Comparing Paired Data

Dave Lorenz

July 9, 2015

These examples demonstrate some of the functions and statistical methods for comparing paired observations that are available in the **smwrQW** package. The examples in this vignette use the Atra dataset from the **NADA** package. The examples in this vignette use the functions `as.lcens` and `as.mcens` to convert those data to a form used by the functions demonstrated. The functions demonstrated in these examples will also accept data of class "qw" without conversion.

```
> # Load the smwrQW package
> library(smwrQW)
> # And the data
> data(Atra, package="NADA")
> head(Atra, 2)


  June Sept JuneCen SeptCen
1 0.38 2.66   FALSE   FALSE
2 0.04 0.63   FALSE   FALSE


> print(head(Atra, 2), digits=11)


          June         Sept JuneCen SeptCen
1 0.37999999520 2.6600000860   FALSE   FALSE
2 0.03999999911 0.6299999952   FALSE   FALSE
```

Note: the data in Atra are not as they appear in the the first display of data; those data are rounded. The actual values are displayed in the second printing.

1 Maximum Likelihood Estimation Method

Comparing paired data using maximum likelihood estimation method extends censored regression as the paired t-test extends ordinary least squares–the test is constructed by comparing the mean of an interval-censored response variable to 0.

An important first step in any parametric statistical analysis is to plot the data. Unfortunately, for this particular type of analysis, computing the difference between two interval-censored values, it is easier to create a probability plot after the analysis. The `censReg` function is used for regression and can be used simply to compute the mean of censored data simply by including only the intercept term in the formula. It functions much like any modeling function, like `lm` in R—it constructs the model from a formula and data and has other options similar to `lm`. Its use for the censored equivalent of the paired t-test is shown below. For this test, it is essential to used interval-censored data because subtracting left-censored values results in indefinite values—for example less than 1 minus less than 1 is the same as less than 1 plus greater than 1; flipping data results in greater-then values. Note, with data of class "qw," conversion to class "mcens" is not necessary, but the subtraction is done using the `add` function.

```
> # First compute the interval-censored values for June and September
> # Set the minimum value to 0 for censored values
> June <- with(Atra, as.mcens(ifelse(JuneCen, 0, June), June))
> Sept <- with(Atra, as.mcens(ifelse(SeptCen, 0, Sept), Sept))
> # The parametric paired-data test, save for plot and print it
> Atra.pt <- censReg(Sept-June ~ 1)
> print(Atra.pt)


Call:
censReg(formula = Sept - June ~ 1)

Coefficients:
            Estimate Std. Error z-score p-value
(Intercept)    3.927      3.672   1.069  0.2747

Estimated residual standard error (Unbiased) = 17.99
Distribution: normal
Number of observations = 24, number left censored = 0,
  number right censored = 0 (0 percent)

Loglik(model) = -146.2 Loglik(intercept only) = -146.2
  Chi-square = 0, degrees of freedom = 0, p-value = 1

Computation method: MLE


> # setSweave required for vignettes, use setPage or setPDF otherwise
> setSweave("graph01", 6 ,6)
> # graph number 2 is the probability plot
> plot(Atra.pt, which=2, set.up=FALSE)
> # Reuired to close graph for the vignette
> graphics.off()
```
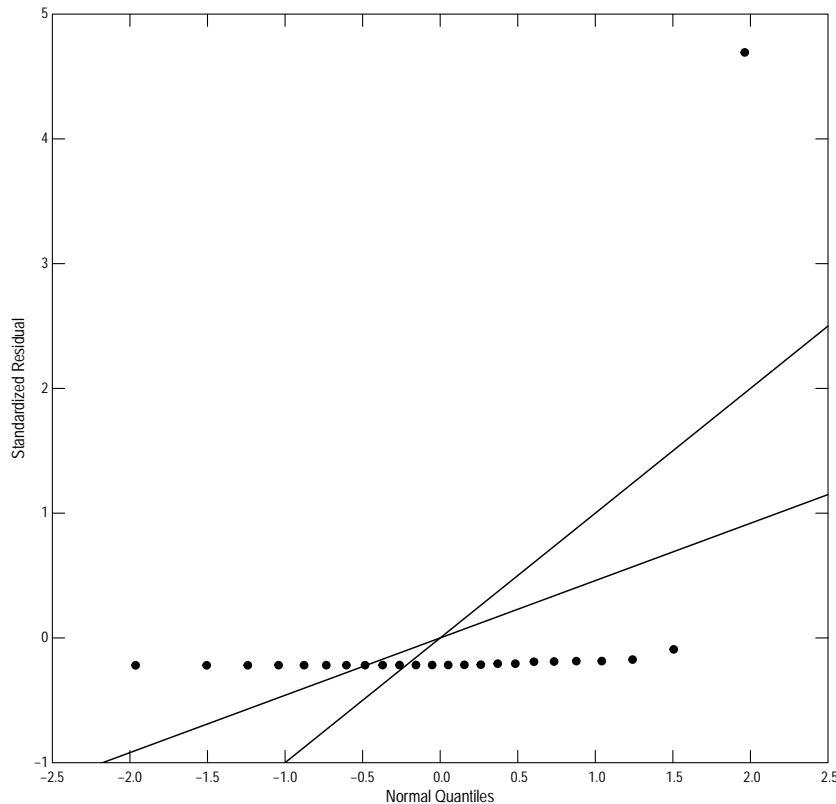
Figure 1. Probability plot to assess the normality of the data.

