2016 IM2C PROBLEM

21 March 2016

Abstract

For many people, the thrill of watching athletics is in the unexpected nature (or not) of witnessing a world record being broken. While world records can create a lot of excitement for the viewer, it also poses a financial problem to the organising committee in the form of some winning bonus. A decision may be made by the committee to purchase insurance over some period of time to finance the bonus should the occasion arise.

In this paper, we developed a model to evaluate the probability of a world record being broken in the near future – and hence determine the viability of signing an insurance contract. The problem was undertaken primarily with respect to the committee's perspective but the insurance company's perspective was also considered in regards to maintaining a reasonable profit margin.

By investigating the distribution of the given data, it was found that a logarithmic function best modelled the decreasing trend in the times. With the advancement of sports technology and training techniques, the trend comes as no surprise. For each edition of the race, the deviations of the actual time from the trend line were taken and combined to determine a general standard deviation for our model. From here when considering how far data was from the mean, the number of these general standard deviations was normally distributed. This claim was justified by investigating the sample size and symmetry of the data.

In considering the Gaussian distribution in the data, the problem was able to be approached by evaluating the expected probability that a world record would be broken by some edition of the race. This model could be used to provide a 'cumulative probability' for the record being broken once after some number of race editions – and thus insight into a favourable term for which the committee should purchase insurance.

The model was extended to the insurance company's perspective. The associated sensitivity in the model depended on the point from which the average cost was calculated. Therefore, it would be reasonable to expect that the insurance company would choose the 'worst-case' average cost and a profit margin could be assumed by 'inflating' the expected number of world record breaks over a given term.

The argument for whether to purchase insurance or to self-insure was interpreted in this paper as the relationship between the average cost for the committee and the quote offered by the insurance company. Ultimately, this will provide a standpoint on what must be done by the committee to save the most money.

Introduction

While breaking world records comes as a spontaneous highlight extravaganza to viewers, the underlying complexity involved with world records is the financial decisions which must be made by the sponsor to purchase insurance or to self-insure. The factors which must be considered by the committee in making this decision are:

- The likelihood of a world record being broken in the near future
- The average cost associated with funding the winning bonus
- The short-term and long-term prospects of the event for the committee

The aim of the model is to determine the conditions necessary such that it is beneficial for insurance to be purchased by the committee. In extending this, the model also provides a guideline as to what must be done so that insurance company maintains a suitable profit margin.

Assumptions

1) The committee's decision to purchase insurance or self-insure is based purely on financial reasons.

2) The insurance company does not follow our model to calculate the average cost – but rather a simplification of the situation - that is:

$$p(breaking\ WR) = \frac{n(WR\ broken)}{n(races)}$$

Take the Zevenheuvelenloop men's race for example. The insurer would calculate average cost to be $\$25000 \div \frac{31}{2}$. In fact because of the 'sensitivity' of this value the insurer would choose a slightly higher value. For example if the world record was broken the next year then the frequency of world record breaks would be $\frac{32}{3}$ which is significantly higher than $\frac{31}{2}$.
Hence take the insurer's average cost to be $\$25000 \div \frac{32}{3} = \$2344$.
Therefore, the insurer charges $\$2344 \times 1.2 = \$2813$ for the event.

3) There are no other insurance options available to the committee. The justification is that a competitive market would mean the price set by the insurance companies will decrease in response to the nature of the market.

4) The world record is broken only at our athletics meet. The reasoning for this is that should the world record be broken elsewhere, we are unable to predict what time this world record would be and hence we would be unable to predict the chance of this world record being broken given that the required time is unknown.

5) There are no accidents, misconduct of the event, or other completely unforeseeable events which would directly influence the running times – i.e. there are no gross outliers in the given data, and future times where this does occur would not be legitimised.

6) Extrapolating the data is allowed. This is because our trend does not come to any gross conclusions (for example as you will see, the trend predicts the time for the 15k race to be 0 seconds after 80 billion years) within in a realistic future number of years (i.e. <30).

7) We're only considering the insurance as carrying until the world record is broken. This is because once a record is broken (similar to assumption 4), we are unable to predict the probability of another world record time as we don't know the time required to beat. This assumption is also because like a car insurance premium after a crash, the insurer would change the premium after a record break.

Question 2:
Three of the most important criteria in deciding how much to add on to the average cost (AC) are profit (P), tax (T) and operating costs (OC). The total cost (TC) is $T + OC + AC$. The average profit margin for insurance companies is around $3 - 8\%$ ['*What is the usual profit margin for a company in the insurance sector?'* Investopedia, 2016, *Accessed 21 Mar 2016*]. Assume that this insurance company aims for a 5% profit margin. In Australia, businesses are taxed 30% of their profit. Hence we assumed that this tax also applies in this case. Many factors affect the operating costs, such as employee wages, computers, office amenities, rental costs etc. Hence it is best to keep operating costs as a variable. Profit margin is calculated using the revenue (R), i.e how much the insurance company charges, and total cost:

$$P = R - TC$$

We also know:

$$TC = T + OC + AC$$
$$T = 0.3P$$
$$P = 0.05R$$

Hence, by substitution:

$$0.05R = R - 0.3 \times 0.05R - OC - AC$$

Thus, by rearranging:

$$R = \frac{OC + AC}{0.935}$$

This gives a required increase of $\left(\frac{1}{0.935} - 1\right) \times 100\% = 6.95\%$ increase on the average cost calculated by the insurer, when compensating on the tax and profit margin. Once operating costs have been taken into account (assuming an increase of 20%), this percentage increase on the average cost is now 26.95%.

