Semiparametric Estimation of PSA Nadir Using Model Averaging

Tanaka S. and Tanaka S.
Kyoto University Hospital
Prostate specific antigen (PSA)

- A tumor marker
  - Declines with anti-cancer therapy
  - Increases as the disease progresses
- A PSA level while the treatment suppresses tumor activity effectively is regarded as a good index for treatment response
- Minimum level of PSA is commonly used for this purpose
Downward bias of PSA nadir estimated by minimum
Models for estimating PSA nadir
  › Penalized spline with join-points
  › Model selection/averaging

Construction of confidence interval
  › C.I. ignoring estimation of join-points is optimistic
Notation

- $Y_i$: Longitudinally measured log PSA level at $i$th visit
- $t_i$: Time at $i$th visit (month)
- $[\tau_0, \tau_1]$: Duration of PSA nadir
Problem formulation

- Join-point models
  \[ f(t) + \varepsilon \quad \text{if} \quad t < \tau_0 \]
  \[ f(\tau_0) = f(\tau_1) \quad \text{if} \quad \tau_0 \leq t \leq \tau_1 \]
  \[ f(\tau_1) \quad \text{if} \quad \tau_1 < t \]

- Goal
  > Estimate \( \theta \) without assuming \( f(t) \) to be known
Our approach is to fit $p$-spline for $t < \tau_0$ or $\tau_1 < t$.

A particular choice of basis function is

$$f(t) = \sum_{k=A}^{B} \beta_k (t - \kappa_k)^2 +$$

Equally-spaced knots: $\kappa_1 < \ldots < \kappa_A < \ldots < \kappa_B < \ldots < \kappa_K$

If we can set $\kappa_A = \tau_0$ and $\kappa_B = \tau_1$, $f(t) = 0$ for $\tau_0 \leq t \leq \tau_1$, as expected.
Examples: $\left(\kappa_A, \kappa_B\right) = (4, 32)$ and $\left(18, 32\right)$
Model selection/averaging

- Actually join-points are unknown
- Model selection
  - Choose $\kappa_A$ and $\kappa_B$ by some criteria
  - AIC, BIC...
- Model averaging
  - Estimates $\theta$ by a weighted average of estimates from models with all possible $\kappa_A$ and $\kappa_B$
Burnham and Anderson (2002)

\[ \sum w_i \sqrt{\text{Var}(\hat{\theta}_i)} + (\hat{\theta}_i - \hat{\theta})^2 \]

- Here \( \hat{\theta} \) is estimator in the \( i \)th sub-model
- Derived by assuming \( \text{Cov}(\hat{\theta}_i, \hat{\theta}_j) = 1 \)

Hjort and Claeskens (2003)

- Derived asymptotic distribution of \( \hat{\theta} \) when misspecification bias is \( O(n^{-1/2}) \)
- Confidence interval with bias correction
AIC selected \((\kappa_A, \kappa_B) = (6, \infty)\)
BIC selected the same model, although assigned larger weights to simpler models.
Estimated trend of PSA levels

PSA (ng/ml, log scale) 0.1 1 10 100 1000

Time from Initiation of Treatment (month) 0 6 12 18 24 30 36 42

Naïve 95% C.I. [1.77 to 1.99]
Burnham and Anderson 95% C.I. [1.65 to 2.11]
Simulation

- Setting
  - For simplicity, we consider only one join-point at median of $t$
  - $N=20, 100$
  - True value of PSA nadir is 0
  - We obtained remaining parameters from the PSA data

- Fitted model
  - P-spline with maximum 10 knots
Results: Join-point selection
Results: Point estimation

- Even when $N=20$, is approximately unbiased and normal.
- Difference between model selection and averaging is small.
### Results: Interval estimation

