INSTITUTE AND FACULTY OF ACTUARIES

Curriculum 2019

AUDIT TRAIL

Subject CP2 – Actuarial Modelling

Paper One

Triathlon model

Objective

Each year on the Island of IFoA a Minister's triathlon is held consisting of three stages, each covering a different sporting discipline: a swim, a cycle ride and a run. The Minister for Sports would like to use the results of a charity triathlon, held earlier in the year, to predict the results for the main event.

The purpose of this spreadsheet is to complete the following calculations:

- Validate the charity race data provided.
- Calculate the speed of each athlete for each stage of the charity triathlon and adjust these in order to predict the speed of each competitor for each stage of the Minister's triathlon.
- Predict the total time it takes for each competitor to complete the Minister's triathlon.
- Determine if the event is profit or loss making.
- Calculate how much swim times would need to change for the event to break even.

NB: Input cells are shown in blue.

Data

The following data has been provided by the Minister for Sports:

- Race data from the charity triathlon event. For each of the 30 competitors registered to enter the Minister's event the total time taken to complete the race is provided. This is further split into the time taken to complete each stage of the race, plus two transition times (T1 = swim to cycle ride and T2 = cycle ride to run).
- Details on the length (in km) of each stage of the charity triathlon and the Minister's triathlon.
- Details of the entrance fee ($50) and the discount of 25% to any competitor finishing in the top 10 of the charity triathlon.
- The fixed and per competitor expense incurred in recent years.
- The prize money offered to the 3 competitors with the best (i.e. shortest) time in the Minister's triathlon.

The boss has proposed a formula to determine the relationship between speed achieved in each stage of the charity triathlon and the corresponding speeds for future races.

Amended data

This worksheet performs a number of checks and adjustments on the charity race data.

The raw data is validated as follows:

- Count of data provided. Should be 30 for each column of data.
- Calculate the minimum, maximum and average of each column.
- Check that the Finish Time is equal to the sum of the individual stages and transition times.

This flagged the following data errors:

- The count of data indicated that 32 records were provided.
- The record for Roberta Johnson had no times for the individual stages. It was assumed that this record was included in error and the record was removed from the data set.
- The record for Bertha Jansch appeared to have times which looked valid for the Swim stage but the times for the other stages look to be corrupted. It was assumed that this record was included in error and / or the record had been corrupted, and therefore the record was removed from the data set.
- The minimum T1 time was zero. This was for Jason Baggaley. It was also identified that the Finish Time did not match the sum of the individual components.

  The time for Jason Baggaley was updated to be:

  Finish Time – Transition Times – Swim Time – Run Time

  (Alternatively it can be set equal to the average of the other T1 times.)
- The maximum T2 time was 1 hr 1 min and 30 secs. This was for Andrew Coe. It was also identified that the Finish Time did not match the sum of the individual components.

  The time for Andrew Coe was updated to be:

  Finish Time – Transition 1 – Swim Time – Cycle Time – Run Time

  (Alternatively it appears that 01:01:30 is a typo and is more likely to be 00:01:30.)
- The minimum run stage time was zero. This was for Richard Bashford. It was also identified that the run stage time did not exceed that of the swim stage time.

  Since the Finish Time does match the sum of the individual components, it is not possible in this case to back solve.

The approach taken is to take the average of the other run stages as Richard's time. The Finish Time is then adjusted to reflect the new sum of the individual components.

All checks are re-applied on the corrected data. No errors are indicated.

Assumptions

- Other than identified errors, the data provided is correct.
- Past performance (charity triathlon) is a good guide to future performance (Minister's triathlon).
- The race time for the Minister's triathlon is only dependent on the race time achieved in the charity triathlon e.g. no further training will impact the results.
- The provided relationship between speeds in the charity triathlon and those in future races gives a good estimate for speeds in the Minister's triathlon.
- Transition times are independent of the length of the preceding stage.
- Transition times in the Minister's triathlon will be the average of T1 and the average of T2 times achieved in the charity triathlon (or other sensible assumption, e.g. individual transition times are the same as those for the charity race).
- The competitors who performed in the charity race performed to their best level.
- The 30 athletes registered for the Minister's triathlon include the top 10 competitors in the charity race i.e. there will be no injuries impacting the ability of competitors to race.
- The performance of each competitor is independent between races and independent of other competitors.
- Expenses incurred will be at the level incurred in recent years.
- If any of the top 3 placed competitors in the Minister's triathlon have times in excess of 3 hours, the prize money awarded to that competitor will be 0.
- The prize money is calculated using the exact time pro-rating between whole minutes. (or other sensible assumption, e.g. rounding down to nearest integer minute).

