THE UNIVERSITY of EDINBURGH

Week 4: More Tests

Univariate Statistics and Methodology using R

Department of Psychology
The University of Edinburgh

Key Points

- Binomial Tests
- χ^2 Distribution
- χ^2 Tests
    - Goodness of Fit
    - Test of Independence

More Tests

- The *t*-test & *z*-test are useful in certain situations

More Tests

- The *t*-test & *z*-test are useful in certain situations
- These tests are used to:

More Tests

- The *t*-test & *z*-test are useful in certain situations
- These tests are used to:
  - Compare *continuous* data from two groups
    - Independent groups (independent-samples t-tests)
    - Dependent groups (paired-samples t-tests)

More Tests

- The *t*-test & *z*-test are useful in certain situations
- These tests are used to:
    - Compare *continuous* data from two groups
        - Independent groups (independent-samples t-tests)
        - Dependent groups (paired-samples t-tests)
    - Compare *continuous* data to a known value
        - A known μ and σ(z-test)
        - A known μ when σ is unknown (one-sample t-test)

More Tests

- The *t*-test & *z*-test are useful in certain situations
- These tests are used to:
  - Compare *continuous* data from two groups
    - Independent groups (independent-samples t-tests)
    - Dependent groups (paired-samples t-tests)
  - Compare *continuous* data to a known value
    - A known μ and σ(z-test)
    - A known μ when σ is unknown (one-sample t-test)
- However, there are times you'll want to conduct a statistical analysis on data that fall outside these parameters

More Tests

- Suppose you want to test data that are not continuous...
- In this case, you're probably interested in whether your data are distributed in the way you would expect
  - Are cat and dog lovers evenly distributed in this course?
  - Does opinion of the monarchy (approve/disapprove) differ across political parties?
  - Do people who vape develop lung disease in similar frequencies to those who smoke?

More Tests

- Suppose you want to test data that are not continuous...
- In this case, you're probably interested in whether your data are distributed in the way you would expect
    - Are cat and dog lovers evenly distributed in this course?
    - Does opinion of the monarchy (approve/disapprove) differ across political parties?
    - Do people who vape develop lung disease in similar frequencies to those who smoke?
- To do this, you can use tests that are specifically meant to compare frequencies of the categories within your data:
    - Binomial Tests
    - Chi-Square (χ^2) Tests

Part 1: Binomial Tests

Binomial tests

- **Binomial distribution:** probability distribution where just two possible outcomes have fixed probabilities

Binomial tests

- **Binomial distribution:** probability distribution where just two possible outcomes have fixed probabilities
- The probability of the two outcomes does not necessarily have to be equal, but must sum to one

Binomial tests

- **Binomial distribution:** probability distribution where just two possible outcomes have fixed probabilities
- The probability of the two outcomes does not necessarily have to be equal, but must sum to one
- The binomial distribution can be used to describe the counts of a variable that has two outcomes and an expectation of a specific proportion between the two

Binomial tests

- **Binomial distribution:** probability distribution where just two possible outcomes have fixed probabilities
- The probability of the two outcomes does not necessarily have to be equal, but must sum to one
- The binomial distribution can be used to describe the counts of a variable that has two outcomes and an expectation of a specific proportion between the two
- The binomial distribution describes data that follow these conditions:
    - The number of trials or observations (n) is fixed
    - Each observation is independent
    - Each observation has one of two outcomes
    - The probability of each outcome is consistent across observations

Can these be described by a binomial distribution?

Rule 1 The number of trials or observations (*n*) is fixed

Rule 2 Each observation is independent

Rule 3 Each observation has one of two outcomes

Rule 4 The probability of each outcome is consistent across observations

Consider the following scenarios, where X = random variable that represents success

- A fair coin is flipped 20 times; X = # of heads

Can these be described by a binomial distribution?

