ACTUARIAL RESEARCH CLEARING HOUSE
1995 VOL. 1

Bivariate Schuette graduation of race-specific mortality rates

Esther Portnoy
University of Illinois at Urbana–Champaign

Abstract

Bivariate Schuette graduation is applied to mortality rates among black male retired workers (Social Security data) at ages 70 to 83, calendar years 1984 to 1992. Graduated values for various choices of the parameters are compared. The Schuette method is also adapted to calculate *regression quantiles*, providing upper and lower envelopes for the graduated values.

Mortality rates u_{ij} as functions of age i and calendar year of experience j can be graduated by a two–dimensional Schuette method, that is, choosing v_{ij} to minimize

$$M = \Sigma\, w_{ij}\ |u_{ij} - v_{ij}| + \alpha\, \Sigma\, |\Delta^2_{\downarrow} v_{ij}| + \beta\, \Sigma\, |\Delta^2_{\rightarrow} v_{ij}| \,. \tag{1}$$

For any given parameters (α,β), the optimal solution $\{v_{ij}\}$ is one of a finite number of basic solutions (in the linear–programming sense); and one can use parametric–programming techniques to find the parameter region in which a given solution is optimal. The method is described in [Portnoy, 1994].

The data set used here covers black male Social Security recipients (retired workers), aged 70 to 83 in 1984 through 1992; see Table 1. (Age is the "vertical" variable, associated with the parameter α.) Solution of (1) requires a design matrix with 332 rows and 126 columns. For a given pair (α,β), a short program in Splus generates the optimal solution in a few seconds, and in a few seconds more the "perspective" plots of graduated values. (Readers interested in specifics of the program should contact the author.) The time required increases somewhat as the parameters increase, presumably because of degeneracy.)

The computer does the number–crunching, leaving to the graduator only one real question: what smoothing parameters to use. As a first step, I calculated the parameter region in which the no–graduation solution is optimal. The technique described in Section 4 of [Portnoy, 1994] works quite easily and gives a pentagonal region with vertices (0,0), (379.5,0), (353.625,103.5), (36.9375,525.75), (0,525.75). The region is of little interest itself, but it gives us an idea of how large we must choose the (α,β) to see the effect of graduation.

It would be useful to know for what parameters the optimal solution is planar (all differences vanishing), but I was not able to carry out that exercise with the programs I had. With α and β each at 10^7 the solution (which came with a warning about non–uniqueness) had $v_{ij} = u_{ij}$ at only 4 of the 126 indices, and the largest residual differences were about 10^{-7}.

With these numbers in mind, I applied the computer program to generate and then plot graduated values for pairs $(\alpha,\beta) = (3^a\times1000, 3^b\times1000)$, where a and b are integers between 0 and 6. A few of the most instructive plots are shown at the end of the text.

Figure 1 shows the ungraduated rates. Note that rates for later years are based on a 10% sample of Social Security beneficiaries; this explains the dramatically more erratic variations.

Figure 2, graduated values for $\alpha = \beta = 1000$, shows the most extreme variations substantially reduced, even though the point (α,β) is not very far from the no–graduation region. It is a characteristic of the l_1 method that instead of moving extreme values gradually toward a smooth function, this graduation replaces values, one by one, by linear interpolants of nearby values.

Figure 3 shows graduated values for $\alpha = 243000$, $\beta = 81000$. The function is almost bi–linear (planar). Note the discernible increase over time in the rates at younger ages. These rates are over–smoothed for many purposes, but not if our aim is to investigate the changing relation between race–specific mortality rates.

The next two figures illustrate the interaction of the two smoothness parameters. When α is large, the graduated rates are nearly linear over age, while if β is large they are nearly linear over time. However, linearity in age seems to bring a fairly smooth (though not linear) variation over time, while forcing linearity over time leaves us with an erratic pattern with respect to age. This may be in large part a quirk of the data set, which has much higher variance in the later calendar years.

Figures 6 through 9 show several reasonable graduations. Which of them is "best" depends largely on the use to be made of the results; for the moment, I only present them as examples. Note, in several of the figures, the irregularities that persist along the boundaries. The fact that the Z–scale (which was not input but determined by the program) varies from one plot to another makes comparison of the graduations a bit difficult, but it does seem that when $\alpha = \beta$ we will have linearity (and maybe over–smoothing) for the later years before we get adequate smoothing for the earlier years. This suggests that we might look at some parameter values with $\alpha \neq \beta$. In Figure 9, with $\alpha = 3000$ and $\beta = 9000$, the 1992 rates are rather smooth but convex (with respect to age); the 1984 rates are not much changed from their ungraduated values, which were not highly irregular. For intermediate years we have a moderately smooth sequence by age.

