

تحلیل علمی انتخابات ریاست جمهوری ۱۳۸۸

سیدمهدی سیدنصراله

پژوهشگر – کارشناس ارشد دانشگاه صنعتی شریف

smsnasrollah@yahoo.com

با توجه با داغ بودن موضوع صحت انتخابات در روزهای اخیر، مقاله ای با عنوان "مستندات علمی دستکاری در آراء انتخابات دهمین دوره انتخابات" در پایگاههای متعدد اینترنتی مرا بر آن داشت تا محاسبات ادعایی در آن مقاله را بررسی کنم که نتیجه این امر چیزی جز رد آن مقاله و اطمینان از نتایج انتخابات نبود. در مقاله ادعایی هیچ نموداری رسم نشده بود تا اصل ماجرا مشخص گردد و البته قسمتهای اصلی مقاله یعنی محاسبه میزان خطا و انحراف، کاملاً اشتباه بود و تعمداً از مقادیر نرمال نشده استفاده شده بود و نتیجه دیگری را به خواننده القا می کرد. اما در مقاله پیش رو سه آزمون قدرتمند برای بررسی علمی نتایج انجام شده است تا به طور کامل صحت داده ها تحلیل گردد. امید است تا خوانندگان گرامی، نگارنده را از نظرات خود بهره مند سازند.

طی چند روز گذشته بحثهای فراوانی در خصوص امکان و احتمال تقلب در انتخابات ایران بعمل آمده است. با توجه به اینکه بازشماری آراء کاری پرزحمت برای دستگاه اجرایی و نظارتی انتخابات خواهد بود باید، کاندیداهای معترض شواهد و مستندات قابل قبولی را به این مراجعه اعلام نمایند تا دستکاری در نتایج انتخابات را آشکارا نشان دهد. راههای علمی زیادی وجود دارد که با استفاده از آمارهای اعلام شده توسط وزارت کشور این دستکاری ها در صورت وجود به نمایش گذاشته شود و یا صحت انتخابات از لحاظ شاخصهای علمی آمار بررسی شود.

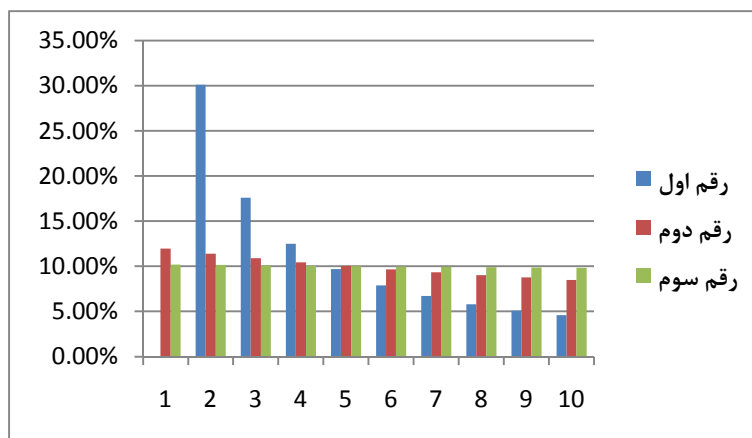
صحبت از تقلب در انتخابات تنها یک موضوع مختص به ایران و حتی کشورهای در حال توسعه نظیر ونزوئلا، مکزیک، و زیمباوه نیست و در سال ۲۰۰۰ شاهد بودیم که بحث تقلب در انتخابات ریاست جمهوری در ایالت فلوریدا به موضوع داغ آنروزها بدل شده بود. به هر حال آنچه مسلم است آنکه اهمیت سیاست و موضوع تقلب در انتخابات سبب شده که دانشمندان علم آمار با استفاده از کمترین داده های موجود از انتخابات روشهای آماری را طراحی کنند که با کمک آن بتوان با صراحت در مورد سلامت انتخابات اظهار نظر نمایند. یکی از معمول ترین

این روشها روشی است که از قانون بن فورد (Benford's Law) پیروی می‌کند. بطور خیلی مختصر و قابل درک برای خوانندگان این قانون به آن اشاره دارد که رقمهای مربوط به یک عدد (خواه یک رقمی یا بیش از یک رقمی) حاصل از شمارش یک پدیده در جهان خارج (جهان واقعی) از توزیع نرمال (Normal distribution) یا توزیع یک فرم (Uniform distribution) پیروی نمی‌کند بلکه از توزیعی شبیه توزیع Chi Square پیروی می‌کند. آقای بن‌فورد احتمال این اعداد را برای حوادثی چون انتخابات محاسبه کرده است. قابل ذکر است که اعداد ذکر شده در قانون بن‌فورد همانند قوانین نیوتن یک واقعیت علمی است و در حال حاضر علاوه بر مسئله انتخابات، برای رسیدگی به احتمال تقلب در سود سهام شرکتها و مسائل مالیاتی نیز از همین روش استفاده می‌شود.

