Lecture 5

Proportion data: chi-square tests

Chi-square tests = proportion tests

- Sometimes data are presented as proportions:

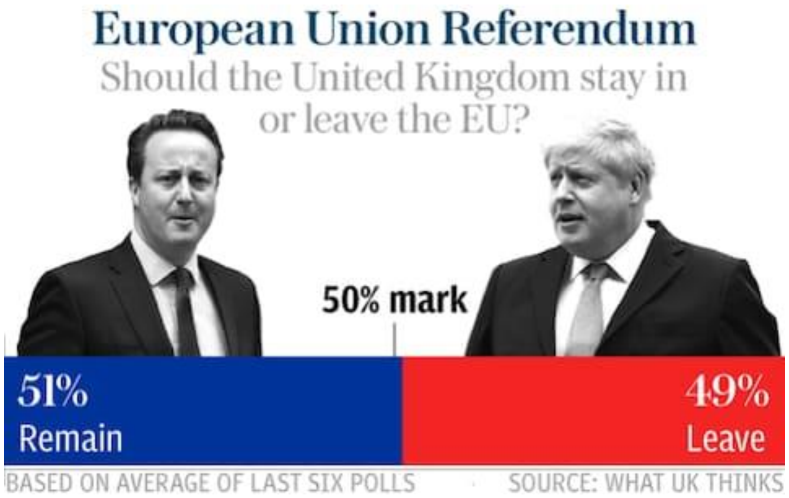
- ***One-sample proportion***
  - Proportion of women and men in the UK

    (**NOTE**: only one 'independent' proportion here: if women are 52%, men must be 48% - they must add up to 100%)
  - Proportion of children vs. adults in a village

- ***Two independent proportions***
  - Proportion of Labour supporters in London vs. Glasgow

  - (**NOTE**: here, proportions do not need to add up to 100%

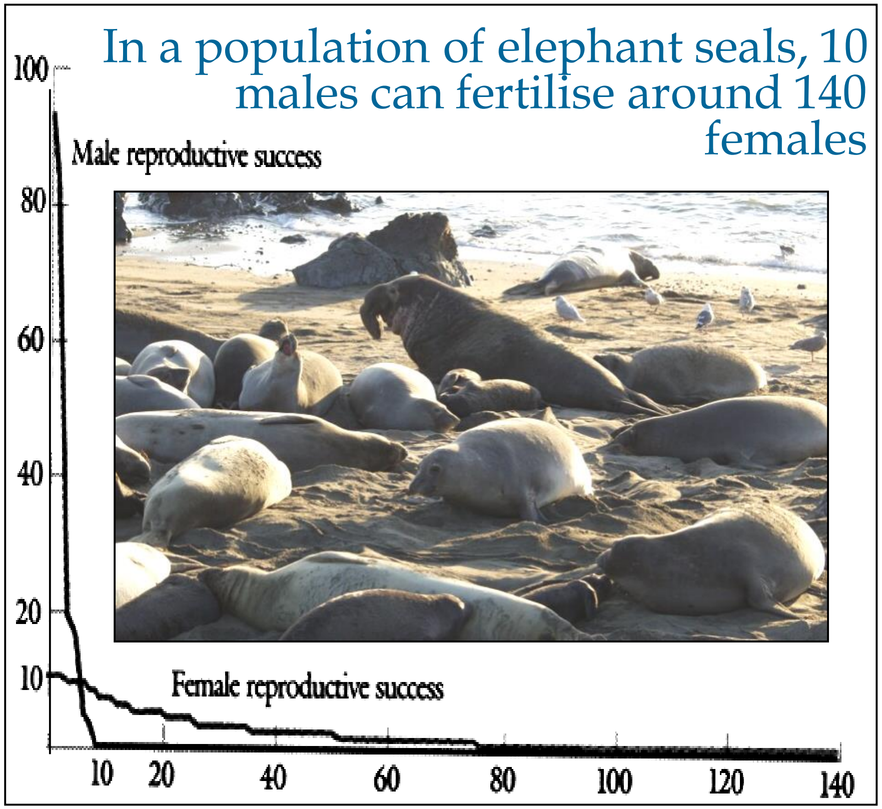
- To analyse proportions, we use *proportion tests* (aka *chi-square tests*)



Proportion test, one sample

- Example: sex ratios at birth

- Sex ratios are intriguing
  - In sexual species, a single male is able to fertilise many (if not all) females in a population
  - In many species, most males never mate!
  - But sex ratios remain ~1; why so many useless males?



