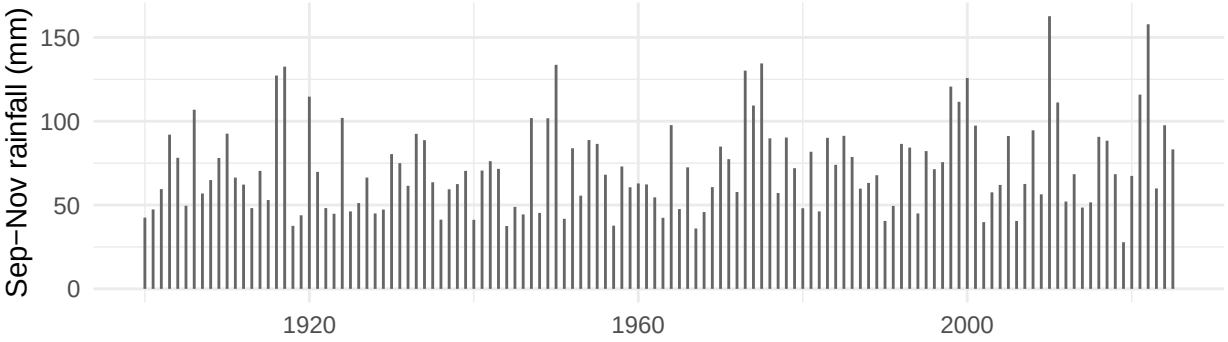
Homework 6: fitting a hidden Markov model

STAT5710 Applied Stochastic Processes



Rainfall over Australia varies a great deal from year to year, and not simply as noise: wet and dry years arrive in runs. One explanation is that the continent sits in one of a few broad climate regimes at any time, each persisting a while before giving way to another. One major driver of such regimes is the El Niño-Southern Oscillation (ENSO). Here you'll fit a hidden Markov model and explore the alignment between ENSO and latent climate states inferred from rainfall data.

The data

Total September–November rainfall, averaged over Australia, 1900–2025, in millimetres, from the Australian Bureau of Meteorology. One number per year, 126 years.

`https://tdruiz-stat5710.share.connect.posit.cloud/data/australia-rainfall.csv`

A second column, `oni`, recorded from 1950, measures sea surface temperature in the central Pacific that spring; it is how climatologists define El Niño (above +0.5) and La Niña (below −0.5).

1. Choose an appropriate emission distribution and fit a three-state hidden Markov model.
2. Inspect the transition matrix and the emission parameters, interpret the states, and quantify their persistence.
3. Compute confidence intervals for the state means. Do the states appear well-differentiated?
4. Order the states from driest to wettest and decode the series.
5. Now compare the distribution of `oni` across decoded states and cross-tabulate decoded state against El Niño, neutral, and La Niña. How accurately can ENSO be inferred from the rainfall data?
6. Interpreting the states as El Niño, neutral, and La Niña, how much time elapses on average between El Niño years according to the model? Does this match the observed data? Repeat for La Niña.
7. How do you explain the results in 6?