精算建模：风险理论参考答案及批改评述（Chap 3）

庄源

日期：2023 年 11 月 21 日

目录

1 Question 3: Properties of compound binomial distribution

1.1 原题

Assume total annual claims S are modeled by a compound binomial distribution where $N \sim B(200, 0.001)$ and all claims are a constant 500. Determine the mean, variance and skewness of S, and find the probability that S exceeds 600 exactly.

1.2 参考答案

可以直接代入教材 81 面式 3.1-3.4，便得到结果。下面我们直接使用一般原理推导这些结论。

$$E(S) = E(N)E(X) = 200 \times 0.001 \times 500 = 100$$

因为 X 恒为常数 500，所以 $Var(X) = 0$。

$$\begin{aligned} Var(S) &= E(N)Var(X) + Var(N)E^2(X) \\ &= 200 \times 0.001 \times 0 + 200 \times 0.001 \times (1-0.001) \times 500^2 \\ &= 49950 \end{aligned}$$

下面求偏度，已知 $S = 500X$，则：

$$\begin{aligned} skew(S) &= \frac{E\left[S - E(S)\right]^3}{Var(S)^{\frac{3}{2}}} = \frac{500^3 \cdot E\left[N - E(N)\right]^3}{500^3 \cdot Var(N)^{\frac{3}{2}}} \\ &= \frac{E\left[N - E(N)\right]^3}{Var(N)^{\frac{3}{2}}} = \frac{npq(1-2q)}{(npq)^{\frac{3}{2}}} \\ &= \frac{0.1994004}{0.0893} = 2.2327 \end{aligned}$$

上述推导中，二项分布的三阶中心矩可通过 cumulant moment generating function 求得。同时，我们也可以看出，$500X$ 的偏度和 X 的偏度是相同的，因为偏度是一个标准化后得到的指标。

最后是 $\Pr(S > 600)$。想要总索赔额大于 600，则需要有两个或两个以上的赔案：

$$\begin{aligned} \Pr(S > 600) &= \Pr(N \geqslant 2) \\ &= 1 - \Pr(N = 0) - \Pr(N = 1) \\ &= 1 - (1-0.001)^{200} - \binom{200}{1}(1-0.001)^{199} \cdot 0.001 \quad = 0.01746 \end{aligned}$$

1.3 给分标准与批改评价

表 1: Question 3 给分标准（共 15 分）

采分点	分值
计算均值	3
计算方差	3
计算偏度	6
计算总索赔额大于 600 的概率	3

本题完成情况很好，部分同学漏掉了 $E(S)$ 和 $Var(S)$。小部分同学使用中心极限定理计算 $\Pr > 600$，如果计算近似正确也给分，但这当中有些同学用错了，还请再查阅数理统计书籍。

2 Question 5：Panjer's recursive formula

注. 本题制作了 EXCEL 解答，可通过查看表格中的公式了解详细步骤。[下载]

本题制作了 R 语言代码，可通过查看代码学会使用 `actuar` 构建复合分布。[下载]

2.1 原题

Total aggregate claims $S = X_1 + \cdots + X_N$ are modeled by a compound binomial distribution where $N \sim B(4, q = 1/2)$ and $P(X = j) \propto j$ for $j = 1,\ 2,\ 3,\ 4,\ 5$. Determine the mean, variance, skewness S and find the exact probability distribution for S using Panjer's recursive formula.

2.2 参考答案

这一题中，均值、方差和偏度可以通过直接代入教材 81 面式 3.1-3.4 解出。首先，需要得出 X 的一阶、二阶和三阶矩 m_1, m_2 和 m_3。因为 $\Pr(X = j) \propto j$[1]，所以：

$$\Pr(X = j) = \frac{j}{\sum_{j=1}^{5} j} = \frac{j}{15} \quad (j = 1,\ 2,\ 3,\ 4,\ 5)$$

所以各阶矩为：

$$m_1 = E(X) = \sum_{j=1}^{5} jP(X = j) = \frac{11}{3}$$

$$m_2 = E\left(X^2\right) = \sum_{j=1}^{5} j^2 P(X = j) = 15$$

$$m_3 = E\left(X^3\right) = \sum_{j=1}^{5} j^3 P(X = j) = \frac{979}{15}$$

因此，

$$E(S) = E(N)E(X) = nqm_1 = 7.3333$$

$$\begin{aligned} Var(S) &= E(N)Var(X) + Var(N)E^2(X) \\ &= nq\left(m_2 - m_1^2\right) + nqpm_1^2 = nq\left(m_2 - qm_1^2\right) = 16.5556 \end{aligned}$$

$$\begin{aligned} skew(S) &= \frac{nqm_3 - 3nq^2 m_2 m_1 + 2nq^3 m_1^3}{\left(nqm_2 - nq^2 m_1^2\right)^{3/2}} \\ &= \frac{1}{\sqrt{n}} \frac{qm_3 - 3q^2 m_2 m_1 + 2q^3 m_1^3}{\left(qm_2 - q^2 m_1^2\right)^{3/2}} = 0.2201 \end{aligned}$$