Why are sex ratios generally balanced?

- If there are 100 offspring in a population, only 1 male (the father of all) and 100 females:
  - average female fitness= 1 offspring per female
  - average male fitness=100 offspring per male

  -> average male fitness is 100 times higher

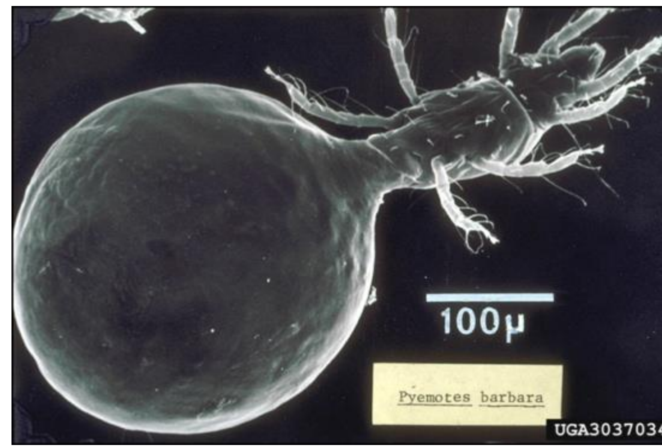
- Selection favours the fittest, i.e. the rarest sex
  - when males are rare, male offspring is favoured
  - when females are rare, female offspring is favoured
  - -> population remains close to a balanced 1:1 sex ratio

- This is an example of *frequency-dependent selection*

Deviations from balanced sex ratio

- If your offspring does not compete with others, sex ratio may change

- Example: viviparous mites
  - brothers inseminate sisters inside mother
  - observed sex ratio: ~ 1 male: 22 females



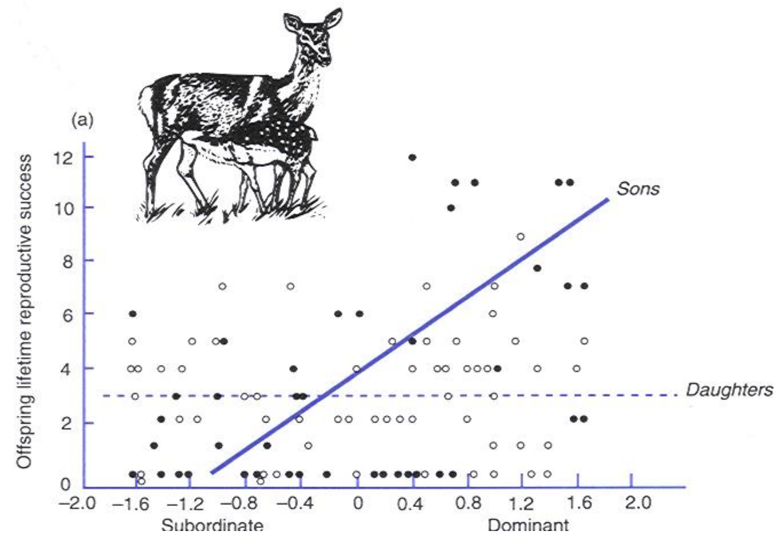
Sex ratio vs. social status

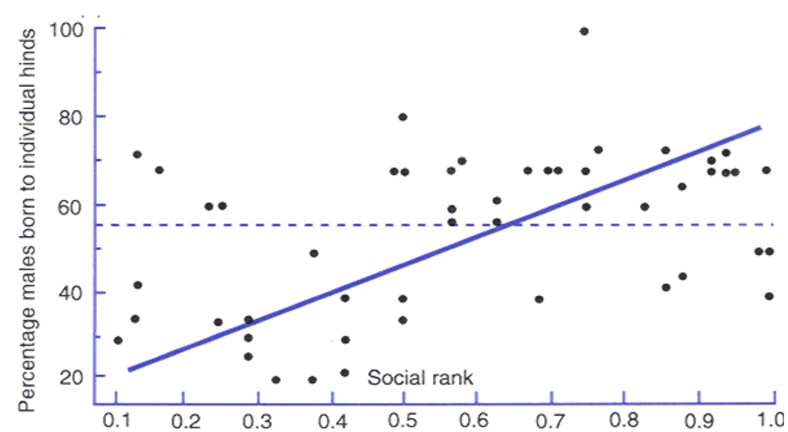
- Offspring fitness also depends on social status

