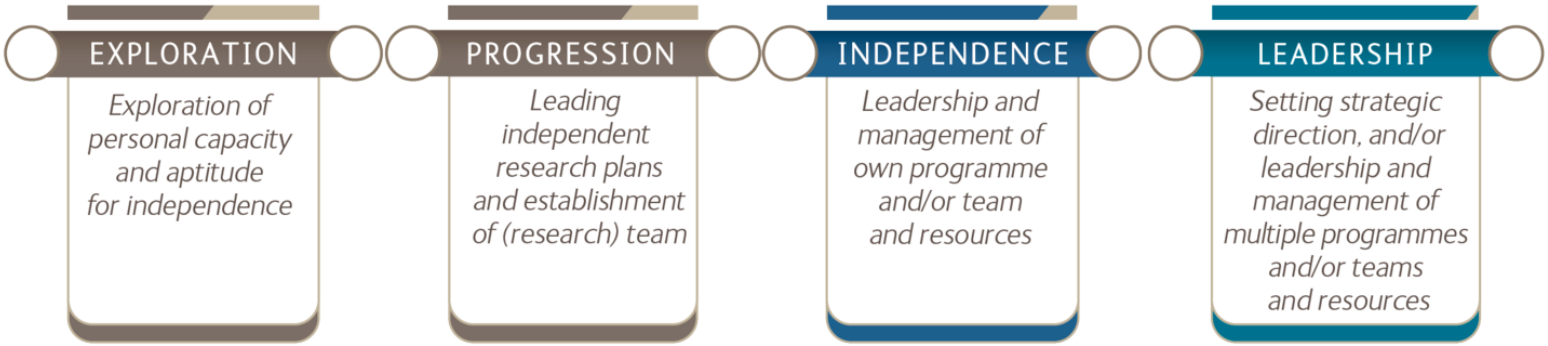
LMS Prospects in Statistics

Christl Donnelly

Department of Statistics
University of Oxford

Department of Infectious Disease Epidemiology
Imperial College London

Where are you now?



https://mrc.ukri.org/skills-careers/fellowships/non-clinical-fellowships/senior-non-clinical-fellowship-sncf/

This is Statistics | Statistics Jobs

https://thisisstatistics.org

Most Visited Getting Started when_does_it_make_s...

THIS IS STATISTICS

ASA

Students Parents Educators Counselors Digital Classroom News

About ASA

Statistics—It's not what you think it is.

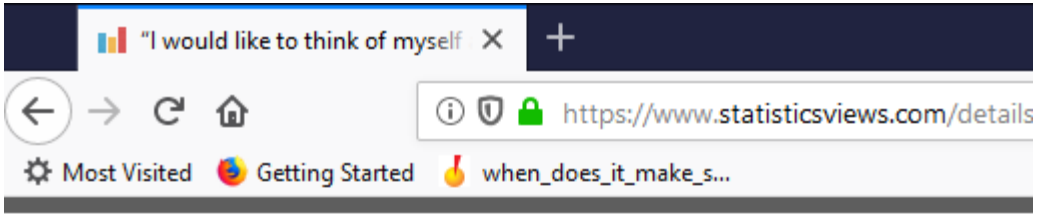
With a career in statistics you can:

Change the World

Have Fun

Satisfy Curiosity

Make Money



STATISTICS VIEWS

HOME | BIOSTATISTICS | BUSINESS & ECONOMICS | ENGINEERING | ENVIRO

Home → Search Results → "I would like to think of myself as a scientist, who happens la

