1 Standardization and Studentization

1.1 Standardization

A random variable is said to be standardized if the difference from its mean is scaled by its standard deviation. The residuals above have mean zero but their variance is unknown, it depends on the true values of θ. Standardization is thus not possible in practice.

1.2 Studentization

Instead, you can compute studentized residuals by dividing a residual by an estimate of its standard deviation.

1.3 Internal and External Studentization

If that estimate is independent of the $i-$th observation, the process is termed *external studentization*. This is usually accomplished by excluding the $i-$th observation when computing the estimate of its standard error. If the observation contributes to the standard error computation, the residual is said to be *internally studentized*.

R Code

- Internally studentized residuals: `rstandard`
- Externally studentized residuals: `rstudent`

2 Standardized and Studentized Residuals

Recall, the standardized residual is the residual divided by its standard deviation.

Plot the standardized residual of the simple linear regression model of the data set faithful against the independent variable waiting.

We apply the `lm` function to a formula that describes the variable eruptions by the variable waiting, and save the linear regression model in a new variable eruption.lm. Then we compute the standardized residual with the rstandard function.

```
eruption.lm = lm(eruptions ~ waiting, data=faithful)
eruption.stdres = rstandard(eruption.lm)
```

We now plot the standardized residual against the observed values of the variable waiting.

```
plot(faithful$waiting, eruption.stdres,
    ylab="Standardized Residuals",
     xlab="Waiting Time",
     main="Old Faithful Eruptions")
abline(0, 0)                    # the horizon
```