

Interpolation

URBPL 5/6010: Urban Research

University of Utah

Pam Perlich

Rev. 10/19/2006



Interpolation

- Procedures used to estimate values in a data set when the data point falls between two known points
- Aggregate Approach
 - Fit an function through a scatter and solve for the point
 - Use with measured data
- Two Point Computation
 - Use when working with precise data points

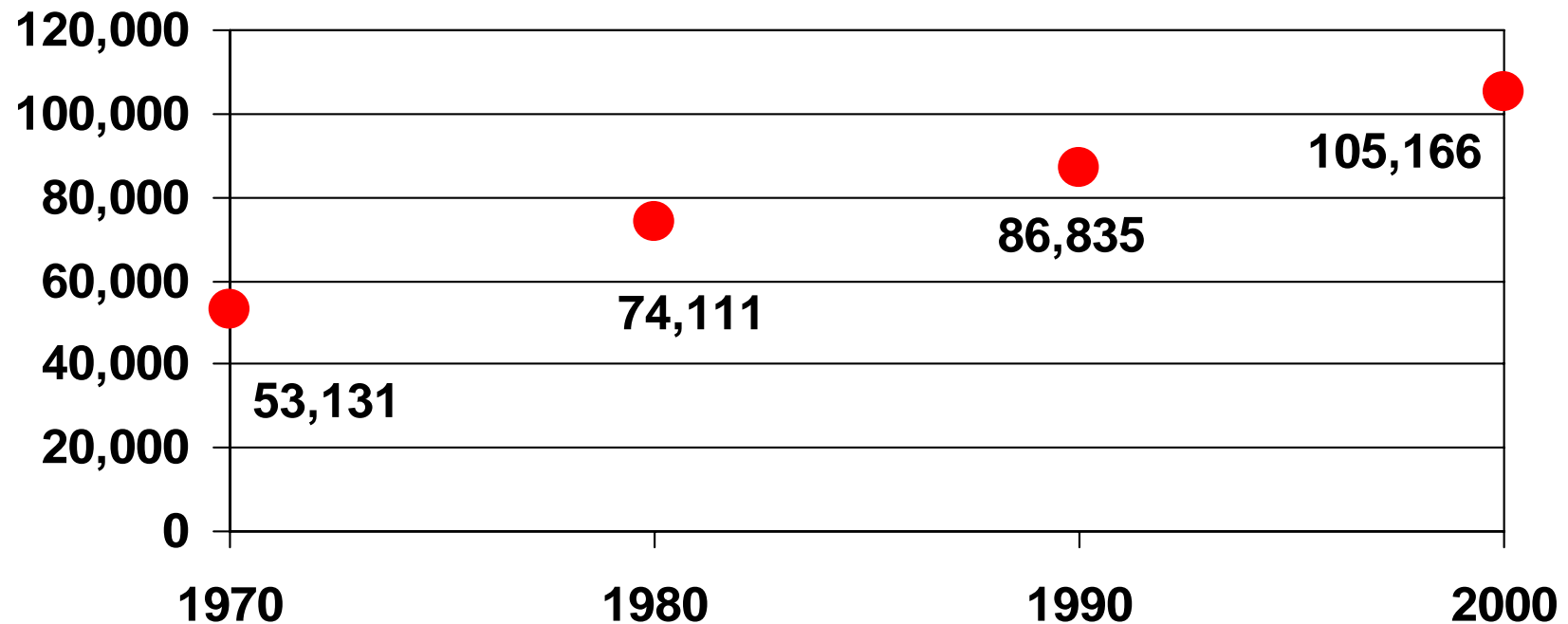


Functions

- Linear Interpolation - Straight Line
- Polynomial Interpolation
- Pareto Interpolation – log transformation
 - Often used with interval income data
- Other functions

