**EPI289 Spring 2024**

**G-estimation worksheet**

**This worksheet is optional and ungraded.**

**For questions that may have multiple correct answers, the question will ask you to select all that apply. For numerical answer questions, please keep all decimal places in the calculations and round the final answer to the number of decimal places requested. Provide the number only (no letters, symbols, or other text) and keep in mind the units, if applicable. For short answer questions, please type your answers in the space provided. You will need to use the NHEFS dataset to complete this worksheet.**

**Estimate the causal effect of smoking cessation *A* (1: yes, 0; no) on weight gain *Y* in the NHEFS data using g-estimation. Assume no structural violations of positivity, and that exchangeability between the treated and the untreated can be achieved conditional on sex (1: female, 0: male), race (1: nonwhite, 0; white), age (continuous variable; linear), smoke intensity (continuous variable; linear), asthma diagnosis in 1971 (1: ever, 0: never), and weak heart medication in 1971 (1: ever, 0: never). Do not include product terms in models. In this simplified example, we disregard potential selection bias due to loss to follow-up.**

1. **[Fill in multiple blanks, ungraded] Provide a point estimate and a conservative 95% confidence interval for the causal effect. Do not use the closed form solution for the point estimate. Round each numeric response to the nearest hundredth (i.e., 2 decimal places)**

**Point estimate (2 decimal places): 3.13**

**Lower bound of the 95% confidence interval (2 decimal places): 2.26**

**Upper bound of the 95% confidence interval (2 decimal places): 4.01**

1. **[Essay question, ungraded] Interpret the point estimate you provided in question 1. Please limit your response to one sentence.**

**Had everyone quit smoking, the mean weight gain would have been 3.13 kg higher compared to had no one quit smoking over the 10-year follow-up period, conditional on sex, race, age, smoke intensity, asthma diagnosis and weak heart medication.**

1. **[Multiple answers, ungraded] Which of the following models did you use to estimate the causal effect of interest? Select all that apply.**
   1. **logit Pr[A=1|sex, race, age, smokeintensity, asthma, weakheart, H(Ψ\*)] = β0 + β1sex + β2race + β3age + β4smokeintensity + β5asthma + β6weakheart + β7H(Ψ\*)**
   2. logit Pr[A=1|sex, race, age, smokeintensity, asthma, weakheart] = β0 + β1sex + β2race + β3age + β4smokeintensity + β5asthma + β6weakheart
   3. E[Ya – Ya=0|A, sex, race, age, smokeintensity, asthma, weakheart] = β1a + β2sex + β3race + β4age + β5smokeintensity + β6asthma + β7weakheart
   4. E[Ya – Ya=0|A, sex, race, age, smokeintensity, asthma, weakheart] = β1a + β2sex + β3race + β4age + β5smokeintensity + β6asthma + β7weakheart
   5. E[Ya=1 – Ya=0|A, sex, race, age, smokeintensity, asthma, weakheart, H(Ψ\*)] = β1a + β2sex + β3race + β4age + β5smokeintensity + β6asthma + β7weakheart + β8H(Ψ\*)
   6. **E[Ya – Ya=0|A, sex, race, age, smokeintensity, asthma, weakheart] = β1a**
   7. E[Y|A] = β0 + β1A
2. **[Essay question, ungraded] What are the modeling assumption(s) that were made? Please limit your response to one sentence per assumption.**

**The contributions of sex, race, age, smokeintensity, asthma, weakheart, and H(Ψ\*) to the log odds of smoking cessation are additive**

**The relationship between age and the log odds of smoking cessation are linear, conditional on sex, race, smokeintensity, asthma, weakheart, and H(Ψ\*).**

**The relationship between smokeintensity and the log odds of smoking cessation are linear, conditional on sex, race, age, asthma, weakheart, and H(Ψ\*).**

**The relationship between H(Ψ\*) and the log odds of smoking cessation are linear, conditional on sex, race, age, smokeintensity, asthma, and weakheart.**

**No effect modification for the effect of smoking cessation on weight gain by sex, race, age, smokeintensity, asthma, weakheart**

1. **[Multiple choice, ungraded] To consistently estimate the average causal effect in the population using g-estimation, are you, in general, forced to make any assumption(s) about effect modification?**
   1. **Yes**
   2. **No**
2. **[Essay question, ungraded] If you selected “No” in question 5, explain why you did not need to make any assumptions. If you selected “Yes” in question 5, explain the assumption(s) in a way that your uncle can understand. Please limit your response to one sentence.**

**Yes, you need to assume no effect modification for the effect of smoking cessation on weight gain by sex, race, age, smoke intensity, asthma and weakheart**