Solution to Problem 1:

Joint posterior distribution is proportional to product of three terms (distributions):

$$f\left(\underline{y}\middle|\underline{x},\eta,\beta,\rho\right)g(\rho) \propto \left(\rho^{(n-2)/2-1}e^{-\frac{1}{2}\rho S_{ee}}\right)\left(\rho^{1/2}e^{-\frac{1}{2}\rho n(\eta-\bar{y})^2}\right)\left(\rho^{1/2}e^{-\frac{1}{2}\rho S_{xx}(\beta-b)^2}\right)$$

The first term is posterior distribution for precision (ρ) which is gamma with

$shape = \frac{n-2}{2} = \frac{21-2}{2} = 14.5$, and $scale = \frac{2}{S_{ee}} = \frac{2}{423.4} = 0.0047$.

Second term is distribution for transformed intercept (η) which is normal with (given ρ and independent of β) $mean = \bar{y} = 32$, and $precision = n\rho = 31\rho$.

Third term is distribution for slope (β) which is normal with (given ρ and independent of η)

$mean = b = 0.608$, and $precision = S_{xx}\rho = 677.419\rho$.

Solution to Problem 2:

Looking at a plot of datapoints along with the estimated line in figure 1, it seems to be reasonable to assume a linear relationship between birthweight and Estriol levels, however it is hard to confirm the constant variance assumption, because it seems that the variance from the line (y=a+bx) is getting bigger as the Estriol level is increased.

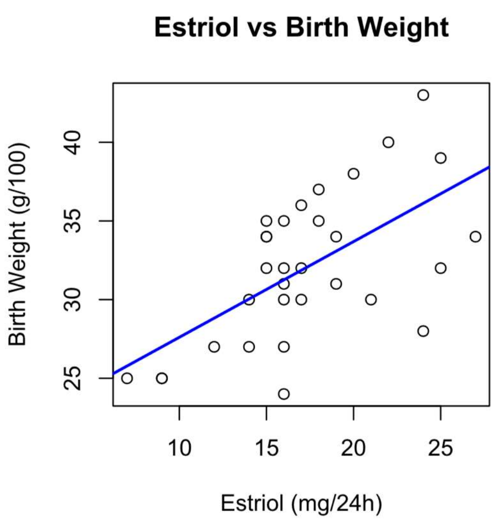


Figure 1

A residual Q-Q plot in figure 2 confirms the assumption of residuals being normally distributed. There is, however, a deviation from Q-Q line in higher quantiles.

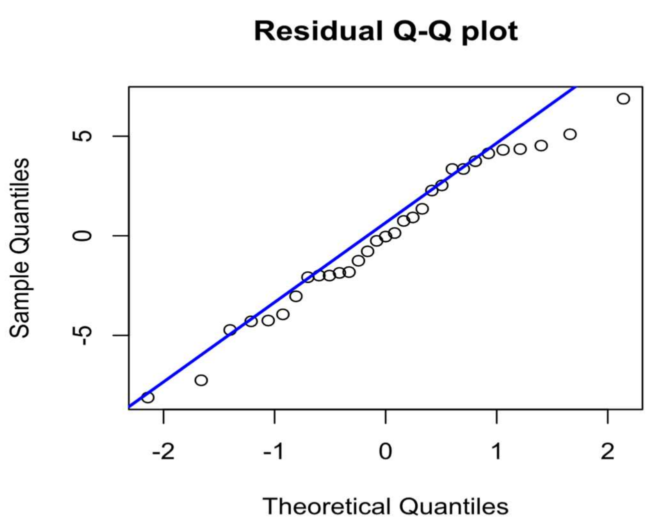


Figure 2

Solution to Problem 3:

Given x_{new}, slope β, transformed intercept η, and precision ρ, birthweight is normally distributed with:

$mean = (x_{new} - \bar{x})\beta + \eta$, and $variance = {}^{1}/_{\rho}$

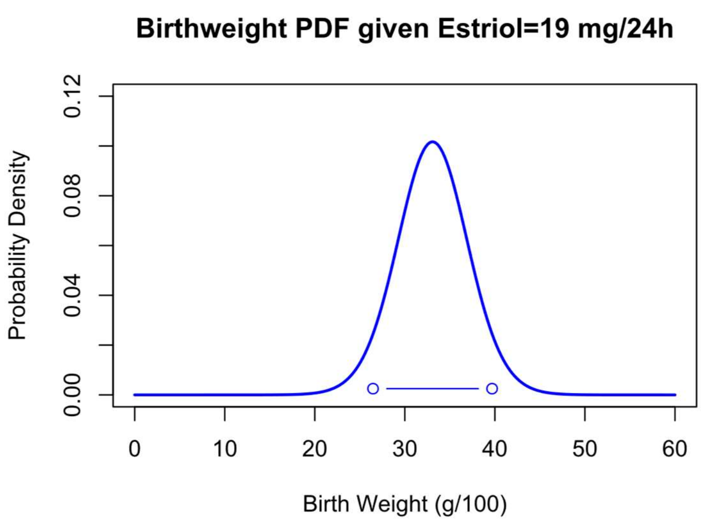
Knowing x_{new}=19, distribution of birthweight can be marginalized into a non-standard t distribution with:

$$center = (x_{new} - \bar{x})b + \bar{y} = (19 - 17.23) \times 0.608 + 32 = 33.079$$

$$spread = \sqrt{\left(\frac{(x_{new}-\bar{x})^2}{S_{xx}} + \frac{1}{n} + 1\right)\left(\frac{S_{ee}}{n-2}\right)} = \sqrt{\left(\frac{(19-1\ \ .32)^2}{677.4} + \frac{1}{31} + 1\right)\left(\frac{423.4}{31-2}\right)} = 3.891$$

$$DOF = n - 2 = 29.$$

The distribution is shown in figure 3 and the 90% credible interval in the following table.

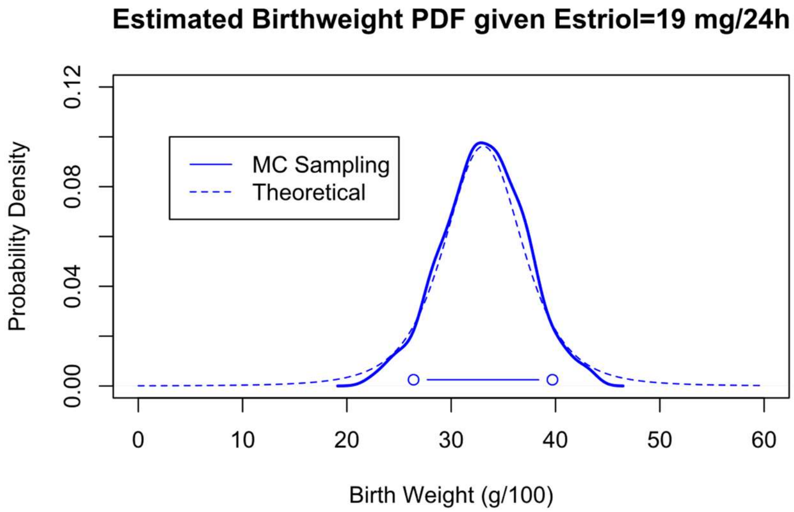


5% quantile	26.47
95% quantile	39.69

Figure 3

Solution to Problem 4:

The resulting distribution from Monte Carlo sampling (based on unit 8- slide 16) is presented in figure 4, and the 90% predictive interval is presented in the following table.



5% quantile	26.39
95% quantile	39.70

Figure 4

From the figure and 90% credible interval, it can be concluded that the predictive distribution resulted from Monte Carlo sampling is compatible with the one resulted from mathematical analysis (from problem 3).

The result of the predicted birthweight (in g/100) from the regression line is:

$$y_{new} = a + bx_{new} = 21.52 + 0.608 \times 19 = 33.079$$

Which is the center of the predictive non-standard t-distribution.

The center of the 90% credible interval from Monte Carlo sampling is:

$$\frac{26.39 + 37.9}{2} = 33.048$$

The difference between two values from analytical distribution and Monte Carlo process seems to be negligible (less than 0.1%).