ACTUARIAL RESEARCH CLEARING HOUSE
1994 VOL. 2

GRADUATION BY OPERATIONAL BOOTSTRAPPING

by William A. Bailey
Kemper National Insurance Companies
1 Kemper Drive
Long Grove, IL 60049

Abstract

Starting with the 3-d array

$$[X, \vartheta_X, N_X]_{X=55,\ldots,100},$$

where X=age, ϑ_X=number of deaths and N_X=exposed lives, a 3-d empirical distribution of

$$[X\cdot N_X, \log(\vartheta_X/N_X), N_X]_{X=55,\ldots,100}$$

is formed, by assigning equal probability $\frac{1}{46} = \frac{1}{100-55+1}$ to each triplet. Forty-six copies of this distribution are convoluted together using 3-d numerical convolutions. The resulting 3-d distribution, call it $f_{Y,Z,T}$, of [Y,Z,T] is transformed into a 2-d distribution of [R,S] where $R = \frac{Y}{T}$ and $S = \frac{Z}{T}$. For each relevant integral value of X, we choose $E[\ S \mid R \in [X-.5, X+.5]\ ]$ as the graduated log value which can be used as an exponent of 10 to obtain the graduated mortality rate corresponding to the original raw mortality rate $\frac{\vartheta_X}{N_X}$.

Section 1. Introduction

The topic is graduation of raw mortality rates. The approach is to use operational bootstrapping, which produces results which could only be obtained by regular bootstrapping if an infinite number of random numbers were generated.

The term "operational bootstrapping" was first introduced in Bailey (1992a). Univariate operational bootstrapping was described in Bailey (1992b), the Appendix of which contained a Unnugeco (Univariate Numerical Generalized Convolution) Algorithm. A Binugeco (Bivariate Numerical Generalized Convolution) Algorithm was described in an Appendix of Bailey (1993). We shall use a Trinugeco (Trivariate Numerical Generalized Convolution) Algorithm, which is a straightforward extension of this Binugeco Algorithm.

Section 2. The Procedure To Graduate By Operational Bootstrapping

Let $[X,\vartheta_X,N_X]_{X=55,\ldots,100}$ be 46 vectors of raw mortality data, where

X=age,

ϑ_X=number of deaths and

N_X=exposed lives.

Original Mortality Data In The Form Of An Empirical Distribution
Taken from TABLE 1 on page 313 of Ramsay (1991)

three d dist,lines=46

X mean and var= 77.500000 176.250000
X mean abs dev = 11.500000
X min and max= 55.000000 100.000000

Y mean and var= 214.108696 37389.836011
Y mean abs dev = 175.823251
Y min and max= 1.000000 581.000000

Z mean and var= 7922.326087 84934881.132798
Z mean abs dev = 7643.541588
Z min and max= 8.000000 29174.000000
cov xy = -491.945652
cov xz = -57176.445652
cov yz = 1329887.225425
cum prob = 1.0000000000000

index	xamount	yamount	zamount	prob	cum
0000)	55	1	84	2.173913e-02	0.021739
0001)	56	2	418	2.173913e-02	0.043478
0002)	57	10	1066	2.173913e-02	0.065217
0003)	58	21	2483	2.173913e-02	0.086957
0004)	59	35	3721	2.173913e-02	0.108696
0005)	60	62	5460	2.173913e-02	0.130435
0006)	61	50	6231	2.173913e-02	0.152174
0007)	62	55	8061	2.173913e-02	0.173913
0008)	63	88	9487	2.173913e-02	0.195652
0009)	64	132	10770	2.173913e-02	0.217391
0010)	65	267	24267	2.173913e-02	0.239130
0011)	66	300	26791	2.173913e-02	0.260870
0012)	67	432	29174	2.173913e-02	0.282609
0013)	68	491	28476	2.173913e-02	0.304348
0014)	69	422	25840	2.173913e-02	0.326087
0015)	70	475	23916	2.173913e-02	0.347826
0016)	71	413	21412	2.173913e-02	0.369565
0017)	72	480	20116	2.173913e-02	0.391304
0018)	73	537	18876	2.173913e-02	0.413043
0019)	74	566	17461	2.173913e-02	0.434783
0020)	75	581	15012	2.173913e-02	0.456522
0021)	76	464	11871	2.173913e-02	0.478261
0022)	77	461	10002	2.173913e-02	0.500000
0023)	78	433	8949	2.173913e-02	0.521739

