الجمهورية الجزائرية الديمقراطية الشعبية
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
SAAD DAHLAB UNIVERSITY BLIDA-1
Faculty of Sciences
Department of Mathematics

MASTER'S THESIS

In: Mathematics
Option: Stochastic and Statistical Models

THEME:

Bayesian claims reserving methods in non life insurance

Presented by:
KACIMI SELMA

Argued before the Jury:

M. TAMI Omar	Saad Dahlab University Blida-1	Chairman
M. FRIHI Redhouane	Saad Dahlab University Blida-1	Supervisor
M.RASSOUL Abdelaziz	Saad Dahlab University Blida-1	Examiner

July 9, 2024

ACKNOWLEDGEMENT

Gratitude , first of all elhamdoulilah , I thank God, the almighty, for giving me the strength and the health to complete this dissertation, as well as the audacity to overcome all difficulties.

My deepest thanks and gratitude go to my parents, for their love , encouragement , guidance and always being there for me.

I would like to thank my supervisor Dr FRIHI Redouane, for his valuable advice and help throughout the work period.

My gratitude goes also to the members of the jury Dr TAMI Omar and RASSOUL Abdelaziz who accepted to examine and evaluate my work.

DEDICATION

To my dear father and mother, whom I can never thank enough, without you the journey would not have been easy for me but with your unconditional support, love I made it, thank you for guiding me to become the person I am.

To my beloved siblings: abdeallah, shaimaa and younes.

ملخص

تركز هذه الأطروحة على طرق حجز المطالبات البايزي في التأمين على غير الحياة، وتستخدم شركات التأمين هذه الأساليب الإحصائية المتقدمة لتقدير الاحتياطيات المالية للمطالبات المستقبلية وتجنب المخاطر. المتقدمة هذه التقنيات الإحصائية لتقدير الاحتياطيات المالية للمطالبات المستقبلية وتجنب المخاطر على الرغم من أن الأساليب التقليدية توفر أساساً، إلا أنها لا تستطيع في كثير من الأحيان معالجة تعقيد عوامل الخطر المعاصرة وعدم القدرة على التنبؤ بها. وعلى العكس من ذلك،تدمج الأساليب البايزية بين المعرفة المسبقة والبيانات الجديدة يوفر هذا النهج تنبؤات أكثر دقة وإطار عمل مرن، مما يعزز قدرة شركات التأمين على الاستعداد للمطالبات المستقبلية المطالبات المستقبلية وتخفيف المخاطر .

كلمات مفتاحية

النهج البايزي ،المطالبات، تامين ،خطر ،النهج الكلاسيكي

Abstract

this thesis focuses on Bayesian claims reserving methods in non-life insurance, comparing them with classical methods.Insurance companies use these advanced statistical techniques to estimate financial reserves for future claims and avoid risk. Although traditional approaches offer a basis, they frequently cannot address the intricacy and unpredictability of contemporary risk factors. Conversely, Bayesian methods, integrate prior knowledge and new data. This approach offers more accurate predictions and a flexible framework, enhancing insurers' ability to prepare for future claims and mitigate risks.

Key words

bayesian approache , claims , insurance ,risk , classical approache .

Résumé

Cette thèse porte sur les méthodes bayésiennes de provisionnement des sinistres en assurance non-vie, en les comparant aux méthodes classiques, Les compagnies d'assurance utilisent ces techniques statistiques avancées pour estimer les réserves financières pour les sinistres futurs et pour éviter les risques. Les compagnies d'assurance utilisent ces techniques statistiques avancées pour estimer les réserves financières pour les sinistres futurs et éviter les risques. Bien que les approches traditionnelles offrent une base, elles ne peuvent souvent pas répondre à la complexité et à l'imprévisibilité des facteurs de risque contemporains. A l'inverse, les méthodes bayésiennes intègrent les connaissances antérieures et les nouvelles données. Cette approche offre des prédictions plus précises et un cadre flexible, améliorant ainsi la capacité des assureurs à se préparer aux futurs sinistres et à atténuer les risques. à se préparer aux sinistres futurs et à atténuer les risques.

Mot-clés

approche bayésienne , revendications , assurance , risque , approche classique.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Insurance mathematics does not appear to be a difficult subject to the general public. In actuality, everyone must deal with insurance-related issues at some point in their lives.

Thats why insurance companies play an important role in society, the idea of insurance is part of our civilized world.

It is based on the mutual trust of the insurer and the insured. In the world of non-life insurance which includes sectors such as automobile, property, liability, and health insurance the ability to manage and predict risk is paramount the word Risk.

The essence of non-life insurance lies in its risk management capabilities. Insurers must anticipate and prepare for future claims based on a myriad of factors that we mentioned.
Traditional reserving methods have provided a foundation for this process, but as the industry evolves, so too must the methods it employs.

The complexity and unpredictability of risks in the modern world necessitate a more flexible and comprehensive approach to claims reserving.

A critical aspect of this thesis is the exploration of risk models and the calculation of ruin probabilities. In insurance, the concept of ruin refers to the scenario where an insurer's liabilities exceed its assets, leading to insolvency.

Understanding the probability of ruin is essential for assessing an insurer's financial health and stability. Bayesian models offer sophisticated tools for estimating these probabilities, incorporating a range of risk factors and scenarios.

This approach not only enhances the accuracy of risk assessments but also provides valuable insights into the potential impact of extreme events.

Bayesian methods represent a paradigm shift in how risks are assessed and managed in the insurance industry. Named after Reverend Thomas Bayes, Bayesian statistics provide a powerful framework for updating the probability of a hypothesis as new evidence or information becomes available.

Unlike classical approaches, Bayesian methods incorporate prior knowledge and expert judgment with historical data, creating a more adaptable and dynamic model.
This integration enables a deeper understanding of underlying uncertainties and offers a prob-

abilistic interpretation of potential outcomes.

By accounting for various sources of uncertainty, Bayesian methods offer a more nuanced and robust framework for predicting future claims.

This thesis also investigates the use of Bayesian techniques outside of the domain of fundamental reserving, such as risk modeling and the determination of ruin probabilities.

When evaluating the entire risk exposure and financial stability of an insurer, these ideas are essential. We are able to obtain a thorough grasp of the risk profile of an insurer by analyzing the distribution of the total amount of claims.This in turn affects capital allocation plans, risk management procedures, and strategic decision-making.

The comparative analysis of classical and Bayesian methods is central to this thesis. While classical methods provide a solid foundation, the flexibility and depth of Bayesian approaches offer significant enhancements.

The objective is not merely to compare these methods but to explore how they can be integrated to provide a more robust and adaptive framework for claims reserving.
This synthesis aims to leverage the strengths of both approaches, providing insurers with the tools they need to navigate the complexities of modern risk management.

Not only is it our goal to compare these approaches, but we also want to investigate how they may be combined to create a more resilient and flexible framework for claims reserving.

By combining the best features of the two methods, this synthesis hopes to give insurers the resources they need to handle the challenges of contemporary risk management. Not only is it our goal to compare these approaches, but we also want to investigate how they may be combined to create a more resilient and flexible framework for claims reserving.

By combining the best features of the two methods, this synthesis hopes to give insurers the resources they need to handle the challenges of contemporary risk management. our work is presented as follow :
In Chapter 1, we introduce the foundational concepts of probability, providing a comprehensive overview of probability spaces, random variables, and common probability distributions .

In Chapter 2, we explore the Bayesian approach, demonstrating its principles and methods for updating probabilities as new information becomes available.

In Chapter 3, we develop a model of risk and ruin probability, integrating both frequentist and Bayesian methods to assess the likelihood of financial insolvency in insurance contexts.

In Chapter 4, we apply the developed models to real-world data, demonstrating their practical utility in predicting aggregate claims and assessing risk within an insurance framework.
We close this study with a General Conclusion.

CHAPTER 1

PRELIMINARIES OF PROBABILITIES

1.1 Introduction

In statistics, engineering, science, and finance, the probability theory is a formal way to express the inherent randomness and uncertainty, Probability theory provides a mathematical framework for quantifying uncertainty and modeling random events.

It begins with defining a sample space, which encompasses all possible outcomes of a random experiment. Within this space, events are subsets of outcomes, and probabilities are assigned to these events to measure their likelihood.

Kolmogorov's axioms establish the foundational rules for probability, ensuring consistency and coherence in calculations. Conditional probability and the concept of independence allow for the analysis of events in relation to each other.

Probability distributions describe how probabilities are distributed over the values of a random variable. Mastering these preliminaries is crucial for applying probability theory to fields such as statistics, finance, insurance, and risk management.

1.1.1 Probability Space

A probability space is a mathematical construct designed for the purpose of dealing with random phenomena and discussing probabilities associated with outcomes.
[4] A probability space is a mathematical construct designed to deal with random phenomena and discuss the probabilities associated with outcomes. A probability space is denoted by a triple where $(\Omega, \mathcal{F}, P)$:

- **Sample Space (Ω) :** The sample space Ω corresponds to the set of possible outcomes of the experiment. For example, in a coin toss, $\Omega = \{\text{Heads}, \text{Tails}\}$.
- **Sigma-Algebra $(\mathcal{F})$:** A collection of subsets of Ω that is closed under countable unions, complementation, and includes the sample space itself. We refer to these subsets as events. For example, in rolling a die we have $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{F}$ could include subsets like $\{1, 3\}$, $\{2, 4, 5\}$, and Ω itself.

- **Probability Measure (P) :** A formula that provides a probability to every event in $\mathcal{F}$.
  The following requirements must be met by this function :
  1. Non-negativity : $P(A) \geq 0$ for any event $A \in \mathcal{F}$.
     $Normalization$:P$(\Omega) = 1$.
  2. Countable Additivity : For any countable sequence of mutually exclusive events A1, A2, A3 the probability of their union is equal to the sum of their probabilities :
  3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$.

1.2 Probability Distribution and Random Variables

1.2.1 Random Variable

Definition 1 *Let $(\Omega, \mathcal{F}, P)$ be a probability space. A random variable is a function X defined on Ω with values in a set E, such that for every subset A of E, the event $X^{-1}(A) \in \mathcal{F}$. The random variable is called real when E is a subset of $\mathbb{R}$.*
For any subset A of $\mathbb{R}$, the event $X^{-1}(A)$ is denoted as $\{X \in A\}$ or simply $(X \in A)$:

$$X^{-1}(A) = (X \in A) = \{\omega \in \Omega : X(\omega) \in A\}.$$

When $A = \{x\}$, the event $X^{-1}(\{x\})$ is more simply denoted as $(X = x)$:

$$(X = x) = \{\omega \in \Omega : X(\omega) = x\}.$$

1.2.2 Discrete Random Variable

Definition 2 *Let $X : \Omega \to E$ be a function, where $(\Omega, \mathcal{F}, P)$ is a probability space. We say that X is a discrete random variable if:*

1. *$X(\Omega)$ is finite or countable, and*
2. *for all $x \in X(\Omega)$, the preimage $X^{-1}(\{x\}) \in \mathcal{F}$.*

When $E = \mathbb{R}$, the random variable X is said to be real. It is said to be finite if $X(\Omega)$ is finite.

1.2.3 Continuous Random Variables:

Continuous random variables can represent any value within a specified range or interval and can take on an infinite number of possible values [18].

1.3 Probability Distributions

An explanation of how probabilities are dispersed throughout a random variables value can be found in its probability distribution.

1.3.1 Discrete Probability Distributions

A Probability Mass Function (PMF) is used to describe the distribution of discrete random variables.

- Probability Mass Function (PMF) :

$$P(X = x_i) = p_i, \quad \text{where } 0 \leq p_i \leq 1 \quad \text{and} \quad \sum_i p_i = 1$$

1.3.2 Continuous Probability Distributions

A Probability Density Function (PDF) is used to describe the distribution of continuous random variables.

- Density Function (PDF) :

  The PDF $f(x)$ gives the relative likelihood that a continuous variable X takes on a value near x.

$$P(a \leq X \leq b) = \int_a^b f(x)\, dx$$

The PDF must satisfy :

$$f(x) \geq 0 \quad \text{and} \quad \int_{-\infty}^{+\infty} f(x)\, dx = 1$$

1.3.3 Expectation and Variance

- **Expectation (Mean) :**

  The mean or expectation of an unknown quantity X is given by

$$E(X) = \sum_i x_i P(X = x_i) \quad \text{if } X \text{ is discrete}$$

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx \quad \text{if} X \text{is continuous.}$$

  The mean is the center of mass of the distribution [19]

1.3.4 Discrete and Continuous Probability Distributions

There are some usual Probability Distributions

Table 1.1: Discrete Distributions

Distribution	**Probability Function (PMF)**	**Expected Value (E(X))**	**Variance (Var(X))**
Bernoulli	$P(X = x) = p^x(1-p)^{1-x}$	p	$p(1-p)$
Binomial	$P(X = k) = \binom{n}{k} p^k(1-p)^{n-k}$	np	$np(1-p)$
Geometric	$P(X = k) = (1-p)^{k-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Poisson	$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$	λ	λ

Table 1.2: Continuous Distributions

Distribution	Probability Function (PDF)	Expected Value (E(X))	Variance (Var(X))
Uniform	$f(x) = \begin{cases} \frac{1}{b-a} & \text{a} \leq x \leq \text{b} \\ 0 & \text{otherwise} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal (Gaussian)	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2
Exponential	$f(x) = \lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gamma	$f(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)}$	$\frac{k}{\lambda}$	$\frac{k}{\lambda^2}$
Beta	$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
Log-Normal	$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}$	$e^{\mu+\sigma^2/2}$	$(e^{\sigma^2} - 1)e^{2\mu+\sigma^2}$
Chi-Square	$f(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2}$	k	$2k$

1.4 Classical (frequentist) statistical approach

1.4.1 Fundamental Assumptions

The frequentist approach to statistical analysis is used in the majority of statistics books for beginners, and is based on the following concepts:

- The population's numerical features, or parameters, are unknown but fixed and unchanging constants.
- Long-run relative frequency is the standard interpretation of probabilities.
- Statistical procedures are judged by how well they perform in the long run over an infinite number of hypothetical repetitions of the experiment is used to evaluate them. [20]

1.5 Parametric Estimation

From sampling data, frequentist statistics utilize estimate techniques to deduce the unknown parameters of a probability distribution the two primary technique for estimating are:

1.5.1 Maximum Likelihood Estimation (MLE):

Definition 3 *Let $X_1, X_2, X_3, \ldots, X_n$ be a random sample from a distribution with a parameter θ. Suppose that we have observed $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$.*

1. *if X_i are discrete, then the likelihood function is defined as:*

$$L(x_1, x_2, \ldots, x_n; \theta) = P_{X_1 X_1, X_2}(x_1, x_2, \ldots, x_n; \theta)$$

2. *If X_i's are jointly continuous, then the likelihood function is defined as:*

$$L(x_1, x_2, \ldots, x_n; \theta) = f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n; \theta)$$

A maximum likelihood estimator (MLE) of the parameter θ, shown by $\hat{\theta}_{ML}$, is a random variable $\hat{\theta}_{ML} = \hat{\theta}_{ML}(X_1, X_2, \ldots, X_n)$ whose value when $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$ is given by $\hat{\theta}_{ML}$.

1.5.2 Method of Moments Estimation:

Definition 4 *Let* $\mathbf{x} = (x_1, \ldots, x_n)$ *be i.i.d realizations (samples) from probability mass function* $p_X(x;\theta)$ *(if* X *is discrete), or from density* $f_X(x;\theta)$ *(if* X *is continuous), where* θ *is a parameter (or vector of parameters).*
We then define the ***method of moments*** *(MoM) estimator* $\hat{\theta}_{MoM}$ *of* $\theta = (\theta_1, \ldots, \theta_k)$ *to be a solution (if it exists) to the* k *simultaneous equations where, for* $j = 1, \ldots, k$*, we set the* j^{th} *(true) moment equal to the* j^{th} *sample moment:*

$$\mathbb{E}[X] = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\vdots$$

$$\mathbb{E}[X^k] = \frac{1}{n}\sum_{i=1}^{n} x_i^k$$

1.5.3 Hypothesis testing:

The frequentist approach is many hypothesis tests are succinctly summarized in this table:

Table 1.3: Hypothesis Testing Summary Table

Type of Test	**Null Hypothesis (H_0)**	**Alternative Hypothesis (H_A)**	**Test Statistic**	**Decision Rule**
Z-Test	$\mu = \mu_0$	$\mu \neq \mu_0$	$Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$	Reject H_0 if
T-Test	$\mu = \mu_0$	$\mu \neq \mu_0$	$T = \frac{\bar{X}-\mu_0}{s/\sqrt{n}}$	Reject H_0 if
Paired T-Test	$\mu_D = 0$	$\mu_D \neq 0$	$T = \frac{D}{s_D/\sqrt{n}}$	Reject H_0 if
Chi-Square Test	Variables are independent	Variables are not independent	$\chi^2 = \sum \frac{(O_i-E_i)^2}{E_i}$	Reject H_0 if
ANOVA (Analysis of Variance)	All group means are equal	At least one group mean is different	$F = \frac{\text{variance between groups}}{\text{variance within groups}}$	Reject H_0 if $F > F_{\alpha,df1,df2}$
Proportion Test	$p = p_0$	$p \neq p_0$	$Z = \frac{p-p_0}{\sqrt{p_0(1-p_0)/n}}$	Reject H_0 if
Two-Proportion Test	$p_1 = p_2$	$p_1 \neq p_2$	$Z = \frac{p_1-p_2}{\sqrt{p(1-p)(\frac{1}{n_1}+\frac{1}{n_2})}}$	Reject H_0 if
F-Test	$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$	$F = \frac{s_1^2}{s_2^2}$	Reject H_0 if $F > F_{\alpha,df1,df2}$

CHAPTER 2

BAYESIAN INFERENCE

2.1 Introduction

To convey our knowledge and opinions about unknown quantities, we frequently use probabilities informally.

But probability can also be formally used to express information: more specifically, it can be demonstrated mathematically that probabilities may be used to numerically represent a collection of rational beliefs, that probability and information are related, and that Bayes' rule offers a logical way to update beliefs in the light of new information [1].

Bayesian inference is a statistical method that combines prior knowledge with new evidence to update our beliefs about a parameter, it offers a flexible framework for decision making in the face of uncertainty. by utilizing prior distributions , Bayesian inference integrates previous beliefs, in contrast to frequentist techniques, which only use observable data.

This chapter will provide you with an extensive understanding of Bayesian inference, including its theoretical foundations and practical applications.

2.2 What's Bayes

2.2.1 Meaning of Bayes statistics

Bayesian statistics refers to a statistical approach named after 18th-century British mathematician Thomas Bayes.
Bayes' rule provides a rational method for updating beliefs in light of new information[1].

Bayesian statistical analysis relies on Bayes' theorem, which tells us to update prior beliefs about parameters and hypotheses in light of data, to yield posterior beliefs.

Bayes theorem itself is utterly uncontroversial and follows directly from the conventional definition of conditional probability, it is adequate to take into account the stylized, graphical representation of Bayesian inference that follows[8]:

$$\textbf{prior beliefs} \rightarrow \textbf{data} \rightarrow \textbf{posterior beliefs}$$

2.2.2 Formula and explanation

Let H be the set of all possible hypotheses and E be the evidence event.

Suppose $\{H_1, ..., H_k\}$ is a partition of H, $P_r(H) = 1$ and E is some specific event. The axioms of probability imply the following [10]:

1. **Rule of conditional probability:**
   The conditional probability $P(E|H_k)$ is the probability of the event E given the hypothesis H_k.
$$P_r(E|H_k) = \frac{P_r(E \cap H_k)}{P_r(H_k)}$$

2. **Rule of total probability:**
$$\sum_{k=1}^{K} P_r(H_k) = 1$$

3. **Rule of marginal probability:**
$$P_r(E) = \sum_{k=1}^{K} P_r(E \cap H_k)$$
$$\sum_{k=1}^{K} P_r(E|H_k)P_r(H_k).$$

4. **Bayes rule:**
$$P_r(H_j|E) = \frac{P_r(E|H_j)P_r(H_j)}{P_r(E)}$$
$$= \frac{P_r(E|H_j)P_r(H_j)}{\sum_{k=1}^{K} P_r(E|H_k)P_r(H_k)}$$

Here's a breakdown of each part of the formula:

- $P_r(H_j|E)$ is the probability of hypothesis H_j given the evidence E. This is the quantity we want to compute, it is the posterior probability.
- $P_r(E|H_j)$ is the probability of observing the evidence E given that H_j is true. This is the likelihood of the evidence under the hypothesis H_j.
- $P_r(H_j)$ is the prior probability of H_j, representing our initial belief in the probability of H_j being true before considering any evidence E.
- $P_r(E)$ is the total probability of observing the evidence E, also called the marginal likelihood.

Example 1

- ***Hypothesis*** *H: an insurance claim is fraudulent. $P_r(H) = 0.05$ (5% of insurance claims are fraudulent).*
- ***Evidence*** *E: the claim involves a high reimbursement amount. $P_r(E|H) = 0.8$ (80% of fraudulent claims involve high reimbursement amounts).*

- ***Evidence in general:*** *the overall probability that any claim involves a high reimbursement amount is* $P_r(E) = 0.15$.

The objective is to find the probability that an insurance claim is fraudulent given that it involves a high reimbursement amount, $P_r(H|E)$.

