

PRACTICE PROBLEMS IN POPULATION GENETICS

1. In a study of the Hopi, a Native American tribe of central Arizona, Woolf and Dukepoo (1959) found 26 albino individuals in a total population of 6000. This form of albinism is controlled by a single gene with two alleles: albinism is recessive to normal skin coloration.

a) Why can't you calculate the allele frequencies from this information alone?

Because you can't tell who might be a carrier just by looking.

b) Calculate the expected allele frequencies and genotype frequencies if the population were in Hardy-Weinberg equilibrium. How many of the Hopi are estimated to be carriers of the recessive albino allele?

If we assume that the population's in H-W equilibrium, then the frequency of individuals with the albino genotype is the square of the frequency of the albino allele. In other words, $\text{freq}(aa) = q^2$. $\text{Freq}(aa) = 26/6000 = 0.0043333$, and the square root of that is 0.0658, which is q , the frequency of the albino allele. The frequency of the normal allele is p , equal to $1 - q$, so $p = 0.934$.

We'd then predict that the frequency of Hopi who are homozygous normal (genotype AA) is p^2 , which is 0.873. In other words, 87.3% of the population, or an estimated 5238 people, should be homozygous normal. The frequency of carriers we'd predict to be $2pq$, which is 0.123. So 12.3%, or 737 people, should be carriers of albinism, if the population is in H-W.

2. A wildflower native to California, the dwarf lupin (*Lupinus nanus*) normally bears blue flowers. Occasionally, plants with pink flowers are observed in wild populations. Flower color is controlled at a single locus, with the pink allele completely recessive to the blue allele. Harding (1970) censused several lupin populations in the California Coast Ranges. In one population of lupins at Spanish Flat, California, he found 25 pink flowers and 3291 blue flowers, for a total of 3316 flowers.

a) Calculate the expected allele frequencies and genotype frequencies if the population were in Hardy-Weinberg equilibrium.

Let B be the blue allele and b be the pink allele, so that p = frequency (B) and q = frequency (b).
 The frequency of the bb genotype = $25/3316 = q^2$, so $q = \sqrt{0.00754} = 0.0868$.
 $p = 1 - q$, so $p = 0.913$
 freq (BB) = $p^2 = 0.834$
 freq (Bb) = $2pq = 0.158$

b) Harding studied the fertility of lupins by counting number of seed pods produced per plant in a subsample of the Spanish Flat population. He found the following:

	mean # pods	number of plants examined
blue	19.33	39
pink	13.08	24

Assume that heterozygotes are as fit as homozygous blue lupins, and that seeds from both pink and blue lupins all suffer about the same mortality rate after germinating. Calculate the relative fitness of each genotype.

Fitness for BB (w_{BB}) = 1
 Fitness for Bb (w_{Bb}) = 1
 Fitness for bb (w_{bb}) = $13.08 / 19.33 = 0.677$

c) Predict quantitatively the effect of natural selection on the frequencies of phenotypes in the next generation of lupins.

First, calculate mean fitness:

$$p^2 (w_{BB}) + 2pq (w_{Bb}) + q^2 (w_{bb}) = \bar{w}$$

$$(0.913 * 0.913 * 1) + (2 * 0.913 * 0.0868 * 1) + (0.0868 * 0.0868 * 0.677) = 0.997$$

Now divide all terms through by \bar{w} to get the predictions for the genotype frequencies after one round of selection:

New frequency (BB) = $(0.913 * 0.913 * 1) / 0.997 = 0.836$

New frequency (Bb) = $(2 * 0.913 * 0.0868 * 1) / 0.997 = 0.159$

New frequency (bb) = $(0.0868 * 0.0868 * 0.677) / 0.997 = 0.00512$

Moral of the story: Natural selection isn't all that efficient at eliminating rare alleles.

3. Cooke and Ryder (1971) studied the nestlings of Ross's goose, a small Arctic nesting goose. Goslings (baby geese) exist in two color morphs, grey or yellow. Cooke and Ryder reported that a population of geese at Karrack Lake, Canada included 263 yellow goslings and 413 grey goslings (676 total). They assumed that color is controlled by two alleles at a single locus.

a) Calculate the frequencies of all three possible genotypes, assuming that grey is dominant and that the population is in Hardy-Weinberg equilibrium. Then repeat, assuming that yellow is dominant.

