Forensic Genetics

Module 16 – Section 4 Answers

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 $G_S = AB$.

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

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

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

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

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

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 $G_S = AB$.

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

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

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

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

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

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

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

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

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 $G_S = AB$.

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

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

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

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

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

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

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

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.

Exercise 3: LR – Continuous Model

The epg for a 3-person mixture at locus vWA is as follows:
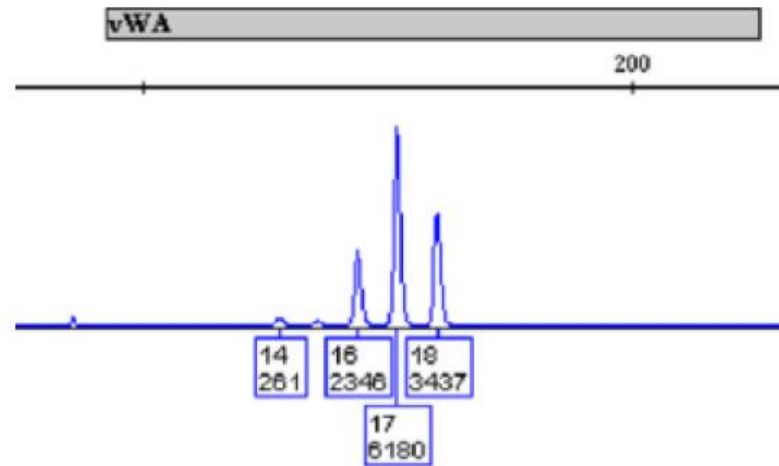
We would like to assess the LR under the hypothesis that:

H_p : $G_S = 17, 18$ and 2U are the source of the sample.

H_d :   3U are the source of the sample.

Exercise 3: LR – Continuous Model

Suppose the following weights have been established for locus vWA:

Genotype Set	Donor 1	Donor 2	Donor 3	Weight
S_1	16, 18	17, 17	14, 14	0.00045
S_2	16, 18	17, 17	14, 15	0.00017
S_3	16, 18	17, 17	14, 16	0.00008
S_4	16, 18	17, 17	14, 17	0.00002
S_5	16, 18	17, 17	14, 18	0.00054
S_6	18, 18	17, 17	14, 16	0.00005
S_7	16, 16	17, 18	14, 14	0.00218
S_8	16, 16	17, 18	Q, 14	0.00010
S_9	16, 16	17, 18	14, 15	0.00207
S_{10}	16, 16	17, 18	14, 16	0.00511
S_{11}	16, 16	17, 18	14, 17	0.02030
S_{12}	16, 16	17, 18	14, 18	0.00279
S_{13}	16, 17	17, 18	14, 14	0.19300
S_{14}	16, 17	17, 18	Q, 14	0.00368
S_{15}	16, 17	17, 18	14, 15	0.15800
S_{16}	16, 17	17, 18	14, 16	0.28700
S_{17}	16, 17	17, 18	14, 17	0.21000
S_{18}	16, 17	17, 18	14, 18	0.11400
S_{19}	17, 17	17, 18	14, 16	0.00016

Exercise 3: LR – Continuous Model

The LR can now be assessed by writing the ratio in the form:

$$\text{LR} = \frac{\Pr(G_C|G_S, H_p, I)}{\Pr(G_C|G_S, H_d, I)}$$

$$= \frac{\sum_j \Pr(G_C|S_j)\Pr(S_j|H_p)}{\sum_{j'} \Pr(G_C|S_{j'})\Pr(S_{j'}|H_d)}$$

$$= \frac{\sum_j w_j \Pr(S_j|H_p)}{\sum_{j'} w_{j'} \Pr(S_{j'}|H_d)}.$$

The two propositions each define sets of genotypes S, and the weights w describe how well these sets fit our observed data G_C. Under H_p all the genotype sets S_j usually include G_S.

Exercise 3: LR − Continuous Model

Use the following allele frequencies (and assume $\theta = 0$):

Allele	Frequency
14	0.1146
15	0.1071
16	0.2044
17	0.2726
18	0.2090

a) What is $\Pr(E|H_p)$?

b) What is $\Pr(E|H_d)$?

c) What is the LR for this locus?

