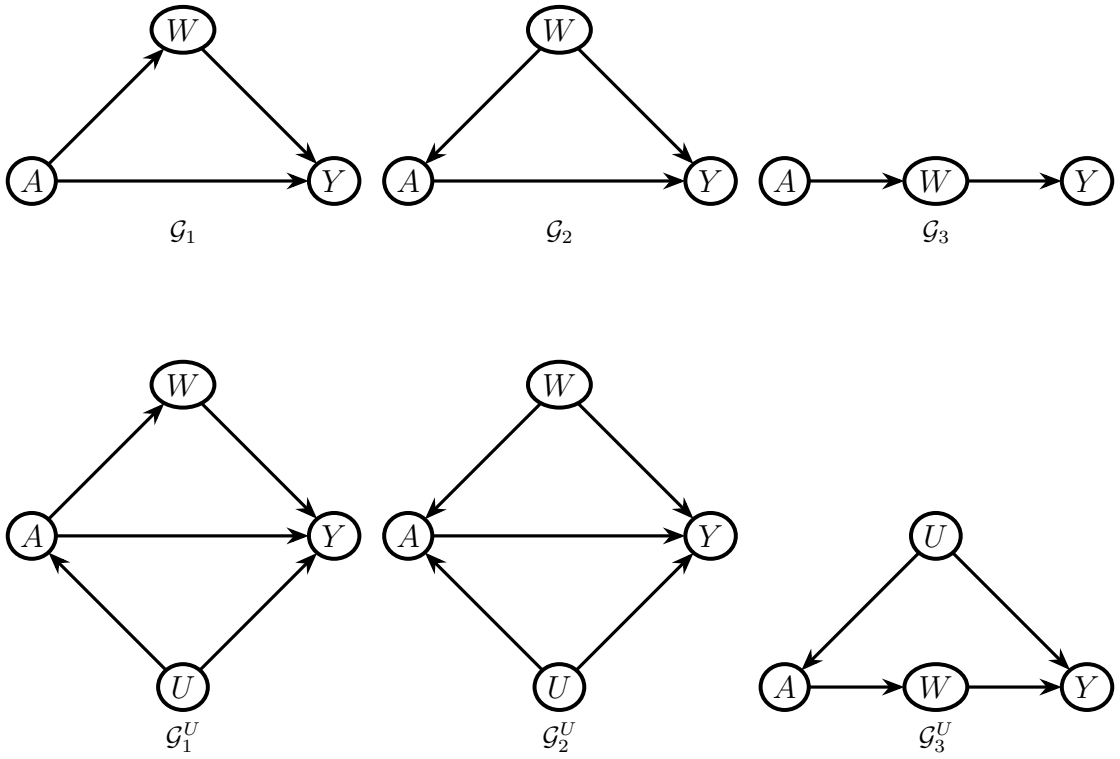
RANDOMIZATION AND CAUSATION (MATH-336)

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Exam-style questions

- This document provides examples of exam-style questions.

Question 1.
Consider six Non-Parametric Structural Equation Models with Independent Errors (NPSEM-IEs) associated with the six directed acyclic graphs (DAGs) $\mathcal{G}_1, \mathcal{G}_2$, $\mathcal{G}_3$, $\mathcal{G}_1^U, \mathcal{G}_2^U$ and $\mathcal{G}_3^U$:



(1) Draw the SWIGs[1] corresponding to an intervention that sets A to $a \in \{0, 1\}$ for:
   (i) $\mathcal{G}_1, \mathcal{G}_2$, and $\mathcal{G}_3$;
   (ii) $\mathcal{G}_1^U, \mathcal{G}_2^U$, and $\mathcal{G}_3^U$.

(2) Using the SWIGs you drew in the previous point, state whether you can read off $Y^a \perp\!\!\!\perp A$ and $Y^a \perp\!\!\!\perp A \mid W$ for every $a \in \{0, 1\}$ from:
   (i) $\mathcal{G}_1, \mathcal{G}_2$, and $\mathcal{G}_3$;
   (ii) $\mathcal{G}_1^U$, $\mathcal{G}_2^U$, and $\mathcal{G}_3^U$.

(3) Suppose:
   (i) *consistency I*

$$A = a \Longrightarrow Y^a = Y, \text{ for all } a \in \{0, 1\},$$

   (ii) *consistency II*

$$W = w \Longrightarrow Y^w = Y, \text{ for all } w \in \mathcal{W},$$
$$A = a \Longrightarrow W^a = W, \text{ for all } a \in \{0, 1\},$$

