

DOCUMENTO METODOLÓGICO

Designação da operação estatística: Projeções de População Residente

Sigla da operação estatística:

Código da operação estatística: 115

Código da atividade estatística - CGA*: 244 Projeções Demográficas

Código de versão do DMET: 4.0

Data de entrada em vigor da versão do DMET: Março 2020

Data da última atualização do DMET: Março 2020

Entidade responsável pela operação estatística: INE/DES/DM

* - Para uso exclusivo nas operações da responsabilidade do INE.

ÍNDICE

I. IDENTIFICAÇÃO DA OPERAÇÃO ESTATÍSTICA

- ☞ I. 1 Designação da operação estatística
- ☞ I. 2 Sigla (ou abreviatura) da operação estatística
- ☞ I. 3 Código da operação estatística
- ☞ I. 4 Código SIGINE (modelo estatístico)
- ☞ I. 5 Código da Atividade Estatística
- ☞ I. 6 Código de Versão do Documento Metodológico
- ☞ I.7 Data de entrada em vigor da versão do Documento Metodológico
- ☞ I.8 Data da última atualização do Documento Metodológico
- ☞ I.9 Entidade responsável pela operação estatística
- ☞ I10. *Outras Entidades Externas relacionadas com a operação estatística*

II. JUSTIFICAÇÃO PARA UMA NOVA VERSÃO DO DOCUMENTO METODOLÓGICO

III. IDENTIFICAÇÃO DAS NECESSIDADES, OBJETIVOS E FINANCIAMENTO

- ☞ III.1 Contexto da operação estatística
- ☞ III.2 Identificação das necessidades de informação estatística que justificam a operação
- ☞ III.3 Objetivos da operação estatística
- ☞ III.4 Financiamento da operação estatística

IV. CARACTERIZAÇÃO GERAL

- ☞ IV.1 Tipo de operação estatística
- ☞ IV.2 Tipo de fonte (s) de informação utilizada (s) na operação estatística
- ☞ IV.3 Periodicidade da operação estatística
- ☞ IV.4 Âmbito geográfico da operação estatística
- ☞ IV.5 Principais utilizadores da informação
- ☞ IV.6 Difusão
 - IV.6.1 Padrão de disponibilização da informação
 - IV.6.2 Revisões
 - IV.6.3 Produtos de difusão regular

V. CARACTERIZAÇÃO METODOLÓGICA

- ☞ V.1 População-alvo
- ☞ V.2 Base de amostragem
- ☞ V.3 Unidade (s) estatística (s) de observação
- ☞ V.4 Desenho da amostra
 - V.4.1 Características da amostra
 - V.4.2. Metodologia para a seleção, distribuição, atualização e dimensionamento da amostra
- ☞ V.5 Construção do (s) questionário (s)
 - V.5.1 Testes de pré-recolha efetuados ao (s) questionário (s)
 - V.5.2 Tempo médio de preenchimento do (s) questionário (s)
- ☞ V.6 Recolha de dados
 - V.6.1 Recolha direta de dados
 - V.6.1.1 Período (s) de recolha
 - V.6.1.2 Método (s) de recolha
 - V.6.1.3 Critério para o fecho da recolha
 - V.6.1.4 Possibilidade de inquirição Proxy
 - V.6.1.5 Sessões Informativas
 - V.6.2 Recolha não-direta de dados
- ☞ V.7 Tratamento de dados

- V. 7.1 Validação e análise
- V.7.2 Tratamento de não respostas
- V.7.3 Obtenção de resultados
- V.7.4 Ajustamentos dos dados
- V.7.5 Comparabilidade e coerência
- V.7.6 Confidencialidade dos dados

VI. SUPORTES DE RECOLHA E VARIÁVEIS DE OBSERVAÇÃO

VII. VARIÁVEIS DERIVADAS

VIII. INDICADORES A DISPONIBILIZAR

IX. CONCEITOS

X. CLASSIFICAÇÕES

XI. SIGLAS E ABREVIATURAS

XII. BIBLIOGRAFIA

I. IDENTIFICAÇÃO DA OPERAÇÃO ESTATÍSTICA

☞ **I. 1 Designação da operação estatística**

Projeções de População Residente

☞ **I. 2 Sigla (ou abreviatura) da operação estatística**

Não aplicável.

☞ **I. 3 Código da operação estatística**

115

☞ **I. 4 Código SIGINE (modelo estatístico)**

Não aplicável.

☞ **I. 5 Código da Atividade Estatística**

31 – População, 314 – Estatísticas de Síntese Demográfica, 244 - Projeções Demográficas

☞ **I. 6 Código de Versão do Documento Metodológico**

4.0

☞ **I.7 Data de entrada em vigor da versão do Documento Metodológico**

Março de 2020

☞ **I.8 Data da última atualização do Documento Metodológico**

Março de 2020

☞ **I.9 Entidade responsável pela operação estatística**

INE

- **Unidade Orgânica (UO):** INE/DES/DM

- Técnico responsável:

- Nome: Cláudia Pina

- E-mail: claudia.pina@ine.pt

- Técnico responsável:

- Nome: Edvigés Coelho

- E-mail: edvigés.coelho@ine.pt

☞ **I.10 Outras Entidades Externas relacionadas com a operação**

Não aplicável.

II. JUSTIFICAÇÃO PARA UMA NOVA VERSÃO DO DOCUMENTO METODOLÓGICO

Assinalar uma cruz (X) nos quadrados correspondentes (resposta múltipla):

Nova versão:

- Alteração de variáveis de observação ☐
- Alteração de conceitos ☐
- Alteração de classificações ☐
- Alteração de produtos de difusão ☐
- Alteração de fontes estatísticas ☐
- Alteração do método de recolha ☐
- Alteração de âmbito na população ☐
- Alteração do desenho amostral ☐
- Outros motivos ☒
 - Especificar: Alteração do ano base (população de partida) e horizonte temporal de projeção, revisão das hipóteses de evolução das componentes de variação da população e reajustamento dos métodos e modelos de projeção.

III. IDENTIFICAÇÃO DAS NECESSIDADES, OBJETIVOS E FINANCIAMENTO

III.1 Contexto da operação estatística

A operação estatística “Projeções de População Residente” é um estudo estatístico, que com base na análise da evolução passada das componentes de dinâmica demográfica, na respetiva modelação, extrapolação e avaliação dos resultados - *expert judgement* -, permite formular hipóteses sobre a evolução futura das componentes de variação da população e, com base na combinação dessas hipóteses, delinear cenários demográficos futuros para a população.

As projeções de população residente são de importância fundamental para a definição das políticas públicas e constituem instrumentos para o planeamento, tanto público como privado, viabilizando análises prospetivas da procura por serviços públicos, para além de serem indispensáveis para o estudo de determinados segmentos da população para os quais são formuladas políticas específicas, como seja para a população mais idosa, jovens e crianças, bem como para o setor privado no que respeita ao dimensionamento dos mercados.

À semelhança de exercícios anteriores, as projeções de população residente 2018-2080 por idades e sexo para Portugal e regiões NUTS II, são produzidas utilizando o método das componentes por coortes. Na aplicação deste método, são definidas hipóteses sobre os níveis futuros da fecundidade, mortalidade e migrações, procedendo-se, de acordo com essas hipóteses, à atualização sucessiva dos efetivos populacionais, por idade e sexo, até atingir o último ano do período de projeção. Enquadram-se no que se pode designar por projeções de população determinísticas, uma vez que os resultados obtidos decorrem de cálculos puramente formais, determinados pela estrutura da população inicial e quando verificados determinados pressupostos relativamente aos níveis futuros da fecundidade, mortalidade e migrações, revestindo-se assim de um carácter condicional, do tipo “se x então y”. A conjugação de diferentes hipóteses alternativas relativas à evolução futura de cada componente possibilita a definição de diferentes cenários e, desta forma, lidar com a incerteza associada aos resultados obtidos.

No presente exercício, procedeu-se à incorporação dos dados observados mais recentes relativos a mortalidade, fecundidade e migrações, à alteração do ano base da projeção e à revisão das hipóteses de evolução destas componentes demográficas, nomeadamente com a introdução de melhorias no ajustamento do modelo de Schmertmann (2003, 2005) para a previsão das taxas de fecundidade específicas por idade e à aplicação do modelo de Hydman et al. (2013) para previsão de emigrantes e imigrantes.

☞ **III. 2 Identificação das necessidades de informação estatística que justificam a operação**

Assinalar uma cruz (X) nos quadrados correspondentes (resposta múltipla):

- **Necessidades** resultantes de **obrigações legais**:
 - Legislação comunitária ☐
Especificar: _____
 - Compromissos perante organizações internacionais ☐
Especificar: _____
 - Legislação nacional ☐
Especificar: _____
- **Pedido direto de informação** por parte do/de:
 - Entidades públicas nacionais ☐
 - Entidades comunitárias ☐
 - Programa Estatístico Europeu (PEE) ☐
 - Acordo informal (“Acordo de Cavalheiros”) ☐
 - Entidades privadas, nacionais ou estrangeiras ☐
Especificar: _____
 - Conselho Superior de Estatística (Recomendações, p.ex.) ☐
Especificar: _____
- Resultado de **inquéritos** às necessidades dos utilizadores ☐
- Necessidades de informação de **outras operações estatísticas** ☐
- Contrato/**Protocolo** específico com Entidade externa ☐
- Outras necessidades ☒
 - Especificar: Necessidades de informação relativas a cenários futuros de evolução da população. As projeções de população constituem uma base para a tomada de decisão informada nos sectores público e privado, nomeadamente no planeamento de infraestruturas (escolas, centros de saúde, lares de idosos e outros equipamentos) e na definição de políticas (por exemplo, relativas às pensões e reformas). São também amplamente utilizadas pelos investigadores para aferir o impacto dos cenários demográficos na sociedade, economia, ambiente, etc.

☞ III.3 Objetivos da operação estatística

A operação estatística “Projeções de População Residente” tem como principal objetivo estimar e disponibilizar valores projetados para a população por idades e sexo, para Portugal e regiões NUTS II, para um conjunto de cenários prováveis de evolução da população.

O ano base da projeção é 2018, correspondendo o horizonte de projeção ao período 2019 – 2080, sendo que o primeiro ano de projeção de referência a divulgar é 2020.

☞ III.4 Financiamento da operação estatística

Assinalar uma cruz (X) nos quadrados correspondentes (resposta única):

• Financiamento total:	
○ da Entidade responsável	<input checked="checked" type="checkbox"/>
○ da União Europeia (EUROSTAT)	<input type="checkbox"/>
○ de outra Entidade	<input type="checkbox"/>
▪ Especificar: _____	
• Cofinanciamento:	
○ Entidade responsável e União Europeia	<input type="checkbox"/>
○ Entidade responsável e outra Entidade (nacional ou externa à União Europeia)	<input type="checkbox"/>
▪ Especificar: _____	

IV. CARACTERIZAÇÃO GERAL

☞ IV.1 Tipo de operação estatística

Assinalar uma cruz (X) nos quadrados correspondentes (resposta única):

• Inquérito amostral	<input type="checkbox"/>
• Recenseamento	<input type="checkbox"/>
• Estudo estatístico	<input checked="" type="checkbox"/>

☞ IV.2 Tipo de fonte(s) de informação utilizada(s) na operação estatística

Assinalar uma cruz (X) nos quadrados correspondentes (resposta múltipla):

• Fonte Direta	<input type="checkbox"/>
• Fonte Não-direta	
○ Fonte administrativa	<input type="checkbox"/>
○ Outra operação estatística	<input checked="" type="checkbox"/>
○ Outra	<input type="checkbox"/>
• Especificar: _____	

☞ IV.3 Periodicidade da operação estatística

Assinalar uma cruz (X) nos quadrados correspondentes (resposta única):

• Mensal	<input type="checkbox"/>
• Trimestral	<input type="checkbox"/>
• Semestral	<input type="checkbox"/>
• Anual	<input type="checkbox"/>
• Bienal	<input type="checkbox"/>
• Trienal	<input checked="" type="checkbox"/>
• Quadrienal	<input type="checkbox"/>
• Quinquenal	<input type="checkbox"/>
• Decenal	<input type="checkbox"/>
• Não periódico	<input type="checkbox"/>
• Outra	<input type="checkbox"/>
Especificar: _____	

☞ IV.4 Âmbito geográfico da operação estatística

Assinalar uma cruz (X) nos quadrados correspondentes (resposta múltipla):

• Continente	<input type="checkbox"/>
• Região Autónoma da Madeira	<input type="checkbox"/>
• Região Autónoma dos Açores	<input type="checkbox"/>
• País	<input checked="" type="checkbox"/>
• Outro	<input type="checkbox"/>
Especificar: _____	

☞ IV. 5 Principais utilizadores da informação

Assinalar uma cruz (X) nos quadrados correspondentes (resposta múltipla):

Utilizadores do Sistema Estatístico Nacional	
○ INE	<input checked="" type="checkbox"/>
○ Banco de Portugal	<input type="checkbox"/>
○ Direção Regional de Estatística da Madeira	<input type="checkbox"/>
○ Serviço Regional de Estatística dos Açores	<input type="checkbox"/>
○ Entidades com delegação de competências	<input type="checkbox"/>
Especificar: _____	
Outros utilizadores nacionais	<input checked="" type="checkbox"/>
Especificar: Administração Central, Regional e Local	
Utilizadores Comunitários e outros Internacionais	<input type="checkbox"/>

☞ IV.6 Difusão

○ IV.6.1 Padrão de disponibilização da informação

Não aplicável.

