

2×2 の検定についての考察

2×2 の検定、2つのセル持と2群の独立した検定で、(有効、無効)、(治癒、未治癒)、(完成、未完成)など統計的解析に使われます。しかしながら検定数も多く、どれが適した検定なのか教科書には具体的に書かれていません、今回この問題について検討してみました。

1. 例題による各種検定の問題点

(1) 次に示すように1群100例ぐらいで、既治療と新治療を比較してみました。既治療では49.5%、新治療では63.3%の治癒率が出ました。これから新治療が既治療に対して有意に治癒率を上げたと言えるかどうかを検討してみましょう。

	治癒	未治癒	合計
既治療	50(49.5%)	51(50.5%)	101
新治療	69(63.3%)	40(36.70%)	109

統計解析は2×2の検定を使います。SASのPROC FREQの中にある統計解析で、結果が数種類も出てきます。

カイ2乗値が $P=0.0438$ 、連続性補正カイ2乗値 $P=0.0606$ 、Fisherの正確検定は $P=0.0514$ と有意差があるような無いようなP値となっています。又P値の差が0.0079開き、有意差について厳密に(規制当局みたいに有意差により、認可条件が変わる)言うときには、問題となります。

統計量	自由度	値	p 値
カイ2乗値	1	4.0644	0.0438
尤度比カイ2乗値	1	4.0747	0.0435
連続性補正カイ2乗値	1	3.5219	0.0606
Mantel-Haenszel のカイ2乗値	1	4.0451	0.0443

Fisher の正確検定

セル (1,1) 度数 (F)	50
左側 $Pr \leq F$	0.0302
右側 $Pr \geq F$	0.9845
表の確率 (P)	0.0147
両側 $Pr \leq P$	0.0514

標本サイズ = 210

(2) 1 群 50 例ではどうでしょう、次の結果の例で確かめると、統計量を見るとカイ 2 乗値が $P=0.0473$ 、連続性補正カイ 2 乗値 $P=0.0779$ 、Fisher の正確検定は $P=0.0769$ と最大で 0.0296 とひらき、統計量結果に対して、どれが正確なのか確信が持てなくなってきました。

	治癒	未治癒	合計
既治療	10 (20.0%)	40 (80.0%)	50
新治療	19 (38.0%)	31 (62.0%)	50

統計量	自由度	値	p 値
カイ 2 乗値	1	4.1667	0.0473
尤度比カイ 2 乗値	1	4.2012	0.0459
連続性補正カイ 2 乗値	1	3.3750	0.0779
Mantel-Haenszel のカイ 2 乗値	1	4.1250	0.0484