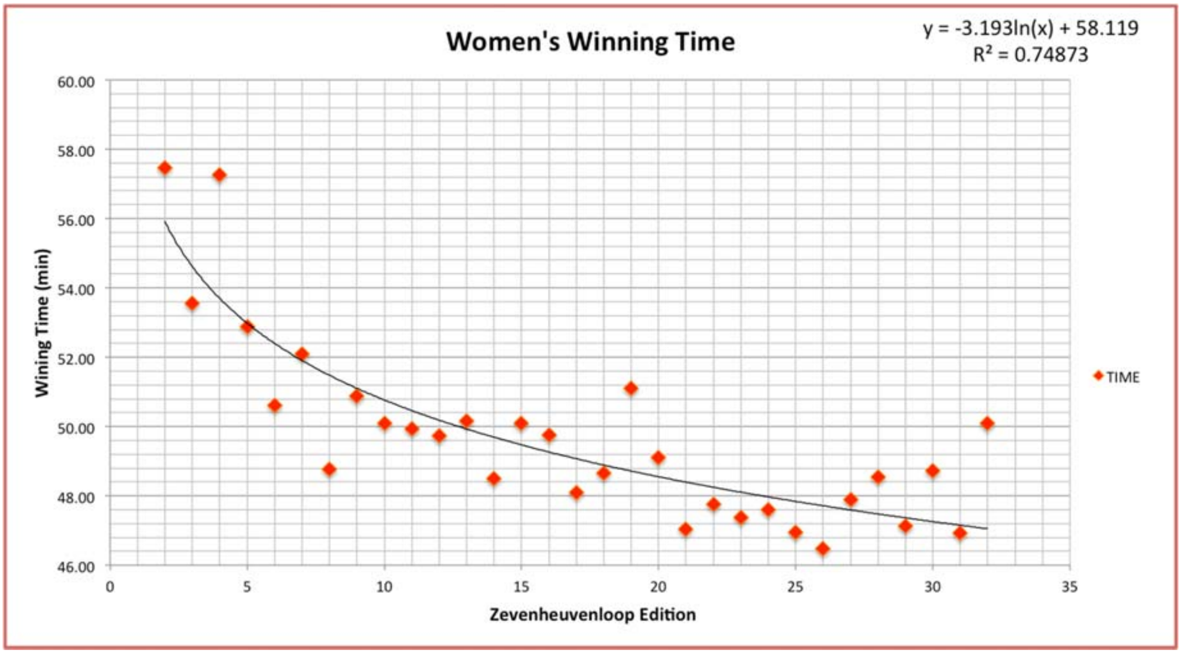
How can we accurately predict the probability of breaking a world record?

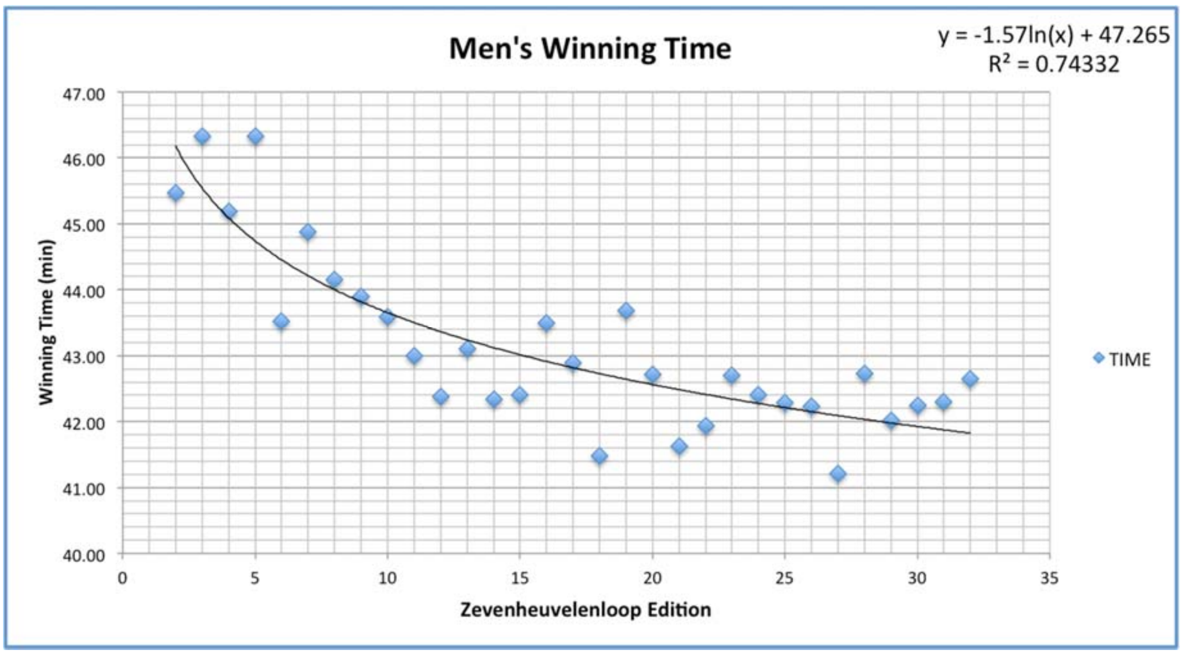
Case Study: Zevenheuvelenloop. Using this case study as an example, all other 38 men's and women's track events can be modelled in a similar way…

