

# ASPECTOS ESTÁTICOS DA ANÁLISE DE POPULAÇÃO

## 6.1 A IDADE COMO CONCEITO DEMOGRÁFICO CENTRAL

Como foi notado no capítulo anterior, as estatísticas demográficas se dividem entre quantidades de estoque (que descrevem uma característica ou situação existente num determinado momento) e de fluxo (que quantificam os processos que mudam tal situação). A maior parte da análise demográfica se ocupa com a dinâmica da população, ou seja, com a forma como as variáveis de fluxo (nascimentos, óbitos e migrações) afetam as variáveis de estoque. Mas antes de entrar nesta análise, vale a pena primeiro olhar alguns conceitos descritivos de população que se baseiam nas variáveis de estoque.

Muitos dos dados analisados na demografia são parecidos com os que se manejam em outras disciplinas. Mas também existem algumas particularidades e alguns instrumentos específicos para ajudar a análise. Uma destas particularidades tem a ver com a centralidade do conceito de “idade” na demografia. Quase não existe qualquer indicador na demografia cujo significado não seja condicionado pelo fator idade. A pergunta “qual é a probabilidade de que uma pessoa atualmente viva continue viva por mais 10 anos ?” tem pouco sentido sem saber a idade (e o sexo) atual da pessoa. Da mesma forma, a pergunta “qual é a probabilidade de que uma mulher tenha um filho durante o próximo ano ?” é difícil de responder sem saber a sua idade. E assim por diante, com a grande maioria dos indicadores demográficos.

A forma como a idade das pessoas é identificada na demografia às vezes causa certa confusão. Na demografia usa-se tanto o conceito de *idade exata* como o conceito de *idade em anos completos*. O conceito de idade exata se refere a um número matemático que não permite nenhuma

ambiguidade: 15 anos é exatamente 15,0000000 anos, não 15,0000001, nem 14,99999999. Mas evidentemente não existe nenhuma pessoa que tenha exatamente 15 anos neste sentido excessivamente restrito. Por isso, o conceito exato de idade não é usado para identificar as idades de pessoas específicas, mas só para delimitar faixas etárias (também chamadas “grupos etários” ou, em Portugal, “escalões etários”) ou para calcular médias. Por exemplo, quando se fala das pessoas com menos de 15 anos, a idade exata de 14,9999..... anos é incluída neste grupo, mas a idade exata de 15,0000 anos não. Por outro lado, quando comumente se fala de uma pessoa com a idade de 14 anos, sem algarismos atrás da vírgula, a ideia subjacente é a de *idade em anos completos*. Em realidade, uma pessoa assim descrita pode ter qualquer idade exata entre 14,0000..... e 14,9999..... anos, mas não pode ter 15,0000 anos. A idade média, em termos exatos, destas pessoas não é 14,0, mas 14,5 anos. Da mesma forma, a faixa etária (grupo etário, escalão etário) convencionalmente descrita como 15-19 anos em realidade se refere a pessoas com 15,0000.....-19,9999..... anos, ou seja, pessoas que já alcançaram o seu 15º aniversário, mas que ainda não alcançaram o 20º. Em termos exatos, a idade média destas pessoas – supondo que estão distribuídas homogeneamente no intervalo – é 17,5 anos. Não é correto descrever esta faixa etária como a de “15-20 anos”, pois isso incluiria pessoas que já passaram o seu 20º aniversário. Tampouco é correto descrever o grupo como o de “pessoas de mais de 15 e menos de 19 anos”, pois ela inclui pessoas de até 19,99999999 anos exatos. O mesmo acontece com anos calendário. O ano calendário 2014 vai desde 2014,0000.... até 2014,9999....., mas não inclui 2015,0000.

Em populações com uma tradição de língua escrita, a idade das pessoas é um dado cronológico amplamente documentado em registros e documentos de identidade, mas em sociedades de tradição oral a idade frequentemente não cumpre esta mesma função e é vista mais como uma característica de status social. Não são raros os casos descritos por antropólogos em que os homens chefes de família em aldeias tradicionais da África sistematicamente descreveram a sua própria idade como 65, a das suas esposas como 25 e dos seus filhos como 5 (Pison e Ohadike, 2006). Nesses casos uma pergunta direta sobre a idade da pessoa não tem muito sentido e pode ser preferível usar outros recursos, como o calendário histórico (em que a data de nascimento das pessoas é relacionada com certos eventos históricos locais que as pessoas podem identificar facilmente) ou o relacionamento entre as idades de diferentes pessoas na comunidade (estabelecendo se a pessoa X nasceu antes ou depois da pessoa Y). O maior problema do calendário histórico reside na dificuldade de encontrar eventos que tenham significado para toda a população, tanto nas cidades como nas áreas rurais mais remotas. Os exemplos mais elaborados de calendários locais se encontram em pesquisas antropológicas detalhadas em pequena escala. Esses exercícios exigem um treinamento bastante intensivo dos entrevistadores, para que sejam aplicados corretamente (Pison, 1980; Kpedekpo, 1982: 41-42; Axinn e Pearce, 2006: Cap. 5). Entretanto, mesmo alguns censos nacionais usam calendários básicos para obter estimativas aproximadas das idades de pessoas cuja idade não pode ser estabelecida de outra forma. O Quadro 6.1 ilustra isso com o calendário usado no Censo de Moçambique de 2017. Os recenseadores também receberam as seguintes instruções adicionais para estimar as idades desconhecidas:

Comparar as características físicas da pessoa cuja idade se desconhece com a de outra pessoa com idade conhecida. Atenção especial deve ser dada às crianças mais pequenas: se elas já andam devem ter mais ou menos 1 ano, se já falam devem ter cerca de 2 anos.

Para determinar a idade de uma criança, se a mãe tiver mais filhos, deve-se proceder da seguinte forma:

- Saber a idade de um dos filhos;
- Seguindo a ordem e intervalo de nascimento dos filhos pode calcular a idade do filho cuja idade se desconhece (Manual do Recenseador: 42-43).

Quadro 6.1: Exemplo de um calendário histórico para aproximar as idades das pessoas com idade desconhecida no Censo de Moçambique de 2017

1960	Massacre de Mueda
1964	Início Luta Armada de Libertação Nacional
1969	Morte de Eduardo Mondlane
1974	Acordos de Lusaka
1975	Independência de Moçambique
1980	I Recenseamento Geral da População e Habitação
1983	Acordo de Nkomati / Operação produção
1986	Morte de Samora Machel
1992	Acordos de Paz / Acordo de Roma
1994	Primeiras eleições presidenciais e legislativas
1997	II Recenseamento Geral da População e Habitação
1998	Primeiras Eleições Autárquicas
1999	Segundas eleições presidenciais e legislativas
2000	Cheias nas Províncias de Sofala, Inhambane, Gaza, Manica e Maputo
2001	Cheias nas Províncias de Sofala, Inhambane e Zambézia
2003	Segundas Eleições Autárquicas
2004	Terceiras eleições presidenciais e legislativas
2008	Terceiras Eleições Autárquicas
2009	Quartas eleições presidenciais, legislativas e provinciais
2013	Quartas Eleições Autárquicas / Greve geral dos médicos
2014	Quintas eleições presidenciais, legislativas e provinciais

Fonte: Manual do Recenseador do Censo de 2017 de Moçambique.

Um outro problema é que algumas culturas, embora tenham o conceito de idade mais internalizado, o contam de uma forma diferente. Por exemplo em Macau, bem como em outros lugares com populações predominantemente chinesas, o sistema ocidental de contagem da idade coexiste com o sistema tradicional chinês em que se conta o número de anos diferentes (de acordo com o calendário chinês) em que a pessoa viveu. Segundo este sistema, uma criança nascida pouco antes do Ano Novo já teria 2 anos de idade logo depois da virada do ano. Entretanto, hoje em dia este sistema é pouco usado, exceto por pessoas de maior idade.

Na notação padronizada de quantidades demográficas usada neste livro, que na maioria dos casos segue as práticas internacionais, a idade é simbolizada por  $x$  e a amplitude de um intervalo etário por  $n$ . Em outros textos também se pode encontrar o símbolo  $a$  para idade (de “age”, em inglês e “âge”, em francês), mas esta convenção parece menos apropriada para a língua portuguesa ou espanhola. Números absolutos são simbolizados por maiúsculas:  $P$  para população,  $D$  para defunções e  $N$  para nascimentos. Alguns autores usam  $O$  (de “óbitos”) para as defunções, mas esta prática distancia-se desnecessariamente da internacional (“deaths”, em inglês, “décès”, em francês,

“defunciones” em espanhol). Por outro lado, o símbolo  $N$  parece mais apropriado do que  $B$  (de “births”) que geralmente se usa na literatura de língua inglesa. Sendo assim, a população contida no intervalo de 5 anos que começa na idade exata de 15 anos se nota como

$${}_5P_{15} = \text{População de 15 a 19 anos inteiros (15,0000.... a 19,9999..... anos exatos)} \quad (6.1)$$

onde  $x=15$  e  $n=5$ . É importante enfatizar que, quando  $x$  e  $n$  são números inteiros, como acontece quase sempre, o intervalo se estende de  $x$  a  $x+n-1$  e não a  $x+n$  anos completos. Para sinalizar que esta informação se refere à situação num determinado momento  $t$  (de “tempo” ou “time”, em inglês), acrescenta-se esta informação entre parênteses:

$${}_5P_{50}(2015) = \text{População de 50-54 anos inteiros existente no 1º de janeiro de 2015} \quad (6.2)$$

Para o último intervalo, que é aberto para cima, geralmente se usa a notação  $x+$ , omitindo o índice da esquerda, ou então se usa  $x$  com um índice  $\omega$  na esquerda:

$${}_{\omega}P_x(t) = P_{x+}(t) \quad (6.3)$$

Quando a quantidade é de fluxo, como nascimentos ou defunções, e não de estoque, é preciso indicar tanto o início como o final do período de observação:

$${}_{5}D_{50}(2015,2020) = \text{Defunções ocorridas entre o 1º de janeiro de 2015 e o 31º de dezembro de 2019, de pessoas que no momento da sua morte tinham 50-54 anos completos} \quad (6.4)$$

Quando se trata de nascimentos, em vez de defunções, os símbolos  $x$  e  $n$  se referem às idades das mães, não dos filhos. Ao tratar-se de números proporcionais, taxas ou probabilidades, geralmente se usam letras minúsculas, por exemplo

$${}_5p_{50}(2015) = \text{Proporção da população com 50-54 anos inteiros no 1º de janeiro de 2015} \quad (6.5)$$

Apesar de ser uma característica básica do indivíduo é preciso ter consciência de que a declaração da idade das pessoas no censo e outras fontes de dados pode ter erros, às vezes até erros significativos. Como se verá no Capítulo 16, existem métodos para avaliar a qualidade da declaração da idade em anos simples, que constituem um dos métodos mais comuns para qualificar a qualidade da informação censitária como um todo. Um dos problemas mais frequentes é que as pessoas (ou o entrevistador), por preguiça ou por realmente não saber, podem arredondar a idade, declarando uma idade de 79 anos como 80 ou 36 como 35. Isso se nota nos resultados dos censos pelo maior número de idades que terminam em “0” ou “5”. O mesmo pode acontecer também porque o entrevistado quer assumir um status legal (por exemplo, de adulto) que em realidade ainda não tem, declarando-se com uma idade de 18 anos quando em realidade tem 17. Finalmente,

não dá para subestimar o interesse do entrevistador em apresentar os dados de certas maneiras para facilitar o seu trabalho. Por exemplo, ele ou ela pode preferir classificar uma mulher de 48 ou 49 anos como 50 pois, dependendo da estrutura do questionário, isso pode evitar a necessidade de fazer uma série de perguntas sobre os filhos que ela teve na sua vida, já que estas perguntas geralmente são feitas só para mulheres de 15-49 anos. Por todas estas razões, além da economia de espaço, os dados sobre idade geralmente são apresentados em faixas de 5 anos, e não por idade simples, o que não acaba com o problema, mas tende a reduzi-lo significativamente. Uma outra maneira para reduzir o problema é pedir a data de nascimento das pessoas, em vez ou além da sua idade, como se faz no censo do Brasil. As datas de nascimento não são livres de erros, mas geralmente estes são menos graves do que os erros na declaração da idade e, no caso onde se dispõe de ambas, é possível verificar a consistência da informação.

Um dos recursos mais conhecidos para representar a distribuição das idades numa população é a *pirâmide etária* (que geralmente é específica por *sexo* também). Alguns exemplos de pirâmides por idade e sexo já foram vistos no Capítulo 2. Trata-se de um diagrama com a população masculina na esquerda e a população feminina na direita. O tamanho da população em cada faixa etária, que pode ser em números absolutos ou em percentuais, é representado por uma série de barras horizontais, com a idade mais baixa em baixo e a idade mais alta em cima. Quando o tamanho das faixas é formulado como porcentagem ou proporção, é importante que tais proporções sejam calculadas sobre a população total, *não a população de cada sexo separadamente*. Este último erro é bastante comum e impossibilita a comparação visual entre os números de homens e mulheres num dado grupo etário. As idades podem ser representadas em termos de idades simples ou como intervalos quinquenais ou decenais. Nos exemplos dos Gráficos 6.1 e 6.2 foram usados intervalos quinquenais. O mesmo procedimento descrito abaixo foi usado para produzir as pirâmides por idades simples mostradas no Gráfico 16.4 do Capítulo 16.

Também é possível usar intervalos desiguais (por exemplo, quinquenais para a maioria das idades e decenais para algumas faixas), mas ao fazer isso é preciso ter em conta que as populações devem ser proporcionais à área de cada barra, não à sua *largura horizontal*. Portanto, quando se usa um intervalo etário decenal em vez de quinquenal, é preciso dividir a escala para a largura horizontal das barras por dois. Essa questão de intervalos desiguais geralmente surge quando se tenta definir o último intervalo aberto, de 70+, 80+ ou 85+ anos. Em muitos casos, esta última faixa tem uma população maior do que as imediatamente anteriores. Ao escolher a mesma espessura que foi escolhida para as outras faixas (por exemplo, 5 anos), isso causa uma impressão um pouco estranha, de uma aparente concentração de população na faixa de 70-74, 80-84 ou 85-89, quando efetivamente esta população não pertence só a esta faixa, mas também às faixas seguintes. Por isso, pode ser melhor dividi-la entre duas ou até três faixas, para que a pirâmide tenha uma aparência mais regular.