○ IV.6.2 Revisões

Não aplicável.

○ IV.6.3 Produtos de difusão regular

Produtos a disponibilizar			
Tipo de produto	Designação do produto	Periodicidade de disponibilização	Nível geográfico (desagregação geográfica máxima)
. Destaque . Indicadores estatísticos	Projeções de População Residente	Trienal	. NUTS II

V. CARACTERIZAÇÃO METODOLÓGICA

☞ V.1 População-alvo

População residente em território nacional.

☞ V.2 Base de amostragem

Não aplicável.

☞ V. 3 Unidade (s) estatística (s) de observação

Indivíduo.

☞ V. 4 Desenho da amostra

Não aplicável.

☞ V.5 Construção do(s) questionário(s)

Não aplicável.

☞ V. 6 Recolha de dados

Não aplicável.

○ **V. 6.2 Recolha não-direta de dados**

Preencher o seguinte quadro:

NOME DA FONTE	TIPO DE FONTE (administrativa, outra operação estatística)	IDENTIFICAÇÃO DA ENTIDADE RESPONSÁVEL DA FONTE	PERÍODO DE REFERÊNCIA DOS DADOS DA FONTE
Estimativas anuais de população residente	Outra operação estatística – Estudo estatístico	INE/DES/DM	Anual
Óbitos	Outra operação estatística - recenseamento	INE/DES/DM	Anual
Nados-vivos	Outra Operação estatística - recenseamento	INE/DES/DM	Anual
Estimativas anuais de emigração	Outra operação estatística – Estudo estatístico	INE/DES/DM	Anual
Estimativas anuais de imigração	Outra operação estatística – Estudo estatístico	INE/DES/DM	Anual

☞ **V.7 Tratamento de dados**

Não aplicável.

○ **V. 7.1 Validação e análise**

Não aplicável.

○ **V. 7.2 Tratamento de não respostas**

Não aplicável.

○ **V.7.3 Obtenção de resultados**

As projeções de população residente são calculadas utilizando o método das componentes por coortes. Na aplicação deste método, são formuladas hipóteses sobre a evolução das componentes de variação da população - fecundidade, mortalidade e migrações – com base nas quais se procede à atualização sucessiva dos efetivos populacionais até se atingir o último ano do período de projeção.

A formulação das hipóteses baseia-se na observação, análise e modelação das tendências passadas de cada uma das componentes, em particular as de anos mais recentes, e na avaliação relativa à evolução futura das mesmas com base na informação disponível até ao momento (*expert judgement*).

Uma vez que os níveis futuros de fecundidade, mortalidade e migrações não são possíveis de prever com exatidão, duas ou mais hipóteses de evolução são consideradas para cada componente. Estas hipóteses alternativas têm como objetivo ilustrar um domínio de resultados futuros possíveis, apesar de não existir certeza sobre a sua eventual realização, ou que valores futuros necessariamente se incluam nesse domínio de valores.

De salientar ainda que na formulação das hipóteses e modelação da evolução de cada componente é apenas considerada a dinâmica temporal de indicadores demográficos, não sendo incorporado nos modelos qualquer outro tipo de variáveis exógenas.

1. FECUNDIDADE

A fecundidade tem uma influência importante na dimensão e na estrutura etária da população, quer pelo impacto direto no número de efetivos populacionais projetados, quer pelo impacto diferido, uma vez que o maior ou menor número de nascimentos, em particular do sexo feminino, num determinado momento, influencia o número de novos nascimentos no futuro.

Sendo uma variável sensível a variações conjunturais, a projeção reveste-se de maior complexidade, particularmente no contexto nacional atual, caracterizado por níveis de fecundidade particularmente reduzidos.

1.1. Hipóteses de Evolução da Componente Fecundidade

Três hipóteses de evolução futura da fecundidade em Portugal são consideradas:

Hipótese central – Nesta hipótese prevê-se uma recuperação moderada dos níveis futuros de fecundidade, com o Índice Sintético de Fecundidade (ISF) a atingir 1,59 crianças por mulher em 2080 (1,41 em 2018).

Hipótese otimista – Nesta hipótese prevê-se uma recuperação mais acentuada dos níveis futuros de fecundidade, com o ISF a atingir 1,82 crianças por mulher em 2080.

Hipótese pessimista – Nesta hipótese prevê-se a manutenção até 2080 dos valores do ISF em torno de 1,42 crianças por mulher.

Em qualquer das hipóteses prevê-se um aumento da idade média da mãe ao nascimento de um filho, entre 0,83 e 1,35 anos entre 2018 e 2080 (31,4 anos em 2018).

1.2. Rationale Subjacente à Definição das Hipóteses de Evolução da Fecundidade

A evolução do ISF, indicador conjuntural que traduz o número médio de crianças nascidas vivas por mulher em idade fértil, tem apresentado uma tendência de decréscimo ao longo dos últimos anos, atingindo um valor mínimo de 1,21 em 2013. Desde 2014 que o ISF tem vindo a recuperar, atingindo, em 2018, 1,41 nados-vivos por mulher em idade fértil.

Na década de sessenta do século XX, cada mulher tinha em média cerca de 3 filhos, valor que tem diminuído desde então, verificando-se desde o início da década de oitenta valores inferiores a 2,1 crianças por mulher, considerado como o nível de substituição de gerações.

Em meados da década de noventa, este indicador reduziu-se até 1,41 crianças por mulher. Assistiu-se posteriormente a uma ligeira recuperação até 2000 (1,56), ano a partir do qual volta a apresentar uma tendência de decréscimo, atingindo em 2013 o valor de 1,21 crianças por mulher, o mais baixo observado até agora em Portugal, tendo posteriormente vindo a apresentar uma ligeira recuperação, atingindo em 2018 1,41 filhos por mulher em idade fértil.

Na projeção da componente fecundidade para Portugal foi analisada informação para o período de 1980 a 2018, nomeadamente a evolução das taxas de fecundidade específicas por idades, do ISF, da idade média da mulher ao nascimento de um filho (calculada com base nas taxas de fecundidade) e da relação de masculinidade à nascença.

As hipóteses de evolução futura da fecundidade são baseadas na análise e modelação das tendências passadas da fecundidade a nível nacional, com particular atenção às tendências mais recentes.

1.3. Metodologia de Projeção da Fecundidade

A projeção da componente fecundidade foi multietápica. Num primeiro momento, foi aplicado para Portugal o modelo de Schmertmann (2003, 2005) às taxas de fecundidade específicas observadas no período 1980-2018, obtendo-se séries temporais para os parâmetros estimados do modelo (α, R, P, H) no referido período. Aos parâmetros estimados foi posteriormente aplicado um Modelo Vetorial Autorregressivo (VAR) para modelização e respetiva previsão que, combinado com o modelo de Schmertmann, permitiu a obtenção de taxas de fecundidade específicas por idade e ISF projetados. Num segundo momento, procedeu-se à análise da plausibilidade das curvas de fecundidade projetadas para Portugal com base numa avaliação *expert judgement*.

Para as regiões NUTSII, à semelhança de Portugal, aplicou-se numa primeira etapa o modelo de Schmertmann (2003, 2005) às taxas de fecundidade específicas observadas por região no período 1992-2018, obtendo-se séries temporais para os parâmetros estimados do modelo. Numa segunda etapa, projetaram-se taxas de fecundidade específicas por região assumindo-se que, partindo cada uma delas de níveis de fecundidade diferentes (iguais aos valores observados em 2018), as trajetórias projetadas para os parâmetros do modelo a nível nacional são acompanhadas a nível regional.

O modelo de Schmertmann

O modelo proposto por Schmertmann (2003, 2005) pretende caracterizar o perfil das taxas específicas de fecundidade por idade, $f(x)$, em termos de três parâmetros que sintetizam determinadas características observadas no comportamento deste fenómeno demográfico (e.g., aumento das idades médias da mãe ao nascimento do primeiro e de um filho). Estes três parâmetros correspondem, nomeadamente, à idade mais jovem em que se registam nascimentos, α , à idade em que o perfil das taxas específicas de fecundidade por idade atinge o valor máximo, P , e à primeira idade acima de P após a qual a fertilidade decresce para 50% do seu valor máximo, H .

Em termos formais, as taxas $f(x)$ entre as idades α e uma idade limite superior β são modeladas através da técnica de *splines* por troços em que:

$$f(x) = R\phi(x) \quad (1)$$

onde

$$\phi(x) = \begin{cases} \sum_{k=0}^4 \theta_k (x - t_k)_+^2, & \alpha \leq x \leq \beta \\ 0, & x \notin [\alpha, \beta] \end{cases}, \quad (2)$$

com "nós" $t_0 < t_1 < \dots < t_k$ no intervalo entre α e β , $(x - t_k)_+ = \max(x - t_k, 0)$ e R denota um escalar. A conjugação da técnica de *splines* com restrições relativamente ao posicionamento dos nós de modo a incorporar a informação relativa aos três eixos acima referidos permite mimetizar o comportamento observado e esperado para as taxas específicas de fecundidade por idade.

Da aplicação das taxas de fecundidade por idades aos efetivos populacionais femininos em idade fecunda resulta um número anual de nados-vivos, que são repartidos em nados - vivos do sexo feminino e masculino pela aplicação das probabilidades relativas à relação de masculinidade dos nascimentos (0,485 - sexo feminino e 0,515 - sexo masculino), e adicionados aos efetivos populacionais anuais.

Modelo Vetorial Autorregressivo (VAR)

Os modelos VAR são sistemas de equações simultâneas que capturam a existência de relações de interdependência entre variáveis, e que permitem avaliar o impacto de choques estocásticos sobre determinada variável do sistema (Tsay, 2010). Uma série temporal multivariada $r_t = (r_{t1}, r_{t2}, \dots, r_{tk})'$ composta por k componentes no tempo é um processo VAR de ordem p , ou simplesmente VAR(p), se para $p > 0$ é definido pelo modelo:

$$r_t = \phi_0 + \phi_1 r_{t-1} + \dots + \phi_p r_{t-p} + \varepsilon_t \quad (3)$$

onde ϕ_0 é um vector de dimensão k , ϕ_j são matrizes $k \times k$ ($j = 1, \dots, p$) e ε_t é um termo residual formado por uma sequência de vetores aleatórios independentes e identicamente distribuídos com média nula e matriz de variância-covariância Σ_ε . Os parâmetros do modelo VAR(p) são estimados pelo método dos mínimos quadrados ordinários (MQO). Para definir a ordem p do modelo VAR(p), são usados três critérios assumindo que ε_t segue uma distribuição Normal multivariada, a saber, critérios AIC (*Akaike's Information Criterion*), BIC (*Bayesian Information Criterion*), e HQC (*Hannan-Quinn Information Criterion*). É feito um diagnóstico ao modelo estimado e testada a existência de relações de causalidade à la Granger (1969).

