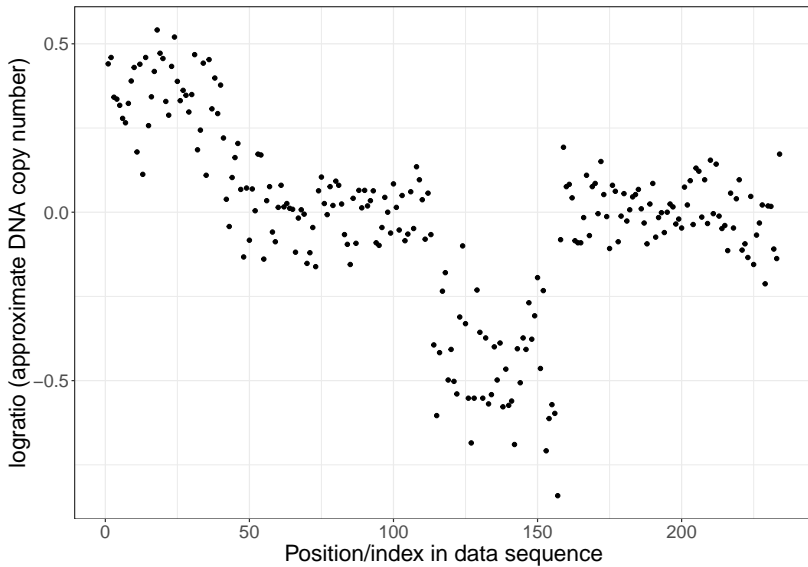


# Segmentation model selection and evaluation

Toby Dylan Hocking

## Background: detecting abrupt changes is important

Example from cancer diagnosis: breakpoints are associated with aggressive disease in neuroblastoma.



# Motivation for segmentation model selection and evaluation

- ▶ In each of the segmentation models we have studied, there is a choice of model size (segments/changepoints or hidden states).
- ▶ Too large model sizes result in false positives (changepoints predicted by algorithm but they are not significant/real).
- ▶ Too small model sizes result in false negatives (no changepoint predicted where there should be).
- ▶ Want to maximize true positive rate (number of correctly predicted changepoints) and true negative rate (number of correct predicted regions without changepoints).

# Model selection via Classic Information Criteria

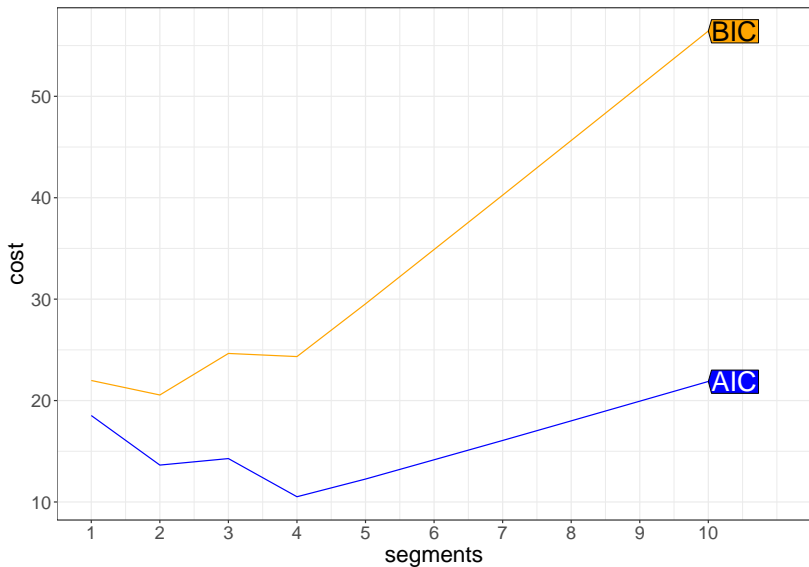
For every model size  $k \in \{1, \dots, K\}$  let  $L_k$  be the loss.

Information criteria choose the model which minimizes the penalized cost, for some non-negative penalty  $\lambda \geq 0$ ,