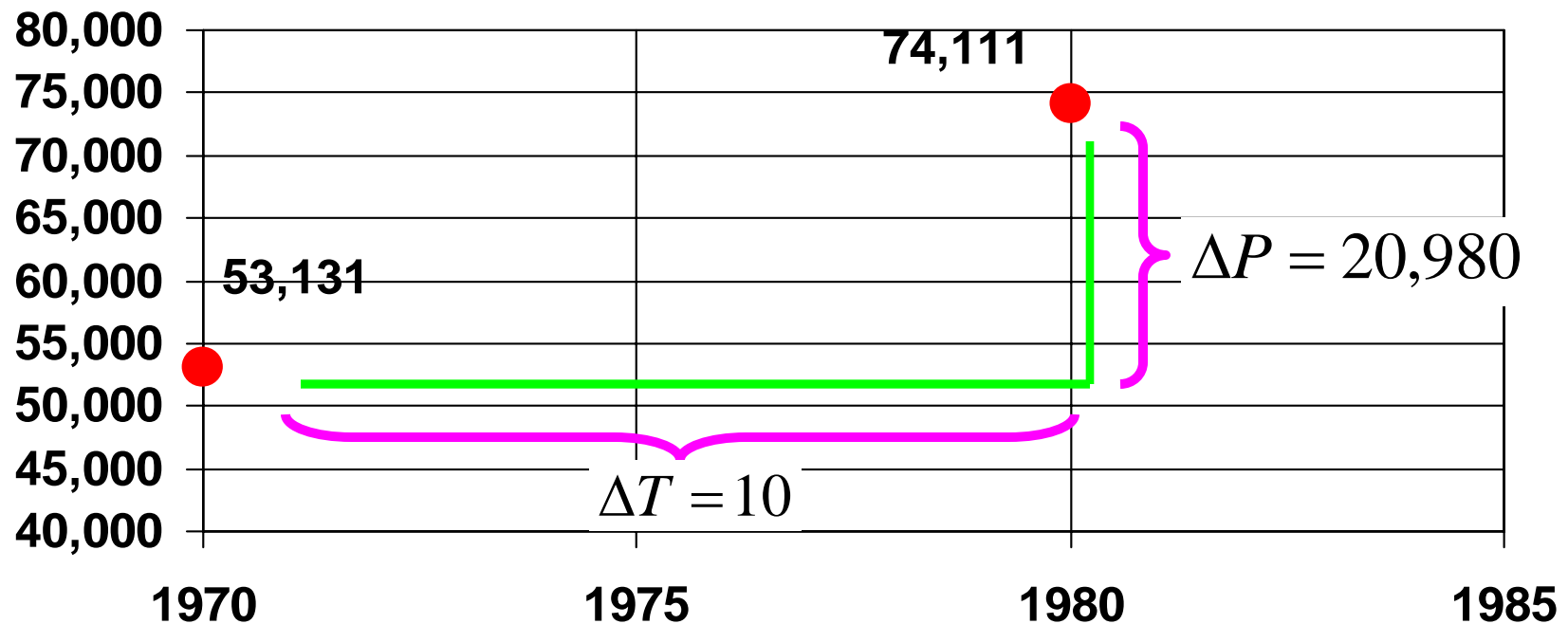
Simplified Example: Linear

Decennial Census Counts: Provo, Utah



Estimate April 1, 1972 Population

Decennial Census Counts: Provo, Utah



Calculation

$$\frac{\Delta P}{\Delta T} = \frac{20,908}{10} = \frac{2,098}{1}$$

$$P_{1972} = P_{1970} + (2,098 \times \Delta T)$$

$$P_{1972} = P_{1970} + (2,098 \times 2)$$

$$P_{1972} = 53,131 + 4,196$$

$$P_{1972} = 57,327$$



Median Income Interpolation

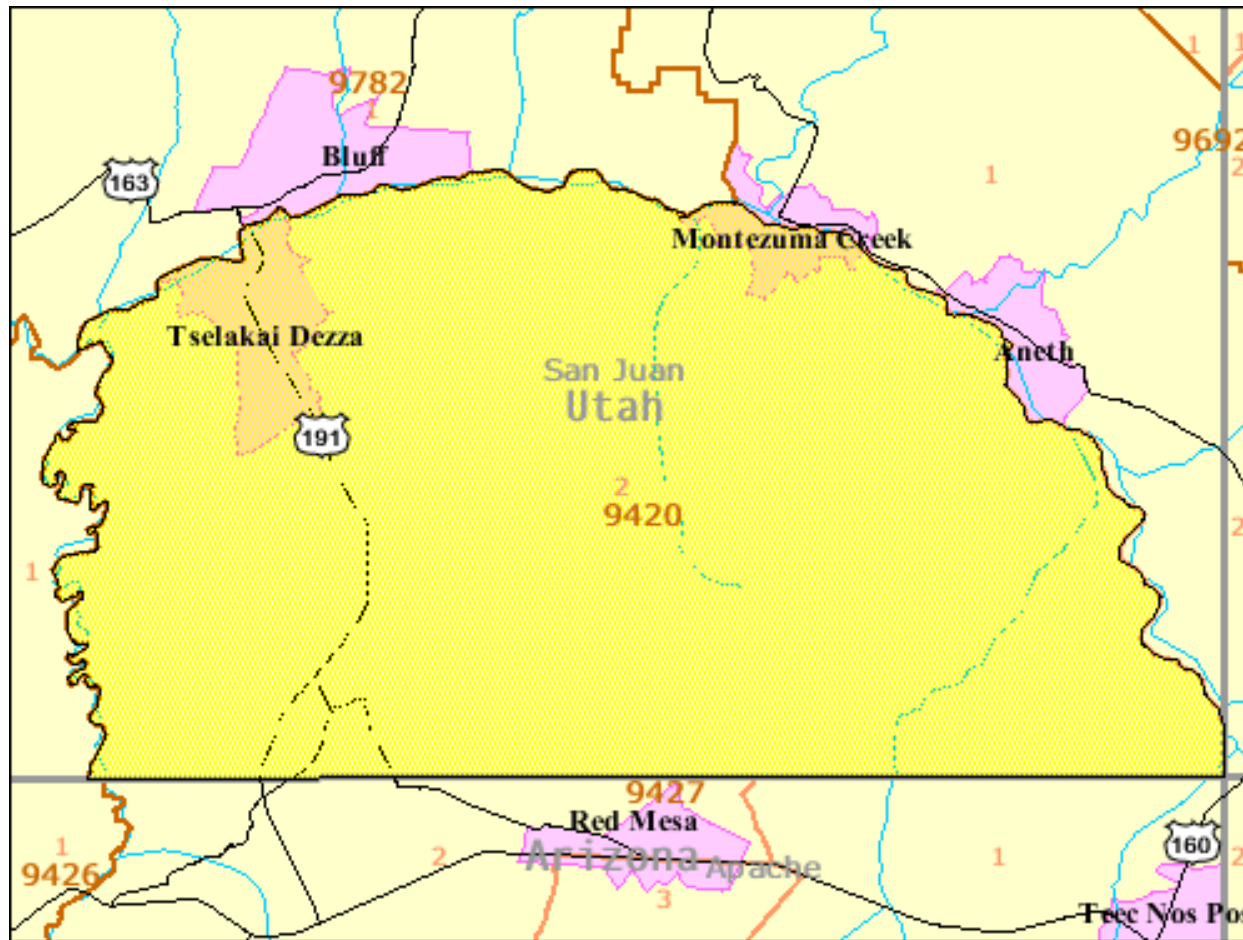
- Common calculation - median income given interval data.
- Methods may include:
 - Linear interpolation
 - Pareto – used by Bureau of the Census for intervals $> \$2,500$
- See Examples in Worksheet in Lab 6



