Chapter 11.9 A Conditional Means Prior

Jim Albert and Monika Hu

Chapter 11 Simple Linear Regression

Recap

- Two methods for constructing a prior on the parameters of a regression model so far:
  - a weakly informative prior (on the original data)
  - an informative prior (on the standardized data)
- A new approach
  - on the original data
  - instead of a prior on $(\beta_0, \beta_1, \sigma)$
  - prior about the expected response value conditional on specific values of the predictor variable

Learning about a gas bill from the outside temperature

- Can one accurately predict one's monthly natural gas bill from the outside temperature?
- Outcome / response variable: montly natural gas bill
- Explanatory / predictor variable: outside temperature

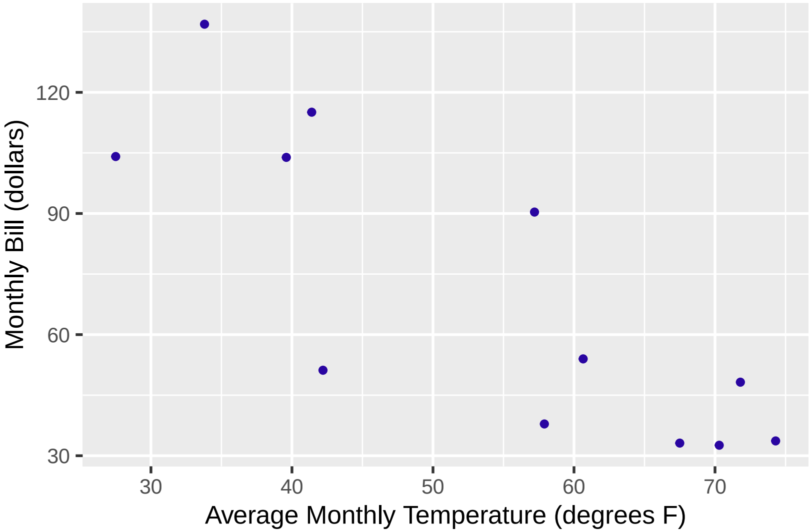
Learning about a gas bill from the outside temperature

- Can one accurately predict one's monthly natural gas bill from the outside temperature?
- Outcome / response variable: montly natural gas bill
- Explanatory / predictor variable: outside temperature

$$Y_i \mid \beta_0, \beta_1, \sigma \sim \mathrm{Normal}(\beta_0 + \beta_1 x_i, \sigma) \tag{1}$$

- x_i and y_i are respectively the average temperature (degrees in Fahrenheit) and the bill amount (in dollars) in month i
- $(\beta_0, \beta_1, \sigma)$ are the unknown regression parameters

Learning about a gas bill from the outside temperature cont'd



A conditional means prior

- Assume independence of parameters, as before

$$\pi(\beta_0, \beta_1, \sigma) = \pi(\beta_0, \beta_1)\pi(\sigma)$$

- Think about formulating prior opinion about the mean values

$$\mu_i^* = \beta_0 + \beta_1 x_i^* \qquad (2)$$

- For two specified values of the predictor x_1^* and x_2^*

A conditional means prior cont'd

Two steps:

1. For the first predictor value x_1^* construct a normal prior for the mean value μ_1^*
   - let the mean and standard deviation values of this prior be denoted by m_1 and s_1, respectively
2. Similarly, for the second predictor value x_2^* construct a normal prior for the mean value μ_2^*
   - let the mean and standard deviation values of this prior be denoted by m_2 and s_2

A conditional means prior cont'd

- Under independence, the joint prior for the vector (μ_1^*, μ_2^*) has the form

$$\pi(\mu_1^*, \mu_2^*) = \pi(\mu_1^*)\pi(\mu_2^*)$$

A conditional means prior cont'd

- Under independence, the joint prior for the vector (μ_1^*, μ_2^*) has the form

$$\pi(\mu_1^*, \mu_2^*) = \pi(\mu_1^*)\pi(\mu_2^*)$$

- This prior on the two conditional means implies a bivariate normal prior on the regression parameters

$$\beta_1 = \frac{\mu_2^* - \mu_1^*}{x_2 - x_1} \tag{3}$$

$$\beta_0 = \mu_1^* - x_1 \left(\frac{\mu_2^* - \mu_1^*}{x_2 - x_1} \right) \tag{4}$$