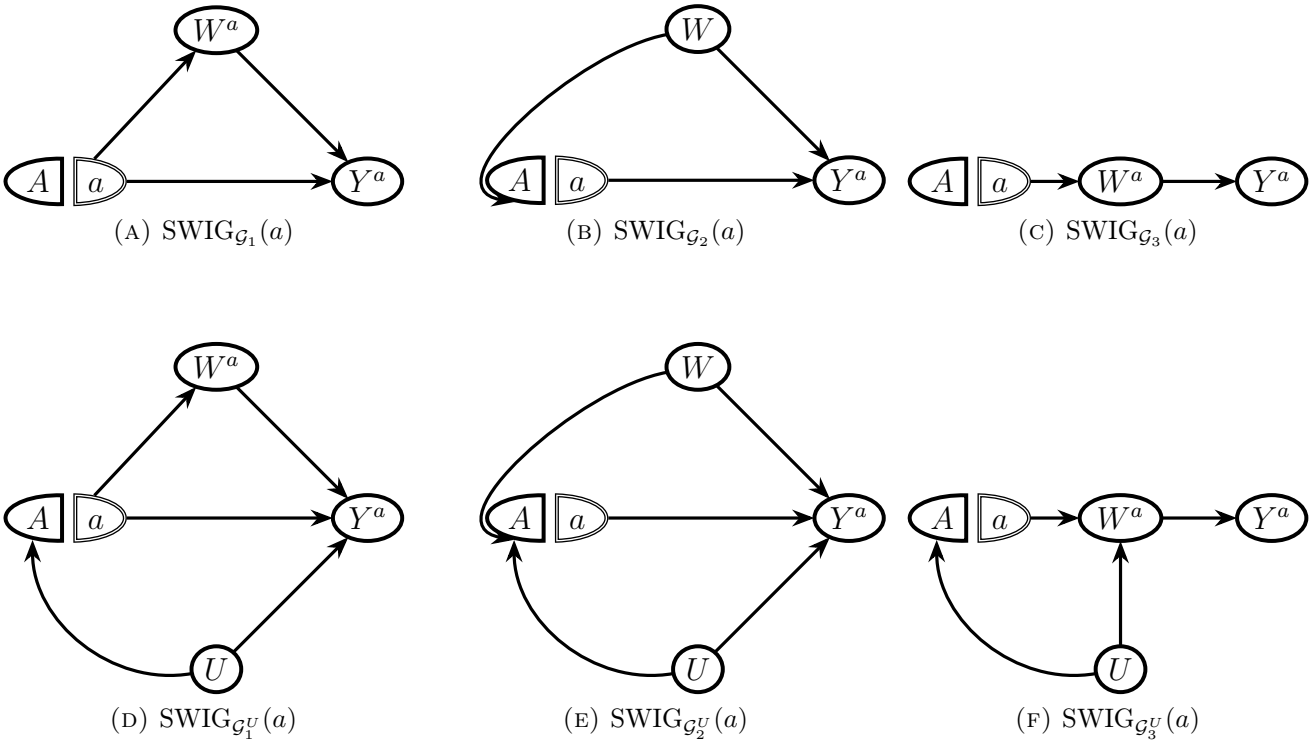
[1]Strictly speaking this is a Single World Intervention Template (SWIT).

(iii) *positivity*

$$P(W = w, A = a) > 0, \text{ for all } w \in \mathcal{W}, a \in \{0, 1\}.$$

Show that $P(Y^{a=1} = y)$ is identifiable for the causal model associated with $\mathcal{G}_3^U$ and derive its identification formula. *Hint: use an exclusion restriction.*

Solution:



(A) $\text{SWIG}_{\mathcal{G}_1}(a)$ (B) $\text{SWIG}_{\mathcal{G}_2}(a)$ (C) $\text{SWIG}_{\mathcal{G}_3}(a)$

(D) $\text{SWIG}_{\mathcal{G}_1^U}(a)$ (E) $\text{SWIG}_{\mathcal{G}_2^U}(a)$ (F) $\text{SWIG}_{\mathcal{G}_3^U}(a)$

(1)

(2) (i)
- $\text{SWIG}_{\mathcal{G}_1}$(a): i) yes; ii) the independence is not implied by the SWIG;
- $\text{SWIG}_{\mathcal{G}_2}$(a): i) no; ii) yes;
- $\text{SWIG}_{\mathcal{G}_3}$(a): i) yes; ii) the independence is not implied by the SWIG;

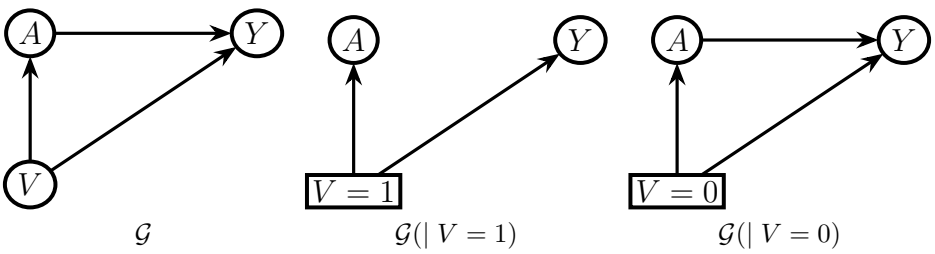
(ii) for all the remaining models none of the conditional independencies hold, $(Y^a \not\perp\!\!\!\perp A \mid L)_{\text{SWIG}_{G(a)}}$ and $(Y^a \not\perp\!\!\!\perp A)_{\text{SWIG}_{G(a)}}$ for every $\mathcal{G} \in \{\mathcal{G}_1^U, \mathcal{G}_2^U, \mathcal{G}_3^U\}$.

(3) See exercise sheet 9, exo 1(b).

Question 2 (Negative populations for the detection of unmeasured confounding and direct effects)**.**
Suppose that we have data from an observational study aimed at assessing the causal effect of penicillin ($A = 1$), an antibiotic that is given to treat a bacterial disease, versus no treatment ($A = 0$) on severe symptoms given by the indicator Y (the lower its value, the better).

(1) The principal investigator (PI) is worried that there is unmeasured confounding, that is, $A \not\perp\!\!\!\perp Y^a$ for some $a \in \{0, 1\}$. Can the PI use the data from the observational study, that is $P(A, Y)$, to test, or check, if $A \not\perp\!\!\!\perp Y^a$? Justify your answer (1-2 sentences).