Fisher の正確検定

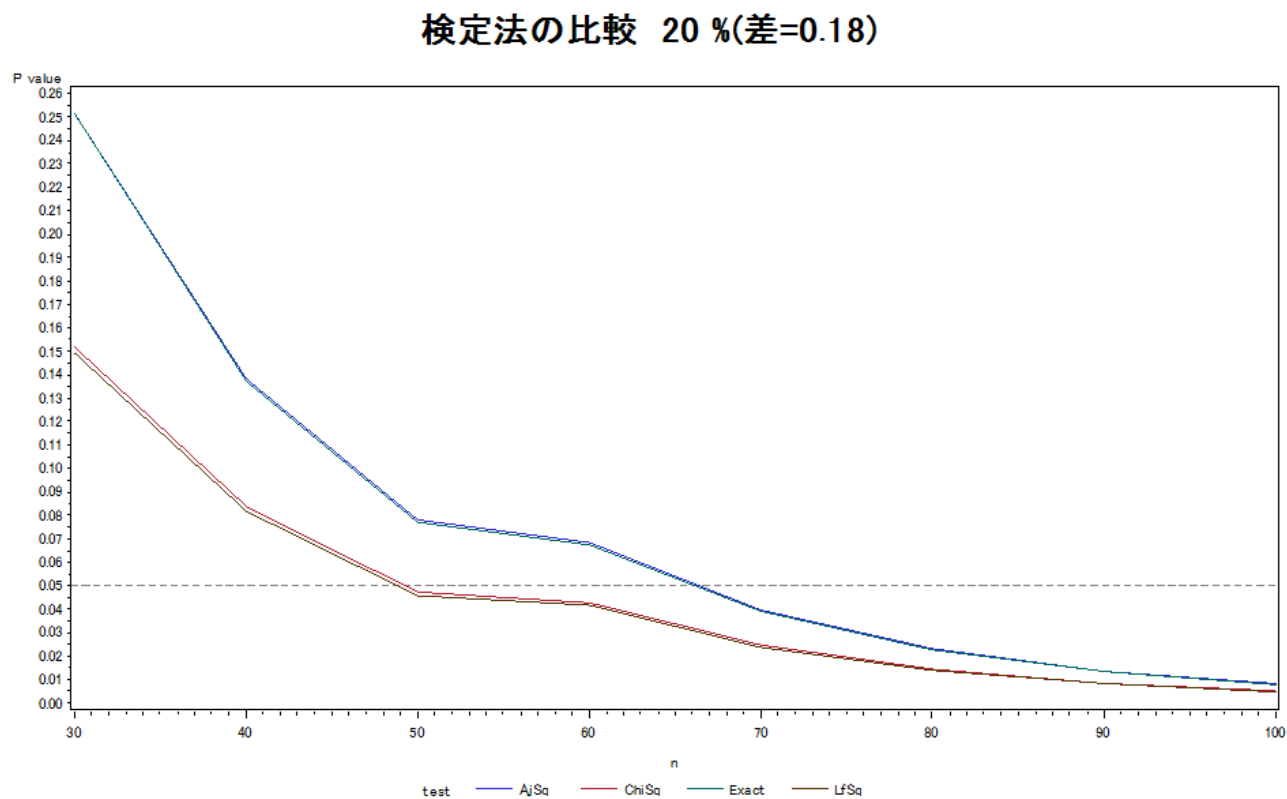
セル (1,1) 度数 (F)	25
左側 $\text{Pr} \leq F$	0.0385
右側 $\text{Pr} \geq F$	0.9867
表の確率 (P)	0.0252
両側 $\text{Pr} \leq P$	0.0769

標本サイズ = 100

2. 二つの結果を踏まえ、次のような試み行ってみました。

(1) コントロール群の治癒率を 10%~80% 10%刻み、対照群との治癒率の差を 18% (50~100 例で有意差が見られる)、1 群の例数を 30~100 例、10 例刻みで、2×2 の検定を行いました。

カイ 2 乗検定 (Pearson)、尤度比カイ 2 乗検定、連続性補正カイ 2 乗検定、Fisher の正確検定の P 値について例数の P 値の推移をグラフで描いてみました。