The results of Schuette graduation, whether in one or two dimensions, tend to be piecewise linear (if smoothness is based on second differences; piecewise quadratic with third differences, and so on). Some consider this a drawback of the method, preferring a smoother transition from one part of the domain to another. But this piecewise linearity is particularly appropriate if we suspect there may be changepoints, where there really is a "corner" that may be smoothed out by other graduation methods.

Presenting a single set of graduated values leaves the "consumer" without any indication of the reliability of these values. This lack can be remedied by giving upper and lower envelopes; the wider the divergence between the envelopes, the more caution is indicated in using the graduated values. The two–dimensional Schuette method can be adapted fairly easily to provide such envelopes.

Alter the problem to one of finding $\{v_{ij}\}$ to minimize

$$\Sigma\, w_{ij}\, \rho_\theta(u_{ij} - v_{ij}) + \alpha \Sigma \left| \Delta^2_{\downarrow} v_{ij} \right| + \beta \Sigma \left| \Delta^2_{\rightarrow} v_{ij} \right| , \tag{2}$$

where

$$\rho_\theta(y) = \left\{ \begin{array}{ll} 2\theta \cdot y & \text{if } y \geq 0 \\ 2(1-\theta)\left| y \right| & \text{if } y \leq 0 \end{array} \right\} = \left| y \right| + (2\theta - 1)y.$$

When $\theta = .5$ this reduces to (1), and the optimal values $\{v_{ij}\}$ provide a "median" in the sense that at most half of the weight lies above the $\{u_{ij}\}$ surface and at most half below — otherwise, simply increasing or decreasing each v_{ij} by the same small amount would decrease the fit measure without changing the smoothness measure.

For values of $\theta < .5$ the fit measure charges a higher penalty when v_{ij} exceeds u_{ij} than when it falls short by the same amount; thus, the optimal values of v_{ij} will exceed the observed data less frequently. In fact one can easily check that at most a fraction θ of the weight will be associated with points where $v_{ij} > u_{ij}$, and at most a fraction $1-\theta$ with points where $v_{ij} < u_{ij}$. This is the meaning of the term *regression quantiles*; it is the generalization to a functional situation of the familiar notion of quantiles of a single random variable.

Figures 10 and 11 show the 0.1 and 0.9 quantiles, respectively, for $\alpha = 3000$, $\beta = 9000$, the "median" for which was shown in Figure 9. The two surfaces are in close agreement for most ages in years 1984 through 1987, then diverge in years 1988 through 1992. In fact at quite a few points the "upper" and "lower" envelopes agree. Obviously this does not mean that we are absolutely certain about the correct mortality rates at those ages, only that the data do not support a significant deviation from the "median" values. Thus it is not correct to speak of the region between the two envelopes as a "confidence region" for the graduated rates. (In fact, occasionally we find a combination (α,β,i,j) for which the quantile solution $v_{ij}(\theta)$ is not monotonic in θ.) A satisfactory interpretation of the quantiles remains as a challenge.

There are several other open questions. Degeneracy (having more variables take on their basic values than the dimension of the space) becomes a problem when one or both smoothing parameters are large; we need to make the program more efficient here. We also need a better way of exhibiting the graduated values so that the eye can quickly interpret the results, at least in general terms.

Finally, note that it was my intention to use the Schuette graduation to *compare* mortality rates among blacks to those among whites, over the decade for which this data is available — thus extending the analysis in [Portnoy, 1991]. Once the technical aspects are in hand, I will do the analysis on the white data also; and I will try to get full data for the later years, rather than the 10% sample that is publicly available. It is to be expected that a smoother set of initial data will behave differently, and may raise some new questions.

References

Portnoy, E. "Recent Trends in Mortality Rates by Race," in ARCH 1991.1, 27–48.

——, "Bivariate Schuette Graduation," in ARCH 1994.1, 127–134.

Annual Statistical Supplement to *Social Security Bulletin,* for years 1983–1992.