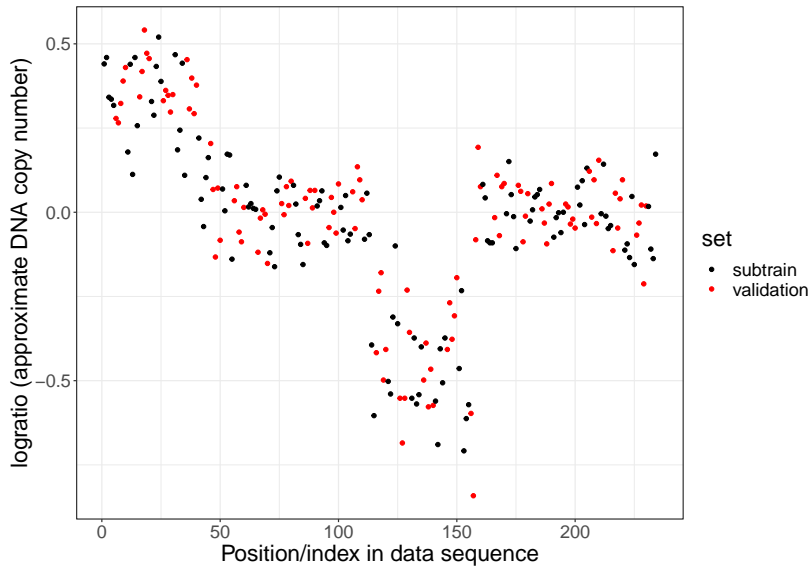
$$C_k = \min_{k \in \{1, \dots, K\}} L_k + \lambda k$$

- ▶ BIC=Bayesian Information Criterion, sometimes referred to as SIC=Schwarz who was the author.  $\lambda = \log n$  where  $n$  is the number of data points.
- ▶ AIC=Akaike Information Criterion:  $\lambda = 2$ .
- ▶ Data viz:  
<http://bl.ocks.org/tdhock/raw/43ac9c6be9188dcb02a7/>

# Model selection criteria plot for binary segmentation



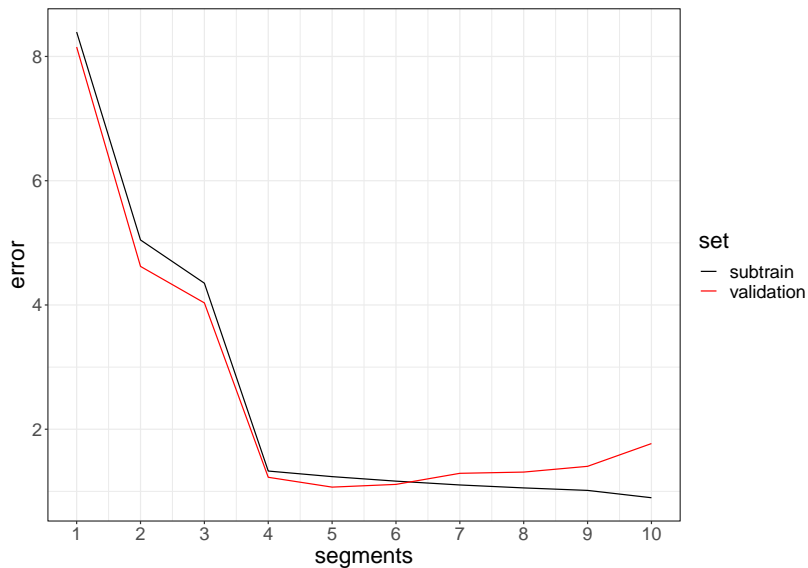
# Cross-validation for model selection



## Idea for cross-validation

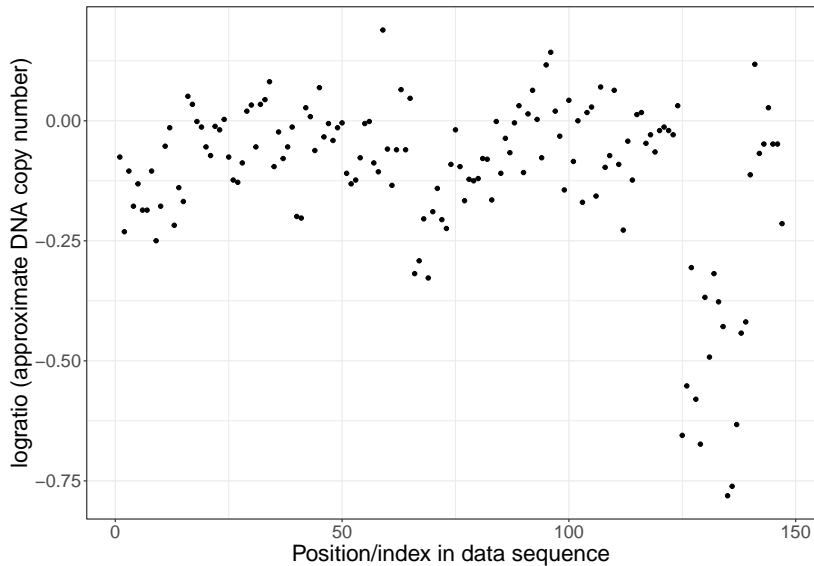
- ▶ Divide full data sequence into subtrain and validation sets.
- ▶ Use subtrain data as input to learning algorithm.
- ▶ Use validation data to choose best model size (min error or negative log likelihood).
- ▶ As model size increases, subtrain error should always decrease, whereas validation error should be U shaped.