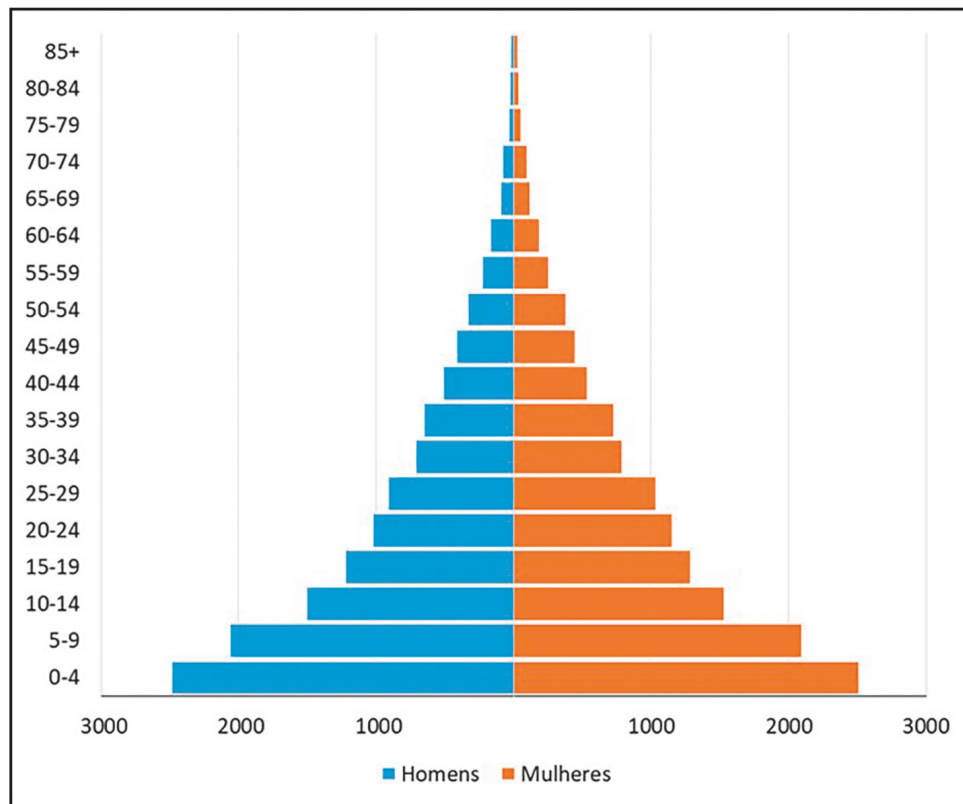
Existem alguns softwares específicos para o desenho de pirâmides etárias, como o programa PYRAMIDS, da Universidade de Wisconsin, feito na década de 80 e alguns programas em “R” que serão discutidos no Capítulo 17. O pacote PASEX (ver seção 17.2 do Capítulo 17) também contém um módulo PYRAMID que executa o procedimento descrito a seguir. É relativamente fácil programar uma pirâmide etária em EXCEL, usando os seguintes passos:

1. Organizar os dados em três colunas: uma com a definição das faixas (0-4, 5-9, 10-14 etc.), uma com as populações masculinas e uma com as populações femininas.
2. A primeira coluna deve ser formatada como “Texto”, para que EXCEL não a leia como datas ou como subtrações de números.
3. Os números da segunda coluna devem ser postos com signo negativo. Depois, na lista de opções de formatação de números, se deve escolher “Personalizado” e digitar a seguinte sequência de caracteres: #####; #####. Se os números são grandes, pode-se acrescentar um ou dois #. O efeito desta especificação é que o EXCEL continua tratando estes números como negativos, mas sem mostrar o seu signo.
4. Selecionar as três colunas e solicitar a sua representação como gráfico de barras horizontais acumuladas.
5. Em Selecionar Dados marcar a primeira série como “Homens” e a segunda como “Mulheres”.
6. Escolher um título que representa o país e o ano ao que se referem os dados.
7. Escolher uma “Largura do Espaçamento” pequena entre as barras, de 5 ou 6%.
8. Escolher a colocação dos rótulos da primeira coluna na esquerda (“Inferior”) e não no meio.
9. Eventualmente ajustar a escala horizontal, para melhor usar o espaço disponível.

O resultado se mostra a seguir no Gráfico 6.1, no caso com os dados do Censo de Angola de 2014.

A pirâmide etária de Angola, com a sua base larga e ápice estreito, é típica de um país com uma alta taxa de crescimento demográfico, alimentado por uma natalidade muito elevada. A pirâmide do Brasil em 1970 era parecida com o Gráfico 6.1, mas em 2015 a parte inferior da pirâmide brasileira já era retangular e a parte triangular característica da pirâmide angolana só aparecia a partir dos 35 anos. A pirâmide de Portugal em 2011 traduzia um processo de evolução demográfica ainda mais adiantada com uma base relativamente estreita, um *aumento* das faixas etárias até o grupo dos 35-39 anos e uma diminuição lenta a partir daí, de tal forma que a largura da barra dos 65-69 anos era parecida com a barra dos 0-4 anos.

Gráfico 6.1: Exemplo de uma pirâmide etária com os dados do Censo de Angola, 2014



Fonte: Recenseamento Geral da População e Habitação: Resultados Definitivos (2016).

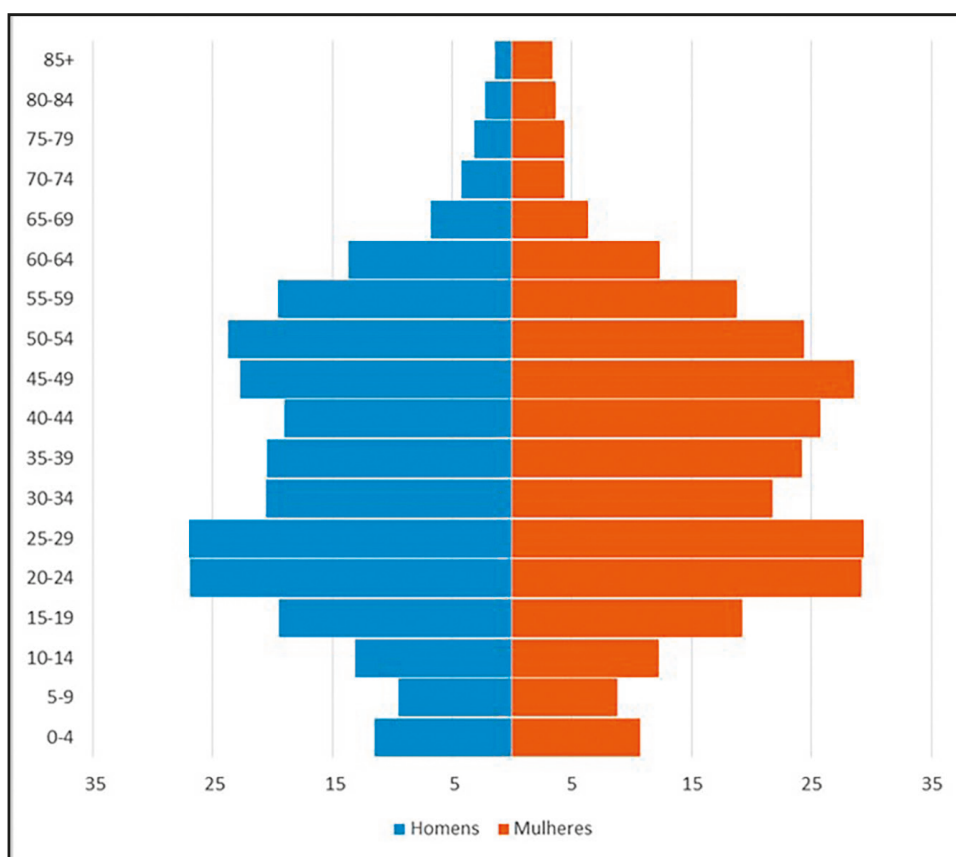
É preciso alertar para uma interpretação equivocada que às vezes se faz de pirâmides como a de Angola, e as diferenças com pirâmides menos triangulares como as do Brasil e especialmente Portugal. Segundo essa interpretação, a forma da pirâmide de Angola seria expressiva de uma *mortalidade* muito alta, na medida em que nascem muitas pessoas, mas aparentemente poucas chegam às idades mais avançadas. O erro deste tipo de interpretação reside em pensar nas diferentes faixas etárias da pirâmide etária como se fossem diferentes fases na vida de um mesmo conjunto de pessoas (uma “coorte”, segundo a terminologia que será introduzida no Capítulo 7). Mas não é assim. As faixas etárias representam pessoas nascidas em distintas épocas quando a população do país tinha tamanhos diferentes. As pessoas na pirâmide de Angola que atualmente têm 60-64 anos, por exemplo, podem ter sofrido um atrito significativo devido à mortalidade elevada, mas a razão principal pelo seu número relativamente reduzido é que já nasceram em números muito menores do que os números atuais de nascimentos, pois 60-65 anos atrás a população como um todo era muito menor do que hoje em dia. Já em Portugal isso não acontece porque não houve tanto crescimento da população entre 1950 e 2011 como houve em Angola devido principalmente a uma *natalidade* muito mais baixa em Portugal. Em realidade o número de nascimentos em Portugal em 1950 era *maior* do que em 2015. Portanto, a natalidade e não a mortalidade é o determinante principal da forma da pirâmide.

Uma situação mais peculiar existe em Macau, cuja pirâmide etária se mostra no Gráfico 6.2. A forma geral da pirâmide é mais ou menos parecida com a de Portugal, no sentido de que a maior parte da população se encontra nas faixas etárias intermédias, entre 20 e 60 anos. Entretanto, esta



característica de alargamento no meio é muito mais extrema no caso do Macau do que em Portugal. Especialmente o número muito reduzido de pessoas de terceira idade, com mais de 65 anos, chama a atenção. Em Portugal, o estreitamento da pirâmide nas idades mais avançadas é muito menos abrupto. Em parte, os determinantes da forma da pirâmide etária em Macau são os mesmos que em Portugal e no Brasil, o seja, um estreitamento da base devido à queda da natalidade. Em 2015, a fecundidade de Macau, em termos do número médio de filhos por mulher, era a segunda mais baixa do mundo, mais baixa ainda do que Portugal. Mas além disso Macau sofreu uma imigração muito significativa da China que se concentrou nas idades economicamente ativas (15-64), o que levou a uma estrutura etária muito enviesada para estas idades. É notável também o desequilíbrio entre o número de homens e mulheres nas idades mais avançadas, algo que ocorre também na pirâmide de Angola, mas é menos visível.

Gráfico 6.2: População por sexo e idade (distribuição relativa): Exemplo de uma pirâmide etária com os dados do Censo de Macau, 2011



Fonte: Censo de População de 2011.

Além da sua representação estática, para um determinado momento no tempo, é muito ilustrativo ver o que acontece quando se junta uma sequência de pirâmides etárias para diferentes momentos no tempo, como se fosse um filme (animação). Isso mostra claramente como a estrutura etária vai se transformando na medida em que diminuem a mortalidade e a natalidade. Não se inclui aqui uma sequência deste tipo, mas existem vários aplicativos na internet que



mostram séries históricas e projeções de países como a França (no sítio de INED), Inglaterra (no sítio web <http://www.neighbourhood.statistics.gov.uk>) e Austrália (no sítio de ABS). Mas talvez o mais interessante seja o sítio [populationpyramid.net](http://populationpyramid.net) que mostra a versão mais recente das estimativas e projeções da Divisão de População das Nações Unidas de 1950 até 2100 para todos os países do mundo. O usuário pode ir para frente ou para trás na história demográfica do país em intervalos de 1 ou 5 anos. Até é possível mostrar os resultados em diferentes idiomas, inclusive o português (PT).

Outro sítio muito didático que mostra estes e outros dados é o de Gapminder, uma ONG fundada pelo médico sueco Hans Rosling ([www.gapminder.org](http://www.gapminder.org)). A animação no sítio de Gapminder não apresenta informação detalhada por sexo, mas sim mostra a evolução da distribuição etária ao longo do tempo. Além disso, o sítio apresenta uma grande variedade de representações de dados demográficos e sociais, das quais os gráficos e animações de “bolhas” são as mais conhecidas. O sítio também contém alguns pequenos vídeos onde o autor tenta demonstrar intuitivamente, com objetos cotidianos, tais como caixas de IKEA ou rolos de papel higiênico, como as estruturas etárias se transformam ao longo do tempo.

## 6.2 ALGUMAS CARACTERIZAÇÕES DA DISTRIBUIÇÃO POR IDADE

A pirâmide etária retrata a estrutura detalhada da distribuição por idades, mas para fins práticos pode ser mais interessante analisar esta distribuição de uma forma mais agregada. Os grupos de idades mais comumente usados são os de crianças, juventude, pessoas em idade economicamente ativa e idosos ou adultos maiores. As definições destes grupos variam. As Nações Unidas definem as crianças como pessoas de até 14 anos completos, a juventude como pessoas na faixa de 15-24 anos e as pessoas idosas ou adultos maiores como pessoas maiores de 60 anos (em países em desenvolvimento) ou 65 anos (em países menos desenvolvidos). Entretanto, como já foi mencionado no Capítulo 3, existem propostas para elevar a idade oficial que marca a terceira idade para 75 anos. A faixa de pessoas em idade ativa é comumente definida como 15-64 anos, mas alguns autores usam 20 anos como o limite inferior e, como será visto no Capítulo 14, as estatísticas do trabalho tendem a usar limites superiores mais altos ou inclusive nenhum limite superior.

Para dimensionar a importância de cada um desses grupos existem alguns índices padrão como os seguintes, ilustrados como os dados do Censo de Angola de 2014:

$$\begin{aligned} \text{Razão de Dependência (Total)} &= \\ &= 100 \text{ (População de 0-14 anos + População de 65+ anos) / População de 15-64 anos} \\ &= 100 (12.196.496 + 612.430) / 12.980.098 = 98,68 \end{aligned} \quad (6.6)$$

A Razão de Dependência Total (RDT) em realidade consiste de duas partes. A primeira é chamada a *Razão de Dependência dos Jovens* (RDJ) e neste caso iguala  $100 \cdot 12.196.496 / 12.980.098 = 93,96$ . A segunda parte se chama a *Razão de Dependência dos Idosos* (RDI) e neste caso iguala  $100 \cdot 612.430 / 12.980.098 = 4,72$ . Esses conceitos já foram usados no Capítulo 3 e especificamente no Gráfico 3.5, para analisar o perfil demográfico do Brasil para efeitos da Previdência Social. No caso de Angola, diferentemente do Brasil, a carga de dependência quase inteiramente vem por conta da sua população extremamente jovem. Já no caso de Portugal, que possui uma estrutura

etária mais envelhecida, a divisão entre as duas partes da Razão de Dependência é mais equilibrada. Em 2015, a RDJ estimada pela Divisão de População das Nações Unidas era 21,57 e a sua Razão de Dependência dos Idosos 22,81.

Os limites de idade (15 e 65 anos) usados na Razão de Dependência são, até certo ponto, arbitrários e sujeitos a definições alternativas. A Divisão de População das Nações Unidas também usa o limite inferior de 20 ou 25 e o limite superior de 70 anos. Combinando esses limites, eles definem cinco índices alternativos, ou seja 1) 0-14 & 65+ / 15-64; 2) 0-19 & 65+ / 20-64; 3) 0-19 & 70+ / 20-69; 4) 0-24 & 65+ / 25-64; e 5) 0-24 & 70+ / 25-69.

Em vez da Razão de Dependência, também se usa o *Razão de Suporte Potencial* e o Índice de Sustentabilidade Potencial, principalmente no contexto das discussões sobre a Previdência Social. A Razão de Suporte Potencial não considera as crianças, apenas a população em idade de trabalhar e a população idosa. Usando os mesmos limites indicados acima, a Divisão de População das Nações Unidas define cinco versões alternativas para a Razão de Suporte, a saber: 1) 15-64 / 65+; 2) 20-64 / 65+; 3) 20-69 / 70+; 4) 25-64 / 65+; e 5) 25-69/70+.

Os economistas também usam a *Razão de Dependência Econômica*. Em vez de referir-se às pessoas em idades *potencialmente* ativas ou inativas, esta se refere aos números de pessoas *efetivamente* ativas ou inativas, ou seja, 100 vezes o número de pessoas inativas dividido pelo número de pessoas ativas. Esses índices, que dependem do conceito de atividade econômica, serão discutidos em mais detalhe no Capítulo 14.

