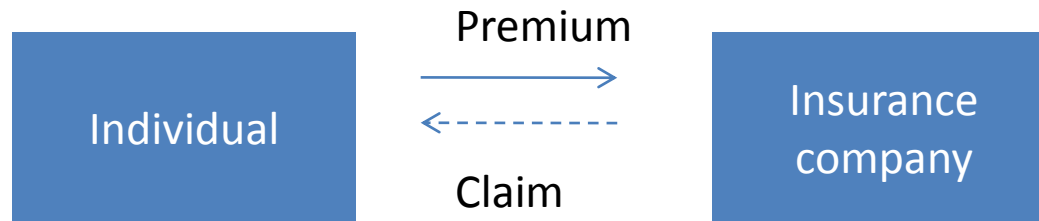




## Recap (1.2.1 in EB)

- Property insurance is economic responsibility for incidents such as fires and accidents passed on to an insurer against a fee
- The contract, known as a policy, releases claims when such events occur
- A central quantity is the total claim  $X$  amassed during a certain period of time (typically a year)

# Overview pricing (1.2.2 in EB)



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

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

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

- 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  $X_1, \dots, X_J$
- The total claim is then

$$\chi = X_1 + \dots + X_J$$

Portfolio claim size

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

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

where  $\varepsilon$  is small for example 1%

- The amount  $q_\varepsilon$  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

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

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

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

- 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

# How are random variables sampled?

Inversion (2.3.2 in EB):

- Let  $F(x)$  be a strictly increasing distribution function with inverse  $X = F^{-1}(U)$  and let

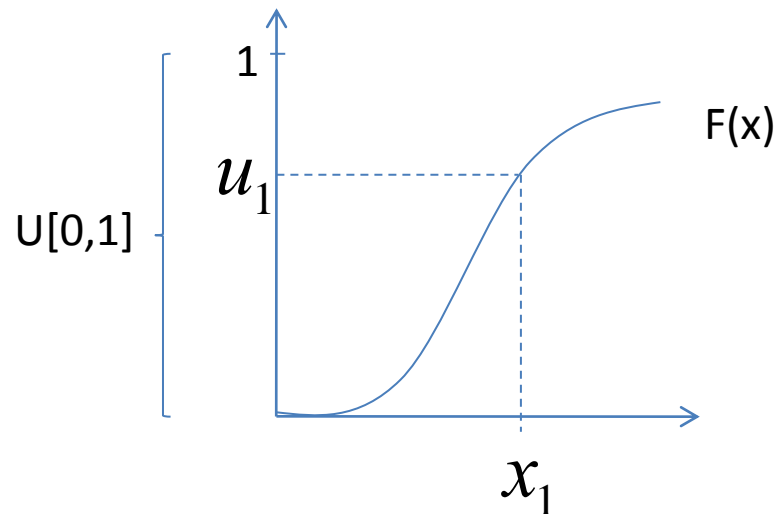
$$X = F^{-1}(U) \text{ or } X = F^{-1}(1-U), \quad U \sim \text{Uniform}$$

- Consider the specification on the left for which  $U=F(X)$
- Note that

$$\Pr(X \leq x) = \Pr(F(X) \leq F(x))$$

$$= \Pr(U \leq F(x)) = F(x)$$

since  $\Pr(U \leq u) = u$



# Outline of the course

|                                  | Models treated   | Curriculum                     | Duration in lectures |
|----------------------------------|--|--------------------------------|----------------------|
| Basic concepts and introduction  |  | EB 1.2, 2.3.1, 2.3.2, 3.2, 3.3 | 1                    |
| How is claim frequency modelled? | Poisson, Compound Poisson, Poisson regression, negative binomial model                 | EB 8.2, 8.3, 8.4               | 2-3                  |
| How is claim reserving modelled? | Delay modelling, chain ladder  | EB 8.5, Note                   | 1-2                  |
| How is claim size modelled?      | Gamma distribution, log-normal distribution, Pareto distribution, Weibull distribution | EB 9                           | 2-3                  |
| How is pricing done?             | Binomial models  | EB 10                          | 1                    |
| Solvency                         | Monte Carlo simulation   | EB 10, Note                    | 1-2                  |
| Credibility theory               | Buhlmann Straub  | EB 10                          | 1                    |
| Reinsurance                      |  | EB 10                          | 1                    |
| Repetition                       |  |                                | 1                    |

# Course literature

Curriculum:

Chapter 1.2, 2.3.1, 2.3.2, 2.5, 3.2, 3.3 in EB

Chapter 8,9,10 in EB

Note on Chain Ladder

Lecture notes by NFH

Exercises

The following book will be used (EB):

Computation and Modelling in Insurance and Finance, Erik Bølviken,  
Cambridge University Press (2013)

- Additions to the list above may occur during the course
- Final curriculum will be posted on the course web site in due time

Assignment must be approved to be able to participate  
in exam



# Overview of this session

The Poisson model (Section 8.2 EB)

Some important notions and some practice too

Examples of claim frequencies

Random intensities (Section 8.3 EB)

