# Supporting Material associated with "Homogeneity/Heterogeneity Hypothese for Standardized Mortality Ratios based on Minimum Power-divergence Estimators", published in *Biometrical Journal*

Leandro Pardo* and Nirian Martín†
Complutense University of Madrid

March 8, 2009

## Contents



*Mathematics Faculty, E-mail: lpardo@mat.ucm.es.
†School of Statistics, E-mail: nirian@estad.ucm.es (Corresponding author).

# 1 Introduction

Table 1: Melanoma death cases / exposed population ($x_{ij}/n_{ij}$), melanoma death rates ($P_i$) in the US during year 2000 (standard population), externally SMRs ($\widehat{Q}_j$) and SRRs ($\widehat{\widehat{Q}}_j$). All of them for white males in $I = 6$ age strata and $J = 9$ counties/states in the US during 2001-2005. Source: National Cancer Institute's SEER*Stat Database.

| Age-group ($i$) | Population ($j$) 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | US standard |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. $< 35$ | 4/1,120,267 | 2/1,065,065 | 6/1,054,133 | 7/3,311,763 | 1/504,363 | 15/5,352,902 | 1/580,619 | 0/414,714 | 3/2,445,686 | 130/56,323,119 |
| 2. $[35, 45)$ | 7/422,429 | 5/390,202 | 5/397,087 | 14/1,009,249 | 4/261,519 | 25/1,676,392 | 5/226,427 | 3/140,592 | 10/718,586 | 345/18,570,881 |
| 3. $[45, 55)$ | 18/408,696 | 12/415,803 | 9/371,416 | 50/1,030,145 | 12/174,403 | 59/1,570,771 | 5/177,269 | 12/143,806 | 18/537,160 | 760/15,794,000 |
| 4. $[55, 65)$ | 21/268,070 | 24/265,949 | 14/225,226 | 51/684,060 | 17/104,390 | 77/954,135 | 10/115,392 | 8/100,615 | 28/326,526 | 851/10,125,675 |
| 5. $[65, 75)$ | 18/142,289 | 30/194,878 | 13/122,440 | 54/461,820 | 3/59,911 | 69/528,036 | 10/44,158 | 7/64,994 | 34/178,348 | 1,047/7,334,391 |
| 6. $\geq 75$ | 32/117,112 | 36/193,377 | 25/109,324 | 71/415,392 | 13/56,453 | 113/378,631 | 15/35,135 | 21/56,582 | 39/127,818 | 1,381/5,569,217 |
| $\widehat{Q}_j$ | 0.9806 | 0.8526 | 0.7895 | 0.8168 | 1.0932 | 0.9714 | 1.1859 | 1.2070 | 1.0198 | 1.0000 |
| rank($\widehat{Q}_j$) | 5 | 3 | 1 | 2 | 8 | 4 | 9 | 10 | 6 | 7 |
| $\widehat{\widehat{Q}}_j$ | 0.9852 | 0.8594 | 0.8017 | 0.8214 | 1.0592 | 0.9893 | 1.3001 | 1.1910 | 1.0684 | 1.0000 |
| rank($\widehat{\widehat{Q}}_j$) | 4 | 3 | 1 | 2 | 7 | 5 | 10 | 9 | 8 | 6 |

Description of the $J = 9$ counties/states:

*Norththern regions*
- 1 = Oakland (Michigan)
- 2 = Allegheny (Pennsylvania)
- 3 = Hennepin (Minnesota)
- 4 = Iowa

*Middle regions*
- 5 = San Francisco (California)
- 6 = Colorado

*Southern regions*
- 7 = Fulton (Georgia)
- 8 = Jefferson (Alabama)
- 9 = Dallas (Texas)

# 2 Derivation of MPDE associated with the SMR

Table 2: Marginal observed death counts and estimated expected death counts based on MPDEs with $\lambda = 0$ and $\lambda = 1$ for the example of Table 1.

| Population | $\sum_{i=1}^{I} x_{ij}$ | $\sum_{i=1}^{I} \widehat{m}_{ij}^{0}$ | $\sum_{i=1}^{I} \widehat{m}_{ij}^{1}$ |
|---|---|---|---|
| $j=1$: Oakland (Michigan) | 169 | 169.000 | 169.931 |
| $j=2$: Allegheny (Pennsylvania) | 287 | 287.000 | 287.710 |
| $j=3$: Hennepin (Minnesota) | 62 | 62.000 | 65.314 |
| $j=4$: Iowa | 55 | 55.000 | 59.218 |
| $j=5$: San Francisco (California) | 560 | 560.000 | 561.272 |
| $j=6$: Colorado | 84 | 84.000 | 84.999 |
| $j=7$: Fulton (Georgia) | 116 | 116.000 | 116.632 |
| $j=8$: Jefferson (Alabama) | 98 | 98.000 | 100.152 |
| $j=9$: Dallas (Texas) | 307 | 307.000 | 308.664 |

