ON THE PROBABILITY AND THE TIME OF INSURANCE RUIN

Jean Tao

Abstract. The estimation of ruin probability has been the central topic in insurance risk theory. In this paper we study the asymptotic behavior of the probability of ruin and the probability that ruin occurs before the end of planning years in the compound Poisson model. The exponential bounds for both probabilities are found to be functions of the rate function in traditional large deviation theory.

1. Introduction In classical risk theory, extensive studies have been done on the aggregate insurance claim amount, the amount of an insurer's surplus at time t and the probability of eventual ruin [2, 5, 6]. In this paper, we study the asymptotic behavior of the claim amount and surplus processes over an extended period of time using large deviation techniques. Let us start with the definitions of three important terms in risk theory.

a) Aggregate Claim Process
Let $X_1, X_2, \ldots$ denote a sequence of identically and independently distributed insurance claim amount random variables with common distribution function $F(x)$ and mean μ. Let $N(t)$ be a homogenous Poisson claim number process with constant intensity λ and be independent of all X_i's. For any $t \geq 0$, define the aggregate claims paid up to time t to be

$$S_t = X_1 + X_2 + \cdots + X_{N(t)}.$$

S_t is said to be a compound Poisson Process.

The following Lemma for S_t is trivial from fundamental probability theory.

Lemma 1.1. Let $m(\theta)$ be the moment generating function of X_1. Then i) $E[S_t] = \lambda\mu t$ and ii) the moment generating function of S_t is given by

$$M_{S_t}(\theta) = e^{\lambda(m(\theta)-1)t}.$$

b) Insurer's Surplus Process
The insurer's surplus at time $t > 0$, denoted by $U(t)$, is defined as the excess of the initial fund plus premiums collected over claims paid through time t. For simplicity, we only consider the case that premiums are collected continuously at a constant rate c and the time value of money is not recognized. Then the insurer's surplus process is written as

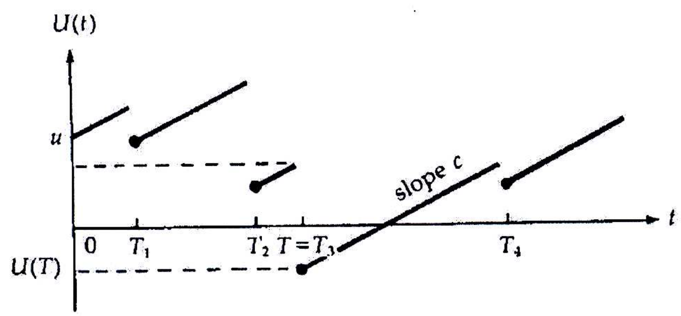
$$U(t) = u + ct - S_t,$$

where u is the initial surplus.

c) Probability of Ruin
We say that ruin occurs when $S_t > ct + u$ for the first time. The time of ruin is denoted by $T(u)$ and

$$T(u) = \inf\{t : S_t > ct + u\}.$$

An example of surplus process and the time of ruin is illustrated below.



The probability of eventual ruin is defined to be

$$\psi(u) = P(T(u) < \infty).$$

It is well-known [2] that when $c \leq \lambda\mu$, $\psi(u) = 1$ for any $u \geq 0$ and when $c > \lambda\mu$, $0 < \psi(u) < 1$ for $u \geq 0$.

In this paper we are concerned with the probability that ruin occurs at the end of a particular planning year, i.e.

$$P\{S_t > u + ct\} \tag{1}$$

and the probability that ruin has already occurred before the end of the planning year, i.e.

$$P\{T(u) < tu\} \tag{2}$$

for positive t and for $c > \lambda\mu$.

It is shown that when t is large, both probabilities are approximately equal to $e^{r(c)t}$, where $r(c) < 0$ and depends only on c and the original process. The $r(a)$ is the so-called exponential rate function for this stochastic process in large deviation theory. It is also shown that two probabilities are related in a natural way.