اجازه بدهید چند خط از این مقاله را به توضیح ساده این قانون پردازیم. در این قانون احتمال اینکه آخرین عدد سمت چپ یک عدد (فرض کنید تعداد آراء اخذ شده به نفع یک کاندیدا) که می‌تواند یکی از اعداد ۱ تا ۹ باشد با هم برابر نیست. بلکه احتمال آنکه رقم سمت چپ یک عدد، ۱ باشد حدود ۳۰٪ است در حالی که این احتمال برای عدد ۹ در حدود ۴.۵٪. همچنین است احتمال وجود عدد ۰ تا ۹ برای دومین رقم سمت چپ تعداد آراء اخذ شده به نفع یک کاندیدا که آن نیز از قانون دیگری پیروی می‌کند.

طبق این قانون احتمال ظهور اعداد در رقم اول از رابطه ۱ تبعیت می‌کند:

$$P_D = \frac{\int_D^{D+1} P(x) dx}{\int_1^{10} P(x) dx} = \log\left(1 + \frac{1}{D}\right) \quad (1)$$



شکل ۱- توزیع ستونی احتمالی توزیع اعداد در رقم اول، دوم و سوم طبق قانون بن فورد

عدد	احتمال رقم اول	احتمال رقم دوم	احتمال رقم سوم
۰	٪۰	٪۱۱.۹۶۸	٪۱۰.۱۷۸
۱	٪۳۰.۱	٪۱۱.۳۸۹	٪۱۰.۱۳۸
۲	٪۱۷.۵۶	٪۱۰.۸۸۲	٪۱۰.۰۹۷
۳	٪۱۲.۵	٪۱۰.۴۳۳	٪۱۰.۰۵۷
۴	٪۹.۷	٪۱۰.۰۳۱	٪۱۰.۰۱۸
۵	٪۷.۹	٪۹.۶۶۸	٪۹.۹۹۷
۶	٪۶.۷	٪۹.۳۳۷	٪۹.۹۴
۷	٪۵.۸	٪۹.۰۳۵	٪۹.۹۰۲
۸	٪۵.۱	٪۸.۷۵۷	٪۹.۸۶۴
۹	٪۴.۶	٪۸.۵	٪۹.۸۲۷

جدول ۱- توزیع اعداد در رقم اول طبق قانون بن فورد

برای توضیح بیشتر به توزیع رقمهای اول، دوم و سوم یک دسته تصادفی می‌پردازیم، برای نمونه یک دسته ۲۰۰۰۰ تایی اعداد تصادفی که به پیوست خواهد آمد را در نظر می‌گیریم و توزیع اعداد ۱ تا ۹ را در ارقام اول تا سوم آن در جدول ۲ بررسی می‌کنیم.

رقم سوم	تعداد	درصد رقم سوم	رقم دوم	تعداد	درصد رقم دوم	رقم اول	تعداد	درصد رقم اول
0	2138	10.69%	0	1948	9.74%	0	0	0.00%
1	1969	9.85%	1	2076	10.38%	1	2205	11.03%
2	1991	9.96%	2	1981	9.91%	2	2291	11.46%
3	2002	10.01%	3	1936	9.68%	3	2269	11.35%
4	1948	9.74%	4	2019	10.10%	4	2251	11.26%
5	2011	10.06%	5	2033	10.17%	5	2193	10.97%
6	1972	9.86%	6	2016	10.08%	6	2172	10.86%
7	1934	9.67%	7	1985	9.93%	7	2210	11.05%
8	2015	10.08%	8	2001	10.01%	8	2160	10.80%
9	2020	10.10%	9	2005	10.03%	9	2249	11.25%
مجموع	20000	100.00%	مجموع	20000	100.00%	مجموع	20000	100.00%

جدول ۲- توزیع یکنواخت اعداد در رقم اول، دوم و سوم در یک دسته تصادفی

همان گونه که در جدول ۲ دیده می‌شود، توزیع اعداد در رقمهای اول، دوم و سوم یک دسته تصادفی کاملاً یکنواخت بوده و این تفاوت، اصلی ترین تفاوت بین داده‌های به دست آمده از طبیعت و داده‌های دستکاری شده

می باشد. چرا که همان گونه که ذکر شد، طبق قانون بن‌فورد برای داده‌های طبیعی این نسبتها یکسان نیستند و مطابق جدول ۱ می‌باشند.