Figure 1: Ungraduated rates. In this and all following figures, X = Age − 69; Y = Calendar year − 1983.

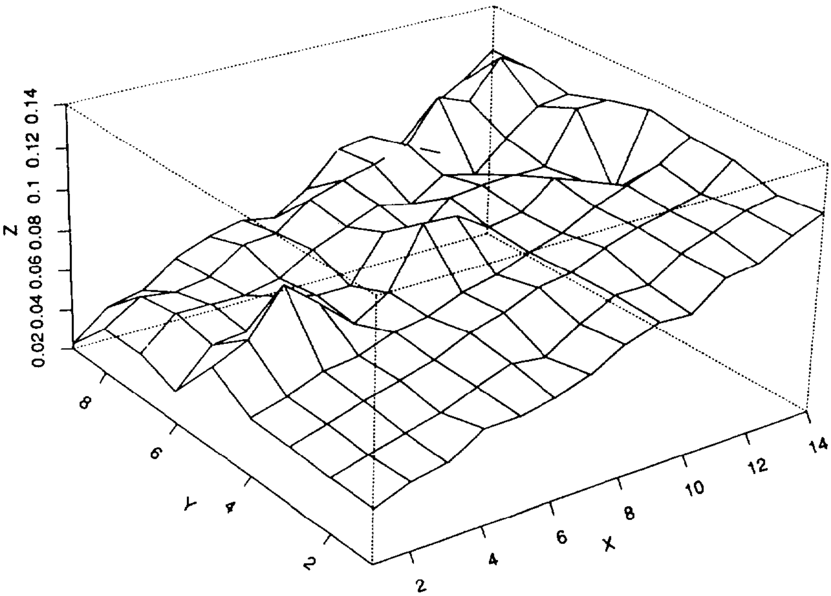

Figure 2: Graduated rates with $\alpha = \beta = 1000$.

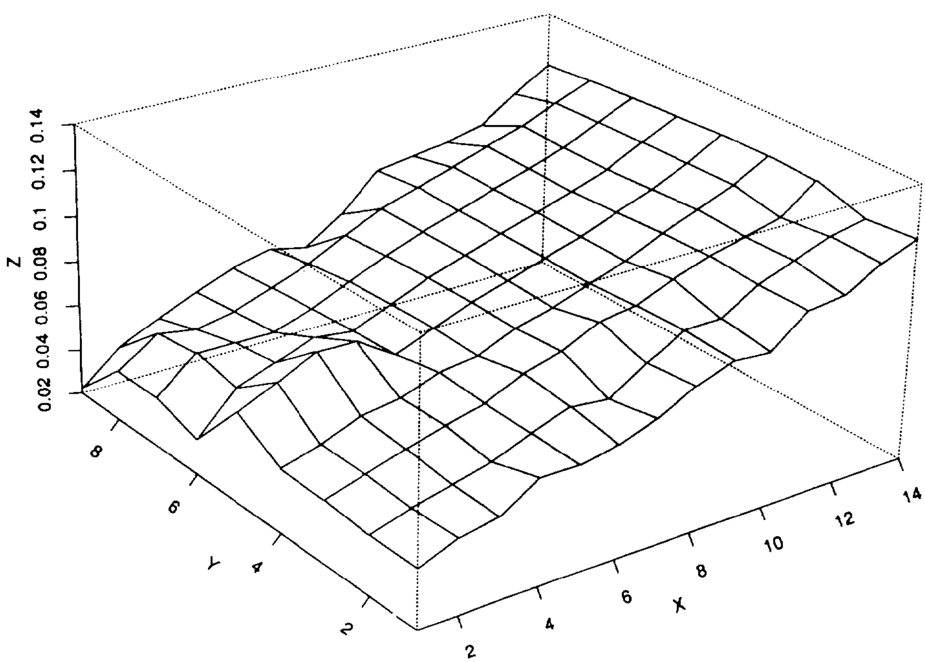

Figure 3: Graduated rates with $\alpha = 243,000$, $\beta = 81,000$.

