Chapter 13: Regression

강의 목표

- 선형 회귀모형
- 최소제곱 추정치
- 선형회귀모형의 베이지안 추론

Regression

- 관심 있는 변수 Y가 있을 때 Y에 대한 추론을 위하여 관련 있는 변수들 $X_1, ..., X_p$에 관한 정보를 이용하는 것이 도움.

$$CarPrice = \beta_0 + \beta_1 Year + \beta_2 CarType + \beta 3 Mileage + \epsilon$$

- Y: 종속변수
- X: 독립변수

선형회귀모형

- 선형 회귀 모형:

$$Y = \beta_0 + \beta_1 X_1 + ... + \beta_p X_p + \epsilon.$$

- ϵ 에 대한 가정: $\epsilon \sim N(0, \sigma^2)$.
  - Normality
  - Constant variance
- 조건부 분포:

$$Y \mid X, \beta, \sigma^2 \sim N(\beta 0 + \beta_1 X_1 + ... + \beta_p X_p, \sigma^2).$$

- 결합밀도함수:

$$f(y \mid \beta, \sigma^2) = \frac{1}{\sqrt{(2\pi\sigma^2)^n}} \exp\left\{-\frac{1}{2\sigma^2}(y - X\beta)^T(y - X\beta)\right\}.$$

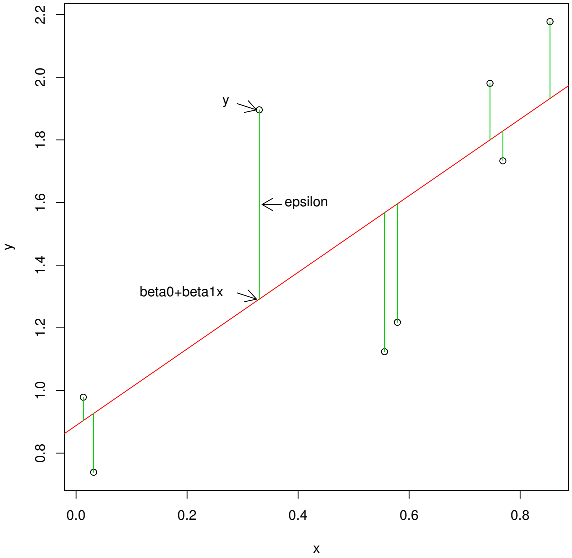
Simple Linear Regression

- One criterion is **least squares**:

$$\min_{\beta_0,\beta_1} \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

- It minimizes vertical distances between observations and the fitted line.

Least Squares Estimate



Estimating β_0, β_1

Differentiate the criterion with respect to β_0, β_1 and set the derivatives equal to 0, we get:

$$
\begin{aligned}
\frac{\partial}{\partial \beta_0} &= (-2)\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) = 0 \\
\frac{\partial}{\partial \beta_1} &= (-2)\sum_{i=1}^{n} x_i(y_i - \beta_0 - \beta_1 x_i) = 0
\end{aligned}
$$

Solving for β_0 and β_1, we have:

$$
\begin{aligned}
\hat{\beta}_1 &= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
\hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \bar{x}
\end{aligned}
$$

"Hat" notation is used for estimates.

최소제곱 추정치

- 잔차제곱합을 최소화 시키는 방법:

$$\hat{\beta}_{\mathrm{LSE}} = \arg\min_{\beta} SSR(\beta) = \arg\min_{\beta} (Y - X\beta)(Y - X\beta)^T.$$

- Estimator:

$$\hat{\beta}_{\mathrm{LSE}} = (X^T X)^{-1} X^T Y$$

- 이는 우도함수를 최대로 만드는 최대우도 추정치이기도 하다.

