generally, for clarifications of derivations & concepts

Force of Mortality $\mu(t)$ / hazard function — conditional deaths

laws: Gompertz $\mu(t)=B\times C^t$; Makehams $\mu(t)=A+B\times C^t$

$\boxed{S_T(t)=\exp\left(-\int_0^t \mu_T(y)dy\right)}$

$\mu_t = \mu_T(t) = -\frac{S_T'(t)}{S_T(t)} = \frac{F_T'(t)}{S_T(t)} = \frac{f_T(t)}{S_T(t)}$

$\mu_x = -\frac{1}{l_x}\frac{\partial}{\partial x}l_x = -\frac{d}{dx}\log(l_x) = \frac{{}_h q_x}{h}$

survival function $S_T(t)=Pr(T>t)$ ⇌ CDF $F_T(t)=Pr(T\le t)$

↳ expected future lifetime

${}_tP_x = S(t)$

$f_T(t) = -S_T'(t) = F_T'(t)$

integrated hazard

$\Lambda_T(t)=\int_0^t \mu_T(y)dy$

sdf (integrated) $S_T(t)=\exp(-\Lambda_T(t))$

PARAMETRIC

parametric Method:

① MOM. $E(X^j)=\frac{\sum_{i=1}^n x_i^j}{n}$ & $E(X^2)=E(X)^2+Var(X)$
$=\mu^2+\sigma^2$

② MLE/LMLE
(check $l''<0$ for maximizer)

↳ property: $\hat\theta_{MLE}\sim N(\theta, \frac{1}{I(\theta)})$ → variance $(\hat\theta_{MLE})$

unbiased

↳ $I(\theta)=-E\left(\frac{d^2}{d\theta^2}\log L(\theta)\right)$

NON PARAMETRIC
(robust but inefficient)

Empirical dist. func.

$\hat F_T(t)=\frac{d(t)}{N}$

$d(t)\sim Bin$, unbiased

$V(\hat F_T(t))=\frac{F_T(t)(1-F_T(t))}{N}$

CENSORING
- R
- L
- Interval
- Random
- Info vs uninfo.

CI: $100(1-\alpha)\%$ CI for $S_T(t)$

$\hat S_T(t)\pm 1.96\sqrt{\hat S(t)^2\sum_{t_j\le t}\frac{d}{r_j(r_j-d_j)}}$

* with log:

$\left(\exp\left(-\log(\hat S_T(t))-1.96\sqrt{\Sigma(\ )}\right), \min\{1, \exp(\cdots+\cdots)\}\right)$

KM

N: #

m: # deaths

$d_1+\cdots+d_k=m$

r_j, d_j $d_j\sim Bin(r_j, 1-p_j)$ $\hat p_j=\frac{r_j-d_j}{r_j}$

$\hat S_T(t)=\prod_{t_j\le t}\frac{r_j-d_j}{r_j}$

Greenwood's Formula: $Var(\text{KM est})=Var[\hat S_T(t)]\approx \hat S_T(t)^2\sum_{t_j\le t}\frac{d_j}{r_j(r_j-d_j)}$

δ-method (est mean & var of a Transf. of r.v)

Taylor expan. of 2nd order.

$U=g(Y)$. μ & σ^2

$g(Y)\approx g(\mu)+(Y-\mu)g'(\mu)$ or $g(Y)-g(\mu)\approx(Y-\mu)g'(\mu)$

DELTA METHOD

$E(g(Y))\approx g(\mu)$ & $Var[g(Y)]\approx g'(\mu)^2\sigma^2$

NA

based on: $\hat\Lambda_T(t)=\sum_{t_j\le t}\frac{d_j}{r_j}\Rightarrow \hat S_T(t)=\exp(-\hat\Lambda_T(t))$

Tie handling { NA: together; FH: separately

$Var(\hat S(t))=(-\hat S(t))^2\cdot\sum_{t_j\le t}\frac{d_j(r_j-d_j)}{r_j^3}$

(NA)

(FH)

KM vs NA

KM	t_j	r_j	d_j	$\frac{r_j-d_j}{r_j}$	$\prod_{t_j\le t}\frac{r_j-d_j}{r_j}$
NA	t_j	r_j	d_j	$\frac{d_j}{r_j}$	$\exp\left(-\sum_{t_j\le t}\frac{d_j}{r_j}\right)$

