Make a world

Slides for break-out group

Daniel Ricard

February 2-3 2021

Section 1

Outline

Outline

- agree on a git workflow
- apples: we will working from the R Markdown Notebook called "orchard.Rmd"
- fishes: we will working from the R Markdown Notebook called "age-and-growth.Rmd"

Section 2

git workflow

git workflow

- clone the sim-make-a-world repository
- pull from the origin regularly (git pull origin main)
- work from your own branch (git checkout -b yourbranchname)
- go nuts with your ideas and code contributions
- add, commit,
- eventually, push
- open a merge request / pull request
- pull from the origin again after the request has been merged

Section 3

Apples

Apples

An orchard with 100 x 100 trees

2 types of orchards: 1 - uniform distribution of apple trees across the orchard (uniform orchard) 2 - an orchard with equal proportions of “good” and “bad” trees (two-tier orchard)

2 sampling strategies with similar time investments: a - sample trees with a ladder (count all the apples in each tree) b - sample trees from the ground (only count the apples that you can reach)

Measure of performance: - the probability of wrongly estimating the number of apples in the orchard by more than x%

Deterministic case - uniform orchard

100 apples in each tree x 10,000 trees = 1,000,000 apples

How many trees do you need to sample to estimate this value?

What is the probability of overestimating the number of apples in the orchard?

Deterministic case - two-tier orchard

100 apples in each tree x 5,000 trees = 500,000 apples 50 apples in each tree x 5,000 trees = 250,000 apples for a total of 750,000 apples

How many trees do you need to sample to estimate this value?

How can we make sure that we are not overestimating the number of apples in the orchard?

What is an appropriate performance measure to capture this?

Stochastic case

Each tree has its own number of apples, drawn from a Poisson distribution

Stochastic case - uniform orchard

The number of apples in each tree is a sample from a Poisson distribution with a mean of 100

The orchard still has 10,000 trees

How many apples are in the orchard?

What is the probability of overestimating the number of apples in the orchard?

Sampling from distributions in R

see ?family in R

Do we really need simulations for this?

probably not, but it is not a bad way to explore things

perhaps with added complexity the simulations would prove useful

Possible expansions (the slippery slope)

anything we would like to explore?

- orchard with areas of different productivity (see what is gained by using stratified sampling)
- add more variability in the number of apples per tree (e.g. use a negative binomial instead of a Poisson)

Section 4

Fishes

Fishes

A fish population where growth is to be estimated

2 types of growth 1 - growth is the same for all individuals in the population 2 - individuals in the population exhibit different growth based on the year they were born (cohort effect)

2 sampling strategies a - use length-age pairs to estimate a growth model b - use individual growth trajectories to estimate a growth model

Measure of performance: - minimize the probability of incorrectly estimating length-at-age matrices

Deterministic case

Stochastic case

Possible expansions (the slippery slope)

Section 5

Final words

Useful nuggets

- the importance of seeding the random number generator (set.seed function in R)
- writing output as your simulations run, R Notebooks are pretty useful for exploratory analyses
- loops vs. vectorized functions
- how long will this simulation take to run?
-

staying the course

1 - operating model realisations 2 - analytic treatments 3 - performance measures

getting lost in the possibilities

Section 6