Insurance and Reinsurance

To insure a risk, it must be financially quantifiable and the person taking out insurance must have a stake in it.

Insurers would also prefer a large number of similar risks with low correlation to allow pooling, the probability of the risk occurring to be low to ease administration, an ultimate limit on the liability to help quantify the risk, and the elimination of moral hazard to stop selection against the scheme.

General insurance typically has fixed periods of cover, during which multiple claims of varying amounts are made. Claims taking a short or long period to settle are called short-tail and long-tail respectively. There are four main types.

Liability — This covers legally mandated compensation to third parties for death, bodily injury or property damage caused by negligence.
Legal expenses are usually covered, illegal acts are usually not.
Claims can be restricted by an excess, a maximum or an aggregate maximum.

- Employers liability - exposure to harmful substances/working practic
- Motor third party liability - most countries make this compulsory.
- Product liability
- Professional liability
- Public liability - covers everything not covered above.

Property — This covers loss or damage to property which the insured owns.
The benefit is usually the amount lost, subject to limits or excesses.
Insurance companies sometimes insist on replacement to deter fraudulent claims.

- Building - Mainly fire, also explosion, theft, storm, flood
- Movable - Mainly theft, accidental damage
- Motor - Accidental/maliscious damage, fire, theft
- Motor fleet
- Marine - Jettison, piracy, perils of the sea.

Fixed Benefit — This covers health and personal injury insurance where fixed sums are paid to help cover costs for specific injuries

Financial — This covers the loss of money

Business interruption cover	Losses because the business was not run
Fidelity guarantee	Fraud or embezzlement
Pecuniary	Bad debts or other third party failure

There are four types of reinsurance considered, which can be split into two classes

Proportional — The direct writer and reinsurer share premiums and the cost of claims

Quota share	The proportion is the same for all risks
Surplus	The proportions vary

Non proportional — The direct writer pays a premium and the reinsurer pays for claims portions falling in a particular layer (say, over 50k)

XOL	The reinsurer pays for individual claims £50k in excess of £50k
Stop loss	The reinsurer pays for a group of policies 90% of excess when claims exceed 105% premiums.

In some ways, insurers writing policies with excesses can be thought of as XOL reinsurers of these people. In practice, the expenses associated with each claim make this view unsustainable.

Basic Distributions

The most useful models for insurance tend to be positively skewed with long tails. We will particularly consider

Exponential	$X \sim Exp(\lambda)$	$f(x) = \lambda e^{-x}$	
Gamma	$X \sim Ga(\alpha, \lambda)$	$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\lambda x)$	$x>0$
Normal	$X \sim N(\mu, \sigma^2)$		
Pareto	$X \sim Pa(\alpha, \lambda)$	$f(x) = \frac{\alpha \lambda^\alpha}{(\lambda+x)^{\alpha+1}}$	$x>0$ $\alpha>0$ λ
Generalised Pareto		$f(x) = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\Gamma(k)} \frac{\lambda^\alpha x^{k-1}}{(\lambda+x)^{\alpha+k}}$	$x>0$ $\alpha>0$ λ
Log normal	$X \sim LN(\mu, \sigma^2)$		
Weibull	$X \sim W(c, \gamma)$	$f(x) = c\gamma x^{\gamma-1} \exp(-c x^\gamma)$	$x>0$ $c>0$ γ
Burr			

Details of these can be found in the formulae and tables book

Let a fixed number of claims be distributed by $X_i \sim Exp(\lambda_i)$, with the parameter for each policyholder distributed by $\lambda_i \sim Ga(\alpha, \delta)$. Then

$$
\begin{aligned}
f_X(x) &= \int f_\lambda(\lambda) f_{X|\lambda}(x|\lambda)\, d\lambda \\
&= \int_0^\infty \delta^\alpha / \Gamma(\alpha)\, \lambda^{\alpha-1} \exp(-\delta x)\, \lambda \exp(-\lambda x)\, d\lambda \\
&= \delta^\alpha / \Gamma(\alpha)\; \Gamma(\alpha+1) / (x+\delta)^{\alpha+1} \quad (a\ G(\alpha+1, x+\delta) \text{ integral} = 1) \\
&= \alpha \delta^\alpha / (x+\delta)^{\alpha+1}
\end{aligned}
$$

And thus $X_i \sim Pa(\alpha, \delta)$

Note that although the related problem $X_i \sim Exp(\lambda)$, $\lambda \sim Ga(\alpha, \delta)$ also gives $X_i \sim Pa(\alpha, \delta)$, in that problem they are not independant, so the overall distribution of the portfolio changes.

