| For office use only | Team Control Number: | For office use only |
|---|---|---|
| T1 _____ | #2016026 | F1 _____ |
| T2 _____ | | F2 _____ |
| T3 _____ | | F3 _____ |
| T4 _____ | | F4 _____ |

# IMMC 2016

**The 2nd Annual International Mathematical Modeling Challenge**

**Summary Sheet**

(Attach a copy of this page to the front of your solution paper.)

## Summary Sheet

Screaming at the top of our lungs, we eagerly witness world-class athletes set world record after world record from the spectators' stand. However, behind the adrenaline-induced races, gripping their armrests tightly and watching the races tensely, are the organizing committee and their insurance company -- the organizing committee bears the risk of paying a significant amount of money as a bonus to the winner if he or she succeeds in setting a new world record, while the insurance company may suffer a loss from insuring such a race.

In order to help the insurance company determine the price of the premium they offer that ensures their financial stability, and the organizing committee to make a the correct decision on when considering purchasing insurance, we looked into the complexity of the probability of record-breaking, premium calculation and decision making under risk in this paper, and came up with three different mathematical models targeting different scenarios when making the above decisions.

The most important factor affecting both parties' decisions is the risk of such an investment, which can be seen as the probability of an athlete breaking the world record. We developed a model using regression to analyze past data collected on athletes and their past records, then computed the standard deviation and normal distribution to predict the next record-breaking as a reference for both the insurance company and the organizing committee. The model was first built around one potential world record-breaker, and was advanced to accommodate multiple potential record-breakers. Since subsequent models are largely based on the result of this model, to access our model, we tested it to ensure its feasibility.

To help the insurance company determine the price of the premium, we developed another model that takes expected claim amount, safety-loading and profit-gaining loading, administrative costs, commission for agents, settlement of compensation and other factors into consideration. After deriving the equation to do so, we applied the eigenvector centrality concept to rank the importance of each factor.

Our third and final model helps the organizing committee decide whether to purchase insurance or not by considering the profit they gain after each competition and the likelihood of paying the bonus across a certain period. The model was later generalised so that it can be applied for multiple sport events using a single equation.

Afterwards we analyzed the strengths and weaknesses of our models, covering areas including but not limited to feasibility, efficiency and ease of usage, through which our team hopes to provide a clearer overview and further understanding of the models and each of their the merits and limitation. With real-life simulating models that consider a variety of scenarios and different factors, it is our sincere wish to provide models that are helpful and useable for the insurance company and the organizing committee to make better decisions for their own benefit.

## Introduction

In athletics, athletes are often given a significant amount of bonus for breaking the world record of an event by the organizing committee of the competition, so as to attract top athletes to join said competition. As the risk of bearing a heavy financial burden if such an event takes place, the organizing committee has to decide whether to purchase insurance or not (in other words, self-insure). To aid the decision-making processes for both the organizing committee and the insurance company, we have developed several mathematical models that are presented in this paper, and test cases to assess the viability and sensibility of our models.

## Problem Restatement

This year's problem requires us to do 5 things:

1. Calculate the average cost of the bonus for the Zevenheuvelenloop.
2. From the insurance company's perspective, determine the insurance premium for the Zevenheuvelenloop by considering different factors such as the average cost of the bonus, operating costs, time value of money, etc.
3. From the organizing committee's perspective, a) identify the criteria used to determine the necessity of purchasing the insurance for the Zevenheuvelenloop, and b) determine whether or not they should.
4. From the organizing committee's perspective, determine the way the organizing committee should weight each factor in deciding whether they should purchase the insurance for each of 40 events.
5. From the organizing committee's perspective, develop a mathematical model for the decision making process.

## Overall Assumptions and Justifications

| **Assumption** | **Justifications** |
|---|---|
| Accidents do not happen before or during the competition, i.e. the runners do not get injured or quit the competition under any circumstances once they are enrolled in the competition. | This is for the ease of calculation and to simplify the problem. |
| The performance of the runners are consistent and do not deviate too much from his or her average performance level. | |
| All runners engaged in the race aim to break the world record. | |
| Each competition's result is independent of others. | |
| Peak period of runners is similar for all athletes participating in the same event. | This assumption is made with references to past medical research[1] on the peak periods of runners, muscle mass declinations and other factors. |
| Only the new record holder can receive the bonus. | This is for the ease of calculation and to simplify the problem, which also resembles real life situations where the organizing committee will do so to avoid a large financial burden. |

[1] Schulz, R., & Curnow, C. (1988). Peak Performance and Age Among Superathletes: Track and Field, Swimming, Baseball, Tennis, and Golf. *Journal of Gerontology, 43*(5).

| | |
|---|---|
| The first few records set in the earliest races are excluded from consideration during calculation. | Since the competitions might not have gained enough popularity to attract top runners, the records are not valid for projecting the probability of breaking records in the future. |
| All the money terms and prices in the paper are calculated in real terms without the effect of inflation. | This is to simplify our calculation and ensure the model can be used in the following years to come. |
| The year of usage of this model is seen as year 0 (or the $0^{th}$ year). | The model is designed to be used for not only this year, but anytime in the future; seeing the year of usage of the model as year 0 will allow calculations to be simplified. |

## Definition of Important Terms

| **Terms in the paper** | **Definition (explanation)** |
|---|---|
| Dynamic information | Variables that change throughout the years |
| Constant | Variables that stay constant throughout the years |
| Top athletes (Potential winners) | Athletes who have the potential to break the world record |
| Peak age | Period in an athlete's life during which his or her performance is at the highest level |
| Average cost of the bonus | The ratio of the amount of bonus and the expected number of times the event is replicated before the current record is broken |
| Insurer | Insurance company, the party responsible for collecting the premium and determining the amount of the premium |
| Insured | The organizing committee in the competition in this paper, the party responsible for paying the premium and making the decision of whether to purchase the premium or not |
| Event | When the world record is broken by an athlete and the organizing committee has to pay him or her a bonus |
| Claim | The amount the insurer can claim back once accident occurs |
| Coverage period | The period in which the insurer is protected by the insurance plan |

## Models

### 1. Finding the expected average cost (Question 1)

*1.1 Additional assumptions*

| Assumptions | Justifications |
|---|---|
| Top athletes have the chance of breaking the world record for more than once in successive years. | This is to simplify the problem and for the ease of our calculation |
| The performance of athletes obey the normal distribution. | |
| Chance of athlete breaking world record is exponentially proportional to the percentage difference of their personal best result and the world record. | |
| The average cost will not remain constant. | As society changes and technology develops, chances of nurturing top athletes also change and so does chance of breaking the world record. |
| The peak performance period of different athletes are the same. | This is to simplify our calculation and ensure the model can be used in the following years to come. |

*1.2 Addition Definition of Terms*

| Term | Definition |
|---|---|
| Average cost of year $y$ | The cost times the chance of recording breaking in the said year |

*1.3 Method analysis*

According to the definition of the average cost (the amount of bonus divided by the expected number of times the event is replicated before the current record is broken), since the former is fixed, the question then becomes finding the latter, which solely depends on the frequency/probability of breaking the world record. Therefore, we develop models to estimate the probability of the world record being broken.