Parecido com a Razão de Dependência é o Índice de Envelhecimento da população, definido da seguinte maneira:

*Índice de Envelhecimento* =

$$100 \text{ População de } 60+ \text{ anos} / \text{População de } 0-14 \text{ anos} = 100 \cdot 968.584 / 12.196.496 = 7,94 \quad (6.7)$$

A relação existente entre a população com 75 e mais anos e a população com 65 e mais anos, muitas vezes chamada o Índice de Longevidade, é um indicador adicional para medir o envelhecimento de uma população. Dois outros indicadores da estrutura etária que serão discutidos mais adiante são as *razões (ou percentagens) de juventude* que expressam o peso do grupo etário de 15-24 anos como porcentagem da população total ou alternativamente da população com 15 anos ou mais.

O envelhecimento da população também pode ser caracterizado de outras maneiras, como a idade média ou mediana da população. Como a idade mediana é mais fácil de calcular, geralmente é a medida preferida. Como todas as medianas, ela é calculada dividindo a população total por 2 e acumulando as populações das faixas etárias a partir de 0 até chegar neste valor. Se o valor é alcançado no meio de um intervalo, se interpola linearmente. Por exemplo, no caso da população de Angola, a metade da população total é  $25.789.024 / 2 = 12.894.512$ . A população acumulada até os 16 anos é 12.757.091 e a acumulada até os 17 anos é 13.294.451. Portanto, a idade mediana é igual a:

$$\text{Idade Mediana} = 16 + (12.894.512 - 12.757.091) / (13.294.451 - 12.757.091) = 16,26 \text{ anos} \quad (6.8)$$

Esta é uma idade mediana muito baixa que implica que a metade da população tem menos de 16,26 anos. Poucos países alcançam valores tão baixos. Hoje em dia, as idades medianas da maioria dos países se encontram entre 25 e 35 anos. Alguns países com estruturas etárias muito envelhecidas, tais como o Japão, têm idades medianas de mais de 40 anos. Portugal, por exemplo, tinha uma idade mediana de 41,99 anos no Censo de 2011. O Gráfico 2.8 do Capítulo 2 mostra a evolução da idade mediana nos países de língua portuguesa ao longo do tempo.

Uma alternativa para a idade mediana é a *idade média*, que é um pouco mais difícil de calcular e sofre uma influência muito grande do tamanho da população idosa. Ela normalmente é maior do que idade mediana. No caso de Angola (2014), o INE a calculou como 20,6 anos, contra 16,26 anos para a idade mediana.

A baixa idade mediana de Angola está relacionada com a alta natalidade no país, embora a relação não seja direta e também passe por outros fatores. Uma medida que expressa esta relação de uma forma um pouco mais direta é a Razão de Crianças / Mulheres, que se calcula da seguinte forma:

$$\begin{aligned} \text{Razão Crianças / Mulheres (RCM)} &= \text{População de 0-4 anos} / \text{Mulheres de 15-49 anos} \\ &= 4.982.427 / 5.995.923 = 0,831 \end{aligned} \quad (6.9)$$

Uma versão alternativa da RCM usa mulheres de 15-44 anos no denominador.

Portugal, com a sua natalidade muito menor, tinha uma razão de 192,2 no Censo de 2011. No Capítulo 10 serão apresentadas outras medidas, mais refinadas, para expressar a natalidade e o nível de reprodução de uma população, mas na ausência destas medidas a Razão de Crianças / Mulheres pode dar uma aproximação bastante razoável do número típico de filhos por mulher. Em realidade, a razão é próxima (com certo viés para baixo) a 4,5 vezes a medida que será introduzida no Capítulo 10 como a Taxa de Fecundidade Geral (TFG).

O que mostra este último parágrafo é que a pirâmide etária ou a distribuição da população por idade e sexo, em que ela se baseia, embora sejam estatísticas de estoque, podem ser usadas para derivar estimativas de certas quantidades de fluxo, como a natalidade, no passado. Esta ideia tem uma aplicabilidade mais geral. Toda a pirâmide etária é, de certa forma, um retrato da história demográfica do país durante as últimas cinco, seis ou sete décadas. A falta de população em certas faixas etárias pode indicar uma queda conjuntural da natalidade – talvez devido a uma recessão econômica – ou uma sobremortalidade devido à guerra ou ainda uma emigração forte durante um período relativamente curto. Entretanto, como todos estes três processos (natalidade, mortalidade e migração) atuam simultaneamente, a pirâmide etária por si só geralmente não é suficiente para reconstruir toda a história, sem a ajuda de elementos adicionais, como séries históricas de taxas de natalidade e/ou mortalidade.

Um exemplo é a distribuição por idades da população de Cabo Verde no Censo de 2010 que mostra uma falta bastante acentuada de população na faixa dos 60-69 anos:

	Homens	Mulheres	Total
50-54	8.162	9.999	18.161
55-59	7.196	7.196	12.143
60-64	2.613	3.580	6.193
65-69	2.499	3.716	6.215
70-74	3.437	5.229	8.666
75-79	2.980	4.454	7.434

Ao comparar o Censo de 2010 com censos anteriores, nota-se uma falta semelhante de população na faixa de 50-59 em 2000, na faixa de 40-49 em 1990 e assim para trás. Na raiz do fenômeno está a conjuntura fortemente adversa que Cabo Verde passou na década de 40, quando a economia local estagnou como consequência da Guerra, a população total diminuiu de 181 mil (1940) para 148 mil (1950), a natalidade caiu num terço e houve crises de fome em 1940-43 e 1946-48, sendo que esta última eliminou 65% da população da Ilha de Santiago. Esta foi a última crise deste tipo entre as muitas que historicamente têm assolado o país (Carreira, 1984 a; Patterson, 1988). Embora os detalhes dos acontecimentos daquela época não possam ser reconstruídos apenas com a informação da pirâmide etária, esta sim transmite sinais claros da calamidade que atingiu o país naqueles anos.

Tabela 6.1: Diversos indicadores da estrutura da população por idade e sexo, baseados nas estimativas e projeções das Nações Unidas para 2020

	Razão de Dependência	Índice de Envelhecimento	Percentagem de Juventude 1	Percentagem de Juventude 2	Idade Mediana	Crianças/Mulher*
Alemanha	55,4	205,2	10,4	12,5	45,7	0,242
Angola	94,5	7,9	19,5	36,4	16,7	0,762
Argentina	55,8	63,6	15,5	20,5	31,5	0,329
Bolívia	60,5	34,3	18,9	27,0	25,6	0,395
Brasil	43,5	67,8	15,7	19,8	33,5	0,253
Cabo Verde	49,0	26,9	17,1	23,7	27,6	0,349
Chile	45,9	90,3	13,8	17,1	35,3	0,241
China	42,2	98,0	11,8	14,3	38,4	0,245
Cuba	46,7	133,5	11,9	14,1	42,2	0,226
EUA	53,9	124,5	13,1	16,1	38,3	0,261
Federação Russa	51,2	122,1	9,4	11,5	39,6	0,273
França	62,4	152,0	11,6	14,1	42,3	0,261
Guiné-Bissau	81,2	10,9	19,7	34,0	18,8	0,620
Guiné Equatorial	64,4	10,3	18,5	29,2	22,3	0,640
Índia	48,7	38,7	18,0	24,4	28,4	0,327
Japão	69,0	275,7	9,3	10,6	48,4	0,192
Macau	35,7	131,5	9,0	10,5	39,3	0,208
México	50,3	43,5	17,2	23,2	29,2	0,314
Moçambique	88,4	9,9	20,6	36,9	17,6	0,682
Portugal	55,8	225,0	10,4	12,0	46,2	0,179
São Tomé & Príncipe	81,0	11,9	19,9	34,2	18,6	0,616
Timor Leste	69,8	17,8	21,4	34,0	20,8	0,548

\* Razão de Dependência =  $100 (\text{População de 0-14 Anos} + \text{População de 65+ anos}) / \text{População de 15-64 Anos}$

Índice de Envelhecimento =  $100 (\text{População de 60+ Anos} / \text{População de 0-14 Anos})$

Percentagem de Juventude 1 =  $100 (\text{População 15-24} / \text{População Total})$

Percentagem de Juventude 2 =  $100 (\text{População 15-24} / \text{População 15 e Mais})$

Idade Mediana = Idade que Divide a População em Duas Metades Iguais

Razão Crianças / Mulheres =  $\text{População de 0-4 Anos} / \text{Mulheres de 15-49 Anos}$

Fonte: Divisão de População das Nações Unidas, Revisão de 2019.

### 6.3 OS DETERMINANTES DA ESTRUTURA ETÁRIA

A pergunta sobre os fatores demográficos aos quais se deve o envelhecimento da população é menos fácil de responder do que parece à primeira vista. Em alguma medida a resposta depende do critério usado. A maioria dos estudiosos do tema usa o critério prospectivo que parte da idade mediana ou algum outro critério de envelhecimento em dois anos distintos  $t_1$  e  $t_2$  e compara a mudança

efetiva do indicador com a mudança contrafactual que teria sido observada caso a mortalidade (ou a fecundidade ou a migração) tivesse permanecido constante. A componente associada à maior diferença entre a mudança observada e a hipotética é considerada a determinante principal. Usando este critério, não há dúvida de que na grande maioria dos países é a fecundidade que pesa mais. No caso do Brasil, mais de 90% do aumento da idade mediana que ocorreu entre 1970 e 2015 foi consequência da queda da fecundidade, enquanto menos de 10% pode ser atribuído ao aumento da esperança de vida. Myrrha, Turra e Wajnman (2017), que fizeram uma decomposição mais detalhada para o período entre 1980 e 2050, concluíram que o declínio da fecundidade será responsável por 70% da variação na estrutura etária, enquanto os outros 30% serão atribuíveis igualmente à variação na mortalidade (10%), à inércia populacional (10%) e a interação entre esses fatores (10%).

Ainda há certa tendência, principalmente no setor da saúde, a confundir a estrutura da população com a estrutura da tábua de vida (que será discutida no Capítulo 9) e conseqüentemente a atribuir o envelhecimento da população a fatores como o aumento da esperança de vida ao nascer e a queda da mortalidade infantil. Conforme o critério adotado acima, esses fatores pesam muito pouco no envelhecimento. Como se mostrará na seção 22.5 do Capítulo 22, dependendo do perfil etário da queda da mortalidade, o resultado pode ser um ligeiro envelhecimento, mas o efeito também pode ser neutro ou até um *rejuvenescimento* da população.

Entretanto, alguns autores (Preston, Himes e Egger, 1989; Preston e Stokes, 2012) usam um critério diferente de avaliação que considera toda a história passada da mortalidade, fecundidade e migração das pessoas vivas em  $t_1$  e  $t_2$ . O raciocínio é bastante técnico e não será mostrado aqui. Basta dizer que os resultados deste tipo de análise tendem a dar mais peso à mortalidade do passado (antes de  $t_1$ ) do que fica evidente de uma análise prospectiva limitada ao intervalo entre  $t_1$  e  $t_2$ . No caso dos países mais desenvolvidos, Preston e Stokes consideram que o envelhecimento entre 2005 e 2010 foi devido predominantemente à queda acumulada da mortalidade. Para os países em desenvolvimento, os resultados sugerem contribuições mais ou menos iguais da mortalidade e da fecundidade.

## 6.4 A RELEVÂNCIA DA ESTRUTURA ETÁRIA PARA AS POLÍTICAS PÚBLICAS

Alguns dos índices mencionados nos parágrafos anteriores encontram-se ao centro de debates sobre questões de política. Em anos recentes houve um interesse crescente em fenômenos relacionados com a estrutura por idade e sexo, especialmente a Razão de Dependência, o Índice de Envelhecimento e as Razões de Juventude. O Índice de Envelhecimento tem uma relevância fundamental para as discussões sobre a Reforma da Previdência Social que foi discutida no Capítulo 3. A Razão de Dependência e a Razão de Suporte cumprem um papel fundamental na discussão do *bônus demográfico*, uma fase de maior potencial de crescimento econômico devido às tendências demográficas, que será discutido no Capítulo 14. Também existe uma controvérsia sobre a Razão de Sexos em alguns países que será explorada na seção 6.5.

A maioria dos países com Razões de Dependência elevadas na Tabela 6.1 ainda está na fase descendente da curva que geralmente é associada com o *bônus demográfico*. A principal exceção é Portugal, onde o valor mínimo da razão (47,7) foi atingido em 2000. Em Cuba, Trindade e Tobago e Macau, o mesmo aconteceu em 2010. O Brasil provavelmente atingirá o valor mínimo (43,3) em 2020. É de notar que o valor excepcionalmente baixo da razão em Macau se deve à forte imigração

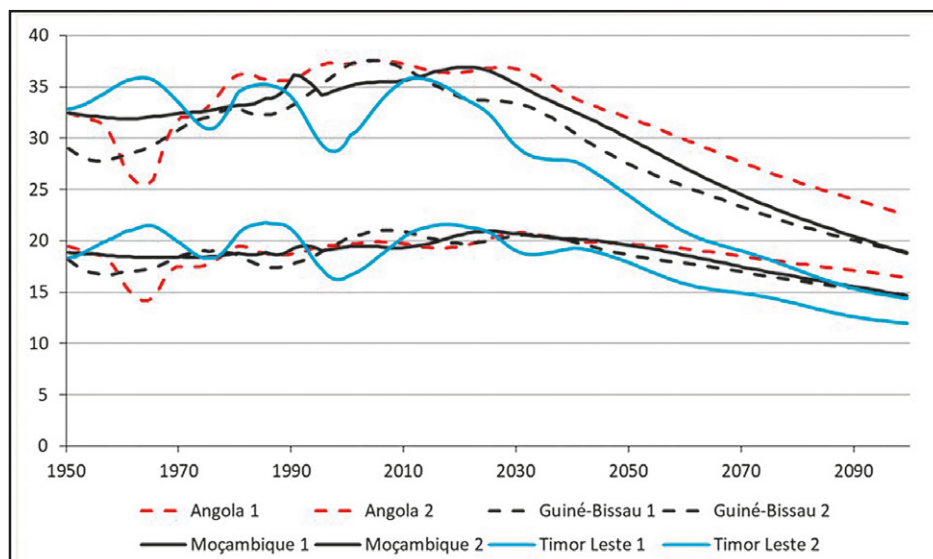


de pessoas nas idades economicamente ativas. No longo prazo, espera-se que a Razão de Dependência em Portugal e do Brasil voltará para valores próximos de 90, como os que caracterizam os países no início da transição (Angola, Moçambique, Timor-Leste) atualmente. Entretanto, muito dependerá da evolução futura fecundidade. A localização exata do período que corresponde à janela de oportunidade demográfica depende de como se define este período. No Brasil, por exemplo, a Razão de Dependência começou a diminuir a partir de 1965 (quando era 89,1) e, depois da queda inicial, continuará subindo até 2100 (quando se espera que chegue a 87,8).

Como já se mencionou acima, na medida em que diminui a fecundidade, as pirâmides etárias dos países se transformam e o maior grupo etário, que antes estava na base da pirâmide, começa a subir. A maior geração na história demográfica do Brasil foi a que nasceu entre 1980 e 1985; em Portugal foi a que nasceu entre 1960 e 1965. Na medida em que essas gerações ficam mais velhas, o modo da distribuição vai subindo, até que no final se chegue a uma situação de envelhecimento significativo em que o maior grupo etário pode ser o de 30-34 anos (como no Brasil em 2015) ou até o de 70-74 anos (como se projeta para Portugal em 2050). Essa situação não é necessariamente permanente e dependendo da evolução da natalidade futura, pode haver uma reversão parcial em que o modo da distribuição etária começará a baixar outra vez. Entretanto, dentro daquilo que se prevê para as próximas décadas, a tendência é que haverá certa redistribuição do peso relativo das pirâmides, de baixo para cima. Como se trata de um processo gradual, este afetará diferentes grupos etários sucessivamente. Primeiro ocorre certo deslocamento do peso relativo da infância para a juventude, depois para o grupo de 25-34 anos, depois para a média idade e assim por diante.