A student of the PI points out that penicillin should not have an effect on severe symptoms in individuals with viral infections, because penicillin is only supposed to work against bacteria. The student refers to the individuals with viral infections as a *negative control population* and wonders if we can use them to rule out the presence of unmeasured confounding. To make progress, she first considers a setting where data are generated by an NPSEM-IE associated with a DAG $\mathcal{G}$ that includes the variables A, Y and an indicator V, taking value $V = 1$ if an individual has a viral infection and $V = 0$ otherwise. Among those in the negative control group ($V = 1$), she further specifies an NPSEM-IE associated with the DAG $\mathcal{G}(| \, V = 1)$. Similarly, she specifies an NPSEM-IE associated with the DAG $\mathcal{G}(| \, V = 0)$ for the population with $V = 0$. The rectangular boxes illustrate that $\mathcal{G}(| \, V = 0)$ and $\mathcal{G}(| \, V = 1)$ describe populations where V only takes one value, i.e., is constant.



$\mathcal{G}$ $\quad$ $\mathcal{G}(| \, V = 1)$ $\quad$ $\mathcal{G}(| \, V = 0)$

(2) Under the causal model $\mathcal{G}(| \, V = 1)$ the following condition holds

$$V = 1 \Longrightarrow Y^{a=1} = Y^{a=0},$$

that is, $Y^{a=1} = Y^{a=0}$ when $V = 1$.
Show that $E\left[Y^a \mid V = 1\right] = E\left[Y \mid V = 1\right]$ for every $a \in \{0, 1\}$. In words, penicillin has no effect on severe symptoms in individuals with viral infection, $V = 1$.

(3) Based on the causal models associated with $\mathcal{G}$, $\mathcal{G}(| \, V = 0)$, and $\mathcal{G}(| \, V = 1)$ can you conclude that $P(A, Y, V)$ is not faithful to $\mathcal{G}$? Justify your answer (1-2 sentences).

In the next questions, suppose that faithfulness holds in the negative control population, $(P(A, Y \mid V = 1), \mathcal{G}(| \, V = 1))$, and in the remaining population, $(P(A, Y \mid V = 0), \mathcal{G}(| \, V = 0))$.

(4) Argue that $Y \perp\!\!\!\perp A \mid V = 1 \Longrightarrow Y^a \perp\!\!\!\perp A \mid V = 1$, for every $a \in \{0, 1\}$.
Hint: use the rules of d-separation in $\mathcal{G}(| \, V = 1)$ *and note that* $V = 1$ *in* $\mathcal{G}(| \, V = 1)$ *is a constant.*

(5) Suppose further that positivity holds in both populations:

$$0 < P(A = 1 \mid V = v) < 1, \text{ for all } v \in \{0, 1\}.$$

Argue that

$$E[Y^{a=1} - Y^{a=0}] = P(V = 0)\left(E[Y \mid A = 1, V = 0] - E[Y \mid A = 0, V = 0]\right).$$

Hint: use the result of the previous point.

(6) Under the same assumptions of Question 5, show that

$$A \perp\!\!\!\perp V \Longrightarrow E[Y^{a=1} - Y^{a=0}] = E[Y \mid A = 1] - E[Y \mid A = 0].$$

(7) Is it possible that $A \perp\!\!\!\perp V$ holds and nevertheless the NPSEM-IE associated with $\mathcal{G}$ is the true causal model?

(8) Suppose now that the student, again, becomes uncertain about whether the NPSEM-IE associated with $\mathcal{G}$ is the correct causal model. She subsequently analyzes the data and finds evidence that $E[Y \mid A = 1, V = 1] \neq E[Y \mid A = 0, V = 1]$. If her analysis is correct, can the NPSEM-IE associated with $\mathcal{G}$ be the true causal model?

Solutions:

(1) No, the dependence between A and Y maybe due to either a direct effect or to a common cause.

(2) By consistency, $Y = AY^{a=1} + (1-A)Y^{a=0}$. By assumption, $Y^{a=1} = Y^{a=0} \equiv Y'$ when $V = 1$. Hence, $Y = Y'$ when $V = 1$. So we conclude.

(3) No. It is easy to show that $Y \perp\!\!\!\perp A | V = 1$ and $Y \not\perp\!\!\!\perp A | V = 0$. Hence $Y \not\perp\!\!\!\perp A | V$ which is consistent with the DAG. So one cannot state that faithfulness does not hold.

(4)

$$\begin{aligned}
&Y \perp\!\!\!\perp A \mid V = 1 \\
\Leftrightarrow &(Y \perp\!\!\!\perp A)_{\mathcal{G}(|V=1)} \\
\Longrightarrow &\text{no paths from } A \text{ to } Y \text{ in } \mathcal{G}(| \ V = 1) \text{ besides } A \leftarrow V \rightarrow Y \\
\Longrightarrow &Y^a \perp\!\!\!\perp A \mid V = 1, \text{ for every } a \in \{0, 1\}
\end{aligned}$$

Arguments: faithfulness, definition of direct effect. Alternatively, one can simply observe that $Y = Y^a$ for every a when $V = 1$ and conclude.

(5)

$$\begin{aligned}
E[Y^{a=1} - Y^{a=0}] &= E[E[Y^{a=1} - Y^{a=0} \mid V]] \\
&= E[E[Y^{a=1} \mid V, A = 1] - E[Y^{a=0} \mid V, A = 0]] \\
&= P(V = 0)(E[Y^{a=1} \mid V = 0, A = 1] - E[Y^{a=0} \mid V = 0, A = 0]) \\
&= P(V = 0)(E[Y \mid V = 0, A = 1] - E[Y \mid V = 0, A = 0])
\end{aligned}$$

where we used LOTP, $Y^a \perp\!\!\!\perp A \mid V = v$ for every v by faithfulness, positivity of the negative control population, and no direct effect in the negative control population, and consistency.

