Solutions

STAT 217: Quiz 23

1. A high respiratory rate is a potential diagnostic indicator of respiratory infection in children. To judge whether a respiratory rate is truly high, a physician must have a clear picture of the distribution of normal respiratory rates. To this end, Italian researchers measured the respiratory rates of 618 children between the ages of 15 days and 3 years.

```
##
## Call:
## lm(formula = log.rate ~ Age, data = kids)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -0.626 -0.132 -0.004  0.135  0.548
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.845119   0.012628   304.5   <2e-16 ***
## Age         -0.019009   0.000736   -25.8   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.196 on 616 degrees of freedom
## Multiple R-squared:  0.52,Adjusted R-squared:  0.519
## F-statistic:  668 on 1 and 616 DF,  p-value: <2e-16
##                2.5 %   97.5 %
## (Intercept)  3.82032  3.86992
## Age         -0.02045 -0.01756
```

(a) Conduct a hypothesis test for the slope coefficient and write your conclusion in the context of the problem. Only write your evidence sentence here.

There is strong evidence that the mean of log rate depends on age (p-value <0.0001 from t-stat $=-25.8$ on 616 df)

(b) Provide an interpretation of b_1 on the original scale.

For a one day increase in age, the median respiratory rate is estimated to change by a factor of $e^{-0.019} = 0.98$, with a 95% CI from $e^{-0.02045} = 0.9798$ to $e^{-0.01756} = 0.9826$.

3. The data come from the United Nations website showing CO2 consumption per capita (in metric tons) between the years of 1990 to 2010. We will examine this relationship.

```
##
## Call:
## lm(formula = Value ~ log.year, data = c02)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -5.13  -4.33  -2.49   1.96  63.75
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -302.8      265.1   -1.14     0.25
## log.year        40.5       34.9    1.16     0.25
##
## Residual standard error: 6.85 on 4270 degrees of freedom
## Multiple R-squared:  0.000315,Adjusted R-squared:  8.13e-05
## F-statistic: 1.35 on 1 and 4270 DF,  p-value: 0.246
##                2.5 % 97.5 %
## (Intercept) -822.53  217.0
## log.year     -27.89  108.9
```

(a) Conduct a hypothesis test for the slope coefficient and write your conclusion in the context of the problem. Only write your evidence sentence here.

There is no evidence that the mean CO_2 consumption per capita depends on log year (p-value = 0.25 from t-stat = 1.16 on 4270 df)

(b) Provide an interpretation of b_1 on the original scale.

For a 10% increase in year, the median CO_2 consumption per capita is estimated to change by a factor of $40.5 \log(1.10) = 3.86$, with a 95% CI from $-27.69 \log(1.10)$ to $108.9 \log(1.10)$.