We are given the following information about winning times for the 15k race for the previous 31 years:

EDITION	YEAR	WINNING TIME (min) **(WR)**	
		WOMENS	**MENS**
2	1985	57.47	45.47
3	1986	53.55	46.33
4	1987	57.27	45.18
5	1988	52.88	46.33
6	1989	50.60	43.52
7	1990	52.10	44.88
8	1991	48.77	44.15
9	1992	50.88	43.90
10	1993	50.10	43.58
11	1994	49.93	43.00
12	1995	49.73	42.38
13	1996	50.15	43.10
14	1997	48.50	42.33
15	1998	50.10	42.40
16	1999	49.75	43.50
17	2000	48.10	42.88
18	2001	48.67	**41.48**
19	2002	51.10	43.68
20	2003	49.10	42.72
21	2004	47.03	41.63
22	2005	47.77	41.93
23	2006	47.37	42.70
24	2007	47.60	42.40
25	2008	46.95	42.28
26	2009	**46.48**	42.23
27	2010	47.88	**41.22**
28	2011	48.55	42.73
29	2012	47.13	42.02
30	2013	48.72	42.25
31	2014	46.93	42.30
32	2015	50.08	42.65

If we graph the trend of winning times (over page), it becomes clearer that the required winning time has been slowly decreasing over the years for both the men's and the women's race:





Both graphs produce a clear decreasing trend in winning times. The correlation coefficient (R) is at its highest when we presume this trend to be logarithmic. This ever present increasing performance of human physically is due to the field of sport science: better training methods, athletic diets, performance enhancing substances and better clothing such as 'bouncier' shoes. This would suggest a decreasing function and one whose magnitude of rate of change is decreasing. It would be decreasing because we approach physical limits.

For example, it would be unreasonable to run faster than 7m/s (i.e. a time of 35 minutes). Our model would show that this occurs when $35 = -3.193\ln(\mathrm{x}) + 58.119$ for the women

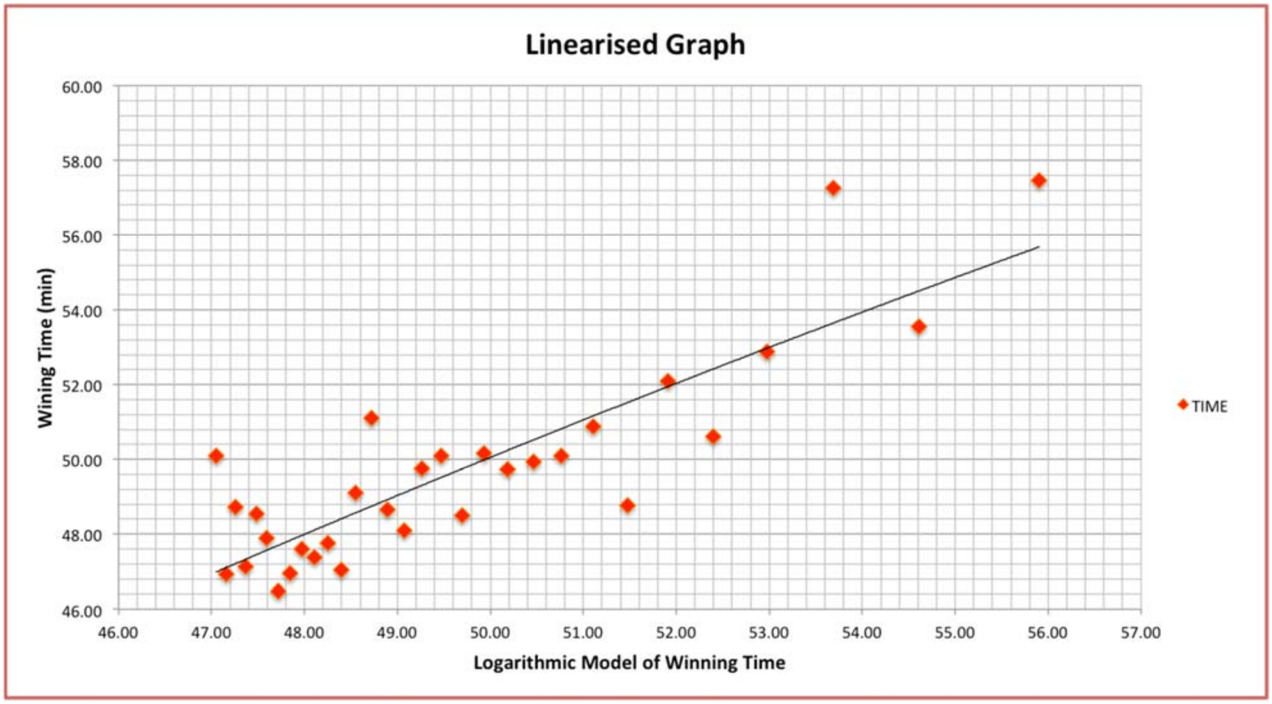
$\therefore \ln(x) = 23.119/3.193$

$\therefore x = e^{23.119/3.193} = 1400$ years.

