Title: Lean on your statistics: The generalization and simplification of the balance intercept problem

Word limit: 1500 words for AJE Research Letter or AJE classroom

Paragraphs:

1. Intro
   1. A strong statistics foundation is necessary even in the computational analytic era.
   2. We demonstrate why statistics foundation is important via visiting the balance intercept problem
   3. Summarize the contribution
      1. Dissect the balance intercept problem from a statistical perspective
      2. Develop a generalized closed-form solution for balance intercepts
      3. An alternative solution that simplifies the balance intercept problem.
2. Previous literature review:
   1. Rudolph et al. (2021)
   2. Robertson, Steingrimsson, and Dahabreh (2021)
   3. Zivich and Ross (2022)
3. Review of statistics knowledge
   1. Reference coding vs effect coding
   2. Link function
4. Explain the balance intercept problem from a statistics perspective
   1. Balance intercept is the conditional probability of the reference level.
5. Generalized solution
   1. Why analytic approximation fails
      1. Inequality between the expectation of a function and a function of expectation.
   2. Calculate the intercept on the response scale is the correct way
      1. A closed-form solution for logarithm link function.
6. Effect coding can be an alternative solution
   1. Explain when to use the effect coding: the conditional probability is known instead of the risk ratios/risk differences.
   2. Explain what the intercept term means in the effect coding system
   3. Discuss the calculation for the balance point using the effect coding system
      1. The marginal mean is the intercept
      2. Derive the marginal mean by taking the average of conditional means and translating it to the link predictor scale.
7. Summarize the contribution of the article
8. Link back to epidemiology education
   1. Bring home statement: statistics knowledge is still very important to understand an analytic problem and prevent computation errors and derive optimal solutions.