Queen Mary
University of London

# MTH5126 Statistics for Insurance

Dr Lei Fang

lei.fang@qmul.ac.uk

## Week 4

Great credit and thanks to D. Boland, G. Ng and F. Parsa, previous lecturers, for their excellent work in producing the original version of these notes.

# Risk models (continued)

Practical applications of risk models

Aggregate claims distributions under proportional reinsurance

- Example

Aggregate claims distributions under excess of loss reinsurance

- Example

Reinsurer's aggregate claims under excess of loss reinsurance

- Example

Exam style question

The individual risk model

- Assumptions
- Differences compared with the collective risk model
- Mean and variance of aggregate claims in the individual risk model
- Special case
- Example

Parameter variability

Variability in a heterogeneous portfolio

- Example 1

Variability in a homogeneous portfolio

- Example 2

# Practical applications of risk models

## *Introduction*

We will look at:

- how the models can be adapted for situations involving reinsurance
- the individual risk model
- parameter variability/uncertainty

# Aggregate claims distributions under proportional reinsurance

- The distribution of the **number** of claims involving the **reinsurer** is the **same** as the distribution of the number of claims involving the insurer, as each pays a defined proportion of every claim.
- For a retention level $\alpha$ $(0 \leq \alpha \leq 1)$, the $i^{th}$ individual claim amount for the insurer is $\alpha X_i$ and for the reinsurer is $(1 - \alpha)X_i$.
- The aggregate claims amounts are $\alpha S$ and $(1 - \alpha)S$ respectively.

Queen Mary
University of London

# Aggregate claims distributions under proportional reinsurance: Example

**Example:**

Show that under a **proportional reinsurance** arrangement where the direct writer retains a proportion, $k$, the MGF $M_Y(t)$ of the net individual claim amount $Y$ paid by the direct insurer is $M_X(kt)$.

Hence find an expression for the MGF of the aggregate claim amount if the number of claims has a $Poisson(\lambda)$ distribution.

**Answer:**

Under this arrangement, if the gross amount of an individual claim is $X$, the net amount paid by the direct insurer will be $Y = kX$. So the MGF will be:

$$M_Y(t) = E(e^{tY}) = E(e^{tkX}) = E(e^{(kt)X}) = M_X(kt)$$

so, the MGF of the aggregate claim amount is:

$$M_{S_{net}}(t) = M_N[logM_Y(t)] = exp\left(\lambda[e^{logM_Y(t)} - 1]\right) = exp(\lambda[M_X(kt) - 1])$$

MGF of Poisson distribution: $M_N(t) = e^{\lambda(e^t - 1)}$
(Slide 20, Week 3)

# Aggregate claims distributions under excess of loss reinsurance

- The amount that an insurer pays on the $i^{th}$ claim under individual excess of loss reinsurance with retention level $M$ is:

$$Y_i = \begin{cases} X_i & X_i < M \\ M & X_i \geq M \end{cases}$$

- The amount that the reinsurer pays is:

$$Z_i = \begin{cases} 0 & X_i < M \\ X_i - M & X_i \geq M \end{cases}$$

- So, the insurer's aggregate claims net of reinsurance can be represented as:

$$S_I = Y_1 + Y_2 + \cdots + Y_N$$

- and the reinsurer's aggregate claims as:

$$S_R = Z_1 + Z_2 + \cdots + Z_N$$

Queen Mary University of London

# Aggregate claims distributions under excess of loss reinsurance

- If, for example, $N \sim Poi(\lambda)$, $S_I$ has a compound Poisson distribution with Poisson parameter $\lambda$ and the $i^{th}$ individual claim amount is $Y_i$ .
- Similarly, $S_R$ has a compound Poisson distribution with Poisson parameter $\lambda$ and the $i^{th}$ individual claim amount is $Z_i$ .
- Note, however, that if $F(M) = P(X \leq M) > 0$, as will usually be the case, then **$Z_i$ may take the value 0**.
- In other words, 0 is counted as a possible claim amount for the reinsurer.
- From a practical point of view, this definition of $S_R$ is rather artificial.
- The insurer will know the observed value of $N$, but the reinsurer will probably know only the number of claims above the retention level $M$ since the insurer may notify the reinsurer only of claims above the retention level.

