

### Question 3

**Separation = A is independent of C given Y**

**Sufficiency = A is independent of Y given C**

Let A = gender and C and Y vary

Example 1 (**Separation** holds but not **Sufficiency**)

Let A = Gender, C = Diabetes, and Y = Hyperlipidemia

**Separation** Argument: **Gender** is independent of **Diabetes** given **Hyperlipidemia**

$P(\text{Gender} \mid \text{Diabetes} = \text{YES}, \text{Hyperlipidemia} = \text{YES}) = [0.571, 0.429]$  (Male, Female)

$P(\text{Gender} \mid \text{Hyperlipidemia} = \text{YES}) = [0.571, 0.429]$  (Male, Female)

Therefore, removing C (Diabetes) from evidence has no effect on the probability and hence independence is shown.

**Sufficiency** Argument (Proof by contradiction): **Gender** is independent of **Hyperlipidemia** given **Diabetes**

$P(\text{Gender} \mid \text{Diabetes} = \text{YES}, \text{Hyperlipidemia} = \text{YES}) = [0.571, 0.429]$  (Male, Female)

$P(\text{Gender} \mid \text{Diabetes} = \text{YES}) = [0.535, 0.465]$  (Male, Female)

Hence, there is **no sufficiency** since Gender (A) is not independent of Hyperlipidemia (Y)

$P(A = \text{Gender}, C = \text{Diabetes}, \text{and } Y = \text{Hyperlipidemia})$

Hyperlipidemia	Gender	Diabetes	P
YES	Male	YES	0.1494073028184638
YES	Male	NO	0.08189405838453621
YES	Female	YES	0.11225172138199818
YES	Female	NO	0.0615281104150018
NO	Male	YES	0.11316005391439124
NO	Male	NO	0.1807298367436088
NO	Female	YES	0.11590888113498378
NO	Female	NO	0.18512003520701634

Example 2 (**Sufficiency** holds but not **Separation**)

Let A = Gender, C = Hyperlipidemia, and Y = Region

**Sufficiency** Argument: **Gender** is independent of **Region** given **Hyperlipidemia**

$P(\text{Gender} \mid \text{Region} = \text{YES}, \text{Hyperlipidemia} = \text{YES}) = [0.571, 0.429]$  (Male, Female)

$P(\text{Gender} \mid \text{Hyperlipidemia} = \text{YES}) = [0.571, 0.429]$  (Male, Female)

Therefore, removing Y (Region) from evidence has no effect on the probability and hence independence is shown.

**Separation** Argument (Proof by contradiction): **Gender** is independent of **Hyperlipidemia** given **Region**

$P(\text{Gender} \mid \text{Region} = \text{YES}, \text{Hyperlipidemia} = \text{YES}) = [0.571, 0.429]$  (Male, Female)

$P(\text{Gender} \mid \text{Region} = \text{YES}) = [0.524, 0.475]$  (Male, Female)

Hence, there is **no separation** since Gender (A) is not independent of Hyperlipidemia (C)

$P(A = \text{Gender}, C = \text{Hyperlipidemia}, \text{and } Y = \text{Region})$

Hyperlipidemia	Gender	Region	P
YES	Male	COUNTRYSIDE	0.11490991692204391
YES	Male	CITY	0.11639144428095609
YES	Female	COUNTRYSIDE	0.08633337015684206
YES	Female	CITY	0.08744646164015787
NO	Male	COUNTRYSIDE	0.13746971262840496
NO	Male	CITY	0.15642017802959507
NO	Female	COUNTRYSIDE	0.14080905787443912
NO	Female	CITY	0.16021985846756095