Dynamic programming for changepoint processes

Dirk Husmeier

Stat Comput (2006) 16: 203–213
DOI 10.1007/s11222-006-8450-8

Exact and efficient Bayesian inference for multiple changepoint problems

Paul Fearnhead

Two priors for changepoints

Prior for the number of changepoints, $\pi(m),$ and a conditional prior on their positions: $\pi(\tau_j|\tau_{j+1}, m)$

Two priors for changepoints

Prior for the number of changepoints, $\pi(m)$, and a conditional prior on their positions: $\pi(\tau_j | \tau_{j+1}, m)$

Point process prior

$g_0(t)$ Probability for the first changepoint

$g(t)$ Probability for the time between two successive changepoints

$$G(t) = \sum_{s=1}^{t} g(s)$$

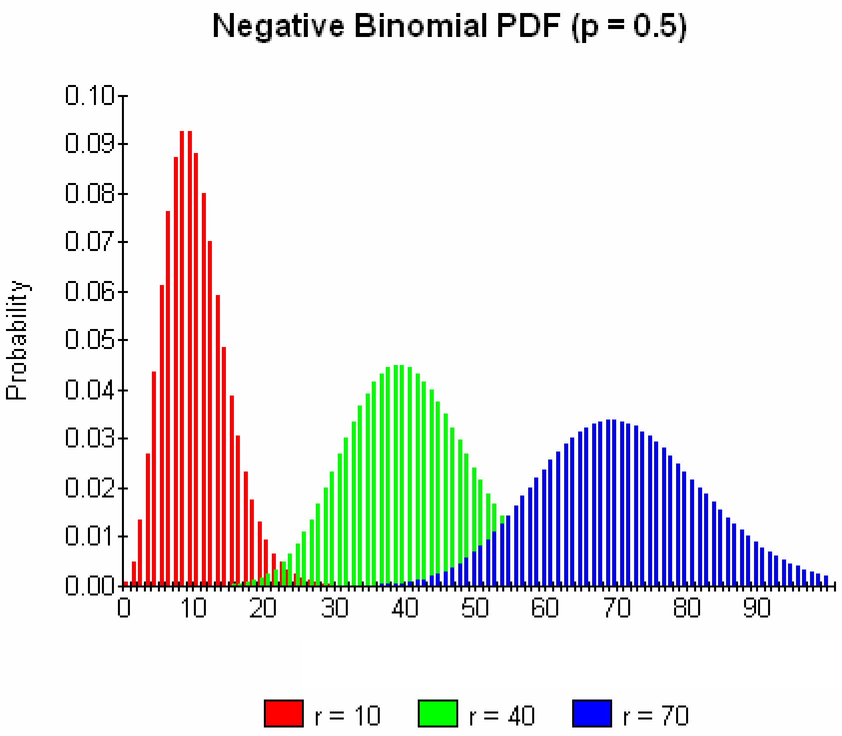
Distribution function for the distance between two successive changepoints