Since the models of the following questions also rely on this probability, we have developed models that calculate not only the average cost of the 15K run as mentioned (i.e. Zevenheuvelenloop), but also that of all other athletics events. Two different models, the historical data model and the real time data model, are developed, and are optimised for different situations.

*1.4 Historical data model*

The historical data model directly calculates the probability of a top athlete breaking the world record using regression and analysis of historical data, and is best used for competitions that have few potential world-record breakers (for example, one in five years).

Data about past world records are collected, along with the year of achieving the record and the information of the athlete that set the record. To come up with a formula that can estimate when the next top athlete will be at peak age, By taking linear, exponential, logarithmic and quadratic regression on the data points of $h(x) =$ *the year the xth world record was broken* and using the one with the least error, we will be able to come up with a formula that can estimate when the next top athlete will be at the peak age.

Suppose the identified regression function is $f(x)$. Then the estimated time is $\mu = f(y+1)$, where $y$ is the number of data points currently on the graph. (Note: If more than one value needs to be estimated,

$f(y+2)$, $f(y+3)$ et cetera can also be calculated, but one should note that the error increases as the number of estimations increase.)

We will calculate the deviation of the regression (denoted by $\sigma$) of the derived function:

$$\sigma = \sqrt{(\sum_{i=1}^{n} |f(i) - h(i)|^2) \div n}$$

After the calculations of the desired values, we will progress to calculate the probability of the said top athlete breaking the record in each of the coming years. According our research[2], it is assumed that the performance of the athletes obey the normal distribution with mean at the peak age (that is, the top athlete has the highest probability of breaking the record at their peak age), the probability of the next top athlete breaking the world record in a certain year is:

$$q(x_0 : \textit{date of calculation}) = \int_{x_0-0.5}^{x_0+0.5} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

The reason why we can take the $\sigma$ as the standard deviation in the above equation is that the physical meaning of $\sigma$ is the error between the estimated peak age of past top athletes and their actual peak age; the more accurate our regression is, the more accurate our estimation of the peak age of athletes is (and thus a smaller $\sigma$), and the higher possibility the peak age of the future athlete can be estimated correctly, and vice versa. Thus, the error in our previous calculations can be used as the standard deviation in the above equation.

Then, the average cost of the competition for a year, $x$, will be $C \cdot q(x)$, where $C$ denotes the prize money of the competition should the world record be broken.

*1.5 Real time data model*

The real time data model analyzes the probability of an athlete with a certain personal best record (PB) breaking the world record, and counts the number of athletes that stand a chance at breaking the world record and their respective probabilities, thus finding the overall probability of having at least one athlete breaking the world record. This model is optimised for events where there are frequently multiple top athletes competing every year.

We first narrow down the pool of athletes who apply to the event to so-called "top athletes", those who stand a higher chance of breaking the world record, for the ease of calculation. This is done by comparing their current performances to the past performance of all world-record holders, as we believe them to be reasonable benchmarks for estimating an athlete's future likelihood for breaking a record. The closer an athlete is to the benchmark, the higher his/her chance. Athletes whose performance is too far off stand a minimal chance and are filtered off, as the possibility of them improving to reach world-record standard in such a short period of time is close to zero. (Due to the nature of this model, this procedure has to be carried out only after the list of applicants has been confirmed, and so we can safely assume the period of time from the application and the actual event is too short to cause significant errors.) All that remains is to quantify the probabilities of the remaining potential winners breaking the record, then the expected value and thus average cost can be calculated.

[2] Schulz, R., & Curnow, C. (1988). Peak Performance and Age Among Superathletes: Track and Field, Swimming, Baseball, Tennis, and Golf. *Journal of Gerontology, 43*(5).

Data about past world records are collected, along with information of the athletes breaking the record. The performance curve of the past athletes is analyzed. First, we define the peak performance period as the age range of world-record breakers when they broke the record, as athletes are in their peak physical status when they break the world records. For example, if the youngest athlete to ever break the world record is 22 years old when he or she broke it and the oldest is 30, then the peak performance period is 22 to 30 years old.

We then compare the performance of all world-record holders in the peak performance period to the world record they hold, and find the average percentage over the period. Finally we average the number over all athletes to obtain the benchmark percentage, denoted by $\Delta p$. For example, suppose there were $n$ past record holders, with record breaking time $a_1, a_2, ..., a_n$ respectively. Also suppose the identified peak performance period is from $y_1$ to $y_2$ years old, a total of $y$ years, and the personal best of athlete $i$ in age $x$ is $b_{ix}$. Then, the benchmark percentage is calculated as:

$$\Delta p = [\sum_{i=1}^{n}(\sum_{j=y_1}^{y_2}\frac{|b_{ij}-a_i|}{a_i}) \div y] \div n$$

Afterwards, current athletes who have achieved good results and can be reasonably deemed as potential winners will be analyzed according to the benchmark. If their personal best result is better than $W \times (1 \pm \Delta p)$, the sign of $\Delta p$ depending on the nature of the event (e.g. in track events it is $1 + \Delta p$, whereas in long jump, high jump etc. it is $1 - \Delta p$),where $W$ is the current world record, then they should be deemed as potential record breakers, or top athletes. The filter is created as there is an average difference between a world record breaker's usual performance and best performance, so there is a need to consider a range so that athletes whose performances fall into the range defined, (i.e. those who have a reasonable chance of breaking the world record) would also be considered when we calculate the probability of the occurrence of the next record breaking.

We will then use the method presented in the historical data model to calculate the performance of a top athlete during a certain year, $q(x)$. The steps of calculations are the same as that in the previous model, just that the samples will include not only past world record breakers, but also past top athletes whose personal best result is within $\Delta p \times 100\%$ of the closest future world record at their time. (A wider range of data allows for a more accurate estimation.) For example, suppose a world record is broken at 2003 and 2008; then athletes whose personal best performance between year 2004 and 2007 is within $\Delta p \times 100\%$ of the 2008 record should also be included in the data set for calculating $q(x)$.

