

準拠点選択

準拠点とは

準拠点とは境界(筆界)復元計算をする際に基準とする点のことです、境界測量には基準点測量と境界標を測る一筆測量があります。基準点とは図面を作成する際に測量の骨格、基準点網といいますがこの骨格を測量したものが基準点です。一筆測量とは境界(筆界)等を測ることで基準点を基におこないます。この二つの測量は別々に行われます。一筆測量は基準点から距離と角度を測量器械で測って計算されます、この時に設置された基準点が使用できれば基準点から距離と角度を使って境界の復元が出来ますので準拠点はいりません。

それ以外のケースでは境界標等から得られた座標データの中から精度の安定したデータを選び、その点を基点に復元計算をします、この基点を準拠点といいます。

準拠点選択とは

準拠点選択とは基準点が無いとか基準点の精度が確認出来ない場合に境界標、引照点などから境界(筆界)復元計算に使用できる点を絞り込むことをいいます。この点を使って各種の計算方法を使って境界復元を行います。

準拠点選択を感覚的にすることもあります、この方法では明らかに確率論から離れた点を除く場合は問題ないのですが、除くか含めるかの迷うレベルでは個人差が出ますので適切とはいえなくなります。

又、公差とか基準を当て嵌めて判断することも考えられますが、対象とした全体のデータ精度が公差等の範囲のどのレベルにあるか解らない場合は適切な方法ではありません。筆界特定書とか判決書で「この辺長は公差内だから復元計算の根拠になりうる・・・」等という判断が示されることがありますが全体の精度(母集団の精度(標準偏差))が把握されていない場合には避けるべきです。又、確率論からいえばどのレベルの値でもかまわないというものではありません。やむを得ず公差等を当て嵌める場合はその公差内はどの程度の不良率を想定したものか知っておく必要があります、例えば国土調査法施行令別表4の筆界特定の位置誤差、距離の誤差、面積の誤差は 3.9 シグマ法、つまり 10,000 分の 1 の不良率を想定してあります、距離の公差が 100mmとすれば標準偏差は $\frac{100}{3.9} = 25.64$ となり、一般的に有意水準 5%、確率 95%で考えますので 95%の確率は 1.96 倍標準偏差となります、 $25.64 \times 1.96 = 50.3\text{mm}$ 以内が通常考えられる誤差ということになります。公差等を使う場合には注意しなければなりません。

準拠点選択に求められることは、準拠点選択の結果に基づいて境界(筆界)復元計算をしたときに復元精度の高い点の組合せを選ぶことができるか、できないかと、ということが重要になります。

復元精度の高い点の組合せを選ぶとは

次のグラフは変数減少法によって、次に示すデータから計算した「実測値－変換値の標準偏差」「真値－変換値の標準偏差」の関係を計算したグラフです。

実際には「実測値－変換値の標準偏差」は計算できますが「真値－変換値の標準偏差」は「真値」が不明ですから計算はできません、ここではサンプルデータですから真値は既知です。

最小二乗法座標変換による座標変換の特徴は①準拠点が多いほど復元精度(標準偏差)が高く、バラツキが小さい。

②準拠点の重心部分ほど復元精度(標準偏差)が高く、バラツキが小さい、周辺部分にいくにしたがって復元精度(標準偏差)は低くなりバラツキも大きい。

③準拠点が正規分布になっていれば復元精度(標準偏差)は高く、異常点が混じると復元精度(標準偏差)が落ち、バラツキも大きくなる。

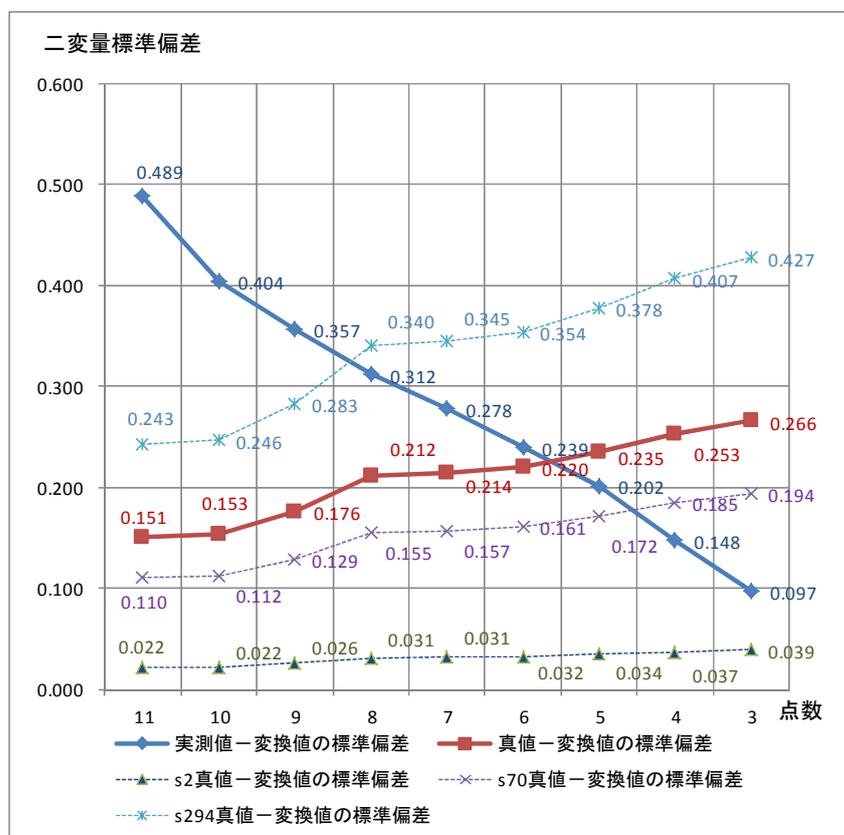
ことが判っています。

異常点のないデータのみで変数減少法により、「実測値－変換値の標準偏差」「真値－変換値の標準偏差」, 「真値－変換値の準拠点のみの標準偏差」, s2 真値との乖離, s22 真値との乖離, s70 真値との乖離, s138 真値との乖離, s224 真値との乖離, s294 真値との乖離 の値を計算したのが下表です。(後掲の点の配置図を参照)

11回の平均 母集団二変量標準偏差0.476

点数	実測値－変換値の標準偏差	真値－変換値の標準偏差	真値－変換値○のみ	s2	s22	s70	s138	s224	s294
11	0.489	0.151	0.147	0.022	0.066	0.110	0.155	0.199	0.243
10	0.404	0.153	0.154	0.022	0.067	0.112	0.157	0.201	0.246
9	0.357	0.176	0.173	0.026	0.077	0.129	0.180	0.232	0.283
8	0.312	0.212	0.208	0.031	0.093	0.155	0.217	0.278	0.340
7	0.278	0.214	0.216	0.031	0.094	0.157	0.220	0.282	0.345
6	0.239	0.220	0.223	0.032	0.097	0.161	0.225	0.290	0.354
5	0.202	0.235	0.229	0.034	0.103	0.172	0.241	0.309	0.378
4	0.148	0.253	0.259	0.037	0.111	0.185	0.259	0.333	0.407
3	0.097	0.266	0.289	0.039	0.117	0.194	0.272	0.350	0.427

「実測値－変換値の標準偏差」「真値－変換値の標準偏差」, s2 真値との乖離, s70 真値との乖離, s294 真値との乖離 の値をグラフにしたのが下図です。



このグラフから読み取れることは

① 「実測値－変換値の標準偏差」は母集団の標準偏差 0.476 に近い値の時に復元精度が 0.151 と高いことです。

② 「実測値－変換値の標準偏差」が一定割合で減少している場合は準拠点が最も多い11点の位置の復元精度(真値－変換値の標準偏差)が 0.151 と小さい値を示す, つまり復元精度が高いと言うことです。

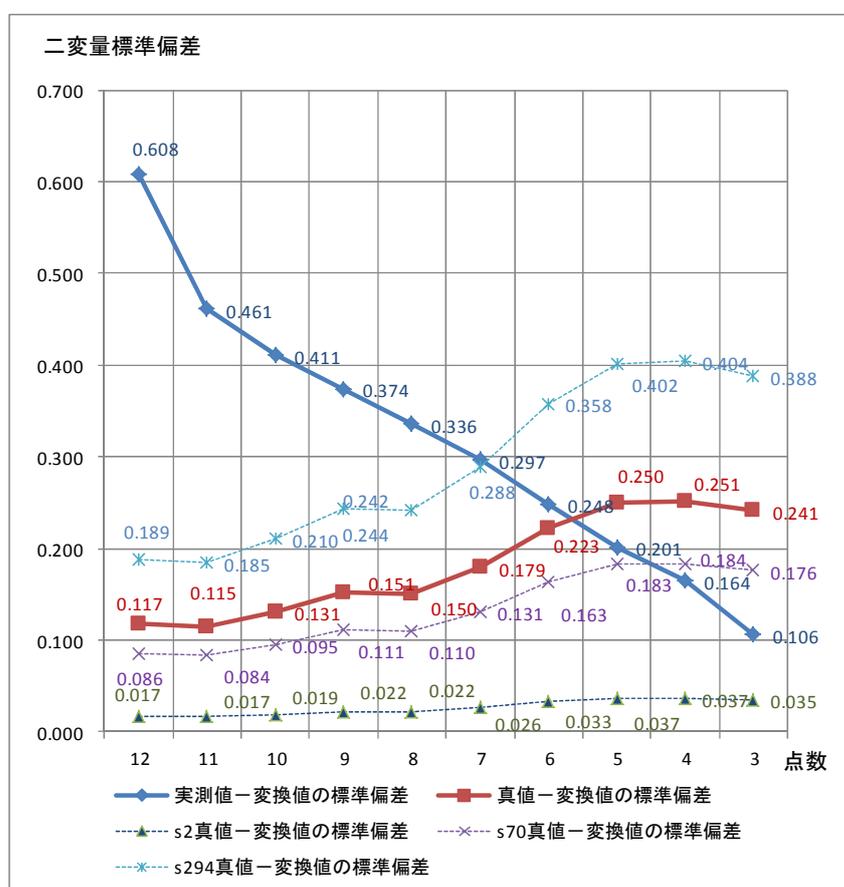
③ 各点の復元精度は11点で s2 が 0.022, s70 が 0.110, s294 が 0.243 と図の重心付近が高く, 周辺に行くにしたがって低くなることです。

したがって, 準拠点選択とは 11 点目の位置を探すことです。

重心部に近い異常点 s2 を追加して計算したのが下表，下図です。

11回の平均 母集団二変量標準偏差0.476 異常点s2

点数	実測値－ 変換値の 標準偏差	真値－変 換値の標 準偏差	真値－変換 値○のみ	s2	s22	s70	s138	s224	s294
12	0.608	0.117	0.111	0.017	0.051	0.086	0.120	0.154	0.189
11	0.461	0.115	0.113	0.017	0.050	0.084	0.118	0.151	0.185
10	0.411	0.131	0.130	0.019	0.057	0.095	0.134	0.172	0.210
9	0.374	0.151	0.154	0.022	0.066	0.111	0.155	0.199	0.244
8	0.336	0.150	0.154	0.022	0.066	0.110	0.154	0.198	0.242
7	0.297	0.179	0.183	0.026	0.079	0.131	0.184	0.236	0.288
6	0.248	0.223	0.220	0.033	0.098	0.163	0.228	0.293	0.358
5	0.201	0.250	0.257	0.037	0.110	0.183	0.256	0.329	0.402
4	0.164	0.251	0.287	0.037	0.110	0.184	0.257	0.331	0.404
3	0.106	0.241	0.285	0.035	0.106	0.176	0.247	0.317	0.388



このグラフから読み取れることは

① 「実測値－変換値の標準偏差」は母集団の標準偏差 0.476 に近い値の時に復元精度が 0.117 と高いことです。

② 「実測値－変換値の標準偏差」が 12 点目で 0.608 から 11 点目で 0.461 と急激に減少しています，その後 11 点目からは一定割合で減少しています，この場合は一定に減少の始まる準拠点が多い 11 点の位置の復元精度(真値－変換値の標準偏差)が 0.115 と小さい値を示す，つまり復元精度が高いということです。

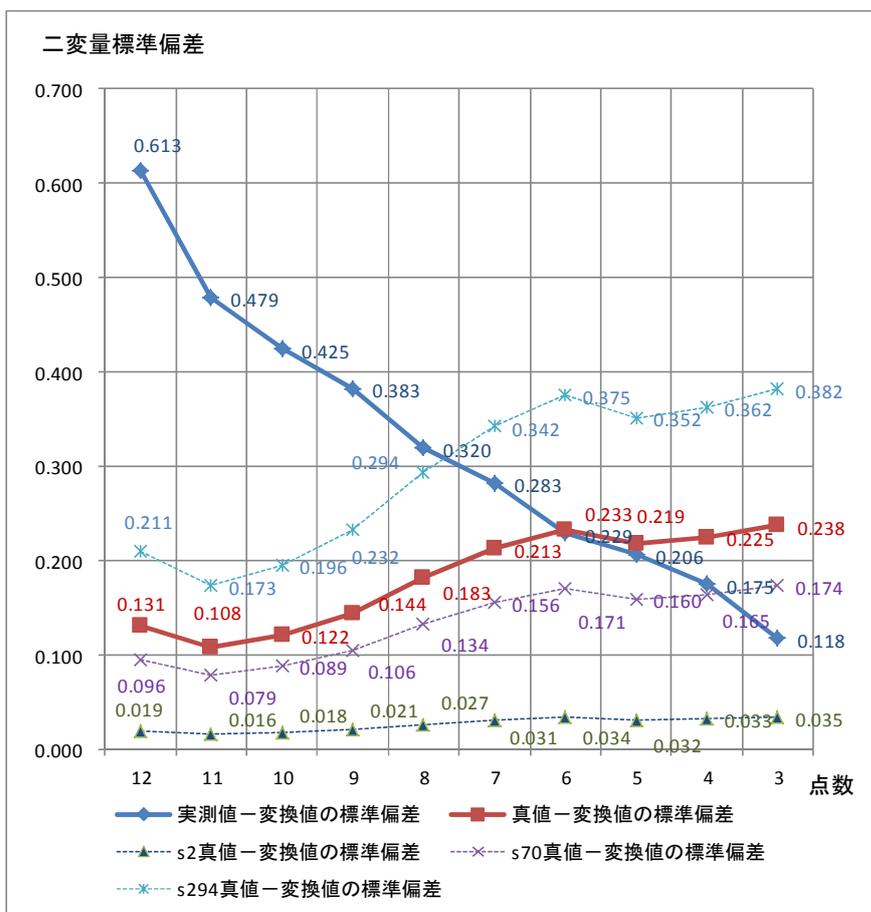
③各点の復元精度は 11 点で s2 が 0.017, s70 が 0.117, s294 が 0.189 と図の重心付近が高く、周辺に行くにしたがって低くなることです。

したがって、準拠点選択とは 11 点目の位置を探すことです。

中間部に近い異常点 s97 を追加して計算したのが下表、下図です。

11回の平均 母集団二変量標準偏差0.476 異常点s97

点数	実測値－変換値の標準偏差	真値－変換値の標準偏差	真値－変換値のみ ○のみ	s2	s22	s70	s138	s224	s294
12	0.613	0.131	0.130	0.019	0.058	0.096	0.134	0.173	0.211
11	0.479	0.108	0.110	0.016	0.047	0.079	0.110	0.142	0.173
10	0.425	0.122	0.121	0.018	0.053	0.089	0.125	0.160	0.196
9	0.383	0.144	0.140	0.021	0.063	0.106	0.148	0.190	0.232
8	0.320	0.183	0.179	0.027	0.080	0.134	0.187	0.241	0.294
7	0.283	0.213	0.208	0.031	0.093	0.156	0.218	0.280	0.342
6	0.229	0.233	0.237	0.034	0.102	0.171	0.239	0.307	0.375
5	0.206	0.219	0.225	0.032	0.096	0.160	0.224	0.288	0.352
4	0.175	0.225	0.234	0.033	0.099	0.165	0.231	0.296	0.362
3	0.118	0.238	0.272	0.035	0.104	0.174	0.243	0.313	0.382



このグラフから読み取れることは

① 「実測値－変換値の標準偏差」は母集団の標準偏差 0.476 に近い値の時に復元精度が 0.108 と高いことです。

② 「実測値－変換値の標準偏差」が 12 点目で 0.613 から 11 点目で 0.479 と急激に減少しています, その後 11 点目からは一定割合で減少しています, この場合は一定に減少の始まる準拠点が多い 11 点の位置の復元精度(真値－変換値の標準偏差)が 0.108 と小さい値を示す, つまり復元精度が高いということです。

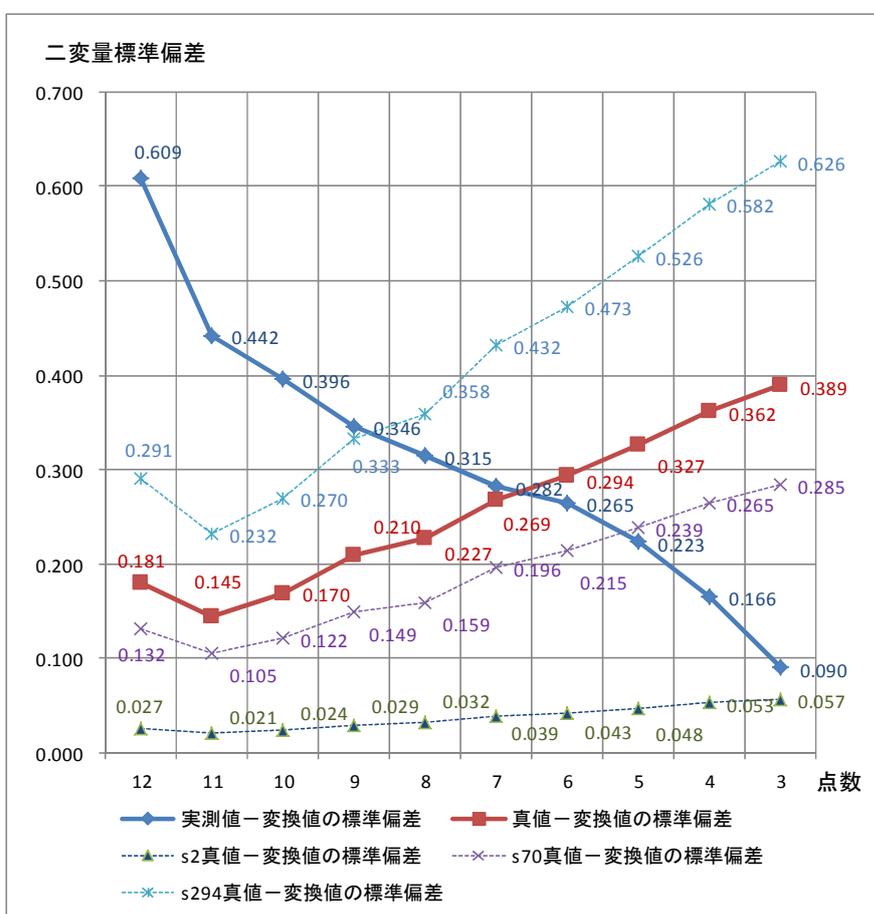
③ 各点の復元精度は 11 点で s_2 が 0.016, s_{70} が 0.079, s_{294} が 0.173 と図の重心付近が高く, 周辺に行くにしたがって低くなることです。