선형 회귀모형의 베이지안 추론

- 선형 회귀모형의 모수 $\boldsymbol{\beta}$와 σ^2의 사전분포:

$$\boldsymbol{\beta} \sim N(\boldsymbol{\beta_0}, \Sigma_0), \quad \sigma^2 \sim IG(a, b)$$

- $\boldsymbol{\beta}$와 σ^2의 결합사후 밀도함수:

$$\begin{aligned}
\pi(\boldsymbol{\beta}, \sigma^2 \mid \boldsymbol{y}) \quad \propto \quad & (\sigma^2)^{-n/2} exp \left\{ -\frac{1}{2\sigma^2} (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^T (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) \right\} \\
& \times \exp \left\{ -\frac{1}{2} (\boldsymbol{\beta} - \boldsymbol{\beta_0})^T \Sigma_0^{-1} (\boldsymbol{\beta} - \boldsymbol{\beta_0}) \right\} \\
& \times (\sigma^2)^{-a-1} e^{-b/\sigma^2}
\end{aligned}$$

선형 회귀모형의 베이지안 추론

- $\boldsymbol{\beta}$가 주어졌을 때 σ^2의 조건부 사후분포:

$$\begin{aligned}
\pi(\sigma^2 \mid \boldsymbol{\beta}, \boldsymbol{y}) &\propto (\sigma^2)^{-n/2-a-1} exp\left[-\frac{1}{\sigma^2}\left\{\frac{1}{2}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})+b\right\}\right] \\
&\sim IG\left(\frac{n}{2}+a, \frac{1}{2}SSR(\boldsymbol{\beta})+b\right)
\end{aligned}$$

선형 회귀모형의 베이지안 추론

- σ^2이 주어졌을 때 β의 조건부 사후분포:

$$
\begin{aligned}
\pi(\boldsymbol{\beta} \mid \sigma^2, \boldsymbol{y}) &\propto \frac{\pi(\boldsymbol{\beta}, \sigma^2 \mid \boldsymbol{y})}{\pi(\sigma^2 \mid \boldsymbol{y})} \\
&\propto exp\left[-\frac{1}{2}\left\{\frac{1}{\sigma^2}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} - 2\frac{1}{\sigma^2}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{y}\right\} - \frac{1}{2}\left\{\boldsymbol{\beta}'\Sigma_0^{-1}\boldsymbol{\beta} - 2\boldsymbol{\beta}'\Sigma_0^{-1}\boldsymbol{\beta_0}\right\}\right] \\
&\propto exp\left[-\frac{1}{2}\left\{\boldsymbol{\beta}'\left(\frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{X} + \Sigma_0^{-1}\right)\boldsymbol{\beta} - 2\boldsymbol{\beta}'\left(\frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{y} + \Sigma_0^{-1}\boldsymbol{\beta_0}\right)\right\}\right] \\
&\propto exp\left[-\frac{1}{2}\left\{\boldsymbol{\beta}'\Sigma_\beta^{-1}\boldsymbol{\beta} - 2\boldsymbol{\beta}'\Sigma_\beta^{-1}\boldsymbol{\mu_\beta}\right\}\right] \\
&\sim N(\boldsymbol{\mu_\beta}, \Sigma_\beta)
\end{aligned}
$$

Where

$$
\Sigma_\beta = \left(\frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{X} + \Sigma_0^{-1}\right)^{-1}, \quad \boldsymbol{\mu_\beta} = \Sigma_\beta\left(\frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{y} + \Sigma_0^{-1}\boldsymbol{\beta_0}\right)
$$

깁스 표본기법

- 깁스 표본기법을 적용하여 β와 σ^2의 사후표본을 얻을 수 있고 β와σ^2에 대한 근사적 사후 추론이 가능.
- 깁스 표본기법의 알고리즘: (k-1번째 스텝에서 (β, σ^2)의 표본 $(\boldsymbol{\beta}^{(k-1)}, \sigma^{2(k-1)})$이 주어졌을 때 k번째 스텝.)