## 2.1 Externally SMR

We are going to clarify the idea of "lack of a common scale" of the MPDEs of externally SMRs through the Rothman's data example (see page 45 in Rothman (1986)) shown in Table 3. Several MPDEs of summary ratios with the external standardization and with/without applying

$$P_I = P_I(\lambda) = \left( \prod_{j=1}^{J} \left( \sum_{i=1}^{I} \frac{n_{ij}p_i}{\sum_{h=1}^{I} n_{hj}p_h} \left( \frac{x_{ij}}{n_{ij}p_i} \right)^{\lambda+1} \right) \right)^{\frac{1}{J(1+\lambda)}}, \tag{1}$$

are in Table 4. We can see that $\widehat{Q}_j^0 < \widehat{Q}_j^1 < \widehat{Q}_j^5$ and it does not make sense to use

$$\widehat{Q}_j^{\lambda} = \frac{1}{P_I} \left( \frac{\sum\limits_{i=1}^{I} n_{ij}p_i \left( \frac{x_{ij}}{n_{ij}p_i} \right)^{\lambda+1}}{\sum\limits_{i=1}^{I} n_{ij}p_i} \right)^{\frac{1}{1+\lambda}}, \quad j = 1, ..., J, \tag{2}$$

with different values of $\lambda$ as as absolute measure of mortality, however by applying (1) we use (2) only focussed on ranking mortality ratios. It should be pointed out that these data are artificial and the people who support the usefulness of SMR say that in real data example are not common the situations in which such extreme population distributions appear (for instance in Table 1, where none normalization has been applied, there are no big differences between externally SMR and SRRs).

# 3 Monte Carlo Study

In Figure 1 some Probability-Plots of $T(1, N, 1)$ based on 10000 observations, jointly the MLE estimators of the parameters $\widehat{\mu}_{1,N}$ and $\widehat{\sigma}_{1,N}$ associated with the Lognormal distribution are shown. They are focussed on one hand on the unrestricted version of the minimum-chi square estimator (MPDE with $\lambda = 1$) for $Q_1$ and on the other hand on the restricted version of the minimum-chi square estimator

Table 3: Rothman's data example

| | Age-groups | |
|---|---|---|
| | young | old |
| Exposure 1 | $\frac{50}{10,000}$ | $\frac{4}{1,000}$ |
| Exposure 2 | $\frac{5}{1,000}$ | $\frac{40}{10,000}$ |
| Standard Population | $\frac{50}{100,000}$ | $\frac{400}{200,00}$ |

Table 4: Summary measures for Rothman's data example in Table 3

| | | without (1) | | | with (1) | | |
|---|---|---|---|---|---|---|---|
| Exposure $j$ | $\widehat{\widehat{Q}}_j$ | $\widehat{Q}_j^0$ | $\widehat{Q}_j^1$ | $\widehat{Q}_j^5$ | $\widehat{Q}_j^0$ | $\widehat{Q}_j^1$ | $\widehat{Q}_j^5$ |
| $j=1$ | 2.8889 | 7.7143 | 8.5189 | 9.4547 | 1.8746 | 1.8393 | 1.3247 |
| $j=2$ | 2.8889 | 2.1951 | 2.5182 | 5.3875 | 0.5334 | 0.5437 | 0.7549 |

for the same parameter, in both cases the same values of $\{P_i\}_{i=1}^6$, $\{Q_j\}_{j=1}^5$ are considered. Because we would like to analyze graphically whether the performance of $T(1,N,1) = ({}^*\widehat{Q}_1^1/Q_{1,0})^{\sqrt{N}}$ is the same for two very large but very different values of $N$, we consider $N \in \{1841.95, 5525.86\}$. We can see that there is no a big difference between the values of $\widehat{\mu}_{1,N}$ and $\widehat{\sigma}_{1,N}$ focussed on the same kind of estimator, either unrestricted or restricted, as it was expected because of the same asymptotic distribution ($\widehat{\mu}_{1,1841.95} = 0.01718 \simeq -0.005056 = \widehat{\mu}_{1,5525.86}$, $\widehat{\sigma}_{1,1841.95} = 2.123 \simeq 2.118 = \widehat{\sigma}_{1,5525.86}$ for unrestricted estimators; $\widehat{\mu}_{1,1841.95} = 0.01709 \simeq 0.004378 = \widehat{\mu}_{1,5525.86}$, $\widehat{\sigma}_{1,1841.95} = 0.8283 \simeq 0.8331 = \widehat{\sigma}_{1,5525.86}$ for restricted estimators). The value of $\widehat{\sigma}_{1,N}$ is smaller for the restricted estimator than the unrestricted one ($2.123 > 0.8283$ for $N = 1841.95$ and $2.118 > 0.8331$ for $N = 5525.86$), which means that more accurate estimators are obtained with restricted estimators. Maintaining the values of the parameters of the multiplicative model, but replacing $\lambda = 1$ by any value $\lambda \in \{-0.5, 0, \frac{2}{3}, 2\}$, almost the same results are obtained. Taking into account that $\widehat{\mu}_{1,5525.86} = 0.004378$ is an approximation of $\mu_1 = 0$ with $N = 5525.86$, we can understand in an easy way that for our model, related with rare events models, $N = 613.98 = 5525.86/9$ is not in fact a very large size.