# Introduction

- Actuarial modelling in general insurance is broken down on claim frequency and claim size
- This is natural due to definition of risk premium:

$$\underbrace{E(X)}_{\text{Risk premium}} = \underbrace{p}_{\text{Probability of claim,}} \underbrace{\int_{\text{all } x} sf(s)ds}_{\text{Expected claim amount given an event}} + (1-p)0 = \pi^{Pu}$$

Expected consequence of claim

•Probability of claim,

•Estimated with claim frequency

- The Poisson distribution is often used in modelling the distribution of claim numbers
- The parameter is  $\lambda = \mu h \cdot T$  (single policy) and  $\lambda = J \cdot \mu h \cdot T$  (portfolios)
- The modelling can be made more sophisticated by extending the model for  $\mu h$ , either by making  $\mu h$  stochastic or by linking  $\mu h$  to explanatory variables

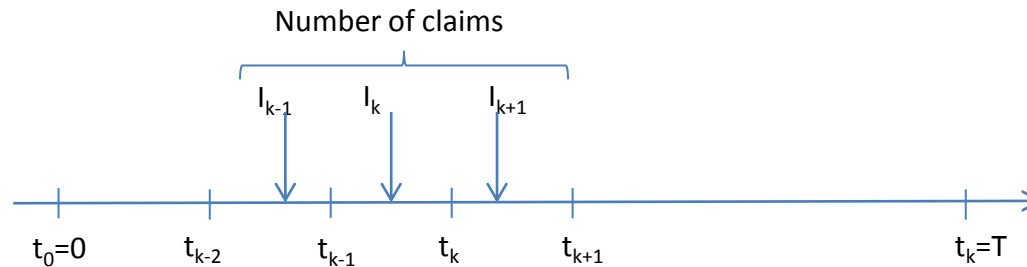
# The world of Poisson (Chapter 8.2)

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  $h=T/K$
- On short intervals the chance of more than one incident is remote
- Assuming no more than 1 event per interval the count for the entire period is

$$N = I_1 + \dots + I_K, \text{ where } I_j \text{ is either 0 or 1 for } j=1, \dots, K$$

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

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

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

# The world of Poisson

Poisson

Some notions

Examples

Random intensities

$$\begin{aligned}
 \Pr(N = n) &= \frac{K!}{n!(K-n)!} p^n (1-p)^{K-n} \\
 &= \frac{K!}{n!(K-n)!} \left(\frac{\mu T}{K}\right)^n \left(1 - \frac{\mu T}{K}\right)^{K-n} \\
 &= \frac{(\mu T)^n}{n!} \underbrace{\frac{K(K-1)\cdots(K-n+1)}{K^n}}_{\substack{\rightarrow 1 \\ K \rightarrow \infty}} \underbrace{\left(1 - \frac{\mu T}{K}\right)^K}_{\substack{\rightarrow e^{-\mu T} \\ K \rightarrow \infty}} \underbrace{\frac{1}{\left(1 - \frac{\mu T}{K}\right)^n}}_{\substack{\rightarrow 1 \\ K \rightarrow \infty}}
 \end{aligned}$$

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

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

# The world of Poisson

- Let us proceed removing the zero/one restriction on  $I_k$ . A more flexible specification is

$$\Pr(I_k = 0) = 1 - \mu h + o(h), \quad \Pr(I_k = 1) = \mu h + o(h), \quad \Pr(I_k > 1) = o(h)$$


Where  $o(h)$  signifies a mathematical expression for which

$$\frac{o(h)}{h} \rightarrow 0 \quad \text{as } h \rightarrow 0$$


It is verified in Section 8.6 that  $o(h)$  does not count in the limit

Consider a portfolio with  $J$  policies. There are now  $J$  independent processes in parallel and if  $\mu_j$  is the intensity of policy  $j$  and  $I_k$  the total number of claims in period  $k$ , then

$$\Pr(I_k = 0) = \prod_{j=1}^J (1 - \mu_j h) \quad \text{and} \quad \Pr(I_k = 1) = \sum_{i=1}^J \left\{ \mu_i h \prod_{j \neq i} (1 - \mu_j h) \right\}$$



No claims



Claims policy  $i$  only

# The world of Poisson

- Both quantities simplify when the products are calculated and the powers of  $h$  identified

$$\begin{aligned}
 \Pr(\mathbf{I}_k = 0) &= \prod_{j=1}^3 (1 - \mu_j h) = (1 - \mu_1 h)(1 - \mu_2 h)(1 - \mu_3 h) \\
 &= (1 - \mu_1 h - \mu_2 h + \mu_2 \mu_1 h^2)(1 - \mu_3 h) \\
 &= 1 - \mu_1 h - \mu_2 h + \mu_2 \mu_1 h^2 - \mu_3 h(1 - \mu_1 h - \mu_2 h + \mu_2 \mu_1 h^2) \\
 &= 1 - \mu_1 h - \mu_2 h - \mu_3 h + o(h) \\
 \Pr(\mathbf{I}_k = 1) &= \left( \sum_{j=1}^J \mu_j \right) h + o(h)
 \end{aligned}$$