2. Preliminaries. For the compound Poisson process $\{S_t\}$, assume that the moment generating function $m(\theta)$ of X_1 exists on some open interval D_m containing origin. Define $h(\theta)$ to be a function which is equal to

$$h(\theta) = \lambda[m(\theta) - 1].$$

In D_m, $h(\theta)$ is infinitely differentiable and convex with $h(0) = 0$. Let function $r(a)$ be the convex conjugate (Legendre transform) of $h(\theta)$, i.e.

$$r(a) = \inf_{\theta}\{h(\theta) - a\theta\}. \tag{3}$$

Note that $r(a)$, the rate function of the compound Poisson process, is determined parametrically for each fixed a by the equations

$$h'(\theta) = a \;\text{ and }\; r(a) = h(\theta) - a\theta. \tag{4}$$

The following lemma summarizes the properties of $r(a)$.

Lemma 2.1. Assume that $m(\theta) < \infty$ for θ in some open interval D_m containing origin and assume that for each a, the solution θ_a to equation $a = h'(\theta)$ exists and lies in the interior of D_m. Then
i) $r(a) = \lambda(m(\theta_a) - 1) - a\theta_a$, where θ_a is the unique solution to the equation

$$a = \lambda m'(\theta);$$

ii) $r(a)$ is strictly concave down and infinitely differentiable with maximum 0 attained at $a = \lambda\mu$;
iii) For any $a \geq \lambda\mu$,

$$\inf_{\theta}\{h(\theta) - a\theta\} = \inf_{\theta\geq 0}\{h(\theta) - a\theta\}.$$

Proof. i) Since $m(\theta)$ is finite in an open interval around origin, $m(\theta)$ is infinitely differentiable in D_m. Furthermore,

$$h''(\theta) = \lambda m''(\theta) > 0. \tag{5}$$

This implies that $h(\theta)$ is strictly concave up and that $a = h'(\theta)$ defines a one-to-one strictly increasing and infinitely differentiable mapping. Therefore, the $\inf_\theta\{h(\theta) - a\theta\}$ is obtained when $\lambda m'(\theta) = a$ and for each a, there exists a unique solution θ_a. Plugging into the definition of $r(a)$, we have Part i).

For Part ii), we use the fact that $h(\theta)$ is strictly concave up and that $a = h'(\theta)$ defines a one-to-one strictly increasing and infinitely differentiable mapping along with the following relations:

$$r(a) = h(\theta_a) - \theta_a a, \;\text{ when }\; a = h'(\theta_a).$$

The inverse mapping is actually given by

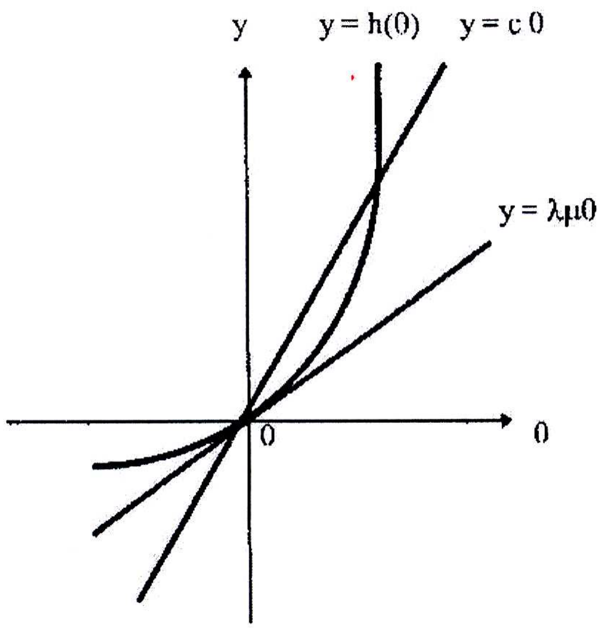
$$\theta = -r'(a).$$

Consequently,

$$r''(a) = -\frac{d\theta}{da} = -\frac{1}{h''(\theta)} < 0,$$

which implies that $r(a)$ is a concave down function.

The proof of Part iii) could be easily obtained from the following picture about the relation between functions $y = h(\theta)$ and $y = a\theta$ with $a > \lambda\mu$.