したがって, 準拠点選択とは 11 点目の位置を探すことです。

外枠部付近の異常点 s291 を追加して計算したのが下表, 下図です。

11回の平均 母集団二変量標準偏差0.476 異常点s291

点数	実測値－ 変換値の 標準偏差	真値－変 換値の標 準偏差	真値－変換値 ○のみ	s2	s22	s70	s138	s224	s294
12	0.609	0.181	0.188	0.027	0.079	0.132	0.185	0.238	0.291
11	0.442	0.145	0.146	0.021	0.063	0.105	0.148	0.190	0.232
10	0.396	0.170	0.170	0.024	0.072	0.122	0.171	0.221	0.270
9	0.346	0.210	0.214	0.029	0.088	0.149	0.210	0.271	0.333
8	0.315	0.227	0.237	0.032	0.093	0.159	0.226	0.292	0.358
7	0.282	0.269	0.265	0.039	0.118	0.196	0.275	0.353	0.432
6	0.265	0.294	0.295	0.043	0.129	0.215	0.301	0.387	0.473
5	0.223	0.327	0.342	0.048	0.143	0.239	0.334	0.430	0.526
4	0.166	0.362	0.377	0.053	0.159	0.265	0.370	0.476	0.582
3	0.090	0.389	0.404	0.057	0.171	0.285	0.398	0.512	0.626



このグラフから読み取れることは

① 「実測値－変換値の標準偏差」は母集団の標準偏差 0.476 に近い値の時に復元精度が 0.145 と高いことです。

② 「実測値－変換値の標準偏差」が 12 点目で 0.609 から 11 点目で 0.442 と急激に減少しています, その後 11 点目からは一定割合で減少しています, この場合は一定に減少の始まる準拠点が多い 11 点の位置の復元精度(真値－変換値の標準偏差)が 0.145 と小さい値を示

す、つまり復元精度が高いと言うことです。

ここまでの4つの計算結果のうち、後ろの3つはそれぞれ同程度の大きな誤差のある点(異常点)を1点混ぜて計算しています、この異常点が復精度(標準偏差)に与える影響は図の重心部ほど小さく、外周に向かうほど大きいことが解ります。

③各点の復元精度は11点でs2が0.021, s70が0.105, s294が0.232と図の重心付近が高く、周辺に行くにしたがって低くなることです。

したがって、準拠点選択とは11点目の位置を探すことです。

復元値のバラツキ具合

復元された値自体にもバラツキがあります、下表は標準偏差のバラツキ(標準偏差)を調べたものです。

異常点 s291 を含んだデータのみですが①準拠点が多いほどバラツキが小さいことから準拠点選択はデータ数が多いこと、異常値が除かれることが重要であることが解ります。

バラツキの計算

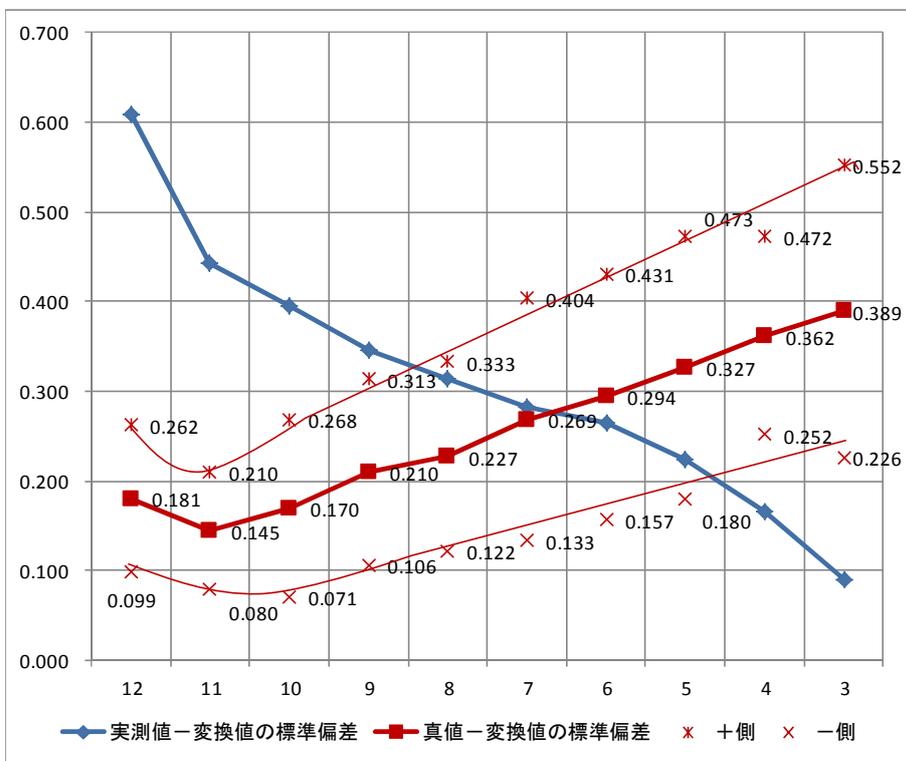
11回の平均 母集団二変量標準偏差0.476 異常点s291

点数	実測値－ 変換値の バラツキ	真値－変 換値のバ ラツキ	真値－変換値 ○のみのバラ ツキ	s2のバラ ツキ	s22のバラ ツキ	s70のバラ ツキ	s138のバ ラツキ	s224のバ ラツキ	s294のバ ラツキ
12	0.148	0.082	0.026	0.040	0.063	0.086	0.110	0.133	0.065
11	0.090	0.065	0.036	0.031	0.049	0.067	0.086	0.104	0.060
10	0.101	0.099	0.051	0.043	0.070	0.097	0.125	0.152	0.091
9	0.095	0.104	0.055	0.041	0.067	0.094	0.122	0.149	0.082
8	0.096	0.106	0.059	0.039	0.063	0.090	0.117	0.145	0.078
7	0.140	0.135	0.066	0.067	0.107	0.148	0.189	0.230	0.122
6	0.131	0.137	0.066	0.065	0.102	0.140	0.178	0.216	0.112
5	0.122	0.146	0.055	0.063	0.098	0.134	0.170	0.205	0.098
4	0.147	0.110	0.078	0.069	0.106	0.144	0.183	0.221	0.112
3	0.191	0.163	0.075	0.080	0.125	0.172	0.218	0.264	0.129

グラフ用+側、－側の計算(グラフは2, 5, 6, 7列)

点数	実測値－ 変換値の 標準偏差			真値－変 換値の標 準偏差		
		+	－		+	－
12	0.609	0.757	0.461	0.181	0.262	0.099
11	0.442	0.532	0.353	0.145	0.210	0.080
10	0.396	0.496	0.295	0.170	0.268	0.071
9	0.346	0.442	0.251	0.210	0.313	0.106
8	0.315	0.411	0.219	0.227	0.333	0.122
7	0.282	0.423	0.142	0.269	0.404	0.133
6	0.265	0.395	0.134	0.294	0.431	0.157
5	0.223	0.345	0.101	0.327	0.473	0.180
4	0.166	0.313	0.019	0.362	0.472	0.252
3	0.090	0.281	-0.100	0.389	0.552	0.226

下図が復元精度精度(真値-復元値)に対してのバラツキの状態(+の細い実線)を示します, 異常点があればバラツキは大きくなります, 準拠点が少なくなるとバラツキは大きくなります。この関係をグラフにしたものです。



検証データ

ここまでのデータとこれから使うのは次のデータです，準拠点選択では異常点を確実に除けば復元精度(標準偏差)は確保されます，この異常点を除く方法を具体的にどうすれば良いか次の下表データを使って調べて見ます。

左から「図面值座標値」「実測座標値」「図面值－実測値の値」「図面值－実測値の中から誤差の大きいs2, s97, s291(以下異常点と言います)を除いた値」の4種類です。

母集団, 291 個のデータの標準偏差は4枠目の 0.476 です。3枠目の標準偏差は誤差の大きい点を含めた時の値で参考です。

図面值座標値			実測値座標値			標準偏差x		標準偏差y		標準偏差x		標準偏差y	
図面值	x	y	実測値	x	y	標準偏差	標準偏差	標準偏差	標準偏差	標準偏差	標準偏差	標準偏差	標準偏差
s1	0.000	0.000	500z1	0.238	-0.221	0.488	0.506	0.484	0.490	0.488	0.506	0.484	0.490
s2	10.000	0.000	500z2	10.457	2.005	標準偏差 σ_{xy}	0.497	標準偏差 σ_{xy}	0.487	標準偏差 σ_{xy}	0.497	標準偏差 σ_{xy}	0.487
s3	10.000	10.000	500z3	10.232	9.921	-0.238	0.221	-0.238	0.221	-0.238	0.221	-0.238	0.221
s4	0.000	10.000	500z4	0.372	10.151	-0.457	-2.005	-0.343	-1.504	-0.343	-1.504	-0.343	-1.504
s5	-10.000	10.000	500z5	-10.716	9.046	-0.232	0.079	-0.232	0.079	-0.232	0.079	-0.232	0.079
s6	-10.000	0.000	500z6	-9.875	-0.355	-0.372	-0.151	-0.372	-0.151	-0.372	-0.151	-0.372	-0.151
s7	-10.000	-10.000	500z7	-10.176	-9.929	0.716	0.954	0.716	0.954	0.716	0.954	0.716	0.954
s8	0.000	-10.000	500z8	0.041	-9.380	-0.125	0.355	-0.125	0.355	-0.125	0.355	-0.125	0.355
s9	10.000	-10.000	500z9	9.243	-10.507	0.176	-0.071	0.176	-0.071	0.176	-0.071	0.176	-0.071
s10	20.000	0.000	500z10	20.118	0.191	-0.041	-0.620	-0.041	-0.620	-0.041	-0.620	-0.041	-0.620
s11	20.000	10.000	500z11	20.590	9.631	0.757	0.507	0.757	0.507	0.757	0.507	0.757	0.507
s12	10.000	20.000	500z12	9.252	19.902	-0.118	-0.191	-0.118	-0.191	-0.118	-0.191	-0.118	-0.191
s13	0.000	20.000	500z13	-0.025	19.805	-0.590	0.369	-0.590	0.369	-0.590	0.369	-0.590	0.369
s14	-10.000	20.000	500z14	-8.948	20.211	0.748	0.098	0.748	0.098	0.748	0.098	0.748	0.098
s15	-20.000	10.000	500z15	-19.484	10.457	0.025	0.195	0.025	0.195	0.025	0.195	0.025	0.195
s16	-20.000	0.000	500z16	-19.753	1.062	-1.052	-0.211	-1.052	-0.211	-1.052	-0.211	-1.052	-0.211
s17	-20.000	-10.000	500z17	-19.521	-10.675	-0.516	-0.457	-0.516	-0.457	-0.516	-0.457	-0.516	-0.457
s18	-10.000	-20.000	500z18	-10.408	-20.214	-0.247	-1.062	-0.247	-1.062	-0.247	-1.062	-0.247	-1.062
s19	0.000	-20.000	500z19	0.103	-20.186	-0.479	0.675	-0.479	0.675	-0.479	0.675	-0.479	0.675
s20	10.000	-20.000	500z20	10.655	-20.187	0.408	0.214	0.408	0.214	0.408	0.214	0.408	0.214
s21	20.000	-10.000	500z21	19.321	-10.123	-0.103	0.186	-0.103	0.186	-0.103	0.186	-0.103	0.186
s22	30.000	0.000	500z22	30.228	0.245	-0.655	0.187	-0.655	0.187	-0.655	0.187	-0.655	0.187
s23	30.000	10.000	500z23	30.274	9.570	0.679	0.123	0.679	0.123	0.679	0.123	0.679	0.123
s24	20.000	20.000	500z24	20.251	19.752	-0.228	-0.245	-0.228	-0.245	-0.228	-0.245	-0.228	-0.245
s25	10.000	30.000	500z25	9.901	29.640	0.726	0.430	0.726	0.430	0.726	0.430	0.726	0.430
s26	0.000	30.000	500z26	0.369	29.835	-0.251	0.248	-0.251	0.248	-0.251	0.248	-0.251	0.248
s27	-10.000	30.000	500z27	-10.596	30.376	0.099	0.360	0.099	0.360	0.099	0.360	0.099	0.360
s28	-20.000	20.000	500z28	-19.858	20.645	-0.369	0.165	-0.369	0.165	-0.369	0.165	-0.369	0.165
s29	-30.000	10.000	500z29	-30.196	9.252	0.596	-0.376	0.596	-0.376	0.596	-0.376	0.596	-0.376
s30	-30.000	0.000	500z30	-29.545	-0.230	-0.142	-0.645	-0.142	-0.645	-0.142	-0.645	-0.142	-0.645
s31	-30.000	-10.000	500z31	-29.504	-10.460	0.196	0.748	0.196	0.748	0.196	0.748	0.196	0.748
s32	-20.000	-20.000	500z32	-19.094	-20.064	-0.455	0.230	-0.455	0.230	-0.455	0.230	-0.455	0.230
s33	-10.000	-30.000	500z33	-10.470	-30.004	-0.496	0.460	-0.496	0.460	-0.496	0.460	-0.496	0.460
s34	0.000	-30.000	500z34	-0.222	-31.480	-0.906	0.064	-0.906	0.064	-0.906	0.064	-0.906	0.064
s35	10.000	-30.000	500z35	10.360	-29.827	0.470	0.004	0.470	0.004	0.470	0.004	0.470	0.004
s36	20.000	-20.000	500z36	20.667	-19.366	0.222	1.480	0.222	1.480	0.222	1.480	0.222	1.480
s37	30.000	-10.000	500z37	29.567	-10.604	-0.360	-0.173	-0.360	-0.173	-0.360	-0.173	-0.360	-0.173
s38	40.000	0.000	500z38	40.138	0.252	-0.667	-0.634	-0.667	-0.634	-0.667	-0.634	-0.667	-0.634
s39	40.000	10.000	500z39	39.748	10.661	0.433	0.604	0.433	0.604	0.433	0.604	0.433	0.604
s40	40.000	20.000	500z40	40.094	19.266	-0.138	-0.252	-0.138	-0.252	-0.138	-0.252	-0.138	-0.252
s41	30.000	20.000	500z41	30.812	20.285	0.252	-0.661	0.252	-0.661	0.252	-0.661	0.252	-0.661
s42	30.000	30.000	500z42	29.254	29.732	-0.094	0.734	-0.094	0.734	-0.094	0.734	-0.094	0.734
s43	20.000	30.000	500z43	19.745	30.364	-0.812	-0.285	-0.812	-0.285	-0.812	-0.285	-0.812	-0.285
s44	20.000	40.000	500z44	19.466	40.619	0.746	0.268	0.746	0.268	0.746	0.268	0.746	0.268
s45	10.000	40.000	500z45	9.831	39.756	0.255	-0.364	0.255	-0.364	0.255	-0.364	0.255	-0.364
s46	0.000	40.000	500z46	0.103	40.311	0.534	-0.619	0.534	-0.619	0.534	-0.619	0.534	-0.619
s47	-10.000	40.000	500z47	-10.004	39.325	0.169	0.244	0.169	0.244	0.169	0.244	0.169	0.244
s48	-20.000	30.000	500z48	-20.239	30.339	-0.103	-0.311	-0.103	-0.311	-0.103	-0.311	-0.103	-0.311
s49	-20.000	40.000	500z49	-20.020	40.465	0.004	0.675	0.004	0.675	0.004	0.675	0.004	0.675
s50	-30.000	30.000	500z50	-29.496	29.658	0.239	-0.339	0.239	-0.339	0.239	-0.339	0.239	-0.339
s51	-30.000	20.000	500z51	-29.095	20.148	0.020	-0.465	0.020	-0.465	0.020	-0.465	0.020	-0.465
s52	-40.000	20.000	500z52	-39.331	19.139	-0.504	0.342	-0.504	0.342	-0.504	0.342	-0.504	0.342
s53	-40.000	10.000	500z53	-40.633	9.711	-0.905	-0.148	-0.905	-0.148	-0.905	-0.148	-0.905	-0.148
s54	-40.000	0.000	500z54	-39.720	-0.028	-0.669	0.861	-0.669	0.861	-0.669	0.861	-0.669	0.861
s55	-40.000	-10.000	500z55	-41.034	-10.062	0.633	0.289	0.633	0.289	0.633	0.289	0.633	0.289
s56	-40.000	-20.000	500z56	-40.378	-19.779	-0.280	0.028	-0.280	0.028	-0.280	0.028	-0.280	0.028
s57	-30.000	-20.000	500z57	-30.404	-20.658	1.034	0.062	1.034	0.062	1.034	0.062	1.034	0.062

s58	-30.000	-30.000
s59	-20.000	-30.000
s60	-20.000	-40.000
s61	-10.000	-40.000
s62	0.000	-40.000
s63	10.000	-40.000
s64	20.000	-30.000
s65	20.000	-40.000
s66	30.000	-20.000
s67	30.000	-30.000
s68	40.000	-20.000
s69	40.000	-10.000
s70	50.000	0.000
s71	50.000	10.000
s72	50.000	20.000
s73	40.000	30.000
s74	30.000	40.000
s75	20.000	50.000
s76	10.000	50.000
s77	0.000	50.000
s78	-10.000	50.000
s79	-20.000	50.000
s80	-30.000	40.000
s81	-40.000	30.000
s82	-50.000	20.000
s83	-50.000	10.000
s84	-50.000	0.000
s85	-50.000	-10.000
s86	-50.000	-20.000
s87	-40.000	-30.000
s88	-30.000	-40.000
s89	-20.000	-50.000
s90	-10.000	-50.000
s91	0.000	-50.000
s92	10.000	-50.000
s93	20.000	-50.000
s94	30.000	-40.000
s95	40.000	-30.000
s96	50.000	-20.000
s97	50.000	-10.000
s98	60.000	0.000
s99	60.000	10.000
s100	60.000	20.000
s101	50.000	30.000
s102	50.000	40.000
s103	40.000	40.000
s104	40.000	50.000
s105	30.000	50.000
s106	20.000	60.000
s107	10.000	60.000
s108	0.000	60.000
s109	-10.000	60.000
s110	-20.000	60.000
s111	-30.000	50.000
s112	-40.000	40.000
s113	-40.000	50.000
s114	-50.000	30.000
s115	-50.000	40.000
s116	-60.000	20.000
s117	-60.000	10.000
s118	-60.000	0.000
s119	-60.000	-10.000
s120	-60.000	-20.000
s121	-50.000	-30.000
s122	-50.000	-40.000
s123	-40.000	-40.000
s124	-40.000	-50.000
s125	-30.000	-50.000
s126	-20.000	-60.000
s127	-10.000	-60.000
s128	0.000	-60.000
s129	10.000	-60.000
s130	20.000	-60.000
s131	30.000	-50.000
s132	40.000	-40.000
s133	40.000	-50.000
s134	50.000	-40.000
s135	50.000	-30.000
s136	60.000	-20.000
s137	60.000	-10.000