Letting claims be distributed by $X_i \sim Ga(k, \lambda_i)$, $\lambda_i \sim Ga(\alpha, \delta)$ gives the generalised Pareto distribution. These are called mixture distributions.

Compound Distributions

A compound distribution S is the sum of independant identically distributed random variables $\{X_i\}_{i=1}^{N}$ where the number summed N is a discrete independant random variable. They have distribution functions, means, variances and moment generating functions as foll

$$G(x) = P(S \leq x)$$
$$= \sum_{n=0}^{\infty} P(N=n)\, P(S \leq x \mid N=n) = \sum_{n=0}^{\infty} P(N=n)\, F^{n*}(x)$$

where $F^{n*}(x)$ is the n-fold convolution of $F(x)$, given by $F^{n*}(x) = \int F^{n*}(x)\, f(x-z)\, dz$

$$E(S) = E(E(S \mid N)$$
$$= E(N\, E(X_i)) = E(N)\, E(X_i)$$

$$Var(S) = E(Var(S \mid N)) + Var(E(S \mid N))$$
$$= E(N\, Var(X_i)) + Var(N\, E(X_i)) = E(N)\, Var(X_i) + E^2(X_i)\, Var(N)$$

$$M_S(t) = E(E(e^{tS} \mid N)$$
$$= E(M_X(t)^N)$$
$$= E(e^{\log M_X(t) N}) = M_N(\log M_X(t))$$

There are three important distributions

Compound Poisson $\quad N \sim P(\lambda)$

$E(S) = \lambda E(X) \quad Var(S) = \lambda E(X)^2$

$M_S(t) = \exp\{\lambda(M_X(t) - 1)\}$

Sums of compound poisson distributions are also compound poisson.

Compound Binomial $\quad N \sim b(n, q)$

$E(S) = nq\, E(X) \quad Var(S) = nq\, E(X^2) - nq^2 E^2(X)$

$M_S(t) = (q\, M_X(t) + 1 - q)^n$

Compound Negative Binomial $\quad N \sim NB(k, p)$

$E(S) = \frac{kq}{p} E(X) \quad Var(S) = \frac{kq}{p} E(X^2) - \frac{kq}{p^2} E^2(X)$

$M_S(t) = p^k / (1 - q M_X(t))^k$

Individual Risk Model

Consider a portfolio with a fixed number of independant risks, each of whom claim at most once for an amount X_i with probability q_i
Let $F_i(x)$, μ_i and σ_i^2 be the distribution factor, mean and variance X_i. Then

$$E(S) = \sum q_i \mu_i$$
$$Var(S) = \sum q_i \sigma_i^2 + q_i(1-q_i)\mu_i^2$$

If the risks are identically distributed, we get back to the compound binomial

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

Finding Distributions

To use distributions, we must estimate their parameters. There are three methods.

Method of moments: Equate population and sample moments $E_\theta(X^m) = \frac{1}{n}\sum x_i^m$ and solve for parameters

Method of percentiles: Choose a percentile for each parameter and equate the population and sample percentiles $F_\theta(p_i) =$ sample $F_\theta(p_i)$ and solve for parameters

Maximum Likelyhood: Set $L(\theta)$ to be the likelyhood of the data, given parameters θ. Solve $d/d\theta\, L(\theta) = 0$ or $d/d\theta \log L(\theta) = 0$ to find the most likely value of the parameters. This is the most useful method when dealing with discontinuities.

If we know nothing about the aggregate claims distribution, using the method of moments with the normal or gamma distributions is usually useful.