1. *Prior Probability:*
$$P_r(H) = 0.05$$

2. *Likelihood:*
$$P_r(E|H) = 0.8$$

3. *Total probability of evidence* $P_r(E)$*:*
$$P_r(E) = 0.15$$

4. *Posterior probability* $P_r(H|E)$*:*
$$\begin{aligned} P_r(H|E) &= \frac{P_r(E|H)P_r(H)}{P_r(E)} \\ &= \frac{0.8 \times 0.05}{0.15} = \frac{0.04}{0.15} \\ &\simeq 0.2667. \end{aligned}$$

given that an insurance claim involves a high reimbursement amount, there is approximately a 26.67% chance it is fraudulent. This shows how Bayes rule updates the likelihood of fraud based on this evidence.

2.2.3 The Philosophy of the Bayesian Approach

posterior law is the foundation for Bayesian statistics. When X as a is observed, the posterior law Can be understood probabilistic summation of the information that is known about θ. The update of a prior information by observing X is some accomplished by the Bayesian approach through $\pi(\theta|X)$.

The diagram below summarizes the Bayesian approach in the context of inferential parametric statistics.
the prior rule demonstrates the stochastic modeling of the x; as realizations of random variables X_i which is feature of inferential statistic it also demonstrates the stochastic modeling of the prior knowledge accessible on the parameter θ[16].

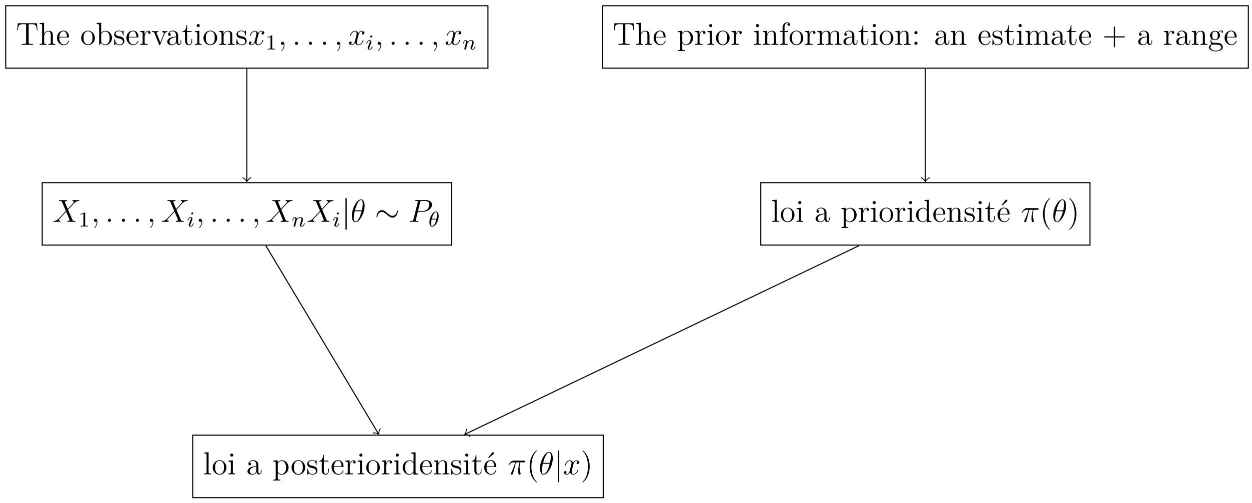

Figure 2.1: the Bayesian approach in the context of inferential parametric statistics

2.3 A prior probability

2.3.1 Definition and Concept

prior probability is the probability of an event determined before any new evidence or data is considered.
It is based on pre-existing information such as historical data, logical analysis, or subjective judgment. This initial probability serves as the baseline for updating beliefs as new evidence becomes available, following Bayes theorem.

Regarding the unknown parameter Θ, we already know certain things. This information is expressed as prior distribution , which is law on the parameter space θ labeled π. The parameter θ becomes a random variable and we note it as $\theta \sim \pi$.

Definition 5 *Let $(f_\theta)_{\theta\in\Theta}$ be a family of probability densities parameterized by θ. A prior distribution π is a probability distribution (probability density) on θ [9].*

2.3.2 How to choose the prior distribution :

One key distinction between frequentist statistics and Bayesian statistics is the selection of prior distribution , which is a basic step in the latter method.

Different points of view may serve as motivation for the various options:

- A decision made by the statistician using their intuition or historical experiences.
- Decision based on how feasible the computations are.
- Decision based on the wish to avoid adding any new details that would skew the estimate.

2.4 Modelling information a priori

2.4.1 Informative priors

Concept of conjugate distributions :

A family F of distributions on Θ is said to be conjugate for the law $f(x/\theta)$ if for all $\pi \in F$, the posterior distribution $\pi(.|x)$ also belongs to F.

The main benefit of conjugate families is that computations are made simpler. Another benefit is that interpretation is frequently made considerably simpler because the law is updated through its parameters [9].

Some examples of conjugate laws[2]:

Table 2.1: Conjugate probability distributions

$f(x \mid \theta)$ likelihood	$\pi(\theta)$ a priori	$\pi(\theta \mid x)$ a posteriori
Normal $\mathcal{N}(\theta, \sigma^2)$	Normal $\mathcal{N}(\mu, \tau^2)$	Normal $\mathcal{N}(\rho(\sigma^2\mu + \tau^2 x), \rho\sigma^2\tau^2)$ $\rho^{-1} = \sigma^2 + \tau^2$
Binomial $\mathcal{B}(n, \theta)$	Beta $\mathcal{B}e(\alpha, \beta)$	Beta $\mathcal{B}e(\alpha + x, \beta + n - x)$
Poisson $\mathcal{P}(\theta)$	Gamma $\mathcal{G}(\alpha, \beta)$	Gamma $\mathcal{G}(\alpha + x, \beta + n)$
Normal $\mathcal{N}(\mu, 1/\theta)$	Gamma $\mathcal{G}a(\alpha, \beta)$	Gamma $\mathcal{G}(\alpha + 0.5, \beta + (\mu - x)^2/2)$
$\mathcal{G}(\nu, \theta)$	$\mathcal{G}(\alpha, \beta)$	$\mathcal{G}(\alpha + \nu, \beta + x)$
Negative Binomial $\mathcal{N}eg(m, \theta)$	Beta $\mathcal{B}e(\alpha, \beta)$	Beta $\mathcal{B}e(\alpha + m, \beta + x)$
Multinomial $\mathcal{M}_k(\theta_1, \ldots, \theta_k)$	Dirichlet $\mathcal{D}(\alpha_1, \ldots, \alpha_k)$	Dirichlet $\mathcal{D}(\alpha_1 + x_1, \ldots, \alpha_k + x_k)$

2.4.2 Non-informative priors

In Bayesian statistics, non-informative priors - also referred to as uninformative or flat priors - are used to indicate a situation in which there is little prior knowledge about the parameters being estimated
They are made with the intention of having the least possible impact on the posterior distribution, letting the data drive the conclusions. Here are a few salient features and examples:

- **Jeffreys Priors**: a substitute was put forth by Jeffreys in 1960.
  Jeffreys prior law is given by:
  $$\pi(\theta)\alpha\sqrt{I(\theta)}$$
  Where $I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\log f(x|\theta)\right]$, $I(\theta)$ is the fisher information which measures the amount of information that an observable random variable carries about the unknown parameter upon which the likelihood depends.

Example 2 *if* $x \sim B(n, P)$*:*

$$f(x|P) = \binom{n}{x} P^x (1-P)^{n-x}$$

$$\frac{\partial^2 \ln f(x|P)}{\partial P^2} = \frac{x}{P^2} + \frac{n-x}{(1-P)^2}$$

So,

$$I(P) = n\left[\frac{1}{P} + \frac{1}{1-P}\right]$$

So Jeffreys law for this model is:

$$\pi^*(P)\alpha[P(1-P)]^{-\frac{1}{2}}$$

- **Uniform Priors**: Uniform priors are considered non-informative as they do not favor any specific value of the parameter. A uniform prior for a parameter θ within $[a, b]$ is defined as:

$$P(\theta) = \frac{1}{b-a} \quad \text{for } \theta \in [a, b]$$

This means that every value of θ within the interval $[a, b]$ is equally probable[11].

- We can also propose the a priori (impropre) Hollander distributions $\pi(\theta) = [\theta(1-\theta)]^{-1} \mathbf{II}_{[0,1[}(\theta)$ arguing that $\mathrm{E}[\theta|X]$ is equal to the maximum likelihood estimator

2.4.3 The Weight of the a prior in the Bayesian response

Using an example , let's examine the question to understand how a prior information and the information Contained in the observation Combine With each other to produce the Bayesian answer . We give the following Bayesian model :$\mathrm{X}_i|\theta \sim$ bernoulli((θ et $\theta \sim Beta(a, b)$).) it is Convenient to reparameterize the Beta law using λ and μ (as above) and working With the following formula (easy to establish)

$$E[\theta|x] = \frac{\lambda}{\lambda+n} E[\theta] + \frac{n}{\lambda+n}\overline{x}$$

the Bayesian estimate ofθ therefore appears as the Weighted average of X. and of the a priori mean $E(\theta)$ the and the a priori mean Weight of X is the size m of the sample , and that of $E(\theta)$ is λ which is interpreted as the precision of the a priori.
Geometrically $E(\theta|X)$ is the barycenter of the coordinates points $E(\theta)$ and $\overline{x}$,assigned respectively to the coefficients $\frac{\lambda}{\lambda+n}$ and $\frac{n}{\lambda+n}$

- If $\lambda > n$: the estimator is closer to $E(\theta)$ than to $\overline{x}$.
- If $\lambda < n$: the estimator is closer to $\overline{x}$ than to $E(\theta)$.
- If $\lambda = n$: the Bayesian estimate of θ is located exactly in the middle of the interval $[E(\theta), \overline{x}]$.

To examine the influence of the a priori on $E(\theta|X)$ it is instructive to look at the limiting cases: $\lambda \to 0$ and $\lambda \to \infty$ (the size n of the sample being fixed, as well as $\mu = E(\theta)$).

- In the first case, the weight of the prior is 0 and $E(\theta|X) \to \overline{x}$ which is the classic answer.

- In the second case, the weight of the data is 0 and $E(\theta|X) \to E(\theta)$ which no longer depends on X.

Translated with DeepL.com (free version) it is also interesting to look at What $E[\theta|X]$ becomes When $n \to +\infty$ $\lambda$and μ being fixed , in this Case , the Weight of the a priori becomes negligible , and the Bayesian response Coincide With the classical response , Cod $\overline{x}$ (the estimate of θ by maximum likelihood).

2.5 A posteriori probability

A posteriori probability is used to describe the revised likelihood of an event or hypothesis following the consideration of fresh data or evidence.
It stands in contrast to the prior probability, which is the initial belief or likelihood before taking we determine the a posteriori distribution of a parameter:

- The joint law of (θ, X): Its density is noted $f(\theta, x)$, $f(\theta, x) = f(x|\theta)\pi(\theta)$.
- The law of the marginal of X: Its density is noted $m(x)$:

$$\begin{aligned} m(x) &= \int_{\Theta} f(\theta, x) d\theta \\ &= \int_{\Theta} f(x|\theta)\pi(\theta) d\theta \\ f(\theta|x) &= \frac{f(x|\theta)\pi(\theta)}{\int_{\Theta} f(x|\theta)\pi(\theta) d\theta} \\ &= \frac{f(\theta, x)}{m(x)}. \end{aligned}$$

Example 3 *Either $X|\theta \sim P(\theta)$ or $\theta \sim \gamma(2,1)$.*
Let us give the a posteriori probability law of the parameter θ as well as the marginal law of X. We have,

$$\begin{aligned} f(X|\theta) &= \frac{e^{-\theta}\theta^x}{x!}, \quad x \in \mathbb{N} \\ \pi(\theta) &= \theta e^{-\theta}, \quad \theta > 0 \end{aligned}$$

$f(\theta, x) = \frac{e^{-2\theta}\theta^{x+1}}{x!}$ *and*

$$\begin{aligned} m(x) &= \int_0^{+\infty} f(x|\theta)\pi(\theta)\, d\theta, \\ &= \int_0^{+\infty} \frac{e^{-2\theta}\theta^{x+1}}{x!} d\theta, \\ &= \int_0^{+\infty} \frac{\Gamma(x+2)}{\Gamma(x+2)} \frac{e^{-2\theta}\theta^{x+1}}{x!} \frac{2^{x+2}}{2^{x+2}} d\theta, \\ &= \frac{\Gamma(x+2)}{x! 2^{x+2}} \int_0^{+\infty} \frac{e^{-2\theta}\theta^{x+1} 2^{x+2}}{\Gamma(x+2)} d\theta, \\ &= \frac{\Gamma(x+2)}{x! \cdot 2^{x+2}}, \\ &= \frac{(x+1)!}{x! \cdot 2^{x+2}}. \end{aligned}$$

(with $\Gamma(x+2) = (x+1)!$*).*
Hence the marginal law of X*:*

$$m(x) = \frac{x+1}{2^{x+2}} \quad \alpha \in \mathbb{N}$$

And the posterior law of θ*:*

$$\begin{aligned} f(\theta|x) &= \frac{e^{-2\theta}\theta^{x+1}}{(x+1)x!}2^{x+2} \\ &= \frac{e^{-2\theta}\theta^{x+1}2^{x+2}}{(x+1)!}, \qquad \theta > 0 \end{aligned}$$

So,

$$\theta|x \sim \gamma(x+2, 2)$$

2.6 Estimation of Parameters

2.6.1 Definitions and Importance

The process of estimating unknown parameters in a statistical model using observed data is known as parameter estimation.
Since parameters are essential to models and define their behavior and structure, it is possible to create models that more accurately reflect the processes that generate the data.
In Bayesian statistics, parameter estimation entails updating prior beliefs about parameters using observed data in order to obtain their posterior distributions.

The parameter estimation is crucial for several reasons:

- **Modeling:**
  To create trustworthy statistical models, precise parameter estimations are necessary. These models can be used to accurately explain complicated systems and help in the understanding of the interactions between variables.

- **Uncertainty Quantification:**
  For risk assessment and well-informed decision-making under uncertainty, Bayesian parameter estimation offers an intuitive means of quantifying uncertainty in parameter estimation by posterior distributions [11].

2.6.2 Bayesian Methods of Parameter Estimation

Bayes Estimator

The Bayes estimator δ of $\theta \in \Theta$, associated with the posterior distribution $\pi(\theta|x)$ and the quadratic cost $C(\theta, d) = (\theta - d)^2$, is the posterior average.

$$\theta_m = E[\theta|x] = \int_\Theta \theta\, \pi(\theta|x)\, d\theta \tag{1.1}$$

Indeed, we have:

$$E[C(\theta, d)] = d^2 - 2dE[\theta] + E[\theta^2] = (d - E[\theta])^2 + \text{Var}[\theta]$$

which is minimized when $d = E[\theta]$.

Example 4 *We consider the following Bayesian model:* $X_i \mid \theta \sim$ *Bernoulli*(θ) *and* $\theta \sim$ *Beta*(a, b).
Recall that: $\theta \mid x \sim$ *Beta*(α, β) *where:*

$$\alpha = a + s \quad \textit{and} \quad \beta = b + n - s \quad \textit{and} \quad s = \sum_{i=1}^{n} x_i$$

from where:

$$E[\theta \mid X] = \frac{a + \sum_{i=1}^{n} x_i}{a + b + n}$$

2.6.3 Properties of the Bayes estimator

- The Bayes estimator is admissible.
- The Bayes estimator is biased.
- The Bayes estimator is convergent in probability (where the size of the sample is $n \longrightarrow +\infty$).
- For large values of n, the posterior distribution may asymptotically (c.a.d) be approximated by a normal law $N(E[\theta|X], var[\theta|X])$.

2.6.4 Maximum a Posteriori Estimator

Such an estimator also takes the form of the extremum of the a posteriori distribution:

$$\theta_{Max} = \arg \max_{\pi(\theta \in \Theta} \{\pi(\theta|x)\} \tag{1.2}$$

It is thus necessary to use integration and/or optimization techniques to such calculations. Regretfully, calculations can typically only be performed analytically in basic circumstances due to the complexity of the estimators equations (1.1) or (1.2), particularly for linear models and standard probability distributions[12].

2.6.5 Bayesian credibility interval

The credibility interval is a real reflection of the confidence that one can have in the value of the parameter concerned.

It emphasizes the original tenet of the Bayesian approach which holds that the parameter is a random variable about which we have a reasonably good judgment rather than an unknown fixed quantity.

By definition, a credibility interval of the level α is an interval such that the probability of the parameter to belong to it according to the posterior distribution is $(1-\alpha)$. There are many ways of building it; for example, we can take the one given by $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ quantiles.
When multiple parameters are of relevance at the same time, we refer to this as a credibility region[12].

Example 5 *Bayesian Inference to Determine the Posterior Credibility Interval of* θ.

To determine the posterior credibility interval of θ, we follow the steps of Bayesian inference. We consider that the observations $X_1, X_2, \ldots, X_n$ are a sample of size n from a random variable X which follows a normal distribution $\mathcal{N}(0, 1)$. With θ following a normal prior distribution $\mathcal{N}(0, 5)$, we aim to determine the posterior credibility interval for θ.

Bayesian Inference Steps

Prior Distribution:
We suppose that the prior distribution of θ is $\theta \sim \mathcal{N}(0, 5)$.
Likelihood Function:

Given the sample $X_1, X_2, \ldots, X_n$ from a normal distribution with mean θ and variance 1, the likelihood function of the observed data is:

$$f(X|\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(X_i - \theta)^2}{2}\right)$$

Posterior Distribution:

Using Bayes's theorem, the posterior distribution of θ given the data $X = (X_1, X_2, \ldots, X_n)$ is:

$$P(\theta|X) \propto f(X|\theta) \cdot P(\theta)$$

Given the prior $\theta \sim \mathcal{N}(0, 5)$ and the likelihood, the posterior distribution $P(\theta|X)$ will also be normal. We need to derive its parameters (mean and variance).

Let's rewrite the posterior distribution:

$$P(\theta|X) \propto \exp\left(-\frac{1}{2}\left(\sum_{i=1}^{n}(X_i - \theta)^2 + \theta^2\right)\right)$$

Combining the exponents:

$$= \exp\left(-\frac{1}{2}\left(\sum_{i=1}^{n} X_i^2 - 2\theta n\bar{X} + (n+1)\theta^2\right)\right)$$

So the posterior distribution is normal with parameters:

$$\theta|X \sim \mathcal{N}\left(\frac{n\bar{X}}{n+1}, \frac{1}{n+1}\right)$$

nCredibility Interval:
For a 95% credibility interval ($\alpha = 0.05$) for a normal distribution, there is approximately ± 1.96. The 95% credible interval for θ is:

$$\left[\frac{n\bar{X}}{n+1} - 1.96\sqrt{\frac{1}{n+1}}, \frac{n\bar{X}}{n+1} + 1.96\sqrt{\frac{1}{n+1}}\right]$$

2.7 Markov Chain Monte Carlo (MCMC) Method:

The general name for computational methods involving random numbers is Monte Carlo. Both conventional and Bayesian statistics can make use of Monte Carlo.
Markov Chain Monte Carlo (MCMC) is a unique type of Monte Carlo that was a major factor in the second part of the 20th century Bayesian statistics resurgence.

One of the main problems of the Bayesian technique prior to MCMC's rise in popularity was that certain calculations were extremely difficult.
We can solve Many different types of Bayesian issues with MCMC that are not amenable to analytical solutions[13].

The fundamental principle behind these techniques is to perform the computations using a sequence of samples $(\theta_1, ..., \theta_n)$ that have been simulated in accordance with the probability law of interest ($\pi(\theta|x)$). Based on a set of data $(\theta_1, ..., \theta_n) \sim \pi(\theta|x)$, the posterior mean estimator provided by (1,1) is calculated using the following formula:

$$\widehat{\theta}_n = \sum_{i=1}^{n} \theta_i.$$

Unfortunately, because there aren't many easily replicated probability laws, modeling still easily places limitations on such an estimating technique. Nevertheless, in order to sample in accordance with the distributions of interest, it is frequently required to use complex simulation of techniques for Markov chain simulation an applied.

The next section utilizes these techniques[12]
The purpose of MCMC techniques is to use a measuring Markov chain invariant g to approximate the g distribution. We can then use this to make an estimator indeed, if $Z_1, ..., Z_n \sim \pi(\theta|x)$ than we can take as an estimator of θ,

$$\hat{\theta}_n := \frac{1}{n}\sum_{i=1}^{n} Z_i \quad \text{(Monte Carlo estimator)}$$
$$\text{or } \hat{\theta}_n := \text{median}(Z_1, ..., Z_n)$$
$$\text{or } \hat{\theta}_n := \arg\max(Z_1, ..., Z_n)$$

The fundamental concept behind MCMC techniques is to take a look at a Markov chain that generates associated samples

$$\theta_0 \to \theta_1 \to \theta_2 ...$$

such that for i sufficiently large, θ_i roughly follows the g law [9].

Marker property:

A Markov chain satisfies the Markov property, meaning the future state depends only on the current state and not the sequence of events that preceded it.

Definition

We denote by χ the state space, in the sequence X is either finite or infinitely countable, or it is R^d.
A Markov chain $(X_0, X_1, \ldots)$ with $X_i \in \chi$ is a sequence of random variables satisfying:

$$f(X_{n+1}|X_i, ..., X_1) = K(X_{i+1}, |X_i)$$

Where $K : \chi \times \chi \to \mathbb{R}^+$ called a Markov Kernel or transition Kernel verifies for all $x \in \chi, x' \longmapsto K(x'|x)$ is a probability density.

