Goodness of Fit Test

- in some situations there are more than two possible outcomes on each trial, and we are given counts in each category
- in this case the multinomial distribution applies, an extension to the binomial
- for example, a roll of a die can give only the outcomes 1,2,3,4,5 or 6
- on 100 rolls we may get the following

Outcome	1	2	3	4	5	6
Count	9	18	13	20	22	18

- the natural question here is whether the die is fair, that is whether each outcome is equally likely, or
- H_0: $p_1 = \ldots = p_6 = 1/6$
- the alternative is that at least one probability is different from 1/6
- H_0: $p \neq 1/6$ for some i
- some variation in counts is to be expected, but how much?
- each cell of a multinomial distribution has a binomial distribution, so in this case the mean count is $100/6 = 16\ 2/3$, and the standard deviation is $\sqrt{100 * 1/6 * 5/6} = 3.73$
- the goodness of fit test statistic compares the observed counts X_i to their means np_i, or what would be expected under the null hypothesis

$$X^2 = \sum_{i=1}^{k} \frac{(X_i - np_i)^2}{np_i}$$

- in this case, the number of categories is $k = 6$ and

$$\begin{aligned} X^2 &= \frac{(9-100/6)^2}{100/6} + \ldots + \frac{(18-100/6)^2}{100/6} \\ &= 3.527 + \ldots + .107 \\ &= 6.92 \end{aligned}$$

- the distribution of the test statistic is approximately χ^2 with degrees of freedom equal to 5, one less than the number of categories
- comparing with tables, we find that the test statistic is smaller than 9.236 so the P value is greater than .1

- there is therefore no evidence against the null hypothesis that the outcomes are equally likely
- what assumptions are needed to do this test?
- we assume that the trials are independent and identical (the probabilities don't change)
- the expected counts in all cells should be at least 5 for the χ^2 approximation to the distribution of the test statistic to hold
- **Another example**: In Mendel's clasic pea experiments to test his genetic theory, he predicted the following proportional breakdown for four types of peas:

shape/colour	round/yellow	wrinkled/yellow	round/green	wrinkled/green
probability	9/16	3/16	3/16	1/16
counts	59	19	14	8
expected	56.25	18.75	18.75	6.25
$(o-e)^2/e$	.134	.003	1.203	.49

- the observed and expected counts are also shown as are the contributions to the goodness of fit statistic
- note that the assumed probabilities are not equal here
- the expected counts are np_i, all are greater than 5
- for example, in the first cell $np_1 = 100 * 9/16 = 56.25$ and the contribution to X^2 is $(59 - 56.25)^2/56.25 = .134$
- the total test statistic is $X^2 = 1.8307$
- comparing to χ^2 tables with 4-1=3 degrees of freedom, we see that 1.8307 is less than 6.251, so $P > .10$
- there is no evidence that Mendel's theories do not hold