

Excel 在随机抽样检测数据统计分析中的运用

吕东明 邹凤 王成军

(无锡市食品安全检验检测中心 江苏无锡 214101)

摘要: Excel 具有强大的计算功能,利用它可以进行专业的统计运算,对相关变量进行回归分析,对定值数据进行质量控制、实验室比对等多有文章介绍,但对随机抽样检测数据的分析介绍较少,本文利用 Excel 中的 NORMDIST 函数,对某第三方检测机构随机抽样油品中某有害成分(地沟油标志物之一)进行随机抽样检测的结果进行统计和分析。

关键词: Excel; NORMDIST; 正态分布

The application of Excel in the statistical analysis of random sample testing data

Lvdongming Zoufeng Wangchengjun

Abstract: Excel has advanced calculation functions, with numerous articles on specialized statistical operations, regression analysis of correlated variables, quality control of fixed-value data, laboratory comparisons, etc., but relatively few analyses of randomly sampled test data have been presented. This paper utilizes the NORMDIST function in Excel to statistically and analytically analyze the results of random sampling of a hazardous ingredient (one of the markers of gutter oil) in a randomly sampled oil product by a third-party testing organization.

Keywords: Excel, NORMDIST, Normal Distribution.

Excel 作为微软 OFFICE 办公软件的重要组成部分，其函数功能强大，对相关数据分析有一定优势，对相关固定的应用，也可以利用 Excel VBA 进行编程开发相应程序，进行自动分析，方便使用者的具体运用，下面以某第三方检测机构随机抽样 200 批油品中某有害成分进行风险分析的结果为例，（部分数据为虚拟）来对该结果进行分析。

10 人次检测结果如下表:

表 1

人员 数值	甲	乙	丙	丁	戊	己	庚	辛	壬	癸
1	0.06	0.14	0.32	0.24	0.15	0.19	0.43	0.56	0.98	1.56
2	0.23	0.07	0.23	0.3	0.26	0.53	0.25	0.59	0.42	3.74
3	0.09	0.24	0.12	0.11	0.32	0.34	1.06	0.43	2.75	205.32
4	0.13	0.12	0.15	0.06	0.11	0.37	0.62	0.31	1.08	5.43
5	0.19	0.09	0.25	0.46	0.49	0.25	0.51	1.89	0.83	0.69
6	0.05	0.32	0.21	0.27	0.24	0.17	0.22	0.45	1.44	0.86
7	0.27	0.23	0.11	0.16	0.28	0.19	0.13	1.05	0.66	2.34
8	0.22	0.17	0.17	0.18	0.15	0.42	0.76	0.67	0.63	4.32
9	0.10	0.08	0.36	0.32	0.42	0.15	0.34	0.88	1.06	7.69

10	0.14	0.37	0.27	0.23	0.35	0.24	0.65	1.34	0.24	0.57
11	0.07	0.12	0.21	0.12	0.16	0.23	0.73	0.26	0.67	4.78
12	0.35	0.26	0.16	0.09	0.13	0.19	0.32	0.93	2.05	0.77
13	0.24	0.15	0.12	0.25	0.27	0.29	0.54	0.42	0.43	5.23
14	0.21	0.06	0.14	0.21	0.32	0.35	0.50	0.37	0.74	2.86
15	0.14	0.06	0.23	0.32	0.17	0.21	0.43	0.84	0.59	12.83
16	0.08	0.16	0.38	0.12	0.20	0.12	0.26	0.56	0.56	0.29
17	0.22	0.22	0.32	0.08	0.26	0.32	0.43	0.97	1.47	108.79
18	0.16	0.08	0.16	0.19	0.33	0.30	1.09	1.60	0.78	3.08
19	0.12	0.19	0.19	0.41	0.12	0.22	0.82	0.73	0.98	1.09
20	0.25	0.23	0.11	0.17	0.19	0.13	0.72	0.45	0.81	3.25
求和	3.32	3.36	4.21	4.29	4.92	5.21	10.81	15.3	19.17	375.49
平均值	0.166	0.168	0.2105	0.2145	0.246	0.2605	0.5405	0.765	0.9585	18.7745
标准差	0.08	0.088	0.083	0.109	0.103	0.104	0.266	0.437	0.59	49.90
方差	139572	472416	758739	134152	993927	653564	250278	426686	377716	388346