Como em muitos países de alta fecundidade onde esta começou a diminuir recentemente, o efeito mais imediato e mais palpável é de um relativo crescimento do grupo de jovens (15-24) comparado com o grupo de crianças (0-14), atualmente se fala muito da chamada “bolha de jovens” (“youth bulge”, em inglês) (Weeks e Fugata, 2012) que se reflete na porcentagem da população total que se encontra na faixa dos 15-24 anos.

Gráfico 6.3: As razões de juventude (15-24) relativas à população total (1) e à população com 15 anos ou mais (2) em Angola, Guiné-Bissau, Moçambique e Timor Leste 1950-2100



Fonte: Calculado com base nas estimativas e projeções da Divisão de População das Nações Unidas, Revisão de 2019.



A “bolha de jovens” é um fenômeno passageiro e de curta duração que na maior parte do mundo de língua portuguesa já ficou para trás. Em Portugal, a participação da juventude na população total diminuiu há muito tempo (Vieira, Ferreira e Rowland, 2015). No Brasil, o mesmo acontece desde a década de 80 quando a porcentagem de jovens atingiu os 20,1%, caindo para 16,4% em 2015. Na América Latina como um todo, a tendência é semelhante e a porcentagem tinha caído para 17,3% em 2015. No mesmo período, a parcela total de pessoas de 15-64 anos no Brasil, da qual os jovens são uma parte minoritária, cresceu de 57,9% para 69,1% da população total. Embora em Cabo Verde o processo seja mais errático, devido à migração internacional, a maior porcentagem de jovens aí se verificou por volta de 1985. Mesmo na Guiné-Bissau e São Tomé & Príncipe as porcentagens de jovens atualmente já diminuem, depois de atingirem o seu máximo na década passada e antepassada, respectivamente.

Os únicos países de língua portuguesa onde ainda se registra um aumento da proporção de jovens em anos recentes são Angola, Moçambique e, de forma bem mais errática, Timor Leste. O Gráfico 6.3 mostra a trajetória do tamanho da população de 15-24 anos nestes países e na Guiné-Bissau, como porcentagem da população total e como porcentagem da população com 15 anos ou mais. O primeiro indicador mostra pouca variação, pelo menos até 2050, o segundo um pouco mais, um aumento de até 27% entre 1955 e 2005, no caso da Guiné-Bissau. Entretanto, o Gráfico 6.3 também deixa claro que esta fase de lento crescimento das últimas décadas quase chegou ao seu fim e que nas próximas décadas a tendência será um rápido declínio da proporção de jovens, na medida em que o peso maior da distribuição etária passa a faixa dos 15-24 anos e sobe para idades maiores.

Alguns autores (Cincotta, Engelman e Anastasion, 2003; Urdal, 2006) associam uma porcentagem muito elevada de jovens de 15-24 ao conflito intergeracional e à maior probabilidade de instabilidade política. As dúvidas que podem ser levantadas são tanto metodológicas como substantivas. Do lado metodológico, é preciso reconhecer que, nesta fase da transição, a porcentagem de jovens de 15-24 na população total aumenta pelo menos em parte devido à redução da faixa de 0-14 anos. Como esta faixa não é relevante para a caracterização de um conflito geracional, o indicador deveria ser limitado à população maior de 15 anos, como é feito no Gráfico 6.3. Do lado substantivo, é preciso notar que – mesmo que haja uma correlação entre o tamanho relativo da população de 15-24 e a instabilidade política – isso não significa que haja uma relação causal. É perfeitamente possível que as mesmas transformações sociais subjacentes à queda da fecundidade e portanto à “bolha de jovens” também sejam responsáveis por um sentimento de anomia e insegurança. Particularmente um aumento acentuado dos níveis de educação que não é acompanhado por um aumento proporcional do emprego para uma nova geração mais capacitada pode gerar um clima de frustração e insatisfação.

## 6.5 A CARACTERIZAÇÃO DA DISTRIBUIÇÃO POR SEXO

Junto com a idade, o sexo é a outra característica básica que define uma população num momento específico no tempo. Uma questão conceitual que precisa ser mencionada aqui se refere à distinção entre as palavras “sexo” e “gênero”. Dado o crescimento do número de estudos sobre temas de gênero, nota-se certa tendência a substituir o termo “sexo” pela terminologia supostamente mais atualizada de “gênero”. Esta tendência é ainda mais forte na língua inglesa, já que a

palavra inglesa “gender” sempre teve um significado mais próximo a “sexo” do que os termos correspondentes em português. Entretanto, se trata de uma prática equivocada. Dentro da análise de gênero, as palavras *sexo* (“sex”) e *gênero* (“gender”) possuem significados claramente diferenciados. O sexo de uma pessoa é uma característica biológica com a qual nasce e que – salvo casos relativamente raros – não muda. O gênero é uma construção social que descreve como as categorias biológicas de masculino ou feminino são representadas como identidades sociais. O gênero pode determinar o que é esperado, permitido e valorizado numa mulher ou num homem num determinado contexto<sup>1</sup>. Por exemplo, os homens da cultura clássica grega biologicamente eram do sexo masculino tanto como os homens africanos de hoje. Mas a identidade de gênero, ou seja, o significado social de ser homem numa cultura ou na outra era bastante diferente. A análise de gênero se baseia em dados que tipicamente estão organizados por *sexo*, mas usa esta informação e a combina com outros elementos para desvendar o conceito de gênero dentro de um determinado contexto sociocultural (ver também Pinelli, 2004).

Como a distribuição por idades, a distribuição da população por sexo também pode ser caracterizada por meio de vários índices resumidos. Os mais conhecidos, ilustrados com os dados de Angola, são os seguintes, começando com a Razão de Sexos para o total da população:

$$\begin{aligned} \text{Razão de Sexos (antigamente também chamada Razão de Masculinidade)} &= RS_t = \\ &= 100 \cdot \text{Número de homens} / \text{Número de mulheres} = 12.499.041 / 13.289.983 = 94,05 \quad (6.10) \end{aligned}$$

Aqui a razão foi calculada para a população como um todo ( $RS_t$ ). Além de informar sobre o relativo equilíbrio populacional segundo sexo, este indicador tem, em princípio, pouco poder explicativo. Por ser uma variável de fácil obtenção, quando há indicações de informação deficiente, antes que erros na declaração, frequentemente se trata de cobertura deficiente, isto é de omissões em determinados grupos. Variações além do intervalo mencionado requerem explicações adicionais; a migração seletiva por sexo, uma excessiva mortalidade por causas maternas entre mulheres ou por causas violentas entre os homens, por exemplo, costumam afetar este indicador.

É de notar que a Razão de Sexos para o total da população ( $RS_t$ ) de Portugal (91,50) é mais baixa do que a de Angola. Em boa parte, a diferença se deve à estrutura etária de Portugal, com a sua preponderância de pessoas de mais de 50 anos. Como nestas idades há muito mais mulheres do que homens, a Razão de Sexos para o conjunto da população acaba sendo muito menor do que seria o caso se Portugal tivesse a estrutura etária de Angola.

A Tabela 6.2 traz alguns exemplos de  $RS_t$ , calculada sempre a partir de dados censitários, para diversos países que ilustram algumas particularidades, por exemplo aquelas associadas à migração internacional. Este é o caso da Argentina no início do século XX e Austrália do XXI. A Argentina com  $RS_t$  altas, inclusive atingiu uma  $RS$  de 115,5 em 1914. Com o fim desta política,  $RS_t$  declina constantemente e quase um século depois situa-se abaixo de 95. A Austrália segue um padrão semelhante, se diferenciando por ter, ainda em 2010, um afluxo migratório masculino de importância o que justificaria, em parte,  $RS_t$  maior que a Argentina nos tempos mais recentes. Já o caso de Cabo Verde nos anos 70, com uma  $RS_t$  de 107,7 (de iure), estaria relacionada com a entrada maciça de mão de obra angolana e nos anos mais recentes à diáspora

<sup>1</sup> Ver: UNWomen - <http://www.un.org/womenwatch/osagi/conceptsanddefinitions.htm>.

que tem caracterizado o país. Também há o caso de países com práticas de valoração do sexo masculino em detrimento das mulheres (ver abaixo). Seria o caso da China, onde estas práticas, impactariam mais a razão de sexos entre crianças e que explicaria os altos valores para a  $RS_t$ . Já para os anos mais recentes, com a guinada das políticas de população a  $RS_t$  tenderia igualmente a declinar até que, num futuro próximo, se aproxime de 95,0. Países como Portugal, França e Japão possuem razões relativamente baixas devido a sua estrutura etária envelhecida e o fato de que existem mais mulheres do que homens nas idades mais avançadas.

Tabela 6.2: Razão de Sexo total ( $RS_t$ ) segundo os censos, para países e períodos selecionados

País	1940-49	1950-59	1960-69	1970-79	1980-89	1990-99	2000-09	2010-17
Argentina	105,1			98,6	96,6	95,6	94,9	94,8
Austrália	100,5	102,4		100,0		98,5	97,5	97,8
China		107,6				106,0	105,2	104,4
França	92,6 *	92,2			95,4	94,8	93,8	
Japão		96,3	96,5	98,1	96,9	96,5	95,8	94,8
Portugal			91,8	90,2	93,0	93,1	93,4	91,5
Angola	90,3	96,3		108,9				98,3
Brasil	99,8	99,3		98,2	99,1	97,5	96,9	96,0
Cabo Verde		83,1	87,2	93,6**	84,8	89,7	93,8	98,0
Guiné-Bissau						93,9	94,2	
Macau							95,4	95,4
Moçambique	90,5	91,7		97,8	94,5	92,0	92,8	
São Tomé & Príncipe							98,4	98,9
Timor-Leste ***							102,4	101,7

\* Valor em 1936.

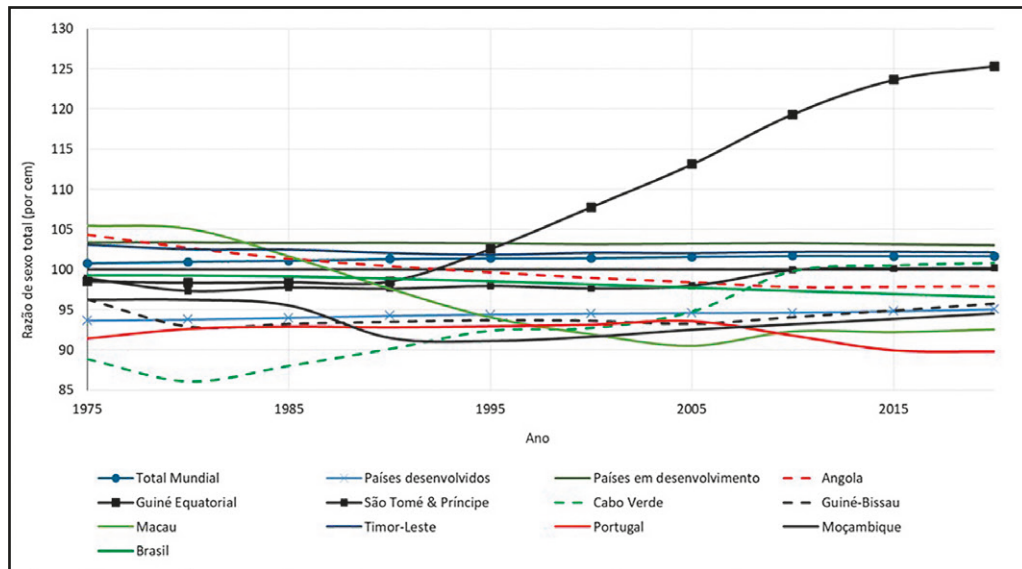
\*\* População *de facto*; segundo o critério *de iure* a RS era 107,7.

\*\*\* O censo da década de 2010 se refere a 2015.

Fontes: *Anuários Demográficos* das Nações Unidas e censos dos respectivos países.

O Gráfico 6.4 apresenta as  $RS_t$  estimada pela Divisão de População para os PALOP, total mundial e para os países mais desenvolvidos e em desenvolvimento (1975-2015). As médias replicam as tendências mencionadas anteriormente: nos países menos desenvolvidos a  $RS_t$  apresentou tendência à diminuição, embora menos acentuada para datas mais recentes. No caso das populações de países desenvolvidos, a  $RS_t$  está em aumento, com o que a média mundial se situa, para 2015, em torno de 102,0. No caso dos PALOP, com exceção do Brasil, e Angola, países de relativamente grande número de habitantes, as  $RS_t$ , no geral, situam-se fora do estreito intervalo antes mencionado, para alguns casos já foram citadas as prováveis causas, para outros, a explicação está ligada muito provavelmente aos movimentos migratórios e instabilidades sociais – o que no passado incluiu Portugal – e à menor qualidade do dado no que se refere à cobertura censitária. Entretanto, o maior desvio se verifica no caso da Guiné Equatorial que a partir dos anos 90 começou a ter uma  $RS_t$  muito elevada em função dos trabalhadores migrantes masculinos.

Gráfico 6.4: Razão de Sexo Total – Total da população mundial, países mais e menos desenvolvidos e PALOP 1975-2015



Fonte: Calculado com base na Divisão de População das Nações Unidas, Revisão de 2019.

### 6.5.1 Razão de Sexos ao Nascer (à Nascimento)

De todas as adaptações que podem ser feitas ao conceito de Razão de Sexos, a mais importante provavelmente é a Razão de Sexos ao Nascer ou à idade de 0 anos ( $RS_0$ ). Na composição por sexo ao nascimento da espécie humana, há uma ligeira tendência ao predomínio do sexo masculino. A razão de sexo *natural* ao nascer, utilizando o termo da OMS, é geralmente considerada como sendo cerca de 105<sup>2</sup>, ou seja, haveria, em média, ao nascer, em torno de 105 homens por cada 100 mulheres<sup>3</sup>. Valores fora do intervalo de 103-107 possuem qualidade deficiente ou são resultado, principalmente de intervenções de ordem:

- Cultural e seletivas por sexo, como a valoração de filhos do sexo masculino seja, provocando interrupção da gravidez, seja omitindo o registro de nascimentos de filhas.
- Sanitária: por corresponder ao sexo masculino maiores riscos de mortalidade, inclusive antes de nascer, avanços médicos que levam a gravidez de risco até um nascido vivo (nado vivo), produziram, proporcionalmente, mais nascimentos masculinos e  $RS_0$  maiores.

<sup>2</sup> [http://www.searo.who.int/entity/health\\_situation\\_trends/data/chi/sex-ratio/en/](http://www.searo.who.int/entity/health_situation_trends/data/chi/sex-ratio/en/).

<sup>3</sup> Antes de 1950, a mortalidade, principalmente nos países em desenvolvimento, situava-se em patamares muito altos, o que devia incluir, provavelmente a fase gestacional com a consequente perda proporcionalmente maior de fetos masculinos, o que levou Henry (1948) a afirmar que valores fora de 100-104 eram inesperados e precisariam ser explicados. É sabido que a proporção de fetos masculinos em idades gestacionais inferiores a 5 meses costuma ser bastante superior a 100% e talvez, próximos de 200% (Kellokumpu-Lehtinen e Pelliniemi, 1984; Byrne et al., 1987; Vatten e Skjaerven, 2004).