正如我在上课时跟大家讲过的，复合分布概率的计算很困难，得要使用 Panjer 递推公式。所谓“递推”，就是从 S 为某个数的概率开始，依次向后推出 S 等于其它数的概率。因此，如果要使用 Panjer 递推，就必须要有一个递推的起点。这个递推的起点非常好找：$\Pr(S = 0)$。既然 $\Pr(X = 0) = 0$，这意味着，只要发生事故，就一定会有损失。所以，总索赔 $S = 0$ 的概率其实就是不发生任何事故的概率：

$$\Pr(S = 0) = \Pr(N = 0) = \left(\frac{1}{2}\right)^5 = 6.25\%$$

[1] $\propto$ 是“正比于”的意思。

下面观察 Panjer 递推公式，我们要用这个公式来计算剩下的概率：

$$f_S(r) = \sum_{j=1}^{r} \left(\alpha + \frac{\beta j}{r}\right) f_X(j) f_S(r-j) \quad \text{for} \quad r = 1, 2, \ldots \tag{1}$$

因为 N 的分布为二项分布，由教材 87 面，$\alpha = -q/(1-q) = -1$，$\beta = (n+1)q/(1-q) = 5$。观察求和号，j 要从 1 一直加到 r。但由题意，实际上 j 只能取 $1 \sim 5$，所以对于 $j > 5$，$f_X(j) = 0$。此外，S 最高为 $4 \times 5 = 20$，所以 $r = 20$。因此，原 Panjer 递推公式可以简化为：

$$f_S(r) = \sum_{j=1}^{5} \left(-1 + \frac{5j}{r}\right) f_X(j) f_S(r-j) \quad \text{for} \quad r = 1,\ 2,\ \cdots,\ 20$$

因此，对于 S，有：

$$\begin{aligned}
\Pr(S=1) &= \left(\frac{5\times 1}{1} - 1\right) \cdot \Pr(X=1) \cdot \Pr(S=0) = \frac{1}{60} \approx 1.67\% \\
\Pr(S=2) &= \left(\frac{5\times 1}{2} - 1\right) \cdot \Pr(X=1) \cdot \Pr(S=1) + \left(\frac{5\times 2}{2} - 1\right) \cdot \Pr(X=2) \cdot \Pr(S=0) \\
&= 3.5\% \\
\Pr(S=3) &= \left(\frac{5\times 1}{3} - 1\right) \cdot \Pr(X=1) \cdot \Pr(S=2) + \left(\frac{5\times 2}{3} - 1\right) \cdot \Pr(X=2) \cdot \Pr(S=1) \\
&\quad + \left(\frac{5\times 3}{3} - 1\right) \cdot \Pr(X=3) \cdot \Pr(S=0) \\
&\approx 5.67\% \\
\Pr(S=4) &= \left(\frac{5\times 1}{4} - 1\right) \cdot \Pr(X=1) \cdot \Pr(S=3) + \left(\frac{5\times 2}{4} - 1\right) \cdot \Pr(X=2) \cdot \Pr(S=2) \\
&\quad + \left(\frac{5\times 3}{4} - 1\right) \cdot \Pr(X=3) \cdot \Pr(S=1) + \left(\frac{5\times 4}{4} - 1\right) \cdot \Pr(X=4) \cdot \Pr(S=0) \\
&\approx 8.38\% \\
\Pr(S=5) &= \left(\frac{5\times 1}{5} - 1\right) \cdot \Pr(X=1) \cdot \Pr(S=4) + \left(\frac{5\times 2}{5} - 1\right) \cdot \Pr(X=2) \cdot \Pr(S=3) \\
&\quad + \left(\frac{5\times 3}{5} - 1\right) \cdot \Pr(X=3) \cdot \Pr(S=2) + \left(\frac{5\times 4}{5} - 1\right) \cdot \Pr(X=4) \cdot \Pr(S=1) \\
&\quad + \left(\frac{5\times 5}{5} - 1\right) \cdot \Pr(X=5) \cdot \Pr(S=0) \\
&\approx 11.82\%
\end{aligned}$$

其余的概率计算和上面的式子相似。这样，我们便能得出 S 的完整分布列，如下表所示：

表 2: S 的具体分布列（使用 EXCEL 或 R 得出）

i	$\Pr(S=i)$	i	$\Pr(S=i)$	i	$\Pr(S=i)$
0	6.25%	7	8.28%	14	3.12%
1	1.67%	8	9.44%	15	1.89%
2	3.50%	9	9.49%	16	0.91%
3	5.67%	10	8.09%	17	0.73%
4	8.38%	11	4.74%	18	0.48%
5	11.82%	12	4.78%	19	0.25%
6	6.25%	13	4.18%	20	0.08%

需要注意的是，我不是通过手算得出上面分布列的，而是利用计算机辅助计算。感兴趣的同学可以翻到这一节的开头，有 EXCEL 和 R 代码的下载链接。这里，再花一点点篇幅介绍 R 语言对于聚合风险模型的实现，这些实现都离不开 `actuar` 这个有名的 R 包。

