ACSC/STAT 4703, Actuarial Models II
Fall 2016
Toby Kenney
Homework Sheet 3
Model Solutions

Basic Questions

1. *An insurance company collects the following claim data (in thousands):*

i	d_i	x_i	u_i	i	d_i	x_i	u_i	i	d_i	x_i	u_i
1	0	0.6	-	8	0.5	5.6	-	15	2.0	2.5	-
2	0	1.3	-	9	1.0	2.8	-	16	2.0	3.9	-
3	0	2.7	-	10	1.0	4.6	-	17	2.0	6.6	-
4	0	-	10	11	1.0	7.7	-	18	2.0	10.4	-
5	0	-	10	12	1.0	11.3	-	19	2.0	-	15
6	0.5	0.9	-	13	1.0	-	10	20	5.0	7.3	-
7	0.5	1.4	-	14	1.5	3.9	-	21	5.0	8.4	-

Using a Kaplan-Meier product-limit estimator:

(a) estimate the probability that a random loss exceeds 10.7.

We compute the Kaplan-Meier estimator in the following table:

x_i	r_i	s_i	$S(x_i)$
0.6	8	1	$\frac{7}{8}=0.875$
0.9	7	1	$\frac{3}{4}=0.75$
1.3	11	1	$\frac{30}{44}=0.6828$
1.4	10	1	$\frac{27}{44}=0.6136$
2.5	15	1	$\frac{63}{110}=0.5727$
2.7	14	1	$\frac{117}{220}=0.5318$
2.8	13	1	$\frac{54}{110}=0.4909$
3.9	12	2	$\frac{9}{22}=0.4090$
4.6	10	1	$\frac{81}{220}=0.3682$
5.6	11	1	$\frac{81}{242}=0.3347$
6.6	10	1	$\frac{729}{2420}=0.3012$
7.3	9	1	$\frac{162}{605}=0.2678$
7.7	8	1	$\frac{567}{2420}=0.2343$
8.4	7	1	$\frac{243}{1210}=0.2008$
10.4	3	1	$\frac{81}{605}=0.1339$
11.3	2	1	$\frac{81}{1210}=0.0669$

So the survival probability is $\frac{81}{605}=0.1339$.

(b) estimate the median of the distribution.

The median is the first value where the survival probability becomes less than 0.5, which is $x_i=2.8$.

(c) Use a Nelson-Åalen estimator to estimate the median of the distribution.

$-\log(0.5) = 0.6931472$, so the median is the solution to $H(x) = 0.6931472$. We compute the following:

x_i	r_i	$r_i - s_i$	$H(x_i)$
0.6	8	1	0.1250000
0.9	7	1	0.2678571
1.3	11	1	0.3587662
1.4	10	1	0.4587662
2.5	15	1	0.5254329
2.7	14	1	0.5968615
2.8	13	1	0.6737845
3.9	**12**	**2**	**0.8404512**
4.6	10	1	0.9404512
5.6	11	1	1.0313603
6.6	10	1	1.1313603
7.3	9	1	1.2424714
7.7	8	1	1.3674714
8.4	7	1	1.5103286
10.4	3	1	1.8436619
11.3	2	1	2.3436619

so the Nelson-Åalen estimate for the median is 3.9.

2. *An insurance company observes the following claim history:*

Number of claims	*Frequency*
0	*2089*
1	*1810*
2	*799*
3	*226*
4	*60*
5	*14*
6	*2*

Use a Nelson-Åalen estimate to obtain a 95% confidence interval for the probability that a random individual makes more than 4 claims.

The Nelson-Åalen estimate gives $H(4) = \frac{2089}{5000} + \frac{1810}{2911} + \frac{799}{1101} + \frac{226}{302} + \frac{60}{76} = 3.303101$. This gives an estimated survival function as $S(4) = e^{-3.303101} = 0.03676895$. The variance of $\hat{H}(4)$ is $\frac{2089}{5000^2} + \frac{1810}{2911^2} + \frac{799}{1101^2} + \frac{226}{302^2} + \frac{60}{76^2} = 0.01382206$, so a 95% confidence interval for $H(4)$ is $3.303101 \pm 1.96 \times \sqrt{0.01382206} = [3.072670, 3.533533]$, so a 95% confidence interval for $S(4)$ is $\left[e^{-3.533533}, e^{-3.072670}\right] = [0.02920156, 0.04629739]$.