# Aggregate claims distributions under excess of loss reinsurance: Example

**Question:**

The annual aggregate claim amount from a risk has a compound Poisson distribution with Poisson parameter 10. Individual claim amounts are uniformly distributed on (0,2000). The insurer of this risk has effected excess of loss reinsurance with retention level 1,600.

a) Calculate the mean, variance and coefficient of skewness of the insurer's aggregate claims under this reinsurance arrangement.
b) Calculate the mean, variance and coefficient of skewness of the reinsurer's aggregate claims under this reinsurance arrangement.
c) What is the variance of $S$, the aggregate claim amount before reinsurance?
d) Why is it that:

$$var(S_I) + var(S_R) \neq var(S)$$

# Aggregate claims distributions under excess of loss reinsurance: Example

**Answer:**

**(a) Mean of insurer's aggregate claims**

Let $S_I$ and $S_R$ be as defined. To find $E(S_I)$ we calculate $E(Y_i)$:

$$E(Y_i) = \int_0^M xf(x)dx + MP(X_i > M)$$

where $f(x) = 0.0005$ is the $U(0, 2000)$ density function and $M = 1,600$. So:

$$E(Y_i) = \left[\frac{0.0005x^2}{2}\right]_0^M + 0.2M = 960$$

And

$$E(S_I) = \lambda E(Y_i) = 10\, E(Y_i) = 9,600$$

Queen Mary University of London

# Aggregate claims distributions under excess of loss reinsurance: Example

**Answer (continued):**

**(a) Variance of insurer's aggregate claims**

To find $var(S_I)$ we calculate $E(Y_i^2)$:

$$E(Y_i^2) = \int_0^M x^2 f(x)dx + M^2 P(X_i > M)$$

$$= \left[\frac{0.0005x^3}{3}\right]_0^M + 0.2M^2 = 1,194,666.7$$

so

$$var(S_I) = \lambda E(Y_i^2) = 10 E(Y_i^2) = 11,946,667$$

Slide 21, Week 3

# Aggregate claims distributions under excess of loss reinsurance: Example

**Answer (continued):**

**(a) Coefficient of skewness of insurer's aggregate claims**

To find the coefficient of skewness of the insurer's claims, we calculate $E(Y_i{}^3)$ from:

$$E(Y_i^3) = \int_0^M x^3 f(x)dx + M^3 P(X_i > M)$$

$$= \left[\frac{0.0005x^4}{4}\right]_0^M + 0.2M^3 = 1,638,400,000$$

so

$$Skew(S_I) = E[(S_I - E(S_I))^3] = \lambda E(Y_i{}^3) = 10 E(Y_i{}^3) = 16,384,000,000$$

Coefficient of skewness of $S_I$ = $Skew(S_I) / [var(S_I)]^{3/2}$

$= 16,384,000,000 / (11,946,667)^{3/2} = 0.397$

# Aggregate claims distributions under excess of loss reinsurance: Example

**Answer (continued):**

**(b) Mean of reinsurer's aggregate claims**

To find $E(S_R)$ note that the expected annual aggregate claim amount from the risk is:

$$E(S) = \lambda\, E(X_i) = 10 \times 1,000 = 10,000$$

since $E(S_I) + E(S_R) = E(S)$,

$$E(S_R) = E(S) - E(S_I) = 10,000 - E(S_I) = 400$$

Queen Mary University of London

# Aggregate claims distributions under excess of loss reinsurance: Example

**Answer (continued):**

**(b) Variance of reinsurer's aggregate claims**

To find $var(S_R)$ we calculate $E(Z_i{}^2)$:

$$E(Z_i^2) = \int_M^{2000} (x - M)^2 f(x) dx$$

$$= \int_0^{2000-M} 0.0005 y^2 dy, \qquad y = x - M$$

$$= \left[\frac{0.0005 y^3}{3}\right]_0^{2000-M}$$

$$= 10{,}666.7$$

so

$$var(S_R) = \lambda E(Z_i{}^2) = 10\, E(Z_i{}^2) = 106{,}667$$

Queen Mary University of London

# Aggregate claims distributions under excess of loss reinsurance: Example

**Answer:**

**(b) Coefficient of skewness of reinsurer's aggregate claims**

To find the coefficient of skewness of the reinsurer's claims, we calculate $E(Z_i{}^3)$ from:

$$E(Z_i^3) = \int_M^{2000} (x - M)^3 f(x)dx = \int_0^{2000-M} 0.0005 y^3 dy, \qquad y = x - M$$
$$= 3,200,000$$

so

$$Skew(S_R) = E[(S_R - E(S_R))^3] = \lambda E(Z_i{}^3) = 10\, E(Z_i{}^3) = 32,000,000$$

and

$$\text{Coefficient of skewness of } S_R = Skew(S_R) / [\, var(S_R)\, ]^{3/2}$$
$$= 32,000,000 / (106,667)^{3/2} = 0.92$$

# Aggregate claims distributions under excess of loss reinsurance: Example

**Answer (continued):**

**(c) Variance of $S$, the aggregate claim amount before reinsurance**

We know that $var(S) = \lambda E(X_i{}^2) = 10\, E(X_i{}^2)$, where:

$$E(X_i{}^2) = \int_0^{2000} \frac{x^2}{2000} dx = \frac{4,000,000}{3}$$

So

$$var(S) = 10\,(1{,}333{,}333.3) = 13,333,333$$

# Aggregate claims distributions under excess of loss reinsurance: Example

**Answer (continued):**

(d) **Why is it that $var(S_I) + var(S_R) \neq var(S)$**

- $var(S) = 13,333,333$ is not equal to the sum of $var(S_I) = 11,946,667$ and $var(S_R) = 106,667$ because $S_I$ and $S_R$ are not independent.

- So, we cannot obtain the variance of the reinsurer's payments by subtraction.
- For similar reasons we can't obtain the skewness by subtraction either.

# Reinsurer's aggregate claims under excess of loss reinsurance

For example, suppose that the risk in the example above gave rise to the following eight claim amounts in a particular year:

$$403 \quad 1,490 \quad 1,948 \quad 443 \quad 1,866 \quad 1,704 \quad 1,221 \quad 823.$$

Recall that we have a retention limit of 1,600. Then the observed value of $N$ is eight and the third, fifth and sixth claims require payments from the reinsurer of:

$$348, \quad 266, \quad 104$$

and the reinsurer makes a "payment" of 0 on the other claims.

The observed value of $N_R$ is then 3 and the observed values of $W_1, W_2, W_3$ are 348, 266, 104 respectively.

- Note that the observed value of $S_R$ is the same (*i.e.* 718) under each definition.

# Reinsurer's aggregate claims under excess of loss reinsurance

We then note that $W_i$ has density function given by:

$$g(w) = \frac{f(w+M)}{1-F(M)}, \qquad w > 0$$

To specify the distribution for $S_R$ we need the distribution of $N_R$. This is found as follows. Define:

$$N_R = I_1 + I_2 + \cdots + I_N$$

where $N$ denotes the number of claims from the risk (as usual). And $I_j$ is an **indicator** random variable which takes the value 1 if the reinsurer makes a (non-zero) payment on the $j^{th}$ claim and takes the value 0 otherwise.

- Thus $N_R$ gives the number of payments made by the reinsurer.

# Reinsurer’s aggregate claims under excess of loss reinsurance

Since $I_j$ takes the value 1 only if $X_j > M$ we have:

$$P(I_j = 1) = P(X_j > M) = \pi$$

$$P(I_j = 0) = 1 - \pi$$

In other words, $I_j$ has a $Bin(1,\pi)$ distribution. This means that $NR$ has a compound Poisson distribution (since $N$ is Poisson).

