## ALLELIC INDEPENDENCE GROUP WORK

## Bayesian Exercise

A rapid test for covid-19 is set up outside a supermarket and is available to anyone who wishes. The test has a false-positive rate of $5 \%$ and a false-negative rate of $30 \%$.

If the disease has a prevalence in that population of $20 \%$. What is the probability a person who tests positive does actually have the disease? i.e. calculate $\operatorname{Pr}(B \mid A)$ if $A$ is the event that a test is positive, and $B$ is the event that a person has the disease. Use Bayes' Theorem.

## Bayesian Exercise

A: positive test; $\bar{A}$ : negative test;
$\operatorname{Pr}(B)=0.20$ : infected; $\operatorname{Pr}(\bar{B})=0.80$ : unaffected.
$\operatorname{Pr}(A \mid \bar{B})=0.05: \operatorname{Pr}(\bar{A} \mid B)=0.30=>\operatorname{Pr}(A \mid B)=0.70$
$\operatorname{Pr}(A)=\operatorname{Pr}(A \mid B) \operatorname{Pr}(B)+\operatorname{Pr}(A \mid \bar{B}) \operatorname{Pr}(\bar{B})=0.18$.

$$
\operatorname{Pr}(B \mid A)=\frac{\operatorname{Pr}(A \mid B) \operatorname{Pr}(B)}{\operatorname{Pr}(A)}=\frac{0.14}{0.18}=0.78
$$

## Permutation Test

Put four folds across the width, of a sheet of paper to mark five genotypess, and then one fold the middle of the page, top to bottom. This gives 10 alleles, six of type $A$ and four of type $a$.

|  | Allele | Allele |
| :---: | :---: | :---: |
| Genotype 1: | A | A |
| Genotype 2: | A | A |
| Genotype 3: | A | A |
| Genotype 4: | a | a |
| Genotype 5: | a | a |

Tear along the folds give a deck of 10 alleles. Shuffle the deck and deal into 5 pairs, to give five new genptypes. How many heterozygotes are there?

## NIST Data

Go to https://strbase.nist.gov/NISTpop.htm and look at "Excel file of revised allele frequencies."

How many loci appear not to be in Hardy-Weinberg Equilibrium?

| Group | No. $p<0.05$ |
| :--- | :---: |
| Total | 8 |
| African American | 0 |
| Caucasian | 3 |
| Hispanic | 1 |
| Asian | 2 |

## Why is $p<0.05$

- Population structure for 'Total' group.
- Multiple testing: each of 29 STRs should be tested with significance level of 0.05/29.
- Possible genotyping error.
- HWE issue in whole database avoided with use of BaldingNichols match probability.