Similarly for the men, $x = e^{(47.265-35)/1.57} = 3100$ years.

Clearly, from the magnitude of such numbers, our model is unlikely to become grossly unreliable any time soon.

Thus, can this model be extrapolated in order to predict when the next world record will be broken? If we firstly look at our linearised graph based on the logarithmic function of the previous 'women's graph', the distribution appears to be evenly spread:



Let $T(x)$ define the recorded winning time and $F(x)$ define the modelled average where x is the race edition from 2 to 32. We can define the standard deviation of our model as:

$$\sigma = \sqrt{\frac{1}{31} \times \sum_{2}^{32} (T(x) - F(x))^2}$$

$$\therefore \sigma = 1.32905 \ldots$$

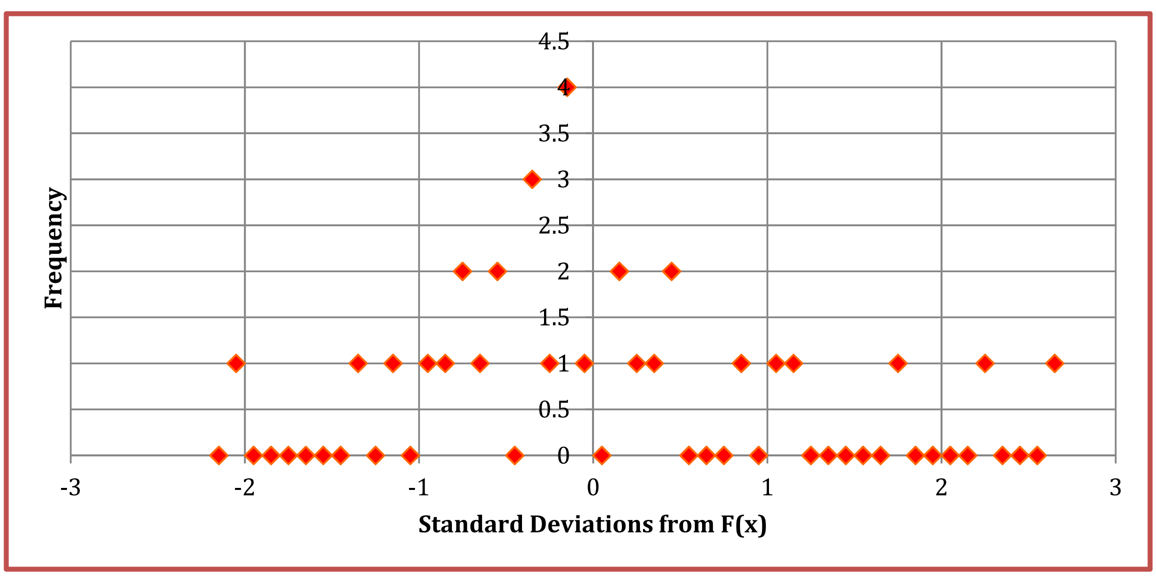
Using this standard deviation, we can calculate (for the women's race) the number of standard deviations each recorded winning time is from the modelled average. This is done by using the formula:

$$S(x) = \frac{T(x) - F(x)}{\sigma}$$

Where $S(x)$ denotes the number of standard deviations from $F(x)$.

Year	Winning Time (min)	S(x)
1985	57.47	1.174
1986	53.55	-0.798
1987	57.27	2.689
1988	52.88	-0.073
1989	50.60	-1.353
1990	52.10	0.146
1991	48.77	-2.041
1992	50.88	-0.165
1993	50.10	-0.502
1994	49.93	-0.398
1995	49.73	-0.340
1996	50.15	0.166
1997	48.50	-0.897
1998	50.10	0.472
1999	49.75	0.364
2000	48.10	-0.732
2001	48.67	-0.168
2002	51.10	1.793
2003	49.10	0.411
2004	47.03	-1.027
2005	47.77	-0.363
2006	47.37	-0.557
2007	47.60	-0.280
2008	46.95	-0.671
2009	46.48	-0.927
2010	47.88	0.217
2011	48.55	0.806
2012	47.13	-0.176
2013	48.72	1.097
2014	46.93	-0.166
2015	50.08	2.280

By grouping the calculated values of S(x) into groups of 0.1 width and determining the number of results in each group, a frequency histogram can be graphed.



It can be assumed that the graph follows the normal distribution model. The points somewhat follow a bell curve. It does not perfectly do so, but this can be attributed to the small amount of data (only 31 recorded times). Not only this, but the graph also approximately follows the 68-95-99.7 rule. 31 of the 31 recorded times were within 3 standard deviations, 29 were within 2 standard deviations, and 24 were within 1 standard deviation.

$$31 \div 31 = 100\% \approx 99.7\%$$
$$29 \div 31 = 93.5\% \approx 95\%$$
$$24 \div 31 = 77.4\% \approx 68\%$$

Because all data points lie well within 3 standard deviations of the predicted value, we can say that there are no outliers in the data set.

