

# Pop Gen HW2

Dylan Sosa

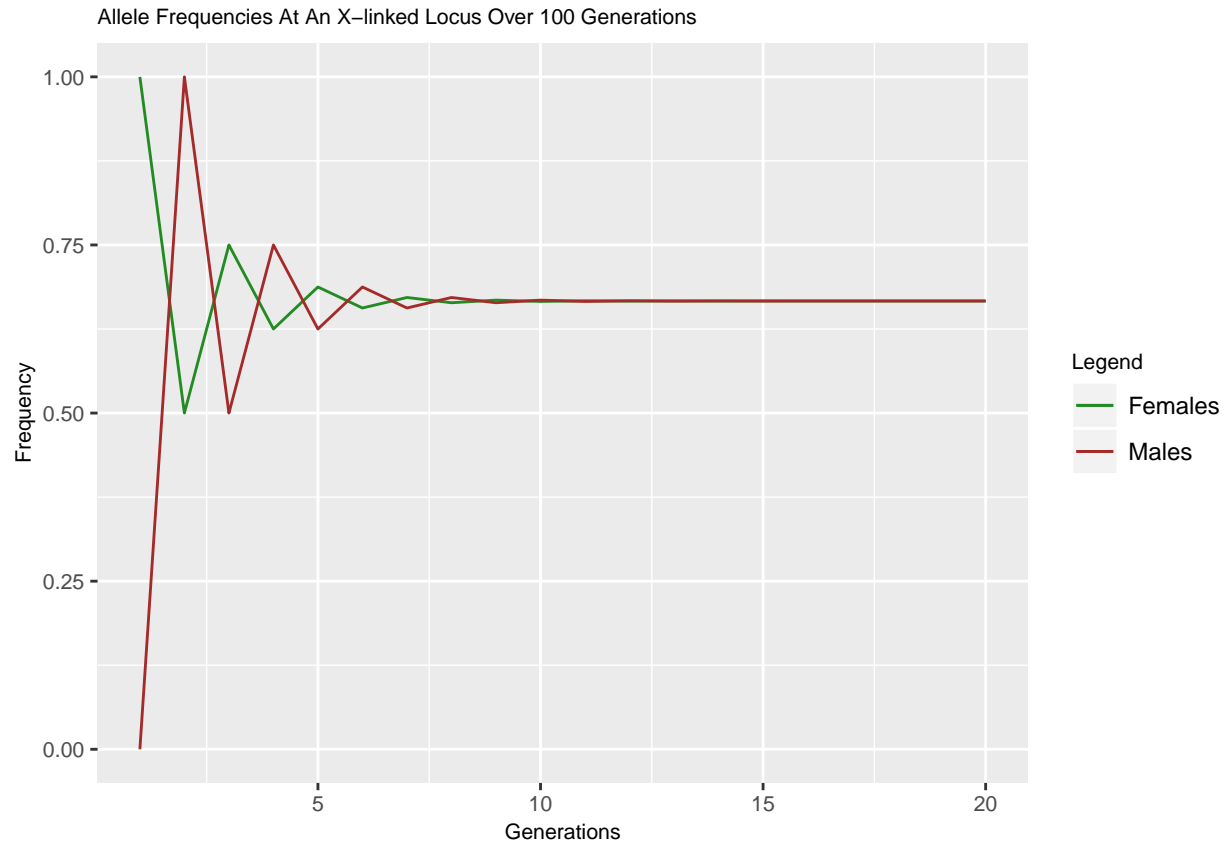
1/28/2020

## Problem 1

1. Using R, write a program to graph the male and female allele frequencies at an X-linked locus over several generations for a population that is started with all A1A1 females ( $pf = 1$ ) and A2 males ( $pm = 0$ ). Show program and output.

```
generations <- 20 #arbitrarily large gen
pf <- c(1)
pm <- c(0)
for( i in c(1:generations-1)){
  females <- (pm[i]+pf[i])/2
  males <- pf[i]
  pf <- c(pf, females)
  pm <- c(pm, males)
}

ggplot() +
  geom_line(aes(x = c(1:generations), y = pf, color = 'Females'), show.legend = TRUE) +
  geom_line(aes(x = c(1:generations), y = pm, color = 'Males'), show.legend = TRUE) +
  xlab('Generations') +
  ylab('Frequency') +
  ggtitle('Allele Frequencies At An X-linked Locus Over 100 Generations') +
  scale_colour_manual("Legend", values = c("forestgreen", "brown")) +
  theme(
    plot.title = element_text(size = 8),
    axis.title = element_text(size = 8),
    axis.text = element_text(size = 8),
    legend.title = element_text(size = 8)
  )
```



## Problem 2

2. What is the equation describing the equilibrium frequency of the mutant allele?

$$p' = \frac{p(1-qhs)}{\bar{w}}(1-u)$$

where

$$\bar{w} = 1 - sq(q + 2ph)$$

- (a) Equation describing equilibrium frequency:

$$p' = \frac{(1-\hat{q})(1-\hat{q}hs)}{1-s\hat{q}(\hat{q}+2(1-\hat{q})h)}(1-u)$$

$$p' = sq^2(2h-1) - s\hat{q}(h+hu) + u = 0$$

- (b)  $h = 0$  :

$$\hat{q} = \sqrt{\frac{u}{s}}$$

$$h = 0.5 :$$

$$\hat{q} = \frac{2u}{su+s}$$

- (c) Ignoring back-mutation is reasonable in this formulation because the mutant allele frequency is very small which makes  $A^* \rightarrow A$  very small as well. This allows us to safely ignore back-mutation.

- (d) Rate of mutation to recessive lethals:  $2e-6$

Equilibrium frequency of lethal alleles in the population:  $2e-4$

```
s <- 1
u <- 2e-6
```

```
q <- 2e-4
h <- (s * q**2 - u) / (2*s*(q**2)-s*q-s*q*u)
cat('Value of h is:', h)
```

```
## Value of h is: 0.009803902
```

## Problem 3

3. What will be the incidence of this anemia 2 generations from now? 20 generations from now? (Assume that the heterozygote has no loss in fitness.) What lesson can be learned from this about eugenics programs intending to eliminate deleterious mutations from the population?

```
generations = 20
p = .5/144
for( i in c(1:generations-1)){
  if(i == 2){
    heterozygotes2 <- 2*p*(1-p)
  }
  heterozygotes <- 2*p*(1-p) # have disease
  p = 2*p*(1-p)*0.5
}
cat('Incidence at 20 generations:',heterozygotes)
```

```
## Incidence at 20 generations: 0.006492084
```

```
cat('Incidence at 2 generations:',heterozygotes2)
```

```
## Incidence at 2 generations: 0.006872606
```

It seems that the incidence of deleterious mutations will decrease slowly, and that eugenics programs should promote the mating of heterozygous individuals.

## Problem 4

4. Assuming a Wright-Fisher infinite sites model, provide an estimate of the effective population size of the population from which this individual has been sampled.  $\pi = \prod_L$

$$\hat{\theta} = 4N_e\mu = \pi$$

$$N_e = \frac{\pi}{4\mu}$$

```
seqLength <- 1e4
mu <- 1.6e-9
pi <- 21/seqLength
Ne <- pi/(4*mu)
cat('Effective population is:', Ne)
```

```
## Effective population is: 328125
```