For both of these calculations, p = frequency of dominant allele, and q = frequency of recessive allele. If grey is dominant:

$q^2 = 263 / 676 = 0.389$
 $q = \sqrt{0.389} = 0.624 =$ frequency of yellow allele
 $p = 1 - q = 0.376 =$ frequency of grey allele
 Predicted frequency of homozygous greys = $0.376 * 0.376 = 0.141$
 Predicted frequency of heterozygous greys = $2 * 0.376 * 0.624 = 0.469$
 Frequency of homozygous yellows = 0.389 .

CHECK: These add up to 1 (well, to 0.999, but that's round-off error)

If yellow is dominant:

$q^2 = 413 / 676 = 0.611$
 $q = \sqrt{0.611} = 0.782$
 $p = 1 - q = 0.218$
 Predicted frequency of homozygous yellows = $0.218 * 0.218 = 0.0475$
 Predicted frequency of heterozygous yellows = $2 * 0.218 * 0.782 = 0.341$
 Frequency of homozygous greys = 0.611

Check: These add up to 1, within round-off error.

b) Assume that grey is dominant. (In real life, Cooke and Ryder were unable to determine which allele was dominant.) There is no difference between yellow and grey goslings once they have matured. However, yellow goslings are at an increased risk of predation by a predatory bird, the Arctic skua. If 303 grey goslings survive to adulthood, but only 150 yellow ones do, calculate the fitness of the yellow phenotype relative to the grey one.

Let G be the gray allele and g be the yellow allele.

We've already figured out that $p = \text{freq}(G) = 0.376$

and $q = \text{freq}(g) = 0.624$.

Survival rate of grey goslings = $303/413 = 0.734$

Survival rate of yellow goslings = $150/263 =$

0.570

We could just use these as estimates of fitness, but remember that life is easiest if fitnesses are normalized so that the highest fitness value gets a value of 1.0, so let

$$\begin{aligned}w_{GG} &= 0.734 / 0.734 = 1.0 \\w_{Gg} &= 0.734 / 0.734 = 1.0 \\w_{gg} &= 0.570 / 0.734 = 0.777\end{aligned}$$

c) Now calculate the mean fitness ("w-bar"). Use that to predict the effect of selection on the next generation.

$$\begin{aligned}p^2w_{GG} + 2pqw_{Gg} + q^2w_{gg} &= \bar{w} \\(0.376 * 0.376 * 1) + (2 * 0.376 * 0.624 * 1) + (0.624 * 0.624 * 0.777) &= \bar{w} \\ \bar{w} &= 0.913\end{aligned}$$

You get the effects of selection by dividing the above equation through by w-bar. So:

$$\text{New frequency of GG genotype} = (0.376 * 0.376 * 1) / 0.913 = 0.155$$

$$\text{New frequency of Gg genotype} = (2 * 0.376 * 0.624 * 1) / 0.913 = 0.514$$

$$\text{New frequency of gg genotype} = (0.624 * 0.624 * 0.777) / 0.913 = 0.331$$

4. A 1970 study of 93 house mice (*Mus musculus*) in a single barn in Texas focused on a single locus (the gene for a certain enzyme) with two alleles, A and A'. The genotype frequencies found were:

AA 0.226
AA' 0.400
A'A' 0.374

a) Calculate the allele frequencies.

Quick and easy way:

$$\begin{aligned}\text{Freq (A)} = p &= 0.226 + (0.400 / 2) = 0.426 \\ \text{Freq (a)} = q &= 0.374 + (0.400 / 2) = 0.574\end{aligned}$$

b) How does this population differ from the predictions of Hardy-Weinberg equilibrium? Show your work.

$$\begin{aligned}\text{Predicted freq (AA)} &= p^2 = 0.181 \\ \text{Predicted freq (AA')} &= 2pq = 0.489 \\ \text{Predicted freq (A'A')} &= q^2 = 0.329\end{aligned}$$

c) In this specific case, what factor or factors are most likely to be causing deviations from Hardy-Weinberg equilibrium? How can you tell?