(6)

$$\begin{aligned}
E[Y^{a=1} - Y^{a=0}] &= E[E[Y^{a=1} - Y^{a=0} \mid V]] \\
&= E[E[Y^{a=1} \mid V, A = 1] - E[Y^{a=0} \mid V, A = 0]] \\
&= P(V = 0)(E[Y^{a=1} \mid V = 0, A = 1] - E[Y^{a=0} \mid V = 0, A = 0]) \\
&+ P(V = 1)(E[Y^{a=1} \mid V = 1, A = 1] - E[Y^{a=0} \mid V = 1, A = 0]) \\
&= P(V = 0 \mid A = 1)E[Y \mid V = 0, A = 1] - P(V = 0 \mid A = 0)E[Y \mid V = 0, A = 0]) \\
&+ P(V = 1 \mid A = 1)E[Y \mid V = 1, A = 1] - P(V = 1 \mid A = 0)E[Y \mid V = 1, A = 0] \\
&= P(V = 0 \mid A = 1)E[Y \mid V = 0, A = 1] + P(V = 1 \mid A = 1)E[Y \mid V = 1, A = 1] \\
&- (P(V = 0 \mid A = 0)E[Y \mid V = 0, A = 0] + P(V = 1 \mid A = 0)E[Y \mid V = 1, A = 0]) \\
&= E[Y \mid A = 1] - E[Y \mid A = 0]
\end{aligned}$$

where we used LOTP, $Y^a \perp\!\!\!\perp A \mid V = v$ for every v by faithfulness, positivity of the negative control population, $A \perp\!\!\!\perp V$, rearranged terms, tower rule.

(7) yes, if $P(A, V, Y)$ is unfaithful to $\mathcal{G}$. Counterexample, $p = P(A = 1 \mid V = 1) = P(A = 1 \mid V = 0)$ for $p \in (0, 1)$. Notice that this does not contradict the fact that $P(A, Y \mid V = 1)$ and $P(A, Y \mid V = 0)$ are faithful to $\mathcal{G}(| \ V = 1)$ and $\mathcal{G}(| \ V = 0)$

respectively. From $\mathcal{G}(|V=0)$ and $\mathcal{G}(|V=1)$ we cannot use faithfulness because it also pertains to the A and Y variables.

(8) If $\mathcal{G}$ is a valid NPSEM-IE and $\mathcal{G}(|\ V=1)$ is correct, then

$$\begin{aligned} E[Y \mid A=1, V=1] &= E[Y^{a=1} \mid A=1, V=1] \\ &= E[Y^{a=1} \mid V=1] \\ &= E[Y^{a=0} \mid V=1] \\ &= E[Y^{a=0} \mid A=0, V=1] \\ &= E[Y \mid A=0, V=1], \end{aligned}$$

where we used consistency, positivity, $Y^a \perp\!\!\!\perp A|V=1$ and $Y^{a=1} = Y^{a=0}$ (no causal effect) if $V=1$. However, we assume that $E[Y \mid A=1, V=1] \neq E[Y \mid A=0, V=1]$, so there is a contradiction. Hence, either

(a) $\mathcal{G}(|V=1)$ is false, which contradicts her assumptions,

(b) $\mathcal{G}(|V=1)$ is correct, but not an NPSEM-IE : the error term of $Y|A=1, V=1$ and of $Y|A=0, V=1$ must be correlated for $E[Y \mid A=1, V=1] \neq E[Y \mid A=0, V=1]$ and hence dependent. However, if $\mathcal{G}$ is a NPSEM-IE, then $\mathcal{G}(|V=1)$ would also be an NPSEM-IE. Thus, $\mathcal{G}$ cannot be a NPSEM-IE. Note, however that $\mathcal{G}$ can still be a valid causal model, but not an NPSEM-IE as she is testing for.

Question 3 (Time-to-event outcomes)**.**
In this exercise, we consider treatments of cancer. Unlike traditional chemotherapy, which attacks all cells, targeted therapy focuses on specific molecular targets that are essential for cancer cell growth.

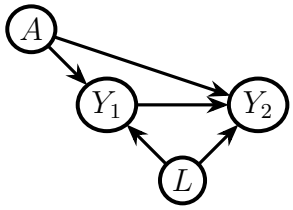
In particular, around 15-20% of breast cancer patients have tumors that overexpress the protein *HER2*. Without appropriate treatment, these tumors are associated with bad prognosis. Trastuzumab is a therapy that targets *HER2*, inhibiting the growth of tumor cells that express this protein. Trastuzumab is expected to be more effective in patients whose breast tumors overexpress *HER2* than in those whose tumors do not or only slightly express *HER2*. It is now used as a standard of care for *HER2*-positive tumors, and its use has significantly improved the prognosis of patients diagnosed with *HER2*-positive tumors.

Now imagine that you are conducting a randomized controlled trial where you enroll patients diagnosed with breast cancer. At enrollment, you randomize the patients into two arms: take trastuzumab (treatment arm) *versus* do not take trastuzumab, irrespectively of *HER2* expression status. You then follow your patients and report deaths (study outcome) every month. For simplicity, we will assume that there is no loss-to-follow-up, and that the follow-up period lasted only 2 months. We also assume consistency and positivity.

We have the following random variables:

- A: treatment assignment (1 if treatment arm, 0 if control arm).
- L: *HER2* expression status, binarized into 1 if the patient's tumor overexpress *HER2* (in this case we expect trastuzumab to be efficient), 0 otherwise. We denote $p_l := Pr(L = 1)$, with $0 < p_l < 1$.
- Y_k: the outcome at month $k \in \{1, 2\}$, with $Y_k = 1$ if the patient died before the end of month k. Notably, $Y_2 = 1$ if $Y_1 = 1$.