برای اینکه اهمیت قانون بن‌فورد بهتر دیده شود، جدول ۳ که شامل مثالهایی از طبیعت و تطابق آن با قانون بن‌فورد است، نشان داده شده است.

تعداد نمونه ها	1	2	3	4	5	6	7	8	9	کمیت
335	31	16.4	10.7	11.3	7.2	8.6	5.5	4.2	5.1	مساحت رودخانه ها
3259	33.9	20.4	14.2	8.1	7.2	6.2	4.1	3.7	2.2	جمعیت
104	41.3	14.4	4.8	8.6	10.6	5.8	1	2.9	10.6	ثابتهای فیزیکی
100	30	18	12	10	8	6	6	5	5	روزنامه ها
1389	24	18.4	16.2	14.6	10.6	4.1	3.2	4.8	4.1	گرمای مخصوص
703	29.6	18.3	12.8	9.8	8.3	6.4	5.7	4.4	4.7	فشار
690	30	18.4	11.9	10.8	8.1	7	5.1	5.1	3.6	افت توان
1800	26.7	25.2	15.4	10.8	6.7	5.1	4.1	2.8	3.2	وزن مولکولی
159	27.1	23.9	13.8	12.6	8.2	5	5	2.5	1.9	فاضلاب
91	47.2	18.7	5.5	4.4	6.6	4.4	3.3	4.4	5.5	وزن اتمی
5000	25.7	20.3	9.7	6.8	6.6	6.8	7.2	8	8.9	مجذور و معکوس اعداد
560	26.8	14.8	14.3	7.5	8.3	8.4	7	7.3	5.6	شاخصهای طراحی
741	32.4	18.8	10.1	10.1	9.8	5.5	4.7	5.5	3.1	هزینه اطلاعات
707	27.9	17.5	14.4	9	8.1	7.4	5.1	5.8	4.8	ولتاژ اشعه ایکس
1458	32.7	17.6	12.6	9.8	7.4	6.4	4.9	5.6	3	مسابقات آمریکا
1165	31	17.3	14.1	8.7	6.6	7	5.2	4.7	5.4	تابشگر کامل
342	28.9	19.2	12.6	8.8	8.5	6.4	5.6	5	5	آدرسها
900	25.3	16	12	10	8.5	8.8	6.8	7.1	5.5	فاکتوریل اعداد
418	27	18.6	15.7	9.4	6.7	6.5	7.2	4.8	4.1	نرخ مرگ و میر
	30.6	18.5	12.4	9.4	8	6.4	5.1	4.9	4.7	میانگین
	0.8	0.4	0.4	0.3	0.2	0.2	0.2	0.3	0.3	خطای احتمالی

جدول ۳- توزیع اعداد در رقم اول در کمیتهای طبیعی و مقایسه با قانون بن‌فورد

همان گونه که در جدول ۳ دیده می‌شود اعداد به دست آمده از طبیعت به طرز شگفت‌انگیزی در انواع کمیتهای از توزیع پیش‌بینی شده بن‌فورد تبعیت می‌کند و البته جمعیت افراد و نتایج به دست آمده بر مبنای جمعیت نیز جزء این کمیت هاست.

