

8: Chi-square goodness of fit

Readings: Zar 22.1, 22.2, 22.4, 22.5, 22.7

Trevor A. Branch, tbranch@uw.edu

School of Aquatic and Fishery Sciences

 [@TrevorABranch](https://twitter.com/TrevorABranch)

Review: covariance and correlation

The variance of $aX + bY$, when a and b are constants is given by

$$\sigma_{aX+bY}^2 = a^2 \sigma_X^2 + b^2 \sigma_Y^2 + 2ab \operatorname{cov}(X, Y)$$

Covariance is high when both X_i and Y_i are above their sample means or both are below their sample means:

$$\operatorname{cov}(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n-1}$$

In R: `cov()`

Correlation standardizes covariance to be between -1 and $+1$:

$$\rho_{XY} = \frac{\operatorname{cov}(X, Y)}{\sigma_X \sigma_Y}$$

In R: `cor()`

Review: Paired-sample t test



The data from the experiment are on the right.

$$\bar{d} = 176$$

Mean of the differences

$$s_{\bar{d}} = \frac{s}{\sqrt{n}} = \frac{190}{\sqrt{12}} = 55$$

Standard error of the differences

$$t = \frac{\bar{d}}{s_{\bar{d}}} = \frac{176}{55} = 3.20$$

t statistic

$$t_{0.05(2),11} = (-2.20, 2.20) \quad v \text{ is } n_{\text{pairs}} - 1, \text{ values from } qt()$$

$$P(t > 3.20) = 0.004 \quad \text{Value from } pt()$$

$$p\text{-value} = 2 \times 0.004 = 0.008$$

No strips	Prairie strips	Differences
2280	1890	390
2380	2410	-30
2290	2390	-100
2570	2300	270
2070	1700	370
2110	1890	220
1850	1680	170
2300	2120	180
2140	1860	280
2280	2430	-150
2300	1880	420
1810	1720	90

Conclusion: water runoff did change significantly by 176 units ($p = 0.008$).

Review: Confidence interval

The confidence interval for the paired-sample t test, for $\alpha = 0.05$, two-tailed, is given by:

$$\begin{aligned} & \bar{d} \pm t_{\alpha(2),v} s_{\bar{d}} \\ & = 176 \pm 2.20 \times 55 \\ & = (55, 297) \end{aligned}$$

This is the 95% confidence interval for the *difference between the means*.



No strips	Prairie strips	Difference
2280	1890	390
2380	2410	-30
2290	2390	-100
2570	2300	270
2070	1700	370
2110	1890	220
1850	1680	170
2300	2120	180
2140	1860	280
2280	2430	-150
2300	1880	420
1810	1720	90

Review: Tennis players

We gather the data for 12 tennis players, calculate the difference between the dominant and non-dominant arms, and assess whether these differences are normally distributed:

```
hist(diffs)
```

```
qqnorm(diffs)
```

```
shapiro.test(diffs)
```

The differences are not normally distributed ($p = 0.006$, Shapiro-Wilk test); we cannot use a paired-sample t test

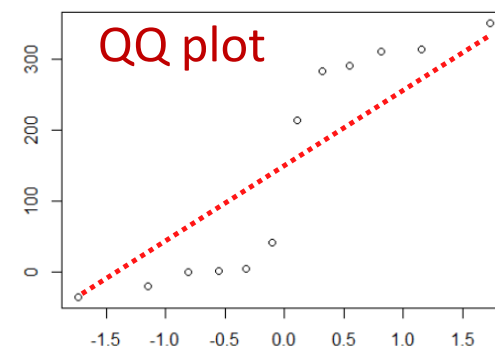
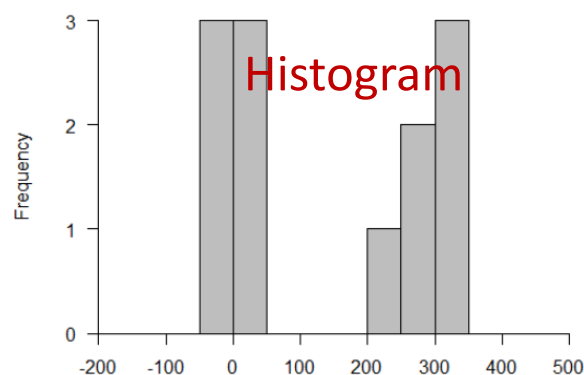
Dominant	Non-dominant	Difference
406	427	-21
989	699	290
454	455	-1
998	784	214
1152	839	313
643	602	41
304	303	1
317	352	-35
884	574	310
879	596	283
966	616	350
477	473	4



Left-handed Rafael Nadal, the top-ranked male tennis player



Right-handed Roger Federer, the greatest of all time.