Parameters sheet

This sheet sets out all parameters used in the spreadsheet model, including:

- Distribution parameters for the 3 stages, i.e. factors, expressed as λ, used in the formula provided, to determine the speed, for each stage, in future races at different distances.

- Course distances for the charity triathlon and the Minister's triathlon for each of the 3 stages.

- The unit of time (referred to as "1 hour" in the formulae below – i.e. 01:00:00), used to scale and format the competitor times when shown in hours and speeds when shown in km per hour.

- Number of competitors, used in determining the cashflows.

- The entry fee.

- The discount rate for the Minister's triathlon entry fee applied to the top competitors from the charity triathlon.

- The number of top competitors in the charity race eligible for a discount on entry.

- Parameters for the prize scaling, i.e. cash sum per minute under the threshold time in the Minister's triathlon.

- Time to compare against, i.e. the threshold time used to determine the prizes awarded.

- Expense parameters, i.e. fixed and variable per-competitor expenses for the Minister's triathlon.

Where named ranges have been defined these are detailed in red italic adjacent to the relevant cells.

Method

Current speeds

In this worksheet the speed each stage was completed in the charity event is calculated.

The amended data in columns D, F and H are linked through from the "Amended data" worksheet. These columns set out the stage times (corrected where necessary) for the swim, cycle and run respectively. The speeds for each stage are determined as follows:

Column E calculates the swim speed, in km/hour, as:
- length (in km) of swim / (swim stage time / 1 hour)

Column G calculates the cycle speed, in km/hour, as:
- length (in km) of cycle / (cycle stage time / 1 hour)

Column I calculates the run speed, in km/hour as:
- length (in km) of run / (run stage time / 1 hour)

The "1 hour" is from the "Parameters" sheet and allows the times in minutes to be calculated as fractions of an hour and thus calculate the speeds in km/hour.

Checks

The minimum, maximum and average of each column has been calculated in rows 3 to 5. The data is counted in row 6. The checks show that the speeds calculated for each stage appear sensible, i.e. the speeds are all humanly possible!

Reasonableness checks

We would expect the cycle stage to take the longest time to complete, the run stage would take the next longest to complete and the swim stage should be the quickest stage to complete. This is due to the relative length of each stage and the speed each discipline can be completed in. The data reflects this pattern – a reasonableness check is carried out in column K.

We would expect the speed competitors can swim will be the slowest, with run next and the cycle the quickest. The data reflects this logic – a reasonableness check is carried out in column M.

Minister triathlon

The aim of the worksheet is to calculate the time it will take each competitor to complete the Minister's triathlon. From this we can determine who we expect to secure the top three places.

The speed in the Minister's triathlon is related to the speed on the charity course using a multiplicative adjustment factor:

$$\frac{1}{1+2e^{\lambda d}}+\frac{2}{3}$$

where λ is a constant which varies by discipline and d is the number of extra kilometres distance in the future triathlon compared to the charity triathlon.

The adjustment factors for each discipline are calculated in cells D9:F9. The speeds on the charity triathlon are multiplied by the adjustment factors to determine the expected speed for the swim, cycle and run stage for each competitor, in columns D, E and F respectively.

In columns G, H and I, the time for each stage of the Minister's triathlon is then calculated as:

Time = Distance of Stage ÷ (Speed for Stage / 1 hour)

The transition times for each competitor are based on the average transition times from the charity triathlon. Namely

Transition Time = Average T1 from Charity triathlon + Average T2 from Charity triathlon

In column J, the total time for each competitor is calculated as the sum of the time taken to complete the three stages plus the two transition times.

In column K, the RANK() Excel function is used to determine the position of each competitor in the race.

The minimum, quartile 1, mean, quartile 3 and maximum are all calculated for the overall competitors' times from the Minister's triathlon using the standard Excel functions MIN(), MAX() and QUARTILE(). See cells D4:J6.

