A1_Q3

Joohyeok

2023-09-26

```
library(ggplot2)

df = read.table("IQ.txt",header=TRUE)

lm_model <- lm(formula = IQ ~ birthweight, data = df)
lm_model
```

```
##
## Call:
## lm(formula = IQ ~ birthweight, data = df)
##
## Coefficients:
## (Intercept)  birthweight
##   87.015672     0.008974
```

```
result <- summary(lm_model)
result
```

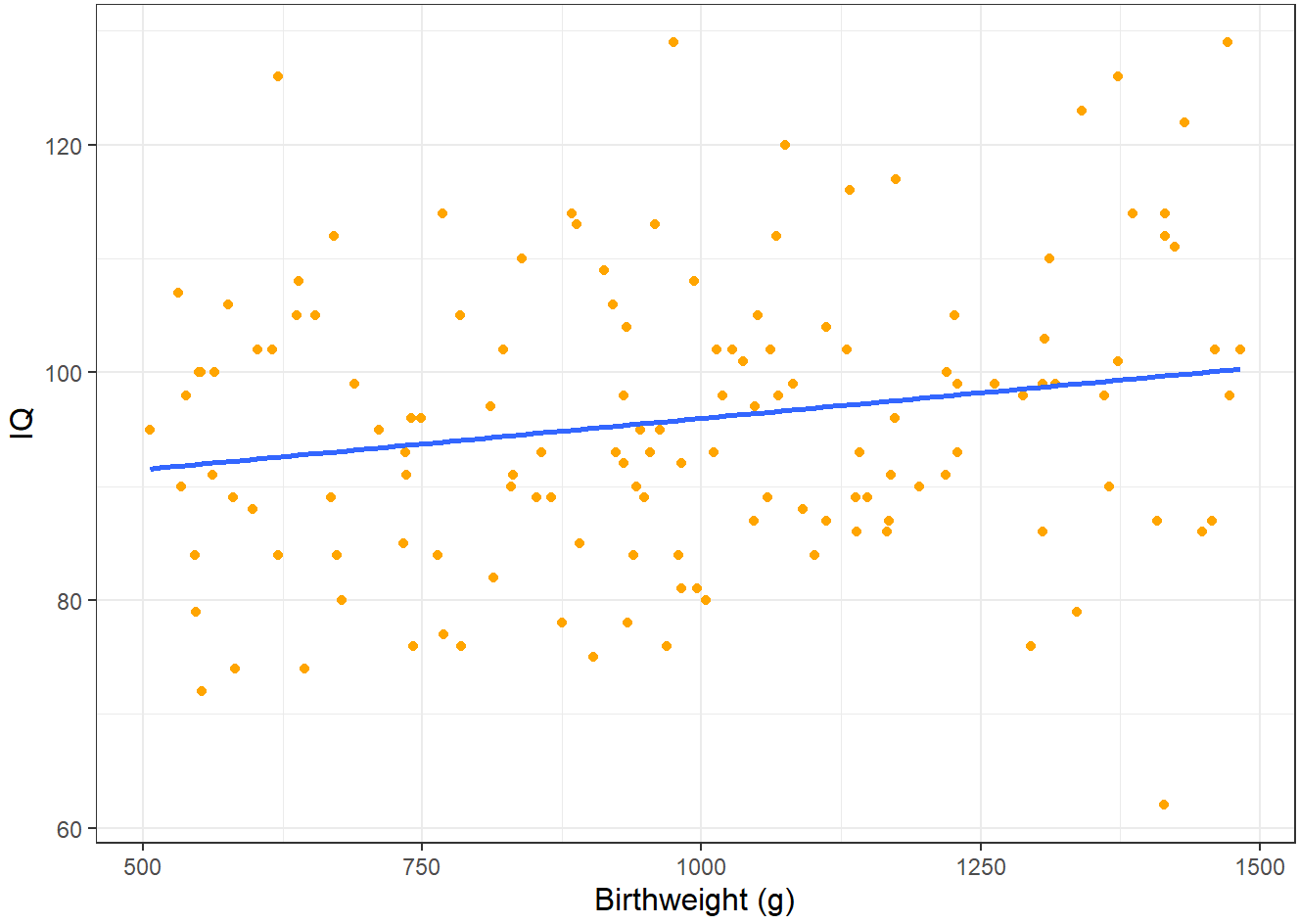
```
##
## Call:
## lm(formula = IQ ~ birthweight, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -37.705  -8.993  -0.635   8.046  33.411
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 87.015672   3.988953  21.814   <2e-16 ***
## birthweight  0.008974   0.003927   2.286   0.0238 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.59 on 136 degrees of freedom
## Multiple R-squared:  0.03699,    Adjusted R-squared:  0.02991
## F-statistic: 5.224 on 1 and 136 DF,  p-value: 0.02383
```

```
#Q3_c)
result$coefficients
```

```
##                 Estimate  Std. Error   t value     Pr(>|t|)
## (Intercept) 87.015672210 3.988952500 21.814166 2.995531e-46
## birthweight  0.008974179 0.003926515  2.285533 2.382829e-02
```

```
#Q3_d)
ggplot(df,aes(y=IQ,x=birthweight))+
  geom_point(colour="orange")+
  geom_smooth(method = 'lm', se = F)+
  theme_bw()+
  theme(axis.title=element_text(size=12))+
  labs(x="Birthweight (g)", y="IQ")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
#Q3_e)
# A 99% CI for the effect of birthweight, beta1, is (-0.00128,0.01923)

confint(lm_model,parm="birthweight",level=0.99)
```

```
##                    0.5 %     99.5 %
## birthweight -0.001283697 0.01923205
```

```
#Q3_f)
# If we were to repeat this study an infinite number of times, each time selecting 138 babies bo
rn at a medical centre, 99% of the 99% CIs that we created would contain beta1.

#Q3_g)
# A 99% CI for mean IQ of babies with a birthweight of 750g is (90.094,97.399)
predict(lm_model,newdata=data.frame(birthweight=750),interval="confidence",level=0.99)
```

```
##        fit      lwr      upr
## 1 93.74631 90.09397 97.39864
```

```
#Q3_h)
# If we were to repeat this study an infinite number of times, each time selecting 138 babies bo
rn at a medical centre, 99% of the 99% CIs that we created would contain the mean IQ of babies w
ith a birthweight of 750g.

#Q3_i)
# If birthweight is not normally distributed, it may have some effect on the validity of the rea
soning made in (e) and (g). This is because the assumptions of linear regression equations deriv
ed from the analysis include the normality, equal variance, and independence of error term. When
this assumption is out of line, the inferred result can lead to biased results, so it is unrelia
ble and the inference can not be said to be valid.

#Q3_j)
# In terms of the way the data were collected, birthweight is a random variable.

#Q3_k)
# In model in (a), birthweight is treated as a fixed variable.

#Q3_l)
# No. In the simple linear regression model, the response is modeled conditional on the predicto
r variable. Therefore, it is the same regardless of whether the predictor is fixed or random.
```