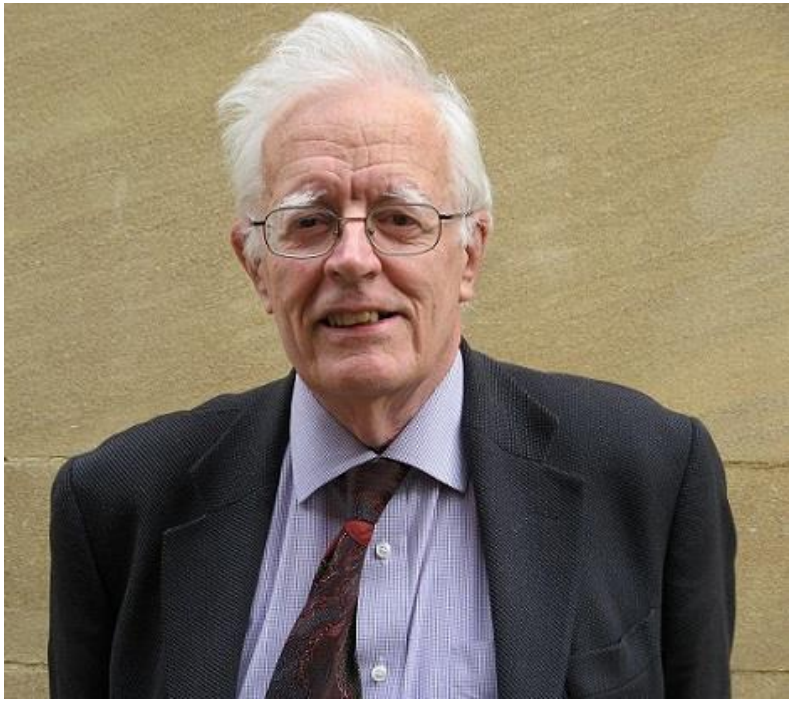
"I would like to think of myself as a scientist, who happens largely to specialise in the use of statistics"– An interview with Sir David Cox

FEATURES

Author: Statistics Views
Date: 24 Jan 2014
Copyright: Image appears courtesy of Sir David Cox

Sir David Cox is arguably one of the world's leading living statisticians. He has made pioneering and important contributions to numerous areas of statistics and applied probability over the years, of which perhaps the best known is the proportional hazards model, which is widely used in the analysis of survival data. The Cox point process was named after him.

Sir David studied mathematics at St John's College, Cambridge and obtained his PhD from the University of Leeds in 1949. He was employed from 1944 to 1946 at the Royal Aircraft Establishment, from 1946 to 1950 at the Wool Industries Research Association in Leeds, and from 1950 to 1955 worked at the Statistical Laboratory at the University of Cambridge. From 1956 to 1966 he was Reader and then Professor of Statistics at



If you had not got involved in the field of statistics, what do you think you would have done? (Is there another field that you could have seen yourself making an impact on?)

I thought I would go into either theoretical physics or pure mathematics but I'm very glad I didn't. I'm not clever enough for either of those fields. They are both fascinating subjects but **statistics is a much more easily satisfying life because there are so many different directions in which to go**.

My career path

- Oberlin College: BA Math 1988
- HSPH: MSc & ScD Biostatistics 1988-1992

- University of Edinburgh (Maths & Stats)
  - Lecturer in Statistics 1992-1995
- University of Oxford (Zoology)
  - Head of Statistics Unit 1995-2000
- Imperial College London (Infectious Disease Epidemiology)
  - Reader then Professor 2000-
- University of Oxford (Stats)
  - Professor of Applied Statistics 2018-
- University of Pretoria (Faculty of Veterinary Science)
  - Extraordinary Professor 2020-

How do I use statistics?

Science/Policy interface

Disease control

Epidemiology

Conservation

Animal welfare

Teaching

1996

BSE/vCJD
in the UK

2001

FMD
in the UK

2003

SARS
in Hong Kong

2009

Pandemic
Influenza

2013/4

MERS
in Saudi
Arabia

2014/6

Ebola
in West
Africa

2016

Zika

2018,
2018/9
2020

Ebola in
DRC

Estimates of the severity of coronavirus disease 2019: a model-based analysis

Robert Verity, Lucy C Okell*, Ilaria Dorigatti*, Peter Winskill*, Charles Whittaker*, Natsuko Imai, Gina Cuomo-Dannenburg, Hayley Thompson, Patrick G T Walker, Han Fu, Amy Dighe, Jamie T Griffin, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Anne Cori, Zulma Cucunubá, Rich FitzJohn, Katy Gaythorpe, Will Green, Arran Hamlet, Wes Hinsley, Daniel Laydon, Gemma Nedjati-Gilani, Steven Riley, Sabine van Elsland, Erik Volz, Haowei Wang, Yuanrong Wang, Xiaoyue Xi, Christl A Donnelly, Azra C Ghani, Neil M Ferguson**

Summary

Background In the face of rapidly changing data, a range of case fatality ratio estimates for coronavirus disease 2019 (COVID-19) have been produced that differ substantially in magnitude. We aimed to provide robust estimates, accounting for censoring and ascertainment biases.

