OUTLINE OF SOLUTION TO EXAM QUESTIONS IN STK 4080/9080

Problem 1

a) Two examples of situations that may be described by the multiplicative intensity model are

- Censored survival data
- Observations from Markov chain models

For details on these and other examples, see the ABG-book, sections 1.4.2, 1.4.3, 1.4.5 and the examples of section 3.1.2 (only a few details are needed for the exam solution)

b) See the argument in section 3.1.5 in the ABG-book [from the start of the section until formula (3.22)]

c) See the argument on the middle of page 88 in the ABG-book

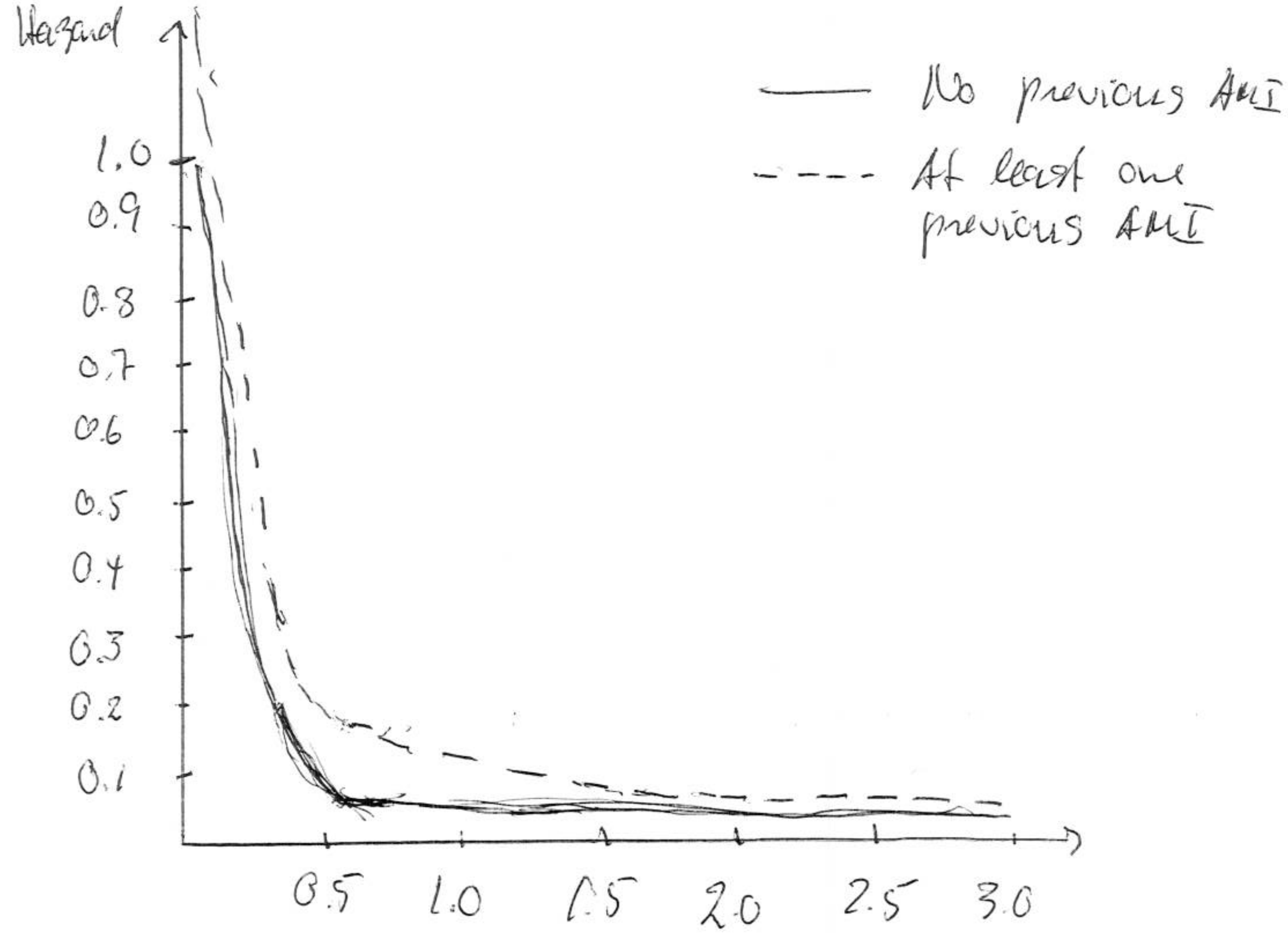
d) See the argument in the last paragraph of page 88 and the top of page 89 in the ABG-book