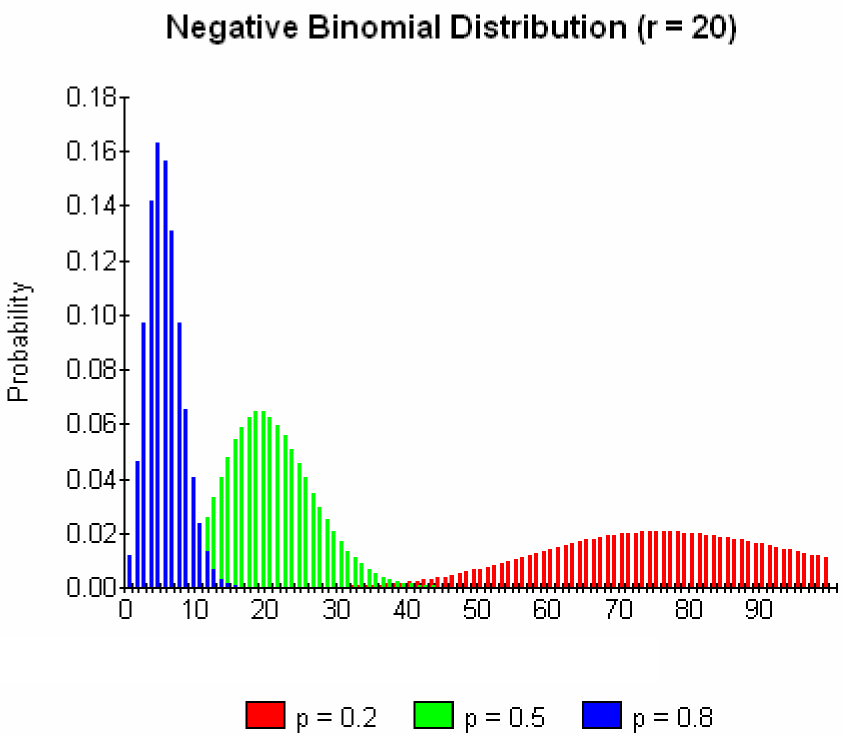
The probability of m changepoints occuring at $\tau_1, \ldots, \tau_m$ is

$$g_0(\tau_1)\left(\prod_{j=2}^{m} g(\tau_j - \tau_{j-1})\right)(1 - G(\tau_{m+1} - \tau_m))$$

Negative binomial distribution

$$g(t) = \binom{t-r}{r-1} p^r (1-p)^{t-r}$$



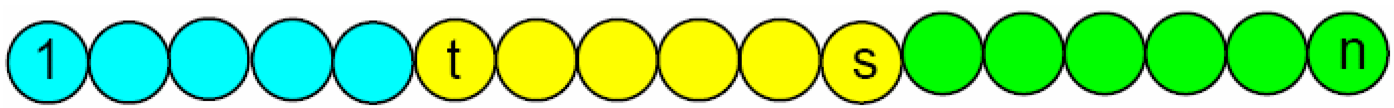


Definitions

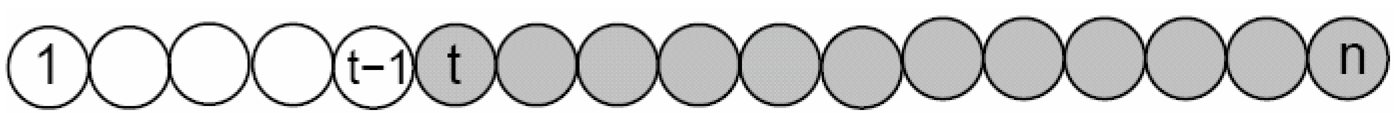
$$P(t, s) = \Pr(y_{t:s} | t, s \text{ in the same segment})$$
$$= \int \prod_{i=t}^{s} f(y_i|\theta)\pi(\theta)\mathrm{d}\theta.$$

Assumes parameters can be integrated out in closed form

Prior



$$Q(t) = \Pr(y_{t:n} | \text{changepoint at } t - 1)$$



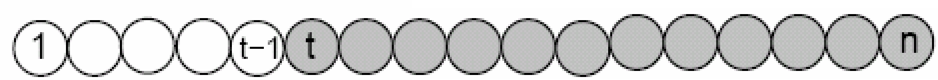
Recursion

$$Q(t) = \sum_{s=t}^{n-1} P(t,s)Q(s+1)g(s+1-t) + P(t,n)(1-G(n-t)),$$

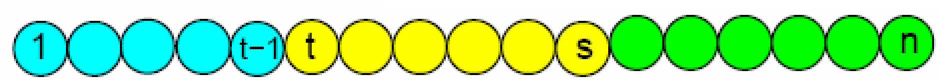
Proof

For notational convenience we drop the explicit conditioning on a changepoint at $t-1$ in the following. Thus,