Suppose the claim numbers follow $p_r = (a + b/r)\, p_{r-1}$ (ie poisson, binomial or negative bin and claims are positive integers following a known but non tractable distribution. Then the distribution $G(x)$ of the aggregate claim amounts follows the recursion

$$g_0 = p_0$$
$$g_r = \sum (a + bj/r)\, f_j\, g_{r-j}$$

To prove this, note that $f_r^{n*} = \sum_{j=1}^{r-1} f_j f_{r-j}^{(n-1)*}$

and that $E(X_1 \mid \sum_{i=1}^{n} X_i = r) = r/n = \sum_{j=1}^{r-1} j \frac{f_j f_{r-j}^{(n-1)*}}{f_r^{n*}}$. Then

$$\begin{aligned} g_r &= \sum_{n=1}^{\infty} p_n f_r^{n*} \\ &= (a+b)p_1 f_r + \sum_{n=2}^{\infty} (a + b/n)\, p_{n-1} f_r^{n*} \\ &= (a+b)p_1 f_r + \sum_{n=2}^{\infty}\sum_{j=1}^{r-1} (a + bj/r)\, f_j f_{r-j}^{(n-1)*} p_{n-1} \\ &= (a+b)p_1 f_r + \sum_{j=1}^{r-1} (a + bj/r) f_j\, g_{r-j} \quad = \sum_{j=1}^{r} (a + bj/r)\, f_j\, g_{r-j} \end{aligned}$$

Generalised Linear Models

In linear regression, we assume a linear relationship between one independent variable and the response variable, with a normally distributed error term. With generalized linear models we relax these restrictions. Firstly, the error term.

A random variable Y belongs to the exponential family if its density has form

$$f_Y(y;\theta,\varphi) = \exp\left[\frac{y\theta - b(\theta)}{a(\varphi)} + c(y,\varphi)\right] \qquad a,b,c \text{ functions}$$

This has mean $b'(\theta)$ and variance $a(\varphi)b''(\theta)$. A number of common distributions belong to the exponential family

Normal $\quad f_Y(y;\theta,\varphi) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right) = \exp\left[\frac{2y\mu-\mu^2}{2\sigma^2} - \frac{y^2}{2\sigma^2} - \frac{1}{2}\log 2\pi\sigma^2\right]$

$\theta=\mu \quad \varphi=\sigma^2 \quad a(\varphi)=\varphi \quad b(\theta)=\theta^2/2 \quad c(y,\varphi)=-\frac{1}{2}(y^2/\varphi + \log 2\pi\varphi)$

Poisson $\quad f_Y(y;\theta,\varphi) = \frac{\mu^y e^{-\mu}}{y!} = \exp\left[y\log\mu - \mu - \log y!\right]$

$\theta=\log\mu \quad \varphi=1 \quad a(\varphi)=1 \quad b(\theta)=e^\theta \quad c(y,\varphi)=-\log y!$

Binomial $\quad$ First we must divide the standard binomial by n, so $y\in[0,1]$

$f_Y(y;\theta,\varphi) = \binom{n}{ny}\mu^{ny}(1-\mu)^{n-ny} = \exp\left[n\left(y\log\frac{\mu}{1-\mu} + \log 1-\mu\right) + \log\binom{n}{ny}\right]$

$\theta=\log\frac{\mu}{1-\mu} \quad \varphi=n \quad a(\varphi)=\frac{1}{\varphi} \quad b(\theta)=\log(1+e^\theta) \quad c(y,\varphi)=\log\binom{\varphi}{\varphi y}$

Gamma $\quad$ First we change the parameters from α,λ to $\alpha,\frac{\alpha}{\mu}$

$f_Y(y;\theta,\varphi) = \frac{\lambda^\alpha}{\Gamma(\alpha)}y^{\alpha-1}e^{-\lambda y} = \exp\left[-\frac{\alpha y}{\mu} - \alpha\log\mu + (\alpha-1)\log y + \alpha\log\alpha - \log\Gamma(\alpha)\right]$

$\theta=-\frac{1}{\mu} \quad \varphi=\alpha \quad a(\varphi)=\frac{1}{\varphi} \quad b(\theta)=-\log(-\theta) \quad c(y,\varphi)=(\varphi-1)\log y + \varphi\log\varphi - \log\Gamma$

Log normal $\quad$ Simply take logs of the data and apply to the normal distribution

Given a value θ and an error term φ, we have determined that the value y has a distribution $f_Y(y;\theta,\varphi)$. However, we still need to find θ for each individua

Let $\theta_i = g^{-1}(\eta_i)$ where g is the link factor and η_i is a linear predictor.