Rule 1 The number of trials or observations (*n*) is fixed

Rule 2 Each observation is independent

Rule 3 Each observation has one of two outcomes

Rule 4 The probability of each outcome is consistent across observations

Consider the following scenarios, where X = random variable that represents success

- A fair coin is flipped 20 times; X = # of heads
- Draw 3 cards at random, one after the other *without replacement*; X = # of diamonds

Can these be described by a binomial distribution?

Rule 1 The number of trials or observations (*n*) is fixed

Rule 2 Each observation is independent

Rule 3 Each observation has one of two outcomes

Rule 4 The probability of each outcome is consistent across observations

Consider the following scenarios, where X = random variable that represents success

- A fair coin is flipped 20 times; X = # of heads
- Draw 3 cards at random, one after the other *without replacement*; X = # of diamonds
- Draw 3 cards at random, one after the other *with replacement*, X = # of diamonds

Can these be described by a binomial distribution?

Rule 1 The number of trials or observations (*n*) is fixed

Rule 2 Each observation is independent

Rule 3 Each observation has one of two outcomes

Rule 4 The probability of each outcome is consistent across observations

Consider the following scenarios, where X = random variable that represents success

- A fair coin is flipped 20 times; X = # of heads
- Draw 3 cards at random, one after the other *without replacement*; X = # of diamonds
- Draw 3 cards at random, one after the other *with replacement*, X = # of diamonds
- The probability of having blood type B is 0.1. Choose 4 people at random; X = # with blood type B

Can these be described by a binomial distribution?

Rule 1 The number of trials or observations (*n*) is fixed

Rule 2 Each observation is independent

Rule 3 Each observation has one of two outcomes

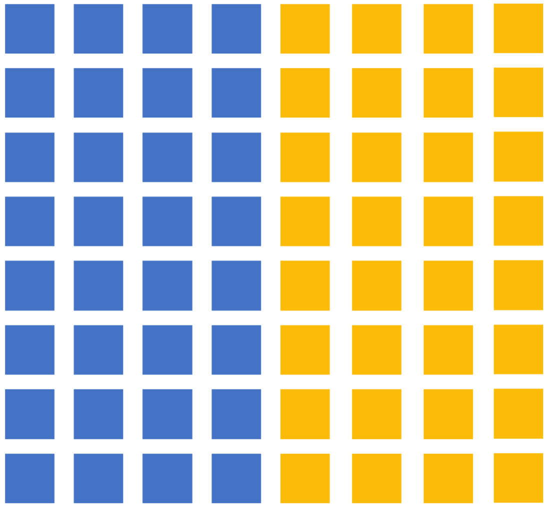
Rule 4 The probability of each outcome is consistent across observations

Consider the following scenarios, where X = random variable that represents success

- A fair coin is flipped 20 times; X = # of heads
- Draw 3 cards at random, one after the other *without replacement*; X = # of diamonds
- Draw 3 cards at random, one after the other *with replacement*, X = # of diamonds
- The probability of having blood type B is 0.1. Choose 4 people at random; X = # with blood type B
- A student answers 10 quiz questions at random; the first five are true/false, the second five are multiple choice, with four options each. X = # of correct answers

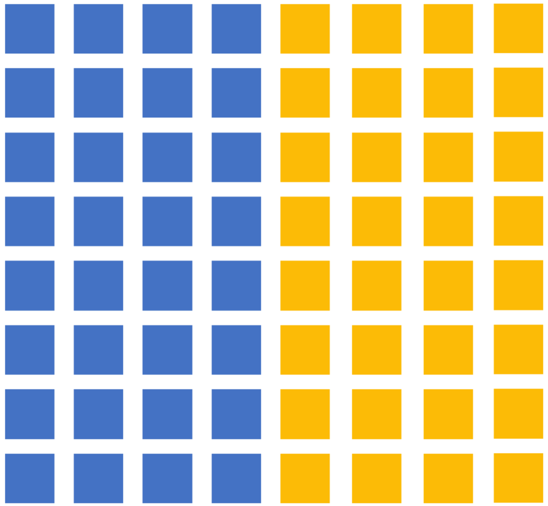
Binomial Test

- Calculates the probability of getting a proportion as extreme as or more extreme than the value measured, given that the expected proportion reflects the ground truth.