对数据进行升序排序见表 2:

人员数值	甲	乙	丙	丁	戊	己	庚	辛	壬	癸
1	0.05	0.06	0.11	0.06	0.11	0.12	0.13	0.26	0.24	0.29
2	0.06	0.06	0.11	0.08	0.12	0.13	0.22	0.31	0.42	0.57
3	0.07	0.07	0.12	0.09	0.13	0.15	0.25	0.37	0.43	0.69
4	0.08	0.08	0.12	0.11	0.15	0.17	0.26	0.42	0.56	0.77
5	0.09	0.08	0.14	0.12	0.15	0.19	0.32	0.43	0.59	0.86
6	0.10	0.09	0.15	0.12	0.16	0.19	0.34	0.45	0.63	1.09
7	0.12	0.12	0.16	0.16	0.17	0.19	0.43	0.45	0.66	1.56
8	0.13	0.12	0.16	0.17	0.19	0.21	0.43	0.56	0.67	2.34
9	0.14	0.14	0.17	0.18	0.20	0.22	0.43	0.56	0.74	2.86
10	0.14	0.15	0.19	0.19	0.24	0.23	0.50	0.59	0.78	3.08
11	0.16	0.16	0.21	0.21	0.26	0.24	0.51	0.67	0.81	3.25
12	0.19	0.17	0.21	0.23	0.26	0.25	0.54	0.73	0.83	3.74
13	0.21	0.19	0.23	0.24	0.27	0.29	0.62	0.84	0.98	4.32
14	0.22	0.22	0.23	0.25	0.28	0.30	0.65	0.88	0.98	4.78
15	0.22	0.23	0.25	0.27	0.32	0.32	0.72	0.93	1.06	5.23
16	0.23	0.23	0.27	0.3	0.32	0.34	0.73	0.97	1.08	5.43
17	0.24	0.24	0.32	0.32	0.33	0.35	0.76	1.05	1.44	7.69
18	0.25	0.26	0.32	0.32	0.35	0.37	0.82	1.34	1.47	12.83
19	0.27	0.32	0.36	0.41	0.42	0.42	1.06	1.60	2.05	108.79

20	0.35	0.37	0.38	0.46	0.49	0.53	1.09	1.89	2.75	205.32
平均	0.16	0.16	0.210	0.240	0.24	0.260	0.540	0.76	0.958	18.774
值	6	8	5	5	6	5	5	5	5	5

对列数据减去平均值后见表 3:

-0.11	-0.10	-0.10	-0.18	-0.13	-0.14	-0.41	-0.50	-0.71	-18.48
6	8	05	05	6	05	05	5	85	45
-0.10	-0.10	-0.10	-0.16	-0.12	-0.13	-0.32	-0.45	-0.53	-18.20
6	8	05	05	6	05	05	5	85	45
-0.09	-0.09	-0.09	-0.15	-0.11	-0.11	-0.29	-0.39	-0.52	-18.08
6	8	05	05	6	05	05	5	85	45
-0.08	-0.08	-0.09	-0.13	-0.09	-0.09	-0.28	-0.34	-0.39	-18.00
6	8	05	05	6	05	05	5	85	45
-0.07	-0.08	-0.07	-0.12	-0.09	-0.07	-0.22	-0.33	-0.36	-17.91
6	8	05	05	6	05	05	5	85	45
-0.06	-0.07	-0.06	-0.12	-0.08	-0.07	-0.20	-0.31	-0.32	-17.68
6	8	05	05	6	05	05	5	85	45
-0.04	-0.04	-0.05	-0.08	-0.07	-0.07	-0.11	-0.31	-0.29	-17.21
6	8	05	05	6	05	05	5	85	45
-0.03	-0.04	-0.05	-0.07	-0.05	-0.05	-0.11	-0.20	-0.28	-16.43
6	8	05	05	6	05	05	5	85	45
-0.02	-0.02	-0.04	-0.06	-0.04	-0.04	-0.11	-0.20	-0.21	-15.91
6	8	05	05	6	05	05	5	85	45
-0.02	-0.01	-0.02	-0.05	-0.00	-0.03	-0.04	-0.17	-0.17	-15.69
6	8	05	05	6	05	05	5	85	45
-0.00	-0.00	-0.00	-0.03	0.014	-0.02	-0.03	-0.09	-0.14	-15.52
6	8	05	05	05	05	05	5	85	45
0.024	0.002	-0.00	-0.01	0.014	-0.01	-0.00	-0.03	-0.12	-15.03
		05	05	05	05	05	5	85	45
0.044	0.022	0.019	-0.00	0.024	0.029	0.079	0.075	0.021	-14.45
		5	05	05	5	5	5	5	45
0.054	0.052	0.019	0.009	0.034	0.039	0.109	0.115	0.021	-13.99
		5	5	05	5	5	5	5	45
0.054	0.062	0.039	0.029	0.074	0.059	0.179	0.165	0.101	-13.54
		5	5	05	5	5	5	5	45
0.064	0.062	0.059	0.059	0.074	0.079	0.189	0.205	0.121	-13.34
		5	5	05	5	5	5	5	45
0.074	0.072	0.109	0.079	0.084	0.089	0.219	0.285	0.481	-11.08
		5	5	05	5	5	5	5	45
0.084	0.092	0.109	0.079	0.104	0.109	0.279	0.575	0.511	-5.944
		5	5	05	5	5	5	5	5
0.104	0.152	0.149	0.169	0.174	0.159	0.519	0.835	1.091	90.015
		5	5	05	5	5	5	5	5

0.184	0.202	0.169 5	0.219 5	0.244	0.269 5	0.549 5	1.125	1.791 5	186.54 55
-------	-------	------------	------------	-------	------------	------------	-------	------------	--------------

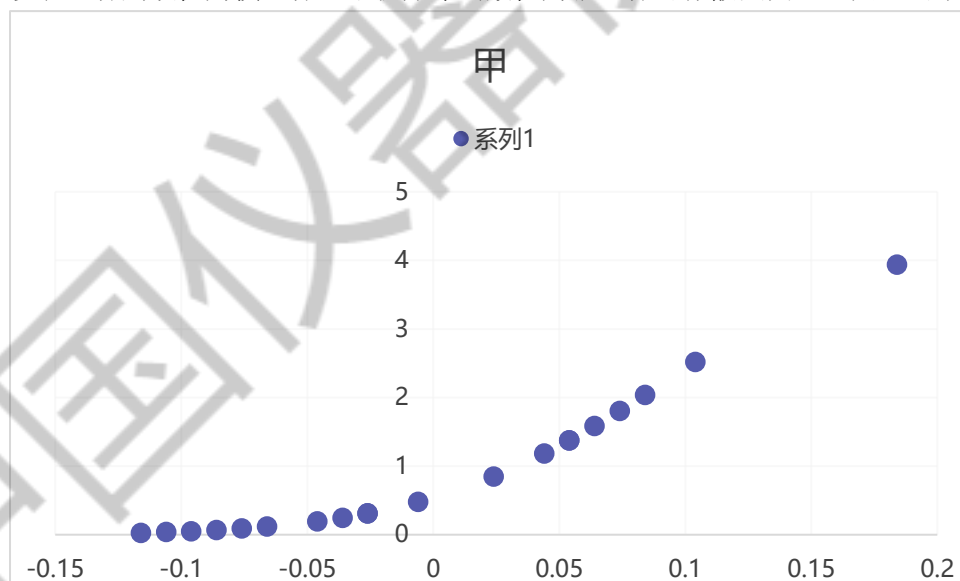
甲-庚平均值取 0.2, 标准差按
实际
辛后平均值, 标准差均按实际
NORMDIST 返回累积分布函数

