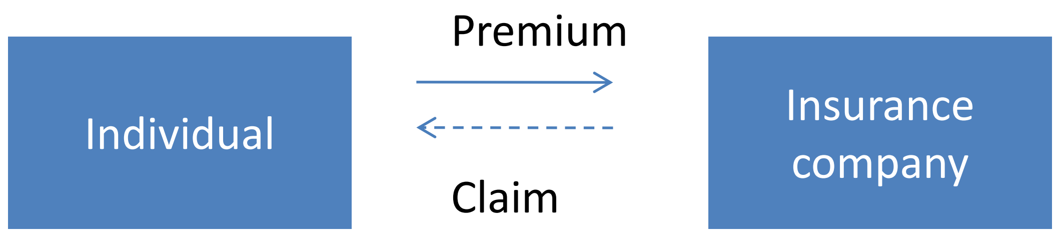
STK 4540 Lecture 3

Uncertainty on different levels
And
Random intensities in the claim frequency

Overview pricing (1.2.2 in EB)



$$\text{Total claim} = \begin{cases} 0 \text{ (no event above deductible), with probability } 1-\text{p} \\ X \text{ (event above deductible) with probability p} \end{cases}$$

Due to the law of large numbers the insurance company is cabable of estimating *the expected claim amount*

$$\underbrace{E(X)}_{\text{Risk premium}} = \underbrace{p\underbrace{\int_{all\,x} s\overbrace{f(s)}^{\text{Distribution of X, estimated with claims data}}ds}_{\text{Expected claim amount given an event}}}_{\text{Expected consequence of claim}} + (1-p)0 = \pi^{Pu}$$

- Probability of claim,
- Estimated with claim frequency
- We are interested in the distribution of the claim frequency
- The premium charged is the risk premium inflated with a *loading* (overhead and margin)

Control (1.2.3 in EB)

- Companies are obliged to aside funds to cover future obligations
- Suppose a portfolio consists of J policies with claims $X_1,\ldots,X_J$
- The total claim is then

$$\chi = X_1 + \ldots + X_J$$

Portfolio claim size

- We are interested in $E(\chi)$ as well as its distribution
- Regulators demand sufficient funds to cover χ with high probability
- The mathematical formulation is in term of q_ε, which is the solution of the equation

$$\Pr\{\chi > q_\varepsilon\} = \varepsilon$$

where ε is small for example 1%

- The amount q_ε is known as the solvency capital or reserve

Insurance works because risk can be diversified away through size (3.2.4 EB)

- The core idea of insurance is risk spread on many units
- Assume that policy risks $X_1,\dots,X_J$ are stochastically independent
- Mean and variance for the portfolio total are then

$$E(\chi) = \pi_1 + \dots + \pi_J \quad \text{and} \quad \text{var}(\chi) = \sigma_1 + \dots + \sigma_J$$

$$\text{and} \quad \pi_j = E(X_j) \quad \text{and} \quad \sigma_j = sd(X_j). \text{ Introduce}$$

$$\bar{\pi} = \frac{1}{J}(\pi_1 + \dots + \pi_J) \quad \text{and} \quad \bar{\sigma}^2 = \frac{1}{J}(\sigma_1 + \dots + \sigma_J)$$

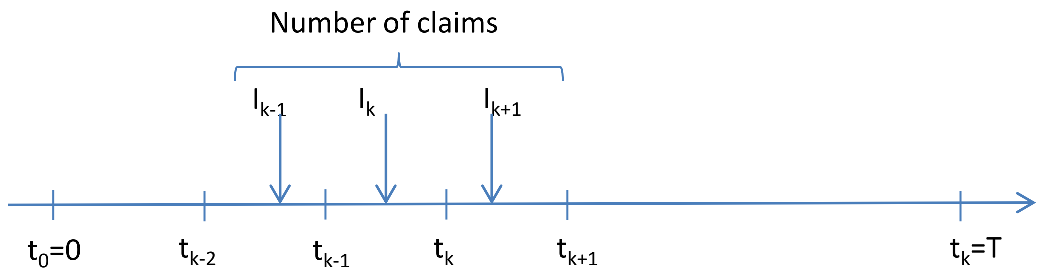
which is average expectation and variance. Then

$$E(\chi) = J\bar{\pi} = \quad \text{and} \quad \text{sd}(\chi) = \sqrt{J}\ \bar{\sigma} \text{ so that } \frac{\text{sd}(\chi)}{\text{E}(\chi)} = \frac{\bar{\sigma}/\bar{\pi}}{\sqrt{J}}$$