Review: Wilcoxon paired-sample test

Calculate the differences d_j and absolute value of differences $|d_j|$, then find ranks of $|d_j|$ (1 = lowest, 12 = highest), and the signs of the ranks. Add the positive ranks to get T_+ and add the negative ranks to get T_- .

Dominant	Non-dominant	d_j	$ d_j $	Rank of $ d_j $	Sign of d_j
406	427	-21	21	4	-
989	699	290	290	9	+
454	455	-1	1	1.5	-
998	784	214	214	7	+
1152	839	313	313	11	+
643	602	41	41	6	+
304	303	1	1	1.5	+
317	352	-35	35	5	-
884	574	310	310	10	+
879	596	283	283	8	+
966	616	350	350	12	+
477	473	4	4	3	+

There is a tie between rank 1 and 2; both are given 1.5

$$T_+ = 9 + 7 + 11 + 6 + 1.5 + 10 + 8 + 12 + 3 = 67.5$$

$$T_- = 4 + 1.5 + 5 = 10.5$$

Our two test statistics are T_+ and T_- .

Review: Wilcoxon paired-sample test (2)

Two-tailed test

H_0 : measures from population 1 = measures from population 2

H_A : measures from population 1 \neq measures from population 2

In this case, reject H_0 if $T_- \leq T_{\alpha(2),n}$ or $T_+ \leq T_{\alpha(2),n}$. These values can be looked up using the function `psignrank()`.

χ^2 goodness of fit

Remember the χ^2 distribution we used for tests on single-sample variance

The χ^2 statistic arises from this calculation:

$$\chi^2 = \frac{vs^2}{\sigma^2}$$

Degrees of freedom, where $v = n-1$ as for the t distribution

Sample variance

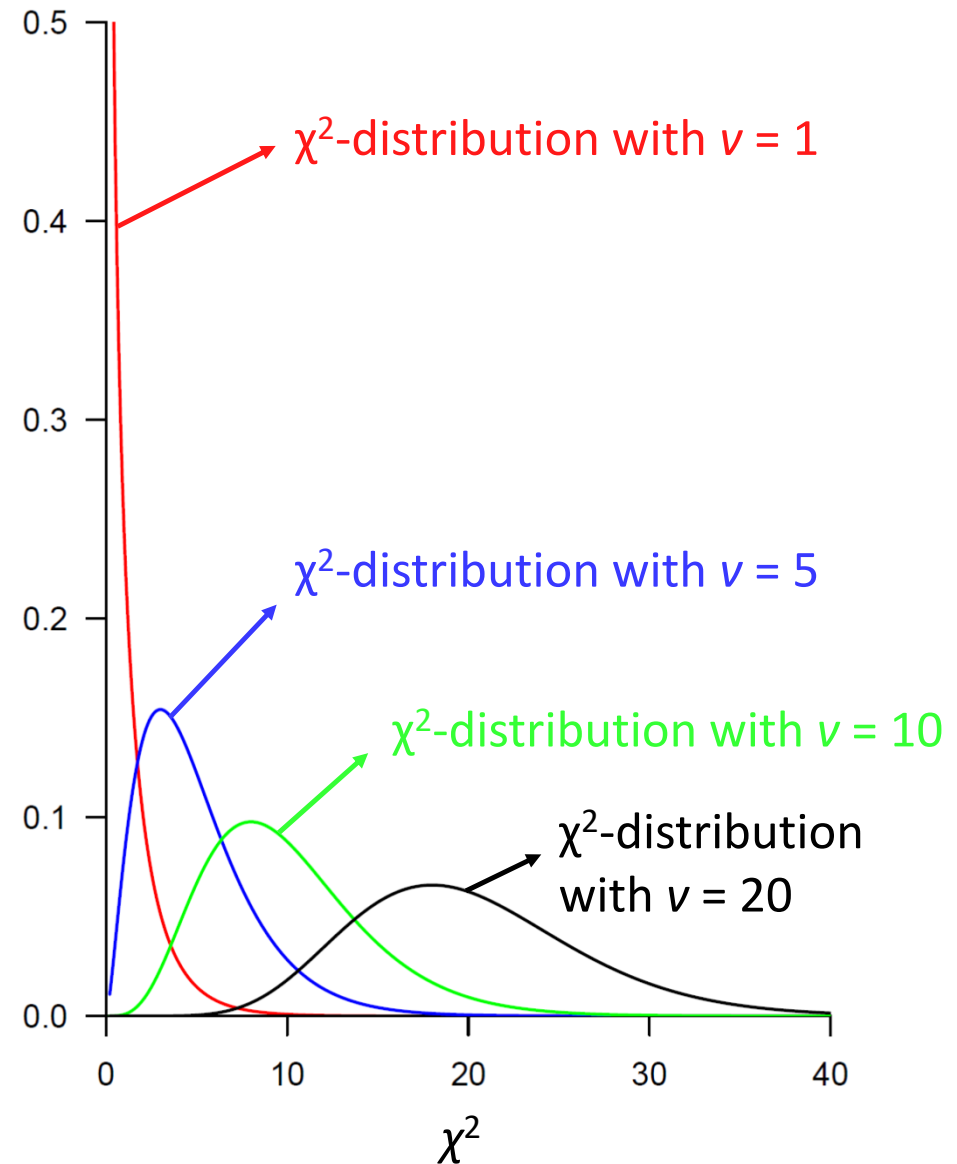
Population variance

For example, if you are testing the hypotheses $H_0: \sigma^2 = 5.6$ and $H_A: \sigma^2 \neq 5.6$, then you calculate:

$$\chi^2 = \frac{vs^2}{5.6}$$

The χ^2 -distribution

- Not symmetrical
- Governed by the degrees of freedom, where $\nu = n - 1$
- Cannot be negative
- With higher values of ν it gets bigger and more spread out
- R functions `pchisq()`, `dchisq()`, `rchisq()`, and `qchisq()`



χ^2 in R

Tests involving the χ^2 distribution are interested in the **right tail only**, answering the question: Are values further away from what would be predicted? To get the right tail in R:

```
qchisq(p=probability, df=v, lower.tail=FALSE)
```

χ^2 goodness of fit test

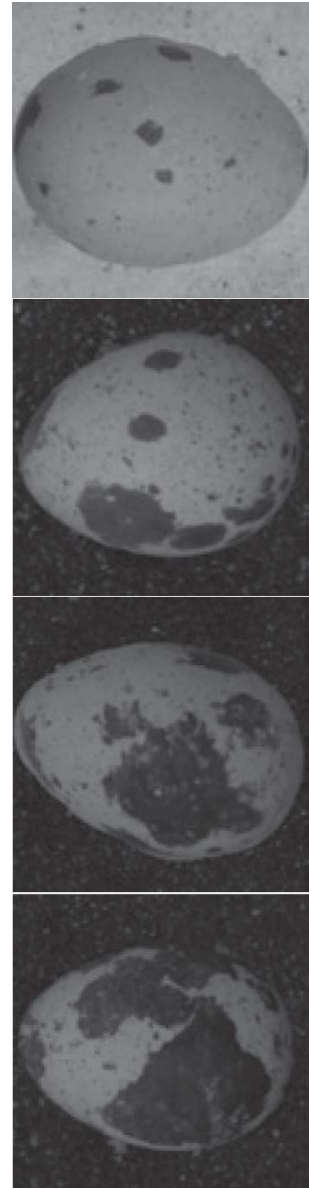
We also use the χ^2 distribution when testing if **counts** of items in particular **categories** follow a particular hypothesis we have. We will be comparing the observed frequencies f_i in each category with the predicted frequencies \hat{f}_i in each category. The circumflex on top of the letter indicates model prediction in statistics. The following expression is the χ^2 test statistic:

$$\chi^2 = \sum_{i=1}^k \frac{(f_i - \hat{f}_i)^2}{\hat{f}_i}$$

Which can be compared to a χ^2 distribution at a given level of α and degrees of freedom $\nu = k - 1$, where k is the number of categories: $\chi_{\alpha(1),\nu}^2$. These tests are always **one-sided** since we are testing whether the test statistic deviates far from the prediction or not; we are not interested in whether observations are closer to the predictions than expected.

Do quails know their egg color?

Japanese quail lay eggs with a wide variation in color from almost completely light, to nearly covered in dark coloration. If the quail know in advance their egg color, they could lay them on the sand color that provides the best camouflage from predators. An experiment was designed with four different colors of sand ranging from light to dark¹. Depending on the color pattern of the egg, these four sand colors can be ranked from 1 (best choice for camouflage) to 4 (worst choice). If the choice of sand is random, eggs should be evenly distributed among the four categories, but if they know in advance their egg color, they might be able to lay them preferentially in rank 1 and not in rank 4.