Linear Interpolation Example

- ❑ Interpolation.xls
- ❑ Calculate the cumulative distribution is ascending order
- ❑ Identify where the middle data element occurs.
- ❑ Calculate how far into the particular income interval the median element occurs.
- ❑ Assume a rectangular distribution.
- ❑ Compute the median income from the assumed share of the interval at which the median element occurs.

Block Group 2, Census Tract 9420, San Juan County, Utah



Block Group 2, Census Tract 9420, San Juan County, Utah

Income Category	Frequency	Cumulative Frequency	
		Count	Percentage
Total:	439		
Less than \$10,000	159	159	36.2%
\$10,000 to \$14,999	99	258	58.8%
\$15,000 to \$19,999	19	277	63.1%
\$20,000 to \$24,999	5	282	64.2%
\$25,000 to \$29,999	7	289	65.8%
\$30,000 to \$34,999	16	305	69.5%
\$35,000 to \$39,999	7	312	71.1%
\$40,000 to \$44,999	28	340	77.4%
\$45,000 to \$49,999	27	367	83.6%
\$50,000 to \$59,999	12	379	86.3%
\$60,000 to \$74,999	34	413	94.1%
\$75,000 to \$99,999	18	431	98.2%
\$100,000 to \$124,999	0	431	98.2%
\$125,000 to \$149,999	0	431	98.2%
\$150,000 to \$199,999	8	439	100.0%
\$200,000 or more	0	439	100.0%

U.S. Census Bureau

Block Group 2, Census Tract 9420, San Juan County, Utah

Income Category	Frequency	Cumulative Frequency	
		Count	Percentage
Total:	439		
Less than \$10,000	159	159	36.2%
\$10,000 to \$14,999	99	258	58.8%
\$15,000 to \$19,999	19	277	63.1%
\$20,000 to \$24,999	5	282	64.2%
\$25,000 to \$29,999	7	289	65.8%

\$30,
\$35,
\$40,
\$45,
\$50,
\$60,
\$75,
\$100,
\$125,
\$150,
\$200

Median is the middle element $\rightarrow 439/2 = 219.5$

This happens in the \$10,000 to \$14,999.99 interval

This interval contains elements 160 through 258 and contains 99 elements.

The median occurs 59.5 elements into the interval or 60.101% into the interval.

The income interval is \$4,999.99.

60.101% times \$4,999.99 or \$3,005.

Add this to the lower bound of the interval to get \$13,005.

U.S. Census Bureau



Resources

- Excel worksheet in Lab 6 and 7

<http://home.business.utah.edu/bebrpsp/URPL5010/Labs/Lab6/Interpolation.xls>

- Siegel, p. 15 + footnote on p. 29

- Shryock and Siegel “Interpolation of Grouped Data,” on ereserve here:

<http://ereserve.lib.utah.edu/webpac-1.2-bin/DoReserve?coursernum=5966&instructor=perlich>