500z58	-30.304	-30.409
500z59	-20.349	-30.330
500z60	-19.703	-40.428
500z61	-10.692	-39.753
500z62	0.226	-40.713
500z63	9.170	-39.613
500z64	19.898	-30.160
500z65	19.374	-40.122
500z66	29.613	-20.395
500z67	29.782	-29.748
500z68	39.878	-20.474
500z69	39.624	-10.263
500z70	50.080	-0.258
500z71	51.010	9.501
500z72	50.629	19.923
500z73	40.886	29.982
500z74	29.747	40.330
500z75	20.022	49.558
500z76	11.034	49.958
500z77	0.748	49.919
500z78	-9.285	49.662
500z79	-19.957	49.971
500z80	-30.205	40.015
500z81	-39.499	29.030
500z82	-50.861	19.598
500z83	-50.207	10.234
500z84	-50.129	0.645
500z85	-50.339	-10.865
500z86	-49.844	-20.269
500z87	-39.306	-30.250
500z88	-30.332	-40.109
500z89	-19.960	-50.011
500z90	-9.987	-50.011
500z91	-0.650	-49.596
500z92	10.216	-49.567
500z93	20.938	-49.337
500z94	29.895	-40.732
500z95	40.193	-29.932
500z96	50.162	-20.603
500z97	49.441	-11.974
500z98	60.084	-0.132
500z99	59.358	10.316
500z100	60.114	20.453
500z101	49.431	30.033
500z102	49.828	41.319
500z103	40.037	40.175
500z104	40.383	49.760
500z105	30.452	50.219
500z106	20.298	60.619
500z107	10.230	59.716
500z108	1.042	60.091
500z109	-9.998	61.021
500z110	-20.405	60.501
500z111	-30.463	50.199
500z112	-39.939	39.400
500z113	-40.054	49.869
500z114	-49.495	29.519
500z115	-50.812	40.263
500z116	-59.703	20.169
500z117	-59.869	10.528
500z118	-60.220	0.499
500z119	-59.912	-10.267
500z120	-60.800	-19.835
500z121	-49.984	-29.709
500z122	-50.681	-39.381
500z123	-40.200	-39.502
500z124	-40.190	-49.390
500z125	-30.123	-50.405
500z126	-20.259	-60.192
500z127	-10.043	-60.334
500z128	-0.092	-60.329
500z129	9.456	-60.425
500z130	19.546	-59.564
500z131	30.362	-50.753
500z132	40.227	-39.710
500z133	40.507	-50.913
500z134	50.085	-40.444
500z135	50.029	-30.311
500z136	59.658	-19.916
500z137	60.118	-8.825

0.304	0.409
0.349	0.330
-0.297	0.428
0.692	-0.247
-0.226	0.713
0.830	-0.387
0.102	0.160
0.626	0.122
0.387	0.395
0.218	-0.252
0.122	0.474
0.376	0.263
-0.080	0.258
-1.010	0.499
-0.629	0.077
-0.886	0.018
0.253	-0.330
-0.022	0.442
-1.034	0.042
-0.748	0.081
-0.715	0.338
-0.043	0.029
0.205	-0.015
-0.501	0.970
0.861	0.402
0.207	-0.234
0.129	-0.645
0.339	0.865
-0.156	0.269
-0.694	0.250
0.332	0.109
-0.040	0.011
-0.013	0.011
0.650	-0.404
-0.216	-0.433
-0.938	-0.663
0.105	0.732
-0.193	-0.068
-0.162	0.603
0.559	1.974
-0.084	0.132
0.642	-0.316
-0.114	-0.453
0.569	-0.033
0.172	-1.319
-0.037	-0.175
-0.383	0.240
-0.452	-0.219
-0.298	-0.619
-0.230	0.284
-1.042	-0.091
-0.002	-1.021
0.405	-0.501
0.463	-0.199
-0.061	0.600
0.054	0.131
-0.505	0.481
0.812	-0.263
-0.297	-0.169
-0.131	-0.528
0.220	-0.499
-0.088	0.267
0.800	-0.165
-0.016	-0.291
0.681	-0.619
0.200	-0.498
0.190	-0.610
0.123	0.405
0.259	0.192
0.043	0.334
0.092	0.329
0.544	0.425
0.454	-0.436
-0.362	0.753
-0.227	-0.290
-0.507	0.913
-0.085	0.444
-0.029	0.311
0.342	-0.084
-0.118	-1.175

0.304	0.409
0.349	0.330
-0.297	0.428
0.692	-0.247
-0.226	0.713
0.830	-0.387
0.102	0.160
0.626	0.122
0.387	0.395
0.218	-0.252
0.122	0.474
0.376	0.263
-0.080	0.258
-1.010	0.499
-0.629	0.077
-0.886	0.018
0.253	-0.330
-0.022	0.442
-1.034	0.042
-0.748	0.081
-0.715	0.338
-0.043	0.029
0.205	-0.015
-0.501	0.970
0.861	0.402
0.207	-0.234
0.129	-0.645
0.339	0.865
-0.156	0.269
-0.694	0.250
0.332	0.109
-0.040	0.011
-0.013	0.011
0.650	-0.404
-0.216	-0.433
-0.938	-0.663
0.105	0.732
-0.193	-0.068
-0.162	0.603
0.419	1.480
-0.084	0.132
0.642	-0.316
-0.114	-0.453
0.569	-0.033
0.172	-1.319
-0.037	-0.175
-0.383	0.240
-0.452	-0.219
-0.298	-0.619
-0.230	0.284
-1.042	-0.091
-0.002	-1.021
0.405	-0.501
0.463	-0.199
-0.061	0.600
0.054	0.131
-0.505	0.481
0.812	-0.263
-0.297	-0.169
-0.131	-0.528
0.220	-0.499
-0.088	0.267
0.800	-0.165
-0.016	-0.291
0.681	-0.619
0.200	-0.498
0.190	-0.610
0.123	0.405
0.259	0.192
0.043	0.334
0.092	0.329
0.544	0.425
0.454	-0.436
-0.362	0.753
-0.227	-0.290
-0.507	0.913
-0.085	0.444
-0.029	0.311
0.342	-0.084
-0.118	-1.175

s138	70.000	0.000
s139	70.000	10.000
s140	70.000	20.000
s141	60.000	30.000
s142	60.000	40.000
s143	50.000	50.000
s144	40.000	60.000
s145	30.000	60.000
s146	20.000	70.000
s147	10.000	70.000
s148	0.000	70.000
s149	-10.000	70.000
s150	-20.000	70.000
s151	-30.000	60.000
s152	-40.000	60.000
s153	-50.000	50.000
s154	-60.000	40.000
s155	-60.000	30.000
s156	-70.000	20.000
s157	-70.000	10.000
s158	-70.000	0.000
s159	-70.000	-10.000
s160	-70.000	-20.000
s161	-60.000	-30.000
s162	-60.000	-40.000
s163	-50.000	-50.000
s164	-40.000	-60.000
s165	-30.000	-60.000
s166	-20.000	-70.000
s167	-10.000	-70.000
s168	0.000	-70.000
s169	10.000	-70.000
s170	20.000	-70.000
s171	30.000	-60.000
s172	40.000	-60.000
s173	50.000	-50.000
s174	60.000	-40.000
s175	60.000	-30.000
s176	70.000	-20.000
s177	70.000	-10.000
s178	80.000	0.000
s179	80.000	10.000
s180	80.000	20.000
s181	70.000	30.000
s182	70.000	40.000
s183	60.000	50.000
s184	60.000	60.000
s185	50.000	60.000
s186	40.000	70.000
s187	30.000	70.000
s188	20.000	80.000
s189	10.000	80.000
s190	0.000	80.000
s191	-10.000	80.000
s192	-20.000	80.000
s193	-30.000	70.000
s194	-40.000	70.000
s195	-50.000	60.000
s196	-60.000	60.000
s197	-60.000	50.000
s198	-70.000	40.000
s199	-70.000	30.000
s200	-80.000	20.000
s201	-80.000	10.000
s202	-80.000	0.000
s203	-80.000	-10.000
s204	-80.000	-20.000
s205	-70.000	-30.000
s206	-70.000	-40.000
s207	-60.000	-50.000
s208	-50.000	-60.000
s209	-40.000	-70.000
s210	-30.000	-70.000
s211	-20.000	-80.000
s212	-10.000	-80.000
s213	0.000	-80.000
s214	10.000	-80.000
s215	20.000	-80.000
s216	30.000	-70.000
s217	40.000	-70.000

500z138	69.018	-0.810
500z139	69.627	9.989
500z140	69.798	20.136
500z141	60.062	30.219
500z142	59.325	41.117
500z143	49.671	50.320
500z144	40.204	59.448
500z145	29.259	60.279
500z146	19.864	70.409
500z147	10.388	71.093
500z148	0.601	69.880
500z149	-9.393	70.472
500z150	-19.113	70.828
500z151	-29.448	60.697
500z152	-40.596	59.704
500z153	-50.429	50.125
500z154	-60.425	39.567
500z155	-59.804	29.474
500z156	-70.099	19.847
500z157	-70.009	10.157
500z158	-69.607	-0.694
500z159	-70.403	-9.858
500z160	-70.559	-19.439
500z161	-59.988	-29.768
500z162	-59.581	-39.313
500z163	-49.352	-48.940
500z164	-40.375	-60.164
500z165	-29.970	-59.739
500z166	-19.891	-70.328
500z167	-11.081	-69.813
500z168	-0.090	-69.279
500z169	10.041	-69.820
500z170	20.573	-70.478
500z171	30.532	-60.145
500z172	39.244	-59.873
500z173	49.780	-50.203
500z174	60.111	-40.031
500z175	59.855	-30.542
500z176	69.909	-20.335
500z177	69.913	-10.351
500z178	80.247	-0.301
500z179	79.780	10.271
500z180	79.990	19.307
500z181	70.572	29.806
500z182	69.321	40.179
500z183	59.896	50.770
500z184	60.505	59.712
500z185	49.946	60.009
500z186	39.946	70.345
500z187	29.824	69.905
500z188	20.326	79.779
500z189	10.205	80.219
500z190	1.019	80.346
500z191	-10.332	79.369
500z192	-19.856	79.643
500z193	-29.646	70.341
500z194	-39.402	69.687
500z195	-49.186	59.621
500z196	-59.729	59.561
500z197	-61.203	49.683
500z198	-69.692	39.841
500z199	-70.786	30.157
500z200	-79.708	19.730
500z201	-80.521	9.760
500z202	-80.567	0.518
500z203	-80.262	-9.318
500z204	-79.896	-20.748
500z205	-71.192	-29.628
500z206	-70.027	-40.133
500z207	-59.980	-50.427
500z208	-48.667	-59.994
500z209	-40.626	-70.239
500z210	-30.031	-70.424
500z211	-19.491	-80.390
500z212	-9.317	-80.011
500z213	0.126	-79.427
500z214	9.680	-79.960
500z215	20.374	-79.575
500z216	30.068	-69.318
500z217	40.378	-70.381

0.982	0.810
0.373	0.011
0.202	-0.136
-0.062	-0.219
0.675	-1.117
0.329	-0.320
-0.204	0.552
0.741	-0.279
0.136	-0.409
-0.388	-1.093
-0.601	0.120
-0.607	-0.472
-0.887	-0.828
-0.552	-0.697
0.596	0.296
0.429	-0.125
0.425	0.433
-0.196	0.526
0.099	0.153
0.009	-0.157
-0.393	0.694
0.403	-0.142
0.559	-0.561
-0.012	-0.232
-0.419	-0.687
-0.648	-1.060
0.375	0.164
-0.030	-0.261
-0.109	0.328
1.081	-0.187
0.090	-0.721
-0.041	-0.180
-0.573	0.478
-0.532	0.145
0.756	-0.127
0.220	0.203
-0.111	0.031
0.145	0.542
0.091	0.335
0.087	0.351
-0.247	0.301
0.220	-0.271
0.010	0.693
-0.572	0.194
0.679	-0.179
0.104	-0.770
-0.505	0.288
0.054	-0.009
0.054	-0.345
0.176	0.095
-0.326	0.221
-0.205	-0.219
-1.019	-0.346
0.332	0.631
-0.144	0.357
-0.354	-0.341
-0.598	0.313
-0.814	0.379
-0.271	0.439
1.203	0.317
-0.308	0.159
0.786	-0.157
-0.292	0.270
0.521	0.240
0.567	-0.518
0.262	-0.682
-0.104	0.748
1.192	-0.372
0.027	0.133
-0.020	0.427
-1.333	-0.006
0.626	0.239
0.031	0.424
-0.509	0.390
-0.683	0.011
-0.126	-0.573
0.320	-0.040
-0.374	-0.425
-0.068	-0.682
-0.378	0.381

0.982	0.810
0.373	0.011
0.202	-0.136
-0.062	-0.219
0.675	-1.117
0.329	-0.320
-0.204	0.552
0.741	-0.279
0.136	-0.409
-0.388	-1.093
-0.601	0.120
-0.607	-0.472
-0.887	-0.828
-0.552	-0.697
0.596	0.296
0.429	-0.125
0.425	0.433
-0.196	0.526
0.099	0.153
0.009	-0.157
-0.393	0.694
0.403	-0.142
0.559	-0.561
-0.012	-0.232
-0.419	-0.687
-0.648	-1.060
0.375	0.164
-0.030	-0.261
-0.109	0.328
1.081	-0.187
0.090	-0.721
-0.041	-0.180
-0.573	0.478
-0.532	0.145
0.756	-0.127
0.220	0.203
-0.111	0.031
0.145	0.542
0.091	0.335
0.087	0.351
-0.247	0.301
0.220	-0.271
0.010	0.693
-0.572	0.194
0.679	-0.179
0.104	-0.770
-0.505	0.288
0.054	-0.009
0.054	-0.345
0.176	0.095
-0.326	0.221
-0.205	-0.219
-1.019	-0.346
0.332	0.631
-0.144	0.357
-0.354	-0.341
-0.598	0.313
-0.814	0.379
-0.271	0.439
1.203	0.317
-0.308	0.159
0.786	-0.157
-0.292	0.270
0.521	0.240
0.567	-0.518
0.262	-0.682
-0.104	0.748
1.192	-0.372
0.027	0.133
-0.020	0.427
-1.333	-0.006
0.626	0.239
0.031	0.424
-0.509	0.390
-0.683	0.011
-0.126	-0.573
0.320	-0.040
-0.374	-0.425
-0.068	-0.682
-0.378	0.381

s218	50.000	-60.000
s219	60.000	-50.000
s220	70.000	-40.000
s221	70.000	-30.000
s222	80.000	-20.000
s223	80.000	-10.000
s224	90.000	0.000
s225	90.000	10.000
s226	90.000	20.000
s227	80.000	30.000
s228	90.000	30.000
s229	80.000	40.000
s230	70.000	50.000
s231	80.000	50.000
s232	70.000	60.000
s233	60.000	70.000
s234	50.000	70.000
s235	50.000	80.000
s236	40.000	80.000
s237	30.000	80.000
s238	30.000	90.000
s239	20.000	90.000
s240	10.000	90.000
s241	0.000	90.000
s242	-10.000	90.000
s243	-20.000	90.000
s244	-30.000	80.000
s245	-30.000	90.000
s246	-40.000	80.000
s247	-50.000	80.000
s248	-50.000	70.000
s249	-60.000	70.000
s250	-70.000	60.000
s251	-70.000	50.000
s252	-80.000	50.000
s253	-80.000	40.000
s254	-80.000	30.000
s255	-90.000	30.000
s256	-90.000	20.000
s257	-90.000	10.000
s258	-90.000	0.000
s259	-90.000	-10.000
s260	-90.000	-20.000
s261	-80.000	-30.000
s262	-90.000	-30.000
s263	-80.000	-40.000
s264	-70.000	-50.000
s265	-70.000	-60.000
s266	-60.000	-60.000
s267	-60.000	-70.000
s268	-50.000	-70.000
s269	-50.000	-80.000
s270	-40.000	-80.000
s271	-30.000	-80.000
s272	-30.000	-90.000
s273	-20.000	-90.000
s274	-10.000	-90.000
s275	0.000	-90.000
s276	10.000	-90.000
s277	20.000	-90.000
s278	30.000	-80.000
s279	30.000	-90.000
s280	40.000	-80.000
s281	50.000	-80.000
s282	50.000	-70.000
s283	60.000	-70.000
s284	60.000	-60.000
s285	70.000	-50.000
s286	70.000	-60.000
s287	80.000	-50.000
s288	80.000	-40.000
s289	80.000	-30.000
s290	90.000	-30.000
s291	90.000	-20.000
s292	90.000	-10.000
s293	100.000	0.000
s294	110.000	0.000