# 4 Test-statistics for Homogeneity / Heterogeneity hypotheses of SMRs

Instaead of showing the formal theory needed for proving such results we are going to provide a short study of the expected exact means of the test-statistics which should approximately coincide with the degrees of freedom of the chi-square distributions (asymptotic distributions). Taking into account the same simulation design considered in the section devoted to the Monte Carlo study in the original paper and focussed only on the expected total size $N = 613.98$, in Table 5 are summarized the values of

$$\widehat{E}(T_h^\lambda) = \frac{1}{R}\sum_{r=1}^{R} T_h^\lambda(r), \quad h = 1, 2'$$

where $T_1^\lambda(r) \in \{X_\lambda^2(\mathcal{M}_1|\mathcal{M}_0), G_\lambda^2(\mathcal{M}_1|\mathcal{M}_0)\}$, $T_{2'}^\lambda(r) \in \{X_\lambda^2(\mathcal{M}_{2'}|\mathcal{M}_1), G_\lambda^2(\mathcal{M}_{2'}|\mathcal{M}_1)\}$ is computed according to the $r$-th replication of the total amount of replications, $R = 10000$. Observe that $\widehat{E}(T_1^\lambda) \simeq$

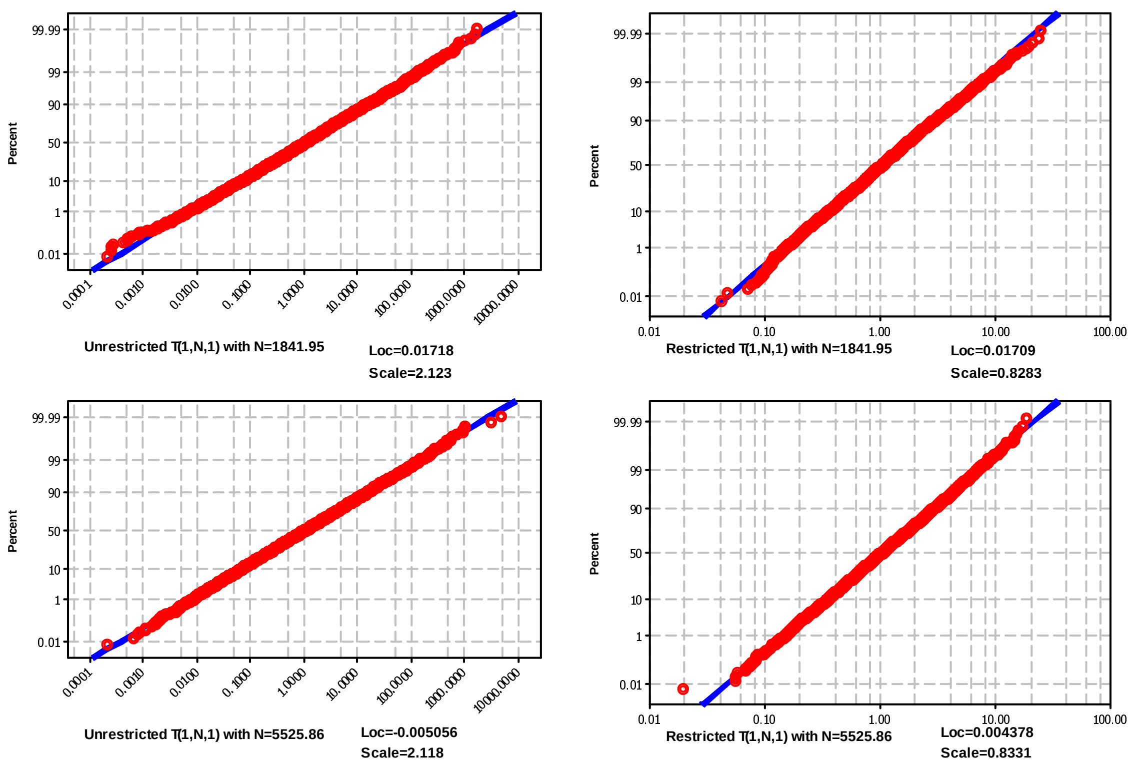


Figure 1: Lognormal Probability-Plots for 10000 observations of unrestricted and restricted $T(1, N, 1) = (\widehat{Q}_1^1 / Q_{1,0}^1)^{\sqrt{N}}$.

$g_1 = J - S = 6 - 3 = 3$ and $\widehat{E}(T_{2'}^{\lambda}) \simeq g_{2'} = S - 1 = 2 - 1 = 1$.