0024)	79	515	7751	2.173913e-02	0.543478
0025)	80	374	6140	2.173913e-02	0.565217
0026)	81	348	4718	2.173913e-02	0.586957
0027)	82	304	3791	2.173913e-02	0.608696
0028)	83	249	2806	2.173913e-02	0.630435
0029)	84	167	2240	2.173913e-02	0.652174
0030)	85	192	1715	2.173913e-02	0.673913
0031)	86	171	1388	2.173913e-02	0.695652
0032)	87	126	898	2.173913e-02	0.717391
0033)	88	86	578	2.173913e-02	0.739130
0034)	89	97	510	2.173913e-02	0.760870
0035)	90	93	430	2.173913e-02	0.782609
0036)	91	75	362	2.173913e-02	0.804348
0037)	92	84	291	2.173913e-02	0.826087
0038)	93	31	232	2.173913e-02	0.847826
0039)	94	75	196	2.173913e-02	0.869565
0040)	95	29	147	2.173913e-02	0.891304
0041)	96	25	100	2.173913e-02	0.913043
0042)	97	20	161	2.173913e-02	0.934783
0043)	98	5	11	2.173913e-02	0.956522
0044)	99	3	10	2.173913e-02	0.978261
0045)	100	2	8	2.173913e-02	1.000000

Form a 3-dimensional empirical distribution

$$f_{U,V,W} = f_{X \cdot N_X, \log(\vartheta_X/N_X) \cdot N_X, N_X}$$

of the 3-dimensional random variable

$$(U,V,W) = (X \cdot N_X, \log(\vartheta_X/N_X) \cdot N_X, N_X)$$

where X takes on the values 55,56,...,100,

by assigning equal probability $\frac{1}{46} = \frac{1}{100-55+1}$ to each triplet.

$f_{U,V,W}$

```
three d dist,lines=46

X mean and var= 556803.826087 401103046179.621928
X mean abs dev = 533140.432892
X min and max= 800.000000 1954658.000000

Y mean and var= -13308.770658 275373573.115291
Y mean abs dev = 13498.689376
Y min and max= -53374.189851  -3.766649

Z mean and var= 7922.326087 84934881.132798
Z mean abs dev = 7643.541588
Z min and max=   8.000000 29174.000000
cov xy = -10273983068.349628
cov xz = 5821023501.143667
cov yz = -151440966.338228
cum prob = 1.0000000000000
```

index	xamount	yamount	zamount	prob	cum
0000)	4620	-162	84	2.173913e-02	0.021739
0001)	23408	-970	418	2.173913e-02	0.043478
0002)	60762	-2162	1066	2.173913e-02	0.065217
0003)	144014	-5147	2483	2.173913e-02	0.086957
0004)	219539	-7541	3721	2.173913e-02	0.108696
0005)	327600	-10619	5460	2.173913e-02	0.130435
0006)	380091	-13058	6231	2.173913e-02	0.152174
0007)	499782	-17460	8061	2.173913e-02	0.173913
0008)	597681	-19284	9487	2.173913e-02	0.195652
0009)	689280	-20588	10770	2.173913e-02	0.217391
0010)	1577355	-47527	24267	2.173913e-02	0.239130
0011)	1768206	-52266	26791	2.173913e-02	0.260870
0012)	1954658	-53374	29174	2.173913e-02	0.282609
0013)	1936368	-50215	28476	2.173913e-02	0.304348
0014)	1782960	-46176	25840	2.173913e-02	0.326087
0015)	1674120	-40705	23916	2.173913e-02	0.347826
0016)	1520252	-36715	21412	2.173913e-02	0.369565
0017)	1448352	-32634	20116	2.173913e-02	0.391304
0018)	1377948	-29181	18876	2.173913e-02	0.413043
0019)	1292114	-26004	17461	2.173913e-02	0.434783
0020)	1125900	-21201	15012	2.173913e-02	0.456522
0021)	902196	-16714	11871	2.173913e-02	0.478261
0022)	770154	-13367	10002	2.173913e-02	0.500000
0023)	698022	-11770	8949	2.173913e-02	0.521739
0024)	612329	-9127	7751	2.173913e-02	0.543478
0025)	491200	-7462	6140	2.173913e-02	0.565217
0026)	382158	-5342	4718	2.173913e-02	0.586957
0027)	310862	-4154	3791	2.173913e-02	0.608696
0028)	232898	-2952	2806	2.173913e-02	0.630435