3. Main Results. Our first result is an upper bound for the probability $P\{S_t > u + ct\}$.

Theorem 2.1. Assume that $m(\theta) < \infty$ for θ in some open interval D_m containing origin. Then for any $t \geq 0$,

$$P(S_t > ct + u) \leq e^{r(c)t}.$$

Proof. For any $\theta > 0$,

$$\begin{aligned} P[S_t > ct + u] &\leq P[S_t > ct] \\ &= P\left[e^{\theta S_t} > e^{\theta ct}\right] \\ &\leq e^{-\theta ct} E(e^{\theta S_t}) \\ &= e^{[-\theta c + \lambda(m(\theta)-1)]t}. \end{aligned} \tag{6}$$

Here, the inequality is Chebycheff's. Taking the infimum for all $\theta > 0$ on the right hand side, by Lemma 2.1 iii) we recognize that the exponent is simply $r(c)t$. This proves the upper bound.

As to the lower bound, we consider the special case when t is a positive integer. Note that by Lemma 1.1,

$$M_{S_t}(\theta) = (e^{\lambda(m(\theta)-1)})^t.$$

Due to the uniqueness of the moment generating function, there exists a sequence of i.i.d. random variables $Y_1, \ldots, Y_t$ with $E(Y_1) = \lambda\mu$ and $M_{Y_1}(\theta) = e^{\lambda(m(\theta)-1)}$ such that $S_t = Y_1 + \cdots + Y_t$. By the Law of Large Numbers,

$$\lim_{t\to\infty} P\left(\frac{S_t}{t} \to \lambda\mu\right) = 1.$$

The event $\{S_t > u + ct\}$ with $c > \lambda\mu$ is an event for S_t to be away from its central mean $\lambda\mu t$ on a large scale (scale of t). Consequently the probability is very small. Furthermore since u is a fixed constant, the change in u is relatively small (scale of constant) compared with the change in c. Therefore, the asymptotic expression for the probability $P(S_t > ct + u)$ in the case of $u = 0$ is almost the same as in the case of $u > 0$. We will provide a proof of the lower bound for the case that $u = 0$.

Theorem 2.2. Assume that $m(\theta) < \infty$ for θ in some open interval D_m containing origin and assume that the solution θ_c to the equation $c = \lambda m'(\theta)$ exists and lies in the interior of D_m. Then for every $0 < \epsilon < c$ and any $0 < \delta < 1$, there exists a number $t_0 > 0$ such that for every $t \geq t_0$,

$$P[S_t > (c-\epsilon)t] \geq (1-\delta)e^{(r(c)-\epsilon\theta_c)t}.$$

Proof. The main idea to prove the lower bound is first to shift the center of the process to ct using Esscher transform and then to use the Law of Large Numbers to estimate the transformed probability.

Let $Y_1, Y_2, \ldots$ be a sequence of i.i.d. random variables whose common moment generating function is equal to $e^{\lambda(m(\theta)-1)}$ and whose distribution function is denoted by $F^*(y)$. Then we have

$$\begin{aligned} P[S_t > (c-\epsilon)t] &= P[Y_1 + \cdots + Y_t > (c-\epsilon)t] \\ &= \int \cdots \int_{y_1+\cdots+y_t > (c-\epsilon)t} dF^*(y_1)\cdots dF^*(y_t). \end{aligned}$$

Define the Esscher transform of $F^*(y)$ by

$$dG(y) = \frac{e^{\theta_c y} dF^*(y)}{e^{\lambda(m(\theta_c)-1)}},$$

where θ_c satisfies Equation: $c = \lambda m'(\theta)$. We claim that under this new distribution $G(y)$, the center of Y_i is shifted to c, i.e. $E_G(Y_1) = c$. This is true because

$$\begin{aligned} E_G(Y_1) &= \frac{\int_0^\infty y e^{\theta_c y} dF^*(y)}{e^{\lambda(m(\theta_c)-1)}} \\ &= \frac{m'_{F^*}(\theta)}{m_{F^*}(\theta)} \\ &= \lambda m'(\theta_c) \\ &= c. \end{aligned}$$

Applying the Law of Large Numbers to $\{Y_i\}$, we conclude that for any $\epsilon > 0$,

$$P_G\left[c - \epsilon < \frac{Y_1 + \cdots + Y_t}{t} < c + \epsilon\right] \to 1 \text{ as } t \to \infty.$$