2. MORTALIDADE

Num contexto de envelhecimento demográfico, como é o caso português, a evolução futura da componente mortalidade tem um importante impacto na dimensão da população idosa (65 anos e mais) e muito idosa (75 anos e mais).

2.1. Hipóteses de Evolução da Mortalidade

As hipóteses incidem sobre a evolução esperada da mortalidade em Portugal, em particular sobre os níveis futuros de esperança de vida à nascença para homens e mulheres. Três hipóteses são consideradas para a evolução da longevidade em Portugal:

Hipótese central – Nesta hipótese prevê-se a continuação das tendências recentes de melhoria na mortalidade. Os valores projetados das taxas de mortalidade são obtidos por extrapolação dos valores passados com base na aplicação do método de Poisson-Lee-Carter (Brouhns, Denuit e Vermunt, 2002) ao período 1980 – 2018. Nesta hipótese consideram-se as estimativas pontuais projetadas das taxas de mortalidade e da esperança de vida à nascença. Assim, sob a hipótese central, a esperança de vida à nascença em 2080, em Portugal, atingirá 87,92 anos para homens e 93,30 para mulheres, um aumento de 10,14 anos para os homens e de 9,87 anos para as mulheres (relativamente aos valores estimados para 2016-2018 que corresponde ao período de referência da última tábua de mortalidade disponível).

Hipótese otimista – Nesta hipótese prevê-se, relativamente à hipótese central, um aumento mais acentuado da esperança de vida à nascença. Os valores projetados para os quocientes de mortalidade correspondem, nesta hipótese, ao limite inferior do intervalo de previsão a 95% do modelo Poisson-Lee-Carter. Assim, sob a hipótese otimista a esperança de vida à nascença em 2080, em Portugal, atingirá 89,60 anos para homens e 94,42 para mulheres, um aumento de 11,82 anos para os homens e de 10,99 anos para as mulheres.

Hipótese pessimista – Nesta hipótese prevê-se, face à hipótese central, um aumento menos acentuado da esperança de vida à nascença. Os valores projetados para os quocientes de mortalidade correspondem, nesta hipótese, ao limite superior do intervalo de previsão a 95% do modelo. Assim, sob a hipótese pessimista a esperança de vida à nascença em 2080, em Portugal, atingirá 85,77 anos para homens e 91,31 para mulheres, um aumento de 7,99 anos para os homens e de 7,88 anos para as mulheres.

2.2. Rationale das Hipóteses de Evolução da Mortalidade

A esperança de vida à nascença reflete os níveis de mortalidade de uma população ao longo de todas as idades. Este indicador, sendo um indicador de síntese das condições de mortalidade de uma população, é amplamente utilizado para aferir as melhorias nos níveis de mortalidade de um país, quer ao longo do tempo, quer por comparação com outros países.

Assiste-se ao longo de várias décadas a um aumento continuado da esperança de vida à nascença nos países desenvolvidos, em particular da Europa Ocidental. O ritmo a que a esperança de vida à nascença tem aumentado apresenta, todavia, variações ao longo do

tempo, observando-se, em geral, menores ritmos de acréscimo da esperança de vida à nascença nos anos mais recentes.

Os aumentos na esperança de vida à nascença resultam dos progressos face à mortalidade em todas as idades. Ao longo do tempo, contudo, as contribuições relativas dos diferentes grupos de idades têm vindo a alterar-se, com impacto no ritmo de crescimento da esperança média de vida. Se no passado distante, a principal contribuição para o aumento da esperança de vida à nascença era proveniente, sobretudo, da redução na mortalidade infantil e juvenil, atualmente as maiores contribuições para o aumento na esperança de vida são provenientes das idades adultas e mais avançadas, com menor impacto no aumento do valor da esperança de vida à nascença. Uma proporção maioritária e crescente de óbitos ocorre em idades cada vez mais avançadas. Neste contexto, não são expectáveis aumentos substanciais no ritmo de acréscimo da esperança de vida à nascença.

Neste contexto, ainda que não seja expectável um aumento substancial no ritmo de acréscimo da esperança de vida à nascença, também não se antecipa que venham a verificar-se retrocessos neste indicador.

2.3. Cálculo dos Quocientes brutos de mortalidade

Um evento demográfico (e.g., um óbito) é caracterizado por três coordenadas que representam as três formas pelas quais o tempo intervém no estudo: (i) a data em que este se produz, (ii) a idade do indivíduo abrangido pelo evento e (iii) a sua data de nascimento. Os pontos referentes aos acontecimentos demográficos são contabilizados em triângulos que representam a dupla classificação, i.e., por idade e por data de nascimento. Para ilustrar esta relação e o método de cálculo dos quocientes, representa-se na Figura 1 a linha de vida de um indivíduo no diagrama de Lexis. Ele nasce no momento t_0 e morre no momento t_1 à idade x_1 , entre as idades exatas (datas de aniversário) x e $x + 1$. O segmento de reta vertical a azul representa o número de indivíduos que no momento t (início do ano civil) têm x anos de idade, i.e., o número de sobreviventes $P_{x,t}$ de idade x no instante t . O segmento de reta horizontal a vermelho representa o número de indivíduos que no ano t cumpriram o $x^{\text{ésimo}}$ aniversário, i.e., o número de sobreviventes à idade x provenientes da geração nascida no ano $g = t - x$.

No decurso do ano t , são registados $d_{x,t}$ óbitos entre os indivíduos com x anos inteiros vividos, localizados num quadrado delimitado pelas coordenadas $ABCD$, provenientes de duas gerações distintas. Este quadrado pode ser decomposto em dois triângulos: um triângulo inferior ACD , que agrega os óbitos $d_{x,t,g}^L$ registados no ano t entre os indivíduos oriundos da geração nascida em $g = t - x$ e um triângulo superior ABC , que agrupa os óbitos $d_{x,t,g-1}^U$ registados no ano t entre os indivíduos oriundos da geração nascida em $g - 1 = t - x - 1$ (Bravo, 2007; Coelho, Bravo e Magalhães, 2010; Bravo, Magalhães e Coelho, 2013).

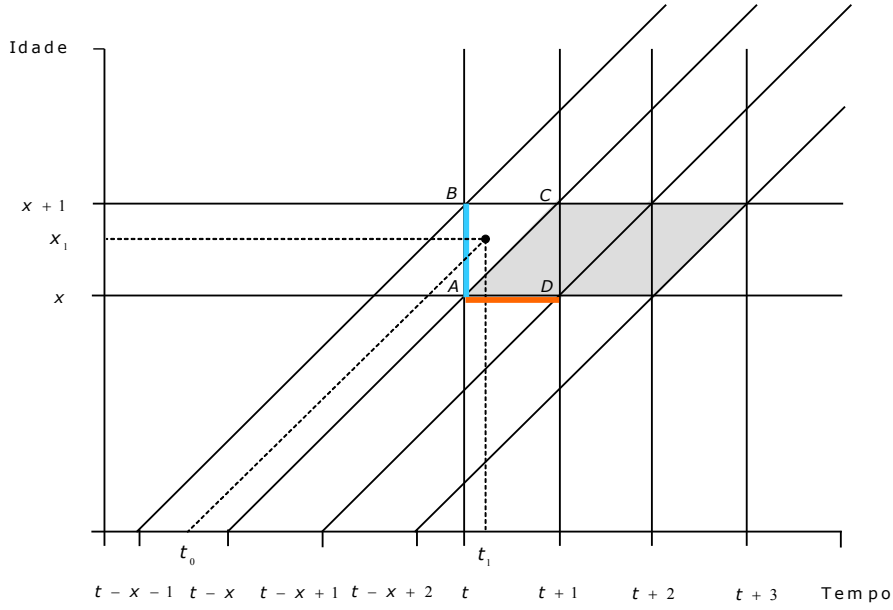


Figura 1: Diagrama de Lexis

O método de cálculo dos quocientes de mortalidade relativos a cada intervalo de idades $[x, x + 1)$ baseia-se na informação recolhida em três anos civis consecutivos, designados por t , $t + 1$ e $t + 2$. A estimativa da probabilidade de óbito q_x ($x = 1, 2, \dots$) é dada pela razão entre o número de óbitos registados à idade x , entre os indivíduos oriundos das gerações $g = t - x$ e $g + 1 = t - x + 1$, e o número de indivíduos das gerações g e $g + 1$ que atingem a idade x . Em termos formais,

$$\hat{q}_x = \frac{d_{x,t,g}^L + d_{x,t+1,g}^U + d_{x,t+1,g+1}^L + d_{x,t+2,g+1}^U}{P_{x,t+1} + d_{x,t,g}^L + P_{x,t+2} + d_{x,t+1,g+1}^L} \quad (4)$$

A probabilidade de óbito no primeiro ano de vida (i.e., à idade $x = 0$), q_0 , é estimada por:

$$\hat{q}_0 = \frac{d_{0,t,t}^L + d_{0,t+1,t}^U + d_{0,t+1,t+1}^L + d_{0,t+2,t+1}^U}{NV_t + NV_{t+1}} \quad (5)$$

onde NV_t denota o número de nados-vivos do ano t .

2.4. Metodologia de Projeção da Componente Mortalidade

Na projeção da componente mortalidade é utilizado o modelo Poisson Lee-Carter (Lee e Carter, 1992; Bravo, 2007; Brouhns, Denuit e Vermunt, 2002) para Portugal. Para as

regiões NUTS II, são aplicados modelos relacionais no quadro de modelos lineares generalizados (GLMs) incluindo como covariáveis a mortalidade da população de referência (população portuguesa masculina ou feminina), a idade atuarial e o tempo cronológico e efeitos de interação (Delwarde e Denuit, 2005), e modelos relacionais não paramétricos também no quadro de GLMs que envolvem a aplicação de técnicas de graduação e de fatores de melhoria por idade estimados com base na população de referência (Tomas & Planchet, 2014).

O modelo Poisson Lee-Carter (PLC) é aplicado às estimativas brutas dos quocientes de mortalidade específicas por idade inteira, estimadas no período 1980 – 2018 para o intervalo de idades compreendido entre os 0 e os 99 anos. Como procedimento de fecho das tábuas de mortalidade e de extrapolação do comportamento dos quocientes de mortalidade nas idades avançadas (acima dos 85 anos) aplica-se o método desenvolvido por Denuit e Goderniaux (2005) tomando como idade limite 115 anos para o período de estimação e para os primeiros três anos da projeção, momento em que aumenta progressivamente um ano para subseqüentes subperíodos de onze anos, atingindo no final do período de projeção a idade limite 125 anos. As estimativas do coeficiente β_x do modelo Lee-Carter, os quocientes de mortalidade $q_{x,t}$ do último ano observado e os $q_{x,t}$ obtidos pelo método Denuit-Goderniaux são graduados em torno da idade focal com recurso à técnica de splines cúbicos (Bravo, Magalhães e Coelho, 2007; Bravo & El Mekkaoui de Freitas, 2018).

Para as regiões NUTS II são aplicados, no quadro de GLMs, modelos relacionais (Brass, 1971, 1973) e métodos não paramétricos considerando os quocientes brutos de mortalidade estimados no período 1991-2018 no intervalo de idades entre os 0 e os 90 anos, e tomando como população de referência para cada NUTS II as tábuas de mortalidade prospetivas derivadas pelo método Poisson Lee-Carter para Portugal.

Numa segunda etapa, é aplicado o método de fecho e extrapolação dos quocientes de mortalidade nas idades mais avançadas desenvolvido por Denuit e Goderniaux (2005). A ideia subjacente é a de encontrar uma relação adequada que associe as características e a dinâmica da mortalidade do país e das regiões NUTS II. Encontrada essa relação, admitimos que esta é estável no tempo, o que nos permite obter os quocientes de mortalidade projetados de cada região NUTS II a partir dos quocientes projetados para a população portuguesa.

