1. The following table on the right gives the mass spectrometer reading for a molecule a scientist thinks is PCl_3. Phosphorus has only one stable isotope with a mass of 31 amu. Chlorine has two stable isotopes, ^{35}Cl and ^{37}Cl, occurring with relative abundance 75.8% and 24.2% respectively. To test if the data fit the assumption answer the following questions.

mass (amu)	frequency
136	181
138	150
140	61
142	8

(a) How many molecules were recorded by the mass spectrometer?
The sum of the frequencies is 400.

(b) If the molecule was PCl_3, what is the probability that it would have exactly 3 ^{35}Cl atoms?
This is a binomial distribution. The probability that a particular chlorine atom is a ^{35}Cl atom is $p = 0.758$. So the probability of getting three is $P(3) = (0.758)^3 = 0.436$.

(c) If the molecule was PCl_3, what is the probability that it would have exactly 2 ^{35}Cl atoms and 1 ^{37}Cl atoms?
Now there are two "successes" and one "failure" so the probability is
$P(2) = \binom{3}{2}(0.758)^2(0.242)^1 = 0.417$.

(d) Find the probabilities for the other two cases. (ie that PCl_3 would have 1 and 0 ^{35}Cl atoms respectively.)
$P(1) = \binom{3}{1}(0.758)^1(0.242)^2 = 0.133$ and $P(0) = \binom{3}{0}(0.242)^3 = 0.0142$.

(e) Use the probabilities above to find the expected frequencies of the mass spectrometer readings.
We multiply each of the probabilities by 400 to get the expected frequencies 174, 167, 53.3 and 5.7, corresponding to PCl_3 molecules with 3,2,1 or 0 ^{35}Cl atoms and mass 136, 138, 140, 142 respectively.

(f) Hence calculate χ^2 and determine if the data is consistent with the molecule being PCl_3.
$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(181-174)^2}{174} + \frac{(150-167)^2}{167} + \frac{(61-53.3)^2}{53.3} + \frac{(8-5.7)^2}{5.7} = 4.05$. There are 4-1=3 degrees of freedom and the critical value of χ^2 at the 0.1 (or 10%) level of significance is 6.25. Since $4.05 < 6.25$ we can say the data is consistent with the molecule being PCl_3.

2. Snapdragons homozygous for an allele R are red, those homozygous for allele r are white. Heterozygotes are pink. The following table shows the results of planting seeds from a heterozygote cross.

phenotype	frequency
white	14
pink	43
red	23

(a) Find the expected frequency using Mendel's laws. Then calculate χ^2 and test if this data is consistent with the theory of independent of assortment at the 0.05 level of significance?
In a Rr×Rr cross we get rr with probability 0.25, rR with probability 0.5 and RR with probability 0.25. Since there are a total of 80 plants we expect 20 white 40 pink and 20 red plants. Hence $\chi^2 = \frac{(14-20)^2}{20} + \frac{(43-40)^2}{40} + \frac{(23-20)^2}{20} = 2.475$. There are 3-1=2 degrees of freedom, and the critical value of χ^2 at the 0.05 level of significance is 5.99. Since $2.475 < 5.99$ we can say the data is consistent with Mendel's laws.

(b) Calculate the relative allele frequencies in the seeds that were planted.
First we find the relative frequency of the phenotypes. For red it's $\frac{23}{80}$ for pink it's $\frac{43}{80}$ and for white it's $\frac{14}{80}$. Now, for the red allele $p = \frac{23}{80} + \frac{1}{2}\frac{43}{80} = 0.5563$ and for the white allele $q = \frac{14}{80} + \frac{1}{2}\frac{43}{80} = 0.4437$. By random chance the allele frequencies are not 50:50.

(c) Assume the seeds from this cross were planted on an island and that their descendants eventually spread over the entire island. Calculate the Hardy Weinberg equilibrium for the relative frequencies of the three different phenotypes assuming the alleles frequencies do not change.
For red we expect relative frequency $p^2 = (0.5563)^2 = 0.3095$ for pink we expect $2pq = 2(0.5563)(0.4437) = .4937$ and for white we expect $q^2 = (0.4437)^2 = 0.1969$.

(d) A botanist collects a random sample from the island. Based on a χ^2 test and the Hardy Weinberg frequencies calculated above, decide if there is evidence for evolution (ie test if the data is significantly different from the expected Hardy-Weinberg frequencies).
Since there are again 80 samples the expected frequencies can be found by taking the relative frequencies times 80. These are shown in the table below, and χ^2 is calculated.

phenotype	observed freq	expected freq	(O-E)^2/E
white	7	15.7	$(7-15.7)^2/15.7$
pink	37	39.5	$(37-39.5)^2/39.5$
red	36	24.8	$(36-24.8)^2/24.8$
total	80	80	10.04

There are 2 degrees of freedom and the critical value at 0.05 level of significance is 5.99. Since our value of $\chi^2 = 10.04$ is greater than the critical value we can say that we are significantly far from the Hardy-Weinberg equilibrium to conclude that evolution has taken place.

3. Suppose you have 5 plants which are heterozygotes in two alleles A and a. Then in the gene pool 50% of alleles are A and 50% are a. Suppose that these 5 plants die and are replaced by 5 offspring, which are produced by random mating.

   (a) What is the probability that in the offspring exactly 50% of alleles are still A? (Hint: We do not need to think about the precise mix of genotypes – we just need to think in terms of a gene pool from which we select a certain number of alleles for the next generation. 5 offspring have 10 alleles. We are asking for the probability of getting exactly 5 A alleles out of 10 trials.
   This is a binomial distribution. The number of trials is 10, the probability of success (ie getting an A allele) is 0.5, and we are asking for the probability of 5 successes. $P(5) = \binom{10}{5}(0.5)^5(0.5)^5 = 0.246$ or about 25% chance. This means there is a 75% chance that the one or the other of the alleles will have a relative frequency less than 50%.

   (b) What is the probability of getting fewer than 50% of allele A?
   This is the probability of getting 0,1,2,3 or 4 A alleles. Which is $P(0) + P(1) + P(2) + P(3) + P(4)$. We can calculate each of these with the binomial distribution – but a faster way is to make use of the symmetry and reason this way. We are just as likely to get more than 50% as less than 50% of A alleles. So the probability of getting less than 50% A alleles is half the probability of not getting exactly 50%, ie $0.5(1 - 0.246) = 0.377$. Check that this is the same answer as you get doing it the long way.

   (c) What is the probability that allele A is completely eliminated from the population?
   This is $P(0) = \binom{10}{0}(0.5)^{10} = 0.00098$ or about 0.1%.

   (d) If the parent generation dies every winter and is replaced by exactly five offspring in spring, how many years would you estimate it would take before one or the other of the alleles is completely eliminated by random chance?
   In any given year there is a 1 in 1000 chance of the A getting eliminated so a rough estimate would be that it would take about 1000 years for the A alleles to be eliminated. There is a 1 in 500 chance that either alleles would be eliminated so we might estimate that one or the other would be eliminated in 500 years. In fact, that is an over estimate, because once an allele become rare the probability of it being eliminated in one step goes up. As we see there is a 75% chance that one or the other alleles will become less common than the other in one year.