Could be several things, but notice in particular that (a) this is a small, restricted population, and (b) the heterozygotes are less common, and BOTH homozygotes are more common, than we'd expect. Sounds like inbreeding is a likely explanation. In fact, we could calculate F by solving the equation:

$$\begin{aligned}\text{actual freq (AA)} &= p^2 + pqF \\ F &= [\text{freq (AA)} - p^2] / pq = (0.226 - 0.181) / (0.426 \\ &\quad * 0.574) = 0.184\end{aligned}$$

5. The geneticist P. M. Sheppard (1959) carried out a selection experiment on a laboratory population of the fruit fly *Drosophila melanogaster*. The stubble allele, which affects bristle shape of the fly, is dominant to the wild-type allele. Flies that are homozygous for stubble always die during embryonic development.

a) Sheppard started out with 86% normal flies and 14% stubble flies. Calculate the allele frequencies.

Let S be the stubble allele and s be the normal allele.

$$\text{Freq (SS)} = 0$$

$$\text{Freq (Ss)} = 0.14$$

$$\text{Freq (ss)} = 0.86$$

$$\text{Freq (S)} = p = 0 + (0.14 / 2) = 0.07$$

$$\text{Freq (s)} = q = 0.86 + (0.14 / 2) = 0.93$$

b) Assuming for now that wild-type and stubble flies do not differ in fitness, use the allele frequencies to calculate the mean fitness. Then predict the percentages of normal and stubble flies in the next generation. Show all work.

$$w_{SS} = 0 \text{ (because all flies with this genotype die)}$$

$$w_{Ss} = 1$$

$$w_{ss} = 1$$

Just because the fitness is 1 doesn't mean that every fly with Ss or SS will survive. What matters is that the fitnesses are the same for both -- and setting them to 1 makes the math easier, although in the end it doesn't change the result.

$$p^2 w_{SS} + 2pq w_{Ss} + q^2 w_{ss} = \bar{w}$$

$$(0.07 * 0.07 * 0) + (2 * 0.07 * 0.93 * 1) + (0.93 * 0.93 * 1) = \bar{w}$$

$$\bar{w} = 0.995$$

Predictions for the next generation: Divide all terms of the equation above by \bar{w} .

$$\text{New freq (SS)} = (0.07 * 0.07 * 0) / 0.995 = 0. \text{ (Duh.)}$$

$$\text{New freq (Ss)} = (2 * 0.07 * 0.93 * 1) / 0.995 = 0.131$$

$$\text{New freq (ss)} = (0.93 * 0.93 * 1) / 0.995 = 0.869$$

c) Sheppard introduced an additional source of selection: he removed 60% of the wild-type flies before they could breed in each generation. Repeat part b taking this into account.

OK, since each wild-type fly now has only a 40% chance of being left in the population to reproduce, we can assign a value of 0.4 to the fitness of the ss genotype.

$$p^2 w_{SS} + 2pq w_{Ss} + q^2 w_{ss} = \bar{w}$$

$$(0.07 * 0.07 * 0) + (2 * 0.07 * 0.93 * 1) + (0.93 * 0.93 * 0.4) = \bar{w}$$

$$\bar{w} = 0.476$$

Divide every term by \bar{w} , and the terms now add up to 1.0, and each term gives the predicted response to selection:

$$\text{New freq (SS)} = (0.07 * 0.07 * 0) / 0.476 = 0. \text{ (Double duh.)}$$

$$\text{New freq (Ss)} = (2 * 0.07 * 0.93 * 1) / 0.476 = 0.273$$

$$\text{New freq (ss)} = (0.93 * 0.93 * 0.4) / 0.476 = 0.727$$

6. The Old German Baptist Brethren, informally known as the "Dunkers", is a small religious denomination founded in Germany in 1708. Beginning in 1719, a number of Dunkers emigrated from Germany to Pennsylvania. As of 1950, there were about 3500 Dunkers in the United States. Dunkers are not as strict about their lifestyle as other similar religious groups, such as the Amish. However, Dunkers usually marry within their community. Dunkers who marry non-Dunkers often leave the community, and converts to the Dunker denomination are relatively rare.

In 1950, geneticist Bentley Glass studied a population of over 200 Dunkers in southern Pennsylvania. Glass used the MN blood group, a blood type controlled by a single gene with two loci. Individuals may be type M (homozygous for the M allele), N (homozygous for the N allele), or MN (heterozygous). The MN blood type has little clinical significance, and as far as is known there is no survival advantage in having one MN blood type over the other.

a) Glass found 102 Dunkers with type M blood, 96 with type MN, and 31 with type N. Calculate the allele frequencies.