O Modelo de Lee-Carter (Lee e Carter, 1992)

O método de Lee-Carter (LC) assenta na decomposição da mortalidade em duas parcelas: um modelo demográfico, específico para cada idade, e um indicador de tendência temporal, sobre o qual são aplicados métodos de séries temporais no sentido de projetar valores futuros para os indicadores de mortalidade.

Formalmente, o método define-se pela seguinte representação log-bilinear para a taxa de mortalidade $m_{x,t}$ (Bravo, 2007):

$$\ln(m_{x,t}) = \alpha_x + \beta_x k_t + \varepsilon_{x,t} \quad (6)$$

onde:

- α_x : define, para cada idade, o nível médio das taxas de mortalidade no tempo;
- k_t : descreve as tendências temporais do nível de mortalidade;
- β_x : simboliza o padrão de oscilação etária nas taxas de mortalidade, ou seja, dá-nos uma ideia da sensibilidade de $m_{x,t}$ às variações em k_t ;
- $\varepsilon_{x,t}$: termo aleatório do tipo ruído branco $\varepsilon_{x,t} \sim N(0, \sigma_\varepsilon^2)$, incluído de modo a refletir influências históricas particulares em cada idade não capturadas pelo modelo.

Para resolver o problema da sobreparametrização do modelo definido por (6), são impostas condições adicionais. Entre as várias alternativas possíveis, consideraram-se as duas seguintes restrições de identificação do modelo

$$\sum_{x=x_{\min}}^{x_{\max}} \beta_x = 1, \quad \sum_{t=t_{\min}}^{t_{\max}} k_t = 0 \quad (7)$$

Uma das virtudes do método LC refere-se à forma como o modelo demográfico está definido, que garante que as taxas de mortalidade exibam um padrão de decréscimo exponencial, sem necessidade de impor qualquer limite assintótico arbitrário ou restrição que demarque os ganhos na esperança de vida.

Por outro lado, o facto de combinar um método demográfico, que é simultaneamente rico e parcimonioso, com um modelo de séries temporais, permite a modelação de um único índice temporal que resume a estrutura de covariâncias entre as m_x no tempo e a dependência dos perfis face ao parâmetro k_t , assegurando um compromisso entre as tendências das taxas correspondentes a cada idade.

A estimação dos parâmetros é efetuada com recurso ao método dos mínimos quadrados ordinários, ou seja, resolvendo o seguinte problema de otimização

$$(\hat{\alpha}_x, \hat{\beta}_x, \hat{k}_t) = \arg \min_{\alpha_x, \beta_x, k_t} \left\{ \sum_{x=x_{\min}}^{x_{\max}} \sum_{t=t_{\min}}^{t_{\max}} (\ln \hat{m}_{x,t} - \alpha_x - \beta_x k_t)^2 \right\} \quad (8)$$

Na versão original, Lee e Carter (1992) propõem a utilização de uma decomposição em valores singulares, mas soluções alternativas mais eficientes podem ser conseguidas mediante algoritmos iterativos.¹

¹ Para mais detalhes sobre os métodos de estimação do modelo veja-se, por exemplo, Bravo (2007) e referências citadas. Para desenvolvimentos recentes na modelação da mortalidade, incluindo a dimensão da heterogeneidade e a sazonalidade, veja-se, por exemplo, Ayuso et al. (2017a,b, 2019), Bravo (2019), Bravo e Nunes (2020), Hunt e Blake (2020), Bravo e Coelho (2019, 2020).

O método LC pertence, desde logo, à categoria dos métodos extrapolativos padecendo, por isso, da sua debilidade fundamental: assume que as tendências históricas se mantêm no futuro, não incorpora elementos de avaliação subjetiva e ignora as implicações de eventuais alterações estruturais. Por outro lado, o método admite um determinado padrão de distribuição etária da mortalidade, em que as taxas de decréscimo nas várias idades mantêm sempre a mesma relação de proporcionalidade no tempo, uma hipótese não confirmada nalguns estudos empíricos. Brouhns *et al.* (2002) e Renshaw e Haberman (2003) criticam a hipótese de homocedasticidade assumida pelo método de Lee-Carter e desenvolvem uma versão baseada no modelo probabilístico definido pela distribuição de Poisson.

O Modelo Poisson Lee-Carter (Brouhns, Denuit e Vermunt, 2002)

O método LC assume que $\varepsilon_{x,t}$ é um termo aleatório do tipo ruído branco, isto é, admite que os erros são normalmente distribuídos com variância constante. A hipótese de variância constante dos erros revela-se empiricamente irrealista, na medida em que devido à existência de um número reduzido de óbitos nas idades avançadas, o logaritmo das taxas de mortalidade específicas por idade é muito mais variável nestas idades do que nas idades mais jovens.

Para contornar este problema, Brouhns *et al.* (2002) desenvolvem um modelo, fundado no método de Lee-Carter, em que consideram que o número de óbitos verificados à idade x no ano t segue uma distribuição de Poisson, i.e.,

$$D_{x,t} \sim \text{Poisson} \left(\mu_{x,t} E_{x,t} \right) \quad (9)$$

com

$$\mu_{x,t} = \exp \left(\alpha_x + \beta_x k_t \right) \quad (10)$$

O modelo preserva a estrutura log-bilinear para $\mu_{x,t}$ (taxa instantânea de mortalidade) mas substitui as hipóteses relativas ao termo erro $\varepsilon_{x,t}$ pela lei de Poisson para $d_{x,t}$. Os parâmetros α_x , β_x e k_t mantêm, na sua essência, o significado originalmente atribuído pelo método Lee-Carter.

O modelo de Poisson tem, contudo, algumas vantagens face ao método original de Lee e Carter que o tornam especialmente atrativo. Em primeiro lugar, o modelo reconhece explicitamente a natureza inteira de $D_{x,t}$ contrariamente ao que acontece no método LC em que se modela o logaritmo da taxa de mortalidade. Em segundo, o modelo deixa cair a hipótese de homocedasticidade do termo erro e reconhece a maior variabilidade de $\mu_{x,t}$ nas idades mais elevadas. Em terceiro, o método permite o recurso ao método de máxima verosimilhança para estimar os parâmetros em detrimento do uso do método dos mínimos quadrados aplicado mediante uma *Singular Value Decomposition* (SVD). Por fim, e não menos importante, o modelo dispensa o procedimento de reestimação do índice temporal k_t necessário na versão original proposta por Lee e Carter.

Em termos formais, os estimadores correspondem aos valores α_x , β_x e k_t que maximizam a seguinte função de log-verosimilhança:

$$L(\alpha_x, \beta_x, k_t) = \sum_{x=x_{\min}}^{x_{\max}} \sum_{t=t_{\min}}^{t_{\max}} \{d_{x,t}(\alpha_x + \beta_x k_t) - E_{x,t} \exp(\alpha_x + \beta_x k_t)\} + c \quad (11)$$

onde c é uma constante.

A presença do termo bilinear $\beta_x k_t$ na especificação (10) de $\mu_{x,t}$ impede a estimação dos parâmetros pelos packages informáticos tradicionais que incluem o modelo de Poisson. Por esta razão, recorremos a um algoritmo iterativo desenvolvido por Goodman (1979), assente no método de Newton-Raphson. O procedimento assenta num princípio simples: em cada iteração v , um único conjunto de parâmetros θ_j é atualizado, mantendo os restantes fixos à sua estimativa anterior, com a ajuda do seguinte esquema de atualização

$$\hat{\theta}_j^{(v+1)} = \hat{\theta}_j^{(v)} - \frac{\partial L^{(v)} / \partial \theta_j}{\partial^2 L^{(v)} / \partial \theta_j^2} \quad (12)$$

onde $L^{(v)} = L(\hat{\theta}_j^{(v)})$.

As estimativas α_x , β_x e k_t geradas pelo método de máxima verosimilhança não garantem, automaticamente, o cumprimento das restrições (7), pelo que para o conseguir é necessário efetuar uma reparametrização final do modelo.

A previsão do modelo baseia-se, à semelhança do que acontece no método LC, na aplicação da metodologia de Box-Jenkins à série temporal das estimativas k_t .

Um ponto delicado nos métodos de LC e de Poisson reside na não simultaneidade dos procedimentos de estimação dos parâmetros, uma vez que a modelação ocorre em duas fases. Numa primeira, os parâmetros α_x , β_x e k_t são estimados. Numa segunda, o parâmetro k_t é extrapolado usando a metodologia de Box-Jenkins.

Método de Extrapolação dos Quocientes de Mortalidade nas Idades Avançadas

Nas idades avançadas, a irregularidade detetada nas estimativas brutas dos quocientes, aconselha a utilização de um método de extrapolação dos quocientes de mortalidade. O método deve ajustar as estimativas brutas dos quocientes de mortalidade e ser, simultaneamente, compatível com as tendências detetadas a nível nacional e internacional no comportamento da mortalidade nas idades avançadas, nomeadamente com a evidência empírica de que existe uma desaceleração na taxa de crescimento a partir de uma certa idade.

Para extrapolação dos quocientes de mortalidade nas idades avançadas aplicou-se o método desenvolvido por Denuit e Goderniaux (2005). Este método trabalha com base nos quocientes de mortalidade q_x e introduz condições referentes à idade limite. Em

termos formais, envolve o ajustamento, pelo método dos mínimos quadrados, do modelo log-quadrático:

$$\ln \hat{q}_x = a + bx + cx^2 + \varepsilon_x, \quad \varepsilon_x \sim N(0, \sigma^2) \quad (13)$$

às observações relativas às idades mais elevadas (e.g., $x \geq 75$), respeitando duas restrições:

$$q_\omega = 1 \quad (14)$$

$$q'_\omega = 0. \quad (15)$$

Em (16), q'_x designa a primeira derivada de q_x com respeito à idade x e ω denota a idade máxima admitida. Demonstra-se que a verificação das restrições (14) e (15) nos permite rescrever o modelo em função de um único parâmetro:

$$\ln \hat{q}_x = (\omega^2 - 2x\omega + x^2)c + \varepsilon_x, \quad \varepsilon_x \sim N(0, \sigma^2) \quad (16)$$

As restrições impõem à curva dos quocientes de mortalidade uma configuração côncava nas idades elevadas e a existência de uma tangente horizontal no ponto $x = \omega$. A restrição (15) visa impedir um eventual decréscimo dos quocientes nas idades avançadas.

Para determinar a idade a partir da qual as estimativas brutas \hat{q}_x são substituídas pelos valores ajustados provenientes do modelo (16) (idade de corte), recorre-se a um procedimento que considera como critério de otimização a maximização do coeficiente de determinação R^2 , fazendo variar a idade de corte, x_0 , num intervalo entre, por exemplo, os 70 e os 95 anos. Para evitar as descontinuidades em torno de x_0 , e de forma a garantir uma transição suave entre as estimativas brutas e os valores ajustados, adota-se um método de graduação da série em torno de x_0 , utilizando para tal a técnica de *splines* cúbicos aplicada aos \hat{q}_x nas idades adjacentes à idade focal x_0 .

Graduação dos Quocientes de Mortalidade Utilizando *Splines* Cúbicos

Na graduação dos quocientes $q_{x,t}$ obtidos pelo método Denuit-Goderniaux em torno da idade focal é utilizada a técnica de *splines* cúbicos. Um *spline* é uma curva definida matematicamente por dois ou mais pontos de controlo, designados por nós (*knots*). A graduação mediante *splines* baseia-se na concatenação de funções polinomiais de determinado grau, definidas por troços, e sujeitas a determinadas condições de modo a garantir a continuidade da função ao longo de todo o intervalo.

Os *splines* podem ser divididos em duas grandes categorias: (i) *splines* de interpolação, que passam por todos os pontos de controlo, e (ii) *splines* de aproximação, que passam

perto de todos os pontos de controlo. Os *splines* podem ainda ser distinguidos segundo as funções de base escolhidas para representar a função interpoladora.

De entre as múltiplas configurações possíveis, uma das escolhas mais populares compreende a utilização de polinómios cúbicos definidos por troços, com primeira e segunda derivadas contínuas nos nós da função. Uma restrição adicional, destinada a garantir a linearidade da função para além dos nós de fronteira, é aquela que exige que a segunda derivada da função nos extremos do intervalo seja nula. Os *splines* assim definidos são denominados de *splines* cúbicos naturais.

