

Practical Three

February 2018

1. Open your workspace from last weeks practical.
2. Attach the data set.

Detecting residual autocorrelation

We are going to examine if the residuals from the working model (`glmFitOD3`) are independent. We will start by generating values from an independent Normal distribution ($\mu = 0, \sigma = 1$), examining the frequency of runs and compare this frequency with the number of runs based on the residuals from our model.

- Generate 50 values at random from a Normal distribution:

```
testVals<- rnorm(50)
```

- Assign each value a positive or negative value, and plot these in sequence:

```
sign(testVals)
plot(sign(testVals), type="l")
```

- Plot the first 800 residuals for the working model in order (just for illustration - you can pick any sequence of residuals as long as the sequence is preserved):

```
#plot residuals for the working model in order:
plot(sign(residuals(glmFitOD3, type="pearson")[1:800]),
     type="l")
```

- Carry out a runs test for the working model using the Pearsons residuals in observation order:

```
require(lawstat)
runs.test(residuals(glmFitOD3, type="pearson"))
```

Diagnosing nonlinearities on the link scale

We are going to examine if the relationships on the link scale are reasonably linear by looking at the Pearsons residuals and by fitting quadratic terms for each covariate also in the model.

- Examine the Pearson's residuals in relation to Depth:

```
require(car)
residualPlots(glmFitOD3, quadratic=T, type = "pearson",
ylim=c(-20,20),terms = ~ Depth, fitted=FALSE)
```

- Re-order the residuals in order of Depth value, view the number of runs for the first 100 residuals and compared this number with 100 independent values:

```
dat$Depth[order(dat$Depth)][1:100]
par(mfrow=c(1,2))
plot(sign(residuals(glmFitOD3, type="pearson")[order(dat$Depth)])[1:100],
type="l", main="100 Residuals in Depth order", ylab="Pearsons Residuals")
plot(sign(rnorm(100)), type="l", main="Random values")
```

- Carry out a runs test for the working model using the Pearsons residuals ordered by Depth value. Use the following:

```
require(lawstat)
runs.test(residuals(glmFitOD3, type="pearson")[order(dat$Depth)])
```

- Examine the Pearson's residuals in relation to x.pos:

```
residualPlots(glmFitOD3, quadratic=T, type = "pearson",
              ylim=c(-20,20), terms = ~ x.pos, fitted=FALSE)
```

- Carry out a runs test for the working model using the Pearsons residuals ordered by x.pos value
- Examine the Pearson's residuals in relation to y.pos:

```
residualPlots(glmFitOD3, quadratic=T, type = "pearson",
              ylim=c(-20,20), terms = ~ y.pos, fitted=FALSE)
```

- Carry out a runs test for the working model using the Pearsons residuals ordered by y.pos value

Questions

1. Which of the following is TRUE?

- ✓ The Pearson's residuals from the working model exhibit less runs than the randomly sampled values (from the Normal distribution) which indicates the residuals are more correlated than the values sampled from the Normal.
- ✗ The Pearson's residuals from the working model exhibit more runs than the randomly sampled values (from the Normal distribution) which indicates the residuals are more correlated than the values sampled from the Normal.
- ✗ The Pearson's residuals from the working model exhibit less runs than the randomly sampled values (from the Normal distribution) which indicates the residuals are less correlated than the values sampled from the Normal.
- ✗ The Pearson's residuals from the working model exhibit more runs than the randomly sampled values (from the Normal distribution) which indicates the residuals are less correlated than the values sampled from the Normal.

2. Which of the following about the residuals from the working model (in observation order) is TRUE?

- ✓ The runs test based on the Pearson's residuals from the working model has a large negative test statistic and an associated small p-value; this signals positive correlation in model residuals.
- ✗ The runs test based on the Pearson's residuals from the working model has a large negative test statistic and an associated small p-value; this signals negative correlation in model residuals (since the test statistic is negative).
- ✗ The runs test based on the Pearson's residuals from the working model has a large negative test statistic and an associated large p-value; this signals positive correlation in model residuals.
- ✗ The runs test based on the Pearson's residuals from the working model has a large positive test statistic and an associated small p-value; this signals positive correlation in model residuals (since the test statistic is positive).
- ✗ The runs test based on the Pearson's residuals from the working model is inconclusive because the p-value is too small.

3. What is the runs test-statistic when the Pearson's residuals are ordered by Depth for the working model? Give your answer to 2 decimal places. **-66.63**

4. TRUE or FALSE? There is evidence for correlation in model residuals when considered in Depth order (and thus evidence for model misspecification for the depth relationship)?

5. What is the runs test-statistic when the Pearson's residuals are ordered by x.pos for the working model? Give your answer to 2 decimal places. **-35.72**

6. TRUE or FALSE? There is evidence for correlation in model residuals when considered in x.pos order (and thus evidence for model misspecification for the x.pos relationship)?

7. What is the runs test-statistic when the Pearson's residuals are ordered by y.pos for the working model? Give your answer to 2 decimal places. **-6.29**

8. TRUE or FALSE? There is evidence for correlation in model residuals when considered in y.pos order (and thus evidence for model misspecification for the y.pos relationship)?

9. TRUE or FALSE? Based on the Pearson's residual plots, there appears to be departures from linearity on the link scale for all three covariates; the Tukey tests for non-additivity for each covariate exhibit small p-values.

10. TRUE or FALSE? Based on the Pearson's residual plots, the quadratic and smoother-based terms relating the Pearson's residuals to each covariate exhibit more similar curvature for the x.pos and y.pos terms when compared with the curvature evidenced for the Depth relationship.