Q1

(a) Let

t_j = time j minutes

n_j = no people in queue after j minutes

d_j = no collecting food after j minutes

c_j = no leaving queue with no food at j mins

λ_j = K-M hazard function = d_j/n_j

Survival function = $\hat{S}_j = \prod_{t \le t_j} (1 - \lambda_j)$

t_j	n_j	d_j	c_j	λ_j	S_j
0	30				
1	30	3	0	0.1	0.9
2	27	6	1	0.222	0.7
3	20	5	0	0.25	0.525
4	15	3	1	0.2	0.42
5	11	2	0	0.182	0.344
6	9	3	2	0.333	0.229
7	4	2	0	0.5	0.115
8	2	2	0	1	0
9	0				

(b) first $S_j < 10\%$ is $j = \underline{\underline{8}}$ minutes

(b) - small number
- only one part of morning
- what happens to censored lives? do they resume queuing later?
- non-independence of observations eg family members queuing together

Q2

(a) baseline applies when all $z_i = 0$

$\therefore$ people in private home in Paris who have just seen their doctor

(b) $z_1 = 1 \quad z_2 = 4 \quad z_3 = 1$

(i) $h_i(t) = h_0(t) \exp(-0.1 + 0.2 + 0.4)$

$= h_0(t) \cancel{\exp}\, 0.5$

(ii) $S(t) = \exp - \int_0^t h_0(t) e^{0.5}\, dt$

(c) for Paris based life

$$S_p(26) = 0.985 = \exp - \int_0^{26} h_0(t) e^{0.1+0.4} \, dt$$

$$= \exp \int_0^{26} h_0(t) e^{0.5} \, dt$$

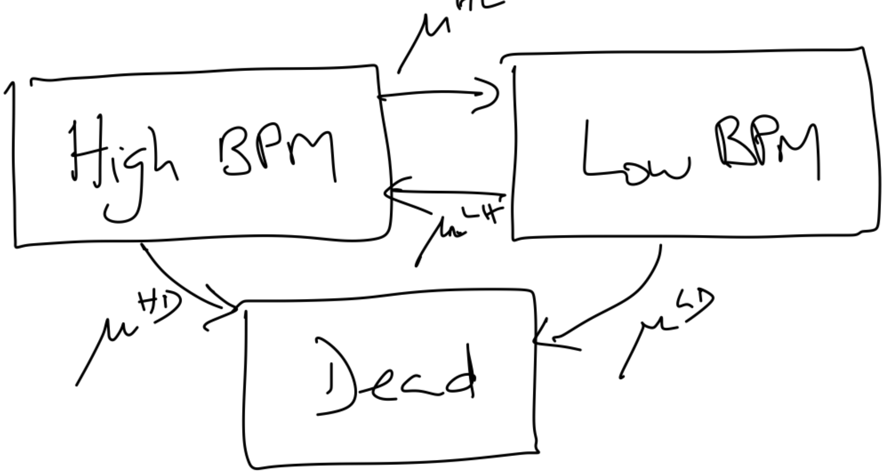
$$\therefore - \int_0^{26} h_0(t) \, dt = \frac{\ln 0.985}{e^{0.5}}$$

We seek $S_G(26)$

$$= \exp - \int_0^{26} h_0(t) \, e^{0.5} \, dt$$

$$= \exp\left(\frac{\ln 0.985 \, e^{0.5}}{e^{0.5}} \right) = \underline{\underline{0.985}}$$

Q3 (a)



where

μ^{HL} is transition intensity from High to Low

μ^{LH} ——— " ——— Low to High

μ^{HD} ——— " ——— High to Dead

μ^{LD} ——— " ——— Low to Dead

(b) If W^H is waiting time in High

W^L is waiting time in Low

and d_{AB} is no. transitions from state A to state B

for d_{HL} d_{LH} d_{HD} d_{LD}

Then the likelihood function is given by

$$L(\mu) = \exp\left(-(\mu^{HL} + \mu^{HD}) W^H\right)$$

$$\times \exp\left(-(\mu^{LH} + \mu^{LD}) W^L\right)$$

$$\times {\mu^{HL}}^{d_{HL}} {\mu^{HD}}^{d_{HD}} {\mu^{LH}}^{d_{LH}} {\mu^{LD}}^{d_{LD}}$$

+ constant.

(c) we seek MLE for μ_{HD}

$$\log L = -\mu^{HL} W^{H} - \mu^{HD} W^{H} - \mu^{2H} W^{L}$$

$$- \mu^{LD} W_L + d_{HL} \log \mu^{HL} + d_{LH} \log \mu^{LH}$$

$$+ d_{HD} \log \mu^{HD} + d_{LD} \log \mu^{LD}$$

differentiating w.r.t. μ^{HD} and setting to zero gives

$$-W_H + \frac{d_{HD}}{\hat{\mu}_{HD}} = 0$$

$$\therefore \hat{\mu}_{HD} = \frac{d_{HD}}{W_H} = \frac{18-8}{1738}$$

$$= \underline{\underline{0.005754}}$$

(d) no transitions H → L and L → H greater than no deaths

Possible heart rate over longer than one hour important

How accurate is device?

Q4

(a) — not so good when don't have a lot of data to graduate

— heterogeneity in the data set will flow through to the graduation

— any extrapolations at highest ages can have large inaccuracies

(b) — add more data from other experiences in this age group

— consider graduating with reference to a standard table

— increase no. parameters from Makeham

(c) We use chi-squared test for overall goodness-of-fit

Null hypothesis H_0: the graduation produces mortality rates which do reflect the underlying mortality experience

$$\sum_{\text{ages } 85-102} z_x^2 = 7.48$$

We have 18 ages of data

Makeham formula $\mu_x = A + Bc^x$ has 3 parameters so we reduce degrees of freedom in χ^2 test by 3 to 15

$$\chi^2_{0.95;15} = 24.996$$

$$\sum z_x^2 = 7.48 < 24.996$$

$\therefore$ do not reject H_0

The graduated values do give good overall fit for experience

(d) chi-sq test will not detect :-

- outliers
- small overall +ve or −ve bias
- run of z_x values of one sign
- correlation between groups of z_x values.

(e) At very high ages need to be aware of :-
- problems of small data sets
- poor results from extrapolating mortality rates
- increased risk of "time selection"

Q5

(a) Initial selection — underwriting that takes account of smoking data

Adverse selection - possibility of policyholders selecting against the company

Spurious selection

(b) ① select distribution from the exponential family
② build linear predictor - a function of the covariates
③ find a link function for distribution and linear predictor

(c) Distribution given as Poisson

Covariates are

age x

no. policies — variable α

smoker status — factor i

sales channel — factor j

With 2 factors we need an interaction term γ_{ij}

Linear predictor in parameters α, β, γ becomes

$$\alpha_j + \beta_i x + \beta_j x + \gamma_{ij}$$

Canonical link function for Poisson

is $g(\mu) = \log \mu$

[other coherently argued linear predictors are acceptable here]

(d) Compare scaled deviance of models with and without smoker factor in linear predictor.

In Poisson models – Scaled Deviance is χ^2