Table 5: Expected exact means of the test-statistics for the simulations carried out through Table 7 in the original paper.

| $\lambda$ | $\widehat{E}(X_\lambda^2(\mathcal{M}_1\|\mathcal{M}_0))$ | $\widehat{E}(G_\lambda^2(\mathcal{M}_1\|\mathcal{M}_0))$ | $\widehat{E}(X_\lambda^2(\mathcal{M}_{2'}\|\mathcal{M}_1))$ | $\widehat{E}(G_\lambda^2(\mathcal{M}_{2'}\|\mathcal{M}_1))$ |
|---|---|---|---|---|
| 0 | 2.953018 | 2.958994 | 1.006857 | 1.007985 |
| 1 | 2.984608 | 2.989591 | 0.999295 | 1.003080 |

# 5 Numerical Application Example

## 5.1 Code for Test 1

```
MODULE ParGlob
  INTEGER, PARAMETER:: AgeG=6, Regi=9, S=3, si(S)=(/4,6,9/)
  INTEGER iter
  DOUBLE PRECISION, PARAMETER :: lam =1.d0, prec = 1.d-12
  DOUBLE PRECISION n(AgeG,Regi), PE(AgeG), QE(Regi), PEI
  LOGICAL canContinue
END MODULE ParGlob

PROGRAM Test1
USE ParGlob
IMPLICIT NONE
DOUBLE PRECISION, PARAMETER:: bigbnd=10.d6
DOUBLE PRECISION x(AgeG,Regi), QE0(Regi), PE0(AgeG),
&   QE1(Regi), PE1(AgeG), testJ, testG, residJ(AgeG*Regi),
&   m0(AgeG*Regi), m1(AgeG*Regi), residG(AgeG*Regi)
INTEGER i,j
n(:,1)=(/1120267.d0,422429.d0,408696.d0,268070.d0,142289.d0,
&         117112.d0/)
n(:,2)=(/1065065.d0,390202.d0,415803.d0,265949.d0,194878.d0,
&         193377.d0/)
n(:,3)=(/1054133.d0,397087.d0,371416.d0,225226.d0,122440.d0,
&         109324.d0/)
n(:,4)=(/3311763.d0,1009249.d0,1030145.d0,684060.d0,461820.d0,
&         415392.d0/)
n(:,5)=(/504363.d0,261519.d0,174403.d0,104390.d0,59911.d0,
&         56453.d0/)
n(:,6)=(/5352902.d0,1676392.d0,1570771.d0,954135.d0,528036.d0,
&         378631.d0/)
n(:,7)=(/580619.d0,226427.d0,177269.d0,115392.d0,44158.d0,
&         35135.d0/)
n(:,8)=(/414714.d0,140592.d0,143806.d0,100615.d0,64994.d0,
&         56582.d0/)
n(:,9)=(/2445686.d0,718586.d0,537160.d0,326526.d0,178348.d0,
&        127818.d0/)
```

```
      x(:,1)=(/4.d0,7.d0,18.d0,21.d0,18.d0,32.d0/)
      x(:,2)=(/2.d0,5.d0,12.d0,24.d0,30.d0,36.d0/)
      x(:,3)=(/6.d0,5.d0,9.d0,14.d0,13.d0,25.d0/)
      x(:,4)=(/7.d0,14.d0,50.d0,51.d0,54.d0,71.d0/)
      x(:,5)=(/1.d0,4.d0,12.d0,17.d0,3.d0,13.d0/)
      x(:,6)=(/15.d0,25.d0,59.d0,77.d0,69.d0,113.d0/)
      x(:,7)=(/1.d0,5.d0,5.d0,10.d0,10.d0,15.d0/)
      x(:,8)=(/0.d0,3.d0,12.d0,8.d0,7.d0,21.d0/)
      x(:,9)=(/3.d0,10.d0,18.d0,28.d0,34.d0,39.d0/)
      OPEN (11, FILE='ResultsTest1.txt')
      DO j=1,Regi
       IF (SUM(x(:,j))<1) THEN
        print*, 'No BD normalization: There is no suitable sample
     &           for region)', j
        stop
       ENDIF
      ENDDO
 9025 FORMAT (G15.8)
 9026 FORMAT ('...........Q(j)s under H0...........')
 9027 FORMAT ('...........P(i)s under H0...........')
 9028 FORMAT ('...........Q(j)s under H1...........')
 9029 FORMAT ('...........P(i)s under H1...........')
 9030 FORMAT ('...........Chi-square test-statistic:')
