March 2014

Methodology for Minnesota County Population Projections by Age and Gender, 2015-2045

by Megan E R Dayton

Minnesota State Demographic Center

This describes the methodology used to produce population projections for the 87 counties of Minnesota by age cohort and by gender for each five-year interval from 2015 to 2045. The method for projecting future population is outlined below. Each step is then discussed in more detail.

1. Project age cohorts and gender of 87 counties within Minnesota using the cohort-component method
    a. Project births and deaths as rates of natural increase/decrease
        i. Exceptions
    b. Project net-migration
        i. Exceptions
2. Control new county totals (sum of age/gender projections) to Minnesota state population projections[1]
3. Check final totals against preliminary regression analysis detailed in Appendix A

Step 1: Cohort-Component Method

The cohort-component method is an approach to demographic projections that accounts for age distribution in a given population. This method consists of segmenting the population into different subgroups (age and gender) that may be differentially exposed to the components of change – births and deaths – by separately calculating the changes over time in each group.[2] The method was applied to each of the 87 counties individually.

The first step in the cohort-component method is to establish a launch-year and calculate the number of persons who will survive to the end of the interval (e.g. 5-years). This is accomplished by applying age/sex survival rates to each age/sex group in the launch-year.[3] A base population of July 1, 2011 was used because 2011 is the most recent year for which corresponding age-specific, county-level vital statistics are available from the Minnesota Department of Health. In subsequent years after 2011, a new birth cohort was then added to form the population under five years of age by applying the projected age-specific fertility rates to the average female population aged 10-54 years and updating the new cohort of the effects of mortality.

[1] This dataset can be found here: http://www.demography.state.mn.us/resource.html?Id=33558
[2] Preston, S. H., P. Hueveline, and M. Guilliot. 2001. *Demography: Measuring and Modeling Population Processes.* Malden, MA: Blackwell Publishing.
[3] M.V George, S. Smith, D. Swanson, J. Tayman. 2004. *The Methods and Materials of Demography.* San Diego, CA: Elsevier Academic Press.

The next section details the methods used to estimate rates of fertility, mortality, and net-migration.

Step 1a: Project Births and Deaths as Natural Increase/Decrease

Births were projected using county-level fertility rates by single year of age among women of childbearing age. A five-year average of age-specific fertility rates was calculated using administrative birth data from years 2007 through 2011. The ratio of male to female births at the state-level was used to predict the sex ratio at birth for all counties.

Deaths were projected using survival rates by single year of age and gender. A three-year average of age-specific mortality rates was calculated using administrative death data from years 2009 through 2011. Survival rates are obtained from a county-specific life table. Regional averages of survival rates were calculated and applied to counties within each EDR.

Step 1ai: Exceptions

Administrative data for Nobles County in EDR 08 and Roseau County in EDR 01 created an unrealistic growth trajectory for a long-term projection. Because of this, both counties are assumed to converge with their respective regional fertility rate in 2020.

It can be difficult to obtain stable measures of fertility for small populations. Because of this, greater Minnesota age-specific fertility rates were averaged from 2009 through 2011 and used for counties with 16,000 or less total population in 2010.

Step 1b: Compute Net Migration

Historic age-specific net-migration rates were observed from 2000 to 2010 using the following formula:

$$P_{x+1} = P_x + (B - D) + M$$

where:

P_x is the county population of any given year
B is county births occurring during P_x

D is county deaths occurring during P_x
M is county net-migration during P_x

Solving for M, yields age-specific net-migration over the decade.

$$M = P_{x+1} - (P_x + (B - D))$$

The product was converted to an annual rate and held constant over the projection window.

Step 1bi: Exceptions

This process for Roseau County in EDR 01 created an unrealistic growth trajectory for a long-term projection. Because of this, age-specific migration rates for this county are increased until 2020 with a gradual return to observed intercensal rates for the duration of the projection horizon.

Step 2: Control Age/Gender County Totals to State of Minnesota Total

The final step involves rounding and making minor adjustments to the existing population in age groups so that all 87 counties sum to the state-level projections. Fertility rates, mortality rates, and net-migration rates at the county level, when applied to a total population, will yield an estimate that is not a whole number. This is the formula used to adjust to whole numbers:

$$_nP_x * (P_t / \Sigma_n P_x)$$

where for any given year:
$_nP_x$ is the population in a single sex/age group
P_t is the total state population
$\Sigma_n P_x$ is the sum of all single sex/age groups

Step 3: Final Totals Check

The final numbers from this process were checked against the numbers created from the process outlined in Appendix A of this document. Controls were not applied to the cohort-component projections.