Further, $I_j$ has MGF:

$$M_I(t) = \pi e^t + 1 - \pi$$

and $N_R$ has MGF:

$$M_{NR}(t) = M_N(\log M_I(t))$$

# Reinsurer's aggregate claims under excess of loss reinsurance: Example

Continuing our example now and using

$$S_R = W_1 + W_2 + \cdots + W_{N_R}$$

as our model for $S_R$ we can see that $S_R$ has a compound Poisson distribution with Poisson parameter,

$$\mu = 0.2 \times 10 = 2.$$

Individual claims, $W_i$, have density function:

$$g(w) = \frac{f(w+M)}{1-F(M)} = \frac{0.0005}{0.2} = 0.0025, \quad 0 < w < 400$$

i.e. $W_i$ is uniformly distributed on (0, 400).

$$E(W_i) = 200$$

$$E(W_i^2) = 53,333.33$$

$$E(W_i^3) = 16,000,000$$

# Reinsurer's aggregate claims under excess of loss reinsurance: Example

Multiplying each of these values by 2 (the Poisson parameter of $S_R$) we get:

$E(S_R) = \mu\, E(W_i) = 2\,(200) = 400$
$var(S_R) = \mu\, E(W_i^2) = 2\,(53,333.33) = 106,667$
$skew(S_R) = \mu\, E(W_i^3) = 2\,(16,000,000) = 32,000,000$

So there are then two ways to specify and evaluate the distribution of $S_R$.

# Exam Style Question

The aggregate claims from a risk have a compound Poisson distribution with parameter $\mu$. Individual claim amounts have a Pareto distribution with parameters $\alpha = 3$ and $\lambda = 1,000$.

The insurer of this risk calculates the premium using a premium loading factor of 0.2 (this means they charge 20% in excess of the risk premium).

The insurer is considering effecting excess of loss reinsurance with a retention limit of £1,000.

The reinsurance premium would be calculated using a premium loading factor of 0.3.

The insurer's profit is defined to be the premium charged by the insurer less the reinsurance premium and less the claims paid by the insurer, net of reinsurance.

a) Show that the insurer's expected profit before reinsurance is $100\mu$.

b) Calculate the insurer's expected profit after effecting the reinsurance and hence find the percentage reduction in the insurer's expected profit.

# Exam Style Question

## *Solution*

**a)** Let $S$ be the aggregate claim amount before reinsurance. Then $S = X_1 + X_2 + \cdots + X_N$

where $X$ is $Pareto(\alpha, \lambda)$ and $\alpha = 3$ and $\lambda = 1000$. So:

$$E(X) = \frac{\lambda}{\alpha - 1} = 500$$

$$var(X) = \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)} = 750,000$$

if the Poisson parameter is $\mu$ then:

$$E(S) = 500\mu$$

$$var(S) = \mu E(X^2) = \mu(750,000 + 500^2) = 1,000,000\mu$$

- The insurer's expected profit without reinsurance is equal to premiums minus expected claims.

But if a loading factor of 0.2 is in use, then the total premium is $1.2 \times 500\mu = 600\mu$. So, the expected profit = $600\mu - 500\mu = 100\mu$.

# Exam Style Question

## *Solution*

**b)** Now consider the effect of reinsurance. For each $X_i$ we have $X_i = Y_i + Z_i$, where

$$Y_i = \begin{cases} X_i & X_i < 1,000 \\ 1,000 & X_i \geq 1,000 \end{cases}$$

$$Z_i = \begin{cases} 0 & X_i < 1,000 \\ X_i - 1,000 & X_i \geq 1,000 \end{cases}$$

so, for the reinsurer, total aggregate claims are given by

$$S_R = Z_1 + Z_2 + \cdots + Z_N$$

- This is a compound Poisson distribution, where each $Z_i$ has the distribution given above.
- The reinsurance premium is given by $1.3E(S_R)$, where

$$E(S_R) = E(Z)E(N) = \mu E(Z)$$

# Exam Style Question

## *Solution*

**b) (continued)** Now

$$E(Z) = \int_{1000}^{\infty} (x - 1000) \frac{3 \times 1000^3}{(1000 + x)^4} dx$$

