priate transformation can be found to stabilize the variance. Table 2.4 shows the transformation for common situations often encountered.

Table 2.4 *Response Distribution-Based Transformations*

Response Distribution	Variance in Terms of Mean μ	Transformation $f(y)$
Binomial	$\frac{\mu(1-\mu)}{n}$	$\sin^{-1}\sqrt{y/n}$ (radians)
Poisson	μ	$\sqrt{y}$ or $\sqrt{y+\frac{1}{2}}$
Lognormal	$c\mu^2$	$\log(y)$

2.6.3 Alternatives to Least Squares Analysis

When the variance of the experimental error is not constant for all levels of the treatment factor, but it is not related to the cell means, a transformation will not be an appropriate way of equalizing or stabilizing the variances. A more general solution to the problem is to use weighted least squares. Using weighted least squares, $\hat{\boldsymbol{\beta}}$ is the solution to the normal equations $\boldsymbol{X'WX\beta} = \boldsymbol{X'Wy}$, where $\boldsymbol{W}$ is a diagonal matrix whose diagonal elements are the reciprocals of the variances within each treatment level. As an illustration of this method, consider the R code below for analyzing the data from the bread rise experiment.

```
> with(bread, { vars <- tapply(height, time, var)
+ weights <- rep( 1/vars, each = 4 )
+ mod3 <- lm( height ~ time, weights = weights, data = bread )
+ anova( mod3 )
+ })
```

In this example, the `with(bread, {...})` function causes all statements within the { } brackets to use the variables from the data frame `bread`. The `tapply (height, time , var)` function is used to calculate the variance of the response at each level of the factor `time`. The weights are calculated as the reciprocal of the variances and the `rep( )` function is used to expand the vector of weights to the number of rows in the data frame `bread`. The `lm` function calculates the weighted least squares estimates and the `anova` function prints the ANOVA table. The results appear on the next page.

```
Analysis of Variance Table

Response: height
          Df Sum Sq Mean Sq F value  Pr(>F)
time       2 14.695  7.3476  7.3476 0.01283 *
Residuals  9  9.000  1.0000
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

With these results, it can be seen that the F-test from the weighted least squares is more sensitive than the unweighted least squares, and the P-value is similar to what was obtained with the Box-Cox transformation shown in Section 2.6.1. An alternate method of analysis is to use the `gls()` function in the `nlme` package as shown in the Rcode on the web page for this book.

When the error distribution is not normal, an alternative to analyzing a transformation of the response is to use a generalized linear model (see McCullagh and Nelder, 1989). In fitting a generalized linear model, the user must specify the error distribution and a link function in addition to the model. The method of maximum likelihood is used to estimate the model parameters and the generalized likelihood ratio tests are used to test the hypotheses. When the link function is the identity and the distribution is normal, the generalized linear model analysis will result in the method of least squares and the ANOVA F-test. There are several R functions to fit the generalized linear models and compute the appropriate likelihood ratio test statistics.

To illustrate the use of one of these functions to analyze experimental data, consider the following example. A professor wanted to compare three different teaching methods to determine how the students would perceive the course. The treatment factor was the teaching method, the experimental unit was a class of students, and the response was the summary of student ratings for the course. The professor taught two sections of the course for three consecutive semesters resulting in a total of six experimental units or classes. He constructed a randomized list so that two classes were assigned to each teaching method. This would reduce the chance that other differences in the classes, or differences in his execution of the teaching methods, would bias the results. At the end of each semester, the students were asked to rate the course on a five-point scale, with 1 being the worst and 5 being the best. Therefore, the response from each class was not a single, normally distributed response, y, but a vector $(y_1, \ldots, y_5)$ response that followed the multinomial distribution. The summary data from the experiment is shown in Table 2.5.