Binomial Test

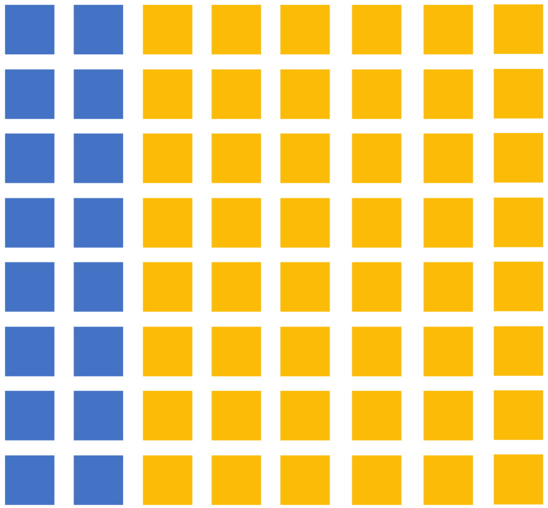
- Calculates the probability of getting a proportion as extreme as or more extreme than the value measured, given that the expected proportion reflects the ground truth.



- How likely are we to draw the following samples, given the above population?

Binomial Test

- Calculates the probability of getting a proportion as extreme as or more extreme than the value measured, given that the expected proportion reflects the ground truth.



- How likely are we to draw the following samples, given the above population?

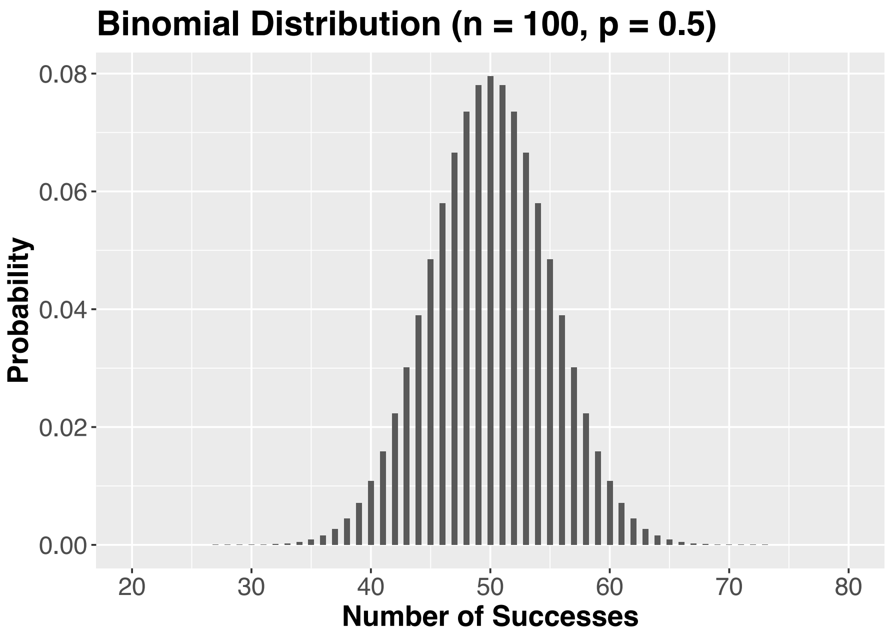
- Binomial test tells you whether your sample deviates significantly from your expectations.

Binomial Distribution

If you toss a coin 100 times, what's the likelihood of success (success = heads)?

Binomial Distribution

If you toss a coin 100 times, what's the likelihood of success (success = heads)?



