# ELL880: Minor Test II

March 24, 2023

Maximum Marks: 15

**Instructions:**

- **Please clearly indicate the question number, and part number if applicable, at the start of each response.**
- **Please read all questions carefully.**
- **Please ensure that your responses are to-the-point and that you write only what is asked for on the answer script you submit.**
- **While the exam is open-notes, all your answers must be written entirely in your own words, without any copying from anywhere.**
- **Please try to be clear and careful with all mathematical notation, so that there is no ambiguity in the expressions/formulae you write down. Try to stick to the kind of notation used in class as far as possible.**

1. Recall the example of the relationship between age ($A$), happiness ($H$), and marital status ($M$) which was discussed in class. Here $A$ denotes age in years, $H$ denotes the happiness level as a real value in the interval $[-2, 2]$, and $M$ is binary ($1 =$ married, $0 =$ unmarried).

   Consider the following models, where $i$ is the index over individuals in the data set:

   **Model 1**:

   $$H_i \sim \mathcal{N}(\mu_i, \sigma_1)$$
   $$\mu_i = \alpha_{1M_i} + \beta_1 A_i$$

   **Model 2**:

   $$H_i \sim \mathcal{N}(\mu_i, \sigma_2)$$
   $$\mu_i = \alpha_2 + \beta_2 A_i$$

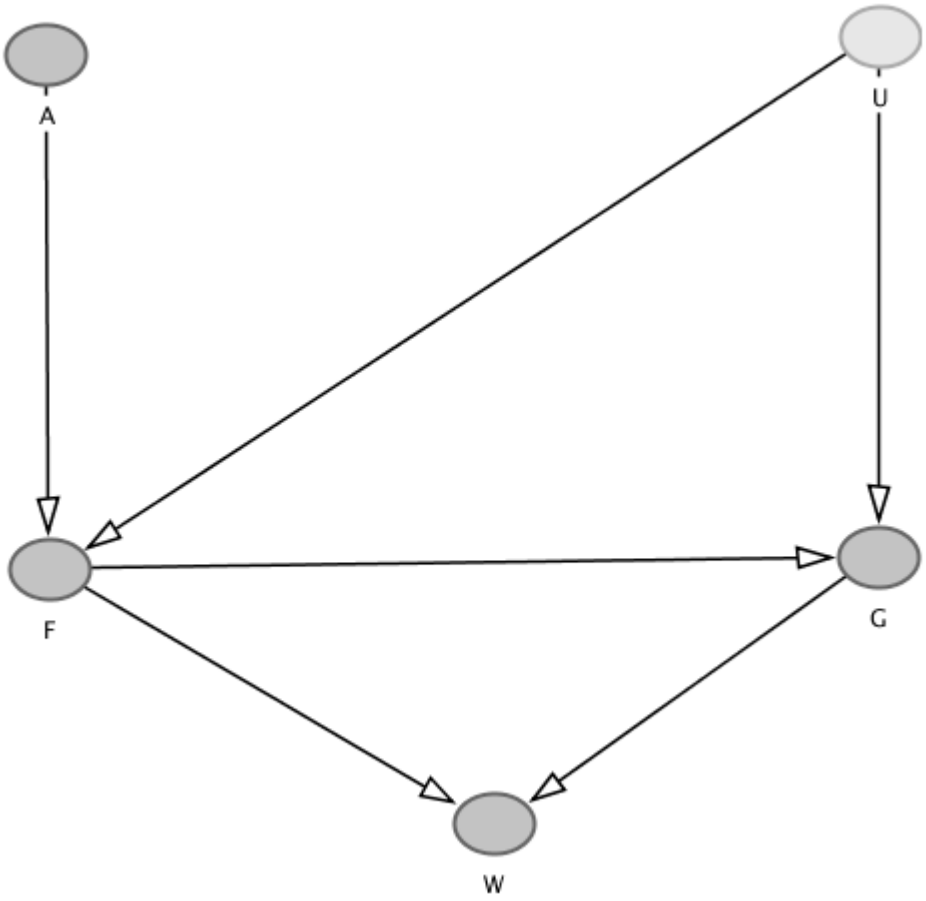
   Let suitable priors be specified over all the model parameters, i.e., $\sigma_1$, $\sigma_2$, $\alpha_{10}$, $\alpha_{11}$, $\alpha_2$, $\beta_1$, and $\beta_2$.

   (a) In general, which of these models would be expected to make better predictions of happiness, as assessed by a criterion like PSIS or WAIC? Explain your answer. **[2]**

   (b) Now we wish to consider the possible causal relationships between these 3 variables. From $A$, clearly there can only be a one-way causal effect on both $H$ and $M$. However, the directionality of any possible causal relationship between $H$ and $M$ is not obvious: your level of happiness could affect your chances of getting married, and your marital status could affect your happiness level. Since in modelling causal relationships as DAGs, we generally have to choose a specific directionality for each link, draw two separate DAGs for this set of variables, representing the two possible choices for the relationship between $H$ and $M$. **[1]**

(c) Now let's say I am interested in inferring the strength of the *direct* causal effect of $A$ on $H$. Which out of Model 1/2 will be better for this purpose, if the correct DAG is the first one you drew above? And which model will be better if the correct DAG is the second one? Explain both your answers clearly, making reference to the kind of confounding that might be involved in each case and why it leads you to your specified model choice. **[4]**

2. Consider the below DAG.




Here $U$ is unobserved and all the other variables are observed.

For each of the following causal inferences, mention whether or not it is possible, given a data set of the observed variables; if it is possible, list the variables required to be stratified by in order to infer that effect. Use the backdoor criterion in all cases involving *total* causal effects. Explain your working/reasoning clearly in all cases. **[8]**

(a) Total causal influence of $A$ on $F$.

(b) Total causal influence of $F$ on $W$.

(c) Total causal influence of $F$ on $W$, if $U$ is removed from the graph.

(d) Direct causal influence of $F$ on $W$.

(e) Direct causal influence of $F$ on $W$, if $U$ is removed from the graph.

(f) Total causal influence of $F$ on $G$.

(g) Total causal influence of $G$ on $W$.

(h) Direct causal influence of $G$ on $W$.