`actuar` 中，`aggregateDist()` 函数可以用来指定相应的聚合风险模型，当方法选为 `recursive` 时，即为使用 Panjer 递推[2]。

```r
# 载入 actuar 包，一个有名的精算包
library(actuar)
# 给出 X（损失强度）的分布
# 一定要有 X 为 0 时的概率（包的规定）
fx <- 0:5 / 15
# 主要函数
# recursive，代表使用 Panjer 递推
# 损失频率（N）分布使用二项分布（由题意可得）
# 还可以使用各种零膨胀分布
# 损失强度（X）分布使用刚才定义的 X
# size 和 prob 是 pbinom 函数中的参数，分别代表 n 和 p
Fs <- aggregateDist("recursive", model.freq = "binomial",
model.sev = fx, size = 4, prob = 0.5)
# 绘制 CDF 的图
plot(Fs)
```

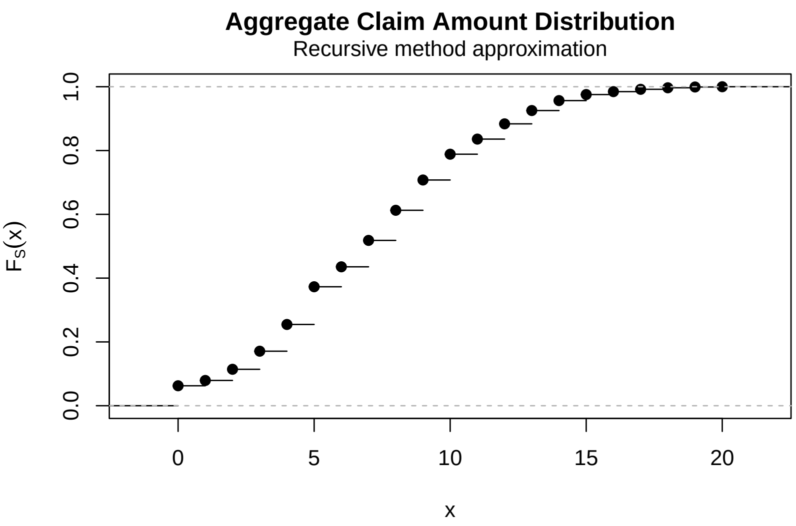


图 1: 总索赔的累积分布函数

[2] `actuar` 中求聚合风险模型确切分布的方法还有卷积法、正态近似法和正态幂近似法，这些不是大纲内容，也就不再涉及。

```r
# scales 包能将概率显示为百分数
library(scales)
S <- 0:20
# 将 cdf 做一阶差分得到分布列
Prob <- diff(Fs)
# 组合成为 dataframe，并更改列名
prob_mass_function <- data.frame(S,
percent(Prob,
accuracy = 0.01))
colnames(prob_mass_function) <- c("S","Probability")
# 查看分布列
prob_mass_function
```

```
##     S Probability
## 1   0       6.25%
## 2   1       1.67%
## 3   2       3.50%
## 4   3       5.67%
## 5   4       8.38%
## 6   5      11.82%
## 7   6       6.25%
## 8   7       8.28%
## 9   8       9.44%
## 10  9       9.49%
## 11 10       8.09%
## 12 11       4.74%
## 13 12       4.78%
## 14 13       4.18%
## 15 14       3.12%
## 16 15       1.89%
## 17 16       0.91%
## 18 17       0.73%
## 19 18       0.48%
## 20 19       0.25%
## 21 20       0.08%
```

除了计算 S 的分布列以外，还可以非常简单地编程得出 S 的均值、方差和偏度：

```r
# S 的均值
mean_S <- sum(S*Prob)
# S 的方差
var_S <- sum(S^2*Prob)-mean_S^2
# S 的偏度
kurtosis_S <- sum((S-mean_S)^3*Prob) / var_S^1.5
mean_S; var_S; kurtosis_S
```

```
## [1] 7.333333
```

```
## [1] 16.55556
```

```
## [1] 0.220148
```

2.3 给分标准与批改评价

表 3: Question 5 给分标准（共 25 分）

采分点	分值
计算均值	3
计算方差	3
计算偏度	6
计算 α 和 β	4
正确写出本题的 Panjer 递推公式	5
得出 S 的完整分布列（即 $\Pr(S=0) \sim \Pr(S=20)$）	4

这个题非常重要，换一个形式出现在期末考试中也不是不可能[3]。批改情况很差，部分同学直接跳过 Panjer 递推没做。错误大多集中在以下几个方面：

- 部分同学把本题中的复合二项分布当成了二项分布。当 N 的分布满足二项分布时，N 才有课本第 87 面那样描述的递推关系。但是这题要求去求 S 的分布，这是一个复合分布，应该使用课本第 88 面 THEOREM 3.2 的方法。
- 本题要求给出 *exact distribution*，但很多同学只给了少数几个点上的概率，这里稍稍扣了分。
- 还有小部分同学把这题当作复合 Poisson，算错了 S 的二阶矩和偏度。