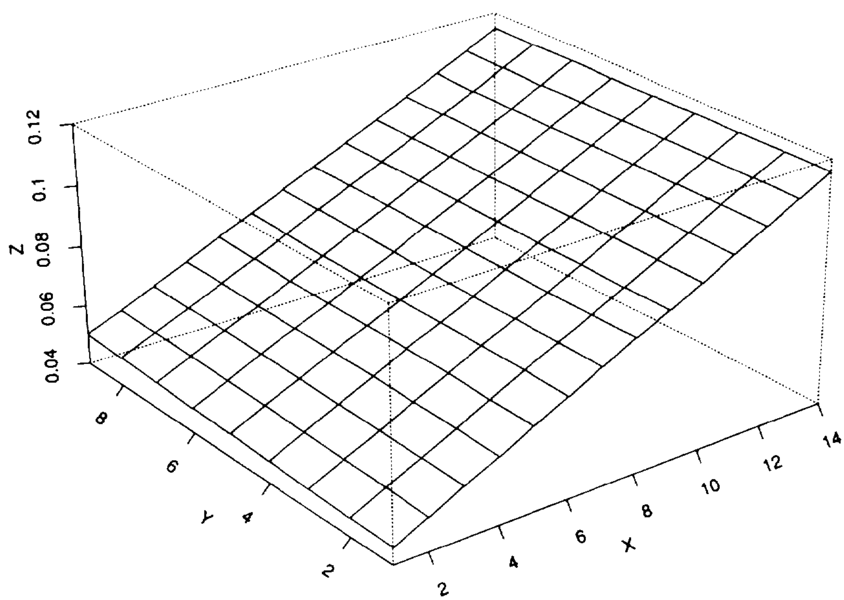

Figure 4: Graduated rates with $\alpha = 243,000$, $\beta = 1000$.

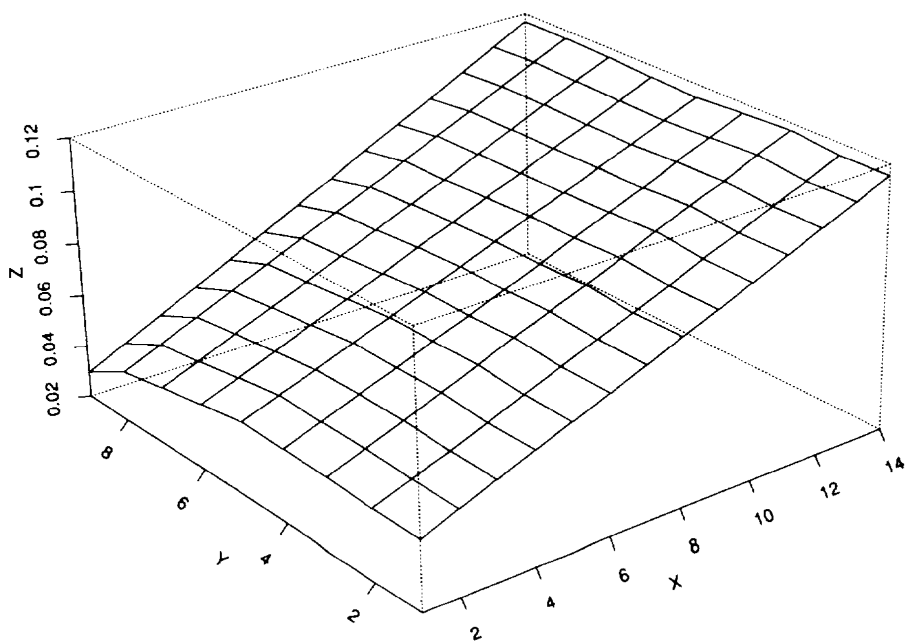

Figure 5: Graduated rates with $\alpha = 1000$, $\beta = 81{,}000$.