<table>
<thead>
<tr>
<th>Model</th>
<th>Method</th>
<th>S.E. (Mean)</th>
<th>Cover. Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC selected</td>
<td>Simulated</td>
<td>0.063</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>B.A.</td>
<td>0.059</td>
<td>0.931</td>
</tr>
<tr>
<td></td>
<td>H.C.</td>
<td>0.072</td>
<td>0.964</td>
</tr>
<tr>
<td></td>
<td>Naive</td>
<td>0.051</td>
<td>0.888</td>
</tr>
<tr>
<td>BIC selected</td>
<td>Simulated</td>
<td>0.063</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>B.A.</td>
<td>0.058</td>
<td>0.929</td>
</tr>
<tr>
<td></td>
<td>H.C.</td>
<td>0.072</td>
<td>0.965</td>
</tr>
<tr>
<td></td>
<td>Naive</td>
<td>0.051</td>
<td>0.888</td>
</tr>
<tr>
<td>AIC weighted</td>
<td>Simulated</td>
<td>0.061</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>B.A.</td>
<td>0.059</td>
<td>0.932</td>
</tr>
<tr>
<td></td>
<td>H.C.</td>
<td>0.072</td>
<td>0.968</td>
</tr>
<tr>
<td></td>
<td>Naive</td>
<td>0.054</td>
<td>0.916</td>
</tr>
<tr>
<td>BIC weighted</td>
<td>Simulated</td>
<td>0.061</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>B.A.</td>
<td>0.058</td>
<td>0.931</td>
</tr>
<tr>
<td></td>
<td>H.C.</td>
<td>0.072</td>
<td>0.969</td>
</tr>
<tr>
<td></td>
<td>Naive</td>
<td>0.054</td>
<td>0.914</td>
</tr>
</tbody>
</table>

- When $N=20$, the Burnham and Anderson S.E. are accurate, but the coverage is slightly low.
- The Hjort and Claeskens C.I. is slightly conservative.
When \( N=100 \), the Burnham and Anderson S.E. and C.I. is OK.

The Hjort and Claeskens C.I. is conservative.
Results: Correlation between estimates in sub-models

<table>
<thead>
<tr>
<th>Knot</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.00</td>
<td>0.95</td>
<td>0.89</td>
<td>0.87</td>
<td>0.87</td>
<td>0.85</td>
<td>0.79</td>
<td>0.72</td>
<td>0.64</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.94</td>
<td>0.91</td>
<td>0.90</td>
<td>0.88</td>
<td>0.82</td>
<td>0.73</td>
<td>0.64</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.98</td>
<td>0.96</td>
<td>0.94</td>
<td>0.88</td>
<td>0.80</td>
<td>0.70</td>
<td>0.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.99</td>
<td>0.97</td>
<td>0.93</td>
<td>0.85</td>
<td>0.76</td>
<td>0.66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.99</td>
<td>0.95</td>
<td>0.88</td>
<td>0.80</td>
<td>0.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.98</td>
<td>0.92</td>
<td>0.85</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.98</td>
<td>0.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

Correlations were high, suggesting the accuracy of the Burnham and Anderson S.E.
Summary

- PSA nadir should be estimated by a modeling approach in clinical trials.
- The P-spline method works well.
- The Burnham and Anderson S.E. and C.I. are reasonably accurate since modeling bias is small.
Let \( df = \text{Trace}\{H(\kappa_A, \kappa_B)\} \) be degree of freedom of smoother.

AIC and weights:

\[
AIC = \sum_{k=1}^K \left( \kappa_k \right) - df
\]

\[
\sum_{k=1}^K \left( \frac{1}{\kappa_k} \right) - df
\]

\[
\sum_{k=1}^K \left( \frac{1}{\kappa_k} \right) - df
\]
Information criteria

- BIC and weights
  \[ BIC = -2l(k_a, k_b) + \log(n)df \]
  \[ = \exp\{l(k_a, k_b) - \log(n)df/2\} \]
  \[ \sum_k \frac{\exp\{l(k_a, k_B) - \log(n)df/2\}}{\exp\{l(k_k, k_l) - \log(n)df/2\}} \]

- BIC is supposed to be a consistent model selector, but difference between AIC and BIC is expected trivial in small sample
**Example: \((\kappa_A, \kappa_B) = (4,32)\) and \((18,32)\)**

<table>
<thead>
<tr>
<th>Model ((\kappa_A, \kappa_B))</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log PSA Nadir ((4,32))</td>
<td>1.91</td>
<td>0.0614</td>
</tr>
<tr>
<td></td>
<td>(18,32)</td>
<td>1.69</td>
</tr>
<tr>
<td>Residual variance ((4,32))</td>
<td>0.107</td>
<td>0.027203</td>
</tr>
<tr>
<td></td>
<td>(18,32)</td>
<td>0.0605</td>
</tr>
</tbody>
</table>