# REVIEW OF REPORT OF COMMITTEE ON MORTALITY FOR DISABLED LIVES

*Gary G. Venter*

*Barbara Schill*

*Jack Barnett*

**Review of Report of**
**Committee on Mortality for**
**Disabled Lives**

Abstract

The problem of what mortality tables to use for injured worker pension reserves is not a new one for casualty actuaries. A study of this issue appeared in the 1945 PCAS. We looked at the data from that study using computer intensive non-linear regression to model the ratio of injured worker to standard mortality.

The methodology and some of the conclusions may still be applicable today. In particular, injured worker mortality after some years comes close to standard mortality, and after some age may actually be lower. Because of this, not much credit can be taken on pension case reserves, even though for younger workers initial mortality is much higher than standard.

Some technical issues in non-linear regression are addressed, including a method to adjust for heteroscedasticity and using the information matrix to measure the significance of the parameters.

## REVIEW OF REPORT OF COMMITTEE ON MORTALITY FOR DISABLED LIVES

Loss reserves for workers compensation cases in the U.S. now are in the area of $50 billion, much of which is tied up in long term cases. Typically standard mortality is used to reserve these cases, but in serious cases a factor (e.g. 10) is applied to the mortality rates on a judgment basis, as in Snader (1987). Some disabled life tables have been calculated from other benefit systems, involving, for example heart disease or cancer cases, but these are probably not appropriate for injured workers.

Faced for the 25 years since the inception of workers compensation insurance with the need for injured worker mortality tables, the CAS decided to take action, and in 1937 appointed a Committee of Three to investigate the feasibility of undertaking a study. Coincidentally, the Committee of Three came up with three conclusions:

1. Very substantial results could not be expected from the data then available.

2. A start should be made in order to get carriers to keep appropriate records.

3. It was as feasible then as it would be at any later time to do a mortality study based on the statistical system in place.

Thus, working with the National Council on Compensation Insurance, a call for disability data was sent out in October 1938. The data used in the study was for accident years or policy years 1930-1935, depending on how carriers reported, and the first year of disability was excluded from each case. Although the first year after the accident was excluded, the data represented fairly new claimants, who might be expected to display higher mortality than more stabilized cases. The results of the study would thus be most applicable to such cases.

This review looks at the data from that study to see if there are any relationships between disabled worker mortality and standard mortality that might endure to the present. A regression methodology is used to explore this question. As the uniform variance assumption of least squares regression is not met, a method for dealing with this heteroscedasticity is developed. The information matrix from the (non-linear) regression is used to test goodness of fit and to develop prediction intervals.

### COMMITTEE REPORT

The report of the committee on mortality for disabled lives produced a mortality table for lives disabled by industrial accidents. The table is based on permanent total cases and nondismemberment permanent partial cases involving 50% or more disability. In total there were 8,598 life years of exposure with 285 claim terminations. The 285 claim terminations included deaths and the few cases where the injured person recovered. These claim terminations did not include cases where permanent partial disability followed permanent total, the benefit period ended, or a lump sum settlement was made. Since the mortality table in workers compensation is primarily used to determine expected claim size it is appropriate to include terminations due to either death or recovery. An alternative method is a multiple decrement model in which deaths and recoveries are measured separately. However the committee chose to consider both types of terminations together.

In the original study, mortality rates for each age were calculated based on the reported data. For those ages with sparse data, below age 22 and over age 73, the reported mortality rates were weighted with the mortality rates from the 1930 U.S. life tables for white males. The resulting mortality rates for ages 10 to 105 were graduated using the Whittaker-Henderson technique. Mortality tables were then constructed with these mortality rates.

The authors state that the mortality rate for these disabled lives is 144% of that for white males in the 1930 U.S. Life Tables. This was determined by comparing the expected number of deaths in the next year under the disabled workers table of mortality rates versus the U.S. Life Table mortality rates. The expected number of deaths is determined by multiplying the number of lives exposed for each age group by the respective mortality rate and summing for all ages. It is clear from the data, however that this 144% varies dramatically and systematically by age.

## RELATIONSHIP BETWEEN DISABLED WORKER MORTALITY AND STANDARD MORTALITY

We projected the mortality rates for disabled workers based on our hypothesis that the ratio, $q_d/q_u$, between the mortality rate for disabled workers, $q_d$, and that of the U.S. population, $q_u$, is a decreasing function of age. This is an alternate method of graduation to the Whittaker-Henderson formula used by the committee. Initially we set the mortality rate of disabled workers equal to a constant plus a power of the mortality rate of the U.S. multiplied by a function of age;

$$q_d = a + q_u^b \times f(\text{age})$$

We found that the constant, $a$, was insignificant. In all regressions attempted of $q_d$ on $q_u$ and age our estimate of the power of $q_u$ was approximately one. Together these suggest that the ratio of $q_d/q_u$ can be adequately expressed as a function of age.

Let $y_t$ be the ratio of observed disabled worker mortality to U.S. population standard mortality at age $t$. A fairly simple model was found to fit quite well:

$$y_t = be^{c/t} + \varepsilon_t \, ; \qquad \text{with } b = 0.32 \text{ and } c = 84$$

The ratio of the parameter to its estimated standard deviation is 3.72 for $b$ and is 10.83 for $c$.

Graph 1 shows three regressions of $y_t$ on $be^{c/t}$ with the parameter $c$ set equal to 1, 40 and 84. The graph illustrates the importance of $c$ in the model.