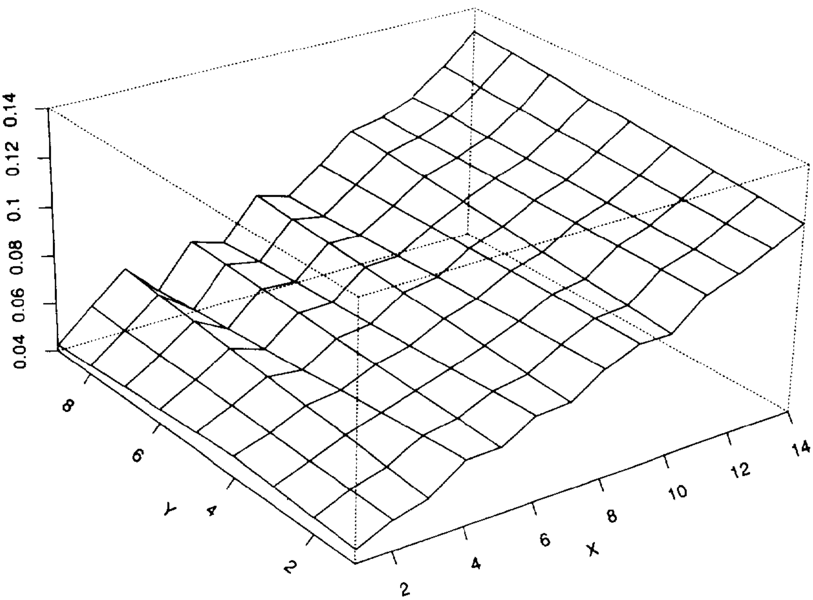

Figure 6: Graduated rates with $\alpha = \beta = 3000$.

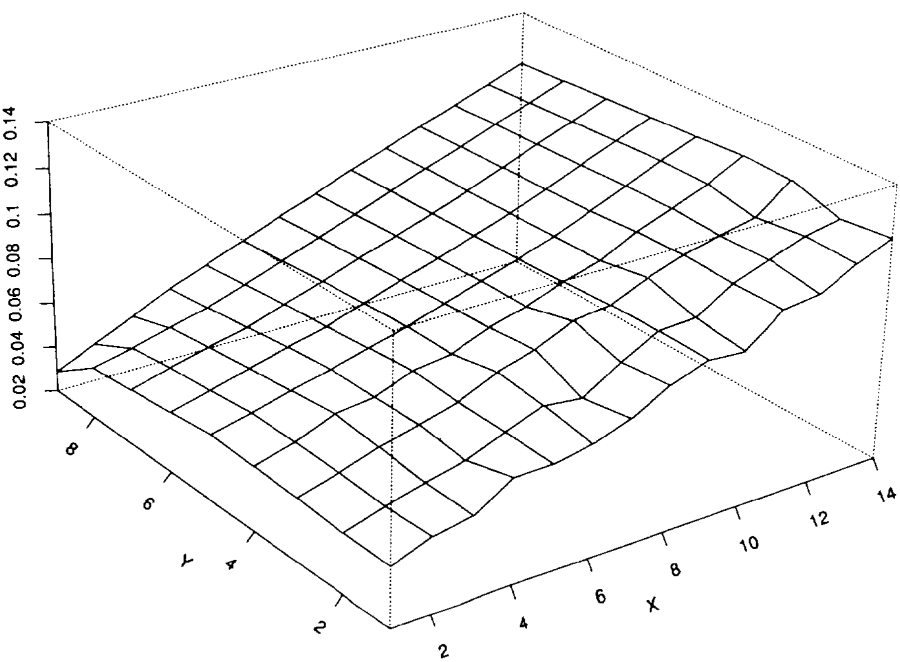

Figure 7: Graduated rates with $\alpha = \beta = 9000$.

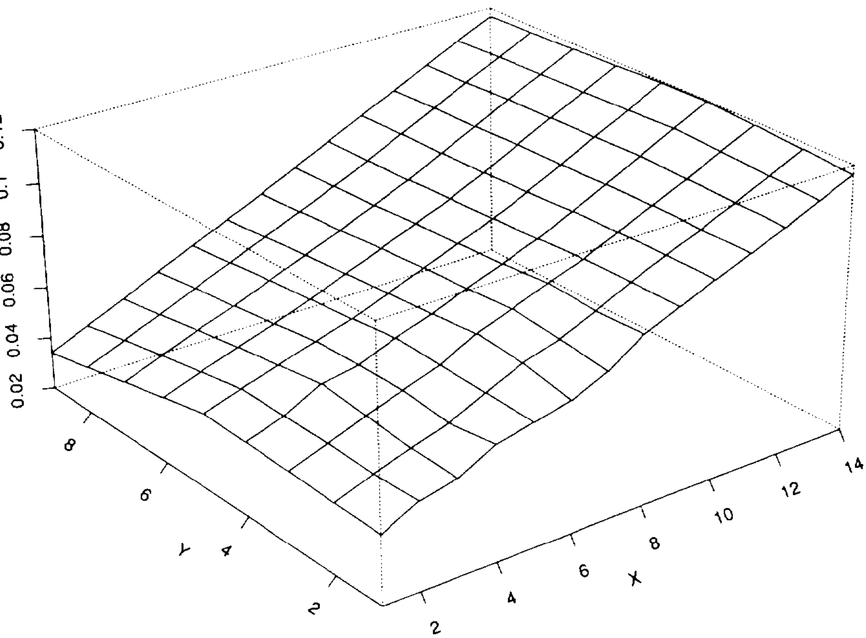

Figure 8: Graduated rates with $\alpha = \beta = 27,000$.