- The *coefficient of variation* (shows extent of variability in relation to the mean) approaches 0 as J grows large (law of large numbers)
- Insurance risk can be diversified away through size
- Insurance portfolios are still not risk-free because
  - of uncertainty in underlying models
  - risks may be dependent

The world of Poisson

Poisson
Some notions
Examples
Random intensities



- What is rare can be described mathematically by cutting a given time period T into K small pieces of equal length h=T/K
- Assuming no more than 1 event per interval the count for the entire period is

$N=I_1+\ldots+I_K$, where I_j is either 0 or 1 for j=1,...,K

- If $p=\Pr(I_k=1)$ is equal for all k and events are independent, this is an ordinary Bernoulli series

$$\Pr(N=n)=\frac{K!}{n!(K-n)!}p^n(1-p)^{K-n}, \quad \text{for} \quad n=0,1,\ldots,K$$

- Assume that p is proportional to h and set $p=\mu h$ where μ is an intensity which applies per time unit

$$\Rightarrow \Pr(N=n) \underset{K\to\infty}{\to} \frac{(\mu T)^n}{n!}e^{-\mu T}$$

In the limit N is Poisson distributed with parameter $\lambda=\mu T$

The Poisson distribution

- Claim numbers, N for policies and **N** for portfolios, are Poisson distributed with parameters

$$\lambda = \mu T \quad \text{and} \quad \lambda = J\mu T$$

Policy level Portfolio level

The intensity μ is an average over time and policies.