In addition, in graph 2 a comparison of the ratio of $q_d/q_u$ to the confidence intervals for the model indicates heteroscedasticity (the variance around the fitted line is not constant over age). The observed $q_d/q_u$ has a much greater variance at younger ages where, on average, $q_d/q_u$ is greater. Therefore rather than assume the constant variance of standard least squares regression it was assumed that errors were normally distributed with mean equal to zero and standard deviation proportional to the mean of the regression. This is referred to as the multiplicative error model and is described further in Appendix 1. The distribution of the error term $\varepsilon_t$ is approximated by a normal distribution:

$$\varepsilon_t = y_t - be^{c/t} \sim N(0, b^2 e^{2c/t} \sigma^2) \qquad \text{where } \sigma^2 = \text{constant of proportionality}$$

In Appendix 1 it is shown that this model can be fit by a standard regression with the "dependent variable" set equal to one , and $y_t/be^{c/t}$ as the independent variable. Then the parameters $b$ and $c$ are found to be, respectively, 0.35 and 88 which are respectively, 6.86 and 13.08 times the estimated parameter standard deviations. Graph 3 shows the observed data along with the confidence intervals for this multiplicative model. This illustrates the basis for the assumption that the standard deviation of $\varepsilon_t$ is proportional to the mean, in that the model confidence intervals more closely approximate the data variations. Table 1 compares the observed $y_t$ and the values from the two fitted models.

To estimate the standard deviations of the parameters for this model we calculated the variance-covariance matrix which is the inverse of the information matrix as described on page 81 of *Loss Distributions* by Robert V. Hogg and Stuart A. Klugman. The calculations of the information matrix and its resulting variance-covariance matrix for both the constant variance and the proportional variance model are described in Appendix 2.

A comparison of mortality rates for 1930 and 1980 from the U.S. Life Tables and the projected mortality rates for disabled workers based on the models is shown in

Table 2. Since the committee used the 1930 U.S. Life Table for white males we used the same 1980 table.

DISCUSSION

The hypothesis that the ratio between the mortality rate for disabled workers versus the population, $q_d/q_u$, is a decreasing function of age is supported by the data analysis described above.

It is possible that the ratio $q_d/q_u$ is closer to one now than is reflected in the 1930's data. The improvements in mortality of the general population may be heavily influenced by a disproportionately larger improvement in the mortality of disabled people. It will require another study of disabled workers mortality to determine if disabled worker mortality is now closer to standard mortality.

At an advanced age, there is a crossover point at which the mortality rate of disabled workers becomes less than that of the general population (Table 2). With the committee's method this occurs at age 81. With the multiplicative error model the crossover occurs at age 85. It is reasonable to assume that since these disabled workers had recently been in the work force at an advanced age they were healthier than the general population. The permanent injuries received were not necessarily serious enough to increase the mortality of these exceptionally healthy individuals to the level of the general population at that age.

In fact a fairly minor injury may be "permanent" at an older age in that the person may not return to work. This may contribute to the existence of a crossover point since permanent disability benefits supplement retirement income for older workers and could thus discourage return to work. Since on average today's workers retire earlier than they did in the 1930's the crossover point may be earlier now.

Below are the annuity values for certain ages calculated with the 1979-81 U.S. Life Tables and with estimated disabled workers' mortalities based on the proportional variance model. These annuity values contain an interest rate assumption of 3.5% and escalating benefits are assumed to increase at 7% per year.

**Lifetime Annuity Values**

| | U.S. Life Table | | Disabled Mortality | |
|---|---|---|---|---|
| Age | Nonescalating | Escalating | Nonescalating | Escalating |
| 25 | 22.756 | 136.298 | 20.272 | 111.229 |
| 45 | 17.776 | 58.464 | 16.631 | 52.366 |
| 65 | 11.009 | 21.442 | 10.507 | 20.364 |
| 85 | 4.606 | 6.117 | 4.811 | 6.486 |

These disabled worker mortalities are created from the general population of permanent total disabled workers and may not apply to the most severely injured workers. As mentioned earlier since the mortality rates are based on recently injured workers they may not be appropriate for claimants who have been disabled for many years. The disabled worker annuity values do not change drastically from those for the general population but they do decrease. However for advanced ages the annuities under the disabled worker mortalities are actually greater than under the U.S. Life Table mortalities.

CONCLUSIONS

1. A model which declines with age seems appropriate for $q_d/q_u$, the ratio between the mortality rate for disabled workers and that of the U.S. population.

2. At some age this ratio goes below unity and this may now occur at an earlier age.

3. The impact of the disabled mortality rates on the annuity values was moderate then and would probably be even less now.

4. These results may not be applicable to the first year of injury when higher mortality rates are likely or to longer period after injury where mortality rates closer to standard are expected.

