PH 250B Week 4, Tab 2 Ungraded quiz - ANSWERS
Topic: Measures of Association

Problem 1. (Fall 2017 250B Practice exam problem set)

A group of investigators used data from a study that recruited 840 women aged 55-65 at a state fair and followed them for 20 years to examine the association between consumption of deep-fried Twinkies during childhood and myocardial infarction (MI). Exactly one-third of women reported having consumed deep-fried Twinkies “frequently” during their childhood. The rest reported rare or no consumption of deep-fried Twinkies. Over the follow-up period, 90 of the women who reported frequent consumption of deep-fried Twinkies had an MI, and 110 of the women who reported rare or no consumption had an MI.

A. Draw a 2 x 2 table to present these findings.

	MI	No MI	Total
Frequent consumption	90	190	280
Rare or no consumption	110	450	560
Total	200	640	840

B. Calculate and interpret the appropriate relative measure of association for these data.

CIR = CI_e/CI_u = (90/280)/(110/560) = 1.636

The 20-year risk of MI in women reporting frequent consumption of deep-fried Twinkies was 1.64 times the risk of those reporting rare or no consumption.

C. Calculate and interpret the RD and APe% for these data.

RD = $CI_e - CI_u$ = (90/280) - (110/560) = 0.125

The risk of MI among women who reported frequent consumption of deep-fried Twinkies was 12.5 percentage points higher compared to women who reported rare or no consumption over 20 years.

Or: Among every 1,000 women who reported frequent consumption of deep-fried Twinkies there were 125 excess cases of myocardial infarction compared with women who reported rare or no consumption over 20 years.

APe% = $[(CI_e – CI_u)/CI_e] * 100 = 38.89\%$

Assuming a causal relationship between deep-fried Twinkie consumption and MI, of the incidence of MI in the women (ages 55-85) who frequently consumed deep-fried Twinkies during childhood, 39% of the incidence is in excess of the incidence of MI in women who did not consume deep-fried Twinkies frequently over 20 years.

D. Using only the data provided and assuming that this group of women is representative of the population of women aged 55-65 in the U.S. at the time of enrollment in the study, calculate <u>and</u> interpret the PRD and APt%.

Could use either:

PRD = $CI_t – CI_u = (200/840) – (110/560) = 0.0417$

PRD = $RD(P_e) = 0.125(1/3) = 0.0417$

In the total population of women (aged 55-65 at enrollment), the risk of MI was 4.17 percentage points higher compared with women who reported rare or no consumption of deep-fried Twinkies over 20 years.

Or: There are 417 excess cases of MI for every 10,000 women in the total population compared to women who reported rare or no consumption of deep-fried Twinkies over the 20 years.

APt% = $(PRD/CI_t) *100 = [0.0417/(200/840)] * 100 = 17.51\%$

((Rt – Ru) / Rt x 100%)

Assuming a causal relationship between deep-fried Twinkie consumption and MI, of the incidence of MI in the total population of women (aged 55-85), 17.5% is in excess of the incidence of MI in women who reported rare or no consumption of deep-fried Twinkies over 20 years.

E. Imagine that instead of an observational study, the data described came from the ideal experiment in which a population is exposed and observed for outcomes over time, and then the same population is unexposed and observed for outcomes over the same time with everything else the same. Recall the causal types (doomed, causal, preventive, immune). Assume that frequent consumption of deep-fried Twinkies never prevents MI. Describe in words what the RD and RR represent in terms of these causal types. You may write down a formula if it helps you, but we are looking for a description *in words.*

[Scoring for *each* RD and RR: 2 points for a correct answer, 1 point for a partially correct answer or formula, 0 points if no correct information].

RD
The RD describes the total incidence of MI caused by frequent consumption of deep-fried Twinkies. The RD describes the proportion of causal types in the population.

125/1000 cases of MI are caused by frequent consumption of deep-fried Twinkies.

RR
The RR describes the incidence of MI caused by frequent consumption of deep-fried Twinkies and other causes (those without frequent consumption of deep-fried Twinkies) relative to incidence due to other causes of MI. The RR describes the proportion of causal and doomed types in the population relative to the proportion of doomed types.