Putting $u = x - 1000$ in this integral we get:

$$E(Z) = \int_0^{\infty} u \frac{3 \times 1000^3}{(2000 + u)^4} du = \left(\frac{1000}{2000}\right)^3 \int_0^{\infty} u \frac{3 \times 2000^3}{(2000 + u)^4} du$$

Recognising this integral as the mean of the *Pareto*(3, 2000) distribution, we see that

$$E(Z) = \left(\frac{1}{2}\right)^3 \times \frac{2000}{3 - 1} = 125$$

So, $E(S_R) = E(Z)E(N) = \mu E(Z) = 125\mu$

and reinsurance premium is $1.3\ E(S_R) = 1.3 \times 125\mu = 162.5\mu$

Queen Mary University of London

# Exam Style Question

## *Solution*

**b) (continued)** The expected profit with reinsurance is

$$E\,(\text{Gross premium} - \text{reinsurance premium} - \text{net claim amounts})$$

$$= 600\mu - 162.5\mu - E(S - S_R)$$

where net claim amounts, $S - S_R$, is what the insurer pays net of reinsurance. We have

$$E(S - S_R) = E(S) - E(S_R) = 500\mu - 125\mu = 375\mu$$

So, the expected profit is $62.5\mu$.

And the percentage reduction in the expected profit (which was $100\mu$ without reinsurance) is 37.5%.

# The individual risk model

- Under this model a portfolio consisting of a fixed number of risks is considered.
- It will be assumed that:
  1. these risks are independent
  2. claim amounts from these risks are not (necessarily) identically distributed random variables
  3. the number of risks does not change over the period of insurance cover
- As before, aggregate claims from the portfolio will be denoted by $S$, so

$$S = Y_1 + Y_2 + \cdots + Y_n$$

  where $Y_j$ denotes the claim amount under the $j^{th}$ risk and $n$ denotes the number of risks.
- It is possible that some risks will not give rise to claims, i.e. some of the observed values of $Y_j$ may $= 0$.
- In fact in most forms of insurance most policies would not give rise to any claims during a given year.
- This approach is referred to as an *individual risk model* because it is considering the claims arising from each individual policy.

# The individual risk model

## *Assumptions*

- For each risk, we make the following assumptions:
  i. the number of claims from the $j^{th}$ risk, $N_j$, is either 0 or 1.
  ii. the probability of a claim from the $j^{th}$ risk is $q_j$.
- The **binary** assumption for the number of claims makes this model particularly appropriate for life insurance policies (since people can **die** at most once during a given period).
- If a claim occurs under the $j^{th}$ risk, the claim amount is denoted by the random variable $X_j$. Let $F_j(x)$ denote the distribution function of $X_j$, $\mu_j$ denotes the mean of $X_j$ and $\sigma_j{}^2$ the variance of $X_j$.
- The assumption saying the number of claims is binary is quite restrictive. It means that a maximum of one claim from each risk is allowed for in the model. This includes risks such as a one-year term assurance, but excludes many types of **general insurance** policies.
- For example, there is usually no restriction on the number of claims that could be made in a policy year under household contents insurance.

Queen Mary
University of London

# The individual risk model

## *Differences compared with the collective risk model*

There are three important differences between this model and the collective risk model.

1) The **number of risks in the portfolio** has been specified. In the collective risk model, there was no need to specify this number, nor to assume that it remained fixed over the period of insurance cover, not even when it was assumed that $N \sim Bin(n, q)$.

2) The **number of claims from each individual risk** has been restricted. There was no such restriction in the collective risk model.

3) It is assumed that **individual risks are independent**. In the collective risk model it was individual claim amounts that were independent.

