Predicting International Adoption Visas in the United State

Noora N. Saleh¹ and Kaitlyn Miner²

¹Department of Mathematics, College of Education, University of Al-Qadisiyah, Al-Diwaniyah, Iraq. noora.saleh@qu.edu.iq, Phone: +964 7821620051

²Department of Mathematics, College of Sciences and Humanities, Ball state University, Indiana, USA. krminer@bsu.edu

Abstract: For Americans who wish to adopt children from other countries, international adoption visas are issued to the child for entrance into the United States. With the laws and regulations are changing so often, the number of international adoptions fluctuating with these changes. We are interested in forecasting future visa values from prior visa value data. Our results support that it is possible to predict future international adoption visa numbers but for only a few years into the future due to the laws and regulations changing frequently.

Keywords: Adoption visas, Correlation matrix, Log-transformation, Outliers, Regression.

Introduction

United States is the highest country in adoption the half of all children across international borders (1). International Adoption Visas given to children being adopted by Americans in the US. Over the years, procedures for international adoption were mainly unregulated and inconsistent, virtually allowing whomever to adopt from any country of their choosing, no matter their qualifications (2). Each year, around 22,000 children join a family in the United State by Intercountry adoption (3). According to US
Department of State, the international adoption rose to 45,000 in 2004 meanwhile the adoption decrease in 2008 to 17,000 and then in 2011 increase to 25,000 (4).

Beginning in 1991 and officially instating on May 29, 1993, The Hague Adoption Convention established firmer rules and procedures to regulate international adoption. This protects children from being adopted by families that are ill prepared or ill-intended as well as reinforcing Article 21 of the UN Convention on the Rights of the Child by making sure international adoptions are always settled with the child’s rights and best interests in mind (5). Based on previous studies, we anticipated that we would be able to forecast future visa values from prior visa value data.

Our data contains of 39 countries of the world that Americans who wish to adopt children from them. Visas issued by the US Immigration and Naturalization Service in 1988, 1991, and 1992. We have two data points missing – Greece and Portugal in 1992. In particular, we are interested in forecasting the 1992 adoption visa data from the 1988 or the 1991 data.

Methods

When building a model to forecast 1992 visas, we concluded that using the 1991 data was the most beneficial. In particular, we built our model using the square root of the 1991 data as shown below:

\[
\text{Visas issued}_{1992} = -62.5 + 17.86 \sqrt{1991}
\]

After trying a variety of transformations, we believe that the square root transformation produced the best resulting model and this conclusion is supported by its Minitab output. The output as shown in Table 1 displays that the square root of 1991 variable is significant with a high F-Value and a P-Value that is significant for alpha<0.000. The Adjusted R Squared for our model is 89.24% meaning that the model we have chosen explains nearly 90% of the variation in our data, which is an satisfactory percentage given real world data. The residual plots for our model shown in Figure 1 reveal that this model also satisfies the regression assumptions of normality, constant variance, and independence indicating that this is a valid model with a good fit.
Table 1: Regression Analysis for Visas issued1992 versus sqrt(1991)

Regression Analysis: Visas issued1992 versus sqrt(1991)

Analysis of Variance

| Source        | DF | Adj SS | Adj MS | F-Value | P-Value |
|---------------|----|--------|--------|---------|---------|
| Regression    | 1  | 250451 | 250451 | 249.89  | 0.000   |
| sqrt(1991)    | 1  | 250451 | 250451 | 249.89  | 0.000   |
| Error         | 29 | 29066  | 1002   |         |         |
| Lack-of-Fit   | 26 | 28199  | 1085   | 3.76    | 0.151   |
| Pure Error    | 3  | 867    | 289    |         |         |
| Total         | 30 | 279517 |        |         |         |

Model Summary

| S     | R-sq | R-sq(adj) | R-sq(pred) |
|-------|------|-----------|------------|
| 31.6586 | 89.60% | 89.24% | 87.45% |

Coefficients

| Term     | Coef | SE Coef | T-Value | P-Value | VIF |
|----------|------|---------|---------|---------|-----|
| Constant | -62.5| 10.8    | -5.78   | 0.000   | 1.00|
| sqrt(1991)| 17.86| 1.13    | 15.81   | 0.000   | 1.00|

Regression Equation

Visas issued1992 = -62.5 + 17.86 sqrt(1991)

Figure 1: Residual Plot for 1992 vs sqrt(1991) without outliers
Analysis of Complete Data Set

Our original data set included information from 39 countries or regions of the world. We ran an initial regression and analyzed the results along with the residual plots. The resulting residual plots are shown below in Figures 2 and 3. Analysis of these residual plots show that nearly none of the regression assumptions of normality, constant variance, or independence hold when modeling this data and that there is clear evidence of outlying data points. Table 2 shows the Minitab output for the 1992 vs. 1991 and clearly displays a low adjusted R Squared value indicating that this model would explain only 35.68% of the variation. While Table 3 showing the Minitab output for the 1992 vs. 1988 data indicates that the variables are significant with low p-values and the adjusted R-Squared value is high at 92.42%, we may not reasonably choose this model with the residuals indicating that none of the regression assumptions are being met.