1. $\boldsymbol{\beta}^{(\boldsymbol{k})}$의 추출:

   (a) $\Sigma_\beta = \left(\frac{1}{\sigma^{2(k-1)}}\boldsymbol{X}'\boldsymbol{X} + \Sigma_0^{-1}\right)^{-1}, \quad \boldsymbol{\mu_\beta} = \Sigma_\beta\left(\frac{1}{\sigma^{2(k-1)}}\boldsymbol{X}'\boldsymbol{y} + \Sigma_0^{-1}\boldsymbol{\beta_0}\right)$

   (b) $\boldsymbol{\beta}^{(\boldsymbol{k})} \sim N(\boldsymbol{\mu_\beta}, \Sigma_\beta)$

2. $\sigma^{2(k)}$의 추출:

   (a) $SSR(\boldsymbol{\beta}^{(\boldsymbol{k})}) = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}^{(\boldsymbol{k})})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}^{(\boldsymbol{k})})$

   (b) $\sigma^{2(k)} \sim IG\left(\frac{n}{2} + a, \frac{1}{2}SSR(\boldsymbol{\beta}^{(\boldsymbol{k})}) + b\right)$

사후분산 Σ_β의 해석

- 사후분산 Σ_β:

$$
\begin{aligned}
\Sigma_\beta^{-1} &= \frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{X} + \Sigma_0^{-1} \\
&= (\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1})^{-1} + \Sigma_0^{-1} \\
&= Var^{-1}(\hat{\boldsymbol{\beta}}_{LSE}) + \Sigma_0^{-1} \\
&= \hat{\boldsymbol{\beta}}_{LSE}\text{의 정확도} + \boldsymbol{\beta}_{\mathbf{0}}\text{의 정확도}
\end{aligned}
$$

표본추정치 $\hat{\boldsymbol{\beta}}_{LSE}$의 정확도와 사전추정치 $\boldsymbol{\beta}_{\mathbf{0}}$의 정확도를 더한 것

사후평균 $\boldsymbol{\mu_\beta}$의 해석

- 사후평균 $\boldsymbol{\mu_\beta}$는

$$
\begin{aligned}
\boldsymbol{\mu_\beta} &= \Sigma_\beta\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y} + \Sigma_0^{-1}\boldsymbol{\beta_0}\right) \\
&= \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \Sigma_0^{-1}\right)^{-1}\left(\frac{1}{\sigma^2}(\mathbf{X}'\mathbf{X})(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{y} + \Sigma_0^{-1}\boldsymbol{\beta_0}\right) \\
&= (Var^{-1}(\hat{\boldsymbol{\beta}}_{\boldsymbol{LSE}}) + \Sigma_0^{-1})^{-1}(Var^{-1}(\hat{\boldsymbol{\beta}}_{\boldsymbol{LSE}})\hat{\boldsymbol{\beta}}_{\boldsymbol{LSE}} + \Sigma_0^{-1}\boldsymbol{\beta_0})
\end{aligned}
$$

- 이는 $\hat{\boldsymbol{\beta}}_{LSE}$와 $\boldsymbol{\beta_0}$의 가중평균으로 가중치는 각 추정치의 정확도에 비례.
- Σ_0^{-1}가 $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$에 비하여 매우 작다면 사후추정치 $\boldsymbol{\mu_\beta}$는 $\hat{\boldsymbol{\beta}_{LSE}}$와 거의 같게 되고, 반대로 σ^2이 매우 커서 $Var^{-1}(\hat{\boldsymbol{\beta}_{LSE}})$가 매우 작으면 $\boldsymbol{\mu_\beta}$는 $\boldsymbol{\beta_0}$와 거의 같게된다.