Hence, for any $\epsilon > 0$ and any $\delta > 0$, there exists a $t_0 > 0$ such that for $t \geq t_0$ we have

$$\int \cdots \int_{c-\epsilon < \frac{y_1+\cdots+y_t}{t} < c+\epsilon} dG(y_1)\cdots dG(y_t) \geq 1 - \delta.$$

So

$$\begin{aligned} &P[S_t > (c-\epsilon)t] \\ &= e^{\lambda(m(\theta_c)-1)t} \int \cdots \int_{y_1+\cdots+y_t > (c-\epsilon)t} e^{-\theta_c(y_1+\cdots+y_t)} dG(y_1)\cdots dG(y_t) \\ &\geq e^{\lambda(m(\theta_c)-1)t} \int \cdots \int_{c-\epsilon < \frac{y_1+\cdots+y_t}{t} < c+\epsilon} e^{-\theta_c(y_1+\cdots+y_t)} dG(y_1)\cdots dG(y_t) \\ &\geq e^{\lambda(m(\theta_c)-1)t} e^{-\theta_c(c+\epsilon)t} \int \cdots \int_{c-\epsilon < \frac{y_1+\cdots+y_t}{t} < c+\epsilon} dG(y_1)\cdots dG(y_t) \\ &\geq (1-\delta) e^{[r(c)-\epsilon\theta_c]t}, \end{aligned}$$

which implies the lower bound.

Combining Theorems 2.1 and 2.2 and letting $\epsilon \to 0$ and $\delta \to 0$, we conclude that $P[S_t > ct + u] \approx e^{r(c)t}$.

The next theorem estimates the probability of ruin occurring before a certain time.

<u>Theorem 2.3</u>. Suppose that the equation $h(\theta) = c\theta$ has a positive root R for every $c > h'(0) = \lambda\mu$. Then we have

i) $-R = \sup_{y>0} \Lambda(y)$, where $\Lambda(y) = yr(c+1/y)$ and the supremum is achieved at Y, i.e. $-R = \Lambda(Y)$.
ii) R and Y are determined by the equations $h(R) = cR$ and $1/Y = h'(R) - c$.
iii)

$$P(T(u) \leq uy) \leq \begin{cases} e^{uyr(c+1/y)} & \text{if } y < Y \\ e^{-uR} & \text{if } y \geq Y; \end{cases} \tag{7}$$

and

$$P(uy \leq T(u) < \infty) \leq \begin{cases} e^{uyr(c+1/y)} & \text{if } y > Y \\ e^{-uR} & \text{if } y \leq Y. \end{cases} \tag{8}$$

Proof. Using Equations (4) along with the facts that $r'(c) = -\theta_c$ and $r''(c) = -1/h''(\theta_c)$, it is not hard to see that

$$\Lambda(y) = y\left[h(\theta) - \theta\left(c + \frac{1}{y}\right)\right] = y(h(\theta) - \theta c) - \theta,$$
$$\Lambda'(y) = r\left(c + \frac{1}{y}\right) - \frac{1}{y}r'\left(c + \frac{1}{y}\right) = h(\theta) - \theta c,$$

and

$$\Lambda''(y) = \frac{1}{y^3}r''\left(c + \frac{1}{y}\right) = -\frac{1}{y^3 h''(\theta)} < 0,$$

with θ given by the solution of $h'(\theta) = c + 1/y$.
So $\Lambda(y)$ is strictly concave down and $h(R) = cR$ with $\theta = R$ determined by $h'(R) = c + 1/Y$ and $\Lambda(Y) = -R$.

In order to prove Equations (7) and (8), we note that

$$E\left(e^{\theta S_t - th(\theta)}\right) = 1 \tag{9}$$

for any fixed $t > 0$.
Wald's relation states that Equation (9) is also valid for any random stopping time such as $T(u)$ if $h'(\theta) > c$, so that

$$E\left(e^{\theta S_{T(u)} - T(u)h(\theta)}, T(u) < \infty\right) = 1.$$

Since $S(T(u)) \geq u + cT(u)$, we have

$$\begin{aligned} 1 &\geq E\left(e^{\theta u - T(u)(h(\theta) - c\theta)}, T(u) \leq uy\right) \\ &\geq e^{\theta u - uy(h(\theta) - c\theta)} P(T(u) \leq uy). \end{aligned}$$

To minimize the exponent in the above line we would like to put $h'(\theta) = c + 1/y$ to obtain $uyr(c + 1/y)$. Since $h'(R) = c + 1/Y$, we have $\theta \geq R$ when $y \leq Y$, and hence, $h(\theta) \geq c\theta$ as required.