![Residual Plots for Visas issued 1992 vs Visas Issued 1988](image)

*Figure 2. Original 1992 Visa Data vs. 1988 Visa Data Minitab Residual Plots*
Figure 3. Original 1992 Visa Data vs. 1991 Visa Data Minitab Residual Plots

Table 2: Regression Analysis for Visas issued1992 versus Visas issued1991

| Source              | DF | Adj SS | Adj MS | F-Value | P-Value |
|---------------------|----|--------|--------|---------|---------|
| Regression          | 1  | 1225091| 1225091| 20.97   | 0.000   |
| Visas issued1991    | 1  | 1225091| 1225091| 20.97   | 0.000   |
| Error               | 35 | 2045086| 58431  |         |         |
| Lack-of-Fit         | 32 | 2044220| 63882  | 221.17  | 0.000   |
| Pure Error          | 3  | 867    | 289    |         |         |
| Total               | 36 | 3270177|        |         |         |

Model Summary

- \( S = 241.725 \)
- \( R^2 = 37.46\% \)
- \( R^2(\text{adj}) = 35.68\% \)
- \( R^2(\text{pred}) = 0.00\% \)

Regression Equation

\[
\text{Visas issued1992} = 71.4 + 0.3638 \times \text{Visas issued1991}
\]
Table 3: Regression Analysis for Visas issued1992 versus Visas issued1988

| Regression Analysis: Visas issued1992 versus Visas issued1988 |
|---------------------------------------------------------------|
| **Analysis of Variance**                                      |
| Source                              | DF | Adj SS   | Adj MS | F-Value | P-Value |
|-------------------------------------|----|----------|--------|---------|---------|
| Regression                          | 1  | 3029041  | 3029041| 439.65  | 0.000   |
| Visas issued1988                    | 1  | 3029041  | 3029041| 439.65  | 0.000   |
| Error                               | 35 | 241136   | 6890   |         |         |
| Lack-of-Fit                         | 31 | 232603   | 7503   | 3.52    | 0.114   |
| Pure Error                          | 4  | 8534     | 2133   |         |         |
| Total                               | 36 | 3270177  |        |         |         |

**Model Summary**

| S     | R-sq | R-sq(adj) | R-sq(pred) |
|-------|------|-----------|------------|
| 83.0036| 92.63%| 92.42% | 54.42%     |

**Regression Equation**

Visas issued1992 = 71.5 + 0.3570 Visas issued1988

After looking at these residuals and ANOVA tables, we ran a correlation matrix to provide insight into which year, 1991 or 1988, would be best to predict the 1992 visa data. In the correlation matrix shown below in Table 4, it is evident that the 1992 visa data is most highly correlated with the 1991 visa data with a correlation coefficient value of .904 and a p-value as low as 0.000. Based on these results we decided to only move forward with models using the 1991 visa data. From the analysis of our original data set we determined that we needed to look for and eliminate outliers in order to improve our residual plots and satisfy the regression assumptions as well as move forward only with models built from the 1991 data.

Table 4: Correlation for Visas issued1992, Visas issued1988, Visas issued1991

| Correlation: Visas issued1992, Visas issued1988, Visas issued1991 |
|---------------------------------------------------------------|
| Visas issued1992                                          |
| Visas issued1988                                          | 0.809 | 0.000 |
| Visas issued1991                                          | 0.904 | 0.885 |

**Cell Contents:** Pearson correlation

P-Value
Determining and Eliminating Outliers

Based on the analysis of these models and their residual plots, we decided to create boxplots of our original data set to determine which outlying points we should investigate. The box plots shown below resulted from the 1992 and 1991 data.

The box plot of the visa data from 1991 show that data entry 34 (Romania) is the largest outlier in this set. This coupled with research that international adoption of Romanian children spiked after the overthrow of the Ceausescu Regime in December 1989 allowed us to reasonably eliminate this outlier (6).

Another significant outlier in the 1992 and 1991 box plot was data entry 35 (South Korea). After the Korean War, US adoption of South Korean children was strong because many had American fathers who were soldiers in the war. Industrialization and urbanization were also increasing during this time and there were many more babies to adopt. These factors drove up adoption numbers in 1991 and 1992, making them larger than most others in this data set and allowing us to reasonably also eliminate this country (7).
Other outliers that we have researched and eliminated in a similar manner include Peru, China, Guatemala and Colombia. Articles with supporting research for these countries can be found under Sources. After eliminating these outliers, we ran our regression using the square root of the 1991 data, which resulted in the residual plots shown in Figure 4. As mentioned in the initial discussion of our model, we attempted several transformations including \( \text{sqrt}(1991) \), \( \ln(1991) \), \( (1991)^2 \), and \( (1991)^{0.25} \). We used the correlation matrix shown in Table 5 as well as the output from running models with the 4 transformations to conclude that the square root of the 1991 data produced the best results. Based on the Normal Probability Plot and the Histogram, our residuals appear to be normally distributed. There does not appear to be a pattern in the residuals, meaning that we can assume independence and the points appear to be spread evenly and not funneling, demonstrating constant variance. Without these six outliers in our data set and using the square root transformation, our residuals now show evidence that all regression assumptions are met.