0029)	188160	-2526	2240	2.173913e-02	0.652174
0030)	145775	-1631	1715	2.173913e-02	0.673913
0031)	119368	-1262	1388	2.173913e-02	0.695652
0032)	78126	-766	898	2.173913e-02	0.717391
0033)	50864	-478	578	2.173913e-02	0.739130
0034)	45390	-368	510	2.173913e-02	0.760870
0035)	38700	-286	430	2.173913e-02	0.782609
0036)	32942	-247	362	2.173913e-02	0.804348
0037)	26772	-157	291	2.173913e-02	0.826087
0038)	21576	-203	232	2.173913e-02	0.847826
0039)	18424	-82	196	2.173913e-02	0.869565
0040)	13965	-104	147	2.173913e-02	0.891304
0041)	9600	-60	100	2.173913e-02	0.913043
0042)	15617	-146	161	2.173913e-02	0.934783
0043)	1078	-4	11	2.173913e-02	0.956522
0044)	990	-5	10	2.173913e-02	0.978261
0045)	800	-5	8	2.173913e-02	1.000000

Regular bootstrapping for the mean vector E[U,V,W] would involve using random numbers to sample (called resampling) with replacement from the 3-d distribution

$$f_{U,V,W} = f_{X \cdot N_X,\, \log(\vartheta_X/N_X) \cdot N_X,\, N_X}$$

to obtain various resample means for resamples of size 46. In order to avoid having to use random numbers and in order to obtain more accurate estimates in the tails of the distributions, we choose to use operational bootstrapping.

In order to graduate the raw mortality data by operational bootstrapping we need to perform the following sequence of convolutions to obtain

$f^{+46}_{U,V,W}$; namely,

$$f^{+2}_{U,V,W} = f_{U,V,W} + f_{U,V,W}$$

where the "+" between two distributions is used to denote "convolute for sums";

$$f^{+4}_{U,V,W} = f^{+2}_{U,V,W} + f^{+2}_{U,V,W}$$

$$f^{+8}_{U,V,W} = f^{+4}_{U,V,W} + f^{+4}_{U,V,W}$$

$$f^{+16}_{U,V,W} = f^{+8}_{U,V,W} + f^{+8}_{U,V,W}$$

$$f^{+32}_{U,V,W} = f^{+16}_{U,V,W} + f^{+16}_{U,V,W}$$

$$f^{+40}_{U,V,W} = f^{+32}_{U,V,W} + f^{+8}_{U,V,W}$$

$$f^{+44}_{U,V,W} = f^{+40}_{U,V,W} + f^{+4}_{U,V,W}$$

$$f^{+46}_{U,V,W} = f^{+44}_{U,V,W} + f^{+2}_{U,V,W}.$$

These trivariate numerical convolutions can be performed according to the Trinugeco Algorithm, which is a straightforward generalization of the Binugeco Algorithm in Bailey (1993).

We can then transform the 3-d distribution $f^{+46}_{U,V,W}$, call it $f_{Y,Z,T}$, of (Y,Z,T) into the 2-d distribution

$$f_{R,S} \text{ where } R = \frac{Y}{T} \text{ and } S = \frac{Z}{T}.$$

For each relevant integral value of X, we choose

$$E[\ S \mid R \in (X-.5, X+.5]\ ]$$

as the graduated log value which can be used as an exponent of 10 to obtain the graduated mortality rate corresponding to the original raw mortality rate $\frac{\vartheta_X}{n_X}$.

The accuracy of the multivariate convolutions for sums can be considerably improved by using a technique described in the next section and described in Anderson (1984).

