Internet & Switzerland
- [subject] => chemistry - computational; chemistry - physical; informatics - applied; informatics - models implementation; mathematics - modeling; mathematics - statistics; research - methodology
- [relatedworks] => N/A
- [file] => [?f=128](#)
- [mime] => application/pdf
- [size] => 196540
- [pubname] => International Journal of Molecular Sciences
- [pubinfo] => [Molecular Diversity Preservation International](#)
- [pubkey] => ISSN 1422-0067
- [workinfo] => 8(4), p. 335-345
- [year] => 2007
- [title] => How Good the Characteristic Polynomial Can Be for Correlations?
- [authors] => Sorana D. BOLBOACĂ, Lorentz JÄNTSCHI
- [abstract] =>

The aim of this study was to investigate the characteristic polynomials resulting from the molecular graphs used as molecular descriptors in the characterization of the properties of chemical compounds. A formal calculus method is proposed in order to identify the value of the characteristic polynomial parameters for which the extremum values of the squared correlation coefficient are obtained in univariate regression models. The developed calculation algorithm was applied to a sample of nonane isomers. The obtained results revealed that the proposed method produced an accurate and unique solution for the best relationship between the characteristic polynomial as molecular descriptor and the property of interest.

- [keywords] => Characteristic polynomial; Graph theory; Structure-Property Relationships; Nonane isomers; Henry's law constant (solubility)
- [acknowledgment] => The research was partly supported by UEFISCSU Romania through project ET46/2006.

Sorana D. BOLBOACĂ, Lorentz JĂNTSCHI, *Data Mining on Structure-Activity/Property Relationships Models*, ECC11- The 11th [Electronic Computational Chemistry Conference](#), online, Monmouth University, New Jersey, USA, N/A, www, Internet, paper #29, Presentation located [here](#), April 2-30, 2007.

- [id] => 126
- [recorddate] => 2007:05:02:12:04:49
- [lastupdate] => 2007:05:02:12:05:01
- [type] => conference
- [place] => www, Internet
- [subject] => chemistry - computational; chemistry - general; informatics - applied; mathematics - applied; mathematics - statistics; medicine - informatics; research - management; research - methodology
- [relatedworks] => N/A
- [file] => [?f=126](#)
- [mime] => application/pdf
- [size] => 15899487
- [pubname] => ECC11- The 11th [Electronic Computational Chemistry Conference](#), online
- [pubinfo] => Monmouth University, New Jersey, USA
- [pubkey] => N/A
- [workinfo] => paper #29, Presentation located [here](#), April 2-30
- [year] => 2007
- [title] => Data Mining on Structure-Activity/Property Relationships Models
- [authors] => Sorana D. BOLBOACĂ, Lorentz JĂNTSCHI
- [abstract] =>
Molecular descriptors family on structure-activity/property relationships studies were carried out in order to identify the link between compounds structure and their activity/property.
A number of fifty-five classes of properties or activities of different compounds sets were investigated. Single and multi-varied linear regression models using molecular descriptors as variables were identified.
The models with estimation and prediction abilities and associated characteristics were stored into a database.
A data mining analysis using classification and clustering were applied on the obtained database for searching and extracting useful information.
The methodology applied in searching and extracting for information and the obtained results are presented.
- [keywords] => Knowledge-Discovery in Database (KDD); Cluster analysis; Structure-Activity/Property Relationships (SAR/SPR); Molecular Descriptors Family (MDF)
- [acknowledgment] => MEC/UEFISCSU Romania, Grant ET.46/2006

Sorana Daniela BOLBOACĂ, Lorentz JÄNTSCHI, *Modeling the Property of Compounds from Structure: Statistical Methods for Models Validation*, Institute of General and Inorganic Chemistry, Bulgarian Academy of Sciences [<http://sizemat.igic.bas.bg>], FP6: EC-INCO-CT-2005-016414 Specific Support Action, Plovdiv, Bulgaria, p. 71, April 19-21, **2007**.

- [id] => 124
- [recorddate] => 2007:04:30:17:08:30
- [lastupdate] => 2007:04:30:17:08:30
- [type] => conference
- [place] => Plovdiv, Bulgaria
- [subject] => chemistry - computational; informatics - applied; informatics - databases; mathematics - statistics; medicine - informatics; research - methodology
- [pubname] => Institute of General and Inorganic Chemistry
- [pubinfo] => Bulgarian Academy of Sciences [<http://sizemat.igic.bas.bg>]
- [pubkey] => FP6: EC-INCO-CT-2005-016414 Specific Support Action
- [workinfo] => p. 71, April 19-21
- [year] => 2007
- [title] => Modeling the Property of Compounds from Structure: Statistical Methods for Models Validation
- [authors] => Sorana Daniela BOLBOACĂ, Lorentz JÄNTSCHI
- [abstract] =>

A molecular descriptors family on structure-property relationships study (MDF-SPR) was conducted in order to model the boiling points of alkanes using the compounds structure information.

