Statistical methods with application to demography and life insurance

(Comments on the book)

Estate Khmaladze

Victoria University of Wellington

26 June 2013

Overview

1. Duration of life as a random variable
   * A note on Stieltjes integral
2. Models of distribution function $F(x)$ and force of mortality $\mu(x)$
3. The empirical distribution function of duration of life
4. Derivation of $\hat{F}_n(x)$ from $F(x)$ as a random process
5. Limit of empirical process: Brownian bridge.
6. Statistical consequences of what we have learned so far

Overview

7. Testing parametric hypotheses. Unexpected example – durations of rule of Roman emperors
   - 7.1 Duration of rule of Roman emperors
   - 7.2 Testing exponentiality
   - * Distribution of Kolmogorov – Smirnov and ω^2 statistics from Brownian motion
8. Estimation of the rate of mortality
9. Censored Observations. Related point processes
10. Kaplan-Meier Estimator for F
    - * A note on Wiener stochastic integral
11. Statistical Inference about F based on Kaplan – Meier estimator
    - 11.1 Testing simple and parametric hypotheses
    - 11.2 Two sample problem for censored observations
    - * A note on Ito integral

Overview

12. Life Insurance and net premiums
    * Remark on mixtures of distributions
13. More on net premiums. Endowments and annuities
14. Annuities certain. Some problems of general theory
15. Right tail behavior of $\hat{F}_n$
    * Asymptotic form of excess life distribution
    * Recurrence formulae for calculation of non-crossing probabilities
16. Population dynamics
    * Age structure of a population

1. Duration of life as a random variable

T - duration of life, x - age, $T - x$ - remaining life; their distributions, moments, etc.

Duration of life – random?

2. Models of $F(x)$ and force of mortality $\mu(x)$

Fitting something to something – is it good?

Much better work starts not with fitting, but with a model: fitting the model is then verification of the assumption.

Classifications according to the behavior of the force of mortality/ failure rate.
E.g., lognormal distribution has decreasing failure rate:

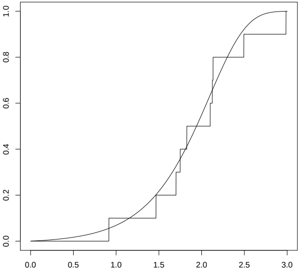
$$\mu(x) \searrow \text{ in } x$$

3. The empirical distribution function of duration of life

Empirical distribution function is a most important object in statistics; it is based solely on observations:

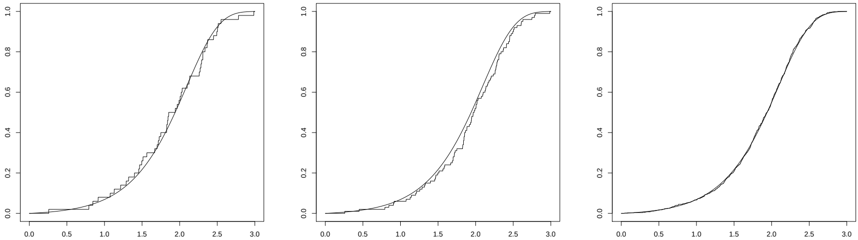
$$\widehat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{\{T_i \le x\}}$$

But does F exist?



$n = 10$ observations from Gompertz distribution

4. Derivation of $\hat{F}_n(x)$ from $F(x)$ as a random process

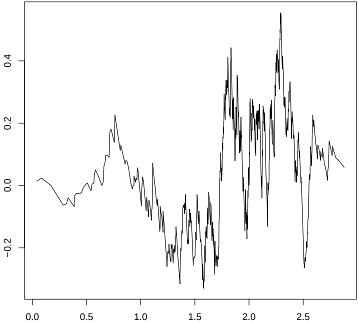


Glivenko – Cantelli Theorem

$$\sup_x |\widehat{F}_n(x) - F(x)| \to 0, \text{ as } n \to \infty$$

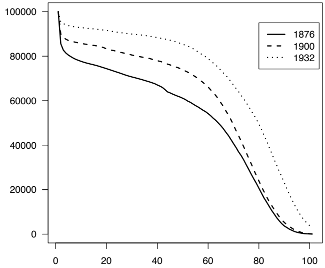
5. Limit of empirical process: Brownian Bridge

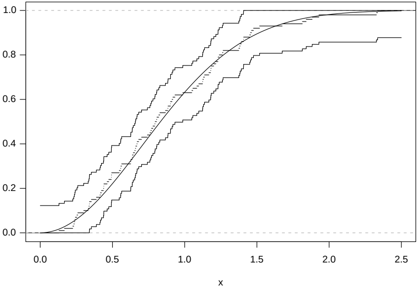
$$\sqrt{n}[\widehat{F}_n(x) - F(x)]$$



Limit distribution of chi-square statistic $\sum_{j=1}^{m} \frac{(\nu_{jn} - np_j)^2}{np_j}$ explained

6. Statistical consequences of what we have learned so far





7. Testing parametric hypotheses. Unexpected example – Roman Emperors

When is it that a model distribution does not depend on a parameter? – never.