[3]可能不会让大家计算偏度，因为公式有点难记。作为考试的话，也不会让大家计算这么多概率，计算 $\Pr(S=3)$ 或 $\Pr(S=4)$ 就已经到头了。要是再出简单点儿，可能就叫大家说说 Panjer 递推是什么。

3 Question 15: Premium, reserve and Poisson approximation

3.1 原题

Total claims S made in respect of a portfolio of fire insurance policies can be modeled by a compound Poisson distribution where the Poisson parameter is λ and the typical claim is X. Let us assume that the claim random variable X is a $40/60$ mixture of claims of type I and II, respectively. Claims of type I are Pareto $(3, 600)$, while those of type II are Pareto $(4, 900)$. Calculate $P(X > 400)$, $E(X)$ and $\mathrm{Var}(X)$. If the security loading of $\theta = 0.15$ is used for determining premiums and $\lambda = 500$, what reserves are necessary in order to be 99.9% sure of meeting claims? What would be the effect of doubling the security loading?

Let Y be a Pareto random variable with the same mean and variance as X. What is $P(Y > 400)$? What would be the reserves necessary to be 99.9% sure all claims will be met (from premium income plus reserves) if we had used Y instead of X in our model?

3.2 参考答案

首先计算 $\Pr(X > 400)$[4]：

$$\begin{aligned}\Pr(X > 400) &= \Pr(X > 400 \mid Type\ I) \times \Pr(Type\ I) + \Pr(X > 400 \mid Type\ II) \times \Pr(Type\ II) \\ &= \left(\frac{600}{600+400}\right)^3 \times 0.4 + \left(\frac{900}{900+400}\right)^4 \times 0.6 = 0.2242\end{aligned}$$

接着是 $E(X)$：

$$\begin{aligned}E(X) &= E(X \mid Type\ I) \times \Pr(Type\ I) + E(X \mid Type\ II) \times \Pr(Type\ II) \\ &= \frac{600}{3-1} \times 0.4 + \frac{900}{4-1} \times 0.6 = 300\end{aligned}$$

然后是 $Var(X)$，它的计算稍微有一点点复杂：

$$Var(X) = E(Var(X \mid Type)) + Var(E(X \mid Type))$$

其中，

$$E(Var(X \mid Type)) = 0.4 \times \frac{3 \times 600^2}{(3-1)^2 \times (3-2)} + 0.6 \times \frac{4 \times 900^2}{(4-1)^2 \times (4-2)} = 216000$$

下面的公式就是我们上课时讲过的 Bernoulli Shortcut。

$$Var(E(X \mid Type)) = \left(\frac{600}{3-1} - \frac{900}{4-1}\right)^2 \times 0.4 \times 0.6 = 0$$

所以 $Var(X) = 216000$。

要算出适当的准备金，就要去计算：什么数额的准备金和保费加在一起，有 99.9% 的概率覆盖未来的损失？令 U 为准备金，所收的保费为安全附加（Security Load）后的总索赔期望[5]。由

[4]对于 $X \sim Pareto(\alpha,\ \lambda)$，有：$\bar{F}_X(x) = \left(\frac{\lambda}{\lambda+x}\right)^\alpha$，$E(X) = \frac{\lambda}{\alpha-1}$，$Var(X) = \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)}$。上述结果见课本第 40-41 面。

[5]即 $(1+\theta)E(S)$。

课本 104 面式 3.15，可得：

$$
\begin{aligned}
99.9\% &= \Pr\left[S < U + (1+\theta)E(S)\right] \\
&= \Pr\left[\frac{S - E(S)}{\sqrt{Var(S)}} < \frac{U + \theta E(S)}{\sqrt{Var(S)}}\right] \\
&\doteq \Pr\left[N(0,1) < \frac{U + \theta E(S)}{\sqrt{Var(S)}}\right]
\end{aligned}
$$

因为总索赔 S 使用了复合 Poisson 模型，因此有

$$E(S) = \lambda E(X) = 150000$$

$$Var(S) = \lambda E(X^2) = 500 \times (216000 + 300^2) = 153000000$$

所以，$U = z_{99.9\%}\sqrt{Var(S)} - \theta E(S) = 15724.05883$。其中，$z_{99.9\%}$ 是标准正态分布的 99.9% 分位数。如果把安全附加翻倍，此时 $U' = z_{99.9\%}\sqrt{Var(S)} - \theta' E(S) = -6775.9412$。此时，不需要准备金，收上来的保费就已经能在给定概率下覆盖未来可能发生的索赔。

对于 Y 来说，其与 X 具有相同的均值和方差，首先估计分布的参数：

$$E(Y) = \frac{\lambda}{\alpha - 1} = 300 \qquad (*)$$

$$Var(Y) = \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)} = 216000 \qquad (**)$$