## CV Error plot

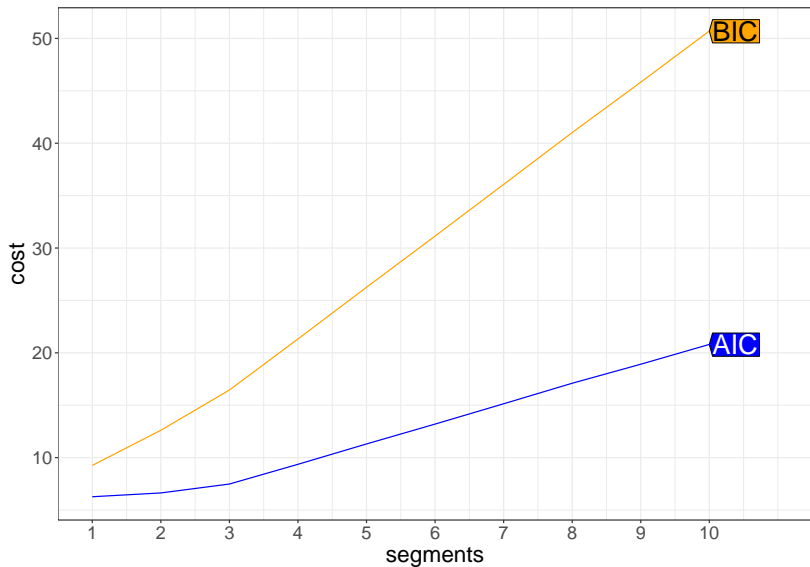




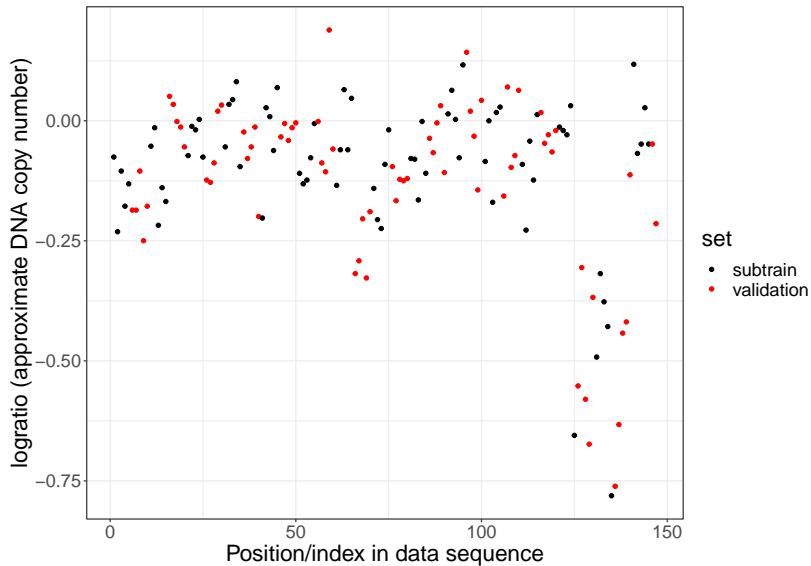
## Another data set



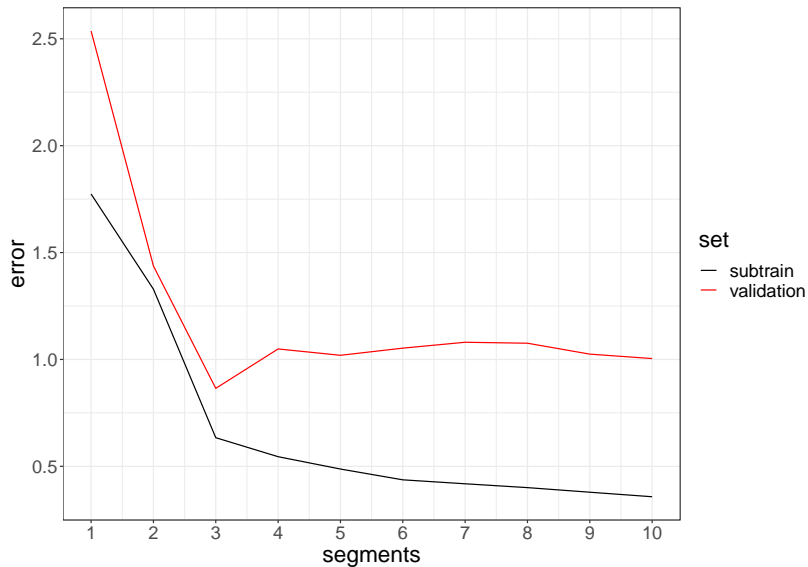
## Model selection plot



# Cross-validation for model selection



## CV error plot



Labeled regions for evaluating accuracy of changepoint predictions

TODO

## Possible exam questions

TODO