OTHER DERIVATIONS:

- $P_x=\frac{l_{x+1}}{l_x}$, $q_x=\frac{d_x}{l_x}$, ${}_tP_x=\frac{l_{x+t}}{l_x}$, ${}_tq_x=\frac{l_x-l_{x+t}}{l_x}$

$\mu_x \longrightarrow {}_nP_x$

${}_nP_x=\exp\left(-\int_0^n \mu_{x+s}ds\right)$

${}_nq_x=\int_0^n {}_sP_x\,\mu_{x+s}\,ds \longrightarrow d_x=\int_0^1 l_{x+s}\mu_{x+s}ds$

${}_nq_x$

EXPECTED FUTURE LIFETIME

① COMPLETE $\mathring{e}_x=E[T_x]=\int_0^\infty t\cdot f(t)dt$
$=\int_0^\infty {}_tP_x\,dt$

② CURTATE $e_x=E[K_x]=\sum_{k=0}^\infty k\cdot P(K_x=k)=\sum_{k=1}^\infty {}_kP_x$

* deferred prob.

${}_{n|m}q_x=P(n<T_x<n+m)=\frac{l_{x+n}-l_{x+n+m}}{l_x}$

life aged x survives n years dies in the subseq. m years.

Non-int ages { (UDD) ${}_tP_x\,\mu_{x+t}$ constant ↳ $T_x\sim Unif(0,1)\Rightarrow {}_sq_x=s\cdot q_x$; ($\mu_{x+t}$ constant)

central rate of mortality (m_x) $m_x=\frac{d_x}{\int_0^1 l_{x+t}dt}=\frac{q_x}{\int_0^1 {}_tP_x\,dt}=\frac{\int_0^1 {}_tP_x\,\mu_{x+t}dt}{\int_0^1 {}_tP_x\,dt}$

Under UDD

$m_x=\frac{d_x}{\int_0^1 l_{x+t}dt}\approx\frac{d_x}{l_x-\frac{1}{2}d_x}\Rightarrow m_x\approx\frac{q_x}{1-\frac{1}{2}q_x}$ & $q_x=\frac{m_x}{1+\frac{1}{2}m_x}$

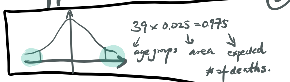
$m_x\ge q_x$

Cox-regression

$\lambda(t;x)=\underbrace{\lambda_0(t)}_{\text{base line}}\exp(\beta^T x)$

$S(t)=S_0^{\exp(\beta^T x)}$, $S_0(t)=\exp(-\int_0^t \lambda_0(s)ds)$

The surival functions for different covariate values cannot cross.



- maximizing the PL.
- $R(t_i)$ the set of i's under study at t_i
- no tied death:

$$L(\beta)=\prod_{i=1}^m \frac{\exp(\beta^T x_{(i)})}{\sum_{j\in R(t_i)}\exp(\beta^T x_j)}$$ m: # of deaths

- with tied death $s_{(i)}=x_{(i),1}+\dots+x_{(i),d}$

$$L(\beta)=\prod_{i=1}^m\frac{\exp(\beta^T s_{(i)})}{[\sum_{j\in R(t_i)}\exp(\beta^T x_j)]^{d_i}}$$ $d_{(i)}$ = # of tied death

- TS for one parameter: $\frac{\hat\beta}{SE(\hat\beta)}$ vs Z
- TS for overall: $-2(LL(H_0)-LL(H_1))\xrightarrow{D}\chi^2_k$

k := # of different parameters.

Stats Tests

① χ^2-test	~ χ^2_m	1	sided
② SD-test	~ $\chi^2_{(\cdot)}$	1	
③ CD-test	~ Z	2	
④ sign-test	~ Z	2	
⑤ runs-test	~ Z	1	

① χ^2-test: H_0: $\underbrace{\text{obs. deaths}}_{\theta_x}\sim Binom(E_x, q_x)$: $TS=\chi^2=\sum_{\text{age groups}}\frac{(\theta_x-E_xq_x)^2}{E_xq_x}$

degree of freedom := number of age groups

rule of thumb: each group ≥5 deaths

+ : can detect a single difference b/w obs. & expected.