The RR measure tells us that incidence of MI due to all causes (including frequent consumption of deep-fried Twinkies and other causes) was 1.64 times the incidence due to causes other than frequent consumption of deep-fried Twinkies.

Problem 2. (Fall 2017 250B Practice exam problem set)

Researchers wish to investigate a potential association between use of in-ear headphones to listen to music and hearing damage among Berkeley students (25,885 undergraduates). They ask for your assistance in thinking through their approach. They have a list of all students in University housing (18,000). They plan to draw a random sample from the housing list and send a survey to invite participants to take a survey assessing in-ear headphone use (binary exposure) and be screened for hearing loss (binary outcome). Based on prior studies, they know ahead of time that 10% of the students in University housing never respond to surveys when invited to participate.

a. The investigators are interested in the net amount of hearing loss created through both harmful and preventive effects of in-ear headphone use. What type of measure do you recommend to isolate the overall causal impact of in-ear headphone use, assuming for the moment an unbiased study design with minimal random error? *Hint: think about which of the four causal types (doomed, immune, etc.) the investigators are interested in.* (1 point)

1 point for any of: absolute measure, risk difference, rate difference, excess risk/rate, CID, IDD, PRD, APe%, APt%

No credit if also or only mentioned relative measures

Problem 3. (Fall 2017 250B Problem Set)

Imagine a counterfactual (i.e. causal) study examining the association between riding public transit at least once per day and hospitalization for seasonal flu among the population of Berkeley (approximately 100,000 people according to the 2000 census) in December 2010.

a. Write an equation to represent the population level causal effect of riding public transportation at least once per day on hospitalization for flu.

$$E[Y_1 - Y_0]$$

where $Y = \begin{cases} 1, riding\ public\ transit \\ 0, not\ riding\ public\ transit \end{cases}$

b. The study finds that when everyone in the population rode public transit at least once a day, the cumulative incidence of first time hospitalization for flu was 250 cases per 100,000 population over 1 month. When no one rode public transit, the cumulative incidence was 95 cases per 100,000 population over 1 month. Calculate the causal risk ratio (RR), the RD and the APe%. Interpret these measures.

$$RR = \frac{R_{exp}}{R_{unexp}}$$

$$RR = \frac{250/100{,}000}{95/100{,}000} = 2.63$$

- The 1-month risk of flu-related hospitalization for flu when everyone rode public transit at least once a day was 2.63 times the risk when no one rode public transportation at least once a day.
- The RR measure tells us that 1-month cumulative incidence [risk] due to all causes (including riding public transit daily and other causes) was 2.63 times the cumulative incidence due to causes other than riding public transit daily.

$$RD = R_{exp} - R_{unexp}$$

$$RD = \frac{250}{100{,}000} - \frac{95}{100{,}000}$$

$$RD = \frac{155}{100{,}000}$$

- There were 155 excess cases of flu-related hospitalization per month due to exposure (only using causal language because this is a counterfactual experiment) when everyone rode public transit at least once a day (out of 100,000 people).
- The RD measure tells us that 155 cases of hospitalization were caused by riding public transit when everyone rode public transit at least daily (out of 100,000 people).
- **Remember that this is a counterfactual experiment. If it were not, we would say: The RD measure tells us that there were 155 excess cases of hospitalization *among those who rode public transit* compared with those who did not ride public transit.

$$APe\% = \frac{R_{exp} - R_{unexp}}{R_{exp}} * 100$$

$$APe\% = \frac{\frac{250}{100,000} - \frac{95}{100,000}}{\frac{250}{100,000}} * 100$$

$$APe\% = 62\%$$

- 62% of the cases of flu-related hospitalization per month were due to exposure (only using causal language because this is a counterfactual experiment) when everyone rode public transit at least once a day.
- APe%: The APe% tells us that 62% of cases of flu-related hospitalization were caused by riding public transit when everyone rode public transit at least once per day.
- **Remember that this is a counterfactual experiment. If it were not, we would say: The APe% tells us that 62% of 1-month cumulative incidence of flu hospitalization among those who rode public transit daily is in excess of the 1-month cumulative incidence of flu hospitalization among those who did not ride public transit. Note that this association (i.e., non-causal) interpretation of the APe% is awkward - typically the APe% and APt% are only calculated when the decision has been made to interpret an association causally.

c. Which measure, RR or RD, is better able to isolate the cases in which riding public transit daily was a cause? Justify your response.