 9031 FORMAT ('...........Likelihood Ratio test-statistic:')
 9032 FORMAT ('...........Externally SMRs with lambda:')
      QE=1.d0
      PEI=1.d0
      PE=1.d0
      canContinue = .TRUE.
      iter=0
      DOWHILE (canContinue)
        CALL algor0(x)
      ENDDO
      QE0=QE
      PE0=PEI*PE
      WRITE (11,9032)
      WRITE (11,9025) lam
      WRITE (11,9026)
      DO j=1,Regi
        WRITE (11,9025) QE0(j)
      ENDDO
      WRITE (11,9027)
      DO i=1,AgeG
        WRITE (11,9025) PE0(i)
      ENDDO
      QE=1.d0
      PEI=1.d0
      PE=1.d0
      canContinue = .TRUE.
```

```
iter=0
DOWHILE (canContinue)
  CALL algor1(x)
ENDDO
QE1=QE
PE1=PEI*PE
WRITE (11,9028)
DO j=1,Regi
  WRITE (11,9025) QE1(j)
ENDDO
WRITE (11,9029)
DO i=1,AgeG
  WRITE (11,9025) PE1(i)
ENDDO
DO i=1,AgeG
  DO j=1,Regi
    m0((j-1)*AgeG+i)=n(i,j)*PE0(i)*QE0(j)
    m1((j-1)*AgeG+i)=n(i,j)*PE1(i)*QE1(j)
  ENDDO
ENDDO
residJ=(m0-m1)/sqrt(m1)
testJ=SUM(residJ**2)
WRITE (11,9030)
WRITE (11,9025) testJ
residG=sqrt(2.d0*abs(m0*log(m0/m1)-m0+m1))
residG=SIGN(residG,(m0*log(m0/m1)-m0+m1))
testG=2.d0*SUM(m0*log(m0/m1)-m0+m1)
WRITE (11,9031)
WRITE (11,9025) testG
CLOSE(11)
END  PROGRAM Test1

SUBROUTINE algor0(x)
USE ParGlob
IMPLICIT NONE
INTEGER i, j
DOUBLE PRECISION np(AgeG,Regi), nq(AgeG,Regi),x(AgeG,Regi),
&               QE2(Regi), PE2(AgeG)
iter=iter+1
QE2=QE
PE2=PE*PEI
DO i=1,AgeG-1
  nq(i,:)=n(i,:)*QE
  PE(i)=SUM(nq(i,:)*(x(i,:)/nq(i,:))**(lam+1.d0))
  PE(i)=(PE(i)/SUM(nq(i,:)))**(1.d0/(lam+1.d0))
  PE(i)=PE(i)/PEI
ENDDO
DO j=1,Regi
  np(:,j)=n(:,j)*PE
```

```
  QE(j)=SUM(np(:,j)*(x(:,j)/np(:,j))**(lam+1.d0))
  QE(j)=(QE(j)/SUM(np(:,j)))**(1.d0/(lam+1.d0))
 ENDDO
 PEI =  (PRODUCT(QE))**(1.d0/Regi)
 QE = QE/PEI
 canContinue = (MAXVAL(ABS(QE-QE2)).GE.prec).OR.
&         (MAXVAL(ABS(PE*PEI-PE2)).GE.prec)
 END SUBROUTINE algor0

 SUBROUTINE algor1(x)
 USE ParGlob
 IMPLICIT NONE
 INTEGER i, j
 DOUBLE PRECISION np(AgeG,Regi), nq(AgeG,Regi), x(AgeG,Regi),
&              QE2(Regi), PE2(AgeG)
 QE2=QE
 PE2=PE*PEI
 DO i=1,AgeG-1
  nq(i,:)=n(i,:)*QE
  PE(i)=SUM(nq(i,:)*(x(i,:)/nq(i,:))**(lam+1.d0))
  PE(i)=(PE(i)/SUM(nq(i,:)))**(1.d0/(lam+1.d0))
  PE(i)=PE(i)/PEI
 ENDDO
 DO j=1,Regi
  np(:,j)=n(:,j)*PE
 ENDDO
 QE(1:si(1))=SUM(np(:,1:si(1))*(x(:,1:si(1))/
&           np(:,1:si(1)))**(lam+1.d0))
 QE(1:si(1))=(QE(1:si(1))/SUM(np(:,1:si(1))))
&           **(1.d0/(lam+1.d0))
 DO j=2,S
  QE(si(j-1)+1:si(j))=SUM(np(:,si(j-1)+1:si(j))*(x(:,si(j-1)
&           +1:si(j))/np(:,si(j-1)+1:si(j)))**(lam+1.d0))
  QE(si(j-1)+1:si(j))=(QE(si(j-1)+1:si(j))/SUM(np(:,si(j-1)
&           +1:si(j))))**(1.d0/(lam+1.d0))
 ENDDO
 PEI =  (PRODUCT(QE))**(1.d0/Regi)
 QE = QE/PEI
 canContinue = (MAXVAL(ABS(QE-QE2)).GE.prec).OR.
&         (MAXVAL(ABS(PE*PEI-PE2)).GE.prec)
 END SUBROUTINE algor1
```