Checks

The minimum, maximum and average of each column has been calculated in rows 4 to 6. The checks show that the speeds calculated for each stage appear sensible, i.e. the speeds are all humanly possible!

Reasonableness checks

We would expect the total time for competitors to complete the Minister's triathlon to be longer than the charity race because each stage of the Minister's triathlon is longer. The results support this assertion.

We would expect the longest stage of the Minister's triathlon for the average competitor to be the swim stage. The data shows that this is the case for 24 out of the 30 competitors, the other 6 had the cycle as the longest stage. See column O and cells P4 and P5.

The average times of the swim, cycle and run are calculated in cells G7, H7 and I7 respectively. Based on these and the average transition times we have calculated the average time for the race in cell J7. This compares closely to the calculated average time from the individual calculations in cell M5.

Also note that setting the distances in the Minister's race to be the same as the charity race yields the same overall time: this is because with $d = 0$ the speed multiplier is then $2/3 + 1/3 = 1$.

Cashflows

The worksheet calculates the expected cashflows and determines if the event is profit or loss making.

Prize money paid

The top 3 competitors in the Minister's triathlon are brought in from the "Rank" column in the "Minister triathlon" sheet. This uses OFFSET and MATCH functions to look up the top 3 finish times. The prize money (cells E5:E7) given to each is determined as:

$$\text{Prize Money} = \text{Multiplier} \times \text{Minutes under 3 hours}$$

where the multipliers are set out in the "Parameters" sheet and vary for the first, second and third placed competitor. The prize amounts are summed to give the total prize money (cell E9).

Expense

Total expenses incurred are given by:

- fixed expenses
- total per competitor expenses; and
- prize money.

These are summed to give the total expenses in cell F17.

Income

Income is derived from the entry fees, which is:

- a fixed price for each competitor.
- a discounted price for the top 10 competitors in the charity race.

The entry fee income is calculated in cells C13 and C14 and is summed to give the total income in cell C17.

Net cashflow

In cell C19, the net cashflow is determined by taking the total income and subtracting the prize money and expenses. The event does make a profit, i.e. total income exceeds total prize money and expenses.

Summary results

The worksheet summarises results and draws charts.

There is a stacked bar chart illustrating the key statistics of the finishing times. These include the minimum, quartile 1, mean, quartile 3 and the maximum. The statistics are brought through from the "Minister triathlon" sheet.

There is a stacked bar chart illustrating the income, outgo and profit. The income bar is split into income from full priced entry fees and income from discounted entry fees. The outgo bar is split into prize money, per competitor expenses and fixed expenses. The income and outgo figures are brought through from the "Cashflows" sheet.

Minister triathlon adj swim

The aim of the worksheet is to calculate the improvements in average swim times necessary for the event to break even.

This sheet is a copy of the "Minister triathlon" sheet with the following additional calculations:

- A new swim λ factor is entered into cell D1.
- The adjustment parameter formula in cell D9 now uses the new λ factor. As a self-check on the spreadsheet, note that changing this cell back to the original λ value yields the original profit value in the base scenario.
- The profit / loss from this additional scenario is brought into cell H1. This is calculated in the "Cashflows adj swim" sheet (see below).
- Cell H1 is targeted to 0 (the event breaks even) by adjusting cell D1 via goalseek.

The reduction in average swim times is calculated in cell K9.

Reasonableness checks

As there is a profit in the original scenario we would expect there to be an improvement in swim times so that more prize money is paid and the profit reduced to zero.

Furthermore, it is reasonable that:

- as λ decreases, the profit should decrease, given the adjustment factor function involving λ.
- all competitors' speeds increase (and times decrease) as λ decreases.
- those competitors taking longest to complete the swim see a greater absolute impact from the decrease in λ.
- the ratio of the λ should be close to the ratio of the average swim times (see cell N9).

Cashflows adj swim

The worksheet calculates the expected cashflows and determines the profit or loss from the event using the new "break even" λ factor for swim times.

This is a copy of the "Cashflows" sheet but it brings through the top 3 times from the "Minister triathlon adj swim" sheet instead of the "Minister triathlon" sheet (see cells C5 to C7).

END OF AUDIT TRAIL