Binomial Test

- If you toss a coin 4 times, what's the probability of it landing on heads at least 3 times?
- 2 possible outcomes for each of the 4 tosses
  - $2^4 = 16$ possible sequences of outcomes
- Of those 16, 5 outcomes include $\geq$ 3 heads
- $p = 5/16 = .3125$

Toss 1	Toss 2	Toss 3	Toss 4	Total Heads
H	H	H	H	4
T	H	H	H	3
H	T	H	H	3
T	T	H	H	2
H	H	T	H	3
T	H	T	H	2
H	T	T	H	2
T	T	T	H	1
H	H	H	T	3
T	H	H	T	2
H	T	H	T	2
T	T	H	T	1
H	H	T	T	2
T	H	T	T	1
H	T	T	T	1
T	T	T	T	0

Binomial Test

```r
binom.test(3, 4, 0.5, alternative="greater")
```

```
##
##      Exact binomial test
##
## data:  3 and 4
## number of successes = 3, number of trials = 4, p-value = 0.3125
## alternative hypothesis: true probability of success is greater than 0.5
## 95 percent confidence interval:
##  0.2486046 1.0000000
## sample estimates:
## probability of success
##                   0.75
```

- Don't be fooled by the the *probability of success* line (which is just 3/4)

Binomial Test

Approximately 9% of the world's population have blue eyes; is the USMR class of 2022-23 a representative sample?

Binomial Test

Approximately 9% of the world's population have blue eyes; is the USMR class of 2022-23 a representative sample?

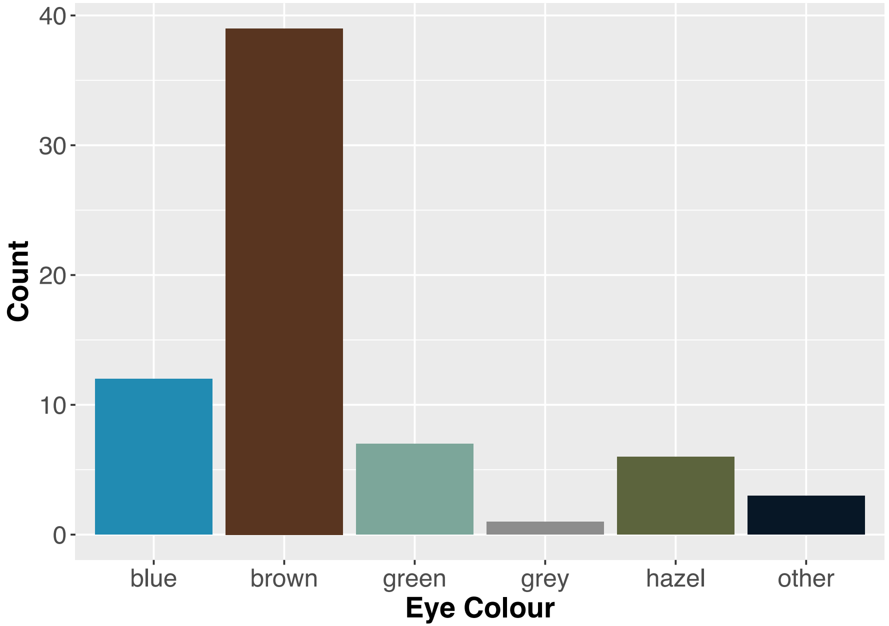
- "Two possible outcomes"
- "An expectation of a specific proportion between the two"
- The number of trials or observations (n) is fixed
- Each observation is independent
- Each observation has one of two outcomes
- The probability of each outcome is consistent across observations

Binomial Test in R

Eye Colours for USMR:

```
eyes <- table(statsClasses$eyecolour[statsClasses$course=='usmr' & sta
eyes
```

```
##
##  blue brown green  grey hazel other
##    12    39     7     1     6     3
```



Binomial Test in R

Approximately 9% of the world's population have blue eyes; is the USMR class of 2022-23 a representative sample?

```
binom.test(eyes['blue'],sum(eyes),0.09,alternative="two.sided")
```

```
##
##      Exact binomial test
##
## data:  eyes["blue"] and sum(eyes)
## number of successes = 12, number of trials = 68, p-value = 0.02
## alternative hypothesis: true probability of success is not equal to 0.09
## 95 percent confidence interval:
##  0.09465 0.28797
## sample estimates:
## probability of success
##                 0.1765
```