Table 1

| Age | Ratio of Observed Mortality Rate to 1930 U.S. Standard Mortality Rate | Fitted Ratio from Constant Variance Model (1) | Fitted Ratio from Proportional Variance Model (2) |
|---|---|---|---|
| 24 | 8.2541 | 10.6254 | 13.7001 |
| 25 | 9.6604 | 9.2373 | 11.8330 |
| 26 | 14.7013 | 8.1175 | 10.3362 |
| 27 | 8.0420 | 7.2020 | 9.1196 |
| 28 | 2.6410 | 6.4446 | 8.1185 |
| 29 | 2.1841 | 5.8113 | 7.2855 |
| 30 | 6.3777 | 5.2764 | 6.5853 |
| 31 | 5.2512 | 4.8207 | 5.9914 |
| 32 | 4.9615 | 4.4293 | 5.4833 |
| 33 | 0.0000 | 4.0907 | 5.0453 |
| 34 | 8.4568 | 3.7956 | 4.6651 |
| 35 | 3.9529 | 3.5369 | 4.3329 |
| 36 | 1.1813 | 3.3088 | 4.0409 |
| 37 | 2.0036 | 3.1066 | 3.7828 |
| 38 | 4.4908 | 2.9264 | 3.5536 |
| 39 | 3.2170 | 2.7652 | 3.3489 |
| 40 | 2.1517 | 2.6202 | 3.1654 |
| 41 | 1.3040 | 2.4894 | 3.0002 |
| 42 | 1.2320 | 2.3709 | 2.8509 |
| 43 | 2.1564 | 2.2631 | 2.7154 |
| 44 | 2.9405 | 2.1648 | 2.5922 |
| 45 | 2.8654 | 2.0749 | 2.4796 |
| 46 | 1.7136 | 1.9924 | 2.3765 |
| 47 | 2.4772 | 1.9165 | 2.2818 |
| 48 | 1.5980 | 1.8464 | 2.1946 |
| 49 | 2.3456 | 1.7816 | 2.1141 |
| 50 | 1.5227 | 1.7216 | 2.0396 |
| 51 | 2.8791 | 1.6658 | 1.9705 |
| 52 | 1.2276 | 1.6139 | 1.9062 |
| 53 | 1.3889 | 1.5654 | 1.8464 |
| 54 | 1.3349 | 1.5201 | 1.7905 |
| 55 | 1.5800 | 1.4778 | 1.7383 |
| 56 | 1.6526 | 1.4380 | 1.6894 |
| 57 | 1.6292 | 1.4006 | 1.6435 |
| 58 | 1.8961 | 1.3655 | 1.6004 |
| 59 | 0.5384 | 1.3324 | 1.5598 |
| 60 | 2.1415 | 1.3012 | 1.5215 |
| 61 | 1.6078 | 1.2716 | 1.4854 |
| 62 | 1.7536 | 1.2437 | 1.4513 |
| 63 | 1.3142 | 1.2172 | 1.4190 |
| 64 | 0.7567 | 1.1921 | 1.3884 |
| 65 | 1.1449 | 1.1683 | 1.3594 |
| 66 | 0.9790 | 1.1457 | 1.3318 |
| 67 | 1.2446 | 1.1241 | 1.3056 |
| 68 | 0.6668 | 1.1036 | 1.2806 |
| 69 | 0.7997 | 1.0840 | 1.2569 |
| 70 | 0.2978 | 1.0653 | 1.2342 |
| 71 | 0.9891 | 1.0474 | 1.2126 |
| 72 | 1.5846 | 1.0304 | 1.1919 |
| 73 | 0.8659 | 1.0140 | 1.1721 |
| 74 | 0.9447 | 0.9984 | 1.1532 |
| 75 | 1.3963 | 0.9834 | 1.1351 |
| 76 | 0.8882 | 0.9690 | 1.1177 |
| 77 | 1.6805 | 0.9552 | 1.1010 |
| 78 | 1.1974 | 0.9419 | 1.0850 |
| 79 | 0.6338 | 0.9292 | 1.0697 |
| 80 | 0.4526 | 0.9169 | 1.0549 |
| 81 | 1.3872 | 0.9051 | 1.0407 |
| 82 | 1.1605 | 0.8937 | 1.0270 |
| 83 | 0.6815 | 0.8828 | 1.0138 |
| 84 | 0.3539 | 0.8722 | 1.0011 |
| 85 | 1.2400 | 0.8620 | 0.9889 |
| 86 | 0.5859 | 0.8521 | 0.9770 |

(1) Y(t)=0.32086e**(84/t)
(2) Y(t)=0.35155e**(87.9074/t)

Table 2

1930

| AGE | U.S. Life Table tQx | Disabled Mortality: Raw Data tQx' | Disabled Mortality: Committee tQx'' | Disabled Mortality: Fit(REG) tQx''' | Disabled Mortality: Fit(MAX) tQx'''' |
|---|---|---|---|---|---|
| 18 | | | | | |
| 19 | | | | | |
| 20 | | | | | |
| 21 | | | | | |
| 22 | | | | | |
| 23 | | | | | |
| 24 | .0037 | .0302 | .0259 | .0389 | .0501 |
| 25 | .0037 | .0358 | .0255 | .0343 | .0439 |
| 26 | .0037 | .0551 | .0250 | .0304 | .0388 |
| 27 | .0038 | .0306 | .0243 | .0274 | .0347 |
| 28 | .0039 | .0103 | .0236 | .0251 | .0317 |
| 29 | .0040 | .0088 | .0227 | .0234 | .0293 |
| 30 | .0041 | .0263 | .0218 | .0218 | .0272 |
| 31 | .0043 | .0224 | .0209 | .0205 | .0255 |
| 32 | .0044 | .0219 | .0201 | .0196 | .0242 |
| 33 | .0046 | .0000 | .0192 | .0189 | .0234 |
| 34 | .0049 | .0411 | .0185 | .0184 | .0227 |
| 35 | .0051 | .0202 | .0178 | .0180 | .0221 |
| 36 | .0053 | .0063 | .0173 | .0177 | .0216 |
| 37 | .0056 | .0113 | .0169 | .0175 | .0213 |
| 38 | .0060 | .0268 | .0166 | .0175 | .0212 |
| 39 | .0064 | .0285 | .0165 | .0176 | .0213 |
| 40 | .0068 | .0146 | .0166 | .0178 | .0215 |
| 41 | .0073 | .0095 | .0169 | .0181 | .0218 |
| 42 | .0078 | .0096 | .0174 | .0184 | .0221 |
| 43 | .0082 | .0178 | .0180 | .0187 | .0224 |
| 44 | .0087 | .0257 | .0187 | .0189 | .0227 |
| 45 | .0093 | .0266 | .0195 | .0193 | .0230 |
| 46 | .0099 | .0169 | .0204 | .0197 | .0235 |
| 47 | .0105 | .0261 | .0214 | .0202 | .0240 |
| 48 | .0112 | .0179 | .0224 | .0207 | .0246 |
| 49 | .9120 | .0281 | .0234 | .0213 | .0253 |
| 50 | .0128 | .0195 | .0245 | .0220 | .0261 |
| 51 | .0136 | .0393 | .0256 | .0227 | .0269 |
| 52 | .0146 | .0179 | .0268 | .0235 | .0278 |
| 53 | .0157 | .0217 | .0281 | .0245 | .0289 |
| 54 | .0169 | .0225 | .0294 | .0256 | .0302 |
| 55 | .0182 | .0287 | .0308 | .0269 | .0316 |
| 56 | .0197 | .0325 | .0322 | .0283 | .0332 |
| 57 | .0212 | .0346 | .0335 | .0298 | .0349 |
| 58 | .0229 | .0434 | .0347 | .0313 | .0366 |
| 59 | .0246 | .0132 | .0358 | .0328 | .0384 |
| 60 | .0264 | .0566 | .0367 | .0344 | .0402 |
| 61 | .0284 | .0456 | .0376 | .0361 | .0422 |
| 62 | .0305 | .0535 | .0383 | .0380 | .0443 |
| 63 | .0330 | .0433 | .0391 | .0401 | .0468 |