Em termos formais, considere-se um intervalo de idades arbitrário $I = [a, b]$, particionado em n sub-intervalos $I_i = [x_i, x_{i+1}]$ ($i = 0, \dots, n-1$) por $n+1$ nós $a = x_0 < x_1 < \dots < x_n = b$. Nos casos em que os nós são distribuídos de forma equidistante no intervalo $[a, b]$, o *spline* diz-se *uniforme*, caso contrário é designado de *não uniforme*. Uma função s é denominada *spline* de grau k com nós nos pontos $\{x_i\}_{i=0}^n$ se é definida por um polinómio de grau k em cada intervalo $[x_i, x_{i+1}]$ e é $(k-1)$ vezes diferenciável em todo o intervalo $[a, b]$.

Dito isto, podemos definir a interpolação de uma função $g(x)$, explicitada em $[a, b]$, por intermédio de *splines* cúbicos naturais com nós em x_0, \dots, x_n , como uma função $s(x)$ que verifica as seguintes condições:

1) $s(x) \in C^2[a, b]$.

2) Em cada intervalo $[x_i, x_{i+1}]$, a função é definida por um polinómio cúbico do tipo:

$$s(x) = s_i(x) = \sum_{l=0}^3 a_l^{(i)} (x - x_i)^l, \quad i = 0, \dots, n-1.$$

3) Em cada um dos nós x_i ($i = 1, \dots, n$)

$$s(x_i) = g(x_i).$$

4) Condições de continuidade nos nós interiores

$$\begin{cases} s_i(x_{i+1}) = s_{i+1}(x_{i+1}) \\ s'_i(x_{i+1}) = s'_{i+1}(x_{i+1}) \\ s''_i(x_{i+1}) = s''_{i+1}(x_{i+1}) \end{cases}$$

5) São respeitadas as seguintes condições limite

$$s''(a) = s''(b) = 0.$$

Dada a série (x_i, y_i) , os coeficientes $a_l^{(i)}$ ($l = 0, \dots, 3$) do polinómio são estimados minimizando uma função objetivo do tipo

$$\hat{a}_l^{(i)} = \arg \min_{a_l^{(i)}} \left\{ \sum_{i=1}^n w_i (y_i - s(x_i))^2 \right\}, \quad (17)$$

onde w_i representa um conjunto de ponderadores pré-definido.

Em face do habitual conflito entre qualidade do ajustamento e alisamento da série, um dos aspetos mais sensíveis na implementação deste método prende-se com a escolha do número de nós e respetiva localização no intervalo. A opção mais simples consiste em escolher apenas o número de nós interiores, posicionando-os de seguida de modo a obter uma distribuição uniforme ao longo do intervalo dos dados. Um esquema alternativo consiste em posicionar os nós nos quartis da variável independente ou em pontos de inversão da concavidade da função.²

Uma forma de obter um compromisso entre os dois objetivos passa pela inclusão de um termo adicional em (Target) de modo a penalizar a sobreparametrização do modelo ou, o que é o mesmo, as oscilações excessivas na curvatura da série. Os *splines* cúbicos possuem, a este propósito, uma propriedade conhecida e interessante: são, de entre todas as funções $s(x)$ com duas derivadas contínuas, aqueles que minimizam a seguinte soma dos quadrados penalizados

$$\hat{a}_l^{(i)} = \arg \min_{a_l^{(i)}} \left\{ \sum_{i=1}^n w_i (y_i - s(x_i))^2 + \lambda \int_a^b (s''(u))^2 du \right\}, \quad (18)$$

onde λ ($\lambda \geq 0$) é um parâmetro de *smoothing* fixo. Na equação (18), o primeiro termo da função objetivo (P-Target) mede o grau de ajustamento face aos valores observados, enquanto que o segundo introduz uma penalização diretamente proporcional às oscilações na curvatura da função interpoladora. No limite, à medida que $\lambda \rightarrow \infty$ o termo correspondente à penalização domina, forçando $s''(x_i) = 0$ em todo o domínio, tendendo a solução para a convencional aproximação pelos mínimos quadrados. No extremo oposto, quando $\lambda \rightarrow 0$, o termo penalizador torna-se irrelevante e a solução tende para uma função interpoladora duplamente diferenciável.

A solução $\hat{s}_\lambda(\cdot)$ do problema de minimização (P-Target) é um *spline* cúbico em que os nós são os pontos x_0, x_1, \dots, x_n , ou seja, $\hat{s}_\lambda(\cdot)$ traduz um polinómio de terceiro grau em cada intervalo $[x_i, x_{i+1}]$, com primeira e segunda derivadas contínuas em cada um dos x_i . O problema (P-Target) equivale, em termos matriciais, a

$$\hat{a} = \arg \min_a \left\{ (y - s)' W (y - s) + \lambda s' P s \right\}, \quad (19)$$

onde

$$P = \Omega' Q^{-1} \Omega$$

e onde Ω é uma matriz tri-diagonal de dimensão $(n-2) \times n$ definida por

²Para uma aplicação da técnica de *splines* em problemas de graduação veja-se, por exemplo, Forfar et al. (1988), McCutcheon (1987) e referências incluídas.

$$\Omega = \begin{pmatrix} \frac{1}{\xi_1} & -\left(\frac{1}{\xi_1} + \frac{1}{\xi_2}\right) & \frac{1}{\xi_2} & \cdots & 0 \\ 0 & \frac{1}{\xi_2} & -\left(\frac{1}{\xi_2} + \frac{1}{\xi_3}\right) & \ddots & 0 \\ 0 & 0 & \frac{1}{\xi_3} & \ddots & \frac{1}{\xi_{n-1}} \\ \vdots & \vdots & \vdots & \ddots & \left(\frac{1}{\xi_{n-2}} + \frac{1}{\xi_{n-1}}\right) \\ 0 & 0 & 0 & \cdots & \frac{1}{\xi_{n-2}} \end{pmatrix}$$

$\xi_i = x_{a+i} - x_{a+i-1}$, e Ω é uma matriz tri-diagonal simétrica de dimensões $(n-2) \times (n-2)$ definida por

$$\Omega = \begin{pmatrix} 2(\xi_1 + \xi_2) & \xi_2 & \cdots & 0 \\ \xi_2 & 2(\xi_2 + \xi_3) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \vdots & 2(\xi_{n-2} + \xi_{n-1}) \end{pmatrix}$$

e W é uma matriz diagonal contendo os ponderadores.

A solução $\hat{s}_\lambda(\cdot)$ deriva-se anulando o gradiente da função objetivo (P-Target M), i.e.,

$$-2W(y - s) + 2\lambda Ps = 0,$$

donde se obtém finalmente

$$\hat{s}_\lambda = (W + \lambda P)^{-1} Wy. \quad (20)$$

O parâmetro λ regula, como vimos, a importância relativa dos objetivos na graduação e assume, como veremos mais adiante, um significado semelhante ao conceito de *kernel bandwidth*. Pode ser determinado de forma subjetiva ou, como explicitaremos em seguida, resultar de um procedimento automático como seja o de validação cruzada. Registe-se, por fim, que o método preconizado pela função objetivo (P-Target) se encontra amplamente difundido na literatura sobre graduação com a designação de fórmula de Whittaker-Henderson.

Modelos Relacionais no quadro de Modelos Lineares Generalizados (NUTS II)

Seja ${}_n p_{x,t}$ a probabilidade de sobrevivência no ano t , isto é, a probabilidade de um indivíduo vivo à idade x sobreviver até à idade $x+n$, e ${}_n q_{x,t}$ a correspondente probabilidade de óbito, também designada quociente de mortalidade, isto é a probabilidade de um indivíduo que atinge a idade x morrer antes de atingir a idade $x+n$.

Estimada a série de quocientes de mortalidade por idade $\{q_{x,t}\}_{x=0}^{\omega-1}$, onde ω denota a idade máxima admissível, a probabilidade de sobrevivência entre as idades x e $x+1$ no ano t é dada por:

$$p_{x,t} = 1 - q_{x,t} \quad (21)$$

para cada uma das idades.

O modelo relacional pressupõe que o número de óbitos verificados à idade x no ano t segue uma distribuição de Poisson

$$D_{x,t} \sim \text{Poisson} \left(\mu_{x,t} E_{x,t} \right) \quad (22)$$

com

$$\mu_{x,t} = \beta_0 + \beta_1 \log q_{x,t}^{ref} + \beta_2 x + \beta_3 t + \beta_4 x \cdot t. \quad (23)$$

onde x e t correspondem ao intervalo de idades e período de observação em comum com a população de referência, e $q_{x,t}^{ref}$ denota os quocientes de mortalidade estimados para a população de referência. A inclusão de efeitos de interação em (23) é justificada pela necessidade de contemplar os casos em que as taxas de mortalidade não se movem de forma paralela em função da dinâmica da população de referência.

Um aspeto crítico na aplicação deste método refere-se à escolha da população *standard*. Neste estudo, para cada região NUTS II é usada como população de referência a população do país do mesmo sexo.

As estimativas dos parâmetros β_0, \dots, β_4 são obtidas através do método de máxima verosimilhança. Uma vez obtidas, as probabilidades de morte estimadas para cada região NUTSII são obtidas mediante:

$$q_{x,t}^{NUTSII} = \exp \left(\hat{\beta}_0 + \hat{\beta}_1 \log q_{x,t}^{ref} + \hat{\beta}_2 x + \hat{\beta}_3 t + \hat{\beta}_4 x \cdot t \right), \quad (24)$$

para $x \in (x_{\min}, x_{\max})$, $t \in (t_{\min}, t_{\max})$.

Modelos Relacionais não paramétricos (NUTS II)

Para a modelação e projecção da mortalidade nas regiões NUTS II foram ainda usados modelos relacionais não paramétricos também no quadro de GLMs. Em termos formais, o método assume que

$$D_{x,t} \sim \text{Poisson} \left(E_{x,t} q_{x,t}^{ref} \exp f(x) \right) \quad (25)$$

Incluindo o número esperado de óbitos de acordo com a população de referência e $f(\cdot)$ é uma função suave não especificada da idade. A seleção dos parâmetros de alisamento é feita mediante um compromisso entre a qualidade do ajustamento e a regularidade da série graduada.

O método relacional é suficientemente flexível para admitir transformações alternativas e modelos não lineares de modo a captar o comportamento de populações com perfis de mortalidade/longevidade mais complexos.

3. MIGRAÇÕES

A desaceleração do crescimento demográfico e, mais recentemente, o decréscimo dos volumes populacionais, a par com um continuado processo de envelhecimento demográfico consubstanciam as principais linhas de caracterização das tendências demográficas dos últimos anos em Portugal, à semelhança do que se verifica também em outros países desenvolvidos.

Num contexto demográfico em que os saldos naturais têm vindo a diminuir, e em que o crescimento demográfico tem sido fortemente influenciado pelos fluxos migratórios, a evolução futura da componente migrações poderá resultar em impactos relevantes quer na dimensão quer na estrutura etária das populações futuras.

No âmbito do presente exercício são apenas consideradas hipóteses sobre as migrações internacionais, elaboradas de forma independente para a imigração e emigração. O saldo migratório projetado é obtido por diferença entre o número projetado de imigrantes e o número projetado de emigrantes.

3.1. Hipóteses de Evolução das Migrações

As hipóteses incidem sobre possíveis evoluções dos fluxos migratórios internacionais, sendo consideradas quatro hipóteses para Portugal:

Hipótese central – Nesta hipótese prevê-se a continuação das tendências, em particular as mais recentes, de evolução da imigração e da emigração, com manutenção de saldo migratórios internacionais anuais positivos ao longo do período de projeção, a atingir em 2080 um saldo migratório de 14 020 (11 570 em 2018).

Hipótese otimista – Nesta hipótese considera-se a possibilidade de uma redução do número de emigrantes e um acréscimo no número de imigrantes superiores aos valores estabelecidos na hipótese central, a atingir em 2080 um saldo migratório positivo de 35 988.