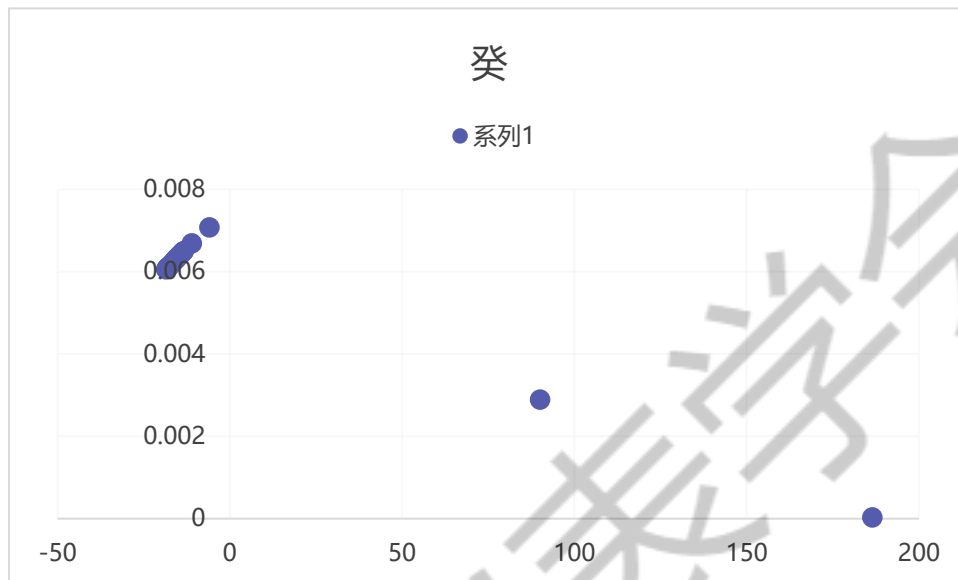
累积分布函数值: 表 4, NORMDIST(项值, 平均值, 标准差, TRUE)

	0.002	0.009	0.007	0.008	0.020	0.019	0.002	0.013	0.012	0.006
1	44075	91684	90051	26716	76452	77760	53149	37921	45342	04981
	5	9	5	9	9	9	4	2	3	5
	0.003	0.009	0.007	0.015	0.028	0.026	0.008	0.018	0.027	0.006
2	92089	91684	90051	42480	19865	81224	00915	53519	99040	07512
	7	9	5	7	6	9	2	3	7	
	0.006	0.014	0.012	0.020	0.037	0.047	0.011	0.026	0.029	0.006
3	20336	66561	00977	80501	94193	95508	46238	93843	19980	08593
	5	8		9	1			1	3	9
	0.009	0.021	0.012	0.036	0.066	0.082	0.012	0.036	0.049	0.006
4	66607	41011	00977	90615	81217	71404	88081	26026	31717	09314
	1	1		5	1	8	4	7	1	3
	0.014	0.021	0.026	0.048	0.066	0.137	0.025	0.038	0.055	0.006
5	83382	41011	59673	53791	81217	58377	18052	42027	28042	10123
	8	1	7	2	1	3	1	7	3	8
	0.022	0.030	0.038	0.048	0.087	0.137	0.031	0.043	0.064	0.006
6	42008	85526	74741	53791	43810	58377	13112	06626	11294	12188
	1	9	5	2	1	3	8	7	8	5
	0.048	0.085	0.055	0.133	0.113	0.137	0.075	0.043	0.071	0.006
7	92682	46960	65473	49239	37845	58377	40459	06626	43879	16388
	9	6	8	7	9	3	2	7	6	7
	0.070	0.085	0.055	0.168	0.185	0.220	0.075	0.077	0.074	0.006
8	64363	46960	65473	33017	41477	69802	40459	72290	02028	23300
		6	8	3	2	1	2	4	3	9
	0.100	0.158	0.078	0.210	0.233	0.275	0.075	0.077	0.094	0.006
9	45686	03716	81458	48058	84510	7435	40459	72290	15007	27866
	4	9	4	3	1		2	4	1	8
	0.100	0.210	0.151	0.260	0.539	0.341	0.138	0.090	0.107	0.006
10	45686	77582	47835	97970	40018	40741	62768	30333	35192	29788
	4	5	2	4	5	4	3		4	
	0.194	0.277	0.275	0.391	0.775	0.418	0.150	0.131	0.118	0.006
11	05946	50717	08992	22808	01608	89139	37432	65684	10255	31268
	5	3	8	2	8			6	8	2
1	0.464	0.360	0.275	0.567	0.775	0.509	0.190	0.170	0.125	0.006

