TS PROJECT TEMPLATE ON BIRTH RATES

ARIMA models are often used to project population growth and decline. Candidates who work with population statistics have much data that can be used for the student project. They are familiar with these time series and they know the trends and changes that can make a good project.

Actuaries examine the components of population. The population in year T is

the population in year T-1 + births – deaths + immigration – emigration.

Even if births, deaths, immigration, and emigration change, the population time series may seem stable. You might model population as an AR(1) process with low stochasticity, which fits reasonably well. But 95% of the population is the same in successive years, so the time series for the total population doesn't say much.

The first differences are the combined effect of births – deaths + immigration – emigration. These four pieces (births, deaths, immigration, and emigration) can all be used for the time series student project. Each has its own process.

This project template outlines a student project on birth rates. The other three pieces are related to government policy (immigration), wars (deaths and emigration), and health care (deaths). You can do a student project on any of the four pieces, but keep in mind that a change in government policy or a major war can disrupt the ARIMA process.

Birth rates are less affected by government policy, wars, and health care. In *developing countries*, health care affects infant mortality. This effect is smaller in advanced economies, where infant mortality is low.

Birth rates have good statistics. Hospitals keep records of live births and most countries tabulate and publish the results.

Birth rates have several characteristics that are useful for the student project. A change in birth rates affects the *long-term population trend*, so it is a good indicator of population.

(1) Most countries show declining birth rates. Japan and Russia have such low birth rates that their populations are declining. Many Western European nations will have negative population growth in the 21^{st} century (Spain, Italy, Germany). In much of the developing world, birth rates are still high but they are well below earlier rates.

Each nation has its own trend, which you estimate from the data. Negative trends are more challenging to model, since the birth rates do not become negative. The trend is neither linear nor logarithmic; you can compare the two types of trend or choose another trend.

(2) Instead of a simple trend, demographers relate birth rates to girls' education, women's participation in the work force, religious commitment of the population, and other factors.

The low birth rates in Western Europe, China, and Japan are characteristic of secular (atheistic) countries with equal education for both sexes and high female participation in the work force. You may look at birth rates in Japan or Western Europe, fit an ARIMA process, and project future rates. China's birth rates reflect government policy, so the ARIMA model differs before and after the single-child policy.

(3) The U.S. shows a different pattern. Birth rates have not declined much in the twentieth century, and they even increased from 1945 to 1960 (the baby boom generation). The U.S. remains a religious country, and birth rates do not show the same decline evident in secular nations.

In the United States, birth rates are positively correlated with economic conditions. We use the U.S. experience for this project template.

(4) Birth rates show a moving average component, which is less common in the other time series on the NEAS web site. We *test for* moving average components in many time series, but interest rates, daily temperatures, and most economic indices do not seem to have moving average components.

We explain the rationale for the moving average component by comparing babies to durable goods. Moving average components occur in sales of expensive, durable goods. Cars are a good example.

Illustration: Suppose the average car lasts five years and all consumers own one car. We add stochasticity and more realistic assumptions in a moment. In this ideal scenario, a consumer has a 20% probability of buying a car each year.

A car may last more or less than five years, and consumers have different views about when the car wears out, so any consumer's purchase is stochastic.

Suppose a consumer buys a car in 20X0. If the car lasts exactly five years and income stays constant, the consumer has a 0% probability of buying a car in 20X1 through 20X4 and a 100% probability of buying a car in 20X5.

But some cars last only 1 year and some cars last 9 years. The probability of buying a car may be 2%, 5%, 10%, 18%, 30%, 18%, 10%, 5%, 2% in the next nine years.

We now add economic conditions. Cars are expensive. Consumers are less likely to buy cars in recessions, when income is low, and more likely to buy cars in prosperous years, when income is high. This is not true for most consumer purchases, such as food, clothing, or books, but it is true for cars, homes, and other durable goods.

We may model car purchases as 20% + β × GDP growth. *GDP growth*, as used here, is the residual from the mean, and β is a positive coefficient. Suppose the mean GDP growth is 2% and β is 3. GDP growth may be –1% in recessions and +5% in prosperous years, so the residuals are –3% and +3%.

~ In recessions, the percentage of consumers buying new cars is 20% + 3 × (–3%) = 11%.
~ In prosperous years, the percentage of consumers buying new cars is 20% + 3 × (+3%) = 29%.

We combine the economic conditions and the probabilities for a consumer.