Total population = $102 + 96 + 31 = 229$. (Yes, I have had students in the past who weren't sure how to calculate this!!)

Long way: Total number of M alleles in the population is two per person with type M blood and one per person with MN blood, so there's $(102 * 2) + 96 = 300$ M alleles in the population, out of $229 * 2 = 458$ alleles overall. $300 / 458 = 0.655$.

Let $p = \text{freq (M)} = 0.655$. Thus $q = \text{freq (N)} = 1 - p = 0.345$.

b) Calculate the expected numbers of people who would have each blood type if the population were in Hardy-Weinberg equilibrium. If the expected figures don't match what is observed, suggest why this might be the case.

Expected freq (MM) = $p^2 = 0.655 * 0.655 = 0.429$, or 98 people
Expected freq (MN) = $2pq = 2 * 0.655 * 0.345 = 0.452$, or 103 people
Expected freq (NN) = $q^2 = 0.345 * 0.345 = 0.119$, or 27 people

This is pretty close to H-W. In this case we really should do a chi-square test to see whether the differences are significant, but in this class I'm not going to make you do that. . .

c) In Germany today, about 30% of the population has type M blood, 50% has type MN, and 20% has type N. In the eastern United States, the figures are almost identical (29% M, 50% MN, 21% N.) Discuss why both of these sets of allele frequencies might differ from the frequencies in the Dunkers. (There could be many reasons, but restrict yourself to the most likely.)

A strong possibility is the founder effect -- the small population that founded the Dunker colony in America may just have had an unusually high level of M alleles by the "luck of the draw". Genetic drift is another; that has the strongest effects in a small population, and the Dunker population is small enough for drift to affect allele frequencies even in the absence of selection.

7. P. D. N. Hebert studied the frequencies of alleles for the gene that codes for the enzyme malate dehydrogenase (*Mdh*) in the "water flea," *Daphnia magna*, living in ponds near Cambridge, England. There are three alleles of the *Mdh* gene, abbreviated S, M and F. Hebert found the following genotypes:

genotype	observed number
SS	3
SM	8
SF	19
MM	15
MF	37
FF	32
total	114

a) Calculate the allele frequencies.

Easy. 114 individuals = 228 alleles. Frequency of S = $(3*2 + 8 + 19)/228$, i.e. 3 individuals with two S alleles and 8+19 individuals with one each, all divided by 228. $\text{freq}(S) = 0.145$. You calculate $\text{freq}(M)$ in the same way: $(15*2 + 8 + 37)/228 = 0.329$, and $\text{freq}(F) = (32*2 + 19 + 37)/228 = 0.526$. Check: They all add to 1.000.

b) Is the population in Hardy-Weinberg equilibrium?

Remember: for three alleles, the H-W equation is: $(p+q+r)^2 = 1$. That expands to:
 $p^2 + 2pq + 2pr + q^2 + 2qr + r^2 = 1$
 and each term gives you a predicted genotype frequency. So let p, q, and r be the frequencies of S, M and F, respectively, and you can plug-and-chug:

genotype	obs.	pred. (rounded)
SS	3	$p^2 = 0.02103 * 114 = 2$
SM	8	$2pq = 0.0954 * 114 = 11$
SF	19	$2pr = 0.153 * 114 = 17$
MM	15	$q^2 = 0.108 * 114 = 12$
MF	37	$2qr = 0.346 * 114 = 39$
FF	32	$r^2 = 0.277 * 114 = 32$

Pretty close to H-W. Might be a little off.

8. *Avena fatua* is a species of wild oat (a type of grass). Jain and Marshall studied wild oat population genetics in California. One of the traits they examined was the pubescence (hairiness) of the leaf sheath, which is controlled by a single locus with two alleles, written L and l. They found that the frequencies of genotypes in one population were:

LL 57.1% Ll 7.1% ll 35.8%

a) Calculate the allele frequencies.

This oughta be easy by now. The quick way:

$$p = \text{freq}(L) = 0.571 + (0.071/2) = 0.606$$

$$q = \text{freq}(l) = 0.358 + (0.071/2) = 0.394$$

b) Predict what the genotype frequencies should be under Hardy-Weinberg equilibrium. If there is a difference between actual and predicted frequencies, explain briefly why the differences might exist.

