

Casos de estudio de estimación de funciones de supervivencia para Ciencias Actuariales

Antonio Fernández Morales

Departamento de Economía Aplicada - Estadística y Econometría
Universidad de Málaga. 2022



Esta obra se encuentra bajo una Licencia Creative Commons Reconocimiento-NoComercial-SinObraDerivada. Puede copiar, distribuir y comunicar públicamente la obra bajo las condiciones siguientes:

- Reconocimiento: Debe reconocer los créditos de la obra citando al autor.
- No comercial: No puede utilizar esta obra para fines comerciales.
- Sin obras derivadas: No se puede alterar, transformar o generar una obra derivada a partir de esta obra.

1. Introducción

Las particularidades técnicas que plantea la estimación de funciones de supervivencia, especialmente en lo tocante a la presencia de datos censurados, demandan una notable atención, tanto en el ámbito actuarial como en otros relacionados [1], [2], [3]. Las técnicas propuestas por Kaplan y Meier [4], Nelson y Aalen [5], [6] o Fleming y Harrington [7] se han convertido en herramientas de uso común en este contexto.

Por otra parte, las técnicas no paramétricas de estimación mediante funciones kernel también son ampliamente utilizadas en diversas aplicaciones actuariales, incluyendo la graduación de tablas de mortalidad y supervivencia [8] y la predicción del riesgo de mortalidad [9].

A los instrumentos analíticos anteriores podemos añadir, además, los modelos de riesgos proporcionales y de Cox, que se constituyen en una alternativa muy eficaz a la estratificación de las carteras para estimar las funciones de supervivencia y estudiar características de las poblaciones con posibles efectos sobre el riesgo de mortalidad que evita la fragmentación de la experiencia [1].

El conjunto de casos de estudio que se incluyen en esta obra, relativos a la estimación de funciones supervivencia, se ha elaborado con el objeto de facilitar la formación actuarial mediante un enfoque aplicado que incentive el desarrollo de las competencias profesionales demandadas por el sector [10], [11], [12], [13], su considerable grado de multidisciplinariedad [14], [15] y los retos que afronta tras la creciente importancia de la analítica predictiva y el *bigdata* [16].

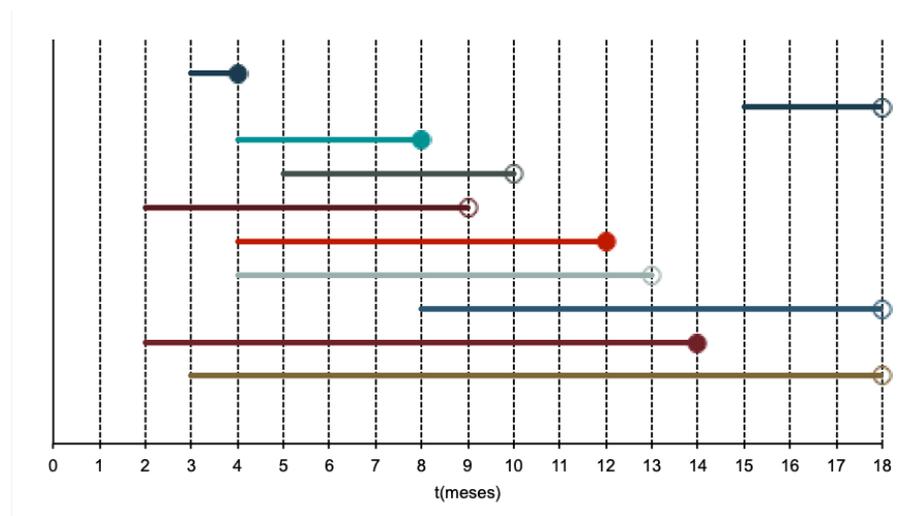
Las características concretas de los casos elaborados permiten su uso para la realización de proyectos individuales o grupales, así como para su inclusión en modelos de enseñanza virtual [17], [18], [19], [20], aptos incluso para experiencias motivadas por circunstancias similares a las acaecidas durante la pandemia de la COVID-19 que obligaron a cambios drásticos en los modelos de aprendizaje y evaluación [21], [22], [23].

Finalmente, este trabajo se ha realizado dentro de un conjunto más amplio de materiales para la formación actuarial [24], [25], que incluye un conjunto de actividades interactivas que pueden servir de complemento en una estrategia de aprendizaje más completa. Estas actividades facilitan la experimentación y simulación de modelos de supervivencia [26], [27], [28], [29], de procesos estocásticos, [30], [31], [32], [33], [34] y de elaboración de tablas de supervivencia dinámicas, [35], [36], [37].