We assume the following DAG $\mathcal{G}$:



$\mathcal{G}$

In oncology (cancer medicine), it is frequent to report the hazard ratio (HR) as a measure of treatment effect in randomized trials. In discrete time, the hazard ratio at month 2 is defined as

$$HR = \frac{Pr(Y_2^1 = 1 | Y_1^1 = 0)}{Pr(Y_2^0 = 1 | Y_1^0 = 0)}.$$

Questions:

(a) **On the hazard of hazard ratios**

(i) Draw the SWIG $\mathcal{G}(a)$ corresponding to the intervention that assigns treatment a.

(ii) Derive an identification formula for the risk ratio RR at time point 1[2], defined as

$$RR = \frac{Pr(Y_1^1 = 1)}{Pr(Y_1^0 = 1)}.$$

State the assumptions you used.

(iii) Derive an identification formula for the hazard ratio HR. State the assumptions you used.

(iv) Table 1 summarises the results you obtained from your randomized controlled trial.

Compute, i.e., estimate,

(A) the risk ratio at time point 1 in the *HER2*-positive population $\frac{Pr(Y_1^1=1|L=1)}{Pr(Y_1^0=1|L=1)}$,

(B) the risk ratio at time point 1 in the *HER2*-negative population $\frac{Pr(Y_1^1=1|L=0)}{Pr(Y_1^0=1|L=0)}$,

(C) the risk ratio at time point 1 in the combined population $\frac{Pr(Y_1^1=1)}{Pr(Y_1^0=1)}$,

(D) the hazard ratio (at time point 2) in the *HER2*-positive population $\frac{Pr(Y_2^1=1|Y_1^1=0,L=1)}{Pr(Y_2^0=1|Y_1^0=0,L=1)}$,

(E) the hazard ratio (at time point 2) in the *HER2*-negative population $\frac{Pr(Y_2^1=1|Y_1^1=0,L=0)}{Pr(Y_2^0=1|Y_1^0=0,L=0)}$,

(F) the hazard ratio (at time point 2) in the combined population $\frac{Pr(Y_2^1=1|Y_1^1=0)}{Pr(Y_2^0=1|Y_1^0=0)}$.

(v) Compare the results you obtained for the first time point, which are risk ratios, and for the second time point, which are hazard ratios. What do you observe? Why could hazard ratios be misleading? Give your answer in 3-5 sentences.

Hazard ratios are controversial in causal inference for at least two reasons: they have a built-in selection bias, and they are non-collapsible. In the rest of the exercise, we will discuss these two caveats.

(b) **Built-in selection bias**

For this question, assume that $0 < Pr(Y_1 = 1|L = l, A = a) < 1$ for all a, l, and define $p \in (0, 1)$ and $\gamma, \alpha, \beta \in \mathbb{R}^{+*}$ as

$$\begin{aligned} p &= Pr(Y_1 = 1|L = 0, A = 0), \\ \gamma &= \frac{Pr(Y_1 = 1|L = 1, A = 0)}{p}, \\ \alpha &= \frac{Pr(Y_1 = 1|L = 1, A = 1)}{Pr(Y_1 = 1|L = 1, A = 0)}, \\ \beta &= \frac{Pr(Y_1 = 1|L = 0, A = 1)}{p}. \end{aligned}$$

(i) Give the probability of overexpressing *HER2* in the treatment arm, $Pr(L = 1|A = 1)$, and control arm, $Pr(L = 1|A = 0)$, as a function of $p_l, \alpha, \beta, \gamma$ and p.

[2]Generally, we set $Y_0^1 \equiv 0$ for convenience, so that the risk ratio equals the hazard ratio for the first time point

(ii) Give the probability of overexpressing *HER2* in survivors of control and treatment arm at the second time point, that is $Pr(L = 1|Y_1 = 0, A = 1)$ and $Pr(L = 1|Y_1 = 0, A = 0)$, as functions of $p_l, \alpha, \beta, \gamma$ and p.

(iii) Show that $Pr(L = 1|Y_1 = 0, A = 1) = Pr(L = 1|Y_1 = 0, A = 0)$ if and only if $\alpha = \frac{\gamma \cdot (1-p\beta)+\beta-1}{\gamma \cdot (1-p)}$.

(iv) Show that $Pr(L = 1|Y_1 = 0, A = 1) = Pr(L = 1|Y_1 = 0, A = 0)$ if the treatment had no effect in any of the subgroups at time point 1.

(v) Show that $Pr(L = 1|Y_1 = 0, A = 1) = Pr(L = 1|Y_1 = 0, A = 0)$ if the following two conditions hold: (i) the effect of the treatment (trastuzumab) on the ratio scale is identical in the two subgroups at time point 1 and (ii) patients who overexpress *HER2* and those who do not overexpress *HER2* have the same risk of death at time point 1 without treatment (trastuzumab).

(vi) In general, the equality $Pr(L = 1|Y_1 = 0, A = 1) = Pr(L = 1|Y_1 = 0, A = 0)$ does not hold. What can you say about the causal interpretation of the hazard ratio in the combined population, $HR := \frac{Pr(Y_2^1=1|Y_1^1=0)}{Pr(Y_2^0=1|Y_1^0=0)}$, in the general case where $Pr(L = 1|Y_1 = 0, A = 1) \neq Pr(L = 1|Y_1 = 0, A = 0)$?

(c) **Non-collapsibility** For this question, let us parametrize the hazard as follows:

$$\log(Pr(Y_k = 1|Y_{k-1} = 0, A = a)) = \mu + \psi a \qquad \text{(marginal hazard model)}$$

$$\log(Pr(Y_k = 1|Y_{k-1} = 0, A = a, L)) = \nu(L) + \kappa a \qquad \text{(conditional hazard model)}$$

for $k \in 1, 2$ and $\mu, \psi, \nu(L), \kappa \in \mathbb{R}$, where we set $Y_0 \equiv 0$. log denotes the natural logarithm (in base e).