¹Lovell et al. (2013) Egg-laying substrate selection for optimal camouflage by quails. Current Biology 23:260-264

Predicted eggs in each category

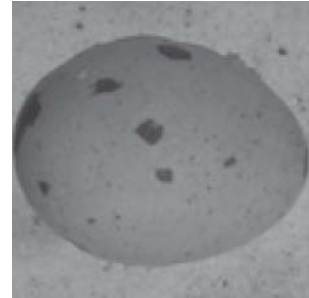
Since there are four sand backgrounds to choose from, if the null hypothesis is true (equal number of eggs on each sand type), then we would expect 25% of the total eggs (178) to be in each rank of the sand backgrounds: $0.25 \times 178 = 44.5$.

Match	Eggs laid	Predicted	χ^2 term
1	86	44.5	
2	43	44.5	
3	25	44.5	
4	24	44.5	
All	178	178.0	

More generally, we can use any kind of model for our predicted values, and obtain a χ^2 value comparing the data to the model prediction.



Do quails know the color of their eggs?



Rank 1 = best camouflage match, Rank 4 = worst camouflage match

Match	Eggs laid	Predicted	χ^2 term
1	86	44.5	38.70
2	43	44.5	0.05
3	25	44.5	8.54
4	24	44.5	9.44
All	178	178.0	56.74

$$\begin{aligned}\chi^2 &= \sum_{i=1}^k \frac{(f_i - \hat{f}_i)^2}{\hat{f}_i} \\ &= \frac{(86 - 44.5)^2}{44.5} + \frac{(43 - 44.5)^2}{44.5} + \frac{(25 - 44.5)^2}{44.5} + \frac{(24 - 44.5)^2}{44.5} \\ &= 56.74\end{aligned}$$

Our χ^2 test statistic is 56.74, with degrees of freedom $\nu = k - 1 = 3$. The critical value is $\chi^2_{0.05(1),3} = 7.81$ so we reject the null hypothesis. The corresponding p value is 2.9×10^{-12} .

```
crit.value <- qchisq(p=0.05, df=3, lower.tail=FALSE)
pvalue <- pchisq(q=56.74, df=3, lower.tail=FALSE)
```

Notes about χ^2 tests

1. The test must be conducted on the **count data** in each category. The counts cannot be multiplied or divided.
2. As you increase the number of categories, the χ^2 test statistic and χ^2 critical value will increase.
3. The χ^2 values for individual categories can be compared, with big values revealing which category departs the furthest from the predictions (for the quail example, this was the first category).
4. The critical values are the right-hand tails from a one-tailed test, i.e., for a significance level of $\alpha = 0.05$, we want the value of χ^2 that has 5% of the area to the **right**, hence the `lower.tail=FALSE` parameter when calling `qchisq()`.



Photo: Ingrid Taylar

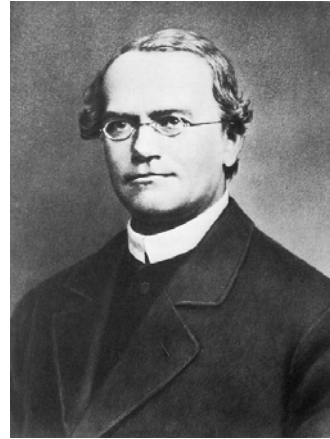
The χ^2 test in R (short way around)

The test can be run using `chisq.test()`. For a goodness-of-fit test, instead of specifying the predicted counts, we specify the predicted **probabilities** of falling in each category under the null hypothesis. In the case of the quail example, this probability is 0.25 (one quarter predicted in each category). The function `rep()` can be used to create a vector of repeated values.

```
chisq.test(x = c(86,43,25,24), p=c(0.25,0.25,0.25,0.25))  
chisq.test(x = c(86,43,25,24), p=rep(x=0.25, times=4))
```

Mendelian genetics

A classic genetics experiment by the monk Gregor Mendel examined what happens when peas with two traits are crossed. Yellow (Y) is a dominant gene for color and green (y) recessive; while round (R) is dominant and wrinkled (r) recessive. The upper-case letters denote dominant and lower-case recessive. Each pea contains two copies of these genes, thus pure-bred yellow-round peas are genetically $YYRR$, while pure-bred green-wrinkled peas are $yyrr$. If these two types of pure-bred peas are crossed, the first generation will all be $YyRr$ (yellow-round), while if these peas are crossed with each other, the second generation should have the ratios 9:3:3:1 for observed type yellow-round : yellow-wrinkled : green-round : green-wrinkled.