- It follows that the portfolio number of claims  $\mathbf{N}$  is Poisson distributed with parameter

$$\lambda = (\mu_1 + \dots + \mu_J)T = J\bar{\mu}T, \quad \text{where } \bar{\mu} = (\mu_1 + \dots + \mu_J) / J$$

- When claim intensities vary over the portfolio, only their average counts

# When the intensity varies over time

- A time varying function  $\mu = \mu(t)$  handles the mathematics. The binary variables  $l_1, \dots, l_k$  are now based on different intensities

$\mu_1, \dots, \mu_K$  where  $\mu_k = \mu(t_k)$  for  $k = 1, \dots, K$

- When  $l_1, \dots, l_k$  are added to the total count  $N$ , this is the same issue as if  $K$  different policies apply on an interval of length  $h$ . In other words,  $N$  must still be Poisson, now with parameter

$$\lambda = h \sum_{k=1}^K \mu_k \rightarrow \int_0^T \mu(t) dt \quad \text{as } h \rightarrow 0$$

where the limit is how integrals are defined. The Poisson parameter for  $N$  can also be written

$$\lambda = T \bar{\mu} \quad \text{where} \quad \bar{\mu} = \frac{1}{T} \int_0^T \mu(t) dt,$$

And the introduction of a time-varying function  $\mu(t)$  doesn't change things much. A time average  $\bar{\mu}$  takes over from a constant  $\mu$

# The Poisson distribution

- Claim numbers,  $N$  for policies and  $\mathbf{N}$  for portfolios, are Poisson distributed with parameters

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

Policy level                      Portfolio level

The intensity  $\mu$  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  $N_1, \dots, N_J$  are independent and Poisson with parameters  $\lambda_1, \dots, \lambda_J$ , then

