Maximizing likelihood

(after slide 46)

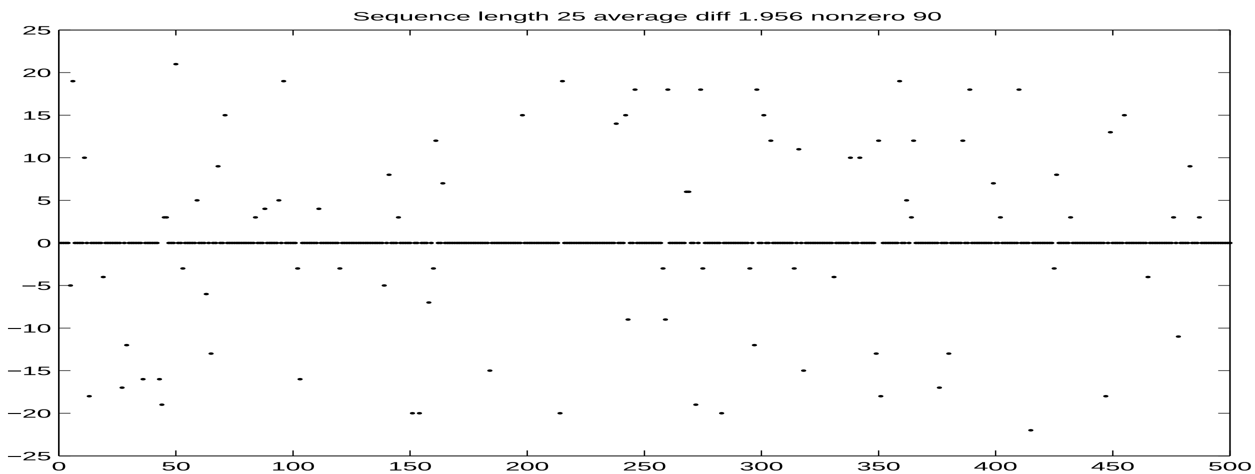
- Points a_i, segment probabilities (levels) p_j
- $r(i)$: the unique j such that i b belongs to segment j
- $L(a_1, a_2, \ldots, a_n|(p_1, \ldots, p_k]) = \prod_{i=1}^{n} L(a_i|p_{r(i)})$
- $L(a_i|p) = p$, if $a_i = 1$, and $L(a_i|p) = 1 - p$, if $a_i = 0$
- $L(a_1, a_2, \ldots, a_n|p) = \prod_{i=1}^{n} p_{r(i)}^{a_i}(1 - p_{r(i)})^{1-a_i}$
- Maximizing this is the same as minimizing the $-$ log likelihood
- $mllh(a_1, a_2, \ldots, a_n|p) = -\sum_{i=1}^{n} a_i \log p_{r(i)} + (1 - a_i)\log(1 - p_{r(i)})$

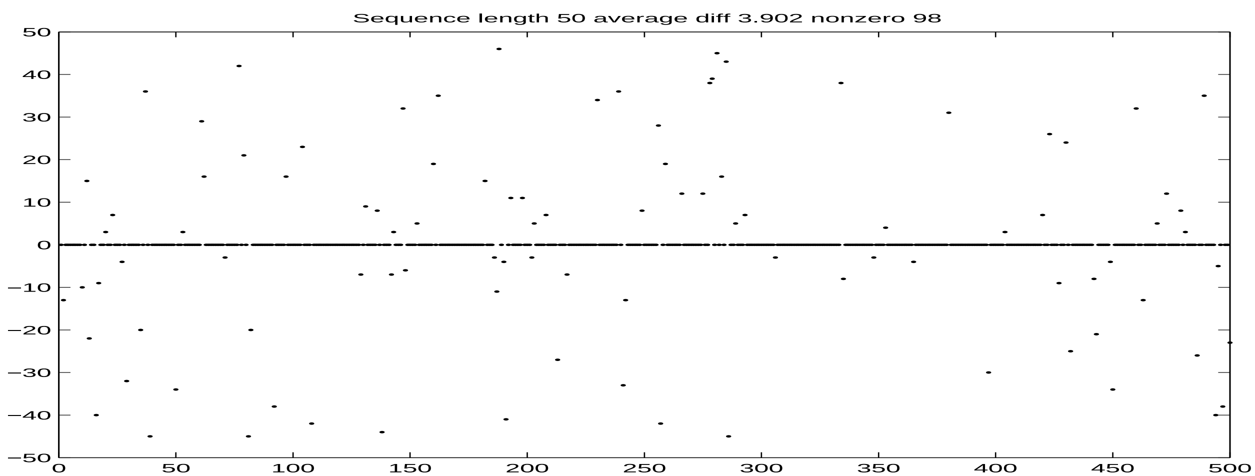
Maximizing likelihood vs. squared error

- When do the two measures prefer different segmentations?
- Sequence A, segmentation into two pieces
- $S_1 = (1, i), (i + 1, n)$ or $S_2 = (1, j), (j + 1, n)$ with $j > i$
- Look for cases where $mllh(S_1) > mllh(S_2)$ but $sqerr(S_1) < sqerr(S_2)$
- Gets fairly complex