표본분산에 비례하는 사전분산

- 사전 정보가 뚜렷이 없는 경우 $\boldsymbol{\beta_0}$와 Σ_0을 정하는 것은 어려운 문제.
- 보통 $\boldsymbol{\beta_0} = \hat{\boldsymbol{\beta}}_{LSE}$로 하고 Σ_0를 $\hat{\boldsymbol{\beta}}_{LSE}$의 분산 $\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$에 비례하도록, 즉 $\Sigma_0 = c\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$로 놓는다.
- 이때 상수 c를 크게 하면 사전정보를 작게하는 것이고 c를 작게 하면 사전정보를 유의하게 사용히는 것.
- Σ_0을 $(\boldsymbol{X}'\boldsymbol{X})^{-1}$에 비례하도록 정하는 것의 장점은 사전분산이 설명변수의 단위에 의존하므로 설명변수의 단위가 비뀌면 Σ_0도 거기에 대응하여 바뀌게 되는 장점.
- 상수 $c = n$으로 놓는 경우를 단일정보 사전분포(unit information prior)라고 한다. 한 개의 관측치가 갖는 정보와 동일한 정보를 사전분포가 갖도록 하는 개념.

표본분산에 비례하는 사전분산을 사용할 경우

- $\hat{\beta}_{LSE}$의 분산이 $\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$인데 $\hat{\beta}_{LSE}$는 n개의 관측치로부터 얻은 추정치이므로 분산의 역행렬 $\frac{1}{\sigma^2}(\boldsymbol{X}'\boldsymbol{X})^{-1}$는 n개의 관측치가 갖는 정보를 나타낸다고 할 수 있고, 따라서 이를 n으로 나눈 값은 $\frac{1}{\sigma^2}(\boldsymbol{X}'\boldsymbol{X})^{-1}$을 Σ_0^{-1}로 놓으면, 즉 $\Sigma_0 = n\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$이면, 대략 한 개의 관측치가 갖는 정보를 사전정보가 갖도록 하는 것.
- 이 경우 Σ_0가 자료에 의존한다는 문제가 있긴 하지만, 관측치가 많을 경우 상대적으로 적은 정보를 사용하므로 영향이 크지 않다.

β의 사후분포

- $\boldsymbol{\beta_0} = \hat{\boldsymbol{\beta}_{LSE}}$, $\Sigma_0 = c\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$로 놓을 경우, $\boldsymbol{\beta}$의 사후평균과 분산:

$$\begin{aligned}
\Sigma_\beta &= \left(\frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{X} + \Sigma_0^{-1}\right)^{-1} = \left(\frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{X} + \frac{1}{c\sigma^2}\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \\
&= \left(1 + \frac{1}{c}\right)^{-1}\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1} = \frac{c}{c+1}\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{\mu_\beta} &= \Sigma_\beta\left(\frac{1}{\sigma^2}(\boldsymbol{X}'\boldsymbol{X})\hat{\boldsymbol{\beta}}_{LSE} + \Sigma_0^{-1}\beta_0\right) \\
&= \Sigma_\beta\left(\frac{1}{\sigma^2}(\boldsymbol{X}'\boldsymbol{X})\hat{\boldsymbol{\beta}}_{LSE} + \frac{1}{c\sigma^2}(\boldsymbol{X}'\boldsymbol{X})\hat{\boldsymbol{\beta}}_{LSE}\right) \\
&= \frac{c}{c+1}\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\left(\frac{c+1}{c}\frac{1}{\sigma^2}(\boldsymbol{X}'\boldsymbol{X})\hat{\boldsymbol{\beta}}_{LSE}\right) \\
&= \hat{\boldsymbol{\beta}}_{LSE}
\end{aligned}$$

β의 사후분포

$$\beta \mid \sigma^2 \boldsymbol{y} \sim N\left(\hat{\beta}_{LSE}, \frac{c}{c+1}\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\right)$$

- 단일정보 사전분포의 경우, $\boldsymbol{\mu}_{\boldsymbol{\beta}} = \hat{\beta}_{LSE}$,
- $\Sigma_\beta = \frac{n}{n+1}\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1} \rightarrow_{n\rightarrow\infty} \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1} = \mathrm{Var}(\hat{\beta}_{LSE})$.