Round



Wrinkled



Yellow



Green

First and second generation crosses

In the first generation, the *YYRR* peas are crossed with *ggww*, resulting in all offspring having a *Y* and *R* from the yellow-round parent, and *g* and *w* from the green-wrinkled parents. Thus, every first generation is *YgRw* and looks yellow and round. The second generation, crossing *YgRw* with *YgRw* is much more interesting. There are $4^2 = 16$ possible genetic outcomes, since the offspring could get genetic material (half the parent's genes) from parent 1 that is either *YR*, *Yw*, *gR* or *gw*; and genetic material from parent 2 that is either *YR*, *Yw*, *gR* or *gw*

	YR	Yw	gR	gw
YR	YYRR	YYwR	gYRR	gYwR
Yw	YYRw	YYww	gYRw	gYww
gR	YgRR	YgwR	ggRR	ggwR
gw	YgRw	Ygww	ggRw	ggww

9 Yellow-Round

YYRR	YYwR	gYRR	gYwR
YYRw	YYww	gYRw	gYww
YgRR	YgwR	ggRR	ggwR
YgRw	Ygww	ggRw	ggww



Round



Yellow

3 Yellow-wrinkled

YYRR	YYwR	gYRR	gYwR
YYRw	YYww	gYRw	gYww
YgRR	YgwR	ggRR	ggwR
YgRw	Ygww	ggRw	ggww



Wrinkled



Yellow

3 green-Round

YYRR	YYwR	gYRR	gYwR
YYRw	YYww	gYRw	gYww
YgRR	YgwR	ggRR	ggwR
YgRw	Ygww	ggRw	ggww



Round



Green

1 green-wrinkled

YYRR	YYwR	gYRR	gYwR
YYRw	YYww	gYRw	gYww
YgRR	YgwR	ggRR	ggwR
YgRw	Ygww	ggRw	ggww



Wrinkled



Green

Predicted peas in each category

The predicted number of peas in each of the four categories should be in the ratio 9:3:3:1, therefore the predicted proportions should be $9/16 : 3/16 : 3/16 : 1/16$. If the experiment involved 4530 second generation offspring, then:



Round



Wrinkled



Yellow



Green

Category	Numbers observed	Numbers predicted
Yellow-Round	2504	$4530 \times 9/16 = 2548.1$
Yellow-wrinkled	853	$4530 \times 3/16 = 849.4$
green-Round	881	$4530 \times 3/16 = 849.4$
green-wrinkled	292	$4530 \times 1/16 = 283.1$
Total (<i>n</i>)	4530	4530

χ^2 test

The null hypothesis is that the predicted number of peas in each of the four categories should be in the ratio 9:3:3:1, while the alternative is that they are not in this ratio.



Round



Wrinkled



Yellow



Green

Category	Observed	Predicted	χ^2
Yellow-Round	2504	2548.1	0.763
Yellow-wrinkled	853	849.4	0.015
green-Round	881	849.4	1.176
green-wrinkled	292	283.1	0.280

$$\begin{aligned}
 \chi^2 &= \sum_{i=1}^k \frac{(f_i - \hat{f}_i)^2}{\hat{f}_i} \\
 &= \frac{(2504 - 2548.1)^2}{2548.1} + \frac{(853 - 849.4)^2}{849.4} + \dots \\
 &= 0.763 + 0.015 + \dots \\
 &= 2.235
 \end{aligned}$$

Running the χ^2 test

The test statistic was 2.235, while the critical value of $\chi^2_{0.05,3}$ is 7.81. The p -value from the test is $p = 0.53$. As a result, we do not reject the null hypothesis, and we conclude that the ratio of the second generation offspring is indeed 9:3:3:1 when crossing offspring of a cross between two dominant traits and two recessive traits.

```
observed <- c(2504, 853, 881, 292)
predprob <- c(9/16, 3/16, 3/16, 1/16)
predicted <- sum(observed) * predprob
chisq.value <- sum((observed-predicted)^2/predicted)
crit.value <- qchisq(p=0.05, df=3, lower.tail=FALSE)
pvalue <- pchisq(q=chisq.value, df=3, lower.tail=FALSE)

chisq.test(x=observed, p=predprob)
```

Yates correction for continuity

The count data used to obtain the χ^2 test values are discrete (only particular values are possible for a given total count). However, the underlying theoretical χ^2 distribution is continuous (any value of χ^2 is possible). For cases where the number of categories is $k \geq 3$ ($\nu \geq 2$) the approximation is very good. But for cases with just 2 categories ($k = 2$, $\nu = 1$), **and only for this case**, the Yates correction for continuity should be used to obtain a modified χ^2 test statistic, which subtracts 0.5 from the absolute value of each term before squaring:

$$\chi_c^2 = \sum_{i=1}^{k=2} \frac{(|f_i - \hat{f}_i| - 0.5)^2}{\hat{f}_i}$$

Mendelian flower colors

In a simpler experiment involving just one trait, Mendel crossed pure-bred peas with purple flowers (PP), with pure-bred peas with white flowers (ww). In the first generation, all of the offspring had purple flowers (Pw), indicating that purple is a dominant gene and white is a recessive gene. When these flowers are self-fertilized, there are four possible outcomes: PP , Pw , wP , and ww , and we predict a 3:1 ratio of purple to white flowers in the second generation (out of every 4 flowers, expect 3 purple and one white). In the actual experiment, Mendel found 705 plants with purple flowers and 224 with white flowers. Do the data follow the expected ratio? Predicted counts:

$(705+224) \times 3/4 = 696.75$ purple and $(705+224) \times 1/4 = 232.25$ white



Purple



White



Purple

χ^2 with Yates correction for continuity (for $k = 2, \nu = 1$)



White

$$\chi_c^2 = \sum_{i=1}^{k=2} \frac{(|f_i - \hat{f}_i| - 0.5)^2}{\hat{f}_i}$$

$$= \frac{(|705 - 696.75| - 0.5)^2}{696.75} + \frac{(|224 - 232.25| - 0.5)^2}{232.25}$$

$$= 0.345$$

Category	Observed	Predicted	χ^2
Purple	705	696.75	0.086
white	224	232.25	0.259
Total	929	929.00	0.345

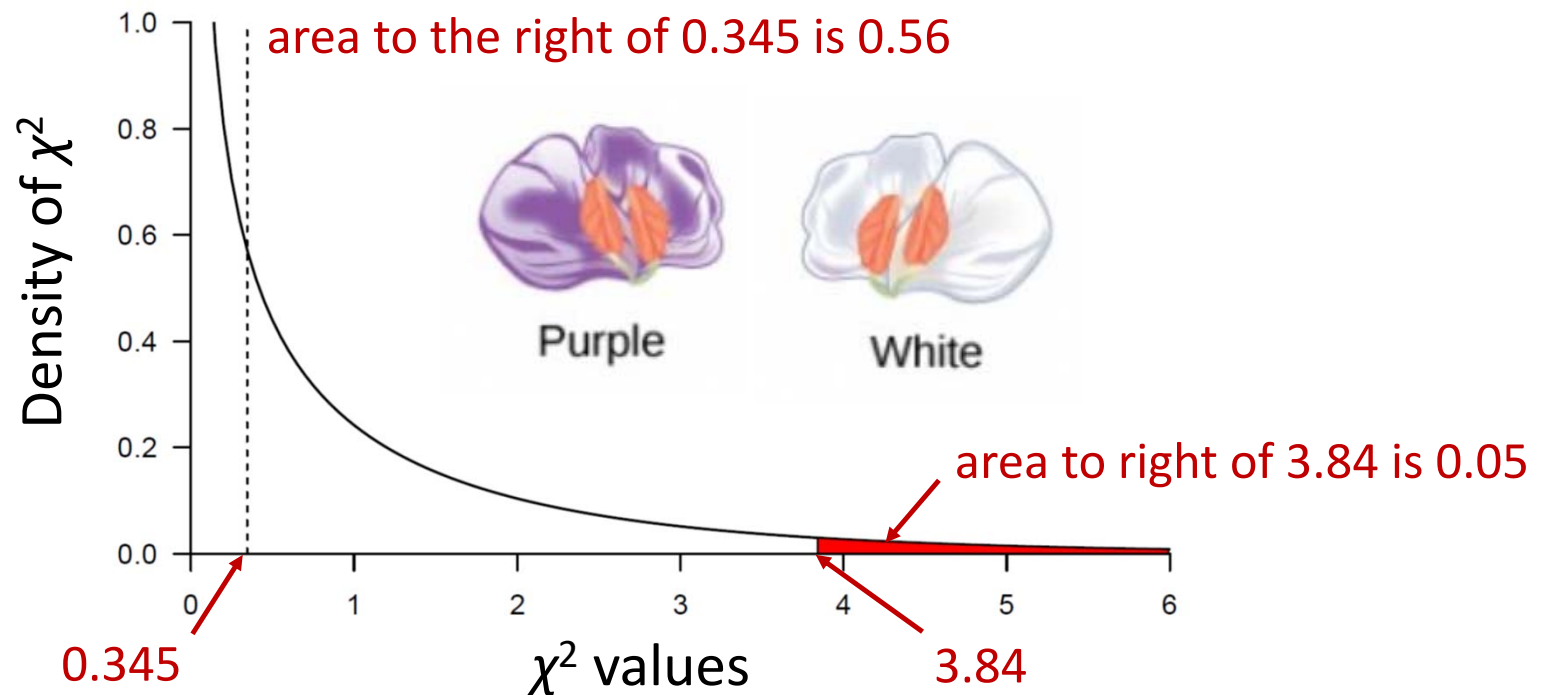
```

predpurple <- (705+224)*3/4
predwhite  <- (705+224)*1/4
chisq.value <- (abs(705-predpurple)-0.5)^2/predpurple +
               (abs(224-predwhite)-0.5)^2/predwhite
crit.value  <- qchisq(p=0.05, df=1, lower.tail=F)
pvalue     <- pchisq(q=chisq.value, df=1, lower.tail=F)
  
