Forensic Genetics

Module 19 – Topics 3 & 4 Answers

Topic 3 Group Work

Allelic Independence

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 $\Pr(B|A)$ 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;
$\Pr(B) = 0.20$: infected; $\Pr(\bar{B}) = 0.80$: unaffected.

$\Pr(A|\bar{B}) = 0.05$: $\Pr(\bar{A}|B) = 0.30 => \Pr(A|B) = 0.70$
$\Pr(A) = \Pr(A|B)\Pr(B) + \Pr(A|\bar{B})\Pr(\bar{B}) = 0.18.$

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

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?

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 Balding-Nichols match probability.

Topic 4 Group Work

LR Calculations

Exercise 1a: LR – Binary Model

Consider a simple two-person mixture profile (e.g. contributors are unrelated, ignoring population structure, no drop-outs/drop-ins), where $G_C = ABCD$. Let K denote a known contributor with observed profile $G_K = CD$, and S the POI with profile $G_S = AB$.

- $G_S = AB$ and $G_K = CD$, with

$$H_p : \text{K + POI (S)} \quad \text{and} \quad H_d : \text{K + Unknown (U)}$$

What are the LRs for $p_A = p_B = p_C = p_D = 0.1$?

Exercise 1a: LR − Binary Model

Consider a simple two-person mixture profile (e.g. contributors are unrelated, ignoring population structure, no drop-outs/drop-ins), where $G_C = ABCD$. Let K denote a known contributor with observed profile $G_K = CD$, and S the POI with profile $G_S = AB$.

- LR $= \dfrac{\text{Pr}(ABCD|H_p:\ K+S)}{\text{Pr}(ABCD|H_d:\ K+U)} = \dfrac{1}{2p_A p_B} = 50$

Exercise 1b: LR − Binary Model

Consider a simple two-person mixture profile (e.g. contributors are unrelated, ignoring population structure, no drop-outs/drop-ins), where $G_C = ABCD$. Let K denote a known contributor with observed profile $G_K = CD$, and S the POI with profile $G_S = AB$.

- $G_S = AB$ and $G_K = CD$, with

  $$H_p : \text{K + POI (S)} \quad \text{and} \quad H_d : \text{K + Unknown (U)}$$

- $G_S = AB$ and $G_K = CD$, with

  $$H_p : \text{K + S} \quad \text{and} \quad H_d : \text{2U}$$

What are the LRs for $p_A = p_B = p_C = p_D = 0.1$?

Exercise 1b: LR − Binary Model

Consider a simple two-person mixture profile (e.g. contributors are unrelated, ignoring population structure, no drop-outs/drop-ins), where $G_C = ABCD$. Let K denote a known contributor with observed profile $G_K = CD$, and S the POI with profile $G_S = AB$.

- $LR = \dfrac{\mathrm{Pr}(ABCD|H_p\text{: } K+S)}{\mathrm{Pr}(ABCD|H_d\text{: } K+U)} = \dfrac{1}{2p_A p_B} = 50$

- $LR = \dfrac{\mathrm{Pr}(ABCD|H_p\text{: } K+S)}{\mathrm{Pr}(ABCD|H_d\text{: } 2U)} = \dfrac{1}{24 p_A p_B p_C p_D} = 417$

Exercise 1c: LR − Binary Model

Consider a simple two-person mixture profile (e.g. contributors are unrelated, ignoring population structure, no drop-outs/drop-ins), where $G_C = ABCD$. Let K denote a known contributor with observed profile $G_K = CD$, and S the POI with profile $G_S = AB$.

- $G_S = AB$ and $G_K = CD$, with

$$H_p : \text{K + POI (S)} \quad \text{and} \quad H_d : \text{K + Unknown (U)}$$

- $G_S = AB$ and $G_K = CD$, with

$$H_p : \text{K + S} \quad \text{and} \quad H_d : \text{2U}$$

- $G_S = AB$ and the second contributor is unknown

$$H_p : \text{S + U} \quad \text{and} \quad H_d : \text{2U}$$

What are the LRs for $p_A = p_B = p_C = p_D = 0.1$?

Exercise 1c: LR − Binary Model

Consider a simple two-person mixture profile (e.g. contributors are unrelated, ignoring population structure, no drop-outs/drop-ins), where $G_C = ABCD$. Let K denote a known contributor with observed profile $G_K = CD$, and S the POI with profile $G_S = AB$.

- $\text{LR} = \dfrac{\Pr(ABCD|H_p\colon K+S)}{\Pr(ABCD|H_d\colon K+U)} = \dfrac{1}{2p_Ap_B} = 50$

- $\text{LR} = \dfrac{\Pr(ABCD|H_p\colon K+S)}{\Pr(ABCD|H_d\colon 2U)} = \dfrac{1}{24p_Ap_Bp_Cp_D} = 417$

- $\text{LR} = \dfrac{\Pr(ABCD|H_p\colon S+U)}{\Pr(ABCD|H_d\colon 2U)} = \dfrac{1}{12p_Ap_B} = 8$

Exercise 2: LR – Binary Model

a) Considering the previous exercise, what do you expect to happen to the LRs if we use match probabilities instead of profile probabilities? *Increase, decrease or stay the same?*

b) Verify your answer by using the appropriate Balding-Nichols formula with $\theta = 0.03$ in Exercise 1a.

c) Ignoring a known contributor under H_d (but not under H_p) is *favorable/unfavorable/irrelevant* to the defendant? (Hint: compare your answers from 1a and 1b).

Exercise 2: LR – Binary Model

a) Considering the previous exercise, what do you expect to happen to the LRs if we use match probabilities instead of profile probabilities? *Increase, **decrease** or stay the same?*

Match probabilities are generally bigger than profile probabilities: having seen a profile once, increases the chance we will see it again. In our simplified setting, the LR is the reciprocal of the match probability. An increase in match probability, will thus lead to a decrease in the LR.

Exercise 2: LR − Binary Model

b) Verify your answer by using the appropriate Balding-Nichols formula with $\theta = 0.03$ in Exercise 1a.

The match probability is 0.03 (compared to a profile probability of 0.02). The LR then decreases from 50 to 33.8.

Exercise 2: LR – Binary Model

c) Ignoring a known contributor under H_d (but not under H_p) is *favorable*/**unfavorable**/*irrelevant* to the defendant? (Hint: compare your answers from 1a and 1b).

If we ignore a known contributor under H_d, the LR will be larger than when assuming K as a known profile under both hypotheses. This is because K will explain many of the observed alleles (especially in case of being a major donor). This will thus be unfavorable to the defendant.