Proposals:

- $U_{mM} = U_n A, n \in N^*$
- $U_n = U_0 A^n, n \in N^*$

Propositions:

$$U_{n+1}(x') = \int_x U_n(x) K(x'|x) dx, \ \ \forall x' \in \chi$$

2.7.1 Metropolis-Hastings algorithm:

When it is difficult to sample directly from a probability distribution, or Marker chain Monte Carlo (MCMC) technique that is utilised is the Metropolis-Hastings algorithm. For Bayesian inference and statistical physics, it is particularly helpful with high-dimensional distributions.

Basic Concepts:

In the target distribution , the goal is to generate samples from a target distribution $\pi(x)$, which is often known only up to normalizing constant.

Proposal distribution, the proposal distribution $q(x'|x)$ is used by the algorithm to create conditional samples. It proposes a new state, x', given the present state x.
Algorithm steps:

Start with an initial state x_0.

For each step $t = 1, 2, .., T$,

draw a candidate x' from the proposal distribution $q(x'|x_t)$.

Calculate the acceptance ratio α:

With probability α, accept the new candidate and set $x_{t+1} = x'$; otherwise, reject the candidate, and set $x_{t+1} = x_t$.

After a sufficient number of iterations, the samples $\{x_t\}$ will approximate the target distribution $\pi(x)$.

Example 6 *Sample from a distribution* $\pi(x)\alpha e^{-\frac{x^2}{2}}$ *a normal distribution* $N(0, 1)$. *We use a normal proposal distribution* $q(x|x') = N(x, 6^2)$.
The algorithm,

- $x_0 = 0$
- *At step* t, *suppose* $x_t = 0$. *Draw a candidate* x' *from* $N(0, 6^2)$.
calculate α*:*

$$\alpha = min\left(1, \frac{e^{-\frac{x^2}{2}}}{e^{-\frac{x_t^2}{2}}}\right)$$

- *If* $x' = 1$, *then* $\alpha = \min(1, e^{-\frac{1}{2}}) \approx 0.606$. *With probability* 0.606, *accept* x' *and set* $x_{t+1} = 1$*; otherwise,* $x_{t+1} = 0$.

2.7.2 Gibbs Algorithm

Gibbs sampling is a Markov Chain Monte Carlo (MCMC) algorithm used to sample from a multivariate probability distribution when direct sampling is difficult.
By sampling iteratively from the conditional distributions of each variable given the others, it creates a series of samples from the joint distribution [15].

The Gibbs Sampling Algorithm

- Initialization: Start with an initial guess for the parameters. Let $\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \ldots, \theta_d^{(0)})$ be the initial values of the parameters.
- Iteration: For each iteration $t = 1, 2, \ldots, T$:

Sample $\theta_1^{(t)}$ from the conditional distribution $P(\theta_1|\theta_2^{(t-1)}, \theta_3^{(t-1)}, \ldots, \theta_d^{(t-1)})$

Sample $\theta_2^{(t)}$ from the conditional distribution $P(\theta_2|\theta_1^{(t)}, \theta_3^{(t-1)}, \dots, \theta_d^{(t-1)})$.

$\vdots$

Sample $\theta_d^{(t)}$ from the conditional distribution $P(\theta_d|\theta_1^{(t)}, \theta_2^{(t)}, \dots, \theta_{d-1}^{(t)})$

- Repeat: Continue the process for a large number of iterations to ensure the samples converge to the target distribution.

2.8 Bayesian Hypothesis testing :

Bayesian hypothesis testing provides a framework to update the probability of a hypothesis as more evidence or information becomes available . in Contrast to frequentist techniques, Which defend on predetermined significance levels and p-values , Bayesian testing uses probability distributions to quantify uncertainty . [11]
Assume , the parameter space θ is divided into θ_0 and θ_1,and $P(\theta \in \theta_k) > 0$ where K=0.1
we consider hypotheses:

- H_0: Null hypothesis
- H_1: Alternative hypothesis

We wish to test $H_0 : \theta \in \theta_0$against the alternative $H_1 : \theta \in \theta_1$ in Bayesian statistic the answer to such a test is based on the posterior probabilities of hypotheses H_0 and H_1

$$\mathrm{P}\left(H_0|\mathbf{x}\right) = \mathrm{P}\left(\theta \in \Theta_0|\mathbf{x}\right) = \int_{\Theta_0} \pi\left(\theta|\mathbf{x}\right) d\theta$$

$$\mathrm{P}\left(H_1|\mathbf{x}\right) = \mathrm{P}\left(\theta \in \Theta_1|\mathbf{x}\right) = \int_{\Theta_1} \pi\left(\theta|\mathbf{x}\right) d\theta$$

note that: $P\left(H_1|\mathbf{x}\right) = 1 - P\left(H_0|\mathbf{x}\right)$.
One way to decide between $H_0$and H_1 is to compare $P(H_0|X)$ and $P(H_1|X)$ and accept the hypothesis with the higher posterior probability. note :
a hypothesis (H_0 or H_1) is accepted as soon as it's the posterior probability is considered sufficiently high (greater than 0.9 or 0.95)
In the event that none of the posterior probabilities surpasses 0.9, two distinct attitudes could exist:

- either we do not make a decision, and we decide for example to collect more observations.
- either we choose the hypothesis whose a posterior probability is the greatest.[16]

CHAPTER 3

MODEL OF CLAIMS IN NON LIFE INSURANCE WITH BAYESIAN APPROACH

3.1 Introduction

Risk theory is concerned with the erratic changes in an insurance company's assets, or risk reserve . The state of an insurance company at any given time may be described by its financial reserves, which result mainly from the continuous trade-off between incoming premiums to the company, and outgoing claims to its clients . A risk theoretical model can be expressed in its most basic form as a stochastic process that involves two random variables that reflect claims and premiums, as well as their mutual relationship and joint evolution over time. Overall, risk models are crucial for ensuring the long-term sustainability and profitability of financial institutions by systematically addressing the uncertainties they face.

3.2 Risk Model and Ruin Probability

3.2.1 Risk Model

Risk in insurance means the possibility of a disaster occurring or any other adverse event likely to impact a company's ability to operate its activities and in respect of which it can make a claim.

Stochastic models are taken into consideration in risk theory, which can be applied to an insurance company risk.

-An insurance Company's financial scenario, which the outcome of a constant trade of between incoming Premium Payments to the Company and outgoing Claims from Clients, can be used to characterize its current situation. A risk theory model can be expressed as a stochastic process in its most basic form, wherein Premiums and claims are represented by two random variables, together with their mutual relationship and combined evolution over time.

The risk reserve, W(t), is the financial reserve of an insurance Company at time t, and can be expressed by the surplus equation:

$$W(t) = u + ct - S(t)$$

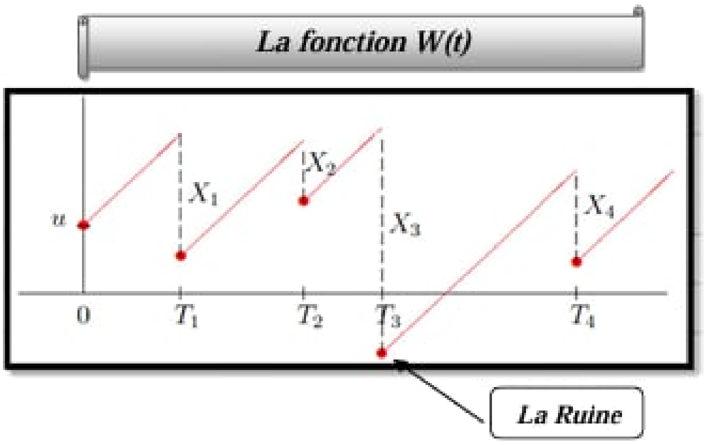

Figure 3.1: Development of the Reserve Risk $W(t)$ Over Time T

where $S(t) = \sum_{i=1}^{N(t)} X_i$ is the aggregate amount of claims
and u is the initial capital of the company, c is the premium density which is assumed to be constant, the function $S(t)$ is the claim process which denotes the total amount that the company has to pay to the customers by time t.

Let $\{N(t); t \geq 0\}$ denote a stochastic process representing the number of claims arising from a portfolio of risks in the time interval $(0, t]$. If the sizes of successive claims are $X_1, X_2, \ldots$, then the aggregate claims in time interval $(0, t]$ is:

$$S(t) = \sum_{i=1}^{N(t)} X_i$$

These schematics below represent the evolution of the number of claims $N(t)$ over the period T, highlighting the cumulative nature and frequency of claim occurrences. Additionally, they illustrate the progression of the aggregate amount of claims $S(t)$ over the same period, demonstrating how the total cumulative amount of claims accumulates as claims occur.[12]

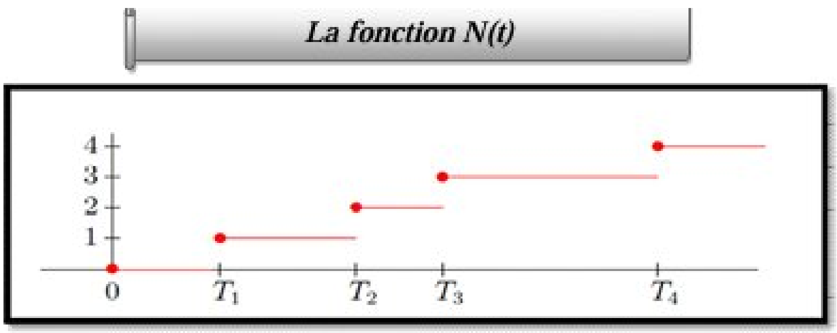

Figure 3.2: Development of the Number of Claims $N(t)$ Over Time T

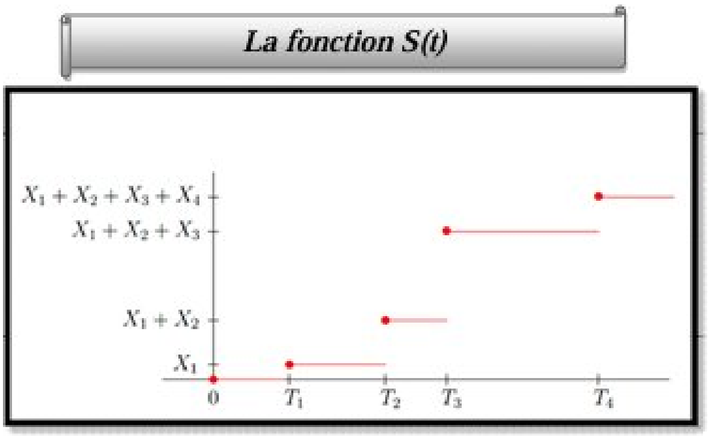

Figure 3.3: Development of the Aggregate Claim Amount $S(t)$ Over Time T

3.2.2 Ruin probability

Definition: The probability of ruin in its primary version corresponds to the point beyond which the initial capital is completely exhausted over a given period following an outcome deficit. The probability of ruin is equal to [7]:

$$\psi(u) = P_u\{R(t) < 0, \text{ for some } t > 0\}$$

3.2.3 The Individual Risk Model

The individual risk model (IRM) is derived by considering the claims on individual policies and summing all policies in the portfolio. The individual risk model represents the aggregate loss as a fixed sum of independent and identically distributed. The basic equation of the individual risk model is:

$$S = \sum_{i=1}^{n} X_i,$$

where S is the total amount, n is the number of individual risks and is fixed, and X_i is the claim amount for the i-th risk.

The mean and the variance of S are given by:

$$\mathbb{E}(S) = nE(X) \quad \text{and} \quad \text{Var}(S) = n\text{Var}(X)$$

-Let the probability of a loss be θ and the probability of no loss be $1-\theta$.
-Where there is a loss, the loss amount is Y which is a positive continuous random variable with mean μ_Y and variance $\sigma^2 Y$.
Thus, $X = Y$ with probability θ and $X = 0$ with probability $1-\theta$. We have $X = IY$, where I is a Bernoulli random variable distributed

independently of Y so that,

$$I = \begin{cases} 0 & \text{with probability } 1-\theta \\ 1 & \text{with probability } \theta \end{cases}$$

So,

$$\begin{aligned} E(X) &= E(I)\,E(Y) = \lambda\mu \\ \text{Var}(X) &= \text{Var}(IY) \\ &= [E(Y)]^2\,\text{Var}(I) + E\left(I^2\right)\text{Var}(Y) \\ &= \mu^2\lambda\theta(1-\theta) + \theta\sigma^2 Y. \end{aligned}$$

3.2.4 The Collective Risk Model

We still compute the distribution of the entire claim amount over a given time period, but as of right now, we think of the portfolio as a collective that makes claims at random intervals we compose[[6]]:

$$S = X_1 + X_2 + \cdots + X_N$$

where N denotes the number of claims and X_i is the i-th claim.

-The Mgf $M_S(t)$ of the aggregate loss S is given by $M_S(t) = M_N(\log M_x(t))$.

If the claim severity takes non-negative discrete values, S is also non-negative and discrete, and its pgf is [17]:

$$P_S(t) = P_N[P_X(t)]$$

The mean of S is:

$$E(S) = E(N)E(X)$$

The variance of S is:

$$\text{Var}(S) = E(N)\text{Var}(X) + \text{Var}(N)[E(X)]^2$$

3.2.5 Common Distributions for Modeling the Number of Claims

Let the random variable X represent the number of claims from individual policy or a policy portfolio. In motor liability insurance different theoretical distributions may be used to model the number of claims (Lemaire, 1995). The following are the most commonly used distributions for modeling the number of claims .

1. **Binomial Distribution**
   The Bernoulli binomial distribution is described with the probability distribution function:
   $$P(X = k) = \binom{n}{k} p^k q^{n-k}, \quad \text{where} \quad k = 0, 1, \ldots, n \quad \text{and} \quad \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

2. **Poisson distribution**
   The Poisson distribution is a distribution with the function of the probability defined by the formula:
   $$P(X = k) = \exp(-\lambda)\frac{\lambda^k}{k!}, \quad k = 0, 1, \ldots.$$
   The random variable X has a negative binomial distribution (Polya) when its probability distribution function has the form[3]:
   $$P(X = k) = \frac{\Gamma(\alpha + k)}{\Gamma(\alpha)k!}\left(\frac{\beta}{1+\beta}\right)^{\alpha}\left(\frac{1}{1+\beta}\right)^{k}.$$
   If the random variable X has a Poisson distribution with parameter λ and the parameter λ has the inverse normal distribution, then the random variable X has a Poisson-inverse normal distribution (Denuit, Marechal, Pitrebois, Walhin, 2007, p.31).
   The probability function of Poisson-inverse normal distribution is given by:
   $$P(X = k) = \sqrt{\frac{2\alpha}{\pi}}\exp(\alpha(1-\theta))\frac{(\alpha\theta/2)^k}{k!}K_{k-1/2}(\alpha) \quad k = 0, 1, \ldots,$$
   where $K_{k-1/2}(\alpha)$ is a modified third kind Bessel function (for positive and real arguments) in the form of:
   $$K_{k-1/2}(\alpha) = \sqrt{\frac{\pi}{2\alpha}}\exp(-\alpha)\left(\sum_{i=0}^{k-1}\frac{(k-1+i)!}{(k-1-i)!i!}\right)(2\alpha)^{-i} \quad k = 1, 2, \ldots.$$

Probability distribution function of the Poisson-Poisson distribution (Neyman type A) is given by:

$$P(X=k)=\exp(-\lambda_1)\sum_{n=0}^{\infty}\frac{\lambda_1^n}{n!}\frac{(\lambda_2 n)^k}{k!}\exp(-\lambda_2 n), \quad k=0,1,2,\ldots.$$

Example 7 *Poisson distribution (uncertainty about the parameter)*
Assume that some car driver causes a Poisson $N(\lambda)$ distributed number of accidents in one year. The parameter λ is unknown and different for every driver. We assume that λ is the outcome of a random variable Λ given $\Lambda=\lambda$. Calculate the marginal distribution of N

$U(\lambda)=P_r(\Lambda\le\lambda)$ denote the distribution function of Λ then we can write the marginal distribution of N as

$$\begin{aligned}P_r[N=n]&=\int_0^{\infty}P_r[N=n/\Lambda=\lambda]dU(\lambda)\\&=\int_0^{\infty}e^{-\lambda}\frac{\lambda^n}{a!}dU(\lambda)\end{aligned}$$

if $\Lambda\to\gamma(\alpha,\beta)$, then

$$\begin{aligned}m_N(t)&=E[E[e^{tN}/\Lambda]]\\&=E[\exp\{\Lambda(e^t-1)\}]=m_\Lambda(e^t-1)\\&=\left(\frac{\beta}{\beta-(e^t-1)}\right)=\left(\frac{\beta}{1-(1-P)e^t}\right)^{\alpha}\end{aligned}$$

where $P=\beta(\beta+1)$. So N has a negative binomial $(\alpha,\beta/(\beta+1))$ distribution [5]

3.3 Distribution of the total claim amount

3.3.1 Pareto distribution of amount claim

The Pareto distribution is among the most heavy-tailed of all models in practical use and this is essential for modeling extreme losses, especially in the more risky types of insurance.[14]

Hence it is a conservative choice when modeling the claim size. The density function of the Pareto distribution is

$$g(x)=\frac{\alpha x_m^{\alpha}}{x^{1+\alpha}} \quad x\ge x_m$$

x_m: the minimum possible value of X

α: a positive parameter.

We assume that the smallest possible value of x is 1, so the density is

$$g(x)=\frac{\alpha}{x^{1+\alpha}} \quad x\ge 1$$

In order to make the density values from zero we let $Z=\beta(x-1)$ such that $x=1+\frac{Z}{\beta}$ by inversion. The probability function of Z is

$$f(Z)=g(x(Z))\left|\frac{\delta x(Z)}{\delta Z}\right|=\frac{\alpha/\beta}{(1+Z/\beta)^{1+\alpha}} \quad Z>0$$

$\alpha>0$ is the shape parameter and $\beta>0$ is the scale parameter.

$$E(Z)=\frac{\beta}{\alpha-1}, \quad \alpha>1$$

$$\text{Var}(Z)=\frac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)} \quad \alpha>2$$

Parameter Estimation using Maximum Likelihood Method :
Likelihood Function:

$$L(\alpha, \beta) = \prod_{i=1}^{n} f(z_i \mid \alpha, \beta) = \left(\frac{\alpha}{\beta}\right)^n \prod_{i=1}^{n} \left(1 + \frac{z_i}{\beta}\right)^{-(1+\alpha)}$$

Log-Likelihood Function:

$$\log L(\alpha, \beta) = n \log(\alpha) - n \log(\beta) - (1 + \alpha) \sum_{i=1}^{n} \log \left(1 + \frac{z_i}{\beta}\right)$$

MLE for Shape Parameter α:
Differentiating $\log L(\alpha, \beta)$ with respect to α and setting the result to zero:

$$\frac{\partial \log L(\alpha, \beta)}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} \log \left(1 + \frac{z_i}{\beta}\right) = 0$$

Solving this gives:

$$\hat{\alpha}_\beta = \frac{n}{\sum_{i=1}^{n} \log \left(1 + \frac{z_i}{\beta}\right)}$$

Log-Likelihood Function Depending Only on β: Insert $\hat{\alpha}$ into the log-likelihood function:

$$\log L(\beta) = n \log \left(\frac{n}{\sum_{i=1}^{n} \log \left(1 + \frac{z_i}{\beta}\right)}\right) - n \log(\beta) - \left(1 + \frac{n}{\sum_{i=1}^{n} \log \left(1 + \frac{z_i}{\beta}\right)}\right) \sum_{i=1}^{n} \log \left(1 + \frac{z_i}{\beta}\right)$$

Simplify:

$$\log L(\beta) = n \left[\log(n) - \log \left(\sum_{i=1}^{n} \log \left(1 + \frac{z_i}{\beta}\right)\right)\right] - n \log(\beta) - \left(1 + \frac{n}{\sum_{i=1}^{n} \log \left(1 + \frac{z_i}{\beta}\right)}\right) \sum_{i=1}^{n} \log \left(1 + \frac{z_i}{\beta}\right)$$

3.3.2 Weibull Distribution

The Weibull distribution is a widely used distribution because of its versatility. The density function is

$$f(Z) = \frac{\alpha}{\beta} \left(\frac{Z}{\beta}\right)^{\alpha - 1} e^{-(Z/\beta)^\alpha}, \quad Z > 0$$

The mean and the variance:

$$E(Z) = \beta \Gamma \left(1 + \frac{1}{\alpha}\right)$$

$$\operatorname{Var}(Z) = \beta^2 \left[\Gamma \left(1 + \frac{2}{\alpha}\right) - \left(\Gamma \left(1 + \frac{1}{\alpha}\right)\right)^2\right]$$

Parameter Estimation using Maximum Likelihood Method :
Likelihood Function:

$$L(\alpha, \beta) = \prod_{i=1}^{n} f(z_i \mid \alpha, \beta) = \frac{\alpha^n}{\beta^{n\alpha}} e^{-\sum_{i=1}^{n} (z_i/\beta)^\alpha} \prod_{i=1}^{n} z_i^{\alpha - 1}$$

Log-Likelihood Function:

$$\log L(\alpha, \beta) = n \log(\alpha) - n\alpha \log(\beta) + (\alpha - 1) \sum_{i=1}^{n} \log(z_i) - \frac{1}{\beta^\alpha} \sum_{i=1}^{n} z_i^\alpha$$

MLE for Scale Parameter β: Differentiate $\log L(\alpha, \beta)$ with respect to β and set the result to zero:

$$\frac{\partial \log L(\alpha, \beta)}{\partial \beta} = -\frac{n\alpha}{\beta} + \frac{\alpha}{\beta^{\alpha+1}} \sum_{i=1}^{n} z_i^\alpha = 0$$

Solving this gives:

$$\hat{\beta}^\alpha = \left(\frac{1}{n} \sum_{i=1}^{n} z_i^\alpha\right)^{1/\alpha}$$

Log-Likelihood Function Depending Only on α: Insert $\hat{\beta}$ into the log-likelihood function:

$$\log L(\alpha) = n\log(\alpha) - n\alpha\log\left(\left(\frac{1}{n}\sum_{i=1}^{n} z_i^{\alpha}\right)^{1/\alpha}\right) + (\alpha-1)\sum_{i=1}^{n}\log(z_i)$$

Simplify:

$$\log L(\alpha) = n\left[\log(\alpha) + \log(n) - \log\left(\sum_{i=1}^{n} z_i^{\alpha}\right) - 1\right] + (\alpha-1)\sum_{i=1}^{n}\log(z_i)$$

The marginal distribution of parameters given insurance claim data is found by integrating over both parameters:

$$P(x_t) = \int_0^{\infty} P(x_t, \beta, \alpha)$$

This integration simplifies to a form involving a new variable ρ, defined as the product of α and the sum of β and Λ:

$$\rho = \alpha\Big(\sum_{t=1}^{n} x_t^{\beta} + \lambda\Big)$$

The likelihood function is then proportional to the marginal function, leading to the derivation of the joint posterior distribution of the two parameters [21]:

$$P(\alpha|x_t) = \frac{\rho(x_t, \beta, \alpha)}{\rho(x_t)}$$

This distribution resembles a Gamma distribution:

$$\frac{\alpha^{n+\psi-1}}{\Gamma(n+\psi)} e^{-\alpha\left(\sum_{t=1}^{n} x_t^{\beta}\right)} \left(\sum_{t=1}^{n} x_t^{\beta} + \lambda\right)^{n+\psi} \sim Gamma\left(n+\psi, \sum_{t=1}^{n} x_t^{\beta} + \lambda\right)$$

From this, the Bayes estimate under the squared error loss function is calculated:

$$\hat{\alpha} = E(\alpha) = \frac{n+\psi}{\sum_{t=1}^{n} x_t^{\beta} + \lambda}$$

3.3.3 Gamma distribution

$X \sim \Gamma(\alpha, \beta)$ with the density

$$f^X(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0.$$

If we choose $\alpha < 1$ we fulfill the property that as much mass of the distribution as possible lies near 0. The Gamma-distribution is often used for insurance contracts without extreme claim heights, such as car insurances.