Ho hum. . .

genotype	observed	predicted
LL	57.1%	$p^2 = (0.606)^2 = 0.367$
Ll	7.1%	$2pq = 2(0.606)(0.394) = 0.478$
ll	35.8%	$q^2 = (0.394)^2 = 0.155$

Looks a lot like inbreeding, doesn't it? Again, you've got that decrease in heterozygotes and increase in both homozygotes.

c) Calculate F.

The simplest formula for me to remember is that $F = 1 - (\text{actual heterozygote frequency} / \text{predicted heterozygote frequency})$. So $F = 1 - (0.071 / 0.478) = 0.851$. Another way to do it would be to plug in one of the formulas in your handout such as $\text{freq}(LL) = p^2 + pqF$ and solve for F; you get the same answer.

9. The biologist B. Battaglia raised the marine copepod *Tisbe reticulata* (a small free-swimming marine crustacean) under crowded conditions. *T. reticulata* has one gene with two alleles, V_v and V_m , showing incomplete dominance. In one of his tanks, Battaglia counted 1751 copepods: 353 V_vV_v , 1069 V_vV_m , and 329 V_mV_m .

a) Show that the population is not in Hardy-Weinberg equilibrium.

Let $p = \text{freq}(V_v)$. Then $p = (353 \cdot 2 + 1069) / (1751 \cdot 2) = 0.507$. And $q = \text{freq}(V_m) = 1 - p = 0.493$. If the population were in H-W, then the frequency of V_vV_v individuals would be p^2 , or 0.257; in reality it is $353/1751 = 0.202$. The frequency of V_mV_m individuals would be q^2 , or 0.243; the actual frequency is 0.188; and the expected frequency of V_vV_m would be $2pq$, or 0.500, vs. the actual frequency of 0.611.

b) Discuss why it might not be in Hardy-Weinberg equilibrium.

Well, we have an excess of heterozygotes here, so it's not inbreeding. My guess would be selection favoring heterozygotes. Another possibility would be disassortative mating: if V_vV_v individuals strongly prefer to mate with V_mV_m individuals, and vice versa, that could drive up the number of heterozygotes.

10. True story: In 1912, the geneticist W. H. Goddard suggested that feeble-mindedness was caused by Mendelian inheritance at a single locus with two alleles. Persons homozygous for the recessive, feeble-minded allele (call it *f*) were dopes, dummies, and dimwits -- "incapable of managing their affairs with ordinary prudence", as Goddard said. Heterozygotes (*Ff*) and homozygous dominants (*FF*) were of normal intelligence. **This is not actually true** -- but pretend that it is, for the purposes of working this problem.

a) According to the 1910 census, the population of the United States was 91,972,266. Goddard estimated that 1% of the population was feeble-minded. Assume that the population of the US was in Hardy-Weinberg equilibrium. Calculate the allele frequencies, and then calculate the percentages of the population that would be heterozygous and homozygous dominant.

Easy enough. If the US is in H-W, then the frequency of *ff* individuals is 0.01. Let the frequency of the normal and the feeble-minded allele be *p* and *q*, respectively. Then *q* is the square root of 0.01, or 0.1; and *p* = 0.9.

b) At one time or another, thirty states had laws mandating the compulsory surgical sterilization of the feebleminded. (As of 1996, Arkansas and nine other states still *did* have such a law on the books.) There were organizations in the early 20th century that lobbied for their enactment nationwide.

Imagine that, in some alternate-reality USA, a mandatory, nationwide law really was put into effect that forced the sterilization of all feebleminded individuals before they could reproduce. Assume that the authorities were so efficient that they were able to track down and sterilize 90% of the feebleminded—and that they never, ever sterilized anyone who wasn't feebleminded. What would be the frequencies of genotypes, and of alleles, after one generation?