Note that the current women's world record is 46.48 minutes (2009). Hence, a time of 46.47 minutes or less is required to break the record. In regards to next year's race, our model calculates $F(33)$ to be $-3.193\ln(33) + 58.119 = 46.95$ min. Thus the required deviation from our model is $46.48 - 46.96 = -0.48$ min. Given the previously calculated standard deviation, this difference is equivalent to $-0.48/1.32905 = -0.36$ standard deviations from the mean (Z). Therefore, the probability of the world record being broken next year can simply be defined as:

$$P(Z < -0.36)\ given\ Z \sim N(0,1)$$

$$= 0.3586$$

Using this method, we can figure out the cumulative probability of a world recorded being broken after some number of years.

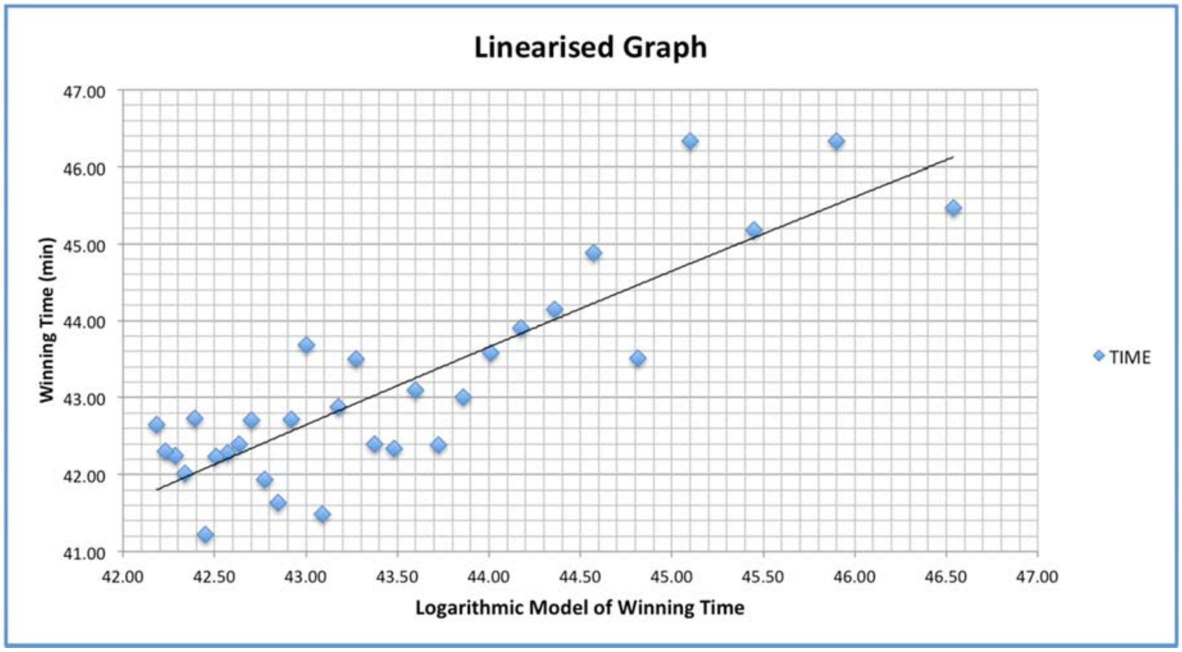
The cumulative probability is the probability the record is broken exactly once over the next *n* Editions of the race. We can define this as:

$$P(cumulative) = \sum_{k=1}^{n}\left[P(T(k+32) < Z_k) \times \prod_{m=1}^{k-1} P(T(m+32) > Z_m)\right]$$

Future Edition (x)	$F(x)$	**Absolute deviation required (from 46.47)**	**(Z) required**	**$P(T(x) < Z))$ i.e. World Record Broken**	**$P(T(x)) > Z))$ i.e. World Record Not Broken**	$P(cumulative)$
33	46.95	-0.48	-0.36	0.36	0.64	0.3586
34	46.86	-0.39	-0.29	0.39	0.61	0.6060
35	46.77	-0.29	-0.22	0.41	0.59	0.7686
36	46.68	-0.20	-0.15	0.44	0.56	0.8702
37	46.59	-0.12	-0.09	0.47	0.53	0.9306
38	46.50	-0.03	-0.02	0.49	0.51	0.9647
39	46.42	0.05	0.04	0.52	0.48	0.9829
40	46.34	0.13	0.10	0.54	0.46	0.9921
41	46.26	0.21	0.16	0.56	0.44	0.9966
42	46.18	0.29	0.22	0.59	0.41	0.9986
43	46.11	0.36	0.27	0.61	0.39	0.9994
44	46.04	0.44	0.33	0.63	0.37	0.9998

Hence, given that $\sigma = 1.32905$, $Z \sim N(0,1)$ and we require $T(x)$ to be smaller than 46.48, we can produce the following table:

Similarly, we can apply the same analysis to the men's race (and ultimately every event that is to be potentially insured…)