~ If the economy enters a recession in the fifth year (20X5), the 30% probability for a particular consumer may drop to 15%. With more old cars on the road, the demand for new cars rises. Even if the economy stays in a recession, the probability of buying a new car in the sixth year (20X6) may rise to 25%.
~ If the economy is prosperous in the fourth year (20X4), the 18% probability may rise to 35%. With fewer old cars on the road, the demand for new cars falls. Even if the economy stays prosperous, the probability of buying a new car in the fifth year (20X5) may drop to 25%.

We summarize the sales of new cars as:

Sales volume depends on economic conditions. If consumers buy many cars one year (autoregressive process), or more cars than expected one year (moving average process), they are likely to buy fewer cars the new year.

Babies

Children are the most expensive and most durable goods. Many women (couples) want two or three years (or more) in between children. Even if they would prefer children half a year apart, biology decrees otherwise.

Illustration: A young, childless couple in 20X0 may want three children. The probability of having a child in 20X1 may be 30%. If the couple does not have a child in 20X1, they may increase their efforts or consult a fertility clinic, and the probability of having a child in 20X2 may rise to 50%. If the couple has a child in 20X1, they may wait a year before trying to have another child, and the probability of having a child in 20X2 may drop to 10%.

This phenomenon may lead to a moving average process. If more women than expected have babies in year X, fewer women than expected may have babies in year X+1.

Children are extremely expensive. Mothers generally lose at least a year of work, and many mothers leave the work force entirely. For middle-income couples, the cost of raising a child may be several hundred thousand dollars.

Recessions cause a dip in the birth rates; prosperous years raise the birth rates. More precisely, they affect the conception rates, for which the birth rates are a proxy. The birth rate in Year T reflects economic conditions in year T-1.

Absolute Values vs Deviations

The relation between birth rates and macroeconomic conditions makes the ARIMA process fit better (and may make the student project more interesting). You may do a student project on birth rates using the absolute rates if you prefer. If you use U.S. data, you may divide the time series into different periods. If you use other countries, the time periods may depend on government policies (as the one child policy in China) or other social conditions, such as the dramatic twentieth century fall in birth rates in secular countries.

Using residuals from a regression analysis on explanatory variables often gives a better fitting ARIMA model and a better understanding of the time series dynamics.

STEP-BY-STEP GUIDE FOR A STRUCTURAL MODEL

For a simple structural model, use the following steps. This regression model is too simplistic for actual demographic analysis, but it gives a good student project.

Step #1: GDP Growth Rates

Form GDP growth rates by taking logarithms and then first differences of GDP. This gives a series of GDP growth rates that you can use as the explanatory variable for birth rates.

Step #2: Regression

Regress the birth rates on a lagged GDP growth rate. If the birth rates are calendar year figures, use the GDP growth rate of the previous calendar year. Even if you have monthly birth rates, the decision to have a baby reflects economic conditions of about a year before the birth.

Step #3: Residuals

Form graphs of the absolute birth rates and the residuals of the birth rates. If the residuals are smoother, fit an ARIMA process to the residuals. The relation of birth rates and GDP growth is not linear, so the residuals are not a white noise process.

You can examine the trends in the absolute rates and the residuals. If you use monthly birth rate figures, you can examine seasonality.

Step #4: ARIMA Modeling

Use the time series techniques to fit an ARIMA process to the absolute birth rates and the birth rate residuals. Use correlograms, regression analysis, the Yule-Walker equations, or other techniques.

Use simple ARIMA processes. The residuals of the birth rates may be a stationary time series, and a simple ARIMA process may fit well and forecast well. It is harder to fit the absolute birth rates to a simple ARIMA process, since the trend in birth rates depends on the time period.

~ Choose a time period with a homogeneous trend, take first differences, form a correlogram, and fit an ARIMA process to the first differences.
~ If your birth rates time series does not have a strong trend, you may not need to take first differences.

Step #5: Testing for White Noise

For each ARIMA process, test if the residuals are a white noise process, using the Box-Pierce Q statistic and Bartlett's test. We don't expect any ARIMA process to fit perfectly, since many other things affect birth rates.

These are the residuals of the ARIMA fitting, or the actual birth rates (or actual residuals from the regression analysis on GDP growth) minus the ARIMA estimates.

Step #6: Comparison

If you use both absolute birth rate and the regression analysis, judge whether the regression on the GDP growth rate improves the ARIMA fitting.