Sterilization removes you permanently from the breeding population, just as surely as execution would. So let's set the fitnesses so that the fitness of the *ff* genotype is only one-tenth of the fitnesses of the other two:

$$w_{FF} = 1.0 \quad w_{Ff} = 1.0 \quad w_{ff} = 0.1$$

Plug these into our favorite formula. . .

$$p^2 w_{FF} + 2pq w_{Ff} + q^2 w_{ff} = \bar{w}$$

$$0.9 \cdot 0.9 \cdot 1 + 2 \cdot 0.9 \cdot 0.1 \cdot 1 + 0.1 \cdot 0.1 \cdot 0.1 = \bar{w}$$

$$\bar{w} = 0.991$$

Now divide all terms by \bar{w} to get the predicted response to selection:

$$\text{freq (FF)} = 0.9 \cdot 0.9 \cdot 1 / 0.991 = 0.817$$

$$\text{freq (Ff)} = 2 \cdot 0.9 \cdot 0.1 \cdot 1 / 0.991 = 0.182$$

$$\text{freq (ff)} = 0.1 \cdot 0.1 \cdot 0.1 / 0.991 = 0.0001$$

We can quickly calculate p after one round as $0.817 + (0.182/2) = 0.908$. So one generation of selection has changed the frequency of the gene for normal intelligence by a factor of 0.8%. End result: Even if you could create a government program that was as efficiently run as this one—which we may well be skeptical of—it would have a tiny effect on allele frequencies. Eugenics isn't efficient.

11. *Fundulus heteroclitus* (common name: mummichog) is a small fish that lives in bays and estuaries along the east coast of North America, from Newfoundland to Florida. It's been extensively used in evolutionary studies.¹

Northern populations of *F. heteroclitus* are virtually completely homozygous for an allele of a liver enzyme called *Ldh-B* (lactate dehydrogenase B), an allele which is just called *Ldh-B^a*. Southern populations are virtually completely homozygous for the other major allele of *Ldh-B*, referred to as *Ldh-B^b*.

Ldh-B has an effect on the survival of fish embryos when exposed to temperature stress. In an experimental test of this survival, *Fundulus heteroclitus* eggs from different populations were incubated at 30°C, under heat stress. 57% of the eggs from a southern population hatched successfully at 30°C. Only 22% of the eggs from a northern population hatched. Interestingly enough, 87% of the eggs that were heterozygous at the *Ldh-B* locus successfully hatched.

a) Suppose you put 500 northern *Fundulus* fish and 500 southern fish in a large tank, allowed them to breed freely, and then incubated all the eggs at 30°C. Predict the frequencies of **both** alleles **and** genotypes after one generation.

For simplicity, I'll call the two alleles A and B.
Northern populations are virtually all AA, and
Southern populations are virtually all BB.

500 AA fish and 500 BB fish means that the frequency of both alleles is 0.5. So let $p = \text{frequency}(A)$ and $q = \text{frequency}(B)$; $p = q = 0.5$.

You could use the raw survival percentages as fitness values and you'd end up with the right answer. But the math is easier, and the meaning is more intuitively right, if you divide them all by the highest. So

$$w_{AA} = 0.22/0.87 = 0.25$$

$$w_{AB} = 0.87/0.87 = 1.00$$

$$w_{BB} = 0.57/0.87 = 0.66$$

Plug these numbers in to

$$p^2w_{AA} + 2pqw_{AB} + q^2w_{BB} = \bar{w}:$$

$$(0.5)^2(0.25) + 2(0.5)(0.5)(1) + (0.5)^2(0.66) = \bar{w}$$

$$0.063 + 0.5 + 0.165 = 0.73$$

¹ Powers, D.A., M. Smith, I. Gonzalez-Villasenor, L. DiMichele, D. Crawford, G. Bernardi, and T. Lauerman. A multidisciplinary approach to the "selectionist/neutralist" controversy using the model teleost, *Fundulus heteroclitus*. *Oxford Surveys in Evolutionary Biology* **9**: 43-107.

Now the new genotype frequencies are given by dividing each term of the above equation by \bar{w} . So:

$$\text{new freq(AA)} = p^2 w_{AA} / \bar{w} = 0.063 / 0.73 = 0.086$$

$$\text{new freq(AB)} = 2pq w_{AB} / \bar{w} = 0.5 / 0.73 = 0.685$$

$$\text{new freq(BB)} = q^2 w_{BB} / \bar{w} = 0.165 / 0.73 = 0.23$$

Check: Do these three add up to 1? Yes, they do, allowing for round-off error.

Now you have to calculate the new allele frequencies. There are several ways you could do that, but here's the long way:

$$\text{new } p = (86 \cdot 2 + 685) / 2000 = 0.428$$

$$\text{new } q = (230 \cdot 2 + 685) / 2000 = 0.572$$

EITHER:

b) Predict the frequencies of **both** alleles **and** genotypes after a second generation.