- The slope β_0 and β_1 are linear functions of the two conditional means μ_1^* and μ_2^*, implying a bivariate normal distribution for β_0 and β_1

Regression analysis of the gas bill example

- Consider two different temperature values, say 40 degrees and 60 degrees
- If $x = 40$, the mean bill $\mu_1^* = \beta_0 + \beta_1(40)$
  - a normal prior with mean \$100 and standard deviation \$20
  - we believe the average gas bill will be relatively high during a cold month averaging 40 degrees
- If $x = 60$, the mean bill $\mu_2^* = \beta_0 + \beta_1(100)$
  - a normal with mean \$50 and standard deviation \$15
  - we believe the gas cost will average \$50 lower than in the first scenario

Regression analysis of the gas bill example

- Consider two different temperature values, say 40 degrees and 60 degrees
- If $x = 40$, the mean bill $\mu_1^* = \beta_0 + \beta_1(40)$
  - a normal prior with mean \$100 and standard deviation \$20
  - we believe the average gas bill will be relatively high during a cold month averaging 40 degrees
- If $x = 60$, the mean bill $\mu_2^* = \beta_0 + \beta_1(100)$
  - a normal with mean \$50 and standard deviation \$15
  - we believe the gas cost will average \$50 lower than in the first scenario

- Under independence

$$\pi(\mu_1^*, \mu_2^*) = \phi(\mu_1^*, 100, 20)\phi(\mu_2^*, 50, 15), \tag{5}$$

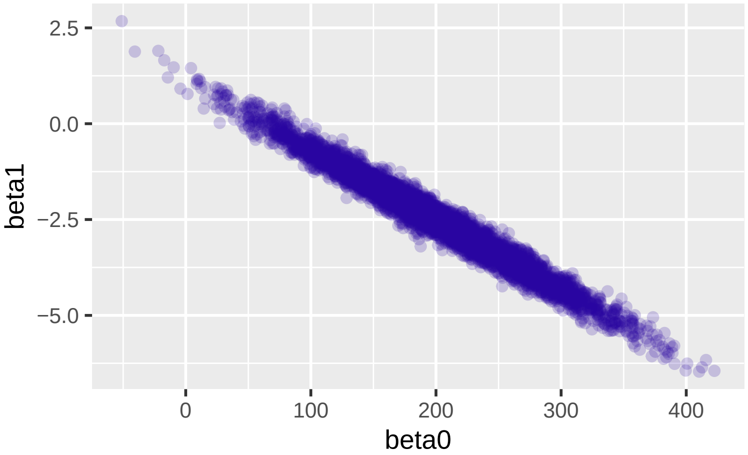
where $\phi(y, \mu, \sigma)$ denotes the normal density with mean μ and standard deviation σ

Regression analysis of the gas bill example cont'd

- The prior on the two means is an indirect way of assessing a prior on the regression parameters β_0 and β_1
- We can simulate pairs (β_0, β_1) from the prior distribution by simulating values of the means μ_1^* and μ_2^* from independent normal distributions and applying Equation (3) and Equation (4)

```
modelString = "
model{
beta1 <- (mu2 - mu1) / (x2 - x1)
beta0 <- mu1 - x1 * (mu2 - mu1) / (x2 - x1)
mu1 ~ dnorm(m1, s1)
mu2 ~ dnorm(m2, s2)
}"
```

Regression analysis of the gas bill example cont'd



- 1000 simulated draws of (β_0, β_1)
- The implied prior on the regression coefficients indicates that β_0 and β_1 are strongly negatively correlated

Regression analysis of the gas bill example cont'd

- Don't forget to assign the precision parameter $\phi = 1/\sigma^2$ a gamma prior with parameters a and b

$$\pi(\beta_0, \beta_1, \sigma) = \pi_{CM}(\beta_0, \beta_1)\pi(\sigma),$$

where π_{CM} is the conditional means prior

Regression analysis of the gas bill example cont'd

- Don't forget to assign the precision parameter $\phi = 1/\sigma^2$ a gamma prior with parameters *a* and *b*

$$\pi(\beta_0, \beta_1, \sigma) = \pi_{CM}(\beta_0, \beta_1)\pi(\sigma),$$

where π_{CM} is the conditional means prior

- Use JAGS for posterior inference
- Compare inferences using conditional means and weakly informative priors