POTENTIAL SOURCES OF ERROR IN PROJECTIONS

There are many potential sources of error in demographic projections. These include:

1. The estimates used for the base population may be inaccurate
    a. The further estimates are from the Census year, the less accurate estimates are likely to be
2. The national projections (that the Minnesota state projections were base on) could be in error
3. The Minnesota state projections (that this dataset is based on) could be in error
4. Fertility rates may change in a pattern other than that assumed by this model
5. Mortality rates may change in a pattern other than that assumed by this model
6. Future migration trends could shift depending upon economic trends, immigration law changes, climate change and many other factors

For more information regarding the methodology described here, please contact Megan Dayton at Megan.Dayton@state.mn.us or 651-201-2461.

Appendix A

A separate regression analysis was conducted to establish the direction of historic and current population trends for the Economic Development Regions and each of the 87 counties of Minnesota. This method was not used as a control in the final dataset. This alternative method for projecting future population is outlined below. Each step is then discussed in more detail.

1. Project total population of eleven Economic Development Regions (EDRs) within Minnesota using regression analysis
2. Control sum of EDRs to official Minnesota state population projections
3. Project total population of 87 counties within Minnesota using regression analysis
4. Control sum of the 87 counties to the EDR projections

Step 1: Project EDR Total Population

The first step was to project the total population of the eleven different Economic Development Regions (ERDs) using regression analysis. The analysis was conducted using the statistical program STATA 12. Four basic models were used:

1) Linear regression of EDR population on year
2) Linear regression of EDR population on year and year-squared
3) Linear regression of log on EDR population on year
4) Linear regression of log on EDR population on year and year-squared

Analysis of the historical trend, in conjunction with larger regional trends, served to guide the choice between different models for each of the EDRs. More recent population totals – since 2000 – are likely to be more indicative of future trends than are previous population totals – from 1960[4]. For this reason, EDR population totals were each 'weighted'[5] twice as heavily for years 2000 and 2010 relative to all previous data points. This weighting procedure was used consistently for each of the models described above, and for each of the historic time windows[6].

[4] For example, the 2020 population of a given EDR will be more affected by the 2010 population than by the 1960 population.
[5] 'Importance weights' in STATA were used to implement this procedure. Using a weight of '2' for each of the following years: 2000 and 2010, while using a '1' for the remaining years, is equivalent to including a second data point with the same value for those years that were weighted double.
[6] For any given county, the 2010 population totals were 'weighted' as being twice as 'important' as the 1980 population totals, regardless of the model selection or the beginning of the time series.

Step 2: Control EDR Projections to State Total

The second step involves rounding and making adjustments to the EDR total population in order to match the official projections for the State of Minnesota[7]. This is the formula used to make these adjustments:

$$ {}_{n}P_{x} * (P_{t} / \Sigma_{n}P_{x}) $$

where for any given year:

$_{n}P_{x}$ is the population in a single EDR
P_{t} is the total state population
$\Sigma_{n}P_{x}$ is the sum of all EDRs

Step 3: Project County Total Population

The next step is similar to the first step of projecting EDR total population. This step involves projecting the total population of the 87 counties of Minnesota using regression analysis. The analysis was conducted using the statistical program STATA. Four basic models were used:

1) Linear regression of EDR population on year
2) Linear regression of EDR population on year and year-squared
3) Linear regression of log on EDR population on year
4) Linear regression of log on EDR population on year and year-squared

Analysis of the historical trend, in conjunction with larger regional, served to guide the choice between different models for each of the counties. More recent population totals – since 2000 – are likely to be more indicative of future trends than are previous population totals – from 1960[8]. For this reason, county population totals were each 'weighted'[9] twice as heavily for years 2000 and 2010 relative to all previous data points. This weighting procedure was used consistently for each of the models described above, and for each of the historic time windows[10].

[7] This dataset can be found here: http://www.demography.state.mn.us/resource.html?Id=33558
[8] For example, the 2020 population of a given EDR will be more affected by the 2010 population than by the 1960 population.
[9] 'Importance weights' in STATA were used to implement this procedure. Using a weight of '2' for each of the following years: 2000 and 2010, while using a '1' for the remaining years, is equivalent to including a second data point with the same value for those years that were weighted double.
[10] For any given county, the 2010 population totals were 'weighted' as being twice as 'important' as the 1980 population totals, regardless of the model selection or the beginning of the time series.

Step 4: Control County Projections to EDRs

This step involves rounding and making adjustments to the county total population in order to match the EDR projections produced in **Step 1** and finalized in **Step 2**. This is the formula used to make these adjustments:

$$ {}_{n}P_{x} * (P_{t} / \Sigma_{n}P_{x}) $$

where for any given year:

- ${}_{n}P_{x}$ is the population in a single county
- P_{t} is the EDR total population
- $\Sigma_{n}P_{x}$ is the sum of all counties in respective EDR