(i) Prove that, with these parametrizations, $HR_{L=1} = HR_{L=0} = e^{\kappa}$, where

$$HR_{L=l} := \frac{Pr(Y_2 = 1|Y_1 = 0, A = 1, L = l)}{Pr(Y_2 = 1|Y_1 = 0, A = 0, L = l)}$$

for $l \in \{0, 1\}$. If the hazard ratio were a collapsible estimand, we would then expect $HR = HR_{L=1} = HR_{L=0}$. We will now show that this equality does not hold in general.

(ii) Prove that

$$Pr(Y_2 = 1|A = 1, L) = f_{\kappa}(Pr(Y_2 = 1|A = 0, L))$$

where $f_{\kappa}(x) = 1 - \left(1 - e^{\kappa} + e^{\kappa} \cdot (1 - x)^{1/2}\right)^2$ defined for $x \in (0, 1)$.

(iii) Derive the second derivative of f_{κ} with respect to x, and discuss its sign with respect to κ.

(iv) Show that $\psi = \log(1 - (1 - p_1)^{1/2}) - \log(1 - (1 - p_0)^{1/2})$, where $p_a = Pr(Y_2 = 1|A = a)$ for $a \in \{0, 1\}$.

(v) Show that $\kappa = \log(1 - (1 - f_{\kappa}(p_0))^{1/2}) - \log(1 - (1 - p_0)^{1/2})$.

(vi) Under what conditions for κ does $HR < HR_{L=0}$, $HR = HR_{L=0}$, and $HR > HR_{L=0}$? Conclude that $HR \neq HR_{L=0}$ in general.

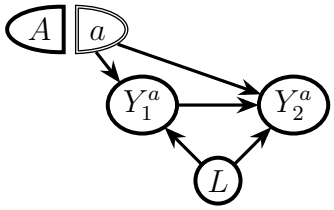
A	L	Y_1	Y_2	n
0	0	0	0	2560
0	0	0	1	640
0	0	1	1	800
0	1	0	0	40
0	1	0	1	160
0	1	1	1	800
1	0	0	0	2560
1	0	0	1	640
1	0	1	1	800
1	1	0	0	80
1	1	0	1	200
1	1	1	1	720

TABLE 1. Number of patients (n) for each group in the randomized controlled trial (Exercise 2).

Solution:

(a) **On the hazard of hazard ratios**

(i) The desired SWIG is shown below:



$\mathcal{G}(a)$

(ii) From the SWIG, we can read off the independency : $Y_1^a \perp\!\!\!\perp A$. From that, and using consistency, we have that:

$$\begin{aligned} Pr(Y_1^a = 1) &= Pr(Y_1^a = 1 | A = a) \\ &= Pr(Y_1 = 1 | A = a), \end{aligned}$$

Finally,

$$RR := \frac{Pr(Y_1^1 = 1)}{Pr(Y_1^0 = 1)} = \frac{Pr(Y_1 = 1 | A = 1)}{Pr(Y_1 = 1 | A = 0)}$$

(iii) From the SWIG, we can read off the independency : $Y_2^a \perp\!\!\!\perp A | Y_1^a$. From that, and using consistency, we have that:

$$\begin{aligned} Pr(Y_2^a = 1 | Y_1^a = 0) &= Pr(Y_2^a = 1 | A = a, Y_1^a = 0) \\ &= Pr(Y_2 = 1 | A = a, Y_1 = 0), \end{aligned}$$

so that

$$HR = \frac{Pr(Y_2 = 1|A = 1, Y_1 = 0)}{Pr(Y_2 = 1|A = 0, Y_1 = 0)}$$

which concludes the identification proof.

(iv) (A) Since $Y_1^a \perp\!\!\!\perp A|L$ for all $a \in \{0, 1\}$, $\frac{Pr(Y_1^1=1|L=1)}{Pr(Y_1^0=1|L=1)} = \frac{Pr(Y_1=1|A=1,L=1)}{Pr(Y_1=1|A=0,L=1)} = \frac{720}{1000} \cdot (\frac{800}{1000})^{-1} = 0.9$

(B) $\frac{Pr(Y_1^1=1|L=0)}{Pr(Y_1^0=1|L=0)} = \frac{Pr(Y_1=1|A=1,L=0)}{Pr(Y_1=1|A=0,L=0)} = \frac{800}{4000} \cdot (\frac{800}{4000})^{-1} = 1$

(C) $\frac{Pr(Y_1^1=1)}{Pr(Y_1^0=1)} = \frac{Pr(Y_1=1|A=1)}{Pr(Y_1=1|A=0)} = \frac{800+720}{5000} \cdot (\frac{800+800}{5000})^{-1} = 0.95$

(D) $\frac{Pr(Y_2^1=1|Y_1^1=0,L=1)}{Pr(Y_2^0=1|Y_1^0=0,L=1)} = \frac{Pr(Y_2=1|Y_1=0,A=1,L=1)}{Pr(Y_2=1|Y_1=0,A=0,L=1)} = \frac{200}{200+80} \cdot (\frac{160}{40+160})^{-1} \approx 0.89$

(E) $\frac{Pr(Y_2^1=1|Y_1^1=0,L=0)}{Pr(Y_2^0=1|Y_1^0=0,L=0)} = \frac{Pr(Y_2=1|Y_1=0,A=1,L=0)}{Pr(Y_2=1|Y_1=0,A=0,L=0)} = \frac{640}{2560+640} \cdot (\frac{640}{2560+640})^{-1} = 1$