Section 3. A Numerical Technique From Anderson's Description of the Principal Component Method

We are interested in the numerical method described on pages 462-465 and exemplified on pages 465-468 in Anderson (1984), because it generates a linear transformation which can be applied to the n-dimensional vectors in an n-dimensional distribution to transform that distribution into one for which each of the 2-way covariances is zero.

The reason we are interested in having the 2-way covariances equal to zero is that the 2-way covariances in the distributions resulting from using the multivariate numerical convolution algorithms gradually tend to decrease toward zero as a significant number of convolutions are performed. By transforming the initial distribution (which we are interested in convoluting a number of times) so that the 2-way covariances are zero, the tendency for the 2-way covariances of the distributions to deteriorate is considerably reduced.

The "C"-language computer program called COCONUT (Command-and Convolute), mentioned above, includes among other commands two commands (anderf1 and anderf2) which accomplish the following:

(1) take an input distribution in what we refer to as space #1, generate the appropriate linear transformation, and apply the transformation to the input distribution in space #1 to obtain a distribution in what we shall refer to as space #2 with the same dimension as the input distribution but with each of the 2-way covariances equal to zero;

(2) take an input distribution in space #2 (i.e. the result of convoluting a series of iid distributions initially generated in (1)), obtain the inverse of the linear transformation originally obtained in (1), and apply that inverse

linear transformation to the input distribution in space #2 to obtain a distribution in space #1.

If we applied (1) to an initial distribution and then immediately applied (2) to the resulting distribution, we would reproduce the initial distribution. The only purpose of doing this would be to confirm that the tranformation and its inverse are operating properly.

Section 4. The Numerical Results

T, the orthonormal linear transformation from Space #1 into Space #2

$$T = \begin{bmatrix} .9995669 & .02787606 & -.009430334 \\ -.02560409 & .9817861 & .1882565 \\ .01450642 & -.1879335 & .9820746 \end{bmatrix}$$

$$f_{U,V,W} \cdot T$$

three d dist,lines=46

X mean and var= 557018.359786 401450694971.590824
X mean abs dev = 533354.849069
X min and max= 799.892896 1955601.252115

Y mean and var= 966.259190 12651055.109059
Y mean abs dev = 2736.671417
Y min and max= -7251.526803 7749.663110

Z mean and var= 24.007084 8607.170170
Z mean abs dev = 63.349557
Z min and max= -220.197222 273.727329
cov xy = 0.000028
cov xz = -0.030139
cov yz = 0.000000
cum prob = 1.0000000000000

index	xamount	yamount	zamount	prob	cum
0000)	4623	-46	8	2.173913e-02	0.021739
0001)	23429	-378	7	2.173913e-02	0.043478
0002)	60806	-629	67	2.173913e-02	0.065217
0003)	144119	-1505	111	2.173913e-02	0.086957
0004)	219691	-1983	164	2.173913e-02	0.108696
0005)	327809	-2319	274	2.173913e-02	0.130435
0006)	380351	-3395	77	2.173913e-02	0.152174
0007)	500130	-4725	-84	2.173913e-02	0.173913
0008)	598054	-4054	50	2.173913e-02	0.195652
0009)	689665	-3023	201	2.173913e-02	0.217391
0010)	1578241	-7252	10	2.173913e-02	0.239130
0011)	1769167	-7058	-203	2.173913e-02	0.260870
0012)	1955601	-3397	170	2.173913e-02	0.282609
0013)	1937228	-673	252	2.173913e-02	0.304348
0014)	1783745	-489	-130	2.173913e-02	0.326087
0015)	1674784	2210	37	2.173913e-02	0.347826
0016)	1520844	2308	-220	2.173913e-02	0.369565
0017)	1448852	4554	-47	2.173913e-02	0.391304
0018)	1378372	6215	50	2.173913e-02	0.413043
0019)	1292473	7207	68	2.173913e-02	0.434783
0020)	1126173	7750	134	2.173913e-02	0.456522
0021)	902405	6509	4	2.173913e-02	0.478261
0022)	770308	6466	44	2.173913e-02	0.500000
0023)	698151	6220	-10	2.173913e-02	0.521739
0024)	612410	6652	119	2.173913e-02	0.543478
0025)	491267	5213	-7	2.173913e-02	0.565217
0026)	382198	4522	24	2.173913e-02	0.586957
0027)	310889	3874	9	2.173913e-02	0.608696
0028)	232913	3067	4	2.173913e-02	0.630435
0029)	188176	2345	-50	2.173913e-02	0.652174
0030)	145779	2140	3	2.173913e-02	0.673913
0031)	119369	1827	-0	2.173913e-02	0.695652
0032)	78125	1257	1	2.173913e-02	0.717391
0033)	50863	840	-2	2.173913e-02	0.739130
0034)	45387	809	4	2.173913e-02	0.760870
0035)	38697	717	4	2.173913e-02	0.782609
0036)	32939	607	-2	2.173913e-02	0.804348
0037)	26769	537	4	2.173913e-02	0.826087
0038)	21575	359	-14	2.173913e-02	0.847826
0039)	18421	396	3	2.173913e-02	0.869565
0040)	13964	260	-7	2.173913e-02	0.891304
0041)	9599	190	-4	2.173913e-02	0.913043
0042)	15616	262	-17	2.173913e-02	0.934783
0043)	1078	24	-0	2.173913e-02	0.956522
0044)	990	21	-0	2.173913e-02	0.978261
0045)	800	16	-0	2.173913e-02	1.000000