2	75347	67786	08992	06367	01608	32001	31020	88215	68309	35512
	8	2	8	4	8	2	2	9	1	8
1	0.770	0.586	0.472	0.674	0.916	1.016	0.335	0.262	0.193	0.006
3	87600	11874	04014	13934	19430	61421	13847	46072	29245	40494
	3	2	8	6	3	7			7	1
1	0.970	1.102	0.472	0.794	1.073	1.181	0.404	0.302	0.193	0.006
4	36799	10692	04014	71636	12214	27113	81588	00342	29245	44410
	6	4	8	2	9	2	7	2	7	7
1	0.970	1.325	0.765	1.076	1.841	1.552	0.598	0.355	0.236	0.006
5	36799	62337	35569	88915	37601	09706	66765	69636	91426	48212
	6	9	4	5	2	6		9	8	
1	1.203	1.325	1.172	1.594	1.841	1.966	0.629	0.401	0.248	0.006
6	00953	62337	54043	75511	37601	66980	51500	64280	57327	49891
	5	9	6	1	2	2	6	9		7
1	1.468	1.574	2.658	1.986	2.059	2.183	0.725	0.499	0.486	0.006
7	86654	01344	13590	54841	33770	88098	74528	39609	18420	68430
	8	5	3	4	1	2	9	6	1	6
1	1.766	2.134	2.658	1.986	2.505	2.620	0.928	0.830	0.505	0.007
8	34802	81663	13590	54841	25454	63455	47053	57719	68390	07165
	8	5	3	4	4	8	4	2	7	6
1	2.440	3.906	3.964	3.519	3.717		1.495	0.901		0.002
9	09172	80381	12421	50453	96266	3.527	33569	27418	0.655	88545
	2	4	8	6	2	07403		1	4	2
2	4.830	4.532	4.446	3.601	3.507	3.052		0.650	0.251	2.806
0	05748	26433	33815	91778	58864	00099	1.498	22388	53203	65E-0
	8	1	5	9	3	3	82682		4	2
										5

以表 3 各列项值为横坐标，累积分布函数值为纵坐标可作散点图：（只画出甲和癸）





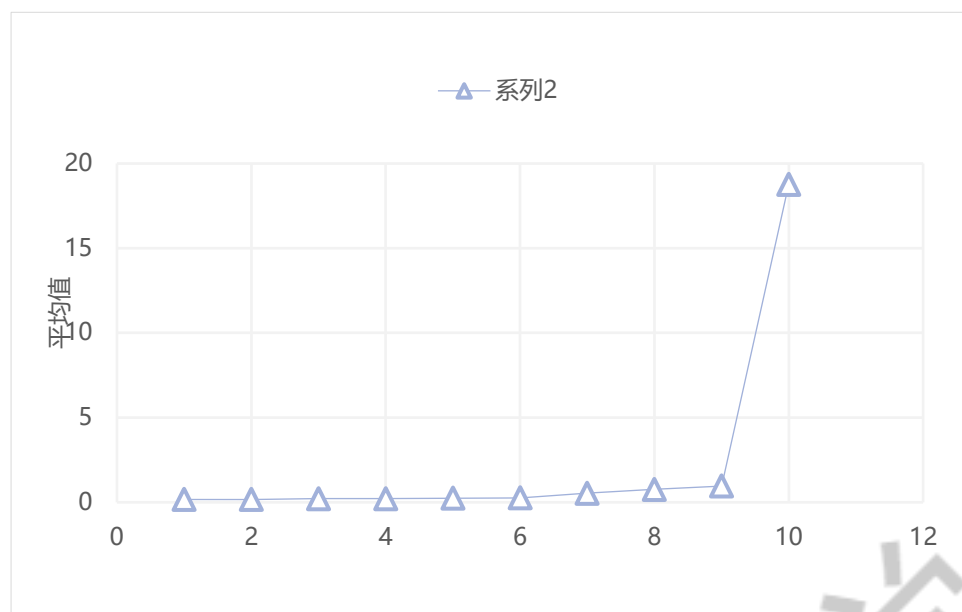
从表 4 及直观的图可看出戊 20, 己 20, 辛 20 , 壬 20, 癸 19, 癸 20, 类积分布值出现下降, (曲线斜率为负值), 代表相关数据脱离正态分布范围, 尤其是辛 20 后 (下降幅度超过 30%) 为可疑数据。