− : cannot detect large group of +/− deviations.

② SD test: TS: $Z_x=\frac{\theta_x-E_xq_x}{\sqrt{E_xq_x(1-q_x)}}\sim N(0,1)$ Steps: ⓐ decide a set of intervals (equal length) ⓑ compute the obs. number of Z_x falling in each interval. ⓒ compute the expected number of Z_x falling in each interval. ⓓ do χ^2 test compare obs. & expected. degree of freedom := # of intervals minus 1

+: can detect a few large deviations being masked by an excessive number of smaller deviations

③ Sign Test: obs # of positive deviations vs Binomial$(m,\frac{1}{2})$, m = # of age groups.

Approx Binomial → $N(0.5m, 0.5^2m)$ Variance ⇒ both too many/few positive/negative deviations are evidence against H_0. Can detect an excessive number of pos/neg deviations.

④ Cumulative Deviation Test: TS: $\frac{\sum_a^b(\theta_x-E_xq_x)}{\sqrt{\sum_a^b(E_xq_xp_x)}}\sim N(0,1)$

⑤ Runs Tests: n_1: # of pos deviations, n_0: # of neg deviations, m: # of age groups.

prob of seeing t or less positive runs: $\sum_{j=1}^t\frac{\binom{n_1-1}{j-1}\binom{n_0+1}{j}}{\binom{m}{n_1}}$, large m → $t\,\dot\sim\,N\left(\frac{n_1(n_0+1)}{n_0+n_1},\frac{(n_0n_1)^2}{(n_0+n_1)^3}\right)$ a small # of pos runs is evidence against H_0

MARKOV MODEL (2-state)

assumptions ① past ↛ future ② short time dt: ${}_{dt}q_{x+t}\approx\mu_{x+t}dt+o(dt)$ ③ μ_{x+t} is a constant in [0,1]

Setup: $x+a_i$ (start), $x+b_i$ (cease), $x+T_i$ (end), V_i (time under observation) waiting time

$\delta_i=0\Rightarrow V_i=b_i-a_i,\ T_i=b_i$

$\delta_i=1\Rightarrow V_i=T_i-a_i$

MLE of μ: $\hat\mu=\frac{\delta}{\nu}$, δ := total number of deaths, ν := total waiting time

$SE(\hat\mu)=\frac{\sqrt\delta}{\nu}$, asymptotically $\hat\mu\sim N(\mu,\frac{\mu^2}{E(\delta)})$ i.e. $N(\hat\mu,\frac{\delta}{\nu^2})$

KERNEL

idea: closer to x^*, more weight; symmetric.

$\hat f(x^*)=\hat y(x^*)=\sum_{j=1}^n w_jy_j$, $w_j=\frac{K(\frac{|x^*-x_j|}{b})}{\sum_{n=1}^n K(\frac{|x^*-x_j|}{b})}$ $t=\frac{|x^*-x_j|}{b}$, b↓ → t↑

common kernel functions: normal $K(t)=\frac{\exp(-t^2/2)}{\sqrt{2\pi}}$, triangle $K(t)=2-4|t|$, $|t|\le0.5$, box $K(t)=1$, $|t|\le0.5$

Spline

assumptions: connected, 1st derivative & 2nd derivative equal at connections. Knots selection by minimizing $AIC=-2\ln(L)+2p$, $BIC=-2\ln(L)+\log(n)\times p$. OR minimizing $\frac{(A_x-E_x^cm_x)^2}{E_x^cm_x'(1-m_x')}$

A_x: obs. # of deaths aged x

E_x^c: central exposed to risk aged x

m_x: graduated value of the central mortality rate at age x

m_x': preliminary value of m_x

Exposed to Risk

$\hat q_x=\frac{d_x}{E_x}$ (initial), $\hat\mu_x=\frac{d_x}{E_x^c}$ (central). Under constant force of mortality:

central mortality = force of mortality

$m_x=\frac{d_x}{\int_0^1 {}_tl_{x+t}dt}=\frac{q_x}{\int_0^1{}_tp_xdt}=\frac{1-e^{-\mu}}{\int_0^1e^{-\mu t}dt}=\frac{1-e^{-\mu}}{\frac{1}{\mu}[e^{-\mu t}]_0^1}=\mu$

under the assumption all deaths happen halfway: $E_x\approx E_x^c+0.5d_x$

But E_x^c requires lots of info, sometimes just ⇒ $E_x^c=\int P(x,t)dt$ ⇒ Use approx $\frac{\text{sum of sides}}{2}\times\text{height}$