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



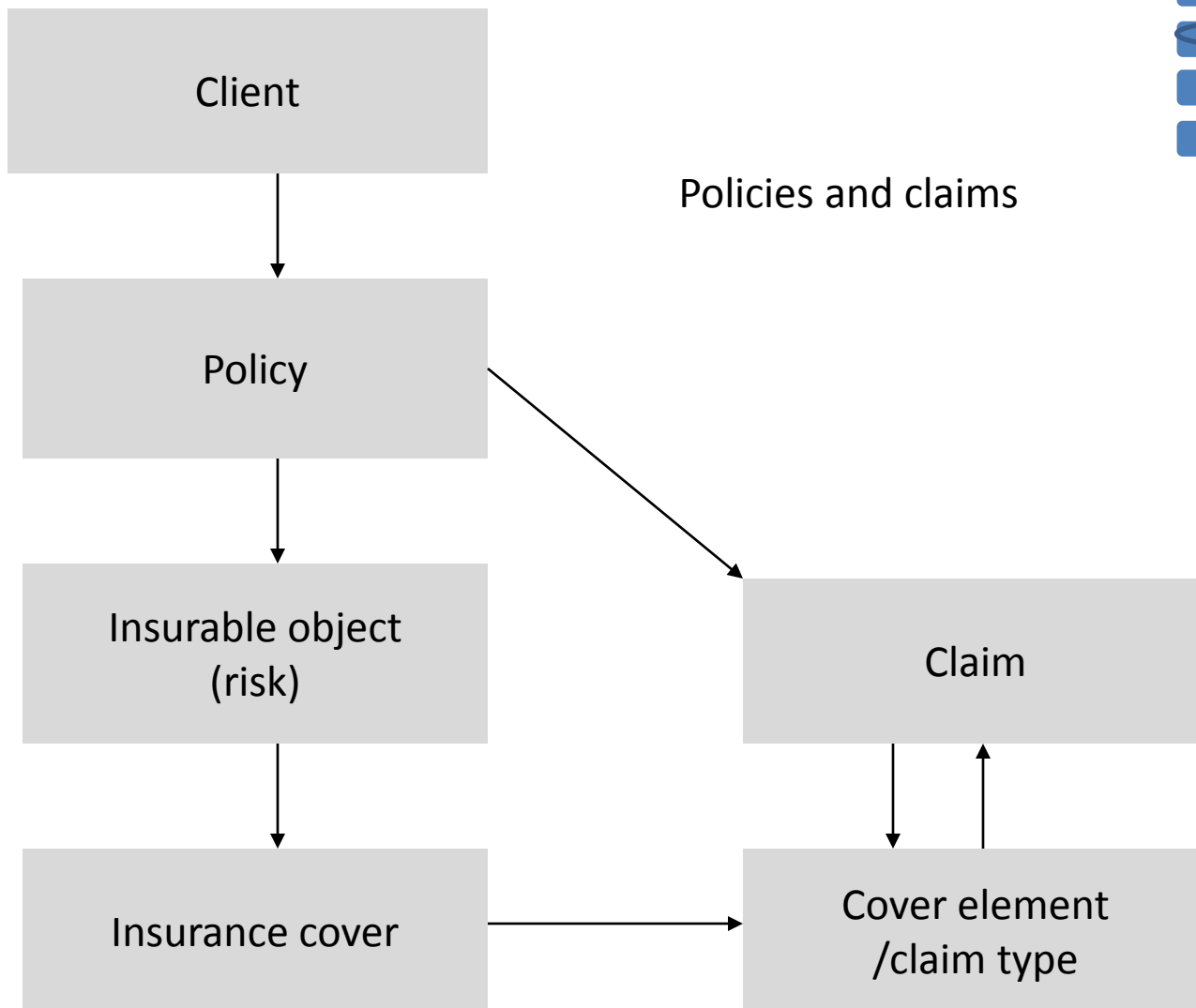
Poisson

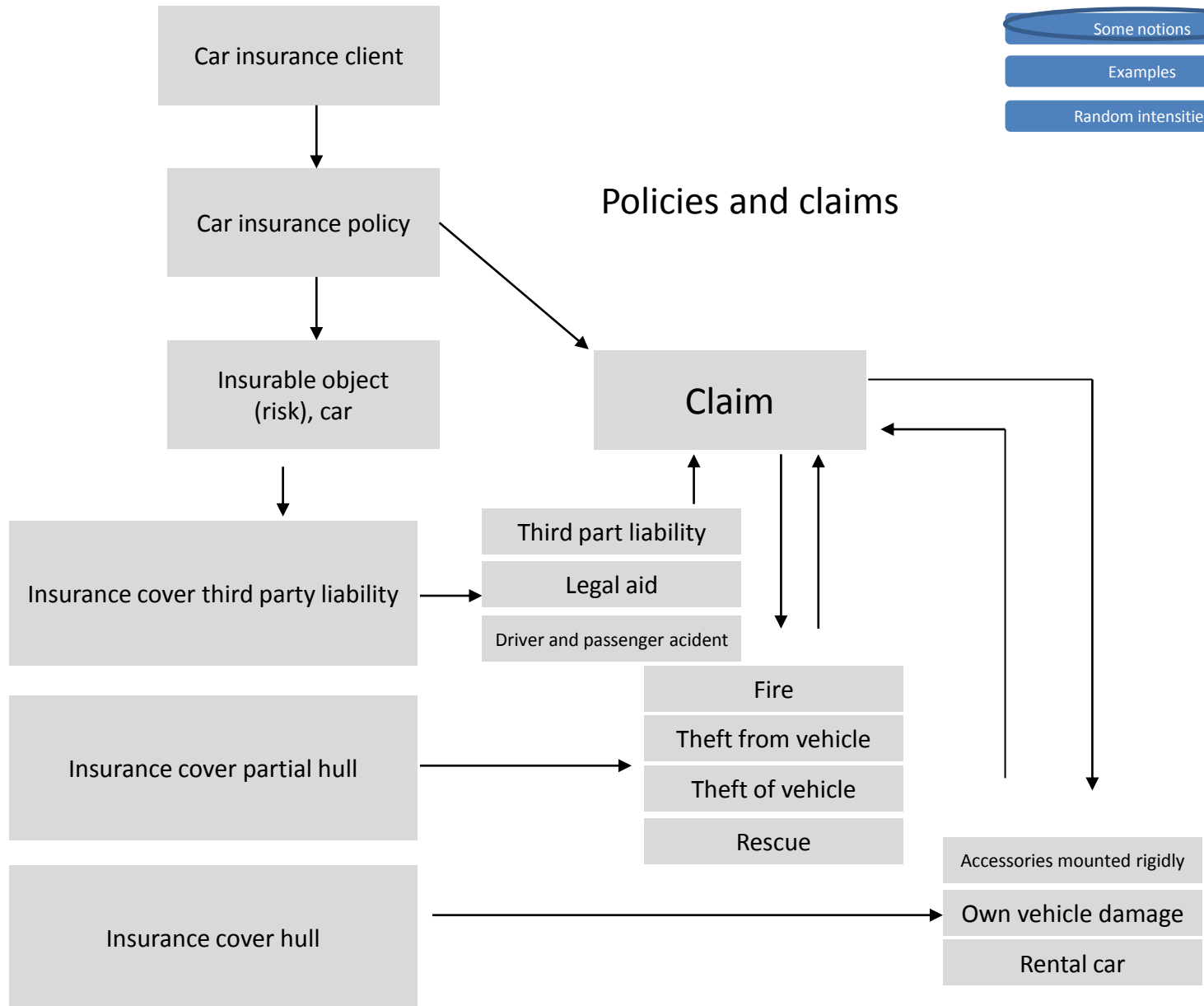
Some notions

Examples

Random intensities

## Policies and claims





## Some notes on the different insurance covers on the previous slide:

**Third part liability** is a mandatory cover dictated by Norwegian law that covers damages on third part vehicles, property and person. Some insurance companies provide additional coverage, as legal aid and driver and passenger accident in-

Poisson

Some notions

Examples

provide

Random intensities

**Partial Hull** covers everything that the third part liability covers. In addition, partial hull covers damages on own vehicle caused by fire, glass rupture, theft and vandalism in association with theft. Partial hull also includes rescue. Partial hull does not cover damage on own vehicle caused by collision or landing in the ditch. Therefore, partial hull is a more affordable cover than the Hull cover. Partial hull also cover salvage, home transport and help associated with disruptions in production, accidents or disease.