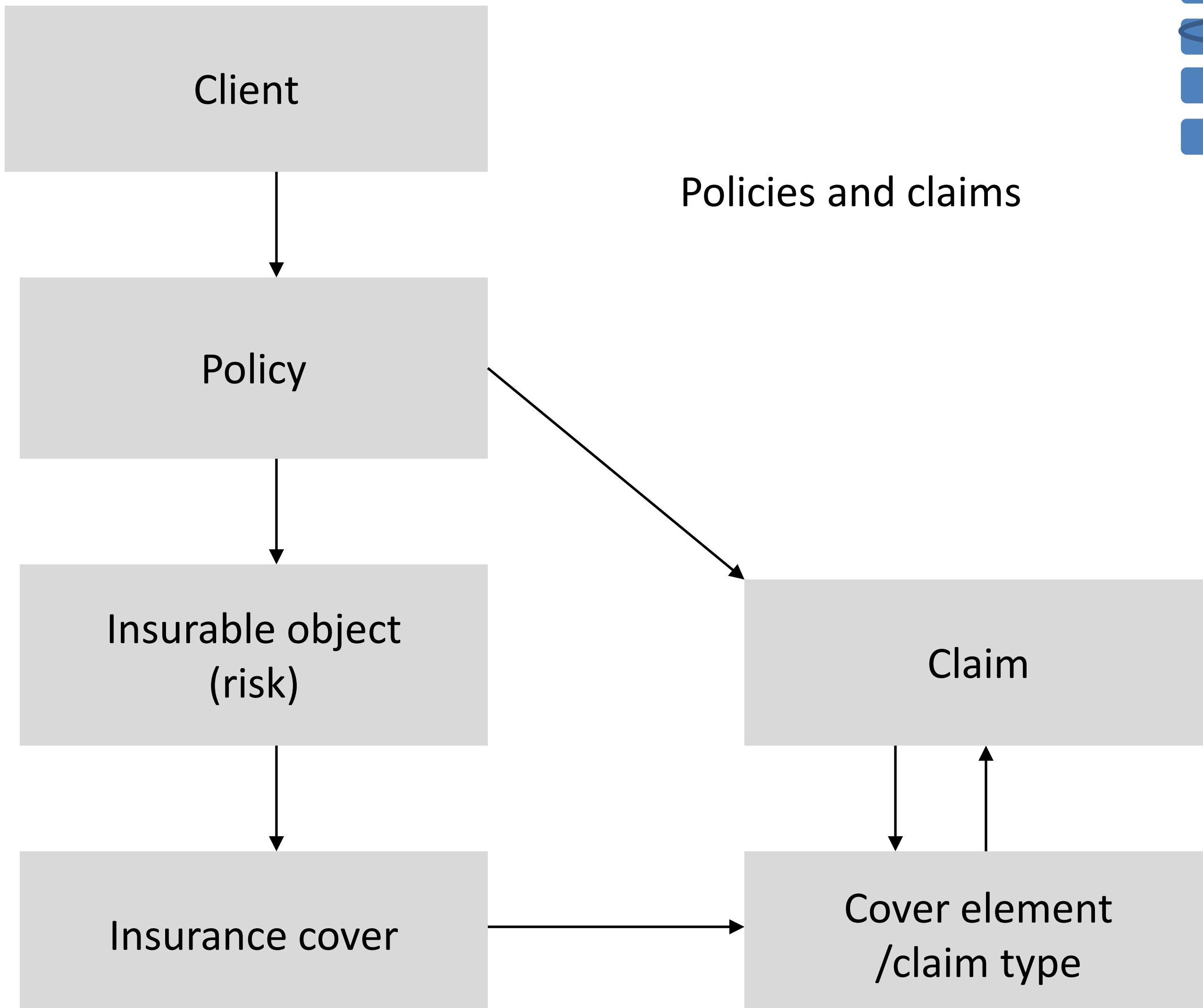
Poisson models have useful operational properties. Mean, standard deviation and skewness are

$$E(N) = \lambda, \quad sd(N) = \sqrt{\lambda} \quad \text{and} \quad skew(\lambda) = \frac{1}{\sqrt{\lambda}}$$