AJS q: 連続性補正カイ 2 乗値

Chisq : カイ 2 乗値

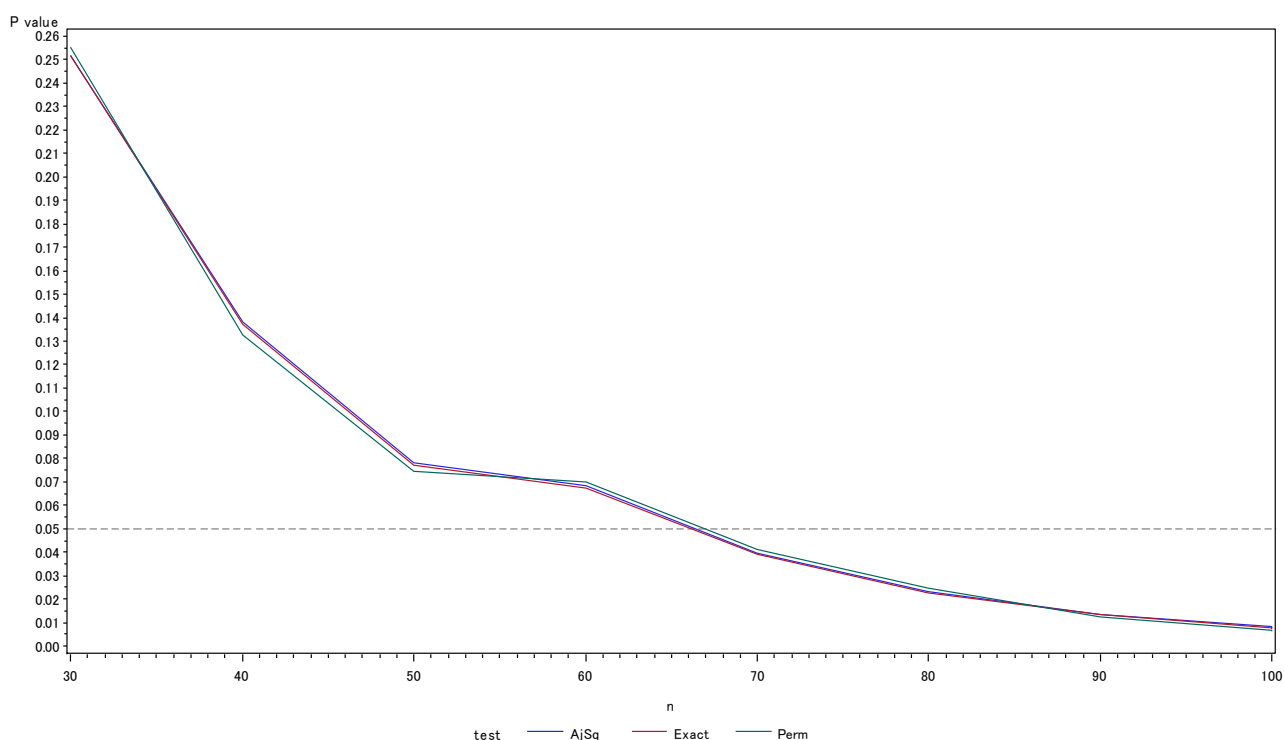
LiSq : 尤度比カイ 2 乗値

Exact : Fisher の正確検定

結果は連続性補正カイ 2 乗検定と Fisher の直接検定がほぼ同じ P 値をとり、カイ 2 乗と尤度比カイが同じ P 値をとりました。30 例~50 例では(カイ 2 乗検定、尤度比カイ 2 乗検定)と(連続性補正カイ 2 乗検定、Fisher の正確検定)との差は大きく、80 例以上になると、二つの検定グループの P 値は近づいてきます。

- (2) (1)のデータについて、Fisher の正確検定と連続性補正カイ検定の精度を見るため、PROC Multtest の非再帰抽出(permutation)法で Sampling を 1 万回で行い、母集団に近い P 値を求め、Fisher の正確検定の P 値、連続修正カイ検定の P 値について、例数毎に P 値を描いてみました。

検定法の比較 20 %(差=0.18)