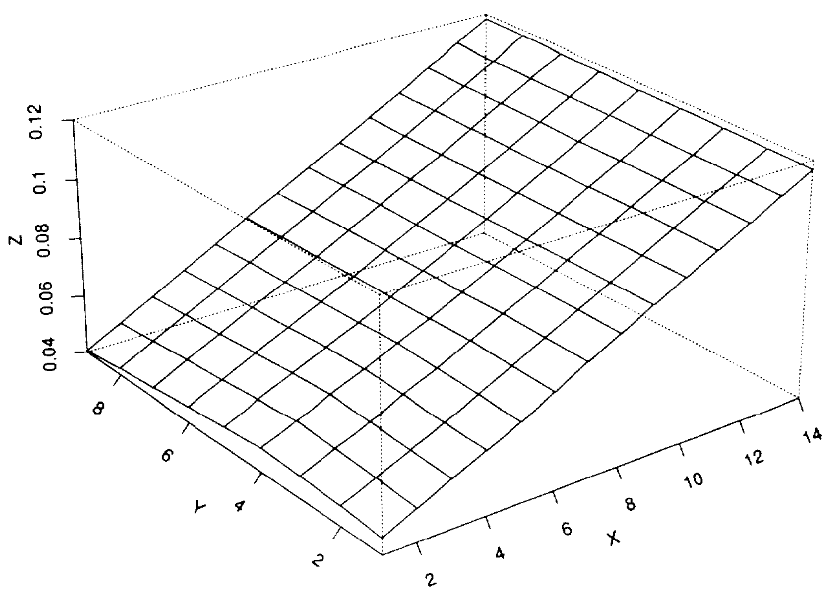

Figure 9: Graduated rates with $\alpha = 3000$, $\beta = 9000$.

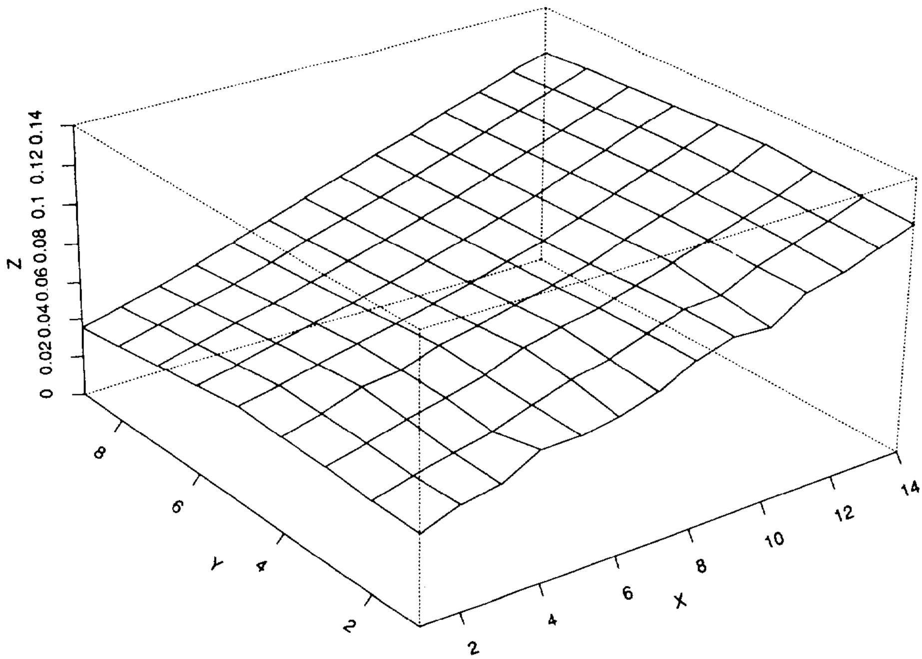

Figure 10: Lower decile ($\theta = 0.1$) for $\alpha = 3000$, $\beta = 9000$.

