**Appendix 12 Future Work**

The development of a modern and more statistically accurate cubic-spline graduation method based on the B-Spline methodology for Irish crude mortality rates provides the opportunity to expand the work undertaken in this project to construct, for the first time:

1. Life Tables for the eight NUTS 2[[1]](#footnote-1) regional areas : Border, Midlands, West, Dublin, Mid-West, Mid-East South-East and South-West, and
2. the development of an Irish Public Sector Pension Mode (IPSPM)l.

**A12.1 Regional Life tables**

Work has already commenced in this area and the 2011 Life Tables for Dublin and the South are presented Tables A9.1, A9.2, A9.3 and A.9.3. This work has the potential to inform both EU and national regional health policies.

In addition, the shape of the curves are compared in *section A9.2*.

**A12.2 Curve matching Hausdorff and Fréchet distances**

Shape matching is an important area of research and two widely use measures are the Hausdorff and Fréchet distances.

Consider 2 point sets and , where the goal is to find the extent of similarity between and . The first question that needs to be answered is: what do one mean by “similar”? Mathematically, one wants to find a mapping , that has one of the following properties:

* minimizes (minimizes maximum distance between mapped points) or,
* minimizes (minimizes sum of distances between mapped points) or,
* minimizes (minimizes sum of squared distances between mapped points)

The mapping does not have be one-to-one. The question is: how does one find ?

**A12.2.1 Hausdorff Distance**

Let and be two sets of points in .

The *directed Hausdorff* distance from to , denoted by , is .

The Hausdorff distance between P and Q, denoted by , is .

Intuitively, the function finds the point that is farthest from any point in and measures the distance from to its nearest neighbour in . *Hausdorff distance is a measure of the mismatch between two point-sets*. For , the Hausdorff distance can be computed in time (where is the number of points), using a *Voronoi* diagram in . In computing a *Voronoi* diagram could take quadratic time, so a different approach is needed to compute in sub-quadratic time.

Let be the disk of radius centred at point . The decision problem for the Hausdorff distance can be written as:

Given , whether ? That is, , .

A point can be mapped to a point and a disk can be mapped to a half-space in such that if and only if . Hence, iff , i.e.,

A point-location data structure for convex polyhedra can be used to answer the decision query. When and are sets of elements other than points (features, for example), a similar formulation is possible.

While the Hausdorff distance is an appropriate measure in many applications, the following figure shows an example where it is not. The two curves have a small Hausdorff distance, but do not look like each other at all.

**Figure A12.1:** Problem with Hausdorff distance

The reason of this discrepancy is that the Hausdorff distance only takes into account the sets of points on both curves and does not reflect the course of the curves.

However, the course of curves is important in many applications, for example in handwriting recognition. In order to overcome this problem, one can use the Fréchet distance.

**A12.2.2 Fréchet distance**

The *Fréchet* distance is a measure that takes the continuity of shapes into account and, hence, is better suited than the *Hausdorff* distance for curve or surface matching. This distance function was introduced by Fréchet (1906).

A popular illustration of the *Fréchet* distance is, as follows. Suppose a man is walking a dog. Assume the man is walking on one curve and the dog on another curve. Both can adjust their speeds but are not allowed to move backwards. The Fréchet distance of the two curves is then the minimum length of leash necessary to connect the man and the dog.

The Fréchet distance between two curves is defined as follows:

where

are parameterizations of the two curves and

range over all continuous and monotone increasing functions. If and are polygonal chains with and line segments respectively, the decision problem for the Fréchet distance can be written as: whether

One can think of the parameter as "time". Then, is the position of the dog and is the position of the dog's owner at time (or vice versa). The length of the leash between them at time is the distance between and . Taking the infimum over all possible reparametrisations of corresponds to choosing the walk along the given paths where the maximum leash length is minimized. The restriction that and be non-decreasing means that neither the dog nor its owner can backtrack.

The Fréchet distance takes into account the ﬂow of the two curves because the pairs of points whose distance contributes to the Fréchet distance sweep continuously along their respective curves. This makes the Fréchet distance a better measure of similarity for curves than alternatives, such as the Hausdorff distance, for arbitrary point sets. It is possible for two curves to have small Hausdorff distance but large Fréchet distance.

As an illustrative example, consider a basic problem where and be two polygonal chains. Define a mapping from the vertices of to those of Q such that:

1. if and and (monotonicity)
2. is minimum

Figure A11.2.: Pair of curves and *Fréchet* distance.

A fundamental study on the computational properties of the *Fréchet* distance was done by Alt and Godau (1992). They give an algorithm that computes the exact *Fréchet* distance between two polygonal curves in time , where and are the number of segments on the polygonal curves. The algorithm uses the parametric search technique.

*Table A9* presents the estimates of the Hausdorff and Fréchet distances for Dublin and the South West Regions for both males and females curves.

|  |  |  |
| --- | --- | --- |
| **Table A12.1: Hausdorff & Fréchet distances, Dublin and the South West regions, by gender** | | |
| **Regions** | Hausdorff distance | *Fréchet* distance |
| **Males** | | |
| South West vs. Total | 0.1504769 | 0.2475473 |
| Dublin vs. Total | 0.1561183 | 0.1665361 |
| Dublin vs. South West | 0.1024405 | 0.2075317 |
| **Females** | | |
| South West vs. Total | 0.1260543 | 0.3106847 |
| Dublin vs. Total | 0.1444131 | 0.1525562 |
| Dublin vs. South West | 0.2704674 | 0.3082828 |

Hausdorff and Fréchet distances were calculated using the associated functions in the Pracma and SimilarityMeasures R-packages.

If curves are identical then the returned distance values will equal zero. This is not the case in any of the comparisons here. Also, the Fréchet in each comparison is larger than the associated Hausdorrff distance.

For both genders: Dublin is closer in shape to the Total curve that of the South West and this is also seen when the Dubiln vs. the South West comparison is undertaken.

1. NUTS - [Nomenclature of Territorial Units for Statistics](https://en.wikipedia.org/wiki/Nomenclature_of_Territorial_Units_for_Statistics) [↑](#footnote-ref-1)