$\boldsymbol{\beta}$의 결합사후분포:

- $\boldsymbol{\beta}$의 결합사후분포:

$$\begin{aligned}\pi(\boldsymbol{\beta},\sigma^2 \mid \boldsymbol{y}) \quad &\approx \quad (\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}(\boldsymbol{y}-\boldsymbol{X\beta})^{'}(\boldsymbol{y}-\boldsymbol{X\beta})\right) \\ &\times \left|\sigma^2(\boldsymbol{X}^{'}\boldsymbol{X})\right|^{-1/2} exp\left[-\frac{1}{2c\sigma^2}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}_{\boldsymbol{LSE}})^{'}(\boldsymbol{X}^{'}\boldsymbol{X})(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}_{\boldsymbol{LSE}})\right] \\ &\times(\sigma^2)^{-a-1}e^{-\frac{b}{\sigma^2}}\end{aligned}$$

위 식의 지수항에서 $\boldsymbol{\beta}$가 포함된 항만 보면, $\boldsymbol{\beta}\hat{_{LSE}}$를 간단히 $\hat{\boldsymbol{\beta}}$라고 할 때

$$\frac{1}{\sigma^2}(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta})+\frac{1}{c\sigma^2}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})^{'}(\boldsymbol{X}'\boldsymbol{X})(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})$$

β의 결합사후분포

$$\frac{1}{\sigma^2}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})+\frac{1}{c\sigma^2}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})'(\boldsymbol{X}'\boldsymbol{X})(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})$$

$$=\frac{1}{\sigma^2}\boldsymbol{y}'\boldsymbol{y}-\frac{2}{\sigma^2}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{y}+\frac{1}{\sigma^2}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}+\frac{1}{c\sigma^2}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}-2\frac{1}{c\sigma^2}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}+\frac{1}{c\sigma^2}\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}$$

$$=\boldsymbol{\beta}'\left(\frac{c+1}{c}\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X}\right)\boldsymbol{\beta}-\frac{2}{\sigma^2}\boldsymbol{\beta}'\left(\boldsymbol{X}'\boldsymbol{y}+\frac{1}{c}\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}\right)+\frac{1}{\sigma^2}\boldsymbol{y}'\boldsymbol{y}+\frac{1}{c\sigma^2}\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}$$

$$=\boldsymbol{\beta}'\Sigma_\beta^{-1}\boldsymbol{\beta}-2\boldsymbol{\beta}'\Sigma_\beta^{-1}\boldsymbol{\beta}+\boldsymbol{\beta}'\Sigma_\beta^{-1}\boldsymbol{\beta}-\boldsymbol{\beta}'\Sigma_\beta^{-1}\boldsymbol{\beta}+\frac{1}{\sigma^2}\boldsymbol{y}'\boldsymbol{y}+\frac{1}{c\sigma^2}\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}$$

$$=(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})'\Sigma_\beta(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})+\frac{1}{\sigma^2}\boldsymbol{y}'\boldsymbol{y}-\frac{1}{\sigma^2}\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}$$

$$=(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})'\Sigma_\beta(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})+\frac{1}{\sigma^2}\boldsymbol{y}'(I-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{y}$$

β의 사후분포

$$\begin{aligned}
\pi(\sigma^2 \mid \mathbf{y}) &= \int \pi(\sigma^2, \boldsymbol{\beta} \mid \mathbf{y}) d\boldsymbol{\beta} \\
&\propto (\sigma^2)^{-\frac{n}{2}-a-1} \left|\sigma^2(\mathbf{X}'\mathbf{X})^{-1}\right|^{-1/2} \left|\Sigma_\beta\right|^{1/2} \\
&\quad \times \int \left|\Sigma_\beta\right|^{-1/2} exp\left[-\frac{1}{2}(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})'\Sigma_\beta(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})\right] d\boldsymbol{\beta} \\
&\quad \times exp\left[-\frac{1}{2\sigma^2}\mathbf{y}'(I-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{y} - \frac{b}{\sigma^2}\right] \\
&\propto (\sigma^2)^{-\frac{n}{2}-a-1} exp\left[-\frac{1}{\sigma^2}\left\{\frac{1}{2}\mathbf{y}'(I-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{y}+b\right\}\right] \\
&\sim IG\left(\frac{n}{2}+a, \frac{1}{2}\mathbf{y}'(I-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{y}+b\right)
\end{aligned}$$