500z218	49.641	-60.266
500z219	60.086	-50.516
500z220	69.895	-40.037
500z221	70.435	-30.041
500z222	79.056	-19.482
500z223	79.845	-10.661
500z224	90.104	0.458
500z225	89.130	9.266
500z226	89.823	19.566
500z227	79.946	30.226
500z228	89.946	29.319
500z229	80.246	39.884
500z230	70.422	49.851
500z231	79.773	49.989
500z232	69.518	59.989
500z233	60.223	70.751
500z234	49.988	69.989
500z235	51.052	79.989
500z236	39.770	79.954
500z237	30.373	79.954
500z238	29.812	90.568
500z239	20.063	89.357
500z240	10.799	89.762
500z241	0.720	90.147
500z242	-9.248	90.979
500z243	-19.508	89.306
500z244	-29.990	79.868
500z245	-31.224	90.120
500z246	-40.187	80.646
500z247	-50.067	80.799
500z248	-50.301	70.294
500z249	-59.984	70.639
500z250	-68.889	59.500
500z251	-69.786	49.692
500z252	-80.685	49.581
500z253	-79.769	39.770
500z254	-80.734	30.417
500z255	-90.473	29.272
500z256	-89.943	20.259
500z257	-90.644	10.518
500z258	-90.054	-1.000
500z259	-90.054	-9.626
500z260	-90.357	-19.160
500z261	-79.686	-30.212
500z262	-89.615	-29.934
500z263	-79.401	-40.551
500z264	-69.260	-49.726
500z265	-70.301	-59.750
500z266	-59.656	-60.443
500z267	-60.272	-69.948
500z268	-49.899	-69.388
500z269	-50.895	-80.684
500z270	-39.806	-80.260
500z271	-29.659	-79.606
500z272	-30.129	-89.280
500z273	-19.875	-89.632
500z274	-9.483	-89.847
500z275	-0.498	-88.831
500z276	9.848	-90.059
500z277	20.067	-89.946
500z278	29.476	-79.429
500z279	29.946	-89.628
500z280	39.469	-79.689
500z281	49.802	-78.887
500z282	49.503	-69.185
500z283	60.089	-70.376
500z284	60.889	-60.531
500z285	68.962	-50.415
500z286	70.282	-59.686
500z287	79.458	-50.245
500z288	79.742	-40.140
500z289	79.890	-30.320
500z290	90.018	-29.945
500z291	91.347	-18.280
500z292	90.335	-9.794
500z293	100.603	-0.629
500z294	109.969	-0.424

0.359	0.266
-0.086	0.516
0.105	0.037
-0.435	0.041
0.944	-0.518
0.155	0.661
-0.104	-0.458
0.870	0.734
0.177	0.434
0.054	-0.226
0.054	0.681
-0.246	0.116
-0.422	0.149
0.227	0.011
0.482	0.011
-0.223	-0.751
0.012	0.011
-1.052	0.011
0.230	0.046
-0.373	0.046
0.188	-0.568
-0.063	0.643
-0.799	0.238
-0.720	-0.147
-0.752	-0.979
-0.492	0.694
-0.010	0.132
1.224	-0.120
0.187	-0.646
0.067	-0.799
0.301	-0.294
-0.016	-0.639
-1.111	0.500
-0.214	0.308
0.685	0.419
-0.231	0.230
0.734	-0.417
0.473	0.728
-0.057	-0.259
0.644	-0.518
0.054	1.000
0.054	-0.374
0.357	-0.840
-0.314	0.212
-0.385	-0.066
-0.599	0.551
-0.740	-0.274
0.301	-0.250
-0.344	0.443
0.272	-0.052
-0.101	-0.612
0.895	0.684
-0.194	0.260
-0.341	-0.394
0.129	-0.720
-0.125	-0.368
-0.517	-0.153
0.498	-1.169
0.152	0.059
-0.067	-0.054
0.524	-0.571
0.054	-0.372
0.531	-0.311
0.198	-1.113
0.497	-0.815
-0.089	0.376
-0.889	0.531
1.038	0.415
-0.282	-0.314
0.542	0.245
0.258	0.140
0.110	0.320
-0.018	-0.055
-1.347	-1.720
-0.335	-0.206
-0.603	0.629
0.031	0.424

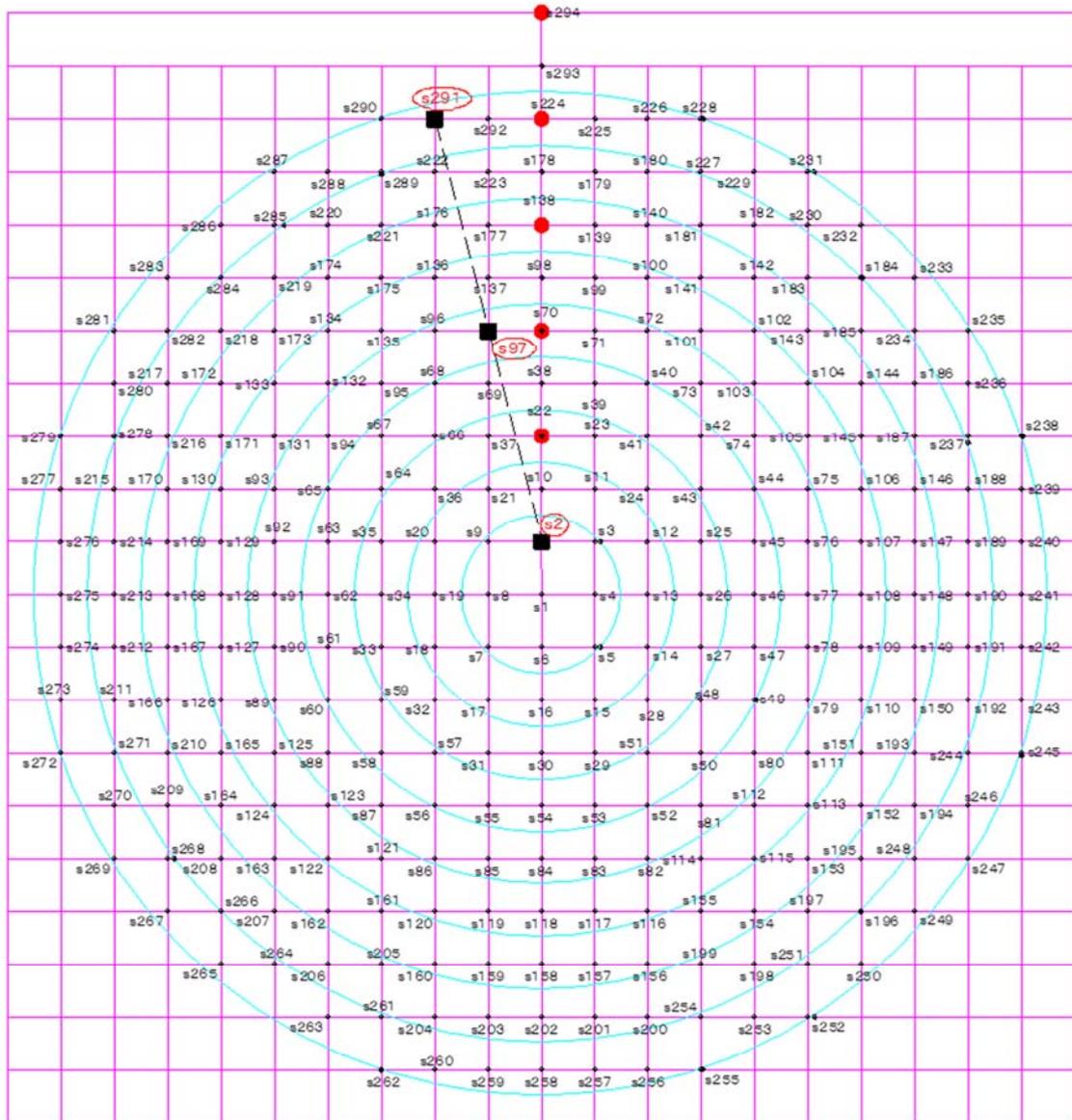
0.359	0.266
-0.086	0.516
0.105	0.037
-0.435	0.041
0.944	-0.518
0.155	0.661
-0.104	-0.458
0.870	0.734
0.177	0.434
0.054	-0.226
0.054	0.681
-0.246	0.116
-0.422	0.149
0.227	0.011
0.482	0.011
-0.223	-0.751
0.012	0.011
-1.052	0.011
0.230	0.046
-0.373	0.046
0.188	-0.568
-0.063	0.643
-0.799	0.238
-0.720	-0.147
-0.752	-0.979
-0.492	0.694
-0.010	0.132
1.224	-0.120
0.187	-0.646
0.067	-0.799
0.301	-0.294
-0.016	-0.639
-1.111	0.500
-0.214	0.308
0.685	0.419
-0.231	0.230
0.734	-0.417
0.473	0.728
-0.057	-0.259
0.644	-0.518
0.054	1.000
0.054	-0.374
0.357	-0.840
-0.314	0.212
-0.385	-0.066
-0.599	0.551
-0.740	-0.274
0.301	-0.250
-0.344	0.443
0.272	-0.052
-0.101	-0.612
0.895	0.684
-0.194	0.260
-0.341	-0.394
0.129	-0.720
-0.125	-0.368
-0.517	-0.153
0.498	-1.169
0.152	0.059
-0.067	-0.054
0.524	-0.571
0.054	-0.372
0.531	-0.311
0.198	-1.113
0.497	-0.815
-0.089	0.376
-0.889	0.531
1.038	0.415
-0.282	-0.314
0.542	0.245
0.258	0.140
0.110	0.320
-0.018	-0.055
-1.010	-1.290
-0.335	-0.206
-0.603	0.629
0.031	0.424

点の位置図

前の座標データの配置図になります。

□の s2, s97, s291 は大きな誤差を与えた点の位置で s2 は図の中心部, s97 は中間部, s291 は図形の外枠部に位置します。

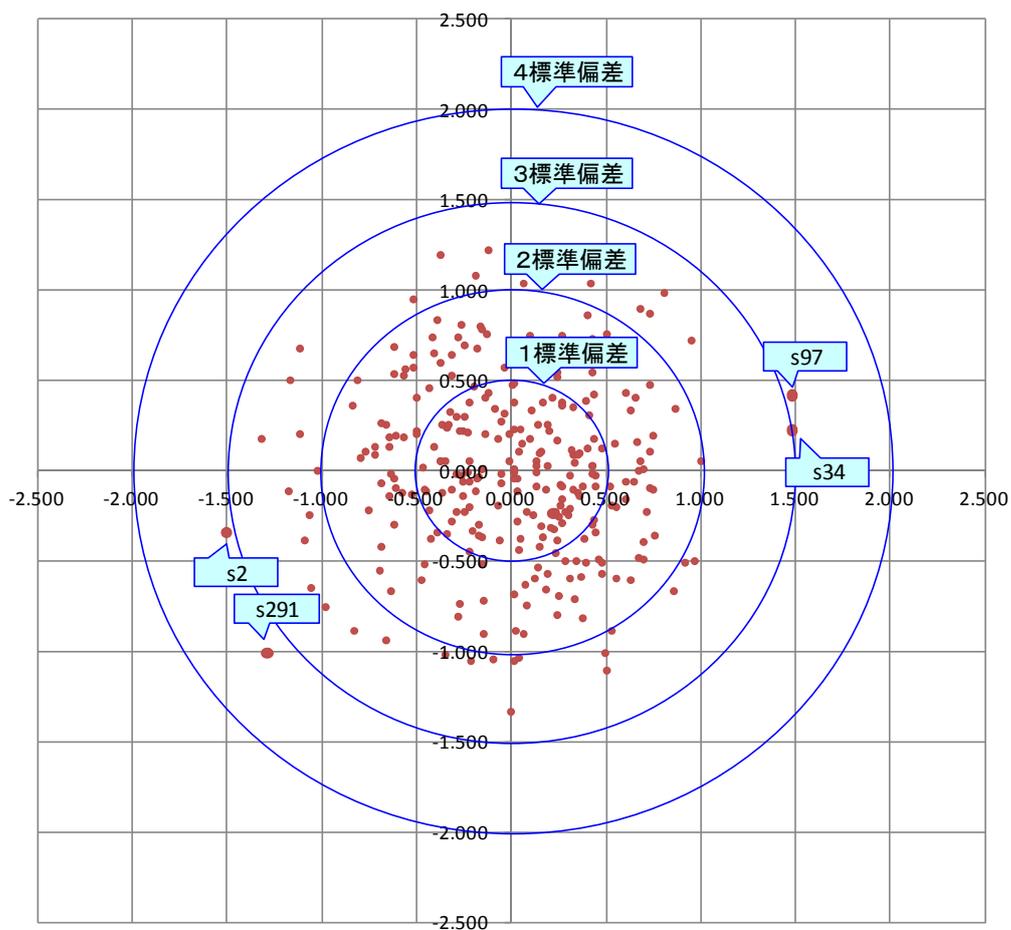
最小二乗法座標変換においてこの位置の違いによって、復元精度、異常点として認識されるレベルが違います、そのことも含めて説明するために図の位置に配置してあります。



異常点の無い散布図

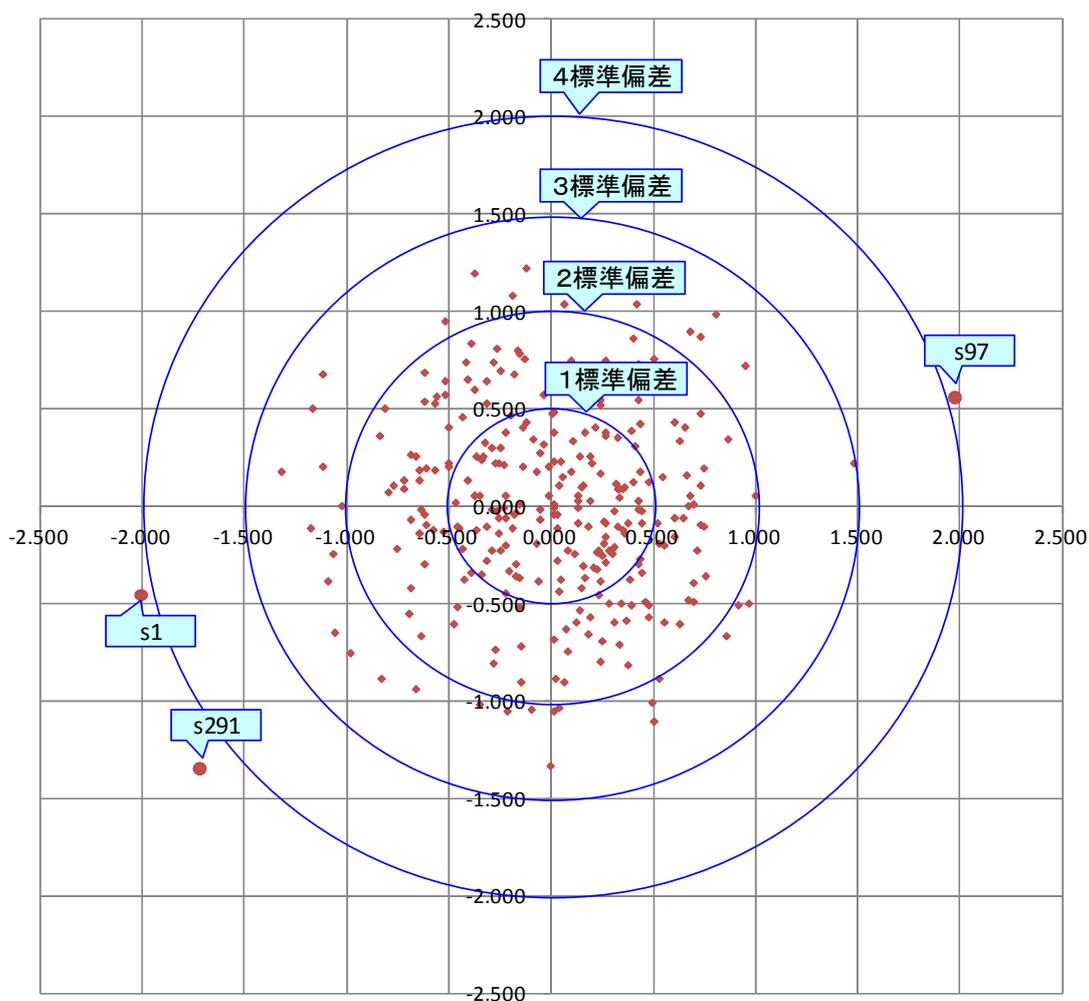
下図が誤差の散布図です、散布図の円は内側から1倍標準偏差～4倍標準偏差の誤差楕円です。

分布上は異常点といわれる点が無い状態ですが、 3σ の円の外側近くに s2, s34, s97, s291 の点があり、これが異常値と判定される可能性があります。



異常点のある散布図

次図は異常点 s2, s97, s291 の 3 個の位置を確認しますと 4 倍標準偏差の外側, 5 倍標準偏差の位置にあります, これは極端に大きな誤差を与えている訳ではありません。このデータは異常点を除けば正規分布になっているデータです。



このデータからランダムに指定した個数を選んで二変量標準偏差を求めます。点の組合せ条件によって標準偏差の値は細かく変わりますが概ね標準偏差で 0.476 以下と理解下さい。判断データをヘルマート変換, アフィン変換結果から取得していることからヘルマート変換では計算上の未知数が2個ある関係で最小の組合せ3個の座標値から1個のデータが取得できます, t検定を使ってそのデータを判断するには連続した3個以上のデータが必要なので異常なデータが無い状態で最低でも5個の座標値が必要になります。

アフィン変換では計算上の未知数が3個ある関係で最小の組合せ4個の座標値から1個のデータが取得できます, t 検定を使ってそのデータを判断するには連続した4個以上のデー

タが必要なので異常なデータが無い状態で最低でも6個の座標値が必要になります。

ここではヘルマート変換を使って検証していますので、検証には 5 個～13 個あった場合を想定し、異常値の無い状態では13個まで、1個の異常値を含んだ場合は14個まで、2個の異常値を含んだ場合は15個までを想定して計算します。

誤差は正規分布になるということを前提にすればデータの誤差を調べて正規分布から外れるデータを異常値として除外すればよいわけですが、あるデータが異常値かどうかを判定して異常値であれば除外して、残ったデータを準拠点として最小二乗法座標変換を行えば適正な結果が得られるということになります。

この異常値を除く方法は①経験と勘から ②t 異常値検定 ③ χ^2 正規分布適合度検定 の方法があります。他にも工夫すれば考えられるのですが①に期待することは適切ではありません、また χ^2 二乗検定はデータ数が31個以上の時に使う検定とされていますので最初に「② t 異常値検定」で次に「③ χ^2 正規分布適合度検定」について検証してみます。

この3個の異常点は極端に悪いデータではないため準拠点対象点が比較的精度の悪い点の集合の場合、異常点として検出されない可能性のある点です、このことを頭の隅に置いておいてください。

検定データはなに

意外と重要なことですが検定に使うデータは何が適切なのでしょうか、このことが重要なのですがほとんど議論されることはありませんでした、そしてこのことが、準拠点選択が一般的に使われてこなかったことの一因になっていましたと言いますか準拠点選択という考えに至らなかった一因です。