(**) / (*)^2，可解得 $\frac{\alpha}{\alpha-2} = 2.4$，解得 $\alpha = 3.42857$，进而解得 $\lambda = 728.5714$。所以：

$$\Pr(Y > 400) = \left(\frac{728.5714}{728.5714 + 400}\right)^{3.42857} = 0.2230$$

因此，使用 Y 来代替 X，概率是有变化的。

但是，如果使用 Y 来代替 X，算出的准备金不会有任何变化。观察刚才计算准备金的式子，里面用到了 X 的一、二阶矩，既然 Y 的一、二阶矩与 X 的相同，那么算出来的准备金也就不会有任何变化。

3.3 给分标准与批改评价

表 4: Question 15 给分标准（共 20 分）

采分点	分值
计算 X 大于 400 的概率	2
计算 X 的均值	2
计算 X 的方差	4
计算 $\theta = 0.15$ 时的准备金	4
计算 θ 翻倍时的准备金	2
估计 Y 的参数	2（一个 1 分）
计算 Y 大于 400 的概率	2
评论/计算使用 Y 作为损失强度随机变量时的准备金	2

这道题的完成情况不错。主要出现的错误是有些同学没有说明 Y 下的准备金，还有些同学把 X 的方差算错了。

4 Question 21: Excess of loss reinsurance

4.1 原题

Assume that aggregate claims are modeled by a compound Poisson process and that the excess of any claim over M is handled by a reinsurer who uses a security loading ξ (while the insurance company uses a loading of θ on policy holders). The typical claim X has a Pareto distribution with parameters (β, δ), that is

$$f_X(x) = \frac{\beta\delta^\beta}{(\delta + x)^{\beta+1}}.$$

Assume that the annual expected number of claims in this process is $\lambda = 300, \beta = 3, \delta = 1200, \theta = 0.2$ and $\xi = 0.3$. Determine the minimum excess level M^* which may be considered by the insurance company if it is desired that expected net profit is nonnegative, and complete the following table for a relationship between possible values of M and expected annual net profit.

Retention limit	Expected annual profit
300	$\star$
800	$\star$
$\star$	28406.25

4.2 参考答案

这道题和课本第 116 面例 3.17 非常相似，但例 3.17 的解答不够详细，因此这里给出更详细的解答。本题使用到的再保险为超额再保险（溢额再保险），保险人对一个赔案最多只需要支付 M。若单案索赔超出 M，则超额索赔由再保险人承担。令 X 为索赔金额，$Y = \min(X, M)$ 为索赔中保险人承担的责任，$Z = \max(X - M, 0)$ 为再保险人承担的责任，则 $X = Y + Z$。如课本第 108 面式 3.18 和式 3.19，保险人的收入来源于客户支付的保费，其为安全附加后的期望总索赔；保险人还需要为再保险支付再保险费，并为可能的保险事故提供赔偿，收入减支出即为净利润：

$$\begin{aligned}
Net\ Profit &= Premium - Reinsurance\ Premium - Claims \\
&= (1+\theta)E(X) - (1+\xi)E(Z) - E(Y) \\
&= \theta E(Y) - (\xi - \theta)E(Z)
\end{aligned}$$

若想使净利润非负，则

$$\frac{E(Y)}{E(Z)} \geqslant \frac{\xi - \theta}{\theta} = \frac{\xi}{\theta} - 1$$

本题的分布是 Pareto 分布，而 Pareto 分布又有一个非常好的性质（见课本第 39 面）：

$$P(X > M + x \mid X > M) = \left(\frac{\lambda}{\lambda + M + x}\right)^\alpha \Big/ \left(\frac{\lambda}{\lambda + M}\right)^\alpha = \left(\frac{\lambda + M}{\lambda + M + x}\right)^\alpha$$

也就是说，对于 $X \sim Pareto(\alpha, \lambda)$，令 $W \sim (X - M)|_{[X>M]}$，则 $W \sim Pareto(\alpha, \lambda + M)$。因

此下面的恒等变换将向 W 推进：

$$\begin{aligned}\frac{E(Y)}{E(Z)} &= \frac{\int_0^M x dF_X(x) + M\bar{F}_X(M)}{\int_M^\infty (x-M)dF_X(x)} \\ &= \frac{E(X) - \int_M^\infty (x-M)dF_X(x)}{\int_M^\infty (x-M)dF_X(x)} \\ &= \frac{E(X) - \bar{F}_X(M)E(W)}{\bar{F}_X(M)E(W)} \\ &= \left[\frac{\delta}{\beta-1} - \left(\frac{\delta}{\delta+M}\right)^\beta \left(\frac{\delta+M}{\beta-1}\right)\right] \Big/ \left[\left(\frac{\delta}{\delta+M}\right)^\beta \left(\frac{\delta+M}{\beta-1}\right)\right] \\ &= \left(1+\frac{M}{\delta}\right)^{\beta-1} - 1 \\ &\geq \xi/\theta - 1\end{aligned}$$