- In polygynous species (most mammals) only a few high-quality males reproduce (competition for harems is tough)

- In red deer:
  - dominant females have more sons (more likely to be future harem leaders) than daughters
  - subordinate females have more daughters (they can still be in a harem and reproduce) than sons

- But total sex ratio (dominant + subordinate) remains close to 1:1



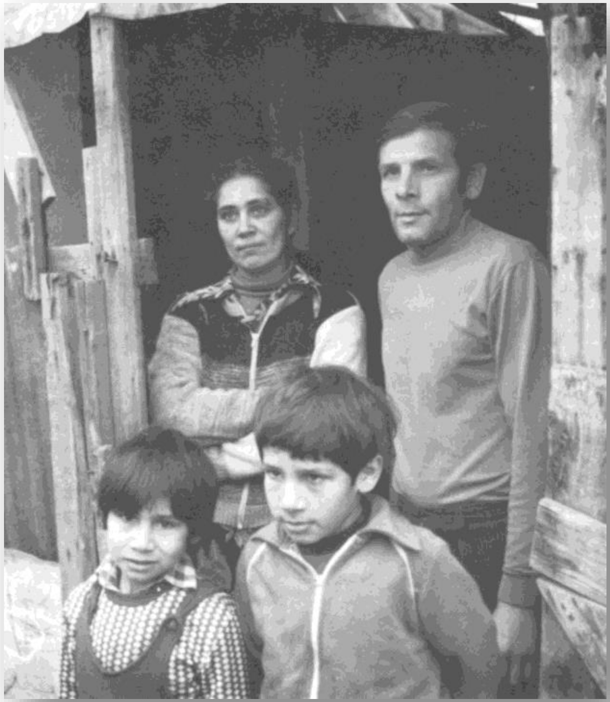


Sex ratios in humans

- As a rule, sex ratio at birth is ~1 in humans
  - slightly male-based at ~1.05, to compensate for higher male infant mortality

- But Berenczi and Dunbar (1997) identified a female-biased ratio at birth among Hungarian gypsies
  - rural and urban gypsies are poorer than neighbour Hungarians
  - daughters often marry richer Hungarians; sons very unlikely to marry out
  - Roma prefer baby girls to baby boys

- their prediction: sex ratio *at birth* should be biased towards females
- (note: sex ratio at birth is unaffected by infanticide etc.)



- But let's re-examine the evidence presented in the study

- In Hungarian gypsies, is sex ratio at birth significantly different from p=0.5?

Table 2. *Sex ratios at birth for each population*

	number of sons per 100 daughters			
	rural populations		urban populations	
	Gypsy	Hungarian	Gypsy	Hungarian
A. all children				
sample size	254	216	239	224
males/100 females	89.3	111.8	89.7	113.3
B. first-born children only				
sample size	87	85	77	102
males/100 females	81.3	157.6	94.3	131.8

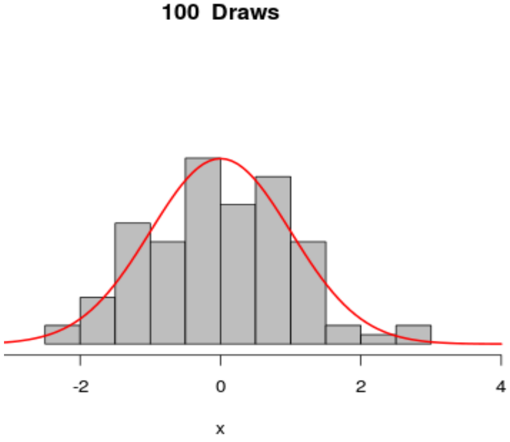
note: in this lecture, *p* is proportion; *P* is significance level
(careful: *R* output does not make the distinction!)