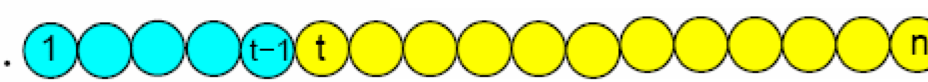
$$Q(t) = \Pr(y_{t:n})$$



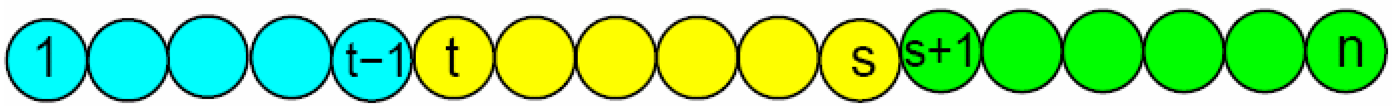
$$= \sum_{s=t}^{n-1} \Pr(y_{t:n}, \text{ next changepoint at } s)$$



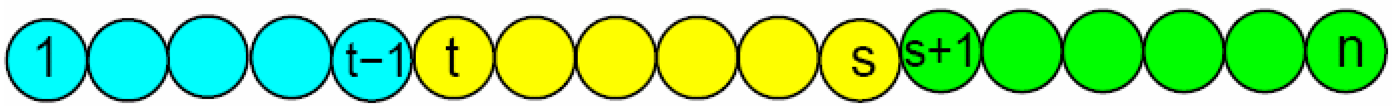
$$+ \Pr(y_{t:n}, \text{ no further changepoints}).$$



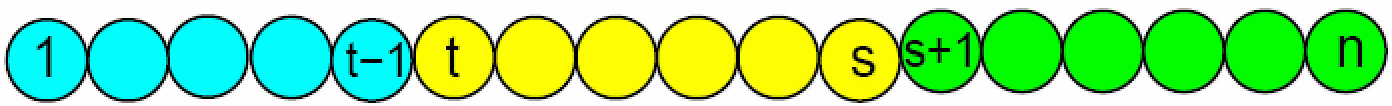
$$\Pr(y_{t:n}, \text{next changepoint at } s)$$
$$= \Pr(\text{next changepoint at } s)$$
$$\Pr(y_{t:s}, y_{s+1:n} | \text{next changepoint at } s)$$



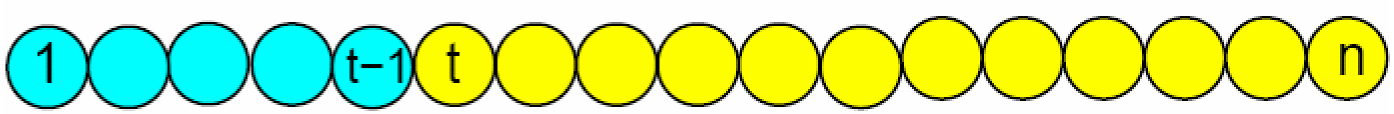
$$
\begin{aligned}
&\Pr(y_{t:n}, \text{next changepoint at } s) \\
&\quad = \Pr(\text{next changepoint at } s) \\
&\qquad \Pr(y_{t:s}, y_{s+1:n} | \text{next changepoint at } s) \\
&\quad = g(s+1-t) \Pr(y_{t:s} | t, s \text{ in same segment}) \\
&\qquad \Pr(y_{s+1:n} | \text{changepoint at } s) \\
&\quad = g(s+1-t) P(t,s) Q(s+1)
\end{aligned}
$$



$$\begin{aligned}
&\Pr(y_{t:n}, \text{next changepoint at } s) \\
&\quad = \Pr(\text{next changepoint at } s) \\
&\quad\quad \Pr(y_{t:s}, y_{s+1:n} | \text{next changepoint at } s) \\
&\quad = g(s+1-t) \Pr(y_{t:s} | t, s \text{ in same segment}) \\
&\quad\quad \Pr(y_{s+1:n} | \text{changepoint at } s) \\
&\quad = g(s+1-t) P(t, s) Q(s+1)
\end{aligned}$$