Next, we assume the probability of an athlete breaking the world record is exponentially proportional to the percentage difference between their personal best result and the world record, as this follows the probability density function and resemble real life situations closely. The rationale behind why we need the following calculation is because there is a difficulty for an athlete to improve his/her performance to reach the world record: a 30-second difference between his/her PB and the world record and a 3-second difference do not imply a 10-time difference in their chances of breaking the world record; the athlete with the 3-second difference has an exponentially higher chance of breaking the record than the one with the 30-second difference. Hence, the probability of a top athlete $i$ with a personal best result $d_i$ breaking the world record in the $x^{th}$ year following year $w$ can be estimated as:

$$p_i(w,x) = e^{\frac{-|d_i-W|}{W}} \cdot q(w+x)$$

Lastly, we will consider the worst case scenario, where an estimated $m$ top athletes will all join the competition, with winning probability $p_1, p_2, ..., p_m$ respectively. The probability of at least one of them breaking the record in the $x^{th}$ year following year $w$ is:

$$p(w,x) = 1 - \prod_{i=1}^{m}(1 - p_i) = 1 - (1 - p_1(w,x))(1 - p_2(w,x))...(1 - p_m(w,x))$$

Naturally, one may question that even in the foreseeable future, the performance of top athletes will deteriorate while new top athletes would appear. Thus, for each top athlete that grows out of the peak performance period, we can assume another athlete with performance $p_{m+1}(w,x) = (\sum_{i=1}^{m} p_i(w,x)) \div m$ with age the young end of the peak performance period will appear for a more accurate model. Since the award money will be given if any one of them breaks the record (the more potential winners there are, their individual chances remain unchanged but the higher the chance of the record itself being broken) , the average cost of the competition will be $C \cdot p(w,x)$, where $C$ denotes the award amount.

*1.6 Comparison of Models*

Both models serve for different functions and are optimised in different situations. In situations where the data set is smaller (for instance, there is only one top athlete, or the database is not as large), the historical data model will be optimal as it utilizes a small amount of information to give a relatively accurate projection of the future. This is especially useful when insurance companies have nothing but data about that particular event on their hands.

On the other hand, when information and data is sufficient, the real time data model will definitely give a more reliable answer compared to the historical data model, as it utilizes data not only from history, but the performance curve of real-time athletes. Although the calculations of the real time data model is more complicated, it can also tackle situations such as multiple potential top athletes, which is in most cases, closer to the real world situation.

As such, the models to the following problems will be based on the real time data model, although the variables in the following models can also be easily replaced to be based on the historical data model.

*1.7 Testing of Models*

The real time data model is used for the test case.

| World Record Breakers | Personal Best of World Record Breakers in Age x (seconds) | | | | | | | | | World Record |
|---|---|---|---|---|---|---|---|---|---|---|
| | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | |
| Robert de Castella | | | | | | 2567 | | | 2726 | 2567 |
| Valdenor dos Santos | | | | 2561 | | | | | | 2561 |
| Philemon Hanneck | | | 2555 | | | | | | | 2555 |
| Josephat Machuka | | 2543 | 2586 | 2543 | | | | | | 2543 |
| Worku Bikila | | | | | | | | | 2540 | 2540 |
| Felix Limo | 2489 | | | | 2612 | | | | | 2489 |
| Leonard Komon | | 2473 | 2486 | 2506 | 2535 | 2605 | | | | 2473 |

The benchmark percentage is calculated as:

$$\Delta p = 0.011532104\ *$$

*Due to insufficient data, we were not able to find out the personal best of world record breakers for every age from 21 to 29. Therefore, $y$ in $\Delta p = [\sum_{i=1}^{n}(\sum_{j=y_1}^{y_2}\frac{|b_{ij}-a_i|}{a_i}) \div y] \div n$ denotes the number of years we have data for the personal best of world record breakers. For example, in the case of Robert de Castella, $y = 2$ , as we only have his personal best data for ages 26 and 29.

Past top athletes whose PB is within $\Delta p \times 100\%$ of the closest future world record at their time:

| | Year | Top Athlete | Time (seconds) | World Record |
|---|---|---|---|---|
| 1 | 1983 | Robert de Castella | 2567 | ✓ |
| 2 | 1984 | Mike McLeod | 2575 | |
| 3 | 1985 | Mike McLeod | 2582 | |
| 4 | 1986 | John Treacy | 2579 | |
| 5 | 1989 | Keith Brantly | 2570 | |
| 6 | 1993 | Valdenor dos Santos | 2561 | ✓ |
| 7 | 1994 | Haile Gebrselassie | 2580 | |
| 8 | 1994 | Philemon Hanneck | 2555 | ✓ |
| 9 | 1995 | Josephat Machuka | 2557 | |
| 10 | 1995 | Josephat Machuka | 2543 | ✓ |
| 11 | 1996 | Jon Brown | 2562 | |
| 12 | 1997 | Worku Bikila | 2540 | ✓ |
| 13 | 2001 | Felix Limo | 2489 | ✓ |
| 14 | 2005 | Sileshi Sihine | 2498 | |
| 15 | 2007 | Samuel Kamau Wanjiru | 2489 | |
| 16 | 2007 | Samuel Kamau Wanjiru | 2490 | |
| 17 | 2007 | Zersenay Tadese | 2494 | |

| | | | | |
|---|---|---|---|---|
| 18 | 2007 | Deriba Merga | 2494 | |
| 19 | 2007 | Patrick Makau Musyoki | 2494 | |
| 20 | 2007 | Evans Kiprop Cheruiyot | 2489 | |
| 21 | 2009 | Deriba Merga | 2489 | |
| 22 | 2009 | Patrick Makau Musyoki | 2490 | |
| 23 | 2009 | Wilson Kipsang Kiprotich | 2495 | |
| 24 | 2010 | Zersenay Tadese | 2493 | |
| 25 | 2010 | Leonard Patrick Komon | 2473 | ✓ |
| 26 | 2011 | Leonard Patrick Komon | 2486 | |
| 27 | 2011 | Samuel Tsegay | 2491 | |

By taking exponential regression on the data points of
$h(x)$ = *the year the xth athlete broke the world record or came close to breaking the world record*,
the formula that can estimate when the next top athlete will be at the peak age is:

$$y = 1.136141636\,x + 1984.390313$$

Then $q(2016) = \int_{2016-0.5}^{2016+0.5} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 0.1569$

If the personal best result of current athletes (2010-2015) is better than
$W \times (1 + \Delta p) = 2473 \times (1 + 0.011532104) = 2501.5$ seconds, then they are deemed as potential record breakers.

| | Year | Potential Record Breakers | Time | World Record |
|---|---|---|---|---|
| 1 | 2010 | Zersenay Tadese | 2493 | |
| 2 | 2010 | Leonard Patrick Komon | 2473 | ✓ |
| 3 | 2011 | Leonard Patrick Komon | 2486 | |
| 4 | 2011 | Samuel Tsegay | 2491 | |

$$p_1(2010, 6) = e^{\frac{-|2493-2473|}{2473}} \cdot q(2010 + 6) = 0.155636213$$
$$p_2(2010, 6) = e^{\frac{-|2473-2473|}{2473}} \cdot q(2010 + 6) = 0.1569$$

$$p_3(2010, 6) = e^{\frac{-|2486-2473|}{2473}} \cdot q(2010 + 6) = 0.156077376$$

$$p_4(2010, 6) = e^{\frac{-|2491-2473|}{2473}} \cdot q(2010 + 6) = 0.155762132$$

The probability of at least one potential record breakers breaking the record in the $6^{th}$ year following year 2010 is:

$$p(2010, 6) = 1 - \prod_{i=1}^{4}(1 - p_i) = 0.492803618$$

The average cost of the 15K race is:

$$C \cdot p(2010, 6) = 25000 \times 0.492803618 = €12320.09046$$

However, the actual probability that at least one of the potential record breakers break the record in the 6th year following 2010 should be lower than 0.49, as the test case demonstrated above is the worst case scenario, where all potential record breakers join the same competition. Moreover, there are two major 15K competitions held each year -- the Zevenheuvelenloop and the Gasparilla Distance Classic, and by using the fact that most potential record breakers only train for one 15K competition per year, the probability can be further reduced by half.