Given a complete sample $X_1, X_2, \ldots, X_n$ from a Gamma distribution with parameters α and β, the likelihood function is:

$$L(\alpha, \beta \mid x) = \beta^{n\alpha}\left[\Gamma(\alpha)\right]^{-n}\left(\prod_{i=1}^{n} x_i^{\alpha-1}\right)\exp\left(-\beta\sum_{i=1}^{n} x_i\right)$$

where $\alpha > 0$ and $\beta > 0$.

By taking the logarithm of the likelihood function and setting the partial derivatives with respect to α and β to zero, we get the likelihood equations:

$$\hat{\beta} = \frac{\hat{\alpha}}{\bar{X}}$$

$$\log\hat{\alpha} - \psi(\hat{\alpha}) = \log\bar{X} - \log\tilde{X}$$

where $\psi(k)$ is the digamma function, $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is the sample mean, and $\tilde{X} = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$ is the geometric mean of the sample.

These equations do not have a closed-form solution and require numerical techniques to solve for the maximum likelihood estimators $\hat{\alpha}$ and $\hat{\beta}$.

The Fisher information matrix for the parameters α and β is:

$$I(\alpha, \beta) = \begin{bmatrix} \psi'(\alpha) & -\frac{1}{\beta} \\ -\frac{1}{\beta} & \frac{\alpha}{\beta^2} \end{bmatrix}$$

where $\psi'(\alpha)$ is the trigamma function, the derivative of the digamma function.

For large sample sizes ($n \to \infty$), the maximum likelihood estimators $\hat{\alpha}$ and $\hat{\beta}$ are approximately normally distributed:

$$\hat{\alpha} \sim N(\alpha, \sigma_1^2) \quad \text{and} \quad \hat{\beta} \sim N(\beta, \sigma_2^2)$$

where $\sigma_1^2 = (\alpha\psi'(\alpha))^{-1}$ and $\sigma_2^2 = \frac{\beta^2\psi'(\alpha)}{\alpha\psi'(\alpha)}$.

If our data $X_1, \ldots, X_n$ are iid Poisson(λ), then a gamma(α, β) prior on λ is a conjugate prior.

Likelihood:

$$L(\lambda|x) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} = e^{-n\lambda}\lambda^{\sum x_i} \prod_{i=1}^{n} \frac{1}{x_i!}$$

Prior:

$$p(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}, \quad \lambda > 0$$

Posterior:

$$\pi(\lambda|x) \propto \lambda^{\sum x_i + \alpha - 1}e^{-(n+\beta)\lambda}, \quad \lambda > 0$$

$$\pi(\lambda|x) \text{ is gamma} \left(\sum x_i + \alpha, n + \beta\right) \text{ (Conjugate!)}$$

The posterior mean is:

$$\hat{\lambda}_B = \frac{\sum x_i + \alpha}{n + \beta} = \frac{\sum x_i}{n + \beta} + \frac{\alpha}{n + \beta} = \left(\frac{n}{n + \beta}\right)\left(\frac{\sum x_i}{n}\right) + \left(\frac{\beta}{n + \beta}\right)\left(\frac{\alpha}{\beta}\right)$$

3.3.4 Exponential Distribution

Exponential distribution:$X \sim$ Exponential(β) with the density

$$f^X(x) = \beta e^{-\beta x}, \quad x > 0.$$

Given a complete sample $X_1, X_2, \ldots, X_n$ from an Exponential distribution with parameter β, the likelihood function is:

$$L(\beta \mid x) = \beta^n \exp\left(-\beta\sum_{i=1}^{n} x_i\right)$$

where $\beta > 0$.

By taking the logarithm of the likelihood function and setting the derivative with respect to β to zero, we get the likelihood equation:

$$\hat{\beta} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\bar{X}}$$

where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is the sample mean.

For large sample sizes ($n \to \infty$), the maximum likelihood estimator $\hat{\beta}$ is approximately normally distributed:

$$\hat{\beta} \sim N\left(\beta, \frac{\beta^2}{n}\right)$$

If our data $X_1, \dots, X_n$ are *i.i.d* Exponential(λ), then a gamma(α, β) prior on λ is a conjugate prior.

Likelihood:

$$L(\lambda|x) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} x_i\right)$$

Prior:

$$p(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}, \quad \lambda > 0$$

Posterior:

$$\pi(\lambda|x) \propto \lambda^{n+\alpha-1} e^{-(\sum_{i=1}^{n} x_i + \beta)\lambda}, \quad \lambda > 0$$

$$\pi(\lambda|x) \text{ is gamma}(n+\alpha, \sum_{i=1}^{n} x_i + \beta) \text{ (Conjugate!)}$$

The posterior mean is:

$$\hat{\lambda}_B = \frac{n+\alpha}{\sum_{i=1}^{n} x_i + \beta} = \frac{n}{\sum_{i=1}^{n} x_i + \beta} + \frac{\alpha}{\sum_{i=1}^{n} x_i + \beta} = \left(\frac{n}{\sum_{i=1}^{n} x_i + \beta}\right)\left(\frac{1}{\bar{X}}\right) + \left(\frac{\beta}{\sum_{i=1}^{n} x_i + \beta}\right)\left(\frac{\alpha}{\beta}\right)$$

3.3.5 Log-normal distribution

Given: - The observed data $x = (x_1, x_2, \dots, x_n)$ are assumed to be log-normally distributed: $x_i \sim \text{LogNormal}(\mu, \sigma^2)$. - The log-normal distribution can be represented in terms of a normal distribution: if $y_i = \log(x_i)$, then $y_i \sim \mathcal{N}(\mu, \sigma^2)$. - We assume a gamma prior on the precision (inverse of variance): $\tau = \frac{1}{\sigma^2} \sim \text{Gamma}(\alpha, \beta)$.

The likelihood of the observed data $y = (\log(x_1), \log(x_2), \dots, \log(x_n))$ under the normal distribution is given by:

$$P(y|\mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2\right)$$

The prior on the precision $\tau = \frac{1}{\sigma^2}$ is:

$$P(\tau) = \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} \exp(-\beta\tau)$$

To find the posterior, we combine the likelihood and the prior:

$$P(\tau|y) \propto P(y|\tau)P(\tau)$$

Since $\sigma^2 = \frac{1}{\tau}$:

$$P(y|\tau) = \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left(-\frac{\tau}{2}\sum_{i=1}^{n}(y_i - \mu)^2\right)$$

Therefore,

$$P(\tau|y) \propto \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left(-\frac{\tau}{2}\sum_{i=1}^{n}(y_i - \mu)^2\right) \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} \exp(-\beta\tau)$$

Combining terms:

$$P(\tau|y) \propto \tau^{n/2} \tau^{\alpha-1} \exp\left(-\tau\left(\frac{1}{2}\sum_{i=1}^{n}(y_i - \mu)^2 + \beta\right)\right)$$

Simplifying:

$$P(\tau|y) \propto \tau^{\alpha+n/2-1} \exp\left(-\tau\left(\beta + \frac{1}{2}\sum_{i=1}^{n}(y_i - \mu)^2\right)\right)$$

This is the kernel of a gamma distribution with updated parameters:

$$\tau|y \sim \text{Gamma}\left(\alpha + \frac{n}{2}, \beta + \frac{1}{2}\sum_{i=1}^{n}(y_i - \mu)^2\right)$$

The posterior distribution of the precision τ (or equivalently the variance σ^2) is:

$$\tau|y \sim \text{Gamma}\left(\alpha + \frac{n}{2}, \beta + \frac{1}{2}\sum_{i=1}^{n}(y_i - \mu)^2\right)$$

Since $\sigma^2 = \frac{1}{\tau}$, the posterior distribution of σ^2 is:

$$\sigma^2|y \sim \text{Inverse-Gamma}\left(\alpha + \frac{n}{2}, \beta + \frac{1}{2}\sum_{i=1}^{n}(y_i - \mu)^2\right)$$

Therefore, the posterior distribution of the log-normal distribution's variance (given a gamma prior on the precision) is an inverse-gamma distribution.

CHAPTER 4

APPLICATION TO AGGREGATE INSURANCE CLAIMS USING BAYESIAN AND CLASSICAL APPROACHES

4.1 Introduction

The objective of this chapter is to simulate aggregate insurance claims using both the Classical and Bayesian approaches, compare the results, and visualize the differences. By conducting this comparison, we aim to understand the impact of incorporating prior knowledge and updating parameters on the estimation of aggregate claims.

4.2 Comparison of Bayesian and Classical Approaches for Simulating Aggregate Insurance Claims

4.2.1 Poisson and Exponential Distributions

We will employ a Poisson Distribution to model the number of claims and an Exponential Distribution to model the amount of each claim. For the Bayesian approach, we will use Gamma Distributions as priors for the Poisson and Exponential parameters. The parameters will be updated based on the observed data, and the posterior Distributions will be used for simulation.

- **Bayesian Approach :**

  Prior Distributions:

  - Poisson parameter (λ): Gamma(2, 1)
  - Exponential parameter (θ): Gamma(2, 1)

  Posterior Distributions:

  - For λ: Gamma$(\alpha + \sum x_i,\, \beta + n)$
  - For θ: Gamma$(\alpha + \sum x_i,\, \beta + n)$

- **Classical Approach :**

  Fixed Parameters:

  - Poisson parameter (λ): 2 ($\alpha/\beta = 2/1$)

– Exponential parameter (θ): 2 ($\alpha/\beta = 2/1$)

Results and Discussion

The table below summarizes the mean and standard deviation of the simulated aggregate claims for both the Bayesian and Classical approaches.

Table 4.1: Mean and standard deviation of Exponential Distribution

Approach	Mean Aggregate Claims	Standard Deviation
Classical	368,99	17,54
Bayesian	421,83	23,97

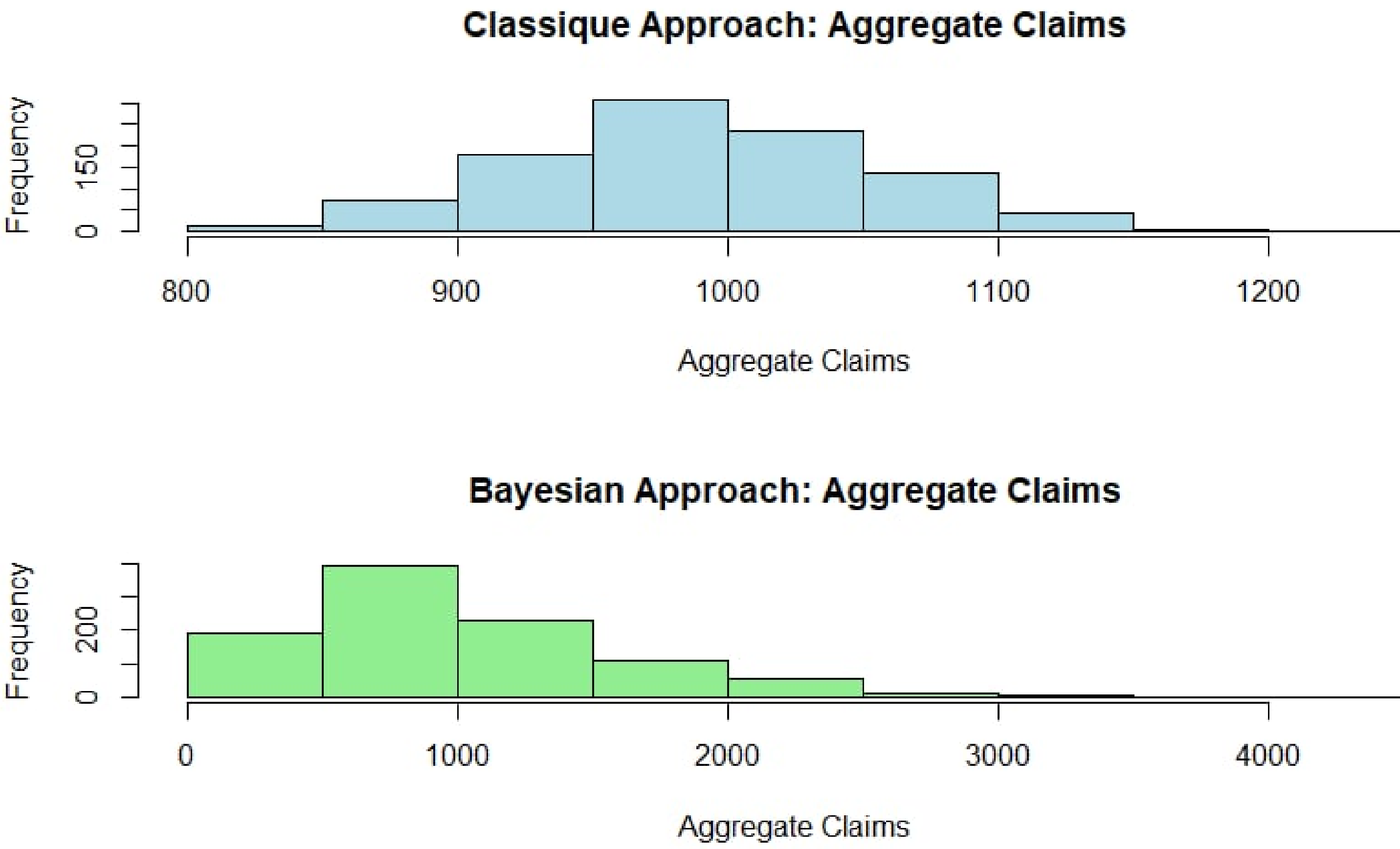

Figure 4.1: Density of aggregate claim with bayesian and Classical approach

4.2.2 Poisson and Gamma Distributions:

Distributions and parameters used:

- **Classical approach:**
  - Poisson Distribution (Λ) : the average number of claims
  $$\Lambda = 3$$
  - Gamma distribution

Parameters:

- Classical shape: $\alpha = 3$
- Classical rate: $\beta = 2$

- **Bayesian Approach:**
  Gamma Distributions are used in claim amount modeling.
  Parameters:
  - Prior Gamma shape: $\alpha = 3$
  - Prior Gamma rate: $\beta = 2$

Table 4.2: Mean of number of claims with Poisson distribution and Gamma prior distribution

Approach	Mean	Standard Deviation
Classical	899.45	1.32
Bayesian	613.90	13.58

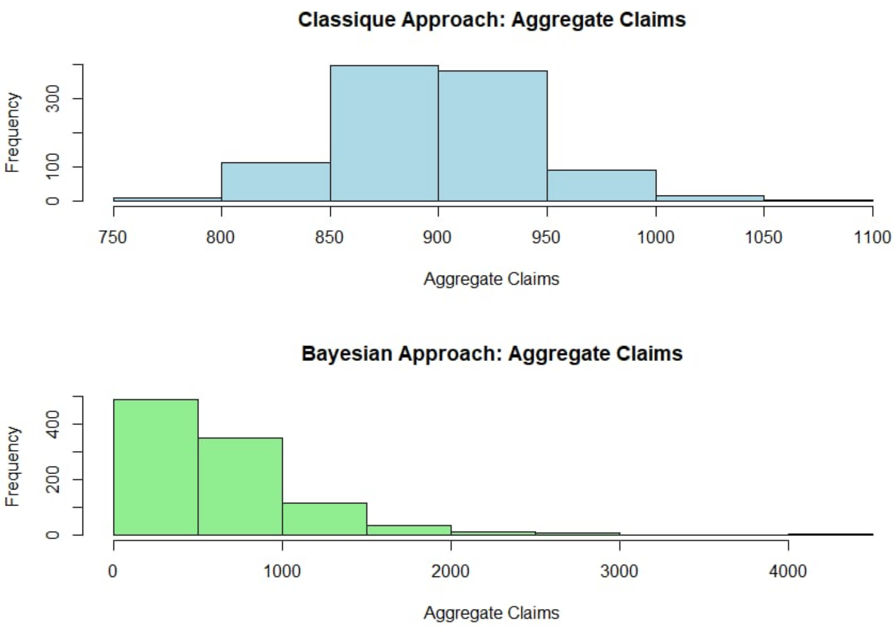

Figure 4.2: frequency of Aggregate claim by Gamma Distribution

Table 4.3: Comparison of Mean and Standard Deviation: Classical vs Bayesian Approach by poisson Distribution

Approach	Mean	Standard Deviation
Classical	598.45	92.45
Bayesian	614.90	361.58

4.2.3 Aggregate with Poisson Distributions

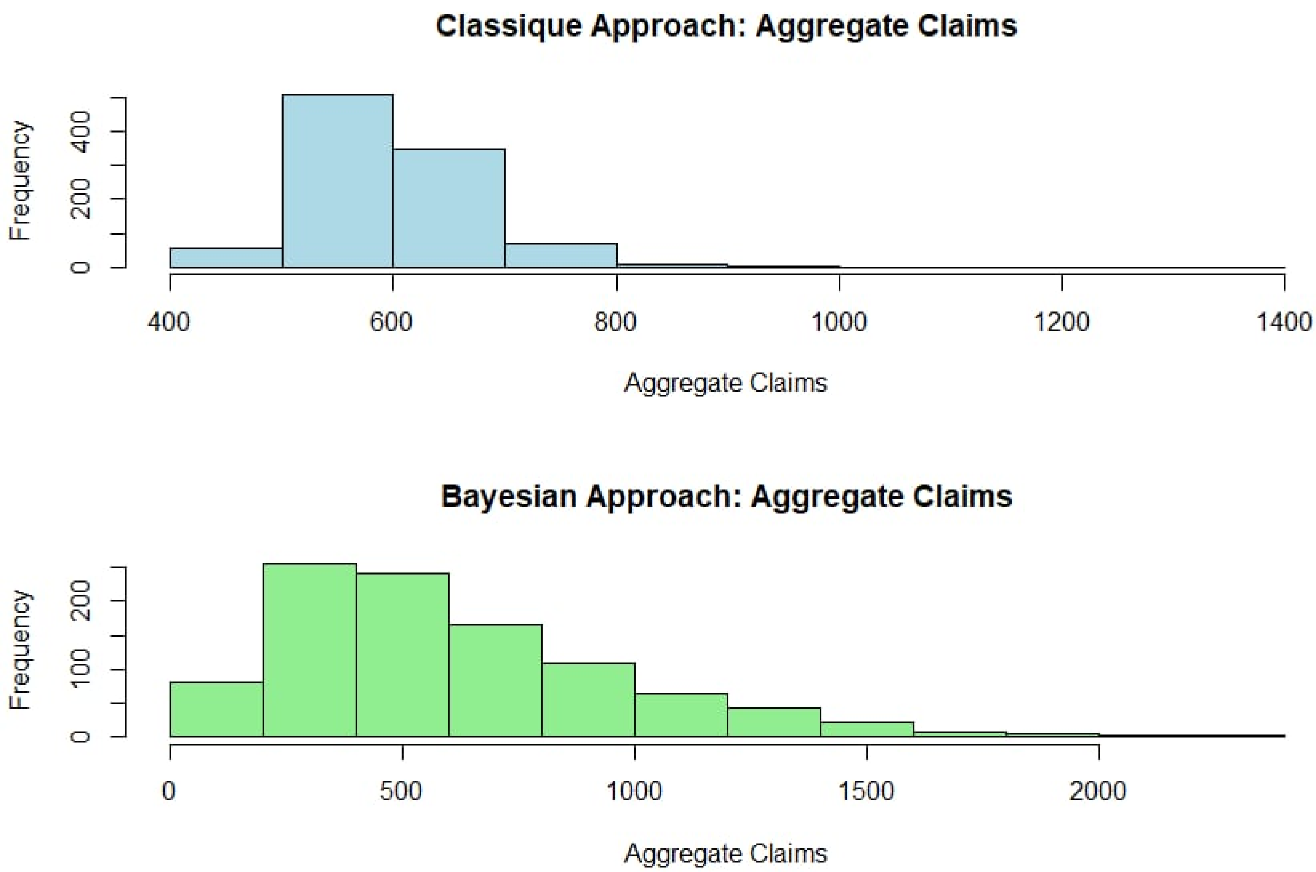

Figure 4.3: frequency of Aggregate claim by poisson Distribution

4.2.4 Aggregate with Log normal Distributions

log normal:

Table 4.4: Comparison of Mean and Standard Deviation: Classical vs Bayesian Approach by log normal Distribution

Approach	Mean	Standard Deviation
Classical	987.40	68.70
Bayesian	1002.62	361.67

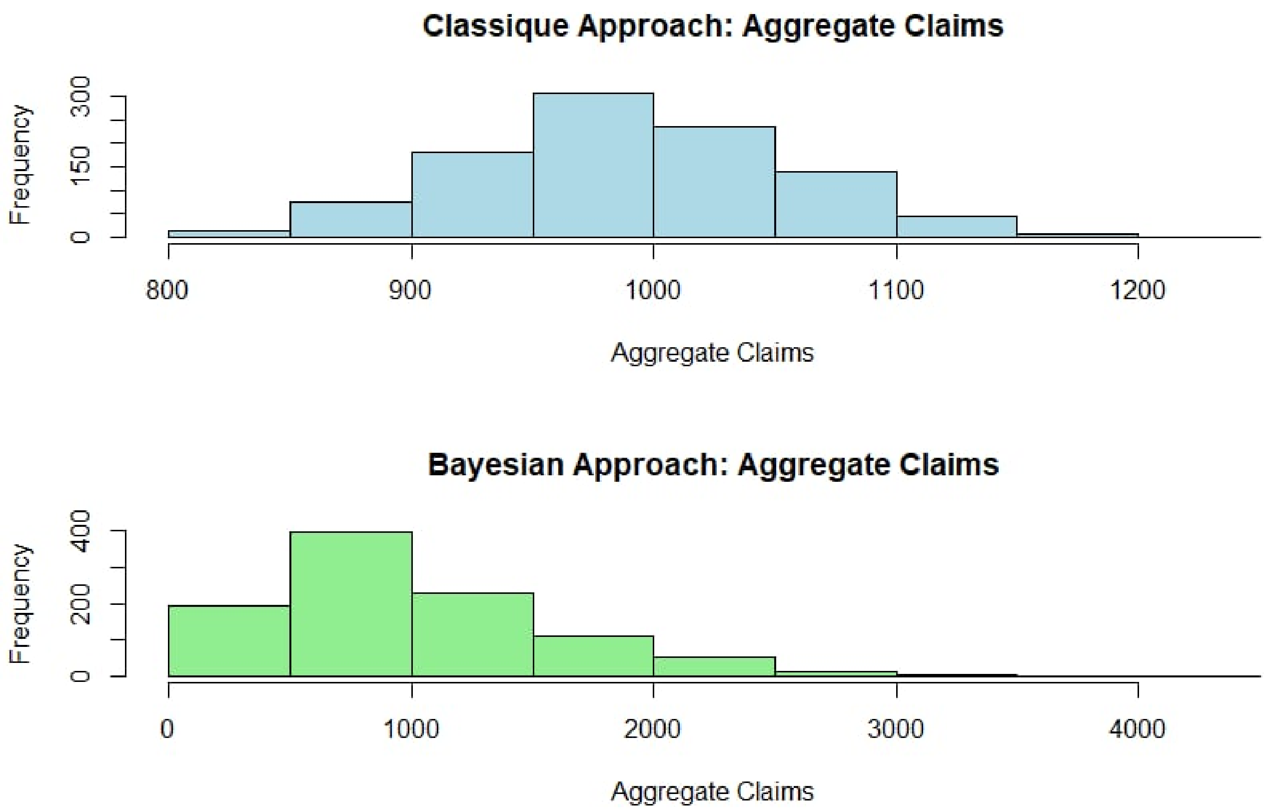

Figure 4.4: frequency of Aggregate claim by log normal Distribution

4.3 Adjustment of Data by Kolmogorov-Smirnov test

Our analysis focuses on the Besecura dataset from the CASdatasets R package, which contains individual automobile loss amounts exceeding 1.2 million euros from 1988 to 2001.

With 371 claims recorded over 14 years, the dataset, adjusted to 2002 euros, provides yearly claim counts and individual loss amounts.

The dataset's small size, with only 14 yearly claim counts, allows for different results between Bayesian and classical methodologies, aligning well with our research.

A summary captures the annual claim counts and the highest and lowest loss amounts for each year in millions of euros. Notably, 1990 had the highest loss at 7.899 million euros, and 1991 the lowest at 1.208 million euros.

The least number of claims occurred in 2001, while 1995 had the most. A graph plots individual loss amounts against their respective years, and a table details sample means and variances for deeper understanding.

Table 4.5: Summary of Bescura dataset.

Year	Number of Claims	Maximum Loss	Minimum Loss	The average between the maximum and minimum loss
1988	13	6.925	1.231	4.0780
1989	15	3.586	1.220	2.4030
1990	20	7.899	1.212	4.5555
1991	37	7.487	1.208	4.3475
1992	31	4.179	1.221	2.7000
1993	29	6.685	1.240	3.9625
1994	20	5.343	1.318	3.3305
1995	44	2.988	1.276	2.1320
1996	36	5.093	1.359	3.2260
1997	36	4.964	1.463	3.2135
1998	33	3.435	1.538	2.4865
1999	25	4.051	1.634	2.8425
2000	25	4.147	1.632	2.8895
2001	7	2.956	1.661	2.3085

4.3.1 Adjustment the number of claims by Data the Poisson Distribution:

Table 4.6: Summary of the Kolmogorov-Smirnov test with Poisson Distribution

Mean of number of claims	**Test statistic (D)**	**P-value**	**Alternative hypothesis**
26.5	0.2075051	0.5829501	two-sided

the p-value is greater than or equal to 0.05, suggesting no significant evidence against the null hypothesis. Therefore, the observed Data could be considered to follow a Poisson Distribution with $\lambda = 26.5$.

Table 4.7: Summary of vector C

Statistic	**Value**
Min.	2.956
1st Qu.	4.051
Median	4.964
Mean	149.805
3rd Qu.	6.685
Max.	3586.000

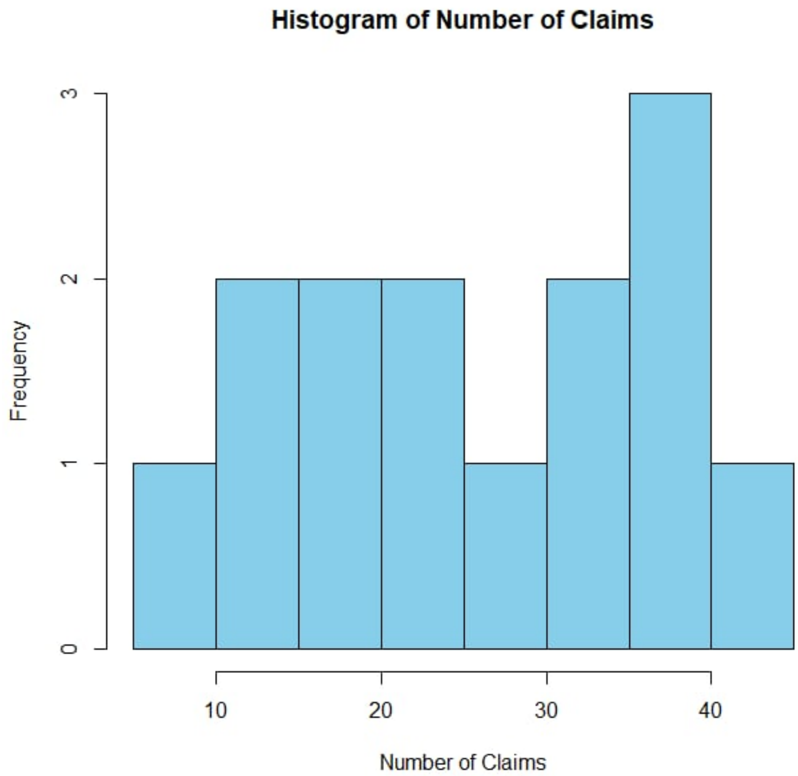

Figure 4.5: Histogram of number of claims

4.3.2 Calculation of the Expected Number of Claims $E(N)$

In insurance risk analysis, the number of claims N is often modeled using a Poisson Distribution. Here, $E(N)$ represents the expected value or mean of the number of claims.

-Expected Value of Poisson Distribution:

The expected value $E(N)$ of a Poisson-distributed random variable N is equal to its parameter λ:

$$E(N) = \lambda$$

This means that for our case, where $E(N) = 26.5$, the average number of claims per interval is 26.5.

Conclusion:

Given $E(N) = 26.5$, we interpret this as the expected number of claims in a specified period or context, assuming a Poisson Distribution where the rate parameter λ is 26.5.

4.3.3 Adjustment of Minimum Loss Values with Different Distributions

- **Exponential Distribution**
- **Estimated Parameters**: λ
- **Gamma Distribution**
- **Estimated Parameters**: α, β
- **Weibull Distribution**

- **Estimated Parameters**: k, λ
- **Lognormal Distribution**
- **Estimated Parameters**: μ, σ

Table 4.8: Summary of Kolmogorov-Smirnov tests for Distributions fitted to vector F generated between number of claims and minimum loss values.

Distribution	**Goodness of fit Results**
Exponential	Test statistic (D): 0.0384 P-value: 0.6433 Interpretation: No significant evidence against the null hypothesis. Therefore, vector F_{exp} could be considered to follow an Exponential Distribution.
Lognormal	Test statistic (D): 0.0259 P-value: 0.9652 Interpretation: No significant evidence against the null hypothesis. Therefore, vector $F_{\text{lognormal}}$ could be considered to follow a Lognormal Distribution.
Gamma	Test statistic (D): 0.2144 P-value: 3.11×10^{-15} Interpretation: Significant evidence against the null hypothesis. Therefore, vector F does not follow a Gamma Distribution.
Weibull	Test statistic (D): 0.2115 P-value: 7.77×10^{-15} Interpretation: Significant evidence against the null hypothesis. Therefore, vector F does not follow a Weibull Distribution.

Table 4.9: Summary statistics of vector F

Summary	**Value**
Min.	2.132
1st Qu.	2.486
Median	3.213
Mean	3.163
3rd Qu.	3.962
Max.	4.556

The vector F was generated as a sample where each minimum loss value Q_i is repeated according to the corresponding number of claims N_i, thus creating an aggregate representation of the loss Data based on the given number of claims and their respective minimum loss values.

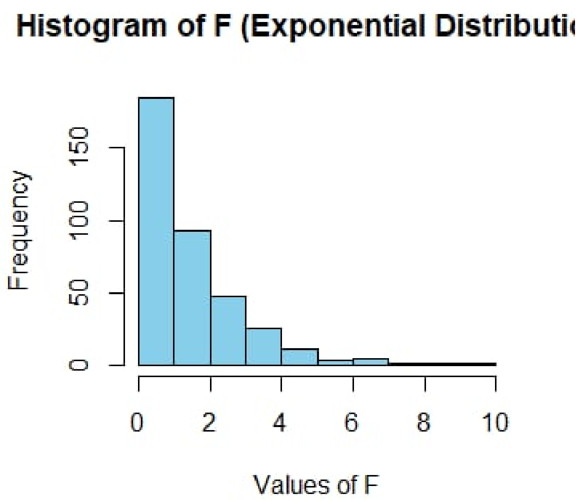

Figure 4.6: Histogram of F with exponential Distribution

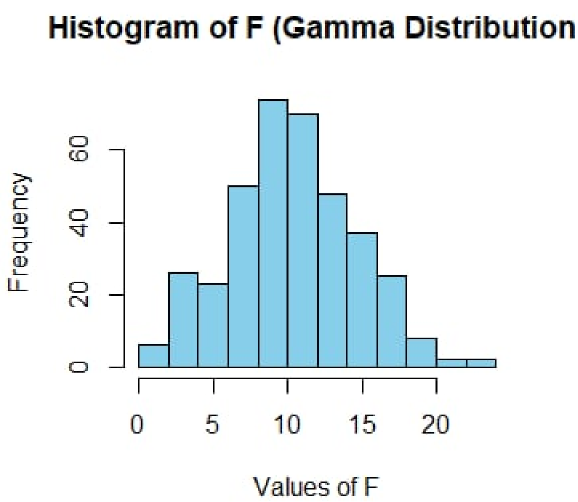

Figure 4.7: Histogram of F Gamma Distribution

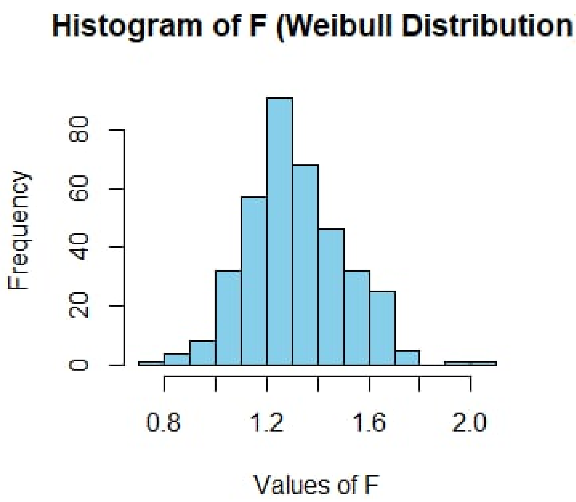

Figure 4.8: Histogram of F weibull Distribution

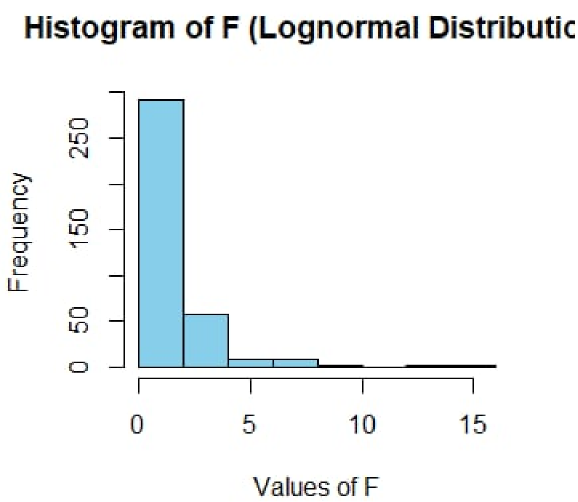

Figure 4.9: Histogram of F lognormal Distribution

4.3.4 Adjustment of maximum Loss Values with Different Distributions

Table 4.10: Summary of Kolmogorov-Smirnov tests for Distributions fitted to vector F generated between number of claims and minimum loss values.

Distribution	**Summary**
Exponential	Test statistic (D): 0.04980428 P-value: 0.3162194 Interpretation: The p-value is greater than or equal to 0.05, suggesting no significant evidence against the null hypothesis. Therefore, vector $F_{\text{exponential}}$ could be considered to follow an Exponential Distribution.
Lognormal	Test statistic (D): 0.06851771 P-value: 0.06140082 Interpretation: The p-value is greater than or equal to 0.05, suggesting no significant evidence against the null hypothesis. Therefore, vector $F_{\text{lognormal}}$ could be considered to follow a Lognormal Distribution.
Gamma	Test statistic (D): 0.004238857 P-value: 0.5176068 Interpretation: The p-value is greater than or equal to 0.05, suggesting no significant evidence against the null hypothesis. Therefore, vector F_{Gamma} could be considered to follow a Gamma Distribution.

Table 4.11: Summary statistics of vector F

Summary Statistics of F	
Minimum (Min.)	2.132
1st Quartile (1st Qu.)	2.486
Median	3.213
Mean	3.163
3rd Quartile (3rd Qu.)	3.962
Maximum (Max.)	4.556

The vector F was generated as a sample where each maximum loss value Q_i is repeated according to the corresponding number of claims N_i, thus creating an aggregate representation of the loss Data based on the given number of claims and their respective maximum loss values.

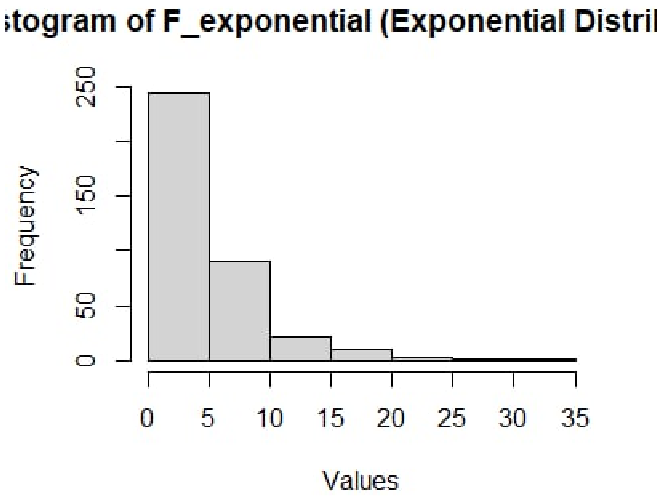

Figure 4.10: Histogram of maximum loss values with exponential Distribution

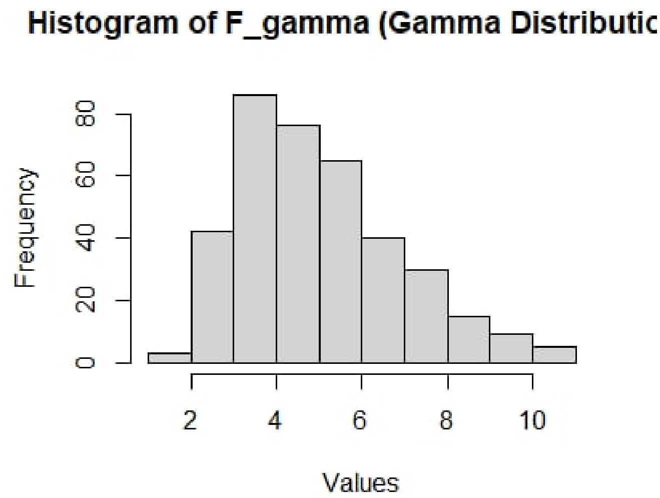

Figure 4.11: Histogram of maximum loss values with Gamma Distribution

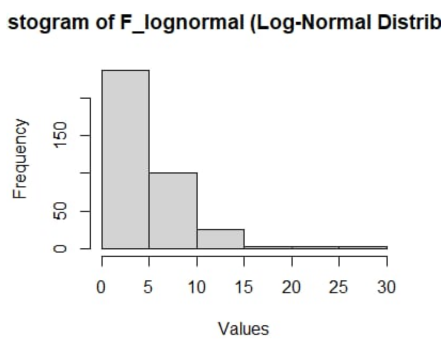

Figure 4.12: Histogram of maximum loss values with lognormal Distribution

4.3.5 Adjustment of Average Loss Values with Different Distributions

Table 4.12: Summary of Kolmogorov-Smirnov tests for Distributions fitted to vector F generated between the number of claims and minimum loss values.

Distribution	Summary
Exponential	Test statistic (D):0.05132519 P-value: 0.2824583 Interpretation: The p-value is greater than or equal to 0.05, suggesting no significant evidence against the null hypothesis. Therefore, vector $F_{\text{exponential}}$ could be considered to follow an Exponential Distribution.
Lognormal	Test statistic (D): 0.0514404 P-value: 0.2800102 Interpretation: The p-value is greater than or equal to 0.05, suggesting no significant evidence against the null hypothesis. Therefore, vector $F_{\text{lognormal}}$ could be considered to follow a Lognormal Distribution.
Gamma	Test statistic (D): 0.02453528 P-value: 0.9788374 Interpretation: The p-value is greater than or equal to 0.05, suggesting no significant evidence against the null hypothesis. Therefore, vector F_{Gamma} could be considered to follow a Gamma Distribution.
Weibull	Test statistic (D): 0.0332375 P-value: 0.8070209 Interpretation: The p-value is greater than or equal to 0.05, suggesting no significant evidence against the null hypothesis. Therefore, vector F_{weibull} could be considered to follow a Weibull Distribution.