β의 사후분포

- σ^2의 주변 사후분포를 알기 때문에 깁스 표본기법을 사용할 필요없이 직접 β와 σ^2의 사후 표본들을 다음 알고리즘을 사용하여 생성가능.
- k번째 스텝에서

$$
\begin{aligned}
&(1) \quad \sigma^{2(k)}\text{의 추출 : } \sigma^{2(k)} \sim IG\left(\frac{n}{2}+a, \frac{1}{2}\boldsymbol{y}'(I-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')+b\right) \\
&(2) \quad \boldsymbol{\beta}^{(\boldsymbol{k})} \mid \sigma^{2(k)} \sim N\left(\hat{\boldsymbol{\beta}}_{LSE}, \frac{c}{c+1}\sigma^{2(k)}(\boldsymbol{X}'\boldsymbol{X})^{-1}\right)
\end{aligned}
$$

$(\beta^{(k)}, \sigma^{2(k)})$은 서로 독립적인 (β, σ^2)의 사후 표본이다. 이 경우 마코브체인 몬테칼로 표본이 아니므로 수렴시간이 필요하지 않으며 표본들은 서로 연관이 없는 독립표본들이다.

예제 13.1

체중과 운동이 HDL 콜레스테롤에 미치는 영향을 알아보기 위하여 26명을 대상으로 다음과 같은 실험을 수행하였다. 8명은 제어그룹으로 운동을 하지 않고, 8명은 유산소운동을 하게 하고, 나머지 10명은 유산소운동과 근력운동을 같이 하도록 하였다. 10주 후에 참가자들의 체중과 HDL 콜레스테롤을 측정한 결과 표 12.1과 같은 자료를 얻었다. (http://home.ewha.ac.kr/msoh/Bayesianbook/HDL.txt) 이 자료는 미국 Virginia Polytech Institute and State University 통계 상담실에서 분석한 것으로 Myers(l990)에 소개되어 있다.

예제 13.1

그룹	체중	HDL	그룹	체중	HDL
1	163.5	75	2	106	57.5
1	180	72.5	2	134	49
1	178.5	62	2	216.5	74
1	161.5	60	3	136.5	54.5
1	127	53	3	142.5	79.5
1	161	53	3	145	64
1	165	65	3	165	69
1	144	63.5	3	226	50.5
2	141	49	3	122	58
2	162	53.5	3	193	63.5
2	134	30	3	163.5	76
2	121	40.5	3	154	55.5
2	145	51.5	3	139	68

예제 13.1

각 그룹에 대하여 독자적인 절편과 기울기를 갖는 다음과 같은 회귀모형을 가정하였다. Y_i를 i번째 참가자의 HDL 콜레스테롤 수치 (mg/decaliter), x_i를 i번째 참가자의 체중(lb)이라고 하고 독립적으로 $\epsilon_i \sim N(0, \sigma^2)$일 때

$$
\begin{aligned}
Y_i &= \beta_0 + \beta_i x_i + \epsilon_i, \;\; i = 1, \ldots, 8 && \text{(그룹I)} \\
Y_i &= \gamma_0 + \gamma_i x_i + \epsilon_i, \;\; i = 9, \ldots, 16 && \text{(그룹II)} \\
Y_i &= \delta_0 + \delta_i x_i + \epsilon_i, \;\; i = 17, \ldots, 26 && \text{(그룹III)}
\end{aligned}
$$

모형을 가정한다.