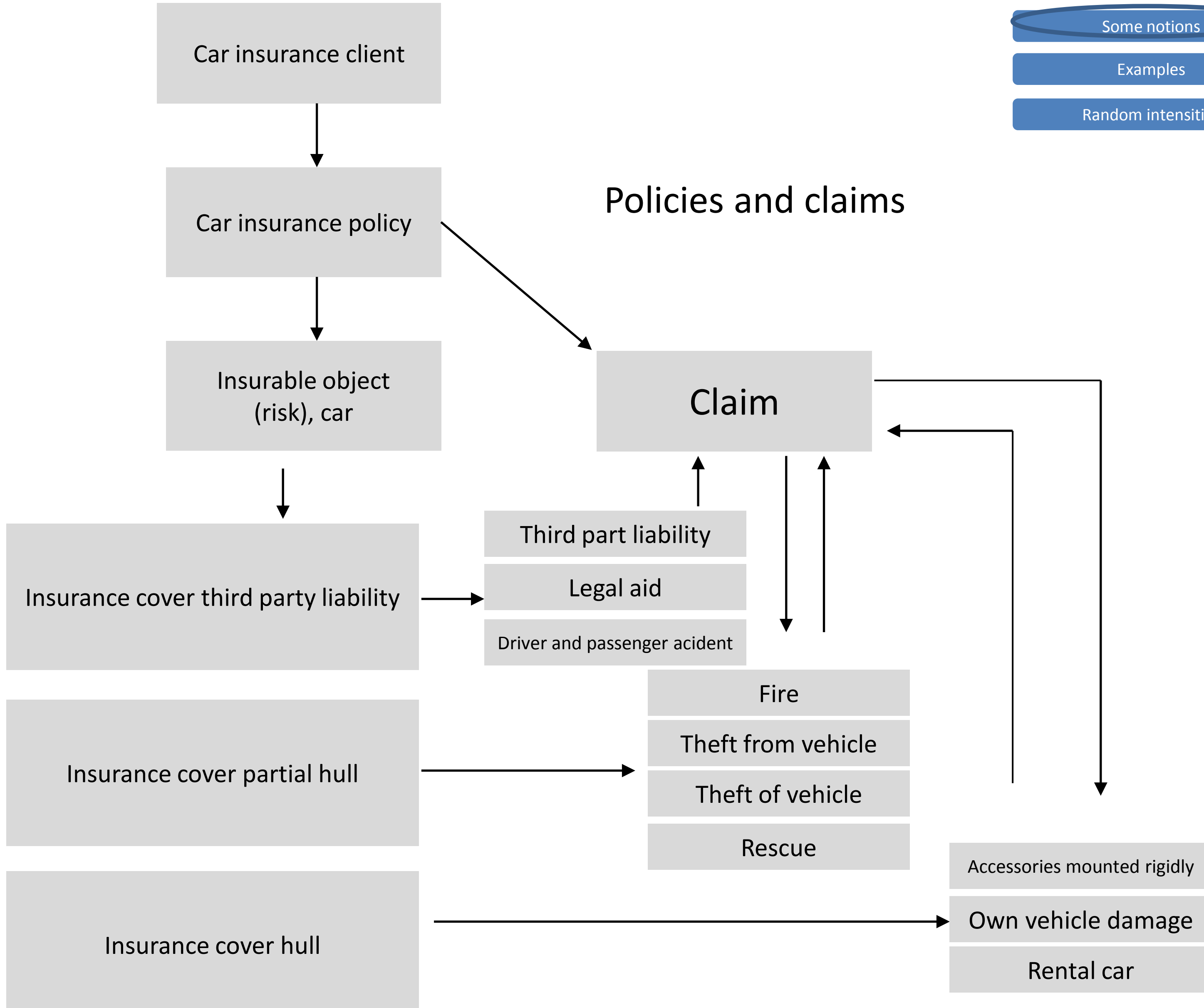
The sums of independent Poisson variables must remain Poisson, if $N_1,\ldots,N_J$ are independent and Poisson with parameters $\lambda_1,\ldots,\lambda_J$ then

$$\mathbf{N} = N_1 + \ldots + N_J \sim \ Poisson(\lambda_1 + \ldots + \lambda_J)$$