Part 2: The χ^2 distribution

But first, Ignaz Semmelweis

- Doctor in maternity clinic in 1846
- Discovered women in clinics staffed by doctors and med students were dying much more often than those in midwives' clinics
- We'll use hypothetical data adapted from Semmelweis's study to demonstrate a χ^2 test

```
## # A tibble: 3 × 3
## # Groups:   Setting [3]
##   Setting           DeathCount LifeCount
##   <chr>                  <dbl>     <dbl>
## 1 Hospital                  13        87
## 2 Teaching Hospital         18        82
## 3 Midwifery Unit             2        98
```

The χ^2 distribution

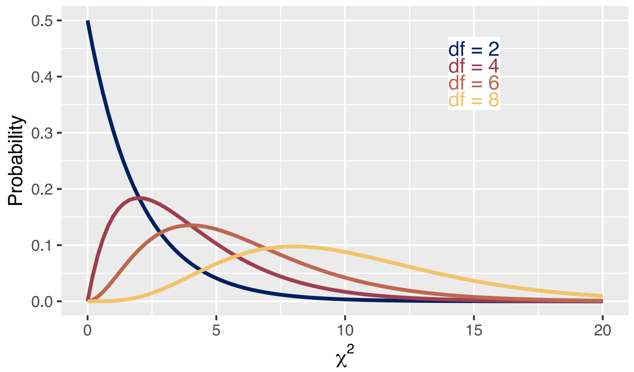
- Sometimes you need to test the frequencies of variables that have more than 2 levels

The χ^2 distribution

- Sometimes you need to test the frequencies of variables that have more than 2 levels
- Similar to previous tests with continuous variables, a test statistic is computed and compared to a cutoff within a larger distribution

The χ^2 distribution

- Sometimes you need to test the frequencies of variables that have more than 2 levels
- Similar to previous tests with continuous variables, a test statistic is computed and compared to a cutoff within a larger distribution
  - Like the *t*-distribution, the comparison distribution is dependent upon the degrees of freedom
  - Unlike the *t*-distribution, the degrees of freedom isn't dependent on sample size, but on the number of comparison groups



χ^2 Statistic

$$\chi^2 = \Sigma \frac{(O-E)^2}{E}$$

- Σ = sum up
- O = Observed Cases
- E = Expected Cases

χ^2 Statistic

$$\chi^2 = \Sigma \frac{(O - E)^2}{E}$$

- Σ = sum up
- O = Observed Cases
- E = Expected Cases

- Once you've computed the statistic, you need to compare it to the proper distribution, but first you'll need the degrees of freedom

χ^2 Degrees of Freedom

- The formula for df depends on the question you're asking and the test you're using...

χ^2 Degrees of Freedom

- The formula for df depends on the question you're asking and the test you're using...

Goodness of Fit

$$\chi^2 = \sum_k \frac{(O-E)^2}{E}$$

- k = the number of categories
- Checks whether your data come from an expected distribution
- $df = k - 1$

χ^2 Degrees of Freedom

- The formula for df depends on the question you're asking and the test you're using...

Goodness of Fit

$$\chi^2 = \sum_k \frac{(O-E)^2}{E}$$

- *k* = the number of categories
- Checks whether your data come from an expected distribution
- *df = k - 1*

Test of Homogeneity/Independence

$$\chi^2 = \sum_{ij} \frac{(O-E)^2}{E}$$

- *i* = # of rows
- *j* = # of columns
- Checks whether different groups have the same distribution of a categorical variable (*homogeneity*) or whether categorical variables are associated with each other within a population (*independence*)
- *df = (i-1)(j-1)*

Which test do we need for our Semmelweis data?

Is maternal mortality rate consistent across healthcare setting?

```
## # A tibble: 3 × 3
## # Groups:   Setting [3]
##   Setting           DeathCount LifeCount
##   <chr>                  <dbl>     <dbl>
## 1 Hospital                  13        87
## 2 Teaching Hospital         18        82
## 3 Midwifery Unit             2        98
```

Goodness of Fit Do the data come from a specific distribution?