Hipótese pessimista – Nesta hipótese considera-se a possibilidade de aumento do número de emigrantes e uma redução do número de imigrantes relativamente aos valores da hipótese central, a atingir em 2080 um valor negativo de 8 268.

Hipótese sem migrações – Apesar da sua forte improbabilidade, foi ainda considerada uma hipótese de evolução assente no pressuposto de ausência de fluxos migratórios com o objetivo de avaliar a influência dos mesmos na dinâmica demográfica.

3.2. Rationale Subjacente à Definição das Hipóteses de Evolução das Migrações

Sendo as migrações uma das componentes da dinâmica demográfica, a par com os nascimentos e os óbitos, esta é a que revela maior volatilidade e incerteza.

Ao longo das últimas décadas, Portugal experienciou períodos de forte emigração, como aconteceu na década de sessenta do século XX, quando a emigração, sobretudo com destino a outros países da Europa, conheceu valores muito elevados, e na década de oitenta, devido a novos fluxos de emigração, ainda que de menor expressão, comparativamente com os registados na década de sessenta, nomeadamente pelos diferentes contextos socioeconómicos que marcaram estes períodos.

Na década de setenta e particularmente nos anos de 1974 e 1975 a tendência alterou-se, registando-se fortes fluxos imigratórios em consequência do retorno de população proveniente das ex-colónias, após a descolonização.

Já nos anos noventa o crescimento migratório voltou a ser positivo, devido sobretudo ao incremento da imigração, com valores que foram aumentando até ao início do novo século, momento a partir do qual se atenuaram, mas mantendo o saldo migratório valores positivos até 2010, tendo passado a valores negativos desde 2011 e até 2016, sobretudo em consequência do aumento dos fluxos emigratórios, que nunca deixaram de ocorrer mas que se acentuaram fortemente, de novo, a partir de 2008, atingindo um valor máximo de 53 786 em 2013, tendo-se posteriormente atenuado para 49 572 em 2014, 40 377 em 2015 e 38273 em 2016. Os fluxos imigratórios, que após atingirem um valor mínimo de 14 606 em 2013, começaram a recuperar. Em conjugação com a redução da emigração, a recuperação do número de imigrantes refletiu-se, em 2017, num valor positivo para o saldo migratório (4 886). Em 2018, o saldo migratório foi de 11 570.

A volatilidade associada à componente migrações, fortemente relacionada com as alterações socioeconómicas, não só em Portugal mas em termos de economia global, torna a projeção da componente migratória particularmente difícil.

3.3. Metodologia de Projeção da Componente Migrações

Na projeção dos fluxos imigratórios e emigratórios internacionais foram adotados modelos funcionais coerentes de séries temporais. Este é o método proposto por Hyndman et al. (2013), uma extensão do modelo de séries temporais funcionais independentes de Hyndman e Ullah (2007) e Hyndman e Booth (2008), assegurando que

as previsões para subpopulações que estão relacionadas, neste caso homens e mulheres, mantenham eventuais relações estruturais baseadas na observação histórica do fenómeno.

Estes métodos são centrados em análise de componentes principais para dados funcionais, uma técnica comum de redução da dimensionalidade em análise estatística. A análise de componentes principais reduz um número intrinsecamente infinito de variáveis a um conjunto de variáveis ortogonais (componentes) que capturam as principais tendências de variação nos dados.

Modelo de dados funcionais

Em termos formais, seja $y_{t,j}^*(x)$ a quantidade a ser modelada por idade x no ano t para a subpopulação j ($j = 1, 2, \dots, J$). Hyndman e Ullah (2007) sugerem, numa primeira fase, a aplicação de uma transformação de Box-Cox para prevenir que a volatilidade aumente com o valor absoluto de $y_{t,j}^*(x)$.

$$y_{t,j}(x) = \begin{cases} \frac{1}{\lambda} \left([y_{t,j}^*(x)]^\lambda - 1 \right) & \text{se } 0 < \lambda \leq 1; \\ \log_e \left(y_{t,j}^*(x) \right) & \text{se } \lambda = 0, \end{cases} \quad (26)$$

onde λ determina a magnitude da transformação.

O seguinte modelo é desenvolvido para a quantidade transformada $y_{t,j}(x)$:

$$y_{t,j}(x) = s_{t,j}(x) + \sigma_{t,j}(x)\varepsilon_{t,x,j}, \quad (27)$$

onde $s_{t,j}(x)$ representa a variável transformada graduada por idade, $\varepsilon_{t,x,j}$ é uma variável aleatória i.i.d. normalmente distribuída e $\sigma_{t,j}(x)$ permite que o ruído varie com a idade. Como método de graduação, Hyndman et al. (2013) aplicam *weighted penalized regression splines*.

Modelo de produto-rácio para homens e mulheres

Hyndman et al. (2013) definem, para homens e mulheres, a função produto $p_t(x)$ e a função rácio $r_t(x)$:

$$\begin{aligned}
p_t(x) &= \left(f_{t,H}(x) f_{t,M}(x) \right)^{\frac{1}{2}} \\
r_t(x) &= \left(\frac{f_{t,H}(x)}{f_{t,M}(x)} \right)^{\frac{1}{2}}
\end{aligned} \tag{28}$$

Estas funções são modelizadas em vez dos dados originais. A vantagem desta abordagem é que o produto e o rácio são funções relativamente independentes para subpopulações com variâncias aproximadamente iguais.

Modelos funcionais de séries temporais (Hyndman e Ullah, 2007) são aplicados a $p_t(x)$ e $r_t(x)$:

$$\log[p_t(x)] = \mu_p(x) + \sum_{k=1}^K \beta_{t,k} \phi_k(x) + e_t(x) \tag{29}$$

$$\log[r_t(x)] = \mu_r(x) + \sum_{l=1}^L \gamma_{t,l} \psi_l(x) + w_t(x) \tag{30}$$

Onde as funções $\{\phi_k(x)\}$ e $\{\psi_l(x)\}$ representam o conjunto de k e l funções de base ortogonais, que são obtidas através da decomposição de componentes principais de $\{p_t(x)\}$ e $\{r_t(x)\}$, respectivamente, e $\beta_{t,k}$ e $\gamma_{t,l}$ representam os conjuntos independentes de coeficientes para cada componente.

A função $\mu_p(x)$ é a média do conjunto de curvas $\{p_t(x)\}$ e $\mu_r(x)$ é a média de $\{r_t(x)\}$. Os termos de erro, dados por $e_t(x)$ e $w_t(x)$, têm média zero e são não correlacionados.

Os modelos são estimados através da aplicação de componentes principais ponderadas seguindo o algoritmo de Shang et al. (2011), que atribuí maior ponderação aos dados mais recentes. A previsão no modelo implica o ajustamento de modelos de séries temporais a cada um dos coeficientes $\{\beta_{t,1}, \dots, \beta_{t,K}\}$ e $\{\gamma_{t,1}, \dots, \gamma_{t,L}\}$.

As previsões pontuais de emigrantes e imigrantes por idade e sexo para Portugal, obtidas pela aplicação do modelo de Hyndman et al. (2013) ao período 1991-2018, foram consideradas para a hipótese central de evolução da emigração e da imigração. Estas previsões representam a continuação das tendências passadas, com maior ponderação atribuída aos valores mais recentemente observados. Assim, sob esta hipótese prevê-se

a manutenção de saldos migratórios positivos ao longo do período de projeção (2019-2080), em resultados de valores projetados para a imigração superiores aos valores da emigração, a atingir, respectivamente, 36 261 imigrantes e 22 241 emigrantes em 2080, e um saldo migratório de 14 020.

Considerando expectativas de maior acréscimo do número de imigrantes e de redução do número de emigrantes, considerou-se uma hipótese otimista de evolução dos fluxos migratórios e consequentemente de um saldo migratório mais favorável. Nesta hipótese, projetam-se valores de emigração inferiores e valores de imigrantes superiores aos considerados na hipótese central, traduzindo-se num saldo migratório a atingir 35 988 em 2080.

Alternativamente, foi considerada uma hipótese pessimista de evolução dos fluxos migratórios, em que se considera um aumento do número de emigrantes e uma redução do número de imigrantes relativamente aos valores estabelecidos na hipótese central, atingindo em 2080 um saldo migratório negativo de 8 268.

Para definição destas hipóteses recorreu-se aos intervalos de previsão obtidos através da aplicação do modelo de Hyndman *et al.* (2013).

A projeção dos fluxos migratórios para as regiões NUTS II foi efetuada em duas etapas. Numa primeira etapa é aplicado o modelo de Hyndman *et al.* (2013) ao número de emigrantes e imigrantes de cada região para o período 1991-2018. As previsões pontuais do número de emigrantes e imigrantes por regiões NUTS II para cada ano de projeção são utilizadas para determinar o peso de cada região nos respetivos fluxos nacionais. A projeção por sexo e idade é obtida com base no pressuposto de que a evolução da estrutura por idades e sexo dos fluxos migratórios é idêntica à projetada para Portugal.

4. PROJEÇÕES DE POPULAÇÃO RESIDENTE

4.1. Cenários

A conjugação das hipóteses atrás descritas, relativas à evolução futura de cada componente, permite definir cenários de projeção da população. No âmbito do presente exercício são definidos quatro cenários para Portugal e regiões NUTS II:

CENÁRIO BAIXO: Neste cenário são consideradas as hipóteses pessimista para a fecundidade, pessimista para a mortalidade e pessimista para as migrações.

CENÁRIO CENTRAL: Neste cenário são consideradas as hipóteses de evolução central da fecundidade, central da mortalidade e central das migrações.

CENÁRIO ALTO: Este cenário resulta da combinação das hipóteses de evolução otimista da fecundidade, otimista da mortalidade e otimista das migrações.

CENÁRIO SEM MIGRAÇÕES: Um cenário idêntico ao cenário central, mas sem migrações.

4.2. Método de cálculo de populações projetadas

As estimativas anuais provisórias de população residente em 31 de dezembro de 2018, calculadas pelo INE e divulgadas em junho de 2019, constituem a população base ou de partida para as projeções de população.

Partindo da população base, a população residente projetada é calculada utilizando o método das componentes por coortes, em que a população inicial é agrupada por sexo e por coortes, definidas pelo ano de nascimento, e continuamente atualizada de acordo com as hipóteses de evolução estabelecidas para cada uma das componentes de mudança da população e pelo natural envelhecimento anual.

O método de cálculo dos efetivos populacionais anuais é expresso pela equação básica:

$$P_t = P_{t-1} + NV_t - Ob_t + I_t - E_t \quad (31)$$

Sendo P_t a população obtida para o ano t (tendo como momento de referência 31 de dezembro), resultante da população de partida no ano $t-1$ (P_{t-1}), a que se adiciona o número de nados-vivos do ano t (NV_t), se subtrai o número de óbitos do ano t (Ob_t), se adiciona o número de imigrantes no ano t (I_t) e se subtrai o número de emigrantes no ano t (E_t).

O cálculo é efetuado para cada sexo em separado e por idades singulares, o que requer algumas adaptações da equação.

É ainda efetuado para cada cenário e região, sendo assegurada a coerência dos dados em cada cenário entre os valores regionais e o valor nacional.

As populações residentes projetadas referem-se a 31 de dezembro do ano de referência da projeção.

Trata-se de resultados condicionais, do tipo “se x então y ”, que ilustram o que se pode obter no futuro em termos de dimensão e de estrutura da população por sexo e idade, caso se verifiquem os pressupostos em que assentam os cálculos, dependendo, por um lado, da estrutura e composição da população no momento de partida (2018) e, por outro, dos diferentes padrões de comportamento da fecundidade, da mortalidade e migrações estabelecidos em cada uma das hipóteses ao longo do período de projeção 2019 - 2080.

○ **V.7.4 Ajustamentos dos dados**

Não aplicável.

○ **V.7.5 Comparabilidade e coerência**

Face ao seu carácter condicional, que depende, por um lado, da dimensão e estrutura por sexo e idade da população no momento de partida e, por outro lado, dos diferentes padrões de comportamento da fecundidade, da mortalidade e migrações delineados em cada cenário, os resultados obtidos no presente exercício não são diretamente comparáveis com os resultados de exercícios anteriores.