c) $\hat{A}(t) - A^{*}(t)$ is a martingale.
By the martingale central limit theorem we obtain that $\hat{A}(t)$ is approximately normally distributed around $A(t)$ [since $A^{*}(t)$ is asymptotically equal to $A(t)$]
Details for the martingale central limit theorem (page 89) is nice, but not needed. By standard results we have that

$$\hat{A}(t) \pm 1.96\, \hat{\sigma}(t)$$

is an approximate 95% confidence interval for $A(t)$, where $\hat{\sigma}^2(t)$ is the variance estimator obtained in d)

e) We may read off the slopes of the Nelson-Aalen estimates to obtain estimates of the hazard rates themselves.
Then we get the rough plots



We see that the hazard is very high just after the AMI, and then drops quickly. The hazard is higher for those with at least one previous AMI than for those with no previous AMI.

Problem 2

a) See the argument in section 6.2.1 in the ABG-book [until formula (6.3)] and the argument in section 6.2.2 [until formula (6.7)]. Note that in this problem $\alpha(t) = \alpha$, so that $A(t) = \alpha t$.

b) See the argument on page 237, formula (6.8)

c) When there is a strong frailty, i.e. σ is large, $\mu(t) = \alpha / (1 + \sigma \alpha t)$ will decrease quickly, just as seen in problem 1e. So frailty may help to explain the results seen there: the patients with a severe AMI (corresponding to high frailty) die quickly, and we are then left with the patients with low frailty who have a much smaller hazard.

Problem 3

a) See page 220 in the ABG-book. (until four lines from below)

b&c) See the last three lines on page 220 and the first four lines below the figure on page 222.

d) The argument is similar to the one on page 224: The likelihood is proportional to

$$\prod_{g=1}^{G} \prod_{k=1}^{K} \left\{ \frac{(\theta_k e^{\beta_g} R_{gk})^{O_{gk}}}{O_{gk}!} e^{-\theta_k e^{\beta_g} R_{gk}} \right\},$$

which is the same likelihood that we would get if we considered the O_{gk} to be independent and Poisson distributed with means $\mu_{gk} = \theta_k e^{\beta_g} R_{gk}$.

e) We have that $\mu_{gk} = e^{\psi_k + \beta_g + \log R_{gk}}$, where $\psi_k = \log \theta_k$. Thus

$$\log \mu_{gk} = \psi_k + \beta_g + \log R_{gk}$$

is linear in the parameters

Thus we may fit the model by Poisson regression using log-link and log R_{jk} as offsets.

f) The occurrence/exposure rates become (per 10 000 person years)

Age group	Occ/exp rates $\cdot 10^4$
40-54	7.11
55-59	26.99
60-64	31.93
65-69	40.94
70-74	45.11

f) By exponentiating the estimates from the Poisson regression, we obtain estimates of the base line hazard rates $\Theta_k = e^{\varphi_k}$ and the hazard rate ratios. For the latter we also obtain the 95% confidence intervals as $\exp(\hat{\beta}_j \pm 1.96 \, se_j)$.

The results are given below

Parameter	Estimate*	95% CI
φ_1 (40-54 years)	8.83	
φ_2 (55-59 years)	26.43	
φ_3 (60-64 years)	46.14	
φ_4 (65-69 years)	51.26	
φ_5 (70-74 years)	56.10	
e^{β_2} (Horsens)	0.83	0.57 - 1.20
e^{β_3} (Kolding)	0.62	0.41 - 0.94
e^{β_4} (Vejle)	0.78	0.52 - 1.16

*) Estimates per 10 000 for baseline rates

We see that the baseline hazard rates are increasing with age, and that they are a bit higher than the occurrence/exposure rates in question e7. This is because the baseline rates are for Fredericia, which tend to have higher rates than the other cities.

The hazard ratios for the other cities (compared to Fredericia) are all below 1, indicating lower lung cancer rates. But only for Kolding is the difference significant (with a confidence interval that does not contain 1).