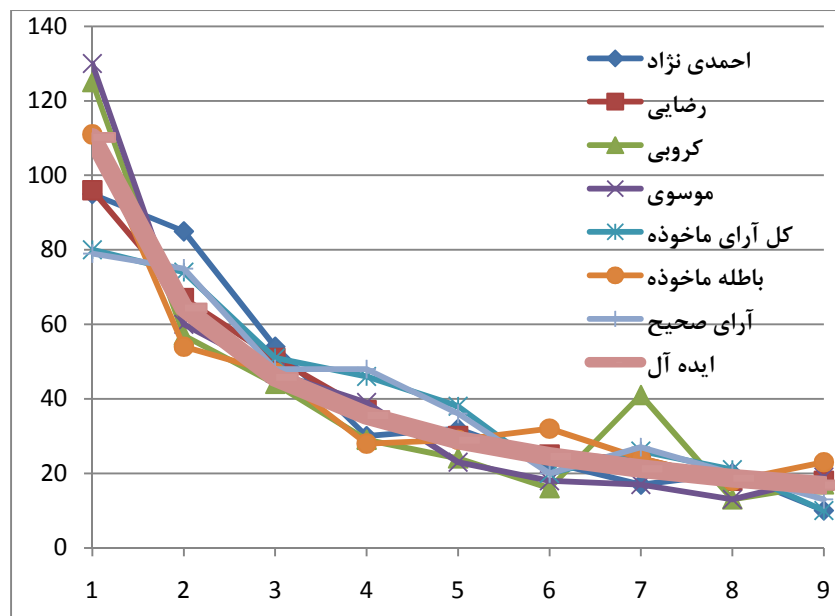
آزمون رقم اول

حال اجازه دهید نتایج انتخابات دوره دهم را از این طریق تحلیل و ارزیابی کنیم. در این قسمت با شمارش تکرار اعداد ۱، ۲، ۳، ۴ در اولین رقم سمت چپ تعداد آرای هر نامزد در ۳۶۶ شهرستان جدول ۴ ارائه می‌گردد.

ایده آل	آرای صحیح	باطله ماخوذه	کل آرای ماخوذه	موسوی	کروبی	رضایی	احمدی نژاد	رقم اول
110.166	79	111	80	130	125	96	95	1
64.416	75	54	74	60	57	67	85	2
45.75	48	47	51	47	44	51	54	3
35.502	48	28	46	39	29	37	30	4
28.914	36	29	38	23	24	30	32	5
24.522	20	32	20	18	16	25	23	6
21.228	27	24	26	17	41	24	17	7
18.666	20	18	21	13	13	18	20	8
16.836	13	23	10	19	17	18	10	9
366	366	366	366	366	366	366	366	مجموع

جدول ۴- تعداد تکرار اعداد ۱، ۲، ۳، ۴ در اولین رقم سمت چپ تعداد آرای هر نامزد در ۳۶۶ شهرستان

نمودار ۱ گویای نتایج جدول ۴ است که مقایسه این اعداد با حالت ایده آل که در واقع معیار بن‌فورد می‌باشد، نشان دهنده تطابق این اعداد با مدل بن‌فورد است.



نمودار ۱- مقایسه توزیع رقم اول نتایج آرای نامزدها با توزیع پیش‌بینی شده بن‌فورد

اگر بخواهیم این اعداد را به صورت درصد نسبی نشان دهیم جدول ۵ ارائه دهنده این اعداد است.

نسبت توزیع اعداد در رقم اول								نسبت ایده آل
1	0.26	0.26	0.342	0.355191	0.21858	0.303279	0.215847	0.301
2	0.232	0.18	0.156	0.163934	0.20219	0.147541	0.204918	0.176
3	0.148	0.14	0.12	0.128415	0.13934	0.128415	0.131148	0.125
4	0.082	0.1	0.079	0.106557	0.12568	0.076503	0.131148	0.097
5	0.087	0.08	0.066	0.062842	0.10383	0.079235	0.098361	0.079
6	0.063	0.07	0.044	0.04918	0.05464	0.087432	0.054645	0.067
7	0.046	0.07	0.112	0.046448	0.07104	0.065574	0.07377	0.058
8	0.055	0.05	0.036	0.035519	0.05738	0.04918	0.054645	0.051
9	0.027	0.05	0.046	0.051913	0.02732	0.062842	0.035519	0.046
مجموع	1	1	1	1	1	1	1	1

جدول ۵- نسبت تکرار اعداد ۱، ۲، ۳، ...، ۹ در اولین رقم سمت چپ تعداد آرای هر نامزد در ۳۶۶ شهرستان

اگر بخواهیم میزان این اختلاف را به صورت علمی حساب کنیم، باید از روشهای معتبر محاسبه خطا مانند-Chi Squared استفاده کنیم، بر مبنای این رابطه داریم:

$$X_s^2 = \sum_{i=1}^k \frac{(m_i - Np_i)^2}{Np_i} \quad (2)$$

در این رابطه m نشان دهنده میزان واقعی کمیت و p احتمال وقوع آن است که با ضرب تعداد کمیتها (N) در آن به تعداد پیش‌بینی شده در توزیع ایده آل می‌رسیم. جدول ۶ بر مبنای این محاسبات تنظیم شده است:

رقم اول	احمدی نژاد	رضایی	کروبی	موسوی	کل آرای ماخوذه	باطله ماخوذه	آرای صحیح
۱	2.087827	1.821574	1.997418	3.570862	8.26014883	0.00631371	8.816872
۲	6.577575	0.103655	0.853779	0.302736	1.42593542	1.68425633	1.739025
۳	1.487705	0.602459	0.06694	0.034153	0.60245902	0.03415301	0.110656
۴	0.852684	0.063208	1.190806	0.344657	3.10427593	1.58526291	4.399752
۵	0.32937	0.04079	0.835145	1.209635	2.85520495	0.00025579	1.736577
۶	0.094466	0.009318	2.961605	1.734625	0.83388321	2.28042101	0.833883
۷	0.842095	0.361974	18.41587	0.842095	1.07273337	0.361974	1.569436
۸	0.095337	0.023763	1.719895	1.719895	0.29184378	0.02376278	0.095337
۹	2.775653	0.080476	0.001598	0.278148	2.77565312	2.25676503	0.874014
میزان خطا	15.14271	3.107217	28.04305	10.03681	21.2221376	8.23316457	20.17555
خطا نسبی	0.041374	0.00849	0.07662	0.027423	0.05798398	0.02249499	0.055124

جدول ۶- محاسبه میزان خطای Chi-Squared برای رقم اول داده‌ها

همان گونه که دیده می‌شود، بیشترین خطای نسبی ۰/۰۷ می باشد، که با توجه به تعداد نه چندان زیاد داده‌ها (۳۶۶ شهرستان) نتیجه بسیار خوبی برای اعتبار نتایج است.

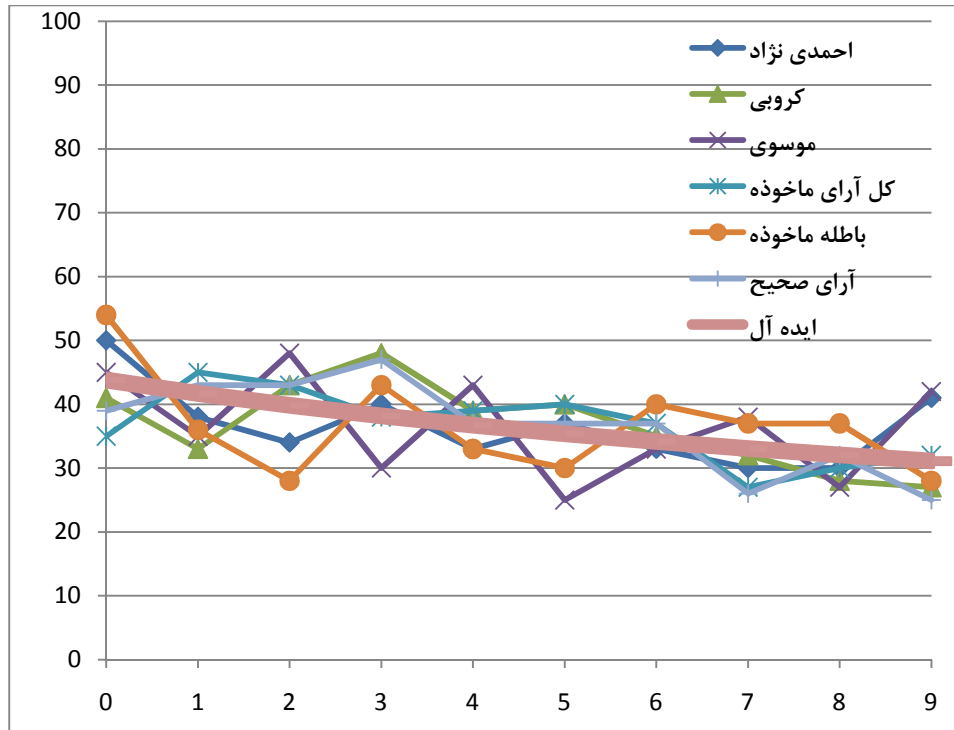
آزمون رقم دوم

برای بررسی دقیق تر صحت اعداد اعلام شده، گاهی اوقات رقم دوم اعداد را نیز بررسی می‌کنند و برای صحت نتایج آن را با پیش‌بینی بنفورد می‌سنجند. اگر آمار ارائه شده وزارت کشور را برای نتایج آرای نامزدها در شهرستانها بررسی کنیم برای توزیع اعداد در رقم دوم، نتایج زیر به دست می‌آید.