## 5.2 Code for Test 2

```
 MODULE ParGlob
  INTEGER, PARAMETER:: AgeG=6, Regi=9, S=3, si(S)=(/4,6,9/)
  INTEGER iter
  DOUBLE PRECISION, PARAMETER :: lam =1.d0, prec = 1.d-12
  DOUBLE PRECISION n(AgeG,Regi), PE(AgeG), QE(Regi), PEI,
&     qi(Regi)
```

```
        LOGICAL canContinue
      END MODULE ParGlob

      PROGRAM Test2
      USE ParGlob
      IMPLICIT NONE
      DOUBLE PRECISION, PARAMETER:: bigbnd=10.d6
      DOUBLE PRECISION x(AgeG,Regi), QE2(Regi), PE2(AgeG),
     &   QE1(Regi), PE1(AgeG), testJ, testG, residJ(AgeG*Regi),
     &   m2(AgeG*Regi), m1(AgeG*Regi), residG(AgeG*Regi)
      INTEGER i,j
      n(:,1)=(/1120267.d0,422429.d0,408696.d0,268070.d0,142289.d0,
     &         117112.d0/)
      n(:,2)=(/1065065.d0,390202.d0,415803.d0,265949.d0,194878.d0,
     &         193377.d0/)
      n(:,3)=(/1054133.d0,397087.d0,371416.d0,225226.d0,122440.d0,
     &         109324.d0/)
      n(:,4)=(/3311763.d0,1009249.d0,1030145.d0,684060.d0,461820.d0,
     &         415392.d0/)
      n(:,5)=(/504363.d0,261519.d0,174403.d0,104390.d0,59911.d0,
     &         56453.d0/)
      n(:,6)=(/5352902.d0,1676392.d0,1570771.d0,954135.d0,528036.d0,
     &         378631.d0/)
      n(:,7)=(/580619.d0,226427.d0,177269.d0,115392.d0,44158.d0,
     &         35135.d0/)
      n(:,8)=(/414714.d0,140592.d0,143806.d0,100615.d0,64994.d0,
     &         56582.d0/)
      n(:,9)=(/2445686.d0,718586.d0,537160.d0,326526.d0,178348.d0,
     &        127818.d0/)
      x(:,1)=(/4.d0,7.d0,18.d0,21.d0,18.d0,32.d0/)
      x(:,2)=(/2.d0,5.d0,12.d0,24.d0,30.d0,36.d0/)
      x(:,3)=(/6.d0,5.d0,9.d0,14.d0,13.d0,25.d0/)
      x(:,4)=(/7.d0,14.d0,50.d0,51.d0,54.d0,71.d0/)
      x(:,5)=(/1.d0,4.d0,12.d0,17.d0,3.d0,13.d0/)
      x(:,6)=(/15.d0,25.d0,59.d0,77.d0,69.d0,113.d0/)
      x(:,7)=(/1.d0,5.d0,5.d0,10.d0,10.d0,15.d0/)
      x(:,8)=(/0.d0,3.d0,12.d0,8.d0,7.d0,21.d0/)
      x(:,9)=(/3.d0,10.d0,18.d0,28.d0,34.d0,39.d0/)
      qi=1.d0
      OPEN (11, FILE='ResultsTest2.txt')
      DO j=1,Regi
       IF (SUM(x(:,j))<1) THEN
        print*, 'No BD normalization: There is no suitable sample
     &           for region)', j
        stop
       ENDIF
      ENDDO
 9025 FORMAT (G15.8)
 9026 FORMAT ('...........Q(j)s under H1...........')
```

```
9027 FORMAT ('...........P(i)s under H1...........')
9028 FORMAT ('...........Q(j)s under H2...........')
9029 FORMAT ('...........P(i)s under H2...........')
9030 FORMAT ('...........Chi-square test-statistic:')
9031 FORMAT ('...........Likelihood Ratio test-statistic:')
9032 FORMAT ('...........Externally SMRs with lambda:')
     QE=1.d0
     PEI=1.d0
     PE=1.d0
     canContinue = .TRUE.
     iter=0
     DOWHILE (canContinue)
       CALL algor1(x)
     ENDDO
     QE1=QE
     PE1=PEI*PE
     WRITE (11,9032)
     WRITE (11,9025) lam
     WRITE (11,9026)
     DO j=1,Regi
       WRITE (11,9025) QE1(j)
     ENDDO
     WRITE (11,9027)
     DO i=1,AgeG
       WRITE (11,9025) PE1(i)
     ENDDO
     CALL algor2(x)
     QE2=QE
     PE2=PE
     WRITE (11,9028)
     DO j=1,Regi
       WRITE (11,9025) QE2(j)
     ENDDO
     WRITE (11,9029)
     DO i=1,AgeG
       WRITE (11,9025) PE2(i)
     ENDDO
     DO i=1,AgeG
       DO j=1,Regi
         m1((j-1)*AgeG+i)=n(i,j)*PE1(i)*QE1(j)
         m2((j-1)*AgeG+i)=n(i,j)*PE2(i)*QE2(j)
       ENDDO
     ENDDO
     residJ=(m1-m2)/sqrt(m2)
     testJ=SUM(residJ**2)
     WRITE (11,9030)
     WRITE (11,9025) testJ
     residG=sqrt(2.d0*abs(m1*log(m1/m2)-m1+m2))
     residG=SIGN(residG,(m1*log(m1/m2)-m1+m2))
```

```
 testG=2.d0*SUM(m1*log(m1/m2)-m1+m2)
 WRITE (11,9031)
 WRITE (11,9025) testG
 CLOSE(11)
 END PROGRAM Test2

 SUBROUTINE algor1(x)
 USE ParGlob
 IMPLICIT NONE
 INTEGER i, j
 DOUBLE PRECISION np(AgeG,Regi), nq(AgeG,Regi), x(AgeG,Regi),
&              QE2(Regi), PE2(AgeG)
 QE2=QE
 PE2=PE*PEI
 DO i=1,AgeG-1
   nq(i,:)=n(i,:)*QE
   PE(i)=SUM(nq(i,:)*(x(i,:)/nq(i,:))**(lam+1.d0))
   PE(i)=(PE(i)/SUM(nq(i,:)))**(1.d0/(lam+1.d0))
   PE(i)=PE(i)/PEI
 ENDDO
 DO j=1,Regi
   np(:,j)=n(:,j)*PE
 ENDDO
 QE(1:si(1))=SUM(np(:,1:si(1))*(x(:,1:si(1))/
&            np(:,1:si(1)))**(lam+1.d0))
 QE(1:si(1))=(QE(1:si(1))/SUM(np(:,1:si(1))))
&            **(1.d0/(lam+1.d0))
 DO j=2,S
   QE(si(j-1)+1:si(j))=SUM(np(:,si(j-1)+1:si(j))*(x(:,si(j-1)
&            +1:si(j))/np(:,si(j-1)+1:si(j)))**(lam+1.d0))
   QE(si(j-1)+1:si(j))=(QE(si(j-1)+1:si(j))/SUM(np(:,si(j-1)
&            +1:si(j))))**(1.d0/(lam+1.d0))
 ENDDO
 PEI =  (PRODUCT(QE))**(1.d0/Regi)
 QE = QE/PEI
 canContinue = (MAXVAL(ABS(QE-QE2)).GE.prec).OR.
&         (MAXVAL(ABS(PE*PEI-PE2)).GE.prec)
 END SUBROUTINE algor1

 SUBROUTINE algor2(x)
 USE ParGlob
 IMPLICIT NONE
 INTEGER i
 DOUBLE PRECISION x(AgeG,Regi), nq(AgeG,Regi)
 DO i=1,AgeG
   nq(i,:)=n(i,:)*qi
   PE(i)=SUM(nq(i,:)*(x(i,:)/nq(i,:))**(lam+1.d0))
   PE(i)=(PE(i)/SUM(nq(i,:)))**(1.d0/(lam+1.d0))
 ENDDO
```

```
QE(Regi)=(PRODUCT(qi)**(1.d0/Regi))
PE=PE*QE(Regi)
QE=qi/QE(Regi)
END SUBROUTINE algor2
```