Há variações de origem biológica entre países e grupos sociais cuja origem não é bem entendida pela ciência genética. Catalano, Bruckner e Smith apresentam algumas evidências baseadas em séries históricas para os países escandinavos entre 1865 e 1914 que sugerem uma influência do clima: em anos mais frios,  $RS_0$  tende a ser mais baixa. Na maioria dos países da África Subsaariana os valores de  $RS_0$  também costumam ser baixos: 102 em Ruanda, Serra Leoa, Togo e Zimbábue, 103 em Angola, Cabo Verde, Guiné-Bissau, Guiné Equatorial, Moçambique, São Tomé & Príncipe e também na África do Sul, no Quênia, Uganda e na RDC, mas 105 em Burquina Faso, Gana e Mali e 106 na Nigéria. Garenne (2004) analisa a consistência destes achados em alguns países, mas não chega a dar uma explicação.

Por outro lado, Razões de Sexo dos nascimentos superiores a 107, como se encontram na China, Índia, Vietnam, Coreia, alguns dos países do Cáucaso e Albânia (Guilmoto, 2009; Bongaarts e Guilmoto, 2015) são consideradas uma aberração que indica uma manipulação proposital devido à forte preferência por filhos homem. Este problema vem se acentuando em anos recentes devido a dois fatores: 1. A maior facilidade de detecção precoce do sexo do feto por meio de ultrassonografia, o que facilita a decisão de abortar caso o feto não seja do sexo desejado; e 2. A diminuição do número médio de filhos por casal, o que reduz a margem de manobra para tentar outra vez a ter um filho do sexo desejado na próxima gravidez. Embora a Razão de Sexos no grupo etário de 0-4 anos (ou 0-6 anos, como se costuma calcular na Índia) seja um indicador imperfeito da Razão de Sexos ao nascer, ela é frequentemente usada para este propósito.

A Tabela 6.3 traz a  $RS_0$  para o período 2015-19, em países selecionados da Europa e América Latina, por um lado, carecendo estas áreas de comprovada preferência nacional pelo sexo da criança. Inclui também países de Ásia onde prevalece a preferência pelo sexo masculino da criança e os países da África Subsaariana, onde as razões de sexo ao nascer são atipicamente baixas. No caso dos países europeus, o maior valor corresponde à Espanha, com uma  $RS_0$  de 106,4. Na região latino-americana, chama a atenção Cuba pelo valor mais alto, de 1,057. Entretanto, o único país no continente com uma  $RS_0$  definitivamente anómala é Suriname ( $RS_0 = 108$ ). Existe uma minoria étnica indiana significativa (responsável por 30% dos nascimentos) naquele país, mas seria temerário afirmar, só com base nesta informação, que existe um problema de aborto seletivo por sexo no Suriname. Nos casos asiáticos os valores parecem atípicos; por um lado, na Coreia do Sul, e em menor medida, no Japão, parece haver uma tendência de aproximação do indicador à marca de 105, enquanto há um claro excesso de nascimentos masculinos na China e na Índia. O valor de Macau parece normal, mas é difícil de saber por certo, devido ao número reduzido de nascimentos.

Tabela 6.3: Razões de sexo (por cem) ao nascer para países selecionados, 2015-2019

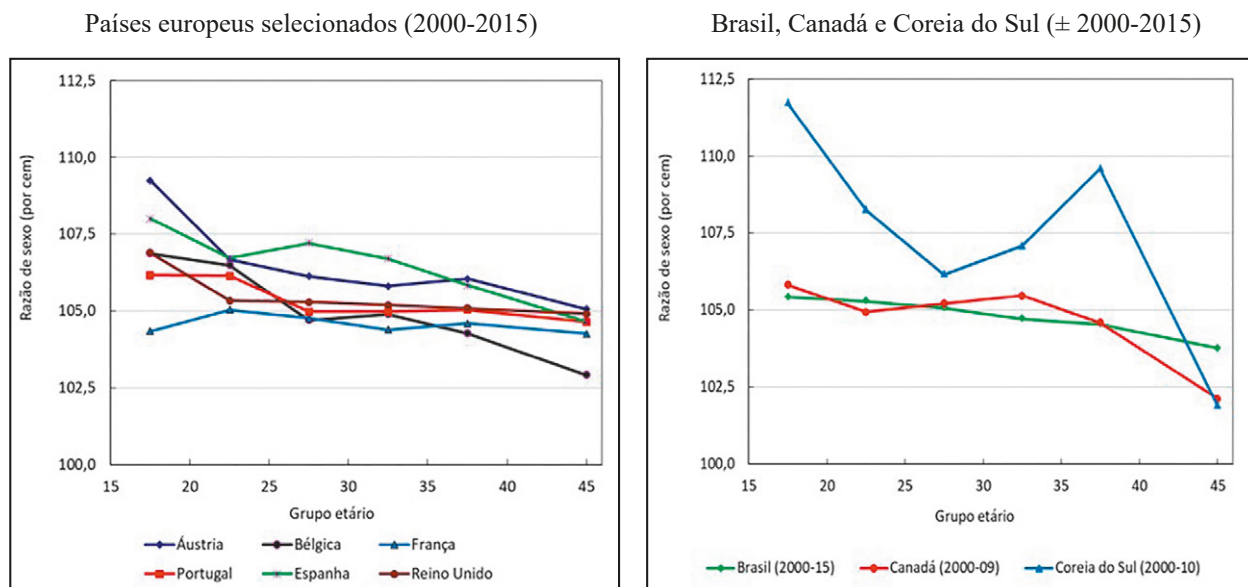
Países	$RS_0$
Alemanha	105,2
Angola	103
Argentina	104
Bolívia	105
Brasil	105
Cabo Verde	103
Chile	104
China	113
Coreia do Sul	105,5
Cuba	105,7
Espanha	106,4
França	105,2
Guiné-Bissau	103
Guiné Equatorial	103
Índia	109,9
Japão	105,6
Macau	105
México	105
Moçambique	101,9
Paraguai	105
Portugal	106
Reino Unido	105,1
São Tomé & Príncipe	103
Suriname	108
Timor-Leste	105
Uruguai	105
Venezuela	105

Fonte: Divisão de População das Nações Unidas, Revisão de 2019.

O comportamento da  $RS_0$  pode ser retratado também considerando sua tendência por idade da mãe ou ainda pela ordem do nascimento. Países sem notado intervencionismo no controle do comportamento reprodutivo têm mostrado uma tendência de leve, mas constante diminuição da razão de sexo ao nascer segundo a idade da mãe, tendendo a ser menor de 105 mas sempre acima de 100. No Gráfico 6.5 se apresentam alguns exemplos para os anos 2000 de uma série de países europeus, além do Canadá e Brasil. Também se inclui Coreia do Sul, pela sua particularidade de ser um país que até 2010 se caracterizava por uma forte seletividade dos nascimentos por sexo, mas onde esta situação se normalizou em anos mais recentes. Consequentemente, o país está adquirindo um padrão de  $RS_0$  caracterizado como dito, decrescente com a idade da mãe.

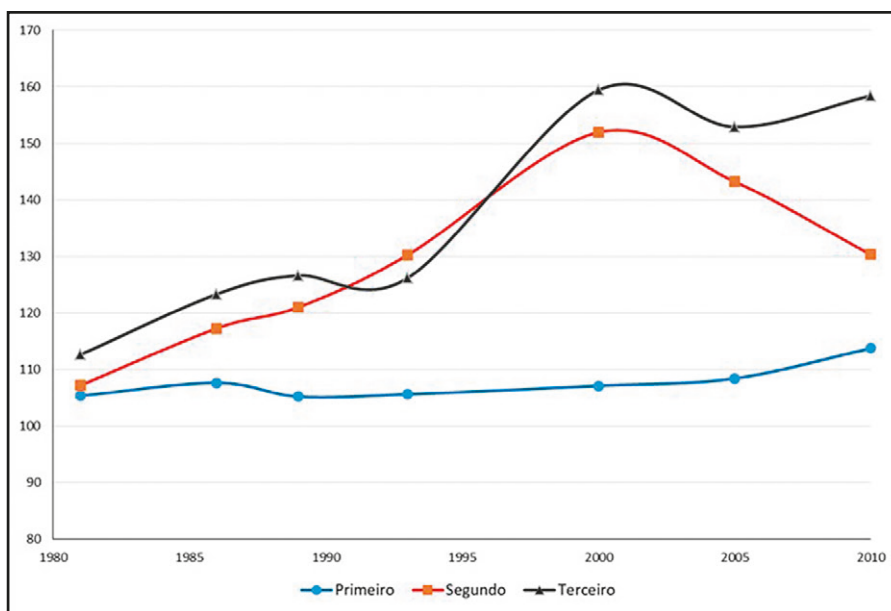


Gráfico 6.5: Média anual da Razão de Sexo ao nascer segundo a idade da mãe



Fonte: United Nations Statistics Division: Demographic Statistics Database.

Gráfico 6.6: Razões de Sexo ao Nascer por ordem de nascimento para China (1980-2010)



Fonte: Jiang et al. (2017): Table 2.

Ainda sobre a razão de sexo e características da mãe, ao poder associar a idade à parturição (o número de filhos que as mulheres têm), seria de esperar que nascimentos de ordem superior tenham uma  $RS_0$  igualmente tendendo a diminuir. Uma tendência oposta é indicativa de qualidade deficiente da informação ou da seletividade por sexo das gestações levadas a termo. Principalmente em países onde a fecundidade é baixa e a preferência por filhos homem é alta, os pais poderiam



não se importar muito com o sexo do primeiro filho, mas aqueles que optam por ter um segundo ou terceiro tenderiam a consistir desproporcionalmente de casais que procuram ter um filho homem. O fenômeno é ilustrado no Gráfico 6.6, com a  $RS_0$  da China, onde se pode perceber uma sequência razoavelmente próxima de 105 no caso de primeiros filhos, mas consideravelmente mais alta no caso de segundos e principalmente terceiros filhos.

### 6.5.2 Razão de Sexos por idade

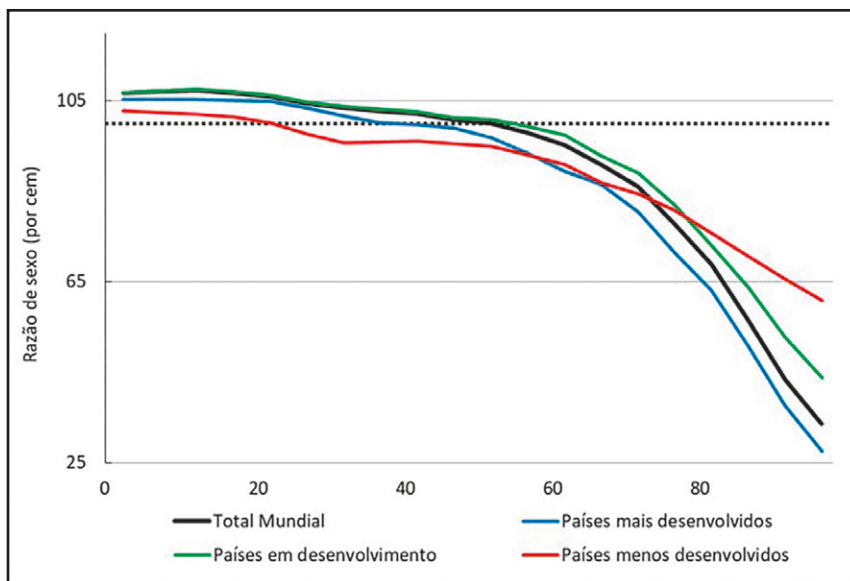
A Razão de Sexos também pode ser calculada por idade da própria pessoa o que define um índice notado como  $RS_x$ . Na Índia, por exemplo, se costuma calcular esta razão para crianças menores de 6 anos ( $RS_{0-5}$ ), para expressar o déficit além do esperado de meninas em comparação com os meninos, devido ao aborto seletivo e os maus tratos de meninas. Esta é uma alternativa para o uso de  $RS_0$ , já que este pode ser mais difícil de calcular, devido às deficiências do sistema de registro civil<sup>4</sup>. No caso do Censo de Angola de 2014, a Razão de Sexos para a faixa etária de 0-4 anos ( $RS_{0-4}$ ) foi  $2.484.583 / 2.513.566 = 98,85$ , um valor muito baixo considerando que normalmente nascem mais meninos do que meninas. No Censo de Portugal de 2011, o valor encontrado foi 104,29, bem mais próximo do esperado.

$RS_0$ , como foi visto, é afetada essencialmente por fatores de ordem biológica com pouca intervenção do ser humano, podendo ser modificada, principalmente, quando há preferência por um bebê de sexo masculino. Esse fator biológico –a mortalidade– tem um papel mais explícito nas idades seguintes, pois a mortalidade masculina geralmente é maior do que a feminina. Desta forma, à medida que se avança na idade e, em se tratando de uma população fechada às migrações (ou seja, sem imigrantes ou emigrantes), o número de homens, segundo a idade tenderá a ser proporcionalmente menor em relação às mulheres e a  $RS_x$  também. O padrão de  $RS_x$  assemelhar-se-ia, conseqüentemente, a uma curva convexa decrescente com o maior valor ao nascer, tal como o apresentado no Gráfico 6.7 para o total da população mundial.

Tem se constatado, ademais, uma associação com os níveis da mortalidade. Embora em condições sociais mais favoráveis às probabilidades de morte sejam relativamente menores, tanto para homens como para mulheres, estas últimas acabam se beneficiando proporcionalmente mais das melhoras que a diminuição da mortalidade traz. Isso explicaria em parte porque nos países desenvolvidos o desequilíbrio na  $RS_x$  é bem maior. Note-se que na média dos países socialmente mais vulneráveis (menos desenvolvidos) a  $RS_x$  diminui, efetivamente, com a idade, mas num ritmo mais lento: ao passo que nestas sociedades, ao fim da vida, haveria pouco mais de 75 homens por 100 mulheres, nos países desenvolvidos a razão equivalente é bem mais baixa que 30.

<sup>4</sup> Uma outra particularidade do uso deste indicador na Índia é que, contrariamente à convenção internacional, lá se costuma colocar o número de mulheres no numerador e o número de homens no denominador da Razão de Sexos.

Gráfico 6.7: Razões de sexo por idade em 2015: Total mundial e países por grau de desenvolvimento

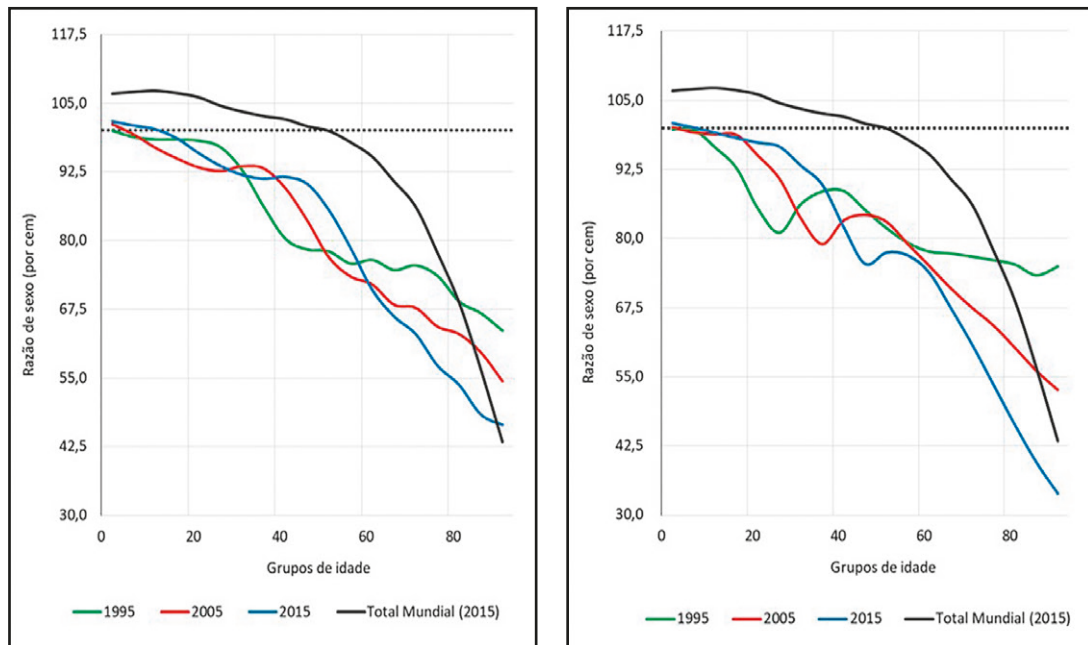


Fonte: Calculado com base na Divisão de População das Nações Unidas, Revisão de 2019.