上面的变换利用了 Pareto 分布的诸多性质，大家可以多体会。此时，解上述不等式，可以知道：

$$M \geq \delta\left[\left(\frac{\xi}{\theta}\right)^{1/(\beta-1)} - 1\right]$$

令 $M^* = \delta\left[\left(\frac{\xi}{\theta}\right)^{1/(\beta-1)} - 1\right]$，这就是获得非负净利润时所需的最低的 M，代入可得 $M^* = 269.6938$。

下面一部分是要计算 M 和预期年利润之间的关系。对于单个赔案来讲，

$$\begin{aligned}Expected\ Net\ Profit\ per\ Claim &= (1+\theta)E(X) - (1+\xi)E(Z) - E(Y) \\ &= \theta E(X) - \xi E(Z) \\ &= \theta E(X) - \xi \bar{F}_X(M)E(W) \\ &= 0.2 \times \frac{1200}{3-1} - 0.3\left(\frac{1200}{1200+M}\right)^3 \times \frac{1200+M}{3-1} \\ &= 120 - \frac{259200000}{(1200+M)^2}\end{aligned}$$

因为使用的总索赔模型是复合 Poisson 模型，所以

$$\begin{aligned}Expected\ Annual\ Profit &= \lambda \times \left[120 - \frac{259200000}{(1200+M)^2}\right] \\ &= 300\left[120 - \frac{259200000}{(1200+M)^2}\right]\end{aligned}$$

由此，便可填写题目中未知的表格元素：

表 5: 索赔限额 M 与总年度利润之间的关系

Retention limit	Expected annual profit
300	1440
800	16560
2000	28406.25

4.3 给分标准与批改评价

表 6: Question 21 给分标准（共 20 分）

采分点	分值
给出单个赔案净利润的表达式	3
给出 Y 和 Z 在利润非负时满足的不等式	3
计算得到 M^*	2
给出 M 与预期年利润之间的关系	3
正确填写问题中的表	9（一个 3 分）

本题做得不错，但部分同学漏做了题目。

5 Question 23: Individual risk model, LEV and stop-loss reinsurance

注. 本题制作了绘制 $L_S(M)$ 曲线的 R 语言代码，可通过查看代码学会绘制一个函数。[下载]

本题制作了计算 L_S^{-1} 的 R 语言代码，可通过查看代码学会使用优化算法解方程。[下载]

5.1 原题

In Example 3.11 total claims S arising from accidents of employees in a large factory were modeled by an individual risk model with mean $E(S) = 33,000$ and variance $\mathrm{Var}(S) = 35,151,000$. Approximating this distribution by a normal distribution with the same mean and variance, plot the limited expected value function $L_S(M)$. The insurance company is presently using a security loading of $\theta = 0.37$, and is considering stop-loss reinsurance with stop-loss level M. Determine the minimal values of the stop-loss level M which should be considered to ensure expected profits are nonnegative when the stop-loss premium loading ξ of the reinsurer is both 0.5 and 0.7 .

5.2 参考答案

由课本第 117 面，对于标准正态分布，其 LEV 为：

$$\begin{aligned} L_{N(0,1)}(M) &= \int_{-\infty}^{M} x\phi(x)dx + M[1-\Phi(M)] \\ &= -\phi(M) + M[1-\Phi(M)] \end{aligned} \tag{2}$$

其中，$\phi(M)$ 和 $\Phi(M)$ 分别是标准正态分布的概率密度函数和累积分布函数。又由第 114 面，可得：

$$L_{aX+b}(M) = aL_X\left(\frac{M-b}{a}\right) + b \tag{3}$$

可以使用 $\mu = 33000$, $\sigma^2 = 35151000$ 的正态分布对本题中的个体风险模型进行近似。一般

的正态分布都可以被表示为诸如 $\mu+\sigma N(0,1)$ 的形式，代入(2)和(3)两式可得：

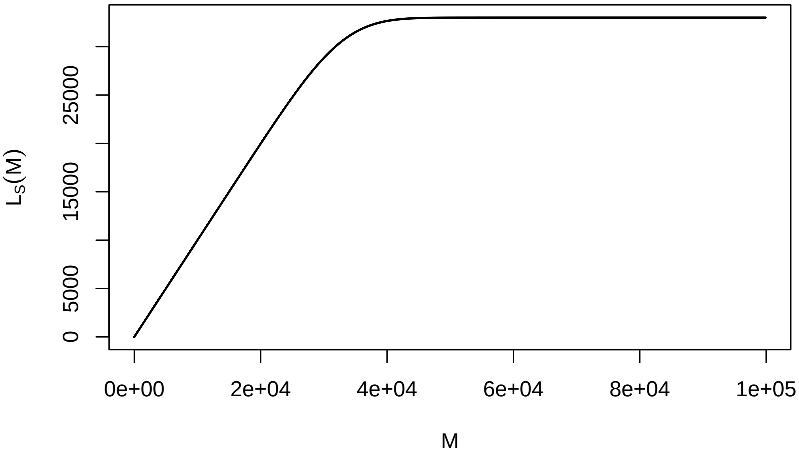
$$\begin{aligned}L_S(M) &\doteq \sigma L_{N(0,1)}\left(\frac{M-\mu}{\sigma}\right)+\mu\\ &=-\sigma\phi\left(\frac{M-\mu}{\sigma}\right)+\sigma\frac{M-\mu}{\sigma}\left[1-\Phi\left(\frac{M-\mu}{\sigma}\right)\right]+\mu\end{aligned}$$