从实验数据计算的数学模型 $X = (C \cdot V) / M$ 看, 除了油品本身含有高含量的有害物质外, 偶然的一些误差不太可能造成相关数据“偏高”, 大幅脱离正态分布范围, 具体见下表:

因 素	偏 离	结 果 影 响	备 注
称样误差	±1%	±1%	
定容误差	±3%	±3%	
标样浓度变化	±1%	±1%	标样化学成分相对稳定
反应不完全	-5%, ~-10%	-5%, ~-10%	结果偏小
反应容器吸附	-3%~-5%	-3%~-5%	结果偏小
内标加入体积变化	±3%	±3%	
滤膜吸附	-5%	-5%	结果偏小
仪器测定误差	±1%	±1%	

那么是否真的出现了“地沟油”? 尽管风险分析“偏好”风险, 但也要具体问题具体分

析，先看 10 人测定结果平均值的变化，如下：



可以看到第 6 人后平均值开始明显变大，第 10 人则显著变大，对应的场景是第 6 人后，因气相顶空出故障，被插入了一批油品溶剂残留分析，第 9 人后又插入了一批溶剂残留分析。因为共用进样部件，因此怀疑是否是进样针，洗针盖等先沾染了溶剂残留瓶盖上的油，然后又沾上了高浓度的标样，最后导致在进样垫上集聚部分高浓度标样，在样品进样时随进样针随机带入样品中！

Excel 中的 NORMDIST 函数为我们的分析提供了异常数值的判断依据，但是，对小概率事件的原因分析，必须结合具体场景分析进行，在时间及器材具备的情况下，最好通过复测进行进一步验证，否则，检测数据存在失稳失真的可能性，在 AI “深度学习” “大模型” 不断发展的今天，知识和经验都有所贬值，但正如 AmehIP 的创始人，负责人洛厄尔·坎贝尔所说：“一根经验的荆棘抵得上忠告的茫茫荒原”。

参考文献

1. 冯灵清编著:Excel 2010 统计分析实例 清华大学出版社, 2015
2. 日本科技联盟“QC 方法研究会”编 新全面质量管理七种工具
3. 王亚奇*, 庞成森, 马妮等 基于 Excelvba 技术构建审核全肠外营养液处方的算法 (# 西安交通大学第一附属医院药学部, 西安 710061) 中国药房 2019 年第 30 卷第 1 期
4. 刘俊桃, 陈晓培, 吴金松等 Excel 在分析化学实验数据处理中的应用 (河南牧业经济学院 理学部, 河南 郑州 450000) 云南化工 2018 第 45 卷第 9 期
5. 黄河清, 储德初, 段路路 Excel 在国际实验室间比对数据处理中的应用 上海化工研究院有限公司, 上海 200062; 2. 上海化学品公共安全工程技术研究中心, 上海 200062 磷肥与复肥 2019 第 34 卷第 1 期

中国仪器仪表学会