2. Casos de estudio

Caso 1

Los datos de 10 individuos sometidos a un tratamiento tras una intervención para evitar la recaída de una enfermedad y observados en un periodo de 18 meses aparecen en la figura. Los círculos llenos indican recaída y los círculos vacíos indican censura derecha.



- Construir la tabla de duraciones observadas y tipo de salida del estudio (recaída o censura).
- Estimar la función de supervivencia mediante el estimador de Kaplan-Meier.
- Estimar la función de riesgo acumulada con el estimador de Nelson-Aalen.
- Obtener una estimación de la función de supervivencia basada en el estimador de Fleming-Harrington.

t_i	d_i	n_i	$\frac{t_i}{d_i}$	$\hat{S}_{KM}(t_i)$	$\hat{H}_{NA}(t_i)$	$\hat{S}_{FH}(t_i)$
-------	-------	-------	-------------------	---------------------	---------------------	---------------------

Caso 2

Una compañía de seguros ha encargado un estudio acerca de la supervivencia con un tratamiento experimental tras un tipo de intervención quirúrgica muy severa. Se ha administrado este tratamiento a 15 pacientes y se va a realizar el seguimiento hasta el octavo fallecimiento que se produzca. El equipo que ha realizado el estudio ha presentado la siguiente información tras 60 meses:

- Un paciente falleció a los 10 meses
 - Tres pacientes fallecieron a los 20 meses
 - Tres pacientes fallecieron a los 40 meses
 - Un paciente falleció a los 60 meses
 - La probabilidad de supervivencia, obtenida con el estimador de Kaplan-Meier es igual a 0,4286.
 - Un paciente abandonó el estudio a los 15 meses.
- a) Calcule la estimación de Kaplan-Meier de la función de supervivencia de la duración a los 60 meses y verifique si coincide con la proporcionada por equipo investigador.
- b) Explique si la estimación ofrecida por el equipo investigador puede estar influida por el tratamiento estadístico (¿dato perdido o dato censurado?) del paciente que abandonó el estudio.

Caso 3

Una importante compañía de seguros ha realizado una investigación sobre la mortalidad de sus aseguradas. Durante el estudio se ha seguido a 50 aseguradas desde que cumplieron 70 años hasta el primero en ocurrir de estos tres sucesos: muerte, cancelación de la póliza o cumplir 75 años exactos. Las edades en las que las aseguradas del estudio murieron o cancelaron sus pólizas se muestran a continuación:

Muerte	Cancelación
70 y 5 meses	70 y 2 meses
72 y 6 meses	70 y 3 meses
73 y 8 meses	71 y 0 meses
73 y 8 meses	71 y 0 meses
74 y 2 meses	72 y 9 meses
	73 y 5 meses
	74 y 3 meses

- Explique qué tipos de censuras hay en este estudio.
- Calcule el estimador de Kaplan-Meier de la función de supervivencia de estas aseguradas.
- Dibuje la función de supervivencia estimada.
- Estime el error estándar de $\hat{S}(2, 5)$ con la fórmula de Greenwood.
- Estime intervalos de confianza de $1 - \alpha = 0,95$ y represéntelos gráficamente.

Caso 4

Para evaluar la eficacia de un nuevo tratamiento experimental para una enfermedad recurrente que requiere baja laboral, se ha realizado un estudio durante 24 meses con 30 pacientes que están en proceso de recuperación. El objetivo del estudio es tratar de conocer si con el tratamiento experimental administrado después de recibir el alta médica se reduce la probabilidad de volver a recaer, necesitando una nueva baja laboral.

Se ha dividido a estos pacientes en dos grupos: (i) 15 pacientes siguieron el tratamiento experimental tras recibir el alta médica y (ii) 15 pacientes no siguieron ningún tratamiento tras recibir el alta médica, como grupo de control.