Figure 4: Residual Plots for 1992 vs. sqrt(1991) with outliers removed
Table 5: Correlation for Visas issued 1992, 1991^2, ln(1991), sqrt(1991), (1991)^.25

|                      | Visas issued 1992 | 1991^2 | ln(1991) | sqrt(1991) | (1991)^.25 |
|----------------------|-------------------|--------|----------|------------|------------|
| sqrt(1991)           |                   | 0.859  | 0.000    |            |            |
| 1991^2               |                   |        | 0.834    | 0.000      |            |
| ln(1991)             |                   |        |          | 0.841      | 0.949      |
| sqrt(1991)           |                   |        |          |            | 0.949      |
| (1991)^.25           |                   | 0.047  | 0.000    | 0.000      | 0.000      |
| 0.988                |                   |        | 0.781    | 0.986      | 0.000      |
| 0.000                |                   |        | 0.000    |            | 0.000      |

The output for this improved model shown in Table 1 shows that taking out these outliers has resulted in an R-Squared as high as 89.24% and low p-values for the square root of 1991 variable as well as the model as a whole. Based on this information and seeing the improvement from all other models that we had created, we decided to make this our final model without using data from the countries of South Korea, Romania, Peru, China, Guatemala and Colombia and using the square root of the 1991 international visa data.

Precautions

We believe that our model is robust, valid, and accurately predicts the 1992 visa data from the 1991 visa data provided. However, we do have some concerns with the application of this model in future use. Firstly, we do not believe that with the volatility of this subject and ever changing legislation that it is possible to accurately predict future visa values with this model more than a couple of years out into the future. As our original analysis of the complete data set indicated, the means and totals of visa values changed drastically over time and it would be increasingly difficult to predict visa values further out in time.

Secondly, it could be a problem that using more current visa value data would be skewed because of the “trendiness” of international adoption. With the prominence of celebrities adopting abroad, it has romanticized the idea of international adoption as a whole and may have encouraged others to do so. This
could possibly cause visa data values in the past to not be comparable to recent years and make future models less predictive unless they take these factors into consideration.

Conclusions

Examining the data from 1988, 1991 and 1992, we were able to draw some general conclusions about international adoption numbers over time. In 1988, the mean value of adoptions was 232 and then decreased to 158.4 in 1992. The output for these findings can be found in Table 6. Similarly, the total number of adoptions in 1988 was 9,033 and decreased to 8,859 in 1991 and then further decreased to 5,860 in 1992. These were found by summing the data in Minitab. What these findings reveal to us is that the international adoption visa numbers are not stable over time but rather variable over time. As time progresses, there is a clear trend that adoption visa numbers are declining. These findings are in accordance with the historical background of international adoption visas and the meeting of The Hague Adoption Convention, with laws being more lenient in the 80’s, adoption numbers were higher because families did not have to go through as much trouble in order to adopt a child abroad. As laws became firmer, we see a decline in the number of adoption visas conceivably because the process of adopting a child internationally became more detailed, time consuming, and required that superior criteria be met (3).

Based on the model that we have created and explained in the Analysis section, we believe that it is possible to predict a few future years of international adoption visa numbers. With laws and regulations changing frequently it is difficult to model how the numbers for international adoption visas may look five or more years out into the future. For example, we see that the changes in the numbers from 1988 to 1991, merely three years into the future, are staggeringly different. Even though the changes in law that occurred during this period were drastic, we believe with the constant volatility of this topic it will only be possible to accurately predict international adoption visa values for a few years at a time.

Table 6: Descriptive Statistics for Visas issued in 1988, Visas issued in 1991, Visas issued in 1992
References

1. Barratt MS. International adoption. Pediatrics in review. 2013 Mar;34(3):145-146.
2. Ahlin L. Writing and Identity in Jane Jeong Trenka’s Life Narratives. International Adoption in North American Literature and Culture 2017 (pp. 121-142). Palgrave Macmillan, Cham.
3. Engel M, Phillips NK, Dellacava FA. International adoption: a sociological account of the US experience. International Journal of Sociology and Social Policy. 2007 Jun 26: 27(5/6), 257-270.
4. Zhang Y. Intercountry Adoption to the United States. Encyclopedia of Family Studies. 2016 Feb 23:1-5.
5. Breuning M, Xi J. The consequences of accession: the Hague Convention on Intercountry Adoption’s Impact on Children’s Rights. Journal of International Relations and Development. 2019 Jun;1-24.
6. Palacios J, Adroher S, Brodzinsky DM, Grotevant HD, Johnson DE, Juffer F, Martínez-Mora L, Muhamedrahimov RJ, Selwyn J, Simmonds J, Tarren-Sweeney M. Adoption in the service of child protection: An international interdisciplinary perspective. Psychology, Public Policy, and Law. 2019 May;25(2):57.
7. Schachter J. Intercountry Adoption/Global Migration: A Pacific Perspective. The Asia Pacific Journal of Anthropology. 2017 Aug 8;18(4):305-22.