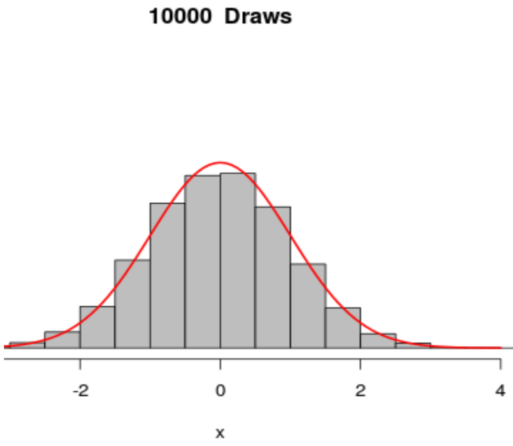
One-sample proportion test

- =tests the likelihood of a proportion p deviating from a test proportion
- often, test proportion is 50%

- Test is based on an approximation to the binary distribution that estimates likelihood of x positives out of n attempts
  - e.g. coin tossing: binary distribution estimates probability of x heads, each one with p(head)=0.5, in n tosses
  - expected mean: np (=10*0.5=5, i.e. you expect 5 heads in 10 tosses); variance: np(1-p)

- Test statistic: u= $\frac{x-np}{\sqrt{np(1-p)}}$
  - u statistic similar to t
  - u distribution approaches normal with large n
  - (alternative u^2, which has a chi-square distribution that also approaches normal with large n)





Syntax: elicit p= 0.5

> prop.test(x=positives, n=total, p=proportion tested)

(default: p=0.5)

Gypsy sex ratios

- So are there fewer gypsy boys than girls at birth?

1) Rural gypsies, *all babies*:

- sex ratio=89.3%=0.893, n=254
  =>120 boys, 134 girls, (120/134=0.893)
  - Proportion of boys = 120/254 = 0.47
  - Test proportion: 0.5 (i.e equal proportions)

  - **Question: is p=0.47 different from 0.5?**

Table 2. *Sex ratios at birth for each population*

	number of sons per 100 daughters			
	rural populations		urban populations	
	Gypsy	Hungarian	Gypsy	Hungarian
A. all children				
sample size	254	216	239	224
males/100 females	89.3	111.8	89.7	113.3
B. first-born children only				
sample size	87	85	77	102
males/100 females	81.3	157.6	94.3	131.8

```
> prop.test(120, 254, 0.5)
1-sample proportions test with continuity correction
data:  120 out of 254, null probability 0.5
X-squared = 0.6654, df = 1, p-value = 0.4147
alternative hypothesis: true p is not equal to 0.5
95 percent confidence interval:
 0.4099875 0.5357417
sample estimates:
      p
0.4724409
```

Null hypothesis: proportion=0.5

- P=0.41
- proportion *p*=0.5 included in 95% CI
- =>no significant deviation from p=5 and balanced 1:1 sex ratio
- ***no evidence of fewer boys than girls in rural Roma***

Note: calculating boys (b) and girls (g)

- Table shows:
  - Sample size n = b + g
  - (males/females) x 100 = b/g x 100 (i.e. shown as percentage)

- If n=254 and male/female = 89.3, then
  - b + g = 254
  - b/g = 0.893

i.e.

- g = n/(1 + sex ratio)

Exercise:

Run the same test with urban gypsies, all children (n=239, ratio=0.897)

Does the sex ratio at birth differ from p=0.5?

prop.test(113,239)
p= 0.47 >>> p>0.05
no significant difference

First-borns only

- But decision to have a second baby or abort may depend on sex of previous offspring
  - if 1st child is boy, Roma parents are more likely to try again (until child is female; if girl, they are more likely to stop

- Solution: analyse only *1st-born* rural babies:
  - Sex ratio=0.813, n=87 => boys=39, girls=48
  - *p*(boys)=39/87=0.448

Table 2. *Sex ratios at birth for each population*

	number of sons per 100 daughters			
	rural populations		urban populations	
	Gypsy	Hungarian	Gypsy	Hungarian
A. all children				
sample size	254	216	239	224
males/100 females	89.3	111.8	89.7	113.3
B. first-born children only				
sample size	87	85	77	102
males/100 females	81.3	157.6	94.3	131.8

```
> prop.test(39, 87, 0.5)
1-sample proportions test with continuity correction
data:  39 out of 87, null probability 0.5
X-squared = 0.7356, df = 1, p-value = 0.3911
alternative hypothesis: true p is not equal to 0.5
95 percent confidence interval:
 0.3427920 0.5583697
sample estimates:
       p
0.4482759
```

- P=0.39
- 95% CI includes test value *p*=0.5
- No evidence of fewer boys in rural Gypsy 1st-borns either

- ***One-sample proportion tests show no evidence that rural gypsies have female-biased sex ratio***

Exercise:

Run the same test with urban gypsies, first-born children only (n=77, ratio=0.943)

Does the sex ratio at birth differ from p=0.5?

