

Methods and Overview of Using EdSurvey for Running Wald Tests*†‡

Developed by Alex Lishinski

Wald Test

The Wald test is a statistical test of estimated parameters in a model, with the null hypothesis being that a set of parameters is equal to some values. In the default case where the null hypothesis value of the parameters is 0, if the test fails to reject the null hypothesis, removing the variables from the model will not substantially harm the fit of that model.

The formula for the test statistic of a single parameter is as follows:

$$W = \frac{(\hat{\beta} - \beta_0)^2}{\text{Var}(\hat{\beta})}$$

The Wald test is similar to the hypothesis tests typically printed in regression output, but the Wald test can be used to test multiple hypotheses about multiple parameters simultaneously.

To test q hypotheses on p estimated parameters, let $\hat{\beta}$ be the vector of estimated coefficients, R be a $q \times p$ hypothesis matrix (this matrix has 1 row per coefficient being tested with a value of 1 in the column corresponding to that coefficient),¹ \hat{V} be the estimated covariance matrix for $\hat{\beta}$, and r be the vector of hypothesized values for $\hat{\beta}$.

Then the Wald test statistic for multiple parameters is equal to:

$$W = (R\hat{\beta} - r)'(R\hat{V}R')^{-1}(R\hat{\beta} - r)$$

The resulting test statistic can be tested against a chi-square distribution or an F distribution. The chi-square distribution is preferable to the F distribution when the number of degrees of freedom is large, whereas the F test is preferable when the number of degrees of freedom is small (Korn & Graubard, 1990).

For the chi-square test, the degrees of freedom value is equal to the number of parameters that are tested, and the test statistic is the unadjusted value of W :

$$W \sim \chi^2(p)$$

For the F test, two scenarios dictate the use of two different adjustments for the test statistics, along with two different values for the denominator degrees of freedom.

In the case of the National Assessment of Educational Progress (NAEP) and other international assessments in EdSurvey, when the data are collected using a multistage stratified sampling design, the test statistic is adjusted based on the sampling parameters:

$$W_{adj} = (d - p + 1)W/(pd)$$

*An overview of the Wald test is available in Draper, N. R., & Smith, H. (1998). *Applied regression analysis*. New York, NY: Wiley.

†This publication was prepared for NCES under Contract No. ED-IES-12-D-0002 with the American Institutes for Research. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government.

‡The author would like to thank Dan Sherman and Mike Cohen for reviewing this document.

¹The Wald test also can be used to test hypotheses about contrasts between parameters, but the `waldTest` function in EdSurvey does not currently support contrasts.

where d is the number of Primary Sampling Units (PSUs) minus the number of strata, and p is the number of parameters being tested (Korn & Graubard, 1990).

The adjusted test statistic is then compared to the F distribution with p and $(d - p - 1)$ degrees of freedom (Korn and Graubard, 1990):

$$W_{adj} \sim F(p, d - p - 1)$$

In cases where the data is collected using single-stage stratified sampling design (e.g., some countries in the PIAAC data), the test statistic is as follows:

$$W_{adj} = W/p$$

The adjusted test statistic is then compared to the F distribution with p and d degrees of freedom, where d is the residual degrees of freedom in the model and p is the number of parameters being tested:

$$W_{adj} \sim F(p, d)$$

Examples

Examples of the `waldTest` function in `EdSurvey`:

```
library(EdSurvey)

sdf <- readNAEP(system.file("extdata/data", "M36NT2PM.dat", package = "NAEPprimer"))

myLogit <- logit.sdf(dsex ~ b017451 + b003501, data = sdf, returnVarEstInputs = T)

summary(myLogit)
```

```
##
## Formula: dsex ~ b017451 + b003501
## Family: binomial (logit)
##
## Weight variable: 'origwt'
## Variance method: jackknife
## JK replicates: 62
## full data n: 17606
## n used: 10551
##
## Coefficients:
##              coef          se          t    dof  Pr(>|t|)
## (Intercept) -0.089703  0.087402 -1.026324  44.963  0.3102286
## b017451Once every few weeks  0.438794  0.071682  6.121417  67.980  5.215e-08
## b017451About once a week    0.528746  0.080104  6.600752  53.547  1.896e-08
## b0174512 or 3 times a week  0.644332  0.065295  9.868046  55.971  7.416e-14
## b017451Every day            0.830757  0.076597  10.845756  66.982  2.220e-16
## b003501Graduated H.S.      -0.246203  0.088380 -2.785732  60.995  0.0071065
## b003501Some ed after H.S.  -0.341405  0.083372 -4.094956  71.782  0.0001094
## b003501I Don't Know       -0.584673  0.092604 -6.313724  65.900  2.650e-08
##
## (Intercept)
## b017451Once every few weeks ***
```

```
## b017451About once a week ***
## b0174512 or 3 times a week ***
## b017451Every day ***
## b003501Graduated H.S. **
## b003501Some ed after H.S. ***
## b003501I Don't Know ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

An integer (or integer vector) can be specified to test coefficients by reference to the order of coefficients in the summary output. For example, to test whether the coefficient for level of **Once every few weeks** of **b017451** is equal to zero, specify `coefficients = 2` because it is the second coefficient in the summary output for `myLogit`. To test all the levels of **b017451**, specify `2:5` to jointly test these four coefficients.