Methods We collected individual-case data for patients who died from COVID-19 in Hubei, mainland China (reported by national and provincial health commissions to Feb 8, 2020), and for cases outside of mainland China (from government or ministry of health websites and media reports for 37 countries, as well as Hong Kong and Macau, until Feb 25, 2020). These individual-case data were used to estimate the time between onset of symptoms and outcome (death or discharge from hospital). We next obtained age-stratified estimates of the case fatality ratio by relating the aggregate distribution of cases to the observed cumulative deaths in China, assuming a constant attack rate by age and adjusting for demography and age-based and location-based under-ascertainment. We also

Lancet Infect Dis 2020; 20: 669–77

Published Online March 30, 2020 https://doi.org/10.1016/S1473-3099(20)30243-7

This online publication has been corrected. The first corrected version first appeared at thelancet.com/infection on April 15, 2020 and the second on May 4, 2020

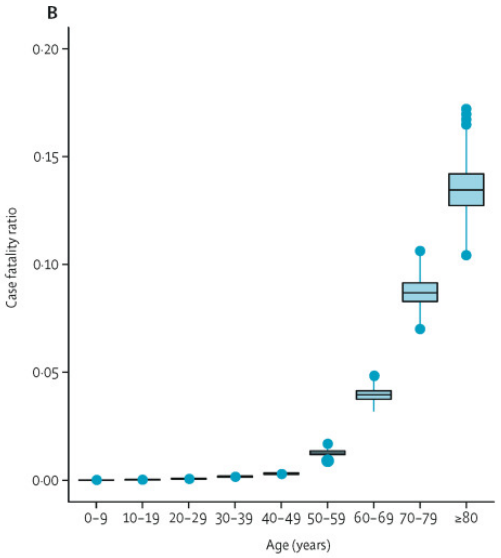
2 Statistical Methods

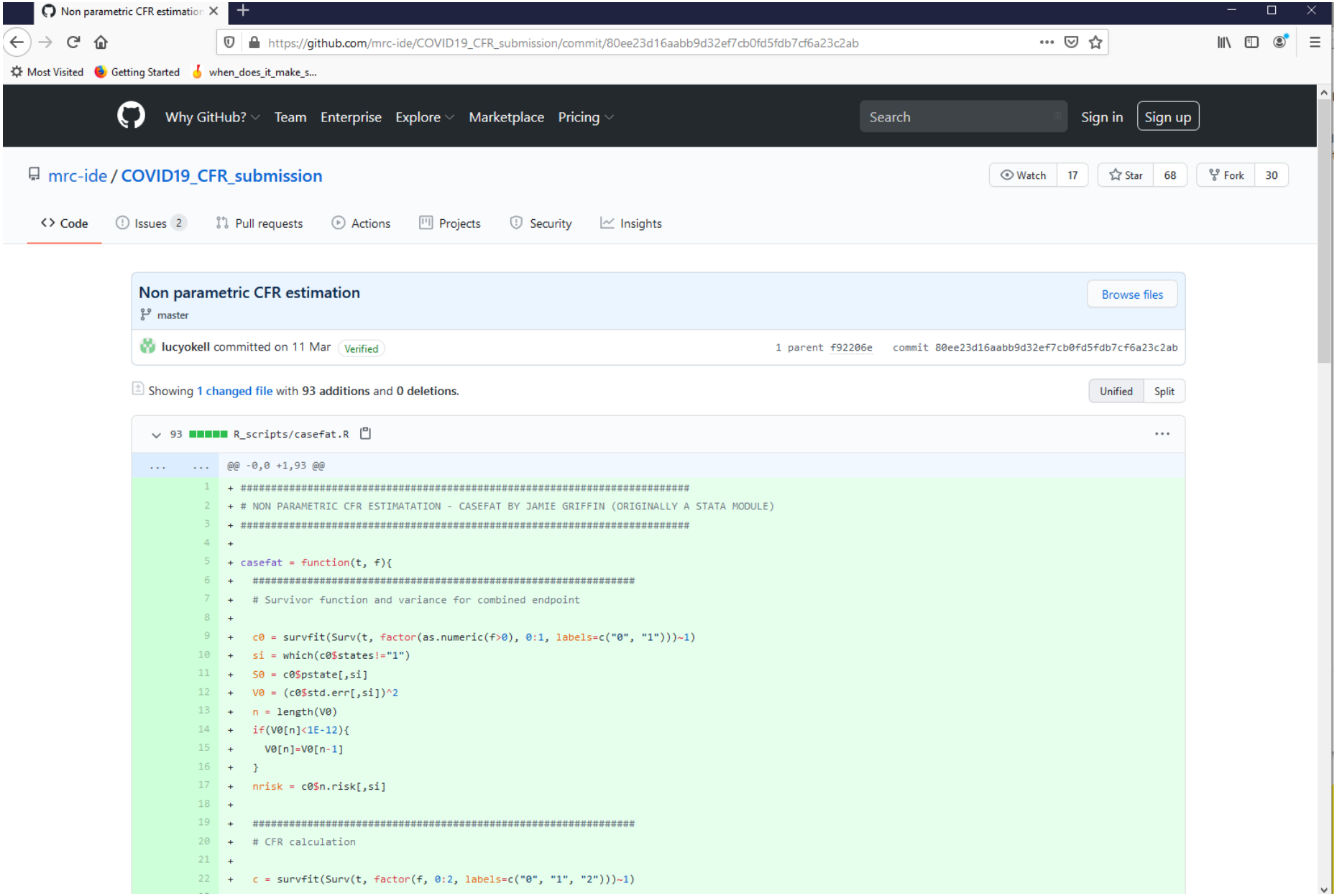
2.1 Intervals between onset of symptoms and death