(F) $\frac{Pr(Y_2^1=1|Y_1^1=0)}{Pr(Y_2^0=1|Y_1^0=0)} = \frac{Pr(Y_2=1|Y_1=0,A=1)}{Pr(Y_2=1|Y_1=0,A=0)} = \frac{640+200}{2560+640+200+80} \cdot (\frac{640+160}{2560+640+40+160})^{-1} = \frac{119}{116} \approx 1.02$

(v) First, note that at both time points, treatment (trastuzumab) does not affect relapse in survivors among *HER2*-negative group (risk ratio / hazard ratio is 1 in this subgroup) and prevents relapse in survivors among *HER2*-positive patients (risk ratio / hazard ratio is below 1 in this subgroup). At the first time point, the marginal risk ratio is the average of the subgroup risk ratios, and this is the behavior we would have expected.
However, at the second time point, the marginal hazard ratio in the combined population is above 1, even though it is 1 in one subgroup (survivors, *HER2*-negative) and below 1 in the other subgroup (survivors, *HER2*-positive). This suggests that HRs cannot be provided with a causal interpretation.

(b) **Built-in selection bias**

(i) By randomization $Pr(L = 1|A = 1) = Pr(L = 1|A = 0) = Pr(L = 1) = p_l$. Indeed, by randomization, we expect the same proportion of *HER2*-positive and *HER2*-negative in the treatment and control arms.

(ii) We have :

$$\begin{aligned}
Pr(L = 1|Y_1 = 0, A = a) &= \frac{Pr(Y_1 = 0|A = a, L = 1) \cdot Pr(L = 1, A = a)}{Pr(Y_1 = 0, A = a)} \\
&= \frac{Pr(Y_1 = 0|A = a, L = 1) \cdot Pr(L = 1) \cdot Pr(A = a)}{Pr(Y_1 = 0|A = a)Pr(A = a)} \qquad (A \perp\!\!\!\perp L) \\
&= \frac{p_l \cdot Pr(Y_1 = 0|A = a, L = 1)}{Pr(Y_1 = 0|A = a)} \\
&= \frac{p_l \cdot Pr(Y_1 = 0|L = 1, A = a)}{p_l \cdot Pr(Y_1 = 0|A = a, L = 1) + (1 - p_l) \cdot Pr(Y_1 = 0|A = a, L = 0)}
\end{aligned}$$

For $a = 0$ this can be written as:

$$Pr(L = 1|Y_1 = 0, A = 0) = \frac{p_l \cdot (1 - \gamma p)}{p_l \cdot (1 - \gamma p) + (1 - p_l) \cdot (1 - p)}$$

For $a = 1$ this can be written as:

$$Pr(L = 1|Y_1 = 0, A = 1) = \frac{p_l \cdot (1 - \alpha\gamma p)}{p_l \cdot (1 - \alpha\gamma p) + (1 - p_l) \cdot (1 - \beta p)}$$

(iii) We have :

$$\begin{aligned}
&Pr(L = 1|Y_1 = 0, A = 0) = Pr(L = 1|Y_1 = 0, A = 0) \\
&\iff \frac{p_l \cdot (1 - \gamma p)}{p_l \cdot (1 - \gamma p) + (1 - p_l) \cdot (1 - p)} = \frac{p_l \cdot (1 - \alpha\gamma p)}{p_l \cdot (1 - \alpha\gamma p) + (1 - p_l) \cdot (1 - \beta p)} \\
&\iff (1 - \gamma p) \cdot (1 + p \cdot (p_l\beta - \beta - p_l\alpha\gamma)) = (1 - \alpha\gamma p) \cdot (1 + p \cdot (p_l - 1 - \gamma p_l)) \\
&\qquad\qquad \text{(We used that } p_l \neq 0) \\
&\iff \gamma \cdot p \cdot (1 - p_l) \cdot (\alpha - \beta) = (1 - p_l) \cdot (1 + \alpha\gamma - \beta\gamma) \qquad \text{(We used that } p \neq 0) \\
&\iff \alpha = \frac{\gamma \cdot (1 - p\beta) + \beta - 1}{\gamma(1 - p)} \qquad (p_l, p \neq 1)
\end{aligned}$$

This concludes the proof.

(iv) This corresponds to the case where $\alpha = \beta = 1$. The right hand side of the equality above becomes :

$$\frac{\gamma \cdot (1 - p)}{\gamma(1 - p)} = 1 = \alpha$$

so that the equality holds trivially.

(v) This corresponds to the case where $\gamma = 1$ and $\alpha = \beta$.The right hand side of the equality above becomes :

$$\frac{1 - p\beta + \beta - 1}{(1 - p)} = \beta = \alpha$$

so that the equality holds trivially as well.

(vi) If the distribution of L differs by treatment arm at the second time point, then the two populations are biased and the benefit of randomization is lost. In fact, by conditioning on $Y_1^1 = 0$ in the numerator and $Y_0^1 = 0$ in the denominator, HR compares two populations that are no longer comparable at the second time point.

(c) **Non-collapsibility**

(i) Using the conditional hazard model, we have

$$Pr(Y_2 = 1|Y_1 = 0, A = 1, L = l) = e^{\nu(l)} \cdot e^{\kappa}$$

and

$$Pr(Y_2 = 1|Y_1 = 0, A = 0, L = l) = e^{\nu(l)}$$

so that $HR_{L=l} = e^{\kappa}$, which does not depend on L.