Test of Homogeneity Do the groups have the same distribution on a variable of interest?

Test of Independence Are the categorical variables associated with each other within a single population?

Computing χ^2

$$\chi^2 = \sum_{ij} \frac{(O-E)^2}{E}$$

$$E = \frac{(row\ total)(column\ total)}{total\ number\ measured}$$

Observed Data

```
##             Setting DeathCount LifeCount RowTotal
## 1          Hospital         13        87      100
## 2 Teaching Hospital         18        82      100
## 3    Midwifery Unit          2        98      100
## 4       ColumnTotal         33       267      300
```

Expected Data

```
##             Setting DeathCount LifeCount
## 1          Hospital         11        89
## 2 Teaching Hospital         11        89
## 3    Midwifery Unit         11        89
```

Computing χ^2

$$\chi^2 = \sum_{ij} \frac{(O-E)^2}{E}$$

Observed

```
##              Setting DeathCount LifeCount
## 1           Hospital         13        87
## 2 Teaching Hospital          18        82
## 3     Midwifery Unit          2        98
```

Expected

```
##              Setting DeathCount LifeCount
## 1           Hospital         11        89
## 2 Teaching Hospital          11        89
## 3     Midwifery Unit         11        89
```

```
##              Setting DeathCount LifeCount
## 1           Hospital     0.3636   0.04494
## 2 Teaching Hospital      4.4545   0.55056
## 3     Midwifery Unit     7.3636   0.91011
```

Computing χ^2

$$\chi^2 = \sum_{ij} \frac{(O-E)^2}{E}$$

Observed

```
##              Setting DeathCount LifeCount
## 1           Hospital         13        87
## 2 Teaching Hospital         18        82
## 3     Midwifery Unit          2        98
```

Expected

```
##              Setting DeathCount LifeCount
## 1           Hospital         11        89
## 2 Teaching Hospital         11        89
## 3     Midwifery Unit         11        89
```

```
##              Setting DeathCount LifeCount
## 1           Hospital     0.3636   0.04494
## 2 Teaching Hospital     4.4545   0.55056
## 3     Midwifery Unit     7.3636   0.91011
```

```
sum(fullDat[,2:3])
```

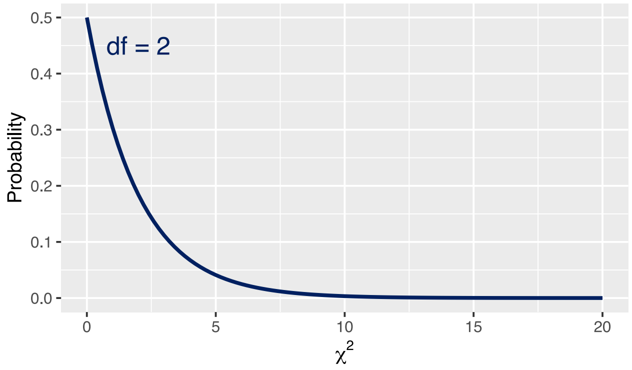
```
## [1] 13.69
```

Interpreting χ^2

- $df = (i-1)(j-1)$
- $df = (3-1)(2-1)$
- $df = 2$

Interpreting χ^2

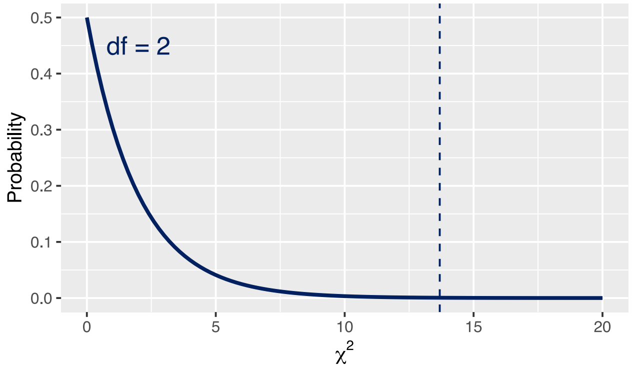
- $df = (i-1)(j-1)$
- $df = (3-1)(2-1)$
- $df = 2$



