Optimizing route choice in Mini-mountain marathons

Peter J. Dodd (Sheffield)
September, 2015

Background

- Mini-mountain marathons are long-course score orienteering races over hilly terrain.
- Checkpoints carry scores reflecting their remoteness.
- Any number can be navigated in any order.
- Not possible to visit all checkpoints.
- Late return is penalized via an escalating points deduction.

Finding the best route is an NP-hard combinatorial optimization problem known as the Orienteering Problem.

Route choices: position 2-5

We used a score function based on the split data from the 1st round of the Rab 2015 event series.

We wondered:
- Did top finishers take similar routes?
- Was speed or route-planning acumen more important?
- How close to optimal is route choice by experienced competitors?

Analysis

- We modelled the split, \( T_{ij} \), over leg \( i \) over competitor \( j \) as:

  \[ T_{ij} = \frac{d_i}{s_j} \]

  where \( d_i \) is a notion of distance for the leg, and \( s_j \) is a notion of speed for the competitor.

- We used linear regression on the log-splits to infer the relative speed of competitors and leg lengths.
- We encoded possible route choices as the sequence appearing between 1 and \( N \) in permutations of 1, ..., \( N \).
- We wrote a score function based on the points accrued minus the penalty associated with the total route duration.
- We used a genetic algorithm to search through the space of permutations to optimize the score for the winner's speed.

Discussion

- Top routes are surprisingly varied (see above).
- Rank and speed related but with increasing scatter due to mishap (see left).
- Genetic algorithm did improve winner's score; but only by ~7% (see below).

Questions

- We considered split data from the 1st round of the Rab 2015 event series.

But...
- Some legs missing, arguably irrelevant ones.
- Within-leg navigation and speed confounded.
- Relatively flat course.

Analysis

- We modelled the split, \( T_{ij} \), over leg \( i \) over competitor \( j \) as:

  \[ T_{ij} = \frac{d_i}{s_j} \]

  where \( d_i \) is a notion of distance for the leg, and \( s_j \) is a notion of speed for the competitor.

- We used linear regression on the log-splits to infer the relative speed of competitors and leg lengths.
- We encoded possible route choices as the sequence appearing between 1 and \( N \) in permutations of 1, ..., \( N \).
- We wrote a score function based on the points accrued minus the penalty associated with the total route duration.
- We used a genetic algorithm to search through the space of permutations to optimize the score for the winner's speed.

Discussion

- Top routes are surprisingly varied (see above).
- Rank and speed related but with increasing scatter due to mishap (see left).
- Genetic algorithm did improve winner's score; but only by ~7% (see below).

Questions

- We considered split data from the 1st round of the Rab 2015 event series.

But...
- Some legs missing, arguably irrelevant ones.
- Within-leg navigation and speed confounded.
- Relatively flat course.

Analysis

- We modelled the split, \( T_{ij} \), over leg \( i \) over competitor \( j \) as:

  \[ T_{ij} = \frac{d_i}{s_j} \]

  where \( d_i \) is a notion of distance for the leg, and \( s_j \) is a notion of speed for the competitor.

- We used linear regression on the log-splits to infer the relative speed of competitors and leg lengths.
- We encoded possible route choices as the sequence appearing between 1 and \( N \) in permutations of 1, ..., \( N \).
- We wrote a score function based on the points accrued minus the penalty associated with the total route duration.
- We used a genetic algorithm to search through the space of permutations to optimize the score for the winner's speed.

Discussion

- Top routes are surprisingly varied (see above).
- Rank and speed related but with increasing scatter due to mishap (see left).
- Genetic algorithm did improve winner's score; but only by ~7% (see below).

Questions

- We considered split data from the 1st round of the Rab 2015 event series.

But...
- Some legs missing, arguably irrelevant ones.
- Within-leg navigation and speed confounded.
- Relatively flat course.

Analysis

- We modelled the split, \( T_{ij} \), over leg \( i \) over competitor \( j \) as:

  \[ T_{ij} = \frac{d_i}{s_j} \]

  where \( d_i \) is a notion of distance for the leg, and \( s_j \) is a notion of speed for the competitor.

- We used linear regression on the log-splits to infer the relative speed of competitors and leg lengths.
- We encoded possible route choices as the sequence appearing between 1 and \( N \) in permutations of 1, ..., \( N \).
- We wrote a score function based on the points accrued minus the penalty associated with the total route duration.
- We used a genetic algorithm to search through the space of permutations to optimize the score for the winner's speed.

Discussion

- Top routes are surprisingly varied (see above).
- Rank and speed related but with increasing scatter due to mishap (see left).
- Genetic algorithm did improve winner's score; but only by ~7% (see below).

Questions

- We considered split data from the 1st round of the Rab 2015 event series.

But...
- Some legs missing, arguably irrelevant ones.
- Within-leg navigation and speed confounded.
- Relatively flat course.