AjSq: 連続修正カイ検定

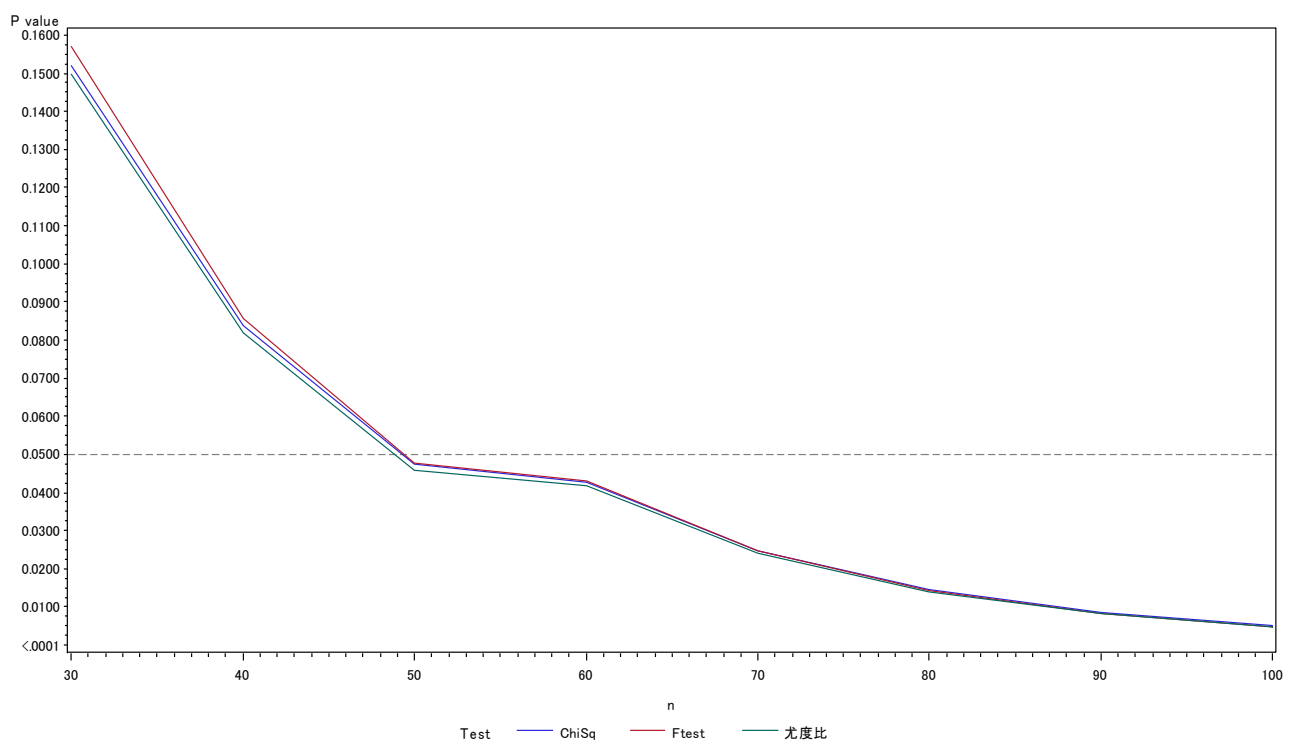
Exact: Fisher の正確検定

Parm: Permutation 検定

結果として Fisher の正確検定、連続修正カイ 2 乗検定、Permutation の結果はほぼ等しくなっています。Permutation で 1 万回以上するとほぼ母集団の検定に等しくなるため、Fisher の正確検定及び連続性補正カイ検定の結果はほぼ母集団に合った確率値となっています。

- (3) (1)のデータを1と0にして分散分析を実施、カイ検定、Logistic 回帰分析を行い、それぞれの p 値を求め、グラフで描いてみました。

検定法の比較 20 %(差=0.18)