2017年現在、使っているのは二変量標準偏差の値です、この二変量標準偏差は**変数減少法**で求めます。

変数減少法とは、例えば 10 個の境界点のデータ(図面值と実測値)がある場合、① 10 個で座標変換をし、二変量標準偏差を求め x_1 とします、② 図面值と実測値は完全に一致しませんので 10 個の中から乖離の一番大きい図面值と実測値 δ_1 を除きます、③ 残った 9 点で再度座標変換して x_2 を求めます。④ この操作を二変量標準偏差が 0 になるまで繰り返します。座標変換がヘルマート変換の場合は残り座標値が2個で 0 になりますので 8 個の検定データが得られます。座標変換がアフィン変換の場合は残り座標値が3個で 0 になりますので 7 個の検定データが得られます。

⑤ 得られたデータを順番に並べて $x_1 x_2 x_3 \cdots x_n$ 座標データと除いた座標値点名はセットになっていますので $x_1 x_2 x_3 \cdots x_n$ の値を検定して 検定値(t検定ではt値、 χ^2 検定では χ^2 値)に有意差がなくなるまで大きい順に該当する点を除いていきます。

手順で③を入れる理由は 1 回目に変換したとき2番目のデータが③の変換を入れときに 1 番目に来ないことがあるからです、このことは座標変換の特徴です、③を繰り返し入れることで誤ったデータを取得しないためです。

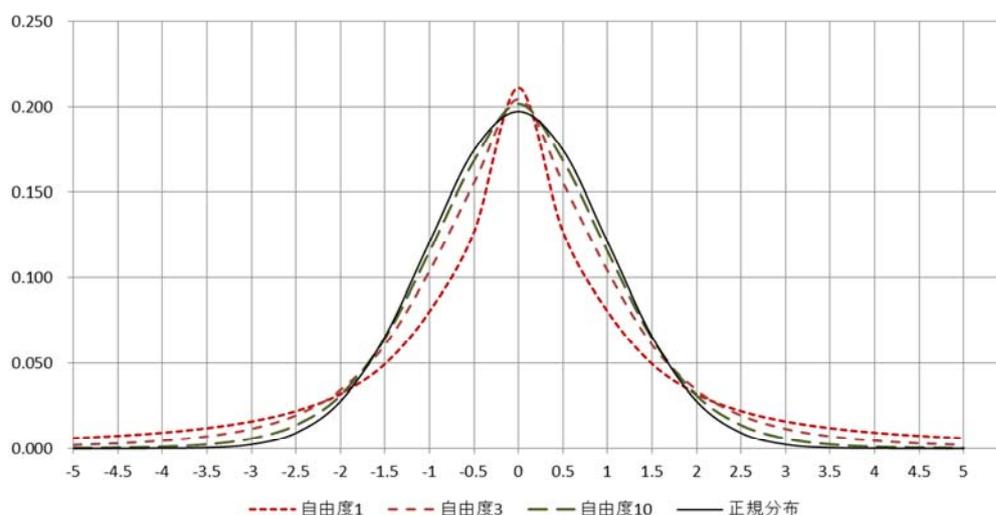
検定注意事項

境界(筆界)復元に伴う準拠点選択の検定で注意しなければならない事項は次のことです、 x_1 x_2 x_3 \cdots x_n と並べて検定した結果 x_1 (有意差あり) x_2 (有意差なし) x_3 (有意差あり) x_4 (有意差なし) \cdots x_n のように 1 回だけ「有意差なし」が表れることがあります、この場合は「有意なし」が 2 回連続するまでは「有意差あり」として処理しています、この場合は x_2 も「有意差あり」としています。偶然に点の組合せの結果一度だけ「有意差なし」が表れることがありますので、少なくとも二度、連続で表れることが必要です。

t検定

t分布という確率分布

データ数が無限に大きいデータは正規分布になります(下図の実線), データ数が少なくなるとベルカーブ(正規分布の曲線を別名ベルカーブという)の裾が広がって, 中心部が尖ってきます, 下図の自由度はデータ数-1で, 自由度1, 自由度3, 自由度10と正規分布の密度曲線を表しています。



各自由度に於ける密度は下表のとおりで。このデータをグラフにしたものです。

目盛 σ	自由度						正規分布
	1	3	10	30	50	1000	
-5	0.0061	0.0021	0.0002	0.0000	0.0000	0.0000	0.0000
-4.5	0.0075	0.0031	0.0005	0.0001	0.0000	0.0000	0.0000
-4	0.0094	0.0046	0.0010	0.0003	0.0002	0.0001	0.0001
-3.5	0.0121	0.0072	0.0025	0.0010	0.0008	0.0005	0.0005
-3	0.0160	0.0116	0.0059	0.0036	0.0031	0.0024	0.0024
-2.5	0.0221	0.0196	0.0139	0.0110	0.0103	0.0093	0.0092
-2	0.0321	0.0342	0.0313	0.0292	0.0287	0.0279	0.0278
-1.5	0.0495	0.0608	0.0645	0.0653	0.0654	0.0656	0.0656
-1	0.0804	0.1039	0.1154	0.1191	0.1198	0.1209	0.1210
-0.5	0.1269	0.1555	0.1686	0.1726	0.1734	0.1746	0.1747
0	0.2110	0.2047	0.2019	0.2011	0.2009	0.2007	0.1974
0.5	0.1269	0.1555	0.1686	0.1726	0.1734	0.1746	0.1747
1	0.0804	0.1039	0.1154	0.1191	0.1198	0.1209	0.1210
1.5	0.0495	0.0608	0.0645	0.0653	0.0654	0.0656	0.0656
2	0.0321	0.0342	0.0313	0.0292	0.0287	0.0279	0.0278
2.5	0.0221	0.0196	0.0139	0.0110	0.0103	0.0093	0.0092
3	0.0160	0.0116	0.0059	0.0036	0.0031	0.0024	0.0024
3.5	0.0121	0.0072	0.0025	0.0010	0.0008	0.0005	0.0005
4	0.0094	0.0046	0.0010	0.0003	0.0002	0.0001	0.0001
4.5	0.0075	0.0031	0.0005	0.0001	0.0000	0.0000	0.0000
5	0.0061	0.0021	0.0002	0.0000	0.0000	0.0000	0.0000

有意水準と自由度とt値の関係は下表のとおりです。正規分布では有意水準 0.05 (確率 95%)

)ではt値は1.960となります,これは正規分布確率95%に於ける標準偏差の倍数1.960倍とおなじです。このことからデータ数が少ないときに正規分布に代わって使われるのがt分布です。詳しいことは統計書等で学習して下さい。

有意水準	0.05	0.05	0.05	0.05	0.05
自由度(n-1)	1	3	10	30	1000
t 値	12.706	3.182	2.228	2.042	1.962

t異常値検定の方法の説明

境界(筆界)復元において準拠点選択で使うt検定はt検定分布におけるt値を使った異常値検定です,観測値 $x_1 x_2 x_3 \dots x_n$ の中で x_k が他の値に比べて大きく離れている,つまり誤差が大きいようだとすると x_k は異常値かどうかを検定する方法です。

異常値と判断すればその後の計算にはこの値は使用せず,異常値の無い状態を調べて,その計算値を採用することになります。

t検定はデータ数が30個以下の場合に使われる方法とされています,方法はt期待値(t理論値)とt観測値(t実測値)を求めてt期待値(t理論値) > t観測値(t実測値)であれば「有意差無し」としその一連のデータには異常なデータが無いとします,t期待値(t理論値) < t観測値(t実測値)であれば「有意差あり」としてその一連のデータの最大値を異常なデータとして除くという方法です。

t期待値= $t_{\text{有意水準, 自由度}}$ でt値の関数から求められます,表計算ソフト(Excel等)で求められます,Excelの関数は $TINV_{\text{有意水準, 自由度}}$ です。

有意水準は「**有意水準を幾つにするか**」を参照して下さい,ここでは0.05(5%)を有意水準として使っています。

自由度はデータ数 -2 (n-2) です。

t観測値は次式から

$$\text{データの平均値 } \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

$$\text{分散 } \sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

$$\tau = \frac{x_k - \bar{x}}{\sigma}$$

$$t(\text{観測値}) = \frac{\sqrt{n-2} \cdot \tau}{\sqrt{n-1-\tau^2}}$$

で求めます,これを次のデータで検定してみます。

データは変数減少法によって計算していますのでデータ数は 13, 12, 11, … 3 までと減少していきま

す。この計算表の方法ではデータを降順(大きい値から小さい値順)に並べる必要はありません。

右欄の「あり」「なし」は有意差があるか無いかの判定をしています。

データ数	データ
13	0.590
12	0.517
11	0.412
10	0.371
9	0.335
8	0.285
7	0.252
6	0.220
5	0.182
4	0.075
3	0.030

1回目の計算 異常点無し

結果はデータ 1~10 迄の一群のデータで「有意差無し」の判定になります, つまり異常値は無いということです。

有意水準	0.05							
自由度	10	n	11					
データ順	t期待値	データ	平均値	標準偏差	τ	t	有意の有無	
1	2.228	0.590	0.297	0.164	1.785	2.051	なし	
2	2.228	0.517	0.297	0.164	1.342	1.405		
3	2.228	0.412	0.297	0.164	0.697	0.677		
4	2.228	0.371	0.297	0.164	0.449	0.430		
5	2.228	0.335	0.297	0.164	0.231	0.220		
6	2.228	0.285	0.297	0.164	-0.074	-0.070		
7	2.228	0.252	0.297	0.164	-0.276	-0.263		
8	2.228	0.220	0.297	0.164	-0.471	-0.452		
9	2.228	0.182	0.297	0.164	-0.700	-0.681		
10	2.228	0.075	0.297	0.164	-1.352	-1.419		
11	2.228	0.030	0.297	0.164	-1.629	-1.803		

2 回目の計算 異常点無し

一番目のデータを除いてデータ 2~10 で計算して見ます, 一群のデータで「有意差無し」の判定になります, つまり異常値は無いということです。

有意水準	0.05							
自由度	9	n	11					
データ順	t期待値	データ	平均値	標準偏差	τ	t		有意の有無
2	2.262	0.517	0.268	0.142	1.756	2.002		なし
3	2.262	0.412	0.268	0.142	1.011	1.012		
4	2.262	0.371	0.268	0.142	0.724	0.706		
5	2.262	0.335	0.268	0.142	0.473	0.454		
6	2.262	0.285	0.268	0.142	0.121	0.114		
7	2.262	0.252	0.268	0.142	-0.113	-0.107		
8	2.262	0.220	0.268	0.142	-0.338	-0.322		
9	2.262	0.182	0.268	0.142	-0.602	-0.582		
10	2.262	0.075	0.268	0.142	-1.356	-1.424		
11	2.262	0.030	0.268	0.142	-1.675	-1.874		

ここまでの解説は統計書などで解説されている内容になります。

これでいいのか

このデータは次の 13 個です, 12 個までは異常値(正規分布から外れるデータ)がないデータ群の中に s291 の異常値(正規分布から外れるデータ)を 1 個混ぜてあるものです。

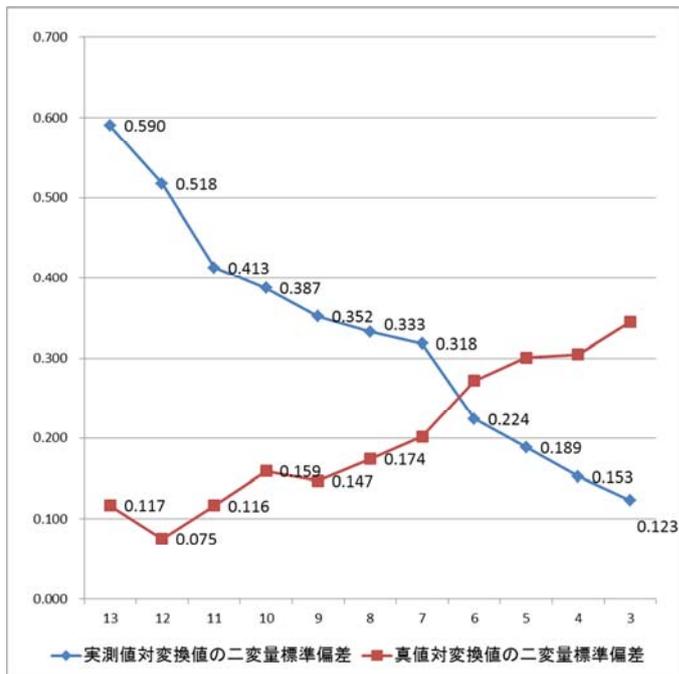
ですから, この s291 は検定で異常値(正規分布から外れるデータ)として除かれて欲しいと言うのが本音です。

対象データ

s148
s179
s209
s9
s178
s259
s114
s102
s85
s80
s60
s240
s291

これを変数減少法で復元精度の関係をみると次のグラフになります。復元精度(真値-変換値の二変量標準偏差)で見れば 12 点目が最適な準拠点選択になります, 13, 12, 11 点のうちのもどれでも大きな差はありません。

母集団標準偏差(実測値-変換値二変量標準偏差)で見れば母集団標準偏差は 0.48 程度ですからこれより小さい最初の値, 11 点目が最適な準拠点選択となります。

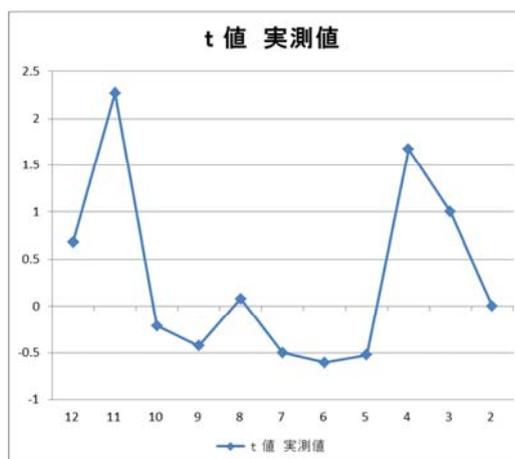
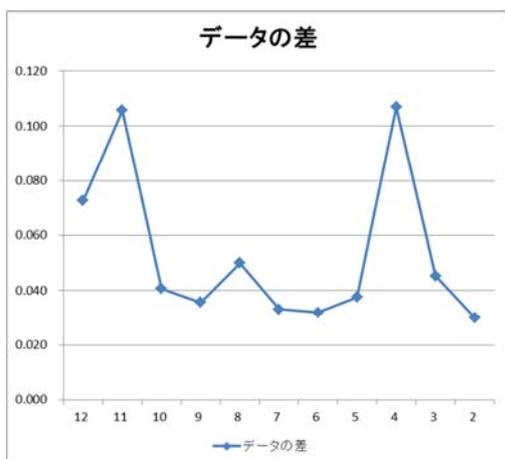


境界（筆界）復元の準拠点選択に使われている χ^2 （カイジジョウ）検定と t 検定は χ^2 （カイジジョウ）検定が最初に実用化になりました，これは統計を知っている方であれば意外なことですが難しい方法が最初に実用化され，比較的簡単な方法が後回しになったわけですから統計書のとおりやっていたら実用化に至らなかったでしょう。

χ^2 （カイジジョウ）検定の問題は準拠点数が少ないと準拠点選択に安定性が無く，信頼性に欠けることでした。そのために何とか実用に耐える方法を考える必要があったのです，最大の理由は境界（筆界）復元に使う準拠点は 10 点前後が圧倒的に多いからです。

多くのグラフの傾向に復元精度（真値－変換値の二変量標準偏差）の減少に他より大きな変化点が生じた個所の復元精度が高いことが解っていたのでこの変化点を変化量として捉えれば可能な訳です。

この差と t 検定が判定の基礎としている t 観測値（実測値）について変数減少法により求めグラフにしたのが下図です。



この結果から、データ差を t 検定のデータとして使えば実用に至ると考えました。

t 異常値検定の「差のデータ」を使うことを簡単に書けば「偶然、この方法に辿りついた」と言うこととなります。前もって χ^2 (カイジジョウ) 検定の方法を確立してありましたのでその結果を踏まえて χ^2 検定と大きな差が生じない方法が必要なことを前提に考えたのです、それと境界(筆界)復元に於ける準拠点は境界の経年変化、境界を測量してきた土地家屋調査士の測量レベル (レベルが低いこと) を考慮すれば通常の検定で得られる結果より厳しい結果が得られる必要があります、特に地租改正地引絵図、地押し調査更正図、平板測量図面を使う場合は重要です。

χ^2 検定では相関関係のない二変量標準偏差にどの式を使うか、偽相関係数の評価をどのようにするかがありますが他の手順は統計的方法をそのまま使って結果が得られます。そのため、 χ^2 検定ではある程度厳しい準拠点選択がなされます、境界測量ではこの偽相関係数が予測できないため準拠点選択結果が通常の検定より厳しい結果が得られることは精度の高い境界(筆界)復元に繋がっている、といいますか本来の境界(筆界)復元に近い値が求められていると言えます (統計学に詳しくない者にとっては偶然の産物でした)。

t 異常値検定で「差のデータ」を使うと通常の検定に比べて 0.23 点程度余分に異常値と判断される傾向があります、これは χ^2 検定と同じ程度の準拠点選択結果になります。

χ^2 検定では準拠点候補が少ないと精度の高い準拠点選択がされないと言う欠点があり、このことが土地家屋調査士のおこなう筆界復元では致命的な欠点になります、これを解消するために t 異常値検定が必要な訳です。

2009年にこの方法を公開したわけですが2007年から2008年の約1年間かけて思いつき、さらに1年間ほどかけて様々なデータで試算しました。

このようなことが可能になった背景には1995年のWindows95とOffice95の発売が大きく影響しています、特にWindows2000とOffice2000の発売が大きかったです、私もこのシステムを使ってプログラム (Book) を作成しました。

差のデータ(差)

データ数	11		
残点数	標準偏差	差	最大値点名
13	0.590	0.073	500z291
12	0.517	0.106	500z102
11	0.412	0.041	500z179
10	0.371	0.036	500z9
9	0.335	0.050	500z85
8	0.285	0.033	500z114
7	0.252	0.032	500z60
6	0.220	0.038	500z178
5	0.182	0.107	500z80
4	0.075	0.045	500z209
3	0.030	0.030	500z148

