Statistical Approaches to Analyzing Trends and Differences in Trends: Cancer Data from Windsor-Essex

Clifford Ekwempe
Ali Artaman
Liz Haugh
Outline

• Introduction
• Joinpoint Regression Model
• JoinPoint Regression Program
• Difference in Trends
Trend

• A **time series** is a collection of observations of well-defined data items obtained through repeated measurements over time.

• An observed time series can be decomposed into three components:
  
  – **seasonal** (systematic, calendar related movements)
  – **irregular** (unsystematic, short term fluctuations)
  – **trend** (long term direction)
Cancer Incidence Projection Programs Used in Provincial/Territorial Cancer Registries

- NordPred
- Linear
- Poisson
- ACS* Joinpoint
- ACS Spatial modeling

*ACS: American Cancer Society

Source: Public Health Agency of Canada presentation June 09, Results of Survey on Cancer Projection Methods Used in Provincial/Territorial Cancer Registries
Trend Analysis

Joinpoint Regression and Software
Joinpoint Regression Model

- Piecewise regression
- Change point regression
- Segmented regression
- Broken line regression
- Multiphase regression
\((x_1, y_1), ..., (x_n, y_n), \text{ where } x_1 \leq ... \leq x_n\)

\[E[y \mid x] = \beta_0 + \beta_1 x + \delta_1 (x - \tau_1)^+ + ... + \delta_k (x - \tau_k)^+\]

\(\beta_0, \beta_1\) regressions coefficients

\(\tau_k\) unknown joinpoints (or break points)

\(\delta_k\) differences in slope

\((x - \tau_k)^+ = (x - \tau_k) \text{ if } (x - \tau_k) > 0\)

\[= 0, \text{ otherwise}\]
A HEALTHY PART OF YOUR LIFE

www.wechealthunit.org
Joinpoint selects the final model using two different methods:

- Permutation Test (PT)
- Bayesian Information Criterion (BIC)

The PT approach controls the error probability of selecting the wrong model at a certain level (i.e. 0.05). It works well for cancer incidence and mortality data.

The BIC approach finds the model with the best fit by penalizing the cost of extra parameters.

The models picked by BIC tend to fit the data well but are less parsimonious.
Permutation Test:

- $H_0$: number of joinpoints = $k_a$
  $H_a$: number of joinpoints = $k_b$

- The procedure begins with $k_a = \text{MIN}$ and $k_b = \text{MAX}$. If the null is rejected, then increase $k_a$ by 1; otherwise, decrease $k_b$ by 1.

- The procedure continues until $k_a = k_b$ and the final value of $\hat{\kappa} = \hat{k}_a = \hat{k}_b$ is the selected number of joinpoints.
Linear or Log-linear

- **Purpose of log transformation**: Cancer rates arise from a Poisson distribution which is skewed especially when the cancer is *rare* or the rates come from a *small* population.

- Rates for *common* cancers or which come from a *large* population can be approximated as arising from a normal distribution without a transformation.
Annual Percentage Change

• To determine significance of trends: **APC**

For \( x_1, \ldots, x_n \) representing years
\( y_1, \ldots, y_n \) representing log of the observed rate

The APC between two joinpoints \( \tau_j \) and \( \tau_{j+1} \) is given by:

\[
APC = 100 \times (e^{\beta_1 + \delta_1 + \delta_2 + \cdots + \delta_j} - 1)
\]

• Positive value of APC suggests an increasing trend
• Negative value of APC suggests a decreasing trend
Cancer Data

Essex County vs. Ontario
Graphs

No Joinpoint!
Joinpoint Regression for the Incidence Rates of all Cancers Combined, Windsor-Essex County, and Ontario 1986-2005

Source: Cancer Care Ontario - SEER*Stat Release 7 - OCRIS (February 2009) released March 2009
Joinpoint Regression of Mortality Rates, All Cancers Combined, Windsor-Essex County and Ontario, 1986-2005

Source: Cancer Care Ontario - SEER*Stat Release 7 - OCRIS (February 2009) released March 2009
Joinpoint Regression of Cervical Cancer Incidence Rates, Windsor-Essex County and Ontario, 1986-2005

Source: Cancer Care Ontario - SEER*Stat Release 7 - OCRIS (February 2009) released March 2009
Joinpoint Regression of Incidence Rates, All Cancers Combined, by Sex, Windsor-Essex County, 1986-2005

Source: Cancer Care Ontario - SEER*Stat Release 7 - OCRIS (February 2009) released March 2009
Joinpoint Regression of Mortality Rates, All Cancers Combined, by Sex, Windsor-Essex County, 1986-2005

Source: Cancer Care Ontario - SEER*Stat Release 7 - OCRIS (February 2009) released March 2009

A HEALTHY PART OF YOUR LIFE

www.wchealthunit.org
Joinpoint Regression of Lung Cancer Incidence Rates, by Sex, Windsor-Essex County, 1986-2005

Source: Cancer Care Ontario - SEER*Stat Release 7 - OCRIS (February 2009) released March 2009
Joinpoint Regression of Lung Cancer Mortality Rates, By Sex, Windsor-Essex County, 1986-2005