1980

| AGE | U.S. Life Table tQx | Disabled Mortality tQx''' | Disabled Mortality tQx''' |
|---|---|---|---|
| 18 | .0015 | .0515 | .0701 |
| 19 | .0016 | .0435 | .0585 |
| 20 | .0017 | .0374 | .0499 |
| 21 | .0019 | .0326 | .0430 |
| 22 | .0019 | .0282 | .0369 |
| 23 | .0019 | .0239 | .0310 |
| 24 | .0019 | .0201 | .0259 |
| 25 | .0018 | .0169 | .0217 |
| 26 | .0018 | .0144 | .0183 |
| 27 | .0017 | .0124 | .0157 |
| 28 | .0017 | .0108 | .0136 |
| 29 | .0017 | .0097 | .0122 |
| 30 | .0017 | .0088 | .0109 |
| 31 | .0016 | .0080 | .0099 |
| 32 | .0017 | .0074 | .0091 |
| 33 | .0017 | .0069 | .0085 |
| 34 | .0017 | .0066 | .0082 |
| 35 | .0018 | .0065 | .0080 |
| 36 | .0020 | .0065 | .0079 |
| 37 | .0021 | .0065 | .0079 |
| 38 | .0022 | .0066 | .0080 |
| 39 | .0024 | .0066 | .0080 |
| 40 | .0026 | .0068 | .0083 |
| 41 | .0029 | .0071 | .0086 |
| 42 | .0032 | .0075 | .0090 |
| 43 | .0035 | .0079 | .0094 |
| 44 | .0038 | .0083 | .0099 |
| 45 | .0042 | .0087 | .0104 |
| 46 | .0046 | .0092 | .0110 |
| 47 | .0051 | .0099 | .0117 |
| 48 | .0057 | .0106 | .0126 |
| 49 | .0064 | .0114 | .0135 |
| 50 | .0071 | .0122 | .0144 |
| 51 | .0077 | .0129 | 0153 |
| 52 | .0085 | .0137 | .0162 |
| 53 | .0093 | .0146 | .0172 |
| 54 | .0103 | .0156 | .0184 |
| 55 | .0112 | .0166 | .0196 |
| 56 | .0123 | .0176 | .0207 |
| 57 | .0134 | .0187 | .0220 |
| 58 | .0146 | .0200 | .0234 |
| 59 | .0160 | .0214 | .0250 |
| 60 | .0176 | .0229 | .0268 |
| 61 | .0193 | .0246 | .0287 |
| 62 | .0212 | .0264 | .0308 |
| 63 | .0232 | .0282 | .0329 |