1回目 異常点無し

有意水準	0.05							
自由度	10	n	11					
データ順	t期待値	データ	平均値	標準偏差	τ	t	有意の有無	
1	2.228	0.073	0.054	0.027	0.697	0.678	なし	
2	2.228	0.106	0.054	0.027	1.910	2.274		
3	2.228	0.041	0.054	0.027	-0.475	-0.455		
4	2.228	0.036	0.054	0.027	-0.657	-0.637		
5	2.228	0.050	0.054	0.027	-0.132	-0.125		
6	2.228	0.033	0.054	0.027	-0.751	-0.734		
7	2.228	0.032	0.054	0.027	-0.794	-0.778		
8	2.228	0.038	0.054	0.027	-0.586	-0.566		
9	2.228	0.107	0.054	0.027	1.954	2.357		
10	2.228	0.045	0.054	0.027	-0.303	-0.288		
11	2.228	0.030	0.054	0.027	-0.864	-0.852		

2回目 異常点あり

有意水準	0.05							
自由度	9	n	11					
データ順	t期待値	データ	平均値	標準偏差	τ	t	有意の有無	
2	2.262	0.106	0.052	0.028	1.935	2.321	あり	
3	2.262	0.041	0.052	0.028	-0.396	-0.378		
4	2.262	0.036	0.052	0.028	-0.574	-0.554		
5	2.262	0.050	0.052	0.028	-0.060	-0.057		
6	2.262	0.033	0.052	0.028	-0.666	-0.646		
7	2.262	0.032	0.052	0.028	-0.708	-0.689		
8	2.262	0.038	0.052	0.028	-0.505	-0.485		
9	2.262	0.107	0.052	0.028	1.978	2.404		
10	2.262	0.045	0.052	0.028	-0.228	-0.217		
11	2.262	0.030	0.052	0.028	-0.776	-0.759		

3回目 異常点無し

有意水準		0.05							
自由度		8							
データ順	t期待値	データ	n	10	平均値	標準偏差	τ	t	有意の有無
3	2.306	0.041	0.046	0.023	-0.224	-0.212	なし		
4	2.306	0.036	0.046	0.023	-0.446	-0.425			
5	2.306	0.050	0.046	0.023	0.192	0.181			
6	2.306	0.033	0.046	0.023	-0.560	-0.538			
7	2.306	0.032	0.046	0.023	-0.612	-0.590			
8	2.306	0.038	0.046	0.023	-0.360	-0.342			
9	2.306	0.107	0.046	0.023	2.722	6.108			
10	2.306	0.045	0.046	0.023	-0.016	-0.015			
11	2.306	0.030	0.046	0.023	-0.697	-0.675			

4回目 異常点無し

有意水準		0.05							
自由度		7							
データ順	t期待値	データ	n	9	平均値	標準偏差	τ	t	有意の有無
4	2.365	0.036	0.046	0.024	-0.448	-0.424	なし		
5	2.365	0.050	0.046	0.024	0.155	0.145			
6	2.365	0.033	0.046	0.024	-0.556	-0.531			
7	2.365	0.032	0.046	0.024	-0.606	-0.580			
8	2.365	0.038	0.046	0.024	-0.367	-0.346			
9	2.365	0.107	0.046	0.024	2.548	5.493			
10	2.365	0.045	0.046	0.024	-0.041	-0.039			
11	2.365	0.030	0.046	0.024	-0.685	-0.661			

この結果から2回目のデータ(点名 500z102)に異常点があったので、1回目のデータ(点名 500z291)も異常点と判断します。(前項の**検定注意事項**から)

計算を一括でおこなえるようにプログラムすると

有意水準		0.05							
データ数	t期待値	データ	平均値	標準偏差	τ	t	有意の有無	最大値点名	
↑	11	2.228	0.073	0.054	0.027	0.697	0.678	なし	500z291
	10	2.262	0.106	0.052	0.028	1.935	2.388	あり	500z102
	9	2.306	0.041	0.046	0.023	-0.224	-0.211	なし	500z179
↓	8	2.365	0.036	0.046	0.024	-0.448	-0.421	なし	500z9
	7	2.447	0.050	0.048	0.025	0.086	0.079	なし	500z85
	6	2.571	0.033	0.048	0.027	-0.532	-0.490	なし	500z114
	5	2.776	0.032	0.050	0.029	-0.641	-0.586	なし	500z60
	4	3.182	0.038	0.055	0.031	-0.570	-0.493	なし	500z178
	3	4.303	0.107	0.061	0.033	1.389	5.222	あり	500z80
	2	12.706	0.045	0.038	0.008	1.000	0.000	なし	500z209
	1		0.030						500z148

判定は全データの約3分の1の範囲内(データ数8~11 ↔の範囲)でおこないます。データと最大値点名、その点数の中の点で最も精度が悪い点とセットにしてあるので右側の点名

500z291 が異常点として判ります。

この標準偏差の差を使う方法では異常点になる割合が高くなります，概ね 0.25 点程度ですが準拠点として使うことを考えればゆるい基準より多少きつい基準が適していますのでこの方法を推奨します。

有意水準を幾つにするか

t検定におけるt値の決定要素は自由度と有意水準です，自由度はデータ数によって決定される要素ですから「データ数-2」 (n-2) を使います。

有意水準は高精度なデータでは 1%，一般的には 5%，底精度では 10% と言われています，統計書とか実務書でもこの値を採用している例が多いようですがこれが果たして正しいのかという疑問もあります。

このように数値を固定する方法はコンピュータの普及していない時代に条件によって有意水準を変化させることが大変な時代考えられた手法ともいえます，現代では適切な方法とは考えられない可能性がありますので，検証が必要です。

そこで有意水準に使う数値の検証をしてみます。有意水準の逆数は信頼率でもあり 10 個のデータから得られた数値と 1000 個のデータから得られた有意水準が同じとしても信頼性も同じか，と考えれば“違う”と誰でも考えるでしょう，つまり有意水準とは検定したデータの信頼性に関係するものであり，5%であれば 20 分の 1 になり 20 回の検定中 1 回は間違っている可能性がありますということになります，これは検定に使うデータ数 n と逆比例する，データ数が多ければ信頼度は高くなりますのでデータ数の逆数 $\frac{1}{n}$ を有意水準としてもよいのではないかと考えます。

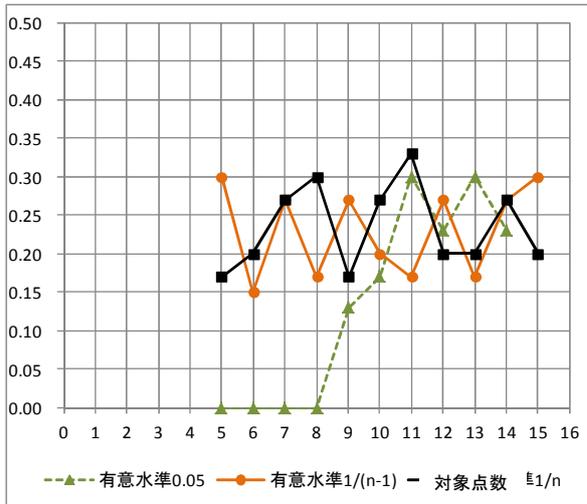
一般的に測量データのような場合は有意水準 5%を使うわけですが $\frac{1}{n}$ を有意水準とした場合と一般的な 5%に固定した場合の 2 通りについて異常値検出の信頼性を検証してみました。n とするか n-1 とするか迷いましたが差がないので n を選択しました。(母集団に異常点(正規分布から外れる点)がない場合の次表を参照)

母集団に異常点(正規分布から外れる点)がない場合

異常点が無い場合の準拠点選択後の数が5個～15 個(横軸)，縦軸が異常点として検出された割合です，これは分布の中に誤差の大きいデータが含まれた場合に検出されることがあるからです，異常点の無い散布図で母集団では異常ではないが 3σ 以上～4σ 未満の間にあるデータが含まれているときに，その周辺から遠くの精度の高いデータとの比較で異常値と判断されることを指しています。概ね 25%前後を想定していますので，平均で 27%では問題ではありません。

この結果から有意水準を 5% に比べて $\frac{1}{n}$ の場合が安定した良い結果が得られている事がわかります, この平均値は 23% で, 最大値は 33% で, 最小値は 17% とバラツキも大きくなく安定した結果が得られています。

グラフから有意水準を 0.05 に固定すると 10 点以上でない安定した結果が得られないことがわかります。



有意水準値の解析

有意水準→ 対象点数	0.05	1/(n-1)	1/n
5	0.00	0.30	0.17
6	0.00	0.15	0.20
7	0.00	0.27	0.27
8	0.00	0.17	0.30
9	0.13	0.27	0.17
10	0.17	0.20	0.27
11	0.30	0.17	0.33
12	0.23	0.27	0.20
13	0.30	0.17	0.20
14	0.23	0.27	0.27
15		0.30	0.20
平均	0.14	0.23	0.23

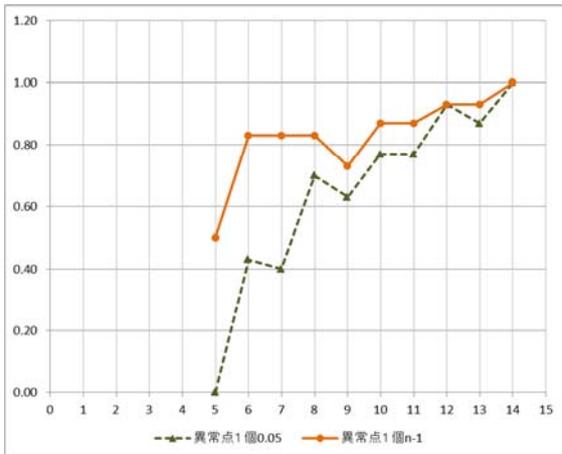
グラフから有意水準を 0.05 に固定すると準拠点は 10 点以上でない安定した結果が得られないことがわかります。

又, 有意水準を n としても $\frac{1}{n}$ としても差がないことがわかります。

母集団に異常点(正規分布から外れる点)がない場合 + 異常点を1点加えた場合

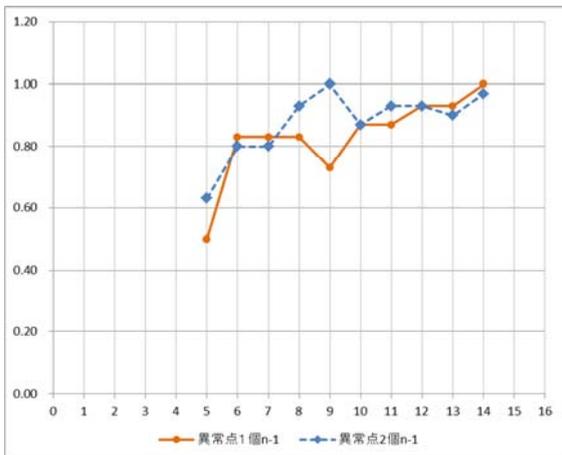
下のグラフは母集団から抜きとったデータに正規分布からは外れるデータを1個混ぜてt検定を実施したときに正規分布から外れるデータが異常とされる割合を示しています。

グラフと表は有意水準 0.05 の結果を示します, このグラフから有意水準を $\frac{1}{n}$ の場合が6個以上で安定した良い結果が得られている事が解ります, したがって準拠点数 6 個以上であれば 87% ですから準拠点選択結果の数が6個以上あれば高い信頼性は得られていると言えます。



準拠点選択後の数 (想定数)	異常点1個, 有意水準0.05	異常点1個, 有意水準(1/n-1)
5	0.00	0.5
6	0.43	0.83
7	0.40	0.83
8	0.70	0.83
9	0.63	0.73
10	0.77	0.87
11	0.77	0.87
12	0.93	0.93
13	0.87	0.93
14	1.00	1
15		
平均		0.87

母集団に異常点(正規分布から外れる点)がない場合 + 異常点を2点加えた場合



準拠点選択後の数 (想定数)	異常点1個, 有意水準(1/n-1)	異常点2個, 有意水準(1/n-1)
5	0.5	0.63
6	0.83	0.8
7	0.83	0.8
8	0.83	0.93
9	0.73	1
10	0.87	0.87
11	0.87	0.93
12	0.93	0.93
13	0.93	0.9
14	1	0.97
15		
平均	0.87	0.90

点数5を除く

念のために母集団から抜きとったデータに正規分布に馴染まないデータを2個含ませて、2個が検出される割合です。検出割合は90%でこのグラフから準拠点選択後の数が6個以上であれば信頼できることが解ります。

6個と11個の差は境界(筆界)復元においては10個以下で準拠点選択をするケースが多いので6~10個は影響のある数値なのです。

t 検定のまとめ

境界(筆界)復元計算に於ける準拠点選択時に使う有意水準は0.05(5%)ではなく $\frac{1}{n}$ を使うとデータ数に影響されない準拠点選択ができることが判りました。
(この結果は2016年10月11日以前のプログラムには使われていません)

χ^2 (カイジジョウ)検定

χ^2 (カイジジョウ)検定

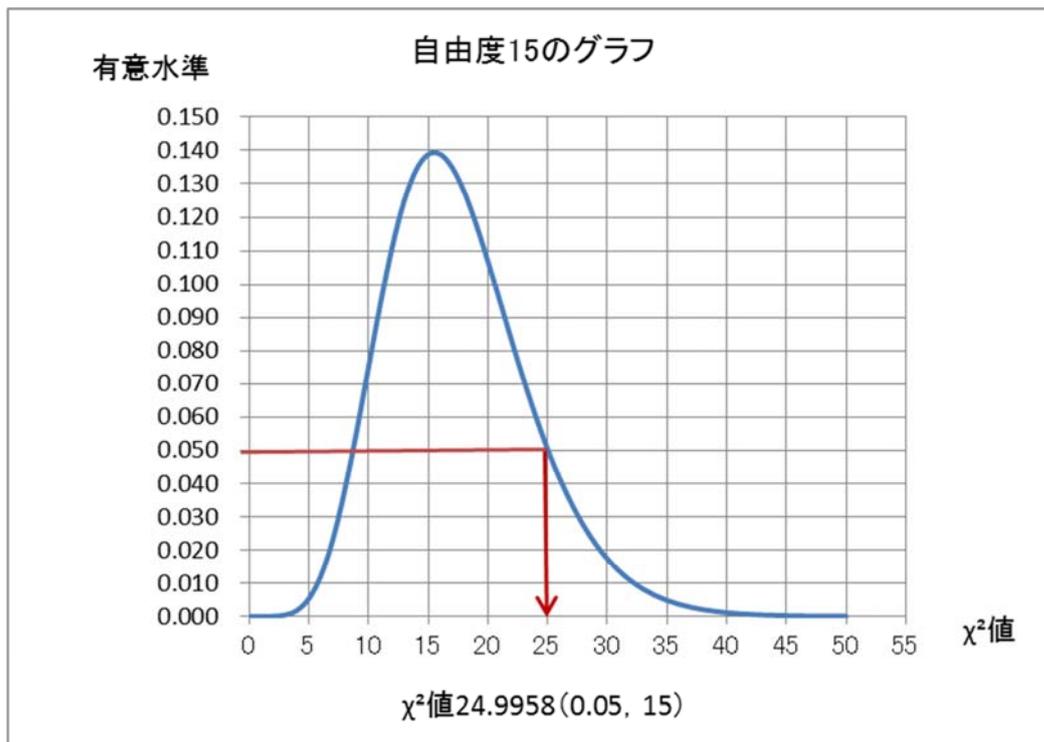
t検定は「観測値 $x_1, x_2, x_3, \dots, x_n$ の中で x_k が他の値に比べて大きく離れている、つまり誤差が大きいようだとすると x_k は異常値かどうかを検定する方法です」ですが、 χ^2 (カイジジョウ)検定は n 個の観測値を数個の小区間(クラス・階級)に分割して、その何番目かの小区間に属する数値の数(観測度数)が期待度数に対して大きいか小さいかで判定する方法です。簡単に言えば、正規分布に近似しているか離れているかの検定です。

χ^2 期待値 = $\chi^2_{\text{有意水準, 自由度}}$ で χ^2 値の関数から求められます, 表計算ソフト(Excel 等)で簡単に求められます, Excelの関数は $CHIINV_{\text{有意水準, 自由度}}$ です。自由度は小区間(クラス・階級)の数 $-r(r=2) - 1$ です。

下表が有意水準 0.05, 自由度 15 の χ^2 (カイジジョウ)値を求めたものです。実際は Excel の関数で計算しますので表を使って求めることはありません。

χ^2 (有意水準, 自由度)

有意水準	0.05	0.025	0.01	0.005
自由度				
1	3.8415	5.0239	6.6349	7.8794
2	5.9915	7.3778	9.2103	10.5966
3	7.8147	9.3484	11.3449	12.8382
4	9.4877	11.1433	13.2767	14.8603
5	11.0705	12.8325	15.0863	16.7496
6	12.5916	14.4494	16.8119	18.5476
7	14.0671	16.0128	18.4753	20.2777
8	15.5073	17.5345	20.0902	21.9550
9	16.9190	19.0228	21.6660	23.5894
10	18.3070	20.4832	23.2093	25.1882
11	19.6751	21.9200	24.7250	26.7568
12	21.0261	23.3367	26.2170	28.2995
13	22.3620	24.7356	27.6882	29.8195
14	23.6848	26.1189	29.1412	31.3193
15	24.9958	27.4884	30.5779	32.8013
16	26.2962	28.8454	31.9999	34.2672
17	27.5871	30.1910	33.4087	35.7185
18	28.8693	31.5264	34.8053	37.1565
19	30.1435	32.8523	36.1909	38.5823
20	31.4104	34.1696	37.5662	39.9968



有意水準は事項「**有意水準を幾つにするか**」を参照して下さい、ここでは 0.05 (5%) を使っています。

χ²(カイジジョウ)正規分布適合度検定の方法

χ²(カイジジョウ)観測値は次の手順で計算します。

データの平均値 $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ (x_i は最小二乗法座標変換で求めていますので 0 です)

分散 $\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$

分散から階級を決めます。

階級 $s = \sqrt{\sigma^2}$ (実際の計算では検定精度を高めるため階級を $s = \frac{\sqrt{\sigma^2}}{2}$ としています)

階級は概ね標準偏差の -4.75±0.25, -4.25±0.25, -3.75±0.25, -3.25±0.25, -2.75±0.25, -2.25±0.25, -1.75±0.25, -1.25±0.25, -0.75±0.25, -0.25±0.25, 0.25±0.25, 0.25±0.25, 0.75±0.25, 1.25±0.25, 1.75±0.25, 2.25±0.25, 2.75±0.25, 3.25±0.25, 3.75±0.25,