Table 4.13: Summary of the descriptive statistics for the vector F.

Statistic	Value
Min	1.208
1st Quartile	1.221
Median	1.276
Mean	1.362
3rd Quartile	1.463
Max	1.661

The vector F was generated as a sample where each average loss value Q_i is repeated according to the corresponding number of claims N_i, thus creating an aggregate representation of the loss Data based on the given number of claims and their respective average loss values.

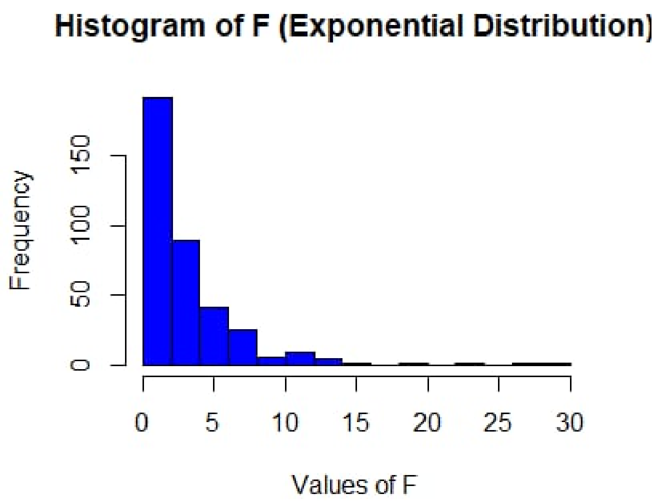

Figure 4.13: Histogram of average loss with exponential Distribution

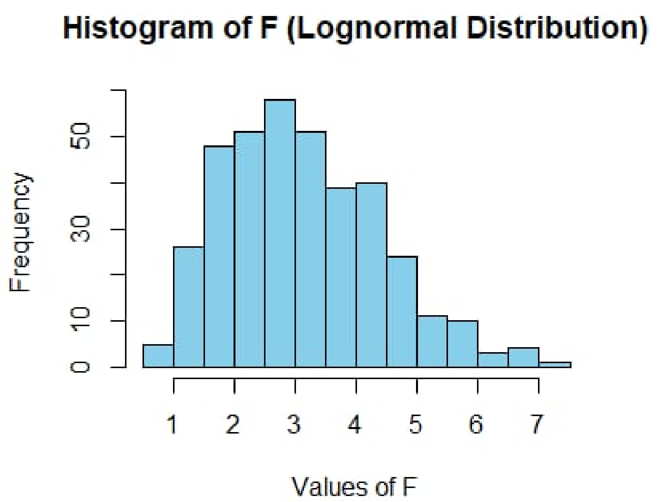

Figure 4.14: Histogram of average loss with lognormal Distribution

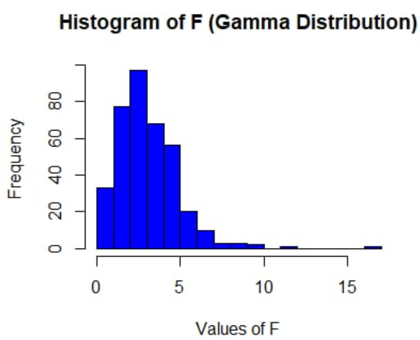

Figure 4.15: Histogram of average loss with Gamma Distribution

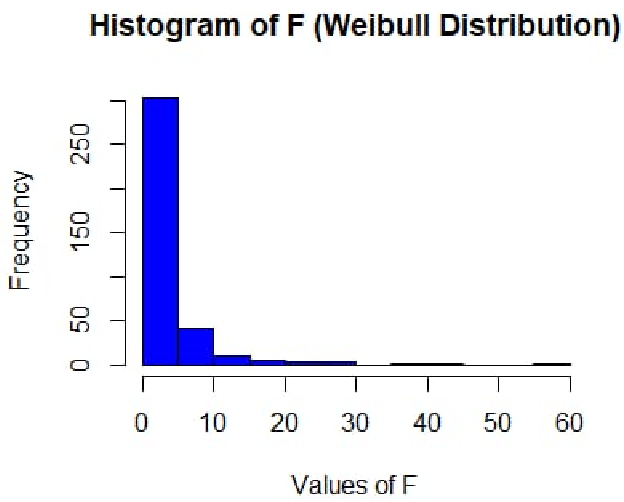

Figure 4.16: Histogram of average loss with weibull Distribution

4.3.6 Mean Calculation for Exponential Distribution

- The mean $E(X)$ for the exponential Distribution, estimated from the Data, is given by:

$$E(X) = \frac{1}{\hat{\lambda}}$$

  where $\hat{\lambda}$ is the estimated rate parameter derived from the Data.

- Computed Mean:

$$E(X) = 2.960179$$

  This value represents the expected mean value of X based on the estimated rate parameter from the exponential Distribution.

4.3.7 Calculation of Aggregate Claim Amount $E(S)$

In insurance risk analysis, the aggregate claim amount S is calculated as the product of the mean number of claims $E(N)$ and the mean claim amount $E(X)$.

-Given:

- $E(N) = 26.5$ (mean number of claims, assumed to follow a Poisson Distribution)
- $E(X) = 2.960179$ (mean of the exponential Distribution of claim amounts)

-Calculation:

$$E(S) = E(N) \times E(X)$$

$$E(S) = 26.5 \times 2.960179$$

$$E(S) \approx 78.47$$

Therefore, the expected aggregate claim amount $E(S)$ is approximately 78.47.

4.4 Simulating aggregate insurance claims with Bayesian approach

4.4.1 Bayesian Inference and Simulation for Comparing Real and Simulated Number of Claims Data Means

We simulate and compare the means of real and simulated Data using Bayesian inference for a Poisson Distribution.
The real Data, representing the number of claims, is generated from a Poisson Distribution with a specified lambda.
The prior Distribution parameters (alpha and beta) are used to generate samples from a Gamma Distribution, which serve as the rate parameter for generating simulated Data from the Poisson Distribution.
We then compare the means of the real and simulated Data and visualize the results using a histogram. Additionally, we calculate the posterior mean of the rate parameter based on the real Data representing the number of claims.

Table 4.14: Comparison of Real and Simulated Data Means for Different Alpha and Beta Parameters

Alpha	Beta	Mean Real Data	Mean Simulated Data
25	1	26.5	26.74
30	12	26.5	2.32

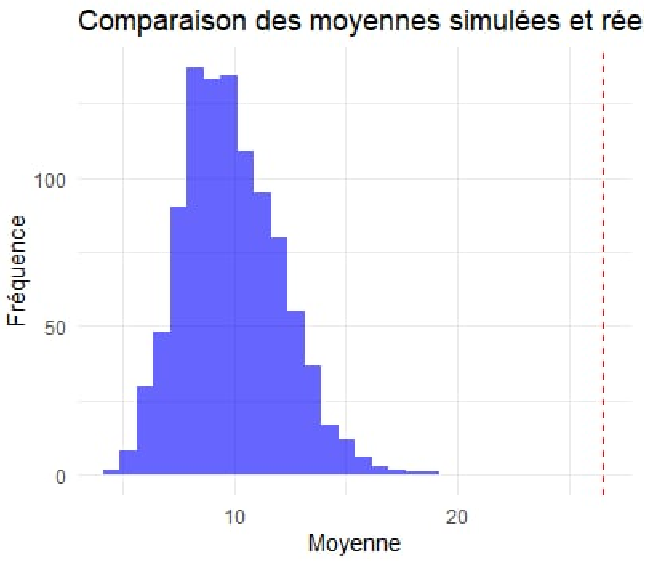

Figure 4.17: Comparison of Real and Simulated Data Means for $\alpha = 20$ and $\beta = 2$

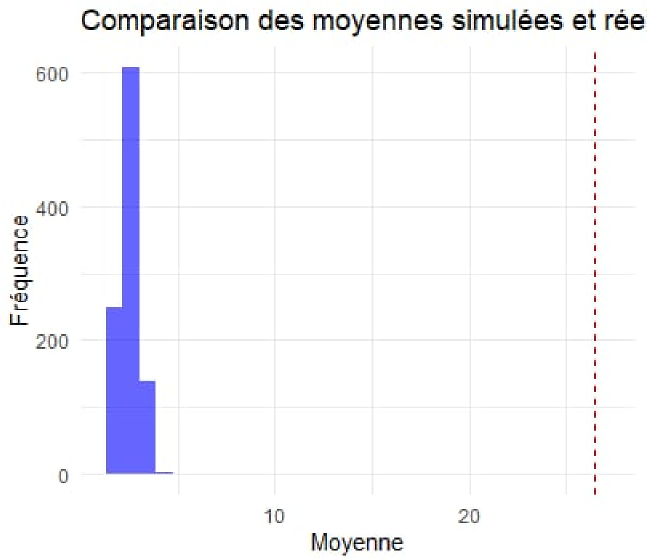

Figure 4.18: Comparison of Real and Simulated Data Means for $\alpha = 30$ and $\beta = 12$

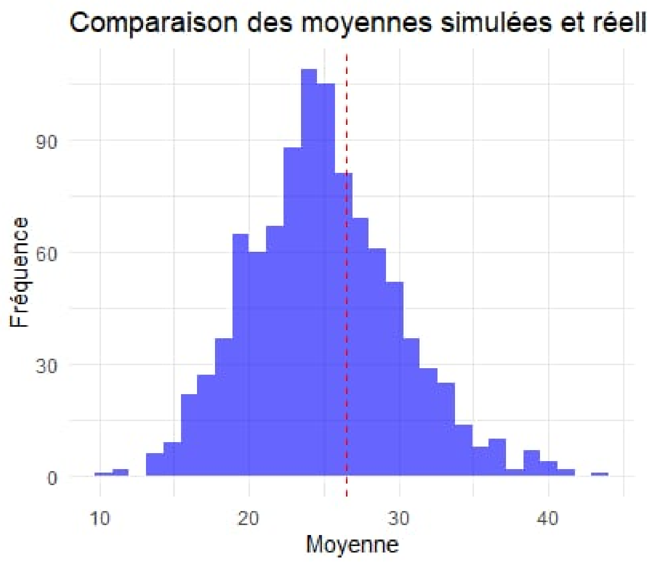

Figure 4.19: Comparison of Real and Simulated Data Means for $\alpha = 25$ and $\beta = 1$

4.4.2 Estimation of Posterior Parameters for Poisson Distribution with MCMC

Table 4.15: Posterior Estimates from Poisson Distribution

Parameter	Value
Posterior Mean of λ	24.86667
Posterior Variance of λ	1.657778

Explanation:

The provided R code estimates the posterior parameters of a Poisson Distribution using Bayesian inference:

- **Posterior Mean of λ:** Represents the estimated mean of the rate parameter λ, which characterizes the Poisson Distribution.

- **Posterior Variance of λ:** Shows the estimated variance of the rate parameter λ.

This Bayesian approach updates prior beliefs (Gamma Distribution as priors for λ) based on observed data (vector `observations`), providing posterior Distributions for the parameter λ.

4.4.3 Estimation of Posterior Parameters for Negative Binomial Distribution with MCMC

Table 4.16: Posterior Estimates from Negative Binomial Distribution

Parameter	Value
Posterior Mean of r	24.86667
Posterior Variance of r	1.657778
Posterior Mean of p	20.72222
Posterior Variance of p	-21.50991

1. **Explanation:**

   The provided R code estimates the posterior parameters of a negative binomial Distribution using Bayesian inference:

   This Bayesian approach updates prior beliefs (Gamma and Beta Distributions as priors for r and p respectively) based on observed data (vector

2. **Observations:** providing posterior Distributions for the parameters r and p.

3. **-Bayesian Posterior Mean of Number of Claims:** The expected value (mean) $E(N)$ using the Bayesian approach with a Gamma prior and Poisson likelihood is calculated as:

   $$E(N) = \frac{\alpha + \sum x_i}{\beta + n}$$

   where:

   - α and β are the shape and rate parameters of the Gamma prior Distribution for λ,
   - $\sum x_i$ is the sum of the observed data points,
   - n is the number of observations.

   This formula represents the expected number of claims N based on the Bayesian estimation incorporating both prior assumptions and observed data. The expected value (mean) $E(N)$ is calculated as:

   $$E(N) = 26.4$$

   .

Table 4.17: Comparison of Mean Real Data and Mean Simulated Data for Different Alpha and Beta Values

Alpha	Beta	Mean Real Data	Mean Simulated Data
2	1	1.361663	1.229292
0.5	3	1.361663	23.44314
1	1	1.361663	4.132548

4.5 Bayesian Inference and Simulation for Comparing Real and Simulated Amount of Claims Data Means

For minimum loss:

1. **For Exponential Distributions:**

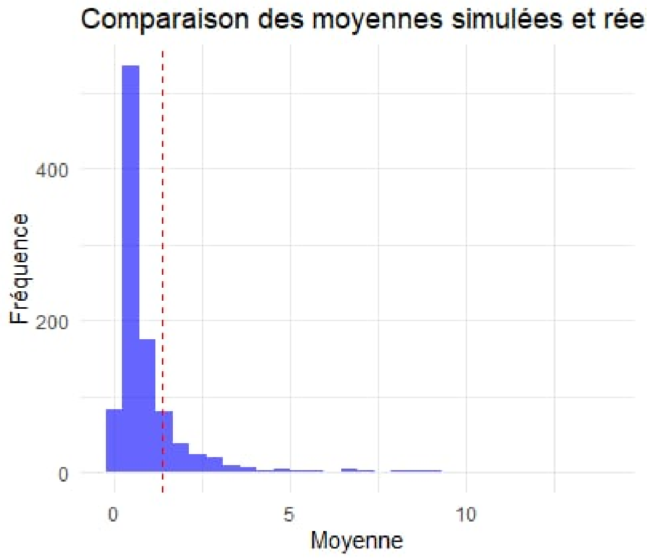

Figure 4.20: Comparison of Real and Simulated Data Means for $\alpha = 2$ and $\beta = 1$

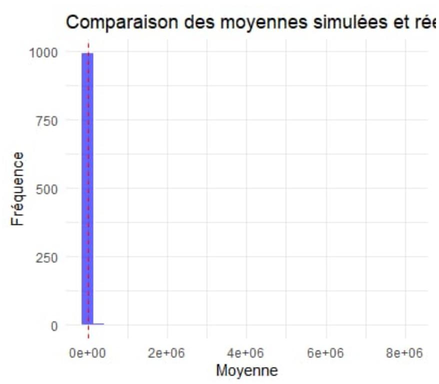

Figure 4.21: Comparison of Real and Simulated Data Means for $\alpha = 0.5$ and $\beta = 3$

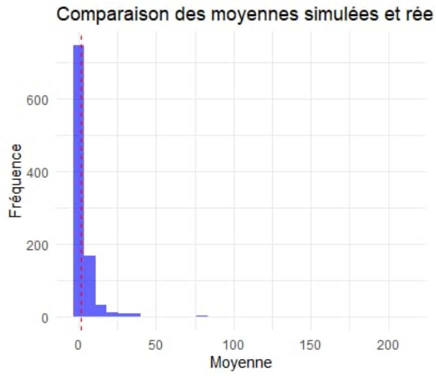

Figure 4.22: Comparison of Real and Simulated Data Means for $\alpha = 1$ and $\beta = 1$

The table above presents a comparison of the mean real data and mean simulated data for different values of the parameters α and β used in the Gamma Distribution prior for Bayesian inference.

For $\alpha = 2$ and $\beta = 1$, the mean of the simulated data (1.229292) is quite close to the mean of the real data (1.361663), indicating a good fit between the model and the observed data.

However, for $\alpha = 0.5$ and $\beta = 3$, there is a significant discrepancy between the mean of the simulated data (23.44314) and the mean of the real data (1.361663). This suggests that this set of prior parameters does not fit the observed data well, leading to a considerable overestimation.

Similarly, for $\alpha = 1$ and $\beta = 1$, the mean of the simulated data (4.132548) is higher than the mean of the real data (1.361663), though the discrepancy is less extreme than in the previous case. This indicates that while the model with these parameters captures the data trend to some extent, it still overestimates the mean number of claims.

These results underscore the importance of selecting appropriate prior parameters in Bayesian inference to ensure that the posterior Distribution accurately reflects the observed data.

2. **Log-normal Distributions:**

Table 4.18: Comparison of Mean Real Data and Mean Simulated Data for Different Alpha and Beta Values

Alpha	Beta	Mean Real Data	Mean Simulated Data
2	1	1.361663	7.09912
2	2	1.361663	1.958708
6	2	1.361663	28.54447

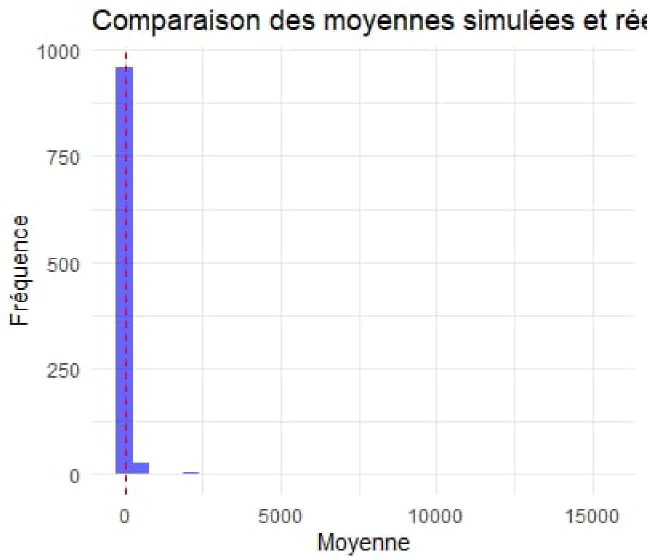

Figure 4.23: Comparison of Real and Simulated Data Means for $\alpha = 6$ and $\beta = 2$

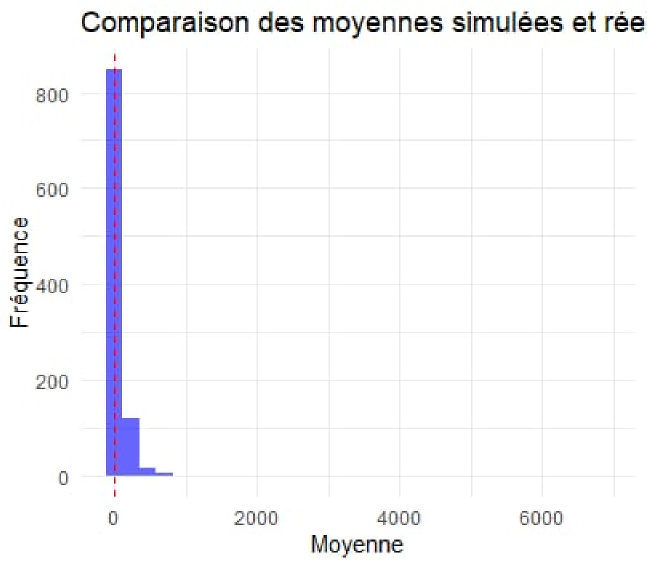

Figure 4.24: Comparison of Real and Simulated Data Means for $\alpha = 2$ and $\beta = 1$

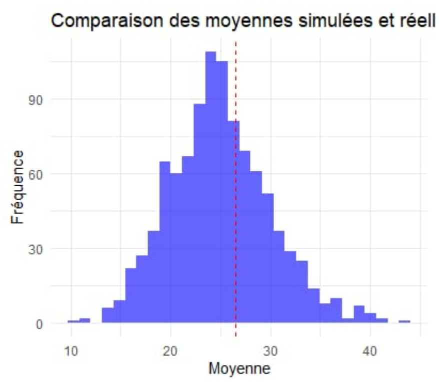

Figure 4.25: Comparison of Real and Simulated Data Means for $\alpha = 2$ and $\beta = 2$

Comments:

1. **Alpha = 2, Beta = 1:** The simulated mean (7.09912) is notably higher than the real mean (1.361663). This discrepancy suggests that with a lower Beta, there is a tendency to overestimate the parameter.

2. **Alpha = 2, Beta = 2:** The simulated mean (1.958708) is closer to the real mean (1.361663) compared to Alpha = 2, Beta = 1. This indicates a better fit of the simulated data with a slightly higher value of Beta.

3. **Alpha = 6, Beta = 2:** The simulated mean (28.54447) is much higher than the real mean (1.361663), highlighting a significant overestimation. This is likely due to the high value of Alpha and a moderate Beta, leading to a wider spread in simulated data.