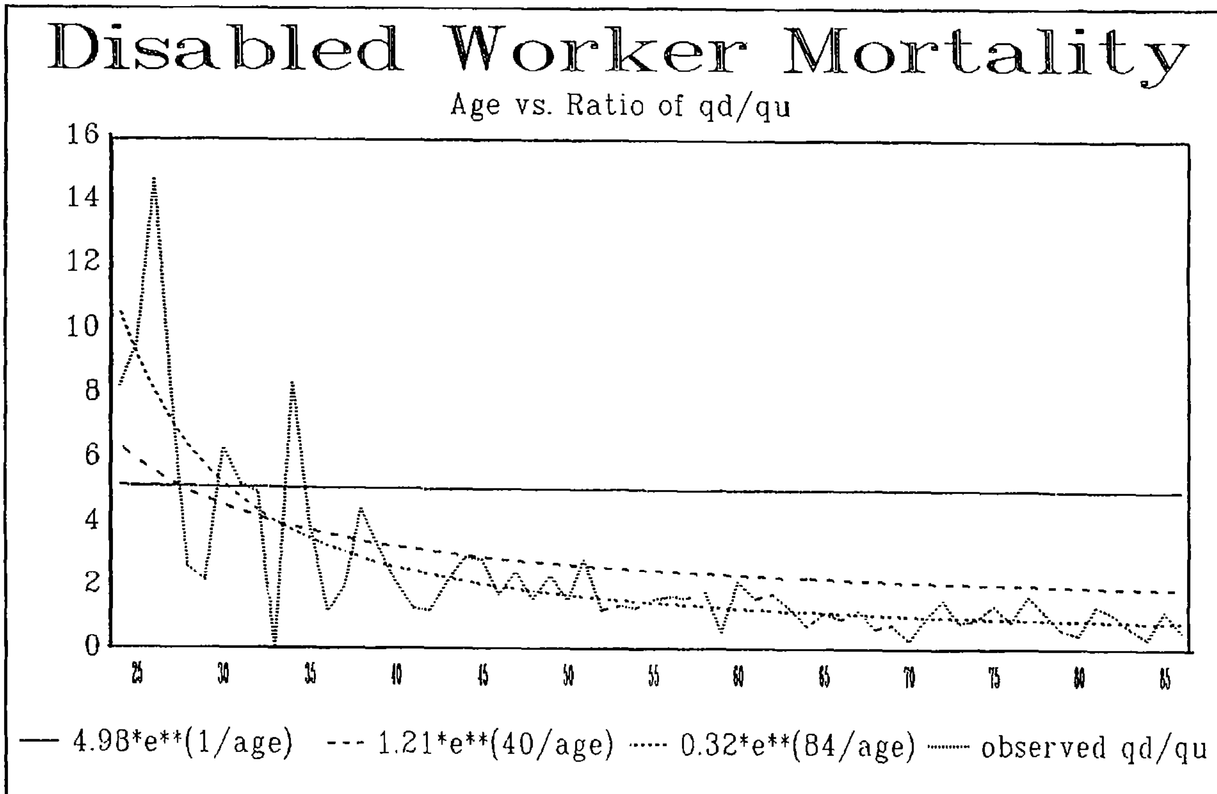
1930

| AGE | U.S. Life Table | Disabled Mortality | | | |
|---|---|---|---|---|---|
| | | Raw Data | Committee | Fit(REG) | Fit(MAX) |
| | tQx | tQx' | tQx'' | tQx''' | tQx'''' |
| 64 | .0357 | .0270 | .0400 | .0425 | .0495 |
| 65 | .0386 | .0442 | .0412 | .0452 | .0525 |
| 66 | .0420 | .0411 | .0428 | .0481 | .0559 |
| 67 | .0456 | .0567 | .0451 | .0512 | .0595 |
| 68 | .0495 | .0330 | .0481 | .0546 | .0634 |
| 69 | .0536 | .0429 | .0519 | .0581 | .0674 |
| 70 | .0580 | .0173 | .0566 | .0617 | .0715 |
| 71 | .0625 | .0618 | .0621 | .0655 | .0758 |
| 72 | .0674 | .1068 | .0682 | .0694 | .0803 |
| 73 | .0727 | .0630 | .0750 | .0737 | .0852 |
| 74 | .0786 | .0743 | .0822 | .0785 | .0907 |
| 75 | .0853 | .1190 | .0898 | .0838 | .0968 |
| 76 | .0927 | .0824 | .0976 | .0899 | .1037 |
| 77 | .1010 | .1698 | .1056 | .0965 | .1113 |
| 78 | .1101 | .1319 | .1137 | .1037 | .1195 |
| 79 | .1199 | .0759 | .1220 | .1114 | .1282 |
| 80 | .1300 | .0588 | .1305 | .1192 | .1371 |
| 81 | .1404 | .1948 | .1393 | .1271 | .1461 |
| 82 | .1512 | .1754 | .1485 | .1361 | .1553 |
| 83 | .1621 | .1105 | .1581 | .1459 | .1644 |
| 84 | .1733 | .0613 | .1681 | .1560 | .1735 |
| 85 | .1847 | .2290 | .1787 | .1662 | .1826 |
| 86 | .1962 | .1149 | .1899 | .1766 | .1917 |
| 87 | .2078 | | .2019 | .1870 | .2007 |
| 88 | .2197 | | .2146 | .1977 | .2097 |
| 89 | .2321 | | .2283 | .2089 | .2191 |
| 90 | .2455 | | .2429 | .2209 | .2292 |
| 91 | .2602 | | .2587 | .2342 | .2403 |
| 92 | .2763 | | .2757 | .2487 | .2525 |
| 93 | .2940 | | .2941 | .2646 | .2660 |
| 94 | .3133 | | .3140 | .2820 | .2820 |
| 95 | .3344 | | .3356 | .3010 | .3010 |
| 96 | .3574 | | .3589 | .3217 | .3217 |
| 97 | .3824 | | .3841 | .3442 | .3442 |
| 98 | .4095 | | .4113 | .3686 | .3686 |
| 99 | .4388 | | .4406 | .3949 | .3949 |
| 100 | .4704 | | .4720 | .4233 | .4233 |
| 101 | .5044 | | .5057 | .4539 | .4539 |
| 102 | .5409 | | .5417 | .4868 | .4868 |
| 103 | .5800 | | .5799 | .5220 | .5220 |
| 104 | .6219 | | .6204 | .5597 | .5597 |
| 105 | .6666 | | .6666 | .5999 | .5999 |
| 106 | | | | | |
| 107 | | | | | |
| 108 | | | | | |
| 109 | | | | | |

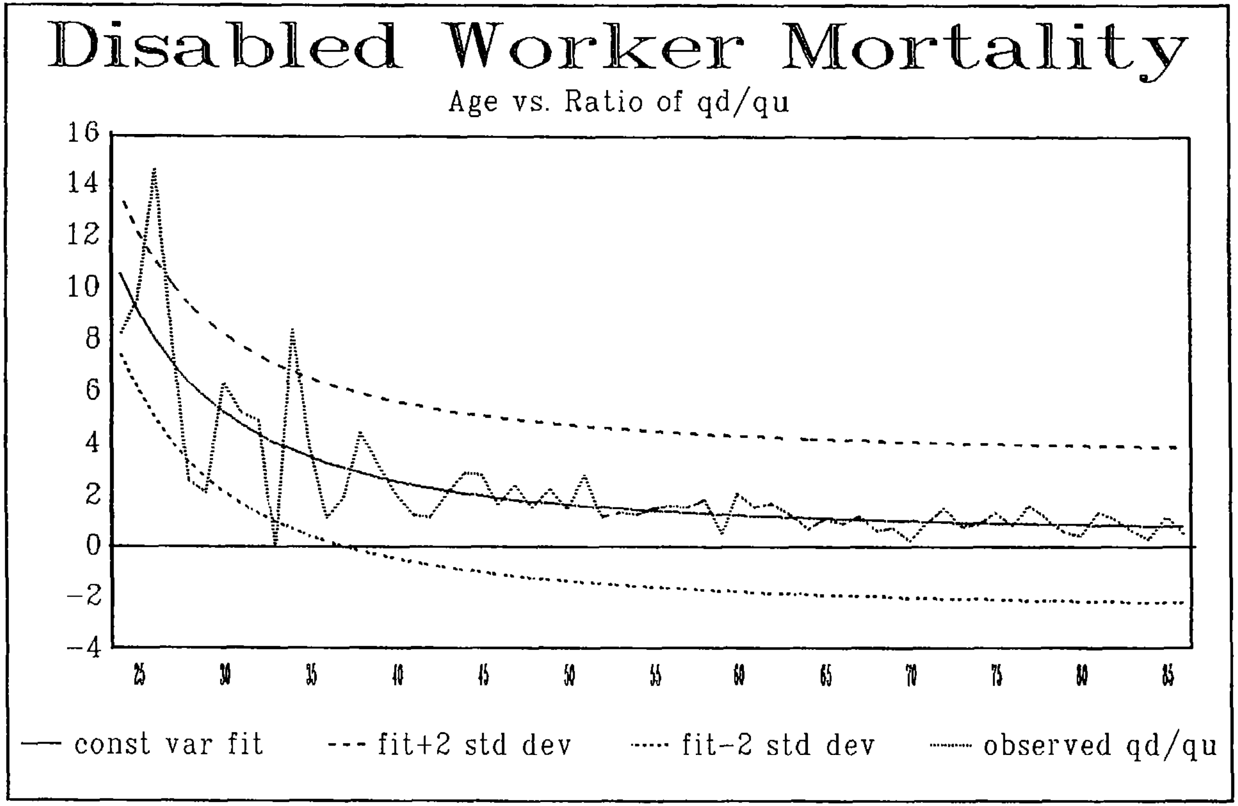
1980.