Let $T(x)$ now define the recorded men's winning time and $F(x)$ define the men's modelled average where x is the race edition from 2 to 32. We can define the standard deviation of our model as:

$$\sigma = \sqrt{\frac{1}{31} \times \sum_{2}^{32} (T(x) - F(x))^2}$$

$$\therefore \sigma = 1.75402 \ldots$$

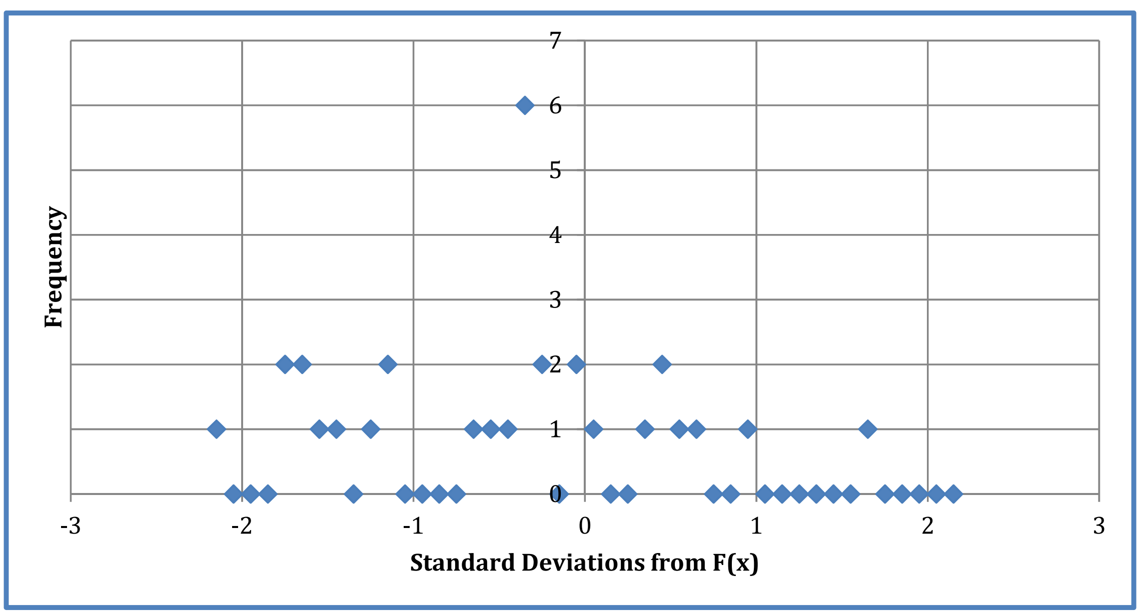
Using this standard deviation, we can calculate, for the men's race, the number of standard deviations each recorded winning time is from the modelled average. This is done by using the formula:

$$S(x) = \frac{T(x) - F(x)}{\sigma}$$

Where S(x) denotes the number of standard deviations from F(x).

Year	Winning Time (min)	S(x)
1985	45.47	-1.419
1986	46.33	0.575
1987	45.18	-0.352
1988	46.33	1.638
1989	43.52	-1.718
1990	44.88	0.416
1991	44.15	-0.279
1992	43.90	-0.365
1993	43.58	-0.566
1994	43.00	-1.141
1995	42.38	-1.778
1996	43.10	-0.661
1997	42.33	-1.523
1998	42.40	-1.291
1999	43.50	0.302
2000	42.88	-0.389
2001	41.48	-2.127
2002	43.68	0.903
2003	42.72	-0.272
2004	41.63	-1.607
2005	41.93	-1.112
2006	42.70	-0.003
2007	42.40	-0.312
2008	42.28	-0.382
2009	42.23	-0.367
2010	41.22	-1.636
2011	42.73	0.451
2012	42.02	-0.427
2013	42.25	-0.047
2014	42.30	0.088
2015	42.65	0.6183

By grouping the calculated values of S(x) into groups of 0.1 width and determining the number of results in each group, a frequency histogram can be graphed.



It can be assumed that the graph follows the normal distribution model. The points somewhat follow a bell curve. It does not perfectly do so, but this can be attributed to the small amount of data (only 31 recorded times). Not only this, but the graph also approximately follows the 68-95-99.7 rule. 31 of the 31 recorded times were within 3 standard deviations, 30 were within 2 standard deviations, and 20 were within 1 standard deviation.

$$31 \div 31 = 100\% \approx 99.7\%$$
$$30 \div 31 = 96.8\% \approx 95\%$$
$$20 \div 31 = 64.5\% \approx 68\%$$

Because all data points lie well within 3 standard deviations of the predicted value, we can say that there are no outliers in the data set.

Note that the current men's world record is 41.22 minutes (2010). Hence, a time of 46.21 minutes or less is required to break the record. In regards to next year's race, our model calculates $F(33)$ to be $-1.57\ln(33) + 47.625 = 42.14$ min. Thus the required deviation from our model is $42.14 - 41.22 = -0.93$ min. Given the above standard deviation, this difference is equivalent to $-0.93/1.75402 = -1.23$ standard deviations from the mean (Z). Therefore, the probability of the world record being broken next year can simply be defined as:

$$P(Z < -1.23)\ given\ Z \sim N(0,1)$$

$$= 0.1090$$

Using this method, we can figure out the cumulative probability of a world recorded being broken after some number of years.
The cumulative probability is the probability the record is broken exactly once over the next *n* Editions of the race. We can define this as:

$$P(cumulative) = \sum_{k=1}^{n}\left[P(T(k+32) < Z_k) \times \prod_{m=1}^{k-1} P(T(m+32) > Z_m)\right]$$