The link factor must be differentiable and invertable to fit a model and make useful predictions. Furthermore, it is useful to have link factors which give an appropriate range for θ. There are a number of canonical link functions for distribution

Normal	identity	$g(\mu) = \mu$
Poisson	log	$g(\mu) = \log(\mu)$
Binomial	logit	$g(\mu) = \log\left(\frac{\mu}{1-\mu}\right)$
Gamma	Inverse	$g(\mu) = 1/\mu$

The linear predictor should be chosen to fit the data being modeled. For example

age	$\beta_0 + \beta_1 x$
sex	α_i
age + sex	$\alpha_i + \beta x$
age + sex + age.sex	$\alpha_i + \beta_i x$
age * sex	$\alpha_i + \beta_i x$
age + age^2	$\beta_0 + \beta_1 x + \beta_2 x^2$

Having determined all this, we substitute everything into the appropriate place and use MLEs to estimate the parameter values α, β, etc.

Accuracy in GLMs

A saturated model is one in which the number of parameters and of observations is equal, so the observations are fitted perfectly. The scaled deviance is defined as twice the difference between the log likelyhood of the model and the saturated model

$$S = \sum_{i=1}^{n} \frac{(y_i - \hat{\theta}_i)^2}{\sigma^2} \quad \text{where } \hat{\theta}_i \text{ is the fitted value for the current model.}$$

This value has approximately a χ^2 distribution (exactly if it is normally distributed). At the 5% level, χ^2_v is significant if it is over $2v$. Thus, with nested models, adding an additional set of parameters significantly improves the result if $S_1 - S_2 > 2$(number parameters added).

It is also useful to analyse the residules, of which there are two

Pearson residules	$y_i - \hat{\theta}_i / \sqrt{\operatorname{Var} \hat{\theta}_i}$	y observed, $\hat{\theta}$ fitted
Deviance residules	$\operatorname{sign}(y_i - \hat{\theta}_i)\sqrt{\frac{(y_i - \hat{\theta}_i)^2}{\sigma^2}}$	

We should examine a histogram of the residules to detect anomalies.

We should also compare each parameter to ensure it is at least two times greater than its standard error. If not, it probably has little effect and can be discarded.

Reinsurance Distributions

The purpose of reinsurance is to reduce the amount and variability of claims that the insurer must pay out by splitting the claims distribution between insurer and reinsurer.

In proportional reinsurance, the insurer pays a fixed portion of the claim αX. The distribution of αX and $(1-\alpha)X$ is simply that of X scaled by α or $1-\alpha$.

Under XOL, the claims distributions are

Insurer $Y = \begin{cases} X & X < M \\ M & X \geq M \end{cases}$ Reinsurer $Z = \begin{cases} 0 & X < M \\ X - M & X \geq M \end{cases}$

For the insurer, although this is less tractable, it can be treated normally.

$$\begin{aligned} E(Y) &= \int_0^M x f(x)\,dx + M\,P(X > M) \\ &= E(X) - \int_0^\infty z\, f(z+M)\,dz \\ M_Y(t) &= \int_0^M e^{tx} f(x)\,dx + e^{tM} P(X > M) \end{aligned}$$

For the reinsurer, we must ignore claims they have no record for.

$$\begin{aligned} P(Z < z) &= P(X < z + M \mid X > M) \\ &= F(z+M) - F(M) \,/\, 1 - F(M) \\ g(z) &= f(z+M) \,/\, 1 - F(M) \qquad z > 0 \end{aligned}$$

All the standard results follow.

Miscellanious Reinsurance

There are formulae to help calculate moments of truncated normal and lognormal distributions. They are as follows

$$\text{Normal} \quad \int_L^U x f_X(x)\, dx = \mu\left[\Phi\left(\frac{U-\mu}{\sigma}\right) - \Phi\left(\frac{L-\mu}{\sigma}\right)\right] - \sigma\left[\phi\left(\frac{U-\mu}{\sigma}\right) - \phi\left(\frac{L-\mu}{\sigma}\right)\right]$$

$$\text{Log Normal} \quad \int_L^U x^k f_X(x)\, dx = e^{k\mu + \frac{1}{2}k^2\sigma^2}\left[\Phi\left(\frac{\log U - \mu}{\sigma} - k\sigma\right) - \Phi\left(\frac{\log L - \mu}{\sigma} - k\sigma\right)\right]$$

If we want to consider the effect of inflation on the claims, where M is not inflation linked, simply substitute X by kX. Note that the new mean is not k times the old mean, etc.