Queen Mary
University of London

# The individual risk model

## *Mean and variance of aggregate claims in the individual risk model*

- From (i) and (ii) we have $N_j \sim Bin(1, q_j)$ (the number of claims from the $j^{th}$ risk). Thus, the distribution of $Y_j$ is compound binomial, with individual claim amount random variable $X_j$. From our compound binomial formulae, we get

$$E(Y_j) = E(N_j)\,E(X_j) = q_j \mu_j$$

$$var(Y_j) = E(N_j)\,var(X_j) + var(N_j)\,[\,E(X_j)\,]^2 = q_j \sigma_j^{\,2} + q_j(1 - q_j)\,\mu_j^{\,2}$$

- $S$ is the sum of $n$ independent compound binomial random variables. The distribution of $S$ can be stated only when the compound binomial random variables are identically distributed, as well as independent. It is possible, but complicated, to compute the distribution function of $S$ under certain conditions. However, it is relatively easy to find the mean and variance of $S$.

$$E(S) = E\left[\sum_{j=1}^{n} Y_j\right] = \sum_{j=1}^{n} E(Y_j) = \sum_{j=1}^{n} q_j \mu_j$$

$$var(S) = var\left[\sum_{j=1}^{n} Y_j\right] = \sum_{j=1}^{n} var(Y_j) = \sum_{j=1}^{n} (q_j \sigma_j^2 + q_j(1 - q_j)\mu_j^2)$$

# The individual risk model

## *Special case*

- In the special case when $Y_j$ is a sequence of independent and identically distributed random variables, then for each policy the values of $q_j$, $\mu_j$ and $\sigma_j^{\ 2}$ are identical, say, $q$, $\mu$ and $\sigma^2$
- Since $F_j(x)$ is independent of $j$ we can refer to it simply as $F(x)$. Hence $S$ is compound binomial, with binomial parameters $n$ and $q$ and individual claims have distribution function $F(x)$.
- In this special case, it reduces to the collective risk model and it can be seen that:

$$E(S) = nq\mu$$
$$var(S) = nq\sigma^2 + nq(1-q)\mu^2$$

# The individual risk model

## *Example*

The probability of a claim arising on any given policy in a portfolio of 1,000 one-year term assurance policies is 0.004. Individual claim amounts have a $Gamma(5, 0.002)$ distribution.

Find the mean and variance of the aggregate claim amount.

**Answer:**

Here, we have an individual risk model. Let $S$ be a r.v. representing aggregate claim amounts, then

$$S = Y_1 + Y_2 + \cdots + Y_n$$

Where the distribution of $Y_j$ is compound binomial, and as $Y_j$ are iid,

$$E(S) = nq\mu$$

$$var(S) = nq\sigma^2 + nq(1-q)\mu^2$$

We then find $q$, $\mu$ and $\sigma^2$ and plug into the formulae above.

# The individual risk model

## *Example*

**Answer (continued):**

$$\mu = 5 / 0.002 = 2,500$$

$$\sigma^2 = 5 / 0.002^2 = 1,250,000$$

$$E(S) = nq\mu = 1,000 \times 0.004 \times 2,500 = 10,000$$

$$\begin{aligned} var(S) &= nq\sigma^2 + nq(1-q)\mu^2 \\ &= 1,000 \times [0.004 \times 1,250,000 + 0.004 \times 0.996 \times 2,500^2] \\ &= (£5,468)^2 \end{aligned}$$

Queen Mary University of London

# Parameter variability

- So far, we have studied risk models assuming that the **parameters**, that is the moments and, in some cases, even the distributions, of the number of claims and of the amount of individual claims, are **known** with certainty.
- In general, these parameters would not be known but would have to be estimated from appropriate sets of data.
- We will now look at how the models we have introduced can be extended to allow for parameter variability / uncertainty.
- We will do this by looking at a series of examples. Most, but not all, of these examples will consider uncertainty in the claim number distribution, since this, rather than the individual claim amount distribution, has received more attention in actuarial literature.
- All our examples will be based on claim numbers having a Poisson distribution.