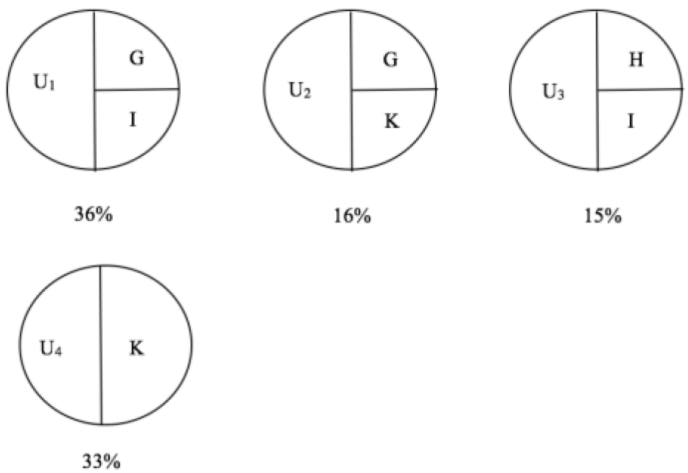
   The RD is better able to isolate cases in which riding public transit daily was a cause because the magnitude of the RR depends on the incidence due to other causes (i.e., the denominator).

d. In another counterfactual experiment, it was found that 15% of flu cases could be prevented by washing hands before eating and 51% of cases could be prevented by getting a flu vaccine. Combined with the 52% of cases that could be prevented by not riding public transit, this adds up to 118% of cases prevented. **How is this possible?**

   If you think about causal pies it helps clarify this. “Blocking” the effect of one cause may prevent one or more sufficient causes from becoming completed, but blocking the effect of another cause might also cause the same or different causes from becoming completed. That is, the same sufficient cause could be prevented by blocking more than one cause. Put simply, this is possible because component causes are being double counted in different sufficient causes.

For example: Riding public transit = G; Not washing hands = H; No vaccine = I; Working in an office building = K; Unknown cause = U

- Blocking G prevents both pies 1 and 2 (52%). Blocking H prevents pie 3 (15%). Blocking I prevents pie 1 and 3 (51%).



Problem 4. (Fall 2017 250B Practice exam problem set)

The same group of researchers is analyzing data from another study in which 918 women aged 55 to 65 were followed-up for 20 years:

	MI	No MI	Total
Blood pressure			
High	145	355	500
Low	55	363	418
Total	200	718	918
Family history			

Yes	120	148	268
No	80	570	650
Total	200	718	918

A. Among those with high blood pressure, calculate the percentage of the MI incidence that is associated with high blood pressure. *1 point*

[Scoring: 1 point if correct answer, 0 points if wrong].

APe% = $[(CI_e – CI_u)/CI_e] * 100 = [((145/500)-(55/418))/(145/500)] *100 = 54.63$

B. Among those with family history, calculate the percentage of the MI incidence that is associated with family history. *1 point*

[Scoring: 1 point for a correct answer, 0 points if wrong].

APe% = $[(CI_e – CI_u)/CI_e] * 100 = [((120/268)-(80/650))/(120/268)] *100 = 72.51$

C. Add up the percentages from (A) and (B). Do you think your answer is plausible? Why or why not? *2 points*

[Scoring: 2 points for a correct answer, 1 point for a partially correct answer, 0 points if totally wrong. You did not have to draw causal pies, but if you did and they made sense, you received credit for those.]

The APe%s add up to >100%. This is plausible because blocking the effect of one component cause may prevent one or more sufficient causes from becoming completed, but blocking the effect of another component cause might also prevent the same or different sufficient causes from becoming completed. That is, the same sufficient cause could be prevented by blocking more than one component cause.

Problem 5. When is the population risk difference (PRD) most useful from a public health perspective?

When the exposure is 1) clearly causally related to the outcome and 2) amenable to intervention.