예제 13.1

이 모형에서

$\boldsymbol{\beta} = (\beta_0, \gamma_0, \delta_0, \beta_1, \gamma_1, \delta_1)', \boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_{26}), \boldsymbol{Y} = (Y_1, \ldots, Y_{26}),$

$$\boldsymbol{X} = \begin{pmatrix} 1 & 0 & 0 & x_1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & x_8 & 0 & 0 \\ 0 & 1 & 0 & 0 & x_9 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & 0 & 0 & x_{16} & 0 \\ 0 & 0 & 1 & 0 & 0 & x_{17} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & 0 & x_{26} \end{pmatrix}$$

이라 놓으면 $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$의 관계가 성립한다.

예제 13.1

표 13.1에 주어진 자료를 사용하여 $\hat{\beta}_{LSE}$를 구하면

$$\hat{\beta}_{LSE} = (23.054, 14.225, 76.880, 0.2495, 0.250, -0.082)^T$$

$$(\boldsymbol{x}'\boldsymbol{x})^{-1} = \begin{pmatrix} 12.166 & 0 & 0 & -0.0752 & 0 & 0 \\ 0 & 22.8334 & 0 & 0 & -0.0187 & 0 \\ 0 & 0 & 3.0721 & 0 & 0 & -0.0187 \\ -0.0752 & 0 & 0 & 0.00047 & 0 & 0 \\ 0 & -0.0187 & 0 & 0 & 0.00013 & 0 \\ 0 & 0 & -0.0187 & 0 & 0 & 0.000118 \end{pmatrix}.$$

예제 13.1

베이지안 추정에 필요한 사전분포로 다음을 가정:

$$\boldsymbol{\beta} \mid \sigma^2 N(\hat{\boldsymbol{\beta}}_L SE, n\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}), \pi(\sigma^2) = 1/\sigma^2$$

R 프로그램을 이용하여 $\boldsymbol{\beta}, \sigma^2$의 사후추론을 수행하여 보자. 참고로, $\pi(\sigma^2) = 1/\sigma^2$일 때 $\boldsymbol{\beta}$의 조건부 사후분포는 앞에서 유도한 것과 동일하고 σ^2의 사후분포는 앞의 식에서 a=0, b=0를 대입한 것과 같다.

σ^2의 실제 사후밀도함수와 표본으로부터 추정된 사후밀도함수

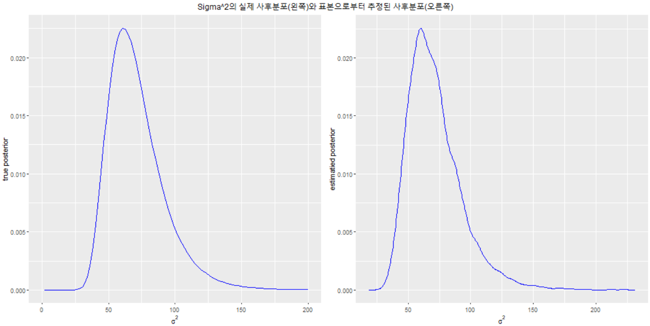


Figure: σ^2의 실제 사후분포(왼쪽)와 σ^2의 표본으로부터 추정된 사후분포(오른쪽)

기울기의 사후분포와 95% HPD 구간

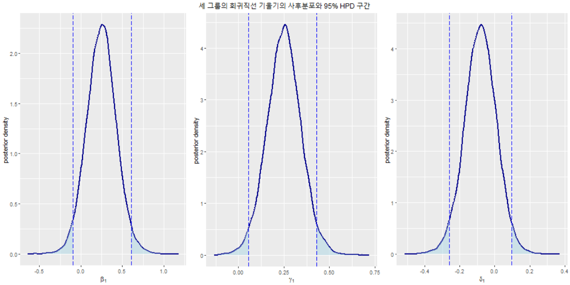
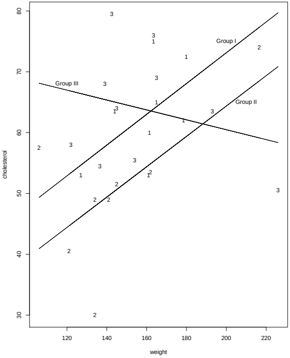