We notice that the 2-way covariances are indeed very close to zero, considering the size of the 2-way covariances of $\mathbf{f}_{U,V,W}$ above in Section 2.

$$(f_{U,V,W} \cdot T)^{+46}$$

```
three d dist,lines=4524

X mean and var= 25622844.550168 18466731968693.582550
X mean abs dev = 3432480.483357
X min and max= 36795.073198 89957657.597305

Y mean and var= 44447.922744 581948535.016732
Y mean abs dev = 23178.848586
Y min and max= -333570.232946 356484.503058

Z mean and var= 1104.325883 395929.827797
Z mean abs dev = 571.979816
Z min and max= -10129.072195 12591.457133
cov xy = 1156779159.382145
cov xz = -25537357.683619
cov yz = -323465.733458
cum prob = 1.0000000000000
```

$$T^{-1}$$

$$T^{-1} = \begin{bmatrix} .9995669 & -.02560409 & .01450642 \\ .02787606 & .9817861 & -.1879335 \\ -.009430334 & .1882565 & .9820746 \end{bmatrix}$$

We note that T^{-1}, the inverse of T, is simply the transpose of **T** here.

$$(f_{U,V,W} \cdot T)^{+46} \cdot T^{-1}$$

```
three d dist,lines=4524

X mean and var= 25612976.000000 18450805070829.184520
X mean abs dev = 3430998.182074
X min and max= 36800.000000 89914268.000000

Y mean and var= -612203.450268 12609153346.155333
Y mean abs dev = 90108.004949
Y min and max= -2455212.733131 -221.558077

Z mean and var= 364426.999998 3900088983.088916
Z mean abs dev = 49808.896875
Z min and max= 368.000000 1342003.999993
cov xy = -471473639326.225590
cov xz = 267525170784.773822
cov yz = -6943971141.832311
cum prob = 1.0000000000000
```

The results of the graduation are shown in the following table.

