

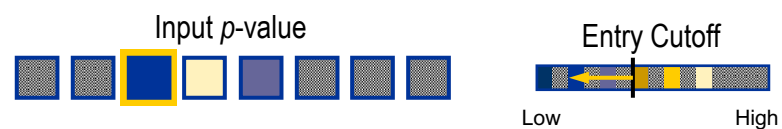
## Sequential Selection – Forward



- Forward selection creates a **sequence of models of increasing complexity**.
  1. The sequence starts with [the baseline model](#), a model predicting the overall average target value for all cases.
  2. The algorithm searches the set of [one-input model](#) and selects the model that most improves on the baseline model.
  3. It then searches the set of [two-input models](#)...
  4. [The sequence terminates when no significant improvement can be made](#) (no  $p$ -value is below a predefined entry cutoff, the forward selection procedure)
- **A small  $p$ -value indicates a significant improvement.**

2

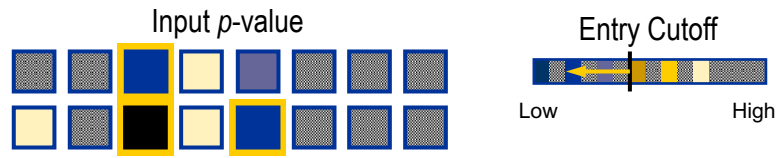
## Sequential Selection – Forward



...

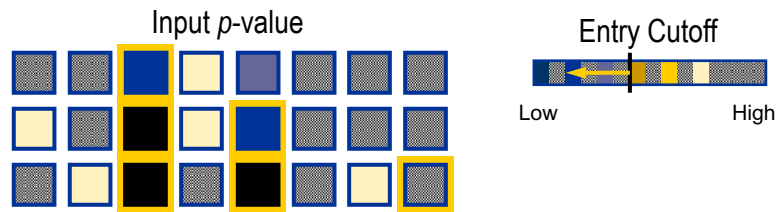
3

## Sequential Selection – Forward



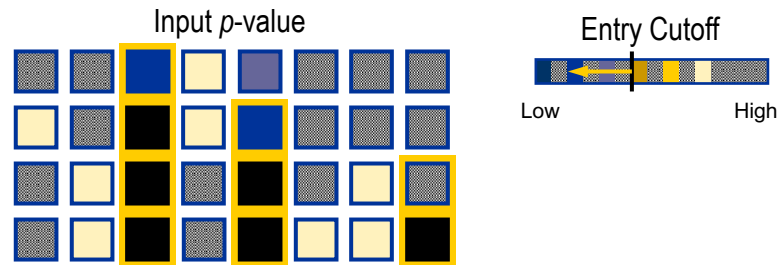
4

## Sequential Selection – Forward



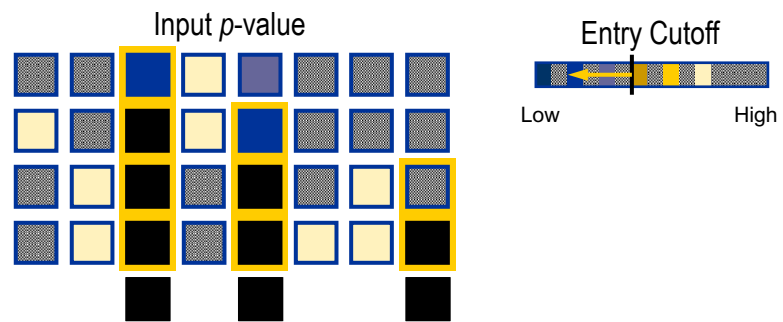
5

## Sequential Selection – Forward



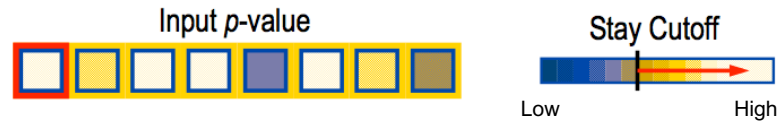
6

## Sequential Selection – Forward



7

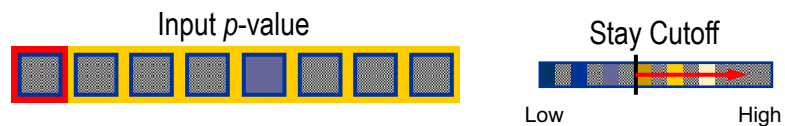
## Sequential Selection – Backward



- Backward selection creates a **sequence of models of decreasing complexity**.
  1. The sequence starts with a saturated model, which is a model that contains **all available inputs**.
  2. Inputs are sequentially removed from the model (**removing the input with the highest  $p$ -value**).
  3. The sequence terminates when all remaining inputs have a  $p$ -value that is less than the predetermined stay cutoff.