Chisq:カイ検定

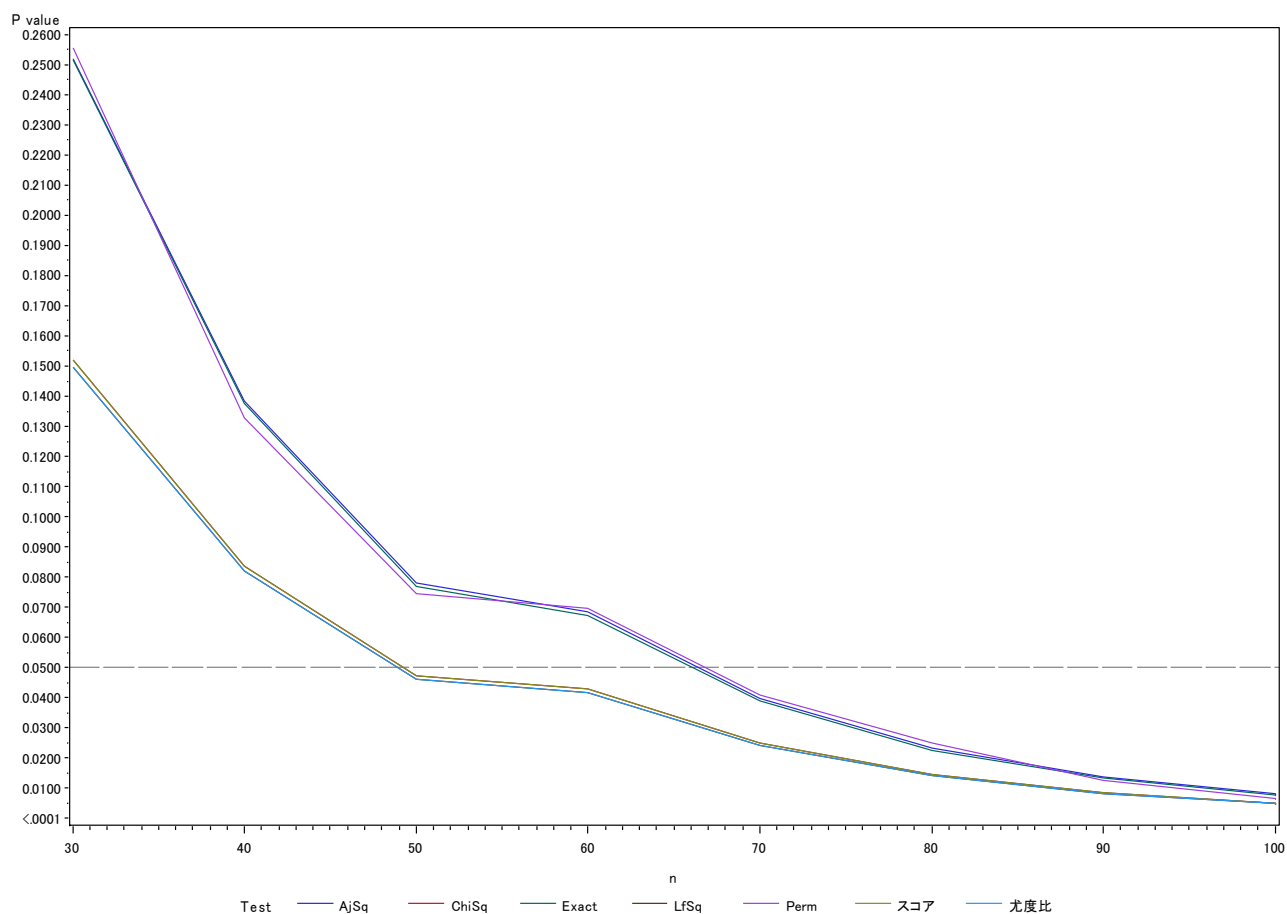
Ftest: 1と0 のして実施した分散分析

尤度比:尤度比カイ2 乗検定

3つの検定のP値は、ほぼ同じ動きをしています。

(4) (1) の デ ー タ で 各 種 検 定 を 実 施 し ま し た 。

検定法の比較 20 %(差=0.18)