Interpreting χ^2

- $df = (i-1)(j-1)$
- $df = (3-1)(2-1)$
- $df = 2$

There is a significant difference between healthcare settings in mortality rate after birth, $\chi^2(2, N = 300) = 13.69, p = .001$




```
pchisq(13.69, 2, lower.tail=FALSE)
```

```
## [1] 0.001065
```

Computing χ^2 in R

```
sDat[,2:3]
```

```
##   DeathCount LifeCount
## 1         13        87
## 2         18        82
## 3          2        98
```

```
chisq.test(sDat[,2:3])
```

```
##
##      Pearson's Chi-squared test
##
## data:  sDat[, 2:3]
## X-squared = 14, df = 2, p-value = 0.001
```

One more small example...

- Imagine that Semmelweis wondered whether there were certain days of the week when maternal mortality was higher than others.
- Which test do we need for these data?

Is the maternal mortality rate consistent throughout the week?

Goodness of Fit Do the data come from a specific distribution?

Test of Homogeneity Do the groups have the same distribution on a variable of interest?

Test of Independence Are the categorical variables associated with each other within a single population?

χ^2 - Goodness of Fit

Let's imagine that Semmelweis spent some time gathering data on daily mortality rates:

```
##         Day weekMR
## 1    Monday     26
## 2   Tuesday     31
## 3 Wednesday     45
## 4  Thursday     37
## 5    Friday     29
## 6  Saturday     32
## 7    Sunday     44
```

If there was absolutely no difference in mortality rates across days of the week, we would expect the values for each day to be equal.

```
gofDat$ExpMR <- rep(round((sum(gofDat$weekMR)/7), 2), 7)
```

```
##         Day weekMR ExpMR
## 1    Monday     26 34.86
## 2   Tuesday     31 34.86
## 3 Wednesday     45 34.86
## 4  Thursday     37 34.86
## 5    Friday     29 34.86
## 6  Saturday     32 34.86
## 7    Sunday     44 34.86
```

Computing χ^2 for Goodness of Fit Test

$$\chi^2 = \sum_k \frac{(O-E)^2}{E}$$

```
##          Day weekMR ExpMR ObsMinExp
## 1     Monday     26 34.86     -8.86
## 2    Tuesday     31 34.86     -3.86
## 3  Wednesday     45 34.86     10.14
## 4   Thursday     37 34.86      2.14
## 5     Friday     29 34.86     -5.86
## 6   Saturday     32 34.86     -2.86
## 7     Sunday     44 34.86      9.14
```

Computing χ^2 for Goodness of Fit Test

$$\chi^2 = \sum_k \frac{(O-E)^2}{E}$$

```
##         Day weekMR ExpMR ObsMinExp SqDiff
## 1    Monday     26 34.86     -8.86  78.50
## 2   Tuesday     31 34.86     -3.86  14.90
## 3 Wednesday     45 34.86     10.14 102.82
## 4  Thursday     37 34.86      2.14   4.58
## 5    Friday     29 34.86     -5.86  34.34
## 6  Saturday     32 34.86     -2.86   8.18
## 7    Sunday     44 34.86      9.14  83.54
```

Computing χ^2 for Goodness of Fit Test

$$\chi^2 = \sum_k \frac{(O-E)^2}{E}$$

```
##         Day weekMR ExpMR ObsMinExp SqDiff SqDiffdivExp
## 1    Monday     26 34.86     -8.86  78.50       2.2519
## 2   Tuesday     31 34.86     -3.86  14.90       0.4274
## 3 Wednesday     45 34.86     10.14 102.82       2.9495
## 4  Thursday     37 34.86      2.14   4.58       0.1314
## 5    Friday     29 34.86     -5.86  34.34       0.9851
## 6  Saturday     32 34.86     -2.86   8.18       0.2346
## 7    Sunday     44 34.86      9.14  83.54       2.3964
```