Policies and claims

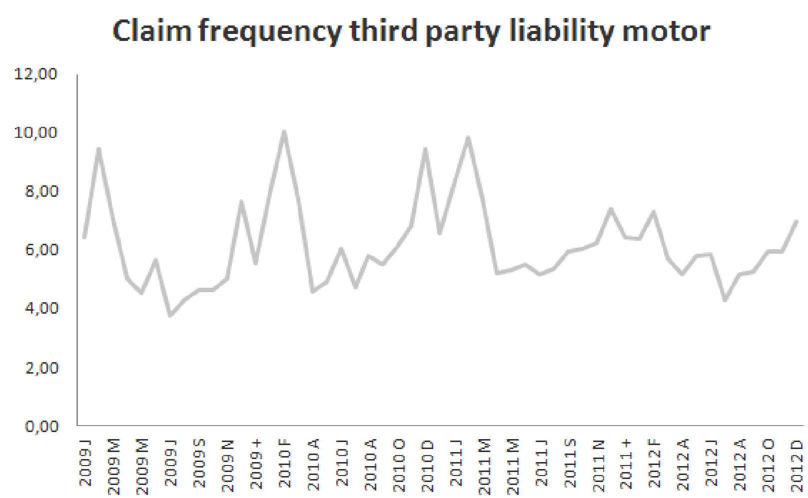


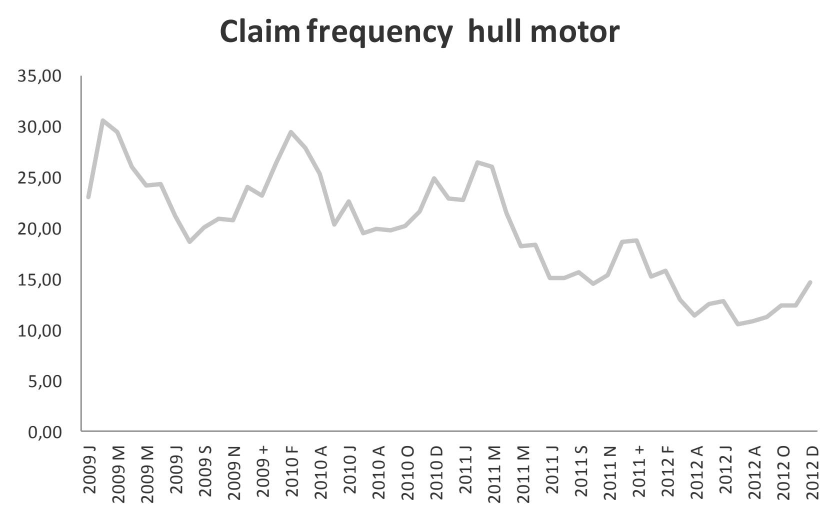
Policies and claims



Key ratios – claim frequency TPL and hull

- The graph shows claim frequency for third part liability and hull for motor insurance





Random intensities (Chapter 8.3)

- How μ varies over the portfolio can partially be described by observables such as age or sex of the individual (treated in Chapter 8.4)
- There are however factors that have impact on the risk which the company can't know much about
  - Driver ability, personal risk averseness,
- This randomeness can be managed by making μ a stochastic variable
- This extension may serve to capture uncertainty affecting all policy holders jointly, as well, such as altering weather conditions
- The models are conditional ones of the form

$$N \mid \mu \sim Poisson(\mu T) \quad \text{and} \quad \mathbf{N} \mid \mu \sim Poisson(J\mu T)$$

Policy level — Portfolio level

- Let $\xi = E(\mu)$ and $\sigma = \text{sd}(\mu)$ and recall that $E(N \mid \mu) = \text{var}(N \mid \mu) = \mu T$

  which by double rules in Section 6.3 imply

$$E(N) = E(\mu T) = \xi T \quad \text{and} \quad \text{var}(N) = E(\mu T) + \text{var}(\mu T) = \xi T + \sigma^2 T^2$$

- Now E(N)<var(N) and N is no longer Poisson distributed

The rule of double variance

Let X and Y be arbitrary random variables for which

$$\xi(x) = E(Y \mid x) \quad \text{and} \quad \sigma^2 = \mathrm{var}(Y \mid x)$$

Then we have the important identities

$$\underbrace{\xi = E(Y) = E\{\xi(X)\}}_{\text{Rule of double expectation}} \quad \text{and} \quad \underbrace{\mathrm{var(Y)} = \mathrm{E}\{\sigma^2(X)\} + \mathrm{var}\{\xi(X)\}}_{\text{Rule of double variance}}$$

Recall rule of double expectation

$$E(E(Y \mid x)) = \int_{\text{all x}} (E(Y \mid x))\, f_X(x)dx = \int_{\text{all x}} \int_{\text{all y}} y f_{Y|X}(y|x)dy\, f_X(x)dx$$

$$= \int_{\text{all y}} \int_{\text{all x}} y f_{X,Y}(x,y)dxdy = \int_{\text{all y}} y \int_{\text{all x}} f_{X,Y}(x,y)dxdy = \int_{\text{all y}} y f_Y(y)dy = E(Y)$$

wikipedia tells us how the rule of double variance can be proved

Law of total variance

From Wikipedia, the free encyclopedia

In probability theory, the **law of total variance**[1] or **variance decomposition formula**, states that if *X* and *Y* are random variables on the same probability space, and the variance of *Y* is finite, then

$$\operatorname{Var}[Y] = \operatorname{E}(\operatorname{Var}[Y \mid X]) + \operatorname{Var}(\operatorname{E}[Y \mid X]).$$