$$\Pr(y_{t:n}, \text{ no further changepoints}) = P(t, n) \cdot (1 - G_0(n-t))$$

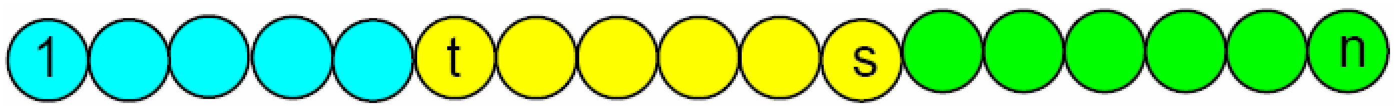


Reminder …

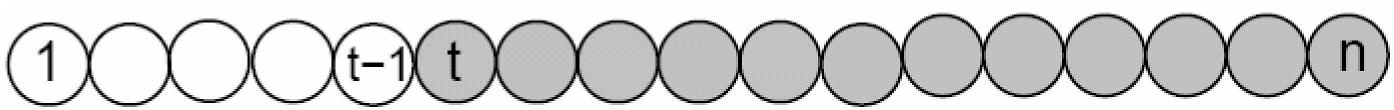
$$P(t,s) = \Pr(y_{t:s} \mid t, s \text{ in the same segment})$$
$$= \int \prod_{i=t}^{s} f(y_i \mid \theta)\pi(\theta)\mathrm{d}\theta.$$

Assumes parameters can be integrated out in closed form

Prior



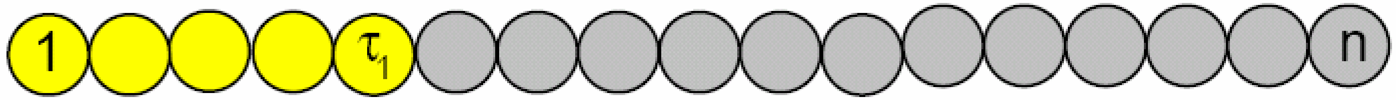
$$Q(t) = \Pr(y_{t:n} \mid \text{changepoint at } t-1)$$



Given the values of $Q(t)$ for $t = 1, \ldots, n$ it is straightforward to simulate from the posterior distribution of the changepoints as follows.

The posterior distribution of the first changepoint is given by

$$
\begin{aligned}
\Pr(\tau_1 | y_{1:n}) &= \Pr(y_{1:n}, \tau_1) / \Pr(y_{1:n}) \\
&= \Pr(\tau_1) \Pr(y_{1:\tau_1} | \tau_1) \Pr(y_{\tau_1+1:n} | \tau_1) / Q(1)
\end{aligned}
$$

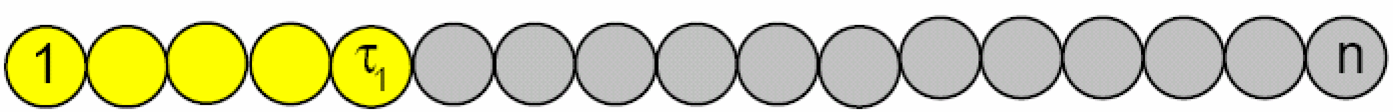


Given the values of $Q(t)$ for $t = 1, \ldots, n$ it is straightforward to simulate from the posterior distribution of the change-points as follows.

The posterior distribution of the first changepoint is given by

$$
\begin{aligned}
\Pr(\tau_1 | y_{1:n}) &= \Pr(y_{1:n}, \tau_1) / \Pr(y_{1:n}) \\
&= \Pr(\tau_1) \Pr(y_{1:\tau_1} | \tau_1) \Pr(y_{\tau_1+1:n} | \tau_1) / Q(1) \\
&= P(1, \tau_1) Q(\tau_1 + 1) g_0(\tau_1) / Q(1),
\end{aligned}
$$

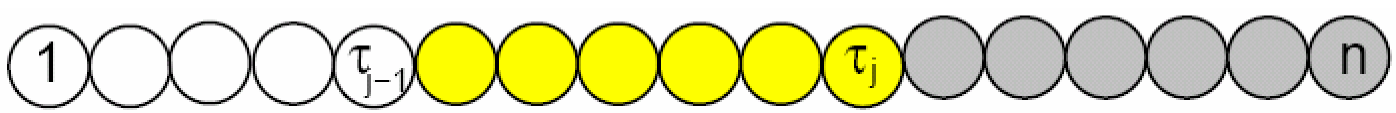
for $\tau_1 = 1, \ldots, n - 1$. The probability of no further change-point being $P(1, n)(1 - G_0(n - 1))/Q(1)$.