4.5.1 For maximum loss:

1. **Exponential Distribution:**

Table 4.19: Summary of Mean Values for different α and β

Alpha	Beta	Mean Real Data	Mean Simulated Data
2	1	6.844756	7.171678
2	2	6.844756	1.071517
10	3	6.990864	0.330146

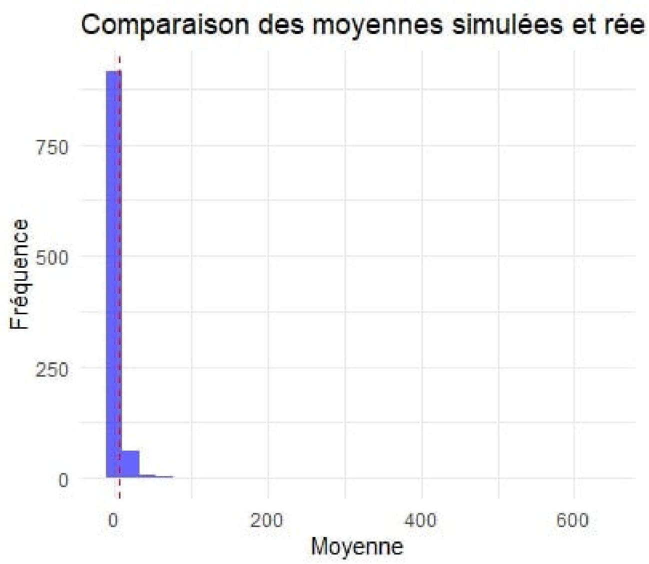

Figure 4.26: Comparison of Real and Simulated Data Means for $\alpha = 2$ and $\beta = 2$

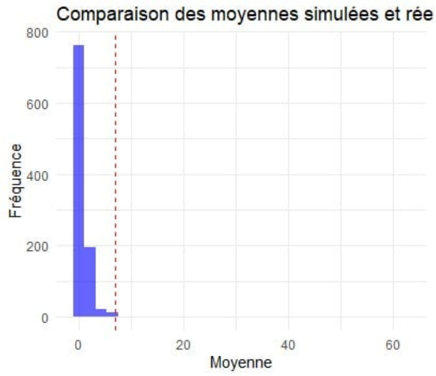

Figure 4.27: Comparison of Real and Simulated Data Means for $\alpha = 2$ and $\beta = 1$

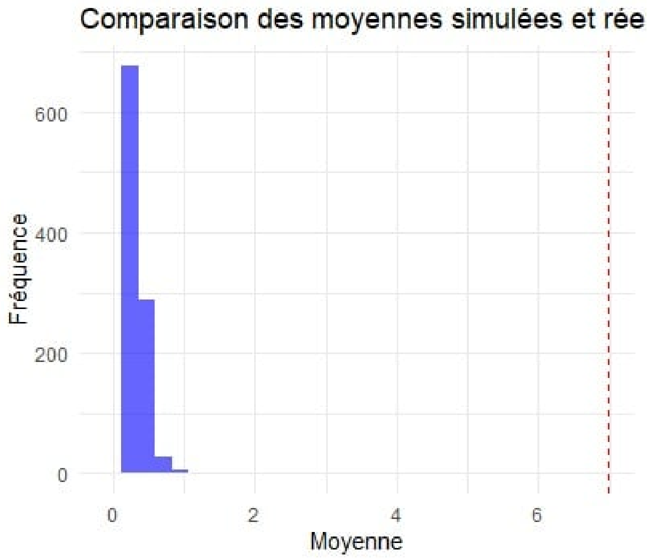

Figure 4.28: Comparison of Real and Simulated Data Means for $\alpha = 10$ and $\beta = 3$

2. **For Gamma Distribution:**

Table 4.20: Summary of Mean Values for Different Parameter Sets

Alpha	Beta	Mean Real Data	Mean Simulated Data
10	3	6.848152	1.120329
1	2	6.916385	4.994823
1	1	6.883371	6.535302

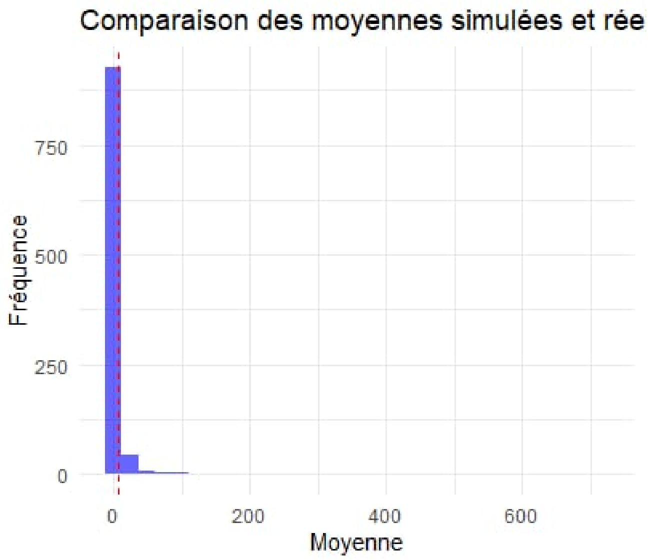

Figure 4.29: Comparison of Real and Simulated Data Means for $\alpha = 1$ and $\beta = 1$

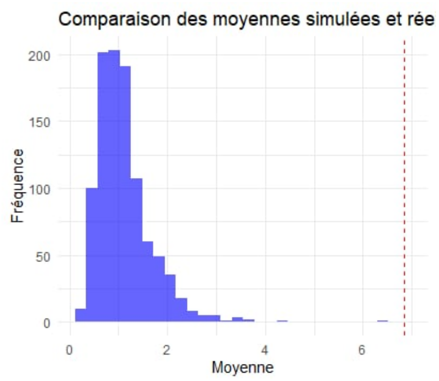

Figure 4.30: Comparison of Real and Simulated Data Means for $\alpha = 10$ and $\beta = 3$

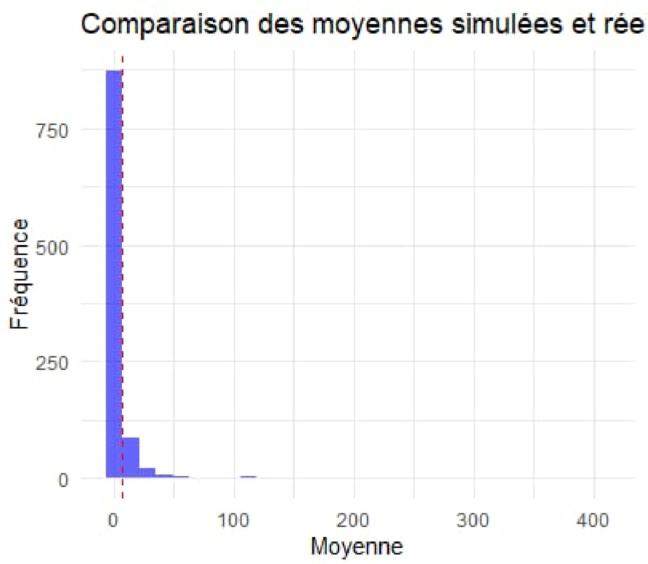

Figure 4.31: Comparison of Real and Simulated Data Means for $\alpha = 1$ and $\beta = 2$

For Average loss :

-Gamma Distribution

Table 4.21: Comparison of Real and Simulated Data Means for Different Alpha and Beta Values

Alpha	Beta	Mean Real Data	Mean Simulated Data
2	1	3.162784	5.782636
20	10	3.162784	3.365112
10	1	3.162784	0.7072573

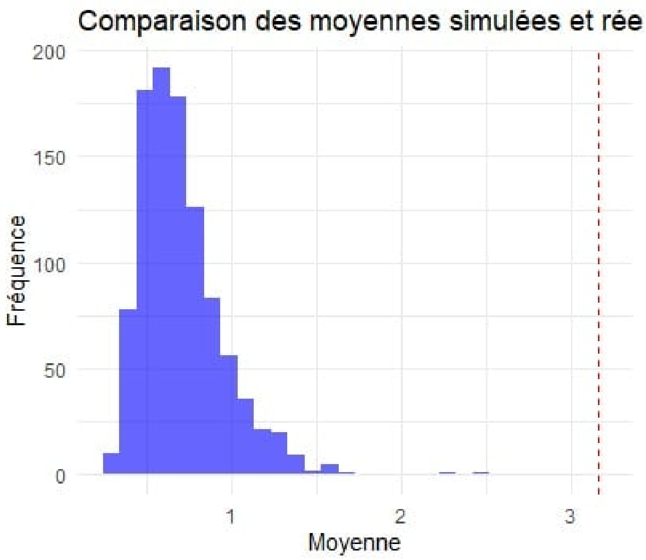

Figure 4.32: Comparison of Real and Simulated Data Means for $\alpha = 10$ and $\beta = 1$

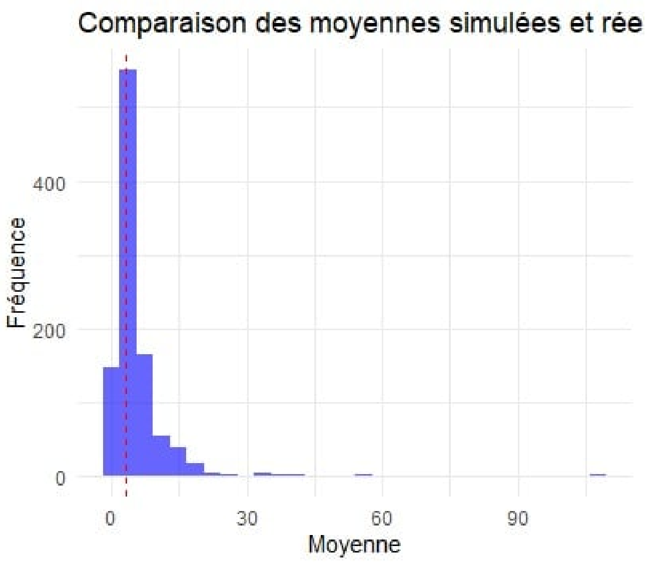

Figure 4.33: Comparison of Real and Simulated Data Means for $\alpha = 2$ and $\beta = 1$

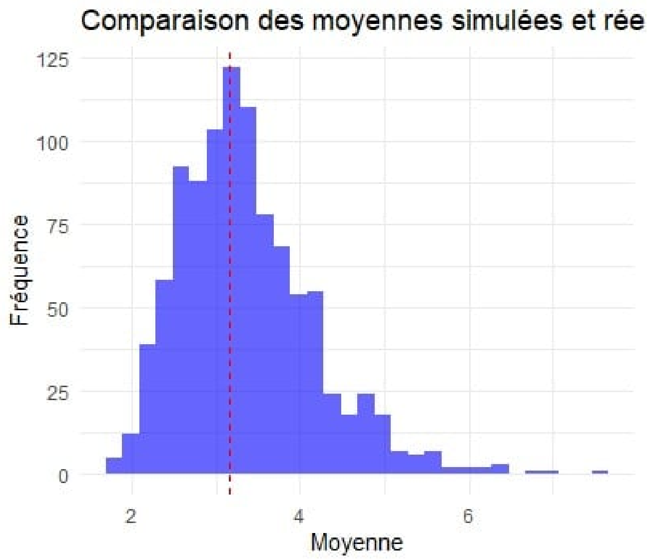

Figure 4.34: Comparison of Real and Simulated Data Means for $\alpha = 20$ and $\beta = 10$

-Exponential disribution

Table 4.22: Summary of Mean Values for Different Parameter Sets

Alpha	Beta	Mean Real Data	Mean Simulated Data
2	1	3.162784	3.130718
1	3	3.162784	11.1317
10	1	3.162784	0.3504516

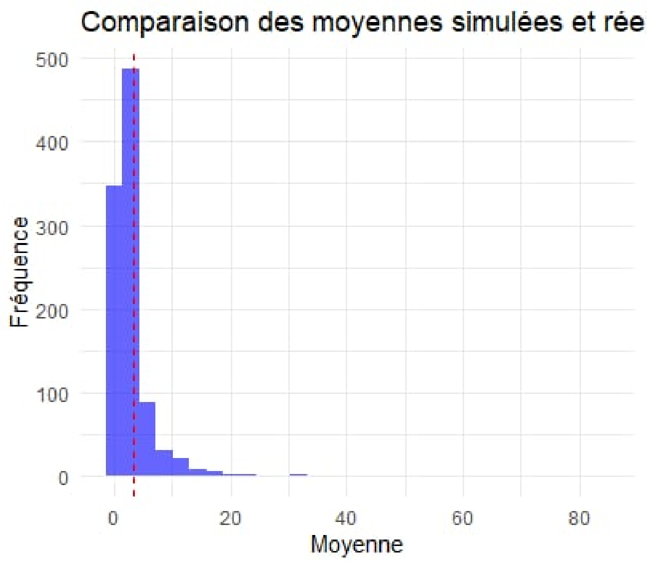

Figure 4.35: Comparison of Real and Simulated Data Means for $\alpha = 2$ and $\beta = 1$

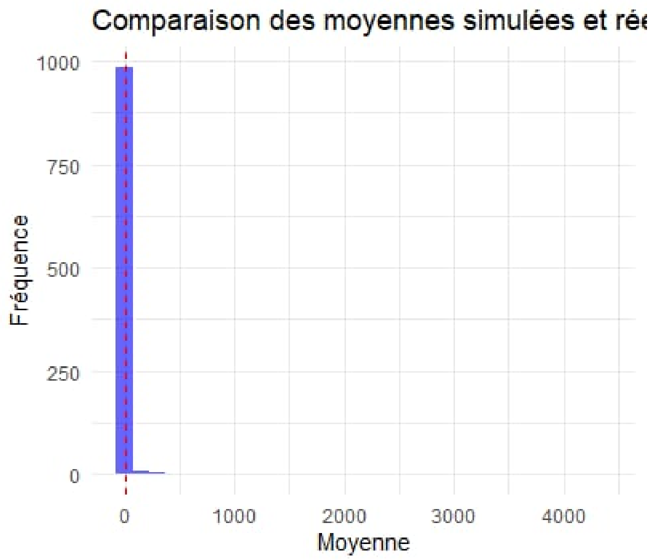

Figure 4.36: Comparison of Real and Simulated Data Means for $\alpha = 1$ and $\beta = 3$

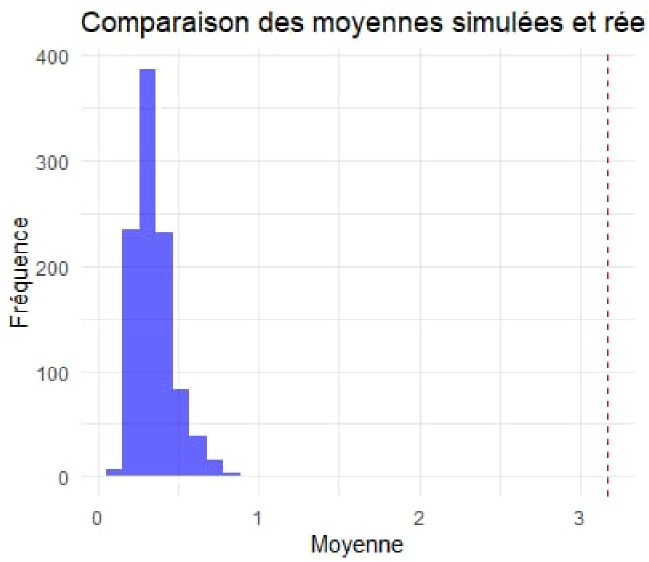

Figure 4.37: Comparison of Real and Simulated Data Means for $\alpha = 10$ and $\beta = 1$

-Bayesian Posterior Mean of amount of Claims:

Calculate Posterior Mean:

- We use a Gamma Distribution as the prior for the rate parameter λ:
$$\lambda \sim \text{Gamma}(\alpha, \beta)$$
Here, α and β are the shape and rate parameters of the Gamma Distribution, respectively.
- Given n observed data points $x_1, x_2, \ldots, x_n$, the likelihood function is:
$$p(x \mid \lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$
- The posterior Distribution is also a Gamma Distribution with updated parameters:
$$\lambda \mid x \sim \text{Gamma}(\alpha + n, \beta + \sum_{i=1}^{n} x_i)$$

Here, n is the number of data points, and $\sum_{i=1}^{n} x_i$ is the sum of the observed data points.

- The mean of the Gamma Distribution is given by:

$$\mathbb{E}[\lambda \mid x] = \frac{\alpha + n}{\beta + \sum_{i=1}^{n} x_i}$$

$\mathbb{E}[\lambda \mid x] = 3.130718$

4.5.2 Calculating the Mean of Aggregate: Expected Value of Total Claims in Insurance

To calculate the expected value $E(S)$ of the total aggregate claims S, we use the following steps:

1. Calculate the expected number of claims per period, $E(N)$:
$$E(N) = 26.4$$

2. Calculate the expected size of each claim, $E(X)$:
$$E(X) = 0.3244229$$

3. Use the formula for the expected aggregate claims:
$$E(S) = E(N) \times E(X)$$

   Substituting the values:
$$E(S) = 26.4 \times 0.3244229 = 8.5605976$$

Therefore, the expected value of the total aggregate claims $E(S)$ is 8.5605976.

4.5.3 Comparison of Expected Aggregate Amount $E(S)$

Table 4.23: Comparison of Expected Aggregate Amount

Approach	$E(N)$	$E(X)$	$E(S)$
Bayesian	26.4	3.130718	82,65
Classical	26.5	2.960179	78.47

Comments on Results:

- **Bayesian Approach**: Estimates $E(S)$ to be approximately 8.55. This approach incorporates prior knowledge and updates it with observed data, resulting in a lower expected aggregate amount.

- **Classical Approach**: Estimates $E(S)$ to be approximately 78.47. This approach assumes independence between the number of claims and the amount per claim, calculating the product of their respective means.

- These results highlight the difference in expectations between the Bayesian and Classical methods, showcasing how prior beliefs and statistical assumptions can influence the estimated aggregate amount $E(S)$.

CONCLUSION

This thesis has successfully traversed the challenging realm of non-life insurance risk management , with a focus on the integration of Bayesian and classical methods and their comparative examination.

Through methodically expanding on fundamental principles of probability, we created a thorough framework for comprehending the random structure of insurance risks.

The creation of risk and ruin probability models, which offered vital insights into the financial stability of insurers, was a key component of our research.

Through the combination of frequentist and Bayesian approaches, we developed an all-encompassing framework that improves the precision and resilience of insolvency risk assessments.

These models' practical usefulness was further shown by applying them to actual data, which showed how well they could anticipate total claims and assess overall risk.

There were notable differences and similarities between the Bayesian and classical techniques compared.

Traditional approaches provide a strong and dependable foundation for risk assessment since they are based on well-established statistical principles.

Nonetheless, the incorporation of previous knowledge and the ability to continuously update with fresh data are transformational edges brought about by the advent of Bayesian techniques.

As a result of this integration, models become more dynamic and sophisticated and can adjust to changing uncertainties.

The process of risk assessment is enhanced by the Bayesian approach's capacity to offer probabilistic interpretations of outcomes, which yield more accurate and thorough forecasts.

Also an important aspect of this thesis was the implementation of simulations using R code. Through these simulations, we compared the expected aggregate amount of claims between Bayesian and classical approaches.

Our findings highlighted the significant impact of incorporating prior knowledge in Bayesian methods, resulting in more conservative and potentially more accurate risk assessments.

In conclusion, this thesis has demonstrated that Bayesian methods represent a significant advancement in the field of insurance risk management.

By integrating these methods with classical approaches, insurers can develop a more resilient and flexible framework for predicting future claims and managing risks.

This synthesis not only enhances the accuracy of risk assessments but also provides valuable insights into the potential impact of extreme events, ultimately contributing to the financial stability and robustness of the insurance industry.

BIBLIOGRAPHY

[1] A First Course in Bayesian Statistical Methods by Peter D. Hoff

[2] Le Choix Bayésien: Principes et Pratique by Christian P. Robert

[3] Modern Actuarial Risk Theory by Rob Kaas ,Marc Goovaerts , Jan Dhaene and Michel Denuit

[4] Actuarial Theory for Dependent Risks Measures, Orders and Models by M. Denuit , J. Dhaene , M. Goovaerts , R. Kaas.

[5] Model Estimation of Claim Risk and Premium for Motor Vehicle Insurance by Using Bayesian Method. To cite this article: Sukono et al., 2018, IOP Conf. Ser.: Mater. Sci. Eng. 300, 012027.

[6] Non-Life Insurance Mathematics: An Introduction with Stochastic Processes by Thomas Mikosch .

[7] the probability of ruin in risk theory by wen XU , B.S .

[8] Bayesian Analysis for the Social Sciences by Simon Jackman.

[9] Introduction aux Statistiques Baytesiennes Yann Traonmilin - Adrien Richou

[10] Introduction to Bayesian Statistics Brendon J. Brewer

[11] "Bayesian Data analysis" by Andrew Gelman, John B.

[12] Mémoire de Diplôme de Magister en Mathématiques intitulé "Probabilité de Ruine et Value-at-Risk en Finance et Assurance" by Hamidane Besma.

[13] Introduction to Bayesian Statistics

[14] Selection of Claim Size Distribution in General Insurance Simen Holter Masters Thesis, Spring 2017

[15] Bayesian claims Reserving Methods in Non-life Insurance with Stan.

[16] Statistqun Bayésiens et Jirome dupuis algorithmes MCMC

[17] Model estimation of claim risk and premium for motor vehicle insurance by using Bayesian method , To cite this article: Sukono et al 2018 IOP Conf. Ser.: Mater. Sci. Eng. 300 012027

[18] random variable , definition , types , how it used and examples by will kenton.

[19] A First Course in Bayesian Statistical Methods by Peter D. Hoff.

[20] introduction to bayesian statistics by Brendon . Brewer.

[21] A Bayesian Approach to Weibull Distribution with Application to Insurance Claims Data by Nuraddeen Yusuf Adamu .