X	ϑ_X	N_X	ungrad ϑ_X/N_X	op.boot. grad. value	op.boot. std.err. of grad. value	cummax of std.err.	dev. = grad. - ungrad.
63	88	9487	0.009276	0.011296	0.000627	0.000627	0.00202
64	132	10770	0.012256	0.011602	0.000634	0.000634	-0.00065
65	267	24267	0.011003	0.012284	0.000489	0.000634	0.001281
66	300	26791	0.011198	0.014078	0.000242	0.000634	0.00288
67	432	29174	0.014808	0.014949	0.000206	0.000634	0.000141
68	491	28476	0.017243	0.017103	0.000497	0.000634	-0.00014
69	422	25840	0.016331	0.018436	0.000903	0.000903	0.002105
70	475	23916	0.019861	0.020121	0.000816	0.000903	0.00026
71	413	21412	0.019288	0.02268	0.000949	0.000949	0.003392
72	480	20116	0.023862	0.025601	0.001646	0.001646	0.001739
73	537	18876	0.028449	0.027007	0.000944	0.001646	-0.00144
74	566	17461	0.032415	0.031918	0.000528	0.001646	-0.0005
75	581	15012	0.038702	0.037102	0.000986	0.001646	-0.0016
76	464	11871	0.039087	0.041001	0.001799	0.001799	0.001914
77	461	10002	0.046091	0.046621	0.001092	0.001799	0.00053
78	433	8949	0.048385	0.05331	0.00191	0.00191	0.004925
79	515	7751	0.066443	0.055023	0.00223	0.00223	-0.01142
80	374	6140	0.060912	0.06778	0.002675	0.002675	0.006868
81	348	4718	0.07376	0.071606	0.001968	0.002675	-0.00215
82	304	3791	0.08019	0.084126	0.002315	0.002675	0.003936
83	249	2806	0.088738	0.08688	0.001838	0.002675	-0.00186
84	167	2240	0.074554	0.099215	0.005372	0.005372	0.024661
85	192	1715	0.111953	0.115262	0.002018	0.005372	0.003309
86	171	1388	0.123199	0.119198	0.004245	0.005372	-0.004
87	126	898	0.140312	0.134533	0.005349	0.005372	-0.00578
88	86	578	0.148789	0.146335	0.005197	0.005372	-0.00245
89	97	510	0.190196	0.159781	0.000073	0.005372	-0.03042

$\frac{\|dev\|}{cummax}$	X
3.221928	63
1.031968	64
2.021142	65
4.542918	66
0.222862	67
0.220174	68

2.33082	69
0.287729	70
3.574026	71
1.056742	72
0.875956	73
0.302003	74
0.972279	75
1.064008	76
0.294729	77
2.57838	78
5.121094	79
2.567457	80
0.805259	81
1.47143	82
0.694736	83
4.590735	84
0.615906	85
0.744759	86
1.075727	87
0.4568	88
5.661779	89

We have used cummax (the maximum so far as we move down the standard error column) as the denominator in the fraction $\frac{|dev|}{cummax}$, so as to avoid understating the standard error in this fraction. The parameters used in performing the numerical convolutions were nax=30, nay=30, naz=30, meshx=3 (log intervals), meshy=meshz=1 (equal intervals) and ε=1.e-299. The operational bootstrap standard error for age 89 looks low and would probably turn out higher if we increase nax, nay and naz significantly.

Section 5. Conclusion

The results from graduation by operational bootstrapping might for some purposes be used as is, or perhaps as a more reliable base (cf the raw data) upon which to perform a further graduation by one of the more usual methods. If smoothness is a priority item, then further graduation of the results may be indicated, since by itself graduation by operational bootstrapping imposes no requirements for smoothness (or fit).

This paper should be considered as an initial safari in the direction of graduation by operational bootstrapping. Further research is needed. For example, it might focus on such aspects as the desirability and possible implications of using initial transformations which

(1) use a base other than a=10 in $\log_a(q_x)$; or

(2) are according to some operators other than logs.

Or it might focus on how appropriate operational bootstrapping is for graduation of mortality rates in the first place.

Upon request the author will furnish (at cost of a diskette and mailing) the "C"-language subroutine which for a discrete finite multivariate distribution finds and applies to that distribution the linear transformation which will cause the 2-way covariances to be zero.

References

Anderson, T. W. (1984) *An Introduction To Multivariate Statistical Analysis*. John Wiley & Sons, Inc.

Bailey, W. A. (1992a) Bootstrapping For Order Statistics Sans Random Numbers. *Exploring The Limits Of Bootstrap.* John Wiley & Sons, Inc., 309-318

Bailey, W. A. (1992b) A Method For Determining Confidence Intervals For Trend. *TRANSACTIONS OF THE SOCIETY OF ACTUARIES*, Vol XLIV, 11-40.

Bailey, W. A. (1993) Six Bridges to ψ's. *ARCH*, 1993.1, 143-228.

Ramsay, C. M. (1991) Minimum Variance Moving-Weighted-Average Graduation. *TRANSACTIONS OF THE SOCIETY OF ACTUARIES*, Vol XLIII, 305-325.