3. *For the data in Question 1, use Greenwood's approximation to obtain a 95% confidence interval for the probability that a random loss exceeds 10.7, based on the Kaplan-Meier estimator.*

(a) Using a normal approximation

Recall from Question 1, that the estimated survival function is $S(10.7) = \frac{81}{605}$.

Using Greenwood's approximation for the product, the variance of the Kaplan-Meier estimator is

$$\left(\frac{81}{605}\right)^2 \left(\frac{1}{8\times 7} + \frac{1}{7\times 6} + \frac{1}{11\times 10} + \frac{1}{10\times 9} + \frac{1}{15\times 14} + \frac{1}{14\times 13} + \frac{1}{13\times 12} + \frac{2}{12\times 10} + \frac{1}{10\times 9} + \frac{1}{11\times 10} + \frac{1}{10\times 9} + \frac{1}{9\times 8} + \frac{1}{8\times 7} + \frac{1}{7\times 6} + \frac{1}{3\times 2}\right) = 0.006251119$$

This means the standard deviation is $\sqrt{0.006251119} = 0.07906402$ which gives a confidence interval of $\frac{81}{605} \pm 1.96 \times 0.07906402 = [-0.02108118, 0.28884978]$.

(b) Using a log-transformed confidence interval.

Using a log-transformed confidence interval, we have $U = e^{\frac{1.96\times 0.07906402}{\frac{81}{605}\log\left(\frac{81}{605}\right)}} = 0.5623523$

The confidence interval is therefore $\left[\left(\frac{81}{605}\right)^{\frac{1}{U}}, \left(\frac{81}{605}\right)^{U}\right] = [0.02799701, 0.32278579]$.

4. *An insurance company records the following data in a mortality study:*

entry	*death*	*exit*	*entry*	*death*	*exit*	*entry*	*death*	*exit*
70.2	-	*73.3*	*70.4*	-	*71.3*	*71.5*	-	*71.9*
68.5	-	*72.3*	*68.7*	*71.4*	-	*70.6*	-	*72.5*
70.9	*71.1*	-	*68.2*	-	*73.5*	*69.4*	-	*73.5*
71.4	-	*72.4*	*68.1*	-	*72.2*	*70.2*	-	*74.3*
69.9	*71.9*	-	*68.4*	-	*72.5*	*69.4*	-	*72.2*
70.1	-	*72.6*	*71.5*	-	*72.2*	*70.0*	-	*72.1*
68.7	-	*74.2*	*70.9*	*71.1*	-	*70.2*	-	*72.4*
68.8	-	*71.4*	*71.4*	-	*74.6*	*69.6*	-	*73.7*
68.4	-	*71.2*	*69.1*	-	*71.3*	*70.6*	-	*73.4*
68.3	-	*71.7*						

Estimate the probability of an individual currently aged exactly 71 dying within the next year using:

(a) the exact exposure method.

The exact exposure is $1+1+0.1+0.6+0.9+1+1+0.4+0.2+0.7+0.3+0.4+1+1+1+0.5+0.1+0.6+0.3+0.4+1+1+1+1+1+1+1+1 = 20.5$. The rate of death is therefore $\lambda = \frac{4}{20.5}$, so the probability of dying is $1 - e^{-\frac{4}{20.5}} = 0.1772657$.

(b) the actuarial exposure method.

Using actuarial exposure, the exposure is $1+1+1+0.6+1+1+1+0.4+0.2+0.7+0.3+1+1+1+1+0.5+1+0.6+0.3+0.4+1+1+1+1+1+1+1+1 = 23$. Therefore the probability of death is $\frac{4}{23} = 0.173913$.

5. *An insurance company observes the following claims (in thousands):*

   2.5 2.9 2.9 3.6 3.8 4.0 4.1 4.8 5.1 5.2 5.9 6.0 6.7 7.8 8.4

   using a kernel density estimate with a uniform kernel with bandwidth 2, estimate the expected loss per claim if the company introduces a deductible of 2.0 on each policy.