Estimation of parameters cannot be done by method of moments, since the models are no longer tractable. It can still be done by method of percentiles of maximum likelyhood estimation.

Basic Ruin Theory

Consider a portfolio with starting capital U, earning c per unit time.
Let $N(t)$ be the number of claims by time t, X_i the amount of each claim and $S(t) = \sum_{i=1}^{N(t)} X_i$ the total amount paid by time t. Then at t we have capital

$$U(t) = U + ct - S(t).$$

Define the probability of ruin in continuous time as

$$\psi(U) = P[\,U(t) < 0 \text{ for some } 0 < t < \infty\,]$$
$$\psi(U,t) = P[\,U(\tau) < 0 \text{ for some } 0 < \tau < t\,]$$

Define the probability of ruin in discrete time as

$$\psi_h(U) = P[\,U(t) < 0 \text{ for some } t = h, 2h, \ldots\,]$$
$$\psi_h(U,t) = P[\,U(\tau) < 0 \text{ for some } \tau = h, 2h, \ldots, t-h, t\,]$$

If $0 < t_1 \leq t_2 < \infty$ and $0 \leq U_1 \leq U_2$ we have the following relationships

$$\psi(U_1,t) \geq \psi(U_2,t) \qquad \psi_h(U_1,t) \geq \psi_h(U_2,t)$$
$$\psi(U,t_1) \leq \psi(U,t_2) \leq \psi(U) \qquad \psi(U,t_1) \leq \psi_h(U,t_2) \leq \psi_h(U)$$
$$\lim_{t\to\infty} \psi(U,t) = \psi(U) \qquad \lim_{t\to\infty} \psi_h(U,t) = \psi_h(U)$$

$$\psi(U,t) \geq \psi_h(U,t)$$
$$\lim_{h\to 0^+} \psi_h(U,t) = \psi(U,t)$$

Poisson Process

Let $N(t)$ follow a poisson process with parameter λ. Let $N(0)=0$, and

$$P(N(t+h) < r \mid N(t)=r) = 0$$
$$P(N(t+h) = r \mid N(t)=r) = 1-\lambda h + o(h)$$
$$P(N(t+h) = r+1 \mid N(t)=r) = \lambda h + o(h)$$
$$P(N(t+h) > r+1 \mid N(t)=r) = o(h)$$

When $s<t$, the claims during $(s,t]$ are independant of those up to time s.

It is called a poisson process because for fixed t, $N(t) \sim Po(\lambda t)$.
We prove this by writing $P(N(t+h)=n)$ in terms of $P(N(t)=n)$ and $P(N(t)=n-1)$.
This gives $\frac{d}{dt} P(N(t)=n) = \lambda P(N(t)=n-1) - \lambda P(N(t)=n)$.
We know $P(N(t)=-1)=0$, so repeatedly substituting and solving gives the result.
$P(N(t)=0) = e^{-\lambda t}$, $P(N(t)=1) = \lambda t e^{-\lambda t}$, $P(N(t)=2) = \frac{(\lambda t)^2}{2} e^{-\lambda t}$, etc...

The inter arrival times are independant with exponential distributions, parameter λ.
This is trivial for the first arrival $P(N(t)=0) = e^{-\lambda t}$, and the rest follows by lack of memo

If we now suppose that X_i are independant identically distributed random variables independant of $N(t)$, then $S(t)$ is a compound poisson process with Poisson parameter λ. Again, for fixed t, $S(t) \sim Comp\ Po(\lambda t)$. Thus

$$M_S(r) = \exp\{\lambda t (M_X(r)-1)\}$$
$$E(S) = \lambda t E(X)$$
$$E(S^n) = \lambda t E(X^n)$$

As we now know the expected claims, set the premium to be slightly more than them

$c = (1+\theta)\lambda E(X)$ $\qquad$ $\theta > 0$ is the premium loading factor.

