patient	before (X)	after (Y)	patient	before (X)	after (Y)
1	102.4	99.6	7	102.5	101.0
2	103.2	100.1	8	103.1	100.1
3	101.9	100.2	9	102.8	100.7
4	103.0	101.1	10	102.3	101.1
5	101.2	99.8	11	101.9	101.3
6	100.7	100.2	12	101.4	100.2

Table 1: Temperatures of $n = 12$ children before and after taking aspirin.

In a pediatric clinical study carried out to see how effective aspirin is in reducing temperature, $n = 12$ five-year-old children suffering from influenza had their temperature taken immediately before (X) and 1 hour after (Y) administration of aspirin. The data are in Table 1, ~~and online at `www.biostat.umn.edu/~brad/data/aspirin_data.txt`.~~

Let $Z = X - Y$, the observed reduction in temperature. The study investigators suspect that the reduction in temperature Z may be related to the initial temperature X. Thus, they would like to fit the simple linear regression (SLR) model,

$$Z_i = \beta_0 + \beta_1(x_i - \bar{x}) + \epsilon_i,\ \epsilon_i \stackrel{iid}{\sim} N(0, \sigma^2),\ i = 1, \ldots, n.$$

(a) Use `R` (or any package) to plot the raw data. Does the linear model appear plausible?

(b) Following the model of Example 2.10, fit the SLR model in `WinBUGS`, using vague priors for all parameters. Find a posterior density estimate and a 95% credible interval for the initial temperature effect β_1. Do your results confirm the investigators' suspicion? Also check the effective model size p_D and DIC score for the fitted SLR.

(c) Suppose a 13th patient arrives with an initial temperature of 100.0 degrees. Find an estimate of the predictive density of Z_{13}, the predicted reduction in temperature for this patient, and a point estimate for $P(Z_{13} > 0|Z_1, \ldots, Z_{12})$, the predictive probability that this new patient's temperature is reduced by aspirin.

(d) Follow the model of CL3, Example 2.16 to investigate approximate residual and CPO values for your model (do not bother with exact cross-validatory results here). Are any data points outliers in any sense?

(e) Suppose we attempt to improve model fit by adding a quadratic term, $\beta_2(x_i - \bar{x})^2$, to our model. Make this change and recheck p_D and DIC, as well as the posterior of β_2. Is this model enhancement well-justified by the data? Also check this change's impact on your answer to part (c) above. Is either Z_{13} prediction well-justified by the data?

Working w/ Deviance

Deviance statistic: $-2\log p(\vec{y}\,|\,\vec{\theta}) + 2\log h(y)$

effective $p_D = \bar{D} - D(\bar{\theta})$

↳ where $D(\bar{\theta})$ is deviance statistic evaluated at mean of posterior samples of model parameters

I think if $Y_i \sim N(\beta_0 + \beta_1 x, \sigma^2)$ then average of β_0, β_1, and σ posterior samples should be used.

↳ and where $\bar{D}$ is defined as $E_{\theta|y}[D]$, which is posterior expectation of the deviance

So maybe evaluate deviance using each sampled value of β_0, β_1 σ and take the average over this deviance distribution