Regression model with an unknown changepoint. Suppose that the density for $y_t$, $t = 1, 2, \ldots, T$, can be expressed as

$$
(y_t|x_t, \gamma) \sim \begin{cases} N(\theta_1 + \theta_2 x_t, \sigma^2) & t \leq \lambda \\ N(\beta_1 + \beta_2 x_t, \tau^2) & t > \lambda \end{cases},
$$

where $\gamma = (\theta_1, \theta_2, \beta_1, \beta_2, \sigma^2, \tau^2, \lambda)'$ and that you employ priors of the form $(\theta_1, \theta_2)' \sim N(\mu_0, V_0)$, $(\beta_1, \beta_2)' \sim N(\mu_\beta, V_\beta)$, $\sigma^2 \sim IG(a_1, b_1)$, $\tau^2 \sim IG(a_2, b_2)$, and $\lambda \sim U\{1, 2, \ldots, T - 1\}$.

Let $X$ be the $n \times 2$ matrix with rows $(1, x_t)$ and $y = (y_1, \ldots, y_n)'$. Also, let $X_{a:b}$ the matrix with rows $a$ to $b$ from $X$, similarly for $y_{a:b}$. Finally, let $\theta = (\theta_1, \theta_2)'$ and $\beta = (\beta_1, \beta_2)'$. The Gibbs sampler scans through the following 5 steps.

**STEP 1:** Sample $\theta$ from $N(\bar{\mu}_\theta, V_\theta)$, where

$$
\bar{V}_\theta^{-1} = V_\theta^{-1} \times X'_{1:1} X_{1:1} / \sigma^2 \quad \text{and} \quad \bar{\bar{\mu}}_\theta = V_\theta^{-1} \mu_\theta + X'_{1:1} y_{1:1} / \sigma^2.
$$

**STEP 2:** Sample $\sigma^2$ from $IG(\bar{a}_1, \bar{b}_1)$, where

$$
\bar{a}_1 = a_1 + \tau / 2 \quad \text{and} \quad \bar{b}_1 = b_1 + (y_{1:1} - X_{1:1} \theta)'(y_{1:1} - X_{1:1} \theta) / 2.
$$

**STEP 3:** Sample $\beta$ from $N(\bar{\mu}_\beta, V_\beta)$, where

$$
\bar{V}_\beta^{-1} = V_\beta^{-1} \times X'_{\lambda:1:n} X_{\lambda:1:n} / \tau^2 \quad \text{and} \quad \bar{\bar{\mu}}_\beta = V_\beta^{-1} \mu_\beta + X'_{\lambda:1:n} y_{\lambda:1:n} / \tau^2.
$$

**STEP 4:** Sample $\sigma^2$ from $IG(\bar{a}_2, \bar{b}_2)$, where

$$
\bar{a}_2 = a_2 + (n - \tau) / 2 \quad \text{and} \quad \bar{b}_2 = b_2 + (y_{\lambda:1:n} - X_{\lambda:1:n} \beta)'(y_{\lambda:1:n} - X_{\lambda:1:n} \beta) / 2.
$$

**STEP 5:** Sample $\lambda$ from $\{1, \ldots, T - 1\}$ with probabilities

$$
Pr(\lambda|\beta_1, \beta_2, \sigma_1^2, \sigma_2^2, y, x) \propto \prod_{t=1}^{\lambda} P_N(y_t; \beta_1 x_t, \sigma_1^2) \prod_{t=\lambda+1}^{T} P_N(y_t; \beta_2 x_t, \sigma_2^2).
$$
**Simulation.** Generate $T = 500$ observations from the changepoint model $y_t|x_t \sim N(2 + x_t, 0.2)$ for $t \leq 85$ and $N(1.5 + 0.8x_t, 0.5)$, for $t > 85$, where $x_t \sim N(0, 1)$ . Select prior hyperparameters as follows: $\mu_\theta = \mu_\beta = 0_2$, $V_\theta = V_\beta = 100I_2$, $a_1 = b_1 = 3$, and $a_2 = b_2 = 0.5$. sample from the joint posterior, $p(\theta_1, \theta_2, \beta_1, \beta_2, \sigma^2, \tau^2, \lambda|y, x)$, via a Gibbs sampler.

**Solution:** See file `finalexam-simulation-R.pdf`.

**Real data.** The data below represents yearly homicide rates, $y_t$, of young men between 15 and 29 years of age (in 100,000 inhabitants) in the city of São Paulo from 1980 to 2009 ($T = 30$).

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Median</th>
<th>95% Credibility Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>88.7</td>
<td>8.7</td>
<td>88.7</td>
<td>[72.3;105.7]</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>8.4</td>
<td>0.7</td>
<td>8.4</td>
<td>[7.0;9.7]</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>18.6</td>
<td>2.7</td>
<td>18.3</td>
<td>[14.3;24.5]</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>847.6</td>
<td>92.6</td>
<td>849.3</td>
<td>[644.4;1031.6]</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-27.4</td>
<td>3.5</td>
<td>-27.4</td>
<td>[-34.2;-19.8]</td>
</tr>
<tr>
<td>$\tau$</td>
<td>24.5</td>
<td>7.2</td>
<td>23.3</td>
<td>[14.6;41.6]</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>2000</td>
<td>1.1</td>
<td>2000</td>
<td>[1998;2001]</td>
</tr>
</tbody>
</table>

MLE of the two regressions, when $\hat{\lambda} = 1998$ is assumed to be the change point, are given by $y_t = 90 + 8x_t$ ($\hat{\sigma} = 20$), for $t \leq 1998$ and $y_t = 822 - 26x_t$ ($\hat{\tau} = 21$), for $t > 1998$, where $x_t = t - 1979$. Assuming the same priors as above, compute the posterior means, standard deviations, medians and 95% posterior credibility intervals of $\theta_1, \theta_2, \sigma, \beta_1, \beta_2, \tau$ and $\lambda$.

**Solution:** See table below and file `finalexam-realdata-R.pdf`.