Let t_o and t_d be the time (in days) of onset of symptoms and death, respectively, and let $\delta_{od} = t_d - t_o$ be the onset-to-death interval. If $f_{od}(\cdot)$ denotes the probability density function (PDF) of time from symptom onset to death, then the probability that a death on day t_d had onset of symptoms on day t_o is

$$g_{od}(t_o \mid t_d) = \frac{\int_{\delta_{od}}^{\delta_{od}+1} f_{od}(\tau) o(t_d - \tau)\, d\tau}{\int_0^{\infty} f_{od}(\tau') o(t_d - \tau') d\tau'},$$

where $o(t)$ denotes the observed number of onsets that occurred at time t. For an exponentially growing epidemic, we assume that $o(t) = o_0 e^{rt}$ where o_0 is the initial number of onsets (at $t = 0$) and r is the epidemic growth rate. Substituting this, we obtain




Non parametric CFR estimation
https://github.com/mrc-ide/COVID19_CFR_submission/commit/80ee23d16aabb9d32ef7cb0fd5fdb7cf6a23c2ab
Most Visited
Getting Started
when_does_it_make_s...
Why GitHub?
Team
Enterprise
Explore
Marketplace
Pricing
Search
Sign in
Sign up
mrc-ide / COVID19_CFR_submission
Watch 17
Star 68
Fork 30
Code
Issues 2
Pull requests
Actions
Projects
Security
Insights
Non parametric CFR estimation
master
Browse files
lucyokell committed on 11 Mar
Verified
1 parent f92206e
commit 80ee23d16aabb9d32ef7cb0fd5fdb7cf6a23c2ab
Showing 1 changed file with 93 additions and 0 deletions.
Unified
Split
93 R_scripts/casefat.R
@@ -0,0 +1,93 @@
+ ##
+ # NON PARAMETRIC CFR ESTIMATATION - CASEFAT BY JAMIE GRIFFIN (ORIGINALLY A STATA MODULE)
+ ##
+
+ casefat = function(t, f){
+   ##
+   # Survivor function and variance for combined endpoint
+
+   c0 = survfit(Surv(t, factor(as.numeric(f>0), 0:1, labels=c("0", "1")))~1)
+   si = which(c0$states!="1")
+   S0 = c0$pstate[,si]
+   V0 = (c0$std.err[,si])^2
+   n = length(V0)
+   if(V0[n]<1E-12){
+   V0[n]=V0[n-1]
+   }
+   nrisk = c0$n.risk[,si]
+
+   ##
+   # CFR calculation
+
+   c = survfit(Surv(t, factor(f, 0:2, labels=c("0", "1", "2")))~1)

What next?

- Statistical analysis …

  will open doors you didn't even know were doors.