## 2. Criteria for determining the cost of insurance (Question 2)

*2.1 Additional assumptions*

- The insurance company is sensible, able to identify high-risk and low-risk investment, aims to gain a profit and avoids losing money.
- The insurance market is imperfectly competitive, and all the insurance policies are homogenous .
- All the clients or insured are able and willing to pay the whole sum of premium at one time.
- All monetary values in the following calculations are using real values without the effect of inflation.
- The insured will report the incident once occurs and the insurer will settle down the payment as soon as the report is made.
- The commission, marketing expenses and other administrative costs are constant .

*2.2 Method analysis*

To analyze the criteria the insurance company should consider and the method they should use to weigh each factor that determines the amount added to the average cost, we first separate the whole premium into different components, then analyze the factors affecting the value of each component and weigh their importance using the eigenvector centrality concept.

*2.3 The model*

To calculate the amount of premium the organizing committee has to pay, the insurance company has to consider a number of factors in order to offer an insurance that is attractive to potential clients while establishing a safety loading (an amount the insurance company charges to lower the probability of loss if the actual number of claims is greater than its expected value) that has the twofold purpose of preventing monetary loss and subsequent bankruptcy in the case of an underestimation of claim amount, and guarantees a profit in the case of an overestimation. Therefore, it is vital for the insurance company to consider both the statistical probability of claiming, which directly affects the expected value of the claim and consequently the value of the premium (according to the equivalence principle), as well as the relevant risk factors (i.e. the accuracy of the probability calculated) that affect the size of the safety loading (i.e. the reserve).

A few additional factors identified that will affect the premium price include but are not limited to:

- Expected interest rate in the economy which affects the proportion of the premium to be used for investment by the insurance company;
- Expenditure of the insurance company, also known as the insurer's expenses, which include administrative costs such as commission for agents, marketing expenses for promoting the offer, etc;
- Market forces, which affect the actual amount of the premium in real life. For example, as the competitiveness of the market increases, the amount of the actual premium may decrease to attract more clients, and vice versa, ceteris paribus.

Figure 2.3.1 shows the typical process for determining the price of the premium:

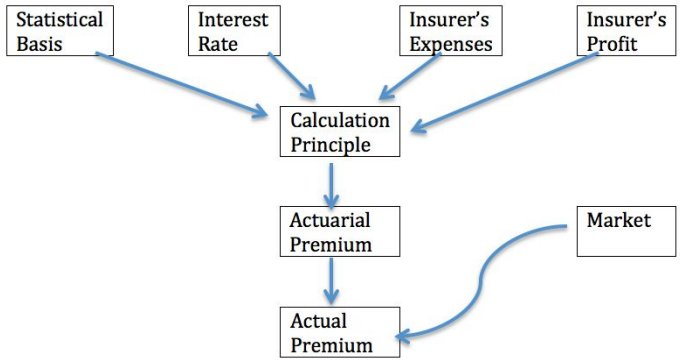



**Figure 2.3.1 process of determining the amount of premium**

There are a number of ways to calculate the actuarial premium. The one presented in this paper is the expense-loaded calculation method, which takes into consideration the administrative cost, the insurer's expenses and profit, as well as the safety loading portions. Thus, this method is a better simulation of real life situations than others (e.g. equivalence premium and net premium calculation, expected which only consider the expected claim value and the risk of each investment.

Before looking at the expense-loaded method, we first look at the calculation of

- the expected aggregate claim amount, which is the expected number of claims times the claim amount per claim (i.e. the total amount claimed);
- the net premium, which, as mentioned, considers the risk of offering such an insurance and the safety-loading portion on top of equivalence premium; finally
- the expenses-loaded premium, which is the net premium plus the added administrative costs and extra components that consider the insurer's expenses and profits.

In pricing the premium, the insurance covers a period of time known as the coverage period, during which the insured can make a maximum of $n$ number of claims. We can express the terms of the plan as $(y, n)$, where $y$ is the duration of the coverage period in terms of years. The claim amount is also fixed according to the terms in the plan, and will not change over the coverage period regardless of the inflation rate.

Assume that all claims are immediately reported and settled in all circumstances, the expected claim amount by the insurer throughout the coverage period is:

$$E[X] = E[N] \cdot X$$

where *E[N]* is the expected number of claims per insured and *X* is the previously fixed claim amount per claim. Note that *E[N]* is *n* times the probability of athletes breaking record over the coverage period, i.e. *p(x)* or *q(x)*, depending on the model used (refer to sections 1.3 and 1.4). The expected claim amount is also part of the consideration for safety loading of the insurance company.

Apart from the expected number of claims and the claim amount, a couple other factors will also affect the value of *E[X]* (the expected claim amount). For example, in practice, the insurance company would take a part of the premium received for investment to gain an income that can at least cover the promised claim amount. This can be represented by the following:

$$R = \Pi + k\Pi \cdot (1+i)^y$$

where $R$ is the total revenue made as a result of the insurance plan, which includes the collected premium $\Pi$ and the profit gained through investing part of the premium received $k\Pi$, $k$ is a fixed percentage, $i$ is the interest rate as determined by market forces, and $y$ is the coverage period in terms of years.

According to the usual practices of the insurance industry, insurers usually aim to generate a profit that is larger than the claim amount to ensure sufficient reserve in the company and avoid monetary loss should the actual claim be larger than the expected claim. In short,

$$k\Pi^{[X]} \cdot (1+i)^y \geq X .$$

In this case, $k\Pi^{[X]}$ should be considered as the safety loading of the premium, as a reward for the risk borne by the insurer, as well as the expected profit of the insurer. To include the risk factor and calculate a safety-loading proportional to the severity of the risk, the net premium is calculated as thus:

$$\Pi = E[X] + \lambda \, Var[X] \ ,$$

Where $\lambda$ is the given intensity, and

$$\begin{aligned} Var[X] &= E[X] - (E[X])^2 \\ &= E[N]Var[X_1] + E[N^2](E[X_1])^2 - (E[X])^2 \\ &= E[N]Var[X_1] + Var[N](E[X_1])^2 \end{aligned} \qquad (3)$$

Aside from considering the risk factor, calculating a safety-loading proportional to severity of the risk, the addition part, $\lambda$ *Var[X]*, variance of the amount the insurer has to pay back to the insured, also takes the time value of the money into consideration while calculating.