4.25±0.25, 4.75±0.25, 20 区分程度です。

各階級の正規分布の確率を正規分布表から求めます, 0.25±0.25s は 0~0.5 標準偏差の確率ですから 0.1915 (19.15%) です。

期待度数 = 各階級の確率×データ数

観測度数 = データの中から, 0.25±0.25s は 0~<0.5 に該当する実際の度数を調べます。(ヒストグラムの作成方法と同じです)

各階級の観測 χ^2 値を計算し増す。

$$\chi^2_i = \frac{(\text{観測度数}_i - \text{期待度数}_i)^2}{\text{期待度数}_i}$$

$$\chi^2(\text{観測値}) = \sum_{i=1}^n \chi^2_i \text{ の総和}$$

χ^2 二乗検定はデータ数が 31 以上の場合に使われる方法とされています, 方法は χ^2 期待値 (χ^2 理論値) と χ^2 観測値 (χ^2 実測値) を求めて χ^2 期待値 (χ^2 理論値) > χ^2 観測値 (χ^2 実測値) であれば「有意差無し」としその一連のデータには正規分布に適合し, 異常なデータが無いとします, χ^2 期待値 (χ^2 理論値) < χ^2 観測値 (χ^2 実測値) であれば「有意差あり」としてその一連のデータは正規分布に適合しないとして最大値を異常なデータとして除くという方法です。

次のデータで計算して見ます。

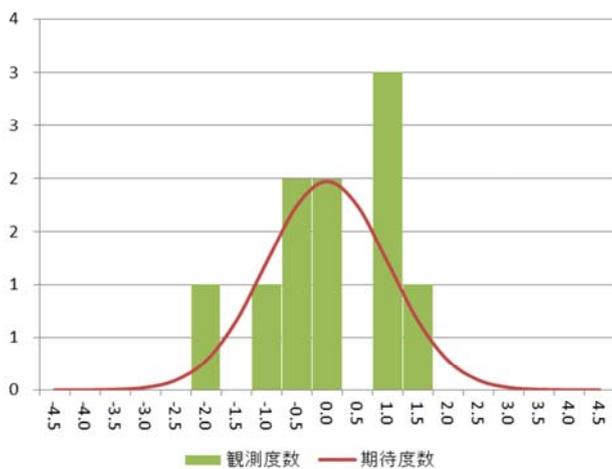
No	x
1	50.002
2	50.005
3	50.001
4	50.003
5	50.005
6	50.004
7	50.003
8	50.004
9	50.005
10	50.003
標準偏差	0.001
データ数	10
平均値	50.003

できればデータは降順に並べます。

n	階級中心	確率	観測度数	期待度数	χ^2 観測値
1	-4.5	0.0000	0	0.0001	0.0001
2	-4.0	0.0001	0	0.0008	0.0008
3	-3.5	0.0005	0	0.0049	0.0049
4	-3.0	0.0024	0	0.0240	0.0240
5	-2.5	0.0092	0	0.0924	0.0924
6	-2.0	0.0278	1	0.2783	1.8710
7	-1.5	0.0656	0	0.6559	0.6559
8	-1.0	0.1210	1	1.2098	0.0364
9	-0.5	0.1747	2	1.7467	0.0367
10	0.0	0.1974	2	1.9741	0.0003
11	0.5	0.1747	0	1.7467	1.7467
12	1.0	0.1210	3	1.2098	2.6492
13	1.5	0.0656	1	0.6559	0.1805
14	2.0	0.0278	0	0.2783	0.2783
15	2.5	0.0092	0	0.0924	0.0924
16	3.0	0.0024	0	0.0240	0.0240
17	3.5	0.0005	0	0.0049	0.0049
18	4.0	0.0001	0	0.0008	0.0008
19	4.5	0.0000	0	0.0001	0.0001
計		1.0000	10	χ^2 観測値	7.6995
				χ^2 期待値	11.0705
				有意水準	0.05
				自由度	5

χ^2 期待値 > χ^2 観測値 で有意差無し このデータは正規分布に適合している。

期待度数と観測度数の関係は下図の通り、以外と1 σ の所が甘い感じはするが全体としてはベルカーブの裾にデータがない・・・とこんな感じです。



一変量であればこれで χ^2 (カイジジョウ)検定が可能です、二変量では異なりますのでこのことを次に説明します。

二変量に於ける χ^2 (カイジジョウ)検定

ここでは、適正な準拠点選択をすることが前提にあります、図面值(期待値)と実測値(観測値)は同一座標系でないことがほとんどです、そのために図面值を実測値に重ねる作業が当然必要になります。

- ① 重心の移動(x軸, y軸への平行移動をおこないます)
- ② 回転(実測の方位に対して図面值の方位が一致しませんので重心を中心に回転させます)
- ③ 図面值と実測値で使う距離スケールが違うため伸縮がありますので伸縮させます。
- ④ 図面に縦とか横の伸び, 歪みがあれば除きます。
- ⑤ 実測点の一部に伸縮があれば修正します。

概ね①～⑤の点検をした上で図面值と実測値の座標差 Δx について χ^2 二乗検定をおこないます。

⑥ この場合, $\Delta x \cdot \Delta y$ の分布は原則, 楕円の分布ですから x 軸座標を 0 度として 0 度から 180° まで全週に渡り χ^2 二乗検定をします, つまり1度単位なら 179 回, 10 度単位なら 17 回ほど必要になります。全週に渡って異常値の無いデータ群が 2 回連続で続いたときに 1 回目のデータ群を準拠点として採用します。

このことを前提に χ^2 (カイジジョウ) 検定をおこなってみます, データは次表のとおりです。(これは t 検定のデータと同じです, 結果を比較しやすくするために同じデータを使用しています)

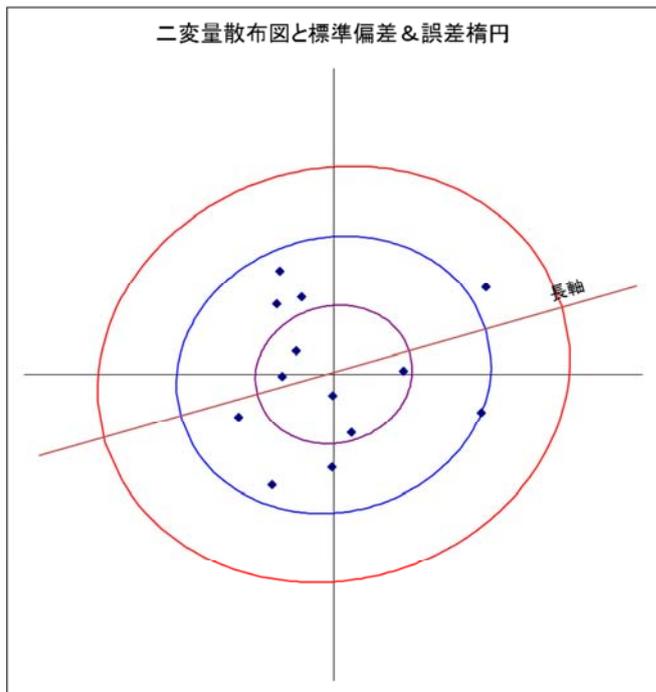
上記手順のうち①～③⑥まで実施, 座標変換はヘルマート変換を使用しています, ④⑤は実験データのため不要ですから実施していません。

対象データ

s148
s179
s209
s9
s178
s259
s114
s102
s85
s80
s60
s240
s291

分布の状態

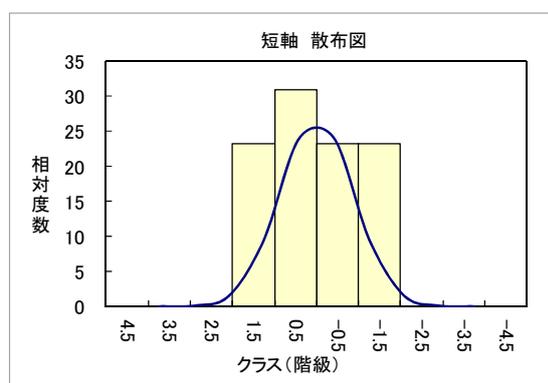
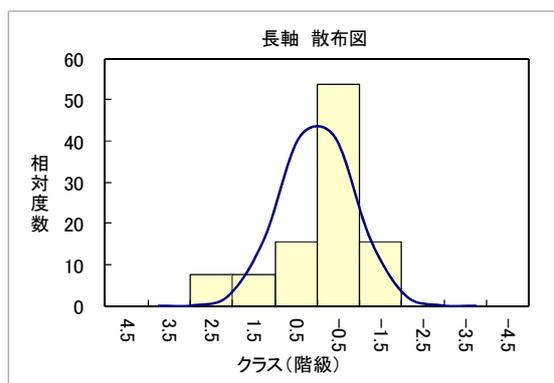
⑥ の分布の状態が下図の散布図のように楕円になっています。内側から 1σ 楕円, 2σ 楕円, 3σ 楕円となります。



(散布図データ:長軸標準偏差 0.629, 短軸標準偏差 0.548, 二変量標準偏差 0.590, 長軸角度 $73^{\circ}23'49''$)

下図が長軸のヒストグラム(棒グラフ)と正規分布曲線(ベルカーブ)です, 上図の散布図を南西方向から北東方向に向かって見た ($114^{\circ}+90^{\circ}=204^{\circ}$) のヒストグラムです。

短軸は南東方向から北西方向に向かって見た (114°) ヒストグラムになります。



次に $0^{\circ} \sim 170^{\circ}$ までの 10° 単位の χ^2 二乗検定値の計算結果を表にしました。
 χ^2 期待値 (χ^2 理論値) < χ^2 観測値 (χ^2 実測値) であれば「有意差あり」です。

χ^2 値計算 階級中心	0 度				
	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	0	0.150	0.150	
1.75	0.0491	0	0.638	0.638	
1.25	0.0919	3	1.194	2.731	
0.75	0.1499	1	1.949	0.462	
0.25	0.1915	2	2.489	0.096	
-0.25	0.1915	2	2.489	0.096	
-0.75	0.1499	3	1.949	0.567	
-1.25	0.0919	1	1.194	0.032	
-1.75	0.0491	1	0.638	0.206	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	5.289	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 値計算		10 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	0	0.150	0.150	
1.75	0.0491	1	0.638	0.206	
1.25	0.0919	1	1.194	0.032	
0.75	0.1499	2	1.949	0.001	
0.25	0.1915	2	2.489	0.096	
-0.25	0.1915	3	2.489	0.105	
-0.75	0.1499	2	1.949	0.001	
-1.25	0.0919	1	1.194	0.032	
-1.75	0.0491	1	0.638	0.206	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	1.140	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 値計算		20 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	0	0.150	0.150	
1.75	0.0491	1	0.638	0.206	
1.25	0.0919	1	1.194	0.032	
0.75	0.1499	2	1.949	0.001	
0.25	0.1915	3	2.489	0.105	
-0.25	0.1915	2	2.489	0.096	
-0.75	0.1499	1	1.949	0.462	
-1.25	0.0919	2	1.194	0.544	
-1.75	0.0491	1	0.638	0.206	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	2.113	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 値計算		30 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	1	0.150	4.810	
1.75	0.0491	0	0.638	0.638	
1.25	0.0919	0	1.194	1.194	
0.75	0.1499	4	1.949	2.159	
0.25	0.1915	2	2.489	0.096	
-0.25	0.1915	2	2.489	0.096	
-0.75	0.1499	1	1.949	0.462	
-1.25	0.0919	2	1.194	0.544	
-1.75	0.0491	1	0.638	0.206	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	10.516	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 値計算		40 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	1	0.150	4.810	
1.75	0.0491	0	0.638	0.638	
1.25	0.0919	0	1.194	1.194	
0.75	0.1499	4	1.949	2.159	
0.25	0.1915	1	2.489	0.891	
-0.25	0.1915	4	2.489	0.918	
-0.75	0.1499	0	1.949	1.949	
-1.25	0.0919	2	1.194	0.544	
-1.75	0.0491	1	0.638	0.206	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	13.619	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 値計算		50 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	1	0.150	4.810	
1.75	0.0491	0	0.638	0.638	
1.25	0.0919	1	1.194	0.032	
0.75	0.1499	1	1.949	0.462	
0.25	0.1915	3	2.489	0.105	
-0.25	0.1915	3	2.489	0.105	
-0.75	0.1499	2	1.949	0.001	
-1.25	0.0919	1	1.194	0.032	
-1.75	0.0491	1	0.638	0.206	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	6.701	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 値計算		60 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	1	0.150	4.810	
1.75	0.0491	1	0.638	0.206	
1.25	0.0919	0	1.194	1.194	
0.75	0.1499	1	1.949	0.462	
0.25	0.1915	2	2.489	0.096	
-0.25	0.1915	4	2.489	0.918	
-0.75	0.1499	2	1.949	0.001	
-1.25	0.0919	1	1.194	0.032	
-1.75	0.0491	1	0.638	0.206	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	8.236	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

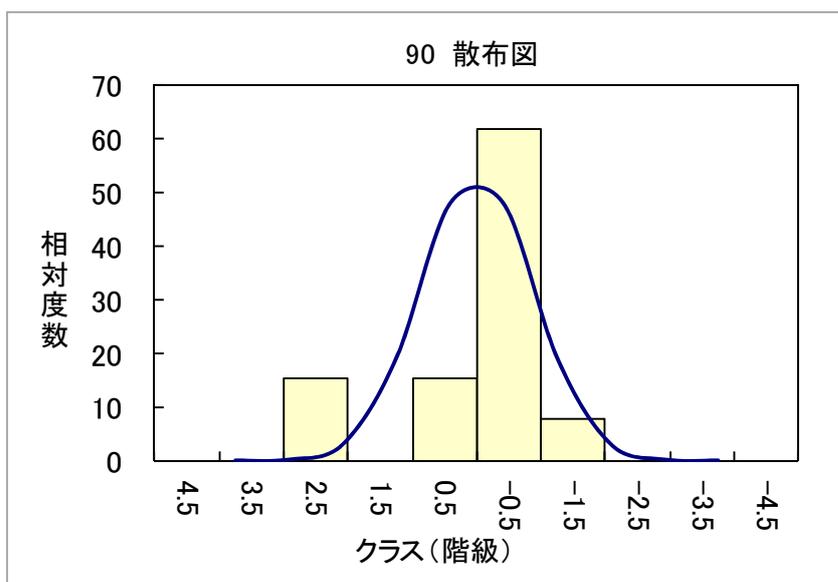
χ^2 値計算		70 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	1	0.150	4.810	
1.75	0.0491	1	0.638	0.206	
1.25	0.0919	0	1.194	1.194	
0.75	0.1499	1	1.949	0.462	
0.25	0.1915	0	2.489	2.489	
-0.25	0.1915	7	2.489	8.177	
-0.75	0.1499	1	1.949	0.462	
-1.25	0.0919	2	1.194	0.544	
-1.75	0.0491	0	0.638	0.638	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	19.292	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 値計算		80 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	1	0.150	4.810	
1.75	0.0491	1	0.638	0.206	
1.25	0.0919	0	1.194	1.194	
0.75	0.1499	1	1.949	0.462	
0.25	0.1915	1	2.489	0.891	
-0.25	0.1915	5	2.489	2.534	
-0.75	0.1499	2	1.949	0.001	
-1.25	0.0919	2	1.194	0.544	
-1.75	0.0491	0	0.638	0.638	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	11.591	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 値計算 階級中心	90 度				
	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	2	0.150	22.790	
1.75	0.0491	0	0.638	0.638	
1.25	0.0919	0	1.194	1.194	
0.75	0.1499	1	1.949	0.462	
0.25	0.1915	1	2.489	0.891	
-0.25	0.1915	3	2.489	0.105	
-0.75	0.1499	5	1.949	4.778	
-1.25	0.0919	1	1.194	0.032	
-1.75	0.0491	0	0.638	0.638	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	31.838	有意差あり
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 期待値(χ^2 理論値)23.685 < χ^2 観測値(χ^2 実測値)31.838 で「有意差あり」,このときのヒストグラムが下図です。

2.5 σ の位置の点が異常点と判断されています,この点は s291 と s102 です。(計算表からベクトル値の一番目と2番目・計算表の掲載はありません)



χ^2 値計算		100 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	1	0.150	4.810	
1.75	0.0491	1	0.638	0.206	
1.25	0.0919	0	1.194	1.194	
0.75	0.1499	1	1.949	0.462	
0.25	0.1915	3	2.489	0.105	
-0.25	0.1915	0	2.489	2.489	
-0.75	0.1499	6	1.949	8.423	
-1.25	0.0919	1	1.194	0.032	
-1.75	0.0491	0	0.638	0.638	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	18.669	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 値計算		110 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	1	0.150	4.810	
1.75	0.0491	1	0.638	0.206	
1.25	0.0919	0	1.194	1.194	
0.75	0.1499	2	1.949	0.001	
0.25	0.1915	2	2.489	0.096	
-0.25	0.1915	1	2.489	0.891	
-0.75	0.1499	3	1.949	0.567	
-1.25	0.0919	3	1.194	2.731	
-1.75	0.0491	0	0.638	0.638	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	11.446	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 値計算		120 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	1	0.150	4.810	
1.75	0.0491	0	0.638	0.638	
1.25	0.0919	1	1.194	0.032	
0.75	0.1499	3	1.949	0.567	
0.25	0.1915	2	2.489	0.096	
-0.25	0.1915	0	2.489	2.489	
-0.75	0.1499	4	1.949	2.159	
-1.25	0.0919	2	1.194	0.544	
-1.75	0.0491	0	0.638	0.638	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	12.284	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 値計算		130 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	0	0.150	0.150	
1.75	0.0491	1	0.638	0.206	
1.25	0.0919	0	1.194	1.194	
0.75	0.1499	4	1.949	2.159	
0.25	0.1915	2	2.489	0.096	
-0.25	0.1915	0	2.489	2.489	
-0.75	0.1499	3	1.949	0.567	
-1.25	0.0919	2	1.194	0.544	
-1.75	0.0491	1	0.638	0.206	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	7.923	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 値計算		140 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	0	0.150	0.150	
1.75	0.0491	1	0.638	0.206	
1.25	0.0919	0	1.194	1.194	
0.75	0.1499	4	1.949	2.159	
0.25	0.1915	2	2.489	0.096	
-0.25	0.1915	2	2.489	0.096	
-0.75	0.1499	1	1.949	0.462	
-1.25	0.0919	2	1.194	0.544	
-1.75	0.0491	1	0.638	0.206	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	5.425	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 値計算		150 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	0	0.150	0.150	
1.75	0.0491	1	0.638	0.206	
1.25	0.0919	1	1.194	0.032	
0.75	0.1499	2	1.949	0.001	
0.25	0.1915	3	2.489	0.105	
-0.25	0.1915	2	2.489	0.096	
-0.75	0.1499	1	1.949	0.462	
-1.25	0.0919	2	1.194	0.544	
-1.75	0.0491	1	0.638	0.206	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	2.113	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 値計算		160 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	0	0.150	0.150	
1.75	0.0491	0	0.638	0.638	
1.25	0.0919	3	1.194	2.731	
0.75	0.1499	1	1.949	0.462	
0.25	0.1915	3	2.489	0.105	
-0.25	0.1915	3	2.489	0.105	
-0.75	0.1499	0	1.949	1.949	
-1.25	0.0919	2	1.194	0.544	
-1.75	0.0491	1	0.638	0.206	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	7.201	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