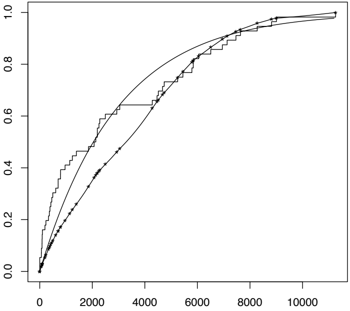
So, there are parameter(s) (say, θ), which need to be estimated (say, by $\hat{\theta}_n$) from the data.

However, behavior of "estimated" empirical process

$$\sqrt{n}[\widehat{F}_n(x) - F_{\hat{\theta}_n}(x)]$$

is very different from that of empirical process with θ known.

. . . *cont.* Unexpected example – Roman Emperors



Empirical distribution function and approxiamting exponential distribution function (+ the compensator)

8. Estimation of the rate of mortality

$$\mu(x) = \frac{f(x)}{1 - F(x)}$$

Smoothing methods

$$\hat{f}_n(x) = \int K_\Delta(x, y) d\widehat{F}_n(y)$$

9. Censored Observations. Related point processes

The observations T we would like to make are screened by "nuisance" censoring variables Y, so that we can only observe $\min(T, Y)$, and we also know whether it was a T or a Y.

10. Kaplan-Meier Estimator

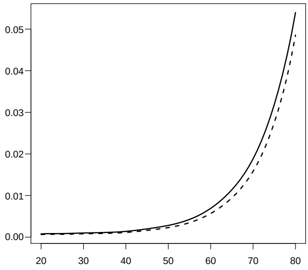
$$\bar{F}_n(x) = 1 - \Pi_{s \leq x} \left[1 - \frac{dN(s)}{Y(s-)} \right]$$

11. Statistical Inference about F, based on Kaplan-Meier estimator

Functional CLT (central limit theorem) for martingales.

12. Life Insurance and net premiums. Remark on mixtures

whole life insurance; whole life insurance with time-limited premium; term life insurance; insurance with variable sum insured and variable premium; group life insurance.



Australian data: population force of mortality and bruto premium per year (dotted line).

13. More on net premiums. Endowments and annuities

Overpayment in the prices of the final moment:
actuarial value of the insurance policy at the moment $t = 0$

$$\pi_0 = cEe^{-\rho\tau};$$

at the (random) moment of payment τ:

$$\pi_0 e^{\rho\tau}$$

and, consequently, its expected value will be

$$\pi_0 Ee^{\rho\tau} = cEe^{-\rho\tau} Ee^{\rho\tau},$$

which is more than c!

pure endowment; life annuity; deferred annuity; term annuity; term deferred annuity; endowment assurance contracts.

14. Annuities certain. Some problems of general theory

If amount c is payable in n contracts at random times $\{\tau_i\}_{i=1}^n$, then

$$c\widehat{F}_n(t) - \int_0^t p(s)(1 - \widehat{F}_n(s))\,ds$$

is nominal accumulated loss. Discounted to the time of inception, its value is

$$c\int_0^t e^{-\rho s}\,d\widehat{F}_n(s) - \int_0^t e^{-\rho s} p(s)(1 - \widehat{F}_n(s))\,ds$$

and at price level of the current time t it becomes

$$c\int_0^t e^{\rho(t-s)}\left[d\widehat{F}_n(s) - \frac{p(s)}{c}(1 - \widehat{F}_n(s))ds\right].$$

15. Right tail behavior of $\hat{F}_n$

When studying long lives (or large claims, etc.), we need to consider behavior of $\widehat{F}_n(z)$ for large z. This behavior, or behavior of the tail

$$1 - \widehat{F}_n(z),$$

is very different from the behavior of the "central" part:

$$n(1 - \widehat{F}_n(z))$$

forms not Gaussian but Poisson process, driven by the tail distribution function of remaining life $T - x$,

$$P(T - x > z | T > x)$$

...*cont.* Right tail behavior of $\hat{F}_n$

$$P(T - x > z | T > x) \to \begin{cases} (1 + \theta z/\sigma)^{-1/\theta}, & \text{if } \theta \neq 0 \\ e^{-z/\sigma}, & \text{if } \theta = 0 \end{cases}$$

16. Population dynamics

A system of populations, first without – then with migrations is studied through the system of stochastic differential equations

$$d\mathbf{P}(t) = \mathbf{C}\mathbf{P}(t)dt + d\mathbf{M}(t),$$

and the system of corresponding martingales ("noise" part) $\mathbf{M}$ is described.

$$\mathbf{P}(t) = \int_0^t e^{\mathbf{C}(t-s)} d\mathbf{M}(s) + \mathbf{P}(0)e^{\mathbf{C}t}$$

Age structure of a population

$$\frac{P(t,x)}{P(t)}$$

and what is involved in its analysis is presented.

Publication details

16 Lectures.
About 75 exercises.
Over 100 references.
About 170 terms in the Index.