| AGE | U.S. Life Table | Disabled Mortality | |
|---|---|---|---|
| | tQx | tQx''' | tQx'''' |
| 64 | .0252 | .0301 | .0350 |
| 65 | .0274 | .0320 | .0372 |
| 66 | .0297 | .0340 | .0395 |
| 67 | .0322 | .0362 | .0420 |
| 68 | .0349 | .0386 | .0448 |
| 69 | .0380 | .0412 | .0478 |
| 70 | .0415 | .0442 | .0512 |
| 71 | .0452 | .0473 | .0548 |
| 72 | .0490 | .0505 | .0584 |
| 73 | .0529 | .0537 | .0621 |
| 74 | .0570 | .0569 | .0658 |
| 75 | .0615 | .0604 | .0698 |
| 76 | .0664 | .0644 | .0742 |
| 77 | .0718 | .0686 | .0791 |
| 78 | .0776 | .0731 | .0842 |
| 79 | .0839 | .0780 | .0898 |
| 80 | .0910 | .0834 | .0960 |
| 81 | .0989 | .0895 | .1029 |
| 82 | .1073 | .0966 | .1102 |
| 83 | .1161 | .1045 | .1177 |
| 84 | .1252 | .1127 | .1254 |
| 85 | .1351 | .1216 | .1336 |
| 86 | .1459 | .1313 | .1426 |
| 87 | .1569 | .1412 | .1515 |
| 88 | .1677 | .1510 | .1601 |
| 89 | .1787 | .1609 | .1687 |
| 90 | .1906 | .1715 | .1779 |
| 91 | .2039 | .1835 | .1883 |
| 92 | .2186 | .1968 | .1998 |
| 93 | .2345 | .2111 | .2122 |
| 94 | .2506 | .2255 | .2255 |
| 95 | .2662 | .2396 | .2396 |
| 96 | .2800 | .2520 | .2520 |
| 97 | .2931 | .2638 | .2638 |
| 98 | .3054 | .2749 | .2749 |
| 99 | .3170 | .2853 | .2853 |
| 100 | .3278 | .2951 | .2951 |
| 101 | .3379 | .3041 | .3041 |
| 102 | .3472 | .3125 | .3125 |
| 103 | .3559 | .3203 | .3203 |
| 104 | .3638 | .3275 | .3275 |
| 105 | .3712 | .3341 | .3341 |
| 106 | .3779 | .3401 | .3401 |
| 107 | .3841 | .3457 | .3457 |
| 108 | .3897 | .3507 | .3507 |
| 109 | .3949 | .3554 | .3554 |

Graph 1

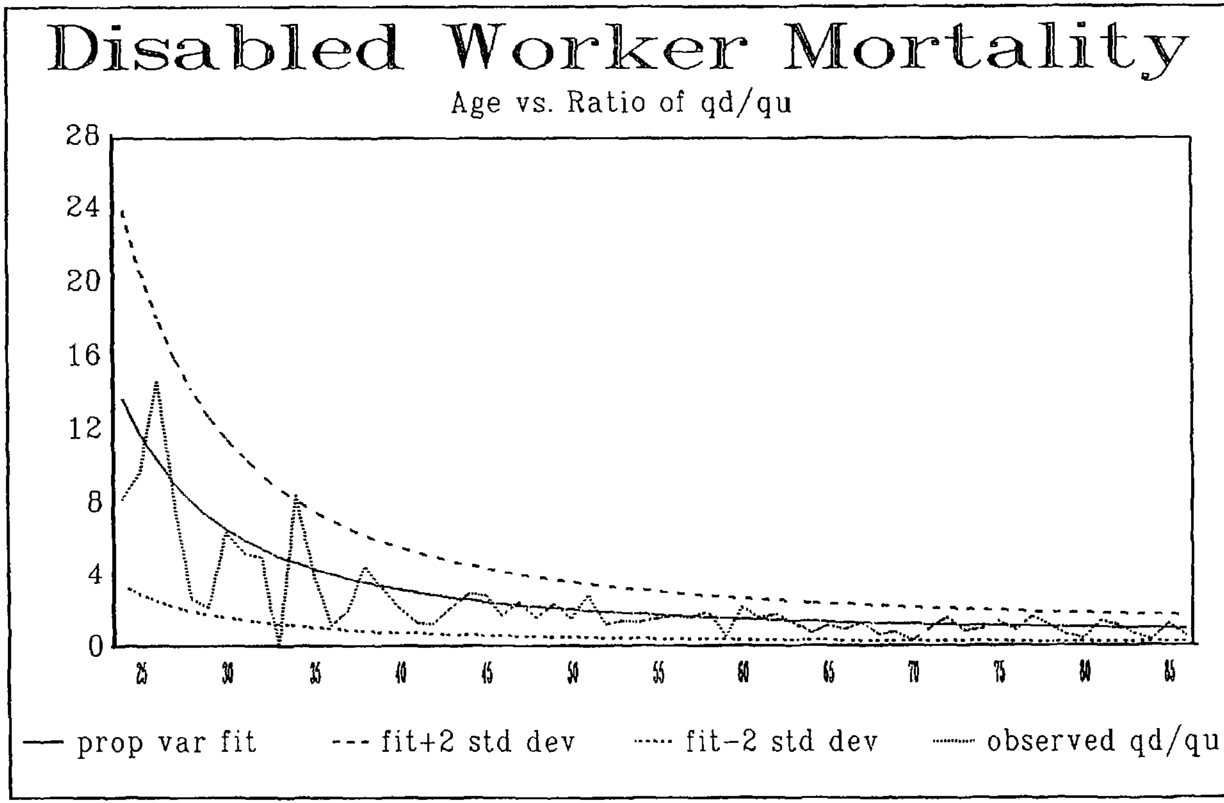

Disabled Worker Mortality
Age vs. Ratio of qd/qu
16
14
12
10
8
6
4
2
0
25
30
35
40
45
50
55
60
65
70
75
80
85
— 4.98*e**(1/age)
--- 1.21*e**(40/age)
······ 0.32*e**(84/age)
observed qd/qu