χ^2 値計算		170 度			
階級中心	確率	観測度数	期待度数	χ^2 観測値	
4.25	0.0001	0	0.001	0.001	
3.75	0.0002	0	0.003	0.003	
3.25	0.0011	0	0.014	0.014	
2.75	0.0049	0	0.063	0.063	
2.25	0.0116	0	0.150	0.150	
1.75	0.0491	0	0.638	0.638	
1.25	0.0919	2	1.194	0.544	
0.75	0.1499	2	1.949	0.001	
0.25	0.1915	3	2.489	0.105	
-0.25	0.1915	2	2.489	0.096	
-0.75	0.1499	1	1.949	0.462	
-1.25	0.0919	2	1.194	0.544	
-1.75	0.0491	1	0.638	0.206	
-2.25	0.0116	0	0.150	0.150	
-2.75	0.0049	0	0.063	0.063	
-3.25	0.0011	0	0.014	0.014	
-3.75	0.0002	0	0.003	0.003	
-4.25	0.0001	0	0.001	0.001	
		13	χ^2 観測値	3.057	有意差無し
			χ^2 期待値	23.685	
			有意水準	0.050	
			自由度	14	

90度の χ^2 二乗検定の判定結果からs291を除いて、同じ操作をおこないます。誌面の関係で省略しますが 2回目は0° ~180° まで18とおりで「有意差無し」、3回目も18とおりで「有意差無し」でした、このデータ群の異常値はs291 となり、対象データからこの点を除いた点の組合せ(12点)が準拠点となります。

有意水準は幾つが適切か

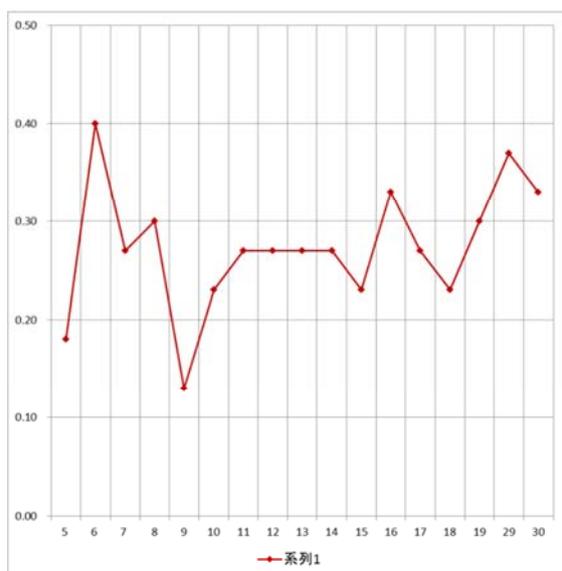
t検定で異常値無しで選択される率が23%なので表の27%を23%まで下げる必要があるのか、選択される率をコントロールする方法は様々ですがここでは短絡的に有意水準を0.045にすれば23%まで下がるはずですが0.05と0.045の差0.005は準拠点数30点で0.15点に相当するので0.05のままとします。

有意水準は5%(4.5%~5%)前後の値であればほぼt検定と差の無い検定結果が得られています、同じデータをt検定、 χ^2 (カイジジョウ)検定との両方で確認する場合も出てきますので両検定での結果に差が少ないことは求められます。

母集団に異常点(正規分布から外れる点)がない場合

異常点の無い散布図で準拠点選択後の数が5個~30個(横軸)、縦軸が異常点として検出された割合です、これは分布の中に誤差の大きいデータが含まれた場合に検出されることがあるからです、母集団では異常ではないが3 σ 以上~4 σ 未満の間にあるデータが含まれているときに、その周辺から遠くの精度の高いデータとの比較で異常値と判断されることがあります。

概ね25%前後を想定していますので、平均で27%では問題ではありません、t検定でも同じです。



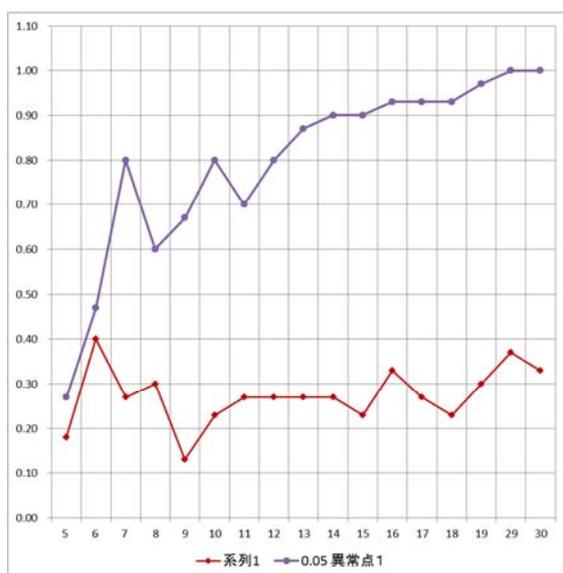
異常点無	
有意水準	0.05
点数	
5	0.18
6	0.40
7	0.27
8	0.30
9	0.13
10	0.23
11	0.27
12	0.27
13	0.27
14	0.27
15	0.23
16	0.33
17	0.27
18	0.23
19	0.30
29	0.37
30	0.33
平均	0.27
平均	

母集団に異常点(正規分布から外れる点)がない場合+異常点を1点加えた場合

下図、下表が異常点1個(s291)を加えた場合の検証結果です、準拠点選択後の点数が14点以上になれば90%異常と安定し検出率が得られ、29点以上で100%になります。

正規分布にどれだけ適合しているかの判断になるため点数が多いと良好な結果が得られます、実用上は87%の13点以上でも使えそうです。

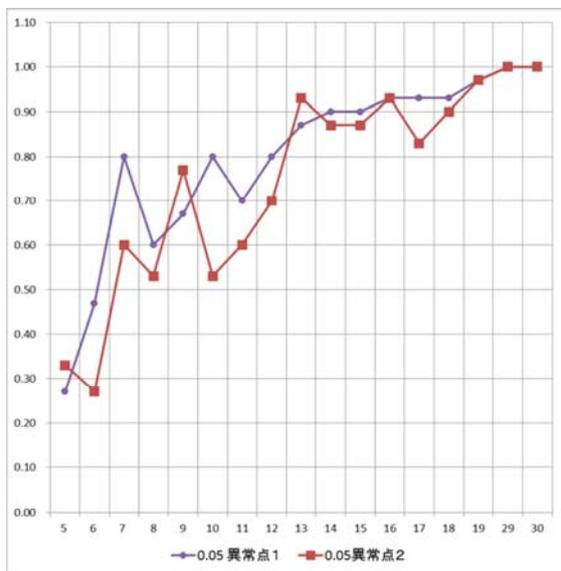
χ^2 (カイジジョウ)検定の解説の多くは31点以上あれば χ^2 (カイジジョウ)検定が適しており、30点以下ではt検定が適していると解説されていますが13点あたりから30点まではどちらでも差がないようです、但しこれ境界(筆界)復元での二変量の検定の場合です。



	異常点無	異常点1
有意水準	0.05	0.05
点数		
5	0.18	0.27
6	0.40	0.47
7	0.27	0.80
8	0.30	0.60
9	0.13	0.67
10	0.23	0.80
11	0.27	0.70
12	0.27	0.80
13	0.27	0.87
14	0.27	0.90
15	0.23	0.90
16	0.33	0.93
17	0.27	0.93
18	0.23	0.93
19	0.30	0.97
29	0.37	1.00
30	0.33	1.00
平均	0.27	0.77
平均	0.92	点数13以上

母集団に異常点(正規分布から外れる点)がない場合+異常点を2点加えた場合

異常点を2個(s97, s291)加えた場合の検証結果です, 異常点が 1 個の場合と検出率, 状態は同じです。



有意水準 点数	異常点1 点 s291	異常点2 点 s291・ s97
	0.05	0.05
5	0.27	0.33
6	0.47	0.27
7	0.80	0.60
8	0.60	0.53
9	0.67	0.77
10	0.80	0.53
11	0.70	0.60
12	0.80	0.70
13	0.87	0.93
14	0.90	0.87
15	0.90	0.87
16	0.93	0.93
17	0.93	0.83
18	0.93	0.90
19	0.97	0.97
29	1.00	1.00
30	1.00	1.00
平均	0.77	0.71
	0.92	0.90

点数13 以上

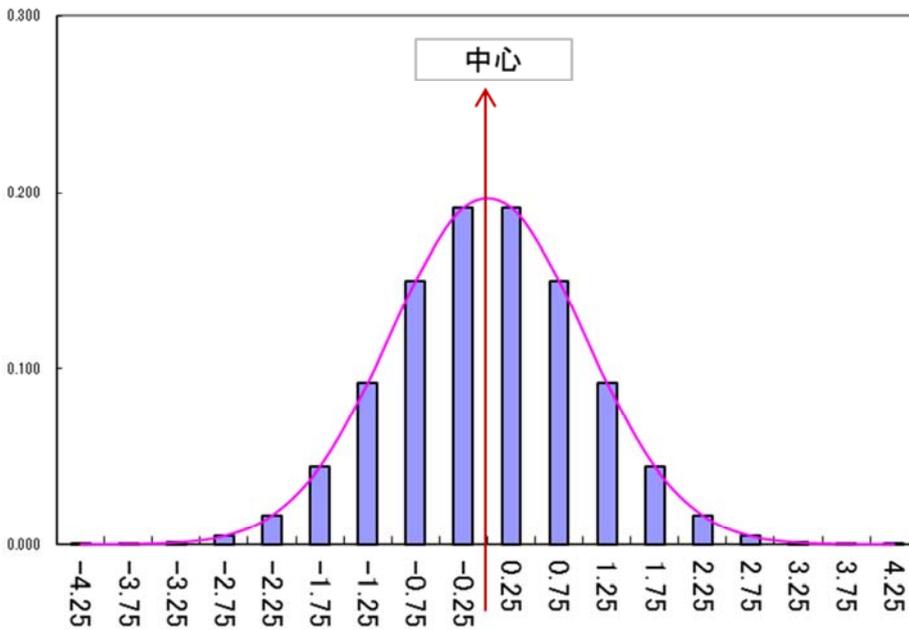
自由度を幾つにするか

χ^2 (カイジジョウ)検定の場合, 自由度を幾つにするのが適正か, ここが解らない点です。クラス(標準偏差の幅)を 1σ にするか 0.5σ にするかでも違って来ますが検定精度を上げるためには 0.5σ が適正です。

統計書の解説でもこの所は具体的に説明されていませんが, 0 の位置を $0.25 \pm 0.25\sigma$ と $-0.25 \pm 0.25\sigma$ の間としたとき, 下図では自由度 13(階級 16)が想定され χ^2 観測値は 0.58 です。

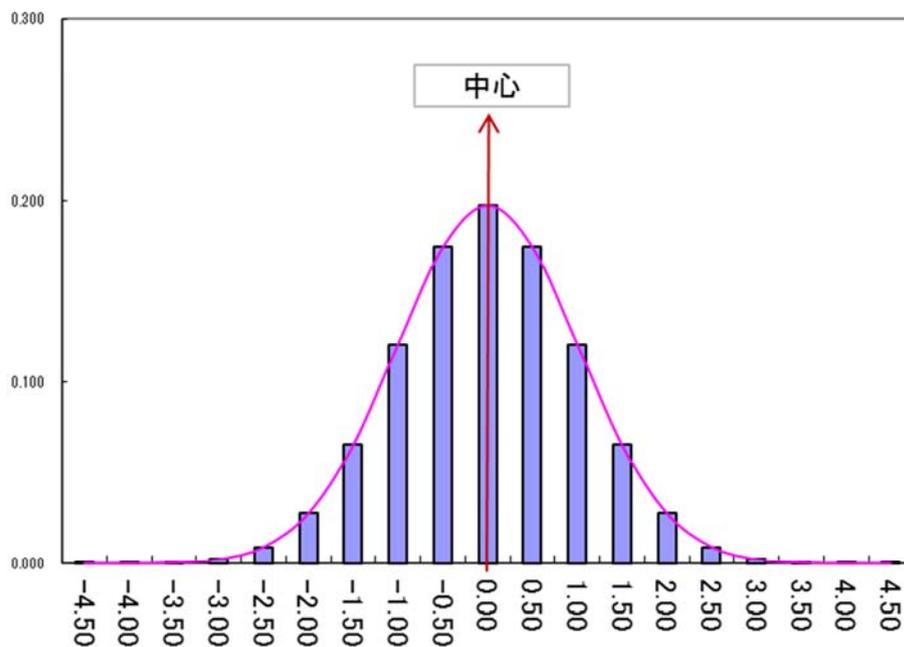
階級	級の中心	確率	観測度数	期待度数	観測値	観測数 10000
1	-4.25	0.0000		0.28	0.2829	0
2	-3.75	0.0002	2	2.01	0.0000	2
3	-3.25	0.0011	11	11.17	0.0027	11
4	-2.75	0.0049	49	48.60	0.0033	49
5	-2.25	0.0165	165	165.40	0.0010	165
6	-1.75	0.0441	441	440.57	0.0004	441
7	-1.25	0.0918	918	918.48	0.0003	918
8	-0.75	0.1499	1499	1498.82	0.0000	1499
9	-0.25	0.1915	1915	1914.62	0.0001	1915
10	0.25	0.1915	1915	1914.62	0.0001	1915
11	0.75	0.1499	1499	1498.82	0.0000	1499
12	1.25	0.0918	918	918.48	0.0003	918
13	1.75	0.0441	441	440.57	0.0004	441
14	2.25	0.0165	165	165.40	0.0010	165
15	2.75	0.0049	49	48.60	0.0033	49
16	3.25	0.0011	11	11.17	0.0027	11
17	3.75	0.0002	2	2.01	0.0000	2
18	4.25	0.0000		0.28	0.2829	0
計		1.0000	10000	χ^2 観測値	0.5813	10000
				χ^2 期待値	22.3620	

16
自由度13



0の位置を 0 ± 0.25 としたとき10000個で自由度14(階級17)が想定され χ^2 観測値は0.33です。

階級	級の中心	確率	観測度数	期待度数	級の観測値	観測数
						10000
1	-4.5	0.0000		0.10	0.0968	0
2	-4.0	0.0001	1	0.78	0.0637	1
3	-3.5	0.0005	5	4.89	0.0026	5
4	-3.0	0.0024	24	24.03	0.0000	24
5	-2.5	0.0092	92	92.45	0.0022	92
6	-2.0	0.0278	278	278.35	0.0004	278
7	-1.5	0.0656	656	655.91	0.0000	656
8	-1.0	0.1210	1210	1209.77	0.0000	1210
9	-0.5	0.1747	1747	1746.66	0.0001	1747
10	0.0	0.1974	1974	1974.13	0.0000	1974
11	0.5	0.1747	1747	1746.66	0.0001	1747
12	1.0	0.1210	1210	1209.77	0.0000	1210
13	1.5	0.0656	656	655.91	0.0000	656
14	2.0	0.0278	278	278.35	0.0004	278
15	2.5	0.0092	92	92.45	0.0022	92
16	3.0	0.0024	24	24.03	0.0000	24
17	3.5	0.0005	5	4.89	0.0026	5
18	4.0	0.0001	1	0.78	0.0637	1
19	4.5	0.0000		0.0968	0.0968	0
計		1.0000	10000	χ^2 観測値	0.3317	10000
				χ^2 期待値	23.6848	



χ^2 期待値に対してほとんど影響のない値です、0 の位置をどちらにしてもよいといえます。
 1σ 単位のヒストグラムを作成する関係で 0 の位置を 0.25 ± 0.25 と -0.25 ± 0.25 としています。
 さらに、自由度も 13 と 14 に変わりますがこちらも影響がないようです。

次の表は有意水準と自由度の関係を調べたものです、有意水準 0.05 であれば自由度は 14 程度が適切のようです。

自由度↓	有意水準				
	0.03	0.04	0.05	0.06	0.07
10	20	19	18	18	17
11	21	20	20	19	19
12	23	22	21	20	20
13	24	23	22	22	21
14	25	24	24	23	22
15	27	26	25	24	24
16	28	27	26	26	25
17	30	28	28	27	26
18	31	30	29	28	28

自由度を固定する旨の説明がありましたので固定できないか様々な条件で試しましたがデータの内容によって固定は無理ではないかと考えていました。

そこで下表のとおり、2,105 個のデータから自由度が幾つで決定されているかを調べると 14 (クラス 17) が 75%, 15 (クラス 18) が 10%, 16 (クラス 19) が 15%との結果が出ています。

その結果、自由度は 14~16 の間で自由に判断するのが適切といえます。

クラス	自由度	個数	割合
17	14	1580	0.75
18	15	219	0.10
19	16	306	0.15
		2105	

χ²(カイジジョウ)検定のまとめ

境界(筆界)復元計算における、準拠点選択、異常な値がデータ中に含まれているかいないかの判断に χ²(カイジジョウ)検定を使う場合、自由度は 14~16 の中から適切な値を使う、有意水準は 4.5~5%で適正と考えられます。

準拠点選択のまとめ

境界(筆界)復元での準拠点選択に χ²(カイジジョウ)検定、t検定を使っていますが厳密には統計書にあるがままの使い方ではありません。

様々な試行錯誤の結果として、現在の使い方に至っております、両検定とも通常に検定される場合よりは少々厳しい検定結果が出るようになっていきます。その結果として精度(標準偏差)の高い境界(筆界)復元が可能となっているからです。

見方を変えれば χ²(カイジジョウ)検定、t検定とも応用が効いているということでしょう。

本稿に記載したデータは一部に過ぎません、要点をコンパクトにまとめたつもりです、それでも相当の頁数になっています、さらなる検証は後の世代に譲るとしてここまでにしておきます。

2017/05/13

土地家屋調査士・測量士 小野孝治