توزیع اعداد در رقم دوم								
رقم دوم	ایده آل	آرای صحیح	باطله	ماخوذه کل آرای	ماخوذه موسوی	کروبی	رضایی	احمدی نژاد
0	43.80288	39	54	35	45	41	37	50
1	41.68374	43	36	45	35	33	47	38
2	39.82812	43	28	43	48	43	23	34
3	38.18478	47	43	38	30	48	38	40
4	36.71346	37	33	39	43	39	34	33
5	35.38488	37	30	40	25	40	51	37
6	34.17342	37	40	37	33	35	37	33
7	33.0681	26	37	27	38	32	35	30
8	32.05062	32	37	30	27	28	36	30
9	31.11	25	28	32	42	27	28	41
مجموع	366	366	366	366	366	366	366	366

جدول ۷- تعداد تکرار اعداد ۰، ۱، ۲، ۳، ...، ۹ در دومین رقم سمت چپ تعداد آرای هر نامزد در ۳۶۶ شهرستان

همان گونه که در جدول ۷ و نمودار ۲ نشان داده شده است نتایج آرای تمامی نامزدها با مدل پیش‌بینی بن‌فورد تطابق بسیار خوبی دارد.



نمودار ۲- مقایسه توزیع رقم دوم نتایج آرای نامزدها با توزیع پیش بینی شده بن فورد

اگر نتایج جدول ۷ را به صورت نسبی نشان دهیم به جدول ۸ می‌رسیم.

نسبت توزیع اعداد در رقم دوم								نسبت ایده آل
0	0.136612	0.101093	0.112022	0.122951	0.095628	0.147541	0.106557	0.11968
1	0.103825	0.128415	0.090164	0.095628	0.122951	0.098361	0.117486	0.11389
2	0.092896	0.062842	0.117486	0.131148	0.117486	0.076503	0.117486	0.10882
3	0.10929	0.103825	0.131148	0.081967	0.103825	0.117486	0.128415	0.10433
4	0.090164	0.092896	0.106557	0.117486	0.106557	0.090164	0.101093	0.10031
5	0.101093	0.139344	0.10929	0.068306	0.10929	0.081967	0.101093	0.09668
6	0.090164	0.101093	0.095628	0.090164	0.101093	0.10929	0.101093	0.09337
7	0.081967	0.095628	0.087432	0.103825	0.07377	0.101093	0.071038	0.09035
8	0.081967	0.098361	0.076503	0.07377	0.081967	0.101093	0.087432	0.08757
9	0.112022	0.076503	0.07377	0.114754	0.087432	0.076503	0.068306	0.085
مجموع	1	1	1	1	1	1	1	1

جدول ۸- نسبت تکرار اعداد ۰، ۱، ۲، ۳، ...، ۹ در دومین رقم سمت چپ تعداد آرای هر نامزد در ۳۶۶ شهرستان

برای محاسبه خطا، این بار نیز با استفاده از محاسبه Chi-Squared به تحلیل نتایج می‌پردازیم، که جدول ۹ به همین منظور تهیه شده است.

رقم دوم	احمدی نژاد	رضایی	کروبی	موسوی	کل آرای ماخوذه	باطله ماخوذه	آرای صحیح
۰	0.876753	1.056533	0.179352	0.032717	1.769078	2.373845	0.526624
۱	0.325545	0.678025	1.809035	1.071698	0.263834	0.775	0.041564
۲	0.852839	7.110193	0.252606	1.676695	0.252606	3.512705	0.252606
۳	0.086292	0.000894	2.522957	1.75438	0.000894	0.607214	2.035054
۴	0.375606	0.200549	0.142407	1.07646	0.142407	0.375606	0.002236
۵	0.073721	6.890852	0.601933	3.047791	0.601933	0.819472	0.073721
۶	0.040292	0.233794	0.019993	0.040292	0.233794	0.993434	0.233794
۷	0.284662	0.112865	0.0345	0.735562	1.113515	0.467515	1.510762
۸	0.1312	0.486655	0.511925	0.79589	0.1312	0.764302	7.99E-05
۹	3.144073	0.3109	0.54298	3.812025	0.025461	0.3109	1.200003
میزان خطا	6.190982	17.08126	6.617688	14.04351	4.534723	10.99999	5.876446
خطا نسبی	0.016915	0.04667	0.018081	0.03837	0.01239	0.030055	0.016056

جدول ۹- محاسبه میزان خطای Chi-Squared برای رقم دوم داده‌ها

همانگونه که دیده می‌شود، بیشترین خطای نسبی برای این آزمون ۰/۰۴ می باشد، که این میزان نیز با توجه به تعداد نه چندان زیاد داده‌ها (۳۶۶ شهرستان) نتیجه بسیار خوبی برای اعتبار نتایج است.