Los datos (en meses) del tiempo en recaer tras recibir el alta médica se muestran a continuación (un * indica un dato censurado):

Pacientes que siguen el tratamiento experimental:

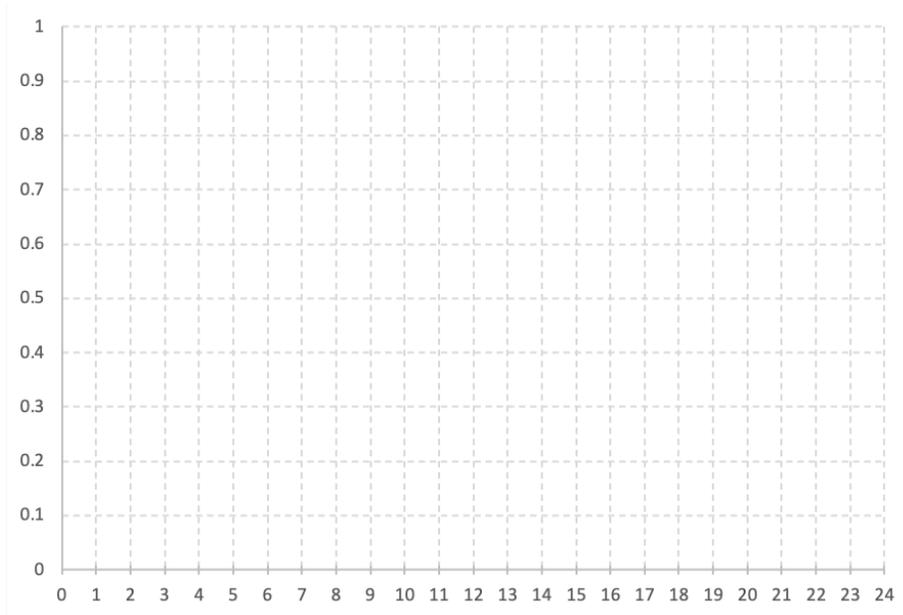
5, 7*, 9, 11*, 12*, 13*, 15, 16*, 17*, 18, 19, 20, 23*, 23*, 23*

Grupo de control:

2, 3, 5*, 6*, 7, 8, 9*, 9*, 9*, 10, 12, 14*, 16, 21, 23*

- a) Calcule, para ambos grupos, el estimador de Kaplan-Meier de la función de supervivencia del tiempo sin recaer tras recibir el alta médica.
- b) Compare los resultados e indique, sin realizar cálculos, cómo podría comprobar si las diferencias los resultados de ambos grupos es estadísticamente significativa.

- c) Dibuje la función de supervivencia $\hat{S}(t)$ estimada con el método de Kaplan-meier, etiquetando con claridad los ejes.



Caso 5

La tabla adjunta muestra los datos de mortalidad del colectivo de mujeres de la población general de un país europeo, para las edades 20 a 50 años. Realice una graduación de las tasas brutas de mortalidad para las edades 33 y 34 mediante un procedimiento no paramétrico usando el kernel de Epanechnikov (utilice el estimador de Nadaraya-Watson [38], [39] y una ventana $b=4$).

x	D_x	E_x	x	D_x	E_x
20	10	47908	36	31	71072
21	5	48381	37	33	70456
22	8	49501	38	41	69543
23	12	50363	39	32	68895
24	5	51382	40	35	68386
25	12	52600	41	47	67770
26	13	54000	42	46	67722
27	10	56093	43	52	67953
28	12	58264	44	82	68870
29	18	60689	45	73	68103
30	15	63151	46	76	66896
31	13	65680	47	92	67761
32	20	67618	48	102	66445
33	33	69040	49	107	62577
34	22	70534	50	103	59960
35	28	71146			

Caso 6

Utilizando la información relativa al colectivo de mujeres de la población general de un país europeo de la tabla, realice una graduación no paramétrica para la edad 39 años, usando el kernel gaussiano, con ventana $b=2$, mediante el estimador de Copas-Haberman [40], [41].

x	D_x	E_x	x	D_x	E_x
20	7	48478	41	59	67876
21	6	49450	42	60	68080
22	7	50387	43	67	68943
23	12	51401	44	75	68229
24	5	52778	45	73	67007
25	17	54233	46	89	67940
26	11	56284	47	88	66553
27	14	58494	48	92	62733
28	15	60835	49	92	60104
29	20	63263	50	93	59425
30	24	65791	51	136	58613
31	19	67789	52	115	57236
32	14	69091	53	124	55940
33	23	70612	54	125	52409
34	28	71269	55	113	49276
35	25	71233	56	159	47413
36	37	70580	57	147	46131
37	36	69614	58	143	45781
38	41	69026	59	174	43928
39	35	68444	60	151	40816
40	44	67859			