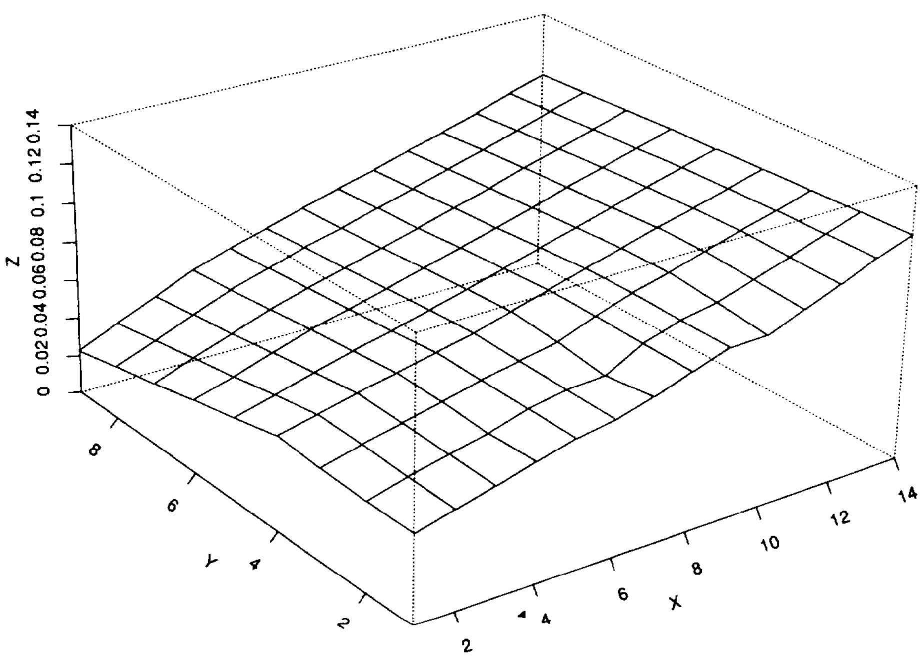

Figure 11: Upper decile ($\theta = 0.9$) for $\alpha = 3000$, $\beta = 9000$.

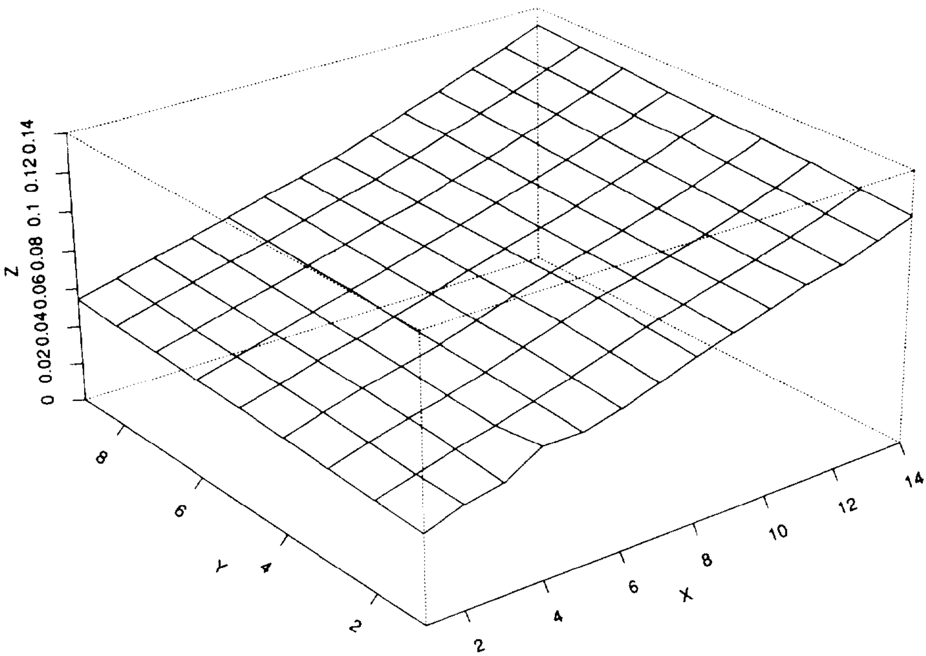

Figure 12: Figures 10 and 11 superimposed.