**Hull** covers everything that partial hull covers. In addition, Hull covers damages on own vehicle in a collision, overturn, landing in a ditch or other sudden and unforeseen damage as for example fire, glass rupture, theft or vandalism. Hull may also be extended to cover rental car.

## Some notes on some important concepts in insurance:

### What is bonus?

Bonus is a reward for claim-free driving. For every claim-free year you obtain a reduction in the insurance premium in relation to the basis premium. This continues until 75% reduction is obtained.

### What is deductible?

The deductible is the amount the policy holder is responsible for when a claim occurs.

### Does the deductible impact the insurance premium?

Yes, by selecting a higher deductible than the default deductible, the insurance premium may be significantly reduced. The higher deductible selected, the lower the insurance premium.

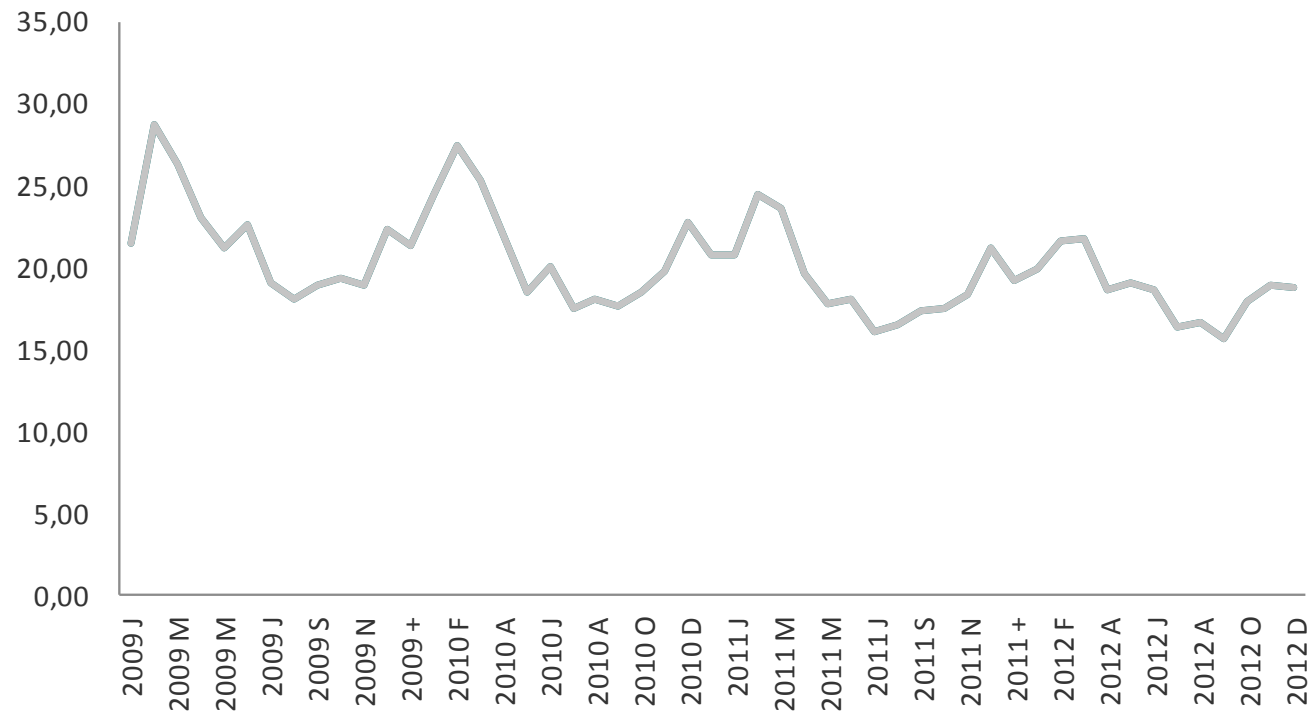
### How is the deductible taken into account when a claim is disbursed?

The insurance company calculates the total claim amount caused by a damage entitled to disbursement. What you get from the insurance company is then the calculated total claim amount minus the selected deductible.