### 5.3 Code for Test 3

```
 MODULE ParGlob
   INTEGER, PARAMETER:: AgeG=6, Regi=9, S=3, si(S)=(/4,6,9/)
   INTEGER iter
   DOUBLE PRECISION, PARAMETER :: lam =1.d0, prec = 1.d-12
   DOUBLE PRECISION n(AgeG,Regi), PE(AgeG), QE(Regi), PEI,
&      qi(Regi)
   LOGICAL canContinue
 END MODULE ParGlob

 PROGRAM Test3
 USE ParGlob
 IMPLICIT NONE
 DOUBLE PRECISION, PARAMETER:: bigbnd=10.d6
 DOUBLE PRECISION x(AgeG,Regi), QE2(Regi), PE2(AgeG),
&   QE1(Regi), PE1(AgeG), testJ, testG, residJ(AgeG*Regi),
&   m2(AgeG*Regi), m1(AgeG*Regi), residG(AgeG*Regi)
 INTEGER i,j
 n(:,1)=(/1120267.d0,422429.d0,408696.d0,268070.d0,142289.d0,
&         117112.d0/)
 n(:,2)=(/1065065.d0,390202.d0,415803.d0,265949.d0,194878.d0,
&         193377.d0/)
 n(:,3)=(/1054133.d0,397087.d0,371416.d0,225226.d0,122440.d0,
&         109324.d0/)
 n(:,4)=(/3311763.d0,1009249.d0,1030145.d0,684060.d0,461820.d0,
&         415392.d0/)
 n(:,5)=(/504363.d0,261519.d0,174403.d0,104390.d0,59911.d0,
&         56453.d0/)
 n(:,6)=(/5352902.d0,1676392.d0,1570771.d0,954135.d0,528036.d0,
&         378631.d0/)
 n(:,7)=(/580619.d0,226427.d0,177269.d0,115392.d0,44158.d0,
&         35135.d0/)
 n(:,8)=(/414714.d0,140592.d0,143806.d0,100615.d0,64994.d0,
&         56582.d0/)
 n(:,9)=(/2445686.d0,718586.d0,537160.d0,326526.d0,178348.d0,
&        127818.d0/)
 x(:,1)=(/4.d0,7.d0,18.d0,21.d0,18.d0,32.d0/)
 x(:,2)=(/2.d0,5.d0,12.d0,24.d0,30.d0,36.d0/)
 x(:,3)=(/6.d0,5.d0,9.d0,14.d0,13.d0,25.d0/)
 x(:,4)=(/7.d0,14.d0,50.d0,51.d0,54.d0,71.d0/)
 x(:,5)=(/1.d0,4.d0,12.d0,17.d0,3.d0,13.d0/)
 x(:,6)=(/15.d0,25.d0,59.d0,77.d0,69.d0,113.d0/)
 x(:,7)=(/1.d0,5.d0,5.d0,10.d0,10.d0,15.d0/)
 x(:,8)=(/0.d0,3.d0,12.d0,8.d0,7.d0,21.d0/)
```

```
      x(:,9)=(/3.d0,10.d0,18.d0,28.d0,34.d0,39.d0/)
      qi=1.d0
      qi(1:4)=3.d0/4.d0
      qi(5:6)=7.d0/8.d0
      OPEN (11, FILE='ResultsTest3.txt')
      DO j=1,Regi
       IF (SUM(x(:,j))<1) THEN
        print*, 'No BD normalization: There is no suitable sample
     &           for region)', j
        stop
       ENDIF
      ENDDO
 9025 FORMAT (G15.8)
 9026 FORMAT ('...........Q(j)s under H1...........')
 9027 FORMAT ('...........P(i)s under H1...........')
 9028 FORMAT ('...........Q(j)s under H2p...........')
 9029 FORMAT ('...........P(i)s under H2p...........')
 9030 FORMAT ('...........Chi-square test-statistic:')
 9031 FORMAT ('...........Likelihood Ratio test-statistic:')
 9032 FORMAT ('...........Externally SMRs with lambda:')
      QE=1.d0
      PEI=1.d0
      PE=1.d0
      canContinue = .TRUE.
      iter=0
      DOWHILE (canContinue)
        CALL algor1(x)
      ENDDO