Note-se, adicionalmente, que o equilíbrio deste índice é atingido mais precocemente na medida em que mais precárias são as condições socioeconômicas. Nos países menos desenvolvidos, o número de homens por 100 mulheres diminui desde o nascimento e alcança a marca de igualdade já por volta da idade 20; nas outras situações, o equilíbrio atinge-se depois dos 40 anos. Isto se deveria às perdas masculinas na população, essencialmente pela mortalidade por causas externas, mais acentuadas nos países menos desenvolvidos; já nas idades mais avançadas, o menor desequilíbrio entre os países menos desenvolvidos estaria associado a causas de morte *exógenas* que atingiriam a população com pouca distinção de sexo, como seria o caso das secas, tragédias socioambientais, deslocamento forçado etc. Estas considerações gerais servem, conseqüentemente, como guia para levantar hipóteses sobre a qualidade do dado por sexo. Os desvios dos valores da Razão de Sexo nas linhas acima demandam maiores análises e explicações; de outra forma, podem ser considerados erros de declaração.

Perfis diferentes ao apresentado no gráfico anterior costumam ocorrer. O Gráfico 6.8 descreve duas situações que ilustram bem situações atípicas dentro dos países PALOP na virada do século XXI: Guiné-Bissau e Moçambique, países que, além de sabidas dificuldades com a qualidade dos dados, experimentaram no passado recentes turbulências sociais de variada intensidade; no gráfico inclui-se, também, a curva que representa o total mundial em 2015. Em ambos países nota-se uma  $RS_0$  próxima de 100, o que indicaria níveis de mortalidade infantil bastante altos e coerentes com o que se sabe deste indicador. A hipótese de subdeclaração de crianças do sexo masculino seria pouco plausível, uma vez que no geral da população africana, prevalece a preferência de por crianças deste sexo.

Gráfico 6.8: Guiné-Bissau, Moçambique e Total mundial (1990-2015) – Razões de Sexo por idade



Fonte: Calculado com base na Divisão de População das Nações Unidas, Revisão de 2019.

O equilíbrio populacional por sexo é atingido a idades relativamente precoces, sendo esta característica, mais acentuada, ainda, em Moçambique. Todavia, neste último, observa-se uma anormal diminuição da  $RS$  já antes da idade 15, cuja causa está relacionada com os conflitos sociais dos anos 90 que ocasionou morte e deslocamento, para fora do país, da população jovem predominantemente masculina. Nas idades seguintes  $RS_x$  tende a aumentar com a idade replicando o perfil esperado de diminuição monótona com a idade. Nos dois períodos seguintes regista-se um padrão similar, diferindo no deslocamento das curvas no eixo das idades e na mais acentuada diminuição da  $RS_x$ .

Em síntese, na ausência de movimentos migratórios ou efeitos perturbadores na sociedade, uma  $RS$  por idade livre de erros ou omissões, deve apresentar, aproximadamente os seguintes rasgos:

- Valores acima de 100 e próximos de 105 ao nascer;
- Tendência monotonamente declinante com a idade até atingir o equilíbrio entre sexos, o que deve ocorrer depois da idade 25-30, dados os níveis de mortalidade do século XXI;
- Nas idades avançadas, 60 e mais, por exemplo, espera-se que a  $RS$  seja definitivamente inferior a 90; e
- A tendência temporal esperada é, por um lado, o adiamento da idade do equilíbrio da  $RS$  e diminuição cada vez mais acentuada nas idades extremas, por exemplo 70 anos e mais.

Uma última consideração a ser feita sobre a informação desagregada por sexo relaciona-se com as possibilidades de ajuste da distribuição da população por sexo. Como dito no início deste capítulo, as respostas sobre esta variável dificilmente estão erradas, sendo que os desvios dos padrões esperados se devem, principalmente a omissões ou a intervenções exógenas à *natural* composição por sexo da população. Mortalidade por causas externas extremamente diferencial por sexo, controle por sexo do produto da concepção, discriminação de gênero, deslocamento populacional seletivo por sexo e similares. Daí que não seja prática corrigir os dados como poderia ser feito com outras variáveis, sendo possível, apenas, emitir um diagnóstico sobre os graus de omissão por sexo.

## 6.6 PADRONIZAÇÃO/ESTANDARIZAÇÃO DIRETA E INDIRETA

Na análise de população encontra-se frequentemente situações em que um determinado indicador é influenciado por outros fatores, além daquilo que se pretende medir. Por exemplo, o número de migrantes de região A para região B obviamente depende do tamanho da região A, do tamanho da região B (ver fórmula 11.13 do Capítulo 11) e do tempo durante o qual a migração foi observada. Sem conhecer estes dados adicionais, o mero número de migrantes é difícil de interpretar. Dada a importância da idade das pessoas em qualquer análise de dados populacionais, em muitos casos a variável interveniente que dificulta a interpretação do resultado é a estrutura por idades.

Por exemplo, no Censo de Portugal de 2011 encontrou-se que 13,22% das mulheres de 15 anos ou mais eram viúvas. No Censo brasileiro de 2010, esta proporção era apenas 8,72%. A diferença poderia explicar-se de diferentes maneiras:

1. Talvez a mortalidade dos homens no Brasil seja mais baixa do que em Portugal;
2. Talvez as mulheres no Brasil tendam a casar-se com homens mais jovens do que em Portugal;
3. Talvez muitas mulheres brasileiras que vivem em uniões consensuais se considerem solteiras e não viúvas quando morre o companheiro;
4. Talvez as mulheres brasileiras tenham maior tendência a casar de novo depois que morre o companheiro; ou
5. Talvez a idade média das mulheres no Censo de Portugal seja mais alta.

A *padronização* (também chamada *estandarização*) oferece um instrumento para avaliar o impacto da quinta alternativa. Ela pode ser aplicada tanto a variáveis de estoque como de fluxo, mas a variável interveniente (geralmente idade, como neste exemplo) deve ser uma variável de estoque. O procedimento tem duas variantes: a direta e a indireta. A padronização direta pode ser usada quando se conhece tanto a distribuição da variável interveniente (neste caso, a idade) como a variação da quantidade analisada (viuvez) por categoria da variável interveniente, ou seja, por idade em ambas as unidades comparadas (aqui Brasil e Portugal). A Tabela 6.4 mostra todos os dados relevantes.

Tabela 6.4: Distribuição da população feminina e percentagens de viúvas por grupo etário em Portugal (2011) e no Brasil (2010), segundo os respectivos censos

Idade	Brasil 2010		Portugal 2011			
	População	% Viúvas	População	% Viúvas	Viúvas Hipotéticas	Padronizado
15-19	8.429.180	0,1	276.725	0,0	2.193	
20-24	8.613.199	0,2	289.042	0,1	7.241	
25-29	8.644.127	0,4	331.228	0,2	16.024	
30-34	8.026.535	0,9	394.833	0,4	33.502	
35-39	7.121.014	1,8	422.376	0,9	64.015	
40-44	6.688.525	3,4	398.136	1,8	122.671	
45-49	6.141.925	5,7	399.305	3,4	211.788	
50-54	5.308.482	9,7	376.112	5,9	312.698	
55-59	4.371.889	14,7	355.556	9,2	401.536	
60-64	3.470.156	22,1	336.195	13,9	482.557	
65-69	2.627.927	31,3	298.697	21,7	569.185	
70-74	2.069.185	42,7	275.977	32,6	673.703	
75-79	1.481.662	54,4	249.575	45,5	674.491	
80+	1.789.772	70,0	343.491	67,1	1.200.492	
Total	74.783.578	8,72	4.747.248	13,22	4.772.095	6,38

As primeiras quatro colunas da Tabela 6.4 (não contando a coluna de idades) simplesmente mostram os dados de partida, ou seja, as populações femininas e as percentagens de viúvas por faixa etária nos dois censos. A coluna crucial é a marcada como “Viúvas hipotéticas”. Aqui se pergunta: “Qual seria o número de viúvas de uma população com a estrutura etária do Brasil, mas a percentagem de viúvas por idade de Portugal?”. Portanto, é o produto da primeira coluna com a quarta (dividido por 100). Neste cenário hipotético haveria 4.772.095 viúvas no Brasil. Dividindo isso pela população brasileira de 74.783.578 obtém-se uma proporção padronizada de 6,38%. O que isso comprova é que a estrutura etária de Portugal efetivamente tem um peso grande na percentagem de viúvas no país. Se Portugal tivesse a estrutura etária feminina do Brasil, que é mais jovem, a sua percentagem de viúvas seria mais baixa do que no Brasil atualmente. Portanto, a ideia de que os homens brasileiros morrem menos do que os portugueses (alternativa 1) provavelmente não é correta.

Também seria possível perguntar: “Qual seria o número de viúvas de uma população com a estrutura etária de Portugal, mas a percentagem de viúvas por idade do Brasil?”. Isso implicaria multiplicar a segunda coluna pela terceira. O resultado seria 16,85%, valor que deve ser comparado com a proporção não corrigida de 8,72%. Portanto, a padronização pela estrutura etária do Brasil reduz a percentagem de Portugal de 13,22% para 6,38% (um fator de 2,07), enquanto a padronização pela estrutura etária de Portugal eleva a percentagem do Brasil de 8,72% para 16,85% (um fator de 1,94). Os dois fatores são próximos, mas não iguais porque o padrão de viuvez por faixa etária não é exatamente o mesmo no Brasil e em Portugal.

Não é preciso sempre usar um dos dois países como padrão. Em vez de usar o Brasil ou Portugal, a padronização direta também poderia ser feita usando como referência a estrutura etária de um terceiro país e calculando qual seria a proporção de viúvas naquele país caso as proporções por grupo etário fossem os do Brasil ou então os de Portugal.

Agora, suponha que por algum motivo as proporções de viúvas por faixa etária em Portugal não são conhecidas, só a percentagem geral de 13,22%. Nestas condições não é mais possível padronizar pela estrutura etária do Brasil para ajustar os 13,22% de Portugal, mas sim ainda é possível padronizar pela estrutura etária de Portugal, para ajustar os 8,72% do Brasil para 16,85%. Mas, que fazer se a padronização que se busca é aquela pela estrutura etária do Brasil? Neste caso, se usa o fator de ajuste da padronização oposta (1,94) para ajustar a proporção de 13,22% de Portugal. O resultado é uma proporção ajustada de 6,84%, próxima, mas não exatamente igual ao valor diretamente padronizado de 6,38%. Este procedimento se chama *padronização indireta*. Ele é um pouco mais aproximado do que a padronização direta, mas possibilita a obtenção de uma resposta mesmo quando faltam os dados completos necessários para uma padronização direta.

Uma técnica parecida com as descritas acima consiste em *decompor* a diferença entre duas taxas nas suas componentes. Essa técnica foi descrita originalmente por Kitagawa (1955) e estendida posteriormente por Das Gupta (1978, 1993). No Brasil, ela foi aplicada, entre outros, por Motta, Fígoli e Wong (2008), para decompor o efeito da estrutura etária nas diferenças de cobertura da Previdência entre as UFs. Dependendo do número de fatores a serem considerados (idade, sexo, região, atividade econômica etc.) e do número de unidades a serem comparadas, a complexidade das fórmulas pode variar, mas no caso mais simples, da comparação de duas unidades *A* e *B* com um fator interveniente (por exemplo, a idade *x*, dividida em *n* grupos), é fácil verificar que

$$\sum_{x=1}^n (m_x^B \frac{P_x^B}{P^B} - m_x^A \frac{P_x^A}{P^A}) = \sum_{x=1}^n (m_x^B - m_x^A) (\frac{P_x^B}{P^B} + \frac{P_x^A}{P^A})/2 + \sum_{x=1}^n (\frac{P_x^B}{P^B} - \frac{P_x^A}{P^A})(m_x^B + m_x^A)/2 \quad (6.11)$$

O primeiro termo quantifica o efeito da diferença entre as taxas ou proporções e o segundo termo o efeito devido à diferença entre as estruturas etárias. Aplicando a fórmula aos dados da Tabela 6.4, o resultado obtido é uma diferença de 4,50 (13,22 – 8,72), que pode ser descomposta em -2,99 (efeito das proporções) e 7,49 (efeito da estrutura etária).

## 6.7 DISTRIBUIÇÃO ESPACIAL

Até agora este capítulo tem focado a importância central da variável “idade” em qualquer análise demográfica. A segunda variável estática mais importante é a localização geográfica das pessoas, ou seja, a distribuição espacial da população. Mais especificamente, os aspectos mais frequentemente considerados desta questão são a densidade demográfica, o grau de urbanização, a distribuição por tamanho de cidades e o grau de concentração da população.

A densidade demográfica talvez seja o aspecto mais conhecido da caracterização da distribuição espacial de uma população. Inclusive, existe certa tendência a exagerar a importância desta característica, principalmente em conexão com a noção de “superpopulação” ou “capacidade de carga”. Cohen (1996), por exemplo, discute um grande número de estimativas históricas do número máximo de pessoas que a Terra poderia sustentar. Muitas destas estimativas se baseiam na



generalização das densidades demográficas observadas em algumas áreas densamente povoadas (começando com os 120 habitantes por km<sup>2</sup> dos Países Baixos, usados por Van Leeuwenhoek em 1679) para o total da superfície terrestre. Na verdade, é bastante difícil definir critérios objetivos e uniformes para associar determinados níveis de densidade demográfica a situações que possam ser qualificadas como “superpopulação”, mesmo em países onde as dimensões reduzidas do território poderiam sugerir uma maior relevância do conceito.

Em 2015, Macau tinha 19.652 habitantes por km<sup>2</sup>, a segunda maior densidade do mundo, superada apenas por Mônaco (25.322,8 por km<sup>2</sup>), enquanto Guiné Equatorial tinha 30 e o mundo como um todo 56,5. Isso significa que Macau era superpovoado e Guiné Equatorial estava numa situação demográfica mais ou menos confortável? Em alguma medida a resposta depende da interpretação exata que se faz do termo “superpopulação” (que não é fácil de definir), mas se for pelo nível de vida ou pela segurança alimentar da população, claramente Guiné Equatorial estava numa situação muito mais precária e vulnerável do que Macau. Para argumentar que Macau esteja efetivamente superpovoado, seria necessário lançar mão de outros argumentos como a constatação de que Macau, para garantir uma boa qualidade de vida para uma população tão concentrada, precisa mobilizar, através do comércio internacional, recursos de um território muito mais vasto do que os seus próprios 115 km<sup>2</sup>. O Relatório do Estudo sobre a Política Demográfica de Macau (Macau, 2015) ilustra esta fluidez conceitual:

De acordo com os dados da consulta pública sobre o “Enquadramento da Política Demográfica da RAEM”, a maioria das opiniões é que a estrutura demográfica futura depende da capacidade de carga demográfica. Simultaneamente, em relação à questão da “capacidade de carga demográfica”, algumas opiniões consideram que os terrenos em Macau são limitados e esperam que a Administração encontre o equilíbrio entre o crescimento demográfico e a falta de recursos terrestres; outras opiniões mencionaram que caso a população cresça, surgirão vários problemas sociais graves como transporte, habitação, saúde, proteção ambiental e outros. (...) Pode-se concluir que, apesar da escassez de recursos, a maior pressão consiste na capacidade de absorção de terrenos, habitação e estradas, pois Macau conseguiu sempre compensar a falta de recursos, através da abertura da economia e cooperação regional, bem como pela utilização conscienciosa dos recursos. Com a exploração da nova zona das Ilhas de Hengqin, a ligação terrestre da Ponte de Hong Kong-Zhuhai-Macau e o funcionamento do posto fronteiriço terrestre durante 24 horas, a ligação com os arredores será mais próxima e prática. Isto aliviará a pressão do crescimento demográfico de Macau, garantindo o conforto dos residentes (Macau, 2015: 76 e 77).