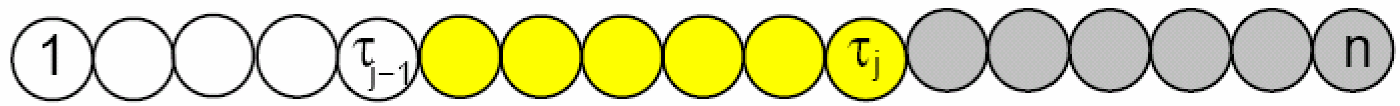
Similarly the posterior distribution of the τ_j given τ_{j-1} is

$$\Pr(\tau_j|\tau_{j-1}, y_{1:n}) = P(\tau_{j-1}+1, \tau_j)Q(\tau_j+1) \times g(\tau_j - \tau_{j-1})/Q(\tau_{j-1}+1),$$

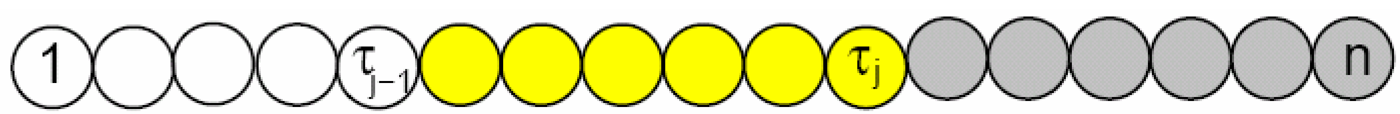
for $\tau_j = \tau_{j-1}+1, \ldots, n-1$, and the probability of no further breakpoint is $P(\tau_{j-1}+1, n)(1 - G_0(n - \tau_{j-1} - 1))/Q(\tau_{j-1}+1)$.



$$
\begin{aligned}
Pr(\tau_j|\tau_{j-1}, y_{1:n}) &= Pr(\tau_j, \tau_{j-1}, y_{1:n})/Pr(\tau_{j-1}, y_{1:n}) \\
&= Pr(y_{1:n}|\tau_j, \tau_{j-1})Pr(\tau_j|\tau_{j-1})Pr(\tau_{j-1})/Pr(y_{1:n}|\tau_{j-1})Pr(\tau_{j-1}) \\
&= Pr(y_{1:n}|\tau_j, \tau_{j-1})Pr(\tau_j|\tau_{j-1})/Pr(y_{1:n}|\tau_{j-1})
\end{aligned}
$$



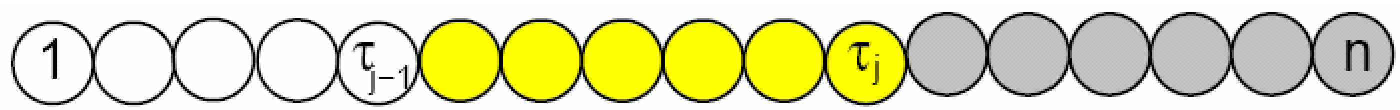
$$
\begin{aligned}
Pr(\tau_j|\tau_{j-1}, y_{1:n}) &= Pr(\tau_j, \tau_{j-1}, y_{1:n})/Pr(\tau_{j-1}, y_{1:n}) \\
&= Pr(y_{1:n}|\tau_j, \tau_{j-1})Pr(\tau_j|\tau_{j-1})Pr(\tau_{j-1})/Pr(y_{1:n}|\tau_{j-1})Pr(\tau_{j-1}) \\
&= Pr(y_{1:n}|\tau_j, \tau_{j-1})Pr(\tau_j|\tau_{j-1})/Pr(y_{1:n}|\tau_{j-1})
\end{aligned}
$$