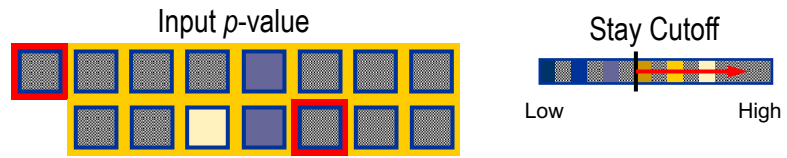
8

## Sequential Selection – Backward



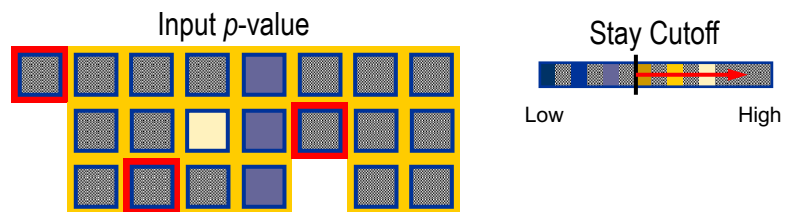
9

## Sequential Selection – Backward

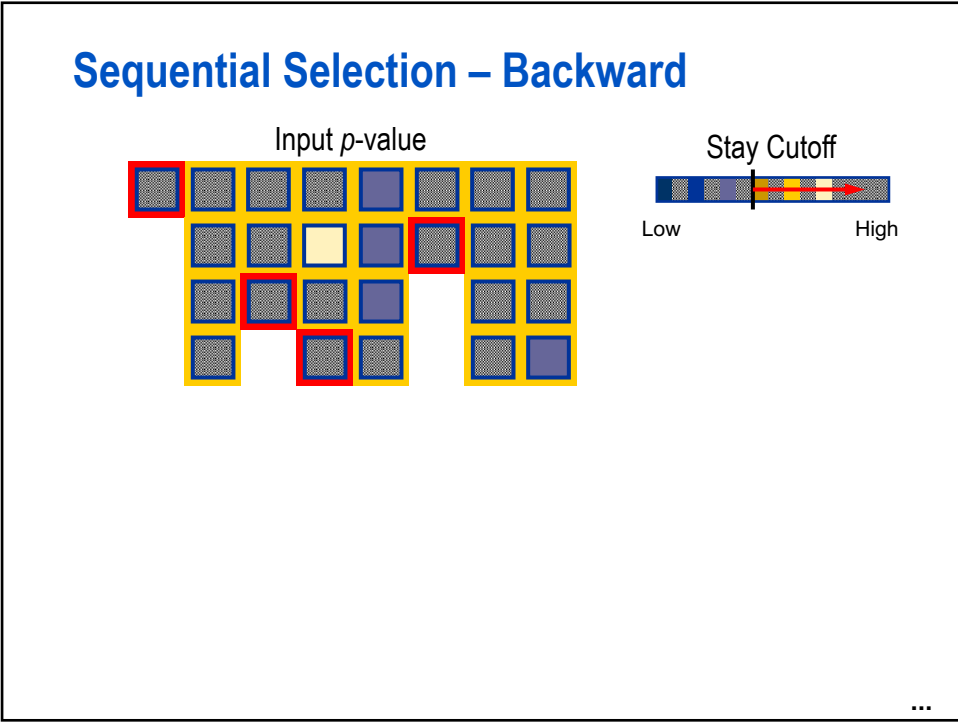


10

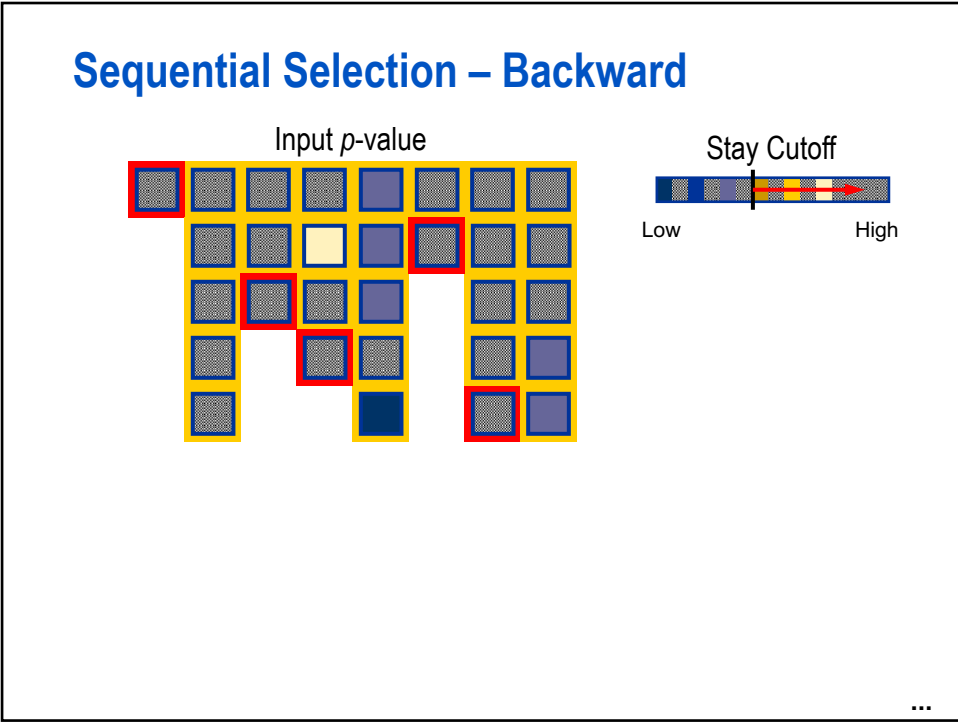
## Sequential Selection – Backward



11

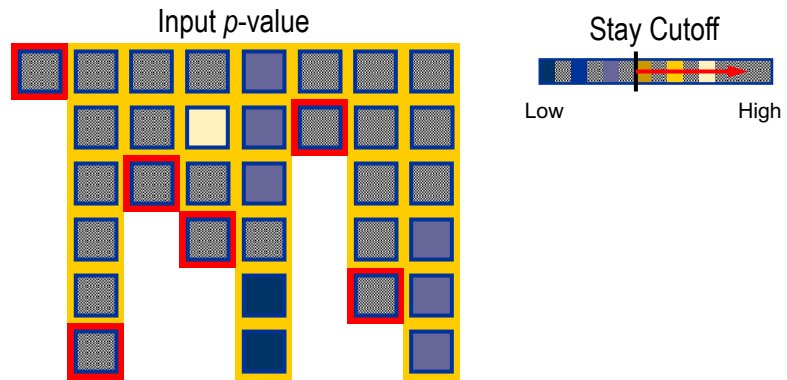


12



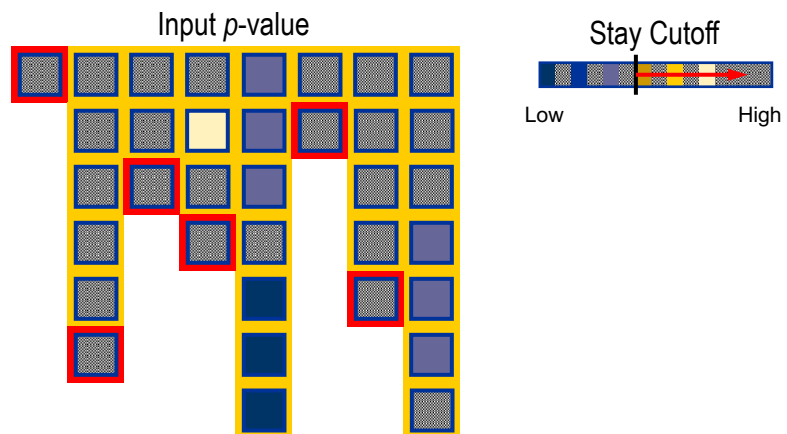
13

## Sequential Selection – Backward



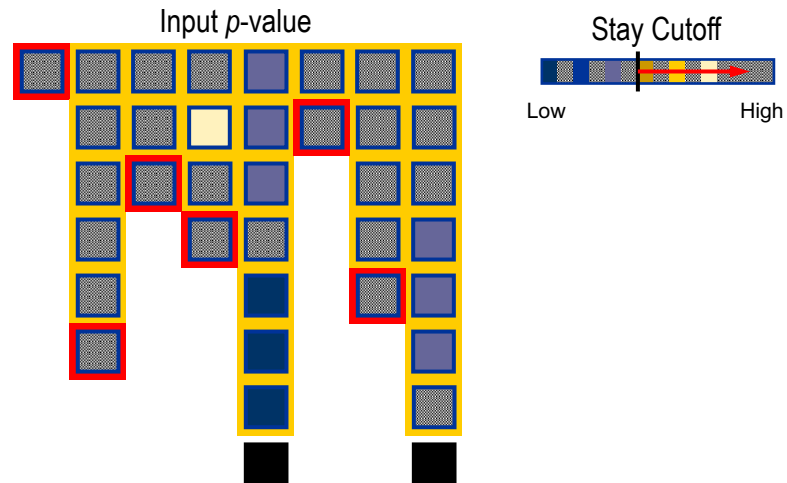
14

## Sequential Selection – Backward



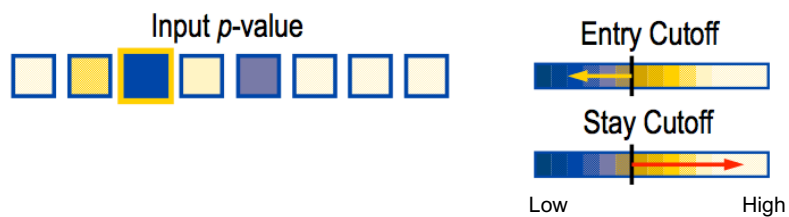
15

## Sequential Selection – Backward



16

## Sequential Selection – Stepwise



- Stepwise selection combines elements from both the forward and backward selection procedures.
  1. The method begins in the same way as the forward procedure – sequentially adding inputs with the **smallest p-value** below the entry cutoff.
  2. After each input is added, the algorithm reevaluates the statistical significance of all included inputs – **inputs exceeding the stay cutoff are removed**.
  3. The process terminates when all inputs available for inclusion in the model have p-values in excess of the entry cutoff and all inputs already included in the model have p-values below the stay cutoff.

17