```
waldTest(model = myLogit, coefficients = 2)
```

```
## Wald test:
## -----
## H0:
## b017451Once every few weeks = 0
##
## Chi-square test:
## X2 = 37.5, df = 1, P(> X2) = 9.3e-10
##
## F test:
## W = 37.5, df1 = 1, df2 = 62, P(> W) = 6.9e-08
```

```
test <- waldTest(model = myLogit, coefficients = 2:5)
```

```
# example of hypothesis matrix R
test$hypoMatrix
```

```
##      (Intercept) b017451Once every few weeks b017451About once a week
## L1           0                1                0
## L2           0                0                1
## L3           0                0                0
## L4           0                0                0
##      b0174512 or 3 times a week b017451Every day b003501Graduated H.S.
## L1                0                0                0
## L2                0                0                0
## L3                1                0                0
## L4                0                1                0
##      b003501Some ed after H.S. b003501I Don't Know
## L1                0                0
## L2                0                0
## L3                0                0
## L4                0                0
```

```
# example of hypothesis vector r
test$H0
```

```
## [1] 0 0 0 0
```

```
test
```

```
## Wald test:
## -----
## H0:
```

```

## b017451Once every few weeks = 0
## b017451About once a week = 0
## b0174512 or 3 times a week = 0
## b017451Every day = 0
##
## Chi-square test:
## X2 = 159.9, df = 4, P(> X2) = 0.0
##
## F test:
## W = 38.0, df1 = 4, df2 = 59, P(> W) = 1.1e-15

```

Coefficient names can be specified to test coefficients by name. The correct coefficient names can be taken from the summary output for the model.

```
waldTest(myLogit, c("b017451Every day", "b017451About once a week"))
```

```

## Wald test:
## -----
## H0:
## b017451Every day = 0
## b017451About once a week = 0
##
## Chi-square test:
## X2 = 124.5, df = 2, P(> X2) = 0.0
##
## F test:
## W = 61.3, df1 = 2, df2 = 61, P(> W) = 2.6e-15

```

The name of a categorical independent variable can also be specified to jointly test all the coefficients corresponding to the levels of that variable.

```
waldTest(myLogit, "b017451")
```

```

## Wald test:
## -----
## H0:
## b017451Once every few weeks = 0
## b017451About once a week = 0
## b0174512 or 3 times a week = 0
## b017451Every day = 0
##
## Chi-square test:
## X2 = 159.9, df = 4, P(> X2) = 0.0
##
## F test:
## W = 38.0, df1 = 4, df2 = 59, P(> W) = 1.1e-15

```

The default value of H0 is 0, but alternatives can be specified using the H0 argument.

```
waldTest(model = myLogit, coefficients = 2, H0 = 0.5)
```

```

## Wald test:
## -----
## H0:
## b017451Once every few weeks = 0.5
##
## Chi-square test:
## X2 = 0.73, df = 1, P(> X2) = 0.39

```

```

##
## F test:
## W = 0.73, df1 = 1, df2 = 62, P(> W) = 0.4
waldTest(model = myLogit, coefficients = 2:5, H0 = c(0.5, 0.6, 0.7, 0.8))

## Wald test:
## -----
## H0:
## b017451Once every few weeks = 0.5
## b017451About once a week = 0.6
## b0174512 or 3 times a week = 0.7
## b017451Every day = 0.8
##
## Chi-square test:
## X2 = 2.2, df = 4, P(> X2) = 0.71
##
## F test:
## W = 0.52, df1 = 4, df2 = 59, P(> W) = 0.72
waldTest(model = myLogit, coefficients = "b017451", H0 = c(0.5, 0.6, 0.7, 0.8))

## Wald test:
## -----
## H0:
## b017451Once every few weeks = 0.5
## b017451About once a week = 0.6
## b0174512 or 3 times a week = 0.7
## b017451Every day = 0.8
##
## Chi-square test:
## X2 = 2.2, df = 4, P(> X2) = 0.71
##
## F test:
## W = 0.52, df1 = 4, df2 = 59, P(> W) = 0.72
waldTest(model = myLogit, coefficients = c("b017451Every day", "b017451About once a week"),
          H0 = c(0.1, 0.2))

## Wald test:
## -----
## H0:
## b017451Every day = 0.1
## b017451About once a week = 0.2
##
## Chi-square test:
## X2 = 91.2, df = 2, P(> X2) = 0.0
##
## F test:
## W = 44.9, df1 = 2, df2 = 61, P(> W) = 1e-12

```

References

Korn, E. L., & Graubard, B. I. (1990). Simultaneous testing of regression coefficients with complex survey data: Use of Bonferroni t statistics. *The American Statistician*, 44(4), 270–276.