EXAM C QUESTIONS OF THE WEEK

S. Broverman, 2006

Week of October 16/06

A particular type of individual health insurance policy models the annual loss per policy as an exponential distribution with a mean that varies with individual insured. A sample of 1000 randomly selected policies results in the following data regarding annual loss amounts in interval grouped form.

Interval	Number of Losses
$[0, 100]$	500
$(100, 200]$	250
$(200, 500]$	150
$(500, 1000]$	60
$(1000, 2000]$	40

It is assumed that the loss amounts are uniformly distributed within each interval.

Apply semiparametric empirical Bayes credibility to estimate the loss in the 3rd year for a particular individual who had annual policy losses of 150 in the first year and 0 in the second year.

Solution can be found below.

Week of October 16/06 - Solution

X is the random variable for annual loss. We are given that the conditional distribution of X given θ is exponential with a mean of θ, where θ has an unspecified distribution. Therefore, the hypothetical mean is $HM = E[X|\theta] = \theta$ and the process variance is $PV = Var[X|\theta] = \theta^2$.

The expected hypothetical mean is $\mu = EHM = E[\,E[X|\theta]\,] = E[\theta] = E[X]$ (using the double expectation rule $E[\,E[X|\theta]\,] = \theta$).
The expected process variance is $v = EPV = E[\,Var[X|\theta]\,] = E[\theta^2]$.
The variance of the hypothetical mean is
$a = VHM = Var[\,E[X|\theta]\,] = Var[\theta] = E[\theta^2] - (E[\theta])^2$.
From this we see that $v - a = (E[\theta])^2 = \mu^2$.

In general, $Var[X] = E[\,Var[X|\theta]\,] + Var[\,E[X|\theta]\,] = v + a$.

From the data set we can use empirical estimation to estimate $E[X]$:

$$\widehat{\mu} = (\tfrac{500}{1000})(\tfrac{0+100}{2}) + (\tfrac{250}{1000})(\tfrac{100+200}{2}) + (\tfrac{150}{1000})(\tfrac{200+500}{2}) + (\tfrac{60}{1000})(\tfrac{500+1000}{2}) + (\tfrac{40}{1000})(\tfrac{1000+2000}{2}) = 220 .$$

From the data set we can use empirical estimation to estimate $E[X^2]$:

$$(\tfrac{500}{1000})(\tfrac{100^3-0^3}{3(100-0)}) + (\tfrac{250}{1000})(\tfrac{200^3-100^3}{3(200-100)}) + (\tfrac{150}{1000})(\tfrac{500^3-200^3}{3(500-200)}) + (\tfrac{60}{1000})(\tfrac{1000^3-500^3}{3(1000-500)}) + (\tfrac{40}{1000})(\tfrac{2000^3-1000^3}{3(2000-1000)}) = 155{,}333 .$$

The empirical estimate of the variance of X is $155{,}333 - (220)^2 = 106{,}933$.

In the semiparametric empirical Bayes credibility model, we use the empirical estimate of $E[X]$ for μ, so that $\widehat{\mu} = 220$. We also know that $Var[X] = v + a$, so using the empirical estimate of $Var[X]$ gives is $106{,}933 = \widehat{v} + \widehat{a}$. But we also know that, for this model, $v - a = \mu^2$, so using our sample estimate of μ, we have $\widehat{v} - \widehat{a} = (220)^2$. We can then solve the two equations $\widehat{v} + \widehat{a} = 106{,}933$ and $\widehat{v} - \widehat{a} = 48{,}400$ to get $\widehat{v} = 77{,}667$ and $\widehat{a} = 29{,}267$.

We can now find the estimated loss in the 3rd year for a policy that had losses of $Y_1 = 150$ in the first year and $Y_2 = 0$ in the second year. The estimate is $\widehat{Z}\overline{Y} + (1 - \widehat{Z})\widehat{\mu}$, where $\widehat{Z} = \frac{2}{2+\frac{\widehat{v}}{\widehat{a}}} = .4298$ and $\widehat{\mu} = 220$. The credibility premium is
$(.4298)(75) + (.5702)(220) = 158$.