Lundberg's Inequality

Define the adjustment coefficient R as the only positive root of the equation

$$E\left(e^{R(S_i - c)}\right) = 1 \qquad S_i = \text{aggregate claims in one unit of time}$$

Then Lundberg's inequality states

$$\psi(u) \leq \exp\{-Ru\}$$

For the compound poisson process, R is the solution to

$$M_X(R) = 1 + (1+\theta)m_1 R$$

It has the following bounds (found by truncating e^{Rx} in $M_X(R)$)

$$\frac{1}{M}\log(c/\lambda m_1) < R < 2(c - \lambda m_1)/\lambda m_2$$

We can also think about things qualitatively. Thus changing λ has no effect on ultimate ruin because we are only rescaling time, but changing u does unless we rescale the claim size with it.

Reinsurance and Ruin

A reinsurance arrangement could be considered optimal if it minimises the possibility of ruin. This is hard to measure, so we'll look for one maximising R.

Under proportional reinsurance, the insurers premium is

$$((1+\theta) - (1+\varepsilon)(1-\alpha))\,\lambda_1 m_1$$

Use this to solve for R as a function of α

$$1 + (\theta + \alpha + \varepsilon\alpha - \varepsilon)\,{}^{1}/_{\lambda}\,R = \int_0^\infty e^{Rx}\,({}^{\alpha}/_{\lambda})\,e^{-\alpha/\lambda\, x}\,dx$$
$$= \frac{1}{1 - \alpha/\lambda\, R}$$

$$\Rightarrow \qquad R = \frac{(\theta + \alpha + \varepsilon\alpha - \varepsilon) - \alpha}{\alpha/\lambda\,(\theta + \alpha + \varepsilon\alpha - \varepsilon)}$$

With $\varepsilon < \theta$ it is possible to push all risk to the reinsurer and make a profit.
With $\varepsilon = \theta$ it is possible to push all risk to the reinsurer and make nothing.
With $\varepsilon > \theta$, we need to solve by finding $d/d\alpha\, R = 0$ with $d^2/d\alpha^2\, R < 0$.
Note that even if reinsurance can reduce the risk, insurers may opt for maximum return

Under XOL, the maths is a lot uglier. Suppose the insurers premium is

$$(1+\theta)\,\lambda_1 m_1 - (1+\varepsilon)\,\lambda\,E(Z)$$

This gives the equation for R as

$$\lambda + (1+\theta)\,\lambda m_1 - (1+\varepsilon)\,\lambda\,E(Z) = \lambda\left[\int_0^M e^{Rx} f(x)\,dx + e^{RM}[1 - F(M)]\right]$$

This can only really be solved numerically. The relations $\varepsilon \sim \theta$ above still hold though, for qualitative reasons.

Rating General Insurance Business

First calculate the pure risk premium, using past data projected to accomodate changes. Then add loadings for commission, expenses, profit and other contingencies to give the office premium.

Internal data is preferable as it will be more relevant and contain more detail.
It should have enough claims for the data to be credible, the period should be long enough to establish trends and the data should be recent enough to be relevant.
If there is insufficient or unsuitable internal data, external data in the form of aggregate market statistics or competitors rates must be used.

We tend to project claim frequency, cost per claim and exposure per policy separately
When projecting, we will need to make suitable adjustments for unusually heavy/lig
experience, exceptional claims and changes in the policy, in reinsurance premium and in investment income.

No claims Discounts

No claims discounts are systems where insured people pay lower premiums during periods in which they have not made claims. They are modelled as markov chains with different states for different levels of no claims discount.

The interesting aspect is considering the probability of not claiming when an accident has occured to keep the no claims discount. This changes the transition probabilities in the chain.

The second aspect is asking whether good drivers get markedly better discounts than bad drivers. This depends on how the no claims discount is set up.

Two Player Zero Sum Games

These are represented as matrices of losses for player A given both players choices

	AI	AII	AIII	AIV
B1	5	1	0	1
B2	3	3	4	4
B3	0	1	5	3

A strategy is said to be dominated by another if it gives an equal or worse result whatever the opponant does. A IV is dominated by A II. It can be removed.

A minimax strategy is one in which you choose the option minimising your maximum loss. A II and B2 are such strategies

A pair of strategies is in equilibrium if no player can do better by changing their strategy. This configuration is called a saddle point. A II and B2 are in equilibri
However if the circled three was a two, they would not be.

