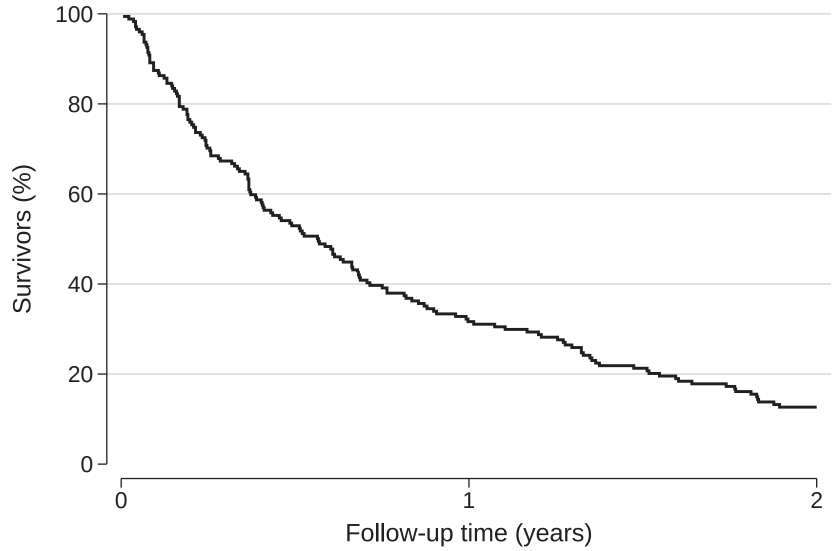
The grreg command for geometric rate regression

Matteo Bottai, Sc.D.

Unit of Biostatistics, Karolinska Institutet, Stockholm, Sweden

The grreg command was developed with Nicola Orsini

Survival in Cancer Patients



Survival in Cancer Patients

```
. strate
   [... output omitted ...]

  +----------------------------------------------+
  |   D          Y      Rate      Lower     Upper |
  |----------------------------------------------|
  | 152   136.9425   1.10996   0.94681   1.30121 |
  +----------------------------------------------+
```

A patient died on average 1.11 times a year.

A Trivial Example

A cohort of 100 subjects is followed up for 2 days.

Day	At risk	Deaths	Survivors	Geometric Rate
1	100	80	20	$80/100 = 0.80$
2	20	4	16	$4/20 = 0.20$

$$\text{Average geometric rate} = 1 - (16/100)^{1/2} = 0.60$$

The average daily risk is 60%.
If 60% died every day, 16 would survive 2 days.

The incidence rate is $84/120 = 0.70$ deaths/person-day.

Geometric Rate Regression

Poisson regression models incidence rates

Geometric rate regression models geometric rates

Geometric rate regression was introduced in Bottai (2015)

Geometric rates over Proportions

Let T be a time variable and $S(t) = P(T > t)$.

The geometric rate over the time interval $(0, t)$ is

$$g(0, t) = 1 - S(t)^{1/t}$$

The geometric rate over the proportion interval $(0, p)$ is

$$g(0, p) = 1 - (1 - p)^{1/Q(p)}$$

where $p = P(T \leq t)$ and $Q(p)$ is the quantile function.

The Primal Idea

Proposition (Bottai, 2015). *The geometric rate*

$$g(0, p) = 1 - (1 - p)^{1/Q(p)}$$

is the $(1 - p)$*-quantile of the transformed time variable*

$$T^* = 1 - (1 - p)^{1/T}$$

That is $P[T^* \leq g(0, p)] = 1 - p$.

The above follows directly from the fact that for a fixed p the function $1 - (1 - p)^{1/t}$ is monotonically decreasing in t for $t > 0$.

Geometric Rate Regression

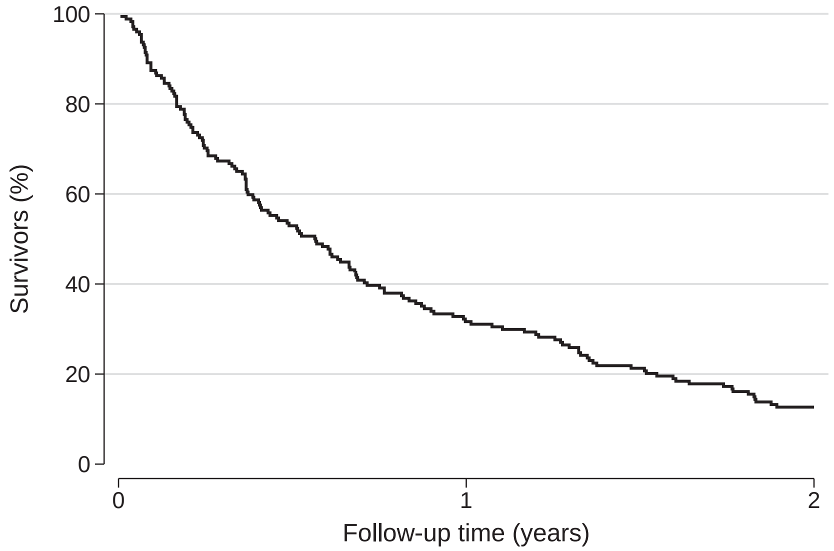
A geometric rate regression model with linear link