AjSq:連続修正カイ検定

Chisq:カイ検定

Exact:Fisher の正確検定

LiSq :尤度比カイ2乗

Parm:Permutation

スコア:Logistic 回帰のスコア

尤度比:Logistic 回帰の尤度比

連続修正カイ検定、Fisher の正確検定、Permutation はほぼ同じ P 値を示しました。

それ以外の検定は、ほぼカイ検定と同じ P 値を示し、70 例以下ならば検定が甘くなってくることがわかりました。

結論として

2×2 の検定は 1 群 70 例以下では、正確性を期待するならば Fisher の正確検定を勧めます。

ただし、背景の要因を検討するときはカイ 2 乗検定、しかし少し割り引いて検討する必要があると思います。

また 100 例以上ではどのような検定を使ってもそれほど大差はありません。

参考までに表 1 に実際の P 値を示しておきます。

表 1. 検定法の比較（差を18%と仮定）

コントロール の%	例数	カイ 2 乗値 (Pearson)	尤度比 カイ 2	連続性 補正カイ	Mantel- Haenszel	Fisher の 正確検定	分散分析	Logistic 回帰			Permutation p-value
								尤度比	スコア	Wald	
10	30	0.0953	0.0903	0.1820	0.0981	0.1806	0.0984	0.0903	0.0953	0.1069	0.1847
10	40	0.0450	0.0416	0.0857	0.0463	0.0834	0.0456	0.0416	0.0450	0.0532	0.0835
10	50	0.0218	0.0197	0.0414	0.0225	0.0395	0.0217	0.0197	0.0218	0.0271	0.0411
10	60	0.0183	0.0166	0.0337	0.0188	0.0321	0.0182	0.0166	0.0183	0.0226	0.0315
10	70	0.0091	0.0081	0.0168	0.0094	0.0157	0.0089	0.0081	0.0091	0.0118	0.0156
10	80	0.0046	0.0040	0.0085	0.0047	0.0077	0.0044	0.0040	0.0046	0.0062	0.0061
10	90	0.0023	0.0020	0.0043	0.0024	0.0038	0.0022	0.0020	0.0023	0.0033	0.0034
10	100	0.0012	0.0010	0.0022	0.0012	0.0019	0.0011	0.0010	0.0012	0.0018	0.0020
20	30	0.1520	0.1497	0.2518	0.1555	0.2516	0.1572	0.1497	0.1520	0.1569	0.2554
20	40	0.0838	0.0819	0.1383	0.0857	0.1374	0.0858	0.0819	0.0838	0.0878	0.1328
20	50	0.0473	0.0459	0.0779	0.0484	0.0769	0.0479	0.0459	0.0473	0.0503	0.0743
20	60	0.0428	0.0416	0.0683	0.0437	0.0674	0.0432	0.0416	0.0428	0.0453	0.0697
20	70	0.0248	0.0239	0.0396	0.0253	0.0389	0.0248	0.0239	0.0248	0.0266	0.0409
20	80	0.0145	0.0139	0.0232	0.0148	0.0226	0.0143	0.0139	0.0145	0.0158	0.0248
20	90	0.0085	0.0081	0.0136	0.0087	0.0132	0.0083	0.0081	0.0085	0.0094	0.0124
20	100	0.0050	0.0048	0.0081	0.0051	0.0078	0.0049	0.0048	0.0050	0.0056	0.0066
30	30	0.1843	0.1830	0.2882	0.1880	0.2882	0.1904	0.1830	0.1843	0.1871	0.2869
30	40	0.1082	0.1070	0.1685	0.1104	0.1681	0.1109	0.1070	0.1082	0.1106	0.1723
30	50	0.0650	0.0641	0.1010	0.0664	0.1004	0.0661	0.0641	0.0650	0.0670	0.1038
30	60	0.0604	0.0597	0.0911	0.0615	0.0906	0.0612	0.0597	0.0604	0.0621	0.0881
30	70	0.0372	0.0366	0.0562	0.0379	0.0557	0.0374	0.0366	0.0372	0.0385	0.0570
30	80	0.0231	0.0226	0.0349	0.0235	0.0345	0.0231	0.0226	0.0231	0.0241	0.0340
30	90	0.0144	0.0141	0.0218	0.0147	0.0215	0.0143	0.0141	0.0144	0.0151	0.0220
30	100	0.0091	0.0088	0.0137	0.0092	0.0135	0.0089	0.0088	0.0091	0.0096	0.0115
40	30	0.1965	0.1954	0.3014	0.2002	0.3015	0.2029	0.1954	0.1965	0.1986	0.3025
40	40	0.1174	0.1165	0.1796	0.1197	0.1793	0.1204	0.1165	0.1174	0.1193	0.1799
40	50	0.0718	0.0710	0.1095	0.0732	0.1091	0.0731	0.0710	0.0718	0.0734	0.1116
40	60	0.0677	0.0671	0.1002	0.0689	0.0998	0.0687	0.0671	0.0677	0.0690	0.1008