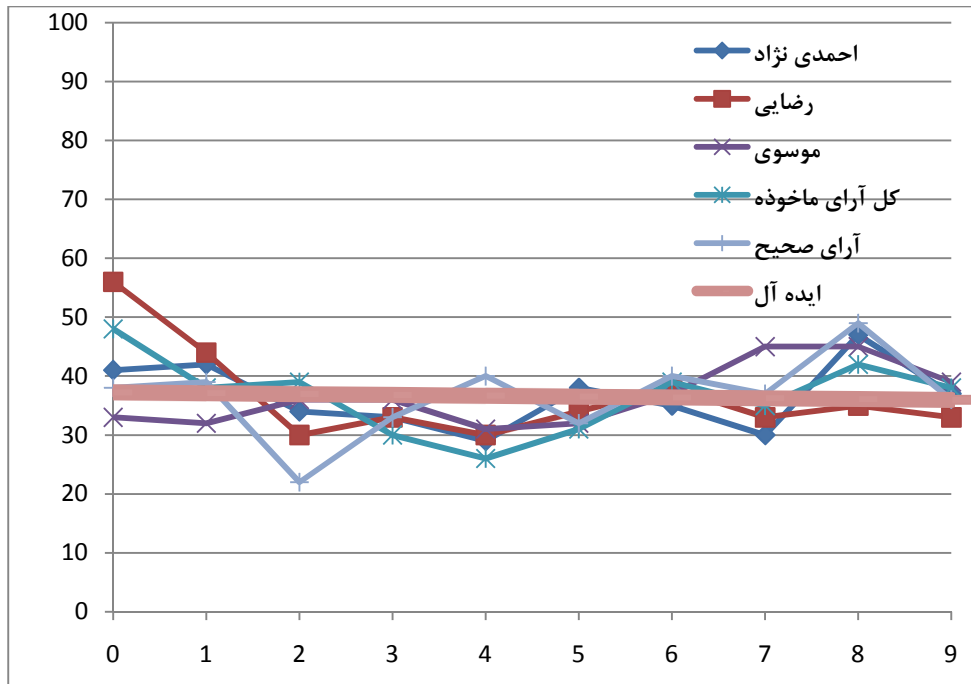
آزمون رقم سوم

اگر بخواهیم حد اطمینان را افزایش دهیم، می‌توان تمامی این محاسبات را برای رقم سوم نیز تکرار کرد، البته برای این منظور باید از معیار رقم سوم بن‌فورد استفاده کرد. با ادامه محاسبات برای بررسی نهایی، جداول و نمودارهای ذیل خواهند آمد. شایان ذکر است در این آزمون از اطلاعات مربوط به یکی از نامزدها به علت اینکه بسیاری از آرا ایشان در شهرستانها دو رقمی و فاقد رقم سوم بود، حذف گردیده است.

توزیع اعداد در رقم سوم							
ایده آل	آرای صحیح	باطله ماخوذه	کل آرای ماخوذه	موسوی	رضایی	احمدی نژاد	رقم سوم
0	38	66	48	33	56	41	0
1	39	35	38	32	44	42	1
2	22	36	39	36	30	34	2
3	33	42	30	36	33	33	3
4	40	30	26	31	30	29	4
5	32	34	31	32	34	38	5
6	40	29	39	37	38	35	6
7	37	31	35	45	33	30	7
8	49	29	42	45	35	47	8
9	36	34	38	39	33	37	9
مجموع	366	366	366	366	366	366	

جدول ۱۰- تعداد تکرار اعداد ۰، ۱، ۲، ۳، ...، ۹ در سومین رقم سمت چپ تعداد آرای هر نامزد در ۳۶۶ شهرستان

نمودار ۳ گویای نتایج آزمون رقم سوم است:



نمودار ۳- مقایسه توزیع رقم سوم نتایج آرای نامزدها با توزیع پیش بینی شده بن فورد

بر همین اساس جدول ۱۱ نسبت تکرار اعداد ۰، ۱، ۲، ۳، ...، ۹ در سومین رقم سمت چپ تعداد آرای هر نامزد را نشان می دهد.