```

0.345
3.84
 $p = 0.56$

Mendelian flower colors

Since our test statistic is 0.345, which is smaller than the critical value of $\chi^2_{0.05(1),1} = 3.84$, we do not reject the null hypothesis, and conclude that the flowers do indeed follow Mendelian genetics ($p = 0.56$).



Notes on the χ^2 goodness of fit test

Assumptions

1. Sample observations are drawn randomly from the population
2. Sample observations are independent from each other

Constraints

1. Data must be categorical (each sample fits in a category).
2. The predicted frequencies can't be "too small" (more on the next slide)

One-tail test with critical values in the right tail

In the χ^2 we subtract expected from observed and then square the result, thus large values of χ^2 (in the right tail) occur whether the observed is much bigger than expected, or much smaller than expected. This is why it is a one-tailed test.

How big do the predicted frequencies need to be?

If the expected frequencies \hat{f}_i in each category are “too small”, the calculated χ^2 is biased high, and the null hypothesis will be rejected more often than α (Type I error). Rules 1–3 below should be followed to ensure that predicted frequencies are large enough.

1. Roscoe & Byars: if all predicted frequencies are equal, ($\hat{f}_i = n / k$, where n is total count, and k is number of predicted categories) then for $\alpha = 0.05$, values as small as 1.0 are fine, or 2.0 if the data depart moderately from a uniform distribution.
2. Koehler & Larntz: three conditions must be met: $k \geq 3$, $n \geq 10$, and $n^2/k \geq 10$.
3. Zar: if $k = 2$ then every $\hat{f}_i > 5.0$.

A new pea example, with different data

Recall Mendel's peas, with an expected ratio of 9:3:3:1 in the second generation after crossing pure-bred dominant Round-Yellow with pure-bred recessive wrinkled-green peas. A new set of data are collected and analyzed with a χ^2 goodness of fit test, with $\nu = 3$, $\chi^2 = 8.972$, and p -value = 0.027, and the χ^2 components below. The null hypothesis is rejected. What do we do now?



Round



Wrinkled



Yellow



Green

Category	f_i	\hat{f}_i	χ^2
Yellow-Round	152	140.625	0.920
Yellow-wrinkled	39	46.875	1.323
green-Round	53	46.875	0.800
green-wrinkled	6	15.625	5.929

Standardized residuals of χ^2

The green-wrinkled category has the highest χ^2 value, which indicates that it is the furthest from the predicted value. We can also calculate the **standardized residual**, which for the χ^2 is the square root of the χ^2 value for each category. As the sample size increases, the standardized residual becomes a standard normal variable $N(0, 1^2)$, and thus values greater than 1.96, as seen for green-wrinkled, are quite unlikely.

Category	f_i	\hat{f}_i	χ^2	$\sqrt{\chi^2}$
Yellow-Round	152	140.625	0.920	0.96
Yellow-wrinkled	39	46.875	1.323	1.15
green-Round	53	46.875	0.800	0.89
green-wrinkled	6	15.625	5.929	2.43

$$\sqrt{\chi^2} = \frac{f_i - \hat{f}}{\sqrt{\hat{f}}}$$

Properties of χ^2

Why should the standardized residual, the square root of χ^2 , be a standard normal? Because the χ^2 distribution with one degree of freedom is obtained by squaring a standard normal distribution.

Subdividing χ^2 goodness of fit



Round



Wrinkled



Yellow



Green

The results suggest that it is an issue with the green-wrinkled category that is responsible for the departure from the null hypothesis. We can subdivide the data and ask whether this is the case. First, run the test just on the 9:3:3 parts of the data, with predicted proportions of 9/15 : 3/15 : 3/15. The resulting test statistic is 2.544 which is less than the critical value $\chi^2_{0.05,2} = 5.991$ and therefore we accept the null hypothesis of 9:3:3.

Category	f_i	\hat{f}_i	χ^2
Yellow-Round	152	146.4	0.214
Yellow-wrinkled	39	48.8	1.968
green-Round	53	48.8	0.362

Subdividing χ^2 goodness of fit

Now compare the numbers in the first three categories combined, with the numbers in the green-wrinkled category.