用 R 或 EXCEL 都可以很好地画出上述 $L_S(M)$ 的图像，我只提供了 R 代码，大家可以跳到本节开头下载学习。下面展示了使用到的 R 代码，首先，下面定义了一个函数，用于生成 LEV 函数的值：

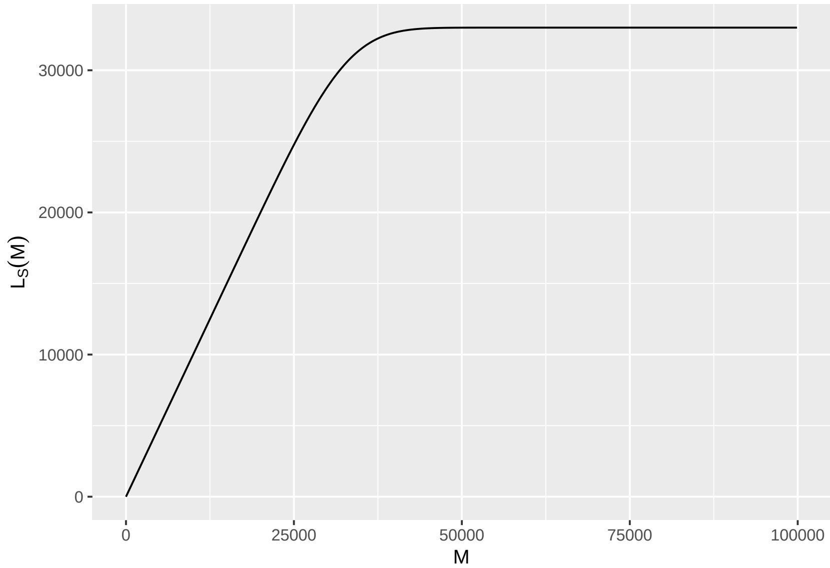
```
# 编程制作 LEV 函数
LEV.normal <- function(mu, sigma, M) {
    # 将 (M-mu)/sigma 视为一个新的变量，使后面代码更加简洁
    Normalized <- (M - mu)/sigma
   # 写出 LEV 的表达式
    lev <- - sigma * dnorm(Normalized)
           + sigma * Normalized * (1 - pnorm(Normalized)) + mu
    return(lev)
}
# 代入 mu、sigma 和一串 M，返回一个向量
mu <- 33000
sigma <- sqrt(35151000)
M <- seq(from = 0, by = 100, length.out = 1000)
LEV <- LEV.normal(mu = mu,sigma = sigma,M = M)
```

接着，可以使用 `plot` 函数或 `ggplot2` 包绘制出 LEV 函数的图像。可以看出，LEV 函数单调递增，但随着 M 变得很大，LEV 函数几乎不再上涨，因为此时尾部的概率已经非常小，LEV 函数近似等于 $E(S)$。

```
# 使用最简单的 plot 函数进行绘图
# 指定 x 和 y
# 将点的大小（cex）设为 0.1，防止曲线过粗
plot(M, LEV, cex = 0.1)
```




```r
library(ggplot2)
result.frame <- data.frame(M, LEV)
ggplot(data = result.frame, mapping = aes(x = M, y = LEV))+
  geom_line()
```



令 S_I 为总索赔中需要保险人承担的那部分，S_R 为总索赔中需要再保险人承担的那部分。对于保险人来说，其利润是保费减去再保险费和未来索赔：

$$\begin{aligned} Net\ Profit &= (1+\theta)E(S) - (1+\xi)E(S_R) - E(S_I) \\ &= \theta E(S_I) - (\xi-\theta)E(S_R) \end{aligned}$$

利润非负时，有 $E\left(S_I\right)/E\left(S_R\right) = L_S(M)/\left[E(S) - L_S(M)\right] \geq \xi/\theta - 1$，也即：

$$L_S(M) \geq \left(1 - \frac{\theta}{\xi}\right) E(S)$$

当利润恰好为非负时，等号成立。由课本 117 面最底下的公式，此时有：

$$L_S(M) = \sigma L_{N(0,1)}\left(\frac{M-\mu}{\sigma}\right) + \mu = \left(1 - \frac{\theta}{\xi}\right)E(S)$$

解得 $L_{N(0,1)} = -\frac{\mu}{\sigma}\frac{\theta}{\xi}$，因此有

$$M^* = \mu + \sigma L^{-1}_{N(0,1)}\left(-\frac{\mu}{\sigma}\frac{\theta}{\xi}\right)$$