Computing χ^2 for Goodness of Fit Test

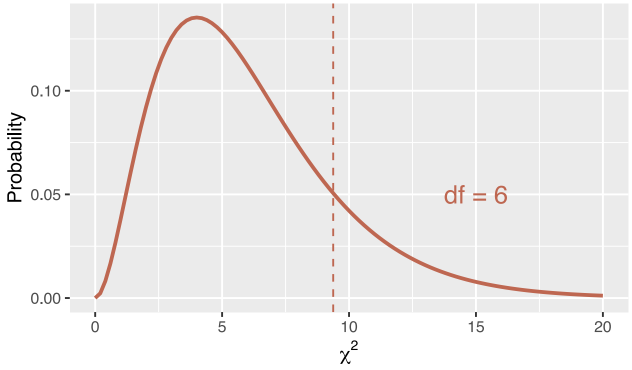
$$\chi^2 = \sum_k \frac{(O-E)^2}{E}$$

```
##         Day weekMR ExpMR ObsMinExp SqDiff SqDiffdivExp
## 1    Monday     26 34.86     -8.86  78.50       2.2519
## 2   Tuesday     31 34.86     -3.86  14.90       0.4274
## 3 Wednesday     45 34.86     10.14 102.82       2.9495
## 4  Thursday     37 34.86      2.14   4.58       0.1314
## 5    Friday     29 34.86     -5.86  34.34       0.9851
## 6  Saturday     32 34.86     -2.86   8.18       0.2346
## 7    Sunday     44 34.86      9.14  83.54       2.3964
```

```
## [1] 9.376
```

Interpreting χ^2 - Goodness of Fit

- $df = k - 1$
- $df = 7 - 1$
- $df = 6$




```r
pchisq(9.37, 6, lower.tail=FALSE)
```

```
## [1] 0.1538
```

Computing χ^2 Goodness of Fit in R

```
chisq.test(gofDat$weekMR, p = rep(1/7, 7))
```

```
##
##     Chi-squared test for given probabilities
##
## data:  gofDat$weekMR
## X-squared = 9.4, df = 6, p-value = 0.2
```

How did it end for Semmelweis?

- He noticed that a major difference in the two settings was that the doctors and residents were doing a lot of autopsies...

How did it end for Semmelweis?

- He noticed that a major difference in the two settings was that the doctors and residents were doing a lot of autopsies...
- AND NOT WASHING THEIR HANDS OR INSTRUMENTS

How did it end for Semmelweis?

- He noticed that a major difference in the two settings was that the doctors and residents were doing a lot of autopsies...
- AND NOT WASHING THEIR HANDS OR INSTRUMENTS
- Semmelweis was a handwashing pioneer - he recommended that doctors use a chlorine solution to cleanse their hands and instruments.
- So it was a great victory for Semmelweis, right?

How did it end for Semmelweis?

- He noticed that a major difference in the two settings was that the doctors and residents were doing a lot of autopsies...
- AND NOT WASHING THEIR HANDS OR INSTRUMENTS
- Semmelweis was a handwashing pioneer - he recommended that doctors use a chlorine solution to cleanse their hands and instruments.
- So it was a great victory for Semmelweis, right?
- No.

How did it end for Semmelweis?

- He noticed that a major difference in the two settings was that the doctors and residents were doing a lot of autopsies...
- AND NOT WASHING THEIR HANDS OR INSTRUMENTS
- Semmelweis was a handwashing pioneer - he recommended that doctors use a chlorine solution to cleanse their hands and instruments.
- So it was a great victory for Semmelweis, right?
- No.
- "He too was attacked widely by the establishment of obstetricians in Europe, who could not believe that they or their midwife colleagues were responsible for the enormous number of deaths." (Chamberlain, 2006)

End