Lecture 29

Motivating example: earthquake data

We have data from the 2015 Gorkha earthquake in Nepal. After the earthquake, a large scale survey was conducted to determine the amount of damage the earthquake caused for homes, businesses and other structures. Variables include:

- `Damage`: the amount of damage suffered by the building (none, moderate, severe)
- `age`: the age of the building (in years)
- `condition`: a de-identified variable recording the condition of the land surrounding the building

Fisher scoring

Wald tests

```
          (Intercept) sqrt(age)  conditiono conditiont
moderate    0.6581163 0.3747641 -0.45376940 -0.5803708
severe      0.1881145 0.4251732  0.04706934 -0.4623774

          (Intercept)  sqrt(age) conditiono conditiont
moderate    0.1208913 0.01684468  0.2305975  0.1155475
severe      0.1243799 0.01725782  0.2292533  0.1180182
```

Suppose we want to know whether there is a relationship between age and the odds of moderate vs. no damage, after accounting for surface condition. What hypotheses would we test?

Wald tests

```
          (Intercept) sqrt(age)  conditiono conditiont
moderate    0.6581163 0.3747641 -0.45376940 -0.5803708
severe      0.1881145 0.4251732  0.04706934 -0.4623774

          (Intercept)  sqrt(age) conditiono conditiont
moderate    0.1208913 0.01684468  0.2305975  0.1155475
severe      0.1243799 0.01725782  0.2292533  0.1180182
```

Suppose we want to know whether the relationship between age and the odds of moderate vs. no damage is the *same* as the relationship between age and the odds of severe vs. no damage. What hypotheses would we test?

Wald tests

```r
diff <- t(c(0, -1, 0, 0, 0, 1, 0, 0)) %*%
  c(t(coef(m1)))
std_err <- sqrt(t(c(0, -1, 0, 0, 0, 1, 0, 0)) %*%
                  vcov(m1) %*%
                  c(0, -1, 0, 0, 0, 1, 0, 0))
(diff - 0)/std_err
```

```
        [,1]
[1,] 4.95677
```

```r
2*pnorm((diff - 0)/std_err, lower.tail = F)
```

```
             [,1]
[1,] 7.167478e-07
```

Likelihood ratio tests

```
          (Intercept) sqrt(age)  conditiono conditiont
moderate    0.6581163 0.3747641 -0.45376940 -0.5803708
severe      0.1881145 0.4251732  0.04706934 -0.4623774

          (Intercept)  sqrt(age) conditiono conditiont
moderate    0.1208913 0.01684468  0.2305975  0.1155475
severe      0.1243799 0.01725782  0.2292533  0.1180182
```

Suppose we want to know whether there is a relationship between surface condition and damage, after accounting for building age. What hypotheses would we test?

Likelihood ratio tests

```
m1 <- multinom(Damage ~ sqrt(age) + condition,
               data = earthquake)
m2 <- multinom(Damage ~ sqrt(age),
               data = earthquake)

pchisq(m2$deviance - m1$deviance, df = 4,
       lower.tail=F)
```

```
[1] 2.452814e-08
```

Class activity

https://sta712-f23.github.io/class_activities/ca_lecture_29.html