For this test, the p-value of the intercept term is used to asses the significance of the difference, not the overall p-value, which will always be 1. For these data, the probability plot indicates that the differences are not normally distributed and the p-value from the test cannot be accepted and no decision can be made to accept or reject the null hypothesis. The `pow` function can be used to transform the data so that the data are more nearly normally distributed. For these data, the probability plot indicates very positively skewed residuals, indicating a very small value for the `lambda` argument in `pow`. A value of 0.04 is used in the revised example. Some iterations may be required to find a good value for `lambda`. Figure 2 shows that most observations fall on or near the theoretical line and more confidence can be placed in the p-value, because of the central-limit theorem, rejecting the null hypothesis at the 0.05 alpha level. Unfortunately, the difference between the September and June values cannot be quantified.

```
> # The revised parametric paired-data test, save for plot and print it
> Atra.pt <- censReg(pow(Sept, .04) - pow(June, .04) ~ 1)
> print(Atra.pt)


Call:
censReg(formula = pow(Sept, 0.04) - pow(June, 0.04) ~ 1)
```

```
Coefficients:
            Estimate Std. Error z-score p-value
(Intercept)    2.053      0.6472   3.172  0.0012

Estimated residual standard error (Unbiased) = 2.74
Distribution: normal
Number of observations = 24, number left censored = 0,
  number right censored = 0 (0 percent)

Loglik(model) = -39.19 Loglik(intercept only) = -39.19
  Chi-square = 0, degrees of freedom = 0, p-value = 1

Computation method: MLE
```

```
> # setSweave required for vignettes, use setPage or setPDF otherwise
> setSweave("graph02", 6 ,6)
> # graph number 2 is the probability plot
> plot(Atra.pt, which=2, set.up=FALSE)
> # Reuired to close graph for the vignette
> graphics.off()
```
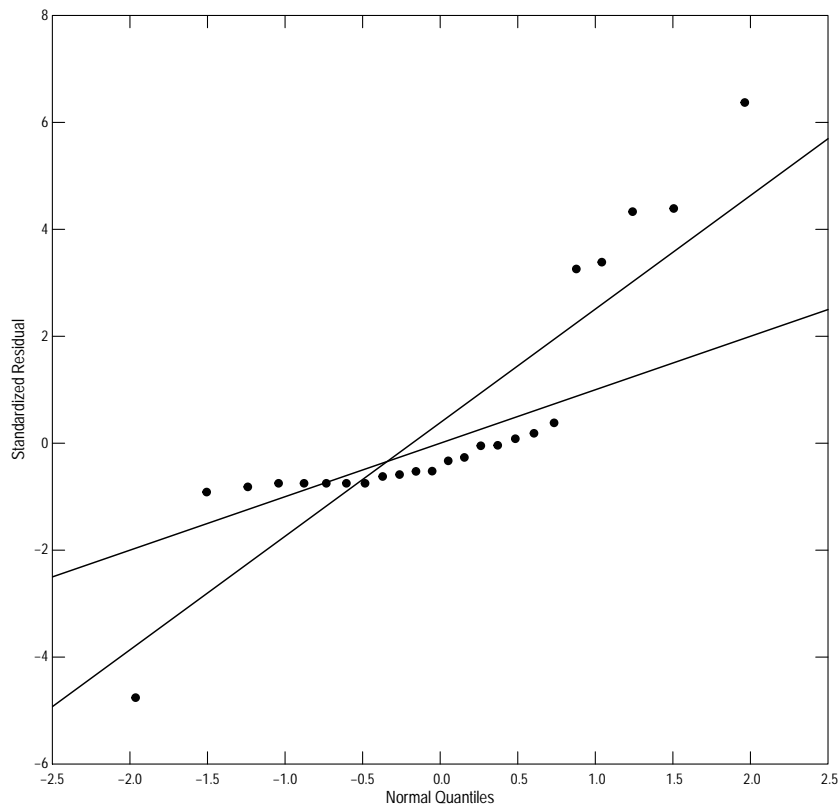
Figure 2. Probability plot to assess the normality of the transformed data.