Caso 7

Se ha realizado una investigación para averiguar el efecto diferencial entre varios grupos poblacionales de un nuevo tratamiento farmacológico para aliviar los síntomas más severos de una enfermedad crónica. Uno de los resultados es la estimación de un modelo de Cox para la función de riesgo $h(t, \mathbf{z})$ hasta la desaparición significativa de los síntomas desde el comienzo del tratamiento (t en días).

$$h(t, \mathbf{z}) = h_0(t)e^{\alpha_1 z_1 + \alpha_2 z_2 + \alpha_3 z_3}$$

El modelo utiliza tres variables explicativas que dependen de las características de cada persona incluida en el estudio:

$z_1 = 0$ si es hombre, 1 si es mujer

$z_2 = 1$ si realiza ejercicio físico regularmente, 0 en caso contrario

$z_3 =$ edad (en años)

Además, conocemos la siguiente información:

- Se ha estimado que para un hombre de 60 años que realiza ejercicio regularmente el riesgo de desaparición de los síntomas es un 50 % mayor que el de un hombre de 20 años que realiza ejercicio físico con regularidad.
- Se ha estimado que para una mujer de 50 años que realiza ejercicio regularmente el riesgo de desaparición de los síntomas es 2 veces el de una mujer de 40 años que no realiza ejercicio físico con regularidad.
- Se ha estimado que para una mujer de 40 años que realiza ejercicio regularmente el riesgo de desaparición de los síntomas es 2 veces el de un hombre de la misma edad que no realiza ejercicio físico con regularidad.

- a) Calcule el valor de los parámetros α_1 , α_2 y α_3 .
- b) Calcule la probabilidad de que los síntomas persistan al menos 30 días desde el comienzo del tratamiento en una mujer de 30 años que realiza ejercicio regularmente, sabiendo que esta probabilidad se estima en 0,6 para un hombre de 40 años que no realiza ejercicio regularmente.

Caso 8

Se ha estimado un modelo de riesgos proporcionales con *baseline* de Gompertz a una cartera de pensionistas de Reino Unido [1] de edades (x) iguales o superiores a 60 años.

$$\mu(x, \mathbf{z}) = AB^x e^{\alpha_1 z_1 + \alpha_2 z_2}$$

Las variables explicativas del modelo son:

$z_1 = 0$ si es mujer, 1 si es hombre

$z_2 = 1$ si la cuantía de la pensión es £10000 o mayor, 0 en caso contrario

Los parámetros estimados son:

$$A = 0,00000299969 \quad B = 1,126154 \quad \alpha_1 = 0,377154 \quad \alpha_2 = -0,238433$$

- a) Represente gráficamente para las edades 60 a 100 años en tanto instantáneo de mortalidad para los cuatro segmentos de la población especificados.
- b) Calcular la reducción proporcional del tanto instantáneo de mortalidad del colectivo de pensiones superiores a £10000 estimada en el modelo respecto al resto de la población investigada.

Referencias

- [1] McDonald, A. S., Richards, S. J., Currie, I. D. (2018). *Modelling Mortality with Actuarial Applications*. Cambridge: Cambridge University Press.
- [2] Bladt, M., Albrecher, H., Beirlant, J. (2020). Combined tail estimation using censored data and expert information. *Scandinavian Actuarial Journal*, 6, 503-525.
- [3] Leung, A. (2021). *Actuarial Principles: Lifetables and Mortality Models*. Academic Press.
- [4] Kaplan, E. L., Meier P. (1958). Nonparametric Estimation from Incomplete Observations. , *Journal of the American Statistical Association* 53 (282), 457-481.
- [5] Nelson, W. (1972). Theory and applications of hazard plotting for censored failure data. *Technometrics* 14, 945–965.
- [6] Aalen, O. (1978). Nonparametric inference for a family of counting processes. *Annals of Statistics* 6, 701–726.
- [7] Fleming, T. R., Harrington, D. P. (1984). Nonparametric Estimation of the Survival Distribution in Censored Data. *Communications in Statistics. Theory and Methods* 13, 2469–2486.
- [8] Chang, L., Shi, Y. (2020). Dynamic modelling and coherent forecasting of mortality rates: a time-varying coefficient spatial-temporal autoregressive approach. *Scandinavian Actuarial Journal* 9, 843-863.
- [9] Gao, G., Shi, Y. (2021). Age-coherent extensions of the Lee–Carter model. *Scandinavian Actuarial Journal* 10, 998-1016.
- [10] Trigo-Martínez, E., Fernández-Morales, A: (2017). Collaborative projects for developing technological and professional competences in Actuarial Science, *EDULEARN17 Proceedings, 9th International Conference on Education and New Learning Technologies*, Valencia: IATED, pp. 2767-2772.
- [11] Fernández-Morales, A., Trigo-Martínez, E., Moreno-Ruiz, R., Gómez-Pérez-Cacho, O. (2019). Involving professionals in curricular activities for developing actuarial skills. *INTED2019 Proceedings*, pp. 9113-9117, 2019. doi: <http://dx.doi.org/10.21125/inted.2019.2267>
- [12] Mayorga-Toledano, M. C., Fernández-Morales, A. (2018). Coordination, transversality and professional skills in actuarial education. *11th*