$$
\begin{aligned}
Pr(y_{1:n}|\tau_{j-1}) &= Pr(y_{1:\tau_{j-1}}|\tau_{j-1})Pr(y_{\tau_{j-1}+1:n}|\tau_{j-1}) \\
&= P(1, \tau_{j-1})Q(\tau_{j-1}+1)
\end{aligned}
$$

$$
\begin{aligned}
Pr(\tau_j|\tau_{j-1}, y_{1:n}) &= Pr(\tau_j, \tau_{j-1}, y_{1:n})/Pr(\tau_{j-1}, y_{1:n}) \\
&= Pr(y_{1:n}|\tau_j, \tau_{j-1})Pr(\tau_j|\tau_{j-1})Pr(\tau_{j-1})/Pr(y_{1:n}|\tau_{j-1})Pr(\tau_{j-1}) \\
&= Pr(y_{1:n}|\tau_j, \tau_{j-1})Pr(\tau_j|\tau_{j-1})/Pr(y_{1:n}|\tau_{j-1})
\end{aligned}
$$

$$
\begin{aligned}
Pr(y_{1:n}|\tau_j, \tau_{j-1}) &= Pr(y_{1:\tau_{j-1}}|\tau_{j-1})Pr(y_{\tau_{j-1}+1:\tau_j}|\tau_{j-1}, \tau_j)Pr(y_{\tau_j+1:n}|\tau_j) \\
&= P(1, \tau_{j-1})P(\tau_{j-1}+1, \tau_j)Q(\tau_j+1)
\end{aligned}
$$

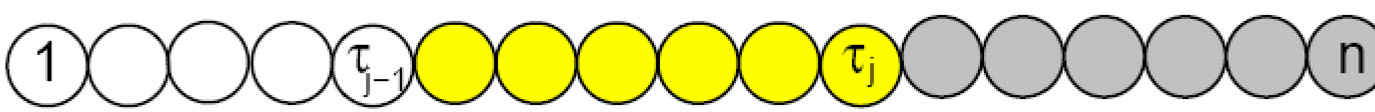


$$
\begin{aligned}
Pr(y_{1:n}|\tau_{j-1}) &= Pr(y_{1:\tau_{j-1}}|\tau_{j-1})Pr(y_{\tau_{j-1}+1:n}|\tau_{j-1}) \\
&= P(1, \tau_{j-1})Q(\tau_{j-1}+1)
\end{aligned}
$$

$$
\begin{aligned}
Pr(\tau_j|\tau_{j-1}, y_{1:n}) &= Pr(\tau_j, \tau_{j-1}, y_{1:n})/Pr(\tau_{j-1}, y_{1:n}) \\
&= Pr(y_{1:n}|\tau_j, \tau_{j-1})Pr(\tau_j|\tau_{j-1})Pr(\tau_{j-1})/Pr(y_{1:n}|\tau_{j-1})Pr(\tau_{j-1}) \\
&= Pr(y_{1:n}|\tau_j, \tau_{j-1})Pr(\tau_j|\tau_{j-1})/Pr(y_{1:n}|\tau_{j-1})
\end{aligned}
$$

$$
\begin{aligned}
Pr(\tau_j|\tau_{j-1}, y_{1:n}) &= P(\tau_{j-1}+1, \tau_j)Q(\tau_j+1)Pr(\tau_j|\tau_{j-1})/Q(\tau_{j-1}+1) \\
&= P(\tau_{j-1}+1, \tau_j)Q(\tau_j+1)g(\tau_j-\tau_{j-1})/Q(\tau_{j-1}+1)
\end{aligned}
$$

$$
\begin{aligned}
Pr(y_{1:n}|\tau_j, \tau_{j-1}) &= Pr(y_{1:\tau_{j-1}}|\tau_{j-1})Pr(y_{\tau_{j-1}+1:\tau_j}|\tau_{j-1}, \tau_j)Pr(y_{\tau_j+1:n}|\tau_j) \\
&= P(1, \tau_{j-1})P(\tau_{j-1}+1, \tau_j)Q(\tau_j+1)
\end{aligned}
$$