O conceito de densidade demográfica pode ser refinado até certo ponto pela limitação do denominador. Por exemplo, Guiné Equatorial possui uma área terrestre de 28.500 km<sup>2</sup>, mas segundo o Departamento de Estatística da FAO, em 2016 só 4,3% desta área consistia de terras cultiváveis<sup>5</sup>. Calculando a densidade demográfica com relação à área cultivável e não à área terrestre total, a densidade demográfica sobe de 42 para 974 habitantes por km<sup>2</sup>. Ainda assim, a densidade continua muito inferior àquela de Macau que teoricamente seria infinita, pois Macau não possui terras cultiváveis, ou de Singapura, onde a área cultivável é apenas 0,8% do total. Mesmo que não fosse assim, é preciso considerar que relacionar a população com a área de terras cultiváveis no país faz muito mais sentido num país predominantemente agrário como Guiné Equatorial do que num país altamente urbanizado como Macau onde a grande maioria da população vive de atividades terciárias.

<sup>5</sup> O termo “terra cultivável” é uma tradução do inglês “arable land”. A FAO alerta para o fato de que não se trata de todas as terras que potencialmente poderiam ser usadas para a agricultura. Mais bem se trata de terras que a curto prazo estão disponíveis para o cultivo de alimentos, excluindo pomares, prados permanentes, plantações de café ou borracha ou terras em descanso.



rias e onde os alimentos são quase todos importados. Ainda com esta ressalva, é útil comparar os dados na Tabela 6.5, que mostra as densidades demográficas de países selecionados conforme os dois critérios. Apesar das limitações do conceito, o quadro permite visualizar a dificuldade que alguns países (Bangladesh, Egito, Japão, Singapura, São Tomé & Príncipe) enfrentam para ser autossuficientes na produção de alimentos, quando comparados com a situação privilegiada de outros (Federação Russa, Brasil, EUA). A densidade de Bangladesh é menor do que a de alguns outros países pequenos, mas é a densidade maior entre países com uma extensão territorial de pelo menos 1.000 km<sup>2</sup>.

Tabela 6.5: Densidade demográfica (habitantes / km<sup>2</sup>) em 2015 de países selecionados conforme o critério de área terrestre e área cultivável

	Por Área Terrestre	Por Área Cultivável
Mundo	56,7	512,9
Angola	22,4	569,1
Argentina	15,7	109,9
Bangladesh	1.200,4	2.013
Brasil	24,5	252,5
Cabo Verde	130,2	1.050
China	49,9	1.183
Egito	92,9	3.317
EUA	35,1	210,7
Federação Russa	8,9	117,8
Guiné-Bissau	61,8	579,1
Guiné Equatorial	41,7	973,8
Índia	440,7	837,4
Japão	351,1	3.059
Macau	20.136,6	-- *)
Moçambique	34,4	478,6
Mongólia	1,9	528,6
Portugal	113,2	1.060
São Tomé & Príncipe	207,7	2.292
Singapura	7.988,8	1.011.438
Timor-Leste	80,5	771,8

\*) Macau não possui terras cultiváveis.

Fontes: Divisão de População das Nações Unidas (Revisão de 2019) para populações em 2015; FAO para áreas cultiváveis em 2016.

Em realidade, a maior utilidade do conceito de densidade demográfica não aparece ao nível de países mas na análise da dinâmica demográfica local (ver, por exemplo, Acioly e Davidson, 1999). Qualquer projeção demográfica municipal ou submunicipal precisa considerar quais são os

padrões de ocupação do espaço, qual é a densidade de ocupação, que áreas estão ficando saturadas e quais são os prováveis vetores de expansão. Mesmo assim, o conceito deve ser manejado com cuidado. Costuma-se afirmar que uma densidade demográfica alta ao nível de bairros ou setores censitários sugere uma precariedade no padrão de assentamento e possíveis riscos para a salubridade. Entretanto, sem informações adicionais não dá para generalizar a respeito. Por exemplo, a favela da Rocinha, no Rio de Janeiro, tem uma das densidades demográficas mais altas do Brasil, com 483 habitantes por hectare em 2010, o que efetivamente sugere precariedade do padrão de assentamento. Mas os bairros de Copacabana e Flamengo, de classe média alta, tinham densidades não muito diferentes: 357 e 304 habitantes por hectare, respectivamente. Por outro lado, o município de Belford Roxo na Baixada Fluminense, que apresenta condições sociais piores do que a Rocinha, tinha uma densidade de apenas 61 habitantes por hectare em 2010.

A associação da densidade demográfica com os riscos para a salubridade encontra alguma justificativa na relação que efetivamente existe entre a densidade demográfica e a rapidez da transmissão de doenças contagiosas. Entretanto, mesmo neste caso é preciso ter muita cautela. Por exemplo, durante a epidemia do coronavírus em 2020 o Presidente Bolsonaro argumentou que a propagação da doença no Brasil seria muito menor do que na Itália, devido à densidade demográfica do Brasil, que é oito vezes menor do que a italiana. Mas em realidade essa diferença nas densidades nacionais tem pouca relevância para os padrões de transmissão. Como se verá abaixo, a distribuição da população brasileira é extremamente desigual. Muito mais importante do que a média nacional é a densidade dos maiores centros de população como São Paulo, Rio de Janeiro e outras cidades grandes, onde a maioria da população mora (Rader et al., 2020). Muito depende também da densidade de ocupação das moradias (número de moradores por dormitório), da composição das famílias e da proximidade das interações sociais. Por exemplo, no caso do coronavírus um fator importante é a frequência de domicílios (agregados familiares) multigeracionais, onde pessoas mais velhas convivem com pessoas mais jovens, o que pode aumentar o risco de transmissão para os mais vulneráveis (ver seção 13.6 do Capítulo 13). Também é importante como as pessoas se deslocam (caminhando, transporte público ou meios individuais de transporte) e trabalham (ambientes individuais fechados ou ambientes coletivos).

Já foi mencionado o termo *grau de urbanização*, cuja definição evidentemente depende de como se entende o “urbano”. O *Dicionário Demográfico Multilíngue* (Macció, 1985: 53-54, tradução livre do espanhol para o português) se refere ao “urbano” nos seguintes termos: “Em demografia é dado um sentido particular às expressões área rural e área urbana pois a classificação numa ou outra categoria é feita de acordo a critérios que variam de um país a outro e que consideram aspectos como: população total, importância do aglomerado principal, proporção de população que depende da agricultura etc. Independentemente do critério usado para determinar a categoria rural ou urbana de uma unidade territorial, por convenção, a população que nela mora denomina-se *população rural* ou *população urbana*, segundo corresponda. Algumas definições de população rural e urbana diferenciam, ademais, outro grupo, a *população semiurbana*, constituída pelos habitantes de localidades cuja população é pouco numerosa e onde a atividade principal é de tipo não agrícola.”

Na prática, os critérios para diferenciar entre o rural e o urbano são de quatro tipos:

1. População: localidades urbanas são aquelas que reúnem certo número mínimo de habitantes e/ou certa densidade demográfica mínima;
2. Função político-administrativa: localidades urbanas são aquelas que são sedes das suas áreas administrativas ou que possuem certos órgãos de administração local;
3. Especialização econômica, ou seja, certa percentagem mínima da força de trabalho dedicada a atividades não agrícolas; e
4. Posse de equipamentos urbanos, tais como escolas, postos de saúde, eletricidade, ruas pavimentadas, água encanada etc.

O Quadro 6.2, baseado predominantemente no *Anuário Demográfico* das Nações Unidas de 2014, ilustra a variedade de critérios usados pelos países.

Quadro 6.2: Critérios oficiais para caracterizar uma área como urbana. Países selecionados, Circa 2014

País	Critério
África do Sul	Usa como critério o tipo de assentamento dominante e o uso da terra. Cidades, vilas, municípios, subúrbios etc., são assentamentos urbanos típicos. As áreas de enumeração que compreendem assentamentos informais, albergues, instituições, áreas industriais e de lazer e pequenas propriedades dentro ou adjacentes a qualquer assentamento urbano formal são classificadas como urbanas.
Argélia	Agrupamento de 100 ou mais construções, com uma distância menor de 200 metros entre elas.
Angola	Capital de província, sede de município e algumas vilas consideradas cidades e aglomerações com 2.000 ou mais habitantes com infraestrutura básica (escolas, estradas, posto médico etc.). Definição do Inquérito Integrado sobre o Bem-Estar da População (IBEP), 2008-2009.
Botsuana	Aglomerado de 5.000 ou mais habitantes onde pelo menos 75% da atividade econômica é não agrícola.
Cabo Verde	As 22 sedes de municípios, mais duas cidades que não são sedes municipais.
Guiné Equatorial	Centro distrital ou localidade com pelo menos 300 moradias ou 1.500 habitantes.
Guiné-Bissau	O censo não fornece um critério claro para separar o urbano do rural, afirmando apenas que “O meio urbano é constituído pelo conjunto de cidades e vilas segundo a divisa administrativa em vigor no País. Por sua vez, as cidades e vilas estão divididas em bairros. Entende-se por meio rural toda a parte do território situada fora do perímetro urbano, isto fora das cidades e das vilas. Do ponto de vista da divisão administrativa o meio rural estrutura-se em tabancas”.
Moçambique	Os municípios, segundo a classificação do Ministério de Administração do Estado (MAE) baseada em critérios políticos, econômicos, sociais e culturais, número e tipo de indústrias, desenvolvimento de atividades comerciais, educação e saneamento.
Quênia	Área com população de pelo menos 2.000 habitantes que possui sistema de transporte, áreas construídas, estruturas industriais/manufatureiras e outras estruturas desenvolvidas.
Cuba	Centro com funções político-administrativas ou mais de 2.000 habitantes e características urbanas.
Estados Unidos	Aglomerado com mais de 2.500 habitantes, geralmente com densidade de pelo menos 1.000 pessoas por milha quadrada.
Argentina	Centro povoado com pelo menos 2.000 habitantes.
Brasil	Áreas correspondentes às cidades (sedes municipais), às vilas (sedes distritais) ou às áreas urbanas isoladas. A situação rural inclui aglomerados rurais de extensão urbana, os povoados e os núcleos.
Chile	Aglomerado com mais de 2.000 habitantes, com 1.001-2.000 habitantes e mais da metade da força de trabalho em atividades secundárias ou terciárias, ou centro turístico com mais de 250 moradias.

País	Critério
Colômbia	Área com uma Prefeitura, definida por um perímetro urbano estabelecido por acordo municipal.
Costa Rica	Capital de província ou cantón.
Peru	Centro povoado com 100 ou mais domicílios (agregados familiares) contíguos.
Suriname	A capital, Paramaribo e o distrito de Wanica.
Venezuela	Centro com uma população de 2.500 ou mais.
Índia	Lugar povoado com corporação municipal, comitê de cidade ou outro órgão de administração local ou com mais de 5.000 habitantes, densidade superior a 400 por km <sup>2</sup> , características nitidamente urbanas e mais de 75% da população masculina adulta empregada fora da agricultura.
Japão	Cidade ( <i>shi</i> ) com mais de 50.000 habitantes com 60% ou mais dos domicílios (agregados familiares) localizados na área de maior concentração dos mesmos, com 60% ou mais da população (e seus dependentes) envolvidos em produção, comércio ou outro tipo de atividades urbanas.
Timor-Leste	Dili (capital) e outros pequenos assentamentos (sucos) designados como urbanos. Esta característica se define com base no número de habitantes em atividades não agrícolas, na existência de algumas comodidades / instalações e em alguns requisitos adicionais.
Espanha	Definição do Eurostat, ou seja, conjunto de população de pelo menos 10.000 pessoas.
França	Comuna onde a maior parte da população faz parte de um aglomerado de pelo menos 2.000 habitantes com menos de 200 metros de distância entre as moradias.
Países Baixos	Municipalidade com mais de 2.000 habitantes.
Portugal	Aglomerado de mais de 2.000 habitantes.
Reino Unido	Assentamento de mais de 10.000 habitantes (Inglaterra), 5.000 habitantes (Irlanda do Norte) ou 3.000 habitantes (Escócia).
Suécia	Comuna de 10.000 ou mais habitantes, incluindo subúrbios.

Fonte: United Nations, *Demographic Yearbook* 2014 e Divisão de População das Nações Unidas, Revisão de 2018.

Como se indicou acima, um dos critérios frequentemente adotados para caracterizar as áreas do país como urbanas ou rurais é o critério político/administrativo (país, capital, sede, província, distritos, condado etc.), de planejamento, fisiográfico, por tamanho de aglomerado etc. Estes critérios buscam agrupar a população sob o entendimento que a classificação usada identifica características comuns que a população, assim agrupada, teria. Isto possibilita, como sugere Welti (1997), definir unidades heterogêneas entre si, mas homogêneas dentro de si. Têm o óbvio objetivo de facilitar a administração de um país, daí que este tipo de classificação dependa de aspectos como necessidades da população, possibilidades de comunicação, condições climáticas ou geográficas etc.

As definições daquilo que oficialmente constitui o urbano e o rural não deixam de ser controversas em certos casos. Por exemplo, depois do Censo brasileiro de 2000, o resultado que mostrava um grau de urbanização de 82% foi contestado (Veiga, 2002), devido ao critério usado e as implicações do resultado para a importância da área rural. A criação de novos municípios, especialmente depois da nova Constituição de 1988, também aumentou o número de sedes de municípios e conseqüentemente o número de áreas urbanas, mas 84% das 5.507 sedes municipais enumeradas no Censo tinham menos de 20.000 habitantes e alguns tinham menos de 100 habitantes. Mais recentemente, o IBGE, com base nos resultados do Censo de 2010, publicou um documento chamado que propõe uma discussão sobre os critérios excessivamente inclusivos para a delimitação das áreas urbanas, de forma a aprimorar o Censo Demográfico de 2021, de

modo a “subsidiar a implementação de políticas públicas e o planejamento em geral no país”. Nessa nova tipologia, que está mais alinhada com as recomendações estatísticas internacionais, a população urbana cai nestes sete anos da data base dos dados utilizados dos 84,4% que vigorava na metodologia até então utilizada para 76%, concentrados em 26% dos municípios. A nova metodologia leva em conta a densidade demográfica, a localização em relação aos principais centros urbanos e o tamanho da população para classificar os municípios como urbanos, intermediários adjacentes, intermediários remotos, rurais adjacentes rurais remotos. Entretanto, a metodologia existente continuará sendo usada para fins legais. Outros países usam conceitos como “área rural urbanizada”, “área rural com estradas” ou “área rural sem estradas” para diferenciar entre diferentes tipos de ruralidade.