${}_{t+dt}P_x^{\overline{gg}}={}_tP_x^{\overline{gg}}\,{}_{dt}P_{x+t}^{\overline{gg}}={}_tP_x^{\overline{gg}}(1-\sum_{r\ne g}{}_{dt}P_{x+t}^{gr})={}_tP_x^{\overline{gg}}(1-\sum_{r\ne g}\mu_{x+t}^{gr}dt)\Rightarrow\frac{\partial}{\partial t}{}_tP_x^{\overline{gg}}=\lim_{dt\to0}\frac{{}_tP_x^{\overline{gg}}(1-\sum\mu_{x+t}^{gr}dt)-{}_tP_x^{\overline{gg}}}{dt}=-{}_tP_x^{\overline{gg}}\sum_{g\ne r}\mu_{x+t}^{gr}$

$\frac{\frac{\partial}{\partial t}{}_tP_x^{\overline{gg}}}{{}_tP_x^{\overline{gg}}}=-\sum_{r\ne g}\mu_{x+t}^{gr}\Rightarrow\frac{\partial}{\partial t}\ln({}_tP_x^{\overline{gg}})=-\sum_{r\ne g}\mu^{gr}_{x+t}\Rightarrow{}_tP_x^{\overline{gg}}=\exp(-\int_0^t\sum_{r\ne g}\mu_{x+s}^{gr}ds)$

MULTISTATE MARKOV

Kolmogorov Forward Equations:

$$\frac{d}{dt}{}_tP_x^{gh}=\sum_{r\ne h}({}_tP_x^{gr}\mu(x+t)^{rh}-{}_tP_x^{gh}\mu(x+t)^{hr})$$

$$\boxed{\frac{d}{dt}{}_tP_x^{\overline{gg}}=-{}_tP_x^{\overline{gg}}\sum_{h\ne r}\mu(x+t)^{gr}}$$

$$\Rightarrow \boxed{{}_tP_x^{\overline{gg}}=\exp(-t\sum_{r\ne g}\mu^{gr})}$$

⇒ MLE $\begin{cases}\hat\mu^{gh}=\frac{\delta^{gh}}{\nu^g}\\ \hat\mu^{gr}\sim N(\hat\mu^{gh},\frac{\delta^{gh}}{(\nu^g)^2})\end{cases}$

$$\Rightarrow \hat p^{\overline{gg}}=\exp(-t\sum_{r\ne g}\hat\mu^{\overline{gr}})$$

⇒ prob not $\overline{gg}$: $1-\exp(-t\sum_{r\ne g}\hat\mu^{\overline{gr}})$

$$\Rightarrow {}_tP_x^{gh}=\frac{\hat\mu^{gh}}{\hat\mu^{gh}+\hat\mu^{gr}}[1-\exp(-t\sum_{r\ne g}\hat\mu^{\overline{gr}})]$$

$$\Rightarrow Var(e^{-t\mu})=[\frac{d}{d\mu}e^{-t\mu}]^2Var(\mu)$$

delta method $=t^2e^{-2t\mu}\cdot Var(\mu)$

SE → CI.

Binomial Markov (est q_x)

$$L=\prod_{i=1}^N {}_{b_i-a_i}q_{x+a_i}^{\delta_i}(1-{}_{b_i-a_i}q_{x+a_i})^{1-\delta_i}$$

assumptions
① UDD ${}_tq_x=tq_x$
② constant hazard ${}_tq_x=1-\exp(-\mu t)$
③ behaviour between (0,1) $\mu_x(t)=\frac{q_x}{1-tq_x}$

e.g. $a_i=b_i=0.5$
ω := total censored
δ := total death

$L=q_x^{\delta}(1-q_x)^{N-\omega-\delta}(1-q_x)^{0.5\omega}$

MLE: $\hat q_x=\frac{\delta}{N-0.5\omega}$