Lastly, administrative costs, commission costs, marketing costs et cetera have to be included in the calculation of the premium. Let $\Theta^A$, $\Theta^G$ be the fixed amount of commission and administrative costs and other expenses of the insurer as a result of offering this insurance service respectively, and $\Theta^S E[N]$ be the amount of expenses for each settlement made, where $E[N]$ is the expected frequency of claims made. Based on the flow chart and the factors to be considered, for a fixed claim amount *X,* the equation for calculating the expense-loaded actuarial premium $\Pi^{[X]}$ is given by:

---

[3] With accordance to a theorem introduced in Olivieri, A., & Pitacco, E. (2011). *Introduction to insurance mathematics: Technical and financial features of risk transfers*. Berlin: Springer Verlag.

$$\Pi^{[X]} = E[X] + \lambda Var[X] + \Theta^{A} + \Theta^{G} + \Theta^{s}E[N]$$

*2.4 Weighting*

Based on section 2.3, we can see that different components involved in the calculation of the price of premium depend on a group of other variables. To determine the importance of each variable, thus the importance of each component, we came up with a network showing the relationship between variables. Note that importance in this paper is determined by how much each component can vary, thus increasing or decreasing the price of the premium.

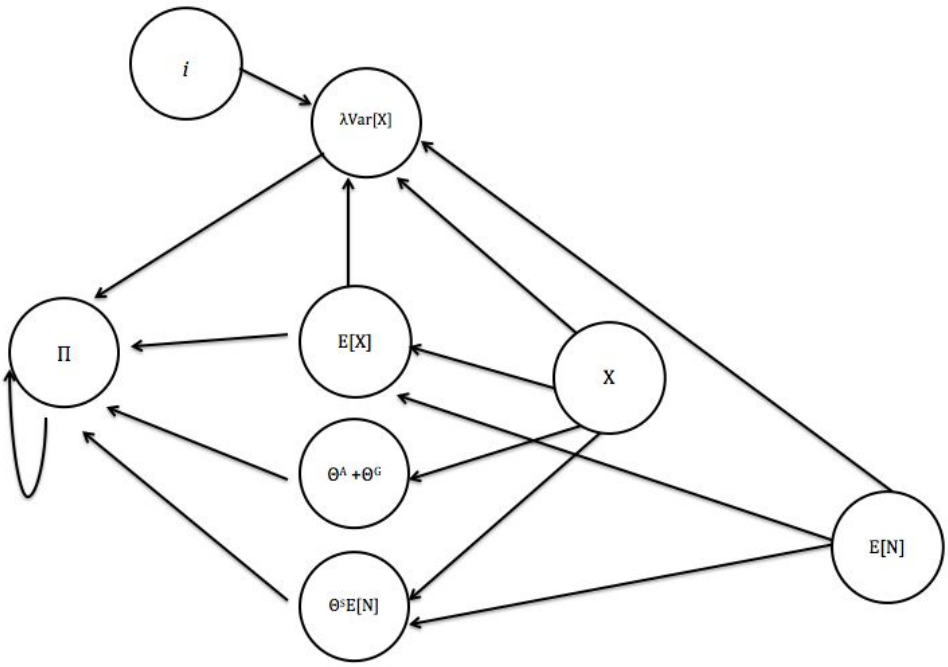



**Figure 2.4.1 Relationship between each component**
*(nodes are not drawn according to the importance yet)*

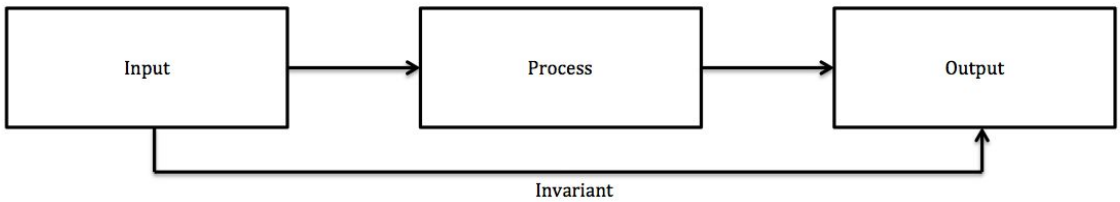
In the above diagram, we can see that a number of components, such as the amount of compensation the insurer promised to pay ($X$), the expected number of claims over the years ($E[N]$), and the interest rate for investment ($i$) affect the different components above according to the calculation of each.

To find out the importance of each factor, we are going to use the eigenvector centrality for the final ranking. The reason we chose the eigenvector centrality instead of other centralities such as the closeness centrality and degree centrality is because it can measure the influence of each node in the whole network as a big picture. According to the diagram, we can develop the following adjacency table:

| | E [N] | X | E [X] | λVAR [X] | $\Theta^{A}+\Theta^{G}$ | $\Theta^{S}$E [N] | Π | *i* |
|---|---|---|---|---|---|---|---|---|
| E [N] | ∅ | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| X | 0 | ∅ | 0 | 0 | 0 | 0 | 0 | 0 |
| E [X] | 1 | 1 | ∅ | 0 | 0 | 0 | 0 | 0 |
| λVAR [X] | 1 | 1 | 1 | ∅ | 0 | 0 | 0 | 1 |
| $\Theta^{A}+\Theta^{G}$ | 0 | 1 | 0 | 0 | ∅ | 0 | 0 | 0 |
| $\Theta^{S}$E [N] | 1 | 1 | 0 | 0 | 0 | ∅ | 0 | 0 |
| Π | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 |
| *i* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ∅ |

**Figure 2.4.2 Adjacency Table**

Where 1 indicates a directed relationship from one component to another; whereas 0 indicates no directed relationship. The process for calculating the eigenvector can be simplified as below:




$$X \rightarrow M \rightarrow M(x) = \lambda x$$

Where *M* is the adjacency matrix, *M(x)* is the re-distribution of nodal value after re-shuffling, λ is the eigenvector number or the eigenvector and *x* is the nodal value.

First, let us set the nodal value of all nodes to 1, that is $x = 1$, then we calculate the nodal value of *q* nodes based on the connection of each node they are connected to, that is

$$M(x) = (value\ of\ connected\ node\ 1)^{-1} + (value\ of\ connected\ node\ 2)^{-1} + ... + (value\ of\ connected\ node\ q)^{-1}$$

After the first round of reshuffle of nodal value, $M(x) \neq \lambda(x)$ , *M(x)* becomes the new nodal value of the node, however it does not reflect the ranking of importance of each node, thus it is not the eigenvector. Therefore, we then conduct another round of reshuffle of nodal value. The nodal value after each shuffle tends to the eigenvector as a limit.