(ii) We have :

$$\begin{aligned}
Pr(Y_2 = 1|A = a, L) &= 1 - Pr(Y_2 = 0|A = a, L) \\
&= 1 - Pr(Y_2 = 0|Y_1 = 0, A = a, L) \cdot Pr(Y_1 = 0|A = a, L)
\end{aligned}$$

$$= 1 - [1 - Pr(Y_2 = 1|Y_1 = 0, A = a, L)] \cdot [1 - Pr(Y_1 = 1|A = a, L)]$$
$$= 1 - \left[1 - e^{\nu(L)+\kappa \cdot a}\right]^2$$

so that

$$Pr(Y_2 = 1|A = 1, L) = 1 - \left[1 - e^{\nu(L)+\kappa}\right]^2$$
$$= 1 - \left[1 - e^{\nu(L)} e^{\kappa}\right]^2$$

and

$$Pr(Y_2 = 1|A = 0, L) = 1 - \left[1 - e^{\nu(L)}\right]^2$$

By evaluating the function f_κ on $x = Pr(Y_2 = 1|A = 0, L)$, we get :

$$f_\kappa(Pr(Y_2 = 1|A = 0, L)) = 1 - \left(1 - e^{\kappa} + e^{\kappa}\left(1 - e^{\nu(L)}\right)^{2 \cdot 1/2}\right)^2$$
$$= 1 - \left(1 - e^{\kappa} e^{\nu(L)}\right)^2$$
$$= Pr(Y_2 = 1|A = 1, L)$$

which concludes the proof.

(iii) First note that f_κ is differentiable twice on $(0, 1)$. let us derive the first derivative

$$f_k'(x) = -2 \cdot \left(1 - e^{\kappa} + e^{\kappa}(1 - x)^{1/2}\right) \cdot (-\frac{1}{2}) e^{\kappa}(1 - x)^{-\frac{1}{2}}$$
$$= e^{\kappa}\left[(1 - e^{\kappa})(1 - x)^{-\frac{1}{2}} + e^{\kappa}\right].$$

From which we get that

$$f_k''(x) = \frac{1}{2} e^{\kappa}(1 - e^{\kappa})(1 - x)^{-\frac{3}{2}}.$$

Now, we have that, for $x \in (0, 1)$, the sign of $f_k''(x)$ is the same as the sign of $(1 - e^{\kappa})$, that is strictly negative if $\kappa > 0$, zero if $\kappa = 0$, and strictly positive if $\kappa < 0$. From that, we conclude that the function f_κ is concave if $\kappa > 1$, linear if $\kappa = 1$ and convex if $\kappa < 1$.

(iv) Using the same arguments than in question (c)(ii), we obtain

$$p_a = Pr(Y_2 = 1|A = a) = 1 - (1 - Pr(Y_2 = 1|Y_1 = 0, A = a))\,(1 - Pr(Y_1 = 1|A = a))\,,$$

which, from the marginal hazard model, could be re-written as

$$p_a = Pr(Y_2 = 1|A = a) = 1 - \left(1 - e^{\mu+\psi a}\right)^2.$$

From that, we have that

$$\log\left((1 - (1 - p_a)^{\frac{1}{2}})\right) = \mu + \psi a$$

and

$$\log\left((1-(1-p_1)^{\frac{1}{2}})\right) - \log\left((1-(1-p_0)^{\frac{1}{2}})\right) = \mu + \psi \cdot 1 - \mu - \psi \cdot 0 = \psi$$

which concludes the proof.

(v) By definition of the function f_κ,

$$\forall x \in (0,1),\ 1-(1-f_\kappa(x))^{\frac{1}{2}} = e^{\kappa}\left(1-(1-x)^{\frac{1}{2}}\right),$$

so that

$$\log(1-(1-f_\kappa(x))^{\frac{1}{2}}) = \kappa + \log\left(1-(1-x)^{\frac{1}{2}}\right)$$

Using $x = p_0 \in (0,1)$ we get :

$$\kappa = \log(1-(1-f_\kappa(p_0))^{\frac{1}{2}}) - \log\left(1-(1-p_0)^{\frac{1}{2}}\right),$$

which concludes the proof.

(vi) With our parametrization, $HR = \psi$. Then, notice that

$$\begin{aligned} p_1 &= Pr(Y_2 = 1|A=1) && \text{(Definition of } p_1\text{)} \\ &= \mathbb{E}\left[Pr(Y_2=1|A=1,L)\right] && \text{(Law of total expectation)} \\ &= \mathbb{E}\left[f_\kappa(Pr(Y_2=1|A=0,L))\right] && \text{(From (b)(ii))} \end{aligned}$$

Using Jensen's inequality and question (b)(iii) we have that

$$p_1 \left\{ \begin{array}{ll} < & \text{if } \kappa > 0 \\ = & \text{if } \kappa = 0 \\ > & \text{if } \kappa < 0 \end{array} \right\} f_\kappa(\mathbb{E}(Pr(Y_2=1|A=0,L]) = f_\kappa(p_0)$$

Finally, as the function $g : x \to \log\left(1-(1-x)^{\frac{1}{2}}\right)$ is strictly non-decreasing for $x \in (0,1)$ (The first derivative of g is $g'(x) = \frac{(1-x)^{-3/2}}{1-\sqrt{1-x}} > 0$ for $x \in (0,1)$), we get that

$$HR = \psi = g(p_1) - g(p_0) \left\{ \begin{array}{ll} < & \text{if } \kappa > 0 \\ = & \text{if } \kappa = 0 \\ > & \text{if } \kappa < 0 \end{array} \right\} g(f_\kappa(p_0)) - g(p_0) = \kappa = HR_{L=l}$$

which means that, when there is a non-null subgroup effect (*i.e.* when $HR_{L=l} \neq 1$), $HR \neq HR_{L=l}$. This proves that hazard ratios are not collapsible in general. Of note, remark that the marginal HR is always closer to the null (1) than the subgroup hazard ratios.