2 The Nonparametric Paired-Prentice Wilcoxon Test

The nonparametric paired-sample test for censored data is known as the Paired-Prentice Wilcoxon (PPW) test. The test is described by Helsel (2012) and can be used for either single or multiple reporting limits. As with the Wilcoxon signed-rank test (Helsel and Hisrch, 2002), it is very sensitive to symmetric differences.

The PPW test is illustrated in the example below using the Atra data. For this test, the data must be nonnegative and left-censored only. The `ppw.test` function performs the test and is demonstrated in the code below. The results from the test, indicate that the null hypothesis should be rejected at the 0.05 level and that the concentration in September are larger than those in June, Z is positive.The graph (fig. 3) indicates that the differences are not symmetric, observation number 24 is a far outlier, but the results are consistent with the revised parametric test, so some confidence can be placed in the results.

```
> # First compute the left-censored values for June and September
> # Set the minimum value to 0 for censored values
> June <- with(Atra, as.lcens(June, censor.codes=JuneCen))
> Sept <- with(Atra, as.lcens(Sept, censor.codes=SeptCen))
> # The PPW test, save for plot and print it
> Atra.ppw <- ppw.test(Sept, June)
> print(Atra.ppw)


        paired Prentice-Wilcoxon test

data:  Sept and June
Paired Prentice Z = 3.0458, n = 24, p-value = 0.002321
alternative hypothesis: true difference is not equal to 0


> # setSweave required for vignettes, use setPage or setPDF otherwise
> setSweave("graph03", 6 ,6)
> # the graphs are the scaled and actual numeric differences
> plot(Atra.ppw, set.up=FALSE)
> # Reuired to close graph for the vignette
> graphics.off()
```
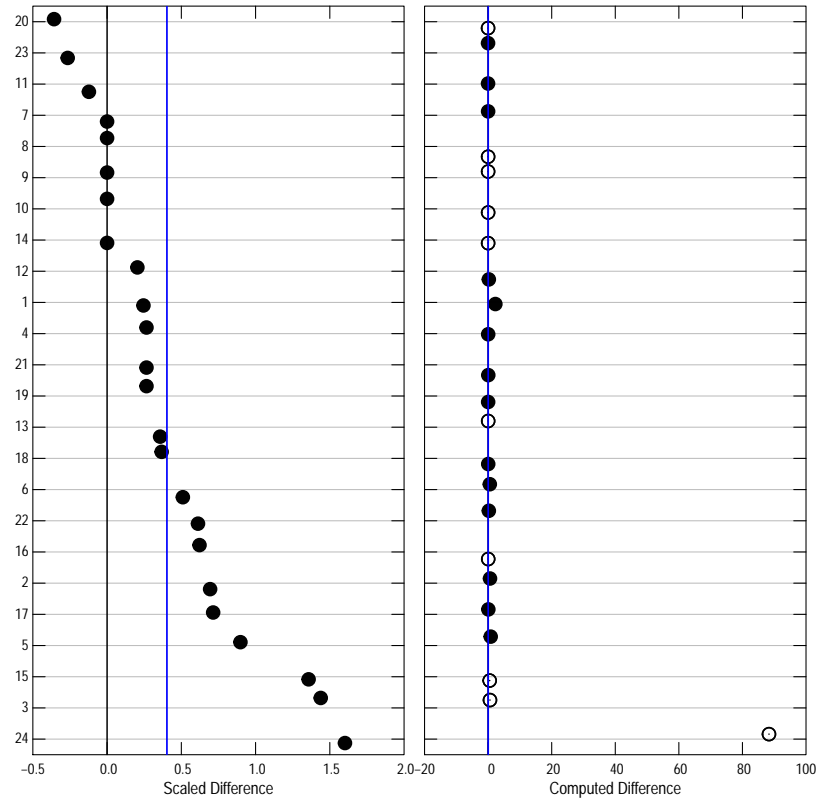
Figure 3. The scaled and numeric differences, sorted by the difference in scaled differences.