Source: Cancer Care Ontario - SEER*Stat Release 7 - OCRIS (February 2009) released March 2009

Source: Cancer Care Ontario - SEER*Stat Release 7 - OCRIS (February 2009) released March 2009
Graphs

Joinpoints
Joinpoint Regression for Prostate Cancer Incidence Rates, Windsor-Essex County and Ontario, 1986-2005

Source: Cancer Care Ontario - SEER*Stat Release 7 - OCRIS (February 2009) released March 2009
Joinpoint Regression of Prostate Cancer Mortality Rates, Windsor-Essex County and Ontario, 1986-2005

Source: Cancer Ontario - SEER*Stat Release 7 - OCRIS (February 2009) released March 2009
Non-Significant Change in Slopes

The permutation test procedure does not require the asymptotic *normality* and maintains the correct Type I error probability level.
Joinpoint Regression for Lung Cancer Incidence Rates, Windsor-Essex County and Ontario, 1986-2005

Source: Cancer Care Ontario - SEER*Stat Release 7 - OCRIS (February 2009) released March 2009

A HEALTHY PART OF YOUR LIFE

www.wechealthunit.org
Joinpoint Regression of Lung Cancer Mortality Rates, Windsor-Essex County and Ontario, 1986-2005

Source: Cancer Care Ontario - SEER*Stat Release 7 - OCRIS (February 2009) released March 2009
Joinpoint Regression of Melanoma Mortality Rates, Windsor-Essex County and Ontario, 1986-2005

Source: Cancer Care Ontario - SEER*Stat Release 7 - OCRIS (February 2009) released March 2009
Joinpoint Regression for Colorectal Cancer Incidence Rates, Windsor-Essex County and Ontario, 1986-2005

Source: Cancer Care Ontario - SEER*Stat Release 7 - OCRIS (February 2009) released March 2009
Joinpoint Regression of Breast Cancer Mortality Rates, Windsor-Essex County and Ontario, 1986-2005

Source: Cancer Care Ontario - SEER*Stat Release 7 - OCRIS (February 2009) released March 2009
Difference in Trends

• Previous question:
  – *Is there a statistically significant increase/decrease of incidence/mortality rates in a single geographic region (WEC, Ontario)?*

• To test whether there are statistically significant differences in time trends in incidence and mortality rates *between* WEC and Ontario as well as *between* males and females within WEC, **negative binomial** regression was performed.
Poisson Regression

- For continuous outcome: *simple* regression

- Response variable in the form of a count: *Poisson* regression

- In Poisson distribution, mean and variance are equal.

- Data that have greater variance than the mean are termed as *Poisson overdispersed*

- The Poisson goodness-of-fit was performed (through *poisgof* command in stata). The test showed that Poisson regression was not a good choice.
Negative Binomial Regression

• To capture any statistically significant difference in trend, that is, to allow slopes of the rates to vary by region (WEC vs. Ontario) as time evolves:
  
  – An *interaction* term consisted of region and time variables (*regtime*) was included in the model.
  
  – An *interaction* term for sex and time (*sextime*) was included.

• A p-value of *regtime* or *sextime* less than 0.05 suggests that the slopes differ significantly over time.
Difference in trends

• In WEC for the period 1986-2005:
  
  – significant difference in incidence rate trends of lung and oral cancers between males and females.

  – significant difference in the trends of mortality rates for all cancers combined and lung cancer.
Acknowledgement

• Dr. Hyune-Ju Kim (Syracuse University)
• Joinpoint Development Team (IMS, Inc.)
• Dr. Eric Holowaty (Cancer Care Ontario)
• Ms. Beth Theis (Cancer Care Ontario)
• Mr. Enayetur Raheem (WEC Health Unit)
Additional Technical Information >>>
Select References

Method:

Software:
Figure 1. Flow chart of the parameter settings and decision tree

Joinpoint Trend Analysis of Cancer Incidence and Mortality using Alberta Data

A HEALTHY PART OF YOUR LIFE
Joinpoint APCs not matching SEER*Stat APCs

The Joinpoint model assumes that the trend is *continuous* at the joinpoint, whereas the APC for the corresponding segment calculated by SEER*Stat does *not* impose the continuity constraint with the consecutive segment.
Joinpoint v. 3.5

• Released in April 2011. Important Changes included:

• The Autocorrelated Errors Options were re-enabled,

• The confidence intervals for the AAPC were modified to follow the t-distribution, and be identical to the CIs for the APC

• A statistical test was added for the comparison of AAPCs between 2 groups when the pairwise differences option is selected.

• In versions prior to 3.5, the default for the maximum number of joinpoints Modified BIC was added as a 3rd option for the Model Selection Method.
Joinpoint v. 3.4

- **Version 3.4.2** was released October 2009.
- **Version 3.4.3** released in April 2010. The changes included:
  - An error was corrected in the p-values for the **comparison test** when the two groups being compared were exactly identical. This error did not affect any comparisons in which the two groups varied, even slightly.
  - An error was corrected that occurred when reading in SEER*Stat export files with the **missing value** set to zero.
  - An error was corrected that could occur when **exporting** data.