When $y \geq Y$ we take $\theta = R$ and get the second part of (7). The proof of (8) is quite similar.

4. Examples. The large deviation rate function is usually very difficult to find. Here, we give two examples where the rate functions could be explicitly calculated.

Example 1. Consider a portfolio of life insurance policies with the same unit amount death benefit. Assume that the number of death claims received by the insurance company in a year is distributed according to a Poisson distribution with mean λ. Let X_i be the amount of the ith death claim. Then we have $X_i = 1$ with probability 1,

$$m_X(\theta) = e^{\theta} \quad \text{and} \quad E(e^{\theta S_t}) = e^{\lambda t(e^{\theta}-1)}.$$

This implies that S_t is a Poisson Process at rate λ. The rate function is calculated by equations:

$$\lambda e^{\theta_a} = a \quad \text{and} \quad r(a) = \lambda(e^{\theta_a} - 1) - \theta_a a.$$

Hence,

$$r(c) = (c - \lambda) - c \ln\left(\frac{c}{\lambda}\right).$$

Let $c = (1 + \alpha)\lambda$ with $\alpha > 0$ the security loading. Then

$$r(c) = \alpha\lambda - (1 + \alpha)\lambda \ln(1 + \alpha).$$

By Taylor expansion, the right hand side is approximately equal to $-\frac{\alpha^2}{2}\lambda$ (assuming $\alpha < 1$).

Example 2. Consider a portfolio of auto insurance policies with claim frequency distribution to be Poisson at a rate of λ per year and severity distribution to be exponential of mean μ. We have $m(\theta) = \frac{1}{1-\mu\theta}$ and $h(\theta) = \frac{\lambda\mu\theta}{1-\mu\theta}$.
A simple calculation gives

$$\theta_c = \frac{1}{\mu} - \sqrt{\frac{\lambda}{\mu c}}$$

and

$$r(c) = -\frac{c}{\mu} + 2\sqrt{\frac{\lambda c}{\mu}} - \lambda.$$

Let $c = (1 + \alpha)\lambda\mu$ with $0 < \alpha < 1$ the security loading. Then

$$r(c) = -(1 + \alpha)\lambda + 2\sqrt{(1 + \alpha)}\lambda - \lambda.$$

Again using Taylor expansion, we can see that

$$P(S_t > (1+\alpha)\lambda\mu t) \approx ke^{-\frac{\alpha^2}{4}\lambda t}.$$

5. Conclusion. Estimating the probability of insurance loss has been a big challenge to researchers in risk theory. This paper provided an approach to analyze the probabilities of insurance losses on a large scale using traditional large deviation techniques.

References

1. R. R. Bahadur and S. L. Zabell, "Large Deviations of the Sample Mean in General Vector Spaces," *Annals of Prob.*, 7 (1979), 587–621.
2. N. Bowers, H. Gerber, J. Hickman, D. Jones, and C. Nesbitt, *Actuarial Mathematics*, Second Edition, The Society of Actuaries, Schaumburg, Illinois, 1997.
3. H. U. Cramer, "On a New Limit Theorem in the Theory of Probability," *Colloquium on the Theory of Probability*, Hermann, Paris, 1937.
4. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. II, John Wiley and Sons, New York, 1966.
5. H. Gerber, *An Introduction to Mathematical Risk Theory*, S. S. Huebner Foundation, University of Pennsylvania, Philadelphia, PA, 1979.
6. H. U. Gerber and E. S. Shiu, "On the Time Value of Ruin," *North American Actuarial Journal*, 2 (1998), 48–78.
7. R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, Princeton, NJ, 1973.

Mathematics Subject Classification (2000): 60F05, 60F10, 60G50

Jean Tao
Department of Mathematics and Computer Science
University of Central Missouri
Warrensburg, MO 64093
email: tao@ucmo.edu