Graph 2



Disabled Worker Mortality
Age vs. Ratio of qd/qu
16
14
12
10
8
6
4
2
0
-2
-4
25
30
35
40
45
50
55
60
65
70
75
80
85
const var fit
fit+2 std dev
fit-2 std dev
observed qd/qu

Graph 3


Disabled Worker Mortality
Age vs. Ratio of qd/qu
28
24
20
16
12
8
4
0
25
30
35
40
45
50
55
60
65
70
75
80
85
prop var fit
fit+2 std dev
fit-2 std dev
observed qd/qu

## Appendix 1 Regression formulas

Regression with additive error structure

This is the standard least squares regression method.

Model is : $y_t = g(x_{1t} \ldots x_{kt}) + \epsilon_t$

where: $y_t$ is the dependent variable

$x_1 \ldots x_k$ are the independent variables

$g$ is the function with parameters to be estimated

$\epsilon_t$ is $\sim \mathcal{N}(0,\sigma^2)$

The additive error structure is appropriate when it can be assumed that the conditional variance = $\text{var}\{y_t \mid g(x_{1t} \ldots x_{kt})\}$ = constant = $\sigma^2$. In other words the variance $\sigma^2$ is independent of t. This is an assumption of least square regression referred to as homoscedasticity.

Assuming a normal distribution of the disturbance term $\epsilon_t$ the maximum likelihood estimates for the parameters of g minimize:

$$\sum_t \epsilon_t^2 = \sum_t \left[y_t - g(x_{1t} \ldots x_{kt})\right]^2$$

The regression function used is: $g(x_{1t}) = be^{c/t}$

where $x_{1t} = t =$ age

Our model becomes : $y_t = be^{c/t} + \epsilon_t$

where $y_t$ is the observed ratio of injured worker mortality to standard mortality at age t.

The regression finds b and c which minimize: $\sum_t \left[y_t - be^{c/t}\right]^2$

## Appendix 1 Regression Formulas

Regression with multiplicative error structure.

Model is : $y_t = g(x_{1t}...x_{kt})(1+\epsilon_t) = g(x_{1t}...x_{kt}) + \epsilon_t \cdot g(x_{1t}...x_{kt})$

where $\epsilon_t$ is $\sim \mathcal{N}(0,\sigma^2)$

Thus the disturbance term increases in size with the function.

This multiplicative error structure is appropriate when it can be assumed that the $\mathrm{var}(\, y_t \mid g(x_{1t}...x_{kt})) = g(x_{1t}...x_{kt})^2\sigma^2$ i.e, the variance increases with the square of the function (the conditional mean).

Also, $$\epsilon_t = \frac{y_t - g(x_{1t}...x_{kt})}{g(x_{1t}...x_{kt})} = \frac{y_t}{g(x_{1t}...x_{kt})} - 1$$

This $\epsilon_t$ satisfies the assumptions of standard least squares regression, that is : $\epsilon_t \sim \mathcal{N}(0,\sigma^2)$, so the maximum likelihood estimates of the parameters of g minimize:

$$\sum_t \left[\frac{y_t}{g(x_{1t}...x_{kt})} - 1\right]^2$$

An alternative model (which we did not use) is : $y_t = g(x_{1t}...x_{kt}) + \epsilon_t\sqrt{g(x_{1t}...x_{kt})}$

Which requires minimization of : $$\sum_t \left[\frac{y_t}{\sqrt{g(x_{1t}...x_{kt})}} - \sqrt{g(x_{1t}...x_{kt})}\right]^2$$

$\mathrm{var}(\, y_t \mid g(x_{1t}...x_{kt})) = g(x_{1t}...x_{kt})\sigma^2$

Here the variance increases linearly with the conditional mean.

## Appendix 1 Regression formulas

Both of these error structures are examples of heteroscedasticity, a common violation of the assumptions of least squares regression.

A multiplicative model was used and eventually chosen as the model that best "fit" our data .

The regression function used is: $g(x_{1t}) = be^{c/t}$

where $x_{1t} = t = \text{age}$

Our model becomes : $y_t = be^{c/t}(1 + \epsilon_t)$

For this model , the regression minimizes: $\sum_t \left[\frac{y_t}{be^{c/t}} - 1\right]^2$

This is equivalent to minimizing the sum of the squares of the proportional errors.

## Appendix 2 Significance of Parameters

Regression can be regarded as fitting a distribution (often a normal distribution) to the error terms $\epsilon_t$ by the method of maximum likelihood.

Variances and covariances of the regression parameters can thus be estimated by the inverse of the information matrix as described in *LOSS DISTRIBUTIONS* by Robert V. Hogg - Stuart A. Klugman (Page 81).

If $f(\epsilon;\theta)$ is the density function for the error terms, and $\theta$ is a vector listing the parameters to be estimated, the ijth element of the information matrix is:

$$a_{ij}(\theta) = -n\mathrm{E}\left[\frac{\partial^2 \ln f(\epsilon;\theta)}{\partial\theta_i\partial\theta_j}\right],$$ Here n is the number of observations.

This is typically estimated by:

$$a_{ij} \approx -\sum_{t=1}^{n}\frac{\partial^2 \ln f(\epsilon_t;\hat{\theta})}{\partial\theta_i\partial\theta_j} = -\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln\prod_{t=1}^{n} f(\epsilon_t;\hat{\theta})$$

Where $\hat{\theta}$ is the vector of parameter estimates and
$\epsilon_t$ = observed deviation from the model for observation t.
Thus the information matrix is estimated by the second partials of the negative loglikelihood.