The alkanes from C3 to C9 were included into study. Two MDF-SPR models, one with one descriptor and other with two descriptors, were identified. The estimation and prediction of the MDF-SPR models were analyzed. The methods used for validation of the obtained MDF-SPR models are presented.

The correlated correlation analysis was using in order to compare the performances of the obtained MDF-SPR models and of the MDF-SPR models comparing with previous reported model. The Steiger's Z test [1] at a significance level of 5% was applied.

The statistical analysis of the obtained MDF-SPR models demonstrated that the model with two descriptors has greater abilities in estimation and prediction compared with the model with one descriptor. More, the MDF-SPR model with two descriptors has greater abilities in estimation comparing with previous reported model. These observations were also sustained by the results of correlated-correlation analysis.

The multi-varied MDF-SPR model can be used in order to predict the property of interest of studied alkanes without any experiments and measurements, by using the MDF SPR Predictor application [2].

Refs

1. J. H. Steiger, *Psychol. Bull.* 87 (1980) 245.
2. ***, MDF SPR-SAR Predictor, © 2005, Virtual Library of Free Software [cited 2006 March]. Available from: URL:

http://vl.academicdirect.org/molecular_topology/mdf_findings/sar

- [keywords] => Structure-activity relationships (SARs); Method validation; Method evaluation; Molecular Descriptors Family (MDF)
- [acknowledgment] => The research was partly supported by UEFISCSU Romania through projects ET46/2006 & ET108/2006.

Sorana-Daniela BOLBOACĂ, Lorentz JÄNTSCHI, *Computer-based testing on physical chemistry topic: A case study*, International Journal of Education and Development using Information and Communication Technology, The University of the West Indies, Barbados, ISSN 1814-0556, www, Internet, [id=242](#), 3(1), 2007.

- [id] => 122
- [recorddate] => 2007:03:21:19:04:19
- [lastupdate] => 2007:03:21:19:04:19
- [type] => article
- [place] => www, Internet
- [subject] => chemistry - general; chemistry - physical; education - evaluation; education - management; education - methodology; education - policy; informatics - applied; informatics - databases; informatics - web programming
- [relatedworks] => N/A
- [file] => [?f=122](#)
- [mime] => application/pdf
- [size] => 130281
- [pubname] => International Journal of Education and Development using Information and Communication Technology
- [pubinfo] => The University of the West Indies, Barbados
- [pubkey] => ISSN 1814-0556
- [workinfo] => [id=242](#), 3(1)
- [year] => 2007
- [title] => Computer-based testing on physical chemistry topic: A case study
- [authors] => Sorana-Daniela BOLBOACĂ, Lorentz JÄNTSCHI
- [abstract] =>
According with national trends in the objective evaluation of undergraduate students' knowledge, an auto-calibrated online evaluation system was developed. The aim of the research was to assess the knowledge on physical chemistry topic of the undergraduate first year students' at the Faculty of Materials Science and Engineering, the Technical University of Cluj-Napoca, Romania by the use of the developed auto-calibrated system. The methodology of multiple-choice questions construction and the evaluation methodology are presented. The students performances in terms of number of correct answers and time needed to give a correct answer were collected and analyzed. The future plans of system development are highlighted.
- [keywords] => auto-calibrated online evaluation; multiple choice questions (MCQs); physical chemistry; undergraduate students
- [acknowledgment] => The research was partly supported by UEFISCSU Romania through the project ET46/2006.

Concluzii

Statistica aplicată oferă un instrument puternic de investigare pe eşantioane de populații. Atunci când volumul eşantionului este redus, asumția distribuirii acestuia după o distribuție Gauss (normală) este eronată. Este necesar să se folosească distribuția binomială.

Optimizarea numerică oferă un răspuns real la problema eşantioanelor distribuite binomial.

Următoarele aplicații oferă soluții reale:

- calculator pentru intervalele de confidență la proporții binomiale:
http://vl.academicdirect.org/applied_statistics/binomial_distribution/
- sistem de evaluare autocalibrată online: http://vl.academicdirect.org/general_chemistry/
- sistem de educație medicală continuă: http://vl.academicdirect.org/medical_informatics/
- sistem de construcție a polinoamelor caracteristice și de numărare pe grafuri:
http://vl.academicdirect.org/molecular_topology/graph_polynomials/
- sistem de generare a tăierilor de vârfuri în grafuri:
http://vl.academicdirect.org/molecular_topology/vertex_cutting/

Activitățile prevăzute anului 2007 au fost realizate. Obiectivele propuse pentru anul 2007 au fost atinse. Rezultatele estimate au fost obținute.

Cluj-Napoca,
la 27.10.2007

Director temă ET46/2006,
Șef. L. Dr., Ing. Lorentz JÄNTSCHI
<http://lori.academicdirect.org>