   Using a uniform kernel with bandwidth 2, if the centre is at point $x \geqslant 3$, then the mean of the excess loss for this uniform distribution is $x - 2$. If the centre is at point $1 \leqslant x \leqslant 3$, then the mean excess of loss is $\int_2^{x+1} \frac{z}{2}\,dz = \left[\frac{z^2}{4}\right]_2^{x+1} = \frac{(x+1)^2}{4} - 1$. The mean excess of loss for the kernel distribution is therefore
   $\frac{1}{15}\left(\frac{3.5^2}{4} + \frac{3.9^2}{4} + \frac{3.9^2}{4} - 3 + 3.6 + 3.8 + 4.0 + 4.1 + 4.8 + 5.1 + 5.2 + 5.9 + 6.0 + 6.7 + 7.8 + 8.4\right) = \frac{73.0675}{15} = 4.8711667$

   The probability that the loss is less than the deductible is $\frac{1}{15}(0.5 + 0.1 + 0.1) = \frac{0.7}{15}$, so the mean payment per loss is $\frac{4.8711667}{1-\frac{0.7}{15}} = \frac{73.0675}{14.3} = 5.109615$.

6. *Using the following table:*

Age	*No. at start*	*enter*	*die*	*leave*	*No. at next age*
61	*0*	*5*	*2*	*1*	*2*
62	*2*	*6*	*0*	*4*	*4*
63	*4*	*7*	*1*	*0*	*10*
64	*10*	*2*	*0*	*8*	*4*
65	*4*	*6*	*2*	*6*	*2*
66	*2*	*7*	*0*	*9*	*0*

   Estimate the probability that an individual aged 62 withdraws from the policy within the next year, conditional on surviving to the end of the years.

   We assume that all events happen in the middle of the year. Using exact exposure, we find that the exposure is $2 + (6-4) \times 0.5 = 3$ at age 62. Therefore the rate of withdrawl at age 62 is $\frac{4}{3}$ and the probability of not withdrawing during the year is $e^{-\frac{4}{3}} = 0.2635971$, so the probability of withdrawl is $1 - 0.2635971 = 0.7364029$.

   If instead we use the actuarial exposure, we have the exposure is $2 + 6 \times 0.5 = 5$. This means the probability of withdrawing is $\frac{4}{5} = 0.8$.

Standard Questions

7. *An insurance company collects the following claim data (in thousands):*

i	d_i	x_i	u_i	i	d_i	x_i	u_i	i	d_i	x_i	u_i
1	0.0	0.1	-	9	1.0	1.8	-	17	2.0	3.6	-
2	0.0	0.7	-	10	1.0	2.2	-	18	2.0	6.4	-
3	0.0	1.8	-	11	1.0	2.6	-	19	2.0	9.6	-
4	0.0	-	5	12	1.0	11.3	-	20	2.0	-	15
5	0.5	-	10	13	1.0	-	5	21	5.0	5.3	-
6	0.5	-	20	14	1.5	4.5	-	22	5.0	7.5	-
7	0.5	-	10	15	1.5	-	10	23	5.0	8.5	-
8	1.0	1.6	-	16	2.0	2.4	-	24	5.0	-	10

It is attempting to choose a deductible for a new policy. The company has set the policy limit to 12.0. Customer satisfaction surveys have shown that at most 20% of claims should exceed the policy limit. Using a Kaplan-Meier product limit estimator, find the largest deductible they can apply while still meeting this criterion.

We want that if L is a random loss $P(L > 12|L > d) < 0.1$.

We form the following table:

x_i	r_i	S_i	$P(L > 12\|L \geqslant x_i)$
0.1	4	3	$\frac{875}{6964}$
0.7	6	5	$\frac{875}{5346}$
1.6	11	10	$\frac{175}{891}$
1.8	12	10	$\frac{35}{162}$
2.2	15	14	$\frac{7}{27}$
2.4	14	13	$\frac{5}{18}$
2.6	13	12	$\frac{35}{117}$
3.6	12	11	$\frac{35}{108}$
4.5	11	10	$\frac{35}{99}$
5.3	12	11	$\frac{7}{18}$
6.4	11	10	$\frac{14}{33}$
7.5	10	9	$\frac{7}{15}$
8.5	9	8	$\frac{14}{27}$
9.6	8	7	$\frac{7}{12}$
11.3	3	2	$\frac{2}{3}$

We see that the first probability in the final column to be more than 0.2 is $\frac{175}{891}$, so the deductible should be less than 1.6.