Additive error structure

For our model: $y_t = be^{c/t} + \epsilon_t$ $\quad \hat{\theta} = \langle b,c,\sigma^2\rangle$ and $f(\epsilon_t;\theta) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\epsilon_t^2/2\sigma^2}$

so that : $\epsilon_t = y_t - be^{c/t}$ $\quad$ Since $\epsilon_t \sim \mathcal{N}(0,\sigma^2)$

Thus $$\ln f(\epsilon_t;\theta) = -\frac{1}{2}\ln 2\pi - \ln\sigma - \frac{\epsilon_t^2}{2\sigma^2}$$

$$= -\frac{1}{2}\ln 2\pi - \ln\sigma - \left[y_t - be^{c/t}\right]^2\frac{1}{2\sigma^2}$$ Since $\epsilon_t = \left[y_t - be^{c/t}\right]$

## Appendix 2 Significance of Parameters

Taking the partial derivatives of $\ln f(\epsilon_t;\theta)$ with respect to b,c and $\sigma^2$ (after some algebra) yields the following estimates of the $a_{ij}$ :

$$a_{11} = \frac{1}{\sigma^2}\sum_t e^{2c/t}$$

$$a_{12} = a_{21} = \frac{1}{\sigma^2}\sum_t \frac{e^{c/t}}{t}\left[2be^{c/t} - y_t\right]$$

$$a_{22} = \frac{b}{\sigma^2}\sum_t \frac{e^{c/t}}{t^2}\left[2be^{c/t} - y_t\right]$$

$$a_{13} = a_{31} = \frac{1}{\sigma^4}\sum_t e^{c/t}\left[y_t - be^{c/t}\right] = \frac{1}{\sigma^4}\sum_t e^{c/t}\epsilon_t$$

$$a_{23} = a_{32} = \frac{b}{\sigma^4}\sum_t \frac{e^{c/t}}{t}\left[y_t - be^{c/t}\right] = \frac{b}{\sigma^4}\sum_t \frac{e^{c/t}}{t}\epsilon_t$$

$$a_{33} = -\frac{n}{2\sigma^4} + \frac{1}{\sigma^6}\sum_t\left[y_t - be^{c/t}\right]^2 = -\frac{n}{2\sigma^4} + \frac{1}{\sigma^6}\sum_t\epsilon_t^2$$

For the data used the sum is from t=24 to t=86.

Appendix 2 Significance of Parameters

For our example the maximum likelihood estimates of the parameters are:

$\hat{b} = .32$ , $\hat{c} = 84$ and $\bar{\sigma}^2 = 2.34$ yielding the

Information Matrix:

$$\begin{bmatrix} 2664.4519 & 28.7613 & .9412 \\ 28.7613 & .3271 & .0104 \\ .9412 & .0104 & 5.0397 \end{bmatrix}$$

Taking the matrix inverse gives us the Variance-Covariance Matrix:

$$\begin{bmatrix} .0074 & -.6493 & 0 \\ -.6493 & 60.1556 & -.0028 \\ 0 & -.0028 & .1984 \end{bmatrix}$$

Our final step is to check the significance of our parameters. We do this by observing the ratio of the estimated parameter values to their standard deviations.

Standard error of parameter b : $\sqrt{.0074} = .086$ $.32/.086 = 3.72$

Standard error of parameter c: $\sqrt{60.16} = 7.76$ $84/7.76 = 10.83$

Parameters b and c appear to be significant.

## Appendix 2 Significance of Parameters

Multiplicative error structure

$\theta = <b,c,\sigma^2>$

$\epsilon_t =$ observed deviation from the model for observation t

Again: $f(\epsilon_t;\theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\epsilon_t^2/2\sigma^2}$ and

$$\ln f(\epsilon_t;\theta) = -\frac{1}{2}\ln 2\pi - \ln\sigma - \frac{\epsilon_t^2}{2\sigma^2}$$

$$= -\frac{1}{2}\ln 2\pi - \ln\sigma - \left[\frac{y_t}{be^{c/t}} - 1\right]^2 \frac{1}{2\sigma^2} \quad \text{,since } \epsilon_t = \left[\frac{y_t}{be^{c/t}} - 1\right]$$

Taking the partial derivatives of $\ln f(\epsilon_t;\theta)$ with respect to b,c and $\sigma^2$ yields the following estimates of the $a_{ij}$ :

$$a_{11} = \frac{1}{b^2\sigma^2}\sum_t (\epsilon_t + 1)(3\epsilon_t + 1)$$

$$a_{12} = a_{21} = \frac{1}{b\sigma^2}\sum_t \frac{1}{t}(\epsilon_t + 1)(2\epsilon_t + 1)$$

$$a_{13} = a_{31} = \frac{1}{b\sigma^4}\sum (\epsilon_t + 1)\epsilon_t$$

$$a_{22} = \frac{1}{\sigma^2}\sum_t \frac{1}{t^2}(\epsilon_t + 1)(2\epsilon_t + 1)$$

$$a_{23} = a_{32} = \frac{1}{\sigma^4}\sum \frac{1}{t}(\epsilon_t + 1)\epsilon_t$$

$$a_{33} = \frac{-n}{2\sigma^4} + \frac{1}{\sigma^6}\sum_t \epsilon_t^2$$

Appendix 2 Significance of Parameters

For our example: $\hat{b} = .35$ , $\hat{c} = 88$ and $\hat{\sigma}^2 = .15$ yielding the

Information Matrix:

$$\begin{bmatrix} 2953.559 & 20.9673 & 17.3812 \\ 20.9674 & .1709 & .1104 \\ 17.3812 & .1104 & 1348.404 \end{bmatrix}$$

Taking the inverse of this matrix gives us the Variance-Covariance Matrix:

$$\begin{bmatrix} .0026 & -.3218 & 0 \\ -.3218 & 45.3341 & .0004 \\ 0 & .0004 & .0007 \end{bmatrix}$$

Standard error of parameter b : $\sqrt{.0026} = .051$ $.35/.051 = 6.86$

Standard error of parameter c: $\sqrt{45.33} = 6.73$ $88/6.73 = 13.08$

Parameters appear to be significant.