We used Excel for the calculation and the result is shown in the figure below:

| | Numbers of nodes connected | $M(x)$ | $M(M(x))$ | $M(M(M(x)))$ | $M(M(M(M(x))))$ | $M(M(M(M(M(x)))))$ | $M(M(M(M(M(M(x))))))$ |
|---|---|---|---|---|---|---|---|
| $E[N]$ | 3 | 2.5 | 2.5 | 3 | 2.5 | 2.5 | 3 |
| $X$ | 4 | 3.5 | 3.5 | 4 | 3.5 | 3.5 | 4 |
| $E[X]$ | 2 | 2 | 1.285714286 | 2 | 2 | 1.285714286 | 2 |
| $\lambda VAR[X]$ | 1 | 1 | 0.285714286 | 1 | 1 | 0.285714286 | 1 |
| $\Theta^A+\Theta^G$ | 1 | 1 | 0.285714286 | 1 | 1 | 0.285714286 | 1 |
| $\Theta^S E[N]$ | 1 | 1 | 0.285714286 | 1 | 1 | 0.285714286 | 1 |
| $\Pi$ | 1 | 3.5 | 3.5 | 1 | 3.5 | 3.5 | 1 |
| $i$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

**Figure 2.4.3 Result table of eigenvector**

From the above result table, we can see that the most important factor that affects the value of the premium is the amount the insurer has to pay back to the insured (ie. $X$), the second most important factor is the expected number of claims the insured can get during the coverage period calculated by the insurer (ie. $E[N]$).

## 3. Risk analysis of self-insuring and general decision scheme (Questions 3-5)

*3.1 Additional Assumptions*

- Insurance company will be responsible for and willing to pay the ensured award amount
- The annual profit of the organizing committee will be able to cover the annual insurance cost
- The organizing committee would prioritise a stable financial status over profit

*3.2 Method analysis*

Based on the results of the models of question 1, we have developed a model that allows the organizing committee to flexibly calculate whether they should purchase insurance.

*3.3 Model for calculating risk factor*

This model allows the organising committee to determine how risky it is to self-insure. Evidently, the organizing company would want to save the added cost (that is, to not buy insurance as much as possible), but would first need to ensure that their financial situation is healthy (i.e. that they will not suffer from a sudden monetary loss as somebody breaks the record). Thus, whether or not insurance should be purchased depends on the organizing committee's financial situation, the amount of bonus and the probability of record-breaking during the coverage period. The problem, then, is to find out the probability of record-breaking during the coverage period, and the number of record breaks the committee's financial situation allows.

The probability of record-breaking of the $xth$ year following year $w$, $p(w,x)$, is already computed in question 1. For example, the probability of the record being broken for 1 time in the $y$ years following year $w$ is:

$$P(w,y,1) = \sum_{i=1}^{y} (p(w,i) \prod_{j \neq i} (1 - p(w,j)))$$

so the expected amount of bonus is:

$$C \cdot \sum_{i=1}^{y} (p(w,i) \prod_{j \neq i} (1 - p(w,j)))$$

the probability of the record being broken for 2 times in the $y$ years following year $w$ is:

$$P(w,y,2) = \sum_{i_1=1}^{y} \sum_{i_2>i_1}^{y} (p(w,i_1)p(w,i_2) \prod_{j\neq i_1, j\neq i_2} (1-p(w,j)))$$

and the expected amount of bonus is:

$$2C \cdot \sum_{i_1=1}^{y} \sum_{i_2>i_1}^{y} (p(w,i_1)p(w,i_2) \prod_{j\neq i_1, j\neq i_2} (1-p(w,j)))$$

In general, the probability of breaking the world record for $z$ times in the $y$ years following year $w$ is:

$$P(w,y,z) = \sum_{i_1=1}^{y} \sum_{i_2>i_1}^{y} \sum_{i_3>i_2}^{y} \dots \sum_{i_z>i_{z-1}}^{y} (p(w,i_1)p(w,i_2)\dots p(w,i_z) \prod_{j\neq i_1, j\neq i_2,\dots, j\neq i_z} (1-p(w,j)))$$

Then, $P(w,y,z)$ is called the **risk factor** of having $z$ record-breaks in the $y$ years following year $w$.

*3.4 Decision on risk-taking*

As mentioned, the organizing committee's financial situation and the risk factor are the most important criteria for determining whether the committee should purchase the insurance.

Suppose the organizing committee can self-insure for a maximum of $k$ times at most without facing bankruptcy. That is, if the committee has a revenue of *$A*, which can be written as a function of year number and past profits, then $k+1 > A \div C > k$.

Say the income of the organizing committee in year $w$ is $T_w$. According to our research and analysis of data collected, if a record is broken at a certain competition, top athletes will be incentivized to go to that competition in hopes of breaking another record. With more top athletes and more advertisement, the income will increase. It is therefore reasonable to assume that for every time that a record is broken, the income will increase by *N%*. On the other hand, if no records are broken during the competition, its credibility and consequently its income will decrease. Suppose that whenever a record is not broken, the income will decrease by *M%* per year. Using conditional probability, we have the $T(w,x)$, the expected value of the income the *xth* year from year $w$:

$$T(w,x) = T_w \prod_{i=1}^{x} (p(w,i)(1+N\%) + (1-p(w,i))(1-M\%))$$

The total expected income in $y$ years from year $w$ is therefore:

$$T_{total}(w,y) = \sum_{i=1}^{y} T(w,i)$$

We first consider a simpler case, where the premium is assumed to be constant throughout the coverage period once the insurance is bought. This implies that we only need to consider whether the insurance should be bought this year, because if there are any future changes the organizing committee can easily start buying insurance. Based on the expected income of the organizing committee and its current assets, we can conclude a few conditions in which the organizing committee should purchase the insurance.

First, if the organizing committee does not have enough money to pay the bonus (even after earning the yearly income), then it is a must for them to purchase the insurance to prevent bankruptcy. In other words, the organizing committee will have to purchase the insurance if:

$$A + T(0, 1) \leq C$$

On the other hand, since insurance companies may choose to invest in other places for profit rather than directly earn from an added cost, the expected cost the organizing committee has to pay may be higher than that of the insurance cost, and in that case the organizing committee should also purchase the insurance to save money. In other words, the organizing committee should purchase the insurance if:

$$\sum_{i=0}^{y} i \cdot C \cdot P(0, y, i) = C \cdot (\sum_{i=1}^{y} p(0, i)) \geq y\Pi$$

Where $\Pi$ is the premium per year and $C$ is the bonus. In addition, $y$ is any value from 1 to the longest number of years that the records were held unbroken in the past. If, for some $y$, there exists a number $z$ that satisfies the above inequality, then purchasing the plan *(y,z)* would be a wise choice. If for some $y$, there exists more than one $z$ that satisfies the above inequality, then the manager should note that the larger the $z$, the safer the plan is, but also the more expensive; which $z$ to insure for is on the discretion of the committee as to whether a safer or cheaper plan is preferred.