Proof [edit source]

The law of total variance can be proved using the law of total expectation.[3] First,

$$\operatorname{Var}[Y] = \operatorname{E}[Y^2] - \operatorname{E}[Y]^2$$

from the definition of variance. Then we apply the law of total expectation to each term by conditioning on the random variable *X*:

$$= \operatorname{E}_X\left[\operatorname{E}[Y^2 \mid X]\right] - \operatorname{E}_X\left[\operatorname{E}[Y \mid X]\right]^2$$

Now we rewrite the conditional second moment of Y in terms of its variance and first moment:

$$= \operatorname{E}_X\left[\operatorname{Var}[Y \mid X] + \operatorname{E}[Y \mid X]^2\right] - \operatorname{E}_X[\operatorname{E}[Y \mid X]]^2$$

Since the expectation of a sum is the sum of expectations, the terms can now be regrouped:

$$= \operatorname{E}_X[\operatorname{Var}[Y \mid X]] + \left(\operatorname{E}_X\left[\operatorname{E}[Y \mid X]^2\right] - \operatorname{E}_X[\operatorname{E}[Y \mid X]]^2\right)$$

Finally, we recognize the terms in parentheses as the variance of the conditional expectation E[Y|X]:

$$= \operatorname{E}_X\left[\operatorname{Var}[Y \mid X]\right] + \operatorname{Var}_X\left[\operatorname{E}[Y \mid X]\right]$$

The rule of double variance

Var(Y) will now be proved from the rule of double expectation. Introduce

$$\hat{Y} = \xi(x) \quad \text{and note that} \quad \mathrm{E}(\hat{Y}) = \mathrm{E(Y)}$$

which is simply the rule of double expectation. Clearly

$$(Y-\xi)^2 = ((Y-\hat{Y})+(\hat{Y}-\xi))^2 = (Y-\hat{Y})^2 + (\hat{Y}-\xi)^2 + 2(Y-\hat{Y})(\hat{Y}-\xi).$$

Passing expectations over this equality yields

$$\mathrm{var}(Y) = B_1 + B_2 + 2B_3$$

where

$$B_1 = E(Y-\hat{Y})^2, \; B_2 = E(\hat{Y}-\xi)^2, \quad B_3 = E(Y-\hat{Y})(\hat{Y}-\xi),$$

which will be handled separately. First note that

$$\sigma^2(x) = E\{(Y-\xi(x))^2 \mid x\} = E\{(Y-\hat{Y})^2 \mid x\}$$

and by the rule of double expectation applied to $(Y-\hat{Y})^2$

$$E\{\sigma^2(x)\} = E\{(Y-\hat{Y})^2 = B_1.$$

The second term makes use of the fact that $\xi = E(\hat{Y})$ by the rule of double expectation so that

The rule of double variance

$$B_2 = \mathrm{var}(\hat{Y}) = \mathrm{var}\{\xi(x)\}.$$

The final term B_3 makes use of the rule of double expectation once again which yields

$$B_3 = E\{c(X)\}$$

where

$$c(X) = E\{(Y - \hat{Y})(\hat{Y} - \xi) \mid x\} = E\{(Y - \hat{Y}) \mid x\}(\hat{Y} - \xi)$$

$$= \{E(Y \mid x) - \hat{Y})\}(\hat{Y} - \xi) = \{\hat{Y} - \hat{Y})\}(\hat{Y} - \xi) = 0$$

And B_3=0. The second equality is true because the factor $(\hat{Y} - \xi)$ is fixed by X. Collecting the expression for B_1, B_2 and B_3 proves the double variance formula

Random intensities

Specific models for μ are handled through the mixing relationship

$$\Pr(N=n)=\int_0^\infty \Pr(N=n\,|\,\mu)g(\mu)d\mu \approx \sum_i \Pr(N=n\,|\,\mu_i)\Pr(\mu=\mu_i)$$

Gamma models are traditional choices for $g(\mu)$ and detailed below

Estimates of ξ and σ can be obtained from historical data without specifying $g(\mu)$. Let $n_1,\dots,n_n$ be claims from n policy holders and $T_1,\dots,T_J$ their exposure to risk. The intensity μ_j of individual j is then estimated as $\hat{\mu}_j = n_j / T_j$.