3 The Nonparametric Paired Pratt Test

An alternative nonparametric paired-sample test for censored data is the Wilcoxon rank-sum test using the Pratt adjustemnt for zero differences. The test was used by Lindsey and Rupert (2012) and can be used for either single or multiple reporting limits. The test was used by Lindsey and Rupert (2012) because it seemed less sensitive to assymmetric differences than the PPW test.

The test is illustrated in the example below using the Atra data. For this test, the data must be nonnegative and left-censored only. The `pairedPratt.test` function performs the test and is demonstrated in the code below. The results from the test, indicate that the null hypothesis should be rejected at the 0.05 level and that the concentration in September are larger than those in June, Z is positive.The graph (fig. 4) indicates that the differences are not symmetric, observation number 24 is a far outlier, but the results are consistent with the other tests, so some confidence can be placed in the results.

```
> # The left-censored values for June and September are from the previous example
> # The Pratt test, save for plot and print it
> Atra.prt <- pairedPratt.test(Sept, June)
> print(Atra.prt)


        Wilcoxon Rank-Sum Test with Pratt Adjustment for Zeros

data:  Sept and June
Z = 3.2925, p-value < 2.2e-16
alternative hypothesis: true difference is not equal to 0


> # setSweave required for vignettes, use setPage or setPDF otherwise
> setSweave("graph04", 6 ,6)
> # the graphs are the scaled and actual numeric differences
> plot(Atra.prt, set.up=FALSE)
> # Reuired to close graph for the vignette
> graphics.off()
```
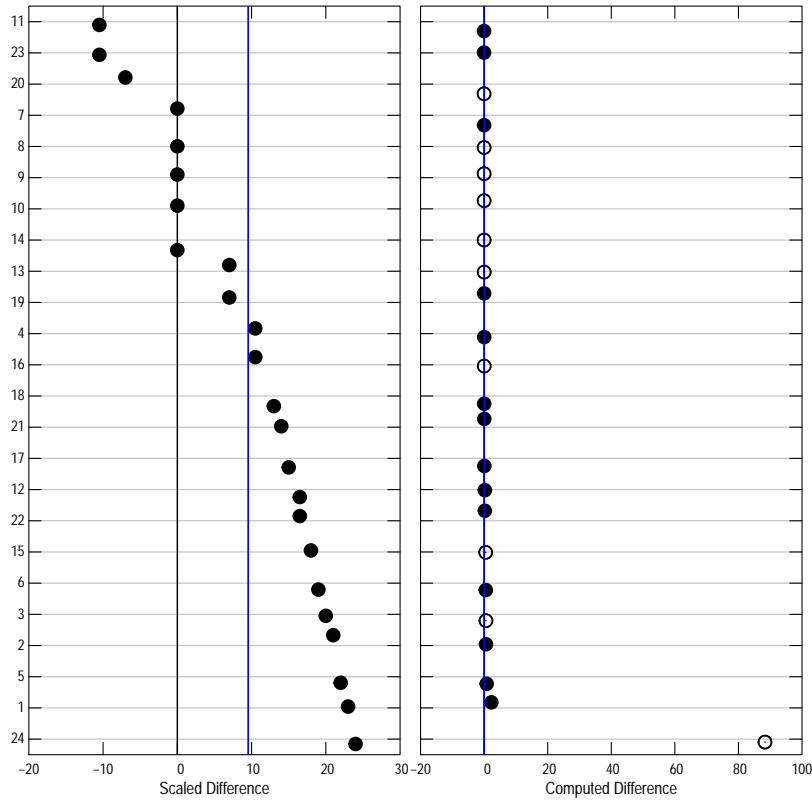
Figure 4. The scaled (signed rank) and numeric differences, sorted by the difference in scaled differences.

References

H12 Helsel, D.R. 2012, Statistics for Censored Environmental Data Using Minitab and R: New York, Wiley, 324 p. HH Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources: U.S. Geological Survey Techniques of Water-Resources Investigations, book 4, chap. A3, 522 p. LR Lindsey, B.D., and Rupert, M.G., 2012, Methods for evaluating temporal groundwater quality data and results of decadal-scale changes in chloride, dissolved solids, and nitrate concentrations in groundwater in the United States, 1988–2010: U.S. Geological Survey Scientific Investigations Report 2012–5049, 46 p.