      QE1=QE
      PE1=PEI*PE
      WRITE (11,9032)
      WRITE (11,9025) lam
      WRITE (11,9026)
      DO j=1,Regi
        WRITE (11,9025) QE1(j)
      ENDDO
      WRITE (11,9027)
      DO i=1,AgeG
        WRITE (11,9025) PE1(i)
      ENDDO
      CALL algor2(x)
      QE2=QE
      PE2=PE
      WRITE (11,9028)
      DO j=1,Regi
        WRITE (11,9025) QE2(j)
      ENDDO
      WRITE (11,9029)
      DO i=1,AgeG
```

```
  WRITE (11,9025) PE2(i)
ENDDO
DO i=1,AgeG
  DO j=1,Regi
    m1((j-1)*AgeG+i)=n(i,j)*PE1(i)*QE1(j)
    m2((j-1)*AgeG+i)=n(i,j)*PE2(i)*QE2(j)
  ENDDO
ENDDO
residJ=(m1-m2)/sqrt(m2)
testJ=SUM(residJ**2)
WRITE (11,9030)
WRITE (11,9025) testJ
residG=sqrt(2.d0*abs(m1*log(m1/m2)-m1+m2))
residG=SIGN(residG,(m1*log(m1/m2)-m1+m2))
testG=2.d0*SUM(m1*log(m1/m2)-m1+m2)
WRITE (11,9031)
WRITE (11,9025) testG
CLOSE(11)
END  PROGRAM Test3

SUBROUTINE algor1(x)
USE ParGlob
IMPLICIT NONE
INTEGER i, j
DOUBLE PRECISION np(AgeG,Regi), nq(AgeG,Regi), x(AgeG,Regi),
&              QE2(Regi), PE2(AgeG)
QE2=QE
PE2=PE*PEI
DO i=1,AgeG-1
  nq(i,:)=n(i,:)*QE
  PE(i)=SUM(nq(i,:)*(x(i,:)/nq(i,:))**(lam+1.d0))
  PE(i)=(PE(i)/SUM(nq(i,:)))**(1.d0/(lam+1.d0))
  PE(i)=PE(i)/PEI
ENDDO
DO j=1,Regi
  np(:,j)=n(:,j)*PE
ENDDO
QE(1:si(1))=SUM(np(:,1:si(1))*(x(:,1:si(1))/
&            np(:,1:si(1)))**(lam+1.d0))
QE(1:si(1))=(QE(1:si(1))/SUM(np(:,1:si(1))))
&            **(1.d0/(lam+1.d0))
DO j=2,S
  QE(si(j-1)+1:si(j))=SUM(np(:,si(j-1)+1:si(j))*(x(:,si(j-1)
&            +1:si(j))/np(:,si(j-1)+1:si(j)))**(lam+1.d0))
  QE(si(j-1)+1:si(j))=(QE(si(j-1)+1:si(j))/SUM(np(:,si(j-1)
&            +1:si(j))))**(1.d0/(lam+1.d0))
ENDDO
PEI =  (PRODUCT(QE))**(1.d0/Regi)
QE = QE/PEI
```

```
canContinue = (MAXVAL(ABS(QE-QE2)).GE.prec).OR.
&         (MAXVAL(ABS(PE*PEI-PE2)).GE.prec)
END SUBROUTINE algor1

SUBROUTINE algor2(x)
USE ParGlob
IMPLICIT NONE
INTEGER i
DOUBLE PRECISION x(AgeG,Regi), nq(AgeG,Regi)
DO i=1,AgeG
  nq(i,:)=n(i,:)*qi
  PE(i)=SUM(nq(i,:)*(x(i,:)/nq(i,:))**(lam+1.d0))
  PE(i)=(PE(i)/SUM(nq(i,:)))**(1.d0/(lam+1.d0))
ENDDO
QE(Regi)=(PRODUCT(qi)**(1.d0/Regi))
PE=PE*QE(Regi)
QE=qi/QE(Regi)
END SUBROUTINE algor2
```

## References

**Rothman, K.J.** (1986). *Modern Epidemiology.* Little, Brown and Company, Boston.
**Surveillance, Epidemiology, and End Results (SEER) Program** (www.seer.cancer.gov) SEER*Stat Database: Incidence - SEER 17 Regs Limited-Use + Hurricane Katrina Impacted Louisiana Cases, Nov 2007 Sub (1973-2005 varying) - Linked To County Attributes - Total U.S., 1969-2005 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2008, based on the November 2007 submission.