Prop.test(37,77)
P= 0.48 >>> p>0.05
No significant
difference

Two independent proportions

- But proportion of boys in rural gypsies vs. Hungarians could still differ *from each other*
  - i.e. we can compare the two independent proportions of boys in rural Gypsies vs. Hungarians
- We use *two-sample proportion tests* in this case
  - Based on *u*-statistic (=difference between proportion divided by pooled variance) and chi-square distribution

Table 2. *Sex ratios at birth for each population*

	number of sons per 100 daughters			
	rural populations		urban populations	
	Gypsy	Hungarian	Gypsy	Hungarian
A. all children				
sample size	254	216	239	224
males/100 females	89.3	111.8	89.7	113.3
B. first-born children only				
sample size	87	85	77	102
males/100 females	81.3	157.6	94.3	131.8

Syntax:

```
> prop.test(c(x1, x2), c(n1, n2))
```

- *xi*= positive cases (=boys) in group 1 and group 2
- *ni*=total cases (=boys+girls) in group 1 and group 2

All babies

```
> prop.test(c(120,114), c(254,216))

2-sample test for equality of proportions with continuity correction

data:  c(120, 114) out of c(254, 216)
X-squared = 1.2171, df = 1, p-value = 0.2699
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.15018442  0.03951076
sample estimates:
   prop 1    prop 2
0.4724409 0.5277778
```

For all babies

Rural gypsies: 254 total, 120 boys; *p*(boys)=0.47

Rural Hungarians: ratio 1.118, 216 total -> 114 boys, 102 girls; *p*(boys)=0.527

Result:
P>0.26
95% CI includes zero (no difference)

Conclusion: no significant difference in sex ratios between rural populations

Exercise:

Run the same test two independent proportions test on ~~rural~~ urban children

Does the sex ratio at birth differ between Gypsies and Hungarians?

First births only

- But now take only first births:
  - Rural gypsies: 39 boys, 48 girls, n=87
  - Rural Hungarians: 52 boys, 33 girls, n=85

```
> prop.test(c(39,52), c(87,85))

2-sample test for equality of proportions
with continuity correction

data:  c(39, 52) out of c(87, 85)
X-squared = 3.9795, df = 1, p-value = 0.04606
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.322272741     -0.004704946
sample estimates:
   prop 1          prop 2
0.4482759    0.6117647
```

Table 2. *Sex ratios at birth for each population*

	number of sons per 100 daughters			
	rural populations		urban populations	
	Gypsy	Hungarian	Gypsy	Hungarian
A. all children				
sample size	254	216	239	224
males/100 females	89.3	111.8	89.7	113.3
B. first-born children only				
sample size	87	85	77	102
males/100 females	81.3	157.6	94.3	131.8

- Finally a significant difference!
  - $P<0.05$ (just about)
  - 95% CI excludes zero
  - difference between p(boys)=0.448 (Gypsies) and p(boys)=0.61 (Hungarians) is significant
- Conclusion: ***rural gypsies show lower proportion of first-born boys than rural Hungarians***

Exercise:

Run the same test two independent proportions test on urban children, first-born only

Does the sex ratio at birth differ between Gypsies and Hungarians?

```
prop.test(c(37,58),c(77,102))
```

Conclusions

- The only test that shows a difference between sex ratios in Roma and Hungarians is a two-independent sample comparing rural, first-born children

- But in that test, the ‘abnormal’ population is rural Hungarians, with a very high number of boys among first-borns!
  - They have 57.6% more boys than girls!!!
  - (but is this ratio significant?)

- Very biased conclusion!
  - Based on testing until result is the one you want
  - we’ll see how to fix this later

Table 2. *Sex ratios at birth for each population*

	number of sons per 100 daughters			
	rural populations		urban populations	
	Gypsy	Hungarian	Gypsy	Hungarian
A. all children				
sample size	254	216	239	224
males/100 females	89.3	111.8	89.7	113.3
B. first-born children only				
sample size	87	85	77	102
males/100 females	8[illegible].3	157.6	94.3	131.8

possible cheating, thus correction for multiple testing >>
limit the amount of testing (ANOVA)

Other cases: more than two independent proportions

- If you want to compare more than two independent proportions
  - e.g. to compare four populations, i.e. proportion of boys among rural gypsies, rural Hungarian, urban gypsies, urban Hungarians ***all at the same time***)

- Just extend *prop.test* to four populations

> prop.test(c(x1, x2, x3, x4), c(n1, n2, n3, n4))

pos total

All values

- Function *chisq.test* is like *prop.test*, but you enter it in matrix form

Syntax: pos neg

> chisq.test(matrix(c(x1, x2, x3, x4, n1, n2, n3, n4), m))

Important! Here n = negatives (in our case, girls), not total!!!