Uncertainty is huge but pooling for portfolio estimation is still possible. One solution is

$$\hat{\xi}=\sum_{j=1}^{n} w_j \hat{\mu}_j \quad \text{where} \quad w_j=\frac{T_j}{\sum_{i=1}^{n} T_i} \qquad (1.5)$$

and

$$\hat{\sigma}^2=\frac{\sum_{j=1}^{n} w_j(\hat{\mu}_j-\hat{\xi})^2-c}{1-\sum_{j=1}^{n} w_j^2} \quad \text{where} \quad c=\frac{(n-1)\hat{\xi}}{\sum_{i=1}^{n} T_i} \qquad (1.6)$$

Both estimates are unbiased. See Section 8.6 for details. 10.5 returns to this.

The negative binomial model

The most commonly applied model for muh is the Gamma distribution. It is then assumed that

$$\mu = \xi G \quad \text{where} \quad G \sim \text{Gamma}(\alpha)$$

Here $\text{Gamma}(\alpha)$ is the standard Gamma distribution with mean one, and μ fluctuates around ξ with uncertainty controlled by α. Specifically

$$E(\mu) = \xi \quad \text{and} \quad \text{sd}(\mu) = \xi / \sqrt{\alpha}$$

Since $\text{sd}(\mu) \to 0$ as $\alpha \to \infty$, the pure Poisson model with fixed intensity emerges in the limit.

The closed form of the density function of N is given by

$$\Pr(N = n) = \frac{\Gamma(n+\alpha)}{\Gamma(n+1)\Gamma(\alpha)} p^{\alpha}(1-p)^{n} \quad \text{where} \quad \text{p} = \frac{\alpha}{\alpha + \xi \text{T}}$$

for n=0,1,.... This is the negative binomial distribution to be denoted $\text{nbin}(\xi, \alpha)$. Mean, standard deviation and skewness are

$$E(N) = \xi T, \quad sd(N) = \sqrt{\xi T(1+\xi T / \alpha)}, \quad \text{skew(N)} = \frac{1 + 2\xi \text{T}/\alpha}{\sqrt{\xi T(1+\xi T / \alpha)}} \qquad (1.9)$$

Where E(N) and sd(N) follow from (1.3) when $\sigma = \xi / \sqrt{\alpha}$ is inserted.
Note that if $N_1, \ldots, N_J$ are iid then $N_1 + \ldots + N_J$ is nbin (convolution property).

Fitting the negative binomial

Moment estimation using (1.5) and (1.6) is simplest technically. The estimate of ξ is simply in (1.5), and $\hat{\xi}$ for α invoke (1.8) right which yields

$$\hat{\sigma} = \hat{\xi} / \sqrt{\hat{\alpha}} \quad \text{so that} \quad \hat{\alpha} = \hat{\xi}^2 / \hat{\sigma}^2 .$$

If $\hat{\sigma} = 0$, interpret it as an infinite $\hat{\alpha}$ or a pure Poisson model.

Likelihood estimation: the log likelihood function follows by inserting n_j for n in (1.9) and adding the logarithm for all j. This leads to the criterion

$$L(\xi, \alpha) = \sum_{j=1}^{n} \log(n_j + \alpha) - n\{\log(\Gamma(\alpha)) - \alpha \log(\alpha)\} +$$

$$\sum_{j=1}^{n} n_j \log(\xi) - (n_j + \alpha) \log(\alpha + \xi T_j)$$

where constant factors not depending on α and ξ have been omitted.