Figure: 세 그룹의 회귀직선 기울기의 사후분포와 95% HPD 구간

γ_1만이 95% HPD 구간이 0을 포함하지 않고 β_1과 δ_1의 95% HPD는 0을 포함하고 있어 유의하지 않음을 알 수 있다.

세 그룹의 추정된 회귀직선



그룹 I(제어 그룹)과 그룹 II(유산소운동 그룹)는 절편만 다를 뿐 기울기는 거의 차이가 없고 체중이 증가할수록 HDL 콜레스테롤 수치는 올라감을 알 수 있다. 그러나 그룹 III(유산소운동 + 근력운동 그룹)은 두 그룹과 매우 다른 양상을 보이는데, 체중 증가와 HDL 콜레스테롤 수치와는 거의 상관이 없는 것으로 보인다.

기울기 차이의 사후분포

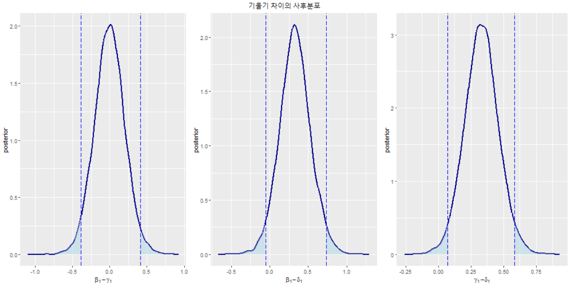


Figure: 기울기 차이의 사후분포

- $\gamma_1 - \delta_1$ HPD 구간은 0을 포함하지 않는다. → 유의한 차이 **존재**.
- $\beta_1 - \gamma_1$ 은 추정치가 거의 0에 가깝다. → 유의한 차이 **부재**.
- $\beta_1 - \delta_1$ HPD 구간의 왼쪽 끝 부분에 0이 위치하고 있다.

  → 확실한 결론을 내리기 어렵다.

```
#Figure 13.3
par(mfrow=c(1,1))
plot(HDL.data[ ,2], HDL.data[,3], xlab="weight", ylab="HDL
cholesterol", type="n")
for(i in 1:n)text(HDL.data[i,2], HDL.data[i,3],HDL.data[i,1])
beta.hat=apply(beta,2,mean)
lines( HDL.data[, 2], beta.hat[1]+ beta.hat[4] * HDL.data[, 2])
lines( HDL.data [, 2], beta.hat[2]+ beta.hat[5] * HDL.data [, 2])
lines ( HDL.data[, 2], beta.hat[3]+ beta.hat[6] * HDL.data [, 2])
text( 200,75, "Group I")
text( 210,65, "Group II")
text( 120,68, "Group III")
```

```
#Figure 13.4
diff.12=beta[,4]-beta[,5]
diff.13=beta[,4]-beta[,6]
diff.23=beta[,5]-beta[,6]
par(mfrow=c(1,3))
plot(density(diff.12), xlab="beta_l-gamma_l", ylab="posterior density", main="")
HPD=HPDsample(diff.12)
abline(v=HPD , lty=2)
plot(density(diff.13), xlab="beta_l-delta_l", ylab="posterior density", main="")
HPD=HPDsample(diff.13)
abline(v=HPD , lty=2)
plot(density(diff.23), xlab="gamma_l-delta_l", ylab="posterior density", main="")
HPD=HPDsample(diff.23)
abline(v=HPD , lty=2)
```