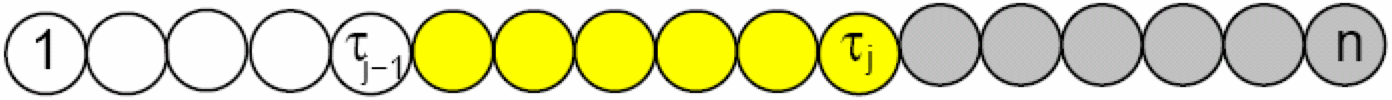


$$
\begin{aligned}
Pr(y_{1:n}|\tau_{j-1}) &= Pr(y_{1:\tau_{j-1}}|\tau_{j-1})Pr(y_{\tau_{j-1}+1:n}|\tau_{j-1}) \\
&= P(1, \tau_{j-1})Q(\tau_{j-1}+1)
\end{aligned}
$$

Similarly the posterior distribution of the τ_j given τ_{j-1} is

$$\Pr(\tau_j|\tau_{j-1}, y_{1:n}) = P(\tau_{j-1}+1, \tau_j)Q(\tau_j+1) \times g(\tau_j - \tau_{j-1})/Q(\tau_{j-1}+1),$$

for $\tau_j = \tau_{j-1}+1, \ldots, n-1$, and the probability of no further breakpoint is $P(\tau_{j-1}+1, n)(1 - G_0(n - \tau_{j-1} - 1))/Q(\tau_{j-1}+1)$.



Summary

Definition

$$Q(t) = \Pr(y_{t:n} | \text{changepoint at } t-1)$$

Recursion

$$Q(t) = \sum_{s=t}^{n-1} P(t,s)Q(s+1)g(s+1-t)$$
$$+P(t,n)(1-G(n-t)),$$

Sampling of changepoints

$$\Pr(\tau_j | \tau_{j-1}, y_{1:n}) =$$

$$P(\tau_{j-1}+1, \tau_j)Q(\tau_j+1)$$
$$\times g(\tau_j - \tau_{j-1})/Q(\tau_{j-1}+1).$$

Point process prior

Definition

$$Q(t) = \Pr(y_{t:n} | \text{changepoint at } t-1)$$

Recursion

$$Q(t) = \sum_{s=t}^{n-1} P(t,s)Q(s+1)g(s+1-t)$$
$$+P(t,n)(1-G(n-t)),$$

Sampling of changepoints

$$\Pr(\tau_j | \tau_{j-1}, y_{1:n}) =$$
$$P(\tau_{j-1}+1, \tau_j)Q(\tau_j+1)$$
$$\times g(\tau_j - \tau_{j-1})/Q(\tau_{j-1}+1).$$

Prior for the number of changepoints, and a conditional prior on their positions

Definition

$$Q_j^{(m)}(t) = \Pr(y_{t:n} | \tau_j = t-1, m \text{ changepoints})$$

Recursion

$$Q_m^{(m)}(t) = P(t,n)\pi_m(\tau_m = t-1)$$
For $j = 1, \ldots, m-1$
$$Q_j^{(m)}(t) = \sum_{s=t}^{n-m+j} P(t,s)Q_{j+1}^{(m)}(s+1)$$
$$\times \pi_m(\tau_j = t-1 | \tau_{j+1} = s)$$

Sampling of changepoints

$$\Pr(\tau_j | \tau_{j-1}, y_{1:n}, m) =$$
$$P(\tau_{j-1}+1, \tau_j)Q_j^{(m)}(\tau_j+1)$$
$$\times \pi_m(\tau_{j-1} | \tau_j)/Q_{j-1}^{(m)}(\tau_{j-1}+1)$$