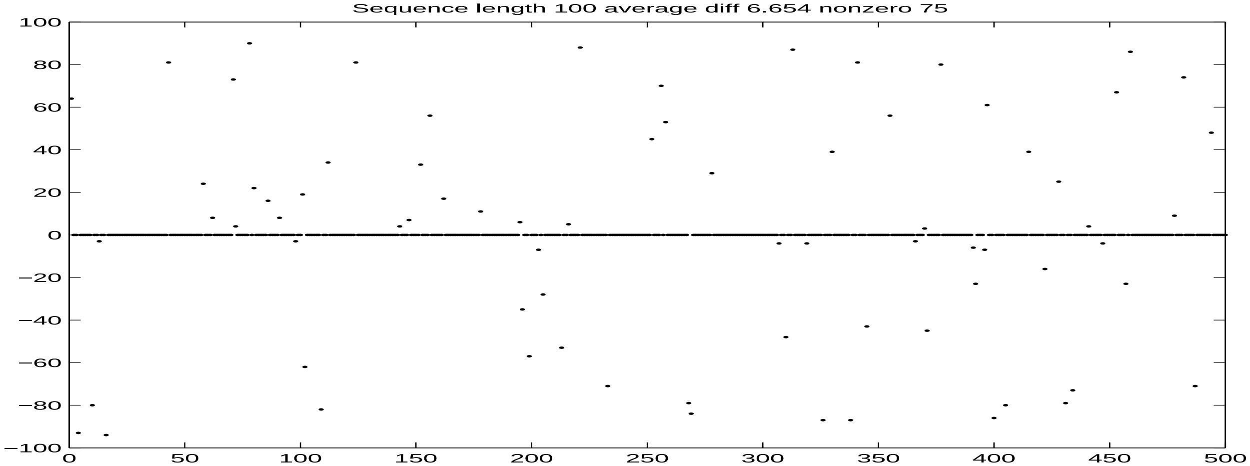
Experimental study

- Generate random sequences of 0s and 1s
- Find the best division into 2 using log likelihood or squared error

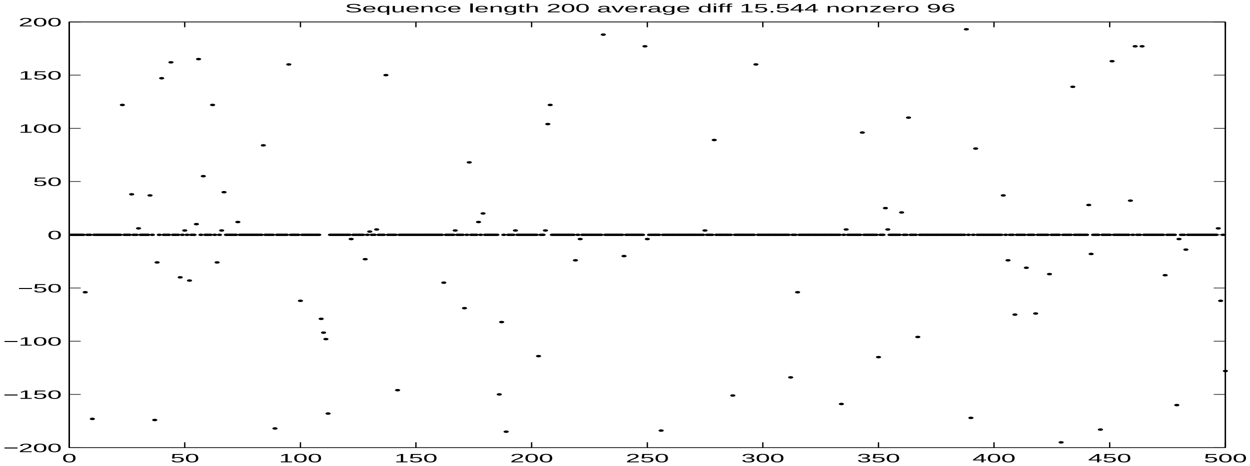




Sequence length 100 average diff 6.654 nonzero 75



Sequence length 200 average diff 15.544 nonzero 96



Comments

- Differences are fairly common
- Can be large in terms of absolute values
- Differences in terms of the length of the sequence?
- Are the differences significant?
- The data was random, so no signal was supposed to be present