当 $\xi = 0.5$ 时，$L^{-1}_{N(0,1)} = -4.118854$，$M^* = 8580.025$；当 $\xi = 0.7$ 时，$L^{-1}_{N(0,1)} = -2.941572$，$M^* = 15559.92$。

上面给出的答案非常简练，但算出 $L^{-1}_{N(0,1)}$ 是一件比较难的事情。因为 $L_{N(0,1)}$ 的公式中有正态分布的累积分布函数，这意味着我们没办法得到反函数的解析解。不过没关系，只要算出数值解，就能够解决这道题。我提供了 R 代码计算 $L^{-1}_{N(0,1)}$，大家可以回到这一节开始处下载学习[6]。下面展示了这些代码。

在 R 中，有一个非常好用的优化函数，名为 optim，其可以计算一元或多元函数的最小值，以及取得最小值时函数自变量的取值。现在设定一个函数：

$$g(M) = \left[L_{N(0,1)}(M) - \left(-\frac{\mu}{\sigma}\frac{\theta}{\xi}\right)\right]^2$$

上述函数的最小值是 0，也就是说，如果能够找到令 $g(M)$ 取得最小值 0 的 M，我们就找到了对应的 $L^{-1}_{N(0,1)}$。这样，一个解方程的问题就被转化为一个优化问题。

```r
# 制作 LEV 函数
LEV.standard.normal <- function(M){
  lev <- - dnorm(M) + M *(1 - pnorm(M))
return(lev)
}
# 写下所有参数
mu <- 33000
sigma <- sqrt(35151000)
theta <- 0.37
xi <- 0.5
# 我们的目标就是让 L(X) = - mu / sigma * theta / xi
LX <- - mu / sigma * theta / xi
# 设定目标函数供 optim 来优化
f <- function(X){
# 这里稍微设计了一下这个函数
# optim 让目标函数最小化
# 如果不平方或者取绝对值的话，无法得出想要结果
  (LEV.standard.normal(X)-LX)^2
}
```

[6]EXCEL 的用法已经在微信群中关于例 3.18 的视频里讲过，大家可以再用一下规划求解。

在优化问题中，初值选取和约束条件的选取非常重要，当 $\xi = 0.5$ 时算出来的 $-\frac{\mu}{\sigma}\frac{\theta}{\xi}$ 为 -4.11886。通过观察课本 119 面 Figure 3.6（如下图所示），我们可以初步断定 $L^{-1}_{N(0,1)}$ 要比 -4 小一些。因此，我们可以在 `optim` 函数中规定 M 的上限和下限，提高算法收敛速度。

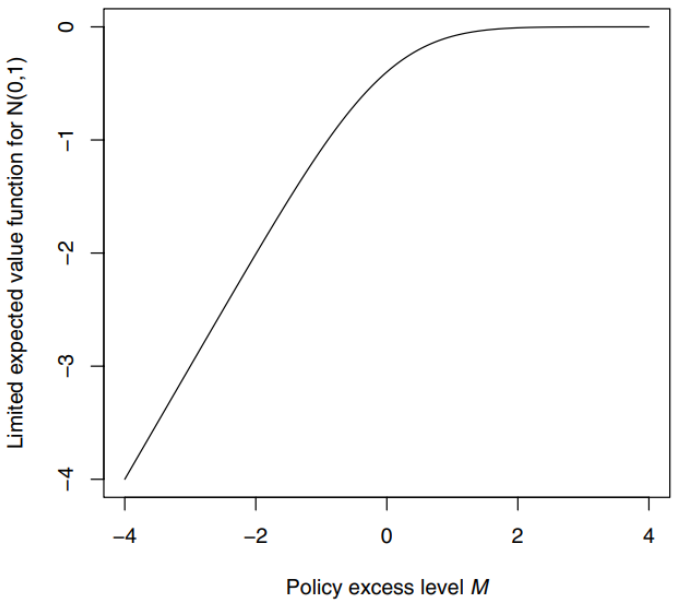


图 2: 标准正态分布的 LEV 函数

```
# 设定初值、方法、限制和最大迭代次数
result <- optim(-4, f,
method = "Brent",
lower = -6, upper = -4,
control= list(maxit = 1000))
# 得出最优的 M
M <- mu + sigma * result[["par"]]
M
```

```
## [1] 8580.025
```

5.3 给分标准与批改评价

本题大多数同学回答得还不错。本题分档次给分：能够写出 M^* 的表达式的，给 12 分；能计算的，给 16 分；可绘制 LEV 函数图像的，20 分（包括手绘图像）。

6 批改评述总结

本次作业共六题，各题分值总结如下：

表 7: 各题分值分布及班级卷面平均分

题号	分值	全班均分（不含迟交、漏交和严重抄袭）
3	15	13.24
5	25	13.40
15	20	15.00
21	20	15.70
23	20	12.41
总分	100	69.76

本次做得最差的题是第 5 题，同学们要好好复习 Panjer 递推，这是期末重点。