# Variability in a heterogeneous portfolio

- Consider a portfolio consisting of *n* independent policies.
- The aggregate claims from the $i^{th}$ policy are denoted by the random variable $S_i$, where $S_i$ has a compound Poisson distribution with parameters $\lambda_i$.
- The CDF of the individual claim amounts distribution is $F(x)$. Notice that, for simplicity, the CDF of the distribution of individual claim amounts, $F(x)$ is assumed to be identical for all the policies.
- In this example the distribution of individual claim amounts, *i.e.* $F(x)$, is assumed to be known but the values of the Poisson parameters (*i.e.* the $\lambda_i$'s) are not known.
- We assume that the $\lambda_i$'s are independent random variables with the same (known) distribution.
- This means that if **a policy is chosen at random** from the portfolio, then it is assumed that the **Poisson parameter** for the policy is **not known** but that probability statements can be made about it.
- For example, "there is a 50 % chance that its Poisson parameter lies between 3 and 5". It is important to understand that the Poisson parameter for a policy chosen from the portfolio is a fixed number.
- The problem is that this number is not known.

# Variability in a heterogeneous portfolio

## *Example 1*

**Question:**

Suppose that the Poisson parameters of policies in a portfolio are not known but are equally likely to be 0.1 or 0.3.

a) Find the mean and variance (in terms of $m_1$ and $m_2$) of the aggregate claims from a policy chosen at random from the portfolio.

b) Find the mean and variance (in terms of $m_1, m_2$ and $n$) of the aggregate claims from the whole portfolio.

**Note:** We will look at the motivation for such a model and then look at the solutions.

# Variability in a heterogeneous portfolio

## *Example 1*

**Motivation:**

- It may be helpful to think of this as a model of part of a motor insurance portfolio.
- The **policies** in the whole portfolio have been **subdivided** according to their values for rating factors such as “age of driver”, “type of car” and even “past claims experience”.
- The policies in the part of the portfolio being considered have identical values for these rating factors. However, there are some factors, such as “**driving ability**” that **cannot easily be measured** and so they cannot explicitly be considered.
- It is supposed that some of the policyholders in this part of the portfolio are “good” drivers and the remainder are “bad” drivers.
- The individual claim amount distribution is the same for all drivers, but **“good” drivers** make fewer claims (0.1 pa on average) than **“bad” drivers** (0.3 pa on average).
- It is assumed that it is known, possibly from national data, that a policyholder in this part of the portfolio is equally likely to be a ”good” driver or a ”bad” driver but that it cannot be known whether a particular policyholder is a “good” driver or a “bad” driver.

# Variability in a heterogeneous portfolio

## *Example 1*

**Answer:**

Let $\lambda_i$, $i = 1, 2, \ldots, n$ be the Poisson parameter of the $i^{th}$ policy in the portfolio. $\lambda_i$'s are independent and identically distributed random variables, each with the following distribution:

$$P(\lambda_i = 0.1) = 0.5$$

$$P(\lambda_i = 0.3) = 0.5$$

so

$$E(\lambda_i) = 0.2$$

$$var(\lambda_i) = 0.01$$

# Variability in a heterogeneous portfolio

## *Example 1*

**Answer:**

**i)** The moments of $S_i$ can be calculated by conditioning on the value of $\lambda_i$. Since $S_i|\lambda_i$ has a straightforward compound Poisson distribution we can write

$$E(S_i) = E(E(S_i|\lambda_i)) = E(\lambda_i m_1) = 0.2\, m_1$$

$$var(S_i) = E(var(S_i|\lambda_i)) + var(E(S_i|\lambda_i))$$

$$= E(\lambda_i m_2) + var(\lambda_i m_1)$$

$$= 0.2\, m_2 + 0.01\, m_1^2$$

# Variability in a heterogeneous portfolio

## *Example 1*

**Answer (continued):**

**ii)** The random variables $S_i$'s are independent and identically distributed, each with the distribution of $S_i$ given in part (i). Hence the results in part (i) above can be used to write

$$E\left[\sum_{i=1}^{n} S_i\right] = nE(S_i) = 0.2nm_1$$

$$var\left[\sum_{i=1}^{n} S_i\right] = nvar(S_i) = 0.2nm_2 + 0.01nm_1^2$$

# Variability in a homogeneous portfolio

- Consider again a portfolio of $n$ independent policies.
- The aggregate claims from a single policy have a compound Poisson distribution with parameters $\lambda$ and the CDF of the individual claim amounts random variable is $F(x)$.
- The Poisson parameters are the same for all policies in the portfolio.
- If the value of $\lambda$ were known, the aggregate claims from different policies would be independent of each other.
- It is assumed that the value of $\boldsymbol{\lambda}$ **is not known**, possibly because it changes from year to year, but that there is some indication of the probability that $\lambda$ will be in any given range of values.
- As in the previous example, it is assumed for simplicity that there is **no uncertainty** about the moments or distribution of the **individual claim amounts** (i.e. about $F(x)$).
- The uncertainty about the value of $\lambda$ can be modelled by regarding $\boldsymbol{\lambda}$ **as a random variable** (with a known distribution).