Exercise 3: LR – Continuous Model

Use HWE to calculate genotype frequencies:

Genotype Set	Donor 1	Donor 2	Donor 3	Weight	$\Pr(U|H_p)$
S_1	16, 18	17, 17	14, 14	0.00045	-
S_2	16, 18	17, 17	14, 15	0.00017	-
S_3	16, 18	17, 17	14, 16	0.00008	-
S_4	16, 18	17, 17	14, 17	0.00002	-
S_5	16, 18	17, 17	14, 18	0.00054	-
S_6	18, 18	17, 17	14, 16	0.00005	-
S_7	16, 16	17, 18	14, 14	0.00218	0.000549
S_8	16, 16	17, 18	Q, 14	0.00010	0.000884
S_9	16, 16	17, 18	14, 15	0.00207	0.001026
S_{10}	16, 16	17, 18	14, 16	0.00511	0.001957
S_{11}	16, 16	17, 18	14, 17	0.02030	0.002610
S_{12}	16, 16	17, 18	14, 18	0.00279	0.002001
S_{13}	16, 17	17, 18	14, 14	0.19300	0.001464
S_{14}	16, 17	17, 18	Q, 14	0.00368	0.002358
S_{15}	16, 17	17, 18	14, 15	0.15800	0.002736
S_{16}	16, 17	17, 18	14, 16	0.28700	0.005221
S_{17}	16, 17	17, 18	14, 17	0.21000	0.006963
S_{18}	16, 17	17, 18	14, 18	0.11400	0.005338
S_{19}	17, 17	17, 18	14, 16	0.00016	0.003481

Exercise 3: LR – Continuous Model

Use HWE to calculate genotype frequencies:

Genotype Set	Donor 1	Donor 2	Donor 3	Weight	$\Pr(U|H_d)$
S_1	16, 18	17, 17	14, 14	0.00045	0.000083
S_2	16, 18	17, 17	14, 15	0.00017	0.000008
S_3	16, 18	17, 17	14, 16	0.00008	0.000297
S_4	16, 18	17, 17	14, 17	0.00002	0.000397
S_5	16, 18	17, 17	14, 18	0.00054	0.000304
S_6	18, 18	17, 17	14, 16	0.00005	0.000297
S_7	16, 16	17, 18	14, 14	0.00218	0.000063
S_8	16, 16	17, 18	Q, 14	0.00010	0.000101
S_9	16, 16	17, 18	14, 15	0.00207	0.000117
S_{10}	16, 16	17, 18	14, 16	0.00511	0.000223
S_{11}	16, 16	17, 18	14, 17	0.02030	0.000297
S_{12}	16, 16	17, 18	14, 18	0.00279	0.000228
S_{13}	16, 17	17, 18	14, 14	0.19300	0.000167
S_{14}	16, 17	17, 18	Q, 14	0.00368	0.000269
S_{15}	16, 17	17, 18	14, 15	0.15800	0.000312
S_{16}	16, 17	17, 18	14, 16	0.28700	0.000595
S_{17}	16, 17	17, 18	14, 17	0.21000	0.000793
S_{18}	16, 17	17, 18	14, 18	0.11400	0.000608
S_{19}	17, 17	17, 18	14, 16	0.00016	0.000397

Exercise 3: LR – Continuous Model

The LR can now be assessed by writing the ratio in the form:

$$\mathsf{LR} = \frac{\Pr(G_C|G_S, H_p, I)}{\Pr(G_C|G_S, H_d, I)}$$

$$= \frac{\sum_j \Pr(G_C|S_j) \Pr(S_j|H_p)}{\sum_{j'} \Pr(G_C|S_{j'}) \Pr(S_{j'}|H_d)}$$

$$= \frac{\sum_j w_j \Pr(S_j|H_p)}{\sum_{j'} w_{j'} \Pr(S_{j'}|H_d)}.$$

The two propositions each define sets of genotypes S, and the weights w describe how well these sets fit our observed data G_C. Under H_p all the genotype sets S_j usually include G_S.

Exercise 3: LR – Continuous Model

Use the following allele frequencies (and assume $\theta = 0$):

Allele	Frequency
14	0.1146
15	0.1071
16	0.2044
17	0.2726
18	0.2090

a) What is $\Pr(E|H_p)$?   0.004365

b) What is $\Pr(E|H_d)$?   0.000498

c) What is the LR for this locus?

Exercise 3: LR – Continuous Model

Use the following allele frequencies (and assume $\theta = 0$):

Allele	Frequency
14	0.1146
15	0.1071
16	0.2044
17	0.2726
18	0.2090

a) What is $\Pr(E|H_p)$?   0.004365

b) What is $\Pr(E|H_a)$?   0.000498

c) What is the LR for this locus? $LR = 8.77$

Exercise 3: LR – Continuous Model

The overall LR is a combination of all loci (here compared with a binary model):

Locus	LR_B	LR_C
D10S1248	0.97	4.69
vWA	1.24	8.21
D16S539	0.45	5.32
D2S1338	2.27	31.22
D8S1179	0.51	7.79
D21S11	0.94	9.98
D18S51	3.85	52.08
D22S1045	4.32	59.18
D19S433	0.92	7.17
TH01	0.97	13.31
FGA	1.39	21.14
D2S441	0.65	4.84
D3S1358	0.93	13.22
D1S1656	5.55	106.14
D12S391	1.42	21.34
SE33	6.23	69.53
Overall LR	356	3.13×10^{19}

Source: The interpretation of single source and mixed DNA profiles (Taylor et al., 2013).