40	70	0.0424	0.0419	0.0629	0.0432	0.0625	0.0428	0.0419	0.0424	0.0435	0.0597
40	80	0.0268	0.0264	0.0398	0.0273	0.0394	0.0268	0.0264	0.0268	0.0276	0.0394
40	90	0.0171	0.0168	0.0253	0.0174	0.0250	0.0169	0.0168	0.0170	0.0176	0.0253
40	100	0.0109	0.0107	0.0162	0.0111	0.0160	0.0107	0.0107	0.0109	0.0113	0.0167
50	30	0.1904	0.1892	0.2949	0.1942	0.2949	0.1967	0.1892	0.1904	0.1929	0.2877
50	40	0.1119	0.1108	0.1730	0.1142	0.1726	0.1147	0.1108	0.1119	0.1141	0.1732
50	50	0.0673	0.0664	0.1038	0.0687	0.1033	0.0684	0.0664	0.0673	0.0691	0.1032
50	60	0.0641	0.0634	0.0956	0.0652	0.0952	0.0650	0.0634	0.0641	0.0655	0.0956
50	70	0.0395	0.0390	0.0591	0.0402	0.0587	0.0398	0.0389	0.0395	0.0407	0.0571
50	80	0.0246	0.0241	0.0368	0.0250	0.0365	0.0245	0.0241	0.0246	0.0255	0.0367
50	90	0.0154	0.0150	0.0231	0.0157	0.0228	0.0152	0.0150	0.0154	0.0160	0.0238
50	100	0.0097	0.0094	0.0145	0.0098	0.0143	0.0095	0.0094	0.0097	0.0101	0.0161
60	30	0.1653	0.1634	0.2670	0.1688	0.2668	0.1708	0.1634	0.1652	0.1691	0.2738
60	40	0.0913	0.0897	0.1478	0.0934	0.1471	0.0935	0.0897	0.0913	0.0948	0.1540
60	50	0.0517	0.0504	0.0837	0.0528	0.0828	0.0524	0.0504	0.0517	0.0544	0.0812
60	60	0.0497	0.0487	0.0774	0.0507	0.0767	0.0503	0.0487	0.0497	0.0518	0.0748
60	70	0.0289	0.0282	0.0452	0.0295	0.0446	0.0290	0.0282	0.0289	0.0305	0.0496
60	80	0.0170	0.0164	0.0266	0.0173	0.0261	0.0168	0.0164	0.0169	0.0181	0.0269
60	90	0.0100	0.0096	0.0157	0.0102	0.0153	0.0098	0.0096	0.0100	0.0108	0.0154
60	100	0.0059	0.0057	0.0094	0.0061	0.0091	0.0058	0.0057	0.0059	0.0065	0.0086
70	30	0.1172	0.1134	0.2100	0.1203	0.2092	0.1212	0.1134	0.1172	0.1255	0.2095
70	40	0.0557	0.0528	0.1010	0.0573	0.0993	0.0568	0.0528	0.0557	0.0624	0.1006
70	50	0.0271	0.0252	0.0495	0.0279	0.0479	0.0272	0.0252	0.0271	0.0318	0.0505
70	60	0.0267	0.0252	0.0461	0.0273	0.0448	0.0267	0.0252	0.0267	0.0303	0.0470
70	70	0.0135	0.0124	0.0235	0.0138	0.0225	0.0133	0.0124	0.0134	0.0159	0.0234
70	80	0.0068	0.0062	0.0120	0.0070	0.0113	0.0066	0.0062	0.0068	0.0084	0.0118
70	90	0.0035	0.0031	0.0062	0.0036	0.0057	0.0033	0.0031	0.0035	0.0045	0.0065
70	100	0.0018	0.0015	0.0032	0.0018	0.0029	0.0017	0.0015	0.0018	0.0024	0.0034
80	30	0.0444	0.0352	0.1077	0.0462	0.1028	0.0452	0.0352	0.0444	0.0756	0.1020
80	40	0.0133	0.0087	0.0338	0.0138	0.0289	0.0129	0.0087	0.0133	0.0362	0.0279
80	50	0.0040	0.0021	0.0106	0.0042	0.0078	0.0037	0.0021	0.0040	0.0192	0.0065
80	60	0.0045	0.0029	0.0105	0.0046	0.0084	0.0042	0.0029	0.0045	0.0120	0.0090
80	70	0.0014	0.0008	0.0035	0.0015	0.0024	0.0013	0.0008	0.0014	0.0059	0.0016
80	80	0.0005	0.0002	0.0011	0.0005	0.0007	0.0004	0.0002	0.0005	0.0031	0.0009
80	90	0.0002	0.0001	0.0004	0.0002	0.0002	0.0001	<.0001	0.0001	0.0017	0.0000
80	100	0.0001	0.0000	0.0001	0.0001	0.0001	0.0000	<.0001	<.0001	0.0009	0.0000