Of course, if $y$ and $z$ are decided by the insurance company and the organizing committee only needs to decide whether or not they will buy the offered plan, the organizing committee should calculate the respective values for two sides of the inequality. If the inequality holds, then the organizing committee should take the offer; else, they should not.

However, in real life, it is impossible for the premium to stay the same each time the organizing committee decides to buy insurance. For instance, when the insurance company knows that the organizing committee is in tight financial situations, it is likely to significantly increase the premium, knowing that the organizing committee has no choice but to buy it. This complicates the problem greatly, because while the above inequalities still apply, they are not the only considerations; say, if after a few years the organizing committee is forced to buy the insurance no matter the price, they may end up paying a lot more than the original premium they would have needed to pay if they had only bought the insurance a few years earlier, when their financial situation allowed self-insurance. To build a more realistic model for this situation, we have the following assumptions:

- When circumstances dictate that the organizing committee buy the insurance, the insurance company will set the price to the highest that the organizing committee can afford in order to maximize profit according to the law of demand in economics. That is, if the organizing committee starts to buy the insurance in year x, the price per year will be $T_x$.
- The yearly income of the organizing committee is not comparable to the bonus; else, we can forgo this case and refer to the previous inequality for decision scheming, as the bonus is of little consequence to the financial situation of the organizing committee and it only needs to consider whether it can save more money.

Let $Y_0$ be the largest number such that $T_{total}(0, Y_0) + A < (k+1) \cdot C$.

Since the income of the organizing committee is not comparable to $C$, it can be safely assumed that $Y_0$ is a number large enough to consider cases for the foreseeable future. (If not so, then $A$ must be very close to $(k+1) \cdot C$, implying that the organizing committee can refer to the previously mentioned decision scheme until A reaches $(k+1) \cdot C$, then consider this approach again.)

The reasoning of this approach is simple: we consider the probability of the record being broken k times before the year reaches $Y_0$, and the total expected bonus caused by this probability. If this total expected bonus is higher than the cost of buying the insurance now, the insurance should be purchased; otherwise, the insurance should not be purchased.

Note that we do not need to consider alternatives such as purchasing the insurance after a number of record-breaks: should the decision now be self-insure, the option of purchasing in the future is always open and can be considered later when the insurance rates and new competition results are confirmed; should the decision now be purchase insurance, due to changing premiums and the case-dependency of repeated insurance purchasing, the safer option will be to purchase now.

The problem then changes to calculating the expected cost of purchasing the insurance after k new records. Since the $k^{th}$ new record can be set at any time from the $k^{th}$ to the $Y_0{}^{th}$ year from now, we will sum up all the possible results to achieve the expected value of the total asset the organizing committee will have, which is:

$$W(k,\ Y_0) = A - C \cdot \sum_{i=0}^{k-1} i \cdot P(0, Y_0, i) - k \cdot C \cdot \sum_{i=k}^{Y_0} P(0, Y_0, i)$$

$$+ \sum_{i_1=1}^{Y_0} \sum_{i_2>i_1}^{Y_0} \sum_{i_3>i_2}^{Y_0} \cdots \sum_{i_k>i_{k-1}}^{Y_0} ((p(0,i_1)p(0,i_2)\ldots p(0,i_k) \prod_{j\neq i_1, j\neq i_2,\ldots, j\neq i_k} (1-p(0,j))) \cdot (\sum_{l=1}^{i_k} T(0,1)\,(1+N\%)^{\alpha}(1-M\%)^{l-\alpha}$$
$$+ T_{total}(i_k, Y_0 - i_k) - (Y_0 - i_k)T_{i_k}))$$

Where according to the above definitions, $T_{i_k} = T(0,1)(1+N\%)^k(1-M\%)^{i-k}$ and $\alpha$ = the largest $u$ such that $i_u \leq l$. If $W(k, Y_0) > A + T_{total}(0,\ Y_0) - Y_0\Pi$ where $\Pi$ is the premium cost, then the insurance should not be purchased; else, the insurance should be purchased.

*3.5 Multiple-event decision*

In the case of multiple events, the above method can be used for multiple times to decide whether insurance should be purchased for each event. This is because the probability of an athlete breaking the record in a event does not affect that of another event and therefore all events can be considered independently.

Since it is assumed that the profit of the organization committee each year is enough to cover insurance for all events, and that the income of each event is insignificant compared to that of the award money, there should not be a limit on the number of events for which insurance is purchased, so the above method can be used to determine whether purchasing an insurance is ideal for each of the events. However, should the organizing committee want to purchase insurance for only a decided number of events (say $v$ events), the following methods can be used to rank the risk of self-insuring for each event, and the first $v$ events with the highest risk should be purchased. We define the following variables:

| Variable | Explanation |
|---|---|
| $\Pi_1,\ \Pi_2, \ldots,\ \Pi_n$ | The premium per year for the $n$ events, respectively |
| $C_1,\ C_2,\ \ldots,\ C_n$ | The bonus for the $n$ events, respectively |
| $T_1(w,x),\ T_2(w,x), \ldots, T_n(w,x)$ | The expected profit by the $n$ events in the $x^{th}$ year following year $w$, respectively |

| | |
|---|---|
| $T^1_{total}(w,y), T^2_{total}(w,y), \ ..., \ T^n_{total}(w,y)$ | The expected total profit by the $n$ events in the $y$ year following year $w$, respectively |
| $p^1(w,x), \ p^2(w,x), ..., p^n(w,x)$ | The probability of the record being broken in the $x^{th}$ y after year $w$ for the $n$ events, respectively |
| $P^1(w,y,z), \ P^2(w,y,z), ..., \ P^n(w,y,z)$ | The probability of $z$ record breaks in the $y$ years following year $w$ for the $n$ events, respectively |

Our strategy is based on the assumption that the most important aim of the organizing committee is to maintain a stable financial status for future events. Naturally, if they do not have enough assets to pay for the bonuses of all the events, they will want to purchase insurance for some events and save their assets for covering the bonuses of the remaining events to avoid bankruptcy. Should they be unable to purchase all the required insurance, they will want to purchase insurance for the events that have the highest expected cost (i.e. the highest risk) to lower the chance of bankruptcy.