# Key ratios – claim frequency

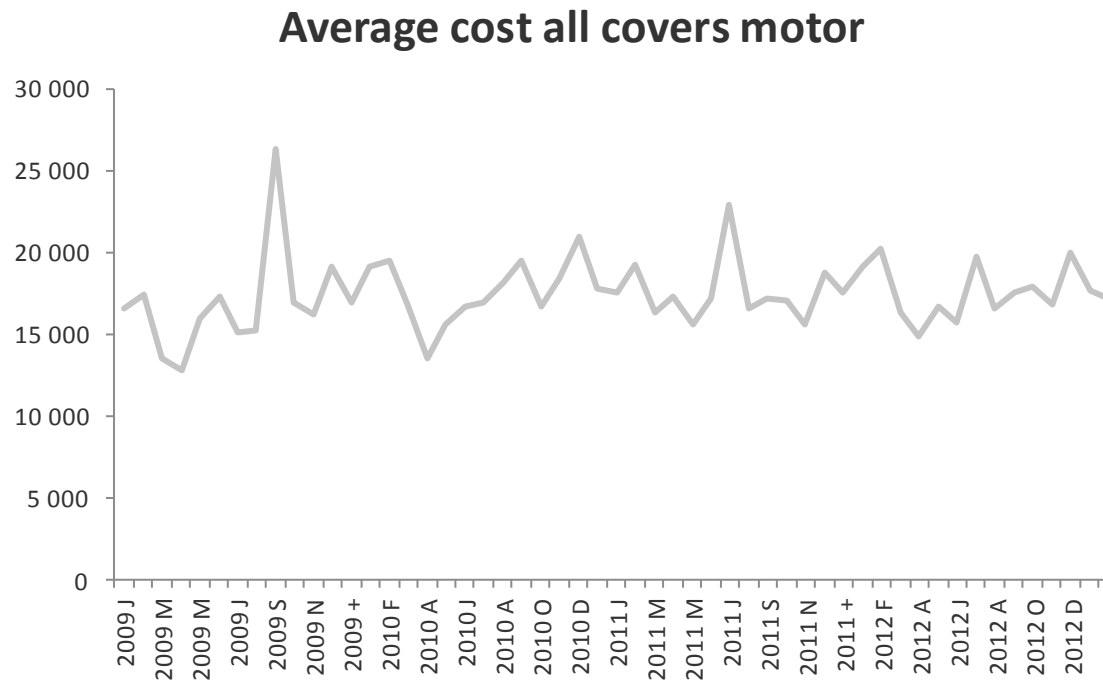
- The graph shows claim frequency for all covers for motor insurance
- Notice seasonal variations, due to changing weather condition throughout the years