نسبت توزیع اعداد در رقم سوم							نسبت ایده آل
0	0.112022	0.153005	0.090164	0.131148	0.180328	0.103825	0.10178
1	0.114754	0.120219	0.087432	0.103825	0.095628	0.106557	0.10138
2	0.092896	0.081967	0.098361	0.106557	0.098361	0.060109	0.10097
3	0.090164	0.090164	0.098361	0.081967	0.114754	0.090164	0.10057
4	0.079235	0.081967	0.084699	0.071038	0.081967	0.10929	0.10018
5	0.103825	0.092896	0.087432	0.084699	0.092896	0.087432	0.09979
6	0.095628	0.103825	0.101093	0.106557	0.079235	0.10929	0.0994
7	0.081967	0.090164	0.122951	0.095628	0.084699	0.101093	0.09902
8	0.128415	0.095628	0.122951	0.114754	0.079235	0.13388	0.09864
9	0.101093	0.090164	0.106557	0.103825	0.092896	0.098361	0.09827
مجموع	1	1	1	1	1	1	1

جدول ۱۱- نسبت تکرار اعداد ۰، ۱، ۲، ۳، ...، ۹ در سومین رقم سمت چپ تعداد آرای هر نامزد در ۳۶۶ شهرستان

رقم سوم	احمدی نژاد	رضایی	موسوی	کل آرای ماخوذه	باطله ماخوذه	آرای صحیح
۰	0.377204	9.436055	0.485218	3.101372	22.18643	0.015041
۱	0.64574	1.281224	0.702379	0.021584	0.119427	0.096772
۲	0.236291	1.308951	0.02468	0.113163	0.02468	6.052023
۳	0.394081	0.394081	0.017764	1.259414	0.732177	0.394081
۴	1.602736	1.211861	0.875533	3.102639	1.211861	0.30318
۵	0.059719	0.174307	0.56016	0.835226	0.174307	0.56016
۶	0.052377	0.072102	0.010552	0.188626	1.497243	0.360125
۷	1.074853	0.289894	2.116768	0.042517	0.758014	0.015882
۸	3.289579	0.033653	2.192942	0.963474	1.397193	4.607809
۹	0.029679	0.244726	0.255796	0.114934	0.107554	3.06E-05
میزان خطا	7.762259	14.44685	7.241794	9.74295	28.20889	12.4051
خطا نسبی	0.021208	0.039472	0.019786	0.02662	0.077073	0.033894

جدول ۱۲- محاسبه میزان خطای Chi-Squared برای رقم سوم داده‌ها

همانگونه که دیده می‌شود، بیشترین خطای نسبی برای این آزمون ۰/۰۳ می باشد، که این میزان نیز با توجه به تعداد نه چندان زیاد داده‌ها (۳۶۶ شهرستان) نتیجه بسیار خوبی برای اعتبار نتایج است.

نتیجه گیری نهایی

همان گونه که بیان شد، امروزه برای اعتبار سنجی علمی نتایج انتخاباتها در سراسر دنیا روشهای خوبی وجود دارد، که یکی از اصلی ترین این روشها روش توزیع *Benford* بر مبنای تعداد ظاهر شدن اعداد ۱ تا ۹ در رقمهای اول تا سوم نتایج تفکیکی است. در این مقاله سه روش قدرتمند بر داده های استخراج شده مربوط به آرای تفکیکی شهرستانها از پایگاه اینترنتی وزارت کشور www.moi.ir اعمال گردید که همگی حاکی از صحت علمی نتایج انتخابات است و بنابراین فرض دستکاری گسترده در آرا و عددسازی مردود است.

منابع و مراجع

۱. پایگاه اینترنتی وزارت کشور www.moi.ir
2. "Detecting Problems in Survey Data using Benford's Law", George Judge, University of California at Berkeley, Laura Schechter, University of Wisconsin at Madison, November 1, 2007

3. “*The Effective Use of Benford’s Law to Assist in Detecting Fraud in Accounting Data*”, Cindy Durtschi, William Hillison, Carl Pacini, Journal of Forensic Accounting, Vol. V 2004, pp 17-34
4. “*Election Forensics: Vote Counts and Benford's Law*”, Walter R. Mebane, Jr., July 17, 2006
5. <http://mathworld.wolfram.com/Chi-SquaredTest.html>
6. <http://mathworld.wolfram.com/BenfordsLaw.html>
7. http://en.wikipedia.org/wiki/Benford's_law
8. http://en.wikipedia.org/wiki/Chi-square_distribution
9. <http://en.wikipedia.org/wiki/P-value>