Aggregate Claims Modelling - Parameter Variability

- Idea:
    - Preceding discussions assumed parameters known
    - In practice parameters only estimated
    - Assume parameters themselves have some distribution
    - Investigate the consequences of "unknown" parameters
    - Quite complex in general

Aggregate Claims Modelling - Parameter Variability

- Specific Example:
  - · Let S be aggregate claim amount based on IRM
  - · S has parameters q_i's and $F_i(x)$
  - · Assume $F_i(x)$ known, but $q_i \overset{iid}{\sim} H(q)$
  - · Define $\eta_1 = E(q_i)$ and $\eta_2 = E(q_i^2)$

    [ASIDE: Assuming known η's rather than known q_i's reduces number of necessary parameters from n to 2]
  - · Examine how this structure effects characteristics of $\tilde{S}$, the Poisson CRM approximation to S
  - · Previously: $E(\tilde{S}) = \sum_{i=1}^n q_i\mu_i$ and $Var(\tilde{S}) = \sum_{i=1}^n q_i(\sigma_i^2 + \mu_i^2)$
  - · Now,

$$\begin{aligned}
E(\tilde{S}) &= E\left(\sum_{i=1}^n \tilde{Y}_i\right) = \sum_{i=1}^n E(\tilde{Y}_i) = \sum_{i=1}^n E\{E(\tilde{Y}_i|q_i)\} \\
&= \sum_{i=1}^n E(q_i\mu_i) = \eta_1 \sum_{i=1}^n \mu_i
\end{aligned}$$

and

$$\begin{aligned}
Var(\tilde{S}) &= Var\left(\sum_{i=1}^n \tilde{Y}_i\right) = \sum_{i=1}^n Var(\tilde{Y}_i) \\
&= \sum_{i=1}^n \left[E\{Var(\tilde{Y}_i|q_i)\} + Var\{E(\tilde{Y}_i|q_i)\}\right] \\
&= \sum_{i=1}^n \left[E\{q_i(\sigma_i^2 + \mu_i^2)\} + Var\{q_i\mu_i\}\right] \\
&= \sum_{i=1}^n \left\{\eta_1(\sigma_i^2 + \mu_i^2) + \mu_i^2(\eta_2 - \eta_1^2)\right\} \\
&= \eta_1 \sum_{i=1}^n \sigma_i^2 + (\eta_2 + \eta_1 - \eta_1^2) \sum_{i=1}^n \mu_i^2,
\end{aligned}$$

NOTE: The following discussion is not in the course notes.

Suppose that the parameter estimates for probability of death (q) are uncertain for each age group, and the estimates for the q are an *iid* sample from an exponential distribution with parameters γ_1, γ_2 and γ_3 for each of the three age groups (using the mean parameterisation for the exponential).

Recall that for an exponentially distributed variable q with mean parameter γ:
$E(q) = \gamma$
$E(q^2) = 2\gamma^2$
$Var(q) = \gamma^2$

Using the notation from page 105 of Section 2 of the lecture notes, it follows that:
$\eta_1 = E(q) = \gamma$
$\eta_2 = E(q^2) = 2\gamma^2$
$\eta_2 - \eta_1^2 = E(q^2) - E(q)^2 = Var(q) = \gamma^2$

Then,
$E(\tilde{S}) = \sum \gamma\mu_i$
$Var(\tilde{S}) = \sum \gamma\sigma_i^2 + \sum(\gamma^2 + \gamma)\mu_i^2$

Since we assume a different exponential distribution for each age group, the mean and variance of $\tilde{S}$ can be written as follows:

$$E(\tilde{S}) = \sum_{i=1}^{2500}(\gamma_1\mu_i) + \sum_{i=2501}^{4500}(\gamma_2\mu_i) + \sum_{i=4501}^{6000}(\gamma_3\mu_i).$$

$$Var(\tilde{S}) = \left(\sum_{i=1}^{2500}\gamma_1\sigma_i^2 + (\gamma_1^2 + \gamma_1)\sum_{i=1}^{2500}\mu_i^2\right) + \left(\sum_{i=2501}^{4500}\gamma_2\sigma_i^2 + (\gamma_2^2 + \gamma_2)\sum_{i=2501}^{4500}\mu_i^2\right)$$
$$+\left(\sum_{i=4501}^{6000}\gamma_3\sigma_i^2 + (\gamma_3^2 + \gamma_3)\sum_{i=4501}^{6000}\mu_i^2\right).$$

For consistency with the previous example, we assume that the means of the exponential are equal to the fixed probabilities of death corresponding with each age group. That is, $\gamma_1 = 0.0007, \gamma_2 = 0.0025$, $\gamma_3 = 0.0085$.

Therefore,

$E(\tilde{S})$
$=\sum_{i=1}^{2500}(0.0007(48000)(0.25)) + \sum_{i=2501}^{4500}(0.0025(65250)(0.23)) + \sum_{i=4501}^{6000}(0.0085(81000)(0.22))$
=323,242.50

$Var(\tilde{S})$
$= \left(\sum_{i=1}^{2500} 0.0007(48000)(0.25^2) + ((0.0007)^2 + 0.0007)\sum_{i=1}^{2500}(48000^2)(0.25^2)\right) +$
$+\left(\sum_{i=1}^{2500} 0.0025(65250)(0.23^2) + ((0.0025)^2 + 0.0025)\sum_{i=1}^{2500}(65250^2)(0.23^2)\right)$
$+\left(\sum_{i=1}^{2500} 0.0085(81000)(0.22^2) + ((0.0085)^2 + 0.0085)\sum_{i=1}^{2500}(81000^2)(0.22^2)\right)$
= 5,464,397,330

The mean of $\tilde{S}$ is identical to the mean when we assumed fixed q*'s*. This is as expected since we chose the means of the exponentials to be identical to the fixed probabilities of death from the earlier example.

However, the variance of $\tilde{S}$ is higher than the variance based on the fixed q's. This is because there is an extra term in the variance of $\tilde{S}$ due to the uncertainty in the q's when we assume that they are no longer fixed but instead follow a distribution.