# Variability in a homogeneous portfolio

## *Example 2*

**Question:**

Suppose that the Poisson parameters of policies in a portfolio are not known but are equally likely to be 0.1 or 0.3.

i. Find the mean and variance (in terms of $m_1$ and $m_2$) of the aggregate claims from a policy chosen at random from the portfolio.

ii. Find the mean and variance (in terms of $m_1, m_2$ and $n$) of the aggregate claims from the whole portfolio.

# Variability in a homogeneous portfolio

## *Example 2*

**Answer:**

Using the same notation as before, let $S_i$ denote the aggregate claims from the $i^{th}$ policy in the portfolio. The situation can be summarised as follows:

- The random variables $S_i$'s are independent and identically distributed, each with a compound Poisson distribution with parameters $\lambda$ and $F(x)$.
- The random variable $\lambda$ has the following distribution:

$$P(\lambda = 0.1) = 0.5$$

$$P(\lambda = 0.3) = 0.5$$

From this:

$$E(\lambda) = 0.2$$

$$var(\lambda) = 0.01$$

# Variability in a homogeneous portfolio

## *Example 2*

**Answer:**

**i)** The moments of $S_i$ can be calculated by conditioning on the value of $\lambda$. Since $S_i|\lambda$ has a straightforward compound Poisson distribution we can write:

$$E(S_i) = E(E(S_i|\lambda)) = E(\lambda m_1) = 0.2\, m_1$$

$$var(S_i) = E(var(S_i|\lambda)) + var(E(S_i|\lambda))$$

$$= E(\lambda m_2) + var(\lambda m_1)$$

$$= 0.2\, m_2 + 0.01\, m_1^2$$

Queen Mary University of London

# Variability in a homogeneous portfolio

## *Example 2*

**Answer (continued):**

**ii)** The random variables $S_i$'s are independent and identically distributed, each with the distribution of $S_i$ given in part (i). Hence the results in part (i) can be used to write:

$$E\left[\sum_{i=1}^{n} S_i\right] = nE(S_i) = 0.2nm_1 \quad , \quad \text{and} \quad var\left[\sum_{i=1}^{n} S_i\right] = E\left[var\left(\sum_{i=1}^{n} S_i|\lambda\right)\right] + var\left[E\left(\sum_{i=1}^{n} S_i|\lambda\right)\right]$$

$$= E[n\lambda m_2] + var[n\lambda m_1]$$

$$= 0.2nm_2 + 0.01n^2m_1^2$$

Compare the answers above with answers from Example 1 (variability in a heterogeneous portfolio):

$$E\left[\sum_{i=1}^{n} S_i\right] = nE(S_i) = 0.2nm_1$$

$$var\left[\sum_{i=1}^{n} S_i\right] = nvar(S_i) = 0.2nm_2 + 0.01nm_1^2$$

# Variability in a homogeneous portfolio

## *Example 2*

**Observations:**

- It is useful to compare the answers to Example 2 to those of Example 1.
- The values of the mean are in all cases the same, as are the variances when a single policy is considered (part (i)).
- The difference occurs when variances for more than one policy are considered (part (ii)), in which case the second example gives the greater variance.
- It is important to understand the differences (and similarities) between the two examples.
- A practical situation where the second example could be appropriate would be a portfolio of **policies insuring buildings in a certain area**.
- The number of claims could depend on, among other factors, the weather during the year.
- An unusually high number of **storms** resulting in a high expected number of claims (*ie* a high value of $\lambda$) and vice versa **for all the policies together**.