- annual International Conference of Education, Research and Innovation, ICERI2018 Proceedings*, pp. 2714-2719, Sevilla: IATED. doi: <http://dx.doi.org/10.21125/iceri.2018.1601>
- [13] Mayorga-Toledano, M. C., Fernández-Morales, A., Moreno-Ruiz, R. (2020). Actions for the contextualization and development of professional competences. *12th Annual International Conference on Education and New Learning Technologies, EDULEARN20 Proceedings*, pp. 5379-5385. doi: <http://dx.doi.org/10.21125/edulearn.2020.1417>
- [14] Fernández-Morales, A., Trigo-Martínez, E., Gómez Pérez-Cacho, O. (2020). Multidisciplinary collaborative projects for a more coordinated Actuarial Education. *12th annual International Conference on Education and New Learning Technologies, EDULEARN20 Proceedings*, pp. 5409-5414. doi: <http://dx.doi.org/10.21125/edulearn.2020.1424>
- [15] Fernández Morales, A., Mayorga Toledano, M.C. (2013). Developing Creativity and Innovation through Collaborative Projects. *Interdisciplinary Studies Journal 2* (3), pp. 70-82.
- [16] Mayorga Toledano, M. C. (2021). Limitaciones legales de la analítica predictiva y el big data en el ámbito asegurador. *Derecho de seguros: nuevas realidades y nuevos retos*. Marcial Pons, 313-327.
- [17] Mayorga-Toledano, M. C., Trigo-Martínez, E., Fernández-Morales, A. (2022). Assessing the students' prospective perceptions of the final master project in Actuarial Science. *15th annual International Conference of Education, Research and Innovation, ICERI2022 Proceedings*. Sevilla: IATED, 2877-2882.
- [18] Mayorga-Toledano, M. C. (2010). Integrating e-learning activities in the teaching and learning of banking and securities market law. *Education and Law Review*, vol. 1, pp. 177-196.
- [19] Mayorga-Toledano, M. C., Fernández-Morales, A. (2004). Learning tools for java enabled phones. An application for actuarial studies. In *Learning with mobile devices. Research and Development* (J. Attewell, C. Savill-Smith, cords.), London: Learning and Skills Development Agency, pp. 95-98.
- [20] Mayorga-Toledano, M. C., Fernández-Morales, A. (2010). Using iPhone Web-Apps to Enhance Learning and Teaching in Actuarial Education, In *Mobile Learning: Pilot Projects and Initiatives* (R. Guy, ed.), Santa Rosa: Informing Science Press, pp. 83-104.
- [21] Mayorga-Toledano, M. C., Fernández-Morales, A. (2020). Assessing the face-to-face to virtual learning adaptation process of the MSc in Actuarial