○ **V.7.6 Confidencialidade dos dados**

<p>1. Indicar se a operação estatística é objeto de tratamento de confidencialidade dos dados (<u>resposta única</u>):</p> <ul style="list-style-type: none"> • Sim <input type="checkbox"/> • Não <input checked="" type="checkbox"/> <p>2. Se responder “sim” completar a resposta de acordo com a instrução de preenchimento.</p>
--

VI. SUPORTES DE RECOLHA E VARIÁVEIS DE OBSERVAÇÃO

1. Variáveis de recolha não-direta:

Fonte (dos dados)	Designação da variável (na origem)	Unidade estatística	Unidade de medida	Informação complementar
Estimativas Anuais da População Residente	População	Indivíduo	N.º	A operação estatística “Estimativas Anuais de População Residente” é um estudo estatístico que se realiza anualmente, de âmbito nacional, tendo por base informação de outras operações estatísticas do INE: dados vivos; óbitos e

				estimativas de fluxos migratórios internacionais (fluxos de entrada em Portugal ou imigração e fluxos de saída de Portugal ou emigração), assim como informação proveniente do Serviço de Estrangeiros e Fronteiras (SEF).
Nados-vivos	Nado-vivo	Indivíduo / nado-vivo	N.º	A operação estatística “Nados-vivos” resulta do aproveitamento de factos obrigatoriamente sujeitos a registo civil (nascimentos de crianças nascidas vivas) para recolha de informação complementar, estatisticamente relevante, junto das Conservatórias do Registo Civil (CRC).
Óbitos	Óbito	Indivíduo/ óbito	N.º	A operação estatística “Óbitos” resulta do aproveitamento de factos obrigatoriamente sujeitos a registo civil (óbito) para recolha de informação complementar, estatisticamente relevante, junto das Conservatórias do Registo Civil (CRC).
Estimativas de emigrantes	Emigrantes	Indivíduo	N.º	As Estimativas de Emigração são um estudo estatístico que disponibiliza informação estatística anual sobre a dimensão e estrutura por sexo, idade, países ou grupos de países de nacionalidade, de nascimento e de residência anterior, dos fluxos migratórios de saída do território nacional (emigração).
Estimativas de imigrantes	Imigrantes	Indivíduo	N.º	As Estimativas de Imigração são um estudo estatístico que disponibiliza informação estatística anual sobre a dimensão e estrutura por sexo, idade, países ou grupos de países de nacionalidade, de nascimento e de residência anterior, dos fluxos migratórios de entrada (imigração).

VII. VARIÁVEIS DERIVADAS

Não aplicável.

VIII. INDICADORES A DISPONIBILIZAR

Indicador		Variável medida			Dimensões de análise					
					Código	Data início vigência	Designação	Classificação/versão associada		
		Código	Data início vigência	Designação				Código	Designação	Nível
13931	População residente (projeções 2018-2080 - N.º) por Local de residência (NUTS - 2013), Sexo, Idade e Cenário; Anual	12736	21-02-2020	População residente (projeções 2018-2080 - N.º)	190	03-08-2005	Período de referência dos dados	-	-	-
					6148	17-07-2015	Local de residência (NUTS - 2013)	V03591	NUTS 2013 (PT, NUTS II) - variante 1	2
					310	01-01-2005	Sexo	V00305	Sexo (difusão-HM)	2
					13640	20-01-2017	Idade	V01493	Idades (0;<=100 - projeções da população)	2
					12744	01-09-2014	Cenário	V03489	Cenários (de projeção de população residente)	1
13932	Esperança de vida à nascença (projeções 2018-2080 - Ano) por Local de residência (NUTS - 2013), Sexo e Cenário; Anual	12737	21-02-2020	Esperança de vida à nascença (projeções 2018-2080 - Ano)	190	03-08-2005	Período de referência dos dados	-	-	-
					6148	17-07-2015	Local de residência (NUTS - 2013)	V03591	NUTS 2013 (PT, NUTS II) - variante 1	2
					6073	16-12-2008	Sexo	V01485	Sexo (difusão)	1
					12744	01-09-2014	Cenário	V03489	Cenários (de projeção de população residente)	1
13933	Índice sintético de fecundidade (projeções 2018-2080 - N.º) por Local de residência (NUTS - 2013) e Cenário; Anual	12738	21-02-2020	Índice sintético de fecundidade (projeções 2018-2080 - N.º)	190	03-08-2005	Período de referência dos dados	-	-	-
					6148	17-07-2015	Local de residência (NUTS - 2013)	V03591	NUTS 2013 (PT, NUTS II) - variante 1	2
					12744	01-09-2014	Cenário	V03489	Cenários (de projeção de população residente)	1
13934	Saldo migratório (projeções 2018-2080 - N.º) por	12739	21-02-2020	Saldo migratório (projeções 2018-2080 - N.º)	190	03-08-2005	Período de referência dos dados	-	-	-

	Local de residência (NUTS - 2013) e Cenário; Anual				6148	17-07-2015	Local de residência (NUTS - 2013)	V03591	NUTS 2013 (PT, NUTS II) - variante 1	2
					12744	01-09-2014	Cenário	V03489	Cenários (de projeção de população residente)	1
13935	Índice de dependência de idosos (projeções 2018-2080 - N.º) por Local de residência (NUTS - 2013) e Cenário; Anual	12740	21-02-2020	Índice de dependência de idosos (projeções 2018-2080 - N.º)	190	03-08-2005	Período de referência dos dados	-	-	-
					6148	17-07-2015	Local de residência (NUTS - 2013)	V03591	NUTS 2013 (PT, NUTS II) - variante 1	2
					12744	01-09-2014	Cenário	V03489	Cenários (de projeção de população residente)	1
13936	Índice de dependência de jovens (projeções 2018-2080 - N.º) por Local de residência (NUTS - 2013) e Cenário; Anual	12741	21-02-2020	Índice de dependência de jovens (projeções 2018-2080 - N.º)	190	03-08-2005	Período de referência dos dados	-	-	-
					6148	17-07-2015	Local de residência (NUTS - 2013)	V03591	NUTS 2013 (PT, NUTS II) - variante 1	2
					12744	01-09-2014	Cenário	V03489	Cenários (de projeção de população residente)	1
13937	Índice de envelhecimento (projeções 2018-2080 - N.º) por Local de residência (NUTS - 2013) e Cenário; Anual	12742	21-02-2020	Índice de envelhecimento (projeções 2018-2080 - N.º)	190	03-08-2005	Período de referência dos dados	-	-	-
					6148	17-07-2015	Local de residência (NUTS - 2013)	V03591	NUTS 2013 (PT, NUTS II) - variante 1	2
					12744	01-09-2014	Cenário	V03489	Cenários (de projeção de população residente)	1
13938	Índice de sustentabilidade potencial (projeções 2018-2080 - N.º) por Local de residência (NUTS - 2013) e Cenário; Anual	12743	27-02-2020	Índice de sustentabilidade potencial (projeções 2018-2080 - N.º)	190	03-08-2005	Período de referência dos dados	-	-	-
					6148	17-07-2015	Local de residência (NUTS - 2013)	V03591	NUTS 2013 (PT, NUTS II) - variante 1	2
					12744	01-09-2014	Cenário	V03489	Cenários (de projeção de população residente)	1
13939	Quociente de mortalidade (projeções 2018 - 2080) por Local de residência (NUTS - 2013), Sexo, Idade e Cenário; Anual	13125	27-02-2020	Quociente de mortalidade (projeções 2018 - 2080)	190	03-08-2005	Período de referência dos dados	-	-	-
					6148	17-07-2015	Local de residência (NUTS - 2013)	V03591	NUTS 2013 (PT, NUTS II) - variante 1	2
					6073	16-12-2008	Sexo	V01485	Sexo (difusão)	1
					4676	03-06-2008	Idade	V01235	Idades (tábua completa de mortalidade)	1

					12744	01-09-2014	Cenário	v03489	Cenários (de projeção de população residente)	1
13940	Taxa de fecundidade (projeções 2018-2080 - %) por Local de residência (NUTS - 2013), Idade e Cenário; Anual	13123	27-02-2020	Taxa de fecundidade (projeções 2018-2080 - %)	190	03-08-2005	Período de referência dos dados	-	-	-
					6148	17-07-2015	Local de residência (NUTS - 2013)	V03591	NUTS 2013 (PT, NUTS II) - variante 1	2
					13124	20-07-2015	Idade	V03634	Idades (14;50 -- projeções da população)	2
					12744	01-09-2014	Cenário	03489	Cenários (de projeção de população residente)	1

IX. CONCEITOS

Código	Designação	Definição
199	Nado vivo	Produto do nascimento vivo (Vide Nascimento Vivo).
202	Óbito	Cessaçã irreversível das funções do tronco cerebral.
3374	Nascimento vivo	É a expulsão ou extração completa, relativamente ao corpo materno e independentemente da duração da gravidez, do produto da fecundação que, após esta separação, respire ou manifeste quaisquer outros sinais de vida, tais como pulsações do coração ou do cordão umbilical ou contração efetiva de qualquer músculo sujeito à ação da vontade, quer o cordão umbilical tenha sido cortado, quer não, e quer a placenta esteja ou não retida.
3800	Esperança de vida à nascença (e_0)	Número médio de anos que uma pessoa à nascença pode esperar viver, mantendo-se as taxas de mortalidade por idades observadas no momento.
180	Idade	Intervalo de tempo que decorre entre a data do nascimento (dia, mês e ano) e as 0 horas da data de referência. A idade é expressa em anos completos, salvo se tratar de crianças com menos de 1 ano, devendo nestes casos ser expressa em meses, semanas ou dias completos.
171	Emigrante permanente	Pessoa (nacional ou estrangeira) que, no período de referência, tendo permanecido no país por um período contínuo de pelo menos um ano, o deixou com a intenção de residir noutro país por um período contínuo igual ou superior a um ano.
185	Imigrante permanente	Pessoa (nacional ou estrangeira) que, no período de referência, entrou no país com a intenção de aqui permanecer por um período igual ou superior a um ano, tendo residido no estrangeiro por um período contínuo igual ou superior a um ano.
192	Migração	Deslocação de uma pessoa através de um determinado limite espacial, com intenção de mudar de residência de forma temporária ou permanente. A migração subdivide-se em migração internacional (migração entre países) e migração interna (migração no interior de um país).
208	População residente	Conjunto de pessoas que, independentemente de estarem presentes ou ausentes num determinado alojamento no momento de observação, viveram no seu local de residência habitual por um período contínuo de, pelo menos, 12 meses anteriores ao