## Claim frequency all covers motor



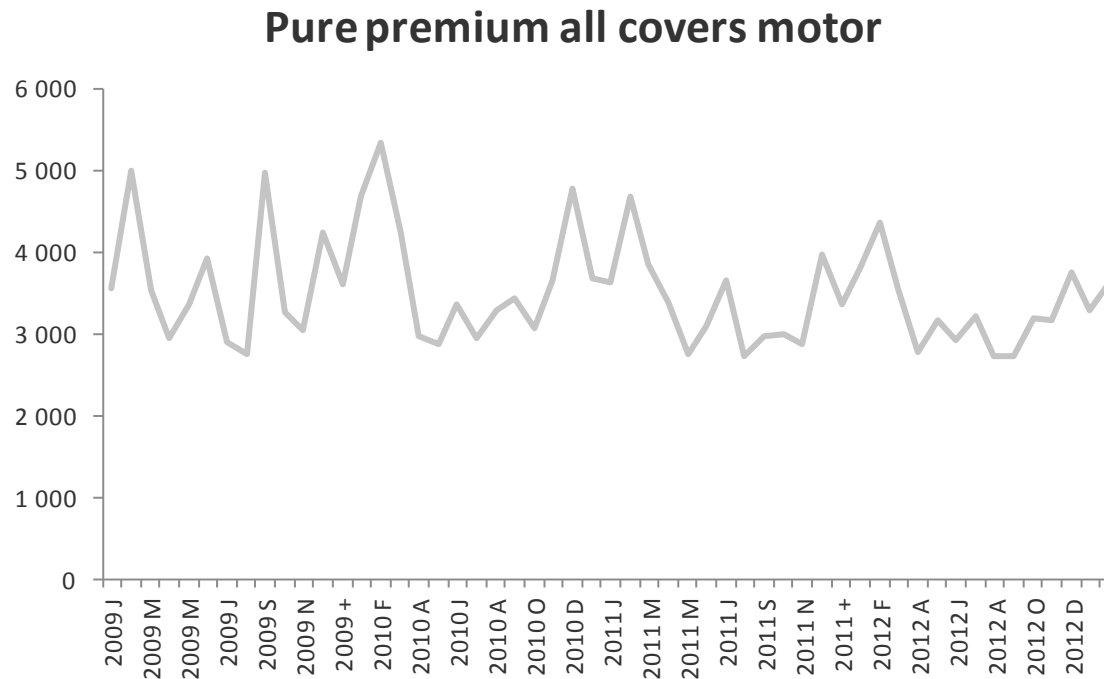
# Key ratios – claim severity

- The graph shows claim severity for all covers for motor insurance



# Key ratios – pure premium

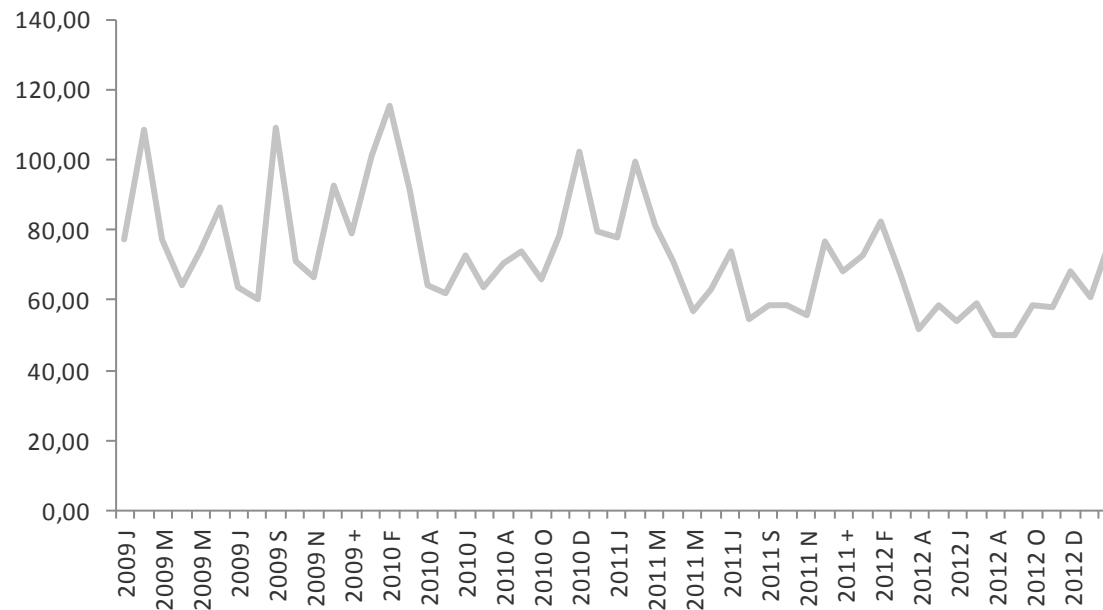
- The graph shows pure premium for all covers for motor insurance