$$g(0, p \mid x) = x'\gamma$$

A geometric rate regression model with log link

$$\log g(0, p \mid x) = x'\gamma$$

Survival in Cancer Patients



The grreg Command

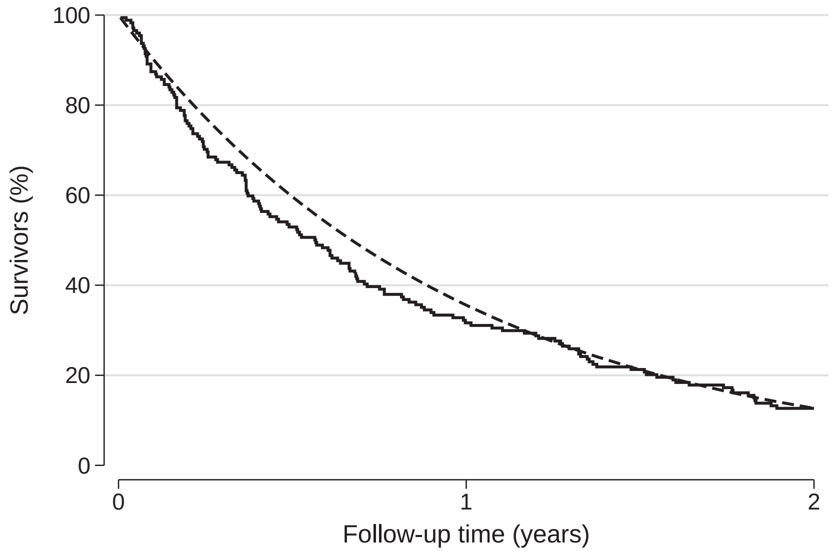
Eighty-nine percent of the patients died during follow up.
I estimate the following geometric regression model

$$\log g(0, 0.89) = \gamma_0$$

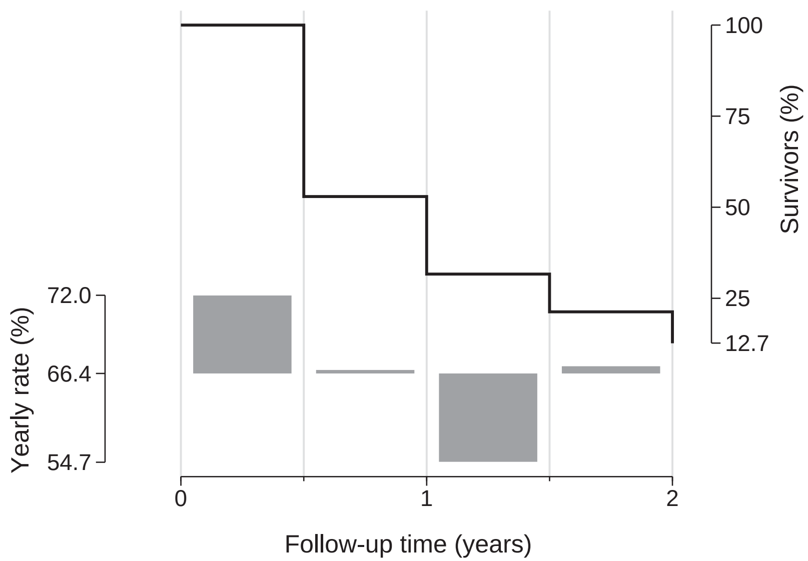
```
. grreg years, fail(died) p(.89)

Geometric rate regression
 Proportion: .89                                  No. of subjects  =        175
       Link: log                                  No. of failures  =        152
------------------------------------------------------------------------------
             |               Robust
       years | Rate Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   .6641147   .0148485   -18.31   0.000     .6356407    .6938642
------------------------------------------------------------------------------
```

The risk of dying in a year was 66.4% (0.3% in a day).

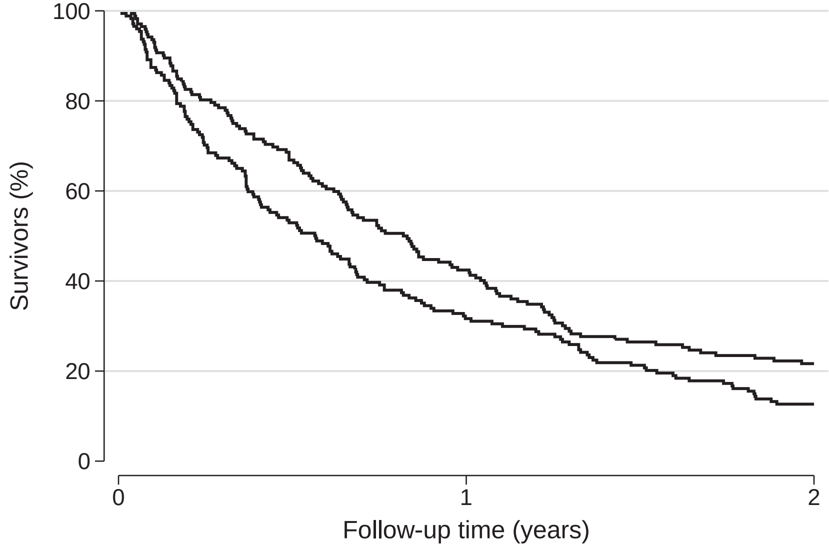
Survival at a Rate of 66.4% a Year (0.3% a Day)