After short-term considerations for avoiding bankruptcy, the organizing committee will look to saving money in the long term. In this case, the model presented in section 3.5 now comes into view. For events that have an expected higher profit for buying insurance now than later, the organizing committee should purchase insurance. Naturally, the priority goes to events that have the highest difference between the expected profit from purchasing now than that of later.

Lastly, after all the above is considered, if the organizing committee is still willing to purchase insurance for more events, the decisions should be made by considering whether purchasing an insurance for a particular event will give the organizing committee a profit over a long period of time. Although this choice will give the organizing committee an expected profit, it does not contribute to maintaining the financial stability of the organizing committee as the small amounts of profit here cannot provide protection from bankruptcy compared to purchasing insurance for the events. Thus this consideration should be put last.

In light of the above, without loss of generality, suppose $C_1 \geq C_2 \geq ... \geq C_n$. The organizing committee should decide on which events to purchase insurance for using the following steps:

1. If $A + T(0,1) > \sum_{i=1}^{n} C_i$, the insurance company has enough assets to pay for the total bonus for all events this year and should look at the long term, and should refer to step 2. If that is not the case, we will find the smallest $u$ such that $A + T(0,1) > \sum_{i=u}^{n} C_i + \sum_{i=1}^{u} \Pi_i$. If $u - 1 > v$, then the organizing committee should calculate the average cost $C_i \cdot p^i(0,1)$ for each event, and purchase the first $v$ events with the largest average cost - although in this case, the organizing committee is highly recommended to purchase at least $u - 1$ insurances for all the aforementioned events. If $u - 1 \leq v$, the organizing committee should purchase insurance for all $u - 1$ events, then go to step 2.

2. We now assume that there are $v_1$ events left to be purchased. We should calculate $W_i(k_i, Y_{0i})$ for each of the events, and purchase the first $v_1$ events that have the smallest $W_i(k_i, Y_{0i}) - A - T^i_{total}(0, \ Y_{0i}) + Y_{0i}\Pi_i$, with the value being negative; should there be any quotas for purchasing left, step 3 can be referred to.
3. For each of the events that no insurance is purchased, the $C_i \cdot (\sum_{i=1}^{y_i} p^i(0,i))$ should be calculated for $y_i$ = *the longest period of years no record is broken for event* $i$. If $C_i \cdot (\sum_{i=1}^{y_i} p^i(0,i)) \geq y_i \Pi_i$, then the

insurance should be purchased for those events as well; in particular, the events with the largest $C_i \cdot (\sum_{i=1}^{y_i} p^i(0,i)) - y_i \Pi_i$ should be purchased as long as the difference is positive. If any quotas are left after step 3, no more insurance should be purchased for any of the events left.

*3.6 General decision plan*

The general decision plan under any circumstance is very similar to the one in the previous section. Although there may be variances, the general approach will also be to calculate the expected bonus, profit and premium cost for each of the events, then analyze first the short term, then the long term impacts according to the equations developed above, and purchase insurance for the events that will give the organizing committee a better protection or profit.

Note: One major question that may be posed is the usability of the above model, due to the complexity of the equations in the above decision plan. Indeed, for larger numbers human-hand calculations cannot come up with a solution within a reasonable time - but computers can. Due to the fact that for sports events, predictions will usually not go out of the range of the foreseeable future (say, 10 or 20 years), the plan above without any strategies to further simplify it has an approximately $O(n \cdot 2^y)$ complexity, where $n$ is the total number of events and $y$ is the number of years for consideration. Since in general $y \leq 20$ and $n$ should be within a reasonably small range (say, $n \leq 1000$), the calculation by computer can be finished in a very reasonable time (within seconds) by a program in accordance to the method utilizing recursion. Therefore, the above method can be used by any organizing committee to obtain a reasonable prediction as to whether insurance should be purchased.

## Strengths and Limitations of the Model

Strengths

The strength of our model is that it is easy to implement after inputting the data for the first time. Through inputting the correct data, we can obtain the result desired that can help the organizing committee and the insurance company to make a decision that could benefit them. With the aid of programming and computer, we are confident that the the model we developed can help the insurer and the insured to make their decision effectively within seconds.

Apart from this, by considering a large variety of scenarios and variables, for example the three different scenarios including the worst case for the organizing committee in section 3, as well as some of the more practical components including administrative costs, agent commission fees, and investment to ensure profit gaining, our model can be made more realistic and we are confident that it can be used in real life.

Limitations

One of the limitations of this model is that there are a large number of variables to be considered in each equation during calculation. The large number of variables thus require a huge amount of information for calculation. This may increase the cost of data collections for both parties, causing variations in the cost of the premium as calculated by our model.

Also, many assumptions are made when we develop the model. The large number of assumptions made for the ease of our calculation may potentially have significant impacts on the accuracy of the output of our model. For example, it is assumed that all the top runners in a sport event will go to the same competition aiming to break the world record and that there would be no accidents that may hinder their performance or cause them to pull out, which may not be true in real life. Thus the result obtained based on each model may not be exactly accurate.

## Conclusion

After preliminary literature review, data collection and research, we have developed three different models in this paper to solve the problems posed.

Firstly, to find the average cost of the bonus (i.e. the expected cost of the premium), we use the concept of regression to predict the frequency of record-breaking in the future based on historical data. Secondly, to determine the price premium, we include a number of components that are included in the premium while considering the amount of reserve and profit margin for the insurance company. Finally, to determine if the organizing committee should purchase insurance in different circumstances, we make use of the concept of probability and inequalities to create a generalised function based on their expected income and the likelihood for record-breaking.

Since sections 2 and 3 of this paper largely depend on the model in section 1, we generate a test case using real-life data for 15K races we collected to assess model 1. The result is reasonable and the model can be used in real life. In sections 2 and 3, we consider multiple scenarios in real life situations. With this, plus our reasonable approach for calculating the probability of record-breaking, we are confident that our model will be able to help the insurance company determine the premium and the organizing committee decide whether to purchase insurance or not for their own benefit.

## References


- Olivieri, A., & Pitacco, E. (2011). *Introduction to insurance mathematics: Technical and financial features of risk transfers*. Berlin: Springer Verlag.
- Harrington, S. E., & Niehaus, G. (2004). *Risk management and insurance*. Boston, MA: McGraw-Hill.
- Karl Borch (1974). Mathematical Models in Insurance . ASTIN Bulletin, 7, pp 192-202 doi:10.1017/S0515036100006036
- Zazanis, M. (2005, July). Stochastic models in risk theory. Retrieved March 25, 2016, from http://www.stat-athens.aueb.gr/~mzazanis/courses/MAP/notes-05.pdf
- Taghavifard, M. T., Damghani, K. K., & Moghaddam, R. T. (2009). Decision Making under Uncertain and Risky Situations. Retrieved March 25, 2016, from http://www.soa.org/library/monographs/other-monographs/2009/april/mono-2009-m-as09-1-damghani.pdf