(m is the number of groups; in the example above, m=4)

Exercise:

1) Run prop.test on the four proportions: rural gypsies, rural Hungarian, urban gypsies, urban Hungarians

2) Run a chi-square test on the same four proportions

Other cases: one sample, n proportions

- You may want to test whether a die (instead of a coin) is loaded
  - now there are six proportions (1/6 for each side) that add up to 100%
  - tested proportion is now p=1/6=0.17

- This test can be done with function *chisq.test*

```
> chisq.test(matrix(c(x1, x2, x3, x4, x5, x6), nrow=1))
```

byrow=T

Chisel.test(matrix(c(x1,x2,x3x,x4,x4),byrow=T, nrow=1))

- x1 = number of times you got a 1 rolling the die, etc.
- nrow= 1 (meaning all 6 values are from the same dice)

Binomial test

- The binomial test is equivalent to a *prop.test*, except that it is based on the binomial distribution itself

- Some prefer *binom.test* because it estimates an exact P value
  - contrary to *prop.test* that calculates P value from a normal approximation to the binomial

```
> prop.test(120, 254, 0.5)
1-sample proportions test with continuity correction
data:  120 out of 254, null probability 0.5
X-squared = 0.6654, df = 1, p-value = 0.4147
alternative hypothesis: true p is not equal to 0.5
95 percent confidence interval:
 0.4099875 0.5357417
sample estimates:  p
0.4724409
```

```
> binom.test(120,254)
                Exact binomial test
data:  120 and 254
number of successes = 120, number of trials = 254, p-
value = 0.4147
alternative hypothesis: true probability of success is not
equal to 0.5
95 percent confidence interval:
 0.4097151 0.5358173
sample estimates: probability of success
          0.4724409
```

Fischer exact test

- Fisher exact test also calculates exact *P* value
- Now you enter the positive *x* cases (e.g. boys) and the ***negative*** cases (e.g girls), ***instead of total n***
- Test is based on odds-ratios not proportions
  - 95% CI looks different
  - if odds ratio (of boys to girls) is different from 1, proportions differ

Syntax:

>fisher.test(matrix(c(pos1, pos2,…, neg1, neg2,…), *m*)) + byrow=T

m= number of compared groups

Using data from urban gypsies vs. Hungarians

```
> fisher.test(matrix(c(39, 52, 48, 33), 2))
Fisher.test(matrix(c(x1,x2,x3,x4),byrow=T,2))
Fisher's Exact Test for Count Data

data:  matrix(c(39, 52, 48, 33), 2)
p-value = 0.03396
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.2684001 0.9885450
sample estimates:
odds ratio
 0.5176459
```

Generalisation

- Chi-square or Fisher exact tests can be generalised for any number of samples and proportions *at the same time*
- For example, does choice of degree (Archaeology, Biology, Engineering) at UCL affect the final grade (1^{st,} 2.1, 2.2, 3rd class degree)?
  - results per degree (number of 1^{st,} 2.1, 2.2, 3rd for each degree, which add up to 1)
  - number of degrees *nrow* (in this case, nrow=3)
- Syntax:

>chisq.test(matrix(c($n_{1st,arc}$, $n_{2.1,arc}$, $n_{2.2,arc}$, $n_{3rd,arc}$, $n_{1st,bio}$, $n_{2.1,bio}$, $n_{2.2,bio}$, $n_{3rd,bio}$, $n_{1st,eng}$, $n_{2.1,eng}$, $n_{2.2,eng}$, $n_{3rd,eng}$), nrow=3, byrow=T))

- *chisq.test* compares observed values (e.g. proportion of Biologists with 2.1) vs. predicted values (e.g. proportion of Biologists times proportion of 2.1 degrees);
- again, advantage of Fisher test is to calculate exact *P* value