We now expect a 15:1 ratio. Since $v = 1$ we apply the Yates correction for continuity. The resulting test statistic is 5.684, which is greater than the critical value of $\chi_{0.05,1}^2 = 3.841$ and we therefore reject the null hypothesis and conclude that the reason for the rejection is that green-wrinkled peas occurred less often than expected.



Category	f_i	\hat{f}_i	χ_c^2
All others	244	234.375	0.355
green-wrinkled	6	15.625	5.329

$$\chi_c^2 = \sum_{i=1}^{k=2} \frac{(|f_i - \hat{f}_i| - 0.5)^2}{\hat{f}_i}$$

Subdividing χ^2 goodness of fit: what we learned

By subdividing the data, we were able to learn that the reason this experiment failed to conform to the Mendelian predictions was because of the observed frequency of the green-wrinkled seeds. The other seeds followed the predictions.

Note that it is not appropriate to test hypotheses **after** examining the data collected for the first hypothesis test. The subdivision tests are not real hypotheses since they are driven by the data. Instead, subdividing gives us information that can be used to collect **new** data to test the **new** proposed hypothesis that green-wrinkled seeds do not obey Mendel's laws.

Log-likelihood ratio for
goodness of fit

Introduction to log-likelihood ratio test

This statistic is often simply called the likelihood ratio test, and at its heart it is the ratio of two likelihoods (the heights or densities of the underlying probability distributions). In this case of the χ^2 test, the first likelihood is the likelihood of the population having the proportions in each category that are observed in the actual data. The second likelihood is the likelihood of the population having the proportions in each category that are stated in the null hypothesis. (Deep background: for the χ^2 test, the derivation of the test statistic is based on a distribution called the multinomial distribution.) The central importance of the likelihood ratio is that it can be applied to any problem with any type of data and any kind of likelihood, not just a χ^2 test of goodness of fit.

Log-likelihood ratio for goodness of fit

The test statistic for the log-likelihood ratio is called G , and is given by:

$$G = 2 \sum_{i=1}^k f_i \ln \frac{f_i}{\hat{f}_i}$$

where these are observed and predicted frequencies in each of k categories, as for the χ^2 test. The resulting value of G will be distributed as χ^2 with ν degrees of freedom, where $\nu = k - 1$ as for the χ^2 test.

Log-likelihood ratio test

This method has greater power to detect differences, but the Type I error is more likely to differ from the stated α . The log-likelihood ratio test can also be applied to a wide variety of situations far beyond goodness-of-fit cases. G is compared to a χ^2 critical value with $\nu = k - 1$ degrees of freedom.



Round



Wrinkled



Yellow



Green

Category	f_i	\hat{f}_i	G
Yellow-Round	152	140.625	11.82
Yellow-wrinkled	39	46.875	-7.17
green-Round	53	46.875	6.51
green-wrinkled	6	15.625	-5.74

$$\begin{aligned}
 G &= 2 \sum_{i=1}^k f_i \ln \frac{f_i}{\hat{f}_i} \\
 &= 2 \left(152 \ln \frac{152}{140.625} + 39 \ln \frac{39}{46.875} + \right. \\
 &\quad \left. 53 \ln \frac{53}{46.875} + 6 \ln \frac{6}{15.625} \right) \\
 &= 10.833
 \end{aligned}$$

Log-likelihood ratio in R

$$G = 2 \sum_{i=1}^k f_i \ln \frac{f_i}{\hat{f}_i}$$

Category	f_i	\hat{f}_i	G
Yellow-Round	152	140.625	11.82
Yellow-wrinkled	39	46.875	-7.17
green-Round	53	46.875	6.51
green-wrinkled	6	15.625	-5.75

```
observed <- c(152,39,53,6)
```

```
n <- sum(observed)
```

```
pexpected <- c(9,3,3,1)/16
```

```
expected <- n*pexpected
```

```
Gvalue <- 2*sum(observed*log(observed/expected))
```

```
crit.value <- qchisq(p=0.05, df=3, lower.tail=FALSE)
```

```
pvalue <- pchisq(q=Gvalue, df=3, lower.tail=FALSE)
```

use chi-square

since the operations are on vectors

$\log()$ is natural logarithm \ln in R, not \log_{10}

Built-in likelihood-ratio test

The built-in test is not part of base R (R that is pre-installed), and needs the library DescTools. The initial installation of this package requires access to the internet.

```
install.packages("DescTools")  
library(DescTools)  
observed <- c(152,39,53,6)  
pexpected <- c(9,3,3,1)/16  
GTest(x=observed, p=pexpected, correct="none")
```