Geometric Rates per Period



Survival by Treatment

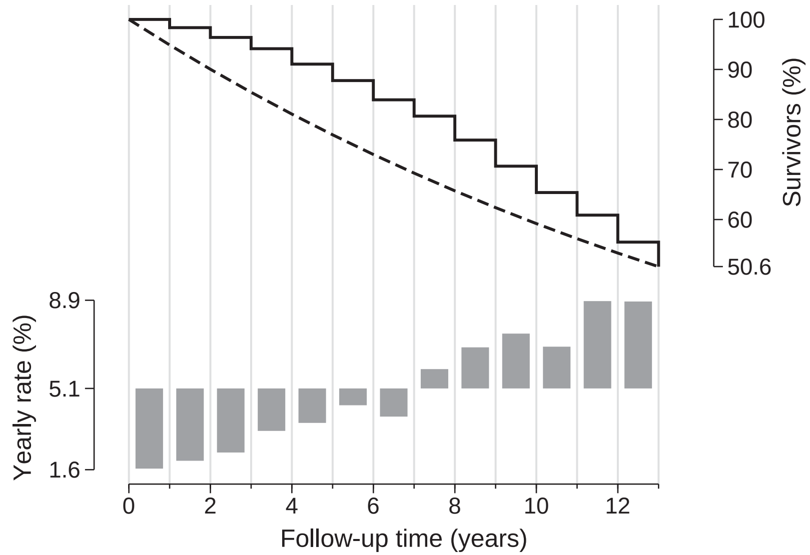


Geometric Rates by Treatment

```
. grreg years trt, f(died)

Geometric rate regression
 Proportion: .5                                     No. of subjects  =        347
       Link: log                                    No. of failures  =        286
------------------------------------------------------------------------------
           |               Robust
     years | Rate Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------+------------------------------------------------------------------
       trt |    .802882   .0724971    -2.43   0.015     .6726537    .9583229
     _cons |   .7054235   .0450773    -5.46   0.000     .6223824    .7995444
------------------------------------------------------------------------------
```

Half of the patients on MPA died at a yearly rate of 0.71.
The risk on IFN was 20% smaller.

Survival in a Cohort of Men from Sweden



Mortality Rates and Physical Activity

I estimated the following geometric rate regression model

$$\log g(0, 0.25 \mid x) = \gamma_0 + \gamma_1 \text{PA2} + \gamma_2 \text{PA3} + \gamma_3 \text{PA4} + \text{covariates}$$

The PA's were indicators of physical activity level.

This example is presented in Bottai (2015).
The data are described in Zheng (2015).

Mortality Rate Ratios

The table shows the estimated rates for the first 25% of deaths.

Physical Activity	Quantile (years)	Annual Rate (%)	Rate Ratio	
			Crude	Adjusted
Very low	6.3 (5.5 7.0)	4.5 (4.1 4.9)	1.00 (referent)	1.00 (referent)
Low	8.7 (7.9 9.4)	3.3 (3.0 3.6)	0.73 (0.64,0.82)	0.74 (0.65,0.84)
High	9.6 (8.9 10.4)	2.9 (2.7 3.2)	0.66 (0.58,0.75)	0.74 (0.65,0.83)
Very high	9.8 (9.1 10.6)	2.9 (2.6 3.2)	0.64 (0.57,0.73)	0.73 (0.64,0.82)

The mortality rate decreased over levels of physical activity.
In the most active it was 36% smaller than in the least active.

Final Remarks

The `grreg` command can estimate geometric rate regression.

Geometric rates have not been used in medical sciences, yet.

They have long been used in demography and finance.

Incidence rates are different from geometric rates.

Assumptions can improve efficiency; for example

$$S(t) = \exp[-(\lambda t)^\theta] \quad \Leftrightarrow \quad g(0,p) = 1 - (1-p)^{\lambda[-\log(1-p)]^{-1/\theta}}$$

References

- Bottai M, A regression method for modelling geometric rates. *Statistical Methods in Medical Research*, 2015 (to appear)
- Medical Research Council Renal Cancer Collaborators. Interferon-α and survival in metastatic renal carcinoma: early results of a randomised controlled trial *Lancet*, 1999, 353:14-17
- Zheng Selin J, Orsini N, Ejdervik Lindblad B, Wolk A. Long-Term Physical Activity and Risk of Age-Related Cataract: A Population-Based Prospective Study of Male and Female Cohorts. *Ophthalmology*, 2014, 122(2):274-80