Em Portugal, apesar da definição oficial que considera aglomerados com mais de 2.000 habitantes como sendo urbanos, frequentemente se usam outros critérios, tais como os aglomerados com mais de 5.000, 10.000 ou 20.000 habitantes. Devido à falta de critérios uniformes, na prática a classificação rural-urbana é pouco usada para fins estatísticos e os dados costumam ser desagregados por região ou NUTS (Nomenclatura das Unidades Territoriais para Fins Estatísticos), um conceito do Eurostat que possui 3 níveis (3 NUTS do nível 1, 7 do nível 2 e 25 do nível 3).

O Quadro 6.3 reproduz os tipos de divisão político-administrativa para os países hispanos e de língua portuguesa. Ele inclui ademais, o tamanho relativo que possui o correspondente maior aglomerado urbano, que na maioria dos casos é a capital (e sede administrativa) do país com relação ao total da população urbana. Relativizando a proporção apresentada por Guiné-Bissau e São Tomé & Príncipe, cujas populações são bastante pequenas, nota-se em geral, a forte primazia que apenas uma cidade tem nesta seleção de países, o que denota o forte desequilíbrio da distribuição espacial da população. Entre aqueles de tamanho relativamente pequeno, sobressaem Angola, Guatemala, República Dominicana, Haiti e Uruguai, nos quais a cidade capital aglomera muito mais da metade de toda a população urbana. Note-se, ademais, que excetuando Uruguai, todos eles são países bastante pobres.

Quadro 6.3: Países selecionados de língua hispana e portuguesa, peso relativo da cidade capital (ou de maior tamanho) com relação à população urbana total e tipo de divisões político-administrativas

Região/País	Cidade Capital (ou de Maior Tamanho) (%)	Nome das Divisões em Ordem Decrescente de Importância		
Angola	39,7	Província	Município	Comuna
Argentina	37,0	Província/ Capital Federal	Departamento/ Partido	
Bolívia	23,8	Departamento	Província	Cantão
Brasil*	11,8	Estado Macrorregião	Município Mesorregião	Distrito Microrregião
Cabo Verde	40,0	Município/Concelho	Freguesia	
Colômbia	25,2	Departamento Intendência Comisaria Distrito Especial	Município	Corregimento
Costa Rica	35,1	Província	Cantão	Distrito
Cuba	24,1	Província	Região	Município
Chile	42,0	Região	Província	Comuna
Equador*	26,8	Província	Cantão	Paróquia
Guiné-Bissau	66,6	Região	Setor	
Guiné Equatorial		Região	Província	Distrito/Município
Haiti	43,4	Département	Arrondissement	Comuna
Macau	100,0	Freguesia		
México	21,4	Estado Distrito Federal	Município	Delegacia
Moçambique	11,4	Província	Distrito/ Município**	Posto Administrativo
Paraguai	73,6	Departamento	Distrito/ Município	Companhia/Povoado
Peru	40,4	Departamento	Província	Distrito
Portugal	43,6	Região/Distrito NUTS 1	Município (antes Concelho) NUTS 2	Freguesia NUTS 3
São Tomé & Príncipe		Distrito		
Timor-Leste		Município/Distrito	Posto Administrativo	Suco
Uruguai	52,4	Departamento	Seção Judicial	
Venezuela	10,6	Estado	Distrito	Município

(\*) Não correspondem às cidades capitais. Trata-se de São Paulo (Brasil) e Guayaquil (Equador).

(\*\*) No Censo de 2017 distinguiram-se 161 Distritos divididos em 446 Postos Administrativos. O município, que agrega alguns distritos, é uma divisão criada mais recentemente (1998) para descrever a parte urbana dos Distritos. Em 2015 existiam 53 municípios em Moçambique.

Fonte: Divisão de População das Nações Unidas. *World Urbanization Prospects* (2018).

As maiores áreas urbanas tendem a transcender os limites dos seus municípios e formar aglomerações urbanas que eventualmente podem ser reconhecidas administrativamente como Áreas ou Regiões Metropolitanas. O objetivo principal de tal reconhecimento administrativo é que desta forma se abre a possibilidade de ações de planejamento coordenadas e unificadas para a região como um todo, evitando a fragmentação das abordagens por diferentes autoridades



municipais ou até – no caso da Região Metropolitana da Cidade de México que pertence a dois Estados diferentes – por diferentes autoridades estaduais. Em Portugal existem duas Áreas Metropolitanas: a de Lisboa, com 18 municípios e 2,82 milhões de habitantes (2011), e a do Porto, com 17 municípios e 1,76 milhões de habitantes (2011). O Brasil atualmente possui 70 Regiões Metropolitanas, das quais 9 já existem há várias décadas (Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba, Porto Alegre). Junto com as Regiões Metropolitanas de Brasília, Goiânia e Manaus, estas compõem as Regiões Metropolitanas de primeiro nível. As demais são de criação mais recente e são consideradas de segundo nível. Do total de 70, 26 Regiões Metropolitanas tinham mais de 1 milhão de habitantes em 2016. No Brasil o termo “aglomeração urbana” também tem um status oficial e em 2015 existiam cinco áreas oficialmente reconhecidas como tal no país: as de Jundiaí, Piracicaba, Litoral Norte de Rio Grande do Sul, Caxias do Sul e Pelotas. Trata-se de conurbações mais incipientes que eventualmente poderão transformar-se em Regiões Metropolitanas no futuro. Nos demais países de língua portuguesa formalmente não existem Áreas Metropolitanas, embora Macau forme parte da Região Metropolitana do Delta do Rio das Pérolas da China que também engloba Hong Kong, Guangzhou e várias outras metrópoles. Luanda, em Angola, pelo seu tamanho, bem poderia ser considerada uma Área Metropolitana, consistindo dos municípios de Luanda propriamente (2,19 milhões de habitantes segundo o Censo de 2014), Cazenga (0,89 milhões), Cacuaco (1,07 milhões), Belas (1,08 milhões) e Viana (1,61 milhões). Em Moçambique, a capital nacional Maputo forma uma conurbanização com a cidade de Matola, a capital da Província de Maputo. A população conjunta das duas cidades já supera os 2,5 milhões de habitantes.

Existem alguns índices que quantificam as características espaciais da população. No Quadro 6.3 já foi introduzida a proporção da população urbana que mora na maior cidade do país, como uma das maneiras para expressar o grau de concentração da população. Outras medidas parecidas poderiam ser a proporção que mora numa das três ou das cinco maiores cidades ou que vive em cidades com mais de 1 milhão de habitantes. Todas estas medidas focalizam o extremo da distribuição populacional, ou seja, os lugares com as maiores concentrações demográficas. Também existem outras medidas, que tratam do assunto de uma forma mais global. Uma destas medidas<sup>6</sup> é o Índice de Gini, que não é um indicador propriamente demográfico, mas que surge mais frequentemente no contexto da medição da desigualdade de renda. Entretanto, ao substituir “renda” por “população” o mesmo pode ser usado de forma análoga para quantificar desequilíbrios na distribuição da população pelo espaço. Existem diferentes maneiras para calcular o Índice de Gini, mas o mais tradicional e o mais intuitivo envolve o uso do *diagrama de Lorentz*.

O Gráfico 6.9 mostra um exemplo deste tipo de diagrama, com dados dos Censos de população do Brasil (2010) e de Moçambique (2007). O eixo vertical do diagrama representa a quantidade sendo distribuída, neste caso a população. O eixo horizontal representa a variável à qual a quantidade distribuída pertence, neste caso o espaço, quantificado como a área geográfica acumulada dos municípios do Brasil ou dos distritos de Moçambique<sup>7</sup>. Estas unidades espaciais estão ordenadas

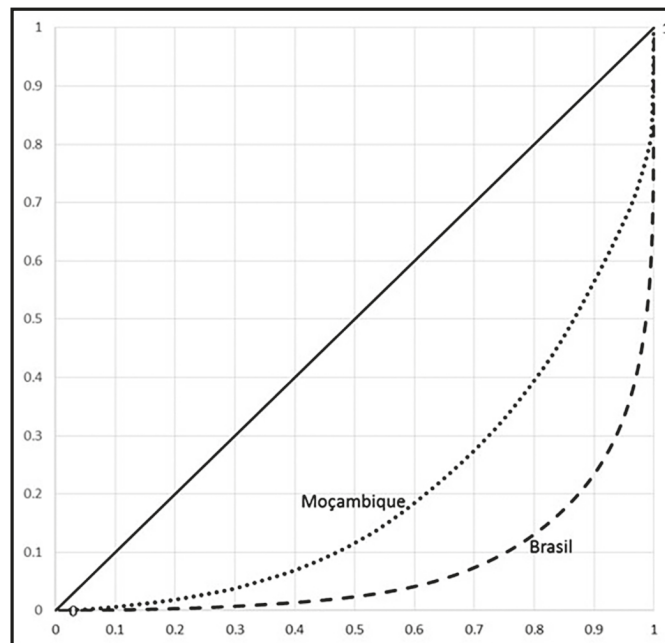
<sup>6</sup> Outra medida é o índice de Theil  $T = (\sum (x_i/\bar{x}) \ln(x_i/\bar{x}))/N$ , onde a soma é sobre  $N$  entidades (indivíduos, municípios), cada uma das quais possui uma quantidade  $x_i$  do bem por distribuir (renda, população) e  $\bar{x}$  é a quantidade média que cada entidade possui. Uma medida ainda mais simples é o índice de Herfindahl-Hirschman que é simplesmente  $H = \sum (x_i/N\bar{x})^2$ . Esses índices são raramente usados na demografia, mas Serrano et al. (2015), por exemplo, usam o índice  $H$  para quantificar o desequilíbrio da distribuição de pessoas altamente escolarizadas no Brasil.

<sup>7</sup> Algumas das unidades territoriais de Moçambique formalmente não são distritos, mas cidades.

da esquerda para direita em ordem de densidade demográfica crescente. Tanto a área das unidades espaciais como a população são divididas pelos seus totais respectivos (área total e população total) para que tanto a escala horizontal como a vertical tenham 1 como o seu valor máximo. O ponto  $(x_n, y_n)$  numa das curvas se refere à área conjunta dos  $n$  municípios ou distritos de menor densidade demográfica ( $x_n$ ) e à população conjunta destes municípios ou distritos ( $y_n$ ). Na parte esquerda do diagrama, onde estão as unidades de menor densidade demográfica, as curvas aumentam lentamente, mas na medida em que se avança para a direita se avança as unidades, são cada vez mais densamente povoadas e o ritmo de aumento cresce cada vez mais.

Se a população estivesse homogeneamente distribuída pelo espaço, a curva resultante seria a linha reta diagonal mostrada no gráfico. Quanto mais desigual for a distribuição, mais a curva tende a aproximar-se do canto direito inferior do diagrama. Como a distribuição da população do Brasil, com suas vastas áreas quase inteiramente despovoadas, é mais desigual do que a de Moçambique, a curva que caracteriza o Brasil se inclina mais para baixo e para a direita. O Índice de Gini agora se define como a superfície contida entre a curva e a linha diagonal como proporção do triângulo inferior direito. No caso do Brasil esta proporção é 0,846 e no caso de Moçambique 0,591. Um valor de 0 indicaria uma distribuição completamente homogênea e um valor de 1 representa a desigualdade total, em que toda a população está concentrada numa área de tamanho quase zero. Outros dados que podem ser derivados do gráfico são que, por exemplo, 90% da população brasileira mora em 25% do território (veja onde a linha horizontal do nível 0,1 corta da curva) e que a metade mora em apenas 1,4% do território (veja onde a linha horizontal do nível 0,5 corta a curva).

Gráfico 6.9: Diagrama de Lorentz da População Municipal do Brasil em 2010 e da População Distrital de Moçambique em 2007



Fontes: Censos Demográficos do Brasil (2010) e Moçambique (2007).

Para construir o diagrama de Lorentz e calcular o Índice de Gini em EXCEL, pode-se seguir os seguintes passos:

1. Organizar os dados em duas colunas: a primeira (A) contém as áreas das unidades territoriais e a segunda (B) as suas populações. Eventualmente se pode acrescentar uma terceira coluna com os nomes das unidades.
2. Calcular uma coluna com as densidades demográficas, dividindo coluna B por coluna A.
3. Ordenar a planilha em ordem crescente das densidades demográficas calculadas em 2.
4. Calcular a soma das áreas da coluna A e definir uma coluna D que divide cada célula da coluna A por esta soma.
5. Calcular a soma das populações da coluna B e definir uma coluna E que divide cada célula da coluna B por esta soma.
6. Calcular colunas F e G que acumulam as colunas D e E de 0 até 1.
7. Definir um gráfico de dispersão usando a coluna F como eixo horizontal e G como eixo vertical. O resultado é o diagrama de Lorentz. Eventualmente pode-se acrescentar uma linha reta diagonal.
8. Calcular uma coluna H com a seguinte fórmula:  $HI = (F2-F1) \cdot (G1+G2)/2$  e assim por diante, até a penúltima célula da coluna.
9. Somar a coluna H, multiplicar o resultado por 2 e subtrair de 1. O resultado é o Índice de Gini.

Como a quantidade de pontos usados para construir o diagrama acima foi muito grande, o processo não é mostrado em detalhe aqui. É preciso assinalar que o resultado da análise depende, até certo ponto, da unidade territorial escolhida. Os distritos de Moçambique em média são maiores do que os municípios brasileiros. Consequentemente, o menor Índice de Gini de Moçambique se deve em alguma medida (pequena) à maior granularidade da unidade de análise. Por outro lado, se a análise do Brasil fosse executada em termos dos distritos dos municípios, que são ainda menores do que os distritos moçambicanos, o Índice de Gini do Brasil seria mais alto ainda.

Existem outros índices, que quantificam o grau de aglomeração de certas características da população em determinados conjuntos de unidades espaciais. Neste caso o objetivo não é tanto quantificar se toda a população com determinadas características se concentra numas poucas unidades, mas se as unidades caracterizadas por determinadas características demográficas (por exemplo, uma alta proporção de idosos ou uma fecundidade elevada) tendem a ser geograficamente contíguas. Ou seja, as unidades espaciais com altos níveis da característica X também tendem a ter áreas vizinhas com altos níveis da característica X (autocorrelação espacial positiva) ou tendem, pelo contrário, a ter áreas vizinhas caracterizadas por uma baixa incidência de X (autocorrelação espacial negativa) ? Como no caso do índice de Gini, a resposta dependerá até certo ponto da definição das unidades espaciais. Se as unidades forem muito pequenas, a autocorrelação espacial

tenderá a ser maior. Por exemplo, setores censitários de alta densidade demográfica tendem a ser localizados no meio de outros setores censitários de alta densidade. Mas a mesma relação não se aplica necessariamente a municípios.

Um dos índices mais usados para quantificar a autocorrelação espacial é o índice de Moran, que pode variar de -1 a 1, da mesma forma como um índice de correlação comum. A análise de autocorrelação espacial não é usada muito frequentemente na demografia, mas tem aplicações importantes em áreas afins, como a epidemiologia, por exemplo para modelar a transmissão de doenças. Alguns exemplos do uso de análise de autocorrelação espacial são Santos et al. (2017), que analisaram os padrões espaciais do suicídio entre idosos, Borges et al. (2016), que analisaram os padrões espaciais da fecundidade adolescente no Brasil, e Gomes et al. (2016), que usaram o índice de Moran, entre outros métodos, para caracterizar os padrões espaciais da queda de fecundidade em Portugal. Os detalhes da construção do índice de Moran não serão discutidos aqui. Para mais informação sobre este tema, consulte-se o artigo original de Moran (1950) ou o artigo mais recente de Assunção e Reis (1999), que introduz algumas alternativas de modificação do índice original.