Hence, given that $\sigma = 1.75402$, $Z \sim N(0,1)$ and we require $T(x)$ to be smaller than 46.21, we can produce the following table:

Future Edition (x)	$F(x)$	Absolute deviation required (from 41.21)	(Z) required	$P(T(x) < Z))$ i.e. World Record Broken	$P(T(x)) > Z))$ i.e. World Record Not Broken	$P(cumulative)$
33	42.14	-0.93	-1.23	0.109	0.891	0.1090
34	42.09	-0.88	-1.17	0.121	0.879	0.2169
35	42.04	-0.84	-1.11	0.134	0.866	0.3215
36	42.00	-0.79	-1.05	0.147	0.853	0.4211
37	41.96	-0.75	-0.99	0.160	0.840	0.5138
38	41.91	-0.71	-0.94	0.174	0.826	0.5985
39	41.87	-0.67	-0.88	0.188	0.812	0.6741
40	41.83	-0.63	-0.83	0.203	0.797	0.7402
41	41.79	-0.59	-0.78	0.218	0.782	0.7968
42	41.76	-0.55	-0.73	0.233	0.767	0.8441
43	41.72	-0.51	-0.68	0.248	0.752	0.8828
44	41.68	-0.48	-0.63	0.263	0.737	0.9377

How does our model and predicted costs compare with the insurers?

The insurer's quote not only covers the average cost they calculated, but it must also cover the insurer's operating cost and the insurer's desired profit margin. The insurer cannot inflate the cost so much as to detract companies but they must also not keep the cost so low as to cut into their profit margin.

Previously, we estimated for the insurer to maintain a 5% profit margin after tax and operating costs, the company would have to add 26.95% to the calculated average cost.

Case 1: An individual event

Takes the men's race for example. The insurer would calculate average cost to be $\$25000 \div \frac{31}{2}$. In fact because of the 'sensitivity' of this value the insurer would probably choose a slightly higher value. For example if the world record was broken the next year then the frequency of world record breaks would be $\frac{32}{3}$ which is significantly higher than $\frac{31}{2}$.
Hence, given the worst case, take the insurer's average cost to be $\$25000 \div \frac{32}{3} = \$2344$
Therefore the insurer charges $\$2344 \times 1.2695 = \$2976$ for the event.
From our model we expect the record to be broken every 1/0.11 repetitions.
Therefore our average cost would be $25000 \times 0.11 = 2750 < 2976$.
Unless the insurer was very generous with their addition of extra money, we would seem to save money by self-insuring.
However the situation is not this simple because our average cost function is always increasing due the decreasing trend line for winning time.
In the second year, we calculate average cost as $\$25000 \times 0.121 = \$3025$ and so over the two years we pay $\$2750 + \$3025 = \$5775 > \$5752 = 2 \times \$2976$ i.e. the organisers would spend less if they took the insurance.
If the organisers are really desperate to save money, they could take the 10% risk that the record won't be broken in the 33rd edition and wait to take the insurance before the 34th edition.
Another point to consider is when we spend more money on insurance than the money we actually receive back upon a record being broken.
$\$25000 \div \$2813 = 8.9$ i.e. after 9 editions, if the record still hasn't been broken then the organising committee is losing money. From the table, the probability of this occurring is 0.7968. If the company doesn't buy the insurance they take a 20.3% risk – probably too high.
In the women's event the insurer's average cost is $\$25000 \div \frac{32}{2} = \$1562.50$ and their charge is $\$1562.50 \times 1.2695 = \$1984 < \$9000 = \$25000 \times 0.36$ as worked out from our model.
Therefore the organizers should definitely buy the insurance for the women's race.

Case 2: Organising many events in a meet.

The first issue we encounter is that we don't know the probabilities of breaking the record in each event. Already from the two events we analysed, we can see that there will be a significant variance in these probabilities – the probability of a record in the women's 15k was thrice the probability of a record in the men's 15k.

Say there is a small risk – 5% for example – a committee takes when they don't buy the insurance for a particular event. When they are only insuring one event, this won't be a problem. But if they are organising many events all of which have a small risk, only a few of the records being broken could result in a financial disaster. We need to find a way for the committee to evaluate the risk for the entire tournament.

As an example, suppose the probability of a record being broken is 0.25 in every event and the organising committee takes the insurance on all 40 events. Assume that the insurer also knows that the probability is 0.25.
Therefore money paid to insurer is (Average Cost + Added Cost) × No. of events which is equal to $\$25000 \times 0.25 \times 1.2 \times 40 = \$300000$
For the insurance to be worthwhile, the insurer needs to pay the committee more than 300k i.e. at least 12 records need to be broken. We can find the probability of that occurrence using the binomial distribution/Bernoulli trials.

$$P(> 11\ events\ broken) = Bin.CDF\ (p = 0.25, n = 40, 12\ to\ 40\ successes) = 0.285$$

Therefore there is only a 28% chance that taking the insurance will be profitable. However, this also means that if the insurance was not taken, then there is 28% chance that the organising committee will be trying to find some extra money than what was anticipated.

In the above example, p = 0.25 was picked arbitrarily as a number roughly between 0.11 and 0.36. A better representation of the 'average probability' is the geometric mean of probabilities for each individual event. The excess 20% added by the insurer for Case 2 was also chosen arbitrarily.
Let g = geometric mean = $\sqrt[n]{\prod_{k=1}^{n} p_k}$ where n = number of events and p_i denotes the probability of there being a world record broken in each of the individual events (as calculated by our regression model).