APPENDICES

Appendix 1: Bayesian Aggregate Claims Simulation

```
# Load necessary library
if(!require("MASS")) install.packages("MASS", dependencies=TRUE)
library(MASS)

# Set prior parameters for Bayesian approach
prior_lambda_shape <- 2
prior_lambda_rate <- 1
prior_exp_shape <- 2
prior_exp_rate <- 1

# Number of replications
num_replications <- 1000

# Function to simulate aggregate claims with given parameters
simulate_aggregate_claims <- function(lambda, rate) {
  # Simulate number of claims from Poisson distribution
  num_claims <- rpois(200, lambda)

  # Simulate the amount of each claim from Exponential distribution
  claim_amounts <- rexp(sum(num_claims), rate = rate)

  # Calculate aggregate claims
  aggregate_claims <- sum(claim_amounts)

  return(aggregate_claims)
}

# Initialize the vector to store aggregate claims samples
aggregate_claims_samples <- numeric(num_replications)

# Simulate 1000 replications of aggregate claims
for (i in 1:num_replications) {
  # Update Poisson parameter (lambda)
  lambda_posterior <- rgamma(1, shape = prior_lambda_shape, rate =
     prior_lambda_rate)

```

```
  # Update Exponential parameter (rate)
  rate_posterior <- rgamma(1, shape = prior_exp_shape, rate = prior_
     exp_rate)

  # Simulate aggregate claims
  aggregate_claims_samples[i] <- simulate_aggregate_claims(lambda_
     posterior, rate_posterior)
}

# Calculate the mean and standard deviation of the aggregate claims
mean_aggregate_claims <- mean(aggregate_claims_samples)
sd_aggregate_claims <- sd(aggregate_claims_samples) / sqrt(num_
   replications)

# Output the results
print(paste("Mean of the 1000 aggregate claims (Bayesian approach):",
   mean_aggregate_claims))
print(paste("Standard deviation of the mean of the 1000 aggregate
   claims (Bayesian approach):", sd_aggregate_claims))
```

Appendix 2: Kolmogorov-Smirnov Test for Poisson Distribution

```
# Given number of claims
number_of_claims <- c(13, 15, 20, 37, 31, 29, 20, 44, 36, 36, 33, 25,
   25, 7)

# Calculate the mean of the Poisson distribution
lambda_est <- mean(number_of_claims)

# Perform the Kolmogorov-Smirnov (KS) test assuming a Poisson
   distribution with estimated lambda
ks_result_poisson <- ks.test(rpois(length(number_of_claims), lambda =
   lambda_est), "ppois", lambda = lambda_est)

# Print the test result
print(ks_result_poisson)

# Summary and Interpretation
cat("\nSummary of the Kolmogorov-Smirnov test with Poisson
   distribution:\n")
cat("Test statistic (D):", ks_result_poisson$statistic, "\n")
cat("P-value:", ks_result_poisson$p.value, "\n")
cat("Alternative hypothesis:", ks_result_poisson$alternative, "\n\n")

# Interpretation based on the p-value
if (ks_result_poisson$p.value < 0.05) {
  cat("The p-value is less than 0.05, suggesting significant evidence
     against the null hypothesis.\n")
```

```r
  cat("Therefore, the observed data does not follow a Poisson distribution with lambda =", lambda_est, "\n")
} else {
  cat("The p-value is greater than or equal to 0.05, suggesting no significant evidence against the null hypothesis.\n")
  cat("Therefore, the observed data could be considered to follow a Poisson distribution with lambda =", lambda_est, "\n")
}

# Plot histogram of number of claims
hist(number_of_claims, breaks = 10, col = "skyblue", xlab = "Number of Claims", main = "Histogram of Number of Claims")
```

Appendix 3: Exponential Distribution Sample Generation and KS Test

```r
# Required libraries
library(fitdistrplus)

# Given number of claims and minimum loss values
number_of_claims <- c(13, 15, 20, 37, 31, 29, 20, 44, 36, 36, 33, 25, 25, 7)
min_losses <- c(6.925, 3.586, 7.899, 7.487, 4.179, 6.685, 5.343, 2.988, 5.093, 4.964, 3.435, 4.051, 4.147, 2.956)

# Function to generate sample with numbers of claims and sum constraint from Exponential distribution
generate_sample_exponential <- function(number_of_claims, rates, target_sum) {
  sample <- numeric(target_sum)
  index <- 1

  for (i in seq_along(number_of_claims)) {
    sample[index:(index + number_of_claims[i] - 1)] <- rexp(number_of_claims[i], rate = rates[i])
    index <- index + number_of_claims[i]
  }

  return(sample)
}

# Set the target sum
target_sum <- sum(number_of_claims)

# Estimate rates for the Exponential distribution
rates_est <- 1 / min_losses

cat("Estimated rates for Exponential distribution:\n")
print(rates_est)

```

```
# Generate a sample from Exponential distribution
F_exponential <- generate_sample_exponential(number_of_claims, rates_est, target_sum)

# Estimate parameters for the Exponential distribution
fit_exponential <- fitdist(F_exponential, "exp")

# Extract estimated parameter
rate_est <- fit_exponential$estimate["rate"]

# Print estimated parameter
cat("Estimated rate parameter (Exponential distribution):", rate_est, "\n")

# Perform the Kolmogorov-Smirnov (KS) test assuming an Exponential distribution with estimated parameter
ks_result_exponential <- ks.test(F_exponential, "pexp", rate = rate_est)

# Print the test result
print(ks_result_exponential)

# Summary and Interpretation
cat("\nSummary of the Kolmogorov-Smirnov test with Exponential distribution:\n")
cat("Test statistic (D):", ks_result_exponential$statistic, "\n")
cat("P-value:", ks_result_exponential$p.value, "\n")
cat("Alternative hypothesis:", ks_result_exponential$alternative, "\n\n")

# Interpretation based on the p-value
if (ks_result_exponential$p.value < 0.05) {
  cat("The p-value is less than 0.05, suggesting significant evidence against the null hypothesis.\n")
  cat("Therefore, vector F_exponential does not follow an Exponential distribution.\n")
} else {
  cat("The p-value is greater than or equal to 0.05, suggesting no significant evidence against the null hypothesis.\n")
  cat("Therefore, vector F_exponential could be considered to follow an Exponential distribution.\n")
}

# Plot histogram of F_exponential
hist(F_exponential, main = "Histogram of F_exponential (Exponential Distribution)", xlab = "Values", ylab = "Frequency")
```

Appendix 4: Kolmogorov-Smirnov Test for Generated Data

```r
# Required libraries
# Function to generate sample with specified numbers and apply
  Kolmogorov-Smirnov (KS) test with exponential distribution
generate_sample_and_ks_test <- function(numbers, target_sum) {
  sample <- numeric(target_sum)
  index <- 1

  for (i in seq_along(numbers)) {
    sample[index:(index + numbers[i] - 1)] <- rexp(numbers[i], rate =
      1)
    index <- index + numbers[i]
  }

  # Perform the Kolmogorov-Smirnov (KS) test assuming an Exponential
    distribution with rate = 1
  ks_result <- ks.test(sample, "pexp", rate = 1)

  return(list(sample = sample, ks_result = ks_result))
}

# Given numbers
numbers <- c(10, 15, 20, 25, 30)

# Set the target sum
target_sum <- sum(numbers)

# Generate sample and perform KS test
result <- generate_sample_and_ks_test(numbers, target_sum)

# Print the test result
print(result$ks_result)

# Summary and Interpretation
cat("\nSummary of the Kolmogorov-Smirnov test with Exponential
  distribution:\n")
cat("Test statistic (D):", result$ks_result$statistic, "\n")
cat("P-value:", result$ks_result$p.value, "\n")
cat("Alternative hypothesis:", result$ks_result$alternative, "\n\n")

# Interpretation based on the p-value
if (result$ks_result$p.value < 0.05) {
  cat("The p-value is less than 0.05, suggesting significant evidence
    against the null hypothesis.\n")
  cat("Therefore, the generated data does not follow an Exponential
    distribution.\n")
} else {
  cat("The p-value is greater than or equal to 0.05, suggesting no
    significant evidence against the null hypothesis.\n")
  cat("Therefore, the generated data could be considered to follow an
    Exponential distribution.\n")
}

# Plot histogram of generated sample
```

```
hist(result$sample, main = "Histogram of Generated Sample (Exponential
    Distribution)", xlab = "Values", ylab = "Frequency")
```

In this continuation, we complete the setup for generating a sample with specified numbers and applying the Kolmogorov-Smirnov (KS) test with an exponential distribution.

Appendix 5: R Code for Bayesian Inference with Poisson Distribution

The following R code implements Bayesian inference for the Poisson distribution:

```
# Bayesian Inference for Poisson Distribution
# Assuming a Poisson distribution with parameter lambda

# Given data
data <- c(5, 3, 7, 2, 4)

# Bayesian analysis using conjugate prior (Gamma prior)
# Prior parameters
alpha <- 1    # Shape parameter
beta <- 1     # Rate parameter

# Posterior parameters
alpha_post <- alpha + sum(data)
beta_post <- beta + length(data)

# Posterior mean and variance
posterior_mean <- alpha_post / beta_post
posterior_var <- alpha_post / (beta_post^2)

# Posterior distribution (Gamma)
posterior_distribution <- rgamma(10000, shape = alpha_post, rate =
    beta_post)

# Summary
cat("Posterior Mean:", posterior_mean, "\n")
cat("Posterior Variance:", posterior_var, "\n\n")

# Plot posterior distribution
hist(posterior_distribution, main = "Posterior Distribution (Gamma)
    for Poisson Parameter", xlab = "Parameter Value", ylab = "Density")
```

This code conducts Bayesian inference for the Poisson distribution using a conjugate Gamma prior.

Appendix 6: R Code for Monte Carlo Simulation with Bayesian Methods

The following R code performs a Monte Carlo simulation using Bayesian methods:

```r
# Monte Carlo Simulation with Bayesian Methods
# Simulating data from a hypothetical Bayesian model

# Parameters
N <- 1000   # Number of simulations
alpha <- 2  # Prior parameter
beta <- 5   # Prior parameter

# Simulate data from a model (e.g., Normal distribution)
# Here we generate data assuming a normal distribution
# with mean alpha and standard deviation beta
simulated_data <- rnorm(N, mean = alpha, sd = beta)

# Bayesian analysis (e.g., updating priors with simulated data)
# Posterior calculations would follow Bayesian updating rules
# For example, updating parameters based on simulated data

# Summary of simulated data
cat("Summary of Simulated Data:\n")
cat("Mean:", mean(simulated_data), "\n")
cat("Standard Deviation:", sd(simulated_data), "\n\n")

# Plot of simulated data
hist(simulated_data, main = "Histogram of Simulated Data", xlab = "
   Values", ylab = "Frequency")
```

This code conducts a Monte Carlo simulation using Bayesian methods, simulating data from a hypothetical Bayesian model and analyzing the results.

These appendices provide detailed R code examples for various statistical simulations and Bayesian analyses.

Appendix 7: R Code for Kolmogorov-Smirnov Test with Exponential Distribution

The following R code performs the Kolmogorov-Smirnov (KS) test with an exponential distribution:

```r
# Kolmogorov-Smirnov Test with Exponential Distribution
# Testing goodness of fit with KS test

# Sample data
data <- c(4.0780, 2.4030, 4.5555, 4.3475, 2.7000, 3.9625, 3.3305,
   2.1320, 3.2260, 3.2135, 2.4865, 2.8425, 2.8895, 2.3085)

# Fit exponential distribution
lambda <- 1 / mean(data)
exp_distribution <- rexp(length(data), rate = lambda)

# Perform KS test
ks_test <- ks.test(data, "pexp", lambda = lambda)
```

```
# Summary
cat("KS Test Summary:\n")
print(ks_test)

# Plot of data and fitted distribution
hist(data, freq = FALSE, main = "Histogram of Data vs. Fitted Exponential Distribution", xlab = "Values", ylab = "Density")
curve(dexp(x, rate = lambda), add = TRUE, col = "blue")
legend("topright", legend = c("Data", "Exponential Fit"), col = c("black", "blue"), lty = 1)
```

This code calculates the Kolmogorov-Smirnov (KS) test statistic to assess the goodness of fit of an exponential distribution to the given data.

Appendix 8: R Code for Markov Chain Monte Carlo (MCMC) Simulation

The following R code implements a Markov Chain Monte Carlo (MCMC) simulation:

```
# Markov Chain Monte Carlo (MCMC) Simulation
# Simulating data using MCMC methods

# Parameters
N <- 1000   # Number of iterations
burn_in <- 100   # Burn-in period

# Initialize
theta <- numeric(N)
theta[1] <- 1   # Initial value

# MCMC simulation (e.g., Gibbs sampling)
for (i in 2:N) {
  # Update theta using a Metropolis-Hastings step (example)
  proposal <- rnorm(1, mean = theta[i - 1], sd = 0.5)

  # Calculate acceptance ratio (in a simple case)
  acceptance <- min(1, dnorm(proposal, mean = 0, sd = 1) / dnorm(theta[i - 1], mean = 0, sd = 1))

  # Accept or reject proposal
  if (runif(1) < acceptance) {
    theta[i] <- proposal
  } else {
    theta[i] <- theta[i - 1]
  }
}

# Remove burn-in period
theta <- theta[(burn_in + 1):N]

```

```
# Summary
cat("MCMC Simulation Summary:\n")
cat("Mean:", mean(theta), "\n")
cat("Standard Deviation:", sd(theta), "\n\n")

# Plot of MCMC samples
hist(theta, main = "Histogram of MCMC Samples", xlab = "Values", ylab
   = "Frequency")
```

This code demonstrates a basic Markov Chain Monte Carlo (MCMC) simulation using Gibbs sampling with a simple Metropolis-Hastings step.

These appendices provide additional R code examples, including the KS test with exponential distribution and a Markov Chain Monte Carlo (MCMC) simulation .

Appendix 9: R Code for Bayesian Inference with Poisson Distribution

The following R code demonstrates Bayesian inference with a Poisson distribution:

```
# Bayesian Inference with Poisson Distribution
# Example of Bayesian updating

# Data (number of claims)
claims <- c(10, 12, 8, 15, 11)

# Prior parameters
alpha <- 1
beta <- 1

# Posterior parameters
alpha_post <- alpha + sum(claims)
beta_post <- beta + length(claims)

# Posterior distribution
posterior <- rgamma(10000, shape = alpha_post, rate = beta_post)

# Summary statistics
cat("Posterior Mean:", mean(posterior), "\n")
cat("Posterior Standard Deviation:", sd(posterior), "\n\n")

# Plot of posterior distribution
hist(posterior, breaks = 30, col = "lightblue", main = "Posterior
   Distribution", xlab = "Values", ylab = "Density")
```

This code illustrates Bayesian updating with a Poisson distribution using conjugate priors and simulates the posterior distribution using a Gamma distribution.

Appendix 10: R Code for Bayesian Linear Regression

The following R code performs Bayesian linear regression:

```
# Bayesian Linear Regression
# Example using Gibbs sampling

# Simulated data
set.seed(123)
n <- 100
x <- seq(0, 10, length.out = n)
true_slope <- 0.5
true_intercept <- 2
epsilon <- rnorm(n, mean = 0, sd = 1)
y <- true_intercept + true_slope * x + epsilon

# Bayesian linear regression model
library(MCMCpack)
model <- MCMCregress(y ~ x)

# Summary of the regression coefficients
summary(model)

# Plot of data and regression line
plot(x, y, main = "Bayesian Linear Regression", xlab = "x", ylab = "y")
abline(coef(model)["(Intercept)"], coef(model)["x"], col = "blue", lwd = 2)
```

This code performs Bayesian linear regression using Gibbs sampling and visualizes the regression line fit to simulated data. These appendices provide additional R code examples for Bayesian inference with a Poisson distribution and Bayesian linear regression .

Appendix 11: R Code for Kolmogorov-Smirnov Test with Exponential Distribution

The following R code demonstrates the Kolmogorov-Smirnov test with an exponential distribution:

```
# Kolmogorov-Smirnov Test with Exponential Distribution
# Example of goodness-of-fit test

# Data (example data, replace with your data)
data <- c(4.0780, 2.4030, 4.5555, 4.3475, 2.7000, 3.9625, 3.3305, 2.1320, 3.2260, 3.2135, 2.4865, 2.8425, 2.8895, 2.3085)

# Fit exponential distribution
lambda <- 1/mean(data)
exp_distribution <- rexp(length(data), rate = lambda)

# Kolmogorov-Smirnov test
```

```r
ks_test <- ks.test(data, "pexp", lambda = lambda)

# Summary of KS test
cat("Kolmogorov-Smirnov Test:\n")
print(ks_test)

# Plot of empirical and fitted distributions
hist(data, freq = FALSE, col = "lightblue", main = "Empirical vs
   Fitted Exponential Distribution", xlab = "Values", ylab = "Density"
   )
curve(dexp(x, rate = lambda), col = "blue", lwd = 2, add = TRUE,
   legend = "Exponential Fit")
legend("topright", legend = c("Empirical", "Exponential Fit"), fill =
   c("lightblue", "blue"))
```

This code fits an exponential distribution to data and performs a Kolmogorov-Smirnov test to assess the goodness of fit.

Appendix 12: R Code for Bayesian Inference with Beta-Binomial Model

The following R code demonstrates Bayesian inference with a beta-binomial model:

```r
# Bayesian Inference with Beta-Binomial Model
# Example using conjugate priors

# Data (number of trials and successes)
trials <- 100
successes <- 60

# Prior parameters
alpha <- 1
beta <- 1

# Posterior parameters
alpha_post <- alpha + successes
beta_post <- beta + trials - successes

# Posterior distribution
posterior <- rbeta(10000, shape1 = alpha_post, shape2 = beta_post)

# Summary statistics
cat("Posterior Mean:", mean(posterior), "\n")
cat("Posterior Standard Deviation:", sd(posterior), "\n\n")

# Plot of posterior distribution
hist(posterior, breaks = 30, col = "lightgreen", main = "Posterior
   Distribution", xlab = "Values", ylab = "Density")
```

This code illustrates Bayesian inference with a beta-binomial model using conjugate priors and simulates the posterior distribution.

These appendices cover additional R code examples for Bayesian inference with a beta-binomial model and Kolmogorov-Smirnov test with an exponential distribution.

Appendix 13: R Code for Bayesian Model of Risk and Ruin Probability

The following R code demonstrates a Bayesian model for risk and ruin probability:

```
# Bayesian Model of Risk and Ruin Probability
# Example using Bayesian methods

# Data (adjust with your specific data)
claims <- c(1000, 1500, 2000, 2500, 3000)  # Example claim amounts
premium <- 1200  # Example premium amount
loss <- claims - premium  # Loss amounts

# Prior distribution (Gamma for claim amounts and Normal for premium)
alpha_prior <- 2
beta_prior <- 1000
mu_prior <- 1200
sigma_prior <- 100

# Bayesian update
alpha_post <- alpha_prior + sum(loss)
beta_post <- beta_prior + length(loss)
mu_post <- (alpha_prior * mu_prior + sum(loss)) / (alpha_prior +
    length(loss))
sigma_post <- sqrt((beta_prior * sigma_prior^2 + sum((loss - mu_prior)
    ^2) + (alpha_prior * length(loss) * (mu_prior - mu_post)^2)) / (
    beta_prior + length(loss)))

# Posterior distributions
posterior_alpha <- rgamma(10000, shape = alpha_post, rate = beta_post)
posterior_mu <- rnorm(10000, mean = mu_post, sd = sigma_post)

# Summary statistics
cat("Posterior Mean of Alpha:", mean(posterior_alpha), "\n")
cat("Posterior Mean of Mu:", mean(posterior_mu), "\n\n")

# Plot of posterior distributions
par(mfrow = c(2, 1))
hist(posterior_alpha, breaks = 30, col = "lightblue", main = "
    Posterior Distribution of Alpha", xlab = "Values", ylab = "Density"
    )
hist(posterior_mu, breaks = 30, col = "lightgreen", main = "Posterior
    Distribution of Mu", xlab = "Values", ylab = "Density")
```

This code exemplifies a Bayesian approach to modeling risk and ruin probability using Gamma and Normal distributions for prior and posterior updates.

Appendix 14: R Code for Simulation with Bayesian Methods

The following R code demonstrates simulation using Bayesian methods:

```
# Simulation with Bayesian Methods
# Example of MCMC simulation

# Data (adjust with your specific data)
n <- 100  # Number of observations
x <- rpois(n, lambda = 3)  # Simulated Poisson data

# Bayesian model (Poisson likelihood, Gamma prior)
alpha <- 2
beta <- 2

# Initial values for MCMC
lambda <- 2
lambda_samples <- numeric(10000)

# MCMC simulation
for (i in 1:10000) {
  lambda_prop <- rnorm(1, mean = lambda, sd = 0.5)
  if (lambda_prop > 0) {
    acceptance_prob <- min(1, exp(sum(dpois(x, lambda_prop, log = TRUE
       )) - sum(dpois(x, lambda, log = TRUE))))
    if (runif(1) < acceptance_prob) {
      lambda <- lambda_prop
    }
  }
  lambda_samples[i] <- lambda
}

# Summary statistics
cat("Posterior Mean of Lambda:", mean(lambda_samples), "\n")
cat("Posterior Standard Deviation of Lambda:", sd(lambda_samples), "\n
   \n")

# Plot of posterior distribution
hist(lambda_samples, breaks = 30, col = "lightblue", main = "Posterior
    Distribution of Lambda", xlab = "Values", ylab = "Density")
```

This code demonstrates a simulation using Bayesian methods, specifically using Markov Chain Monte Carlo (MCMC) to estimate parameters.

These appendices cover additional R code examples for Bayesian modeling of risk and ruin probability, as well as simulation with Bayesian methods.