The main strategy used is to randomise the choice in such a way that no matte what the opponant chooses, the expected loss is the same minimum value.

Suppose now between B and As choice, A is given information about what B picked, and instead of options he has a choice of strategies using the information to follow. Simple probabilities give expected loss which we can put into a matrix

	d_1	d_2	d_3
$\theta 1$	1	$\frac{1}{2}$	0
$\theta 2$	0	$\frac{1}{2}$	2

If we believe B is trying to maximise the loss, follow the techniques above. Otherwise use minimax or guess what Bs distribution is and minimise loss according to that.

Baysion Statistics

Bayes Theorem states that if $B_1, \dots, B_k$ partition a somple space S, then for any event A in S.

$$P(B_r \mid A) = \frac{P(A \mid B_r)\, P(B_r)}{P(A)} \qquad \text{where} \quad P(A) = \sum_{i=1}^{k} P(A \mid B_i)\, P(B_i)$$

Baysian statistics extends this to probility distributions. Suppose we take a sample $\underline{X}$ from a distribution $f(x, \theta)$ in order to estimate θ. Suppose further that before taking the sample we believed θ had distribution $f(\theta)$. Now

$$f(\theta \mid \underline{X}) = \frac{f(\underline{X} \mid \theta)\, f(\theta)}{f(\underline{X})} \qquad \text{where} \quad f(X) = \int f(\underline{X} \mid \theta)\, f(\theta)\, d\theta$$

It is often simpler to use a statistic and usually this gives the same result

$$f(\theta \mid \bar{X}) = \frac{f(\bar{X} \mid \theta)\, f(\theta)}{f(\bar{X})} \qquad \text{where} \quad f(X) = \int f(\bar{X} \mid \theta)\, f(\theta)\, d\theta$$

The new distribution of θ, $f(\theta \mid X)$ is called the posterior distribution. If it is in the same form as the prior, it is also called the conjugate prior.

To get an estimate of θ from its distribution, we choose a loss function and find the value of θ that minimises it. There are three main ones

Quadratic loss	$L(g(x), \theta) = [g(x) - \theta]^2$	Gives the mean
Absolute error loss	$L(g(x), \theta) = \lvert g(x) - \theta \rvert$	Gives the median
All or nothing loss	$L(g(x), \theta) = \delta(g(x), \theta)$	Gives the mode

Baysian Credibility

Let $\bar{X}$ be an estimate based soley from the data from the risk. Let μ be an estimate based soley on collateral data (from risks similar to but not identical to the risk). The credibility premium formula is

$$Z\bar{X} + (1-Z)\mu \qquad Z \in (0,1)$$

In general terms, the more data from the risk, the higher Z will be; the more reliable the collateral data, the lower Z will be. We will consider two ways of estimating Z - the first is baysian credibility.

Choose a prior distribution for the parameter under consideration, and use the sample (or more simply, an appropriate sample statistic) to get the posterior distribution. Now use a loss function to estimate θ.

At this stage, it is normally possible to rewrite the formula estimating θ to put it into the form of the credebility premium formula, and thus get Z.

This approach works provided we can sensibly put the risks into prior and posterior distributions, and that in the end we get a formula in the form $Z\bar{X} + (1-Z)\mu$.

Empirical Bayes Credibility Theory

$X_1, \ldots, X_{n+1}$ are independant identically distributed random variables depending on unknown fixed θ
Define $m(\theta) = E(X_{n+1} | \theta)$ and $s^2(\theta) = V(X_{n+1} | \theta)$. Given $\underline{X} = X_1, \ldots, X_n$ we will try to estimate $m(\theta)$.

Let our estimate be of the form $a_0 + a_1 X_1 + \ldots + a_n X_n$ and choose the result minimising $E((m(\theta) - a_0 - a_1 X_1 - \ldots - a_n X_n)^2)$.