		<p>momento de observação, ou que chegaram ao seu local de residência habitual durante o período correspondente aos 12 meses anteriores ao momento de observação, com a intenção de aí permanecer por um período mínimo de um ano.</p> <p>Notas: este conceito é utilizado no Recenseamento Geral da População (CENSO), pelo que o momento de observação se reporta ao momento censitário e é extensível às Estimativas de População Residente, cuja população de partida se reporta também ao momento censitário.</p>
3642	Residência principal /habitual	Alojamento que constitui a residência de pelo menos um agregado familiar durante a maior parte do ano, ou para onde um agregado tenha transferido a totalidade ou maior parte dos seus haveres.
217	Saldo migratório	<p>Diferença entre o número de entradas e saídas por migração, internacional ou interna, para um determinado país ou região, num dado período de tempo.</p> <p>Notas: O saldo migratório pode também ser calculado pela diferença entre a variação populacional e o saldo natural</p>
218	Saldo natural	Diferença entre o número de nados vivos e o número de óbitos, num dado período de tempo.
187	Índice de dependência de idosos	Relação entre a população idosa e a população em idade ativa, definida habitualmente como o quociente entre o número de pessoas com 65 ou mais anos e o número de pessoas com idades compreendidas entre os 15 e os 64 anos (expressa habitualmente por 100 (10^2) pessoas com 15-64 anos).
188	Índice de dependência de jovens	Relação entre a população jovem e a população em idade ativa, definida habitualmente como o quociente entre o número de pessoas com idades compreendidas entre os 0 e os 14 anos e o número de pessoas com idades compreendidas entre os 15 e os 64 anos (expressa habitualmente por 100 (10^2) pessoas com 15-64 anos).
190	Índice de envelhecimento	Relação entre a população idosa e a população jovem, definida habitualmente como o quociente entre o número de pessoas com 65 ou mais anos e o número de pessoas com idades compreendidas entre os 0 e os 14 anos (expressa habitualmente por 100 (10^2) pessoas dos 0 aos 14 anos).
191	Índice sintético de fecundidade (ISF)	Número médio de crianças vivas nascidas por mulher em idade fértil (dos 15 aos 49 anos de idade), admitindo que as mulheres estariam submetidas às taxas de fecundidade observadas no momento. Valor resultante da soma das taxas de fecundidade por idades, ano a ano ou grupos quinquenais, entre os 15 e os 49 anos, observadas num determinado período (habitualmente um ano civil).
158	Variação populacional	Diferença entre os efetivos populacionais em dois momentos do tempo (habitualmente dois fins de ano consecutivos). A variação populacional pode ser calculada pela soma algébrica do saldo natural e do saldo migratório.
3808	Índice de sustentabilidade potencial	Relação entre a população em idade ativa e a população idosa, definida habitualmente como o quociente entre o número de pessoas com idades compreendidas entre os 15 e os 64 anos e o número de pessoas com 65 ou mais anos (expressa habitualmente por cada pessoa (10^2) com 65 ou mais anos).
211	Relação de masculinidade à nascença	Quociente entre os nados-vivos do sexo masculino e os do sexo feminino, ocorridos num determinado período (habitualmente expresso por 100 (10^2) nados-vivos do sexo feminino).
7499	Quociente de mortalidade	Probabilidade de um indivíduo com idade x falecer antes de alcançar a idade (x+1).
228	Taxa de fecundidade geral	Número de nados-vivos observado durante um determinado período de tempo, normalmente um ano civil, referido ao efetivo médio de mulheres em idade fértil (entre os 15 e os 49 anos) desse período (habitualmente expressa em número de nados-vivos por 1000 (10^3) mulheres em idade fértil).

X. CLASSIFICAÇÕES

- Lista de classificações utilizadas:

Código (versão)	Designação (versão)	Sigla
V00305	Sexo (difusão-HM)	
V01235	Idades (tábua completa de mortalidade)	
V03591	NUTS 2013 (PT, NUTS II) - variante 1	
V01493	Idades (0;<=100 - projeções da população)	
V03489	Cenários (de projeção de população residente)	
V03634	Idades (14;50 - projeções da população)	
V01485	Sexo (difusão)	

XI. SIGLAS E ABREVIATURAS

- Lista de abreviaturas e acrónimos (siglas) utilizados:

Código	Designação	Extensão
7185	CGA	Classificação Geral das Atividades
7744	CRC	Conservatória do Registo Civil
5868	DES/DM	Departamento de Estatísticas Demográficas e Sociais / Serviço de Estatísticas Demográficas
4492	DMET	Documento Metodológico
4134	EUROSTAT	Serviço de Estatística da União Europeia
10444	GLM	Modelo Linear Generalizado
4172	INE	Instituto Nacional de Estatística, I. P.
8209	ISF	Índice Sintético de Fecundidade
10446	LC	Lee-Carter
11342	MQO	Mínimos Quadrados Ordinários
4201	NUTS	Nomenclatura das Unidades Territoriais para Fins Estatísticos
10445	PLC	Poisson Lee-Carter
4225	SEF	Serviço de Estrangeiros e Fronteiras

4229	SIGINE	Sistema de Informação de Gestão do INE
10447	SDV	Singular Value Decomposition
4578	UO	Unidade Orgânica
11340	VAR	Modelo Vetorial Autorregressivo

XII. BIBLIOGRAFIA

- Ayuso, M., Bravo, J. M. & Holzmann, R. (2017a). Addressing Longevity Heterogeneity in Pension Scheme Design. *Journal of Finance and Economics*. Volume 6, No. 1 (2017), 1-21. doi: 10.12735/jfe.v6n1p1
- Ayuso, M., Bravo, J. M. & Holzmann, R. (2017b). On the Heterogeneity in Longevity among Socioeconomic Groups: Scope, Trends, and Implications for Earnings-Related Pension Schemes. *Global Journal of Human Social Sciences - Economics*, 17 (1) Version 1.0, 31-57.
- Ayuso, M., Bravo, J. M. & Holzmann, R. (2019). Getting Life Expectancy Estimates Right for Pension Policy: Period versus Cohort Approach. *Journal of Pension Economics and Finance*. In Press.
- Brass, W. (1971). *On the scale of mortality*. In: *Biological Aspects of Demography*, London Taylor and Francis.
- Brass, W. (1973). *Mortality models and their uses in demography*. *Transactions of the Faculty of Actuaries*, 33, 123-132.
- Bravo, J. M. (2007). *Tábuas de Mortalidade Contemporâneas e Prospectivas: Modelos Estocásticos, Aplicações Actuarias e Cobertura do Risco de Longevidade*. Dissertação de Doutoramento em Economia pela Universidade de Évora. DOI: 10.13140/RG.2.1.3907.3041.
- Bravo, J. M. (2019). Funding for Longer Lives: Retirement Wallet and Risk-Sharing Annuities. *EKONOMIAZ Basque Economic Review*, N° 96 (II-2019), 268–291.
- Bravo, J. M. & Coelho, E. (2020). Forecasting small population monthly fertility and mortality data with seasonal time series methods. In: Linhares, W. (Ed.). *As Ciências Sociais Aplicadas e a Interface com vários Saberes 2*, Atena Editora, 158-176. doi: 10.22533/at.ed.79020280112.
- Bravo, J. M. & Coelho, E. (2020). Modelling monthly births and deaths using Seasonal Forecasting Methods as an input for population estimates. In: *Demography of Population Health, Aging and Health Expenditures*. The Springer Series on Demographic Methods and Population Analysis, Springer. In Press.
- Bravo, J. M. & Coelho, E. (2019). Forecasting Subnational Demographic Data using Seasonal Time Series Methods. In: *Proceedings of the 19th Portuguese Association of Information Systems Conference: digital disruption: living between data science, IoT and ... people* (pp. 40). Associação Portuguesa de Sistemas de Informação. ISSN 2183-489X

- Bravo, J. M., & El Mekkaoui de Freitas, N. (2018). Valuation of longevity-linked life annuities. *Insurance: Mathematics and Economics*, 78, 212–22.
- Bravo, J. M., & Nunes, J. P. (2019). Pricing Longevity Swaps via Option Decomposition and Fourier Transforms. Preprint submitted to *Insurance: Mathematics and Economics*.
- Bravo, J. M., Ayuso, M., Holzmann, R. & Palmer, E. (2019). Addressing Life Expectancy Gap in Pension Policy. Preprint submitted to *Insurance: Mathematics and Economics*.
- Bravo, J. M., Magalhães, M. G. & Coelho, E. (2007). Mortality and Longevity Projections for the Oldest-Old in Portugal. In: EUROSTAT - European Commission (eds.), Working Session on Demographic Projections, Bucharest, Romania, EUROSTAT-EC Collection: Methodologies and working papers, Theme: Population and Social Conditions, ISBN: 978-92-79-04759-6, ISSN: 1977-0375, pp. 117-132.
- Bravo, J. M., Magalhães, M. G. & Coelho, E. (2013). Dinâmica e Estrutura da População Humana: Medidas e Modelos Matemáticos. In: *Matemática do Planeta Terra*. In: J. Buescu, F.P. da Costa e J.T. Pinto (Eds), Instituto Superior Técnico Press, pp. 503-541.
- Brouhns, N., Denuit, M. e Vermunt, J. (2002). A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics*, 31, 373-393.
- Brouhns, Natacha, Denuit, Michel and Van Keilegom, Ingrid. (2005). Bootstrapping the Poisson log-bilinear model for mortality forecasting', *Scandinavian Actuarial Journal*, 2005:3, 212 — 224.
- Coelho, E., Bravo, J. M. & Magalhães, M. G. (2010). Mortality projections in Portugal. In: EUROSTAT - European Commission (eds.), Work Session on Demographic Projections, EUROSTAT-EC Collection: Methodologies and working papers, Theme: Population and Social Conditions, pp. 241-252. ISBN 978-92-79-15862-9, doi:10.2785/50697.
- Delwarde, A. e Denuit, M. (2005), Construction de tables de mortalite periodiques et prospectives. *Assurance Audit Actuariat*. Economica.
- Denuit, M. and Goderniaux, A. (2005). Closing and projecting life tables using log-linear models. *Bulletin de l'Association Suisse des Actuaries*, 1, 29-49.
- Eilers, P. e Marx, B. (1996). Flexible smoothing using B-Splines and penalized likelihood (with comments and Rejoinder). *Statistical Science*, 11(2), 89-121.
- Forfar, D., McCutcheon, J. e Wilkie, A. (1988). On Graduation by Mathematical Formula. *Journal of the Institute of Actuaries*, 115, 1-149.
- Goodman, L. (1979). Simple models for the analysis of association in cross classifications having ordered categories. *Journal of the American Statistical Association*, 74, 537-552.
- Granger, Clive WJ. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica Journal of the Econometric Society*, 424–38.
- Hunt, A., and Blake, D. (2020). On the structure and classification of mortality models. *North American Actuarial Journal*. DOI: 10.1080/10920277.2019.1649156.

- Hyndman, R. J., and Booth, H. (2008). Stochastic population forecasts using functional data models for mortality, fertility and migration. *International Journal of Forecasting*, 24(3), 323–342.
- Hyndman, R. J., and Ullah, S. (2007). Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics Data Analysis*, 51(10), 4942–4956.
- Hyndman, R. J., Booth, H., & Yasmeen, F. (2013). Coherent mortality forecasting: the product-ratio method with functional time series models. *Demography*, 50(1), 261-283.
- Lee, R. e Carter, L. (1992). *Modelling and forecasting the time series of US mortality*. *Journal of the American Statistical Association*, 87, 659-671.
- Lindsey, J. (1997). *Applying Generalized Linear Models*. Springer-Verlag, New York.
- McCullagh, P e Nelder, J. (1989). *Generalized Linear Models*. Second Edition, Chapman & Hall, London
- McCutcheon, J. (1987). *Experiments in graduating the data for the English Life Tables N.º 14*. *Transactions of the Faculty of Actuaries*, 40, 135-147.
- R.J. Hyndman. “Demography: Forecasting Mortality, Fertility, Migration and Population Data.” R Package version 1.22. With Contributions from Heather Booth and Leonie Tickle and John Maindonald. Available online.
- Renshaw, A. e Haberman, S. (2003). *Lee-Carter mortality forecasting with age specific enhancement*. *Insurance: Mathematics and Economics*, 33, 255-272.
- Schmertmann C. (2003). A system of model fertility schedules with graphically intuitive parameters. *Demographic Research*, 9:81-110.
- Schmertmann C. (2005). Quadratic spline fits by nonlinear least squares. *Demographic Research*, 12:105-106.
- Shang, H. (2013). *ftsa: An R package for analyzing functional time series*. *The R Journal* Vol. 5/1, June 2013. ISSN 2073-4859.
- Shang, H. L., Booth, H., & Hyndman, R. J. (2011). Point and interval forecasts of mortality rates and life expectancy: A comparison of ten principal component methods. *Demographic Research*, 25, 173-214.
- Siegel, J. S. & Swanson (eds.) (2004). *The Methods and Materials of Demography*, 2nd edition, San Diego: Elsevier, Academic Press.
- Tomas, J. & Planchet, F. (2014), *Constructing entity specific prospective mortality table: Adjustment to a reference*, *European Actuarial Journal* 2/2014.
- Tsay, R. S. (2010). *Analysis of Financial Time Series*. John Wiley & Sons.
- Wilmoth, J.R., Andreev, K., Jdanov, D. e Gleit, D. (2007). *Methods Protocol for the Human Mortality Database*. Human Mortality Database.