Are there other factors to consider?

For the committee planning an event in the near future (for example they are asking for insurance two-weeks prior to the event), there are many possible short-term implications that could skew the reliability of our model. One possibility is that the committee has been told that some highly regarded athletes (who may have previously won their event) are unable to participate. Consequently, the probability of a record being broken is reduced as a key competitor is no longer able to break the record at this particular event. Similarly, if the committee knows that the track conditions or the weather conditions are unfavourable, the probability of a record breaking time would further decrease. On the other hand, if the committee were aware of new recent sport science advancements (such as 'bouncier shoes', more aerodynamic swimsuits, improved training methods, stronger performance enhancing drugs or other secret concoctions…) would consequently cause an immediate shift in the winning time required.
In this case, we are required to multiply some constant to the probability calculated such that this new probability is more 'realistic'. For this reason we have defined the following constants:

<u>Unfavourable conditions resulting in slower times:</u>
Weather is favourable: 1
Weather is slightly unfavourable: 0.9
Weather is highly unfavourable: 0.6
Track is favourable: 1
Track is unfavourable: 0.9
Track is heavily sloped: 0.6

<u>Scientific Advancements tested professionally for the first time:</u>

1. Better sportswear: 1.1
2. Improved performance enhancing drugs: 1.2
3. Better training methods: 1.1

These factors would obviously change the geometric mean by the same amount and consequently would change the average cost of our model. Thus, when comparing our model to the insurer's model, the decision as to whether to insure or not may potentially change. Therefore, within the following decision making scheme, it may be necessary to add on these constants in the short term to produce a more reliable insurance plan.

However, if the committee is focused on the long-term perspective of insurance, these factors would cancel out because as previously mentioned, the cause for our logarithmic progression is in fact these scientific and human advancements. Weather conditions and other such factors would also be too hard to predict in the long term.

Full Generalisation (Decision Making Scheme)

In terms of both long-term and short term insurance plans, we can define our 'decision making scheme' as the following generalisation:

1. Calculate the probability of a world record for each of the n events using our model.
2. Take the geometric mean, g, of these n probabilities.
3. (multiply g by the necessary constants if short-term factors are present)
4. Decide some level of acceptable risk, r, which might be about 5% for example
5. Get a quote from the insurer for insuring all events. Call this value I.
6. Calculate $\lceil I/m \rceil$ where m is the prize money per event ($25000 in Zevenheuvelenloop). This represents the number of records that need to be broken.
7. If $r < Bin.CDF(\text{probability} = g, \#\text{trials} = n, \left\lceil \frac{I}{m} \right\rceil$ to n successes) then insure all events.
8. If the above condition was not satisfied then don't insure the lowest probability event.
9. Now if $r < Bin.CDF(\text{probability} = g, \#\text{trials} = n, \left\lceil \frac{I}{m} \right\rceil$ to n successes) is satisfied, then insure all remaining events. Note that I has changed because we are insuring fewer events.

Strengths and weaknesses

The strength in the model is that it is provides a concise and reliable solution to the problem at hand. Due to the high correlation coefficient of the trend line, extrapolation for future instances of breaking the world record was a valid approach.

The main limitation of the model is that previous data must be known in order to ultimately determine the viability of self-insurance. This, however, was a fair assumption as it should be expected that the organising committee had prior records to the events if they were concerned with the possibility of a new world record in the first place.

Several simplifications were made in various areas of uncertainty – for instance, the operating costs and taxes imposed on the client. The uncertainty is due to the fact these costs can vary widely depending on the individual company. In the event that a company, in reality, uses a more accurate model, then they should be able to apply the necessary information, rather than using our simplification.

Conclusion

After extensive preliminary modelling, we proposed that a logarithmic function was a suitable model of the trends in the given data. We adopted the method of finding deviations of the true data from the trend line to produce parameters for a general data distribution. By using Microsoft Excel to generate the cumulative probabilities for setting a new world record, an average cost was determined. Thus, the question of whether to self-insure or to purchase insurance was interpreted as an analysis on the acceptability of the insurer's fee with respect to the average cost. This first component of the model presented little adversity in the extrapolation of data – as there were no significant outliers and the correlation coefficient was high.

In assuming that the outcomes of all the events were independent, the model could be extended to a wider range of events. Because of the independent nature of the outcomes, a mathematical analysis based on Bernoulli trials was employed to determine whether or not the organising committee should self-insure or not. The model incorporated the notion of 'equivalence' or 'break-even point' and is somewhat representative of a practical real-life scenario. Although, this was somewhat ambiguous as it also depended on the short-term environmental factors or scientific advancements.

A large assumption that had to be made was that the record couldn't be broken elsewhere leading up to the event. This would have a detrimental effect as the time required could have potentially significantly decreased, thus greatly reducing the probability of a world record being broken. Unfortunately, we were unable to quantitatively address this problem and perhaps, with more time, this could have been overcome.

However, we are confident that the model provides a logical and practical approach to the given problem. The reliability of the model seemed high and any modifications in the model would be to address assumptions and simplifications of the problem.