- Science at the University of Malaga. *13th International Conference of Education, Research and Innovation, ICERI2020 Proceedings*, pp. 6232-6240.
- [22] Fernández-Morales, A., Mayorga-Toledano, M. C. (2020). Using COVID19 outbreak data to engage students in the learning process. *13th International Conference of Education, Research and Innovation, ICERI2020 Proceedings*, pp. 6205-6211.
- [23] Mayorga-Toledano, M. C., Fernández-Morales, A. (2021). Students' perceptions of the teaching and learning mode adopted in the MSc in Actuarial Science at the University of Malaga during COVID-19 first wave. *115th International Technology, Education and Development Conference, INTED2021 Proceedings*, pp. 9650-9655.
- [24] Fernández-Morales, A. (2022). Casos de estudio de modelos de supervivencia para Ciencias Actuariales. RIUMA, Universidad de Málaga, <https://hdl.handle.net/10630/25377>.
- [25] Fernández-Morales, A. (2022). Casos de estudio de tablas de mortalidad y supervivencia para Ciencias Actuariales. RIUMA, Universidad de Málaga, <https://hdl.handle.net/10630/25501>.
- [26] Fernández-Morales, A. (2011). Learning survival models with on-line simulation activities in the Actuarial Science Degree. *International Journal of Emerging Technologies in Learning*, Vol. 6, no. 1, pp. 15–19.
- [27] Fernández-Morales, A. (2017). Simulating lifetimes with actuarial survival models. *9th International Conference on Education and New Learning Technologies, EDULEARN17 Proceedings*, pp. 725-731, Valencia: IATED, 2017. <http://dx.doi.org/10.21125/edulearn.2017.1161>
- [28] Fernández-Morales, A., Mayorga-Toledano, M. C. (2021). Using an on-line interactive graphical simulator to experiment with late-life mortality models. *115th International Technology, Education and Development Conference, INTED2021 Proceedings*, pp. 9656-9662.
- [29] Fernández-Morales, A. (2020). Micro simulación de modelos de supervivencia en edades avanzadas para Ciencias Actuariales. RIUMA, Universidad de Málaga, <https://hdl.handle.net/10630/20543>.
- [30] Fernández-Morales, A. (2017). Enriching the statistics learning experience with D3.js interactive animations: Insurance applications of Markov chains. *International Journal of Educational Research and Innovation*, vol. 7, pp. 25-39, 2017. URI: <http://hdl.handle.net/10433/4923>

- [31] Fernández-Morales, A. (2015). Application of a Discrete-time Markov Chain Simulation in Insurance. *International Journal of Recent Contributions from Engineering, Science & IT*, vol.3, no.3, pp. 27-32, 2015. doi: <http://dx.doi.org/10.3991/ijes.v3i3.4929>
- [32] Mayorga-Toledano, M. C., Fernández-Morales, A. (2019). Enhancing Actuarial education with interactive online resources. *13th International Technology, Education and Development Conference, INTED2019 Proceedings*, pp. 9139-9145, 2019. doi: <http://dx.doi.org/10.21125/inted.2019.2270>
- [33] Mayorga-Toledano, M. C., Fernández-Morales, A. (2017). Interactive resources based on serious gaming for Actuarial Education, *EDULEARN17 Proceedings, 9th International Conference on Education and New Learning Technologies*, Valencia: IATED, pp. 3245-3251.
- [34] Fernández-Morales, A., Mayorga-Toledano, M. C. (2018). Using serious games un Higher Education. An Application in actuarial studies. *11th annual International Conference of Education, Research and Innovation, ICERI2018 Proceedings*, pp. 2727-2734, Sevilla: IATED. doi: <http://dx.doi.org/10.21125/iceri.2018.1603>
- [35] Fernández-Morales, A. (2021). Tablas actuariales de supervivencia y mortalidad dinámicas con hoja de cálculo. RIUMA, Universidad de Málaga, <https://hdl.handle.net/10630/23265>.
- [36] Fernández-Morales, A. (2016). Tablas de mortalidad dinámicas con hoja de cálculo en la práctica actuarial. RIUMA, Universidad de Málaga, <http://hdl.handle.net/10630/10922>.
- [37] Fernández-Morales, A. (2016). Tutorial para la construcción de tablas de mortalidad dinámicas PERM/F 2000 con hoja de cálculo. RIUMA, Universidad de Málaga, <http://hdl.handle.net/10630/5682>.
- [38] Selingerova, I., Katina, S., Horova, I. (2021). Comparison of parametric and semiparametric survival regression models with kernel estimation. *Journal of Statistical Computation and Simulation*, 91:13, 2717-2739, DOI: 10.1080/00949655.2021.1906875
- [39] Peristera, P., Kostaki, A. (2005). An evaluation of the performance of kernel estimators for graduating mortality data. *Journal of Population Research* 22, 185–197. <https://doi.org/10.1007/BF03031828>
- [40] Copas, J., Haberman, S. (1983). Non-parametric graduation using kernel methods. *Journal of the Institute of Actuaries*, 110, 135–156.

- [41] Gavin, J., Haberman, S., Verrall, R. (1993). Moving weighted average graduation using kernel estimation. *Insurance: Mathematics and Economics*, 12(2), 113-126.