Start with

$$\text{new } p = (86 \cdot 2 + 685) / 2000 = 0.428$$

$$\text{new } q = (230 \cdot 2 + 685) / 2000 = 0.572$$

Use the fitness values from the previous problem, and calculate \bar{w} :

$$p^2 w_{AA} + 2pq w_{AB} + q^2 w_{BB} = \bar{w}:$$

$$(0.428)^2(0.25) + 2(0.428)(0.572)(1) + (0.572)^2(0.66) = \bar{w}$$

$$0.046 + 0.49 + 0.22 = 0.76$$

Now the new genotype frequencies are given by dividing each term of the above equation by \bar{w} . So:

$$\text{new freq(AA)} = p^2 w_{AA} / \bar{w} = 0.046 / 0.76 = 0.06$$

$$\text{new freq(AB)} = 2pq w_{AB} / \bar{w} = 0.49 / 0.76 = 0.64$$

$$\text{new freq(BB)} = q^2 w_{BB} / \bar{w} = 0.22 / 0.76 = 0.29$$

Check: Do these add up to 1? Yes, they do, allowing for rounding off.

And now you have to calculate the new allele frequencies:

$$\text{new } p = (6 \cdot 2 + 64) / 200 = 0.38$$

$$\text{new } q = (29 \cdot 2 + 64) / 200 = 0.61$$

OR:

c) Suppose that, when you actually do this experiment, you discover that the actual heterozygote frequency is 0.55. Calculate F .

The right way to calculate F in this situation is one I didn't put on the handout: $F = 1 - (H_o/H_e)$, where H_o is the observed frequency of heterozygotes and H_e is the expected frequency of heterozygotes. Here, H_o is 0.55, and H_e is the predicted heterozygote frequency. If it's the second generation after the founding generation that we're talking about, $H_e = 0.68$. $H_o/H_e = 0.81$, so $F = 0.19$.

You might have also remembered the formula that $\text{freq}(AB) = 2pq - 2pqF$. If you plug in the numbers there, you end up with a negative value for F . Technically, that's not the right way to do it -- that formula works only if the only thing that's pushing the population away from H-W Equilibrium is inbreeding. Here, we have selection going on as well, and that makes the math messier.

12. In 1973, Francisco Ayala and colleagues sampled a population of deep-sea starfish of the species *Nearchaster aciculosus*, living at a depth of 1244 meters off the coast of San Diego, California. They used electrophoresis to determine the frequencies of various alleles. One gene that they looked at, for the enzyme tetrazolium oxidase (abbreviated *To*), had four alleles (abbreviated *To-100*, *To-107*, *To-114*, and *To-120*.) They found that the frequency of each allele was as follows:

<i>To-100</i>	0.206
<i>To-107</i>	0.147
<i>To-114</i>	0.353
<i>To-120</i>	0.294

Ayala et al. observed that in this population, 70.6% of the starfish were heterozygous. Does this observation match the prediction for a population in Hardy-Weinberg equilibrium?

Oooh! Four alleles! The formula for the four-allele Hardy-Weinberg equilibrium is

$$(p + q + r + s)^2 = 1$$

which comes to

$$p^2 + 2pq + 2pr + 2ps + q^2 + 2qr + 2qs + r^2 + 2rs + s^2 = 1$$

Each term of the equation gives the predicted frequency for a particular genotype; p^2 is the predicted frequency for the *To-100/To-100* genotype, $2pq$ is the predicted frequency for the *To-100/To-107* genotype. . . and so on. p^2 , q^2 , r^2 , and s^2 are the predicted frequencies of the four possible homozygous genotypes. All the other terms in the above equation are the predicted frequencies for the six possible heterozygous genotypes. So if the population is in Hardy-Weinberg equilibrium, then the frequency of all heterozygotes would be

$$2pq + 2pr + 2ps + 2qr + 2qs + 2rs$$

Plug in the numbers. . .

$$2(0.206)(0.147) + 2(0.206)(0.353) + 2(0.206)(0.294) + 2(0.147)(0.353) + 2(0.147)(0.294) + 2(0.353)(0.294) = 0.725.$$

The actual frequency of heterozygotes is 0.706. Close to the predicted value, but maybe a little low.