Differentiating with respect to $a_0, a_1, \ldots, a_n$ gives (taking $X_0 = 1$)

$$E(X_k (m(\theta) - a_0 - \sum a_j X_j)) = 0$$

$$\Rightarrow \quad a_0 = E(m(\theta)) (1 - \sum a_j)$$
$$a_k = \frac{(1-\sum a_j) E(m^2(\theta)) - a_0 E(m(\theta)}{E(s^2(\theta))} \qquad [\text{note all } a_k \text{ are equal}]$$

Now choosing

$$Z = n V(m(\theta)) / n V(m(\theta)) + E(s^2(\theta))$$

we get an estimate of

$$Z \sum X_j / n + (1 - Z) E(m(\theta))$$

All that remains is to find values of $E(m(\theta))$, $V(m(\theta))$ and $E(s^2(\theta))$.
Suppose we have a table of N different risks over a period of n years. Then

$$E(m(\theta)) = \bar{X}$$
$$E(s^2(\theta)) = \frac{1}{N(n-1)} \sum_{i=1}^{N} \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2$$
$$V(m(\theta)) = \frac{1}{N-1} \sum_{i=1}^{N} (\bar{X}_i - \bar{X})^2 - \frac{1}{Nn(n-1)} \sum_{i=1}^{N} \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2$$

These formulae are in the 'Formula and tables' book.

Empirical Bayes Credibility Theory 2

Let $Y_1, \dots, Y_n$ be random variables representing claims in successive years and let $P_1, \dots, P_n$ be the risk volume. Set $X_j = Y_j / P_j$. Define $m(\theta) = E(X_j | \theta)$ and $s^2(\theta) = P_j V(X_j | \theta)$. Following the same method as before, we get

$$Z = \sum P_j V(m(\theta)) \Big/ \sum P_j V(m(\theta)) + E(s^2(\theta))$$

which gives an estimate of

$$Z \sum P_j X_j / \sum P_j + (1 - Z) E(m(\theta)).$$

Estimating the parameters is a little more complex than before

$$\begin{array}{ll}
E(m(\theta)) & \bar{X} \\
E(s^2(\theta)) & \frac{1}{N(n-1)} \sum_{i=1}^{N} \sum_{j=1}^{n} P_{ij} (X_{ij} - \bar{X}_i)^2 \\
V(m(\theta)) & \frac{1}{\sum \bar{P}_i (1 - \bar{P}_i / \bar{P})} \sum \sum P_{ij} (X_{ij} - \bar{X})^2 - \frac{Nn-1}{N(n-1)} \frac{1}{\sum \bar{P}_i (1 - \bar{P}_i / \bar{P})} \sum \sum P_{ij} (X_{ij} - \bar{X}_i)^2
\end{array}$$

Run-off Triangles

These arise in insurance where it takes some time before the extent of claims to be paid are known. Data is presented as follows.

Development year		0	1	2	3
Accident year	2000	100	179	201	203
	2001	97	160	191	-
	2002	121	181	-	-
	2003	190	-	-	-

The data in the table above is cumulative and does not account for inflation. If we were concerned about inflation, the first thing we would have to do is inflation adjust it. Namely, put it in non cumulative form and multiply the values in each diagonal appropriately for the inflation between then and now.

We can now estimate the values marked by dashes. The simplest technique is the chain ladder. Firstly put the table into cumulative form. Then consider the first pair of development years. The ratio for these is the sum of all known values in the later year divided by the sum of the corresponding values in the earl
Our estimates for the dashes are the values in the earlier year times the ratio.

We may then need to inflation unadjust the result by finding the non cumulativ table and multiplying the diagonals by the expected future inflation.

To check the values, we could pretend we don't know any of the data beyond development year zero, and estimate it. Comparing the non cumulative tables resulting should give a good idea of discrepancies.

The average cost per claim method splits the triangle into two - one triangle for claim numbers, the second for average cost per claim.

For each triangle we use the following method. Look at each incomplete accident year in turn, from earliest to latest. Use the latest development year to estimate the ultimate value using the average ratio over each completed year. Now calculate the ratio each other value has to the ultimate value. We now have ultimate claim numbers and ultimate average cost per claim. Multiply them together to get ultimate loss estimates.

The Bornhuetter - Furguson method estimates the reserves that have to be kept assuming the ultimate claims are Premium × Estimated Loss Ratio. Essentially, follow the usual chain ladder method. Then find f, the ratio between the ultimate loss and the latest loss under chain ladder. Then

$$\text{Further claims development} = \text{Premium} \times \text{Estimated Loss Ratio} \times (1 - 1/f).$$

Adding this to the amount we have paid so far gives the estimate of the ultimate liability.