# Key ratios – pure premium

- The graph shows loss ratio for all covers for motor insurance

**Loss ratio all covers motor**



# Key ratios – claim frequency TPL and hull

Poisson

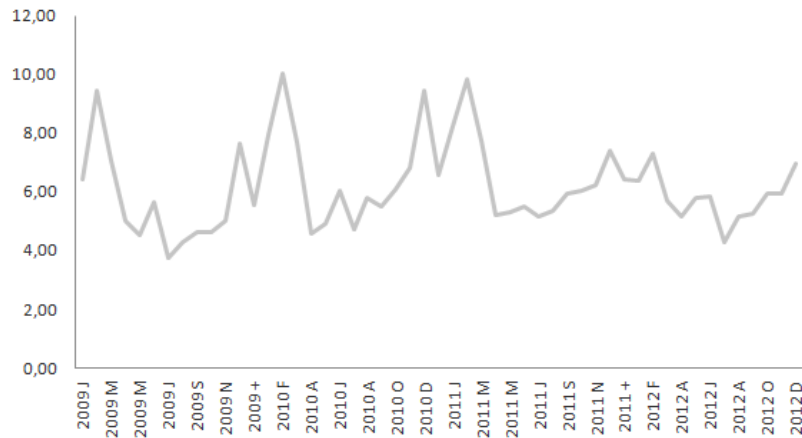
Some notions

Examples

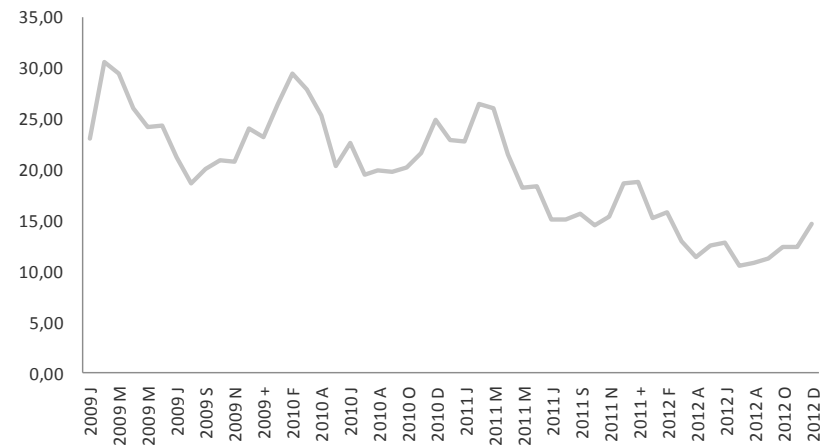
Random intensities

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

**Claim frequency third party liability motor**



**Claim frequency hull motor**





# Key ratios – claim frequency and claim severity

Poisson

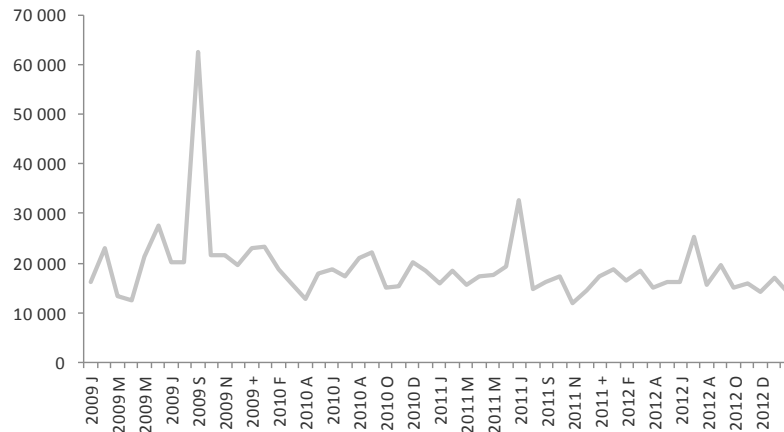
Some notions

Examples

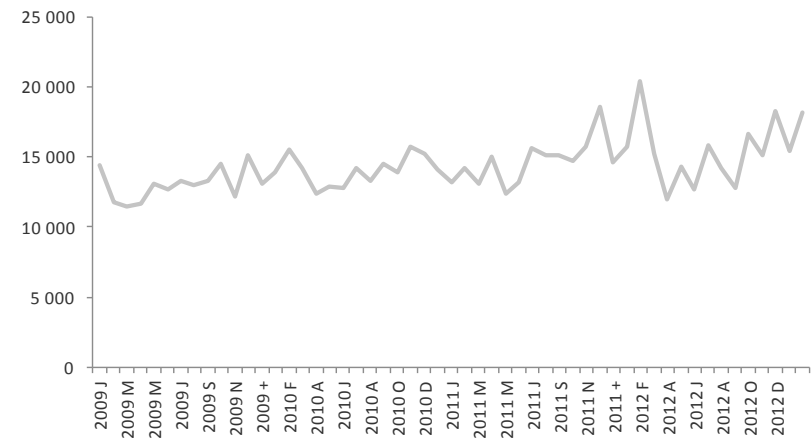
Random intensities

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

**Average cost third part liability motor**



**Average cost hull motor**



# Random intensities (Chapter 8.3)

- How  $\mu$  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 randomness can be managed by making  $\mu$  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

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

Policy level

Portfolio level

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

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  $E(N) < \text{var}(N)$  and  $N$  is no longer Poisson distributed

# The rule of double variance

Let  $X$  and  $Y$  be arbitrary random variables for which

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

Then we have the important identities

$$\xi = E(Y) = E\{\xi(X)\} \quad \text{and} \quad \text{var}(Y) = E\{\sigma^2(X)\} + \text{var}\{\xi(X)\}$$

Rule of double expectation Rule of double variance

Recall rule of double expectation

$$\begin{aligned} E(E(Y | x)) &= \int_{\text{all } x} (E(Y | x)) f_X(x) dx = \int_{\text{all } x} \int_{\text{all } y} y f_{Y|X}(y/x) dy f_X(x) dx \\ &= \int_{\text{all } y} \int_{\text{all } x} y f_{X,Y}(x, y) dx dy = \int_{\text{all } y} y \int_{\text{all } x} f_{X,Y}(x, y) dx dy = \int_{\text{all } y} y f_Y(y) dy = E(Y) \end{aligned}$$

# 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**<sup>[1]</sup> 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

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

### Proof [\[ edit source \]](#)

The law of total variance can be proved using the [law of total expectation](#).<sup>[3]</sup> First,

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

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

$$= \mathrm{E}_X [\mathrm{E}[Y^2 \mid X]] - \mathrm{E}_X [\mathrm{E}[Y \mid X]]^2$$

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

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

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

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

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

$$= \mathrm{E}_X [\mathrm{Var}[Y \mid X]] + \mathrm{Var}_X [\mathrm{E}[Y \mid X]]$$

# The rule of double variance

$\text{Var}(Y)$  will now be proved from the rule of double expectation. Introduce

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

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

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

where

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

which will be handled separately. First note that

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

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

double expectation so that

by the rule of double expectation

# The rule of double variance

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

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

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

where

$$\begin{aligned} c(X) &= E\{(\hat{Y} - Y)(\hat{Y} - \hat{\xi}) \mid x\} = E\{(\hat{Y} - Y) \mid x\}(\hat{Y} - \hat{\xi}) \\ &= \{E(\hat{Y} \mid x) - Y\}(\hat{Y} - \hat{\xi}) = \{\hat{Y} - Y\}(\hat{Y} - \hat{\xi}) = 0 \end{aligned}$$

And  $B_3=0$ . The second equality is true because the factor  $(\hat{Y} - \hat{\xi})$  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  $\mu$  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  $\mu$  and are detailed below

Estimates of  $\xi$  and  $\sigma$  can be obtained from historical data without specifying  $g(\mu)$ . Let  $n_1, \dots, n_n$  be claims from  $n$  policy holders and  $T_1, \dots, T_n$  their exposure to risk. The intensity if individual  $j$  is then estimated as  $\mu_j$ .  $\mu_j = n_j / T_j$

Uncertainty is huge. One solution is

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

and

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

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  $\mu$  is the Gamma distribution. It is then assumed that

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

Here  $\text{Gamma}(\alpha)$  is the standard Gamma distribution with mean one, and  $\mu$  fluctuates around  $\xi$  with uncertainty controlled by  $\alpha$  specifically

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

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

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

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

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

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

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



# Fitting the negative binomial

Moment estimation using (1.5) and (1.6) is simplest technically. The estimate of  $\xi$  is simply  $\bar{n}$  in (1.5), and for  $\alpha$  invoke (1.8) 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  $\alpha$  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  $\xi$  and  $\alpha$  have been omitted.